Can the Use of Genotype Probabilities with Selective Genotyping Reduce Bias in Genetic Marker Estimates?

M.L.A.N.R. Deepani[1,2], J.P. Gibson[2] and B.P. Kinghorn[2]

[1]*Faculty of Veterinary Medicine and Animal Science, University of Peradeniya*
[2]*University of New England, Australia*

Incorporation of genetic marker information in animal breeding programs is useful in improving traits with low heritabilities, traits that are hard to measure, expressed in only one sex or later in life, rare alleles etc. The marker genotype information of such traits is used in estimation of breeding values of animals, in order to select the best animals for breeding purposes. However, obtaining good estimates for the markers are necessary for this to be more effective. Genotyping of the entire population is usually not practicable to obtain a good estimate for the markers. Therefore, most of the time, only a proportion from a population is genotyped. If the pedigree information is available, the utility of the genotype information can be improved with the use of genotype probabilities of the non-genotyped individuals.

Due to the comparisons of individuals in two phenotypic extremities of a trait, selective genotyping has shown more power over random selection in detecting quantitative trait loci (QTL) effects. However, when a smaller proportion is genotyped from the population, the estimates are expected to be biased upwards, and the assumptions of the formula for correction of this bias tend to deviate. Use of genotype probabilities reduces this estimation bias due to the inclusion of additional genotype information in the analysis. Since the correction formula available in the literature tends to over correct the "less biased" estimates obtained, a modification was required to accommodate the additional information from animals with genotype probabilities in the formula.

This study attempts to fill this gap and compares the use of genotype probabilities with selective genotyping in order to estimate a bi-allelic single nucleotide polymorphism (SNP) marker segregating in a population, with an allele frequency of 0.5 and allele substitution effect (α) = 6. A random mating population [phenotypic mean (μ) =100, standard deviation (σ) =30, heritability (h^2) =0.25] was simulated for this purpose. The marker genotype or the probability was fitted as a fixed effect in Best Linear Unbiased Prediction (BLUP). Marker estimates obtained from BLUP was compared when 50% or 35% of the population was genotyped, with and without using genotype probabilities under selective genotyping and random selection. Each experiment was replicated 20 times.

Results show that the corrected estimates for bias under selective genotyping with 50% or 35% genotyped were more accurate (smaller standard error, $p<0.05$), compared to random selection under similar proportion genotyped. At 35% genotyping, there was a significant deviation of the marker estimate ($P<0.05$) from the true value (6) under selective genotyping, when the individuals with genotype probabilities were not used in the analysis. This suggests that inclusion of individuals with genotype probabilities can increase the utility of genotype data, and facilitate obtaining accurate and unbiased estimates with selective genotyping at extreme levels of genotyping.