

Evaluation of genotype imputation for ancient genomes from regions underrepresented in the reference panel

Hande Çubukcu¹, İdil Yet¹, Gülşah Merve Kılınç^{1*}

¹Department of Bioinformatics, Graduate School of Health Sciences, Hacettepe University, 06100, Ankara, Turkey

*Correspondence: gulsahkilinc@hacettepe.edu.tr

Keywords: imputation, ancient DNA, GLIMPSE

Genotype imputation decreases the need for producing high coverage genomes since it allows to generate diploid genotypes from low coverage genome sequencing data. GLIMPSE is a recently developed software for imputing low-coverage genomes and it outperforms other softwares regarding runtime and accuracy. Imputation of low coverage ancient genomes using GLIMPSE has revealed high levels of accuracy for alleles with frequency of minimum 0.02 for genomes well represented in the reference panel. However the accuracy of the imputation for the genomes from various populations that are under-represented in the reference panel is largely unknown. Here, we impute 56 low to medium coverage ancient genomes (0.25X-10X) from Turkey and neighboring regions encompassing Greece, Iran, Israel and Jordan and use 1000 Genomes Project data including genotypes of modern populations mainly from Europe, West Africa, East Asia and North America as the reference panel. We downsample two genomes with coverages >9X to coverages 0.25X, 0.5X and 1X and compute the imputation accuracy by comparing imputed data with genotyped data. The accuracy is over 0.80 for MAF over 0.05 in 0.5 and 1X imputations. We further run PCA and compute allele frequency for a set of selected imputed variants. Our analysis shows that imputation of ancient genomes which are not represented completely in the reference panel can still be promising for analysis based on alleles with minor allele frequency over 0.05. For rare alleles, to increase the accuracy as well as efficiency to use the imputed genomes in downstream population genetics analysis, addition of samples which might better represent the imputed population to the reference panel could be promising and needs further investigation.