

# AMP-DCC Data Analysis Report

## METSIM

### Phase 2

07/10/2019 (13:05)

Prepared by: Ryan Koesterer and Maria Costanzo on behalf of Broad Institute

Contact: AMP-DCC Analysis Team ([amp-dcc-dat@broadinstitute.org](mailto:amp-dcc-dat@broadinstitute.org))

This document was generated using Loamstream [18] and the AMP-DCC Data Analysis Pipeline [19]

## Contents

<b>1</b>	<b>Data</b>	<b>3</b>
<b>2</b>	<b>Strategy</b>	<b>5</b>
2.1	Sample structure and pipeline . . . . .	5
2.1.1	Cohort-level analysis . . . . .	5
2.1.2	Merged results . . . . .	5
2.2	Ancestry Adjustment and Outlier Removal . . . . .	5
<b>3</b>	<b>Serum Creatinine (SERUM_CREATININE)</b>	<b>7</b>
3.1	Summary . . . . .	7
3.2	Calibration . . . . .	8
3.3	Top associations . . . . .	9
3.4	Previously identified risk loci . . . . .	9
<b>4</b>	<b>Systolic Blood Pressure (SBP15)</b>	<b>10</b>
4.1	Summary . . . . .	10
4.2	Calibration . . . . .	11
4.3	Top associations . . . . .	12
4.4	Previously identified risk loci . . . . .	12

<b>5</b>	<b>HDL Cholesterol (HDL)</b>	<b>14</b>
5.1	Summary . . . . .	14
5.2	Calibration . . . . .	15
5.3	Top associations . . . . .	16
5.4	Previously identified risk loci . . . . .	17
<b>6</b>	<b>Estimated Glomerular Filtration Rate (eGFR)</b>	<b>19</b>
6.1	Summary . . . . .	19
6.2	Calibration . . . . .	20
6.3	Top associations . . . . .	21
6.4	Previously identified risk loci . . . . .	21
<b>7</b>	<b>Diastolic Blood Pressure (DBP10)</b>	<b>23</b>
7.1	Summary . . . . .	23
7.2	Calibration . . . . .	24
7.3	Top associations . . . . .	25
7.4	Previously identified risk loci . . . . .	25
<b>8</b>	<b>LDL Cholesterol (LDL_DIRECT)</b>	<b>27</b>
8.1	Summary . . . . .	27
8.2	Calibration . . . . .	28
8.3	Top associations . . . . .	29
8.4	Previously identified risk loci . . . . .	31
<b>9</b>	<b>Body Mass Index (BMI)</b>	<b>33</b>
9.1	Summary . . . . .	33
9.2	Calibration . . . . .	34
9.3	Top associations . . . . .	35
9.4	Previously identified risk loci . . . . .	35
<b>10</b>	<b>Acknowledgements</b>	<b>37</b>
<b>11</b>	<b>References</b>	<b>38</b>

# 1 Data

In order to run the data we received through our analysis pipeline in an efficient manner, the genotype arrays were each given a short code name; EX, and OMNI. In Table 1, we list the corresponding filename of the data set we received, the format of the file set (*note: 'bfile' refers to binary Plink format [1]*), and a liftOver [2] chain file if it was required to remap the variants to GRCh37 / hg19 coordinates

See Figures 1 and 2 for intersection counts of samples and variants available for analysis. The counts for each genotype array have been broken down by inferred ancestry as well.

Table 1: Genotype array information

ID	Filename	Format	LiftOver
EX	metsim_exomechip_portal	bfile	N/A
OMNI	metsim_omniexpress_portal	bfile	N/A

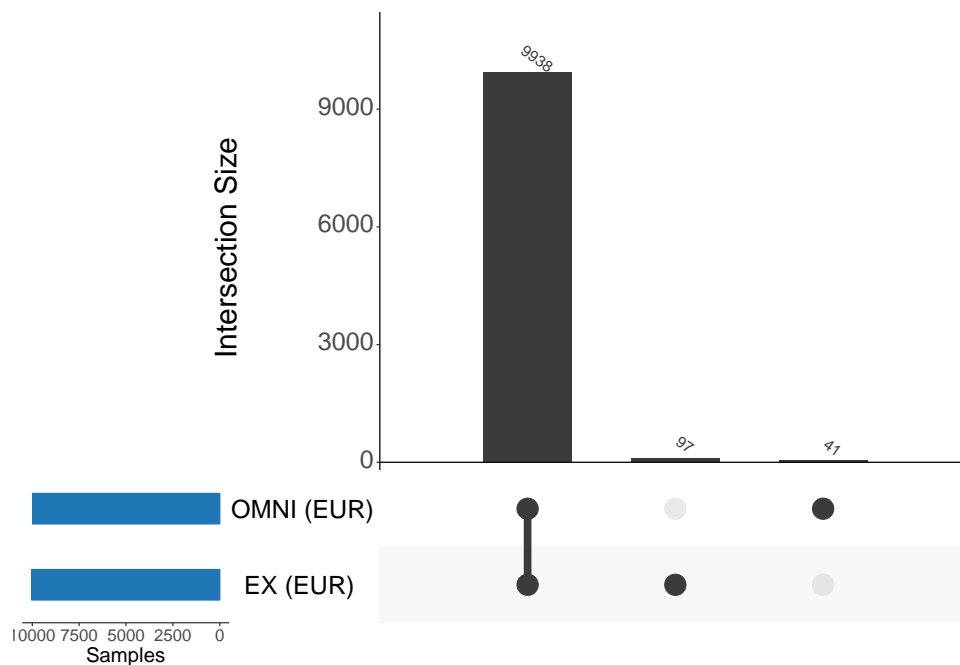


Figure 1: Samples remaining for analysis after quality control

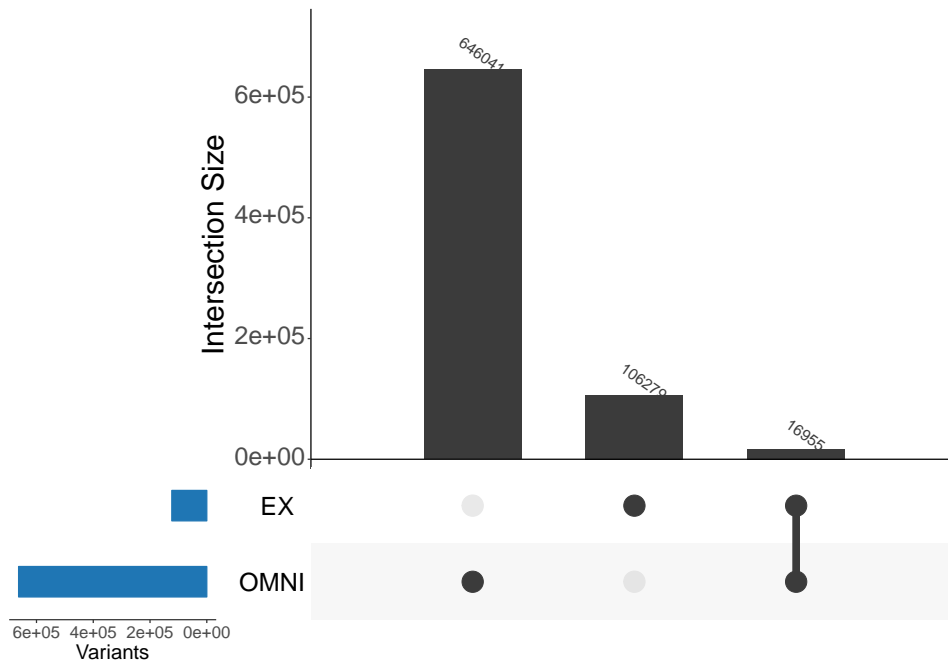


Figure 2: Variants remaining for analysis after quality control

## 2 Strategy

### 2.1 Sample structure and pipeline

The strategy we used to perform association testing can be found below. The 'ID' columns are the names used to identify each set of association test results in this document. The 'Report' columns indicate whether or not that particular set of association results will be presented in the tables and plots of the proceeding sections.

#### 2.1.1 Cohort-level analysis

In Table 2, all of the cohorts available for analysis are defined. Each cohort was defined by a single array and one or more ancestral populations.

Table 2: Cohort-level analysis

ID	Array	Ancestry	Report
EX_EUR	EX	EUR	NO
OMNI_EUR	OMNI	EUR	NO

#### 2.1.2 Merged results

In order to present results in a comprehensive way, we identified a single reference set of results as the default and merged in results from other arrays where either the variant failed to provide a  $p$ -value or did not exist in the reference set. Table 3 describes the merges performed. The '>' symbol in the 'Cohorts/Metas' column implies the strategy used to combine the results. The left-most results set was kept as reference, while variants from the following set were merged in where applicable. This merge was repeated (ie. additively) for all sets listed from left to right.

Table 3: Merged results

ID	Cohorts/Metas	Report
MERGE	EX_EUR>OMNI_EUR	YES

### 2.2 Ancestry Adjustment and Outlier Removal

Adjusting the statistical models for underlying ancestry is often crucial to reduce or eliminate Type 1 error. Often analysts include principal components of ancestry as covariates in their models as a matter of convention. In our

case, we undertook a more nuanced approach. First, the top 10 PC's were calculated for each cohort using the PC-AiR method [3]. Then, the phenotype of interest was regressed on the covariates to be used in the model and all of the PC's. If the  $N$ th PC exhibited a statistically significant  $p$ -value ( $p \leq 0.05$ ), we selected PC's 1 –  $N$  to be included in association testing. Once determined, any sample lying outside 6 standard deviations from the mean on any of the  $N$  PC's was marked as an outlier and removed from the sample set. This process was repeated up to a maximum of ten times until no outliers were found, resulting in more homogeneous sample sets for each particular analysis. For this project, a hard minimum of 0 PC's to be included in analysis was set by the analyst.

### 3 Serum Creatinine (SERUM\_CREATININE)

#### 3.1 Summary

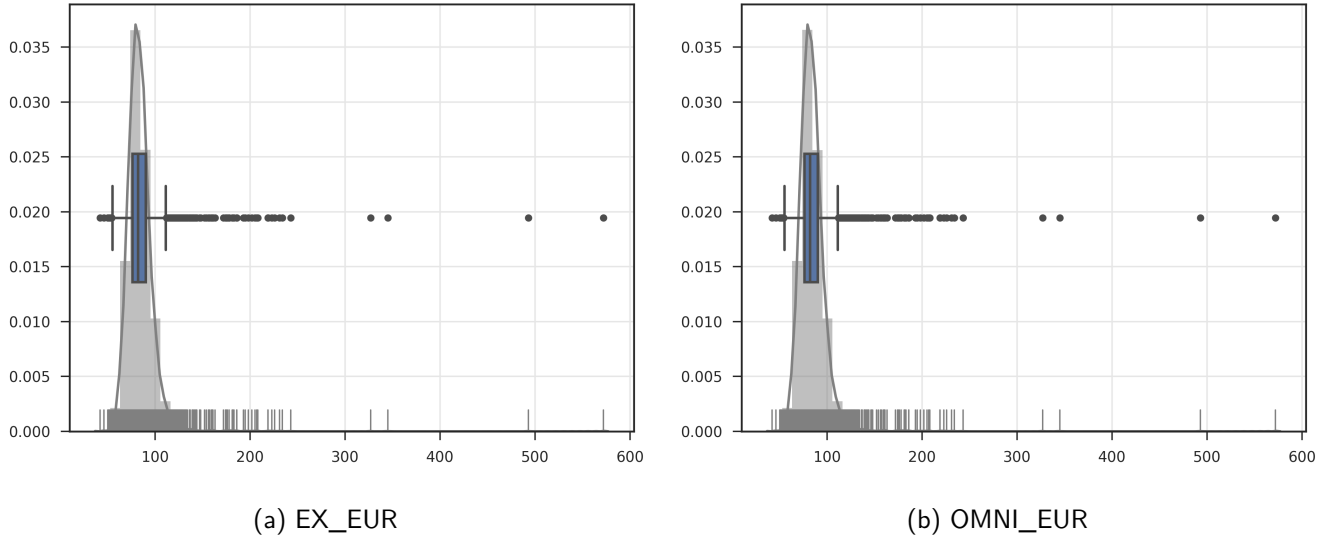


Figure 3: Distribution of SERUM\_CREATININE in cohort-level analyses

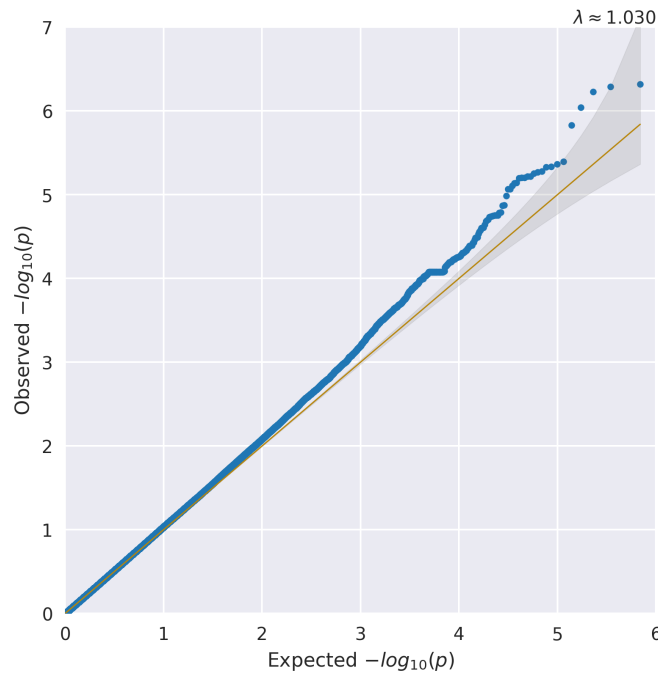
Table 4: Summary of samples removed from Serum Creatinine analysis by cohort and model

Cohort	Array	Ancestry	Trans	Covars	Total	-SampleQc	-missObs	-Kinship	-PcOutlier
EX_EUR	EX	EUR	invn	Age	10071	36	0	1518	1
OMNI_EUR	OMNI	EUR	invn	Age	10048	69	0	1497	1

Table 5: Summary of samples remaining for Serum Creatinine analysis by cohort and model

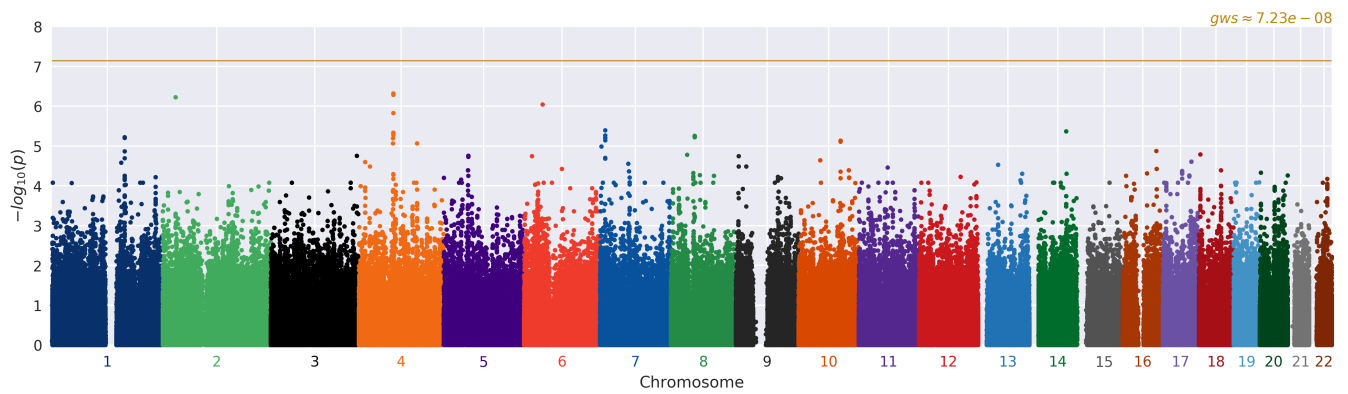
Cohort	Array	Ancestry	Trans	Covars	PCs	N	Male	Female	Max	Min	$\mu$	$\bar{x}$	$\sigma$
EX_EUR	EX	EUR	invn	Age	2	8516	8516	0	572.0	42.0	83.902	82.0	15.138
OMNI_EUR	OMNI	EUR	invn	Age	10	8481	8481	0	572.0	42.0	83.921	83.0	14.728

### 3.2 Calibration



(a) invn Adjusted Age

Figure 4: QQ plots for SERUM\_CREATININE in the MERGE analysis



(a) invn Adjusted Age

Figure 5: Manhattan plots for SERUM\_CREATININE in the MERGE analysis



### 3.3 Top associations

Table 6: Top variants in the MERGE invn Adjusted Age model (**bold** variants indicate previously identified associations)

CHR	POS	ID	EA	OA	GENE <sub>CLOSEST</sub>	COHORT	N	MALE	FEMALE	MAC	FREQ	EFFECT	STDERR	P
4	77398015	<b>rs10032549</b>	G	A	<b>SHROOM3</b>	OMNI_EUR	8,471	8,471	0	8,037	0.474	$7.74 \cdot 10^{-2}$	$1.54 \cdot 10^{-2}$	$4.8 \cdot 10^{-7}$
2	29268375	rs2276551	A	C	FAM179A	OMNI_EUR	8,479	8,479	0	1,596	$9.41 \cdot 10^{-2}$	0.13	$2.61 \cdot 10^{-2}$	$5.92 \cdot 10^{-7}$
6	42084146	rs10456507	A	C	C6orf132	OMNI_EUR	8,480	8,480	0	1,098	$6.47 \cdot 10^{-2}$	0.154	$3.14 \cdot 10^{-2}$	$9.07 \cdot 10^{-7}$
7	11581003	rs2072139	G	A	THSD7A	OMNI_EUR	8,478	8,478	0	2,566	0.151	$9.87 \cdot 10^{-2}$	$2.14 \cdot 10^{-2}$	$4.01 \cdot 10^{-6}$
14	82510077	rs2217448	T	C	SEL1L	OMNI_EUR	8,478	8,478	0	2,037	0.12	0.108	$2.36 \cdot 10^{-2}$	$4.31 \cdot 10^{-6}$
8	53355822	rs17228239	A	G	ST18	OMNI_EUR	8,480	8,480	0	8,376	0.506	$6.96 \cdot 10^{-2}$	$1.53 \cdot 10^{-2}$	$5.55 \cdot 10^{-6}$
1	163744838	rs10799961	G	A	NUF2	OMNI_EUR	8,480	8,480	0	6,725	0.603	$7.06 \cdot 10^{-2}$	$1.56 \cdot 10^{-2}$	$6.03 \cdot 10^{-6}$
4	77269262	rs10002448	A	G	CCDC158	OMNI_EUR	8,437	8,437	0	6,709	0.602	$7.11 \cdot 10^{-2}$	$1.57 \cdot 10^{-2}$	$6.32 \cdot 10^{-6}$
10	94517064	rs11187167	C	T	HHEX	OMNI_EUR	8,467	8,467	0	2,985	0.176	$9.03 \cdot 10^{-2}$	$2.01 \cdot 10^{-2}$	$7.28 \cdot 10^{-6}$
10	94538960	rs2497338	T	C	EXOC6	OMNI_EUR	8,481	8,481	0	3,664	0.784	$8.32 \cdot 10^{-2}$	$1.86 \cdot 10^{-2}$	$7.85 \cdot 10^{-6}$
4	130824678	rs3111741	G	A	C4orf33	OMNI_EUR	8,481	8,481	0	4,295	0.747	$7.8 \cdot 10^{-2}$	$1.75 \cdot 10^{-2}$	$8.53 \cdot 10^{-6}$
4	77228724	rs3796491	C	A	STBD1	OMNI_EUR	8,395	8,395	0	6,673	0.603	$6.99 \cdot 10^{-2}$	$1.57 \cdot 10^{-2}$	$8.58 \cdot 10^{-6}$
7	3060397	rs10950998	A	G	CARD11	OMNI_EUR	8,481	8,481	0	8,047	0.526	$6.79 \cdot 10^{-2}$	$1.54 \cdot 10^{-2}$	$1.03 \cdot 10^{-5}$
16	76021509	rs8055817	A	G	AC025287	OMNI_EUR	8,373	8,373	0	5,227	0.688	$7.19 \cdot 10^{-2}$	$1.65 \cdot 10^{-2}$	$1.33 \cdot 10^{-5}$
18	3655393	rs1675240	G	A	DLGAP1	OMNI_EUR	8,479	8,479	0	6,693	0.395	$6.81 \cdot 10^{-2}$	$1.58 \cdot 10^{-2}$	$1.63 \cdot 10^{-5}$
8	36995874	rs543946	T	C	KCNU1	OMNI_EUR	8,476	8,476	0	543	0.968	0.187	$4.34 \cdot 10^{-2}$	$1.65 \cdot 10^{-5}$
3	193811162	rs10933711	C	T	HES1	OMNI_EUR	8,480	8,480	0	6,161	0.363	$6.86 \cdot 10^{-2}$	$1.6 \cdot 10^{-2}$	$1.76 \cdot 10^{-5}$
5	55509576	rs16885032	C	T	ANKRD55	OMNI_EUR	8,481	8,481	0	561	$3.31 \cdot 10^{-2}$	0.184	$4.28 \cdot 10^{-2}$	$1.76 \cdot 10^{-5}$
6	18143854	rs17839843	G	A	TPMT	OMNI_EUR	8,478	8,478	0	22	$1.3 \cdot 10^{-3}$	0.914	0.213	$1.78 \cdot 10^{-5}$
9	6533092	rs138640017	G	C	GLDC	EX_EUR	8,516	8,516	0	153	$8.98 \cdot 10^{-3}$	0.349	$8.13 \cdot 10^{-2}$	$1.8 \cdot 10^{-5}$

### 3.4 Previously identified risk loci

Table 7 shows statistics from the MERGE cohort for 10 loci that were shown to be significantly associated with Serum Creatinine in the 2016 Nature Communications paper by Kettunen et al [8]. Where a previously reported variant was not genotyped in the study (indicated by  $\bar{R}^2 < 1$ ), if available, a tagging variant in LD with the reported variant ( $\bar{R}^2 \geq 0.7$  and within 250kb) was provided. Tags were identified using 1000 Genomes data. There are 9 variants that show at least nominal significance ( $p < 0.05$ ) in this study. Out of the 10 variants in both studies, 10 exhibit the same direction of effect with the known result (binomial test  $p = 0.000977$ ).

Table 7: Top known loci in MERGE model invn Adjusted Age (**bold** variants indicate matching direction of effect)

CHR	POS	ID	EA	OA	N	FREQ	EFFECT	STDERR	P	COHORT	GENE <sub>CLOSEST</sub>	R <sup>2</sup>	ID <sub>KNOWN</sub>	N <sub>KNOWN</sub>	EFFECT <sub>KNOWN</sub>	STDERR <sub>KNOWN</sub>	P <sub>KNOWN</sub>
15	45641225	<b>rs2453533</b>	A	C	8,515	0.583	$4.33 \cdot 10^{-2}$	$1.54 \cdot 10^{-2}$	$4.93 \cdot 10^{-3}$	EX_EUR	GATM	1	rs2453533	24,925	$7.37 \cdot 10^{-2}$	$9.39 \cdot 10^{-3}$	$5.03 \cdot 10^{-15}$
15	45695695	<b>rs1153849</b>	A	G	8,481	0.69	$4.3 \cdot 10^{-2}$	$1.65 \cdot 10^{-2}$	$9.2 \cdot 10^{-3}$	OMNI_EUR	SPATASL1	1	rs1153849	24,925	$8.02 \cdot 10^{-2}$	$1.03 \cdot 10^{-2}$	$9.14 \cdot 10^{-15}$
17	59456589	<b>rs9895661</b>	C	T	8,517	0.196	$7.73 \cdot 10^{-2}$	$1.94 \cdot 10^{-2}$	$6.75 \cdot 10^{-5}$	EX_EUR	BCAS3	1	rs9895661	24,925	$8.79 \cdot 10^{-2}$	$1.23 \cdot 10^{-2}$	$8.25 \cdot 10^{-13}$
15	45729123	<b>rs1974981</b>	G	A	8,479	0.679	$3.48 \cdot 10^{-2}$	$1.64 \cdot 10^{-2}$	$3.42 \cdot 10^{-2}$	OMNI_EUR	C15orf48	1	rs1974981	24,925	$7.03 \cdot 10^{-2}$	$1.03 \cdot 10^{-2}$	$9.04 \cdot 10^{-12}$
7	151415041	<b>rs10224002</b>	G	A	8,515	0.775	$6.66 \cdot 10^{-2}$	$1.84 \cdot 10^{-2}$	$2.86 \cdot 10^{-4}$	EX_EUR	PRKAG2	1	rs10224002	24,925	$7.07 \cdot 10^{-2}$	$1.11 \cdot 10^{-2}$	$1.89 \cdot 10^{-10}$
15	45577783	<b>rs1719236</b>	C	A	8,480	0.403	$2.42 \cdot 10^{-2}$	$1.56 \cdot 10^{-2}$	0.12	OMNI_EUR	SLC28A2	1	rs1719236	24,925	$5.86 \cdot 10^{-2}$	$9.54 \cdot 10^{-3}$	$8.74 \cdot 10^{-10}$
4	77368847	<b>rs17319721</b>	A	G	8,517	0.599	$7.07 \cdot 10^{-2}$	$1.56 \cdot 10^{-2}$	$6.17 \cdot 10^{-6}$	EX_EUR	SHROOM3	1	rs17319721	24,925	$5.71 \cdot 10^{-2}$	$9.31 \cdot 10^{-3}$	$9.1 \cdot 10^{-10}$
6	160693107	<b>rs3119311</b>	C	T	8,479	0.893	$5.14 \cdot 10^{-2}$	$2.48 \cdot 10^{-2}$	$3.81 \cdot 10^{-2}$	OMNI_EUR	SLC22A2	1	rs3119311	24,925	$9.08 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$	$5.29 \cdot 10^{-9}$
17	59483766	<b>rs8068318</b>	C	T	8,517	0.274	$4.59 \cdot 10^{-2}$	$1.72 \cdot 10^{-2}$	$7.52 \cdot 10^{-3}$	EX_EUR	TBX2	1	rs8068318	24,925	$6.21 \cdot 10^{-2}$	$1.08 \cdot 10^{-2}$	$8.69 \cdot 10^{-9}$
15	45801035	<b>rs950027</b>	C	T	8,475	0.521	$3.24 \cdot 10^{-2}$	$1.54 \cdot 10^{-2}$	$3.49 \cdot 10^{-2}$	OMNI_EUR	SLC30A4	1	rs950027	24,925	$5.28 \cdot 10^{-2}$	$9.32 \cdot 10^{-3}$	$1.59 \cdot 10^{-8}$

## 4 Systolic Blood Pressure (SBP15)

### 4.1 Summary

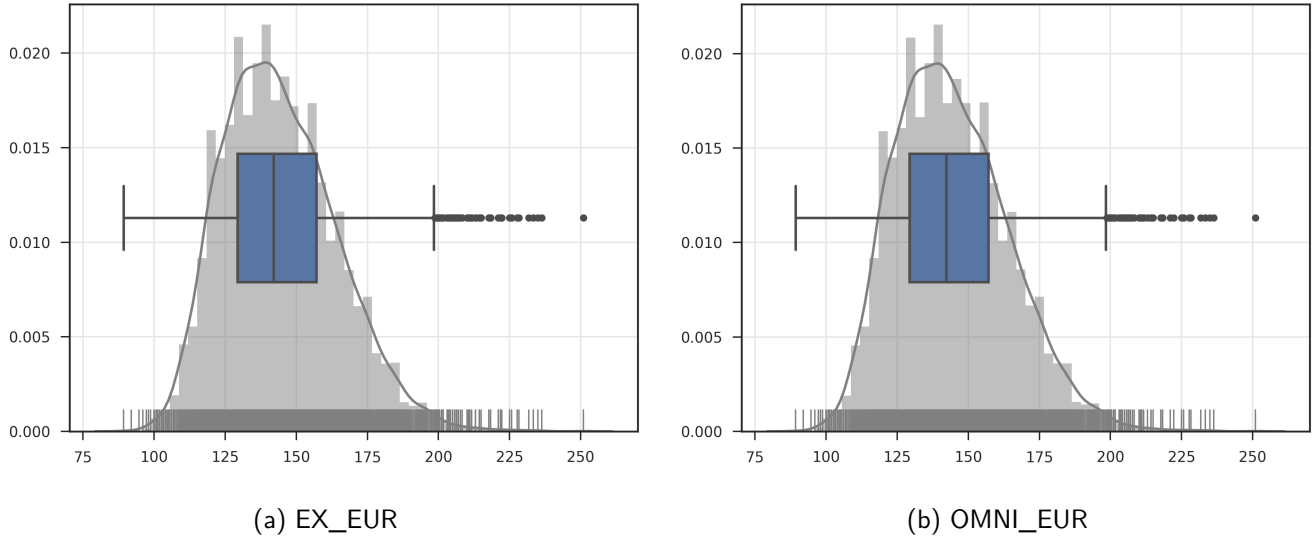


Figure 6: Distribution of SBP15 in cohort-level analyses

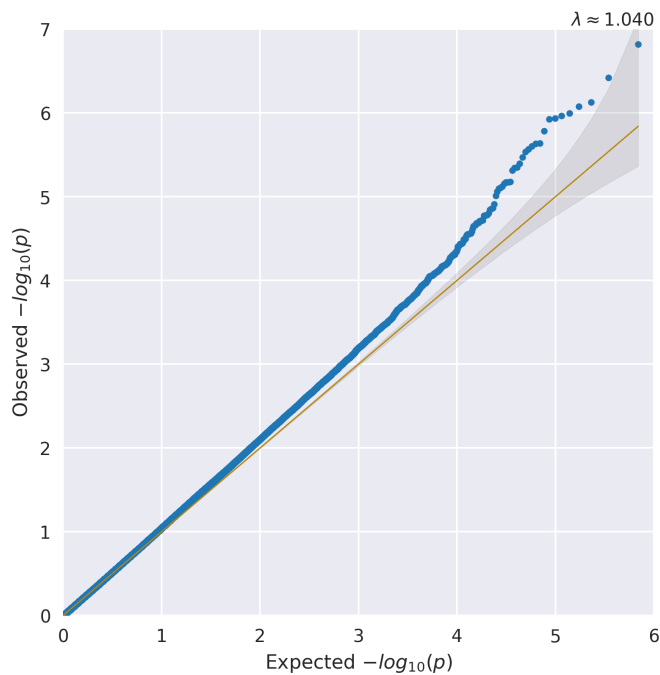
Table 8: Summary of samples removed from Systolic Blood Pressure analysis by cohort and model

Cohort	Array	Ancestry	Trans	Covars	Total	-SampleQc	-missObs	-Kinship	-PcOutlier
EX_EUR	EX	EUR	invn	Age+Age2+BMI	10071	36	5	1514	2
OMNI_EUR	OMNI	EUR	invn	Age+Age2+BMI	10048	69	5	1494	1

Table 9: Summary of samples remaining for Systolic Blood Pressure analysis by cohort and model

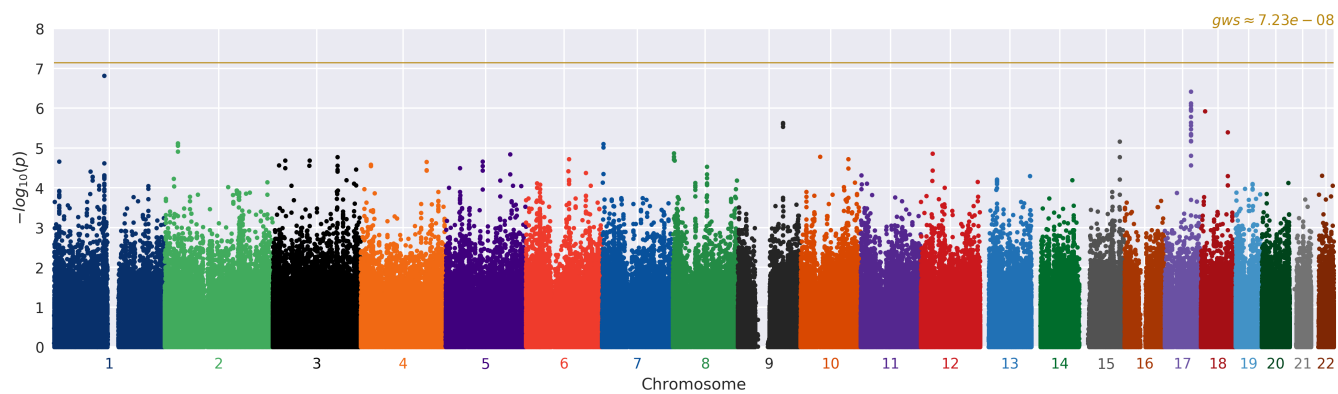
Cohort	Array	Ancestry	Trans	Covars	PCs	N	Male	Female	Max	Min	$\mu$	$\bar{x}$	$\sigma$
EX_EUR	EX	EUR	invn	Age+Age2+BMI	10	8514	8514	0	251.0	89.3	144.242	142.3	19.909
OMNI_EUR	OMNI	EUR	invn	Age+Age2+BMI	8	8479	8479	0	251.0	89.3	144.263	142.3	19.903

### 4.2 Calibration



(a) invn Adjusted Age+Age2+BMI

Figure 7: QQ plots for SBP15 in the MERGE analysis



(a) invn Adjusted Age+Age2+BMI

Figure 8: Manhattan plots for SBP15 in the MERGE analysis

### 4.3 Top associations

Table 10: Top variants in the MERGE invn Adjusted Age+Age2+BMI model (**bold** variants indicate previously identified associations)

CHR	POS	ID	EA	OA	GENE <sub>CLOSEST</sub>	COHORT	N	MALE	FEMALE	MAC	FREQ	EFFECT	STDERR	P
1	113161350	rs11102516	A	G	ST7L	OMNI_EUR	8,479	8,479	0	3,480	0.205	$9.93 \cdot 10^{-2}$	$1.89 \cdot 10^{-2}$	$1.52 \cdot 10^{-7}$
17	59255231	rs917571	A	G	BCAS3	OMNI_EUR	8,478	8,478	0	3,255	0.808	$9.86 \cdot 10^{-2}$	$1.94 \cdot 10^{-2}$	$3.81 \cdot 10^{-7}$
18	9958623	rs29092	C	A	VAPA	OMNI_EUR	8,478	8,478	0	43	$2.54 \cdot 10^{-3}$	0.742	0.153	$1.19 \cdot 10^{-6}$
9	101619403	rs7864581	G	A	GALNT12	OMNI_EUR	8,478	8,478	0	4,745	0.72	$8 \cdot 10^{-2}$	$1.69 \cdot 10^{-2}$	$2.33 \cdot 10^{-6}$
18	60661358	rs7238945	G	T	PHLPP1	OMNI_EUR	8,374	8,374	0	1,862	0.111	0.114	$2.47 \cdot 10^{-2}$	$4.03 \cdot 10^{-6}$
15	91404788	rs4932371	C	T	FURIN	OMNI_EUR	8,393	8,393	0	4,748	0.283	$7.79 \cdot 10^{-2}$	$1.73 \cdot 10^{-2}$	$6.95 \cdot 10^{-6}$
2	29575197	rs6716914	T	G	ALK	OMNI_EUR	8,475	8,475	0	1,824	0.892	0.113	$2.51 \cdot 10^{-2}$	$7.59 \cdot 10^{-6}$
7	2920067	rs17132769	C	T	CARD11	OMNI_EUR	8,479	8,479	0	2,032	0.12	0.105	$2.35 \cdot 10^{-2}$	$7.96 \cdot 10^{-6}$
8	2304723	rs13259957	G	A	MYOM2	OMNI_EUR	8,479	8,479	0	688	$4.06 \cdot 10^{-2}$	0.17	$3.91 \cdot 10^{-2}$	$1.36 \cdot 10^{-5}$
12	26344726	rs1480036	C	T	SSPN	OMNI_EUR	8,477	8,477	0	4,723	0.721	$7.44 \cdot 10^{-2}$	$1.71 \cdot 10^{-2}$	$1.4 \cdot 10^{-5}$
5	144939151	rs17103387	C	T	PRELID2	OMNI_EUR	8,470	8,470	0	1,528	$9.02 \cdot 10^{-2}$	0.117	$2.69 \cdot 10^{-2}$	$1.43 \cdot 10^{-5}$
10	43747330	rs12244413	T	G	RASGEF1A	OMNI_EUR	8,476	8,476	0	1,144	$6.75 \cdot 10^{-2}$	0.132	$3.06 \cdot 10^{-2}$	$1.65 \cdot 10^{-5}$
3	145477425	rs16857193	C	T	AC107021	OMNI_EUR	8,477	8,477	0	1,200	$7.08 \cdot 10^{-2}$	0.129	$3 \cdot 10^{-2}$	$1.68 \cdot 10^{-5}$
6	96999725	rs1127175	T	C	UFL1	OMNI_EUR	8,476	8,476	0	5,616	0.331	$7 \cdot 10^{-2}$	$1.64 \cdot 10^{-2}$	$1.91 \cdot 10^{-5}$
10	107526521	rs4437963	T	C	SORCS3	OMNI_EUR	8,477	8,477	0	2,252	0.867	$9.67 \cdot 10^{-2}$	$2.26 \cdot 10^{-2}$	$1.92 \cdot 10^{-5}$
3	82701950	rs936035	C	A	GBE1	OMNI_EUR	8,467	8,467	0	3,764	0.778	$7.81 \cdot 10^{-2}$	$1.83 \cdot 10^{-2}$	$2.05 \cdot 10^{-5}$
8	3896853	rs12677236	T	C	CSMD1	OMNI_EUR	8,475	8,475	0	3,560	0.21	$7.96 \cdot 10^{-2}$	$1.87 \cdot 10^{-2}$	$2.07 \cdot 10^{-5}$
3	27652096	rs11710219	T	C	EOMES	OMNI_EUR	8,479	8,479	0	5,783	0.659	$6.93 \cdot 10^{-2}$	$1.63 \cdot 10^{-2}$	$2.08 \cdot 10^{-5}$
5	83329010	rs1428925	C	T	EDIL3	OMNI_EUR	8,476	8,476	0	573	$3.38 \cdot 10^{-2}$	0.182	$4.28 \cdot 10^{-2}$	$2.18 \cdot 10^{-5}$
1	12194451	rs1148472	C	T	TNFRSF8	OMNI_EUR	8,475	8,475	0	5,770	0.66	$6.92 \cdot 10^{-2}$	$1.63 \cdot 10^{-2}$	$2.21 \cdot 10^{-5}$

### 4.4 Previously identified risk loci

Table 11 shows statistics from the MERGE cohort for 22 loci that were shown to be significantly associated with Systolic Blood Pressure in the 2011 Nature paper by Ehret et al [17]. Where a previously reported variant was not genotyped in the study (indicated by  $\bar{R}^2 < 1$ ), if available, a tagging variant in LD with the reported variant ( $\bar{R}^2 \geq 0.7$  and within 250kb) was provided. Tags were identified using 1000 Genomes data. There are 10 variants that show at least nominal significance ( $p < 0.05$ ) in this study. Out of the 20 variants in both studies, 11 exhibit the same direction of effect with the known result (binomial test  $p = 0.412$ ).

Table 11: Top known loci in MERGE model invn Adjusted Age+Age2+BMI (**bold** variants indicate matching direction of effect)

CHR	POS	ID	EA	OA	N	FREQ	EFFECT	STDERR	P	COHORT	GENE <sub>CLOSEST</sub>	R <sup>2</sup>	ID <sub>KNOWN</sub>	N <sub>KNOWN</sub>	EFFECT <sub>KNOWN</sub>	STDERR <sub>KNOWN</sub>	P <sub>KNOWN</sub>
12	90060586	<b>rs17249754</b>	G	A	8,512	0.923	6.46 · 10 <sup>-2</sup>	2.9 · 10 <sup>-2</sup>	2.61 · 10 <sup>-2</sup>	EX_EUR	ATP2B1	1	rs17249754	2 · 10 <sup>5</sup>	0.955	0.134	9.73 · 10 <sup>-13</sup>
1	11862778	rs17367504	A	G	8,512	0.832	7.31 · 10 <sup>-2</sup>	2.04 · 10 <sup>-2</sup>	3.46 · 10 <sup>-4</sup>	EX_EUR	MTHFR	1	rs17367504	2 · 10 <sup>5</sup>	-0.861	0.136	2.11 · 10 <sup>-10</sup>
1	11883731	rs12567136	C	T	8,479	0.832	7.65 · 10 <sup>-2</sup>	2.05 · 10 <sup>-2</sup>	1.92 · 10 <sup>-4</sup>	OMNI_EUR	CLCN6	1	rs12567136	2 · 10 <sup>5</sup>	-0.847	0.135	3.41 · 10 <sup>-10</sup>
15	75077367	<b>rs1378942</b>	C	A	8,512	0.447	5.15 · 10 <sup>-2</sup>	1.55 · 10 <sup>-2</sup>	8.89 · 10 <sup>-4</sup>	EX_EUR	CSK	1	rs1378942	2 · 10 <sup>5</sup>	0.632	0.101	3.43 · 10 <sup>-10</sup>
10	104846178	<b>rs11191548</b>	T	C	8,506	0.925	5.36 · 10 <sup>-2</sup>	2.93 · 10 <sup>-2</sup>	6.73 · 10 <sup>-2</sup>	EX_EUR	CNNM2	1	rs11191548	2 · 10 <sup>5</sup>	1.083	0.174	5.03 · 10 <sup>-10</sup>
10	104594507	rs1004467	A	G	8,512	0.91	1.6 · 10 <sup>-2</sup>	2.71 · 10 <sup>-2</sup>	0.555	EX_EUR	CYP17A1	1	rs1004467	2 · 10 <sup>5</sup>	-1.01	0.164	6.61 · 10 <sup>-10</sup>
10	104906211	<b>rs11191580</b>	T	C	8,512	0.924	5.14 · 10 <sup>-2</sup>	2.93 · 10 <sup>-2</sup>	7.91 · 10 <sup>-2</sup>	EX_EUR	NT5C2	1	rs11191580	2 · 10 <sup>5</sup>	1.058	0.173	9.16 · 10 <sup>-10</sup>
12	112007756	rs653178	C	T	8,512	0.398	3.9 · 10 <sup>-2</sup>	1.55 · 10 <sup>-2</sup>	1.19 · 10 <sup>-2</sup>	EX_EUR	ATXN2	1	rs653178	2 · 10 <sup>5</sup>	-0.605	9.88 · 10 <sup>-2</sup>	9.3 · 10 <sup>-10</sup>
12	89942390	rs11105328	A	G	8,479	0.927	5.94 · 10 <sup>-2</sup>	2.94 · 10 <sup>-2</sup>	4.33 · 10 <sup>-2</sup>	OMNI_EUR	POC1B-GALNT4	1	rs11105328	2 · 10 <sup>5</sup>	-0.838	0.137	1.08 · 10 <sup>-9</sup>
10	104660004	rs11191454	A	G	8,512	0.923	3.77 · 10 <sup>-2</sup>	2.9 · 10 <sup>-2</sup>	0.194	EX_EUR	BORCS7-ASMT	1	rs11191454	2 · 10 <sup>5</sup>	-1.043	0.171	1.12 · 10 <sup>-9</sup>
12	111884608	<b>rs3184504</b>	T	C	8,511	0.388	3.92 · 10 <sup>-2</sup>	1.56 · 10 <sup>-2</sup>	1.18 · 10 <sup>-2</sup>	EX_EUR	SH2B3	1	rs3184504	2 · 10 <sup>5</sup>	0.598	9.93 · 10 <sup>-2</sup>	1.69 · 10 <sup>-9</sup>
4	81164723	<b>rs1458038</b>	T	C	8,478	0.692	7.54 · 10 <sup>-3</sup>	1.65 · 10 <sup>-2</sup>	0.647	OMNI_EUR	FGF5	1	rs1458038	2 · 10 <sup>5</sup>	0.662	0.111	2.12 · 10 <sup>-9</sup>
11	16902268	<b>rs381815</b>	T	C	8,439	0.767	2.67 · 10 <sup>-2</sup>	1.82 · 10 <sup>-2</sup>	0.142	OMNI_EUR	PLEKHA7	1	rs381815	2 · 10 <sup>5</sup>	0.655	0.11	2.45 · 10 <sup>-9</sup>
10	104546284	<b>rs486955</b>	T	C	8,479	0.111	4.82 · 10 <sup>-3</sup>	2.47 · 10 <sup>-2</sup>	0.845	OMNI_EUR	WBP1L	1	rs486955	2 · 10 <sup>5</sup>	0.895	0.156	9.47 · 10 <sup>-9</sup>
12	112072424	<b>rs11065987</b>	G	A	8,512	0.624	3.87 · 10 <sup>-2</sup>	1.56 · 10 <sup>-2</sup>	1.34 · 10 <sup>-2</sup>	EX_EUR	BRAP	1	rs11065987	2 · 10 <sup>5</sup>	0.57	0.102	2.12 · 10 <sup>-8</sup>
15	75115895	<b>rs7162232</b>	G	A	8,478	0.332	4.48 · 10 <sup>-2</sup>	1.63 · 10 <sup>-2</sup>	6.07 · 10 <sup>-3</sup>	OMNI_EUR	LMAN1L	1	rs7162232	2 · 10 <sup>5</sup>	0.606	0.109	2.33 · 10 <sup>-8</sup>
15	75057203	rs4886406	G	T	8,477	0.283	4.63 · 10 <sup>-2</sup>	1.7 · 10 <sup>-2</sup>	6.43 · 10 <sup>-3</sup>	OMNI_EUR	CYP1A2	1	rs4886406	2 · 10 <sup>5</sup>	-0.599	0.108	3.06 · 10 <sup>-8</sup>
12	112486818	<b>rs17696736</b>	G	A	8,512	0.6	2.95 · 10 <sup>-2</sup>	1.54 · 10 <sup>-2</sup>	5.54 · 10 <sup>-2</sup>	EX_EUR	NAA25	1	rs17696736	2 · 10 <sup>5</sup>	0.549	9.96 · 10 <sup>-2</sup>	3.43 · 10 <sup>-8</sup>
12	112906415	rs11066320	A	G	8,477	0.399	2.36 · 10 <sup>-2</sup>	1.55 · 10 <sup>-2</sup>	0.127	OMNI_EUR	PTPN11	1	rs11066320	2 · 10 <sup>5</sup>	-0.544	9.96 · 10 <sup>-2</sup>	4.56 · 10 <sup>-8</sup>
10	104652323	rs11191447	C	T	8,512	0.923	3.75 · 10 <sup>-2</sup>	2.9 · 10 <sup>-2</sup>	0.197	EX_EUR	AS3MT	1	rs3740390	2 · 10 <sup>5</sup>	-1.005	0.172	4.61 · 10 <sup>-9</sup>
12	112486818	<b>rs17696736</b>	G	A	8,512	0.6	2.95 · 10 <sup>-2</sup>	1.54 · 10 <sup>-2</sup>	5.54 · 10 <sup>-2</sup>	EX_EUR	TRAFD1	0.922	rs17630235	2 · 10 <sup>5</sup>	0.569	0.1	1.45 · 10 <sup>-8</sup>
12	112486818	<b>rs17696736</b>	G	A	8,512	0.6	2.95 · 10 <sup>-2</sup>	1.54 · 10 <sup>-2</sup>	5.54 · 10 <sup>-2</sup>	EX_EUR	HECTD4	0.913	rs11066188	2 · 10 <sup>5</sup>	0.567	0.101	1.72 · 10 <sup>-8</sup>

## 5 HDL Cholesterol (HDL)

### 5.1 Summary

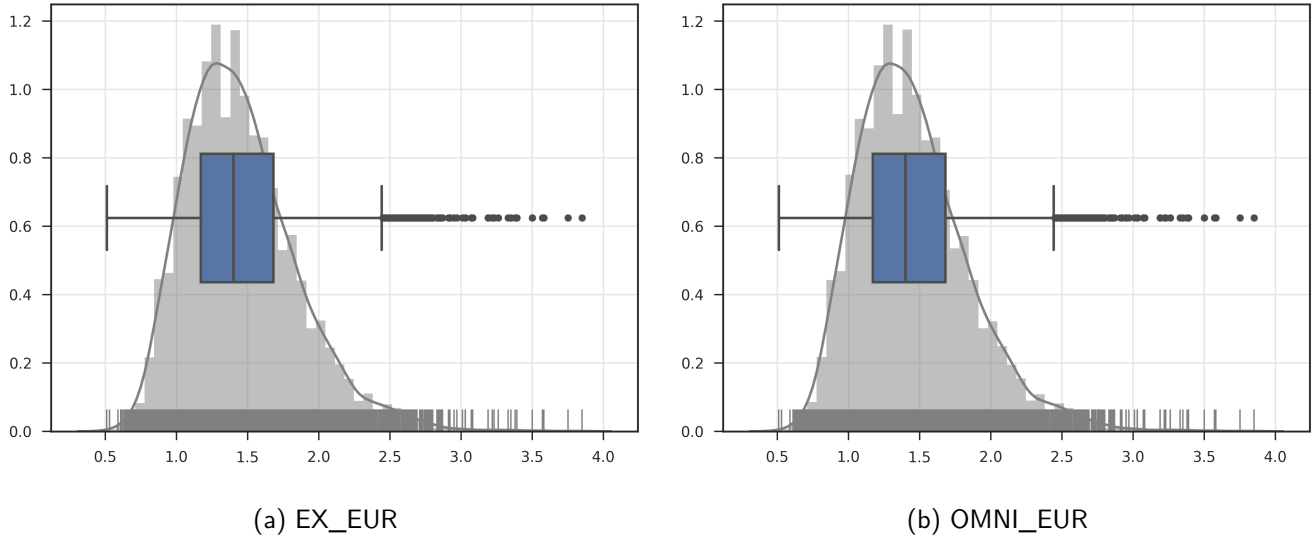


Figure 9: Distribution of HDL in cohort-level analyses

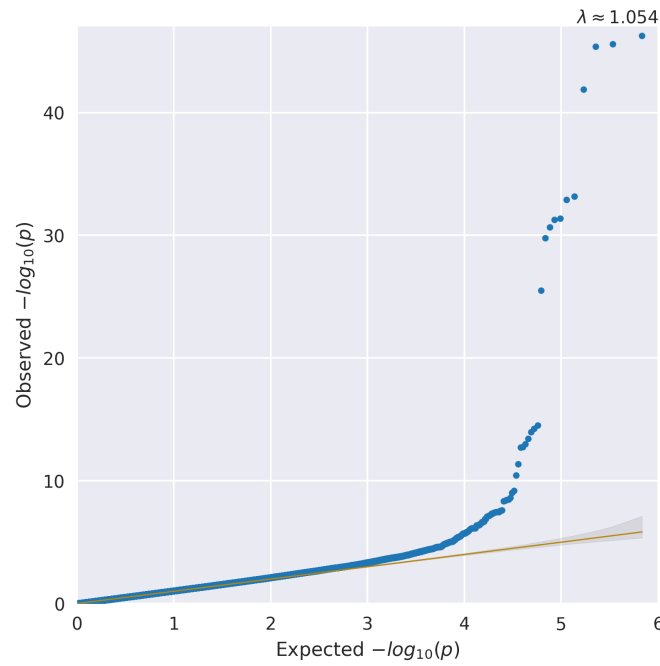
Table 12: Summary of samples removed from HDL Cholesterol analysis by cohort and model

Cohort	Array	Ancestry	Trans	Covars	Total	-SampleQc	-missObs	-Kinship	-PcOutlier
EX_EUR	EX	EUR	invn	Age+Age2+BMI	10071	36	2847	1104	2
OMNI_EUR	OMNI	EUR	invn	Age+Age2+BMI	10048	69	2840	1077	0

Table 13: Summary of samples remaining for HDL Cholesterol analysis by cohort and model

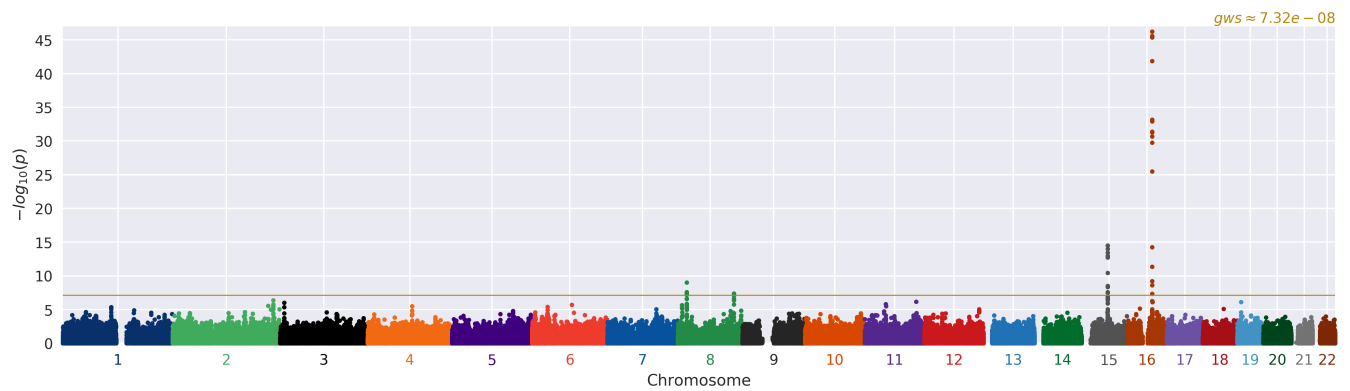
Cohort	Array	Ancestry	Trans	Covars	PCs	N	Male	Female	Max	Min	$\mu$	$\tilde{x}$	$\sigma$
EX_EUR	EX	EUR	invn	Age+Age2+BMI	1	6091	6091	0	3.85	0.51	1.452	1.4	0.396
OMNI_EUR	OMNI	EUR	invn	Age+Age2+BMI	3	6078	6078	0	3.85	0.51	1.453	1.4	0.398

### 5.2 Calibration



(a) invn Adjusted Age+Age2+BMI

Figure 10: QQ plots for HDL in the MERGE analysis



(a) invn Adjusted Age+Age2+BMI

Figure 11: Manhattan plots for HDL in the MERGE analysis

### 5.3 Top associations

Table 14: Top variants in the MERGE invn Adjusted Age+Age2+BMI model (**bold** variants indicate previously identified associations)

CHR	POS	ID	EA	OA	GENE <sub>CLOSEST</sub>	COHORT	N	MALE	FEMALE	MAC	FREQ	EFFECT	STDERR	P
16	56988044	<b>rs173539</b>	T	C	<b>CETP</b>	EX_EUR	6,091	6,091	0	3,311	0.272	0.29	1.99 · 10 <sup>-2</sup>	5.67 · 10 <sup>-47</sup>
15	58678512	<b>rs10468017</b>	T	C	<b>LIPC</b>	EX_EUR	6,091	6,091	0	4,050	0.332	0.15	1.9 · 10 <sup>-2</sup>	3.03 · 10 <sup>-15</sup>
16	56985139	<b>rs9989419</b>	G	A	<b>HERPUD1</b>	EX_EUR	6,088	6,088	0	4,548	0.626	0.145	1.85 · 10 <sup>-2</sup>	5.77 · 10 <sup>-15</sup>
8	19816934	<b>rs301</b>	C	T	<b>LPL</b>	EX_EUR	6,090	6,090	0	2,603	0.214	0.135	2.2 · 10 <sup>-2</sup>	9.42 · 10 <sup>-10</sup>
8	126507389	<b>rs2954038</b>	A	C	<b>TRIB1</b>	EX_EUR	6,090	6,090	0	3,130	0.743	0.113	2.06 · 10 <sup>-2</sup>	4.08 · 10 <sup>-8</sup>
2	227020653	<b>rs7578326</b>	G	A	<b>NYAP2</b>	EX_EUR	6,091	6,091	0	4,376	0.359	9.6 · 10 <sup>-2</sup>	1.89 · 10 <sup>-2</sup>	4.13 · 10 <sup>-7</sup>
16	57275033	rs16968135	A	G	RSPRY1	OMNI_EUR	6,078	6,078	0	504	4.15 · 10 <sup>-2</sup>	0.225	4.54 · 10 <sup>-2</sup>	7.1 · 10 <sup>-7</sup>
11	116648917	<b>rs964184</b>	C	G	<b>ZPR1</b>	EX_EUR	6,091	6,091	0	1,558	0.872	0.134	2.7 · 10 <sup>-2</sup>	7.27 · 10 <sup>-7</sup>
16	57181134	rs2291406	C	T	<b>CPNE2</b>	OMNI_EUR	6,077	6,077	0	504	4.15 · 10 <sup>-2</sup>	0.225	4.54 · 10 <sup>-2</sup>	7.38 · 10 <sup>-7</sup>
19	8429323	rs116843064	A	G	<b>ANGPTL4</b>	EX_EUR	6,091	6,091	0	289	2.37 · 10 <sup>-2</sup>	0.292	5.91 · 10 <sup>-2</sup>	7.76 · 10 <sup>-7</sup>
3	9381122	rs2648549	G	A	THUMPD3	OMNI_EUR	6,076	6,076	0	6,034	0.503	8.83 · 10 <sup>-2</sup>	1.8 · 10 <sup>-2</sup>	9.33 · 10 <sup>-7</sup>
11	47357479	rs35078470	C	T	<b>MYBPC3</b>	EX_EUR	6,091	6,091	0	344	2.82 · 10 <sup>-2</sup>	0.263	5.47 · 10 <sup>-2</sup>	1.47 · 10 <sup>-6</sup>
2	227068080	exm-rs2943634	A	C	<b>IRS1</b>	EX_EUR	6,091	6,091	0	4,332	0.644	9.08 · 10 <sup>-2</sup>	1.9 · 10 <sup>-2</sup>	1.84 · 10 <sup>-6</sup>
6	89967489	rs61739705	C	T	GABRR2	EX_EUR	6,091	6,091	0	10	8.21 · 10 <sup>-4</sup>	1.503	0.316	1.98 · 10 <sup>-6</sup>
8	9183596	<b>rs4841132</b>	G	A	<b>PPP1R3B</b>	EX_EUR	6,091	6,091	0	2,127	0.825	0.113	2.39 · 10 <sup>-2</sup>	2.2 · 10 <sup>-6</sup>
2	215661710	rs10498023	T	C	BARD1	OMNI_EUR	6,070	6,070	0	12	9.88 · 10 <sup>-4</sup>	1.354	0.288	2.7 · 10 <sup>-6</sup>
4	100301048	rs1826906	T	C	ADH1C	OMNI_EUR	6,077	6,077	0	4,010	0.33	8.99 · 10 <sup>-2</sup>	1.93 · 10 <sup>-2</sup>	3.12 · 10 <sup>-6</sup>
11	48149491	rs61739179	G	A	<b>PTPRJ</b>	EX_EUR	6,091	6,091	0	346	2.84 · 10 <sup>-2</sup>	0.255	5.47 · 10 <sup>-2</sup>	3.31 · 10 <sup>-6</sup>
6	35033854	rs820077	G	A	<b>ANKS1A</b>	OMNI_EUR	6,077	6,077	0	926	0.924	0.158	3.42 · 10 <sup>-2</sup>	3.83 · 10 <sup>-6</sup>
1	109817590	<b>rs12740374</b>	T	G	<b>CELSR2</b>	EX_EUR	6,091	6,091	0	2,778	0.228	9.97 · 10 <sup>-2</sup>	2.17 · 10 <sup>-2</sup>	4.39 · 10 <sup>-6</sup>



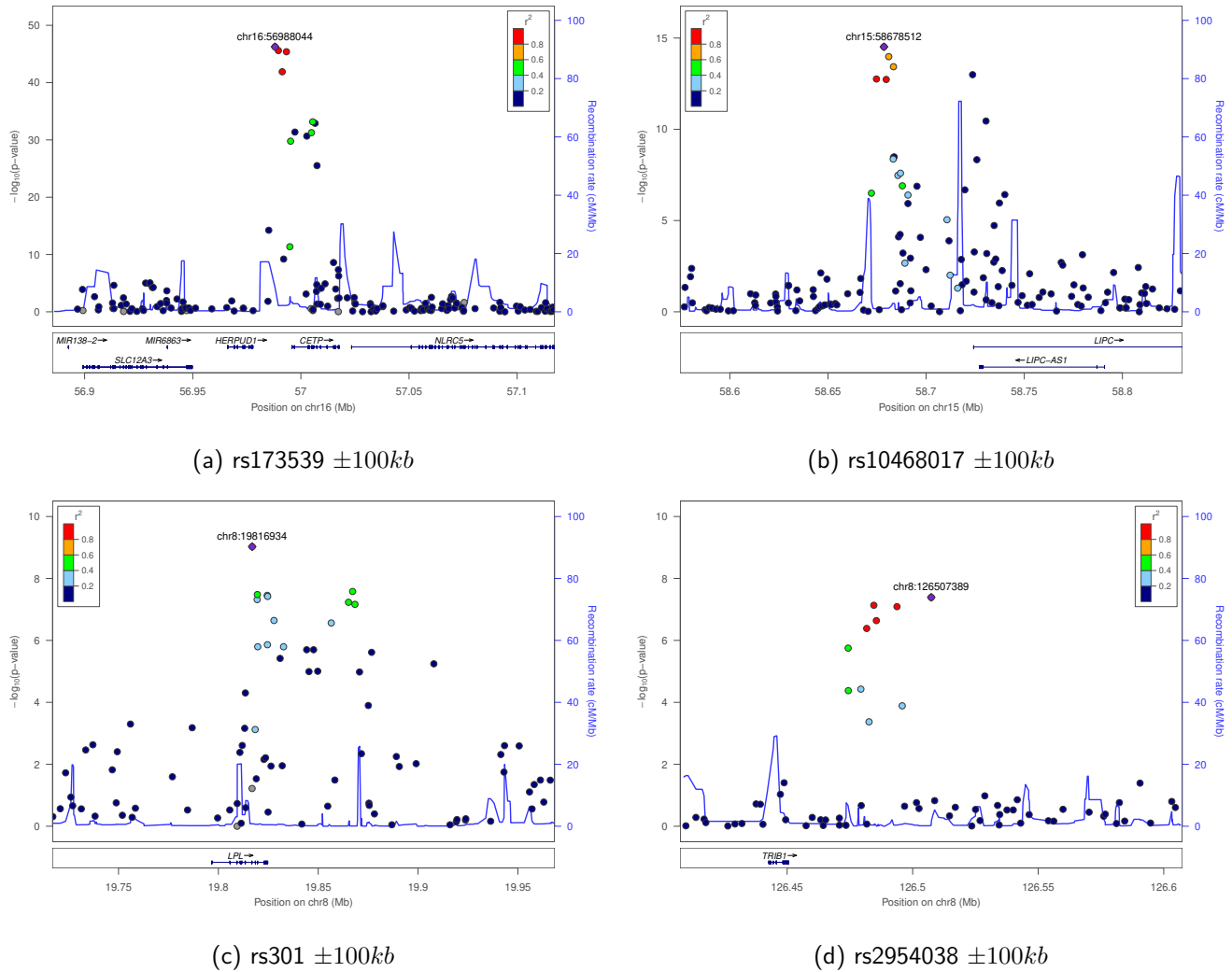


Figure 12: Regional plots for cohort MERGE model invn Adjusted Age+Age2+BMI

### 5.4 Previously identified risk loci

Table 15 shows statistics from the MERGE cohort for 50 loci that were shown to be significantly associated with HDL Cholesterol in the 2013 Nature Genetics paper by Willer et al [11]. Where a previously reported variant was not genotyped in the study (indicated by  $\bar{R}^2 < 1$ ), if available, a tagging variant in LD with the reported variant ( $\bar{R}^2 \geq 0.7$  and within 250kb) was provided. Tags were identified using 1000 Genomes data. There are 44 variants that show at least nominal significance ( $p < 0.05$ ) in this study. Out of the 50 variants in both studies, 49 exhibit the same direction of effect with the known result (binomial test  $p = 4.53e - 14$ ).



## 6 Estimated Glomerular Filtration Rate (eGFR)

### 6.1 Summary

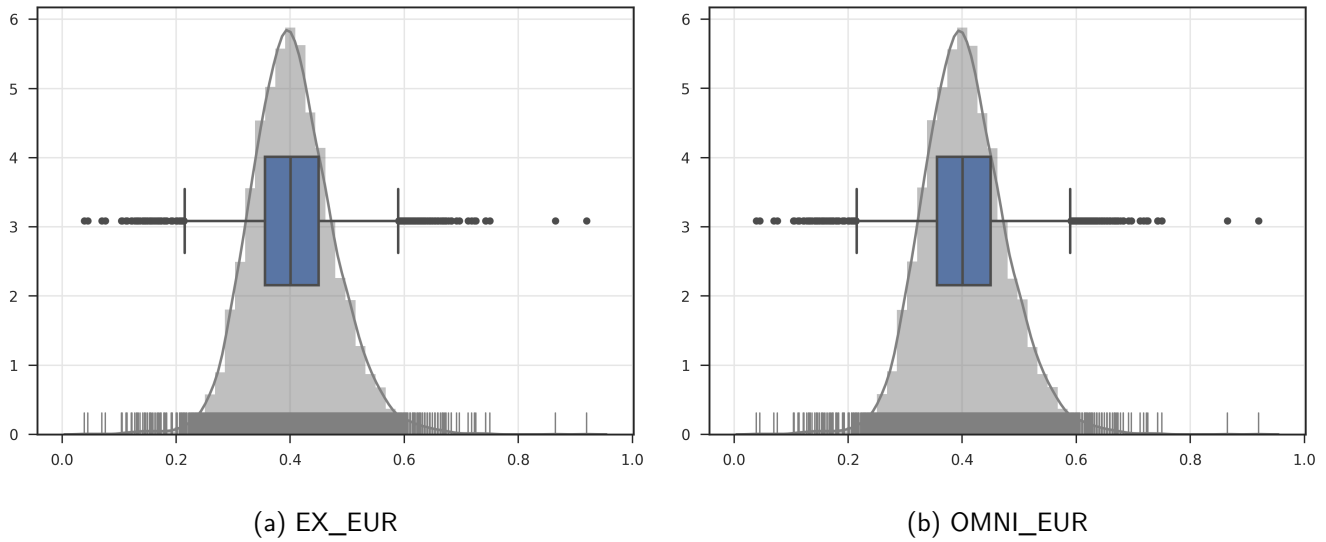


Figure 13: Distribution of eGFR in cohort-level analyses

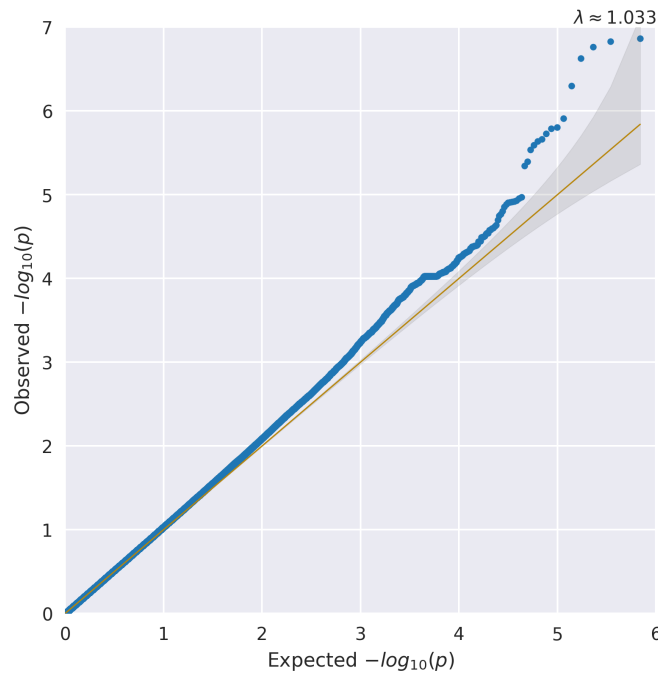
Table 16: Summary of samples removed from Estimated Glomerular Filtration Rate analysis by cohort and model

Cohort	Array	Ancestry	Trans	Covars	Total	-SampleQc	-missObs	-Kinship	-PcOutlier
EX_EUR	EX	EUR	invn		10071	36	0	1515	2
OMNI_EUR	OMNI	EUR	invn		10048	69	0	1497	1

Table 17: Summary of samples remaining for Estimated Glomerular Filtration Rate analysis by cohort and model

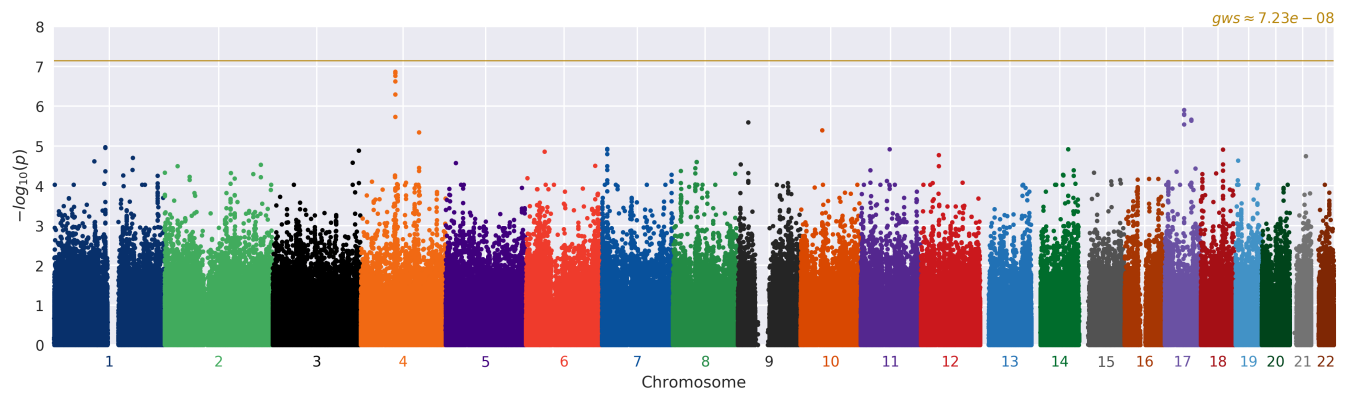
Cohort	Array	Ancestry	Trans	Covars	PCs	N	Male	Female	Max	Min	$\mu$	$\tilde{x}$	$\sigma$
EX_EUR	EX	EUR	invn		1	8518	8518	0	0.92	0.039	0.404	0.4	0.074
OMNI_EUR	OMNI	EUR	invn		8	8481	8481	0	0.92	0.039	0.403	0.399	0.074

## 6.2 Calibration



(a) invn Unadjusted

Figure 14: QQ plots for eGFR in the MERGE analysis



(a) invn Unadjusted

Figure 15: Manhattan plots for eGFR in the MERGE analysis

### 6.3 Top associations

Table 18: Top variants in the MERGE invn Unadjusted model (**bold** variants indicate previously identified associations)

CHR	POS	ID	EA	OA	GENE <sub>CLOSEST</sub>	COHORT	N	MALE	FEMALE	MAC	FREQ	EFFECT	STDERR	P
4	77398015	<b>rs10032549</b>	A	G	<b>SHROOM3</b>	OMNI_EUR	8,471	8,471	0	8,035	0.526	$8.1 \cdot 10^{-2}$	$1.54 \cdot 10^{-2}$	$1.37 \cdot 10^{-7}$
17	43483774	rs1230101	G	A	ARHGAP27	OMNI_EUR	8,477	8,477	0	5,852	0.655	$7.82 \cdot 10^{-2}$	$1.61 \cdot 10^{-2}$	$1.24 \cdot 10^{-6}$
17	59456589	<b>rs9895661</b>	T	C	<b>BCAS3</b>	EX_EUR	8,518	8,518	0	3,338	0.196	$9.2 \cdot 10^{-2}$	$1.94 \cdot 10^{-2}$	$2.19 \cdot 10^{-6}$
9	23656678	rs1329035	T	C	ELAVL2	OMNI_EUR	8,475	8,475	0	3,779	0.223	$8.69 \cdot 10^{-2}$	$1.85 \cdot 10^{-2}$	$2.55 \cdot 10^{-6}$
10	48438147	rs1902724	T	G	GDF10	OMNI_EUR	8,480	8,480	0	4,168	0.246	$8.21 \cdot 10^{-2}$	$1.78 \cdot 10^{-2}$	$4.02 \cdot 10^{-6}$
4	130824678	rs3111741	A	G	C4orf33	OMNI_EUR	8,481	8,481	0	4,293	0.253	$8.03 \cdot 10^{-2}$	$1.75 \cdot 10^{-2}$	$4.55 \cdot 10^{-6}$
1	115738343	rs1076754	T	C	NGF	OMNI_EUR	8,480	8,480	0	4,740	0.721	$7.44 \cdot 10^{-2}$	$1.69 \cdot 10^{-2}$	$1.07 \cdot 10^{-5}$
7	11604426	rs7805116	A	G	THSD7A	OMNI_EUR	8,481	8,481	0	2,612	0.846	$9.32 \cdot 10^{-2}$	$2.13 \cdot 10^{-2}$	$1.18 \cdot 10^{-5}$
11	65338363	rs143562975	A	C	SSSCA1	EX_EUR	8,518	8,518	0	4	1	2.183	0.499	$1.21 \cdot 10^{-5}$
14	82510077	rs2217448	C	T	SEL1L	OMNI_EUR	8,479	8,479	0	2,040	0.88	0.103	$2.36 \cdot 10^{-2}$	$1.21 \cdot 10^{-5}$
18	49570773	rs4318302	A	C	DCC	OMNI_EUR	8,478	8,478	0	4,327	0.255	$7.73 \cdot 10^{-2}$	$1.77 \cdot 10^{-2}$	$1.24 \cdot 10^{-5}$
3	193811162	rs10933711	T	C	HES1	OMNI_EUR	8,480	8,480	0	6,158	0.637	$6.97 \cdot 10^{-2}$	$1.6 \cdot 10^{-2}$	$1.31 \cdot 10^{-5}$
6	42084146	rs10456507	C	A	C6orf132	OMNI_EUR	8,480	8,480	0	1,098	0.935	0.137	$3.14 \cdot 10^{-2}$	$1.4 \cdot 10^{-5}$
12	40273222	rs4277185	T	G	SLC2A13	OMNI_EUR	8,442	8,442	0	3,846	0.772	$7.85 \cdot 10^{-2}$	$1.82 \cdot 10^{-2}$	$1.69 \cdot 10^{-5}$
21	36172283	rs2073354	T	G	RUNX1	OMNI_EUR	8,463	8,463	0	932	0.945	0.145	$3.38 \cdot 10^{-2}$	$1.78 \cdot 10^{-5}$
1	177179432	rs17380506	G	A	BRINP2	OMNI_EUR	8,480	8,480	0	1,377	0.919	0.12	$2.8 \cdot 10^{-2}$	$2 \cdot 10^{-5}$
19	5711930	rs11085147	T	C	LONP1	EX_EUR	8,518	8,518	0	1,476	0.913	0.115	$2.71 \cdot 10^{-2}$	$2.31 \cdot 10^{-5}$
1	90601121	rs6428572	C	T	ZNF326	OMNI_EUR	8,478	8,478	0	1,149	0.932	0.129	$3.06 \cdot 10^{-2}$	$2.4 \cdot 10^{-5}$
8	53355822	rs17228239	G	A	ST18	OMNI_EUR	8,480	8,480	0	8,376	0.494	$6.46 \cdot 10^{-2}$	$1.53 \cdot 10^{-2}$	$2.51 \cdot 10^{-5}$
3	179529649	rs141827659	C	A	PEX5L	EX_EUR	8,518	8,518	0	2	1	2.965	0.705	$2.62 \cdot 10^{-5}$

### 6.4 Previously identified risk loci

Table 19 shows statistics from the MERGE cohort for 50 loci that were shown to be significantly associated with Estimated Glomerular Filtration Rate in the 2016 Nature Communications paper by Pattaro et al [12]. Where a previously reported variant was not genotyped in the study (indicated by  $\bar{R}^2 < 1$ ), if available, a tagging variant in LD with the reported variant ( $\bar{R}^2 \geq 0.7$  and within 250kb) was provided. Tags were identified using 1000 Genomes data. There are 32 variants that show at least nominal significance ( $p < 0.05$ ) in this study. Out of the 50 variants in both studies, 46 exhibit the same direction of effect with the known result (binomial test  $p = 2.23e - 10$ ).



## 7 Diastolic Blood Pressure (DBP10)

### 7.1 Summary

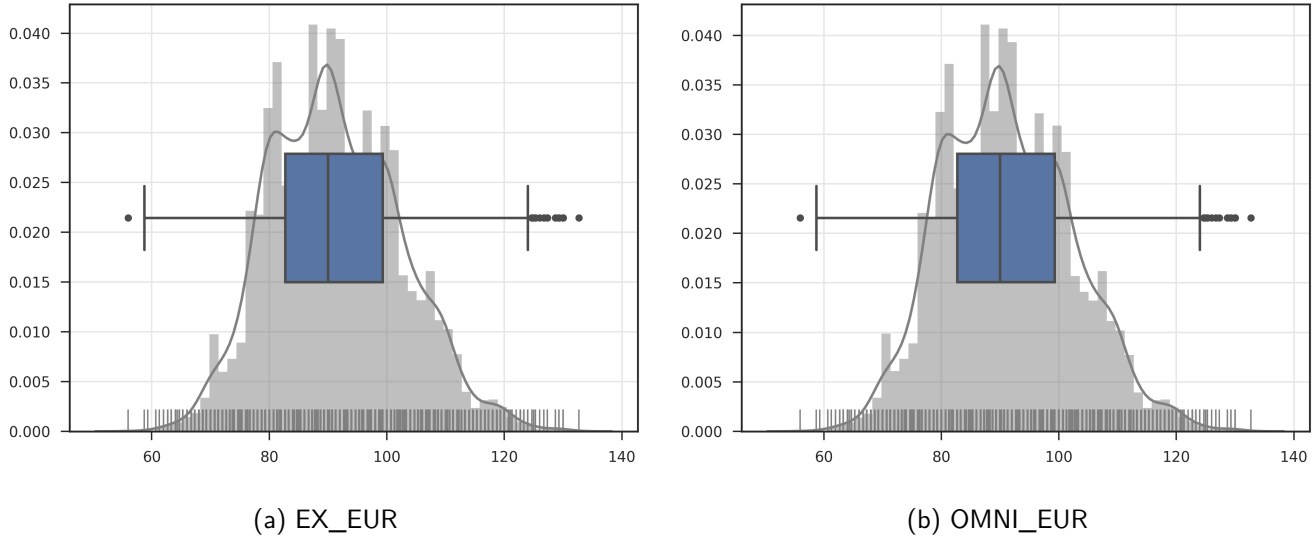


Figure 16: Distribution of DBP10 in cohort-level analyses

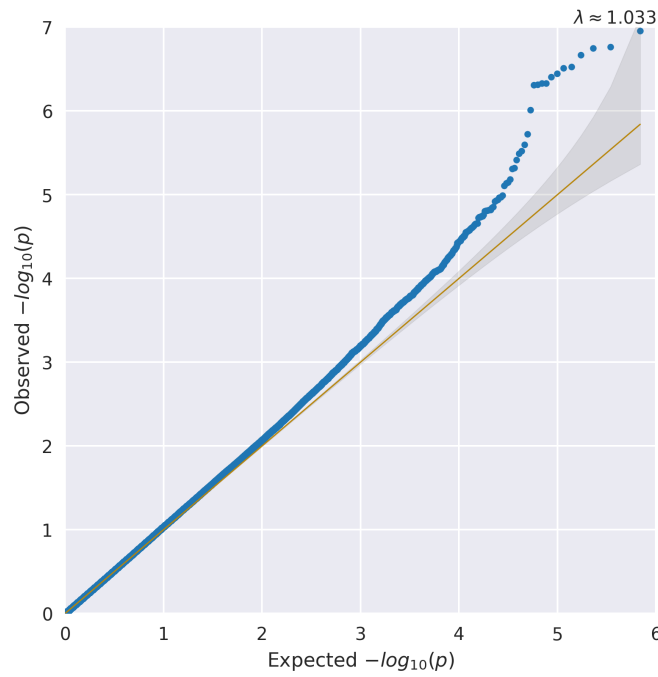
Table 20: Summary of samples removed from Diastolic Blood Pressure analysis by cohort and model

Cohort	Array	Ancestry	Trans	Covars	Total	-SampleQc	-missObs	-Kinship	-PcOutlier
EX_EUR	EX	EUR	invn	Age+Age2+BMI	10071	36	5	1517	1
OMNI_EUR	OMNI	EUR	invn	Age+Age2+BMI	10048	69	5	1494	0

Table 21: Summary of samples remaining for Diastolic Blood Pressure analysis by cohort and model

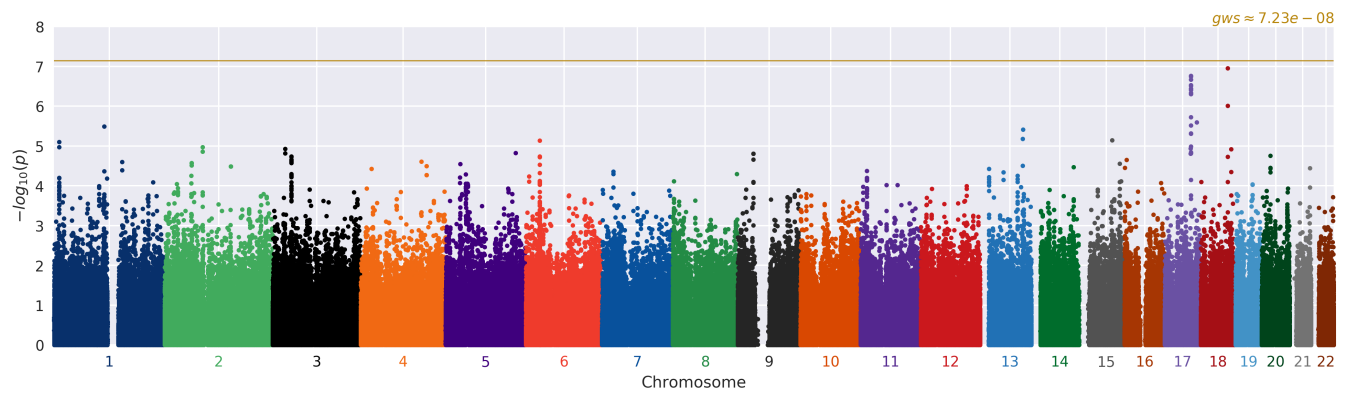
Cohort	Array	Ancestry	Trans	Covars	PCs	N	Male	Female	Max	Min	$\mu$	$\bar{x}$	$\sigma$
EX_EUR	EX	EUR	invn	Age+Age2+BMI	10	8512	8512	0	132.7	58.7	91.318	90.7	11.119
OMNI_EUR	OMNI	EUR	invn	Age+Age2+BMI	4	8480	8480	0	132.7	56.0	91.289	90.15	11.145

## 7.2 Calibration



(a) invn Adjusted Age+Age2+BMI

Figure 17: QQ plots for DBP10 in the MERGE analysis



(a) invn Adjusted Age+Age2+BMI

Figure 18: Manhattan plots for DBP10 in the MERGE analysis



### 7.3 Top associations

Table 22: Top variants in the MERGE invn Adjusted Age+Age2+BMI model (**bold** variants indicate previously identified associations)

CHR	POS	ID	EA	OA	GENE <sub>CLOSEST</sub>	COHORT	N	MALE	FEMALE	MAC	FREQ	EFFECT	STDERR	P
18	60661358	rs7238945	G	T	PHLPP1	OMNI_EUR	8,375	8,375	0	1,860	0.111	0.131	$2.47 \cdot 10^{-2}$	$1.1 \cdot 10^{-7}$
17	59087808	rs9652858	A	G	BCAS3	OMNI_EUR	8,476	8,476	0	2,709	0.16	0.109	$2.08 \cdot 10^{-2}$	$1.73 \cdot 10^{-7}$
17	72126787	rs4789620	C	T	RPL38	OMNI_EUR	8,479	8,479	0	570	0.966	0.2	$4.26 \cdot 10^{-2}$	$2.54 \cdot 10^{-6}$
1	113161350	rs11102516	A	G	ST7L	OMNI_EUR	8,480	8,480	0	3,482	0.205	$8.82 \cdot 10^{-2}$	$1.89 \cdot 10^{-2}$	$3.24 \cdot 10^{-6}$
13	96013885	rs7992512	A	G	ABCC4	OMNI_EUR	8,473	8,473	0	2,586	0.153	$9.89 \cdot 10^{-2}$	$2.14 \cdot 10^{-2}$	$3.83 \cdot 10^{-6}$
15	74486444	rs974456	C	T	STRA6	OMNI_EUR	8,473	8,473	0	6,071	0.358	$7.2 \cdot 10^{-2}$	$1.6 \cdot 10^{-2}$	$7.15 \cdot 10^{-6}$
6	31177503	rs3132505	G	T	POU5F1	OMNI_EUR	8,480	8,480	0	2,670	0.157	$9.48 \cdot 10^{-2}$	$2.11 \cdot 10^{-2}$	$7.29 \cdot 10^{-6}$
1	11905974	rs5068	A	G	NPPA	OMNI_EUR	8,480	8,480	0	2,001	0.118	0.107	$2.4 \cdot 10^{-2}$	$7.86 \cdot 10^{-6}$
1	11876662	rs17376328	G	A	<b>CLCN6</b>	OMNI_EUR	8,477	8,477	0	2,003	0.118	0.106	$2.4 \cdot 10^{-2}$	$1.07 \cdot 10^{-5}$
2	85480715	rs17763853	G	A	TCF7L1	OMNI_EUR	8,479	8,479	0	4,442	0.262	$7.67 \cdot 10^{-2}$	$1.74 \cdot 10^{-2}$	$1.08 \cdot 10^{-5}$
17	58359510	rs11652788	G	A	USP32	OMNI_EUR	8,480	8,480	0	1,707	0.101	0.112	$2.55 \cdot 10^{-2}$	$1.15 \cdot 10^{-5}$
3	27651544	rs12715123	T	C	EOMES	OMNI_EUR	8,475	8,475	0	5,228	0.692	$7.33 \cdot 10^{-2}$	$1.67 \cdot 10^{-2}$	$1.18 \cdot 10^{-5}$
18	68050814	rs3943675	G	A	SOCS6	OMNI_EUR	8,478	8,478	0	8,256	0.513	$6.71 \cdot 10^{-2}$	$1.53 \cdot 10^{-2}$	$1.21 \cdot 10^{-5}$
5	157523151	rs10045606	G	A	CLINT1	OMNI_EUR	8,480	8,480	0	4,064	0.24	$7.7 \cdot 10^{-2}$	$1.78 \cdot 10^{-2}$	$1.51 \cdot 10^{-5}$
9	35374982	rs2149368	G	A	UNC13B	OMNI_EUR	8,480	8,480	0	3,472	0.205	$8.22 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$	$1.55 \cdot 10^{-5}$
17	58504201	rs11655643	C	T	C17orf64	OMNI_EUR	8,479	8,479	0	1,708	0.101	0.11	$2.55 \cdot 10^{-2}$	$1.56 \cdot 10^{-5}$
20	19280659	rs6081560	A	G	SLC24A3	OMNI_EUR	8,479	8,479	0	5,351	0.316	$7.11 \cdot 10^{-2}$	$1.65 \cdot 10^{-2}$	$1.75 \cdot 10^{-5}$
6	31206979	rs2394895	T	C	HLA-C	EX_EUR	8,511	8,511	0	2,665	0.157	$9.12 \cdot 10^{-2}$	$2.13 \cdot 10^{-2}$	$1.81 \cdot 10^{-5}$
3	41853161	rs7632387	T	C	ULK4	OMNI_EUR	8,446	8,446	0	3,850	0.228	$7.83 \cdot 10^{-2}$	$1.83 \cdot 10^{-2}$	$1.82 \cdot 10^{-5}$
18	60878883	rs9955190	A	G	BCL2	OMNI_EUR	8,478	8,478	0	5,944	0.351	$6.85 \cdot 10^{-2}$	$1.6 \cdot 10^{-2}$	$1.85 \cdot 10^{-5}$

### 7.4 Previously identified risk loci

Table 23 shows statistics from the MERGE cohort for 26 loci that were shown to be significantly associated with Diastolic Blood Pressure in the 2011 Nature paper by Ehret et al [10]. Where a previously reported variant was not genotyped in the study (indicated by  $\bar{R}^2 < 1$ ), if available, a tagging variant in LD with the reported variant ( $\bar{R}^2 \geq 0.7$  and within 250kb) was provided. Tags were identified using 1000 Genomes data. There are 11 variants that show at least nominal significance ( $p < 0.05$ ) in this study. Out of the 23 variants in both studies, 8 exhibit the same direction of effect with the known result (binomial test  $p = 0.953$ ).

Table 23: Top known loci in MERGE model invn Adjusted Age+Age2+BMI (**bold** variants indicate matching direction of effect)

CHR	POS	ID	EA	OA	N	FREQ	EFFECT	STDERR	P	COHORT	GENE <sub>CLOSEST</sub>	R <sup>2</sup>	ID <sub>KNOWN</sub>	N <sub>KNOWN</sub>	EFFECT <sub>KNOWN</sub>	STDERR <sub>KNOWN</sub>	P <sub>KNOWN</sub>
12	112007756	rs653178	C	T	8,512	0.398	$3.6 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$	$2.06 \cdot 10^{-2}$	EX_EUR	ATXN2	1	rs653178	$2 \cdot 10^5$	-0.48	$6.26 \cdot 10^{-2}$	$1.64 \cdot 10^{-14}$
12	111884608	<b>rs3184504</b>	T	C	8,511	0.388	$3.85 \cdot 10^{-2}$	$1.56 \cdot 10^{-2}$	$1.35 \cdot 10^{-2}$	EX_EUR	SH2B3	1	rs3184504	$2 \cdot 10^5$	0.48	$6.29 \cdot 10^{-2}$	$2.33 \cdot 10^{-14}$
4	81164723	<b>rs1458038</b>	T	C	8,479	0.692	$1.48 \cdot 10^{-2}$	$1.65 \cdot 10^{-2}$	0.37	OMNI_EUR	FGF5	1	rs1458038	$2 \cdot 10^5$	0.503	$7.02 \cdot 10^{-2}$	$7.91 \cdot 10^{-13}$
12	112072424	<b>rs11065987</b>	G	A	8,512	0.624	$3.38 \cdot 10^{-2}$	$1.57 \cdot 10^{-2}$	$3.1 \cdot 10^{-2}$	EX_EUR	BRAP	1	rs11065987	$2 \cdot 10^5$	0.449	$6.46 \cdot 10^{-2}$	$3.43 \cdot 10^{-12}$
15	75077367	<b>rs1378942</b>	C	A	8,512	0.447	$4.02 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$	$9.58 \cdot 10^{-3}$	EX_EUR	CSK	1	rs1378942	$2 \cdot 10^5$	0.445	$6.4 \cdot 10^{-2}$	$3.47 \cdot 10^{-12}$
12	112486818	<b>rs17696736</b>	G	A	8,512	0.6	$2.25 \cdot 10^{-2}$	$1.54 \cdot 10^{-2}$	0.146	EX_EUR	NAA25	1	rs17696736	$2 \cdot 10^5$	0.422	$6.34 \cdot 10^{-2}$	$2.8 \cdot 10^{-11}$
12	112906415	rs11066320	A	G	8,478	0.399	$1.62 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$	0.295	OMNI_EUR	PTPN11	1	rs11066320	$2 \cdot 10^5$	-0.413	$6.32 \cdot 10^{-2}$	$6.32 \cdot 10^{-11}$
15	75057203	rs4886406	G	T	8,478	0.283	$3.26 \cdot 10^{-2}$	$1.7 \cdot 10^{-2}$	$5.52 \cdot 10^{-2}$	OMNI_EUR	CYP1A2	1	rs4886406	$2 \cdot 10^5$	-0.426	$6.85 \cdot 10^{-2}$	$4.83 \cdot 10^{-10}$
15	75125645	rs6495122	A	C	8,512	0.53	$2.61 \cdot 10^{-2}$	$1.54 \cdot 10^{-2}$	$9.06 \cdot 10^{-2}$	EX_EUR	CPLX3	1	rs6495122	$2 \cdot 10^5$	-0.383	$6.23 \cdot 10^{-2}$	$8.41 \cdot 10^{-10}$
15	75115895	<b>rs7162232</b>	G	A	8,479	0.332	$3.47 \cdot 10^{-2}$	$1.63 \cdot 10^{-2}$	$3.39 \cdot 10^{-2}$	OMNI_EUR	LMAN1L	1	rs7162232	$2 \cdot 10^5$	0.416	$6.89 \cdot 10^{-2}$	$1.58 \cdot 10^{-9}$
12	90008959	rs2681472	A	G	8,510	0.924	$6.33 \cdot 10^{-2}$	$2.92 \cdot 10^{-2}$	$3 \cdot 10^{-2}$	EX_EUR	ATP2B1	1	rs2681472	$2 \cdot 10^5$	-0.492	$8.36 \cdot 10^{-2}$	$3.9 \cdot 10^{-9}$
15	75140854	<b>rs3765066</b>	G	A	8,480	0.432	$2.93 \cdot 10^{-2}$	$1.58 \cdot 10^{-2}$	$6.34 \cdot 10^{-2}$	OMNI_EUR	SCAMP2	1	rs3765066	$2 \cdot 10^5$	0.382	$6.51 \cdot 10^{-2}$	$4.17 \cdot 10^{-9}$
10	63524591	rs1530440	C	T	8,511	0.801	$1.49 \cdot 10^{-2}$	$1.93 \cdot 10^{-2}$	0.441	EX_EUR	C10orf107	1	rs1530440	$2 \cdot 10^5$	-0.459	$7.92 \cdot 10^{-2}$	$6.71 \cdot 10^{-9}$
1	11883731	rs12567136	C	T	8,480	0.832	$7.91 \cdot 10^{-2}$	$2.05 \cdot 10^{-2}$	$1.17 \cdot 10^{-4}$	OMNI_EUR	CLCN6	1	rs12567136	$2 \cdot 10^5$	-0.488	$8.56 \cdot 10^{-2}$	$1.15 \cdot 10^{-8}$
1	11862778	rs17367504	A	G	8,512	0.832	$6.67 \cdot 10^{-2}$	$2.05 \cdot 10^{-2}$	$1.12 \cdot 10^{-3}$	EX_EUR	MTHFR	1	rs17367504	$2 \cdot 10^5$	-0.49	$8.61 \cdot 10^{-2}$	$1.29 \cdot 10^{-8}$
12	89942390	rs11105328	A	G	8,480	0.927	$5.85 \cdot 10^{-2}$	$2.94 \cdot 10^{-2}$	$4.65 \cdot 10^{-2}$	OMNI_EUR	POC1B-GALNT4	1	rs11105328	$2 \cdot 10^5$	-0.487	$8.66 \cdot 10^{-2}$	$1.83 \cdot 10^{-8}$
12	111798553	rs3742004	A	G	8,479	0.79	$3.97 \cdot 10^{-2}$	$1.89 \cdot 10^{-2}$	$3.6 \cdot 10^{-2}$	OMNI_EUR	FAM109A	1	rs3742004	$2 \cdot 10^5$	-0.429	$7.66 \cdot 10^{-2}$	$2.15 \cdot 10^{-8}$
15	75234610	<b>rs11072518</b>	T	C	8,480	0.472	$2.74 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$	$7.8 \cdot 10^{-2}$	OMNI_EUR	COX5A	1	rs11072518	$2 \cdot 10^5$	0.355	$6.45 \cdot 10^{-2}$	$3.78 \cdot 10^{-8}$
6	26107463	rs198846	A	G	8,512	0.101	$1.9 \cdot 10^{-2}$	$2.55 \cdot 10^{-2}$	0.456	EX_EUR	HIST1H1T	1	rs198846	$2 \cdot 10^5$	-0.487	$8.85 \cdot 10^{-2}$	$3.8 \cdot 10^{-8}$
6	26091179	rs1799945	G	C	8,512	0.9	$1.83 \cdot 10^{-2}$	$2.56 \cdot 10^{-2}$	0.474	EX_EUR	HFE	1	rs1799945	$2 \cdot 10^5$	-0.482	$8.82 \cdot 10^{-2}$	$4.78 \cdot 10^{-8}$
12	111788402	rs10219736	C	T	8,476	0.791	$4.08 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$	$3.14 \cdot 10^{-2}$	OMNI_EUR	CUX2	1	rs10219736	$2 \cdot 10^5$	-0.414	$7.58 \cdot 10^{-2}$	$4.88 \cdot 10^{-8}$
6	26107463	rs198846	A	G	8,512	0.101	$1.9 \cdot 10^{-2}$	$2.55 \cdot 10^{-2}$	0.456	EX_EUR	HIST1H2BC	1	rs198833	$2 \cdot 10^5$	-0.485	$8.88 \cdot 10^{-2}$	$4.58 \cdot 10^{-8}$
15	75189930	rs1130741	A	G	8,512	0.572	$1.87 \cdot 10^{-2}$	$1.57 \cdot 10^{-2}$	0.234	EX_EUR	MPI	1	rs7495739	$2 \cdot 10^5$	-0.335	$6.15 \cdot 10^{-2}$	$5.02 \cdot 10^{-8}$
12	112486818	<b>rs17696736</b>	G	A	8,512	0.6	$2.25 \cdot 10^{-2}$	$1.54 \cdot 10^{-2}$	0.146	EX_EUR	TRAFD1	0.922	rs17630235	$2 \cdot 10^5$	0.447	$6.4 \cdot 10^{-2}$	$2.92 \cdot 10^{-12}$
12	112486818	<b>rs17696736</b>	G	A	8,512	0.6	$2.25 \cdot 10^{-2}$	$1.54 \cdot 10^{-2}$	0.146	EX_EUR	HECTD4	0.913	rs11066188	$2 \cdot 10^5$	0.447	$6.41 \cdot 10^{-2}$	$3.06 \cdot 10^{-12}$
7	2508072	rs2906166	T	C	8,479	0.261	$3.28 \cdot 10^{-2}$	$1.75 \cdot 10^{-2}$	$6.14 \cdot 10^{-2}$	OMNI_EUR	GRIFIN	0.819	rs2969070	$2 \cdot 10^5$	-0.386	$6.47 \cdot 10^{-2}$	$2.57 \cdot 10^{-9}$

## 8 LDL Cholesterol (LDL\_DIRECT)

### 8.1 Summary

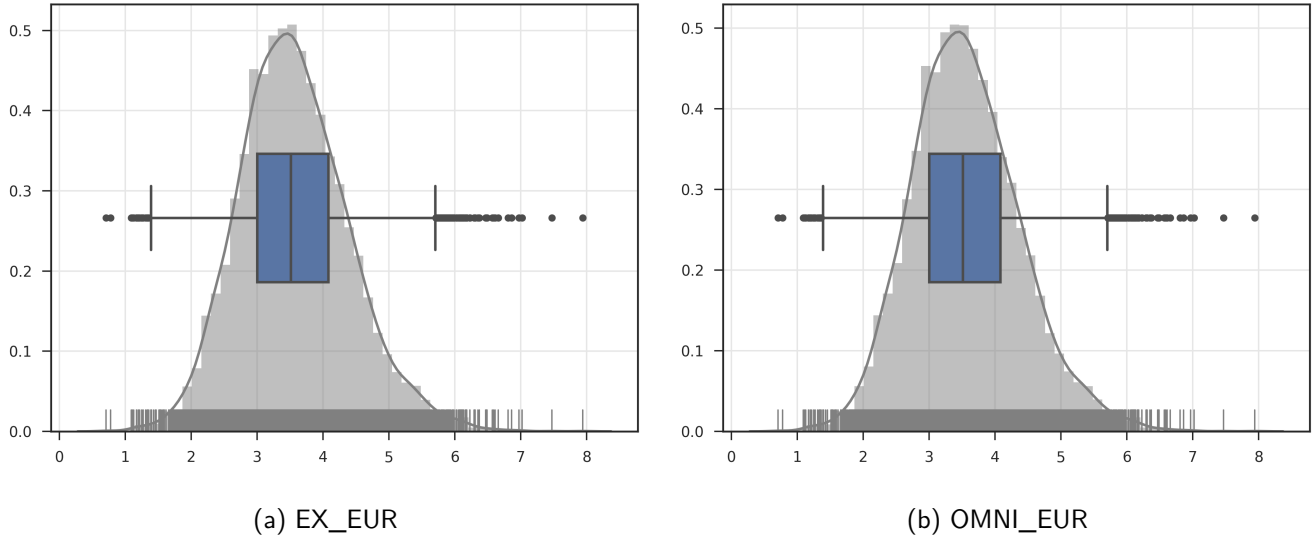


Figure 19: Distribution of LDL\_DIRECT in cohort-level analyses

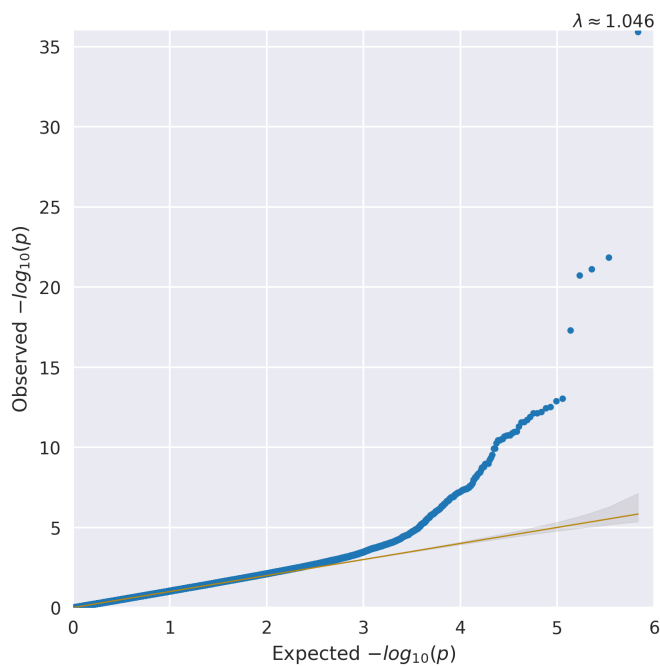
Table 24: Summary of samples removed from LDL Cholesterol analysis by cohort and model

Cohort	Array	Ancestry	Trans	Covars	Total	-SampleQc	-missObs	-Kinship	-PcOutlier
EX_EUR	EX	EUR	invn	Age+Age2+BMI	10071	36	2848	1104	2
OMNI_EUR	OMNI	EUR	invn	Age+Age2+BMI	10048	69	2841	1074	0

Table 25: Summary of samples remaining for LDL Cholesterol analysis by cohort and model

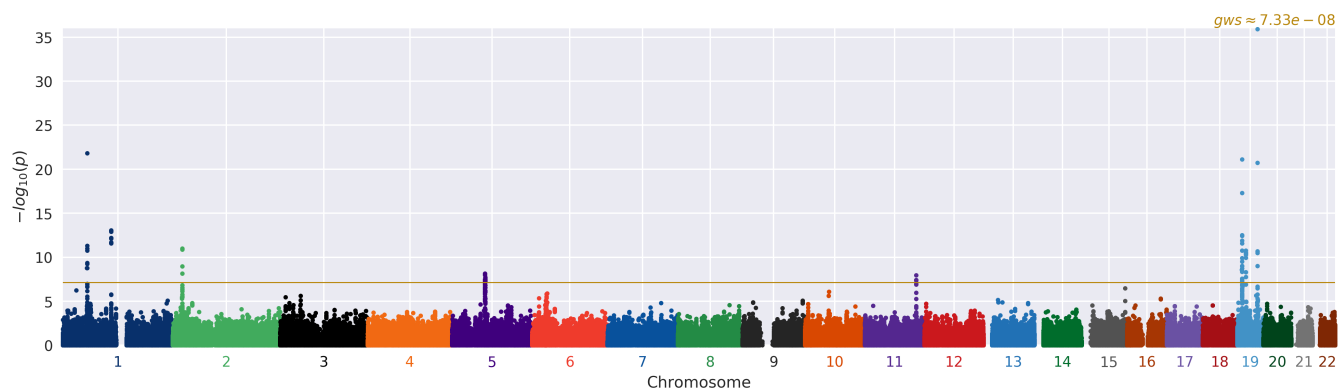
Cohort	Array	Ancestry	Trans	Covars	PCs	N	Male	Female	Max	Min	$\mu$	$\hat{x}$	$\sigma$
EX_EUR	EX	EUR	invn	Age+Age2+BMI	10	6090	6090	0	7.47	0.71	3.564	3.51	0.83
OMNI_EUR	OMNI	EUR	invn	Age+Age2+BMI	4	6080	6080	0	7.94	0.71	3.564	3.52	0.831

## 8.2 Calibration



(a) invn Adjusted Age+Age2+BMI

Figure 20: QQ plots for LDL\_DIRECT in the MERGE analysis



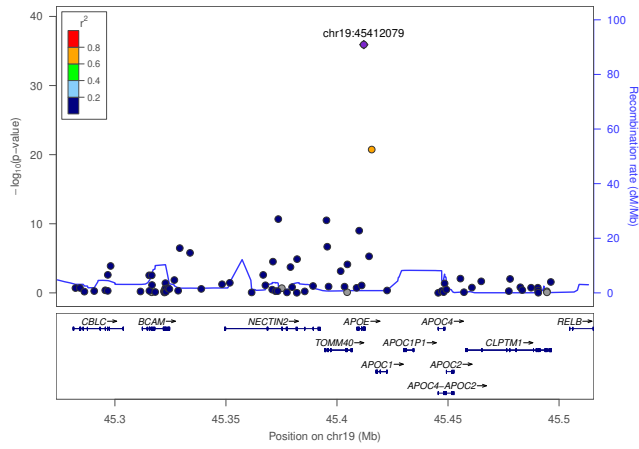
(a) invn Adjusted Age+Age2+BMI

Figure 21: Manhattan plots for LDL\_DIRECT in the MERGE analysis

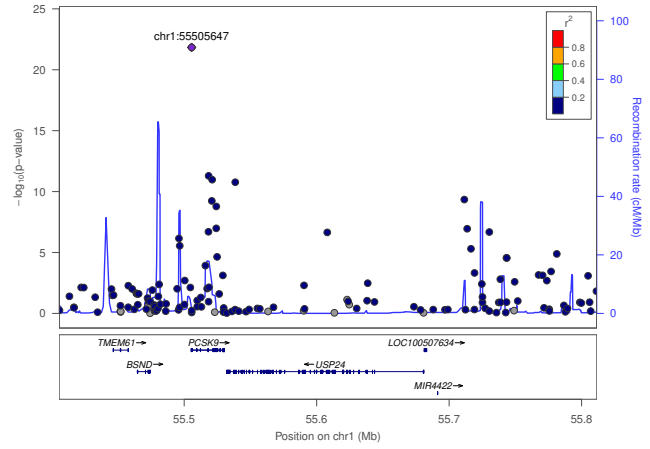
### 8.3 Top associations

Table 26: Top variants in the MERGE invn Adjusted Age+Age2+BMI model (**bold** variants indicate previously identified associations)

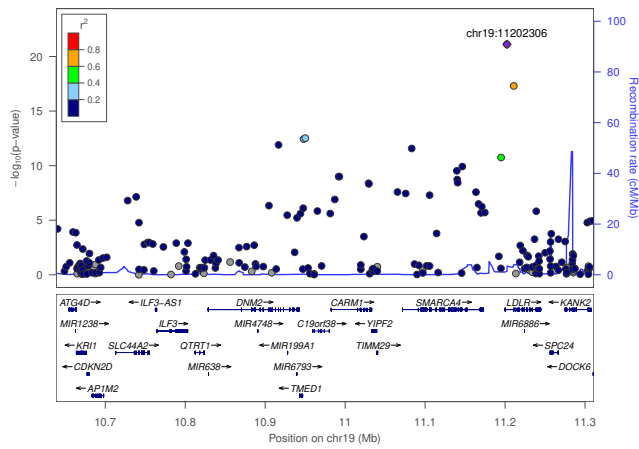
CHR	POS	ID	EA	OA	GENE <sub>CLOSEST</sub>	COHORT	N	MALE	FEMALE	MAC	FREQ	EFFECT	STDERR	P
19	45412079	<b>rs7412</b>	C	T	<b>APOE</b>	EX_EUR	6,089	6,089	0	611	0.95	0.529	$4.15 \cdot 10^{-2}$	$1.2 \cdot 10^{-36}$
1	55505647	<b>rs11591147</b>	G	T	<b>PCSK9</b>	EX_EUR	6,090	6,090	0	576	0.953	0.414	$4.22 \cdot 10^{-2}$	$1.46 \cdot 10^{-22}$
19	11202306	<b>rs6511720</b>	G	T	<b>LDLR</b>	EX_EUR	6,090	6,090	0	1,570	0.871	0.261	$2.71 \cdot 10^{-2}$	$7.63 \cdot 10^{-22}$
19	45415640	<b>rs445925</b>	G	A	<b>APOC1</b>	EX_EUR	6,090	6,090	0	819	0.933	0.348	$3.64 \cdot 10^{-2}$	$1.82 \cdot 10^{-21}$
1	109817192	<b>rs7528419</b>	A	G	<b>CELSR2</b>	OMNI_EUR	6,076	6,076	0	2,739	0.775	0.162	$2.17 \cdot 10^{-2}$	$8.94 \cdot 10^{-14}$
19	10950125	<b>rs11881156</b>	C	T	<b>C19orf38</b>	OMNI_EUR	6,079	6,079	0	2,001	0.835	0.178	$2.44 \cdot 10^{-2}$	$3.01 \cdot 10^{-13}$
19	10916684	<b>rs2287029</b>	C	T	<b>DNM2</b>	OMNI_EUR	6,076	6,076	0	1,990	0.836	0.174	$2.45 \cdot 10^{-2}$	$1.25 \cdot 10^{-12}$
1	109821511	<b>rs602633</b>	G	T	<b>PSRC1</b>	EX_EUR	6,086	6,086	0	2,779	0.228	0.153	$2.17 \cdot 10^{-2}$	$1.95 \cdot 10^{-12}$
19	11083210	<b>rs4804554</b>	G	T	<b>SMARCA4</b>	OMNI_EUR	6,079	6,079	0	2,608	0.785	0.155	$2.21 \cdot 10^{-2}$	$2.66 \cdot 10^{-12}$
2	21277922	<b>rs6548010</b>	G	A	<b>APOB</b>	OMNI_EUR	6,080	6,080	0	3,433	0.718	0.138	$2.02 \cdot 10^{-2}$	$1.01 \cdot 10^{-11}$
1	55538552	<b>rs10493176</b>	T	G	<b>USP24</b>	OMNI_EUR	6,080	6,080	0	2,914	0.76	0.144	$2.14 \cdot 10^{-2}$	$1.7 \cdot 10^{-11}$
19	19379549	<b>rs58542926</b>	C	T	<b>TM6SF2</b>	EX_EUR	6,089	6,089	0	735	0.94	0.255	$3.79 \cdot 10^{-2}$	$1.7 \cdot 10^{-11}$
19	45373565	<b>rs395908</b>	G	A	<b>NECTIN2</b>	OMNI_EUR	6,078	6,078	0	766	0.937	0.251	$3.73 \cdot 10^{-2}$	$2.09 \cdot 10^{-11}$
19	19723215	<b>rs10500212</b>	C	T	<b>PBX4</b>	OMNI_EUR	6,080	6,080	0	742	0.939	0.25	$3.75 \cdot 10^{-2}$	$2.97 \cdot 10^{-11}$
19	45395266	<b>rs157580</b>	A	G	<b>TOMM40</b>	EX_EUR	6,090	6,090	0	3,321	0.273	0.135	$2.03 \cdot 10^{-2}$	$3.16 \cdot 10^{-11}$
19	19662220	<b>rs17216525</b>	C	T	<b>CILP2</b>	EX_EUR	6,090	6,090	0	755	0.938	0.248	$3.73 \cdot 10^{-2}$	$3.55 \cdot 10^{-11}$
19	19407718	<b>rs10401969</b>	T	C	<b>SUGP1</b>	EX_EUR	6,090	6,090	0	731	0.94	0.252	$3.79 \cdot 10^{-2}$	$3.56 \cdot 10^{-11}$
19	19329924	<b>rs2228603</b>	C	T	<b>NCAN</b>	EX_EUR	6,090	6,090	0	786	0.935	0.236	$3.66 \cdot 10^{-2}$	$1.16 \cdot 10^{-10}$
19	10991827	<b>rs11880628</b>	G	A	<b>CARM1</b>	OMNI_EUR	6,077	6,077	0	2,884	0.763	0.131	$2.14 \cdot 10^{-2}$	$9.98 \cdot 10^{-10}$
2	21320516	<b>rs2337383</b>	G	A	<b>TDRD15</b>	OMNI_EUR	6,080	6,080	0	3,067	0.748	0.128	$2.09 \cdot 10^{-2}$	$1.06 \cdot 10^{-9}$



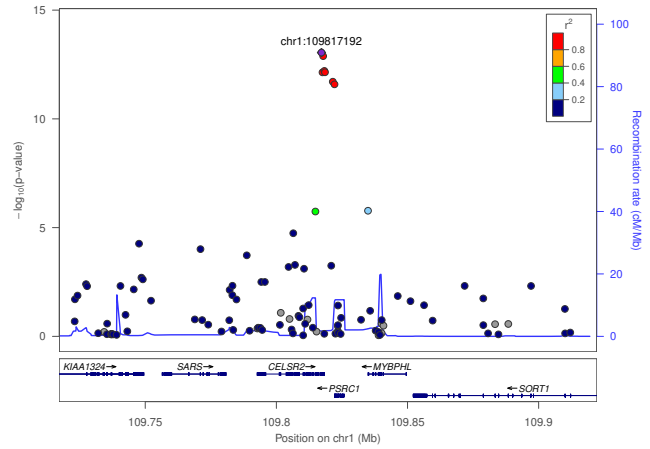
(a) rs7412 ±100kb



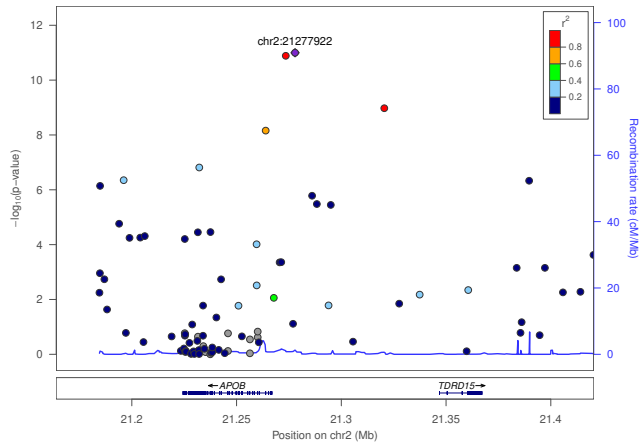
(b) rs11591147 ±100kb



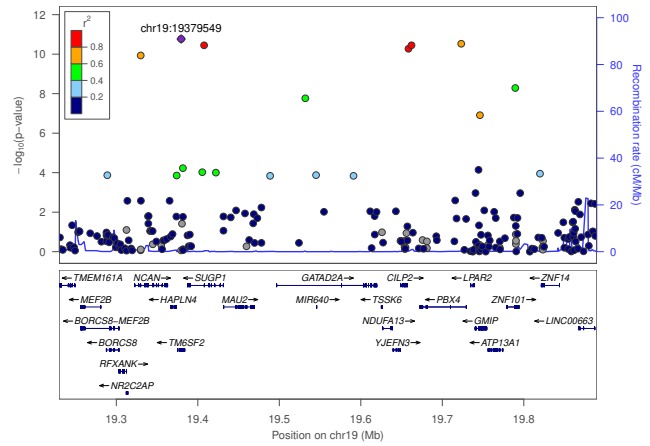
(c) rs6511720 ±100kb



(d) rs7528419 ±100kb

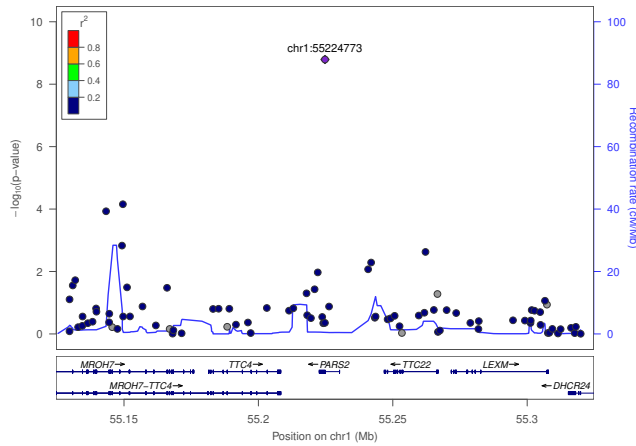


(e) rs6548010 ±100kb

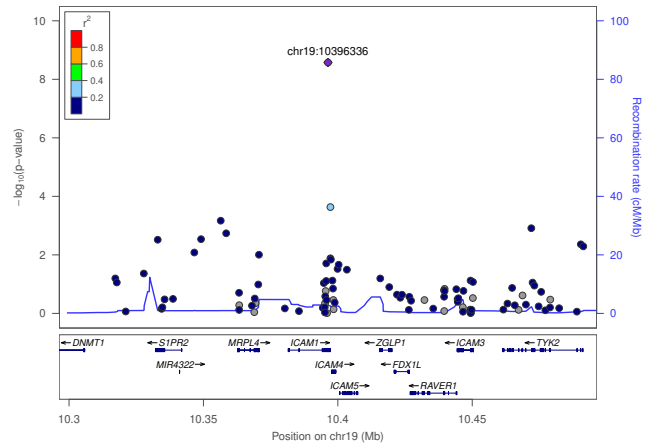


(f) rs58542926 ±100kb

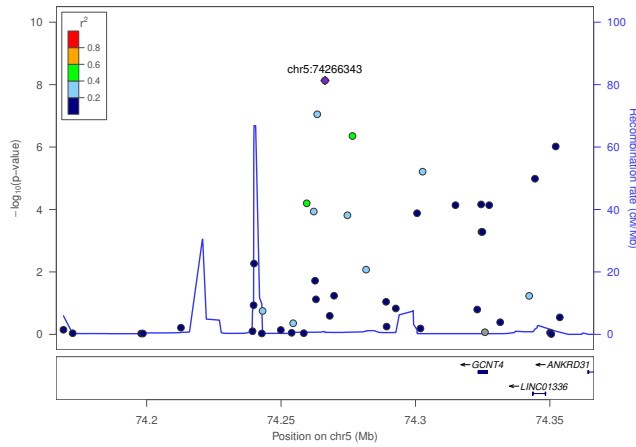
Figure 22: Regional plots for cohort MERGE model invn Adjusted Age+Age2+BMI (Continued on next page)



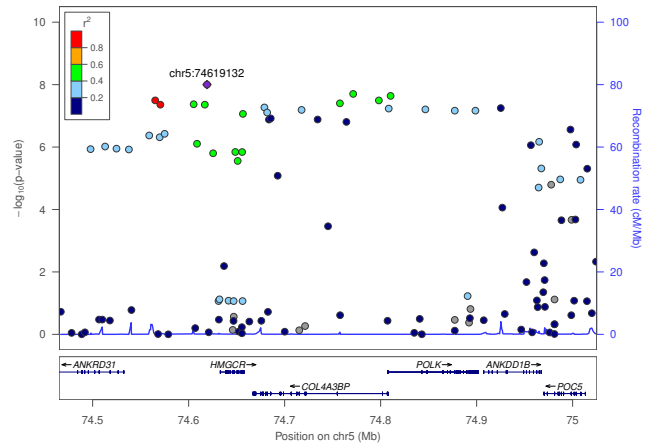
(g) rs116816976 ±100kb



(h) rs3093032 ±100kb



(i) rs10043960 ±100kb



(j) rs13354746 ±100kb

Figure 22: Regional plots for cohort MERGE model invn Adjusted Age+Age2+BMI (Continued)

### 8.4 Previously identified risk loci

Table 27 shows statistics from the MERGE cohort for 50 loci that were shown to be significantly associated with LDL Cholesterol in the 2013 Nature Genetics paper by Willer et al [14]. Where a previously reported variant was not genotyped in the study (indicated by  $\bar{R}^2 < 1$ ), if available, a tagging variant in LD with the reported variant ( $\bar{R}^2 \geq 0.7$  and within 250kb) was provided. Tags were identified using 1000 Genomes data. There are 49 variants that show at least nominal significance ( $p < 0.05$ ) in this study. Out of the 50 variants in both studies, 50 exhibit the same direction of effect with the known result (binomial test  $p = 8.88e - 16$ ).





## 9 Body Mass Index (BMI)

### 9.1 Summary

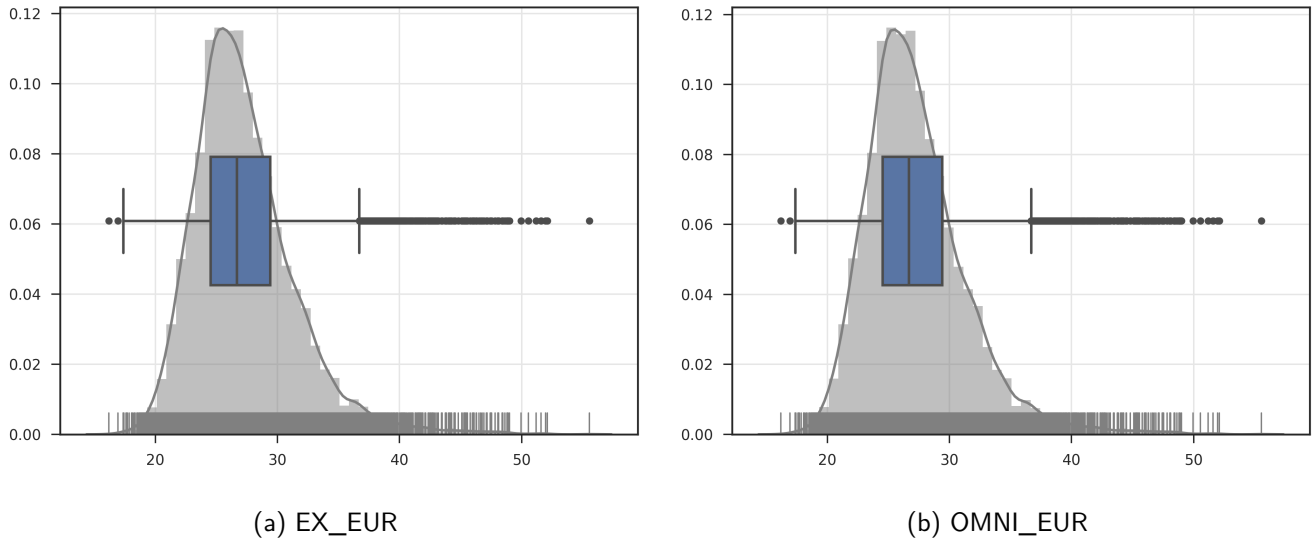


Figure 23: Distribution of BMI in cohort-level analyses

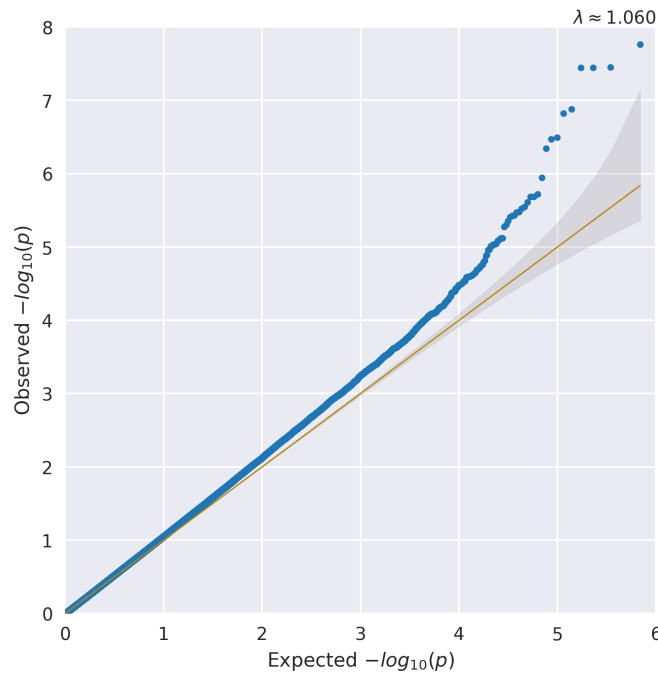
Table 28: Summary of samples removed from Body Mass Index analysis by cohort and model

Cohort	Array	Ancestry	Trans	Covars	Total	-SampleQc	-missObs	-Kinship	-PcOutlier
EX_EUR	EX	EUR	invn	Age+Age2	10071	36	4	1518	1
OMNI_EUR	OMNI	EUR	invn	Age+Age2	10048	69	4	1495	1

Table 29: Summary of samples remaining for Body Mass Index analysis by cohort and model

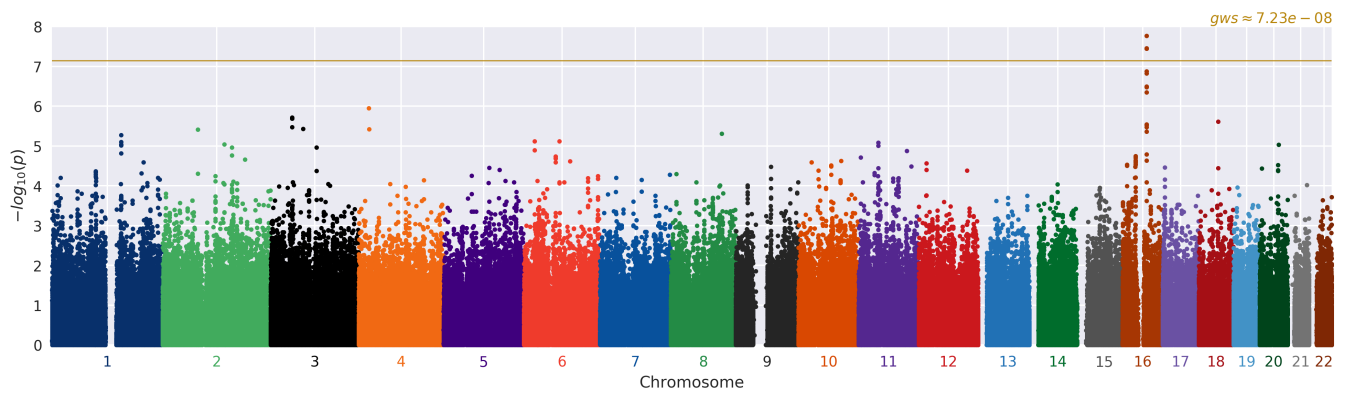
Cohort	Array	Ancestry	Trans	Covars	PCs	N	Male	Female	Max	Min	$\mu$	$\bar{x}$	$\sigma$
EX_EUR	EX	EUR	invn	Age+Age2	5	8512	8512	0	55.54	16.947	27.293	26.644	4.205
OMNI_EUR	OMNI	EUR	invn	Age+Age2	9	8479	8479	0	55.54	16.947	27.284	26.673	4.192

### 9.2 Calibration



(a) invn Adjusted Age+Age2

Figure 24: QQ plots for BMI in the MERGE analysis



(a) invn Adjusted Age+Age2

Figure 25: Manhattan plots for BMI in the MERGE analysis

### 9.3 Top associations

Table 30: Top variants in the MERGE invn Adjusted Age+Age2 model (**bold** variants indicate previously identified associations)

CHR	POS	ID	EA	OA	GENE <sub>CLOSEST</sub>	COHORT	N	MALE	FEMALE	MAC	FREQ	EFFECT	STDERR	P
16	53825488	<b>rs9941349</b>	T	C	<b>FTO</b>	EX_EUR	8,510	8,510	0	6,978	0.59	$8.81 \cdot 10^{-2}$	$1.56 \cdot 10^{-2}$	$1.71 \cdot 10^{-8}$
4	23165631	rs11932311	C	T	GBA3	OMNI_EUR	8,470	8,470	0	5,274	0.689	$8 \cdot 10^{-2}$	$1.64 \cdot 10^{-2}$	$1.13 \cdot 10^{-6}$
3	48487338	rs9876781	G	A	ATRIP	EX_EUR	8,512	8,512	0	6,561	0.615	$7.38 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$	$1.91 \cdot 10^{-6}$
3	48419897	rs6442117	T	C	FBXW12	EX_EUR	8,512	8,512	0	6,548	0.385	$7.35 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$	$2.06 \cdot 10^{-6}$
18	43948285	rs6507697	C	T	RNF165	OMNI_EUR	8,479	8,479	0	5,130	0.303	$7.93 \cdot 10^{-2}$	$1.68 \cdot 10^{-2}$	$2.44 \cdot 10^{-6}$
3	72690645	rs4676890	T	C	SHQ1	OMNI_EUR	8,478	8,478	0	3,904	0.77	$8.42 \cdot 10^{-2}$	$1.82 \cdot 10^{-2}$	$3.68 \cdot 10^{-6}$
2	79039355	rs896530	A	G	REG3G	OMNI_EUR	8,479	8,479	0	4,781	0.282	$7.83 \cdot 10^{-2}$	$1.7 \cdot 10^{-2}$	$3.9 \cdot 10^{-6}$
8	115157455	rs11988087	T	C	CSMD3	OMNI_EUR	8,478	8,478	0	728	0.957	0.172	$3.76 \cdot 10^{-2}$	$4.96 \cdot 10^{-6}$
1	155508882	rs6696888	A	G	ASH1L	EX_EUR	8,511	8,511	0	7,063	0.415	$7.15 \cdot 10^{-2}$	$1.57 \cdot 10^{-2}$	$5.29 \cdot 10^{-6}$
6	23998157	rs554400	A	G	NRSN1	OMNI_EUR	8,472	8,472	0	1,802	0.894	0.112	$2.49 \cdot 10^{-2}$	$7.59 \cdot 10^{-6}$
6	79835010	rs6454096	A	G	PHIP	OMNI_EUR	8,443	8,443	0	6,825	0.596	$7.05 \cdot 10^{-2}$	$1.58 \cdot 10^{-2}$	$7.62 \cdot 10^{-6}$
1	155873314	rs2993207	C	T	RIT1	OMNI_EUR	8,458	8,458	0	6,295	0.372	$7.1 \cdot 10^{-2}$	$1.59 \cdot 10^{-2}$	$7.98 \cdot 10^{-6}$
11	44098000	rs178524	C	T	ACCS	OMNI_EUR	8,476	8,476	0	5,205	0.693	$7.4 \cdot 10^{-2}$	$1.66 \cdot 10^{-2}$	$8.2 \cdot 10^{-6}$
1	155860091	rs1749405	A	G	SYT11	OMNI_EUR	8,479	8,479	0	6,534	0.385	$7 \cdot 10^{-2}$	$1.58 \cdot 10^{-2}$	$9.04 \cdot 10^{-6}$
2	138716177	rs62168710	A	G	HNMT	EX_EUR	8,512	8,512	0	8,166	0.48	$6.84 \cdot 10^{-2}$	$1.54 \cdot 10^{-2}$	$9.1 \cdot 10^{-6}$
20	42720099	rs932415	A	C	JPH2	OMNI_EUR	8,476	8,476	0	3,744	0.779	$8.21 \cdot 10^{-2}$	$1.85 \cdot 10^{-2}$	$9.25 \cdot 10^{-6}$
1	155891045	rs625658	C	A	KIAA0907	OMNI_EUR	8,479	8,479	0	6,317	0.373	$7.03 \cdot 10^{-2}$	$1.59 \cdot 10^{-2}$	$9.66 \cdot 10^{-6}$
11	44142934	rs7935138	C	T	EXT2	OMNI_EUR	8,478	8,478	0	5,416	0.681	$7.27 \cdot 10^{-2}$	$1.64 \cdot 10^{-2}$	$9.81 \cdot 10^{-6}$
3	103057272	rs9862353	C	A	ZPLD1	OMNI_EUR	8,478	8,478	0	4,017	0.763	$7.88 \cdot 10^{-2}$	$1.79 \cdot 10^{-2}$	$1.09 \cdot 10^{-5}$
2	155423435	rs4461221	T	C	GALNT13	OMNI_EUR	8,479	8,479	0	6,209	0.366	$7.01 \cdot 10^{-2}$	$1.59 \cdot 10^{-2}$	$1.1 \cdot 10^{-5}$

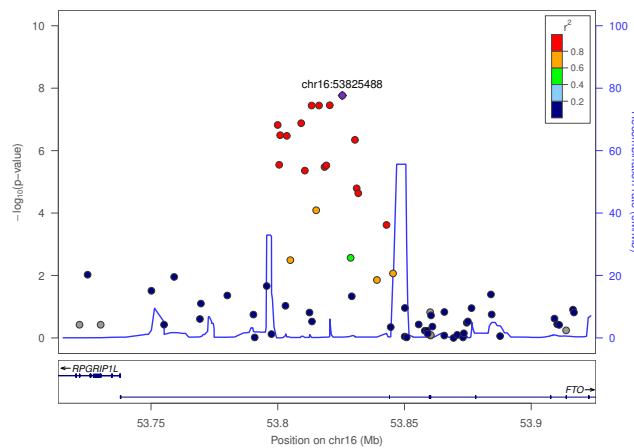


Figure 26: Regional plot for cohort MERGE model invn Adjusted Age+Age2: rs9941349 ±100kb

### 9.4 Previously identified risk loci

Table 31 shows statistics from the MERGE cohort for 50 loci that were shown to be significantly associated with Body Mass Index in the 2015 Nature paper by Locke et al [13]. Where a previously reported variant was not genotyped in the study (indicated by  $\bar{R}^2 < 1$ ), if available, a tagging variant in LD with the reported variant



## 10 Acknowledgements

We would like to acknowledge the following people for their significant contributions to this work.

Ryan Koesterer

Maria Costanzo

Lizz Caulkins

Noel Burt

Jason Flannick

Miriam Udler

Alisa Manning

Jose Florez

## 11 References

- [1] Plink1.9, <https://www.cog-genomics.org/plink2>.
- [2] Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res.* 2002 Jun;12(6):996-1006. LiftOver [http://hgdownload.soe.ucsc.edu/admin/exe/linux.x86\\_64/liftOver](http://hgdownload.soe.ucsc.edu/admin/exe/linux.x86_64/liftOver).
- [3] Conomos MP. GENetic ESTimation and Inference in Structured samples (GENESIS): Statistical methods for analyzing genetic data from samples with population structure and/or relatedness, <https://www.rdocumentation.org/packages/GENESIS/versions/2.2.2>.
- [4] <http://people.virginia.edu/~wc9c/KING/>.
- [5] 1000 Genomes Phase 3 v5, [https://mathgen.stats.ox.ac.uk/impute/1000GP\\_Phase3.html](https://mathgen.stats.ox.ac.uk/impute/1000GP_Phase3.html).
- [6] Seed C, Bloemendal A, Bloom JM, Goldstein JI, King D, Poterba T, Neale BM. Hail: An Open-Source Framework for Scalable Genetic Data Analysis. In preparation. <https://github.com/hail-is/hail>.
- [7] Morris A, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet.* 2012 Sep; 44(9): 981-990. Published online 2012 Aug 12. doi: 10.1038/ng.2383
- [8] Kettunen J, et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. *Nature Communications.* 2016; 11122(2016). doi:10.1038/ncomms11122
- [9] Scott RA, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet.* 2012 Sep;44(9):991-1005. doi: 10.1038/ng.2385. Epub 2012 Aug 12
- [10] Ehret G, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature.* 2011 Oct; 478: 103-109. doi:10.1038/nature10405
- [11] Willer C, et al. Discovery and Refinement of Loci Associated with Lipid Levels. *Nature Genetics.* 2013 Nov; 45(11): 1274-1283. doi:10.1038/ng.2797
- [12] Pattaro C, et al. Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. *Nat Comm.* 2016 Jan; 7:10023; Published online 2016 Jan 21. doi: 10.1038/ncomms10023
- [13] Locke A, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature.* 2015 Feb; 518(7538): 197-206. doi:10.1038/nature14177
- [14] Willer C, et al. Discovery and Refinement of Loci Associated with Lipid Levels. *Nature Genetics.* 2013 Nov; 45(11): 1274-1283. doi:10.1038/ng.2797

- [15] Scott RA, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet.* 2012 Sep;44(9):991-1005. doi: 10.1038/ng.2385. Epub 2012 Aug 12
- [16] Soranzo N, et al. Common variants at 10 genomic loci influence hemoglobin A1(C) levels via glycemic and nonglycemic pathways. *Diabetes.* 2010 Dec;59(12):3229-39. doi: 10.2337/db10-0502. Epub 2010 Sep 21
- [17] Ehret G, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature.* 2011 Oct; 478: 103-109. doi:10.1038/nature10405
- [18] Gilbert C, Ruebenacker O, Koesterer R, Massung J, Flannick J. Loamstream. loamstream 1.4-SNAPSHOT (1.3-211-g92f442f) branch: master commit: 92f442fb6952fbc26e474fbf49b6db15a7370677 built on: 2019-05-03T16:01:23.778Z. <https://github.com/broadinstitute/dig-loam-stream>.
- [19] Koesterer R, Gilbert C, Ruebenacker O, Massung J, Flannick J. AMP-DCC Data Analysis Pipeline. dig-loam-2.5.19. <https://github.com/broadinstitute/dig-loam>.