Genomic selection for any dairy breeding program via optimized investment in phenotyping and genotyping

Jana Obšteter[1*], Janez Jenko[2], and Gregor Gorjanc[3,4]

[1]Department of Animal Science, Agricultural Institute of Slovenia, Hacquetova ulica 17, 1000 Ljubljana, Slovenia

[2]Geno Breeding and A.I. Association, Storhamargata 44, 2317 Hamar, Norway

[3]The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, EH259RG, United Kingdom

[4]Biotechnical Faculty, University of Ljubljana, Jamnikarjeva 101, 1000 Ljubljana, Slovenia

*Jana Obšteter, Agricultural Institute of Slovenia, Department of Animal Science, Hacquetova ulica 17, 1000 Ljubljana, Slovenia

E-mail addresses:

JO: jana.obsteter@kis.si

JJ: janez.jenko@geno.no

GG: gregor.gorjanc@roslin.ed.ac.uk

Abstract

Background: This paper evaluates the potential of maximizing genetic gain in dairy cattle breeding by optimizing investment into phenotyping and genotyping. Conventional breeding focuses on phenotyping selection candidates or their close relatives to achieve desired selection accuracy for breeders and quality for producers. Genomic selection decoupled phenotyping and selection and through this increased genetic gain per year compared to the conventional selection. However, genomic selection requires a large initial investment, which limits the adoption of genomic selection for some breeding programmes.

Methods: We simulated a case-study of a small dairy population with a number of scenarios under equal available resources. The conventional progeny testing scenario had 11 phenotype records per lactation. In genomic scenarios, we reduced phenotyping to between 10 and 1 phenotype records per lactation and invested the saved resources into genotyping. We tested these scenarios at different relative prices of phenotyping to genotyping and with or without initial training population for genomic selection.

Results: Reallocating a part of phenotyping resources for repeated milk records to genotyping increased genetic gain compared to the conventional scenario regardless of the amount and relative cost of phenotyping, and the availability of an initial training population. Genetic gain increased by increasing investment in genotyping, despite reduced phenotyping, and with high-genotyping scenarios not even using all the available resources. Compared to the conventional scenario, genomic scenarios also increased accuracy for young non-phenotyped male and female candidates, and cows.

Conclusions: This study shows that breeding programmes should optimize investment into phenotyping and genotyping to maximise return on investment. Our results suggest that any dairy

breeding programme using conventional progeny testing with repeated milk records can implement genomic selection without increasing the level of investment.

Background

This paper evaluates the potential of maximizing genetic gain in dairy cattle breeding by optimizing investment into phenotyping and genotyping. All breeding programmes strive to maximize genetic gain, which is a function of selection intensity, accuracy of selection, genetic variation, and generation interval. The conventional dairy breeding programme uses a long and expensive progeny test, which limits selection intensity. This programme allocates most of resources into phenotyping to achieve sufficient accuracy of sire selection, since this is the main driver of genetic gain in conventional selection. Genomic selection [1, 2], on the other hand, achieves genetic gain mainly through substantially reduced generation interval, increased selection intensity on the male side, and increased accuracy of selection for young animals [2, 3]. Despite lower accuracy of sire selection compared to the conventional selection, genomic selection doubles the rate of genetic gain per year in dairy cattle [4].

All breeding programmes operate with a set amount of resources allocated to breeding activities with the aim to maximise return on investment. Genomic selection is now a de-facto standard in well-resourced breeding programmes, but is still challenging to implement in others. The major hurdle is the large initial investment in genotyping to establish a training population, though updating this population can also be challenging. We hypothesise that these breeding programs need to evaluate priorities and could optimize phenotyping and genotyping to maximise return on investment. We base this hypothesis on the following simple examples (see Additional file 1).

The accuracy of conventional (pedigree-based) estimates of breeding values increases with increasing heritability and increasing number of phenotype records per animal or its closest relatives [e.g. 5]. Assume a female-expressed trait with 0.25 heritability and progeny testing 100 sires each on 100 daughters (10,000 cows in total). Collecting 10 phenotype records per daughter (100,000 phenotypes) gives the accuracy of 0.98 for sires, 0.89 for cows, and 0.66 for non-phenotyped progeny. If we decrease the number of phenotype records per daughter to five, two,

or one, the accuracy respectively decreases to 0.97, 0.96, or 0.93 for sires; to 0.81, 0. 70, or 0.62 for cows; and to 0.64, 0.59, or 0.56 for non-phenotyped progeny. This example shows diminishing returns with repeated phenotype records and a scope for optimizing return on investment. Namely, at the extreme we reduced phenotyping 10x, but reduced accuracy only for 0.05 in sires and 0.10 in non-phenotyped progeny.

We could invest the saved resources from reducing the number of phenotype records per daughter into phenotyping more daughters. Still assuming resources for 100,000 phenotypes and decreasing the number of phenotype records per daughter to five, two, or one respectively enables phenotyping 200, 500, or 1,000 daughters per sire for 100 sires. Compared to the previous example this change increases accuracy for sires to 0.99 in all cases, slightly increases accuracy for cows, and respectively increases accuracy for non-phenotyped progeny to 0.64, 0.61, or 0.59. This example shows that genetic evaluation can be sufficiently accurate with fewer phenotype records per cow and more phenotyped cows.

The accuracy of genome-based estimates of breeding values also increases with increasing heritability and increasing number of phenotype records per genotyped animal, but also with increasing training population, decreasing genetic distance between training and prediction individuals, and decreasing number of effective genome segments [6–10]. Following the previous example, assume 10,000 effective genome segments, 0.25 heritability, and a training population of 10,000 cows. Recording 10 phenotype values per cow gives the heritability of training population phenotype of 0.78 and genomic prediction accuracy of 0.76 for non-phenotyped progeny [6]. Reducing the number of phenotype records per cow to five, two, or one respectively reduces the heritability of training population phenotype to 0.67, 0.49, or 0.38, and genomic prediction accuracy to 0.71, 0.63, or 0.56. This example again shows diminishing returns with repeated phenotyping and a scope for optimizing return on investment in genomic breeding programmes. Namely, at the extreme we reduced repeated phenotyping 10x and reduced genomic prediction accuracy for 0.20.

We could invest the saved resources from reducing the number of phenotype records per daughter into genotyping. If we could increase the number of genotyped and phenotyped cows from 10,000 to 20,000, 50,000, or 100,000, each respectively phenotyped with five, two, or one record, we would respectively increase the genomic prediction accuracy to 0.84, 0.90, or 0.93. While these genomic prediction accuracies are lower than with progeny testing, shorter generation interval enables larger genetic gain per unit of time [2]. Previous studies also explored the value of adding females to the training population [11, 12]. They concluded, that accuracy has diminishing returns with increasing the number of females in the training population, hence additional female is most valuable when the training population is small.

The above examples suggest that repeated phenotyping could be an internal reserve to enable dairy breeding programmes to implement genomic selection. In dairy breeding the most repeatedly and extensively recorded phenotypes are milk production traits. There are different milk recording methods that differ in the recording responsibility, sampling scheme, recording and sampling frequency, and the number of milkings per day [13]. The recording interval ranges from daily recording to recording every nine weeks, which translates to between 310 and 5 records per lactation. The different recording methods have different costs, which also vary considerably between recording systems, countries, and even their regions. For example, some organizations require payment of a participation fee plus the cost per sample, while others include the fee in the sample cost, or cover the costs in other ways.

The examples assumed simplified scenarios where cows form a training population and focused only on accuracy. Real breeding programmes involve overlapping generations, individuals with a mix of phenotype, pedigree, and genotype information, various selection intensities, and other dynamic components. While we can use the single-step genomic prediction to combine all phenotypic, pedigree, and genomic information [14–16] evaluating the optimal allocation of resources into phenotyping and genotyping is beyond simple examples.

The aim of this study was to evaluate the potential of maximizing genetic gain by optimizing investment into phenotyping and genotyping in dairy breeding programmes. Since milk recording is an example of a repeated phenotype with diminishing returns, we aimed to optimize investment into milk recording and genotyping. To this end we have compared a dairy breeding programme with conventional progeny testing and genomic testing under equal available resources. To implement genomic selection, we reduced the number of milk records per lactation and invested the saved resources into genotyping. We compared these strategies in a case-study of a small cattle breeding programme where implementing genomic selection is challenging. The results show that reallocating a part of phenotyping resources to genotyping increases genetic gain regardless of the cost and amount of genotyping, and the availability of an initial training population.

Methods

The study aimed to evaluate the effect of different investment into phenotyping and genotyping with a simulation of a case-study of a small dairy breeding programme. The simulation mimicked a real dairy cattle population of ~30,000 animals analysed in our previous study [3]. We evaluated 36 genomic scenarios against the conventional scenario, all with equal available resources, but varying extent of phenotyping and genotyping. The conventional scenario implemented progeny testing and collected 11 phenotype records per lactation, while genomic scenarios reduced phenotyping and invested saved resources to genotyping. The genomic scenarios differed in i) the number of phenotype records per lactation; ii) the relative cost of phenotyping and genotyping; and iii) the availability of an initial training population. All tested scenarios were compared based on their genetic gain and accuracy of selection.

Simulation of the base population, phenotype and historical breeding

The simulation mimicked a small dairy cattle breeding programme of ~30,000 animals with ~10,500 cows, where introduction of effective genomic selection is challenging. We use this

population as a case-study to optimize investment into phenotyping and genotyping. The breeding programme aimed to improve dairy performance, which we simulated as a single polygenic trait. For this we used a coalescent process to simulate genome comprised of 10 cattle-like chromosomes, each with 10^8 base pairs, 1,000 randomly chosen causal loci, and 2,000 randomly chosen marker loci. We sampled the effects of causal loci from a normal distribution and use them to calculate animal's breeding value (a_i) for dairy performance (y_{ijkl}), which was affected also by a permanent environment (p_i), herd (h_j), herd-year (hy_{jk}), herd-test-day (htd_{jkl}), and residual environment (e_{ijkl}) effects:

$$y_{ijkl} = a_i + p_i + h_j + hy_{jk} + htd_{jkl} + e_{ijkl}.$$

We sampled permanent environment effects from a normal distribution with zero mean and variance equal to a base population additive genetic variance (σ^2_A). We sampled herd, herd-year, and herd-test-day effects each from a normal distribution with zero mean and variance of 1/3 σ^2_A. Finally, we sampled residual environment effects from a normal distribution with zero mean and variance of σ^2_A. This sampling scheme gave a trait with 0.25 heritability and 0.50 repeatability. With the simulated genome and phenotype architecture we have initiated a dairy cattle breeding programme and ran it for 20 years of conventional selection with progeny-testing based on 11 cow phenotype records per lactation. The detailed parameters of the simulation are described in [3]. In summary, in the breeding programme we selected 3,849 out of 4,320 new-born females as cows and 139 as bull dams over their second, third, and fourth lactation. We generated 45 male calves from elite matings and out of these chose 8 for progeny testing of which 4 were eventually selected as elite sires. We made all selection decisions based on pedigree-based estimates of breeding values. The 20 years represented historical breeding and provided a starting point for evaluating future breeding scenarios, which we ran for additional 20 years.

Scenarios

We evaluated 36 genomic scenarios with varying the extent of phenotyping and genotyping against the conventional scenario. All scenarios had equal available resources. The conventional scenario continued the breeding scheme from the historical breeding. It used progeny testing and 11 phenotype records per lactation (named C11), corresponding to the standard ICAR recording interval of 4 weeks [13]. We assumed that this scenario represented the total resources available for generating the data. We then created genomic scenarios by distributing resources between phenotyping and genotyping - we reduced phenotyping and invested the saved resources into genotyping. In the genomic scenarios we selected females as in the conventional scenario and males on genomic prediction. We varied the number of genomically tested male candidates with the resources and selected the best 5 as elite sires on genomic prediction. We evaluated the genomic scenarios with varying number of phenotype records per lactation, relative cost of phenotyping to genotyping, and the availability of an initial training population.

Genomic scenarios reduced the number of phenotype records per lactation to between 10 and 1. The scenarios followed ICAR standards of 9, 8, and 5 records per lactation, corresponding to recording intervals of 5, 6, and 9 weeks. Additionally, we created three non-standard recording systems collecting 10, 2, and 1 record per lactation. We named the scenarios as “GX” with X being the number of records per lactation. The reduction in phenotyping and the relative cost of phenotyping to genotyping dictated the saved resources and the number of genotyped animals (Table 1). We invested the saved resources into genotyping females and males in ratio 7:1 based on our previous work [3]. We genotyped first parity cows. This maximized the accuracy of genomic prediction, since it reduced genetic distance between training and prediction population, prevented the loss of investment with culled heifers, and minimized the time to obtain a phenotype linked to a genotype. If the resources for genotyping females were larger than the cost of genotyping all first parity cows, we did not reallocate the excess of resources to genotyping males for consistency. To maximise the genetic gain, we genotyped male calves from elite matings and other high parent average matings.

Genomic scenarios next varied the relative cost of phenotyping ($P) to genotyping ($G). We compared the cost of one genotype to the cost of 11 phenotype records per lactation. Based on a survey of several breeding programmes, milk recording organizations, and genotyping providers we have considered three cost ratios of $P:$G: 2:1, 1:1, and 1:2. Following the survey, we implemented decreasing price of repeated milk recording - the first recording was the most expensive and the cost of each subsequent recording was 95% of the preceding recording.

Lastly, we created scenarios with and without an initial training population for genomic prediction. When we assumed an initial training population was available, we genotyped all active cows (10,653) and progeny tested sires (100) before the first genomic evaluation. When initial training population was not available, we yearly genotyped a designated number of first parity cows until the training population reached 2,000 cows. Once we reached this goal, we started to genotype both females and males as specified in Table 1. At that point we started genomic selection of males.

Estimation of breeding values

We selected the animals based on their breeding values estimated from a pedigree or single-step genomic repeatability model with breeding value, permanent environment, and herd-year as random effects. We did not fit the herd-test-day effect as data structure of this small population did not enable its accurate estimation. We estimated breeding values once a year with blupf90 [17] with default settings. In the estimation we included all available phenotype and pedigree records for all active, phenotyped, or genotyped animals and additional three generations of their ancestors. We used at most 25,000 genotype records due to a limit in the academic software version. When we accumulated more than 25,000 genotyped animals, we removed genotypes of the oldest animals in favour of the latest genotyped cows and male selection candidates.

Table 1 Number of genotyped animals per year by scenario and relative cost of phenotyping to genotyping.

	Scenario					
Relative cost	G10	G9	G8	G5	G2	G1
$P:$G = 1:2	160 F 22 M	350 F 50 M	590 F 85 M	1,610 F 235 M	3,230 F 465 M	3,850 F 565 M
$P:$G = 1:1	310 F 45 M	700 F 100 M	1,180 F 165 M	3,230 F 465 M	3,850 F 925 M	3,850 F 1,125 M
$P:$G = 2:1	620 F 90 M	1,400 F 295 M	2,360 F 335 M	3,850 F 925 M	3,850 F 1,845 M	3,850 F 2,245 M

Scenarios are named “G” for genomic, followed by the number of phenotype records per lactation. The number of phenotype records and the relative cost of phenotyping to genotyping ($P:$G) dictated the number of genotyped animals. We genotyped females (F) and males (M) in 7:1 ratio. We genotyped the females to update and increase the training population and males for selection.

Analysis of scenarios

All scenarios had equal available resources. We compared the scenarios based on their final genetic gain, which indicated return on investment, and accuracy of selection. We measured the genetic gain as an average true breeding value by year of birth and standardized it to have zero mean and unit standard genetic deviation in the first year of comparison. We measured the accuracy of breeding values as the mean correlation between true and estimated breeding values. We measured the accuracy separately for four groups of animals: i) male candidates (genotyped and non-phenotyped); ii) sires (currently used in artificial insemination); iii) female candidates (non-genotyped and non-phenotyped); and iv) cows (all active phenotyped cows and bull dams). We repeated simulation of the base population and each scenario 10 times and summarised them with mean and standard deviation across the replicates. We used Tukey’s multiple comparison test to test the significance of the difference between means.

Results

Genomic scenarios increased the genetic gain compared to the conventional scenario regardless of the number of phenotype records per lactation, relative cost of phenotyping to genotyping, and the

availability of an initial training population. Genomic scenarios with an initial training population increased the genetic gain of the conventional scenario by up to 143%, despite reduced phenotyping. Genetic gain increased with increasing investment into genotyping. Genomic scenarios increased accuracy for non-phenotyped male and female candidates, and cows. Scenarios without an initial training population showed the same trends for genetic gain and accuracy. Although these scenarios had a slightly smaller genetic gain due to delayed implementation of genomic selection, they still increased the genetic gain of the conventional scenario by up to 134%.

Genetic gain with an initial training population

With the same available resources, genomic scenarios with an initial training population increased the genetic gain of the conventional scenario between 79% and 143%. The genetic gain increased with the increasing investment in genotyping, despite reduced phenotyping. We show this in Figure 1 and Additional file 2 with genetic gain by scenario and relative cost of phenotyping to genotyping with an initial training population. We show the corresponding intensities of sire selection in Additional file 3. When phenotyping costed the same as genotyping ($P:$G = 1:1), the genomic scenarios increased the genetic gain of the conventional scenario between 79% and 143%. By reducing the number of phenotype records from 11 (C11) to 10 per lactation (G10), we saved resources for genotyping 355 animals per year (310 cows and 45 male candidates). This small change increased the male selection intensity from 0.80 to 1.71 and coupled with a shorter generation interval increased the genetic gain by 79% (from 3.01 to 5.41). By reducing the phenotype records to nine or eight per lactation (G9 or G8), we respectively saved resources to genotype 800 or 1,345 animals per year, of which 100 or 165 were male candidates. This respectively increased the male selection intensity to 2.06 or 2.27, and genetic gain by 109% or 120% (from 3.01 to 6.30 or 6.62). We achieved the highest genetic gain, between 135% and 143% of the conventional scenario (between 7.07 and 7.33), when we collected five, two, or one phenotype records per lactation. In these three scenarios we saved resources for genotyping between

3,230 and 3,850 (all) cows and between 465 and 1,125 male candidates per year, and achieved the male selection intensity between 2.63 and 2.93.

We observed a similar trend for genetic gain when phenotyping costed half or twice the genotyping. Changing the relative cost of phenotyping to genotyping had the largest effect in the scenario with the smallest amount of genotyping (G10). In this scenario, when phenotyping costed twice or half the genotyping, we respectively saved resources for genotyping 182 or 710 animals, of which 22 or 90 were males, and increased the genetic gain for 80% (from 3.01 to 5.43) or 116% (from 3.01 to 6.50). When we maximized the investment into genotyping (G1), we genotyped all females at all three price ratios and between 565 and 2,245 male candidates. Correspondingly, we achieved a comparable genetic gain, between 136% and 143% of the conventional scenario, regardless of the relative cost of phenotyping to genotyping and male selection intensities.

The high-genotyping scenarios achieved the observed genetic gain without using all the available resources (marked bold in Additional file 2). In these scenarios the resources designated to genotyping females exceeded the cost of genotyping all females. The savings could cover between 42 and 23,800 additional phenotypes or between 85 and 11,900 additional genotypes.

In Figure 1 we also show the growth of the training population for genomic prediction. The training population started with ~10,000 individuals and grew until reaching 25,000 individuals. The increase was not linear through all generations, since the procedure for choosing the training animals changed when the training population exceed 25,000 (only latest females and male candidates included).

Accuracy with an initial training population

Compared to the conventional scenario, genomic scenarios increased accuracy for young non-phenotyped and genotyped male and non-phenotyped and non-genotyped female candidates, and cows, but decreased accuracy for sires. We show this in Figure 2 with the accuracy for different groups of individuals with an initial training population and equal cost of phenotyping and

genotyping. In Additional file 4 we compare accuracies at all three relative costs of phenotyping to genotyping. When the cost of phenotyping was equal to the cost of genotyping, the accuracy for young genomically tested male candidates ranged between 0.90 and 0.91 and did not depend on the amount of phenotyping and genotyping. This was between 0.53 and 0.54 higher compared to the pre-selection for progeny testing in the conventional scenario. However, this was between 0.03 and 0.04 lower compared to the sire selection in the conventional scenario. In contrast to stable accuracy for young genomically tested male candidates, the accuracy for already selected sires decreased with reallocating phenotyping resources into genotyping. We observed the lowest accuracy for sires, 0.63, when we invested the most into genotyping (G1), and the highest, 0.75, when we invested the most into phenotyping (G10). Compared to the conventional scenario, the accuracy for already selected sires in the genomic scenarios was between 0.11 and 0.23 lower.

The accuracy for non-genotyped female candidates increased with increasing genotyping, despite reduced phenotyping. We observed the highest accuracy for female candidates, between 0.55 and 0.57, when we recorded five, two, or one phenotype record per lactation and invested the rest into genotyping. Compared to the conventional scenario, the genomic scenarios increased the accuracy for female candidates between 0.03 and 0.11. The accuracy for cows followed the same trend, but with higher values. We observed the highest accuracy for cows, between 0.77 and 0.79, when we collected five, two, or one phenotype record per lactation and investing the rest in genotyping. Compared to the conventional scenario, genomic scenarios increased the accuracy for cows between 0.11 and 0.29.

Changing the relative cost of phenotyping to genotyping affected primarily the accuracy for female candidates and cows. We observed that in the majority of scenarios the accuracy increased with decreasing the relative cost of genotyping, which enabled more genotyping. We observed the largest difference of 0.06 for female candidates and 0.12 for cows when we changed the relative cost of phenotyping from half to twice the cost of genotyping. Changing the relative costs, however, did not change the trends.

Genetic gain and accuracy without an initial training population

Genetic gain

When an initial training population was not available, we increased the genetic gain of the conventional scenario between 31% and 134% by optimizing investment in phenotyping and genotyping. We show this in Figure 3 with the genetic gain, training population size, and accuracy by scenario at equal cost of phenotyping and genotyping without an initial training population. The trends were in line with what we observed with an initial training population, that is, increasing genotyping increased genetic gain despite reduced phenotyping. However, all corresponding scenarios achieved between 2% and 28% smaller genetic gain than when an initial training population was available (Additional file 2).

When phenotyping costed the same as genotyping, genomic scenarios increased the genetic gain of the conventional scenario between 51% and 131%. Compared to when we had an initial training population, the corresponding scenarios achieved between 2% and 16% lower genetic gain. We observed the largest difference in the scenario that invested the least into genotyping (G10). In this scenario we needed six years to build a training population of 2,000 cows and implement genomic selection, since we only genotyped 355 cows per year. Increasing the investment into genotyping decreased this difference. We observed the smallest difference in the scenario that collected two phenotype records per lactations (G2) and implemented genomic selection already in the first year.

Changing the relative cost of phenotyping to genotyping did not change the overall trend, only the level of genetic gain in the low-genotyping scenarios. When phenotyping costed half the genotyping, the genomic scenarios increased genetic gain of the conventional scenario between 31% and 126%. The corresponding scenarios achieved between 4% and 28% lower genetic gain than when we had an initial training population. When phenotyping costed twice the genotyping, the genomic scenarios increased the genetic gain of the conventional scenario between 86% and

133%. The corresponding scenarios achieved between 3% and 14% lower genetic gain than when we had an initial training population.

Accuracy

As when we had an initial training population, genomic scenarios without an initial training population increased the accuracy for non-phenotyped male and female candidates, and cows. We show this in Figure 3 with the accuracy without an initial training population and equal cost of phenotyping and genotyping. In Additional file 4 we compare the accuracies of all scenarios. When phenotyping costed the same as genotyping, the accuracy for male candidates ranged between 0.84 and 0.91. In contrast to scenarios with initial training population, the accuracy increased with increasing the investment into genotyping, hence was significantly lower in the scenario that invested the least into genotyping. The accuracy for sires ranged between 0.64 and 0.74. Contrary to when we had an initial training population, we observed no clear trend of either increasing or decreasing accuracy. For female candidates the accuracy ranged between 0.47 and 0.56, and for cows between 0.56 and 0.76. For female candidates and cows the accuracies followed the trends of when we had an initial training population, where increasing genotyping increased the accuracy.

As in the scenarios with an initial training population, changing the relative cost of phenotyping to genotyping affected the accuracy for non-genotyped female candidates and cows, but also male candidates. Decreasing the relative cost of genotyping to phenotyping increased the accuracy in the majority of the scenarios, particularly the low-genotyping ones.

Discussion

Our results show that any dairy breeding programme using conventional progeny testing with repeated milk records can implement genomic selection without extra costs by optimizing the investment into phenotyping and genotyping. The estimation of breeding values requires continuous investment in data collection. While breeding programmes have established funding for

phenotyping, not all of them have well established funding for genotyping. We show that by reallocating a part of phenotyping resources into genotyping, breeding programmes can implement genomic selection and substantially increase genetic gain regardless of the amount and cost of genotyping, and availability of an initial training population. The results raise four discussion points: 1) how optimizing the investment in phenotyping and genotyping affects genetic gain; 2) how optimizing the investment in phenotyping and genotyping affects accuracy; 3) implications for dairy breeding programmes; and 4) limitations of the study. In the following we first discuss the results under equal cost of phenotyping and genotyping, and initial training population available. We then discuss changes at different costs and no initial training population.

Genetic gain

Genomic vs. conventional selection

Implementing genomic selection by optimizing the investment in phenotyping and genotyping increased genetic gain compared to the conventional selection. With an initial training population of 10,000 cows, all genomic scenarios outperformed the conventional scenario, mainly due to reduced generation interval in sire selection paths. This is in agreement with previous modelling and real data studies. Modelling showed increased genetic gain with genomic selection due to reduced generation interval compared to progeny test, despite reduced selection accuracy [2, 3, 18]. Real data confirmed that the main driver of genetic gain with genomic selection is the reduced generation interval in sires of sires and sires of dams paths. In the US Holstein population, these generation intervals recently decreased between 25% and 50% compared to the conventional selection [19]. Van Grevenhof et al. [11] computed a break-even size of a training population to achieve a comparable response with genomic and conventional selection. They showed, that if the generation interval is not reduced and the number of phenotypes is limited, genomic selection cannot compete with conventional selection. But when generation interval is halved, a training population with ~2,000 individuals with own performance or ~3,500 individuals with ten progeny gives comparable response as conventional selection for a trait with intermediate heritability. While the assumption of

an available initial training population might not be realistic for some populations, it can be achieved through international collaboration [20].

The genomic scenarios had higher genetic gain also because the reduced number of phenotype records did not proportionally reduce selection accuracy. While genomic scenarios slightly decreased sire selection accuracy, they increased cow and dam selection accuracy. We discuss this in detail below.

Another major advantage of the genomic scenarios was increased intensity of sire selection. A costly and lengthy progeny-testing limits the number of tested sires in conventional selection. Genomic selection significantly reduces the cost of testing [2] and thus allows for testing more sires. In the US Holstein population, genomic selection improved the selection differential for all traits, particularly for traits with low heritability, such as health and fertility [19].

Increasing the investment into genotyping

Genetic gain increased with increased investment into genotyping. This was mainly due to higher intensity of sire selection, since more resources for genotyping allowed us to test more male candidates while selecting the same number. A larger investment into genotyping also increased update and size of the training population, which assisted in achieving genetic gain. This is in agreement with Thomasen et al. [21] who showed that adding more cows yearly to the training population increases genetic gain.

The genetic gain had diminishing relationship with investment into genotyping. This has important implications for dairy breeding programmes, since they use phenotypes also for management, and we discuss this separately. The results showed that investing resources of more than six phenotype records into genotyping did not significantly improve the genetic gain. There are four reasons for this. First, the accuracy of sire selection in genomic scenario was high regardless of the amount of genotyping when there were at least 10,000 animals in the training population. Second, the intensity of sire selection had diminishing relationship with increasing genotyping. This agrees with

Reiner-Benaim et al. [22] that showed an increased genetic gain with increasing the number of tested male candidates, but with a diminishing return. While they achieved the maximum profit with four selected sires out of 1,721 tested candidates, they achieved 99% or 90% of the maximum profit with respectively 740 or 119 tested candidates. Third, increasing female training population has diminishing relationship with genetic gain [11, 12]. Since our scenarios with initial training population started with ~10,000 genotyped and phenotyped cows, enlarging the training population had a marginal effect. And fourth, increasing investment into genotyping did not proportionally increase the size of the training population due to limited number of animals in the studied population and limited size of the training population. Once the investment sufficed to genotype all the females or when the size of the training population hit 25,000, investing more into genotyping did not increase the training population. The same four reasons enabled comparable maximum genetic gain regardless of the relative price of phenotyping to genotyping. In general, selecting less than 2% of the tested males and updating the training population with more than 35% of first-parity cows resulted in the maximum genetic gain.

While genetic gain increases with the number of cows in training population, repeated records do not have the same relationship. As we increased the number of cows in the training population, the number of repeated records decreased (Additional file 5). The scenarios with the largest genetic gain therefore had a training population with many cows and few repeated records. However, since we used the single-step genomic prediction, the phenotypes of the non-genotyped animals contributed to the estimation as well. Effectively, all scenarios thus operated with the same number of phenotyped animals.

We should emphasize, that some of the high-genotyping scenarios achieved the observed genetic gain at a lower total cost, since they could not use all the saved resources for genotyping females in the studied population. The saved resources could be invested back into phenotyping females for milk production or novel traits, genotyping more male candidates, or other breeding actions.

Scenarios without an initial training population

We also considered that some small populations do not have access to an initial training population and have to initialize one themselves. These genomic scenarios still increased genetic gain compared to the conventional scenario, but achieved lower genetic gain than corresponding scenarios with an initial training population available. This was mainly due to a delay in implementing genomic selection and a smaller training population. Consequently, increasing the investment into genotyping compensated for starting without a training population in two ways. First, it shortened the time to obtain the targeted 2,000 genotypes required to implement genomic selection down to one year in high-genotyping scenarios. Second, it shortened the time to build a training population in which an additional record had negligible effect on accuracy [12].

When implementing genomic selection with a delay due to building the training population, we ran a conventional selection with reduced phenotyping until we accumulated the targeted 2,000 genotypes. In this period, we did not observe decreased genetic gain compared to the conventional scenario with full phenotyping. This suggests that breeding programmes can run a conventional breeding programme with reduced phenotyping until they accumulate genotypes to initiate genomic selection, without harming the genetic gain in the accumulation or transition period.

Accuracy

Despite reduced phenotyping, genomic scenarios increased the accuracy for young non-phenotyped calves and cows. In general, genomic prediction increases the accuracy of the Mendelian sampling term. This is the main reason for increase in accuracy with genomic prediction when the accuracy of parent average is already high. But when the accuracy of parent average is low, such as for animals with non-phenotyped parents or parents with little own or progeny information, genomic information increases accuracy both for the parent average and the Mendelian sampling term [23, 24].

Accuracy for males with initial training population

For male candidates, genomic prediction more than doubled the accuracy compared to the parent average used in pre-selection for progeny testing in conventional scenario. This is in agreement with two-fold accuracy increase in dairy [2] and layers [23]. Within the genomic scenarios, the accuracy for male candidates was high regardless of the amount of genotyping and phenotyping for two reasons. First, the accuracy of their parent average was high, since we tested offspring of elite matings. Second, starting with an initial 10,000 training population gave an adequate accuracy that was additionally boosted by using all available information jointly through the single-step genomic prediction. Using single-step genomic prediction also removed the bias due to pre-selection [25].

In contrast, reducing phenotyping decreased the accuracy for selected sires. We believe this is due to two reasons. First is the fact that sire breeding values are in the tail of a distribution. Each additional phenotypic record only marginally changes the overall accuracy of individuals breeding values, but that affects distinguishing the very best sires in the tail of a distribution. Second, as we invested more into genotyping, the training population grew quicker and reached the limit of 25,000. At this point we removed sires' in favour of cows' genotypes, hence prediction for sires depended only on daughters data and no longer on own genotype. However, since this is the accuracy after the selection has already been made, it is not of great interest for breeding.

Accuracy for females with initial training population

Genomic scenarios increased the accuracy for cows compared to the conventional scenario. Besides increasing the accuracy of Mendelian sampling term, using genomic information increases genetic connectedness between individuals from different management units [26, 27]. This in turn increases the accuracy of prediction regardless of the heritability and the number of causal loci or markers [28].

The accuracy for cows increased with increasing investment into genotyping, despite reduced phenotyping. This had important implications, since we selected bull dams for elite mating from

cows. Increasing the number of genotyped cows affected the accuracy in three ways. First, more cows had both genomic and phenotypic information available, which increased the accuracy of their estimated breeding values. Second, more genotyped cows increased genetic connectedness [28]. And third, investing more into genotyping translated into larger training population and its yearly update. As shown by previous studies [11, 12], the accuracy of genomic prediction increases with increasing the size of a female training population. They showed that the accuracy of 0.70 is achieved with ~20,000 animals as in our study. However, these studies did not account for varying the degree of genetic distance between a training and prediction population. We can increase the accuracy in the evaluation population with a higher relationship to the training population [8, 9, 29]. Increasing the investment into genotyping allowed us to genotype more recent cows, which decreased genetic distance between our training and prediction populations. As with genetic gain, accuracy had a diminishing return relationship with the size of the training population [11, 12]. We observed plateau in accuracy when we invested more than six phenotype records into genotyping.

Accuracy for female candidates followed the accuracy trend for the dams, but at lower values. Female candidates were not genotyped nor phenotyped, hence their accuracy mainly reflected the accuracy of their parent average. Increasing genotyping increased the accuracy for dams and in turn increased the accuracy of female candidate`s parent average. The benefit of this increase was not large, since the intensity of cow selection was low. However, there is potential for this benefit to be larger with sexed semen and embryo transfer.

Accuracy without an initial training population

Accuracy in scenarios without an initial training population closely followed the trends of the corresponding scenarios with an initial training population available. We observed minor differences in the low genotyping scenarios that had reduced accuracy for male candidates and sires. We attribute this to a smaller training population. Buch et al. [30] showed that for new traits and large scale recording, we can achieve 75% of the maximum genomic accuracy within first two to

three years of recording. In our study we shortened this period even more by including the historical data through the single-step genomic prediction.

Implications

The results suggest that any dairy breeding programme using conventional progeny testing with repeated milk records can implement genomic selection without extra costs by optimizing the investment of resources into breeding actions. Here we propose funding the genotyping with a part of resources for milk recording, since we can manipulate the number of repeated records. Breeding programmes could reduce phenotyping for a different trait that they record repeatedly and is perhaps less crucial for management. They could also reallocate the funds from another breeding action.

Another alternative could be to optimize which individuals to genotype and phenotype, which we did not consider in this study. We expect this would further increase the genetic gain for the same level of investment or require less investment for the same genetic gain, but increase the complexity of optimization. Such optimizations were shown to increase the accuracy of genomic prediction up to 20% with small sample sizes in plant breeding [31, 32]. Similarly, selective genotyping of cows from the distribution tails has been shown to increase the accuracy of genomic prediction by 15% [33]. Furthermore, there are proposals for phenotyping farms, which would be paid to provide records [34].

The economic efficiency of breeding programmes strongly depends on which stakeholders fund which breeding action. Different programmes have different investment schemes, often intricate. The scenarios presented in this paper are of little value for programmes where funding for phenotyping and genotyping is disconnected. Similarly, optimizing the investment into phenotyping is not of interest for breeding programmes with abundant use of automated milking systems. With automated systems the cost of phenotyping does not depend on the number of records. But in populations with small herds the use of automated system is limited. Further on, genomic selection

could benefit some settings more than others. For example, genomic information is especially important for generating genetic connectedness in systems with small herd sizes, geographically dispersed farms, and limited use of artificial insemination, often found in low to mid income countries [27]. The same benefits are expected for small ruminant programmes that do not actively exchange of sires between herds [35].

When breeding programmes do not have access to high performance computing needed for genomic evaluation of large genotyped populations, they can also optimize computational costs. As shown in our study, we can achieve large genetic gain with a relatively small training population of recent genotypes. This implies that breeding programmes do not have to use all historical genotypes for prediction. The problem of a large number of genotypes can be alternatively solved by using methods with reduced computational costs, such as algorithm for proven and young [36] or singular value decomposition of the genotype matrix [37].

We did not account for the benefits of genotyping besides predicting genomic breeding values and selection. Genomic information has additional value for parentage verification or parentage discovery [38], management of monogenic diseases and traits, and better monitoring and control of inbreeding [39] and optimization of matings [3]. These additional uses of genotypes increase the return on investment in genomic selection beyond what we measured in this study.

Limitations of the study

Reducing the number of phenotype records

This study on balancing phenotyping and genotyping can lead to conflicts between managing production (short-term goal) and achieving genetic gain (long-term goal). Producers use phenotype records to manage animals' health and feed composition, which affect milk yield and its composition. Besides managing production, milk recording is also important from an environmental perspective [40], but so is genetic improvement. In general, about half of phenotypic improvement is due to management and half due to selection [41]. Until recently, data that drives dairy selection

has largely been collected for management and used for free in genetic evaluation. With the advent of genomic selection, the new genotype data is largely used for selection, but the same data could also serve in management (predicting feed requirements, disease liability, etc.). Therefore, evaluating the value of phenotype and genotype data is complex and beyond the scope of this study. One possible way forward would be to compare variance between herd-test day effects and genetic variance to contrast the value of managing production and genetic gain in addition to comparing phenotypic and genetic trends [41].

The longest sampling interval tested in our study and still approved by ICAR was nine weeks, which yielded five records per lactation and invested the resources of six records into genotyping. In most settings this sufficed to achieve the maximum genetic gain and selection accuracy. Previous studies also showed a good predictive ability of such scheme for estimating the 305-day milk yield [42,43].

Single additive trait

We simulated milk yield as a single polygenic trait with additive genetic as well as herd, permanent environment, and residual environmental effects. In reality, non-additive genetic effects also affect dairy performance [44–47]. We did not simulate nor account for these effects, but we note that permanent environment effects capture non-additive genetic effects and other individual specific environmental effects [46]. We also simulated milk yield in different lactations as a single trait with constant heritability through the lactation, whereas genetic correlation between different lactations and through the lactation is not unity [48–50].

Genomic selection of females

We did not use genomic selection for females nor did we use reproductive technologies such as sexing semen or embryo transfer. This would further decrease the generation interval, increase selection intensity on female side, and in turn increase genetic gain of genomic scenarios even more [18, 19]. Such an implementation of genomic selection requires only a minor modification of the

design used in this study - genotyping heifers instead of first-parity cows. However, reproductive technologies require a larger modification and investment. Some of the scenarios saved resource and could invest into these technologies.

Conclusion

This study suggests that any dairy breeding programme using conventional progeny testing with repeated milk records can implement genomic selection with no extra costs by optimizing the investment into milk phenotyping and genotyping. We showed, that allocating some phenotyping resources into genotyping increased both genetic gain and selection accuracy for non-phenotyped candidates, despite reduced phenotyping. The increase was observed regardless of the amount and cost of genotyping, and availability of an initial training population. However, increasing investment in genotyping has diminishing returns, which suggests that breeding programmes should optimize the investment into phenotyping and genotyping to maximise return on investment for selection and management.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Competing interests

Not applicable

Funding

GG acknowledges support from the BBSRC to The Roslin Institute (BBS/E/D/30002275) and The University of Edinburgh's Data-Driven Innovation Chancellor's fellowship.

Authors' contributions

JO designed the study, ran the simulation, analyzed the data, interpreted the results, and wrote the manuscript. JJ participated in study design, results interpretation, and manuscript revision. GG has initiated and supervised the work and contributed to all its stages.

Acknowledgements

The authors would like to thank Ivan Pocrnić from The Roslin Institute, University of Edinburgh, for his help in interpreting the results and comments on the manuscript.

Author's information (optional)

Not applicable.

References

[1] Meuwissen TH, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. Genetics. 2001;157:1819–1829.

[2] Schaeffer LR. Strategy for applying genome-wide selection in dairy cattle. J Anim Breed Genet. 2006;123:218–223.

[3] Obšteter J, Jenko J, Hickey JM, Gorjanc G. Efficient use of genomic information for sustainable genetic improvement in small cattle populations. J Dairy Sci. 2019;102:9971–9982.

[4] Wiggans GR, Cole JB, Hubbard SM, Sonstegard TS. Genomic Selection in Dairy Cattle: The USDA Experience. Annu Rev Anim Biosci. 2017;5:309–327.

[5] Mrode RA. Linear Models for the Prediction of Animal Breeding Values. 3rd edition. Wallingford, UK: CABI, 2014.

[6] Daetwyler HD, Villanueva B, Woolliams JA. Accuracy of Predicting the Genetic Risk of Disease Using a Genome-Wide Approach. PLoS ONE. 2008;3:e3395.

[7] Goddard M. Genomic selection: prediction of accuracy and maximisation of long term response. Genetica. 2009;136:245–257.

[8] Habier D, Tetens J, Seefried F-R, Lichtner L, Thaller G. The impact of genetic relationship information on genomic breeding values in German Holstein cattle. Genet Sel Evol. 2010;42:5.

[9] Clark SA, Hickey JM, van der Werf JHJ. Different models of genetic variation and their effect on genomic evaluation. Genet Sel Evol. 2011;43:18.

[10] Goddard ME, Hayes BJ, Meuwissen THE. Using the genomic relationship matrix to predict the accuracy of genomic selection. J Anim Breed Genet. 2011;128:409–421.

[11] Van Grevenhof EM, Van Arendonk JAM, Bijma P. Response to genomic selection: the Bulmer effect and the potential of genomic selection when the number of phenotypic records is limiting. Genet Sel Evol. 2012;44:26.

[12] Gonzalez-Recio O, Coffey MP, Pryce JE. On the value of the phenotypes in the genomic era. J Dairy Sci. 2014;97:7905–7915.

[13] International Committee for Animal Recording. Section 02 - Cattle Milk Recording. Overview. In: ICAR Guidelines. Rome: ICAR, 2017.

[14] Gao H, Christensen OF, Madsen P, Nielsen US, Zhang Y, Lund MS, et al. Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. Genet Sel Evol. 2012;44:8.

[15] Gray KA, Cassady JP, Huang Y, Maltecca C. Effectiveness of genomic prediction on milk flow traits in dairy cattle. Genet Sel Evol. 2012;44:24.

[16] Lourenco D a. L, Tsuruta S, Fragomeni BO, Masuda Y, Aguilar I, Legarra A, et al. Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. J Anim Sci. 2015;93:2653–2662.

[17] Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee DH. BLUPF90 and related programs (BGF90). In: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production: 19-23 August 2002; Montpellier, France.

[18] Pryce JE, Goddard ME, Raadsma HW, Hayes BJ. Deterministic models of breeding scheme designs that incorporate genomic selection. J Dairy Sci. 2010;93:5455–5466.

[19] García-Ruiz A, Cole JB, VanRaden PM, Wiggans GR, Ruiz-López FJ, Tassell CPV. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. PNAS. 2016;113:E3995–E4004.

[20] Jorjani H. Genomic evaluation of BSW populations, InterGenomics: Results and Deliverables. Interbull Bulletin. 2012;42:5-8.

[21] Thomasen JR, Liu H, Sørensen AC. Genotyping more cows increases genetic gain and reduces rate of true inbreeding in a dairy cattle breeding scheme using female reproductive technologies. J Dairy Sci. 2020;103:597–606.

[22] Reiner-Benaim A, Ezra E, Weller JI. Optimization of a genomic breeding program for a moderately sized dairy cattle population. J Dairy Sci. 2017;100:2892–2904.

[23] Wolc A, Arango J, Settar P, Fulton JE, O'Sullivan NP. Persistence of accuracy of genomic estimated breeding values over generations in layer chickens. Genet Sel Evol. 2011;43:23.

[24] Daetwyler HD, Villanueva B, Bijma P, Woolliams JA. Inbreeding in genome-wide selection. Journal of Animal Breeding and Genetics. 2007;124:369–376.

[25] Jibrila I, ten Napel J, Vandenplas J, Veerkamp RF, Calus MPL. Investigating the impact of preselection on subsequent single-step genomic BLUP evaluation of preselected animals. Genet Sel Evol. 2020;52:42.

[26] Yu H, Spangler ML, Lewis RM, Morota G. Genomic Relatedness Strengthens Genetic Connectedness Across Management Units. G3 (Bethesda). 2017;7:3543–3556.

[27] Powell O, Mrode R, Gaynor RC, Johnsson M, Gorjanc G, Hickey JM. Genomic data enables genetic evaluation using data recorded on LMIC smallholder dairy farms. BioRxiv. 2019; doi: https://doi.org/10.1101/827956.

[28] Yu H, Spangler ML, Lewis RM, Morota G. Do stronger measures of genomic connectedness enhance prediction accuracies across management units? J Anim Sci. 2018;96:4490–4500.

[29] Pszczola M, Strabel T, Mulder HA, Calus MPL. Reliability of direct genomic values for animals with different relationships within and to the reference population. J Dairy Sci. 2012;95:389–400.

[30] Buch LH, Kargo M, Berg P, Lassen J, Sørensen AC. The value of cows in reference populations for genomic selection of new functional traits. Animal. 2012;6:880–886.

[31] Heslot N, Feoktistov V. Optimization of selective phenotyping and population design for genomic prediction. bioRxiv. 2017; https://doi.org/10.1101/172064.

[32] Akdemir D, Isidro-Sánchez J. Design of training populations for selective phenotyping in genomic prediction. Sci Rep. 2019;9:1446.

[33] Jenko J, Wiggans GR, Cooper TA, Eaglen SAE, Luff WGL. Cow genotyping strategies for genomic selection in a small dairy cattle population. J Dairy Sci. 2017;100:439–452.

[34] Coffey M. Dairy cows: in the age of the genotype, #phenotypeisking. Anim Fron. 2020;10:19–22.

[35] Kasap A, Mioc B, Hickey JM, Gorjanc G. Genetic connectedness in the U.S. sheep industry. In: Book of Abstracts of the 69th Annual Meeting of the European Federation of Animal Science: 27-31 August 2018; Dubrovnik, Croatia.

[36] Misztal I, Legarra A, Aguilar I. Using recursion to compute the inverse of the genomic relationship matrix. J Dairy Sci. 2014;97:3943–3952.

[37] Ødegård J, Indahl U, Strandén I, Meuwissen THE. Large-scale genomic prediction using singular value decomposition of the genotype matrix. Genet Sel Evol. 2018;50:6.

[38] ICAR DNA Working Group. ICAR Guidelines for Parentage Verification and Parentage Discovery Based on SNP Genotypes. The International Committee for Animal Recording, November 2017.

[39] Sonesson AK, Woolliams JA, Meuwissen THE. Genomic selection requires genomic control of inbreeding. Genet Sel Evol. 2012;44:27.

[40] Verbič J, Jenko J, Jeretina J, Babnik D. Milk urea concentration as a tool to reduce the nitrogen footprint of milk production in conditions of small scale farming. In: Towards Precision livestock husbandry and its potential to mitigate ammonia and GHG emissions: Abstracts' Leaflet: 17 – 21 February 2019; Galilee, Israel.

[41] Dekkers JCM, Hospital F. The use of molecular genetics in the improvement of agricultural populations. Nat Rev Genet .2002;3:22–32.

[42] Pool MH, Meuwissen THE. Prediction of Daily Milk Yields from a Limited Number of Test Days Using Test Day Models. J Dairy Sci. 1999;82:1555–1564.

[43] Berry DP, Olori VE, Cromie AR, Veerkamp RF, Rath M, Dillon P. Accuracy of predicting milk yield from alternative milk recording schemes. Animal Science. 2005;80:53–60.

[44] Fuerst C, Sölkner J. Additive and Nonadditive Genetic Variances for Milk Yield, Fertility, and Lifetime Performance Traits of Dairy Cattle. J Dairy Sci. 1994;77:1114–1125.

[45] Ertl J, Legarra A, Vitezica ZG, Varona L, Edel C, Emmerling R, et al. Genomic analysis of dominance effects on milk production and conformation traits in Fleckvieh cattle. Genet Sel Evol. 2014;46:40.

[46] Aliloo H, Pryce JE, González-Recio O, Cocks BG, Hayes BJ. Accounting for dominance to improve genomic evaluations of dairy cows for fertility and milk production traits. Genet Sel Evol. 2016;48:8.

[47] Jiang J, Shen B, O'Connell JR, VanRaden PM, Cole JB, Ma L. Dissection of additive, dominance, and imprinting effects for production and reproduction traits in Holstein cattle. BMC Genomics. 2017;18:425.

[48] Meyer K. Estimates of genetic parameters for milk and fat yield for the first three lactations in British Friesian cows. Animal Science. 1984;38:313–322.

[49] Dong MC, van Vleck LD. Correlations Among First and Second Lactation Milk Yield and Calving Interval. J Dairy Sci. 1989;72:1933–1936.

[50] Swalve H, Vleck LDV. Estimation of Genetic (Co) Variances for Milk Yield in First Three Lactations Using an Animal Model and Restricted Maximum Likelihood. J Dairy Sci. 1987;70:842–849.

Figures

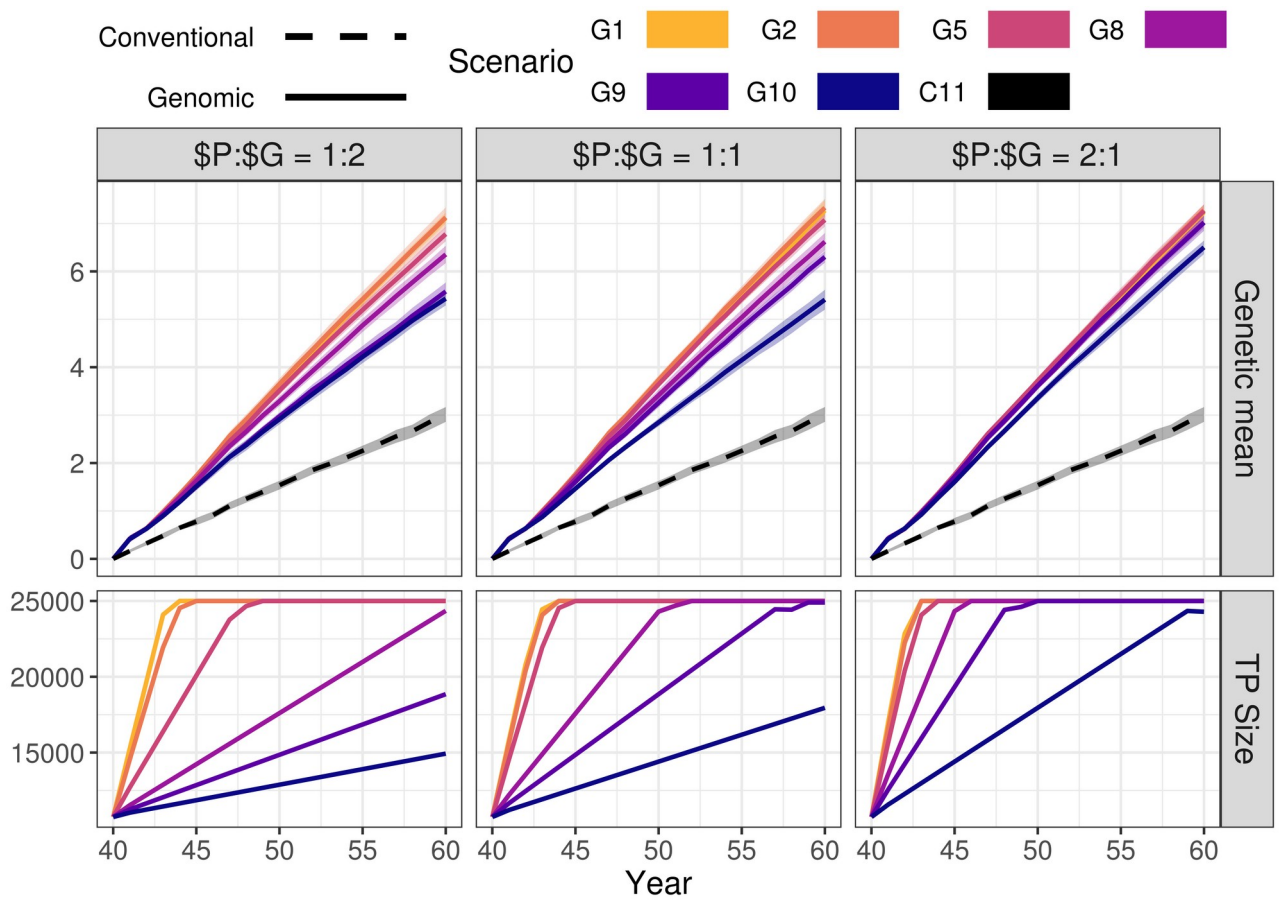


Figure 1 Genetic gain and training population size by scenario and relative cost of phenotyping to genotyping ($P:$G) with an initial training population (TP).

The figure presents the means (lines) and 95% confidence intervals (polygons) across 10 replicates for the conventional (C) and genomic (G) scenarios, with numbers indicating the number of phenotype records per lactation.

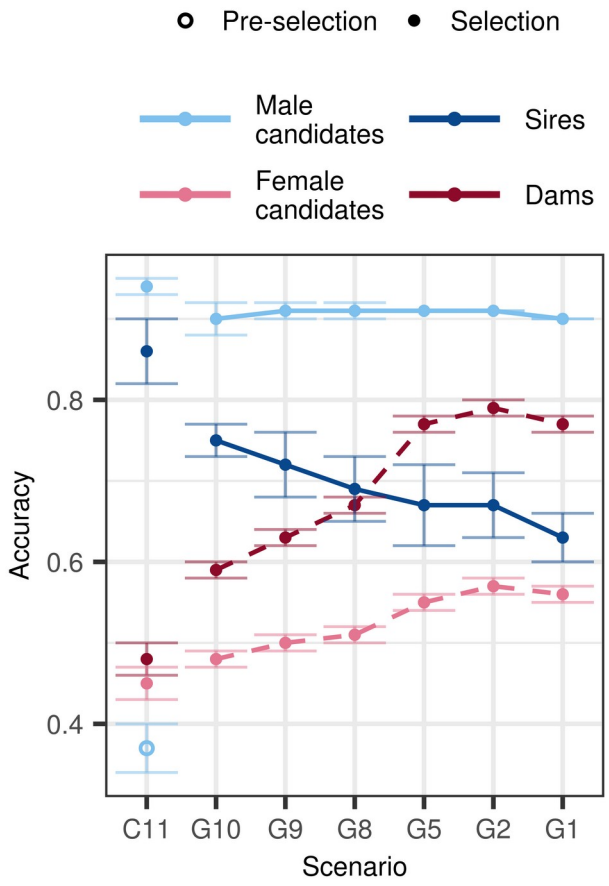
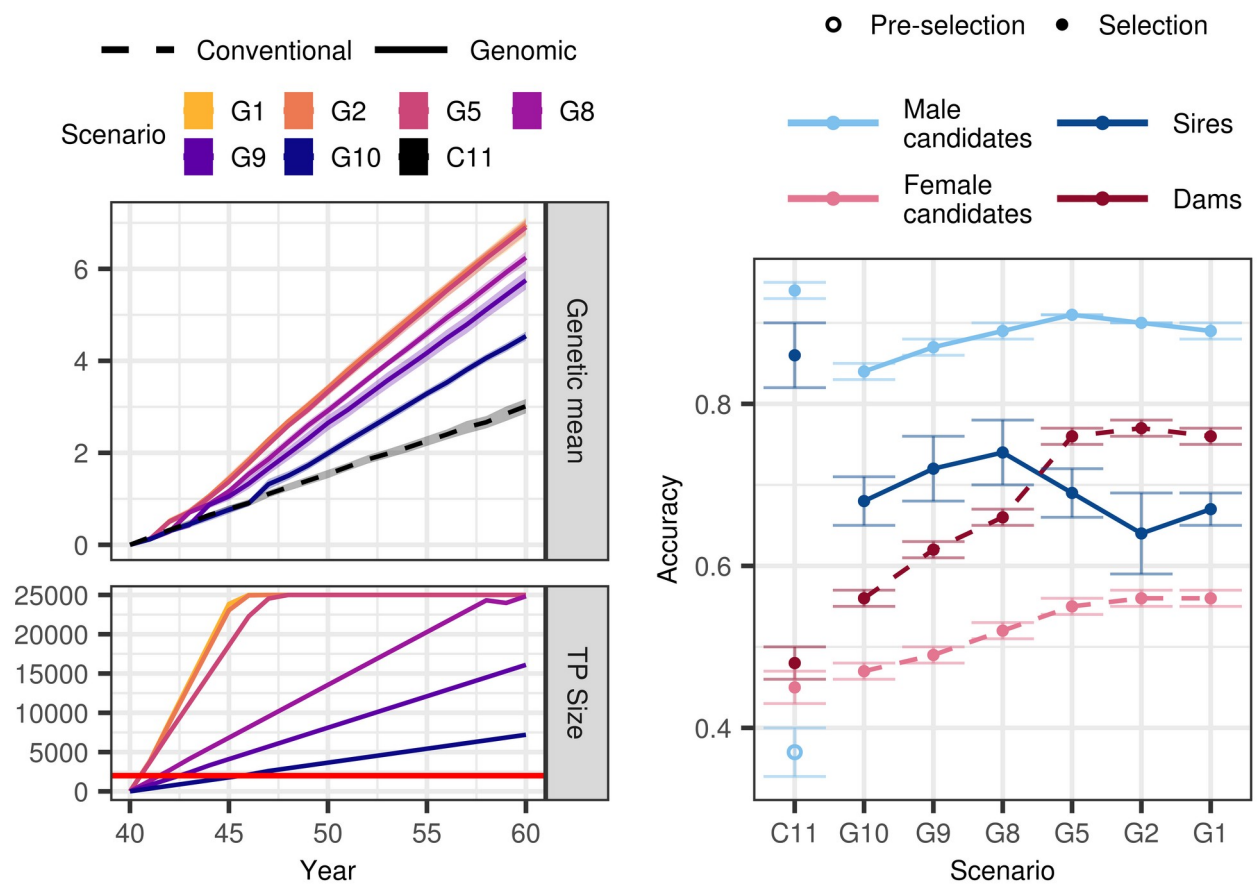


Figure 2 Accuracy by scenario with an initial training population and equal cost of phenotyping and genotyping.

The figure presents the means (lines) and 95% confidence intervals (error bars) across 10 replicates for the conventional (C) and genomic (G) scenarios with numbers indicating the number of phenotype records per lactation. Conventional selection implemented two-stage selection for males, hence we present the accuracy of pre-selection for progeny testing (empty point) and the accuracy of sire selection (solid point).



636

637 **Figure 3 Genetic gain, training population size, and accuracy by scenario without an initial**
638 **training population (TP) and equal cost of phenotyping and genotyping.**

639 The figure presents the means (lines or points) and 95% confidence intervals (polygons or error-
640 bars) across 10 replicates for the conventional (C) and genomic (G) scenarios with numbers
641 indicating the number of phenotype records per lactation. The red line marks the condition of
642 required 2,000 training animals to start genomic selection. Conventional selection implemented
643 two-stage selection for males, hence we present the accuracy of the pre-selection for progeny
644 testing (empty point) and the accuracy of sire selection (solid point).

645

Additional files

Additional file 1

Table S1 Accuracy of conventional and genomic selection with varying number of phenotypes and phenotyped animals.

NoPheno	NoDaughters	r_{sires}	r_{cows}	$r_{non\text{-}pheno}$	NoPhenoCows	NoPhenoTotal
Conventional selection, 100 sires						
Variable resources for phenotyping						
1	100	0.93	0.62	0.56	10,000	10,000
2	100	0.96	0.70	0.59	10,000	20,000
5	100	0.97	0.81	0.64	10,000	50,000
10	100	0.98	0.89	0.66	10,000	100,000
Fixed resources for phenotyping						
1	1000	0.99	0.63	0.59	100,000	100,000
2	500	0.99	0.71	0.61	50,000	100,000
5	200	0.99	0.82	0.64	20,000	100,000
10	100	0.98	0.89	0.66	10,000	100,000
Genomic selection						
Variable resources for phenotyping						
1	-	-	0.62	0.56	10,000	10,000
2	-	-	0.70	0.63	10,000	20,000
5	-	-	0.81	0.71	10,000	50,000
10	-	-	0.89	0.76	10,000	100,000
Fixed resources for phenotyping						
1	-	-	0.63	0.93	100,000	100,000
2	-	-	0.71	0.90	50,000	100,000
5	-	-	0.82	0.84	20,000	100,000
10	-	-	0.89	0.76	10,000	100,000

Description: NoRec = Number of phenotypic records per lactation, NoDaughters = number or daughters per sire, r_{sire} = accuracy for sires, r_{cows} = accuracy for cows, $r_{non\text{-}pheno}$ = accuracy for non-

phenotyped animals, NoPhenoCows = number of phenotyped cows, NoPhenoTotal = total number of phenotypes (number of phenotypes per lactation times the number of phenotyped cows).

Additional file 2

Table S2 Genetic gain by scenario, relative cost of phenotyping to genotyping, and availability of an initial training population.

		Relative cost of phenotyping (P) to genotyping (G)		
	Scenario	$P:$G = 1:2	$P:$G = 1:1	$P:$G = 2:1
	C11	$3.01_{0.22}^{a,A}$	$3.01_{0.22}^{a,A}$	$3.01_{0.22}^{a,A}$
With initial training population	G10	$5.43_{0.20}^{b,A}$	$5.41_{0.29}^{b,A}$	$6.50_{0.20}^{b,B}$
	G9	$5.58_{0.26}^{b,A}$	$6.30_{0.17}^{c,B}$	$7.02_{0.24}^{c,C}$
	G8	$6.35_{0.25}^{c,A}$	$6.62_{0.25}^{d,B}$	$7.02_{0.17}^{c,C}$
	G5	$6.78_{0.21}^{d,A}$	$7.07_{0.20}^{e,B}$	$\mathbf{7.26_{0.19}^{c,B}}$
	G2	$7.13_{0.29}^{e,A}$	$\mathbf{7.33_{0.26}^{e,A}}$	$\mathbf{7.28_{0.17}^{c,A}}$
	G1	$\mathbf{7.11_{0.16}^{e,A}}$	$\mathbf{7.27_{0.28}^{e,A}}$	$\mathbf{7.24_{0.22}^{c,A}}$
Without initial training population	G10	$3.93_{0.22}^{b,A}$	$4.54_{0.14}^{b,B}$	$5.61_{0.25}^{b,C}$
	G9	$4.64_{0.18}^{c,A}$	$5.75_{0.28}^{c,B}$	$6.52_{0.17}^{c,C}$
	G8	$5.61_{0.28}^{d,A}$	$6.24_{0.19}^{d\,B}$	$6.70_{0.25}^{cd,C}$
	G5	$6.43_{0.21}^{e,A}$	$6.90_{0.22}^{e,B}$	$\mathbf{7.05_{0.27}^{de,B}}$
	G2	$6.81_{0.28}^{f,A}$	$\mathbf{6.96_{0.17}^{e,A}}$	$\mathbf{7.00_{0.30}^{de,A}}$
	G1	$\mathbf{6.78_{0.29}^{f,A}}$	$\mathbf{6.92_{0.26}^{e,A}}$	$\mathbf{7.01_{0.23}^{e,A}}$

The table presents the means and standard deviations (subscript) across 10 replicates for the conventional (C) and genomic (G) scenarios, with numbers indicating the number of phenotype records per lactation. The scenarios in bold did not spend all the available resources. Lower-case letters denote statistically significant differences between scenarios within the same $P:$G and upper-case letters between different $P:$G within the same scenario.

Additional file 3

Table S3 Intensity of sire selection by scenario and relative cost of phenotyping to genotyping.

	Relative cost of phenotyping (P) to genotyping (G)		
Scenario	$P:$G = 1:2	$P:$G = 1:1	$P:$G = 2:1
C11	0.80	0.80	0.80
G10	1.32	1.71	2.02
G9	1.76	2.06	2.48
G8	1.99	2.27	2.52
G5	2.40	2.63	2.85
G2	2.63	2.86	3.11
G1	2.70	2.93	3.14

$P:$G = relative cost of phenotyping ($P) to genotyping ($G). The scenarios are named C/G for conventional/genomic with numbers indicating the number of phenotype records per lactation.

Additional file 4

Table S4 Selection accuracy by scenario, relative cost of phenotyping to genotyping ($P:$G), and the availability of an initial training population.

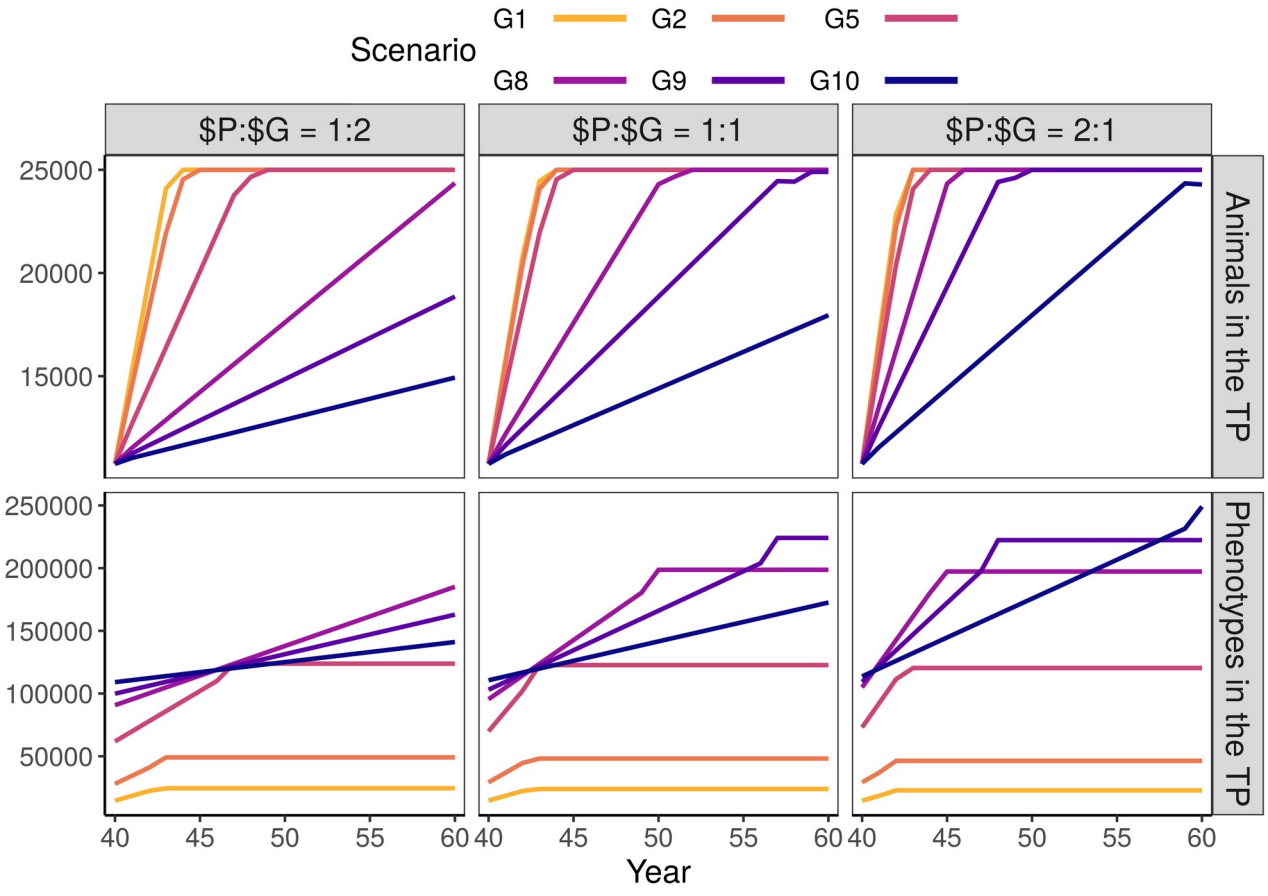
	With initial training population			Without initial training population		
Scenario	\$P:\$G=1:2	\$P:\$G=1:1	\$P:\$G=2:1	\$P:\$G=1:2	\$P:\$G=1:1	\$P:\$G=2:1
Male candidates						
C11, S1	$0.37_{0.04}^{a,A}$	$0.37_{0.04}^{a,A}$	$0.37_{0.04}^{a,A}$	$0.37_{0.04}^{a,A}$	$0.37_{0.04}^{a,A}$	$0.37_{0.04}^{a,A}$
C11, S2	$0.94_{0.01}^{b,A}$	$0.94_{0.01}^{b,A}$	$0.94_{0.01}^{b,A}$	$0.94_{0.01}^{b,A}$	$0.94_{0.01}^{b,A}$	$0.94_{0.01}^{b,A}$
G10	$0.89_{0.03}^{c,A}$	$0.90_{0.02}^{bc,AB}$	$0.91_{0.01}^{bc,B}$	$0.81_{0.03}^{b,A}$ *	$0.84_{0.01}^{b,B}$ *	$0.87_{0.01}^{b,C}$ *
G9	$0.90_{0.03}^{bc,A}$	$0.91_{0.02}^{bc,A}$	$0.91_{0.01}^{bc,A}$	$0.85_{0.02}^{c,A}$ *	$0.87_{0.01}^{bc,B}$ *	$0.90_{0.01}^{bc,C}$ *
G8	$0.91_{0.01}^{bc,A}$	$0.91_{0.01}^{bc,A}$	$0.91_{0.01}^{bc,A}$	$0.86_{0.01}^{cd,A}$ *	$0.89_{0.01}^{c,B}$ *	$0.90_{0.01}^{bc,B}$
G5	$0.91_{0.01}^{bc,A}$	$0.91_{0.00}^{bc,A}$	$0.91_{0.01}^{bc,A}$	$0.90_{0.01}^{d,A}$	$0.91_{0.01}^{c,A}$	$0.91_{0.01}^{c,A}$
G2	$0.91_{0.01}^{bc,A}$	$0.91_{0.00}^{bc,A}$	$0.90_{0.01}^{bc,A}$	$0.90_{0.01}^{d,A}$	$0.90_{0.01}^{c,A}$	$0.90_{0.01}^{bc,A}$
G1	$0.89_{0.01}^{c,A}$	$0.90_{0.01}^{c,A}$	$0.89_{0.01}^{c,A}$	$0.89_{0.01}^{cd,A}$	$0.89_{0.01}^{c,A}$	$0.89_{0.01}^{bc,A}$
Sires						
C11	$0.86_{0.05}^{a,A}$	$0.86_{0.05}^{a,A}$	$0.86_{0.05}^{a,A}$	$0.86_{0.05}^{a,A}$	$0.86_{0.05}^{a,A}$	$0.86_{0.05}^{a,A}$
G10	$0.75_{0.04}^{b,A}$	$0.75_{0.03}^{b,A}$	$0.73_{0.05}^{b,A}$	$0.67_{0.08}^{bc,A}$ *	$0.68_{0.05}^{cde,A}$ *	$0.67_{0.06}^{b,A}$ *
G9	$0.76_{0.04}^{b,A}$	$0.72_{0.06}^{bc,AB}$	$0.69_{0.05}^{c,A}$	$0.70_{0.05}^{b,A}$ *	$0.72_{0.05}^{bc,A}$	$0.71_{0.05}^{b,A}$
G8	$0.76_{0.03}^{b,A}$	$0.69_{0.05}^{cd,B}$	$0.68_{0.06}^{c,B}$	$0.71_{0.05}^{b,A}$ *	$0.74_{0.05}^{b,A}$ *	$0.70_{0.07}^{b,A}$
G5	$0.68_{0.07}^{c,A}$	$0.67_{0.08}^{de,A}$	$0.69_{0.04}^{c,A}$	$0.68_{0.05}^{bc,A}$	$0.69_{0.05}^{cd,A}$	$0.69_{0.03}^{b,A}$
G2	$0.67_{0.05}^{c,A}$	$0.67_{0.05}^{de,A}$	$0.67_{0.04}^{c,A}$	$0.65_{0.06}^{c,A}$	$0.64_{0.07}^{e,A}$	$0.69_{0.05}^{b,A}$
G1	$0.66_{0.06}^{c,A}$	$0.63_{0.05}^{e,A}$	$0.67_{0.04}^{c,A}$	$0.67_{0.04}^{bc,A}$	$0.67_{0.03}^{de,A}$	$0.69_{0.05}^{b,A}$
Female candidates						
C11	$0.45_{0.02}^{a,A}$	$0.45_{0.02}^{a,A}$	$0.45_{0.02}^{a,A}$	$0.45_{0.02}^{a,A}$	$0.45_{0.02}^{a,A}$	$0.45_{0.02}^{a,A}$
G10	$0.48_{0.01}^{ab,A}$	$0.48_{0.01}^{ab,A}$	$0.51_{0.01}^{b,B}$	$0.46_{0.02}^{ab,A}$ *	$0.47_{0.02}^{ab,AB}$	$0.49_{0.01}^{b,B}$ *
G9	$0.49_{0.02}^{b,A}$	$0.50_{0.01}^{b,B}$	$0.52_{0.01}^{b,C}$	$0.47_{0.02}^{ab,A}$ *	$0.49_{0.02}^{bc,B}$	$0.52_{0.01}^{bc,C}$
G8	$0.51_{0.01}^{b,A}$	$0.51_{0.01}^{b,A}$	$0.54_{0.01}^{bc,B}$	$0.49_{0.02}^{bc,A}$ *	$0.52_{0.01}^{cd,B}$	$0.53_{0.01}^{cd,C}$
G5	$0.51_{0.01}^{bc,A}$	$0.55_{0.01}^{c,B}$	$0.57_{0.01}^{c,C}$	$0.52_{0.01}^{cd,A}$	$0.55_{0.01}^{de,B}$	$0.57_{0.01}^{d,C}$
G2	$0.55_{0.01}^{cd,A}$	$0.57_{0.01}^{c,B}$	$0.57_{0.01}^{c,B}$	$0.55_{0.01}^{d,A}$	$0.56_{0.02}^{e,AB}$	$0.57_{0.01}^{d,B}$
G1	$0.56_{0.01}^{d,A}$	$0.56_{0.01}^{c,A}$	$0.56_{0.01}^{c,A}$	$0.55_{0.01}^{d,A}$	$0.56_{0.01}^{e,A}$	$0.56_{0.01}^{d,A}$
Cows						
C11	$0.48_{0.03}^{a,A}$	$0.48_{0.03}^{a,A}$	$0.48_{0.03}^{a,A}$	$0.48_{0.03}^{a,A}$	$0.48_{0.03}^{a,A}$	$0.48_{0.03}^{a,A}$
G10	$0.56_{0.02}^{b,A}$	$0.59_{0.02}^{b,B}$	$0.63_{0.01}^{b,C}$	$0.53_{0.01}^{b,A}$ *	$0.56_{0.01}^{b,B}$ *	$0.61_{0.01}^{b,C}$ *
G9	$0.59_{0.03}^{bc,A}$	$0.63_{0.02}^{c,B}$	$0.70_{0.01}^{c,C}$	$0.57_{0.02}^{bc,A}$ *	$0.62_{0.02}^{c,B}$	$0.68_{0.02}^{c,C}$ *
G8	$0.62_{0.02}^{c,A}$	$0.67_{0.02}^{c,B}$	$0.74_{0.02}^{d,C}$	$0.60_{0.02}^{c,A}$ *	$0.66_{0.01}^{d,B}$	$0.73_{0.02}^{d,C}$
G5	$0.70_{0.02}^{d,A}$	$0.77_{0.01}^{d,B}$	$0.79_{0.02}^{e,C}$	$0.69_{0.02}^{d,A}$	$0.76_{0.01}^{e,B}$	$0.78_{0.02}^{e,B}$
G2	$0.76_{0.02}^{e,A}$	$0.79_{0.02}^{d,B}$	$0.78_{0.01}^{e,AB}$	$0.76_{0.01}^{e,A}$	$0.77_{0.02}^{e,A}$ *	$0.77_{0.01}^{de,A}$
G1	$0.77_{0.02}^{e,A}$	$0.77_{0.02}^{d,A}$	$0.77_{0.01}^{de,A}$	$0.76_{0.01}^{e,A}$	$0.76_{0.02}^{e,A}$	$0.76_{0.02}^{de,A}$

The table presents the means and standard deviations (subscript) across 10 replicates for the conventional (C) and genomic (G) scenarios, with numbers indicating the number of phenotype records per lactation. Conventional selection implemented two-stage selection for males, hence we

present the accuracy of pre-selection for progeny testing (S1) and the accuracy sire selection (S2). In genomic scenarios the male candidates were genotyped and non-phenotyped. We also present the accuracy for sires currently used in artificial insemination (sires), for non-genotyped and non-phenotyped females (female candidates), and for all active phenotyped cows and bull dams (cows). Lower-case letters denote statistically significant differences between scenarios within the same $P:$G and upper-case letters between different $P:$G within the same scenario. Stars denote statistically significant difference between corresponding scenarios with and without an initial training population.

Additional file 5

Figure S1 The number of animals and repeated phenotypes in the training population.



The figure presents the results for three relative costs of phenotyping to genotyping ($P:$G). In our simulation, scenarios traded repeated phenotype records for genotypes. Hence, the scenarios with

688 the largest training population collected the least repeated records. These were also the scenarios

689 that achieved the highest genetic gain.