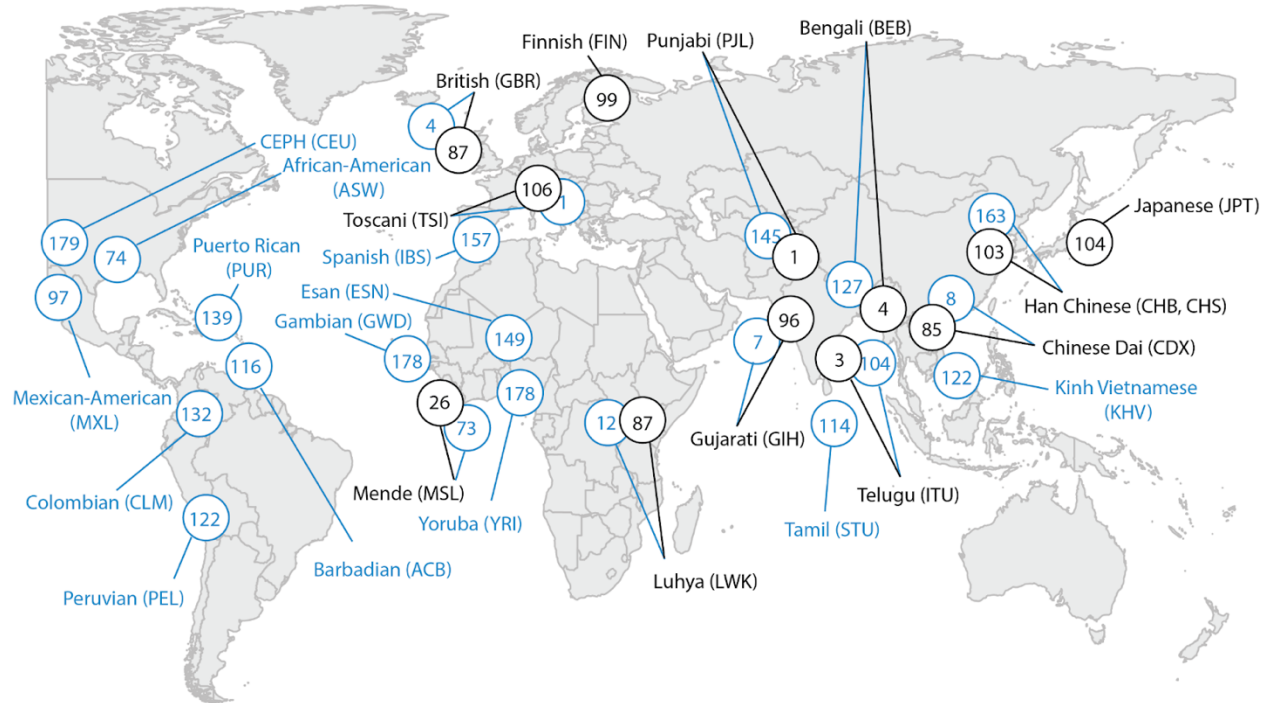


Supplementary Material

Supplementary Figures



Supplementary Fig. 1 | Samples used as imputation targets (black) and reference (blue) in the 1000 Genomes 30x Dataset.

Supplementary Tables

Supplementary Table 1. Overall imputation accuracy in the 1KG, TOPMed and UK Biobank dataset for Beagle5.4, Minimac4, IMPUTE5 and Selphi.

Method	Sum error	P0	R2	F-score	Dataset (N)
Selphi	300826681	0.99451	0.95196	0.9973	1KG (801)
Beagle5.4	324516831	0.99407	0.94834	0.9971	
IMPUTE5	336514775	0.99386	0.94648	0.9970	
Minimac4	343309686	0.99376	0.94469	0.9969	
Selphi	169093080	0.99870	0.81106	0.9991	TOPMed (5,000)
Beagle5.4	185309357	0.99851	0.79708	0.9990	
IMPUTE5	187626533	0.99849	0.79557	0.9990	
Minimac4	198005705	0.99841	0.78776	0.9989	
Selphi	9393488241	0.99780	0.97912	0.9999	UKB (50,000)
Beagle5.4	11878351998	0.99774	0.97381	0.9998	

Supplementary Table 2. Imputation accuracy for different MAF intervals in the 1KG dataset for Beagle5.4, Minimac4, IMPUTE5 and Selphi.

MAF (%)	Method	Sum error	Mean error	Variants	P0	R2	F-score
0.05 - 0.1	Selphi	6598303	0.7331	8999974	0.9991	0.0775	0.9995
	Beagle5.4	8049474	0.8944	8999974	0.9989	0.1077	0.9994
	Minimac4	7398909	0.8221	8999974	0.9990	0.0994	0.9995
	IMPUTE5	8486635	0.9430	8999974	0.9988	0.1081	0.9994
0.1 - 0.2	Selphi	10617239	1.3736	7729390	0.9983	0.1818	0.9991
	Beagle5.4	12007756	1.5535	7729390	0.9981	0.2216	0.9990
	Minimac4	11276290	1.4589	7729390	0.9981	0.2085	0.9991
	IMPUTE5	12571856	1.6265	7729390	0.9980	0.2205	0.9990
0.2 - 0.5	Selphi	17378413	2.4423	7115603	0.9970	0.3729	0.9985
	Beagle5.4	19137117	2.6894	7115603	0.9967	0.3971	0.9983
	Minimac4	18326241	2.5755	7115603	0.9968	0.3830	0.9984
	IMPUTE5	19959034	2.8050	7115603	0.9965	0.3928	0.9983
0.5 - 1	Selphi	15419809	3.8176	4039109	0.9953	0.5465	0.9976
	Beagle5.4	16991224	4.2067	4039109	0.9948	0.5472	0.9974
	Minimac4	16492572	4.0832	4039109	0.9950	0.5373	0.9975
	IMPUTE5	17660234	4.3723	4039109	0.9946	0.5421	0.9973
1 - 2	Selphi	18996525	5.7424	3308106	0.9929	0.6370	0.9965
	Beagle5.4	20902705	6.3186	3308106	0.9922	0.6307	0.9961
	Minimac4	20603432	6.2282	3308106	0.9923	0.6202	0.9962
	IMPUTE5	21665033	6.5491	3308106	0.9919	0.6253	0.9960
2 - 5	Selphi	30015434	9.1326	3286632	0.9888	0.7325	0.9945
	Beagle5.4	32578779	9.9125	3286632	0.9878	0.7248	0.9941
	Minimac4	33393736	10.1605	3286632	0.9876	0.7098	0.9940
	IMPUTE5	33730514	10.2629	3286632	0.9874	0.7197	0.9939
5 - 10	Selphi	29536778	14.8988	1982490	0.9818	0.8337	0.9916
	Beagle5.4	31634330	15.9569	1982490	0.9805	0.8273	0.9910
	Minimac4	33777032	17.0377	1982490	0.9792	0.8120	0.9904
	IMPUTE5	32777685	16.5336	1982490	0.9798	0.8230	0.9907
10 - 20	Selphi	46527577	22.0513	2109968	0.9731	0.8909	0.9886
	Beagle5.4	49394567	23.4101	2109968	0.9714	0.8861	0.9880
	Minimac4	54381112	25.7734	2109968	0.9687	0.8734	0.9867
	IMPUTE5	51200934	24.2662	2109968	0.9704	0.8828	0.9876
20 - 50	Selphi	105532014	28.5553	3695701	0.9654	0.9248	0.9871
	Beagle5.4	110892254	30.0057	3695701	0.9636	0.9211	0.9865
	Minimac4	125722639	34.0186	3695701	0.9591	0.9118	0.9846
	IMPUTE5	114867683	31.0814	3695701	0.9623	0.9184	0.9861

Supplementary Table 3. Imputation accuracy for different MAF intervals in the TOPMed dataset for Beagle5.4, IMPUTE5, Minimac4 and Selphi.

MAF (%)	Method	Sum error	Mean error	Variants	P0	R2	F-score
0.05 - 0.1	Selphi	815634	5.2200	156251	0.9994	0.3594	0.9995
	Beagle5.4	970834	6.2133	156251	0.9993	0.3683	0.9994
	IMPUTE5	999554	6.3971	156251	0.9993	0.3676	0.9994
	Minimac4	1067481	6.8318	156251	0.9993	0.3578	0.9993
0.1 - 0.2	Selphi	1124191	9.0146	124707	0.9991	0.4164	0.9991
	Beagle5.4	1355363	10.8684	124707	0.9988	0.4154	0.9989
	IMPUTE5	1391653	11.1594	124707	0.9988	0.4146	0.9989
	Minimac4	1489211	11.9417	124707	0.9987	0.4025	0.9988
0.2 - 0.5	Selphi	2396118	17.2333	139040	0.9982	0.4808	0.9983
	Beagle5.4	2917884	20.9859	139040	0.9978	0.4743	0.9979
	IMPUTE5	3003702	21.6032	139040	0.9977	0.4720	0.9978
	Minimac4	3239915	23.3020	139040	0.9975	0.4571	0.9977
0.5 - 1	Selphi	3078075	33.7323	91250	0.9964	0.5355	0.9967
	Beagle5.4	3775645	41.3769	91250	0.9956	0.5268	0.9959
	IMPUTE5	3887404	42.6017	91250	0.9954	0.5243	0.9958
	Minimac4	4235933	46.4212	91250	0.9950	0.5055	0.9954
1 - 2	Selphi	5050647	64.3190	78525	0.9931	0.5612	0.9937
	Beagle5.4	6227147	79.3014	78525	0.9914	0.5527	0.9923
	IMPUTE5	6398597	81.4848	78525	0.9911	0.5508	0.9920
	Minimac4	6966310	88.7145	78525	0.9902	0.5320	0.9914
2 - 5	Selphi	10560295	136.8357	77175	0.9853	0.5727	0.9870
	Beagle5.4	13021888	168.7319	77175	0.9815	0.5646	0.9840
	IMPUTE5	13395608	173.5744	77175	0.9808	0.5629	0.9835
	Minimac4	14541236	188.4190	77175	0.9791	0.5453	0.9822
5 - 10	Selphi	13591477	280.8272	48398	0.9692	0.5953	0.9743
	Beagle5.4	16480263	340.5153	48398	0.9614	0.5863	0.9694
	IMPUTE5	16908940	349.3727	48398	0.9602	0.5848	0.9686
	Minimac4	18217330	376.4067	48398	0.9570	0.5684	0.9663
10 - 20	Selphi	23206974	463.7407	50043	0.9463	0.6622	0.9608
	Beagle5.4	26834050	536.2198	50043	0.9351	0.6508	0.9560
	IMPUTE5	27376943	547.0684	50043	0.9334	0.6491	0.9552
	Minimac4	29340482	586.3054	50043	0.9284	0.6313	0.9523
20 - 50	Selphi	106154687	1004.1497	105716	0.8380	0.6073	0.9348
	Beagle5.4	109992405	1040.4518	105716	0.8291	0.5956	0.9334
	IMPUTE5	110392171	1044.2333	105716	0.8276	0.5934	0.9333
	Minimac4	114805306	1085.9785	105716	0.8223	0.5776	0.9304

Supplementary Table 4. Imputation accuracy for different MAF intervals in the UK Biobank dataset for Beagle5.4 and Selphi.

MAF (%)	Method	Sum error	Mean error	Variants	P0	R2	F-score
0.05 - 0.1	Selphi	195241576	34.39620144	5676254	0.9996	0.1059	0.9996
	Beagle5.4	224017164	39.46566944	5676254	0.9995	0.1339	0.9996
0.1 - 0.2	Selphi	258188230	62.03295489	4162114	0.9991	0.2015	0.9994
	Beagle5.4	289180242	69.4791738	4162114	0.9990	0.2320	0.9993
0.2 - 0.5	Selphi	485746254	132.2627309	3672586	0.9977	0.3697	0.9987
	Beagle5.4	536055465	145.9613104	3672586	0.9974	0.3861	0.9985
0.5 - 1	Selphi	474402503	260.0186698	1824494	0.9936	0.5789	0.9974
	Beagle5.4	535200676	293.3419765	1824494	0.9930	0.5692	0.9971
1 - 2	Selphi	538887052	360.4586021	1495004	0.9860	0.7371	0.9965
	Beagle5.4	631961022	422.7152717	1495004	0.9851	0.7142	0.9959
2 - 5	Selphi	826383016	439.0504839	1882205	0.9684	0.8577	0.9958
	Beagle5.4	1018628896	541.1891351	1882205	0.9673	0.8341	0.9949
5 - 10	Selphi	893364496	574.1557287	1555962	0.9324	0.9155	0.9949
	Beagle5.4	1129592078	725.9766485	1555962	0.9311	0.8966	0.9936
10 - 20	Selphi	1399944986	724.3714982	1932634	0.8742	0.9430	0.9941
	Beagle5.4	1814800220	939.0294386	1932634	0.8727	0.9271	0.9924
20 - 50	Selphi	3410133462	898.962059	3793412	0.7788	0.9597	0.9937
	Beagle5.4	4549890356	1199.418981	3793412	0.7772	0.9464	0.9916

Supplementary Table 5. TOPMed Studies and Consensus Groups utilized for the construction of the TOPMed reference panel.

	Study Name	Study ID	Consensus Group	Consensus Code	Version
1	AfricanAm Sarcoidosis	phs001207	c1	DS-SAR-IRB	v2.p1
2	ARIC WholeGene	phs001211	c1	HMB-IRB	v3.p2
3	ARIC WholeGene	phs001211	c2	DS-CVD-IRB	v3.p2
4	Asthma Barbados	phs001143	c1	GRU-IRB	v4.p1
5	Asthma CostaRica	phs000988	c1	DS-ASTHMA-IRB-MDS-RD	v5.p1
6	BrazilSCD	phs001468	c1	GRU-IRB-PUB-NPU	v1.p1
7	CARDIA	phs001612	c1	HMB-IRB	v1.p1
8	CARDIA	phs001612	c2	HMB-IRB-NPU	v1.p1
9	CardioHealth	phs001368	c1	HMB-MDS	v2.p2
10	CardioHealth	phs001368	c2	HMB-NPU-MDS	v2.p2
11	CardioHealth	phs001368	c4	DS-CVD-NPU-MDS	v2.p2
12	CardioHealth Amish	phs000956	c2	HMB-IRB-MDS	v4.p1
13	Cleveland Family	phs000954	c1	DS-HLBS-IRB-NPU	v3.p2
14	Cleveland Afib	phs001189	c1	GRU-IRB	v4.p1
15	COPD	phs000951	c1	HMB-MDS	v4.p4
16	CVHealth	phs000993	c2	DS-CVD-IRV-MDS	v5.p2
17	CVHealth	phs000993	c1	HMB-IRB-MDS	v5.p2
18	Framingham WholeGene	phs000974	c1	HMB-IRB-MDS	v4.p3
19	Framingham WholeGene	phs000974	c2	HMB-IRB-NPU-MDS	v4.p3
20	GeneSTAR	phs001218	c2	DS-CVD-IRB-NPU-MDS	v2.p1
21	GENOA	phs001345	c1	DS-ASC-RF-NPU	v2.p1
22	GenSalt	phs001217	c1	DS-HCR-IRB	v2.p1
23	GOLDN	phs001359	c1	DS-CVD-IRB	v3.p1
24	HCHS SOL	phs001395	c1	HMB-NPU	v2.p1
25	HCHS SOL	phs001395	c2	HMB	v2.p1

26	HyperGEN	phs001293	c1	GRU-IRB	v2.p1
27	HyperGEN	phs001293	c2	DS-CVD-IRB-RD	v2.p1
28	Jackson Heart	phs000964	c1	HMB-IRB-NPU	v4.p1
29	Jackson Heart	phs000964	c2	DS-FDO-IRB-NPU	v4.p1
30	Jackson Heart	phs000964	c3	HMB-IRB	v4.p1
31	Jackson Heart	phs000964	c4	DS-FDO-IRB	v4.p1
32	MESA	phs001416	c1	HMB	v2.p1
33	MESA	phs001416	c2	HMB-NPU	v2.p1
34	MGH Afib	phs001062	c1	HMB-IRB	v5.p2
35	MGH Afib	phs001062	c2	DS-AF-IRB-RD	v5.p2
36	MyLifeOurFuture Hemophilia	phs001515	c1	HMB-PUB	v2.p2
37	Novel Risk Afib Women	phs001040	c1	HMB	v5.p1
38	Partners HCBiobank	phs001024	c1	HMB	v4.p1
39	PharmHU	phs001466	c1	HMB	v1.p1
40	PharmHU	phs001466	c2	DS-SCD-RD	v1.p1
41	PharmHU	phs001466	c3	DS-SCD	v1.p1
42	SAFHS CVD	phs001215	c1	DS-DHD-IRB-PUB-MDS-RD	v3.p2
43	SARP	phs001446	c1	GRU	v1.p1
44	Vanderbilt AFablation	phs000997	c1	HMB-IRB	v4.p2
45	Vanderbilt Afib	phs001032	c1	GRU-IRB	v6.p2
46	Venous Throm	phs001402	c1	GRU	v3.p1
47	WHI	phs001237	c1	HMB-IRB	v2.p1
48	WHI	phs001237	c2	HMB-IRB-NPU	v2.p1

Supplementary Table 6. Reference and Target Markers across datasets. Number of reference and target markers for each chromosome in our benchmarking panel obtained from the 1000 Genomes Project (1KG) Dataset and UK Biobank.

Chr	1000 Genomes Project		TOPMed		UK Biobank	
	Reference	Target	Reference	Target	Reference	Target
1	5,769,087	49,977	/	/	48,448,838	53,087
2	6,095,976	52,869	/	/	50,617,221	52,599
3	4,986,824	43,376	/	/	43,235,013	44,223
4	4,878,537	40,435	/	/	39,818,105	41,193
5	4,539,890	37,574	/	/	36,802,991	38,728
6	4,317,093	45,433	/	/	35,576,415	45,844
7	4,140,924	34,937	/	/	35,081,174	35,910
8	3,888,893	32,669	/	/	33,856,820	33,628
9	3,169,328	26,871	/	/	26,935,065	29,101
10	3,499,286	31,064	/	/	29,351,162	32,610
11	3,425,446	31,148	/	/	29,825,047	33,110
12	3,335,036	30,267	/	/	28,658,800	31,413
13	2,512,948	22,288	/	/	21,123,329	22,131
14	2,294,933	19,904	/	/	19,395,462	21,192
15	2,111,611	19,197	/	/	17,802,705	20,846
16	2,366,114	20,815	/	/	19,960,418	23,773
17	2,075,523	18,196	/	/	17,334,935	22,032
18	1,965,907	18,085	/	/	16,663,631	19,243
19	1,672,929	14,481	/	/	13,181,709	18,934
20	1,647,102	15,171	17,900,635	15,365	13,780,208	17,180
21	1,004,437	8,521	/	/	7,949,357	9,785
22	1,070,401	9,283	/	/	8,255,855	10,792

Supplementary Table 7. Phenotypic traits used in GWAS analysis. QT = quantitative trait, Disease = binary trait.

Trait	Code	Type
Age at menopause	AAM	QT
Apolipoprotein A1	APOEA	QT
Apolipoprotein B	APOEB	QT
Body mass index	BMI	QT
Calcium	CAL	QT
Docosahexaenoic acid	DOA	QT
Glycated hemoglobin	HBA1C	QT
High density lipoprotein cholesterol	HDL	QT
Height	HEIGHT	QT
Intraocular pressure	IOP	QT
Low density lipoprotein cholesterol	LDL	QT
Omega-6 fatty acids	OSFA	QT
Omega-3 fatty acids	OTFA	QT
Phosphatidylcholines	PDCL	QT
Polyunsaturated fatty acids	PFA	QT
Phosphoglycerides	PHG	QT
Resting heart rate	RHR	QT
Remnant cholesterol (Non-HDL and Non-LDL cholesterol)	RMNC	QT
Sphingomyelins	SGM	QT
Total cholesterol	TCH	QT
Total fatty acids	TFA	QT
Total triglycerides	TTG	QT
Alzheimer's disease	AD	Disease
Atrial fibrillation	AF	Disease
Age-related macular degeneration	AMD	Disease
Asthma	AST	Disease
Breast cancer	BC	Disease

Bipolar disorder	BD	Disease
Coronary artery disease	CAD	Disease
Crohn's disease	CD	Disease
Coeliac disease	CED	Disease
Bowel cancer	CRC	Disease
Cardiovascular disease	CVD	Disease
Epithelial ovarian cancer	EOC	Disease
Hypertension	HT	Disease
Ischaemic stroke	ISS	Disease
Melanoma	MEL	Disease
Multiple sclerosis	MS	Disease
Osteoporosis	OP	Disease
Prostate cancer	PC	Disease
Parkinson's disease	PD	Disease
Primary open angle glaucoma	POAG	Disease
Psoriasis	PSO	Disease
Rheumatoid arthritis	RA	Disease
Schizophrenia	SCZ	Disease
Systemic lupus erythematosus	SLE	Disease
Type 1 diabetes	T1D	Disease
Type 2 diabetes	T2D	Disease
Ulcerative colitis	UC	Disease
Venous thromboembolic disease	VTE	Disease

Supplementary Table 8. Studies used for GWAS meta-analysis for each phenotype in PRS analysis.

Trait	Studies
Height	Yengo et al. 2022
Calcium	Dennis et al. 2021 , Sakaue et al. 2021
Atrial fibrillation	Christophersen et al. 2017 , Sakaue et al. 2021 , Kurki et al. 2023
Asthma	Demenais et al. 2018 , Sakaue et al. 2021 , Kurki et al. 2023
Apolipoprotein B	Kettunen et al. 2016 , Sun et al. 2018 , Gudjonsson et al. 2022 , Thareja et al. 2023
Hypertension	Wojcik et al. 2019 , Kurki et al. 2023
Type 2 diabetes	Wojcik et al. 2019 , Vujkovic et al. 2020 , Kurki et al. 2023