

# Supplementary Figures and Tables: Leveraging polygenic functional enrichment to improve GWAS power

Gleb Kichaev<sup>1,†</sup>, Gaurav Bhatia<sup>2</sup>, Po-Ru Loh<sup>2,3</sup>, Steven Gazal<sup>2,3</sup>, Kathryn Burch<sup>1</sup>, Malika Freund<sup>4</sup>, Armin Schoech<sup>2,3</sup>, Bogdan Pasaniuc<sup>1,4,5,†,\*</sup>, and Alkes L Price<sup>2,3,6,†,\*</sup>

<sup>1</sup>Interdepartmental Program in Bioinformatics, University of California, Los Angeles, California 90095, USA.

<sup>2</sup>Dept. of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, Massachusetts 02115, USA

<sup>3</sup>Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts 02142, USA

<sup>4</sup>Dept. of Human Genetics, University of California, Los Angeles, California 90095, USA.

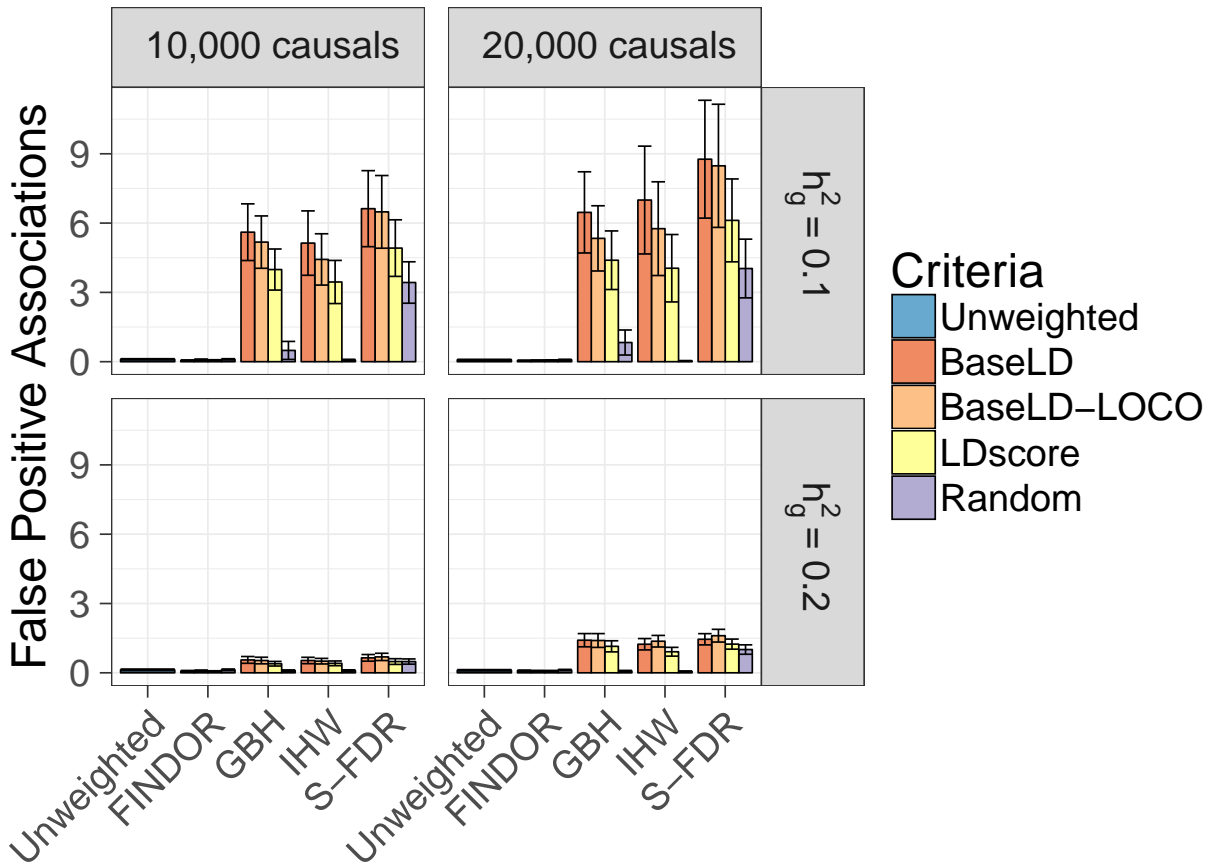
<sup>5</sup>Dept. Pathology and Laboratory Medicine, University of California, Los Angeles, California 90095, USA

<sup>6</sup>Dept. of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, Massachusetts 02115, USA

\*These authors jointly supervised this work

†Correspondence should be addressed to G.K. (gkichaev@ucla.edu), B.P. (pasaniuc@ucla.edu) or A.L.P. (aprice@hsph.harvard.edu).

## Figures



**Supplementary Figure 1: Null mis-calibration for GBH, IHW and S-FDR is worse at lower effective sample size (50K).** We report the average number of independent, genome-wide significant ( $p < 5 \times 10^{-8}$ ) associations on null chromosomes in simulations with 50K individuals (vs. 100K in Figure 1). Results are averaged across 500 simulations. Error bars represent 95% confidence intervals.

# Tables

Annotation	Proportion of SNPs	$h_g^2$ Enrichment	Tau
All SNPs	1	NaN	-1.011
Coding	0.014	4.634	1.248
Coding + 500bp	0.064	1.523	-1.303
Conserved (GERP NS)	-	1.801	0.3244
Conserved (GERP RS $\geq 4$ )	0.8	12.99	7.121
Conserved (Lindblad-Toh)	0.026	9.353	6.166
Conserved (Lindblad-Toh) + 500bp	0.33	1.669	-0.3263
CTCF	0.024	0.3465	-1.074
CTCF + 500bp	0.071	0.7298	-0.3127
DGF	0.136	2.062	0.1708
DGF + 500bp	0.538	1.367	0.05873
DHS Peaks	0.111	2.272	0.08154
DHS	0.166	2.017	-0.5108
DHS + 500bp	0.496	1.3	-0.1541
FANTOM5 Enhancer	0.4	1.296	-1.391
FANTOM5 Enhancer + 500bp	0.019	1.723	-0.3048
Enhancer	0.042	2.724	0.6413
Enhancer + 500bp	0.09	2.113	-0.1927
Fetal DHS	0.084	2.493	0.1489
Fetal DHS + 500bp	0.283	1.581	-0.1934
H3K27ac (Hnisz)	0.389	1.526	-0.4566
H3K27ac (Hnisz) + 500bp	0.42	1.534	0.5073
H3K27ac (PGC2)	0.269	1.716	-0.546
H3K27ac (PGC2) + 500bp	0.335	1.708	0.5095
H3K4me1 Peaks	0.17	2.254	0.5072
H3K4me1	0.424	1.71	0.5112
H3K4me1 + 500bp	0.606	1.338	-0.2583
H3K4me3 Peaks	0.042	2.936	0.3316
H3K4me3	0.133	2.378	0.3035
H3K4me3 + 500bp	0.255	1.756	-0.08552
H3K9ac Peaks	0.038	3.261	0.5507
H3K9ac	0.125	2.592	0.6866
H3K9ac + 500bp	0.23	1.915	-0.232
Intron	0.387	1.11	2.252
Intron + 500bp	0.397	1.177	-2.379
Promoter Flanking	0.8	1.797	-0.06156
Promoter Flanking + 500bp	0.033	1.373	-0.8093
Promoter	0.031	1.961	1.882
Promoter + 500bp	0.038	1.5	-2.017
Repressed	0.461	0.7185	0.05893
Repressed + 500bp	0.719	0.7835	0.1844
Super Enhancer (Vahedi)	0.021	2.076	2.19
Super Enhancer (Vahedi) + 500bp	0.021	2.017	-2.265
Super Enhancer (Hnisz)	0.167	1.814	-1.809
Super Enhancer (Hnisz) + 500bp	0.17	1.878	1.835
Typical Enhancer	0.022	1.698	0.995
Typical Enhancer + 500bp	0.026	1.653	-0.9213
TFBS	0.131	2.439	0.9492
TFBS + 500bp	0.341	1.499	-0.1196
Transcribed	0.346	1.173	0.309
Transcribed + 500bp	0.762	0.965	-0.09606
TSS	0.018	3.469	0.7095
TSS + 500bp	0.034	2.916	0.1618
3 UTR	0.011	2.905	0.3657
3 UTR + 500bp	0.026	1.991	-0.04583
5 UTR	0.5	3.271	0.6998
5 UTR + 500bp	0.027	1.581	-0.3904
Weak Enhancer	0.021	1.69	-0.4957
Weak Enhancer + 500bp	0.089	1.41	-0.4051
MAF bin 1	0.102	0.6522	0.5037
MAF bin 2	0.1	0.7087	0.5734
MAF bin 3	0.1	0.8438	0.7306
MAF bin 4	0.101	0.7483	0.6469
MAF bin 5	0.098	0.9875	0.9017
MAF bin 6	0.1	1.088	0.985
MAF bin 7	0.1	1.093	1.028
MAF bin 8	0.1	1.168	1.102
MAF bin 9	0.101	1.18	1.109
MAF bin 10	0.098	1.291	1.222
MAF-Adjusted Allele Age	-	NA	-0.2584
LLD-AFR	-	NA	-0.2012
Recombination Rate (10kb)	-	0.891	-0.08077
Nucleotide Diversity (10kb)	-	0.8324	-0.05295
Background Selection Statistic	-	1.238	0.6152
CpG-Content (50kb)	-	1.162	41.6

$h_g^2$	# Causals	Criteria	FINDOR	GBH	IHW	S-FDR	Unweighted
0.1	10,000	BaseLD	0.066 (0.0086)	0.4 (0.035)	0.38 (0.031)	0.47 (0.035)	NA
		BaseLD-LOCO	0.073 (0.009)	0.37 (0.033)	0.36 (0.033)	0.48 (0.035)	NA
		LDscore	0.07 (0.0087)	0.29 (0.03)	0.29 (0.024)	0.4 (0.029)	NA
		Random	0.11 (0.011)	0.072 (0.0088)	0.071 (0.0089)	0.37 (0.029)	NA
		Unweighted	NA	NA	NA	NA	0.091 (0.0099)
0.1	20,000	BaseLD	0.061 (0.013)	1.6 (0.24)	1.3 (0.18)	1.7 (0.21)	NA
		BaseLD-LOCO	0.073 (0.014)	1.6 (0.25)	1.5 (0.22)	1.8 (0.23)	NA
		LDscore	0.066 (0.012)	1.2 (0.19)	1.1 (0.16)	1.4 (0.21)	NA
		Random	0.1 (0.016)	0.086 (0.015)	0.076 (0.014)	1.2 (0.16)	NA
		Unweighted	NA	NA	NA	NA	0.1 (0.016)
0.2	10,000	BaseLD	0.06 (0.008)	0.067 (0.0084)	0.051 (0.0075)	0.066 (0.0082)	NA
		BaseLD-LOCO	0.058 (0.0079)	0.062 (0.0081)	0.058 (0.0077)	0.068 (0.0085)	NA
		LDscore	0.064 (0.0081)	0.055 (0.0078)	0.048 (0.0071)	0.059 (0.0078)	NA
		Random	0.093 (0.0099)	0.081 (0.0092)	0.081 (0.0092)	0.096 (0.01)	NA
		Unweighted	NA	NA	NA	NA	0.09 (0.0096)
0.2	20,000	BaseLD	0.059 (0.0076)	0.081 (0.0093)	0.064 (0.0083)	0.09 (0.01)	NA
		BaseLD-LOCO	0.063 (0.008)	0.081 (0.0091)	0.075 (0.0093)	0.093 (0.01)	NA
		LDscore	0.055 (0.0073)	0.061 (0.008)	0.042 (0.0065)	0.069 (0.0084)	NA
		Random	0.097 (0.0098)	0.087 (0.0092)	0.089 (0.0094)	0.12 (0.011)	NA
		Unweighted	NA	NA	NA	NA	0.098 (0.0098)

**Supplementary Table 2: Numerical results for simulations of null loci (Figure 1).** We report the average number of independent, genome-wide significant ( $p < 5 \times 10^{-8}$ ) associations on null chromosomes. Results are averaged across 1000 simulations. Standard errors are reported in parentheses.

$h_g^2$	# Causals	log10 threshold	BaseLD	BaseLD (LOCO)	LDscore	Random	Unweighted
0.1	10000	-8	1.07 (0.79)	0.96 (0.74)	0.89 (0.68)	0.52 (0.32)	0.51 (0.32)
		-7	3.9 (1.14)	3.99 (1.05)	3.66 (1.03)	3.45 (0.8)	3.49 (0.82)
		-6	27.77 (1.77)	28.27 (1.72)	25.91 (1.61)	26.16 (1.34)	26.27 (1.35)
		-5	257.06 (4.68)	260.59 (4.47)	246.15 (4.18)	243.82 (2.94)	244.22 (2.98)
		-4	2497.25 (16.01)	2534.26 (15.07)	2471.05 (15.68)	2425.97 (9.73)	2421.79 (9.93)
0.1	20000	-8	0.23 (0.08)	0.26 (0.08)	0.29 (0.09)	0.29 (0.08)	0.26 (0.07)
		-7	2.22 (0.35)	2.12 (0.31)	2.02 (0.31)	2.48 (0.28)	2.45 (0.28)
		-6	25.41 (1.73)	25.34 (1.67)	23.31 (1.43)	23.86 (1.07)	24.01 (1.1)
		-5	253.37 (6.05)	258.33 (6)	239.84 (5.62)	241.9 (3.75)	241.82 (3.8)
		-4	2505.8 (25.29)	2552.36 (24.37)	2450.59 (23.69)	2429.73 (15.65)	2424.49 (15.96)
0.2	10000	-8	0.57 (0.25)	0.59 (0.26)	0.51 (0.26)	0.33 (0.09)	0.31 (0.08)
		-7	2.89 (0.4)	2.81 (0.4)	2.57 (0.39)	2.67 (0.35)	2.65 (0.35)
		-6	27.3 (1.85)	27.16 (1.64)	26.69 (1.92)	24.99 (0.81)	25.05 (0.82)
		-5	253.01 (4.7)	258.05 (4.63)	260.4 (5)	247.23 (2.89)	247.72 (2.93)
		-4	2433.43 (14.02)	2485.62 (13.81)	2467.72 (14.75)	2444.87 (10.06)	2439.6 (10.15)
0.2	20000	-8	0.24 (0.06)	0.25 (0.06)	0.2 (0.06)	0.22 (0.05)	0.22 (0.05)
		-7	2.13 (0.23)	2.25 (0.24)	2.15 (0.2)	2.29 (0.17)	2.31 (0.18)
		-6	24.09 (1.01)	24.73 (0.99)	24.37 (1.03)	24.28 (0.69)	24.36 (0.71)
		-5	241.74 (3.9)	249.05 (3.76)	246.46 (4.15)	245.44 (2.67)	246.35 (2.71)
		-4	2414.67 (14.36)	2468.78 (13.93)	2468.45 (14.56)	2431.91 (9.74)	2426.62 (9.81)

**Supplementary Table 3: FINDOR is well-calibrated at less stringent significance thresholds in simulations of null loci.** We report the average *total number of associated SNPs* on null chromosomes at various significance thresholds. (In contrast to our main simulations, we do not report the average number of independent associations, due to problems with clumping using PLINK at less significant thresholds.) Results are averaged across 1000 simulations. Standard errors are reported in parentheses.

$h_g^2$	# Causals	Criteria	FINDOR	Unweighted	% Improve
0.1	10,000	Baseline	8.92 (0.1)	NA	20.00
		BaseLD	9.15 (0.1)	NA	23.00
		LDscore	7.95 (0.095)	NA	6.90
		Random	7.59 (0.092)	NA	2.00
		Unweighted	NA	7.44 (0.092)	0
0.1	20,000	Baseline	5.28 (0.12)	NA	34.00
		BaseLD	5.45 (0.12)	NA	38.00
		LDscore	4.5 (0.11)	NA	14.00
		Random	4.02 (0.1)	NA	2.00
		Unweighted	NA	3.94 (0.1)	0
0.2	10,000	Baseline	55.4 (0.23)	NA	6.70
		BaseLD	56.4 (0.23)	NA	8.70
		LDscore	52.2 (0.23)	NA	0.58
		Random	52.3 (0.23)	NA	0.77
		Unweighted	NA	51.9 (0.23)	0
0.2	20,000	Baseline	31.5 (0.16)	NA	16.00
		BaseLD	32.3 (0.17)	NA	19.00
		LDscore	29.1 (0.16)	NA	7.00
		Random	27.4 (0.16)	NA	0.74
		Unweighted	NA	27.2 (0.15)	0

**Supplementary Table 4: Numerical results for simulations of causal loci (Figure 2).** We report the average number of independent, genome-wide significant ( $p < 5 \times 10^{-8}$ ) associations on causal chromosomes. Results are averaged across 1000 simulations. Standard errors are reported in parentheses.

Trait	Baseline	BaseLD	LDscore	Random	Unweighted
Eosinophil Count	198	200	189	188	187
Mean Corpular Hemoglobin	247	248	233	237	237
Red Blood Cell Distribution Width	205	212	201	199	198
Red Blood Cell Count	201	206	191	192	192
White Blood Cell Count	158	165	148	148	148
Heel T Score	308	308	295	302	300
Balding Type I	96	100	92	96	96
Body Mass Index	126	132	119	117	117
Height	685	690	668	675	674
Waist-hip Ratio	102	104	100	99	98
Systolic Blood Pressure	105	106	98	96	98
Years of Education	19	24	18	17	17
Smoking Status	22	24	21	19	18
Auto Immune Traits	15	18	14	14	14
Eczema	43	46	39	34	35
Cardiovascular Diseases	47	49	39	41	38
Hypothyroidism	27	30	27	27	27
Respiratory and Disease	26	29	24	25	24
Type 2 Diabetes	16	14	13	15	14
FEV1-FVC Ratio	178	185	172	174	174
Forced Vital Capacity (FVC)	99	99	94	91	90
Neuroticism	15	16	10	11	11
Morning Person	12	14	13	14	14
Hair Color	142	143	139	139	140
Sunburn Occasion	24	25	23	22	23
Age at Menarche	56	56	54	52	52
Age at Menopause	17	18	17	17	18
Total	3189	3261	3051	3061	3054
Difference	135	207	-3	7	0
Jackknife SE	17.22	20.41	14.56	5.88	0

**Supplementary Table 5: Results for FINDOR with different stratification criteria in the 145K UK Biobank release.** For each trait, we report the number of independent, genome-wide significant loci identified by the Unweighted approach and by FINDOR with various stratification criteria in the 145K UK Biobank release.

**Supplementary Table 6: List of independent, genome-wide significant loci for all 27 traits in 145K and 460K UK Biobank releases.** We report independent, genome-wide significant loci for both Unweighted and FINDOR. See Excel file.



Class	145K				459K			
	# Loci Unweighted	# Loci FINDOR	Overall % Increase	Average % Increase	# Loci Unweighted	# Loci FINDOR	Overall % Increase.	Average % Increase
Anthropometric	1285	1334	4%	6%	5227	5353	2%	4%
Blood Cell	962	1031	7%	8%	3669	3831	4%	4%
Disease	152	186	22%	20%	860	946	10%	10%
Other	655	710	8%	15%	3527	3736	6%	7%
Overall	3054	3261	7%	13%	13283	13866	4%	7%

**Supplementary Table 7: Results for each phenotype class in 145K and 459K UK Biobank releases.** For each phenotype class, we report the number of independent, genome-wide significant loci identified by the Unweighted approach and by FINDOR in the 145K and 459K UK Biobank releases.

Trait	Baseline	BaseLD	LDscore	Random	Unweighted
Eosinophil Count	166	164	157	159	159
Mean Corpular Hemoglobin	209	210	200	203	203
Red Blood Cell Distribution Width	165	170	159	160	160
Red Blood Cell Count	164	167	153	153	154
White Blood Cell Count	119	120	112	114	112
Heel T Score	244	250	238	240	239
Balding Type I	79	83	75	75	76
Body Mass Index	79	82	76	78	78
Height	563	568	539	538	538
Waist-hip Ratio	79	75	74	71	70
Systolic Blood Pressure	75	75	73	73	71
Years of Education	10	11	10	10	10
Smoking Status	11	13	8	7	7
Auto Immune Traits	11	11	10	9	10
Eczema	27	28	27	24	24
Cardiovascular Diseases	30	30	29	28	29
Hypothyroidism	22	23	23	23	24
Respiratory and Disease	21	22	18	20	19
Type 2 Diabetes	9	9	7	10	10
FEV1-FVC Ratio	135	138	131	134	134
Forced Vital Capacity (FVC)	62	63	58	57	57
Neuroticism	5	8	5	5	5
Morning Person	6	5	5	5	5
Hair Color	117	122	119	121	121
Sunburn Occasion	20	20	18	18	18
Age at Menarche	39	39	37	39	38
Age at Menopause	14	14	13	14	14
Total	2545	2582	2438	2452	2450
Difference	95	132	-12	2	0
Jackknife SE	15.83	16.72	10.14	5.28	0

**Supplementary Table 8: Results for FINDOR with different stratification criteria with p-value threshold of  $5 \times 10^{-9}$  in the 145K UK Biobank release.** For each trait, we report the number of independent,  $p < 5 \times 10^{-9}$  loci identified by the Unweighted approach and by FINDOR with various stratification criteria in the 145K UK Biobank release.

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.0000	0.0001	-0.46	0.6468
Both Methods	0.9114	0.0039	234.92	$< 1 \times 10^{-20}$

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.0005	0.0003	-1.83	0.0685
FINDOR only	0.6613	0.0179	37.04	$< 1 \times 10^{-20}$

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.0015	0.0007	-2.30	0.0262
Unweighted Only	0.5724	0.0427	13.42	$< 1 \times 10^{-20}$

**Supplementary Table 9: Numerical results for UK Biobank replication analysis (Figure 4, left panel).** For loci detected using Both Methods, FINDOR only, or Unweighted only, respectively, we report results of a regression of standardized effect sizes ( $\frac{Z}{\sqrt{N}}$ ) at lead SNPs in UK Biobank replication data vs. UK Biobank discovery data.

Phenotype	Replication Study	Discovery N	Replication N (max)
BMI	Locke et al. (Nature 2015) <sup>2</sup>	145,209	322,091
Height	Wood et al. (Nature 2014) <sup>3</sup>	145,368	253,237
WHR adjusted BMI	Shugin et al (Nature 2015) <sup>4</sup>	145,375	210,039
Edu Years	Rietveld et al. (Science 2013) <sup>5</sup>	144,204	126,559
Eczema	Paternoster et al. (Nature Genetics 2015) <sup>6</sup>	145,416	40,835
Type II Diabetes	Morris et al. (Nature Genetics 2012) <sup>7</sup>	145,298	29,842
Ever Smoked	Furberg et al. (Nature Genetics 2010) <sup>8</sup>	145,227	74,035
Age at Menarche	Perry et al. (Nature 2014) <sup>9</sup>	74,944	182,416
Age at Menopause	Day et al (Nature Genetics 2015) <sup>10</sup>	44,410	69,360

**Supplementary Table 10: List of nine traits used for non-UK Biobank replication analysis.** We report the non-UK Biobank replication reference, UK Biobank discovery sample size and non-UK Biobank replication sample size for each trait.

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.0001	0.0003	-0.39	0.6970
Both Methods	0.6717	0.0126	53.39	$< 1 \times 10^{-20}$

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.0006	0.0015	-0.38	0.7054
FINDOR only	0.6884	0.1021	6.74	$9.52 \times 10^{-8}$

**Supplementary Table 11: Numerical results for non-UK Biobank replication analysis (Figure 4, right panel).** For loci detected using Both Methods or FINDOR only, respectively, we report results of a regression of standardized effect sizes ( $\frac{Z}{\sqrt{N}}$ ) at lead SNPs in non-UK Biobank replication data vs. UK Biobank discovery data. We do not report results for Unweighted only, which contained only a single locus.

Trait	Baseline	BaseLD	LDscore	Random	Unweighted
Eosinophil Count	710	731	686	700	699
Mean Corpular Hemoglobin	791	791	758	766	765
Red Blood Cell Distribution Width	677	674	641	651	652
Red Blood Cell Count	878	885	834	839	840
White Blood Cell Count	744	750	710	713	713
Heel T Score	1148	1149	1113	1127	1130
Balding Type I	346	346	335	335	334
Body Mass Index	930	950	913	907	908
Height	2397	2402	2354	2395	2395
Waist-hip Ratio	496	506	472	458	460
Systolic Blood Pressure	694	703	661	664	666
Years of Education	302	315	293	287	286
Smoking Status	169	178	164	154	154
Auto Immune Traits	84	86	72	75	75
Eczema	191	198	179	182	181
Cardiovascular Diseases	304	314	286	285	285
Hypothyroidism	151	153	141	140	139
Respiratory and Disease	108	109	98	105	104
Type 2 Diabetes	87	86	80	78	76
FEV1-FVC Ratio	703	714	684	684	684
Forced Vital Capacity (FVC)	559	565	541	543	544
Neuroticism	143	149	136	128	128
Morning Person	161	165	159	156	156
Hair Color	433	436	427	429	428
Sunburn Occasion	82	82	74	79	78
Age at Menarche	326	338	318	318	318
Age at Menopause	89	91	85	86	85
Total	13703	13866	13214	13284	13283
Difference	420	583	-69	1	0
Jackknife SE	39.95	40.64	33.81	10.02	0

**Supplementary Table 12: Results for FINDOR with different stratification criteria in the 459K UK Biobank release.** For each trait, we report the number of independent, genome-wide significant loci identified by the Unweighted approach and by FINDOR with various stratification criteria in the 459K UK Biobank release.

		% of lead SNPs within 95% credible set		
Data	Method	GTE <sub>x</sub> _GE	BLUEPRINT_GE	BLUEPRINT_H3K27ac
145K	FINDOR-only	17.6% (2.4%)	10.2% (1.9%)	21.2% (2.6%)
	Unweighted-only	0.0% (0.0%)	0.0% (0.0%)	2.0% (1.9%)
459K	FINDOR-only	13.6% (1.2%)	7.1% (0.9%)	23.4% (1.5%)
	Unweighted-only	5.4% (1.7%)	3.0% (1.3%)	10.7% (2.4%)

		Average Posterior Probability		
Data	Method	GTE <sub>x</sub> _GE	BLUEPRINT_GE	BLUEPRINT_H3K27ac
145K	FINDOR-only	0.052 (0.012)	0.029 (0.009)	0.039 (0.009)
	Unweighted-only	0 (0)	0 (0)	0.015 (0.015)
459K	FINDOR-only	0.037 (0.006)	0.025 (0.005)	0.055 (0.006)
	Unweighted-only	0.025 (0.011)	0.008 (0.006)	0.032 (0.011)

**Supplementary Table 13: Novel loci identified by FINDOR are more likely to be molecular QTL.** Top panel: for lead SNPs at loci detected using FINDOR only or Unweighted only, in both 145K and 459K UK Biobank releases, we report the % of lead SNPs that lie inside 95% causal sets for three molecular QTL, as described in ref.<sup>11</sup>. Bottom panel: for lead SNPs at loci detected using FINDOR only or Unweighted only, in both 145K and 459K UK Biobank releases, we report the average causal posterior probabilities for three molecular QTL, as described in ref.<sup>11</sup>.

## References

1. Steven Gazal, Hilary K Finucane, Nicholas A Furlotte, Po-Ru Loh, Pier Francesco Palamara, Xuanyao Liu, Armin Schoech, Brendan Bulik-Sullivan, Benjamin M Neale, Alexander Gusev, et al. Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nature Genetics*, 49(10):1421–1427, 2017.
2. Adam E Locke, Bratati Kahali, Sonja I Berndt, Anne E Justice, Tune H Pers, Felix R Day, Corey Powell, Sailaja Vedantam, Martin L Buchkovich, Jian Yang, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*, 518(7538):197–206, 2015.
3. Andrew R Wood, Tonu Esko, Jian Yang, Sailaja Vedantam, Tune H Pers, Stefan Gustafsson, Audrey Y Chu, Karol Estrada, Jian'an Luan, Zoltán Kutalik, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature genetics*, 46(11):1173–1186, 2014.
4. Dmitry Shungin, Thomas W Winkler, Damien C Croteau-Chonka, Teresa Ferreira, Adam E Locke, Reedik Mägi, Rona J Strawbridge, Tune H Pers, Krista Fischer, Anne E Justice, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature*, 518(7538):187–196, 2015.
5. Cornelius A Rietveld, Sarah E Medland, Jaime Derringer, Jian Yang, Tõnu Esko, Nicolas W Martin, Harm-Jan Westra, Konstantin Shakhbazov, Abdel Abdellaoui, Arpana Agrawal, et al. Gwas of 126,559 individuals identifies genetic variants associated with educational attainment. *science*, 340(6139):1467–1471, 2013.
6. Lavinia Paternoster, Marie Standl, Johannes Waage, Hansjörg Baurecht, Melanie Hotze, David P Strachan, John A Curtin, Klaus Bønnelykke, Chao Tian, Atsushi Takahashi, et al. Multi-ethnic genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nature genetics*, 47(12):1449, 2015.
7. AP Morris, BF Voight, TM Teslovich, T Ferreira, AV Segre, V Steinthorsdottir, RJ Strawbridge, H Khan, H Grallert, A Mahajan, I Prokopenko, HM Kang, C Dina, T Esko, RM Fraser, S Kanoni, A Kumar, V Lagou, C Langenberg, J Luan, CM Lindgren, M Muller-Nurasyid, S Pechlivanis, NW Rayner, LJ Scott, S Wiltshire, L Yengo, L Kinnunen, EJ Rossin, S Raychaudhuri, AD Johnson, AS Dimas, RJ Loos, S Vedantam, H Chen, JC Florez, C Fox, CT Liu, D Rybin, DJ Couper, WH Kao, M Li, MC Cornelis, P Kraft, Q Sun, RM van Dam, HM Stringham, PS Chines, K Fischer, P Fontanillas, OL Holmen, SE Hunt, AU Jackson, A Kong, R Lawrence, J Meyer, JR Perry, CG Platou, S Potter, E Rehnberg, N Robertson, S Sivapalaratnam, A Stancakova, K Stirrups, G Thorleifsson, E Tikkanen, AR Wood, P Almgren, M Atalay, R Benediktsson, LL Bonnycastle, N Burt, J Carey, G Charpentier, AT Crenshaw, AS Doney, M Dorkhan, S Edkins, V Emilsson, E Eury, T Forsen, K Gertow, B Gigante, GB Grant, CJ Groves, C Guiducci, C Herder, AB Hreidarsson, J Hui, A James, A Jonsson, W Rathmann, N Klopp, J Kravic, K Krjutskov, C Langford, K Leander, E Lindholm, S Lobbens, S Mannisto, G Mirza, TW Muhleisen, B Musk, M Parkin, L Rallidis, J Saramies, B Sennblad, S Shah, G Sigurethsson, A Silveira, G Steinbach, B Thorand, J Trakalo, F Veglia, R Wennauer, W Winckler, D Zabanah, H Campbell, C van Duijn, AG Uitterlinden, A Hofman, E Sijbrands, GR Abecasis, KR Owen, E Zeggini, MD Trip, NG Forouhi, AC Syvanen, JG Eriksson, L Peltonen, MM Nothen, B Balkau, CN Palmer, V Lyssenko, T Tuomi, B Isomaa, DJ Hunter, L Qi, AR Shuldiner, M Roden, I Barroso, T Wilsgaard, J Beilby, K Hovingh, JF Price, JF Wilson, R Rauramaa, TA Lakka, L Lind, G Dedoussis, I Njolstad, NL Pedersen, KT Khaw, NJ Wareham, SM Keinanen-Kiukkaanniemi, TE Saaristo, E Korpi-Hyovalti, J Saltevo, M Laakso, J Kuusisto, A Metspalu, FS Collins, KL Mohlke, RN Bergman, J Tuomilehto, BO Boehm, C Gieger, K Hveem, S Cauchi, P Froguel, D Baldassarre, E Tremoli, SE Humphries, D Saleheen, J Danesh, E Ingelsson, S Ripatti, V Salomaa, R Erbel, KH Jockel, S Moebus, A Peters, T Illig, U de Faire, A Hamsten, AD Morris, PJ Donnelly, TM Frayling, AT Hattersley, E Boerwinkle, O Melander, S Kathiresan, PM Nilsson, P Deloukas, U Thorsteinsdottir, LC Groop, K Stefansson, F Hu, JS Pankow, J Dupuis, JB Meigs, D Altshuler, M Boehnke, and MI McCarthy. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet*, 44(9):981–990, 2012.
8. Tobacco, Genetics Consortium, et al. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nature genetics*, 42(5):441–447, 2010.
9. John RB Perry, Felix Day, Cathy E Elks, Patrick Sulem, Deborah J Thompson, Teresa Ferreira, Chunyan He, Daniel I Chasman, Tõnu Esko, Gudmar Thorleifsson, et al. Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature*, 514(7520):92–97, 2014.
10. Felix R Day, Katherine S Ruth, Deborah J Thompson, Kathryn L Lunetta, Natalia Pervjakova, Daniel I Chasman, Lisette Stolk, Hilary K Finucane, Patrick Sulem, Brendan Bulik-Sullivan, et al. Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and brca1-mediated dna repair. *Nature genetics*, 47(11):1294–1303, 2015.
11. Farhad Hormozdiani, Steven Gazal, Bryce van de Geijn, Hilary Finucane, Chelsea J-T Ju, Po-Ru Loh, Armin Schoech, Yakir Reshef, Xuanyao Liu, Luke O'Connor, et al. Leveraging molecular qtl to understand the genetic architecture of diseases and complex traits. *bioRxiv*, page 203380, 2017.