Supplementary Notes and Figures for: Comprehensive Benchmarking of Genotype Imputation Tools Using a Large-Scale Chinese Reference Panel

The detailed description of evaluation metrics

To evaluate the reliability of data imputed from different platforms, especially rare SNPs, we introduced Imputation Quality Score (IQS) [20]. IQS is computed by subtracting the chance agreement from the observed agreement and dividing by the maximum possible value of the numerator:

$$IQS = \frac{P_o - P_c}{1 - P_c} \tag{1}$$

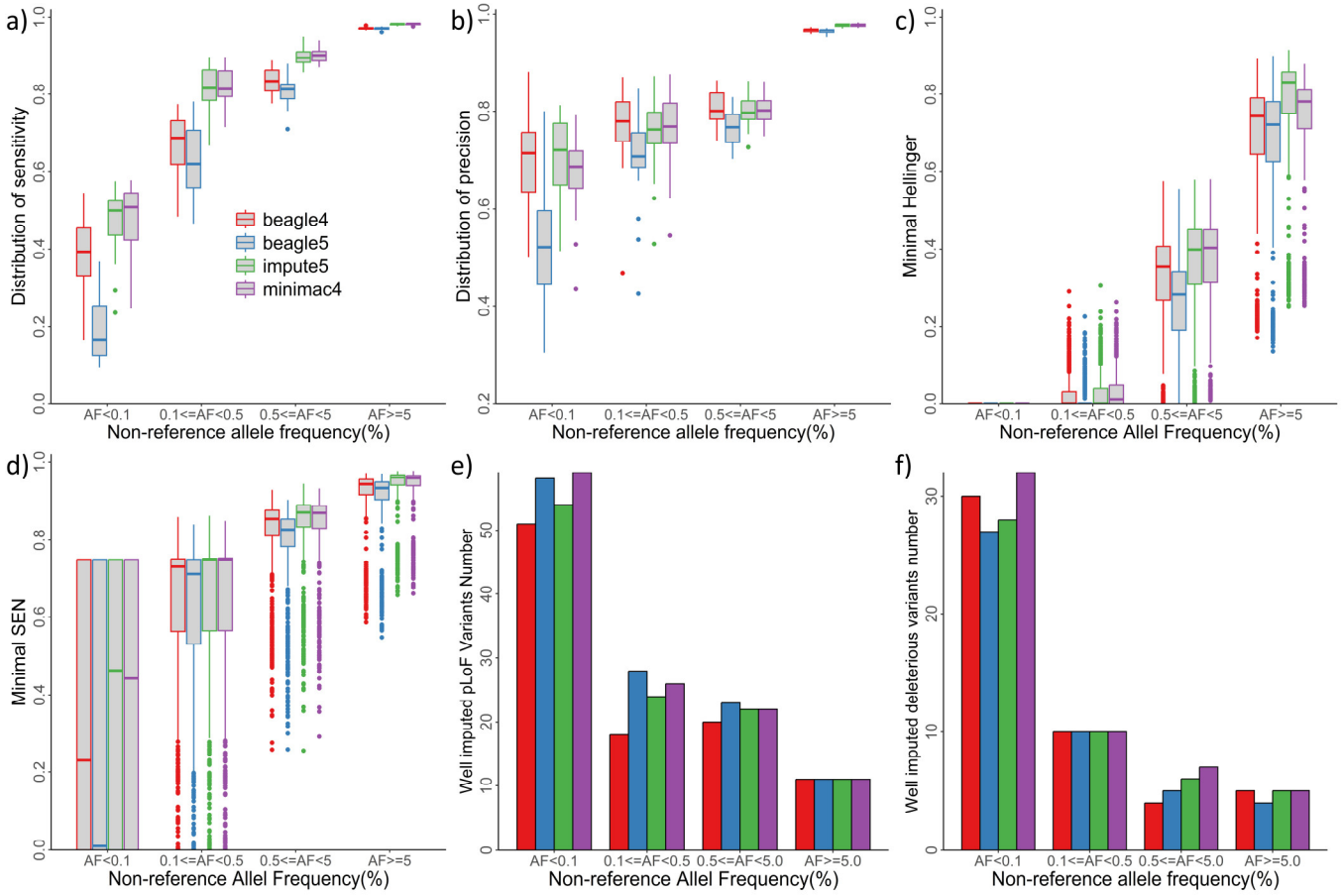
where the calculation method for observed proposal of agreement (P_o) and chance agreement (P_c) are:

$$P_o = \frac{\sum_i n_{ii}}{n_{..}} \tag{2}$$

$$P_c = \frac{\sum_i n_{i.} n_{.i}}{n_{..}^2} \tag{3}$$

The n_{ij} at each position in the following table represents the number of individuals with actual genotype i and impute genotype j, used to calculate observed agreement and chance agreement:

	True genotypes			
Imputed Genotypes	AA	AB	BB	Total
AA	n_{11}	n_{12}	n_{13}	$n_{1.}$
AB	n_{21}	n_{22}	n_{23}	$n_{2.}$
BB	n_{31}	n_{32}	n_{33}	$n_{3.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.3}$	$n_{..}$



Supplementary Fig. 1 The benchmarking results of the four imputation tools on the evaluation metrics that not mentioned in main text. The evaluation results respectively on **a)** distribution of sensitivity, **b)** distribution of precision, **c)** min Hellinger score, **d)** min SEN score, **e)** the number of well-imputed possible Loss of Function (pLoF) variants, and **f)** the number of well-imputed deleterious variants.