



HAL
open science

Maximising polygenic response to selection with an optimal frequency path for a selected major gene

Léopoldo Sanchez, Armando Caballero, Enrique Santiago

► **To cite this version:**

Léopoldo Sanchez, Armando Caballero, Enrique Santiago. Maximising polygenic response to selection with an optimal frequency path for a selected major gene. 8. World Congress on Genetics Applied to Livestock Production, Aug 2006, Belo Horizonte, Brazil. hal-02754201

HAL Id: hal-02754201

<https://hal.inrae.fr/hal-02754201>

Submitted on 3 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MAXIMISING POLYGENIC RESPONSE TO SELECTION WITH AN OPTIMAL FREQUENCY PATH FOR A SELECTED MAJOR GENE

L. Sanchez^{1,2}, A. Caballero¹ and E. Santiago³

¹ Dep. Bioquímica, Genética e Inmunología, Universidad de Vigo, 36310 Vigo, Spain

² INRA-Orléans, UAGPF. B.P. 20619 - 45166 Olivet Cedex, France

³ Dep. Biología Funcional, Universidad de Oviedo, 33071 Oviedo, Spain

INTRODUCTION

The use of gene-assisted selection can provide extra gains in short periods of time (Gibson 1994). However, the selection of a gene with a major effect on the selected trait reduces the intensity of selection and the effective population size applied at the level of unidentified genes (polygenes). Thus, the increased gain from the major gene occurs at the cost of reduced response from the polygenic variation (Gibson 1994; Pong-Wong and Woolliams 1998; Villanueva *et al.* 1999). In the context of a selected population where fixation of a known gene is desired, the question arises as to which is the optimal path of increase in frequency of the gene so that the selective sweep of polygenic variation resulting from its fixation is minimised.

In the frequency path recently proposed by Meuwissen and Sonesson (2004), the selection intensity on the major gene remains constant over generations. The reasoning behind this choice was the simulation evidence by which optimally weighted major gene information over generations leads to approximately constant selection intensity at the major gene, first noticed by Dekkers and van Arendonk (1998). Using theoretical arguments we propose an alternative frequency path that maximises simultaneously the effective population size and the selection intensity applicable to the polygenes by minimising the average squared selection intensity on the major gene over generations up to a given fixation time.

OPTIMAL FREQUENCY PATH FOR THE MAJOR GENE

The frequency path for the major gene should be derived taking into account the reductions in effective selection intensity and effective population size (Ne) of the polygenes. Let us consider the former, assuming an infinite population where a QTL has to be fixed in n generations. Selecting the QTL at generation t is equivalent to assign it an effect a_t . We can give a_t as a function of the increase in frequency $\Delta_t = (q_{t+1} - q_t)$ of the QTL, $a_t \approx (\Delta_t \sigma_p / i q_t (1 - q_t))$, where σ_p and i are the phenotypic standard deviation and the selection intensity for polygenes respectively. Considering the above QTL effect, then the expected response of polygenes will be reduced to $R_t^* \approx R[1 - (\Delta_t^2 / 2q_t(1 - q_t)i^2)]$, where R is the expected response if the QTL were absent. The total loss of response due to the increase in the QTL frequency until its fixation at generation n is then

$$\sum_{t=1}^n (R - R_t^*) \approx \frac{R}{2i^2} \sum_{t=1}^n \left(\frac{\Delta_t^2}{q_t(1 - q_t)} \right) \quad [1]$$

Note that Δ_t is the selection differential at the major gene and $q_t(1 - q_t)$ is the variance of allele frequencies. Hence, the ratio $\Delta_t^2 / q_t(1 - q_t)$ is equivalent to the squared selection intensity applied to the QTL, or the variance of the contributions in QTL copies by individuals of generation t (V_t). Thus, in order to minimise the loss of response in the polygenic background, the average squared selection intensity at the QTL across generations has to be minimised.

Considering now the reduction in N_e , we assume for simplicity that the cumulative effect of selection on N_e is ignored (Santiago and Caballero 1995). The expected N_e for an unlinked neutral locus at generation t can be given as a function of the population size (N) and V_t (Robertson 1961), $N_e = N/(1+V_t)$. Over the selection process on the QTL, the expected N_e is the harmonic mean given as

$$\overline{N_e} = N / (1 + \overline{V}) \quad \text{and} \quad \overline{V} = \frac{1}{n} \sum_{i=1}^n \frac{\Delta_i^2}{q_i(1-q_i)}. \quad [2]$$

Therefore, in order to maximise $\overline{N_e}$, \overline{V} must be minimised. The method presented here (hereafter PS path) shares the same principle as that of Meuwissen and Sonneson (2004) (denoted M&S path), thus to decide the progression of the QTL allele up to its fixation. However, our principle is slightly but functionally different. Our rationale is that the impact of the QTL fixation on the reduction of selection intensity and that on the reduction of effective population size are both minimised through the minimisation of the average squared selection intensity at the QTL across generations, rather than keeping a strictly constant value.

SIMULATION METHODS

Simulations were carried out modelling a diploid population, with $N = 40$ founder genotypes (half of each sex) assumed to be in linkage equilibrium. The population was reproduced for 40 additional non-overlapping generations at constant breeding size. $N/2$ random mating couples were established each generation, and each dam produced $n = 10$ full-sibs (half of each sex). Directional selection proceeded on a quantitative trait, which comprised a biallelic QTL of large effect, and 199 biallelic loci with minor effects (polygenes). Additive gene action was assumed within and between loci, and mutation was absent. All alleles of polygenic loci were set at frequencies of 0.5 at $t = 0$, while only one copy of the favourable QTL was present at $t = 0$. All selected loci were assumed to be equally spaced across a single chromosome L cM long. Two genome lengths were investigated, $L = 20$ to $L = \infty$. Crossing-overs occurred without interference. The genotypic value of the QTL was known without error, while the phenotypic value from polygenes was obtained as the sum of a genotypic value and an environmental deviation. The latter was sampled across generations as a normal deviate with mean zero and variance $\sigma_E^2 = [\sigma_A^2(1-h^2)/h^2]$, where h^2 is the narrow-sense heritability at $t = 0$, and σ_A^2 is the additive genetic variance at $t = 0$. The probability of identity-by-descent was estimated from an additional set of 200 neutral loci located uniformly between the selective loci.

Classic truncation selection (TS) based on the aggregate of QTL genotype and the polygenic phenotype was compared to the two path selection schemes, M&S and PS. The frequency path was expressed in terms of the number of favourable copies for the QTL in the selected candidates per generation up to a given fixation horizon ($Tfix$). The PS path was obtained by applying the simulated annealing technique (Press *et al.* 1992). The objective in M&S was to find the optimal allocation of gene copies from generation $t = 1$ to $Tfix - 1$, so that all the terms involved in the sum across generations in [1-2] were equal, while for the PS path the objective was to minimise [1-2] across generations. Once the paths were obtained, selection proceeded by assigning the required number of carriers each generation, observing simultaneously the maximisation of the polygenic phenotype among selected carriers and non-carriers of the QTL.

RESULTS AND DISCUSSION

The two paths are plotted in Figure 1a, together with the path described by a QTL allele under TS, as a reference. The M&S path describes an approximately symmetrical curve from the initial p to $p = 1$, where an equal number of generations occur before and after $p = 0.5$. The PS

path, however, is asymmetrical relative to $p = 0.5$, showing a slower rate of increase of the QTL allele during the initial two thirds of the QTL segregation period, with faster increases during the last generations before fixation. Due to the counteracting effect between the selection pressure on the QTL and that on the polygenes, less selection pressure on the QTL corresponds to a greater emphasis of the selection pressure on the polygenes. Thus, the PS path favours the progression of the polygenes relaxing the pressure on the QTL during the first stages of the process, when the risk of loss caused by drift and hitchhiking is highest (Caballero *et al.* 1996), and increases the pressure on the QTL, relaxing that on the polygenes, when most of the favourable polygenic variability has been efficiently assured.

Figures 1b and c show the difference in polygenic gain and coancestry, respectively, between PS and M&S paths, together with filled symbols to indicate when the difference was significantly larger than zero. Without linkage ($L = \infty$), PS obtained 0.3% higher polygenic gain than M&S once the QTL became fixed (Figure 1b), accumulating also 0.4% less coancestry (Figure 1c). With $L = 20\text{cM}$, the advantage in polygenic gain of PS over M&S was of 1.1%. This advantage, however, happened at the cost of non significant differences in coancestry at the time of QTL fixation. The temporal reduction of the advantage of PS at T_{fix} , was not enough to reverse the PS's cumulated advantage attained earlier.

A question that remains partially unresolved for these methods is the choice of the selection horizon (T_{fix}). In our model, however, we found little advantage of $T_{fix} = 36$ over 16 (not shown), and our results suggest that most of the long-term benefits behind PS rely on the shape of the QTL path, rather than on the choice of a longer T_{fix} .

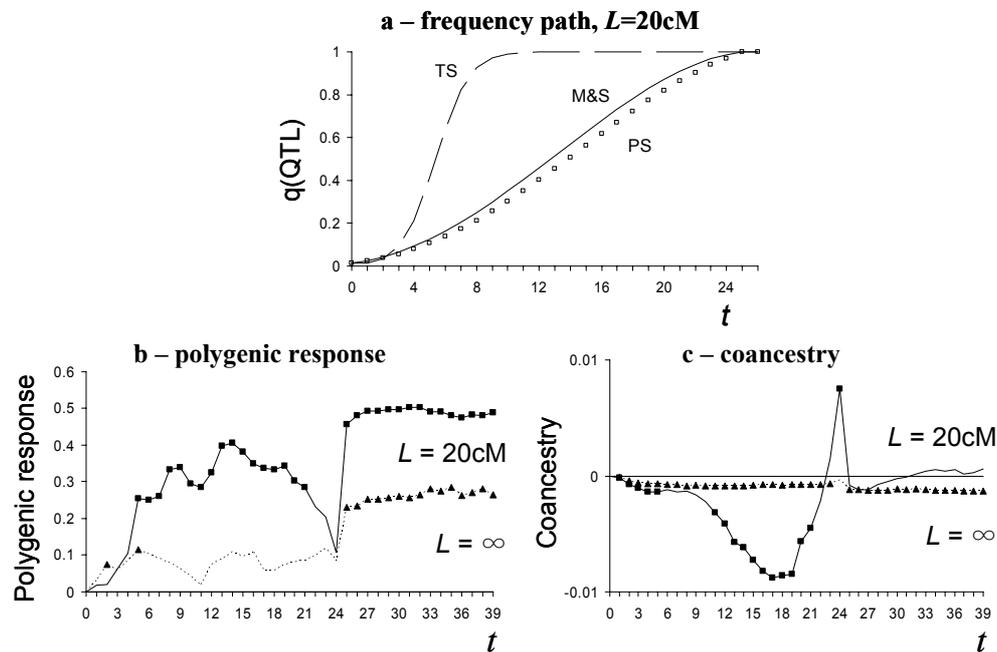


Figure 1. Frequency paths (a) for TS with a QTL effect of one standard polygenic phenotypic deviation, and for PS and M&S methods, both with $T_{fix} = 26$. Difference in polygenic response (b) and coancestry (c) between PS and M&S methods across generations, for two levels of linkage. Initial heritability 0.1. Filled symbols in (b) and (c) indicate differences significantly larger than zero at a 5% level.

CONCLUSION

The path proposed is based on the theoretical argument by which both the effective selection intensity and the Ne applicable to the selected polygenes are maximised by minimising the average squared selection intensity on the major gene over generations up to a given selection horizon. The results show that the two collateral effects of a selective sweep of variation caused by the fixation of a major gene, *i.e.* the reduction of selection response from linked polygenes and the increase in inbreeding in the region close to the major gene, can be efficiently reduced by the implementation of the path.

REFERENCES

- Caballero, A., Wei, M. and Hill, W.G. (1996) *J. Genetics* **75** : 63-80.
- Dekkers, J.C.M. and van Arendonk, J.A.M. (1998) *Genet. Res.* **71** : 257-275.
- Gibson, J.P. (1994) *Proc. 5th WCGALP* **21** : 201-204.
- Meuwissen T.H.E. and Sonesson, A. (2004) *Genet. Res.* **84** : 109-116.
- Pong-Wong, R. and Woolliams, J.A. (1998) *Genet. Sel. Evol.* **30** : 313-337.
- Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1989) «Numerical Recipes: The Art of Scientific Computing». Cambridge University Press, Cambridge, UK.
- Robertson, A. (1961) *Genet. Res.* **2** : 189-194.
- Santiago, E. and Caballero, A. (1995) *Genetics* **139** : 1013-1030.
- Villanueva, B., Pong-Wong, R., Grundy, B. and Woolliams, J.A. (1999) *Genet. Sel. Evol.* **31** : 115-133.