

Chapter 8.15.1: Principle of genomic selection

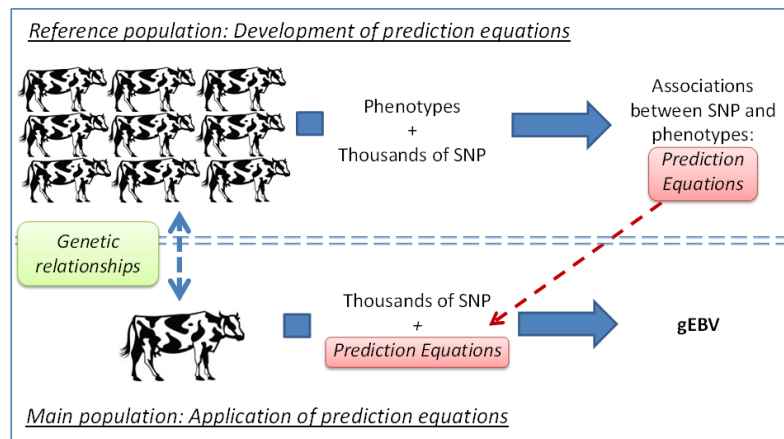
Animal breeding

Exported on 03/15/2021

Table of Contents

No headings included in this document

In figure 7 the general principle of genomic selection is illustrated. First a lot of information needs to be collected on a select group of animals: the *reference population*. All animals in this reference population are genotyped for a very large number of SNP that are nicely spread across the entire genome. How many, is still under debate, but at least multiple thousand (e.g. 60,000). Genotyping for more SNP is more expensive, but will also result in more accurately estimated associations between the SNP and the phenotypes (i.e. the SNP effects). It is also still under debate what would be the optimum number of animals in the reference population. A larger population obviously is more expensive as the phenotyping and detailed genotyping of these animals is costly. But a larger population also allows for more accurate estimates of the SNP effects. Like many aspects in animal breeding the choice of size of the reference population and number of SNP will be a matter of cost-benefit analysis.



Given the phenotypes and the genotypes of the reference population, the associations between genotype and phenotype will be estimated for each of the genetic markers. Subsequently, the estimated effects are combined into so-called *prediction equations*. These are just summations (the estimated SNP effects are additive!) of the effect of the first SNP + that of the second SNP + ... + that of the last SNP, so that the end result is the sum of all estimated SNP effects. Because each SNP has 2 alleles, for each SNP there are 3 possible genotypes. The prediction equations are established such that for each SNP the effects of all genotypes that are present in the reference population are estimated. This is one of the reasons why you need a large reference population: to estimate all these SNP effects accurately, each genotype need to be represented by a sufficiently large number of animals. Now we have a set of equations with estimated SNP effects. The breeding value of animals outside the reference population can now be estimated by applying these equations to their SNP genotypes. These breeding values based on genomic information only are called genomic breeding values, or gEBV.

Thus:

Genomic selection is based on estimation of detailed associations between a very dense set of genetic markers (SNP) and phenotypes on a select group of animals: the reference population.

The resulting prediction equations are then applied to SNP genotypes of the rest of the population to estimate their genomic breeding value (gEBV), without the need for additional phenotypes.