

Day 4 Genomic Prediction and Genomic Selection

Exercise 1 Accuracy of Genomic Predictions

Use the spreadsheet ‘*GSaccuracy 2015.xls*’ to investigate and explain the impact of the following parameters on the effective number of chromosome segments (M_e), the proportion of variance explained by markers (q^2), the accuracy with which marker effects are estimated (r_{Qhat}), and the accuracy of the genomic prediction or molecular breeding value, MBV (r_{MBV}):

- number of markers (M)
- Effective population size (N_e)
- Heritability of phenotypes (h^2)
- Number of training individuals (N)

Set $L=1$ and $k=30$ for a genome of 30 chromosomes of 1 Morgan

- a) What is the minimum number of markers that is needed to achieve near maximum genome coverage ($q^2=0.99$) when $N_e=100$ versus 340 versus 1000 versus 10,000? Enter this in the table below.
- b) Set the number of markers $M = 1,000,000$ to get nearly complete coverage regardless of N_e . Set $h^2=0.9$.
Now evaluate the size of the training set (N) needed to reach an MBV accuracy of 0.8 for $N_e=100$, versus 340 versus 1,000 versus 10,000. Enter the results in the table below.
- c) Repeat b) for heritabilities equal to 0.5 and 0.2

	$N_e = 100$	$N_e = 340$	$N_e = 1000$	$N_e = 10,000$
Min. # markers				
$h^2 = 0.9$				
$h^2 = 0.5$				
$h^2 = 0.2$				

- d) Test the genome scaling argument that if the size of the simulated genome is reduced by a factor C, then the size of the training population also has to be reduced by the same factor C in order to maintain the same accuracy of MBV.

Exercise 2 EBVs combining phenotype and GBV

Consider selection for an additive trait with (total) heritability 0.2 and phenotypic standard deviation equal to 50. For this trait, you have been able to derive genomic breeding values with accuracy (r_{MBV}) equal to 0.5. The following sources of information are available for selection:

- own MBV
- own phenotype
- MBV of the individual's sire
- Phenotype of the individual's sire

Assume the top 20% individuals are selected. The base scenario is having own phenotype and phenotype of the sire only (no genotypes). Evaluate the impact of the following on the accuracy of selection and the genetic superiority of the selected individuals, using the spreadsheet 'STEBVaccuracy.xls':

- Adding genotyping of the individual to the base information of own phenotype and phenotype of the sire.
- Adding genotyping of the sire to the information from question a).
- Derive the elements of the P matrix and the G vector for question b)
- Explain the weight on the MBV of the sire in question b)
- Convert this MBV blending problem into a two-trait problem consisting of the original trait and the MBV as a correlated trait with heritability = 1. I.e. derive the genetic and phenotypic correlations between the two traits. Enter the results into MTindex.xls and check that it gives the same answers for question b).

Exercise 3 GS with SelAction

Load input file 'Day4 Single trait GS.d1s' into SelAction. This represents a scenario for a trait with $h^2=0.4$ and genomic EBV with accuracy = 0.8. Explore the impact of the following scenarios on accuracies of selection, response to selection, and rate of inbreeding:

- Adding genotyping males only, females only, or both, to the phenotypic records that are already available.
- Adding genotypes of relatives to scenario a).
- Now consider a sex-limited trait, for which phenotypes are only recorded on females. Now evaluate the impact of genotyping males only, females only, or both. Compare results to a)
- For scenario c), evaluate the impact of genotyping male selection candidates when they already have 30 progeny with phenotypes

- e) For scenario a) evaluate the impact of the Bulmer effect on accuracy and genetic variance and response, with or without genotyping both male and female selection candidates. To evaluate accuracy without selection, change the proportion selected to 0.99999

Exercise 4 Multiple trait MBV with SelAction

Load input file 'Day 4 Multi trait GS.d1s' into SelAction under 'Discrete Generations' '1-stage selection'.

It represents the 2-trait example with GS from the notes.

Explore the impact on responses of phenotype-based BLUP selection, selection on MBV, and combined selection on MBV and phenotype-based BLUP, for different scenarios, e.g

- having no own or sib phenotypes available prior to selection on males, females, for trait 1, trait 2, or both, with or without MBV.

Exercise 5 Inbreeding with GS

Load input file 'Day 4 Single trait GS.d1s' into SelAction under 'Discrete Generations' '1-stage selection' and evaluate the impact of including MBV on rates of inbreeding.

Exercise 6

Consider a purebred line producing pigs for crossbred performance.

Use CROSSBRGS.XLS to work out selection accuracy of the Index.

The breeding goal is 100% CB information.

Selection in the purelines is based on own performance, sire, dam, 3FS and 40HS.

In addition here are 40 Xbred HS performances recorded.

Trait heritability 0.25; c-squared 0.15.

Work out the accuracy of the index for 2 values of r_{PC} (correlation between purebred and crossbred performance):

Cases	$r_{PC} = 0.7$	$r_{PC} = 0.9$
-------	----------------	----------------

No Genomic Selection

No information from Xbreds

With HS Information from Xbreds

Genomic Selection ($N_e=100$)

With HS Information from Xbreds

RefPopSize = 3000 PB

Genomic Selection

With HS Information from Xbreds

RefPopSize = 1500PB and 1500XB

Genomic Selection

With HS Information from Xbreds

RefPopSize = 3000 XB