



## Imputation of non-genotyped parents using genotype information of their relatives

E. Dehnavi<sup>a,\*</sup>, S. Ansari Mahyari<sup>a</sup>, F. S. Schenkel<sup>b</sup>, M. Sargolzaei<sup>b,c</sup>

<sup>a</sup> Department of Animal Science, College of Agriculture, Isfahan University of Technology, Isfahan, 84156-83111, Iran

<sup>b</sup> Department of Animal Biosciences, Centre for Genetic Improvement of Livestock, University of Guelph, Guelph, N1G 2W1, Canada

<sup>c</sup> Semex Alliance, Guelph, N1H 6J2, Canada

\* e.dehnavi@ag.iut.ac.ir

**Abstract:** In genomic evaluation, animals with reliable proofs are usually used in the training population. Even though historical proofs are normally publicly available, genotypes may not be in some circumstances. Imputing non-genotyped parents using information of their relatives can increase information content in the training population especially for countries that do not have access to genotypes of older animals. Therefore, the objectives of this study were to compare the imputation accuracy of non-genotyped parents from using genotype information of different relatives and to assess the accuracy of genomic prediction using imputed parents. First, an animal population was simulated using QMSim software [2]. Two traits, representing production ( $h^2=0.3$ ) and functional traits ( $h^2=0.05$ ), were independently simulated and 20 replicates were carried out. The different sources of information consisted of genotypes from an alternate number of offspring (i.e., 4 to 20 offspring), both parents, and mates, which were used to impute non-genotyped animals using FImpute software [4]. Concordance rate was measured as the average proportion of correctly imputed genotypes per individuals. Animal-specific imputation accuracy, was computed as Pearson correlation coefficient between imputed and true genotypes of an individual, which adjusts for MAF distribution. The imputed animals were then included in the training population to evaluate the impact of imputation on the bias and accuracy of genomic predictions [3]. Using mates plus 7 paternal half sibs resulted in an average genotype concordance rate of 0.99 and an average adjusted animal-specific imputation accuracy of 0.95. Using as many as 20 progeny without using parents' and mates' genotypes, the concordance rate and animal-specific imputation accuracy were 0.97 and 0.95, respectively. The average concordance rate and animal-specific imputation accuracy of 0.95 and 0.89 were observed using genotype information of both parents plus 7 progeny and no mates, respectively. With real sheep data, similar result for genotype concordance rate (0.99) was found when genotype information of 7 progeny plus mates were considered [1]. Regarding to genomic predictions, using imputed individuals based on mate information in addition to progeny information yielded similar GEBV accuracy and bias to those obtained when true simulated genotypes were used. This study illustrated that high accuracy of imputation for non-genotype individuals can be achieved when mates' and parents' genotypes, and a sufficient number of progeny genotypes (i.e. 7 progeny) are available. Since a progeny receives one chromosome from its sire and one from its dam, knowing the mate genotype besides the progeny genotype can lead to more accurate haplotype inference. Having a large number of genotyped offspring helps to capture better the Mendelian sampling. Scenarios presented in this study are applicable especially in small ruminants (e.g. sheep and goat) and in some circumstances in dairy and beef cattle.

**Keywords:** Genomic selection; imputation; mates; sources of information; non-genotyped parents.

### References

- [1] DP. Berry, N. McHugh S. Randles, E. Wall, K. McDermott M. Sargolzaei, AC. O'Brien, Imputation of non-genotyped sheep from the genotypes of their mates and resulting progeny, *Animal*, (2017), 1-8.