

Figure S1. Distribution of the expected (pink) and observed (blue) minor allele fractions of the non-dbSNP heterozygous sites in DS1-II-2. The expected distribution was generated assuming that the allele fractions follow a binomial sampling process with $P = 0.5$. (A) Genotyping with GATK followed by its standard filters; (B) Genotyping with GATK followed by our error filters; (C) Genotyping with our Bayesian genotyper followed by our error filters. The discrepancy between the observed and expected distributions diminished when our error filters were applied, with average Kolmogorov-Smirnov test's P-values 0, 0.42 and 0.07 in 100 permutations for panels A, B and C respectively, suggesting the efficiency of our filters in removing technical artifacts that could not be removed by probabilistic genotypers.

