Abstract

Genotype imputation from low- density to high-density SNP chips is an important step before applying genomic selection, because denser chips can provide more reliable genomic predictions. The objectives of this thesis were to study the accuracy of genotype imputation from the low and intermediate-density panels (5K and 50K) to the high density panels (777K) and compare the accuracy of genomic estimated breeding values (GEBVs) or EBVs before and after the genotype imputation in the different marker densities scenarios of reference population selection and methods of genomic breeding values estimation. Here we used simulated purebred and crossbred Holstein populations. The simulated populations included two purebred (lines A and B) and two crossbred populations (cross and backcross). Different scenarios for selecting animals in the reference population were evaluated to impute genotypes of the validation population including: 1) high relationship with validation population, 2) randomly and 3) highest inbreeding. The subsets of reference population were used to impute genetically from the low and intermediate-density panels (5K and 50K) to the high-density panels (777K). The genotype imputation was performed using FImpute software. The accuracies of imputation were compared through two methods of Pearson Correlation Coefficient (PCC) and Concordance Rate (CR). GEBVs in the validation population were estimated by using BLUP with the different combinations of marker and pedigree information. Three combinations of the pedigree and marker were used including: ABLUP in which only numerator relationship matrix A was used, GBLUP in which only genomic relationship matrix G was used and BLUP|GA which combined matrices A and G through a weighting parameter (λ). The accuracy of genomic evaluation was evaluated as a correlation between GEBV/EBV and the true phenotype values in the validation population. The results showed that the accuracy of genotype imputation in the purebred populations (lines A and B) was higher than the crossbred populations (cross and backcross). In the purebred populations, the accuracy of genotype imputation was the highest in the scenario of reference population selection based on the high relationships. In these populations, there was no significant difference between the imputation from 5K density and 50K density to 777K density when the reference population had been selected based on a high relationship, but there was a significant difference in the other scenarios ($P<0.01$). In the crossbred population, the highest accuracy of genotype imputation was obtained in the scenario of random reference population selection and in 50K density to 777K density. The results also demonstrated that the highest accuracy of breeding values prediction was obtained in the purebred populations by GBLUP method and in the scenario of the highly related reference population with the validation population; but in the cross and backcross populations for the trait with intermediate heritability (h^2 =0.25), the highest accuracy of breeding values prediction was obtained in the model with the weighting parameter equal to 0.2. In the scenario of the related reference population selection when 50K panel was used for the genotype imputation to 777K density, the accuracy of genome breeding values increased, but in the most scenarios of inbred and random reference population selection there was no significant difference in the accuracy of genomic breeding values prediction between 5k and 50k densities after their imputation to 777K density.

Key words: Accuracy, Genomic evaluation, Genotype imputation, Marker density, Reference population.

University of Zabol

Graduate school

Faculty of Agriculture

Department of Animal Science

Thesis Submitted for the Degree of Ph.D

In Genetic and Animal Breeding

Effect of genotype imputation on genomic prediction in Holstein cattle population of Iran

Supervisors:

Dr. Gholam Reza Dashab

Dr. Mohammad Rokouei

Advisors:

Dr. Mohammad Mahdi Shariati

Dr. Mehdi Vafaye Valleh

By:

Shirin Barjasteh

Summer 2019