EFFICIENCY OF A TACTICAL PHENOTYPING STRATEGY FOR MULTI-STAGE SELECTION

C Massault and J.H.J van der Werf

School of Environmental and Rural Science, University of New England, Armidale, NSW 2351

SUMMARY

Accurate genetic evaluation relies on measurements, which can be difficult to achieve for some economical important traits (hard and /or costly to measure). We developed a strategy that can select an optimised subset of animals to phenotype based on pedigree relationship, prior information (previously phenotyped animals) and diversity to maximise genetic gain under inbreeding and cost constraints.

We simulated a two-stage two-trait selection scenario for a small population of 10 paternal half-sib families of size 10 (pilot study). One trait was phenotyped for all animals (parents and selection candidates) and the second trait was phenotyped on only a selected set of 20 selection candidates based on a prior decision on phenotyping (stage 1). Phenotyping decisions were made either based on maximizing diversity of the set chosen to be phenotyped (DIVERSITY) or simply based on breeding values at stage 1 (MERIT). After phenotyping, the second stage selection of animals as parents for the next generation was based on optimum contributions. The DIVERSITY strategy was most useful when there was limited prior information about the Mendelian sampling term of predicted breeding value. When parents of selection candidates have not been phenotyped, DIVERSITY does not provide any advantage over truncation selection (MERIT). However, when sires or both parents have been previously phenotyped, DIVERSITY resulted in higher genetic gain for similar level of inbreeding. From this study, we conclude that an optimized phenotyping strategy can have potential long term benefits in breeding programs but more work is needed to investigate under which conditions benefits are largest.

INTRODUCTION

Trait measurement provides the necessary information to perform accurate genetic evaluation, whether it is based on phenotype on the animal itself or on its relatives. However, it can be costly and/or difficult to achieve trait measurement for a large number of animals (e.g. carcass traits, methane emission). Reducing the number of animals phenotyped is a simple and efficient way to cut cost and/or allow economically important traits to be part of the selection criterion, but the question is how phenotyping costs can be reduced with minimal impact on genetic gain.

Various efforts to manage cost of phenotyping have been made over the years. The first attempt to manage measurement was made by Robertson (1957) who proposed a theory to optimise the family size in a progeny testing breeding program by optimising the product of expected selection differential and accuracy. Wade and James (1990) developed a theory to manage the cost of testing while limiting a reduction in genetic gain. They mainly optimised the proportion of selection candidates to be phenotyped. More recently, Okeno et al. (2014) found that using knowledge on previously estimated breeding values was better than phenotyping randomly selected animals and that a phenotyping 80% of the animals provides the same gain as when all animals were phenotyped. Previous studies were therefore mainly concerned with determining an optimal proportion of animals to be phenotyped, but did not give any insight about which particular individuals should be measured.

In a previous study, we developed a phenotyping strategy for a single trait measured on an optimised subset of animals with no prior information on the candidates to be considered (Massault *et al.* 2013). A set of individuals was phenotyped that maximised the information and thus the accuracy of genetic evaluation. While this strategy proved to be efficient, it does not reflect a practical breeding program where multi-trait and multi stage selection is common and prior information on candidates exists in the form of estimated breeding values.

In the current study, we present a phenotyping strategy that uses prior knowledge for the case of two-stage two-trait selection. We use a small paternal half-sib population structure and extreme parameters to assess the potential usefulness of such a strategy.

MATERIALS AND METHODS

Population simulation. To explore the efficiency of our selection criterion, we simulated a small pedigree of 10 paternal half-sib families, comprising 10 offspring each. We simulated genetic and environmental values for 2 traits with both a heritability of 0.3 and a phenotypic variation σ^2_P of 100 (e.g. body weight and feed efficiency) with a correlation between traits of r_A = 0.5 and r_E = 0.25. Traits 1 and 2 have economic value of 0.01 and 0.1 respectively (the most important trait being the one with restricted phenotyping). We therefore have a breeding objective G of 0.01 * EBV_{Trait1} + 0.1 * EBV_{Trait2} (EBV = Estimated Breeding Value). We used selection in stage 1 to determine which selection candidates to phenotype for Trait 2, and offspring were selected to become parents at stage 2. All animals were measured for Trait 1 before stage 1 selection. We had three different scenarios for prior information on Trait 2; NOT2, where no parents of selection candidates have been phenotyped for Trait 2; ST2, where sires have been phenotyped for Trait 2 and PT2, where both parents have been phenotyped for Trait 2. These scenarios differ in the amount of information known about the between and within family components of EBV, and hence the correlation in EBV among relatives We calculated EBVs at stage 1 using multi-trait BLUP based on all available information on both traits. EBVs in stage 2 were calculated using BLUP after phenotyping 20% of selection candidates for Trait 2. We used optimum contribution selection at stage 2 (Sonesson and Meuwissen, 2000), where contribution of animals to the next generation are optimised and balanced with diversity:

$$Optimum\ Contribution = \boldsymbol{x}'\boldsymbol{G} + \lambda_2\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$$

where $\boldsymbol{G}$ is a vector of breeding values, $\boldsymbol{x}$ a vector of contribution to the next generation and $\boldsymbol{A}$ the numerator relationship matrix. We used 6 different values for λ_2 (0, -10,-100, -1000, -9999). We then compared the genetic gain (**x'G**) for the same level of inbreeding (F = **x'Ax** / 2) between different phenotyping strategies.

Selection criteria for phenotyping (stage 1 selection). We used 2 strategies to select 20% of selection candidates to phenotype: DIVERSITY, where animals are selected for phenotyping based on merit as well as diversity, and MERIT, where animals are selected simply based on merit (highest EBVs). We also simulated an ALL strategy where all selection candidates were phenotyped. We propose a selection criterion based on the average (EBVs) of ‘would-be’ phenotyped animals and their genetic diversity:

$$Selection\ criterion = \boldsymbol{x}'\boldsymbol{G} + \lambda_1\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$$

where $\boldsymbol{G}$ is the vector of expected breeding value of phenotyped animal, $\boldsymbol{x}$ a vector indicating for each animal 1/**n** (number of phenotyped animals) when phenotyped or 0 otherwise and $\boldsymbol{A}$ the numerator relationship matrix. The first term reflects the average breeding values of phenotyped animals while the second term reflects the diversity between phenotyped animals. We use the extreme value of -9999 for λ_1 for DIVERSITY. Note that the MERIT scenario where the phenotyped set is chosen based on merit alone is equal to setting λ_1 to zero, while a scenario where diversity was the overriding criterion equal one where λ_1 = 9999. The optimality of the result will also depend on the importance of genetic diversity at the final selection stage, hence we also varied λ_2. We simulated 100 replicates.

Differential algorithm. To determine which subset of animals is best to phenotype, we used a differential algorithm (DE, Storn and Price 1997). The DE creates 16 subsets of animals to phenotype (solutions). Each subset of phenotyped animals is then evaluated using a selection criterion (described above) and the solutions are ranked. The DE creates a challenger for each of the 16 solutions by crossing-over and mutating solutions (i.e. a different set of animals to be phenotyped). If the challenger performs better, the current solution is discarded and the challenger enters the pool for the next generation of solutions. Subsequently challengers are evaluated with the selection criterion over another 50,000 generations. At the end, the DE will have evolved to a best (or at least close to best) set animals to phenotype, for a given λ_1 value.

RESULTS AND DISCUSSION

Genetic gain and diversity. Fig. 1 shows the possible selection points at stage 2 (plotting genetic gain ($\mathbf{x'G}$) and level of inbreeding F ($\mathbf{x'Ax}/2$)) for the two selection options at stage 1, with MERIT (i.e. $\lambda_1 = 0$) and DIVERSITY($\lambda_1 = 9999$) for NOT2, ST2 and PT2 scenarios. ALL always out-performs the 2 other strategies, as expected, due to the fact that all selection candidates have been phenotyped. The ALL strategy gave by far the highest gain under the same level of inbreeding at stage 2 selection. The DIVERSITY strategy performed generally better than the MERIT strategy. The advantage is not so important in the case of NOT2, but DIVERSITY is significantly higher when phenotyping sire (ST2) and both parents (PT2). The performance of DIVERSITY increases as the level of information on parents increases.

Between and within family information. The difference seen between DIVERSITY and MERIT can be explained by the additional information given by phenotyping parents of selection candidates. In the case of NOT2, where no parents are phenotyped, there is no information other than the selection candidates own phenotype for Trait 1. EBVs from sibs are lowly correlated (correlation = 0.06, see Stanish and Taylor (1983) for calculation) which reduces the co-selection of relatives. Therefore, in this case, a phenotyping strategy that emphasises diversity is not really advantageous. MERIT distributed phenotypes to best animals across 8 different families (Fig 2.A) while DIVERSITY phenotyped 2 good animals for each family (Fig 2.B) resulting in DIVERSITY slightly better than MERIT. On the other hand, when sires have been phenotyped for Trait 2, sibs will have a higher correlation (0.40) among their EBVs. Selection on merit would emphasize family selection, which becomes restrictive at stage 2 selection (unless $\lambda_2 = 0$) and the distributing phenotypes over more families will allow more emphasis on within family selection at stage 2. Fig. 2.C shows that MERIT allocated phenotypes for the 5 best families while Fig 2.D shows that DIVERSITY phenotyped 2 good animals per family. This permits a relatively high genetic gain and also maximises the diversity and show the benefit of choosing specific individuals to phenotype rather than a random proportion to maintain diversity. The same principle applies when both parents have been phenotyped and the additional information brought more information on the selection candidates itself as each dam has a single progeny.

Further work. The results showed in this study proved that in selecting animals for phenotyping there is an advantage to emphasize diversity of the set to be measured. We concluded that phenotyping good animals across a larger number of families resulted in higher genetic gain than phenotyping the best animals of few families for same level of inbreeding. An optimal solution is likely to be less extreme than the DIVERSITY strategy, hence it is important to find the λ_1 value that optimises the subset of animals to phenotype that allows an optimal solution between merit and diversity at the second selection stage. It is also pertinent to explore the benefit of an optimal strategy, over a long period of time (e.g. 10 years of selection) and vary the parameter values such as trait heritabilities, genetic correlation and economic weights.

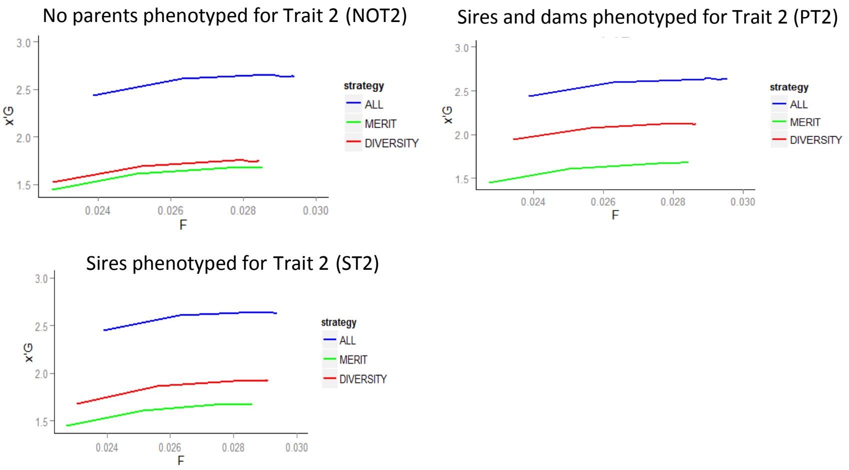

Fig 1. Genetic gain (x'G) versus diversity (F, level of inbreeding) plot for 3 different measurement scenarios NOT2, ST2 and PT2 with the three phenotyping strategies ALL, MERIT and DIVERSITY

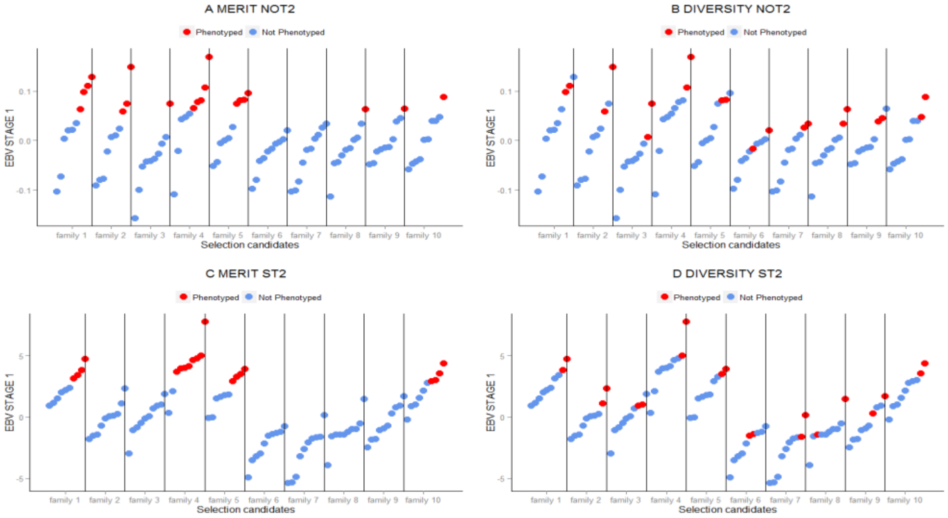

Fig 2. Phenotyped individuals in one replicate with no information on Trait 2 for MERIT(A) and DIVERSITY (B) and when sires are phenotyped for Trait 2 for MERIT (C) and DIVERSITY (D) . Individuals are classified by EBV at stage 1 and families.

REFERENCES

Massault, C, Kinghorn, B., van der Werf, J.H.J (2013) *Proc. Assoc.Advmt. Anim. Genet.* **20**: 416
Okeno, T. O., Henryon , M., Sørensen, A. C. (2014) *Proc. WCGALP*. **10**
Robertson, F.W. (1957) *Biometrics* **13**: 442
Sonesson, A.K., Meuwissen, T.H.E. (2000) *Genetics Selection Evolution* **32**: 231
Stanish, W.M., Taylor, N. (1983) *The American Statistician*, **37**: 221
Storn, R., Price, K. (1997) *Journal of global optimization*, **11**: 341
Wade, C.M., James, J.W. (1990) *Proc. Assoc. Advmt. Anim. Genet*. **8**:115