final report

Project code: B.GEN.0012

Prepared by: Julius van der Werf and Cecile Massault
                          University of New England

Date published: 8 August 2019

Tactical approach to phenotyping and genotyping

<Project code - Project Title, keep hyphen>

Abstract

Background
Phenotyping cost can be prohibitive when resources are limited to perform accurate genetic evaluation and can impact both traditional and genomic selection. A sensible approach is to use a two stage selection where some candidates are selected for phenotyping to reduce the number of phenotypes, hence to reduce the cost. There is potential to improve this approach by optimising which candidate should be selected for phenotyping.

Methods
A phenotyping strategy DIVERSITY is proposed which optimises through a differential evolution algorithm the best set of candidates to be phenotyped in a two stage two trait selection scheme through the balance of merit and diversity. The strategy was tested by simulation using three populations with various degrees of relatedness and genetic parameters in a two stage selection scenario.

Results
DIVERSITY performed slightly (6.86%) better than selection only based on estimated breeding value (EBV) at stage 1. The efficiency of the strategy was related to the number of families phenotyped, the correlation of index among relatives and the variance between families in EBV at the first stage of selection, which was influenced by the genetic parameters such as trait heritabilities, relatedness among selection candidates and the level of constraint on inbreeding in the breeding program.

Conclusions
The results of this study showed the advantages of a phenotyping strategy that considers merit as well as diversity. The correlation of breeding values between candidates from the same families and the number of families with at least one candidate phenotyped with truncation selection could be used as predictors of the merit of the DIVERSITY strategy. The implementation of such a strategy could be beneficial to farmers and breeders as it will reduce measurement cost while maintaining genetic gain and diversity.

Keywords
Genetic evaluation, phenotyping strategy, evolutionary algorithm

<Project code - Project Title, keep hyphen>

Table of contents

<Project code - Project Title, keep hyphen>

1   Background

Phenotyping is an essential requirement for genetic evaluation as it provides information to estimate breeding values (EBVs) of selection candidates in breeding programs and guides decisions on which animals should be selected as parents for future generations. The aim of most breeding programs is to maximize genetic gain while maintaining genetic diversity (i.e. controlling inbreeding). Genetic evaluations are often based on Best Linear Unbiased Prediction (BLUP – [1]) where information is given by phenotypes and pedigree relationships and this could be augmented by genomic information (e.g. SNPs, [2]) leading to genomic BLUP (GBLUP).

Collecting phenotypic records on animals is not always a cheap and a trivial operation. Henryon et al. [3] argued that, firstly, there are usually not enough resources to phenotype all animals for all economically important traits (especially the difficult and expensive traits such as disease resistance, feed intake or methane production) and, secondly, that there is a diminishing marginal return, meaning that there is often little additional gain when phenotyping a large number of candidates.

To reduce phenotyping cost, the number of phenotypes measured per trait could be reduced. Wade and James [4] explored deterministic ways to allocate resources in a two-stage selection scheme, where all candidates are phenotyped for an inexpensive trait at stage 1 and then only a proportion are being phenotyped for a more expensive/difficult to measure trait at stage 2. The problem is then to decide on the best subset of selection candidates to be phenotyped. The problem of choosing which candidates to phenotype can be extrapolated to the problem of which candidates to genotype in a genomic selection scenario as genotyping remains relatively expensive.

Wade and James [4] showed the relationship between measurement cost and selection accuracy, which determined the optimal proportion of males and females to be phenotyped. Efficiency of this approach was evaluated by Jopson et al. [5] in a practical breeding scheme for terminal sire sheep, showing best cost-benefit when 5 to 20% of males candidates were phenotyped. This work mainly focused on optimizing the proportion of animals to be selected for phenotyping at stage 2 ([4, 5]), but not at the genetic diversity among the selection candidates. In a two stage selection scheme, candidates will have prior information coming from previous genetic evaluation (EBVs or index) at stage 1. If candidates were to be chosen on merit at stage 1, i.e. based on their estimated breeding values, candidates from few families might be selected for phenotyping for stage 2. As animals with more information have more chance to be selected, this narrows down the choices that can be made at the final selection of parents after stage 2, possibly jeopardizing the potential genetic gain that can be made at stage 2 of selection if a certain level of diversity is to be maintained.

Using a fixed proportion of candidates to phenotype, Massault et al. [6] developed a phenotyping strategy where a set of candidates to be phenotyped was optimised using a genetic algorithm to maximise genetic gain and maintain diversity. The selection of candidates was based on the balance of average accuracy and diversity to simultaneously maximise both gain and diversity in single-trait and single-stage selection with no prior information (e.g. no known EBVs of parents or on other traits). In this scenario, phenotyping is directly related to selection accuracy and aiming for higher selection accuracy among selection candidates also resulted in phenotyping of candidates from more families, i.e. higher diversity.

In practice, previous information is often available and a two-stage selection scenario is often applied where candidates selected at stage 1 are typically chosen based on their EBV, which we will refer to as "index". We hypothesize that the distribution of indexes among candidates would be an important feature in deciding which candidates should be phenotyped. The distribution of indexes can be divided into two components: 1) the correlation of indexes among family members and 2) the

variance between families in mean index value. A large family mean variance and a large correlation among sibs will produce distinct clusters of candidates belonging to the same families and a strategy based on merit will likely phenotype candidates from the few best clusters and increase the relatedness among phenotyped candidates.

In this study, we develop and investigate a phenotyping strategy called DIVERSITY that optimises a set of candidates to phenotype at the second stage in a two-stage two-trait scenario of a breeding program. We propose to use a differential evolution algorithm (DE, [7]) that finds an optimal subset of candidates to be phenotyped at stage 2 based on a selection criterion that balances genetic gain and diversity. We investigate a number of possible predictors of the efficiency of DIVERSITY compared to a strategy where phenotyping is decided based on truncation selection on merit only.

2 Project objectives

- To develop and evaluate a framework for multi-trait phenotyping strategy by optimizing which animal to phenotype using differential evolution algorithm that optimizes two-stage selection processes.
- To develop and evaluate a genotyping strategy framework that optimizes investment in low and high density genotyping.
- To develop and evaluated a complete prototype that combines genotyping and phenotyping strategies
- To engage with breeders, breed societies and breeder's groups on optimal genotyping to fine tune case studies and establish practical feasibility of a tactical method to optimise a combined phenotyping and genotyping strategy
- To implement the prototype, either as stand-alone software or a web application, for tactical decision method and render it accessible through BREEDPLAN and SHEEP GENETICS platforms

3 Methodology

3.1 Materials and methods

3.1.1 DIVERSITY strategy

We used a selection criteria at stage 1 based on Meuwissen and Sonesson [8] where a high genetic gain is balanced with low relatedness among selected candidates:

$$Selection\ criterion = \ \boldsymbol{x'g} + \lambda \boldsymbol{x'Ax}\ \ \textbf{[Eq. 1]}$$

Where $\boldsymbol{g}$ is the vector of an index for merit for candidates at stage 1, $\boldsymbol{x}$ a vector indicating for each animal 1/$\boldsymbol{n}$ (number of phenotyped animals) when phenotyped or 0 otherwise and $\boldsymbol{A}$ is the numerator relationship matrix. The first term reflects the average breeding value of animals to be phenotyped while the second term reflects the average co-ancestry between these animals, which is weighted by the factor λ.

It was not possible to compare the selection criteria of every possible set of phenotyped individuals even in very small pedigree. If only 10 animals were being phenotyped among 100, there would be 1.73 X 1013 possible solutions to explore. To overcome this obstacle, we used a differential algorithm ([7]), e.g. such as applied in mate selection problems (Kinghorn [8]). The DE algorithm

<Project code - Project Title, keep hyphen>

finds the best, or close to best, set of individuals to be phenotyped efficiently. The algorithm is initiated by drawing randomly a set of 16 possible solutions, each solution being a subset of animals to be phenotyped, forming the 16 members of a first generation of solutions. A new generation of solutions was then created from existing solutions, using mutation and crossing over operations and a new solution would replace an existing solution if it was deemed better, where solutions were evaluated using an objective function as defined in equation [eq. 1]. Repeating this procedure for 10,000 generations evolved a set of solutions towards a maximum value for the objective function, i.e. each solution providing a set of individuals to phenotype with the highest value for the objective function [Eq. 1] for given a λ.

3.1.2 Population designs

We choose to simulate 10 paternal half-sib families with various degrees of relatedness and various family sizes. The number of animals were kept low to be able to demonstrate the efficiency of DIVERSITY strategy in a replicated stochastic simulation study.

The degree of relatedness was varied by changing the maximum number of possible male offspring selected from grand-sires. These offspring were subsequently used as sires of the 10 paternal half-sib families:

- POP1: the 10 sires descended from 10 different grand-sires (no relatedness between the sires of the half-sib families)
- POP2: up to 2 sires can share the same grand-sire (creating some relatedness among between the sires of the half-sib families)
- POP5: up to 5 sires can share the same sire (creating a large amount of relatedness between the sires of the half-sib families)

The simulation started with generating 10 grandsires, and these were given male progeny to become the sires of the selection candidates. EBVs of grand-sires were simulated in order to determine the size of their family. A larger number of offspring was assigned to sires with the highest EBVs to represent the fact that sires with high breeding values will reproduce more than sires with a lower breeding values and their offspring will also be more likely to be selected in practice. Table 1 shows the family distribution used for the simulation. The true breeding value (TBV) of the grandsires were drawn from a normal distribution $N(0, \sqrt{6_A^2})$, where 6_A^2 is the additive genetic variance. Ten sires were selected from their progeny according to the grandsires' EBV and the population design (POP1, POP2 and POP5).

The EBV of the grandsires was assumed to have a reliability r^2 = 0.5. The EBV of grand sire *i* was derived by adding an error to the true breeding values

$$EBV_i = r^2\, TBV_i + e_i \quad \text{[eq. 2]}$$

where the error (e_i) was simulated as follows:

$$e_i = r.\sqrt{(1 - r^2).VA_i} \quad \text{[eq. 3]}$$

To create the 10 half-sib families, the 10 sires were each mated to ~ 20 unrelated dams, each dam producing one offspring. This study focussed solely on male selection as candidates for the second stage trait measurement due to the higher selection intensity in males. Therefore, the number of candidates for a measurement in the second stage was 100.

3.1.3   Two-stage selection

To explore the efficiency of the DIVERSITY strategy, we designed a two-stage selection scenario where a genetic evaluation (multi trait BLUP) was performed after the phenotyping of a cheap trait for all animals (heritability trait 1 is h^2t_1= 0.1) and phenotyping the sires for the expensive trait (heritability trait 2 is h^2t_2= 0.3). Then only a proportion of selection candidates, i.e. the sires' progeny, were phenotyped for trait 2, following the results of the phenotyping strategy applied. A second genetic evaluation was then performed and optimum contribution selection ([10]) was used to select sires and dams for the next generation. A maximum rate of increase of inbreeding was fixed at 10% (coancestry x'Ax <0.2). The inbreeding rate has been set up so high as the number of families was small and relatedness could be high. This corresponded to selecting roughly 5% of sires for the next generation. With more animals, the rate of inbreeding can be reduced to a more appropriate level (e.g. 1%).

The two traits had genotypic variance of 10 and 30 respectively, a genetic correlation r_a of 0.1 and an environmental correlation of 0.25. The breeding objective was determined by the equation $w_1.A_1 + A_2$ *[eq. 4]*, where A_i is the breeding value for trait i, and the relative economic value w_1 was assumed to be 0.5 in the base scenario.

To determine if DIVERSITY was valuable, we compared three phenotyping strategies:
> ALL: all males offspring were phenotyped for trait 2
> MERIT: 10% of males with highest EBV at stage 1 were phenotyped for trait 2
> DIVERSITY: 10% males selected using [Eq. 1] that balances merit and diversity

3.1.4   Relatedness, λ at stage 1 and genetic parameters

The value of λ in Eq. *(1)* controls the weight given to diversity in choosing which male candidates to phenotype at stage 2. A value of λ close to zero will result into a phenotyping pattern equal to MERIT (with no weight given on diversity), while a high negative value (e.g. -500), the phenotyping pattern will tend to choose candidates from more diverse families. To find the optimal λ, 500 replicates were simulated for each of various values of λ (-1, -2, -3, -4, -5, -10, -50, -100, -500) at stage 1 and for the 3 different population structures POP1, POP2 and POP5 in the two stage selection scenario described in the section above. The genetic parameters used were those of scenario 1 shown in Table 2.

To investigate the impact of genetic parameters on the efficiency of DIVERSITY, h^2t_1, h^2t_2, r_a and w_1 were varied from the base values (Table 2) as the amount of information at stage 1 can impact on the efficiency of DIVERSITY through correlation of indexes among relatives and family index variance. The two-stage selection was replicated 500 times for each scenario using POP5 population with a λ value of -50.

3.1.5   Measurements of efficiency of DIVERSITY

To assess the efficiency of DIVERSITY, we measured the difference in genetic gain at stage 2 between DIVERSITY and MERIT. The test was the combination of the percentage of extra genetic gain for the same level of inbreeding of DIVERSITY over MERIT (called Difference):

$$Difference \ = \frac{\Delta G(DIVERSITY)-\Delta G(MERIT)}{\Delta G(MERIT)} \times 100 \ \ \textit{[eq. 6]}$$

<Project code - Project Title, keep hyphen>

Additionally, the 95 percent confidence interval (95CI) gave a range of the Difference. The upper and lower bound correspond to ±1.96 Standard Deviations of outcomes among replicates.

3.1.6   Prediction of efficiency of DIVERSITY

Three potential predictors were calculated to evaluate the possibility of predicting the success of DIVERSITY: i) the number of different families phenotyped by MERIT (Nfamtop10%); ii) the family index variance (EBVvar) and iii) the correlation of index among sibs candidates (ICC). The key parameters driving these are i) heritability of Trait 1 and Trait 2 (h^2t_1 and h^2t_2); ii) genetic correlation between the traits (r_g) and iii) relative economic weight of the traits in the breeding objective (w1). The genetic parameters h^2t_1, h^2t_2 and r_g) and economic weight of breeding values in the breeding objective (w1) define an important single predictor of each case, which is the accuracy of the index at stage 1 (ACC). We will therefore use only the accuracy as a predictor of the outcomes in our analysis. The information used at stage 2 is the same as that at stage 1 plus the additional information given by selectively phenotyping potential selection candidates for Trait 2. The difference in accuracy at stages 1 and 2 will affect the added value of a DIVERSITY strategy at stage 1 selection.

The family index variance ICC was calculated as the variance of the mean index value of male candidates for each family, while the correlation of index among sibs was calculated as follows [11]:

$$t = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2} \ [eq.\ 6]$$

With σ_B^2 the between half sib family variation of EBVs and σ_W^2 the within family variation of EBVs.

To assess the best predictor(s) that could potentially predict the additional benefit of the DIVERSITY strategy (Difference used explanatory variable), we start with a full model with all predictors (Nfamtop10%, EBVvar, ICC and ACC) to model the Difference between the DIVERSITY strategy and the MERIT strategy:

$$Difference = Nfamtop10\% + EBVvar + ICC + ACC\ [eq.\ 7] + e$$

We used backward stepwise model selection to obtain the best prediction of the Difference. This avoid examining every single combination of predictors while testing for most likely predictors. The reliability and p-value of reliability of each model was then calculated to assess how much each model explains the success of DIVERSITY.

We also tested the DIVERSITY approach on a real dataset of Angus cattle (courtesy of P. Parnell, Angus Australia). We used the 500 best young bulls of the 2013 cohort with known sire and grand-sires. The index (Angus Breeding Index) was provided as well as the accuracy for each of the 500 bulls and their sires, paternal and maternal grand-sires. The average accuracy of indexes was 0.36. We compared the DIVERSITY and the MERIT selection strategies to select xxx candidates to be phenotyped for an expensive second trait.

4   Results

4.1   Relatedness and optimal values for λ

Figure 1 shows the Difference between MERIT and DIVERSITY in percentage (as per Eq. [5]) for a range of λ values from -1 to -500 for the three population structures POP1 (unrelated sires), POP2 (some relatedness among sires) and POP5 (a large amount of relatedness among sires). The difference observed with λ = -1 is small. As the values of λ became more negative (and the emphasis on diversity increased), the difference between the 2 strategies generally increased, with the largest difference around 6% at a lambda value around -10 (POP1) and -50 (POP2 and POP5). From that point, the difference decreased as the values of λ became further negative for all 3 population structures. The lower bound of the 95% confidence interval showed that DIVERSITY was significantly more efficient than MERIT for all λ values and all 3 population structures. Therefore, DIVERSITY selection was more efficient than MERIT selection, but the optimal efficiency is achieved when the weight on diversity was chosen appropriately (i. e. there is an optimal λ).

 The advantage of DIVERSITY selection over MERIT selection was not evidently different between the three population structures. However, the population where sires were more related (POP5 and POP2) required a more negative value for λ to achieve the higher difference between the selection methods. Hence, with more related selection candidates, more emphasis was needed on diversity when selecting candidates for further phenotyping.

4.2   Genetic parameters

The base scenario (scenario 1, Table 2) used here is a POP5 scenario (high relatedness) with h^2t_1 = 0.1, h^2t_2 = 0.3, r_a = 0.1, w_1 = 0.5 and using a λ value of -50, which was found optimal in this case. This value has also been used for the other 7 scenarios. (see previous results). The Difference and the lower confidence interval are reported in Table 3.

The base scenario resulted in a difference between the 2 strategies of 6.03%, with a lower bound of 95CI of 5.2%, showing that DIVERSITY was more efficient than MERIT. Increasing the heritability of TRAIT 1 resulted in a decrease of the difference to -1.60% with h^2t_1 = 0.3 (scenario 2) and to -3.21% with h^2t_1 = 0.5 (scenario 3). With a high heritability for TRAIT1, MERIT is more efficient than DIVERSITY. Decreasing the heritability of trait 2 to 0.1 also resulted in a negative difference between DIVERSITY and MERIT (-4.04% - scenario 4), meaning that there is no benefit in using DIVERSITY. An increased heritability of trait 2 to 0.5 gave a similar result than when heritability was 0.3 (6.86% - scenario 5). As the genetic correlation between the traits was increased to 0.3 and to 0.5, the difference between the 2 strategies decreased to 3.35% and 1.48% respectively (scenario 6 and 7 respectively). Finally, an increase of economic value of trait 2 (by reducing w1 to 0.25 – scenario 8) resulted in a similar difference than for the base scenario (6.14%). More economic weighting on trait 2resuts in a lower index accuracy at stage 1 and more benefit in using a DIVERSITY strategy. Scenarios 1,5, 6, 7 and 8 all have positive values for CI95min showing a significant benefit for the DIVERSITY strategy over MERIT.

4.3   Measure of efficiency

Figure 2 shows the average Difference across the 500 replicates and percentage of success (i.e. percentage of times that DIVERSITY gave more gain than MERIT) as a function of the number of families phenotyped by MERIT and the correlation of index among half sibs candidates (ICC). We also

explored the index family variance (EBVvar), but as it was highly correlated with ICC (0.83). Results are plotted for the 8 different scenarios that were investigated, i.e. across a range of parameter values. There was a clear relationship between the number of families that would be phenotyped under a MERIT scenario (i.e. the number of families represented in the top10% at stage 1), and the superiority of the DIVERSITY strategy, with a larger difference as the number of families selected under MERIT decreased. The difference also increased as the ICC increased.

Table 3 reports ICC, ACC, EBVvar and the Nfamtop10% for the 8 scenarios with genetic parameters varied. A substantial Difference between MERIT and DIVERSITY was found for Scenarios 1, 5, 6, 7 and 8. each having a high ICC and low NFamtop10%. The EBVvar should be also higher in the scenario where the DIVERSITY performs better. A large family variance index indicates large differences between families and more similarity of individuals from the same family(high ICC), a case where DIVERSITY can make better phenotyping decision than MERIT to maintain diversity. EBVvar and ICC depend on the heritabilities of the traits, the genetic correlation and the economic weights of the traits in the breeding objective For example, a low heritability for trait 1 will result in EBVs being very similar between individuals from different families because a lot of family info is used to estimate it. A large ICC indicates that candidates have highly correlated indexes and therefore individuals of the same family are more likely to be selected when ranking on EBV. The case of high ICC and high family index variance is illustrated by an hypothetical population in Figure 3.A where candidates from same families cannot be differentiated from each other, while the family index are very different from one family to the other. Therefore, if the selection was performed by MERIT, only individuals from the same family would be selected. In the other hand, if EBVvar is small (i.e. 0.30 scenario 4), the difference between the family is small and if ICC is low as well, the candidates within a family would have very different indexes and therefore selecting by MERIT would often result into the selection of candidates from various family (see Figure 3.B), achieving high diversity among selected candidates.

4.4   Efficiency of DIVERSITY

Table 4. shows the coefficients, reliability and p-value associated with predicting the Difference between DIVERSITY and MERIT scenarios for different subsets of predictors. The full model yielded a reliability of 0.95 with a p-value of 0.008. When ICC was removed the reduced model gave a similar r2 (0.96) and a smaller p-value (0.001). The reduced model with 2 predictors (removing also EBVvar) had the same reliability as the reduced model with 3 predictors and even smaller p-value (0.0001). The reduced models with one predictor (either with nfamtop10% or ACC) had considerably lower predictive value (r2 of 0.78 and 0.23 respectively and large p-values). The reduced model with 2 predictors (accuracy and the number of family phenotyped with MERIT) was therefore the most accurate in predicting Difference and can be used to determine whether DIVERSITY has additional benefit over MERIT.

4.5   Analysis of real data

In a real dataset from Angus Australia We used the 500 best young bulls born in 2013 and selected xxx of these to be measured for a second trait.
The accuracy of index at stage 1 was 0.36. If MERIT is used, 14 families out of the 73 paternal half-sibs families will have at least one member phenotyped, or 19% of the families would have phenotypic information for the additional trait. To adjust for the prediction equation [eq. 5] which is based on 10 families, we used a value of   1.9 for Nfamtop10%. The correlation between indexes of sibs (ICC) was 0.12.

<Project code - Project Title, keep hyphen>

Based on these parameters and Eq. 5 (using ACC and nfamtop10%) the predicted Difference between DIVERSIY and MERIT would be only 0.372%.

This result indicates that a strategy such as DIVERSITY would not give additional gain in this particular case. The number of family being phenotyped using the 10% top male being large (14 family phenotyped) by MERIT, this prediction was expected. When using DIVERSITY with λ =1, 5 and 50, the number of family phenotyped increases to 18, 20 and 21 (standardized to 2.5, 2.7 and 2.8). The correlations between phenotyped individuals with MERIT and DIVERSITY (λ=1) was 0.82, decreasing to 0.77 with λ=5 and λ = 0.62, showing that MERIT captures a lot of diversity without the necessity of a complex strategy. Maintaining diversity does not increase the number of families phenotyped by much and it was expected, as the prediction equation revealed, that there is no additional benefit to use DIVERSITY in this case.

5 Discussion

5.1 Overview

This study proposed a phenotyping strategy based on the balance between genetic gain and diversity at stage 1, so that the selected candidates for the next generation also exhibit the best genetic gain when undertaking optimal contribution selection after stage 2. We showed that the success of DIVERSITY strategy applied at stage 1 selection depends on the existing relatedness among selection candidates, the intra class correlation among their indexes, the accuracy of index at stage 1 and the number of families phenotyped if using MERIT strategy. Note that the efficiency of DIVERSITY selection also depended on choosing the right emphasis on genetic diversity at stage 1 selection.

5.2 Additional benefit of selecting on diversity at stage 1

While Wade and James [4] proposed a deterministic method to phenotype an optimal proportion of selection candidates, our phenotyping strategy relied on actual pedigree to determine which individuals to phenotype given the optimal proportion. The DIVERSITY strategy can be viewed as an extension of Wade and James' [4] method, where in a first step the optimal proportion is determined and subsequently which individuals should be part of that set. The algorithm to decide which proportion to phenotype can also be added in the selection criteria, creating a more complex but more complete strategy. This would require an assessment of the cost of additional phenotyping and the benefit of genetic gain and genetic diversity.

It is difficult to compare DIVERSITY strategy to others as they are usually deterministic ([4,5]) or trials of several parameters to find an optimal solution (Massault et al. [6]). The DIVERSITY strategy is a novel strategy that is capable of finding an optimal subset of individual to phenotype for an expensive/hard to measure trait. The evolutionary algorithm (EA) is a very powerful approach in finding the best subset among many possible solution in nonlinear systems. The methodology has been applied to mate selection problems (Kinghorn reference) and Akdemir et al.[12] used it to design training populations in a genomic selection scheme. The DIVERSITY strategy can also be adapted to the training population problem using genomic information as a phenotypic trait. It can then also be extended to a phenotyping and genotyping optimising strategy.

We showed that DIVERSITY does not add value for all population structures. The method will be more beneficial in population structures where inbreeding is more likely (e.g. some fish population with small effective population size but a large number of progeny- [13,14]), whereas the benefit maybe smaller in large cattle populations where many ancestors contribute to the current gene pool, as in the real data used in this study. It is important to have a good understanding of the population structure before applying a phenotyping strategy. Knowing ICC and ACC will inform on the usefulness of DIVERSITY strategy as a selection criterion (Table 4). We used extreme values for the genetic parameters and small populations to demonstrate the principle of a DIVERSITY startgey strategy. Therefore, the results showed here are for illustration purposes, and the prediction of the extra benefit based on some key parameters maybe not be easily extendable to real data, depending on the distribution of these key parameters. A wider exploration of the parameters space might be needed to obtain a better idea of the benefit of DIVERSITY in practice. In our real data exploration, the predicted benefit seemed to be small and this was more or less illustrated by some empirical observations about the number of tested families with MERIT, and the similarity of the subsets selected to be phenotyped at stage 2.

5.3 Impact of the information propagation through relatedness

In the current settings of the study, only male candidates are chosen for phenotyping. However, each male candidate phenotyped brings additional information on their relatives (their half sibling males and females). Moreover, sires of each family are also phenotyped and give additional information to all male and female selection candidates. Therefore, in a case where MERIT would only phenotype male candidates from 1 or 2 families, additional information will be transferred to their female siblings, increasing the accuracy of the index of these individuals within the tested families, increasing the likelihood of co-selection of relatives and increase of inbreeding. Therefore, the DIVERSITY strategy is useful for both male and female. The relationship between selecting for diversity and increasing the index accuracy over a wider range of selection candidates was earlier illustrated by Massault et al., 2013.

5.4 Finding the optimal λ

While the results showed that extreme values for λ are very unlikely to be optimal, it was necessary to use exhaustive search to determine the optimal λ by using a large number of simulations with 11 values of λ. Results in Figure 1 show that the curves are concave with a maximum value.
A potential strategy to optimise the value for λ could be a three point parabola approach, where genetic gain at stage 2 can be calculated for three different values of λ (e.g. 0, 10 and 100). The equation [$y = ax^2 + bx + c$] can then be solved and the maximum $\frac{-b}{2a}$ can be deduced, which would correspond to an optimal value for λ. Iterations can be further added to obtain a more precise value for λ.

5.5 Prediction

The paper describes a criterion that can predict whether the DIVERSITY strategy is beneficial prior to run the algorithm. The best equation prediction [Eq. 7] comprised only 2 predictors ACC and nfamtop10% with a training r^2 of 0.96. Prediction accuracy did not increase when adding ICC and EBVvar. This is due to the high correlation between predictors: EBVvar and ICC are highly positively correlated (0.83), and nfamtop10% and ICC are highly negatively correlated (-0.998), and therefore EVBvar and ICC are also highly negatively correlated. The correlated terms EBVvar and ICC can

therefore be removed to simplify the prediction equation. It is advantageous to only keep ACC, a value that will be very likely available in practice prior the decision of applying a phenotyping strategy, and nfamtop10%, a number very easy to obtain. ICC and EBVvar require slightly more calculations.

5.6   Long term efficiency of DIVERSITY

This paper has demonstrated that DIVERSITY strategy could be more efficient than a strategy solely based on merit in the next generation, in some cases, particularly when ICC and EBVvar are both high where EBVs of relative are similar while EBVs of families are very different. This manuscript was mainly intended to describe the method and explore the paramters that make the DIVERSITY approach is more efficient   using simple cases with small populations and favourable parameters (i.e. low initial accuracy, high ICC). However, it is important to know whether the benefits also holds in long term selection.   In an extension of this study we compared the for 3 breeding programs (DIVERSITY, MERIT and all) over a 10 year time horizon Each year, selection candidates were males aged from 2 to 4 years old. The mating design was identical to the one used in this study with high merit sires producing more offspring. Results showed that a significant difference between DIVERSITY and MERIT could be observed already at generation 2, and hat difference increased over generation almost linearly with time (Supp. Figure).   Therefore, DIVERSITY could be a very useful long term strategy as well.

6   Conclusions/recommendations

This study demonstrated the usefulness of a phenotyping strategy that balances merit and diversity in a two-stage selection scheme through simulation of various scenarios. Through the exploration of a number of measurements, we found that the efficiency depended on relatedness between individuals, optimal λ and genetic parameters such as heritabilities of traits and genetic correlation between the traits, which in turn influenced the distribution of indexes, ICC and accuracy.

7   Key messages

- Phenotyping strategies should balance both merit and diversity
- DIVERSITY could be a very useful long term strategy
- The implementation of such a strategy could be beneficial to farmers and breeders as it will reduce measurement cost while maintaining genetic gain and diversity.
- There is potential benefit in a framework for multi-trait phenotyping strategy by optimizing which animal to phenotype using differential evolution algorithm that optimizes two-stage selection processes.

8 Bibliography

8.1 Heading

8.1.1 Sub heading

1. Henderson CR: **Best linear unbiased estimation and prediction under a selection model**. *Biometrics* 1975, **31**:423-447

2. Meuwissen THE, Hayes BJ, Goddard ME: **Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps**. *Genetics* 2001, **157**:1819-1829

3. Henryon M, Berg P, Sorenson AC: (2014) **Animal-breeding plans designed to maximise long-term genetic gains**. *Livestock Science* 2014, **166**:38-47

4. Wade CM, James JW: **Optimum allocation of resources considering two sexes and selection in two stages.** *Genet. Sel. Evol.* 1996, **28**: 3-21

5. Jopson NB, Amer PR, Mcewan JC: **Comparison of two-stage selection breeding programmes for terminal sire sheep**. *Proceedings of the New Zealand Society of Animal Production* 2004, **64**:212-216

6. Massault C, Kinghorn B, van der Werf JHJ: **Managing phenotyping cost**. *Proc. Assoc. Advmt. Anim. Genet.* 2013, **20**: 416

7. Storn R, Price K: **Differential evolution- a simple and efficient heuristic for global optimization over continuous spaces.** *Journal of global optimization* 1997, **11:**341-359

8. Kinghorn BP: **An algorithm for efficient constrained mate selection.** *Genet. Sel. Evol.* 2011, **43**: 4

9. Ollivier L: **Mass selection in livestock using limited testing facilities.** *Genet. Sel. Evol.* 1990, **22**: 109-117

10. Meuwissen THE, Sonesson AK: **Maximising the response to selection with apredefined rate of inbreeding: overlapping generations**. *J. Anim. Sci*. 1998, **76**: 2575-2583

11. Falconer DS, Mackay TF: **Introduction to quantitative genetics**. 4th ed. England, Pearson

<Project code - Project Title, keep hyphen>

12. Akdemir D, Sanchez JI, Jannick JL: **Optimization of genomic selection training populations with a genetic algorithm**. *Genet. Sel. Evol*. 2015, **47**:38

13. Brown, C.R., Wooliams, J.A., McAndrew, B.J. **Factors influencing effective population size in commercial population of gilthead seabream *Sparus Aurata***. *Aquaculture*, 2005, **247**:219-225

14. Fessahaye, Y., El-Bialy, Z., Rezk, M.A., Croojmans, R., Bovenhuis, H., Komen, H. **Mating systems and male reproductive success in Nile tilapia (*Oreochromis nilticus*) in breeding hapas: a microsatellite analysis.** *Aquaculture*, 2006, **256**:148-158

9 Appendix

9.1 Tables and figures

9.1.1 Tables

Table 1. Distribution of progeny number per sire depending on their EBV ranking.

# families	1	2	1	6
size	20-25	11-20	3-10	2
EBV rank	1	2-3	4	5-10

Table 2. Genetic parameters used for the 8 scenarios explored in this study

Scenario	h^2t_1	h^2t_2	r_a	w_1
1	0.1	0.3	0.1	0.5
2	0.3	0.3	0.1	0.5
3	0.5	0.3	0.1	0.5
4	0.1	0.1	0.1	0.5
5	0.1	0.5	0.1	0.5
6	0.1	0.3	0.3	0.5
7	0.1	0.3	0.5	0.5
8	0.1	0.3	0.1	0.25

h^2t_1: heritability of Trait 1 r_a: genetic correlation between the 2 traits

<Project code - Project Title, keep hyphen>

h^2t_2: heritability of Trait 2 w_1: weight of Trait 1 in the breeding objective (weight of Trait 2 always equals 1

Table 3. Observed Difference (%) between DIVERSITY and MERIT, lower confidence interval of difference (CI95min) and value of various predictive parameters with λ = -50.

Scenario	Difference (%)	CI95min	EBVvar	Nfamtop10%	icc	acc
1	6.03	5.19	2.02	2.1	0.94	0.61
2	-1.60	-2.17	2.43	3.3	0.53	0.61
3	-3.21	-3.56	3.28	4.1	0.26	0.63
4	-4.04	-4.80	0.30	2.9	0.73	0.40
5	6.86	6.06	4.63	1.8	0.97	0.66
6	3.35	2.55	2.12	2.3	0.87	0.61
7	1.48	0.74	2.11	2.6	0.79	0.62
8	6.14	5.24	2.11	1.9	0.97	0.64

sc: scenario number (see table 2)   Nfam10%top: Number of families with at least one individual phenotyped by MERIT

icc: correlation of index among relatives

acc: accuracy of index at stage 1

EBVvaer: variance of the mean index of families

<Project code - Project Title, keep hyphen>

Table 4. Full and reduced models to predict the additional benefit of DIVERSITY over MERIT

Model Description	Model	R^2	p-value
Full model (4 predictors)	Difference = 12.4 -6.4 nfamtop10% + 21.4 ACC – 6.6 ICC -0.9 EBVvar	0.95	0.008
Reduced model 1 (3 predictors)	Difference = 0.28 – 4.22 nfamtop10% + 23.8 ACC – 0.8 EBVvar	0.96	0.001
Reduced model 2 (2 predictors)	Difference = 1.3 – 4.6 nfamtop10% + 21.7 ACC	0.96	0.0001
Reduced model 3 (1 predictor)	Difference = 15.5 – 5.1 nfamtop10%	0.78	0.22
Reduced model 4 (1 predictor)	Difference = -16.3 + 30.9 ACC	0.23	0.13

Difference = Difference in % between MERIT and DIVERSITY
Nfamtop10% = the number of family with at least on individual phenotyped by MERIT
ACC = Accuracy at stage 1
ICC = The correlation of index among sibs candidates
EBVvar = The family index variance

9.1.2 Figures

Figure 1. Difference between MERIT and DIVERSITY. The Difference between the 2 phenotyping strategies is plotted for with different values of λ and for different level of relatedness within the population (POP1, POP2 and POP5). The darker lines with dots represents the average difference for each λ. The lighter lines represent the lower bound of 95% confidence interval when below the average and the upper bound of 95% confidence interval when above.

Figure 2. Plot of the two predictors from the best model as a function of the percentage of success of DIVERSITY and the difference. Figures 2.A and 2.B represent the percentage of success and Difference as a function of number of families phenotyped by MERIT (Nfamtop10%), Figure 2.C and 2.D represent the percentage of success and the Difference as a function the correlation of index among relatives (ICC) lustered by nfamtop10% with λ = -50.

Figure 3. Distribution of index of an hypothetical population of 50 individuals with low/high correlation of index among relatives and low/high variance of family index (EBVvar). Figure 3.A represent the distribution of EBV of candidates with high ICC and high variance of family index, Figure 3.B represent the distribution of EBV of candidates with low ICC and low variance of family index.

<Project code - Project Title, keep hyphen>

Supplementary Figure. Long term genetic gain using ALL, MERIT and DIVERSITY phenotyping strategies. The genetic gain of ALL (blue), MERIT (green) and DIVERSITY (red) is plotted as a function of the generation. DIVERSITY is roughly 30% better than MERIT after 10 years.
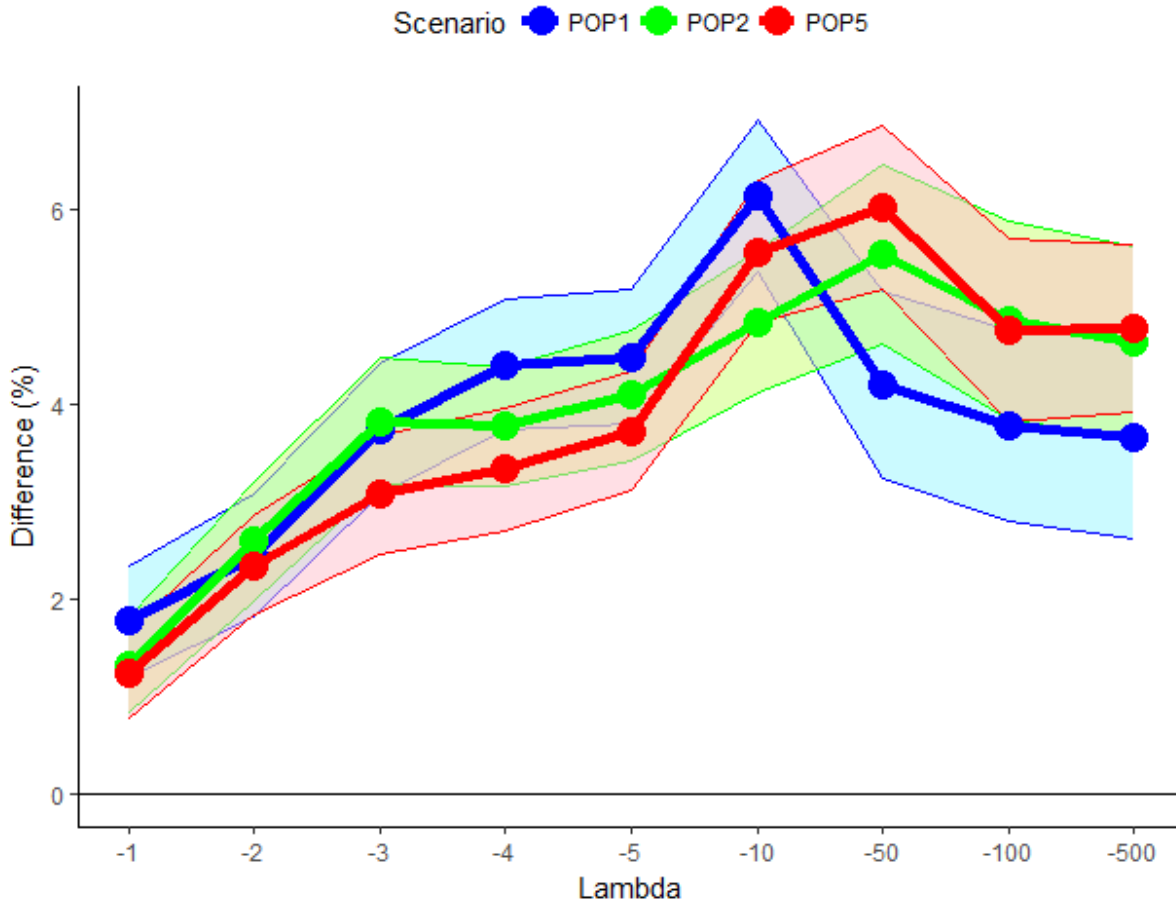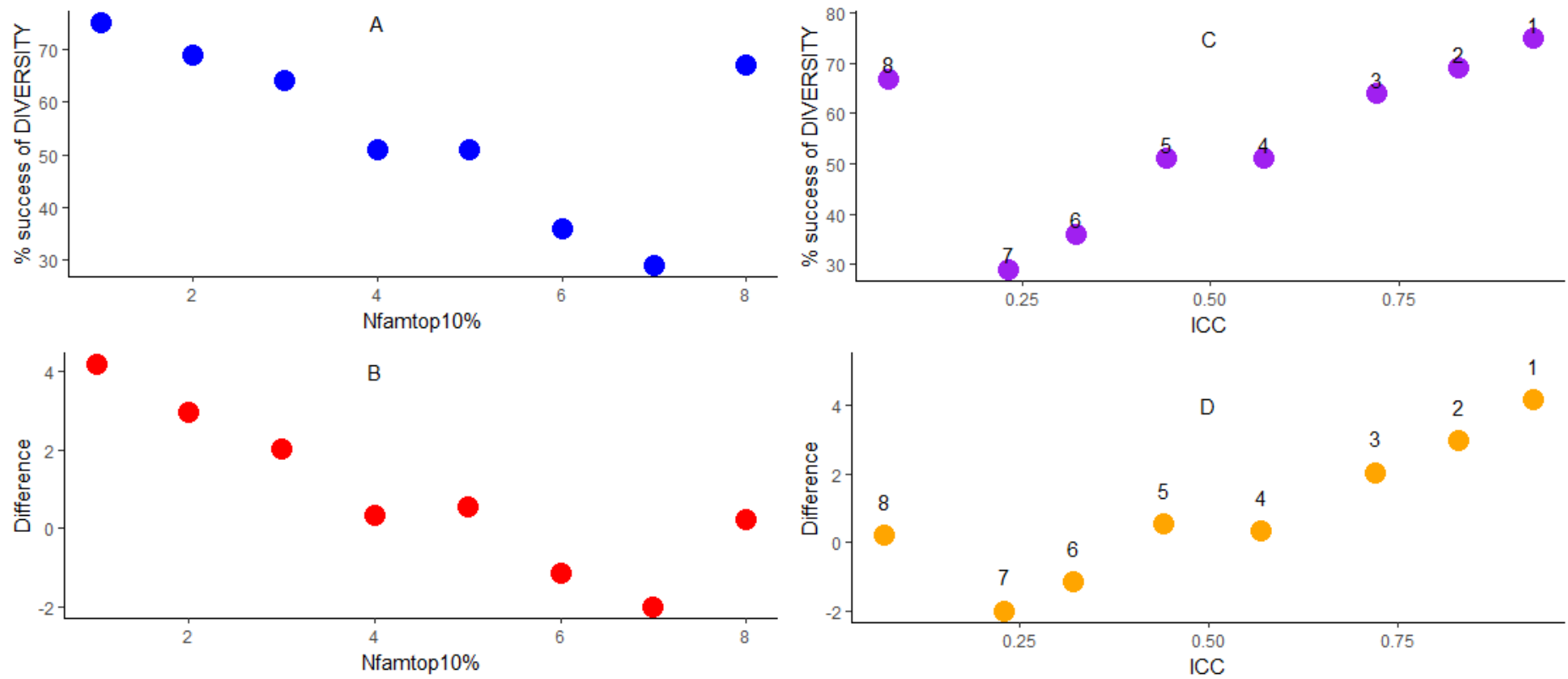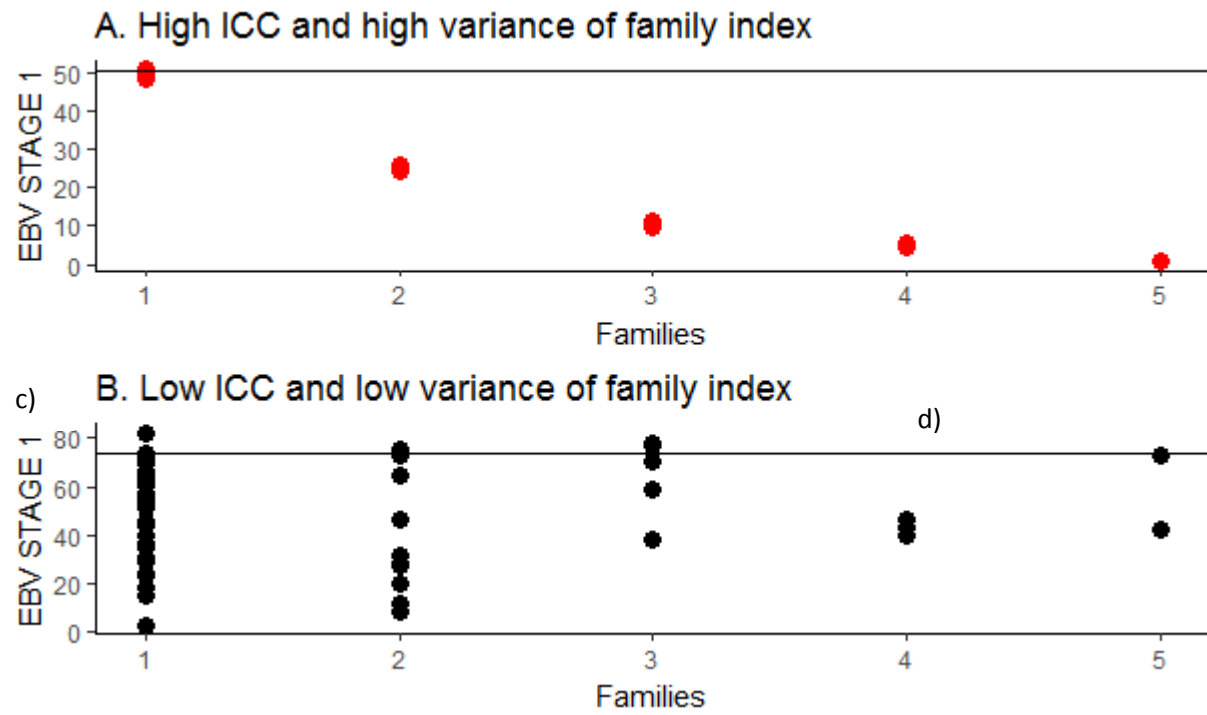
Figure 1.

Figure 2.

<Project code - Project Title, keep hyphen>

Figure 3.

A. High ICC and high variance of family index

B. Low ICC and low variance of family index

Figure 4.