

Supplementary Table I. Imputation accuracy rates (in percent) of all imputed SNPs using HapMap II (HMII) and 1000 Genomes (1KG) data across the MAF spectrum. Red boldfaced are highest accuracy rates for each masked data.

Masking Rate	Public Data	Strategy	MAF spectrum			Polymorphic SNPs			Overall Accuracy
			Rare	Low Freq	Common	CEU-specific	YRI-specific	In both panels	
5%	HMII	YRI	14.4	57.9	79.3		71.7	78.3	77.1
		CEU	15.1	35.0	53.9	21.7		53.1	46.2
		INT	29.9	67.8	81.9			80.7	80.7
		UNI	35.4	63.2	82.6	44.1	73.5	82.4	74.2
		MER	15.2	45.9	78.4	21.7	71.7	78.3	66.7
	1KG	YRI	17.2	61.1	78.3		73.7	78.1	77.1
		CEU	15.4	34.0	54.7	23.5		54.9	47.8
		INT		67.2	80.8			80.2	80.2
		UNI	37.5	60.5	80.5	46.7	74.1	82.0	74.0
		MER	15.5	43.2	77.0	21.7	71.7	78.3	66.7
50%	HMII	YRI	12.6	49.6	71.0		65.6	70.4	69.6
		CEU	14.4	28.7	46.6	27.3		45.9	44.3
		INT	27.8	58.4	73.8			72.9	72.9
		UNI	30.7	55.5	75.0	48.8	67.8	75.1	72.4
		MER	14.5	39.0	70.4	27.3	65.6	70.4	66.4
	1KG	YRI	7.2	52.2	71.1		65.1	71.6	70.4
		CEU	13.4	28.5	49.2	29.4		49.2	46.8
		INT		57.0	74.2			73.7	73.7
		UNI	30.4	52.5	74.1	48.0	66.2	75.9	71.4
		MER	13.3	36.1	70.2	29.4	65.1	71.6	66.2
80%	HMII	YRI	4.9	29.8	49.6		43.8	49.3	48.4
		CEU	7.8	19.4	31.7	20.0		31.2	30.3
		INT	7.2	35.8	52.0			51.2	51.2
		UNI	17.1	36.3	53.9	35.3	45.8	54.1	51.8
		MER	7.9	24.7	49.2	20.0	43.8	49.3	46.4
	1KG	YRI		32.1	49.6		42.9	50.4	49.0
		CEU	8.0	19.6	34.7	20.8		34.7	33.1
		INT		38.0	52.2			51.8	51.8
		UNI	16.1	32.7	52.1	30.8	44.6	53.8	50.1
		MER	7.8	23.1	48.9	20.8	42.9	50.4	46.3

Rare variants with $MAF \leq 0.01$; Low frequency variants with $0.01 < MAF \leq 0.05$; Common variants with $MAF > 0.05$

INT refers the intersection strategy; UNI refers the union strategy; MER refers the merge strategy

CEU-specific SNPs are polymorphic only in CEU; YRI-specific SNPs are polymorphic only in YRI; SNPs in both panels are polymorphic in both CEU and YRI

Supplementary Table II. Numbers of imputed SNPs and filtered SNPs (with MACH R_{sq} over 0.3) in the masking data sets that were used for computing imputation accuracy.

Masking Rate	Public Data	Strategy	Imputed SNPs			Filtered SNPs		
			Rare	Low Freq	Common	Rare	Low Freq	Common
5%	HMII	YRI	13	61	903	9	54	880
		CEU	159	94	774	71	76	724
		INT	6	43	752	4	41	740
		UNI	163	111	898	111	105	889
		MER	166	112	925	78	92	900
	1KG	YRI	5	63	1129	3	58	1,068
		CEU	149	115	929	58	82	858
		INT		31	890		28	863
		UNI	153	147	1168	85	127	1,125
		MER	153	147	1168	60	115	1,102
50%	HMII	YRI	13	578	9063	5	483	8,503
		CEU	159	851	7842	60	637	6,967
		INT	6	412	7655	2	377	7,331
		UNI	163	1,007	8979	83	882	8,711
		MER	166	1,017	9250	64	779	8,667
	1KG	YRI	5	411	10032		317	9,036
		CEU	149	903	8615	50	611	7,568
		INT		213	8265		179	7,787
		UNI	153	1,101	10385	61	850	9,648
		MER	153	1,101	10382	51	761	9,331
80%	HMII	YRI	13	904	14,554		441	10,516
		CEU	159	1,301	12,607	32	684	8,629
		INT	6	613	12,303		347	9,416
		UNI	163	1,571	14,418	36	897	11,482
		MER	166	1,592	14,858	34	778	10,715
	1KG	YRI		613	16,040		222	9,464
		CEU	149	1,363	13,839	20	555	8,661
		INT		311	13,269		156	9,178
		UNI	153	1,665	16,613	18	597	10,663
		MER	153	1,665	16,612	22	628	9,792

Rare variants with $MAF \leq 0.01$; Low frequency variants with $0.01 < MAF \leq 0.05$; Common variants with $MAF > 0.05$

INT refers the intersection strategy; UNI refers the union strategy; MER refers the merge strategy