nature genetics

LETTERS

https://doi.org/10.1038/s41588-018-0157-1

Genome-wide association and HLA fine-mapping studies identify risk loci and genetic pathways underlying allergic rhinitis

Allergic rhinitis is the most common clinical presentation of allergy, affecting 400 million people worldwide, with increasing incidence in westernized countries[1,2]. To elucidate the genetic architecture and understand the underlying disease mechanisms, we carried out a meta-analysis of allergic rhinitis in 59,762 cases and 152,358 controls of European ancestry and identified a total of 41 risk loci for allergic rhinitis, including 20 loci not previously associated with allergic rhinitis, which were confirmed in a replication phase of 60,720 cases and 618,527 controls. Functional annotation implicated genes involved in various immune pathways, and fine mapping of the HLA region suggested amino acid variants important for antigen binding. We further performed genome-wide association study (GWAS) analyses of allergic sensitization against inhalant allergens and nonallergic rhinitis, which suggested shared genetic mechanisms across rhinitis-related traits. Future studies of the identified loci and genes might identify novel targets for treatment and prevention of allergic rhinitis.

Allergic rhinitis, an inflammatory disorder of the nasal mucosa, is mediated by allergic hypersensitivity responses to environmental allergens[1]. It has substantial adverse effects on quality of life and health-care expenditures. The underlying causes of allergic rhinitis are still not understood, and prevention of the disease is not possible. The heritability of allergic rhinitis has been estimated to be more than 65% (refs [3,4]). Seven loci have been associated with allergic rhinitis in GWAS on allergic rhinitis, and other loci have been suggested from GWAS on related traits such as self-reported allergy, asthma plus hay fever, or allergic sensitization[5–9]; however only a few of these findings have been replicated.

We carried out a large-scale meta-GWAS on allergic rhinitis including a discovery meta-analysis of 16,531,985 genetic markers from 18 studies comprising 59,762 cases and 152,358 controls of primarily European ancestry (Supplementary Table 1 and cohort-recruitment details in Supplementary Note). We report the genetic heritability on the liability scale of allergic rhinitis to be at least 7.8% (assuming 10% disease prevalence), with a genomic inflation of 1.048 (Supplementary Fig. 1). We identified 42 genetic loci with index markers below genome-wide significance ($P < 5 \times 10^{-8}$), of which 21 have previously been reported in relation to allergic rhinitis or other inhalant allergy[6–9] (Fig. 1, Tables 1 and 2, and Supplementary Figs. 2 and 3).

One study (23andMe) had a proportionally large weight (~80%) in the discovery phase. Overall, there was good agreement between 23andMe and the other studies with respect to effect size and direction and regional association patterns (Supplementary Table 2 and Supplementary Figs. 4 and 5), and the genetic correlation was 0.80 ($P < 2 \times 10^{-17}$). The heterogeneity between 23andMe and the remaining studies was statistically significant ($P < 0.05$) for 7 of 42 loci, in most cases because of a smaller effect size in 23andMe. This finding was probably due to many non-23andMe studies using a more robust phenotype definition of doctor-diagnosed allergic rhinitis (Supplementary Table 3), which tended to result in larger effect sizes (Supplementary Table 4).

The index markers from a total of 25 loci that had not previously been associated with allergic rhinitis or other inhalant allergy were carried forward to the replication phase. These markers included 16 loci that showed genome-wide-significant association in the discovery phase and evidence of association ($P < 0.05$) in both 23andMe and non-23andMe studies (Supplementary Table 2), and an additional nine loci that were selected from the *P*-value stratum between 5×10^{-8} and 1×10^{-6}, on the basis of enrichment of gene sets involved in immunological signaling (Supplementary Table 5). We sought replication in another ten studies with 60,720 cases and 618,527 controls. Of the 25 loci, 20 reached a Bonferroni-corrected significance threshold of 0.05 ($P < 0.0019$) in a meta-analysis of replication studies (Fig. 1 (blue) and Table 2) and also reached genome-wide significance in the combined fixed-effect meta-analysis of discovery and replication studies (Table 2). Evidence of heterogeneity was seen for one of these loci (rs1504215), which did not reach statistical significance in the random-effects model (0.95 (0.92; 0.97), $P = 2.83 \times 10^{-7}$; Supplementary Fig. 3).

A conditional analysis of top loci identified 13 additional independent variants at $P < 1 \times 10^{-5}$, four of which were genome-wide significant (near *WDR36*, *HLA-DQB1*, *IL1RL1*, and *LPP*; Supplementary Table 6 and Supplementary Fig. 5).

To gain insight into the functional consequences of known and novel loci, we used a number of data sources, including: (i) 11 expression quantitative trait locus (eQTL) sets and one methylation quantitative trait locus (meQTL) set from blood and blood subsets; (ii) two eQTL sets and one meQTL set from lung tissue; and (iii) data on enhancer–promoter interactions in 15 different blood subsets. We found support for regulatory effects on coding genes for 33 of the 41 loci. Many loci showed evidence of regulatory effects across a wide range of immune-cell types (including B and T cells), whereas other loci seemed to be cell-type specific (Supplementary Table 7). Calculating the 'credible set' of markers for each locus through a Bayesian approach that selected markers likely to contain the causal disease-associated markers (Supplementary Table 8) and then looking up those markers in the Variant Effect Predictor database generated a list of 17 markers producing amino acid changes, including deleterious changes in NUSAP1, SULT1A1 and PLCL, as predicted by SIFT (Supplementary Table 9).

The major histocompatibility complex (MHC) on chromosome 6p contained some of the strongest association signals in the GWAS, including independent signals located around *HLA-DQB* and *HLA-B*. The top variant at *HLA-DQB* was an eQTL for several HLA genes, including *HLA-DQB1*, *HLA-DQA1*, *HLA-DQA2*, and *HLA-DRB1* in immune and/or lung tissue, and the top variant at *HLA-B* was

A full list of authors and affiliations appears at the end of the paper.

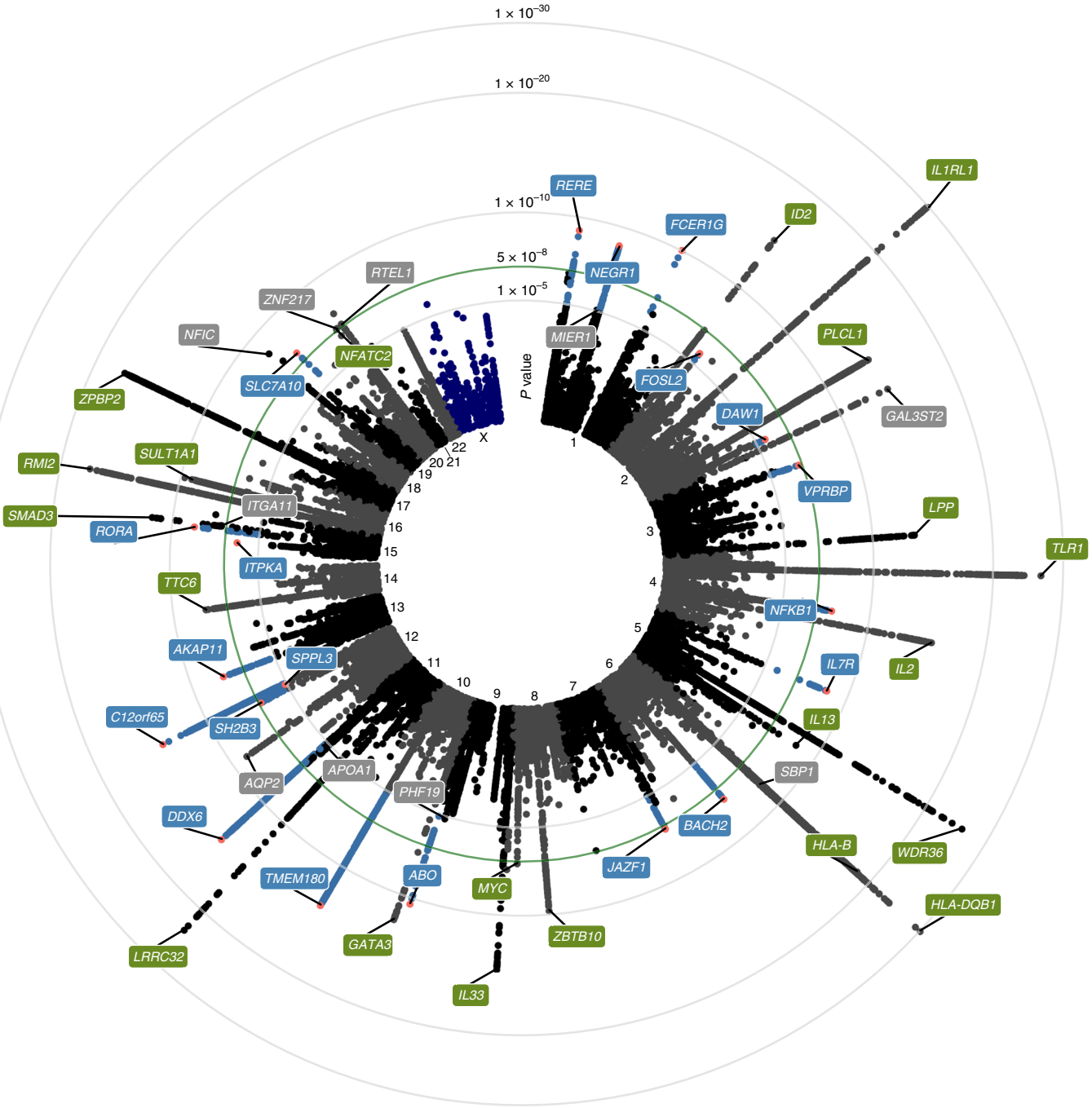


Fig. 1 | Manhattan plot of the meta-GWAS discovery phase. Circular plot of *P* values, calculated from the two-tailed *z* score from an inverse-variance-weighted fixed-effect meta-analysis of association of 16,531,985 genetic markers with allergic rhinitis from the discovery phase, including 212,120 individuals. Only markers with $P<1\times10^{-3}$ are shown. Labels indicate the nearest gene name for the index marker in the locus (the marker with the lowest *P* value). Green labels indicate loci previously associated with allergy; blue labels indicate novel allergic rhinitis–associated loci; gray labels indicate novel loci that were not carried forward to the replication phase. Green line indicates the level of genome-wide significance ($P=5\times10^{-8}$).

an eQTL for *MICA* (Supplementary Table 7). In addition, we found associations with several classical HLA alleles, including *HLA-DQB1*02:02*, *HLA-DQB1*03:01*, *HLA-DRB1*04:01*, and *HLA-C*04:01*, which were in weak linkage disequilibrium (LD) ($r^2<0.1$) with the GWAS top SNPs (Supplementary Tables 10 and 11), and strong associations with well-imputed amino acid variants, including HLA-DQB1 His30 ($P=2.06\times10^{-28}$, odds ratio (OR) = 0.91) and HLA-B Asp116/His116-/Leu116 ($P=6.00\times10^{-13}$, OR = 1.06) (Supplementary Tables 12 and 13). Within HLA-DQB1, the amino acid variant was in moderate LD ($r^2=0.71$) with the GWAS top SNP and accounted for most of the SNP association (rs34004019, $P=2.18\times10^{-28}$, OR = 0.88, conditional $P=1.35\times10^{-3}$). Within *HLA-B*, the strongest associated amino acid variant was only in weak LD ($r^2=0.23$) with the top SNP and accounted for a small part of the SNP association (rs2428494, $P=3.99\times10^{-15}$, OR = 1.07, conditional $P=3.23\times10^{-10}$). Importantly, the strongest associated amino acid variants in HLA-DQB1 and HLA-B, respectively, were both located in the peptide-binding pockets with a high likelihood of affecting MHC–peptide interaction (Fig. 2). MHC class II molecules, including HLA-DQ, are known for their roles in allergen binding and type

Table 1 | Association results of index markers (variant with lowest *P* value for each locus) previously reported in relation to allergic rhinitis or other inhalant allergy

			Discovery						
Variant	Locus	Nearest genes	EA/OA	EAF	*n* (studies)	OR	95% CI	*P*	Het. *P*
Known									
rs34004019	6p21.32	*HLA-DQB1, HLA-DQA1*	G/A	0.27	196,951 (11)	0.89	0.87–0.90	1.00×10^{-30}	0.41
rs950881	2q12.1	*IL1RL1, IL1RL1*	T/G	0.15	212,120 (18)	0.88	0.87–0.90	1.74×10^{-30}	0.91
rs5743618	4p14	*TLR1, TLR10*	A/C	0.27	210,652 (17)	0.90	0.89–0.92	4.38×10^{-27}	0.70
rs1438673	5q22.1	*CAMK4, WDR36*	C/T	0.50	212,120 (18)	1.08	1.07–1.10	3.15×10^{-26}	0.26
rs7936323	11q13.5	*LRRC32, C11orf30*	A/G	0.48	212,120 (18)	1.08	1.06–1.09	6.53×10^{-24}	0.0001
rs2428494	6p21.33	*HLA-B, HLA-C*	A/T	0.42	195,753 (12)	1.08	1.06–1.09	7.01×10^{-19}	0.25
rs11644510	16p13.13	*RMI2, CLEC16A*	T/C	0.37	212,120 (18)	0.93	0.92–0.95	1.58×10^{-17}	0.65
rs12939457	17q12	*GSDMB, ZPBP2*	C/T	0.44	212,120 (18)	0.94	0.92–0.95	2.35×10^{-17}	0.02
rs148505069	4q27	*IL21, IL2*	G/A	0.33	212,120 (18)	1.07	1.05–1.08	2.54×10^{-15}	0.02
rs13395467	2p25.1	*ID2, RNF144A*	G/A	0.28	212,120 (18)	0.94	0.92–0.95	9.93×10^{-15}	0.61
rs9775039	9p24.1	*IL33, RANBP6*	A/G	0.16	212,120 (18)	1.08	1.06–1.10	2.22×10^{-14}	0.40
rs2164068	2q33.1	*PLCL1*	A/T	0.49	212,120 (18)	0.94	0.93–0.96	4.21×10^{-14}	0.82
rs2030519	3q28	*TPRG1, LPP*	G/A	0.49	212,120 (18)	1.06	1.04–1.07	1.83×10^{-13}	0.12
rs11256017	10p14	*CELF2, GATA3*	T/C	0.18	212,120 (18)	1.07	1.05–1.09	2.72×10^{-12}	0.60
rs17294280	15q22.33	*AAGAB, SMAD3*	G/A	0.25	212,120 (18)	1.07	1.05–1.09	5.97×10^{-12}	0.07
rs7824993	8q21.13	*ZBTB10, TPD52*	A/G	0.37	212,120 (18)	1.05	1.04–1.07	1.86×10^{-10}	0.56
rs9282864	16p11.2	*SULT1A1, SULT1A2*	C/A	0.33	208,761 (16)	0.94	0.93–0.96	4.69×10^{-10}	0.03
rs9687749	5q31.1	*IL13, RAD50*	T/G	0.44	207,604 (16)	1.06	1.04–1.09	1.84×10^{-9}	0.19
rs61977073	14q21.1	*TTC6*	G/A	0.22	212,120 (18)	1.06	1.04–1.08	5.78×10^{-9}	0.05
rs6470578	8q24.21	*TMEM75, MYC*	T/A	0.28	212,120 (18)	1.05	1.03–1.07	4.36×10^{-8}	0.02
rs3787184	20q13.2	*NFATC2, KCNG1*	G/A	0.19	207,604 (16)	0.94	0.93–0.96	4.76×10^{-8}	0.69

'Nearest genes' denotes the nearest up- and downstream gene (for intergenic variants with two genes listed) or surrounding gene (for intronic variants with one gene listed), with the exception of rs5743618, an exonic missense variant within *TLR1*. EA/OA, effect allele/other allele; EAF, effect-allele frequency. *P* values were calculated from two-tailed *z* scores from an inverse-variance-weighted fixed-effect meta-analysis. Het. *P*, *P* value for heterogeneity obtained from Cochrane's *Q* test.

2 helper T cell (T_H2)-driven immune responses[10]; our results therefore suggest that the GWAS signal at this locus involves structural changes related to allergen binding properties. This effect might be in addition to gene-regulatory effects similar to those found for autoimmune disease[11,12]. Most of the 20 loci not previously associated with allergic rhinitis implicate genes with a known role in the immune system, including *IL7R*[13,14], *SH2B3* (ref. [15]), *CEBPA-CEBPG*[16,17], *CXCR5* (ref. [18]), *FCER1G*, *NFKB1* (ref. [19]), *BACH2* (refs [20,21]), *TYRO3* (ref. [22]), *LTK*[23], *VPRBP*[24], *SPPL3* (ref. [25]), *OASL*[26], *RORA*[27], and *TNFSF11* (ref. [28]). Other loci implicated genes with no clear function in allergic rhinitis pathogenesis. These included one of the strongest associated loci in this meta-analysis at 12q24.31, with the top signal located between *CDK2AP1* and *C12orf65*, and containing cis-eQTLs in blood and lung tissue for several genes and evidence of enhancer–promoter interaction with *DDX55* in various immune cells (Supplementary Table 14 and further locus description in Supplementary Note). Concomitantly with the current study, a GWAS combining asthma, eczema, and allergic rhinitis was conducted[29]. Most (15/20) of the identified allergic rhinitis–associated loci in our study were also suggested in the previous, less specific GWAS[29] (as indicated in Table 2), whereas many suggested loci from the previous GWAS were not identified in our study. Asthma, eczema, and allergic rhinitis are related but distinct disease entities, often with separate disease mechanisms; for example, allergic sensitization is present in only 50% of children with asthma[30] and 35% of children with eczema[31]. Our results therefore complement those from the less specific 'atopic phenotype' GWAS[29] by pinpointing loci specifically associated and replicated in relation to allergic rhinitis.

Allergic rhinitis–associated loci were significantly enriched ($P<1\times10^{-5}$) in variants reported to be associated with autoimmune disorders. Reported autoimmune variants were located within 1 Mb of 31 (76%) of the 41 allergic rhinitis–associated loci. For 24 of these, an autoimmune top SNP was also associated with allergic rhinitis, and for 12 of these, the autoimmune top SNP was in LD ($r^2>0.5$) with the allergic rhinitis top SNP (Supplementary Table 15). For approximately half of these, the direction of the effect was the same for the autoimmune and allergic rhinitis top SNP, in line with findings from a previous study[32], thus underlining the complex genetic relationship between allergic rhinitis and autoimmunity, which might involve shared as well as diverging molecular mechanisms.

Assessment of the enrichment of allergic rhinitis–associated variant burden in open chromatin, as defined by DNase-hypersensitive sites, showed a clear enrichment in several blood and immune-cell subsets, with the largest enrichment in T cells (CD3 expressing), B cells (CD19 expressing), and T and NK cells (CD56 expressing) (Fig. 3, Supplementary Table 16 and Supplementary Fig. 6). We also probed tissue enrichment by using gene-expression data from a wide number of sources and observed enrichment of allergic rhinitis–associated genes in blood and immune-cell subsets, as well as in tissues of the respiratory system, including the oropharynx and the respiratory and nasal mucosa (Supplementary Table 17).

To explore biological connections and identify new pathways associated with allergic rhinitis, we combined all genes suggested from the eQTL and meQTL analyses, enhancer–promoter interactions, and localization within the top loci. The resultant prioritized gene set consisted of 255 genes, of which 89 (~36%) were

Table 2 | Association results of index markers (variant with lowest *P* value for each locus) not previously associated with allergic rhinitis reaching a Bonferroni-corrected significance threshold of 0.05 in the replication phase

				Discovery						Replication					Combined				
Variant	Locus	Nearest genes	EA/OA	EAF	*n* (studies)	OR	95% CI	*P*	Het. *P*	*n* (studies)	OR	95% CI	*P*	FWER	*n* (studies)	OR	95% CI	*P*	Het. *P*
rs7717955[a]	5p13.2	*CAPSL, IL7R*	T/C	0.27	212,120 (18)	0.95	0.93–0.96	1.50×10^{-9}	0.24	679,247 (10)	0.93	0.91–0.94	4.09×10^{-25}	1.06×10^{-23}	891,367 (28)	0.94	0.93–0.95	3.78×10^{-32}	0.09
rs63406760[a]	12q24.31	*CDK2AP1, C12orf65*	G/–	0.26	210,652 (17)	0.93	0.91–0.95	5.12×10^{-14}	0.91	675,338 (7)	0.95	0.93–0.96	3.27×10^{-12}	8.51×10^{-11}	885,990 (24)	0.94	0.93–0.95	2.54×10^{-24}	0.89
rs1504215[a]	6q15	*BACH2, GJA10*	A/G	0.34	207,604 (16)	0.95	0.94–0.97	1.49×10^{-8}	0.02	679,247 (10)	0.95	0.94–0.97	1.99×10^{-11}	5.17×10^{-10}	886,851 (26)	0.95	0.94–0.96	1.54×10^{-18}	0.05
rs28361986[a]	11q23.3	*CXCR5, DDX6*	A/T	0.20	212,120 (18)	0.93	0.91–0.95	1.81×10^{-14}	0.87	675,919 (8)	0.94	0.93–0.96	7.92×10^{-11}	2.06×10^{-9}	888,039 (26)	0.94	0.92–0.95	2.32×10^{-23}	0.91
rs2070902[a]	1q23.3	*AL590714.1, FCER1G*	T/C	0.25	212,120 (18)	1.06	1.04–1.08	1.03×10^{-10}	0.18	679,247 (10)	1.05	1.03–1.06	7.27×10^{-10}	1.89×10^{-8}	891,367 (28)	1.05	1.04–1.06	6.19×10^{-19}	0.23
rs111371454[a]	15q15.1	*ITPKA, RTF1*	G/A	0.21	212,120 (18)	1.06	1.03–1.08	1.65×10^{-7}	0.17	675,338 (7)	1.04	1.03–1.06	8.47×10^{-9}	2.20×10^{-7}	887,458 (25)	1.05	1.03–1.06	1.28×10^{-14}	0.22
rs12509403[a]	4q24	*MANBA, NFKB1*	T/C	0.32	212,120 (18)	0.95	0.94–0.97	9.97×10^{-9}	0.27	679,247 (10)	0.96	0.95–0.97	1.86×10^{-8}	4.84×10^{-7}	891,367 (28)	0.96	0.95–0.97	1.17×10^{-15}	0.39
rs9648346[a]	7p15.1	*JAZF1, TAX1BP1*	G/C	0.22	207,604 (16)	1.05	1.03–1.07	3.62×10^{-8}	0.74	679,247 (10)	1.04	1.03–1.06	1.39×10^{-7}	3.63×10^{-6}	886,851 (26)	1.05	1.03–1.06	3.30×10^{-14}	0.48
rs35350651[a]	12q24.12	*ATXN2, SH2B3*	C/–	0.49	206,136 (15)	1.04	1.03–1.06	6.63×10^{-8}	0.60	672,701 (6)	1.04	1.02–1.05	1.41×10^{-7}	3.66×10^{-6}	878,837 (21)	1.04	1.03–1.05	5.82×10^{-14}	0.43
rs2519093[a]	9q34.2	*ABO, OBP2B*	T/C	0.20	212,120 (18)	1.06	1.04–1.09	4.96×10^{-11}	0.38	675,919 (8)	1.04	1.03–1.06	2.96×10^{-7}	7.68×10^{-6}	888,039 (26)	1.05	1.04–1.07	2.79×10^{-16}	0.61
rs62257549	3p21.2	*VPRBP*	A/G	0.20	212,120 (18)	0.95	0.93–0.97	7.13×10^{-8}	0.45	677,615 (9)	0.96	0.94–0.97	3.37×10^{-7}	8.76×10^{-6}	889,735 (27)	0.95	0.94–0.97	1.84×10^{-13}	0.53
rs11677002	2p23.2	*FOSL2, RBKS*	C/T	0.45	212,120 (18)	0.96	0.95–0.98	3.80×10^{-7}	0.21	679,247 (10)	0.97	0.96–0.98	3.54×10^{-7}	9.20×10^{-6}	891,367 (28)	0.97	0.96–0.97	7.08×10^{-13}	0.36
rs35597970[a]	10q24.32	*ACTR1A, TMEM180*	–/A	0.45	210,652 (17)	1.06	1.04–1.07	1.34×10^{-13}	0.96	676,970 (8)	1.03	1.02–1.05	4.37×10^{-7}	1.14×10^{-5}	887,622 (25)	1.04	1.03–1.05	5.42×10^{-18}	0.53
rs2815765	1p31.1	*LRRIQ3, NEGR1*	T/C	0.37	212,120 (18)	0.95	0.94–0.97	1.18×10^{-9}	0.59	679,247 (10)	0.97	0.95–0.98	6.16×10^{-7}	1.60×10^{-5}	891,367 (28)	0.96	0.95–0.97	9.45×10^{-15}	0.52
rs11671925[a]	19q13.11	*CEBPA, SLC7A10*	A/G	0.17	206,136 (15)	0.94	0.92–0.96	1.80×10^{-8}	0.97	677,551 (9)	0.96	0.94–0.98	2.80×10^{-6}	7.29×10^{-5}	883,687 (24)	0.95	0.94–0.96	5.91×10^{-13}	0.60
rs2461475[a]	12q24.31	*SPPL3, ACADS*	C/T	0.47	212,120 (18)	1.04	1.02–1.05	9.19×10^{-7}	0.97	677,551 (9)	1.03	1.02–1.04	6.52×10^{-6}	0.0002	889,671 (27)	1.03	1.02–1.04	3.81×10^{-11}	0.83
rs6738964[a]	2q36.3	*SPHKAP, DAW1*	G/T	0.24	212,120 (18)	0.96	0.94–0.97	4.51×10^{-7}	0.72	679,247 (10)	0.97	0.96–0.98	4.96×10^{-5}	0.0013	891,367 (28)	0.96	0.95–0.97	1.86×10^{-10}	0.87
rs10519067[a]	15q22.2	*RORA*	A/–	0.13	212,120 (18)	0.93	0.91–0.96	1.78×10^{-9}	0.37	442,354 (7)	0.93	0.90–0.96	7.53×10^{-5}	0.0020	654,474 (25)	0.93	0.92–0.95	5.53×10^{-13}	0.36
rs138050288[a]	1p36.23	*RERE, SLC45A1*	–/CA	0.29	210,652 (17)	1.05	1.04–1.07	5.96×10^{-10}	0.71	675,338 (7)	1.03	1.01–1.04	0.0002	0.0046	885,990 (24)	1.04	1.03–1.05	6.62×10^{-12}	0.63
rs7328203	13q14.11	*TNFSF11, AKAP11*	G/T	0.46	212,120 (18)	1.05	1.03–1.06	5.94×10^{-9}	0.90	677,551 (9)	1.02	1.01–1.04	0.0005	0.0134	889,671 (27)	1.03	1.02–1.04	1.28×10^{-10}	0.78

'Nearest genes' denotes the nearest up- and downstream gene (for intergenic variants with two genes listed) or surrounding gene (for intronic variants with one gene listed), with the exception of rs1504215, an exonic synonymous variant within *BACH2*. EA/OA, effect allele/other allele; EAF, effect-allele frequency. *P* values were calculated from two-tailed *z* scores from an inverse-variance-weighted fixed-effect meta-analysis. Het. *P*, *P* value for heterogeneity obtained from Cochrane's Q test. FWER, family-wise error rate. [a]Variants also reported associated with a combined asthma/eczema/hay fever phenotype by Ferreira et al.[29] (within ±1 Mb).

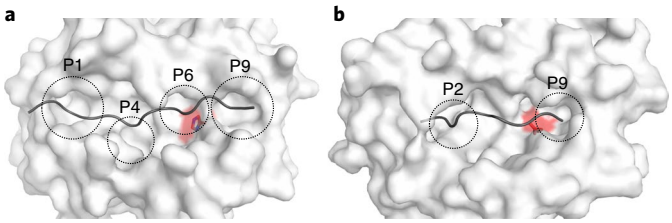


Fig. 2 | Structural visualization of amino acid variants associated with allergic rhinitis. The surface of the MHC molecule is shown in white, and the backbone of the bound peptide is shown in dark gray. The amino acid variant in focus is highlighted in red, and the peptide-binding pockets of the MHC molecule are indicated with dashed circles and are annotated P1–P9. **a**, The amino acid variant with the strongest association with allergic rhinitis was HLA-DQB1 His30 (MHC class II), located close to P6 with a distance of 6 Å to the peptide (excluding the peptide side chain). The protective amino acid variant at this location in relation to allergic rhinitis was histidine, whereas the risk variant was serine. Histidine is positively charged and has a large aromatic ring, whereas serine is not charged and not aromatic. Therefore, this alteration results in a substantial change in the binding-pocket environment. **b**, The strongest allergic rhinitis–associated amino acid variants in HLA-B (MHC class I) were HLA-B Asp116/His116-/Leu116, located close to P9 with a distance of 7 Å to the peptide (excluding the peptide side chain). The proximity to the bound peptide for both variants indicates that they are likely to affect the MHC–peptide interaction and thereby affect which peptides are presented.

present in more than one set (Supplementary Fig. 7). Overall, the full set was enriched in pathways involved in type 1 helper T cell (T_H1) and T_H2 activation (Fig. 4), antigen presentation, cytokine signaling, and inflammatory responses (Supplementary Table 18).

Using the 255 prioritized genes in combination with the STRING database to identify proteins interacting with the proteins encoded by the high-priority genes, we demonstrated a high degree of interaction at the protein level, and several of these proteins, including TNFSF11, NDUFAF1, PD-L1, IL-5, and IL-13, are targets of approved drugs or drugs in development (Fig. 4).

Allergic rhinitis is strongly correlated with allergic sensitization (presence of allergen-specific IgE), but sensitization is often present without allergic rhinitis, thus suggesting that specific mechanisms determine the progression from sensitization to disease. We therefore conducted a GWAS on sensitization to inhalant allergens, comprising 8,040 cases and 16,441 controls from 13 studies (Supplementary Table 1), in what is, to our knowledge, the largest GWAS on allergic sensitization to date[7]. A total of ten loci reached genome-wide significance, including one novel locus near *FASLG* (Supplementary Table 19). The genetic heritability on the liability scale was 17.75% (10% prevalence), a value considerably higher than the heritability of allergic rhinitis and consistent with a more homogeneous phenotype. Look-up of top allergic rhinitis–associated loci in the allergic-sensitization GWAS demonstrated high agreement: 40 of the 41 allergic rhinitis markers showed the same direction of effect, and 28 also showed nominal significance for allergic sensitization (Supplementary Table 20). This result suggests that allergic

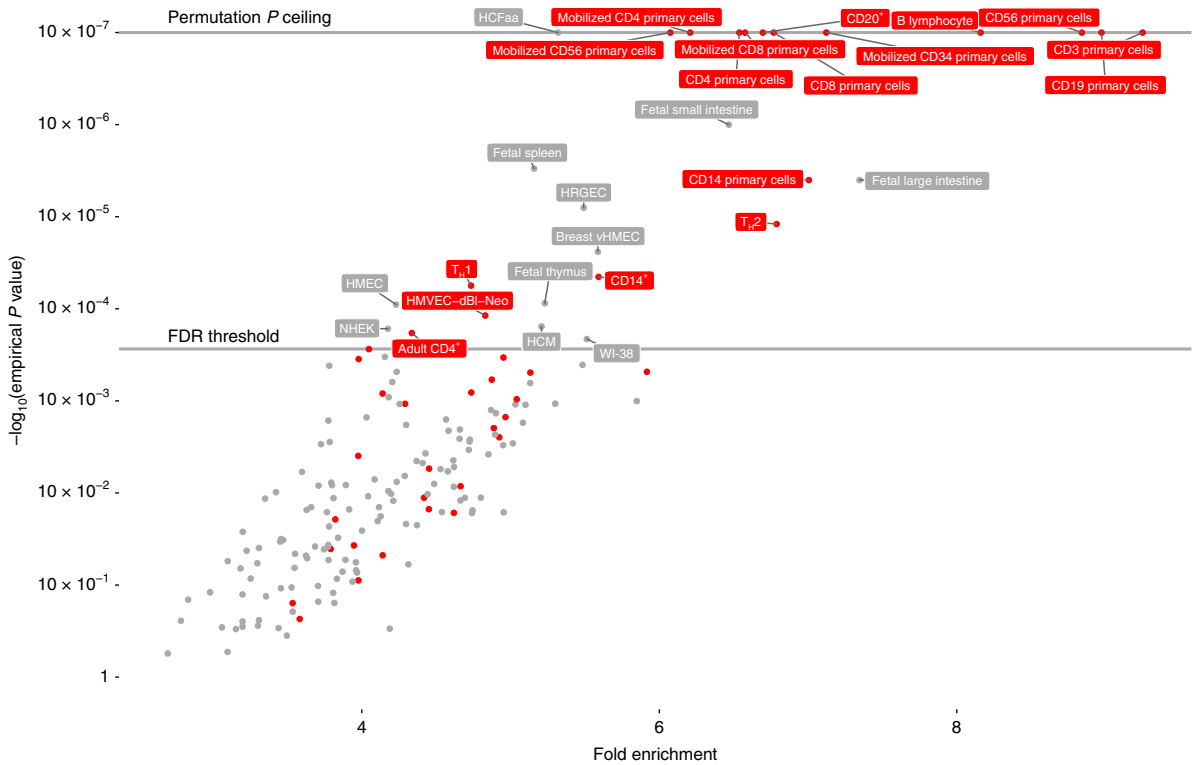


Fig. 3 | Enrichment of allergic rhinitis-associated variants in tissue-specific open chromatin. Enrichment of 16,531,985 genetic variants associated with allergic rhinitis in 212,120 individuals (at $P < 1 \times 10^{-8}$ as a threshold for marker association) in 189 cell types from ENCODE and Roadmap epigenomics data. Enrichment and *P* values were calculated empirically against a permuted genomic background with the GARFIELD tool. Red labels indicate blood and blood-related cell types, and gray labels indicate other cell types. Owing to the number of permutations (1×10^7), empirical *P* values reached a minimum ceiling of $1/(1 \times 10^7)$. The false discovery rate (FDR) threshold was 0.00026. For Epstein–Barr virus–transformed B-lymphocyte cell types (cell type GM****), only the most enriched instance is shown ('B lymphocyte'). NHEK, normal human epidermal keratinocytes; HMEC/vHMEC, mammary epithelial cells; HCM, human cardiac myocytes; WI-38, lung fibroblast derived; HRGEC, human renal glomerular endothelial cells; HCFaa, human cardiac fibroblast–adult atrial cells; HMVEC−dBl−Neo, human microvascular endothelial cells.

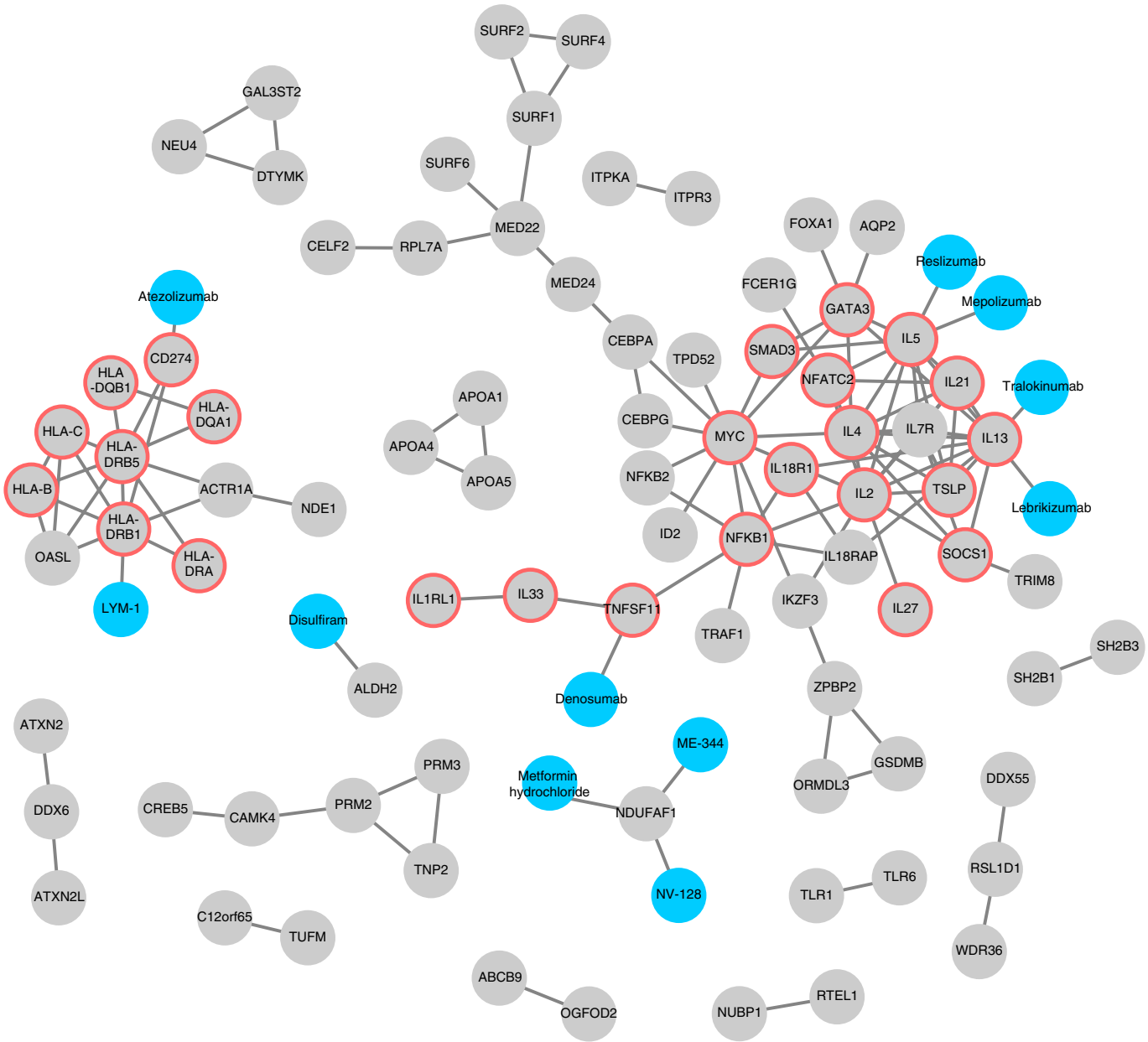


Fig. 4 | Interaction network among drugs and proteins from genes associated with allergic rhinitis. Gray nodes represent locus genes as well as genes prioritized from eQTL/meQTL and PCHiC sources, on the basis of genetic association of 16,531,985 markers with allergic rhinitis in 212,120 individuals. Blue nodes represent drugs from the ChEMBL drug database. Edges represent very high-confidence interactions from the STRING database (for locus–locus interactions) and drug-target evidence (for drug–locus interactions). Red borders indicate genes with protein products that were significantly enriched in the 'T_H1 and T_H2 activation' pathway ($-\log(P) > 19.1$) from ingenuity pathway analysis.

rhinitis and allergic sensitization share biological mechanisms, and that allergic-sensitization loci generally affect systemic allergic sensitization. We compared genetic pathways of allergic rhinitis and allergic sensitization by using the DEPICT tool, which showed overlap in enriched pathways but also differences among the top gene sets: allergic rhinitis gene sets were characterized by B cell, T_H2, and parasite responses, whereas allergic-sensitization gene sets were characterized by a broader activation of cells (Supplementary Fig. 8 and Supplementary Tables 21 and 22).

Nonallergic rhinitis, defined as having rhinitis symptoms without evidence of allergic sensitization, is a common but poorly understood disease entity[33]. We performed what is, to our knowledge, the first GWAS on this phenotype, with the goal of identifying specific rhinitis mechanisms. The analysis included 2,028 cases and 9,606 controls from nine studies but did not identify any risk loci at the genome-wide-significant level. Comparison with the allergic rhinitis results suggested some overlap in susceptibility loci (Supplementary Note and Supplementary Table 23).

We estimated the proportion of allergic rhinitis in the general population that was attributable to the 41 identified allergic rhinitis–associated loci and obtained a conservative population-attributable risk-fraction estimate of 39% (95% confidence interval (CI) 26–50%), by considering the 10% of the population with the lowest genetic-risk scores to represent an 'unexposed' group. The allergic rhinitis prevalence plotted by genetic-risk score (Supplementary Fig. 9) showed a prevalence approximately two times higher in the 7% of the population with the highest risk score than the 7% with the lowest risk score.

Finally, we investigated the genetic correlation of allergic rhinitis with allergic sensitization, asthma[34], and eczema[35] by using LD-score regression. There was a strong correlation between allergic rhinitis and allergic sensitization ($r^2 = 0.73$, $P < 2 \times 10^{-34}$), a moderate correlation with asthma ($r^2 = 0.60$, $P < 3 \times 10^{-14}$), and a weaker correlation with eczema ($r^2 = 0.40$, $P < 2 \times 10^{-7}$).

We tested the identified allergic rhinitis–associated loci for association with allergic rhinitis in non-European cohorts, which showed only a nominally significant association for one locus, but this analysis had limited statistical power, owing to population sizes (Supplementary Table 24).

In conclusion, we expanded the number of established susceptibility loci for allergic rhinitis and highlighted the involvement of allergic rhinitis–susceptibility loci in diverse immune-cell types and both innate and adaptive IgE-related mechanisms. Future studies of novel allergic rhinitis–associated loci might identify targets for the treatment and prevention of this disease.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41588-018-0157-1.

Received: 30 June 2017; Accepted: 10 May 2018;
Published online: 16 July 2018

References

1. Greiner, A. N., Hellings, P. W., Rotiroti, G. & Scadding, G. K. Allergic rhinitis. *Lancet* **378**, 2112–2122 (2011).
2. Björkstén, B., Clayton, T., Ellwood, P., Stewart, A. & Strachan, D. Worldwide time trends for symptoms of rhinitis and conjunctivitis: phase III of the International Study of Asthma and Allergies in Childhood. *Pediatr. Allergy Immunol.* **19**, 110–124 (2008).
3. Willemsen, G., van Beijsterveldt, T. C. E. M., van Baal, C. G. C. M., Postma, D. & Boomsma, D. I. Heritability of self-reported asthma and allergy: a study in adult Dutch twins, siblings and parents. *Twin Res. Hum. Genet.* **11**, 132–142 (2008).
4. Fagnani, C. et al. Heritability and shared genetic effects of asthma and hay fever: an Italian study of young twins. *Twin Res. Hum. Genet.* **11**, 121–131 (2008).
5. Ramasamy, A. et al. A genome-wide meta-analysis of genetic variants associated with allergic rhinitis and grass sensitization and their interaction with birth order. *J. Allergy Clin. Immunol.* **128**, 996–1005 (2011).
6. Hinds, D. A. et al. A genome-wide association meta-analysis of self-reported allergy identifies shared and allergy-specific susceptibility loci. *Nat. Genet.* **45**, 907–911 (2013).
7. Bønnelykke, K. et al. Meta-analysis of genome-wide association studies identifies ten loci influencing allergic sensitization. *Nat. Genet.* **45**, 902–906 (2013).
8. Ferreira, M. A. R. et al. Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. *J. Allergy Clin. Immunol.* **133**, 1564–1571 (2014).
9. Bunyavanich, S. et al. Integrated genome-wide association, coexpression network, and expression single nucleotide polymorphism analysis identifies novel pathway in allergic rhinitis. *BMC Med. Genom.* **7**, 48 (2014).
10. Jahn-Schmid, B., Pickl, W. F. & Bohle, B. Interaction of allergens, major histocompatibility complex molecules, and T cell receptors: a 'ménage à trois' that opens new avenues for therapeutic intervention in type I allergy. *Int. Arch. Allergy Immunol.* **156**, 27–42 (2011).
11. Cavalli, G. et al. MHC class II super-enhancer increases surface expression of HLA-DR and HLA-DQ and affects cytokine production in autoimmune vitiligo. *Proc. Natl Acad. Sci. USA* **113**, 1363–1368 (2016).
12. Hayashi, M. et al. Autoimmune vitiligo is associated with gain-of-function by a transcriptional regulator that elevates expression of HLA-A*02:01 in vivo. *Proc. Natl Acad. Sci. USA* **113**, 1357–1362 (2016).
13. Puel, A., Ziegler, S. F., Buckley, R. H. & Leonard, W. J. Defective IL7R expression in $T^-B^+NK^+$ severe combined immunodeficiency. *Nat. Genet.* **20**, 394–397 (1998).
14. Lundmark, F. et al. Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. *Nat. Genet.* **39**, 1108–1113 (2007).
15. Mori, T. et al. Lnk/Sh2b3 controls the production and function of dendritic cells and regulates the induction of IFN-γ-producing T cells. *J. Immunol.* **193**, 1728–1736 (2014).
16. Scott, L. M., Civin, C. I., Rorth, P. & Friedman, A. D. A novel temporal expression pattern of three C/EBP family members in differentiating myelomonocytic cells. *Blood* **80**, 1725–1735 (1992).
17. Gao, H., Parkin, S., Johnson, P. F. & Schwartz, R. C. C/EBP gamma has a stimulatory role on the IL-6 and IL-8 promoters. *J. Biol. Chem.* **277**, 38827–38837 (2002).
18. León, B. et al. Regulation of T_H2 development by $CXCR5^+$ dendritic cells and lymphotoxin-expressing B cells. *Nat. Immunol.* **13**, 681–690 (2012).
19. Lawrence, T. The nuclear factor NF-kappaB pathway in inflammation. *Cold Spring Harb. Perspect. Biol.* **1**, a001651 (2009).
20. Shinnakasu, R. et al. Regulated selection of germinal-center cells into the memory B cell compartment. *Nat. Immunol.* **17**, 861–869 (2016).
21. Roychoudhuri, R. et al. BACH2 regulates $CD8^+$ T cell differentiation by controlling access of AP-1 factors to enhancers. *Nat. Immunol.* **17**, 851–860 (2016).
22. Rothlin, C. V., Ghosh, S., Zuniga, E. I., Oldstone, M. B. A. & Lemke, G. TAM receptors are pleiotropic inhibitors of the innate immune response. *Cell* **131**, 1124–1136 (2007).
23. Chan, P. Y. et al. The TAM family receptor tyrosine kinase TYRO3 is a negative regulator of type 2 immunity. *Science* **352**, 99–103 (2016).
24. Kassmeier, M. D. et al. VprBP binds full-length RAG1 and is required for B-cell development and V(D)J recombination fidelity. *EMBO J.* **31**, 945–958 (2012).
25. Hamblet, C. E., Makowski, S. L., Tritapoe, J. M. & Pomerantz, J. L. NK cell maturation and cytotoxicity are controlled by the intramembrane aspartyl protease SPPL3. *J. Immunol.* **196**, 2614–2626 (2016).
26. Andersen, J. B., Strandbygård, D. J., Hartmann, R. & Justesen, J. Interaction between the 2′-5′ oligoadenylate synthetase-like protein p59 OASL and the transcriptional repressor methyl CpG-binding protein 1. *Eur. J. Biochem.* **271**, 628–636 (2004).
27. Halim, T. Y. F. et al. Retinoic-acid-receptor-related orphan nuclear receptor alpha is required for natural helper cell development and allergic inflammation. *Immunity* **37**, 463–474 (2012).
28. Anderson, D. M. et al. A homologue of the TNF receptor and its ligand enhance T-cell growth and dendritic-cell function. *Nature* **390**, 175–179 (1997).
29. Ferreira, M. A. et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat. Genet.* **49**, 1752–1757 (2017).
30. Pearce, N., Pekkanen, J. & Beasley, R. How much asthma is really attributable to atopy? *Thorax* **54**, 268–272 (1999).
31. Böhme, M., Wickman, M., Lennart Nordvall, S., Svartengren, M. & Wahlgren, C. F. Family history and risk of atopic dermatitis in children up to 4 years. *Clin. Exp. Allergy* **33**, 1226–1231 (2003).
32. Kreiner, E. et al. Shared genetic variants suggest common pathways in allergy and autoimmune diseases. *J. Allergy Clin. Immunol.* **140**, 771–781 (2017).
33. Bousquet, J. et al. Important research questions in allergy and related diseases: nonallergic rhinitis: a GA2LEN paper. *Allergy* **63**, 842–853 (2008).
34. Moffatt, M. F. et al. Genetic variants regulating *ORMDL3* expression contribute to the risk of childhood asthma. *Nature* **448**, 470–473 (2007).
35. Paternoster, L. et al. Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nat. Genet.* **47**, 1449–1456 (2015).

Acknowledgements

Detailed acknowledgements and funding details for each contributing study are provided in the Supplementary Note.

Author contributions

Study design and management: K.B., J.W., M.S., and D.P.S. Meta-analyses: M.S. and J.W. Manuscript writing: K.B., J.W., M.S., J.A.C., J.T., L.E.J., and M.A.F. Systems biology analyses: J.W., J.A.C., J.T., L.E.J., J.M.M., S.B.-G., and D.T. Data collection, analysis, and design in the individual contributing studies: K.B., J.W., M.S., J.A.C., C.F., A. Abdellaoui, T.S.A., A.C.A., A.F.S.A., J.M.A., A. Arnold, A.B.-L., H. Baurecht, C.E.M.v.B., E.R.B., D.I.B., S. Bunyavanich, E.B., Z.C., I.C., A.C., H.T.d.D., S.C.D., J.D., L.D., M.J.E., W.J.G., C.G., F.G., R.G., H.G., T.H., J. Heinrich, J. Henderson, N.H.-P., D.A.H., P.H., M.I., V.W.V.J., M.-R.J., D.L.J., I.J., M.K., J.K., A.K., Y.-A.L., A.M.L., X.L., F.L.-D., E.M., D.A.M., R.M., D.L.N., E.A.N., T.P., L.P., C.E.P., G.P., M.P.-Y., N.M.P.-H., F.R., A.S., K.S., J.S., G.S., E.T., P.J.T., C.T., M.T., J.Y.T., C.A.W., S. Weidinger, S. Weiss, G.W., L.K.W., C.O., M.A.F., H. Bisgaard, D.P.S., The 23andMe Research Team and AAGC collaborators. Immunological interpretation: N.S. and S. Brix. Gene-expression analysis: M.G. and J.D. Protein modeling: K.K.J.

Competing interests

G.S., I.J., and K.S. are affiliated with deCODE genetics/Amgen and declare competing financial interests as employees. C.T., D.A.H., J.Y.T., and the 23andMe Research Team are employees of and hold stock and/or stock options in 23andMe, Inc. L.P. has received a fee for participating in a scientific-input engagement meeting from Merck Sharp & Dohme Limited, outside of this work.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41588-018-0157-1.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to K.B.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Johannes Waage[1,74], Marie Standl[2,74], John A. Curtin[3], Leon E. Jessen[1], Jonathan Thorsen[1], Chao Tian[4], Nathan Schoettler[5], The 23andMe Research Team[6], AAGC collaborators[6], Carlos Flores[7,8], Abdel Abdellaoui[9,10], Tarunveer S. Ahluwalia[1], Alexessander C. Alves[11], Andre F. S. Amaral[12], Josep M. Antó[13,14,15,16], Andreas Arnold[17], Amalia Barreto-Luis[7], Hansjörg Baurecht[18], Catharina E. M. van Beijsterveldt[9], Eugene R. Bleecker[19], Sílvia Bonàs-Guarch[20], Dorret I. Boomsma[9,21], Susanne Brix[22], Supinda Bunyavanich[23], Esteban G. Burchard[24,25], Zhanghua Chen[26], Ivan Curjuric[27,28], Adnan Custovic[29], Herman T. den Dekker[30,31,32], Shyamali C. Dharmage[33], Julia Dmitrieva[34], Liesbeth Duijts[30,32,35], Markus J. Ege[36], W. James Gauderman[26], Michel Georges[34], Christian Gieger[2,37], Frank Gilliland[26], Raquel Granell[38], Hongsheng Gui[39], Torben Hansen[40], Joachim Heinrich[2,41], John Henderson[38], Natalia Hernandez-Pacheco[7,42], Patrick Holt[43], Medea Imboden[27,28], Vincent W. V. Jaddoe[30,31,44], Marjo-Riitta Jarvelin[11,45,46,47], Deborah L. Jarvis[12], Kamilla K. Jensen[48], Ingileif Jónsdóttir[49,50], Michael Kabesch[51], Jaakko Kaprio[52,53,54], Ashish Kumar[27,28,55], Young-Ae Lee[56,57], Albert M. Levin[58], Xingnan Li[59], Fabian Lorenzo-Diaz[42], Erik Melén[55,60], Josep M. Mercader[20,61,62], Deborah A. Meyers[19], Rachel Myers[5], Dan L. Nicolae[5], Ellen A. Nohr[63], Teemu Palviainen[53], Lavinia Paternoster[38], Craig E. Pennell[64], Göran Pershagen[55,65], Maria Pino-Yanes[7,8,42], Nicole M. Probst-Hensch[27,28], Franz Rüschendorf[56], Angela Simpson[3], Kari Stefansson[49,50], Jordi Sunyer[13], Gardar Sveinbjornsson[49], Elisabeth Thiering[2,66], Philip J. Thompson[67], Maties Torrent[68], David Torrents[20,69], Joyce Y. Tung[4], Carol A. Wang[64], Stephan Weidinger[18], Scott Weiss[70], Gonneke Willemsen[9], L. Keoki Williams[39,71], Carole Ober[5], David A. Hinds[4], Manuel A. Ferreira[72], Hans Bisgaard[1], David P. Strachan[73] and Klaus Bønnelykke[1*]

[1]COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark. [2]Institute of Epidemiology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [3]Division of Infection, Immunity and Respiratory Medicine, University of Manchester, Manchester Academic Health Science Centre, Manchester University NHS Foundation Trust, Manchester, UK. [4]23andMe, Inc., Mountain View, CA, USA. [5]Department of Human Genetics, University of Chicago, Chicago, IL, USA. [6]A list of members and affiliations appears in the Supplementary Note. [7]Research Unit, Hospital Universitario N.S. de Candelaria, Universidad de La Laguna, Tenerife, Spain. [8]CIBER de Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, Madrid, Spain. [9]Department of Biological Psychology, Netherlands Twin Register, VU University, Amsterdam, The Netherlands. [10]Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands. [11]Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment & Health, School of Public Health, Imperial College London, London, UK. [12]Population Health and Occupational Disease, National Heart and Lung Institute, Imperial College London, London, UK. [13]ISGlobal, Barcelona, Spain. [14]IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain. [15]Universitat Pompeu Fabra (UPF), Barcelona, Spain. [16]CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain. [17]Clinic and Polyclinic of Dermatology, University Medicine Greifswald, Greifswald, Germany. [18]Department of Dermatology, Venereology and Allergology, University-Hospital Schleswig-Hostein, Campus Kiel, Kiel, Germany. [19]Divisions of Pharmacogenomics and Genetics, Genomics and Precision Medicine, Department of Medicine, University of Arizona College of Medicine, Tucson, AZ, USA. [20]Barcelona Supercomputing Center (BSC), Joint BSC-CRG-IRB Research Program in Computational Biology, Barcelona, Spain. [21]APH Amsterdam Public Health, Amsterdam, The Netherlands. [22]Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark. [23]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [24]Department of Medicine, University of California San Francisco, San Francisco, CA, USA. [25]Department of Bioengineering & Therapeutic Sciences, University of California San Francisco, San Francisco, CA, USA. [26]Department of Preventive Medicine, University of Southern California, Keck School of Medicine, Los Angeles, CA, USA. [27]University of Basel, Basel, Switzerland. [28]Swiss Tropical and Public Health Institute, Basel, Switzerland. [29]Department of Paediatrics, Imperial College London, London, UK. [30]The Generation R Study Group, Erasmus Medical Center, Rotterdam, The Netherlands. [31]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands. [32]Department of Pediatrics, Division of Respiratory Medicine, Erasmus Medical Center, Rotterdam, The Netherlands. [33]Allergy and Lung Health Unit, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia. [34]Laboratory of Animal Genomics, Unit of Medical Genomics, GIGA Institute, University of Liège, Liège, Belgium. [35]Department of Pediatrics, Division of Neonatology, Erasmus Medical Center, Rotterdam, The Netherlands. [36]LMU Munich, Dr von Hauner Children's Hospital, Munich, and German Center for Lung Research (DZL), Munich, Germany. [37]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [38]MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. [39]Center for Health Policy and Health Services Research, Henry Ford Health System, Detroit, MI, USA. [40]Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Department of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [41]Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, University of Munich Medical Center, Ludwig-Maximilians-Universität München, Munich, Germany. [42]Genomics and Health Group, Department of Biochemistry, Microbiology, Cell Biology and Genetics, Universidad de La Laguna, La Laguna, Tenerife, Spain. [43]Telethon Kids Institute (TKI), Perth, Western Australia, Australia. [44]Department of Pediatrics, Erasmus Medical Center, Rotterdam, The Netherlands. [45]Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland. [46]Biocenter Oulu, University of Oulu, Oulu, Finland. [47]Unit of Primary Care, Oulu University Hospital, Oulu, Finland. [48]Department of Bio and Health

Informatics, Technical University of Denmark, Kongens Lyngby, Denmark. [49]deCODE genetics/Amgen Inc, Reykjavik, Iceland. [50]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [51]Department of Pediatric Pneumology and Allergy, University Children's Hospital Regensburg (KUNO), Regensburg, Germany. [52]Department of Public Health, University of Helsinki, Helsinki, Finland. [53]Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland. [54]National Institute for Health and Welfare, Helsinki, Finland. [55]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [56]Max-Delbrück-Center (MDC) for Molecular Medicine, Berlin, Germany. [57]Clinic for Pediatric Allergy, Experimental and Clinical Research Center, Charité Universitätsmedizin Berlin, Berlin, Germany. [58]Department of Public Health Sciences, Henry Ford Health System, Detroit, MI, USA. [59]Divisions of Genetics, Genomics and Precision Medicine, Department of Medicine, University of Arizona College of Medicine, Tucson, AZ, USA. [60]Sachs' Children's Hospital, Stockholm, Sweden. [61]Programs in Metabolism and Medical & Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [62]Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA. [63]Institute of Clinical Research, University of Southern Denmark, Department of Obstetrics & Gynecology, Odense University Hospital, Odense, Denmark. [64]School of Medicine and Public Health, Faculty of Medicine and Health, University of Newcastle, Callaghan, New South Wales, Australia. [65]Centre for Occupational and Environmental Medicine, Stockholm County Council, Stockholm, Sweden. [66]Ludwig-Maximilians-University of Munich, Dr. von Hauner Children's Hospital, Division of Metabolic Diseases and Nutritional Medicine, Munich, Germany. [67]Institute for Respiratory Health, Harry Perkins Institute of Medical Research, University of Western Australia, Nedlands, Western Australia, Australia. [68]Ib-Salut, Area de Salut de Menorca, Institut d'Investigacio Sanitaria Illes Balears (IdISBa), Palma de Mallorca, Spain. [69]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain. [70]Channing Division of Network Medicine, Brigham & Women's Hospital and Harvard Medical School, Boston, MA, USA. [71]Department of Internal Medicine, Henry Ford Health System, Detroit, MI, USA. [72]QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [73]Population Health Research Institute, St George's, University of London, London, UK. [74]These authors contributed equally: Johannes Waage, Marie Standl. *e-mail: kb@copsac.com

Methods

Phenotype definition. *Allergic rhinitis.* Cases were defined as individuals who had ever received a diagnosis or experienced symptoms of allergic rhinitis, on the basis of available phenotype definitions in the included studies (Supplementary Table 3 and cohort-recruitment details in Supplementary Note). All relevant ethical regulations were followed, as specified in relation to the individual studies in the Supplementary Note. To maximize numbers and optimize statistical power, we did not require doctor-diagnosed allergic rhinitis or verification by allergic sensitization. This approach was confirmed by a sensitivity analysis in 23andMe on the basis of association with known risk loci for allergic rhinitis (data not shown). Controls were defined as individuals who had never received a diagnosis or experienced symptoms of allergic rhinitis.

Allergic sensitization. We considered specific IgE production against inhalant allergens without restricting by assessment method or type of inhalant allergen. Cases were defined as individuals with objectively measured sensitization against at least one of the inhalant allergens tested for in the respective studies, and controls were defined as individuals who were not sensitized against any of the allergens tested for. We included sensitization assessed by skin reaction after puncture of the skin with a droplet of allergen extract (SPT) and/or by detection of the levels of circulating allergen-specific IgE in the blood. The SPT wheal-diameter cutoffs were ≥ 3 mm for cases and < 1 mm for controls. To optimize case specificity and the correlation between methods, we chose a high cutoff of specific IgE levels for cases (0.7 IU/ml) and a low cutoff for controls (0.35 IU/ml).

Nonallergic rhinitis. Cases were defined as individuals with current allergic rhinitis symptoms (within the previous 12 months) and no allergic sensitization (negative specific IgE (<0.35 IU/ml) and/or negative skin-prick test (<1 mm) for all allergens and time points tested).

Controls were defined as individuals who had never had symptoms of allergic rhinitis and showed no allergic sensitization (negative specific IgE (<0.35 IU/ml) and/or negative skin-prick test (<1 mm) for all allergens and time points tested)

For all three phenotypes, we combined data from children and adults but chose a lower age limit of 6 years, because allergic rhinitis and sensitization status at younger ages show poorer correlation with status later in life, owing to both transient symptoms/sensitization status and frequent development of symptoms/sensitization during late childhood.

GWAS quality control and cohort summary-data harmonization. For allergic rhinitis, allergic sensitization, and nonallergic rhinitis, data were imputed separately for each cohort with the 1000 Genomes Project (1KGP) phase 1, version 3 release, and the genome-wide association analysis was adjusted for sex and, if necessary, for age and principal components (Supplementary Table 3). All studies included individuals of European descent, except Generation R and RAINE, which comprised a mixed, multiancestral population. We used EasyQC v. 9.2 (ref. [36]) for quality control and marker harmonization for cohort-level meta-GWAS summary files. Cohort data were harmonized to genome build GRCh37 and were checked against the 1KGP phase 3 reference-allele frequencies for processing problems. GWAS summary 'karyograms' were visually inspected to catch cohorts with incomplete data. Distributions of estimate coefficients and errors, as well as plots of standard error versus sample size and of *P* value versus *Z* score, were inspected for each cohort for systematic errors in statistical models. Ambiguous markers that were nonunique in terms of both genomic position and allele coding were removed. A minimum imputation score of 0.3 (r^2) or 0.4 (proper_info) was required for markers. A minimum minor-allele count of 7 was required for each marker in each cohort, as suggested by the GIANT consortium and EasyQC.

Meta-analysis. For allergic rhinitis, allergic sensitization, and nonallergic rhinitis, meta-analysis for the discovery phase was conducted with GWAMA[37] with an inverse-variance-weighted fixed-effect model with genomic-control correction of the individual studies. Each locus is represented by the variant showing the strongest evidence within a 1-Mb buffer. Loci were inspected visually by plotting genomic neighborhood and coloring for 1KGP r^2 values. From the pool of genome-wide-significant markers in the discovery, one locus with index marker rs193243426 without a credible LD structure was removed from further analysis (Supplementary Fig. 10). Heterogeneity was assessed with Cochran's *Q* test.

Meta-analysis of replication candidates from the allergic rhinitis discovery phase was carried out in R version 3.4.0 and the *meta* package version 4.8-2 with an inverse-variance-weighted fixed-effect model. For a subset of markers, cohorts reported suitable proxies ($r^2 > 0.85$), and followed-up markers were not present or had insufficient imputation or genotyping quality (Supplementary Table 25).

Gene-set-overrepresentation analysis, discovery phase. To facilitate the selection of biologically relevant discovery candidates in the sub-genome-wide significant stratum ($5 \times 10^{-8} < P < 1 \times 10^{-6}$), we used a custom algorithm for gene-set-overrepresentation analysis implemented in R, with a scoring and permutation regime modeled after MAGENTA[38]. Genes with lengths less than 200 bp, with copies on multiple chromosomes, and with multiple copies on the same chromosome more than 1 Mb apart were removed from analysis. Gene models (GENCODE v 19) were downloaded from the UCSC Table Browser[39] and expanded 110 kb upstream and 40 kb downstream, similarly to the MAGENTA procedure. The HLA region was excluded from analysis (chromosome 6: 29691116–33054976). Similarly to the MAGENTA procedure, gene scores were adjusted for the number of markers per gene, gene width, recombination hotspots, genetic distance, and number of independent markers per gene, all with updated data from the UCSC Table Browser. For the gene-set-overrepresentation permutation calculation, gene sets from the MSigDB collections c2, c3, c5, c7, and hallmark were included[40]. A MAGENTA-style enrichment cutoff at 95% was used. Gene sets with FDR <0.05 were considered.

Conditional analyses. To identify additional independent markers at each discovery genomic region, we used genome-wide complex trait analysis (GCTA) v. 1.26.0 (ref. [41]). Within a window of ±1 Mb of each discovery-phase index marker, all markers were conditioned on the index by using the --cojo-cond feature of GCTA with default parameters. Plink v. 1.90b3.42 (ref. [42]) was used to calculate r^2 for GCTA with the UK10K full genotype panel[43] as a reference. A total of 42 of 52 markers from the full discovery phase were present in UK10K. Because a MAF-dependent inflation of conditional *P* values was observed (data not shown), only conditional markers with MAF ≥10% were selected.

Locus definition and credible sets for VEP annotation. Discovery loci were defined as index markers extended with markers in LD ($r^2 \geq 0.5$), on the basis of 1KGP phase 3. Protein-coding gene transcript models (GENCODE v. 24) were downloaded from the UCSC Table Browser, and the nearest upstream and downstream genes, as well as all genes within the extended loci, were annotated. Credible sets for each locus were calculated with the method of Morris[44].

LD was calculated for each discovery index variant within ±500 kb, and markers with $r^2 < 0.1$ were excluded. For the remaining markers, the Bayesian factor (ABF) values and the posterior probabilities (PostProb) were calculated, and cumulative posterior probability values were generated on the basis of ranking markers on ABF. Finally, variants were included in the 99%-credible set until the cumulative posterior probability was ≥0.99.

Credible sets for each locus were annotated with information on mutation effects in coding regions by using the Variant Effect Predictor (VeP) REST API[45], exporting only the nonsynonymous substitutions.

GWAS-catalog lookup. For annotation of markers with identification in previous GWA studies, the GWAS catalog was downloaded from NHGRI-EBI (v.1.0.1, 28 November 2016). For this analysis, allergic rhinitis–associated loci were lifted from genomic build GRCh37 to GRCh38 and extended with ±1 Mb in each direction before being overlapped with GWAS-catalog annotations. Relevant GWAS-catalog overlap traits were binned into the trait groups 'allergic rhinitis', 'asthma', 'autoimmune', 'eczema', 'infectious diseases', 'lung-related traits', and 'other allergy'. One million random genomic intervals of the same length (2 Mb) were obtained to generate a background overlap distribution, and *P* values were calculated from this background.

Classical HLA-allele analysis. Analyses of imputed classical HLA alleles were performed in the 23andMe study (AR discovery population) comprising 49,180 individuals with allergic rhinitis and 124,102 controls.

HLA imputation was performed with HIBAG v. 1.2.3 (ref. [46]). We imputed allelic dosage for *HLA-A*, *HLA-B*, *HLA-C*, *HLA-DPB1*, *HLA-DQA1*, *HLA-DQB1*, and *HLA-DRB1* loci at four-digit resolution with the default settings of HIBAG, for a total of 292 classical HLA alleles.

Using an approach suggested by de Bakker[47], we downloaded the files mapping HLA alleles to amino acid sequences from https://www.broadinstitute.org/mpg/snp2hla/ and mapped our imputed HLA alleles at four-digit resolution to the corresponding amino acid sequences; in this way, we translated the imputed HLA allelic dosages directly to amino acid dosages. We encoded all amino acid variants in the 23andMe European samples as 2,395 biallelic amino acid polymorphisms, as previously described[48].

Similarly to the SNP imputation, we measured imputation quality by using r^2, which is the ratio of the empirically observed variance of the allele dosage to the expected variance, assuming Hardy–Weinberg equilibrium.

To test associations among imputed HLA alleles, amino acid variants, and phenotypes, we performed logistic regression using the same set of covariates used in the SNP-based GWAS. We applied a forward stepwise strategy within each type of variant to establish statistically independent signals in the HLA region. Within each variant type, we first identified the most strongly associated signals (lowest *P* value) and performed forward iterative conditional regression to identify other independent signals. All analyses were controlled for sex and five principal components of genetic ancestry. The *P* values were calculated with a likelihood-ratio test.

Structural visualization of amino acid variants. Structural visualization of amino acid variants was performed for the strongest associated variants in HLA-DQB1 (position 30) and HLA-B (position 116), respectively (Supplementary Table 10) and was based on X-ray structures from the Protein Data Bank (PDB)[49]. To find

the best structure, we used the specialized search function in the Immune Epitope Database[50], selecting only X-ray-crystalized structures for the specific MHC classes HLA-DQB1 (class II) and HLA-B (class I). Using this criterion, we found 17 crystallized structures for HLA-DQB1 and 164 structures for HLA-B. From these lists, we selected the structure with the lowest resolution and the amino acids encoded by the reported top SNPs. The accession codes for the selected structures were PDB 4MAY[51] for HLA-DQB1 and PDB 2A83 (ref. [52]) for HLA-B. Both structures were visualized in PyMOL v. 1.8.2.1 (http://www.pymol.org/). Furthermore, we used PyMOL to measure intramolecular distances from the side chains of the amino acids associated with allergic rhinitis to the Cα atoms in the peptide. This distance measure was chosen to accommodate the possibility of different amino acids in the peptide. For two amino acids to interact, the distance should be approximately 4 Å or less. We measured distances of 6 Å (HLA-DQB1) and 7 Å (HLA-B). However, these distances do not include the peptide side chains, which range from 1.5 to 8.8 Å. Therefore, we estimate that physical interaction between the amino acids associated with allergic rhinitis and the peptide is likely.

Genetic heritability and genetic correlation. For calculating genetic heritability and genetic correlation between allergic rhinitis and allergic sensitization, as well as between clinical cohorts and 23andMe within allergic rhinitis, we used the LD-score regression-based method, as implemented in LDSC v. 1.0 (refs [45,53]). The population prevalence was set to 10% for allergic rhinitis and allergic sensitization. Genetic correlation analysis among allergic rhinitis, allergic sensitization, and published GWAS studies was carried out in LDHUB platform v. 1.3.1 (ref. [54]) against all traits but excluding metabolites[55].

eQTL sources and analysis. From GTEx V6p[56], all significant variant-gene cis-eQTL pairs for whole blood, lung, and EBV-transformed lymphocytes were downloaded from https://gtexportal.org/ and used in analysis. From Westra et al.[57], we downloaded both cis and trans eQTLs in whole blood, then used variant-gene pairs with FDR <0.1 in subsequent analyses. From Fairfax et al.[58], we downloaded cis-eQTLs from monocytes and B cells, then used variant-gene pairs with FDR < 0.1 in subsequent analyses. From Bonder et al.[58], we downloaded meQTLs from whole blood and used variant-probe pairs with FDR <0.05 in subsequent analyses. From Nicodemus-Johnson et al.[59], we downloaded cis-eQTLs and meQTLs from the lung and used variant-gene pairs with FDR <0.1 in subsequent analyses. From Momozawa et al.[60], we downloaded cis-eQTLs from blood cell types CD14, CD15, CD19, CD4, and CD8 used and variant-gene pairs with a weighted correlation ≥0.6 in analyses. For Supplementary Table 14 priority genes, protein-coding information was downloaded from the UCSC Table Browser, with the 'transcriptClass' field from the 'wgEncodeGencodeAttrsV24lift37' table.

Promoter-capture Hi-C gene prioritization. To assess spatial promoter interactions in the discovery set, we performed a capture Hi-C gene prioritization (CHIGP) as described in Javierre et al.[61] and https://github.com/ollyburren/CHIGP/ by using the following recommended settings and data sources: 0.1 cM recombination blocks, 1KGP EUR reference population, coding markers from the GRCh37 Ensembl assembly, and the CHICAGO-generated[62] promoter-capture Hi-C peak matrix data from 17 human primary blood cell types supplied in the original paper. The resulting protein-coding prioritized genes (gene score >0.5) were used in the downstream network analysis, from cell types 'fetal thymus', 'total CD4 T cells', 'activated total CD4 T cells', 'nonactivated total CD4 T cells', 'naive CD4 T cells', 'total CD8 T cells', 'naive CD8 T cells', 'total B cells', 'naive B cells', 'endothelial precursors', 'macrophages M0', 'macrophages M1', 'macrophages M2', 'monocytes', and 'neutrophils'.

Gene-set-overrepresentation analysis of known and replicating novel loci. All high-confidence gene symbols from eQTL and meQTL sources, PCHiC, as well as genes (models extended 110 kb upstream, and 40 kb downstream) within each r^2-based locus definition from known and replicating novel loci were input into the pathway-based set over-representation analysis module in ConsensusPathDB (CPDB) database and tools[63] with 229 of 277 gene identifiers translated. In addition, these same symbols were used for ingenuity pathway analysis (IPA; http://www.ingenuity.com/; a curated database of the relationships among genes obtained from published articles, and genetic and expression data repositories), to identify biological pathways common to genes. IPA determines whether the associated genes are significantly enriched in a specific biological function or network by assessing direct interactions. We assigned significance if the right-tailed Fisher's exact test $P < 0.05$. eQTL/meQTL, PCHiC and locus–gene intersections were visualized with the UpSetR package (v1.3.2)[64].

Tissue overrepresentation. To assay the enrichment of variants associated with allergic rhinitis in tissue-specific gene-expression sets, we used the DEPICT enrichment method[65], using a P-value threshold of 1×10^{-5} and standard settings.

Enrichment of regulatory regions. To assay the enrichment of variants associated with allergic rhinitis in regions of open chromatin and specific histone marks, we used the GWAS analysis of regulatory or functional information enrichment with LD correction (GARFIELD v. 1) method[66]. In essence, GARFIELD performs greedy pruning of GWAS markers (LD $r^2 > 0.1$) and then annotates them according to functional information overlap. Next, it quantifies fold enrichment at various GWAS significance cutoffs and assesses them by permutation testing while adjusting for the minor allele frequency, distance to nearest transcription start site and number of LD proxies ($r^2 > 0.8$). GARFIELD was run with 10,000,000 permutations and otherwise default settings.

Population-attributable-risk fraction. Population-attributable-risk fractions (PARFs) were estimated from B58C, a general-population sample with participant ages 44–45 years also contributing to the discovery stage. The genetic-risk score was calculated by applying the pooled per-allele coefficients (ln(OR) values) from the allergic rhinitis discovery set to the number of higher-risk alleles of each of the 41 established loci (known genome-wide-significant and novel replicated loci), with one SNP per locus. Because there were no individuals observed with zero higher-risk alleles, the prevalence of sensitization for individuals in the lowest decile of the genetic-risk score distribution was used to derive PARF estimates on the assumption that this 10% of the population was unexposed. This method has the advantage that it does not predict beyond the bounds of the data, but its results are conservative. The PARF was then derived (with 95% confidence intervals) by expressing the difference between the observed prevalence and the predicted (unexposed) prevalence as a percentage of the observed prevalence. PARFs were estimated by using the 41 allergic rhinitis–associated loci in relation to allergic rhinitis, allergic sensitization, and nonallergic rhinitis.

Protein-network and drug interactions. To analyze protein–protein–drug interaction networks, STRING (V10)[67] was used. Protein-network data (9606.protein.links.v10.txt.gz) and protein alias data (9606.protein.aliases.v10.txt) files were downloaded from the string db website (http://string-db.org/). GWAS hits stratified on 'all', 'blood', and 'lung' were converted to Ensembl protein IDs by using the protein alias data. The interactors were subsequently identified by using the link data at a 'high-confidence cutoff of >0.7', as described in the STRING FAQ. The interactor Ensembl protein IDs were then converted to UniProt gene names, and both hits and interactors were then analyzed for interactions with FDA approved drugs by using the ChEMBL Database v. 22 (ref. [68]) API via Python (v. 2.7.12). Finally, stratified networks consisting of GWAS hits connected to interactors and drugs connected to both GWAS hits and interactors were visualized with GGraph (v. 1.0.0), iGraph (v. 1.0.1), and TidyVerse (v. 1.1.1) in R (v. 3.3.2).

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability. Genome-wide results, excluding 23andMe, are available from the corresponding author on request and will be available on LDhub after July 2018. The full GWAS summary statistics for the 23andMe discovery dataset will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please contact D.A.H. (dhinds@23andme.com) for more information and to apply for access to the 23andMe data.

References

36. Winkler, T. W. et al. Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* **9**, 1192–1212 (2014).
37. Mägi, R. & Morris, A. P. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinforma.* **11**, 288 (2010).
38. Segrè, A. V., Groop, L., Mootha, V. K., Daly, M. J. & Altshuler, D. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet.* **6**, e1001058 (2010).
39. Karolchik, D. et al. The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* **32**, D493–D496 (2004).
40. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
41. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
42. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
43. UK10K Consortium. et al. The UK10K project identifies rare variants in health and disease. *Nature* **526**, 82–90 (2015).
44. Wellcome Trust Case Control Consortium. et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat. Genet.* **44**, 1294–1301 (2012).
45. McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
46. Zheng, X. et al. HIBAG—HLA genotype imputation with attribute bagging. *Pharm. J.* **14**, 192–200 (2014).
47. Jia, X. et al. Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One* **8**, e64683 (2013).

48. Tian, C. et al. Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. *Nat. Commun.* **8**, 599 (2017).
49. Berman, H. M. et al. The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
50. Vita, R. et al. The immune epitope database (IEDB) 3.0. *Nucleic Acids Res.* **43**, D405–D412 (2015).
51. Sethi, D. K., Gordo, S., Schubert, D. A. & Wucherpfennig, K. W. Crossreactivity of a human autoimmune TCR is dominated by a single TCR loop. *Nat. Commun.* **4**, 2623 (2013).
52. Rückert, C. et al. Conformational dimorphism of self-peptides and molecular mimicry in a disease-associated HLA-B27 subtype. *J. Biol. Chem.* **281**, 2306–2316 (2006).
53. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
54. Zheng, J. et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2017).
55. Kettunen, J. et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat. Genet.* **44**, 269–276 (2012).
56. Lucas, A. O. Surveillance of communicable diseases in tropical Africa. *Int. J. Epidemiol.* **5**, 39–43 (1976).
57. Westra, H.-J. et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* **45**, 1238–1243 (2013).
58. Fairfax, B. P. et al. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat. Genet.* **44**, 502–510 (2012).
59. Nicodemus-Johnson, J. et al. DNA methylation in lung cells is associated with asthma endotypes and genetic risk. *JCI Insight* **1**, e90151 (2016).
60. Momozawa, Y. et al. IBD risk loci are enriched in multigenic regulatory modules encompassing putative causative genes. *Nat. Commun.* **9**, 2427 (2018).
61. Javierre, B. M. et al. Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell* **167**, 1369–1384. e19 (2016).
62. Cairns, J. et al. CHiCAGO: robust detection of DNA looping interactions in Capture Hi-C data. *Genome Biol.* 17, (127 (2016).
63. Kamburov, A. et al. ConsensusPathDB: toward a more complete picture of cell biology. *Nucleic Acids Res.* **39**, D712–D717 (2011).
64. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938–2940 (2017).
65. Pers, T. H. et al. Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6**, 5890 (2015).
66. Iotchkova, V. et al. GARFIELD: GWAS analysis of regulatory or functional information enrichment with LD correction. https://www.biorxiv.org/content/early/2016/11/07/085738 (2016).
67. Szklarczyk, D. et al. STRINGv10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447–D452 (2015).
68. Bento, A. P. et al. The ChEMBL bioactivity database: an update. *Nucleic Acids Res.* **42**, D1083–D1090 (2014).

nature research

Corresponding Author: Klaus Bønnelykke

Date: Apr 28, 2018

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work we publish. This form is published with all life science papers and is intended to promote consistency and transparency in reporting. All life sciences submissions use this form; while some list items might not apply to an individual manuscript, all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   We aimed to include all available data-sets on allergic rhinitis worldwide. In this study we achieved a sample size that was several times larger than previous GWAS- studies, which was sufficent for providing genome-wide significant findings.

2. Data exclusions

   Describe any data exclusions.

   Other than the preliminary analyses/experimental optimization, no data were excluded.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   No experimental studies were performed.

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   No experimental studies were performed. Cases and controls for genome-wide association were analyzed based upon predefined criteria.

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   No blinding was performed and not considered relevant in this meta-Genome-wide association study.

   Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

   For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly.
☒	☐	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. *p* values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A summary of the descriptive statistics, including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	This is described in details in the methods section. In short, for AR, AS, and NAR, meta-analysis for the discovery phase was conducted using GWAMA, while meta-analysis of replication candidates from the AR discovery phase was carried out using R version 3.4.0, and the meta package version 4.8-2 with an inverse variance weighted fixed-effect model. Additional software include EasyQC (v9.2), STRING (v10), ChEMBL Database (v22), GGraph (v1.0.0), iGraph (v1.0.1), TidyVerse (v1.1.1), LDHUB platform (v1.3.1), GENCODE (v19 and v24, Genome-wide Complex Trait Analysis (GCTA) (v. 1.26.0), Plink (v1.90b3.42), PyMOL (v1.8.2.1), LDSC (v.1.0), LDHUB platform (v1.3.1), GTEx (V6p), UpSetR package (v1.3.2), GWAS Analysis of Regulatory or Functional Information Enrichment with LD correction (GARFIELD) method (v1), Python (v2.7.12), HIBAG (v1.2.3).

For all studies, we encourage code deposition in a community repository (e.g. GitHub). Authors must make computer code available to editors and reviewers upon request. The *Nature Methods* guidance for providing algorithms and software for publication may be useful for any submission.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	No unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	No antibodies were used.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	No eukaryotic cell lines were used.
b. Describe the method of cell line authentication used.	No eukaryotic cell lines were used.
c. Report whether the cell lines were tested for mycoplasma contamination.	No eukaryotic cell lines were used.
d. If any of the cell lines used in the paper are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No eukaryotic cell lines were used.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	No animals were used.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	Relevant information on research participants in the individual studies is provided in the supplementary material.

nature genetics

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41588-018-0157-1

In the format provided by the authors and unedited.

Genome-wide association and HLA fine-mapping studies identify risk loci and genetic pathways underlying allergic rhinitis

Johannes Waage[1,74], Marie Standl[2,74], John A. Curtin[3], Leon E. Jessen[1], Jonathan Thorsen[1], Chao Tian[4], Nathan Schoettler[5], The 23andMe Research Team[6], AAGC collaborators[6], Carlos Flores[7,8], Abdel Abdellaoui[9,10], Tarunveer S. Ahluwalia[1], Alexessander C. Alves[11], Andre F. S. Amaral[12], Josep M. Antó[13,14,15,16], Andreas Arnold[17], Amalia Barreto-Luis[7], Hansjörg Baurecht[18], Catharina E. M. van Beijsterveldt[9], Eugene R. Bleecker[19], Sílvia Bonàs-Guarch[20], Dorret I. Boomsma[9,21], Susanne Brix[22], Supinda Bunyavanich[23], Esteban G. Burchard[24,25], Zhanghua Chen[26], Ivan Curjuric[27,28], Adnan Custovic[29], Herman T. den Dekker[30,31,32], Shyamali C. Dharmage[33], Julia Dmitrieva[34], Liesbeth Duijts[30,32,35], Markus J. Ege[36], W. James Gauderman[26], Michel Georges[34], Christian Gieger[2,37], Frank Gilliland[26], Raquel Granell[38], Hongsheng Gui[39], Torben Hansen[40], Joachim Heinrich[2,41], John Henderson[38], Natalia Hernandez-Pacheco[7,42], Patrick Holt[43], Medea Imboden[27,28], Vincent W. V. Jaddoe[30,31,44], Marjo-Riitta Jarvelin[11,45,46,47], Deborah L. Jarvis[12], Kamilla K. Jensen[48], Ingileif Jónsdóttir[49,50], Michael Kabesch[51], Jaakko Kaprio[52,53,54], Ashish Kumar[27,28,55], Young-Ae Lee[56,57], Albert M. Levin[58], Xingnan Li[59], Fabian Lorenzo-Diaz[42], Erik Melén[55,60], Josep M. Mercader[20,61,62], Deborah A. Meyers[19], Rachel Myers[5], Dan L. Nicolae[5], Ellen A. Nohr[63], Teemu Palviainen[53], Lavinia Paternoster[38], Craig E. Pennell[64], Göran Pershagen[55,65], Maria Pino-Yanes[7,8,42], Nicole M. Probst-Hensch[27,28], Franz Rüschendorf[56], Angela Simpson[3], Kari Stefansson[49,50], Jordi Sunyer[13], Gardar Sveinbjornsson[49], Elisabeth Thiering[2,66], Philip J. Thompson[67], Maties Torrent[68], David Torrents[20,69], Joyce Y. Tung[4], Carol A. Wang[64], Stephan Weidinger[18], Scott Weiss[70], Gonneke Willemsen[9], L. Keoki Williams[39,71], Carole Ober[5], David A. Hinds[4], Manuel A. Ferreira[72], Hans Bisgaard[1], David P. Strachan[73] and Klaus Bønnelykke[1*]

[1]COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark. [2]Institute of Epidemiology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [3]Division of Infection, Immunity and Respiratory Medicine, University of Manchester, Manchester Academic Health Science Centre, Manchester University NHS Foundation Trust, Manchester, UK. [4]23andMe, Inc., Mountain View, CA, USA. [5]Department of Human Genetics, University of Chicago, Chicago, IL, USA. [6]A list of members and affiliations appears in the Supplementary Note. [7]Research Unit, Hospital Universitario N.S. de Candelaria, Universidad de La Laguna, Tenerife, Spain. [8]CIBER de Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, Madrid, Spain. [9]Department of Biological Psychology, Netherlands Twin Register, VU University, Amsterdam, The Netherlands. [10]Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands. [11]Department of Epidemiology and Biostatistics, MRC–PHE Centre for Environment & Health, School of Public Health, Imperial College London, London, UK. [12]Population Health and Occupational Disease, National Heart and Lung Institute, Imperial College London, London, UK. [13]ISGlobal, Barcelona, Spain. [14]IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain. [15]Universitat Pompeu Fabra (UPF), Barcelona, Spain. [16]CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain. [17]Clinic and Polyclinic of Dermatology, University Medicine Greifswald, Greifswald, Germany. [18]Department of Dermatology, Venereology and Allergology, University-Hospital Schleswig-Hostein, Campus Kiel, Kiel, Germany. [19]Divisions of Pharmacogenomics and Genetics, Genomics and Precision Medicine, Department of Medicine, University of Arizona College of Medicine, Tucson, AZ, USA. [20]Barcelona Supercomputing Center (BSC), Joint BSC–CRG–IRB Research Program in Computational Biology, Barcelona, Spain. [21]APH Amsterdam Public Health, Amsterdam, The Netherlands. [22]Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark. [23]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [24]Department of Medicine, University of California San Francisco, San Francisco, CA, USA. [25]Department of Bioengineering & Therapeutic Sciences, University of California

San Francisco, San Francisco, CA, USA. [26]Department of Preventive Medicine, University of Southern California, Keck School of Medicine, Los Angeles, CA, USA. [27]University of Basel, Basel, Switzerland. [28]Swiss Tropical and Public Health Institute, Basel, Switzerland. [29]Department of Paediatrics, Imperial College London, London, UK. [30]The Generation R Study Group, Erasmus Medical Center, Rotterdam, The Netherlands. [31]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands. [32]Department of Pediatrics, Division of Respiratory Medicine, Erasmus Medical Center, Rotterdam, The Netherlands. [33]Allergy and Lung Health Unit, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia. [34]Laboratory of Animal Genomics, Unit of Medical Genomics, GIGA Institute, University of Liège, Liège, Belgium. [35]Department of Pediatrics, Division of Neonatology, Erasmus Medical Center, Rotterdam, The Netherlands. [36]LMU Munich, Dr von Hauner Children's Hospital, Munich, and German Center for Lung Research (DZL), Munich, Germany. [37]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [38]MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. [39]Center for Health Policy and Health Services Research, Henry Ford Health System, Detroit, MI, USA. [40]Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Department of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [41]Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, University of Munich Medical Center, Ludwig-Maximilians-Universität München, Munich, Germany. [42]Genomics and Health Group, Department of Biochemistry, Microbiology, Cell Biology and Genetics, Universidad de La Laguna, La Laguna, Tenerife, Spain. [43]Telethon Kids Institute (TKI), Perth, Western Australia, Australia. [44]Department of Pediatrics, Erasmus Medical Center, Rotterdam, The Netherlands. [45]Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland. [46]Biocenter Oulu, University of Oulu, Oulu, Finland. [47]Unit of Primary Care, Oulu University Hospital, Oulu, Finland. [48]Department of Bio and Health Informatics, Technical University of Denmark, Kongens Lyngby, Denmark. [49]deCODE genetics/Amgen Inc, Reykjavik, Iceland. [50]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [51]Department of Pediatric Pneumology and Allergy, University Children's Hospital Regensburg (KUNO), Regensburg, Germany. [52]Department of Public Health, University of Helsinki, Helsinki, Finland. [53]Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland. [54]National Institute for Health and Welfare, Helsinki, Finland. [55]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [56]Max-Delbrück-Center (MDC) for Molecular Medicine, Berlin, Germany. [57]Clinic for Pediatric Allergy, Experimental and Clinical Research Center, Charité Universitätsmedizin Berlin, Berlin, Germany. [58]Department of Public Health Sciences, Henry Ford Health System, Detroit, MI, USA. [59]Divisions of Genetics, Genomics and Precision Medicine, Department of Medicine, University of Arizona College of Medicine, Tucson, AZ, USA. [60]Sachs' Children's Hospital, Stockholm, Sweden. [61]Programs in Metabolism and Medical & Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [62]Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA. [63]Institute of Clinical Research, University of Southern Denmark, Department of Obstetrics & Gynecology, Odense University Hospital, Odense, Denmark. [64]School of Medicine and Public Health, Faculty of Medicine and Health, University of Newcastle, Callaghan, New South Wales, Australia. [65]Centre for Occupational and Environmental Medicine, Stockholm County Council, Stockholm, Sweden. [66]Ludwig-Maximilians-University of Munich, Dr. von Hauner Children's Hospital, Division of Metabolic Diseases and Nutritional Medicine, Munich, Germany. [67]Institute for Respiratory Health, Harry Perkins Institute of Medical Research, University of Western Australia, Nedlands, Western Australia, Australia. [68]Ib-Salut, Area de Salut de Menorca, Institut d'Investigacio Sanitaria Illes Balears (IdISBa), Palma de Mallorca, Spain. [69]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain. [70]Channing Division of Network Medicine, Brigham & Women's Hospital and Harvard Medical School, Boston, MA, USA. [71]Department of Internal Medicine, Henry Ford Health System, Detroit, MI, USA. [72]QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [73]Population Health Research Institute, St George's, University of London, London, UK. [74]These authors contributed equally: Johannes Waage, Marie Standl. *e-mail: kb@copsac.com

Supplementary Information

Supplementary Note	**3**
Supplementary Results	3
Description of novel loci	3
Non-allergic rhinitis	4
Association in non-European populations	4
Cohort recruitment details, acknowledgements and funding statements	5
Supplementary References	24
Supplementary Figures	
Supplementary Figure 1**:** QQ-plot for the discovery phase of the allergic rhinitis GWAS	27
Supplementary Figure 2**:** Forest plots from the discovery phase – known loci	28
Supplementary Figure 3: Forest plots from the discovery and replication phase – novel loci	50
Supplementary Figure 4**:** Scatter plot of beta coefficients between 23andme and non-23andme cohorts	77
Supplementary Figure 5**:** Regional locus plots of allergic rhinitis discovery phase markers	78
Supplementary Figure 6**:** Enrichment plot **(**GARFIELD**)**	130
Supplementary Figure 7: "UpSet" set visualisation plot	131
Supplementary Figure 8: Overlap enriched gene sets allergic rhinitis and allergic sensitization	132
Supplementary Figure 9: Allergic rhinitis genetic risk score and disease prevalence	133
Supplementary Figure 10: Regional locus plot for rs193243426	134
Supplementary Tables	see separate file
Supplementary Table 1: Overview of participating cohorts	ST1
Supplementary Table 2: Extended association results for allergic rhinitis	ST2
Supplementary Table 3: Information for each cohort participating	ST3
Supplementary Table 4. Association results stratified by "doctor diagnosed" or "self-reported" allergic rhinitis	ST4
Supplementary Table 5: Results from MAGENTA analysis	ST5
Supplementary Table 6: List of conditional association results of secondary loci (GCTA)	ST6
Supplementary Table 7: List of locus-to-gene mappings from functional associations	ST7
Supplementary Table 8: List of markers part of the 99% credible set for each locus	ST8
Supplementary Table 9: List of coding and amino acid consequences from the credible sets	ST9
Supplementary Table 10: Association with classical HLA alleles (conditional analyses)	ST10
Supplementary Table 11: Association with classical HLA alleles (unconditional analyses)	ST11

Supplementary Table 12: Association of allergic rhinitis with HLA amino acid variants after iterative conditional analyses ST12

Supplementary Table 13: Association of allergic rhinitis with HLA amino acid variants (univariable analyses) ST13

Supplementary Table 14. Functional description of known and novel replicating loci ST14

Supplementary Table 15: Overlap with reported autoimmune disease variants ST15

Supplementary Table 16: Results from the GARFIELD regulatory region enrichment analysis ST16

Supplementary Table 17: Tissue enrichment results (DEPICT) ST17

Supplementary Table 18: Results from the Ingenuity Pathway enrichment analysis ST18

Supplementary Table 19: Table of index markers for allergic sensitisation ST19

Supplementary Table 20: Replication of allergic rhinitis top loci in the allergic sensitization GWAS ST20

Supplementary Table 21: DEPICT gene set enrichment analysis for allergic rhinitis ST21

Supplementary Table 22: DEPICT gene set enrichment analysis for allergic sensitization ST22

Supplementary Table 23: Replication of allergic rhinitis top loci in the non-allergic rhinitis GWAS ST23

Supplementary Table 24: Results of novel, replicating loci in the random-effects meta-analysis (non-European populations) ST24

Supplementary Table 25: List of proxies reported for the replication phase ST25

Supplementary Note

Supplementary Results

Description of novel loci

Three of the 20 loci not previously associated with AR have previously been associated with other atopic traits. These include the locus at 5p13.2 previously associated with eczema[1]. This index SNP is in strong LD with a missense variation in *IL7R* and with meQTLs in blood and lung tissue. IL-7R is necessary for V(D)J recombination of T and B cell receptors, for cellular activation by the type 2 immune inducer TSLP, and *IL7R* polymorphisms have been associated with the autoimmune disease multiple sclerosis[2, 3]. A second locus at 12q24.12 has been associated with blood eosinophil count[4]. A possible functional candidate at this locus, supported by a blood cis-eQTL and a correlated missense variation, is *SH2B3* which is involved in inflammatory pathways and T Cell activation[5]. The third locus at 19q13.11 has previously been associated with eosinophilic esophagitis.[6] We found evidence of enhancer-promotor interaction with *CEBPA* and *CEBPG* in macrophages, monocytes and neutrophils. CEBPA is essential for myeloid cell lineage differentiation[7] while CEBPG function as an enhancer binding protein in *IL4, IL6*, *IL8* and *IGH*, the immunoglobulin heavy chain locus[8].

The majority of the remaining novel loci imply genes with a known role in the immune system. The locus at 11q23 is located near *CXCR5,* a plausible causal candidate supported by cis eQTL and enhancer-promotor interaction in B-cells. *CXCR5* encodes a chemokine receptor that is present on B cells and in a subset of follicular T cells and is involved in B cell migration and facilitates B and T cell interactions within the lymph node[9]. This locus has previously been associated with autoimmune diseases[10,11] and lymphoma[12]. The lead signal at 1q23.3 is located near *FCER1G* and associated with expression of FCER1G in blood and lung tissue. *FCER1G* encodes the gamma chain of the high affinity IgE receptor, a central component in allergic responses, and is thereby a plausible candidate conferring risk of AR at this locus. The locus at 4q24 harbors cis-eQTLs for NFKB1 in monocytes and whole blood and meQTLs in lung tissue. *NFKB1* encodes a subunit of the NF-kappaB complex playing an important role for activation of multiple inflammatory pathways.[13] NF-kappaB activation might also be implicated in the association signal at 10q24.32 with evidence of enhancer-promoter interaction with *NFKB2*, encoding another subunit of the NF-kappaB complex. Other immune related candidates include *BACH2* at 6q15 with a role in antigen-induced formation of memory B cells and memory T cells[1415], *TYRO3* and *LTK* at 15q15.1 modulating TLR signaling[16] and T helper 2 immunity[17], respectively, *VPRBP* at 3p21.2 required for optimal T cell proliferation after antigen recognition and involved in V(D)J recombination during B cell development[18], *SPPL3* and *OASL* at 12q24.31 with potential roles for NK cell maturation[19] and IFN-alpha signaling[20] respectively, *RORA* at 15q22.2 involved in type 2 innate lymphoid cell development and allergic inflammation[21], and *TNFSF11* at 13q14.11 with a role in T cell activation by dendritic cells[22].

Finally, a number of the novel AR loci imply genes with no clear function in AR pathogenesis. These include one of the strongest associated loci in this meta-analysis at 12q24.31 with the top-signal located between *CDK2AP1* and *C12orf65*. This locus harbors cis-eQTLs in blood and/or lung tissue for several genes (*ABCB9, ARL6IP4, C12orf65, CDK2AP1, MPHOSPH9, OGFOD2, PITPNM2, RILPL2, SBNO1, SETD8,* and *SNRNP35*) and shows evidence for enhancer-promoter interaction with *DDX55* in various immune cells, but none of these genes have an obvious function related to AR. The same is the case for *NEGR1* at 1p31.1, *JAZF1* at 7p15.1,

FOSL2 at 2p23.2, and *RERE* at 1p36.32, all supported by cis-eQTLs or me-QTLs in blood and/or lung tissue. No cis-eQTL was identified for the locus at 2q36.3 near *DAW1*. Further studies of these genes and loci might help to understand AR pathogenesis and identify novel drug targets.

Non-allergic rhinitis

Non-allergic rhinitis (NAR), defined as rhinitis symptoms without evidence of allergic sensitization, is a common but poorly understood disease entity.[23] We performed the first GWAS on this phenotype hypothesizing that this might reveal specific rhinitis mechanisms. The analysis included 2,028 cases and 9,606 controls from 9 studies but did not identify any risk loci at the genome-wide significance level. Look-up of the 41 AR top-loci in the non-allergic rhinitis GWAS indicated considerable overlap of susceptibility loci with 33 of 41 loci showing consistent direction of effect and 8 AR loci showing nominal significance ($p<0.05$) in the non-allergic rhinitis GWAS (Supplementary Table 23). In line with this, we found that calculating the population-attributable risk fraction (PARF) for NAR using the risk score from AR showed a highly statistically significant PARF of 48% (95% CI 23%-65%), indicating that a significant proportion of NAR cases can be attributed to AR susceptibility loci. This suggests some genetic commonality between the two phenotypes that can be due to methodological issues, e.g. that some cases of non-allergic rhinitis misdiagnosed due to sensitization not captured by the tests performed, or might have a biological explanation, e.g. the presence of local allergic mechanisms in the nose without presence of systemic allergic sensitization.

Association in non-European populations

We only had data from 3 populations of non-European descent, including populations of African-American (SAPPHIRE, SAGEII) and Latin-American (GALAII) descent. The combined number of individuals was 6,313 providing insufficient statistical power to perform genome-wide association analysis. Look up of the identified novel AR loci was possible for 19 variants. Of these, 13 showed the same direction of effect and two (near RERE and BACH2) reached nominal statistical significance in the random-effects model (Supplementary Table 24). However, the limited statistical power in the non-European sample set does not allow solid interpretations about the potential overlap of AR loci between ethnicities.

Cohort recruitment details, acknowledgements and funding statements

23andme

Recruitment

All individuals included in the analyses provided informed consent and answered surveys online according to our human subjects protocol, which was reviewed and approved by Ethical & Independent Review Services, a private institutional review board (http://www.eandireview.com). Samples were drawn from 23andMe research participants who reported via web-based questionnaires and answered questions about allergic symptoms.

Case/Control definition

23andMe research participants were able to fill out web-based questionnaires whenever they logged into their 23andMe accounts.

For GWAS:

The ”allergic rhinitis” phenotype combines the questions from multiple surveys

(returning all consistent responses, inconsistent responses are set to NA)

1. Survey: "Asthma"

Question: "Have you ever had any of the following? Please check all that apply: "Allergic rhinitis (stuffed or dripping nose caused by allergies)"

2. Samples from ‘Allergies and Asthma’ and “Allergies” surveys that controls are those with no allergies to a broad range of allergens (e.g seasonal_allergies, environmental_allergy, food_allergy and drug_allergy); cases report the allergy to any of these:

[‘grasses’, ‘trees’, ‘weeds’, ‘cats’, ‘dogs’, ‘dust_mite’, ‘mold’],

and had symptoms ‘itchy or runny nose’.

2.1 Survey: "Allergies and Asthma"

Question: "Have you had an allergic reaction to any of the following types of plants?" "Grasses"

2.2 Survey: "Allergies and Asthma"

Question: "What type of reaction did you have after being exposed to grasses? Please check all that apply: "Itchy or runny nose"

2.3 Survey: “Allergies (allergies)"

Question: "Plant Allergies"/"Have you ever had an allergic reaction to any of these types of plants?" "Grasses"

2.4 Survey: "Allergies(allergies)",

Question: "What type of reaction did you have after being exposed to grass? Please check all that apply." : "Itchy or runny nose"

For replication analysis:

The “allergic rhinitis” phenotype combines the questions from multiple surveys

(returning all consistent responses, inconsistent responses are set to NA)

1. Survey: "Asthma"

Question: "Have you ever had any of the following? Please check all that apply: "Allergic rhinitis (stuffed or dripping nose caused by allergies)"

2. New Survey: "Asthma(asthma_background)"

Question: "Have you ever had any of the following? Please check all that apply: “Allergic rhinitis (stuffed or dripping nose caused by allergies)"

3. New Survey: "Asthma(asthma_meds)",

Question: "Have you ever had any of the following? Please check all that apply: "Allergic rhinitis (stuffed or dripping nose caused by allergies)"

4. Samples from allergy surveys that controls are those with no allergies to a broad range of allergens (e.g seasonal_allergies, environmental_allergy, food_allergy and drug_allergy); cases report the allergy to any of these:

['grasses', 'trees', 'weeds', 'cats', 'dogs', 'dust_mite', 'mold'],

and had symptoms 'itchy or runny nose'.

4.1 Survey: "Allergies and Asthma"

Question: "Have you had an allergic reaction to any of the following types of plants?" "Grasses"

4.2 Survey: "Allergies and Asthma"

Question: "What type of reaction did you have after being exposed to grasses? Please check all that apply: "Itchy or runny nose"

4.3 Survey: “Allergies (allergies)"

Question: "Have you ever had an allergic reaction to any of these types of plants?" "Grasses"

4.4 Survey: "Allergies(allergies)",

Question: "What type of reaction did you have after being exposed to grass? Please check all that apply." : "Itchy or runny nose"

Acknowledgements and Funding

We thank the research participants of 23andMe who answered surveys, as well as the employees of 23andMe, who together made this research possible. We particularly thank the following members of the 23andMe Research Team: Michelle Agee, Babak Alipanahi, Adam Auton, Robert K. Bell, Katarzyna Bryc, Sarah L. Elson, Pierre Fontanillas, Nicholas A. Furlotte, Bethann S. Hromatka, Karen E. Huber, Aaron Kleinman, Nadia K. Litterman, Matthew H. McIntyre, Joanna L. Mountain, Elizabeth S. Noblin, Carrie A. M. Northover, Steven J. Pitts, J. Fah Sathirapongsasuti, Olga V. Sazonova, Janie F. Shelton, Suyash Shringarpure, Vladimir Vacic, and Catherine H. Wilson. This work was supported in part by the National Heart, Lung, and Blood Institute of the US National Institutes of Health under grant 1R43HL115873-01, and the National Human Genome Research Institute of the National Institutes of Health under grant number R44HG006981.

AAGC (QIMR)

Recruitment

As part of the AAGC, we performed a GWAS of allergic rhinitis in unrelated individuals of European ancestry ascertained from the Australian population. Participants provided informed consent to participate in this study, which was approved by the respective ethics committees.

Acknowledgements and Funding

The AAGC was funded by a project grant from the NHMRC (613627).

AAGC members

Graham Jones[1], Patrick Danoy[2], Svetlana Baltic[3], Desiree Mészáros[4], Catherine Hayden[5], Sarah E Medland[6], Andrew J. Kemp[7], Faang Cheah[3], Dale R. Nyholt[6], Melissa C. Southey[8], Mary Roberts[9], Scott D. Gordon[6], Euan R. Tovey[1], Loren Price[3], Margaret J. Wright[6], James Markos1[0], Anjali K. Henders[6], Graham Giles[11], Li P. Chung[3], Paul S. Thomas[12], Ian Feather[13], Pamela A. Madden[14], Suzanna Temple[3], Stephen Morrison[15], Chalermchai Mitrpant[3], Brad Shelton[3], Andrew C. Heath[14], Mark Jenkins[2], Warwick J. Britton[16], John L. Hopper[17], Stephen R. Leeder[18], Haydn Walters[4], Michael J. Abramson[19], Matthew A Brown[2], Shyamali C. Dharmage[17]

[1] Woolcock Institute of Medical Research, University of Sydney, Sydney, Australia.
[2] University of Queensland Diamantina Institute, Princess Alexandra Hospital, Brisbane, Australia.
[3] Lung Institute of WA and Centre for Asthma, Allergy and Respiratory Research, University of WA, Perth, Australia.
[4] Menzies Research Institute, Hobart, Australia.
[5] School of Paediatrics and Child Health, Princess Margaret Hospital for Children, Perth, Australia.
[6] The Queensland Institute of Medical Research, Brisbane, Australia.
[7] The Children's Hospital, Westmead, Sydney, Australia.
[8] Department of Pathology, The University of Melbourne, Melbourne, Australia.
[9] Department of Respiratory Medicine, Royal Children's Hospital, Parkville, Australia.
[10] Launceston General Hospital, Lauceston, Australia.
[11] Cancer Epidemiology Centre, The Cancer Council Victoria, Melbourne, Australia.
[12] Faculty of Medicine, University of New South Wales, Sydney, Australia.
[13] Gold Coast Hospital, Southport, Australia.
[14] Washington University School of Medicine, St Louis, United States.
[15] University of Queensland, Brisbane, Australia.
[16] Centenary Institute of Cancer Medicine & Cell Biology, Royal Prince Alfred Hospital, Camperdown, Australia.
[17] Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of Melbourne, Melbourne, Australia.
[18] Australian Health Policy Institute, University of Sydney, Sydney, Australia

ALSPAC

Recruitment

ALSPAC recruited 15,247 pregnant women resident in Avon, UK with expected dates of delivery 1st April 1991 to 31st December 1992, resulting in 14,775 live births and 14,701 children who were alive at 1 year of age. Enrolment is described in more detail in the cohort profile paper[24] and via the website http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/. Biological samples including DNA have been collected for 10,121 of the children from this cohort. Ethical approval for the study was obtained

from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees and written and informed consent was provided by the parents.

Acknowledgements and Funding

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and the Wellcome Trust (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. GWAS data was generated by Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. FLG genotyping was funded by an MRC centenary award (awarded to L Paternoster).

B58C (British 1958 birth cohort)

Recruitment

The British 1958 birth cohort is an ongoing follow-up of all persons born in England, Scotland and Wales during one week in 1958. At ages 7, 11, 16, 23, 33 and 42 years, a history of hay fever or allergic rhinitis was obtained by interview with a parent (aged 7-16) or the cohort member (ages 23-42), At the age of 44-45 years, the cohort were followed up with a biomedical examination and blood sampling, from which a DNA collection was established as a nationally representative reference panel, and the presence of circulating IgE antibodies to house dust mite, mixed grass pollen and cat fur was ascertained by Hytec enzyme immunoassay, with a detection threshold of 0.35 kU/L.

Acknowledgements and Funding

We acknowledge use of phenotype and genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Genotyping for the B58C-WTCCC subset was funded by the Wellcome Trust grant 076113/B/04/Z. The B58C-T1DGC genotyping utilized resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. B58C-T1DGC GWAS data were deposited by the Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research (CIMR), University of Cambridge, which is funded by Juvenile Diabetes Research Foundation International, the Wellcome Trust and the National Institute for Health Research Cambridge Biomedical Research Centre; the CIMR is in receipt of a Wellcome Trust Strategic Award (079895). The B58C-GABRIEL genotyping was supported by a contract from the European Commission Framework Programme 6 (018996) and grants from the French Ministry of Research.

The funding agencies did not have any influence on design and conduct of the study; collection, management, and interpretation of the data; or preparation, review, or approval of the manuscript

BAMSE

Recruitment

BAMSE is a Swedish birth cohort study. A total number of 4,089 newborn infants were recruited between 1994 and 1996 in the Stockholm area[19]. The first questionnaire data, dealing with parental allergic diseases, socio-economic status and residential characteristics, was obtained when the children were about 2 months. Similar questionnaires, with a focus on the children's symptoms related to asthma and allergic diseases

including eczema, were answered by the parents when the children were approximately 1, 2, 4, 8 and 12 years old. At 8 years of age, all children were invited to clinical testing, and blood samples were obtained from 2,480 children (~60%). DNA was extracted from 2,033 samples after exclusion of samples with too little blood, lack of questionnaire data, or if parental consent to genetic analysis of the sample was not obtained.

Acknowledgements and Funding

BAMSE was supported by The Swedish Research Council, The Swedish Heart-Lung Foundation, Stockholm County Council (ALF) and the Strategic Research Programme (SFO) in Epidemiology at Karolinska Institutet. Genome-wide genotyping in BAMSE was funded by the European Commission as part of GABRIEL (A multidisciplinary study to identify the genetic and environmental causes of asthma in the European Community) contract number 018996 under the Integrated Program LSH-2004-1.2.5-1.

CAMP

Recruitment

The Childhood Asthma Management Program (CAMP) population is composed of non-Hispanic white subjects from a multicenter clinical trial that followed 1,041 children with asthma for four years and 84% of the original participants for 12 years. Stringent inclusion criteria ensured that participants had mild to moderate asthma, which was assessed as having asthma symptoms at least twice per week, using asthma medication daily, or using an inhaled bronchodilator twice per week for six or more months of the year prior to recruitment. CAMP subjects had increased airway responsiveness, as established by a bronchoprovocation test of up to 12.5mg/dl of methacholine resulting in 20% or greater forced expiratory volume in one second (FEV1) reduction.

Acknowledgements and Funding

We thank all CAMP subjects for their ongoing participation in this study. We acknowledge the CAMP investigators and research team, supported by NHLBI, for collection of CAMP Genetic Ancillary Study data. All work on data collected from the CAMP Genetic Ancillary Study was conducted at the Channing Laboratory of the Brigham and Women's Hospital under appropriate CAMP policies and human subject's protections. The CAMP Genetics Ancillary Study is supported by U01 HL075419, U01 HL65899, P01 HL083069 R01 HL 086601, from the National Heart, Lung and Blood Institute, National Institutes of Health, and Supinda Bunyavanich is funded by the US National Institutes of Health (NIH) R01 AI118833 and NIH K08 AI093538. In addition Supinda Bunyavanich is funded by US National Institutes of Health (NIH) R01 AI118833 and NIH K08 AI093538.

CEDAR

Recruitment

The CEDAR cohort comprises 323 healthy individuals of European descent visiting the University of Liège Hospital as part of a national screening for colon cancer. They provided blood samples and intestinal biopsies under full informed consent. The experimental protocol was approved by the ethics committee of the University of Liège Academic Hospital. Informed consent was obtained prior to donation in agreement with the recommendations of the declaration of Helsinki for experiments involving human subjects. We refer to this cohort as CEDAR for Correlated Expression and Disease Association Research.

Acknowledgements and funding

Assembling the CEDAR cohort was supported by grants from WELBIO (CAUSIBD), BELSPO (BeMGI), and Horizon 2020 (SYSCID). Computational resources at ULg have been provided by GIGA and the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique

(F.R.S.-FNRS) under Grant No. 2.5020.11. We thank Benoit Charloteaux, Wouter Coppieters, François Crins, Valérie Deffontaine, Elisa Docampo, Mahmoud Elansary, Anne-Stefan Gori, Rob Mariman, Myriam Mni, Yukihide Momozawa, Emilie Théâtre and the members of the GIGA genomics platform for their contributions.

COPSAC2000

Recruitment

The COPSAC2000 birth cohort study is a prospective clinical study of a birth cohort of 411 infants born to mothers with a history of asthma. The newborns were enrolled at the age of 1 month, the recruitment of which was previously described in detail.[25] The study was approved by the Ethics Committee for Copenhagen (KF 01- 289/96) and The Danish Data Protection Agency (2008-41-1754) and informed consent was obtained from both parents. For this study, only the parents were included.

Acknowledgements and Funding

COPSAC is funded by private and public research funds all listed on www.copsac.com. The Lundbeck Foundation (Grant no R16-A1694); The Danish Ministry of Health (Grant no 903516); Danish Council for Strategic Research (Grant no 0603-00280B); The Danish Council for Independent Research (Grant no 10-082884 and 271-08-0815); The Capital Region Research Foundation (No grant number) and NIH-NHLBI R01 HL129735 have provided core support for COPSAC.

deCODE

Recruitment

Icelandic patients with allergic rhinitis, based on physicians diagnosis, who attended an asthma clinic or emergency room at the National University Hospital of Iceland or the Icelandic Medical Center (Laeknasetrid) during the years 1977 to 2016 were recruited.[2627] All participating subjects who donated biospecimens provided informed consent. Personal identities of the participants and biological samples were encrypted by a third-party system approved and monitored by the Icelandic Data Protection Authority. The study, Genetics of asthma and allergies, was approved by the National Bioethics Committee (VSN-14-099, VS00-135) and the Data Protection Authority (no. PV_2014060841/ÞS) in Iceland.

Acknowledgements and Funding

We thank all the participants who provided data and samples for this study. We acknowledge the clinical investigators and staff of the Patient Recruitment Center for collection clinical information and biological samples. We thank the scientists at deCODE genetics and the staff deCODE's core facilities for genotyping, processing and analysis of genome sequences data. [28] The study is funded by deCODE genetics.

ECRHS

Study description

The European Community Respiratory Health Survey (ECRHS) is an international cohort study designed to assess the prevalence of asthma and allergic disease and identify their risk factors (Burney et al. The European Community Respiratory Health Survey.[29] Young adults of European descent were randomly recruited from community-based sampling frames in the ECRHS I (1991-1993) and followed up twice in the 20 years after the first assessment (ECRHS II: 1998-2002; ECRHS III: 2008-2013). Genotyping was performed on blood samples collected at ECRHS II. The current analysis was restricted to participants from centres that took part in the GABRIEL genotyping initiative and were representative of the random sample taken at baseline (i.e. this sample is not enriched with asthmatics). Ethical approval from local research ethics committees and written consent from participants were obtained.

Acknowledgements and Funding

The authors would like to thank the participants, field workers and researchers who have participated in the ECRHS study for their time and cooperation. This work was supported by a contract from the European Commission (018996), Fondo de Investigación Sanitaria (91/0016-060-05/E, 92/0319, 93/0393, 97/0035-01, 99/0034-01 and 99/0034-02), Hospital General de Albacete, Hospital General Ramón Jiménez, Consejería de Sanidad del Principado de Asturias, CIRIT (1997SGR 00079, 1999SGR 00241), and Servicio Andaluz de Salud, SEPAR, Public Health Service (R01 HL62633-01), RCESP (C03/09), Red RESPIRA (C03/011), Basque Health Department, Swiss National Science Foundation, Swiss Federal Office for Education and Science, Swiss National Accident Insurance Fund (SUVA), GSF-National Research Centre for Environment and Health, Deutsche Forschungsgemeinschaft (DFG) (FR 1526/1-1, MA 711/4-1), Programme Hospitalier de Recherche Clinique-DRC de Grenoble 2000 no. 2610, Ministry of Health, Direction de la Recherche Clinique, Ministere de l'Emploi et de la Solidarite, Direction Generale de la Sante, CHU de Grenoble, Comite des Maladies Respiratoires de l'Isere. UCB-Pharma (France), Aventis (France), Glaxo France. Estonian Science Foundation, and Asthma UK (formerly known as National Asthma Campaign UK).

EVE (CAG/CSGA)

Recruitment

Subjects were from the Chicago Asthma Genetics Study (CAG) and the NHLBI Collaborative Studies on the Genetics of Asthma (CSGA).

CAG included subjects recruited from a) families ascertained through asthma affected sib pairs, b) asthma affected children, c) adults and children with severe persistent asthma, and d) non-asthma controls over 18 years of age. Asthma was defined as 1) a physicals diagnosis, 2) the presence of at least two self-reported symptoms, 3) current use of medications, and 4) bronchial hyperresponsiveness to methacholine or reversibility to inhaled bronchodilator. Controls had no self-reported personal or family history (1st degree relatives) of asthma.

CSGA subjects were recruited from the NHLBI Collaborative Studies on the Genetics of Asthma (CSGA). CSGA cases were recruited and diagnosed as described above for CAG.

Acknowledgements and Funding

EVE was funded by an NIH grant RC2 HL101651.

EVE (CHS)

Recruitment

Two cohorts in the Children's Health Study (CHS) including a total of 3000 children (1512 from cohort 1, and 1488 from cohort 2) who were either Hispanic (n=1398) or non-Hispanic white (n=1602) were included in the analysis. The first cohort included children between the ages of 8-14 years recruited from schools in 12 southern California communities in 1993 and 1996. The second cohort included children between the ages of 5-8 years recruited from schools in 13 communities in 2002. A total of 1249 cases were classified as having asthma if the adult completing the questionnaire reported that a doctor had "ever diagnosed the child as having asthma". A total of 1751 controls were selected using frequency matching with cases on sex, ethnicity, cohort, and follow-up time from children without the asthma outcome.

Acknowledgements and Funding

The CHS was supported by the Southern California Environmental Health Sciences Center (grant P30ES007048); National Institute of Environmental Health Sciences (grants 5P01ES011627, ES021801,

ES023262); P01ES009581, P01ES011627, P01ES022845, R01 ES016535, R03ES014046, P50 CA180905, R01HL061768, R01HL076647, R01HL087680, and RC2HL101651), the Environmental Protection Agency (grant #s RD83544101, R826708, RD831861, and R831845), and the Hastings Foundation.

FINNTWIN (FTC)

Recruitment

The total of 2373 subjects originate within the Finnish Twin Cohort (FTC).[3031] Altogether 1161 of the subjects belong to the Older Finnish Twin Cohort consisting of adult twins born in 1938-1957.[32] Altogether 687 of the subjects belong to the FinnTwin12 cohort, a population-based longitudinal study of five consecutive birth cohorts (1983-1987) of Finnish twins. Altogether 525 of the subjects belong to the FinnTwin16 cohort, a population-based longitudinal study of five consecutive birth cohorts (1975-1979) of Finnish twins and their families.

Acknowledgements and Funding

We warmly thank the participating twin pairs and their family members for their contribution. We would like to express our appreciation to the skilled study interviewers A-M Iivonen, K Karhu, H-M Kuha, U Kulmala-Gråhn, M Mantere, K Saanakorpi, M Saarinen, R Sipilä, L Viljanen and E Voipio. Anja Häppölä and Kauko Heikkilä are acknowledged for their valuable contribution in recruitment, data collection, and data management.

Phenotyping and genotyping of the Finnish twin cohorts has been supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (grants 213506, 129680), the Academy of Finland (grants 100499, 205585, 118555, 141054, 265240, 263278 and 264146 to J. Kaprio), National Institute of Alcohol Abuse and Alcoholism (grants AA-12502, AA-00145, and AA-09203 to R. J. Rose and AA15416 and K02AA018755 to D. M. Dick), Sigrid Juselius Foundation (to J. Kaprio), Global Research Award for Nicotine Dependence, Pfizer Inc. (to J. Kaprio), and the Welcome Trust Sanger Institute, UK. Antti-Pekka Sarin and Samuli Ripatti are acknowledged for genotype data quality controls and imputation.

The funding agencies did not have any influence on design and conduct of the study; collection, management, and interpretation of the data; or preparation, review, or approval of the manuscript.

GABRIELA

Recruitment

The GABRIELA study (GABRIEL Advanced Studies) was a cross-sectional population-based study in 5 rural regions of Europe: Baden-Württemberg and Bavaria in Germany, North/Central Switzerland, Tyrol in Austria, and Lower Silesia in Poland. In phase 1, a recruitment questionnaire was sent out to 132,518 children age 5 to 13 years through their primary schools. Completed questionnaires were returned by 79,888 (60.3%). With a documented parental consent to dust sampling, blood sampling, and genetic analyses 34,491 children were eligible for phase 2. A stratified random selection process was applied to enrich informative observations. In a first step, 9668 children were selected for phase 2 within 3 exposure strata per center, i.e. farm children, non-farm children with and without farm exposure. In the second step, 1708 children from the Austrian, German, and Swiss centers were selected for genotyping within the mentioned exposure strata and simultaneously stratified for outcome, i.e. asthma, atopic sensitization, and unaffected controls.[33]

Acknowledgements and Funding

The GABRIELA study was supported by the European Commission as part of GABRIEL (a multidisciplinary study to identify the genetic and environmental causes of asthma in the European Community) contract

number 018996 under the Integrated Program LSH-2004-1.2.5-1. Michael Kabesch was supported by the German Ministry for Education and Science (BMBF) with the fund SYSINFAME, 01ZX1306E.

The GABRIELA study group

Silvia Apprich, PhD, Andrzej Boznanski, MD, PhD, Charlotte Braun-Fahrländer, MD,

Gisela Büchele, MPH, William Cookson, MD, DPhil, Paul Cullinan, MD, Hanna Danielewicz, MD, Anna Debinska, Martin Depner, PhD, Markus J. Ege, MD, Urs Frey, MD, PhD, Oliver Fuchs, MD, Jon Genuneit, Dick Heederik, PhD, Elisabeth Horak, MD, Anne Hyvärinen, PhD, Sabina Illi, PhD, Michael Kabesch, MD, Katalin Kovacs, Aleksandra Kosmeda, PhD, Wolfgang Kneifel, PhD, Philipp Latzin,MD, PhD, Roger Lauener, MD, Georg Loss, Stephanie MacNeill, MSc, Bernhard Morass, MD, Anne-Cécile Normand, Ilka Noss, PhD, Renaud Piarroux, MD, PhD, Helena Rintala, PhD, Mascha K. Rochat, MD, Nikolaos Sitaridis, Barbara Sozanska, MD, David Strachan, MD, Christine Strunz-Lehner, MPH, Bertrand Sudre, MD, PhD, Erika von Mutius, MD, MSc, Marco Waser, PhD, Juliane Weber, MD, and Inge Wouters, PhD.

From Innsbruck and Vienna, Austria; Munich, Regensburg, and Ulm, Germany; Kuopio, Finland; Besancon and Lausanne, France, Utrecht, the Netherlands; Wroclaw, Poland; Basel, Bern, Davos, and St. Gallen, Switzerland; London, UK.

GALA II

Recruitment

The Genes-environments & Admixture in Latino Americans (GALA II) study is an ongoing case-control study of asthma in Latinos that are recruited in five centers throughout the U.S. (Chicago, Illinois; Bronx, New York; Houston, Texas; San Francisco Bay Area, California; and San Juan, Puerto Rico).[34] Participants were eligible if they were aged between 8 to 21 years, had <10 pack-years of smoking history and self-identified all four grandparents as Latino. Subjects were excluded if they reported any of the following: (1) 10 or more pack-years of smoking; (2) any smoking within one year of recruitment date; (3) pregnancy in the third trimester; or (4) history of one of the following conditions: sickle cell disease, cystic fibrosis, sarcoidosis, cerebral palsy, or history of heart or chest surgery. Both children and young adults with physician-diagnosed asthma and non-allergic, non-asthmatic controls were recruited. The presence of allergic rhinitis was defined as a positive response to the question “Has a doctor ever diagnosed the child with hay fever or allergic rhinitis?” According to this, a total of 755 individuals were classified as cases and 2,929 as controls.

All local institutional review boards approved the study and all subjects/parents provided written assent/consent, respectively.

Acknowledgements and Funding

The authors acknowledge the families and patients for their participation and thank the numerous health care providers and community clinics for their support and participation in GALA II. In particular, the authors thank the study coordinator Sandra Salazar and the recruiters who obtained the data: Duanny Alva, MD, Gaby Ayala-Rodriguez, Ulysses Burley, Lisa Caine, Elizabeth Castellanos, Jaime Colon, Denise DeJesus, Iliana Flexas, Blanca Lopez, Brenda Lopez, MD, Louis Martos, Vivian Medina, Juana Olivo, Mario Peralta, Esther Pomares, MD, Jihan Quraishi, Johanna Rodriguez, Shahdad Saeedi, Dean Soto, Ana Taveras, Emmanuel Viera. Some computations were performed using the UCSF Biostatistics High Performance Computing System. We also acknoledge the prinicipal investigators of the centers where the GALA II and SAGE II participants were recruited: Kelley Meade, Harold J. Farber, Pedro C. Avila, Denise Serebrisky, Shannon M. Thyne, Emerita

Brigino-Buenaventura, William Rodriguez-Cintron, Saunak Sen, Rajesh Kumar, Michael Lenoir, Luisa N. Borrell, and Jose R. Rodriguez-Santana.

The GALA II study was funded by the Sandler Family Foundation, the American Asthma Foundation, the RWJF Amos Medical Faculty Development Program, Harry Wm. and Diana V. Hind Distinguished Professor in Pharmaceutical Sciences II, National Institutes of Health 1R01HL117004, R01Hl128439, R01HL135156, 1X01HL134589, National Institute of Health and Environmental Health Sciences R01ES015794, R21ES24844, the National Institute on Minority Health and Health Disparities 1P60MD006902, RL5GM118984, 1R01MD010443, and the Tobacco-Related Disease Research Program under Award Number 24RT-0025 to E.G.B. MP-Y was supported by the Ramón y Cajal Program (RYC-2015-17205) by the Spanish Ministry of Economy, Industry and Competitiveness; an award (AC15/00015) by the ISCIII through Strategic Action for Health Research (AES) and European Community (EC) within the Active and Assisted Living (AAL) Programme framework (MP-Y); and the SysPharmPedia grant from the ERACoSysMed 1st Joint Transnational Call from the European Union under the Horizon 2020.

The content of this publication is solely the responsibility of the authors and does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government and the funding agencies.

GENEVA-KORA

Recruitment

A total of 1895 individuals were obtained from an atopic dermatitis case collection recruited from tertiary dermatology clinics based at three centres (Technische Universität Munich, as part of the GENEVA study, University of Kiel, University of Bonn) as well as from two German population-representative cohorts, the PopGen biorepository[35] and the KORA study.[36]

Acknowledgements and Funding

The project received infrastructure support through the DFG Clusters of Excellence “Inflammation at Interfaces” (grants EXC306 and EXC306/2), and was supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (sysINFLAME, grant # 01ZX1306A), and the PopGen 2.0 network (01EY1103). The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

GENR

Recruitment

The Generation R Study is a population-based prospective cohort study of pregnant women and their children from fetal life onwards in Rotterdam, The Netherlands. All children were born between April 2002 and January 2006, and currently followed until young adulthood. Of all eligible children in the study area, 61% were participating in the study at birth. Cord blood samples including DNA have been collected at birth. The study protocol was approved by the Medical Ethical Committee of the Erasmus Medical Centre, Rotterdam (MEC 217.595/2002/20). Written informed consent was obtained from parents of all participants.

Acknowledgements and Funding

The Generation R Study gratefully acknowledges the contributions of the children and their parents, the general practitioners, the hospitals and the midwives and pharmacies in Rotterdam. We would like to thank A. Abuseiris, K. Estrada, Dr. T. A. Knoch, and R. de Graaf as well as their institutions Biophysical Genomics, Erasmus MC Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German D-Grid, both funded by the German Bundesministerium fuer Forschung und Technology under grants #01 AK 803 A-H and # 01 IG 07015 G, for access to their grid resources. We thank M. Jhamai, M. Ganesh, P. Arp, M. Verkerk, L. Herrera and M. Peters for their help in creating, managing and performing quality control for the genetic database. Also, we thank K. Estrada and C. Medina-Gomez for their support in the creation and analysis of imputed data. The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with the School of Law and the Faculty of Social Sciences of Erasmus University Rotterdam, the Municipal Health Service, Rotterdam area, the Rotterdam Homecare Foundation and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR-MDC; Rotterdam). The generation and management of genotype data for the Generation R Study were performed at the Genetic Laboratory of the Department of Internal Medicine at Erasmus Medical Center. The Generation R Study is made possible by financial support from the Erasmus Medical Center (Rotterdam), the Erasmus University Rotterdam and the Netherlands Organization for Health Research and Development (ZonMw; 21000074). The project received funding from the European Union's Horizon 2020 research and innovation programme (LIFECYCLE project, grant agreement no 733206; 2016).Dr Vincent Jaddoe received an additional grant from the Netherlands Organization for Health Research and Development (VIDI 016.136.361) and the European Research Council (Consolidator Grant, ERC-2014-CoG-648916). Dr Liesbeth Duijts received an additional grant from the Lung Foundation Netherlands (no 3.2.12.089; 2012) and from the European Union's Horizon 2020 co-funded programme ERA-Net on Biomarkers for Nutrition and Health (ERA HDHL) (ALPHABET project (no 696295; 2017), ZonMW The Netherlands (no 529051014; 2017)). The researchers are independent from the funders. The study sponsors had no role in the study design, data analysis, interpretation of data, or writing of this report.

GENUFAD

Recruitment

All patients were recruited at Charité Universitätsmedizin Berlin for the GENUFAD study (GEnetic analysis of NUclear Families with Atopic Dermatitis) which has been described in a previous GWAS.[37] Briefly, 270 complete nuclear German families were recruited through two siblings affected by eczema with an age of onset below two years. For all family members, eczema, allergic rhinitis, and asthma status was determined by a physician at recruitment. To ensure independent, unrelated cases, all parents with allergic rhinitis were included in this study. If the parents were unaffected, only one child with allergic rhinitis per family was included. Allergic sensitization status was determined for all cases.

German control individuals originated from the Study of Health in Pomerania (SHIP), a population-based cohort from North-Eastern Germany.[38]

Acknowledgements and Funding

We thank all individuals and families for their participation in this study. We thank all physicians and nurses involved in patient recruitment for their valuable contribution to the study. We are grateful to the laboratory technicians C. Flachmeier and T. Thuss for their work. The study was funded by the German Ministry of Education and Research (BMBF) through the Clinical Research Group for Allergy at Charité Berlin, the National Genome Research Network (NGFN).

SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI_MED) funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012).

GINI/LISA

Recruitment

The influence of Life-style factors on the development of the Immune System and Allergies in East and West Germany PLUS the influence of traffic emissions and genetics (LISAplus) Study is a population based birth cohort study. A total of 3094 healthy, full-term neonates were recruited between 1997 and 1999 in Munich, Leipzig, Wesel and Bad Honnef. The participants were not preselected based on family history of allergic diseases.[39] A total of 5991 mothers and their newborns were recruited into the German Infant study on the influence of Nutrition Intervention PLUS environmental and genetic influences on allergy development (GINIplus) between September 1995 and June 1998 in Munich and Wesel. Infants with at least one allergic parent and/or sibling were allocated to the interventional study arm investigating the effect of different hydrolysed formulas for allergy prevention in the first year of life.[40] All children without a family history of allergic diseases and children whose parents did not give consent for the intervention were allocated to the non-interventional arm. Detailed descriptions of the LISAplus and GINIplus studies have been published elsewhere1,2. DNA was collected at the age 6 and 10 years and 1511 children from the Munich study center from both studies were genotyped. For both studies, approval by the local Ethics Committees and written consent from participant's families were obtained.

Acknowledgements and Funding

The authors thank Martina Müller-Nurasyid for her help with imputation and analysis, all families for participation in the studies and the LISAplus and GINIplus study teams for their excellent work. The GINIplus study team wishes to acknowledge the following: Institute of Epidemiology I, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg (Heinrich J, Brüske I, Schulz H, Flexeder C, Zeller C, Standl M, Schnappinger M, Ferland M, Thiering E, Tiesler C); Department of Pediatrics, Marien-Hospital, Wesel (Berdel D, von Berg A); Ludwig-Maximilians-University of Munich, Dr von Hauner Children's Hospital (Koletzko S); Child and Adolescent Medicine, University Hospital rechts der Isar of the Technical University Munich (Bauer CP, Hoffmann U); IUF- Environmental Health Research Institute, Düsseldorf (Schikowski T, Link E, Klümper C). The LISAplus study team wishes to acknowledge the following: Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Epidemiology I, Munich (Heinrich J, Schnappinger M, Brüske I, Ferland M, Lohr W, Schulz H, Zeller C, Standl M); Department of Pediatrics, Municipal Hospital "St. Georg", Leipzig (Borte M, Gnodtke E); Marien Hospital Wesel, Department of Pediatrics, Wesel (von Berg A, Berdel D, Stiers G, Maas B); Pediatric Practice, Bad Honnef (Schaaf B); Helmholtz Centre of Environmental Research – UFZ, Department of Environmental Immunology/Core Facility Studies, Leipzig (Lehmann I, Bauer M, Röder S, Schilde M, Nowak M, Herberth G , Müller J, Hain A); Technical University Munich, Department of Pediatrics, Munich (Hoffmann U, Paschke M, Marra S); Clinical Research Group Molecular Dermatology, Department of Dermatology and Allergy, Technische Universität München (TUM), Munich (Ollert M).

GOA

Recruitment

Subjects were drawn from the Genetics of Asthma study (GOA) in the Spanish population that have been previously described.[41,42] Cases with rhinitis were diagnosed by a physician, aged ≥5 years, taking medications to control their asthma symptoms, and had an extensive clinical characterization, including allergic sensitization by skin prick tests and specific IgE measurements. Control subjects self-reported absence of personal or familiar history of chronic diseases, and were obtained from the National Genotyping Center (CEGEN-PRB2-ISCIII), Universidad de Santiago de Compostela Node.

Acknowledgements and Funding

This work was supported by Instituto de Salud Carlos III (FIS PI11/00623 and FI12/00493) and co-financed by the European Regional Development Funds, "A way of making Europe" from the European Union, and by a grant from the 7th Framework Programme (FP7-REGPOT-2012-CT2012-31637-IMBRAIN).

GOYA

Recruitment

The GOYA study (Genetics of Overweight Young Adults) is derived from the genome wide association study (GWAS) on obesity with a case-cohort (obese-non obese) design comprising young men and women of Danish origin.[43]

The GOYA (Male) cohort is a longitudinal case-cohort (obese, non-obese) study comprising a randomly (1%) selected control group and all extremely overweight men identified among 362,200 Caucasian men examined at the mean age of 20 years at the draft boards in Copenhagen and its surrounding areas during 1943–1977. Obesity was defined as 35% overweight relative to a local standard in use at the time (mid 1970's), corresponding to a BMI ≥31.0 kg/m2, which proved to be above the 99th percentile. All of the obese and 50% of the random sampled controls, who were still living in the region, were invited to a follow-up survey in 1992–94 at the mean age of 46 years, at which time the blood samples were taken and genotyping were performed for a total of 673 extremely overweight and 792 controls. With a sampling fraction of 0.5% (50% of 1%), the controls represent about 158,000 men among whom the case group was the most obese.

The GOYA (Female) cohort cohort is nested within the Danish National Birth Cohort (DNBC[43]) which is a collection of data on 92,274 pregnant women recruited between 1996 and 2002, from their first antenatal visit to their general practitioner. Women participated in four telephone interviews (16 and 30 weeks gestation and 6 and 18 months after birth). They also provided two blood samples during pregnancy. The GOYA females were drawn from the 67,863 women within the DNBC who provided information about prepregnancy BMI, gave birth to a live born infant and provided a blood sample during pregnancy. A case sample of the 3.6% most obese women (n=2451) was defined as those with the largest residuals from the regression of BMI on age and parity (all entered as continuous variables). The BMI for these women ranged from 32.6 to 64.4. From the remaining cohort we selected a random sample of similar size (n=2450). Of these, 1,960 extremely overweight and 1,948 control women were genotyped and passed quality control. Information on mother's year of birth and age at first birth derived from national registers. The study was approved by the regional scientific ethics committee and by the Danish Data Protection Board.

Acknowledgements and Funding

This GOYA male study was conducted as part of the activities of the Gene-diet Interactions in Obesity project (GENDINOB, www.gendinob.dk) and the MRC Centre for Causal Analyses in Translational Epidemiology (MRC CAiTE). TSA received his Post-Doctoral Research Grant from the GENDINOB project and acknowledges the same.

The GOYA female part of the study was sampled as a case-cohort study within the Danish National Birth Cohort. The establishment of this cohort was based on a major grant from the Danish National Research Foundation. Additional support for the Danish National Birth Cohort was obtained from the Pharmacy Foundation, the Egmont Foundation, the March of Dimes Birth Defects Foundation and the Augustinus Foundation. Genotyping for the GOYA female study was funded by the Wellcome Trust (Grant ref: 084762MA).

We thank all the participants of the GOYA study.

INMA

Recruitment

Population-based birth cohorts were established as part of the INMA – INfancia y Medio Ambiente [Environment and Childhood] Project in several regions of Spain following a common protocol. This analysis uses the INMA subcohort of Menorca established between 1997 and 1998 in Menorca island. This project aims to study the associations between pre- and postnatal environmental exposures and growth, health, and development from early fetal life until adolescence and has been described previously in detail.[44] Pregnant women were enrolled during the 1st trimester of pregnancy at public primary health care centers or public hospitals and children were followed after birth until 18y of age. Informed consent was obtained from all participants and the study was approved by the Hospital Ethics Committee.

Acknowledgements and Funding

This study was funded by grants from Instituto de Salud Carlos III (97/0588, 00/0021-02, G03/176, 04/1705, 06/1756, 09/1958 and PS09/00432), Fundació La Marató de TV3, Obra Social la Caixa (IV Convocatoria de ayudas a la Investigación en Enfermedades Neurodegenerativas) and European Community (Contract No. QLK4-CT-2000-00263). DNA extractions and genotyping was performed at the Spanish National Genotyping Centre (CEGEN-Barcelona). The authors would particularly like to thank all the participants for their generous collaboration. The authors are grateful to Mireia Garcia, Maria Victoria Estraña, Maria Victoria Iturriaga, Cristina Capo and Josep LLuch for their assistance in contacting the families and administering the questionnaires. A full roster of the INMA Project Investigators can be found at http://www.proyectoinma.org/presentacion-inma/listado-investigadores/en_listado-investigadores.html.

MAAS

Recruitment

The Manchester Asthma and Allergy Study is an unselected (i.e. population-based), birth cohort study. The setting is the maternity catchment area of Wythenshawe and Stepping Hill Hospitals, comprising of 50 square miles of South Manchester and Cheshire, UK, a stable mixed urban-rural population. Study was approved by the Local Research Ethics Committee. Informed consent was obtained from all parents.

Screening & Recruitment

All pregnant women were screened for eligibility at antenatal visits (8th-10th week of pregnancy). The study was explained to the parents, and informed consent for initial questionnaires and skin prick testing was obtained. Both parents completed a questionnaire about their and their partner's history of asthma and allergic diseases and smoking habits. If the pregnant woman's partner was not present at the antenatal clinic visit, an invitation was sent for him to attend an open-access evening clinic for skin prick testing and questionnaire. Once both parents had completed questionnaires and skin prick testing, a full explanation of the proposed future follow-up for the child was given. Of the 1499 couples who met the inclusion criteria

(<10 weeks of pregnancy, maternal age >18 years, questionnaire and skin test data available for both parents), 288 declined to take part in the study. A total of 1185 participants had at least some evaluable data.

Follow-up

The children have been followed prospectively, and attended review clinics at ages 1, 3, 5, 8 and 11 years.

Acknowledgements and Funding

We would like to thank the children and their parents for their continued support and enthusiasm. We greatly appreciate the commitment they have given to the project. We would also like to acknowledge the hard work and dedication of the study team (post-doctoral scientists, research fellows, nurses, physiologists, technicians and clerical staff). This report includes independent research supported by National Institute for Health Research Respiratory Clinical Research Facility at Manchester University NHS Foundation Trust (Wythenshawe). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health. MAAS was supported by the Asthma UK Grants No 301 (1995-1998), No 362 (1998-2001), No 01/012 (2001-2004), No 04/014 (2004-2007), the BMA James Trust and Medical Research Council, UK (G0601361) and The Moulton Charitable Foundation (2004-current); the Medical Research Council (MRC) Grants G0601361, MR/K002449/1 and MR/L012693/1, and Angela Simpson is supported by the NIHR Manchester Biomedical Research Centre. The authors would like to acknowledge the North West Lung Centre Charity for supporting this project.

NFBC1966

Recruitment

The Northern Finland Birth Cohort 1966 is a prospective follow-up study of children from the two northernmost provinces of Finland. Women with expected delivery dates in 1966 were recruited through maternity health centers. Cohort members living in northern Finland or in the capital area were invited to a clinical examination as well as questionnaire at age 31 years. DNA was extracted from blood samples given at the clinical examination. Informed consent for the use of the data including DNA was obtained from all subjects. The study was approved by the ethics committees in Oulu (Finland) and Oxford (UK) universities in accordance with the Declaration of Helsinki.

Acknowledgements and Funding

We thank the late Professor Paula Rantakallio (launch of NFBC1966), and Ms Outi Tornwall and Ms Minttu Jussila (DNA biobanking). The authors would like to acknowledge the contribution of the late Academian of Science Leena Peltonen.

The NFBC resource has been supported by grants from the Academy of Finland (project grants 104781, 120315, 129269, 1114194, 24300796, Center of Excellence in Complex Disease Genetics and SALVE), University Hospital Oulu, Biocenter, University of Oulu, Finland (75617), NHLBI grant 5R01HL087679-02 (1RL1MH083268-01), NIH/NIMH (5R01MH63706:02), ENGAGE project and grant agreement HEALTH-F4-2007-201413, EU FP7 EurHEALTHAgeing -277849, the Medical Research Council, UK (G0500539, G0600705, G1002319, PrevMetSyn/SALVE) and the MRC, Centenary Early Career Award. H2020 DynaHEALTH (European Union's Horizon 2020 research and innovation programme under grant agreement No 633595); Exposomic, Genomic and Epigenomic Approach to Prediction of Metabolic and Cardiorespiratory function and Ill-Health (EGEA), Academy of Finland, Grant No 285547; ALEC Study (funded by the European Union's Horizon 2020 Research and Innovation programme under grant agreement No. 633212); H2020 / Marie Skłodowska-Curie Actions, CAPICE (Marie Curie Grant agreement Number 721567); National Public Health Institute, Biomedicum Helsinki, Finland.

NTR

Recruitment

In the Young Netherlands Twin Register (YNTR), twins are followed from birth onwards. Around every two to three years, a survey is sent out inquiring about physical and mental health. At ages 1, 2, 3, 5, 7, 10 and 12 years, the surveys are completed by the parents and/or teachers, from age 14 onward, twins and their non-twin siblings are asked to complete the surveys by themselves. At age 5 years of the twins, parents are asked a series of questions about allergies and rhinitis.[45] Adult and adolescent twins (ANTR) have been recruited through city councils and media, and complete a self-rating survey every few years. Surveys contain a series of questions about doctor diagnosed diseases and disorders. Subsamples of twin families have taken part in biobank and genotyping studies.

Acknowledgements and Funding

The Netherlands Organization for Scientific Research (NWO 480-15-001/674): Netherlands Twin Registry Repository: researching the interplay between genome and environment; Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007); Amsterdam Public Health (APH); KNAW Academy Professor Award **(**PAH/6635); the Avera Institute, Sioux Falls, South Dakota (USA) and Grand Opportunity grants 1RC2 MH089951 and 1RC2 MH089995).

RAINE

Recruitment

The Western Australian Pregnancy Cohort (Raine) Study is a prospective pregnancy cohort where 2900 were recruited from King Edward Memorial Hospital between 1989 and 1991. Data were collected throughout pregnancy and the children have been followed-up at ages 1, 2, 3, 5, 8, 10, 14, 17, 18, 20, and 23. Ethics approval for this study was obtained from King Edward Memorial Hospital and Princess Margaret Hospital. Participants were consented to being involved in this study prior to each follow-up.

Acknowledgements and Funding

This study was supported by the National Health and Medical Research Council of Australia [grant numbers 572613, 403981 and 003209] and the Canadian Institutes of Health Research [grant number MOP-82893]. The authors are grateful to the Raine Study participants and their families, and to the Raine Study research staff for cohort coordination and data collection. The authors gratefully acknowledge the NH&MRC for their long term contribution to funding the study over the last 29 years and also the following Institutions for providing funding for Core Management of the Raine Study: The University of Western Australia (UWA), Curtin University, Raine Medical Research Foundation, UWA Faculty of Medicine, Dentistry and Health Sciences, The Telethon Kids Institute, Women and Infants Research Foundation (King Edward Memorial Hospital), Murdoch University, The University of Notre Dame (Australia), and Edith Cowan University. The authors gratefully acknowledge the assistance of the Western Australian DNA Bank (National Health and Medical Research Council of Australia National Enabling Facility).

This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia.

SAGE II

Recruitment

The Study of African Americans, Asthma, Genes & Environments (SAGE II) is an ongoing study of asthma in children and young adults coordinated from the University of California San Francisco. Recruitment protocols and the definition of allergic rhinitis were similar to GALA II with the only differences being that the

recruitment was restricted to San Francisco Bay Area and participants self-identified all four grandparents as African American.[46] In the current study, 385 cases of allergic rhinitis and 978 controls were analyzed.

Acknowledgements and Funding

The authors acknowledge the patients, families, numerous recruiters, health care providers and community clinics for their participation in the SAGE II study.

N.H-P was funded by a fellowship (FI16/00136) from the Instituto de Salud Carlos III (ISCIII) (NH-P) and co-funded by the European Social Funds from the European Union (ESF) "ESF invests in your future". The SAGE II study was funded by the Sandler Family Foundation, the American Asthma Foundation, the RWJF Amos Medical Faculty Development Program, Harry Wm. and Diana V. Hind Distinguished Professor in Pharmaceutical Sciences II, National Institutes of Health 1R01HL117004, R01Hl128439, R01HL135156, 1X01HL134589, National Institute of Health and Environmental Health Sciences R01ES015794, R21ES24844, the National Institute on Minority Health and Health Disparities 1P60MD006902, U54MD009523, 1R01MD010443, and the Tobacco-Related Disease Research Program under Award Number 24RT-0025 to E.G.B. The content of this publication is solely the responsibility of the authors and does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. We also acknoledge the prinicipal investigators of the centers where the GALA II and SAGE II participants were recruited: Kelley Meade, Harold J. Farber, Pedro C. Avila, Denise Serebrisky, Shannon M. Thyne, Emerita Brigino-Buenaventura, William Rodriguez-Cintron, Saunak Sen, Rajesh Kumar, Michael Lenoir, Luisa N. Borrell, and Jose R. Rodriguez-Santana.

SAPALDIA

Recruitment

SAPALDIA is a population-based cohort from Switzerland that recruited subjects aged 18 to 60 from population registries in eight communities, representing the three largest language groups (German, French, Italian) as well as different levels of air pollution and degrees of urbanization**.** Subjects underwent spirometry and answered a detailed questionnaire on respiratory health, allergies, smoking history, and lifestyle factors in the baseline (year 1991) and follow-up (year 2002) examination. 6,055 subjects participated in both examinations and agreed to provide blood for genetic analysis. All subjects that were part of a nested asthma case-control sample subjected for genomewide genotyping in the context of the GABRIEL genome-wide association study on asthma formed the basis of the current study. In the current study, subjects were considered to be a case if they ever reported to have had allergic rhinitis at the baseline or follow-up examination. The analysis included 526 cases of allergic rhinitis and 921 controls.

Acknowledgements

Study directorate: NM Probst-Hensch (PI; e/g); T Rochat (p), C Schindler (s), N Künzli (e/exp), JM Gaspoz (c)

Scientific team: JC Barthélémy (c), W Berger (g), R Bettschart (p), A Bircher (a), C Brombach (n), PO Bridevaux (p), L Burdet (p), Felber Dietrich D (e), M Frey (p), U Frey (pd), MW Gerbase (p), D Gold (e), E de Groot (c), W Karrer (p), F Kronenberg (g), B Martin (pa), A Mehta (e), D Miedinger (o), M Pons (p), F Roche (c), T Rothe (p), P Schmid-Grendelmeyer (a), D Stolz (p), A Schmidt-Trucksäss (pa), J Schwartz (e), A Turk (p), A von Eckardstein (cc), E Zemp Stutz (e).

Scientific team at coordinating centers: M Adam (e), I Aguilera (exp), S Brunner (s), D Carballo (c), S Caviezel (pa), I Curjuric (e), A Di Pascale (s), J Dratva (e), R Ducret (s), E Dupuis Lozeron (s), M Eeftens (exp), I Eze (e),

E Fischer (g), M Foraster (e), M Germond (s), L Grize (s), S Hansen (e), A Hensel (s), M Imboden (g), A Ineichen (exp), A Jeong (g), D Keidel (s), A Kumar (g), N Maire (s), A Mehta (e), R Meier (exp), E Schaffner (s), T Schikowski (e), M Tsai (exp).

(a) allergology, (c) cardiology, (cc) clinical chemistry, (e) epidemiology, (exp) exposure, (g) genetic and molecular biology, (m) meteorology, (n) nutrition, (o) occupational health, (p) pneumology, (pa) physical activity, (pd) pediatrics, (s) statistics.

Acknowledgements

The study could not have been done without the help of the study participants, technical and administrative support and the medical teams and field workers at the local study sites.

Local fieldworkers : Aarau: S Brun, G Giger, M Sperisen, M Stahel, Basel: C Bürli, C Dahler, N Oertli, I Harreh, F Karrer, G Novicic, N Wyttenbacher, Davos: A Saner, P Senn, R Winzeler, Geneva: F Bonfils, B Blicharz, C Landolt, J Rochat, Lugano: S Boccia, E Gehrig, MT Mandia, G Solari, B Viscardi, Montana: AP Bieri, C Darioly, M Maire, Payerne: F Ding, P Danieli A Vonnez, Wald: D Bodmer, E Hochstrasser, R Kunz, C Meier, J Rakic, U Schafroth, A Walder.

Administrative staff: N Bauer Ott, C Gabriel, R Gutknecht.

Funding

The Swiss National Science Foundation (grants no 33CS30-148470/1&2, 33CSCO-134276/1, 33CSCO-108796, 324730_135673, 3247BO-104283, 3247BO-104288, 3247BO-104284, 3247-065896, 3100-059302, 3200-052720, 3200-042532, 4026-028099, PMPDP3_129021/1, PMPDP3_141671/1), the Federal Office for the Environment, the Federal Office of Public Health, the Federal Office of Roads and Transport, the canton's government of Aargau, Basel-Stadt, Basel-Land, Geneva, Luzern, Ticino, Valais, and Zürich, the Swiss Lung League, the canton's Lung League of Basel Stadt/ Basel Landschaft, Geneva, Ticino, Valais, Graubünden and Zurich, Stiftung ehemals Bündner Heilstätten, SUVA, Freiwillige Akademische Gesellschaft, UBS Wealth Foundation, Talecris Biotherapeutics GmbH, Abbott Diagnostics, European Commission 018996 (GABRIEL), Wellcome Trust WT 084703MA, Exposomics EC FP7 grant(Grant agreement No: 308610).

SAPPHIRE

Recruitment

The Study of Asthma Phenotypes and Pharmacogenomic Interactions by Race-ethnicity (SAPPHIRE) is an ongoing study that was approved by the Institutional Review Board of Henry Ford Health System. Study individuals included in the current study were members of a large health system, which serves southeast Michigan and all of the Detroit metropolitan statistical area. Individuals were 12-56 years of age and no prior diagnosis of congestive heart failure or chronic obstructive pulmonary disease. Patient recruitment included patients with and without a clinical diagnosis of asthma. Written informed consent was required at the time of enrolment as a condition for study participation. The exam at the time of enrollment included both a staff-administered questionnaire and lung function testing. For the current study, we analyzed only those SAPPHIRE participants who identified themselves as African American.

Acknowledgements and Funding

This work was supported by grants from the Fund for Henry Ford Hospital, the American Asthma Foundation, and the following institutes of the National Institutes of Health: National Institute of Allergy and Infectious Diseases (R01AI079139 and R01AI061774), the National Heart Lung and Blood Institute (R01HL118267 and R01HL079055), and the National Institute of Diabetes and Digestive and Kidney diseases (R01DK064695 and R01DK113003).

UK Biobank

Recruitment

The UK Biobank study is a prospective study of 502,682 volunteer participants aged 40-69 recruited at 22 centers across the UK between 2006 and 2010.[47] A subset of 152,566 individuals with available genotype data were included in the analysis. All particpiants have given informed consent for the storage and use of their data for research.

Acknowledgements and Funding

This research has been conducted using the UK Biobank Resource. L Paternoster was supported by a UK Medical Research Council Fellowship (MR/J012165/1) and works in a unit which receives funding from the UK Medical Research Council (MC_UU_12013/1 & MC_UU_12013/4).

Barcelona Supercomputing Center

Acknowledgements and Funding

Josep M. Mercader was supported by Sara Borrell Fellowship from the Instituto Carlos III and Beatriu de Pinós fellowship from the Agency for Management of University and Research Grants (AGAUR). Sílvia Bonàs was FI-DGR Fellowship from FI-DGR 2013 from Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR, Generalitat de Catalunya). We acknowledge PRACE for awarding us access to MareNostrum supercomputer, based in Spain at Barcelona.

Supplementary References

1. Paternoster, L. *et al.* Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nat. Genet.* **47,** 1449–1456 (2015).
2. Puel, A., Ziegler, S. F., Buckley, R. H. & Leonard, W. J. Defective IL7R expression in T(-)B(+)NK(+) severe combined immunodeficiency. *Nat. Genet.* **20,** 394–397 (1998).
3. Lundmark, F. *et al.* Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. *Nat. Genet.* **39,** 1108–1113 (2007).
4. Gudbjartsson, D. F. *et al.* Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat. Genet.* **41,** 342–347 (2009).
5. Mori, T. *et al.* Lnk/Sh2b3 controls the production and function of dendritic cells and regulates the induction of IFN-γ-producing T cells. *J. Immunol.* **193,** 1728–1736 (2014).
6. Sleiman, P. M. A. *et al.* GWAS identifies four novel eosinophilic esophagitis loci. *Nat. Commun.* **5,** 5593 (2014).
7. Scott, L. M., Civin, C. I., Rorth, P. & Friedman, A. D. A novel temporal expression pattern of three C/EBP family members in differentiating myelomonocytic cells. *Blood* **80,** 1725–1735 (1992).
8. Gao, H., Parkin, S., Johnson, P. F. & Schwartz, R. C. C/EBP gamma has a stimulatory role on the IL-6 and IL-8 promoters. *J. Biol. Chem.* **277,** 38827–38837 (2002).
9. León, B. *et al.* Regulation of T(H)2 development by CXCR5+ dendritic cells and lymphotoxin-expressing B cells. *Nat. Immunol.* **13,** 681–690 (2012).
10. Zhang, J. *et al.* Three SNPs in chromosome 11q23.3 are independently associated with systemic lupus erythematosus in Asians. *Hum. Mol. Genet.* **23,** 524–533 (2014).
11. International Multiple Sclerosis Genetics Consortium *et al.* MANBA, CXCR5, SOX8, RPS6KB1 and ZBTB46 are genetic risk loci for multiple sclerosis. *Brain* **136,** 1778–1782 (2013).
12. Charbonneau, B. *et al.* CXCR5 polymorphisms in non-Hodgkin lymphoma risk and prognosis. *Cancer Immunol. Immunother.* **62,** 1475–1484 (2013).
13. Lawrence, T. The nuclear factor NF-kappaB pathway in inflammation. *Cold Spring Harb. Perspect. Biol.* **1,** a001651 (2009).
14. Shinnakasu, R. *et al.* Regulated selection of germinal-center cells into the memory B cell compartment. *Nat. Immunol.* **17,** 861–869 (2016).
15. Roychoudhuri, R. *et al.* BACH2 regulates CD8(+) T cell differentiation by controlling access of AP-1 factors to enhancers. *Nat. Immunol.* **17,** 851–860 (2016).
16. Rothlin, C. V., Ghosh, S., Zuniga, E. I., Oldstone, M. B. A. & Lemke, G. TAM receptors are pleiotropic inhibitors of the innate immune response. *Cell* **131,** 1124–1136 (2007).
17. Chan, P. Y. *et al.* The TAM family receptor tyrosine kinase TYRO3 is a negative regulator of type 2 immunity. *Science* **352,** 99–103 (2016).
18. Kassmeier, M. D. *et al.* VprBP binds full-length RAG1 and is required for B-cell development and V(D)J recombination fidelity. *EMBO J.* **31,** 945–958 (2012).
19. Hamblet, C. E., Makowski, S. L., Tritapoe, J. M. & Pomerantz, J. L. NK Cell Maturation and Cytotoxicity Are Controlled by the Intramembrane Aspartyl Protease SPPL3. *J. Immunol.* **196,** 2614–2626 (2016).
20. Andersen, J. B., Strandbygård, D. J., Hartmann, R. & Justesen, J. Interaction between the 2’-5' oligoadenylate synthetase-like protein p59 OASL and the transcriptional repressor methyl CpG-binding

protein 1. *Eur. J. Biochem.* **271,** 628–636 (2004).

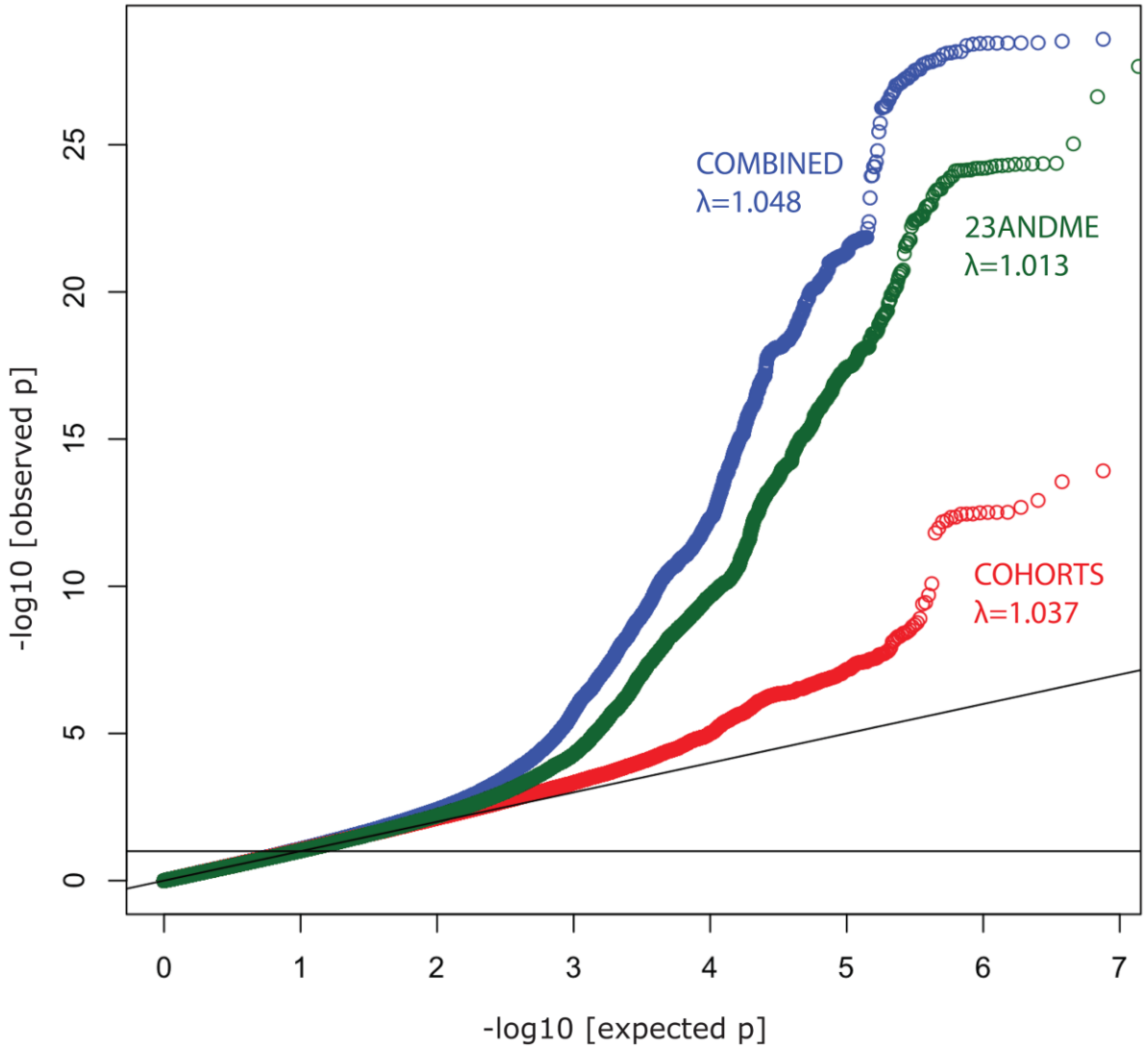
21. Halim, T. Y. F. *et al.* Retinoic-acid-receptor-related orphan nuclear receptor alpha is required for natural helper cell development and allergic inflammation. *Immunity* **37,** 463–474 (2012).
22. Anderson, D. M. *et al.* A homologue of the TNF receptor and its ligand enhance T-cell growth and dendritic-cell function. *Nature* **390,** 175–179 (1997).
23. Bousquet, J. *et al.* Important research questions in allergy and related diseases: nonallergic rhinitis: a GA2LEN paper. *Allergy* **63,** 842–853 (2008).
24. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42,** 111–127 (2013).
25. Bisgaard, H. The Copenhagen Prospective Study on Asthma in Childhood (COPSAC): design, rationale, and baseline data from a longitudinal birth cohort study. *Ann. Allergy Asthma Immunol.* **93,** 381–389 (2004).
26. Gudbjartsson, D. F. *et al.* Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat. Genet.* **41,** 342–347 (2009).
27. Smith, D. *et al.* A rare IL33 loss-of-function mutation reduces blood eosinophil counts and protects from asthma. *PLoS Genet.* **13,** e1006659 (2017).
28. Gudbjartsson, D. F. *et al.* Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47,** 435–444 (2015).
29. Burney, P. G., Luczynska, C., Chinn, S. & Jarvis, D. The European Community Respiratory Health Survey. *Eur. Respir. J.* **7,** 954–960 (1994).
30. Kaprio, J. Twin studies in Finland 2006. *Twin Res. Hum. Genet.* **9,** 772–777 (2006).
31. Kaprio, J. The Finnish Twin Cohort Study: an update. *Twin Res. Hum. Genet.* **16,** 157–162 (2013).
32. Kaprio, J. & Koskenvuo, M. Genetic and environmental factors in complex diseases: the older Finnish Twin Cohort. *Twin Res.* **5,** 358–365 (2002).
33. Ege, M. J. *et al.* Gene-environment interaction for childhood asthma and exposure to farming in Central Europe. *J. Allergy Clin. Immunol.* **127,** 138–44, 144.e1–4 (2011).
34. Pino-Yanes, M. *et al.* Genetic ancestry influences asthma susceptibility and lung function among Latinos. *J. Allergy Clin. Immunol.* **135,** 228–235 (2015).
35. Krawczak, M. *et al.* PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet.* **9,** 55–61 (2006).
36. Wichmann, H.-E., Gieger, C., Illig, T. & MONICA/KORA Study Group. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67 Suppl 1,** S26–30 (2005).
37. Esparza-Gordillo, J. *et al.* A common variant on chromosome 11q13 is associated with atopic dermatitis. *Nat. Genet.* **41,** 596–601 (2009).
38. Völzke, H. *et al.* Cohort profile: the study of health in Pomerania. *Int. J. Epidemiol.* **40,** 294–307 (2011).
39. Heinrich, J. *et al.* Allergens and endotoxin on mothers' mattresses and total immunoglobulin E in cord blood of neonates. *Eur. Respir. J.* **20,** 617–623 (2002).
40. Berg, A. v. *et al.* Impact of early feeding on childhood eczema: development after nutritional intervention compared with the natural course - the GINIplus study up to the age of 6 years. *Clin. Exp. Allergy* **40,** 627–636 (2010).
41. Barreto-Luis, A. *et al.* Genome-wide association study in Spanish identifies ADAM metallopeptidase

with thrombospondin type 1 motif, 9 (ADAMTS9), as a novel asthma susceptibility gene. *J. Allergy Clin. Immunol.* **137,** 964–966 (2016).

42. Barreto-Luis, A. *et al.* A pathway-based association study reveals variants from Wnt signalling genes contributing to asthma susceptibility. *Clin. Exp. Allergy* **47,** 618–626 (2017).
43. Paternoster, L. *et al.* Genome-wide population-based association study of extremely overweight young adults--the GOYA study. *PLoS One* **6,** e24303 (2011).
44. Guxens, M. *et al.* Cohort Profile: the INMA--INfancia y Medio Ambiente--(Environment and Childhood) Project. *Int. J. Epidemiol.* **41,** 930–940 (2012).
45. van Beijsterveldt, C. E. M. & Boomsma, D. I. Genetics of parentally reported asthma, eczema and rhinitis in 5-yr-old twins. *Eur. Respir. J.* **29,** 516–521 (2007).
46. White, M. J. *et al.* Novel genetic risk factors for asthma in African American children: Precision Medicine and the SAGE II Study. *Immunogenetics* **68,** 391–400 (2016).
47. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12,** e1001779 (2015).

Supplementary Figure 1

QQ-plot for the discovery phase of the allergic rhinitis GWAS, based on p-values from an inverse variance weighted fixed-effect meta-analysis of genetic marker association of 16,531,985 genetic markers to allergic rhinitis from the discovery phase, including 212,120 individuals). λ indicates the genomic inflation rate.



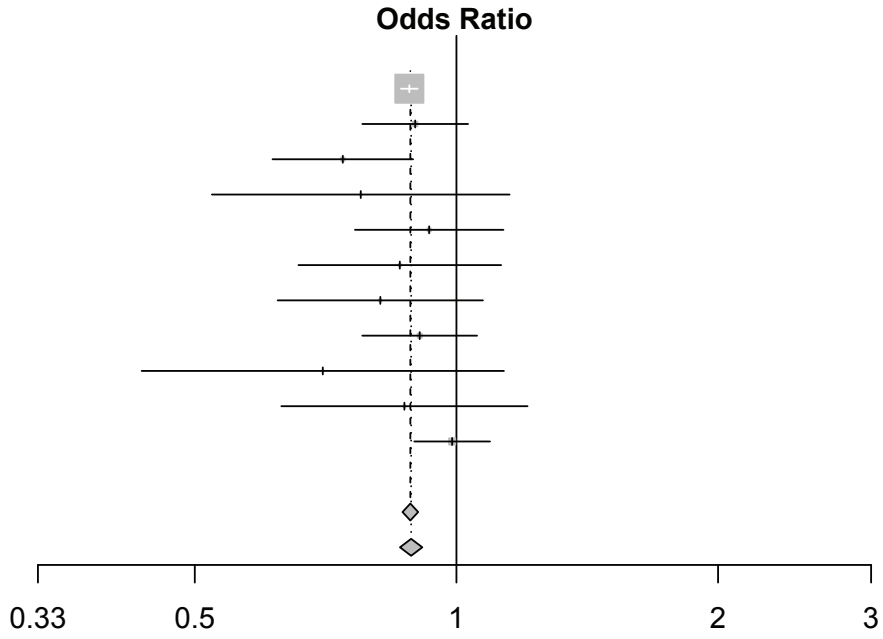
Supplementary Figure 2

Forest plots from the discovery phase for markers previously associated to allergy, including weights, estimates and 95% confidence intervals for inverse variance weighted fixed and random effects meta-analyses from genetic marker association of 16,531,985 genetic markers to allergic rhinitis from the discovery phase, including 212,120 individuals.

rs34004019[G]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	87.3%	75.0%	0.88	[0.86; 0.90]
AAGC(QIMR)	2.2%	4.3%	0.90	[0.78; 1.03]
ALSPAC	1.2%	2.5%	0.74	[0.61; 0.89]
EVE (CAG/CSGA)	0.3%	0.6%	0.78	[0.52; 1.15]
EVE (CHS)	1.1%	2.2%	0.93	[0.76; 1.13]
GENR	0.6%	1.2%	0.86	[0.66; 1.13]
GINI/LISA	0.6%	1.2%	0.82	[0.62; 1.07]
GOYA	1.9%	3.7%	0.91	[0.78; 1.06]
INMA	0.2%	0.4%	0.70	[0.43; 1.13]
MAAS	0.4%	0.8%	0.87	[0.63; 1.21]
NFBC1966	4.3%	8.2%	0.99	[0.89; 1.09]
Fixed effect model	**100.0%**	**−−**	**0.89**	**[0.87; 0.90]**
Random effects model	**−−**	**100.0%**	**0.89**	**[0.86; 0.91]**

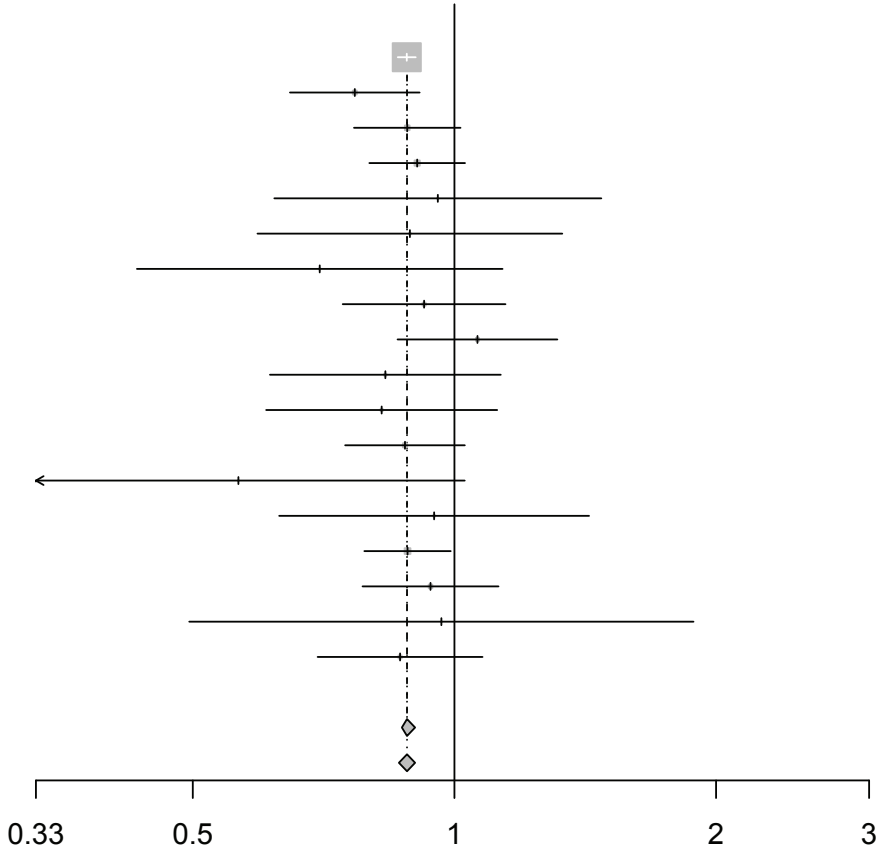
Heterogeneity: I^2 = 4%, p = 0.41



rs950881[T]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	81.4%	81.4%	0.88	[0.86; 0.90]
AAGC(QIMR)	1.6%	1.6%	0.77	[0.65; 0.91]
ALSPAC	2.3%	2.3%	0.88	[0.77; 1.02]
B58C	2.9%	2.9%	0.91	[0.80; 1.03]
BAMSE	0.2%	0.2%	0.96	[0.62; 1.48]
COPSAC2000	0.3%	0.3%	0.89	[0.59; 1.33]
EVE (CAG/CSGA)	0.2%	0.2%	0.70	[0.43; 1.14]
EVE (CHS)	1.0%	1.0%	0.92	[0.74; 1.14]
GENR	1.0%	1.0%	1.06	[0.86; 1.31]
GENUFAD	0.5%	0.5%	0.83	[0.61; 1.13]
GINI/LISA	0.5%	0.5%	0.82	[0.61; 1.12]
GOYA	1.8%	1.8%	0.88	[0.75; 1.03]
INMA	0.1%	0.1%	0.56	[0.31; 1.03]
MAAS	0.3%	0.3%	0.95	[0.63; 1.43]
NFBC1966	3.5%	3.5%	0.88	[0.79; 0.99]
NTR (adults)	1.4%	1.4%	0.94	[0.78; 1.12]
NTR (children)	0.1%	0.1%	0.97	[0.50; 1.88]
RAINE	1.0%	1.0%	0.87	[0.70; 1.08]
Fixed effect model	**100.0%**	**−−**	**0.88**	**[0.87; 0.90]**
Random effects model	**−−**	**100.0%**	**0.88**	**[0.86; 0.90]**

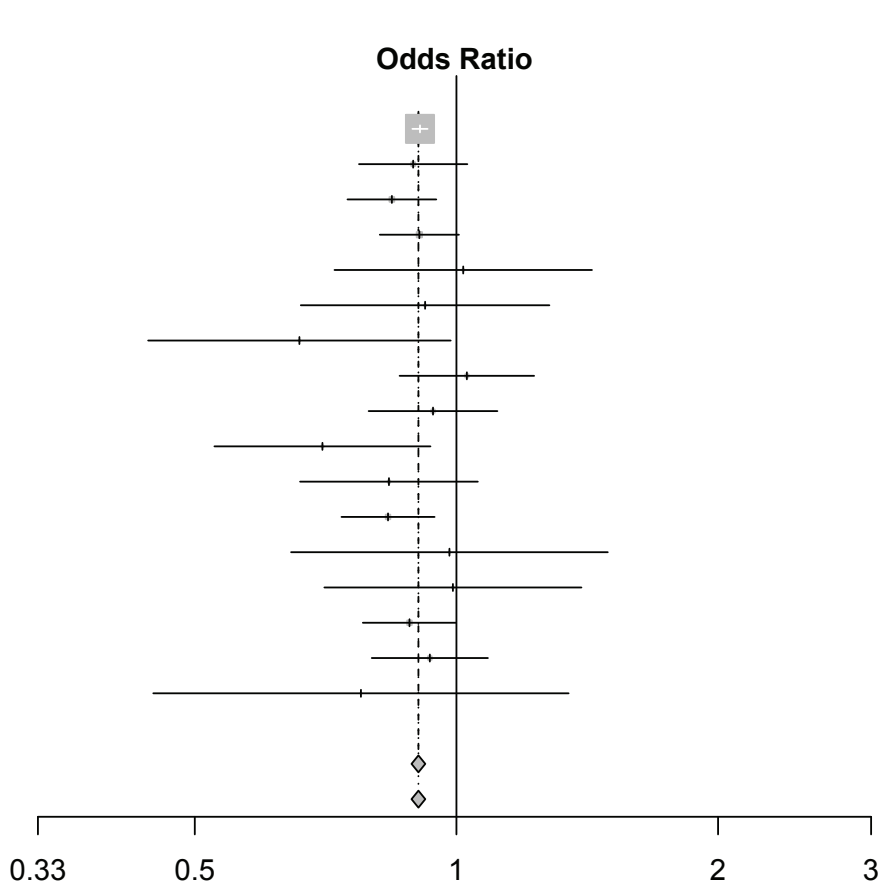
Heterogeneity: I^2 = 0%, p = 0.90



rs5743618[A]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	82.5%	82.5%	0.91	[0.89; 0.93]
AAGC(QIMR)	1.6%	1.6%	0.89	[0.77; 1.03]
ALSPAC	2.4%	2.4%	0.84	[0.75; 0.95]
B58C	3.0%	3.0%	0.91	[0.82; 1.01]
BAMSE	0.3%	0.3%	1.02	[0.72; 1.43]
COPSAC2000	0.3%	0.3%	0.92	[0.66; 1.28]
EVE (CAG/CSGA)	0.2%	0.2%	0.66	[0.44; 0.98]
EVE (CHS)	1.1%	1.1%	1.03	[0.86; 1.23]
GENR	1.2%	1.2%	0.94	[0.79; 1.11]
GENUFAD	0.4%	0.4%	0.70	[0.53; 0.93]
GINI/LISA	0.6%	0.6%	0.84	[0.66; 1.06]
GOYA	2.2%	2.2%	0.83	[0.74; 0.94]
INMA	0.2%	0.2%	0.98	[0.65; 1.49]
MAAS	0.3%	0.3%	0.99	[0.70; 1.39]
NFBC1966	2.2%	2.2%	0.88	[0.78; 1.00]
NTR (adults)	1.4%	1.4%	0.93	[0.80; 1.09]
NTR (children)	0.1%	0.1%	0.78	[0.45; 1.35]
Fixed effect model	**100.0%**	**−−**	**0.90**	**[0.89; 0.92]**
Random effects model	**−−**	**100.0%**	**0.90**	**[0.89; 0.92]**

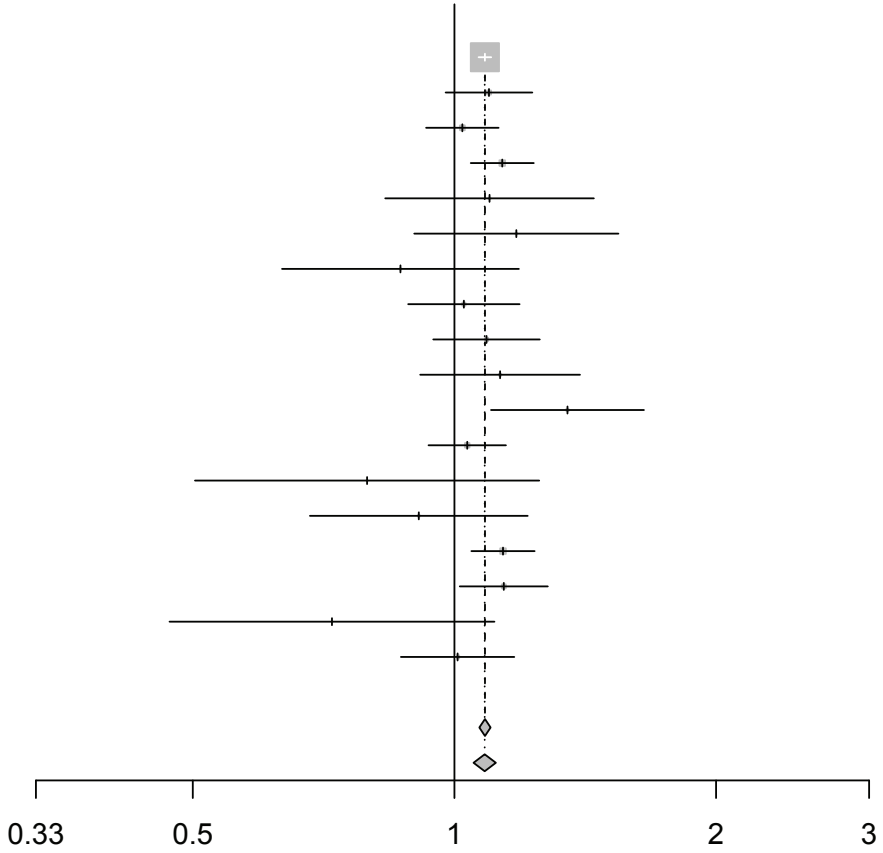
Heterogeneity: I^2 = 0%, p = 0.69



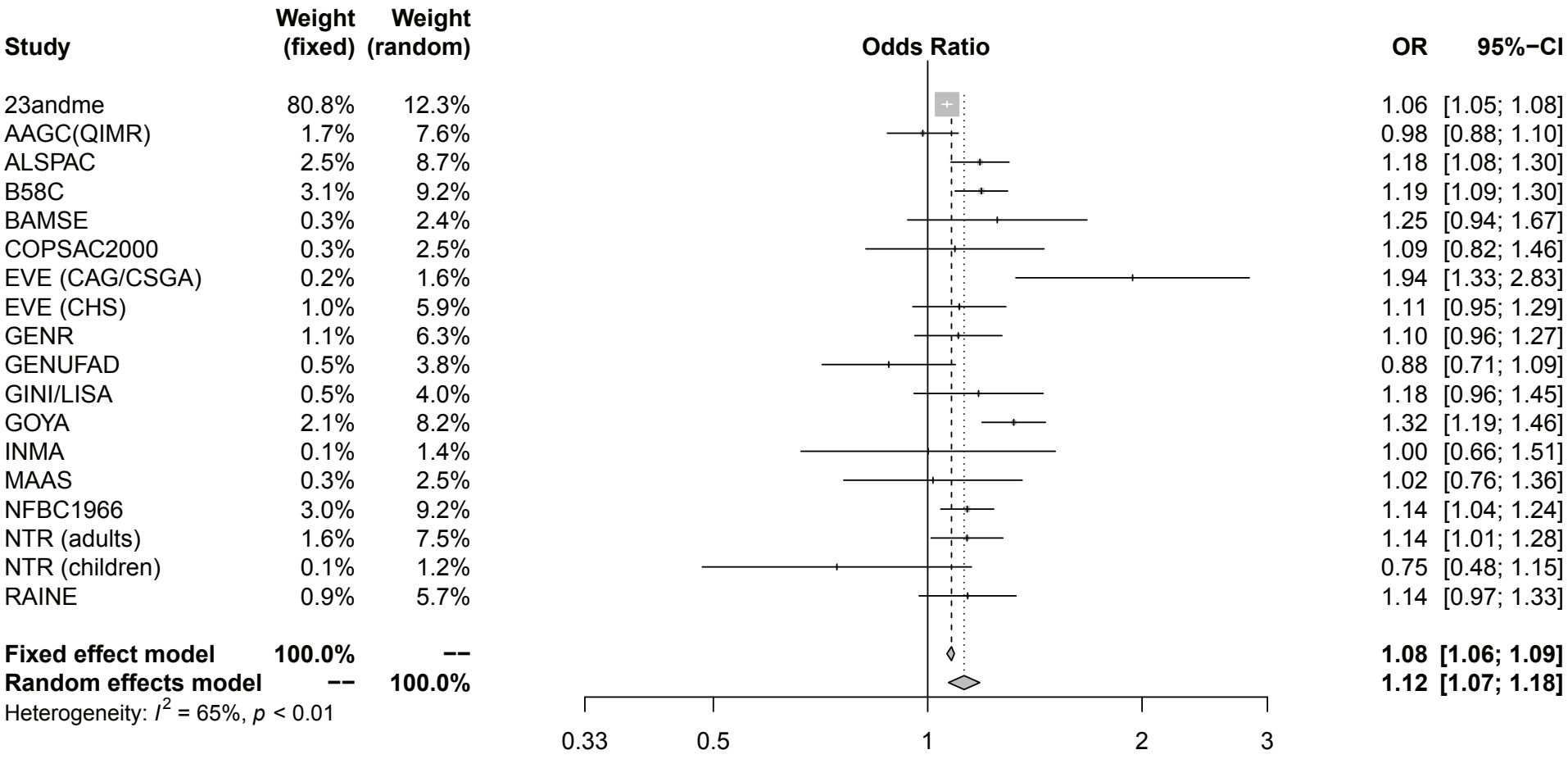
rs1438673[C]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	80.2%	35.3%	1.08	[1.07; 1.10]
AAGC(QIMR)	1.7%	5.7%	1.10	[0.98; 1.23]
ALSPAC	2.4%	7.6%	1.02	[0.93; 1.12]
B58C	3.2%	9.5%	1.14	[1.04; 1.23]
BAMSE	0.3%	1.1%	1.10	[0.83; 1.45]
COPSAC2000	0.3%	1.2%	1.18	[0.90; 1.54]
EVE (CAG/CSGA)	0.2%	0.9%	0.87	[0.63; 1.19]
EVE (CHS)	1.0%	3.6%	1.03	[0.88; 1.19]
GENR	1.1%	3.9%	1.09	[0.95; 1.25]
GENUFAD	0.5%	1.9%	1.13	[0.91; 1.39]
GINI/LISA	0.5%	2.0%	1.35	[1.10; 1.65]
GOYA	2.1%	6.9%	1.03	[0.93; 1.15]
INMA	0.1%	0.4%	0.79	[0.50; 1.25]
MAAS	0.3%	1.0%	0.91	[0.68; 1.21]
NFBC1966	3.2%	9.5%	1.14	[1.05; 1.24]
NTR (adults)	1.6%	5.5%	1.14	[1.01; 1.28]
NTR (children)	0.1%	0.5%	0.72	[0.47; 1.11]
RAINE	1.0%	3.5%	1.01	[0.87; 1.17]
Fixed effect model	**100.0%**	**−−**	**1.08**	**[1.07; 1.10]**
Random effects model	**−−**	**100.0%**	**1.08**	**[1.05; 1.12]**

Heterogeneity: I^2 = 17%, p = 0.25

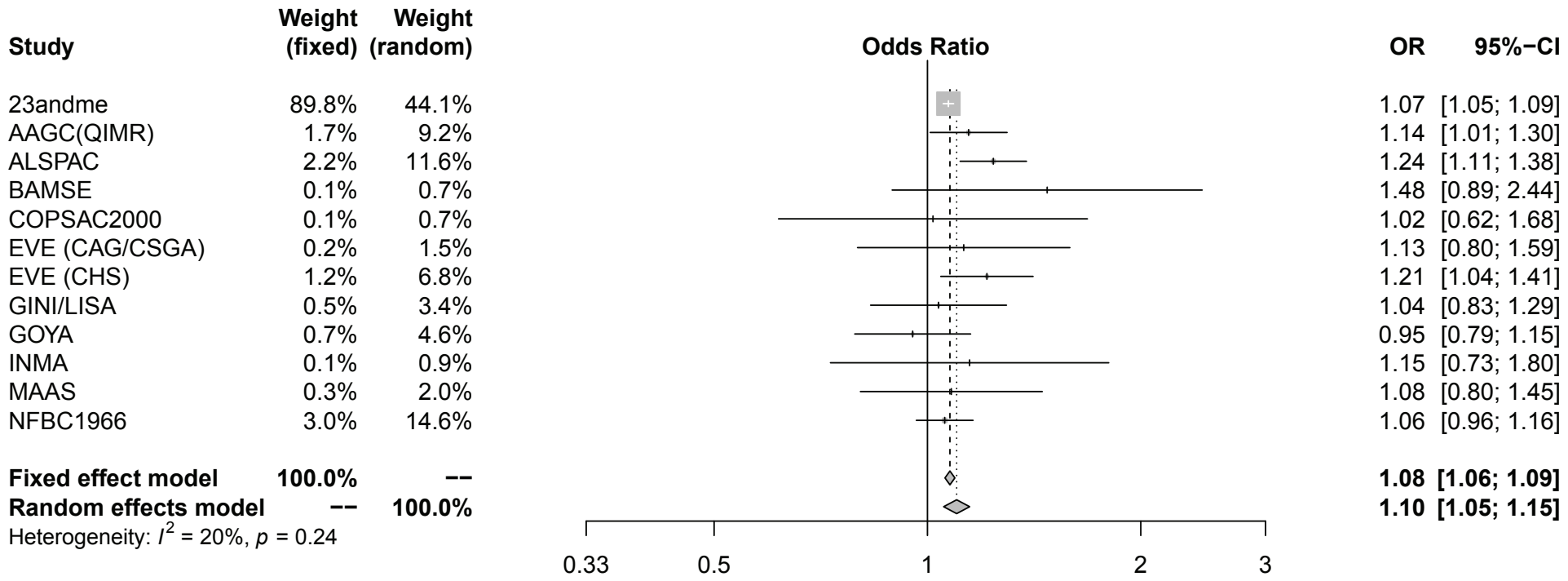


rs7936323[A]

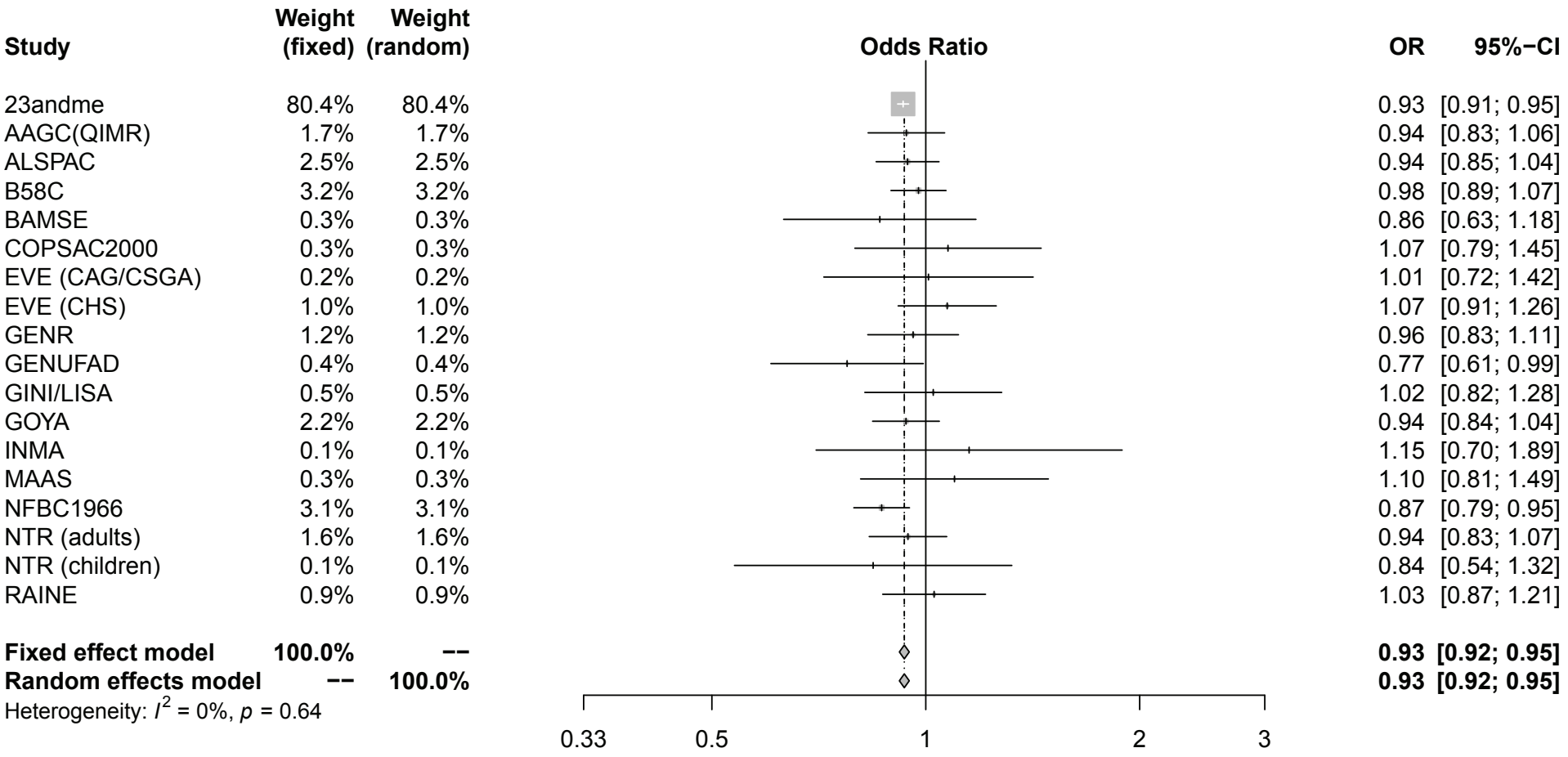


Heterogeneity: I^2 = 65%, p < 0.01

rs2428494[A]


Study
Weight (fixed)
Weight (random)
Odds Ratio
OR
95%−CI
23andme 89.8% 44.1% 1.07 [1.05; 1.09]
AAGC(QIMR) 1.7% 9.2% 1.14 [1.01; 1.30]
ALSPAC 2.2% 11.6% 1.24 [1.11; 1.38]
BAMSE 0.1% 0.7% 1.48 [0.89; 2.44]
COPSAC2000 0.1% 0.7% 1.02 [0.62; 1.68]
EVE (CAG/CSGA) 0.2% 1.5% 1.13 [0.80; 1.59]
EVE (CHS) 1.2% 6.8% 1.21 [1.04; 1.41]
GINI/LISA 0.5% 3.4% 1.04 [0.83; 1.29]
GOYA 0.7% 4.6% 0.95 [0.79; 1.15]
INMA 0.1% 0.9% 1.15 [0.73; 1.80]
MAAS 0.3% 2.0% 1.08 [0.80; 1.45]
NFBC1966 3.0% 14.6% 1.06 [0.96; 1.16]
Fixed effect model 100.0% −− 1.08 [1.06; 1.09]
Random effects model −− 100.0% 1.10 [1.05; 1.15]
Heterogeneity: I^2 = 20%, p = 0.24
0.33
0.5
1
2
3

rs11644510[T]

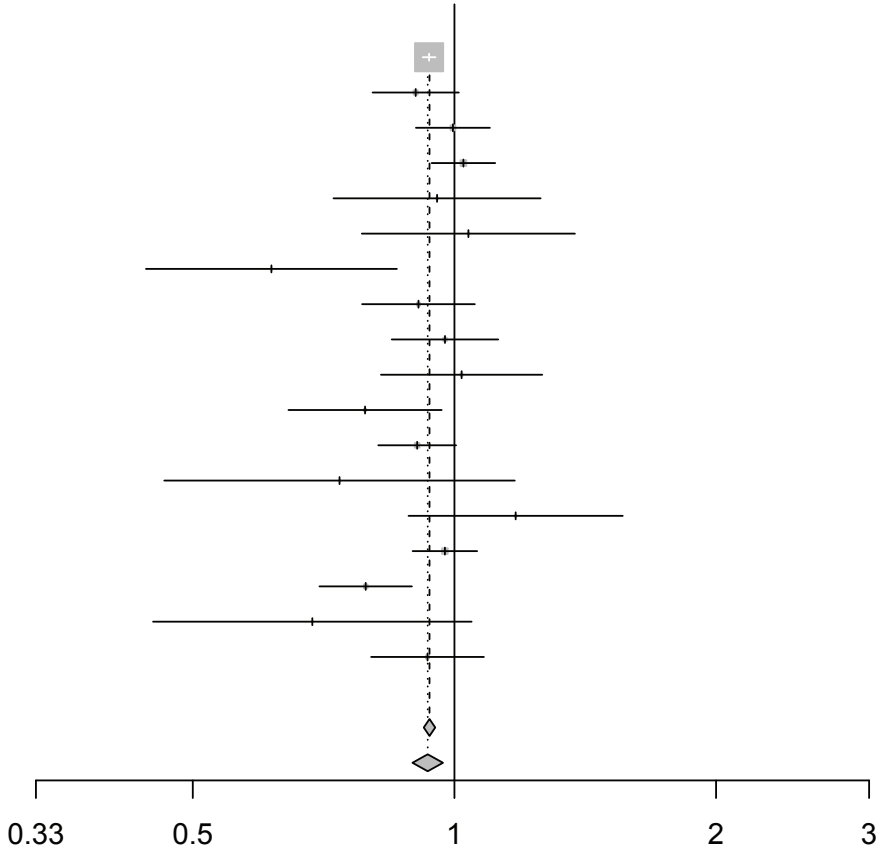


Heterogeneity: I^2 = 0%, p = 0.64

rs12939457[C]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	79.9%	17.5%	0.94	[0.92; 0.95]
AAGC(QIMR)	1.8%	7.4%	0.90	[0.81; 1.01]
ALSPAC	2.4%	8.8%	1.00	[0.90; 1.10]
B58C	3.3%	10.2%	1.02	[0.94; 1.11]
BAMSE	0.3%	1.9%	0.96	[0.73; 1.26]
COPSAC2000	0.3%	1.8%	1.04	[0.78; 1.38]
EVE (CAG/CSGA)	0.2%	1.4%	0.62	[0.44; 0.86]
EVE (CHS)	1.0%	5.2%	0.91	[0.78; 1.06]
GENR	1.2%	5.7%	0.98	[0.85; 1.12]
GENUFAD	0.5%	3.0%	1.02	[0.82; 1.26]
GINI/LISA	0.6%	3.3%	0.79	[0.64; 0.97]
GOYA	2.2%	8.3%	0.91	[0.82; 1.00]
INMA	0.1%	0.7%	0.74	[0.46; 1.17]
MAAS	0.3%	1.8%	1.18	[0.89; 1.56]
NFBC1966	3.2%	10.0%	0.98	[0.90; 1.06]
NTR (adults)	1.5%	6.8%	0.79	[0.70; 0.89]
NTR (children)	0.1%	0.9%	0.69	[0.45; 1.05]
RAINE	1.0%	5.2%	0.93	[0.80; 1.08]
Fixed effect model	**100.0%**	**−−**	**0.94**	**[0.92; 0.95]**
Random effects model	**−−**	**100.0%**	**0.93**	**[0.89; 0.97]**

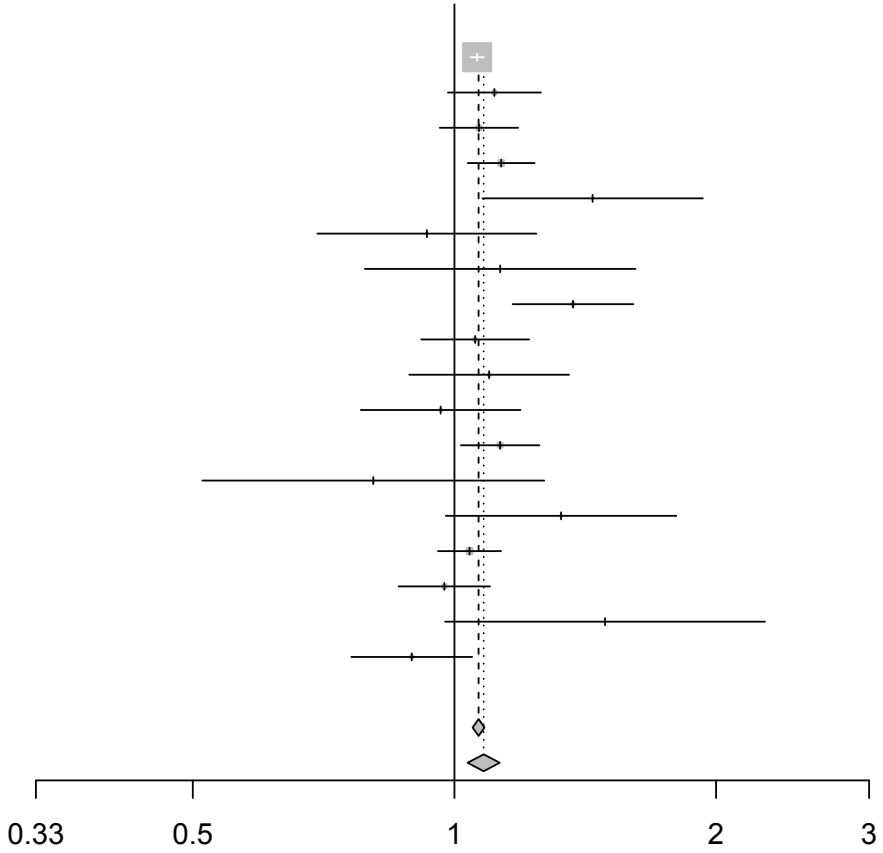
Heterogeneity: I^2 = 45%, p = 0.02



rs148505069[G]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	79.7%	17.0%	1.06	[1.04; 1.08]
AAGC(QIMR)	1.6%	7.1%	1.11	[0.98; 1.26]
ALSPAC	2.3%	8.6%	1.07	[0.96; 1.18]
B58C	3.2%	9.9%	1.13	[1.04; 1.24]
BAMSE	0.3%	1.9%	1.44	[1.08; 1.93]
COPSAC2000	0.3%	1.9%	0.93	[0.70; 1.24]
EVE (CAG/CSGA)	0.2%	1.3%	1.13	[0.79; 1.62]
EVE (CHS)	1.0%	5.0%	1.37	[1.17; 1.61]
GENR	1.2%	5.9%	1.06	[0.92; 1.22]
GENUFAD	0.6%	3.3%	1.10	[0.89; 1.35]
GINI/LISA	0.6%	3.3%	0.96	[0.78; 1.19]
GOYA	2.3%	8.5%	1.13	[1.02; 1.25]
INMA	0.1%	0.8%	0.81	[0.51; 1.27]
MAAS	0.3%	1.7%	1.33	[0.98; 1.80]
NFBC1966	3.6%	10.4%	1.04	[0.96; 1.13]
NTR (adults)	1.7%	7.2%	0.97	[0.86; 1.10]
NTR (children)	0.1%	1.0%	1.49	[0.98; 2.28]
RAINE	1.0%	5.0%	0.89	[0.76; 1.05]
Fixed effect model	**100.0%**	**--**	**1.07**	**[1.05; 1.08]**
Random effects model	**--**	**100.0%**	**1.08**	**[1.04; 1.13]**

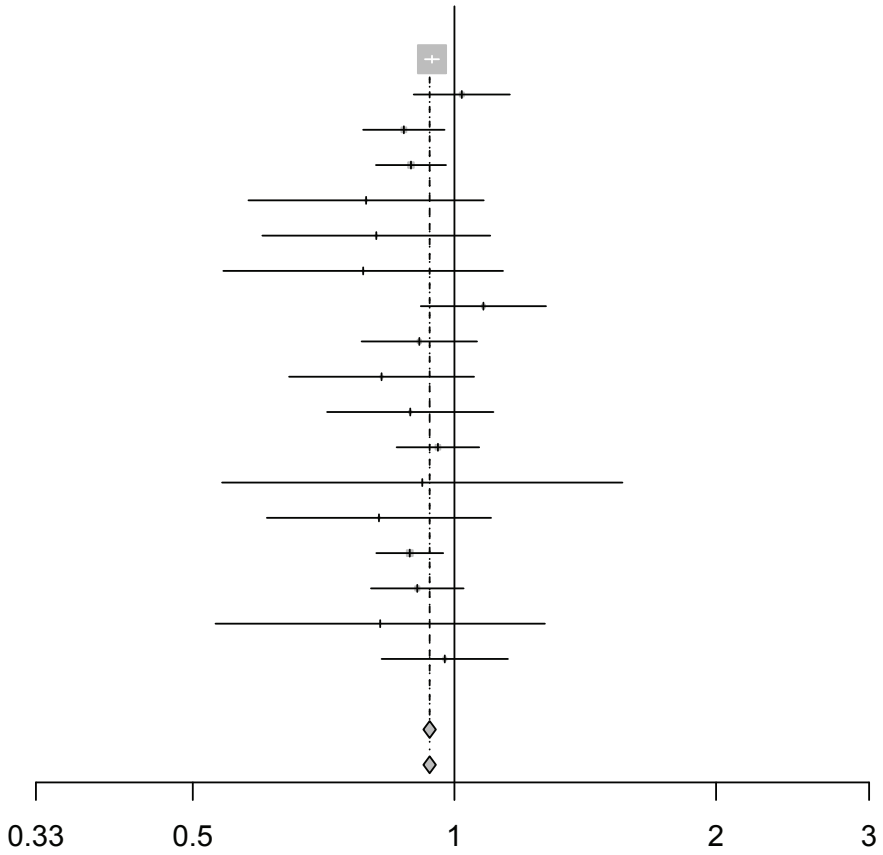
Heterogeneity: I^2 = 47%, p = 0.02



rs13395467[G]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	79.6%	79.6%	0.94	[0.92; 0.96]
AAGC(QIMR)	1.7%	1.7%	1.02	[0.90; 1.16]
ALSPAC	2.4%	2.4%	0.87	[0.79; 0.97]
B58C	3.2%	3.2%	0.89	[0.81; 0.98]
BAMSE	0.3%	0.3%	0.79	[0.58; 1.08]
COPSAC2000	0.3%	0.3%	0.81	[0.60; 1.10]
EVE (CAG/CSGA)	0.2%	0.2%	0.79	[0.54; 1.14]
EVE (CHS)	1.0%	1.0%	1.08	[0.91; 1.27]
GENR	1.2%	1.2%	0.91	[0.78; 1.06]
GENUFAD	0.5%	0.5%	0.82	[0.65; 1.05]
GINI/LISA	0.6%	0.6%	0.89	[0.71; 1.11]
GOYA	2.3%	2.3%	0.96	[0.86; 1.07]
INMA	0.1%	0.1%	0.92	[0.54; 1.56]
MAAS	0.3%	0.3%	0.82	[0.61; 1.10]
NFBC1966	3.5%	3.5%	0.89	[0.81; 0.97]
NTR (adults)	1.8%	1.8%	0.91	[0.80; 1.02]
NTR (children)	0.1%	0.1%	0.82	[0.53; 1.27]
RAINE	1.0%	1.0%	0.97	[0.82; 1.15]
Fixed effect model	**100.0%**	**−−**	**0.94**	**[0.92; 0.95]**
Random effects model	**−−**	**100.0%**	**0.94**	**[0.92; 0.95]**

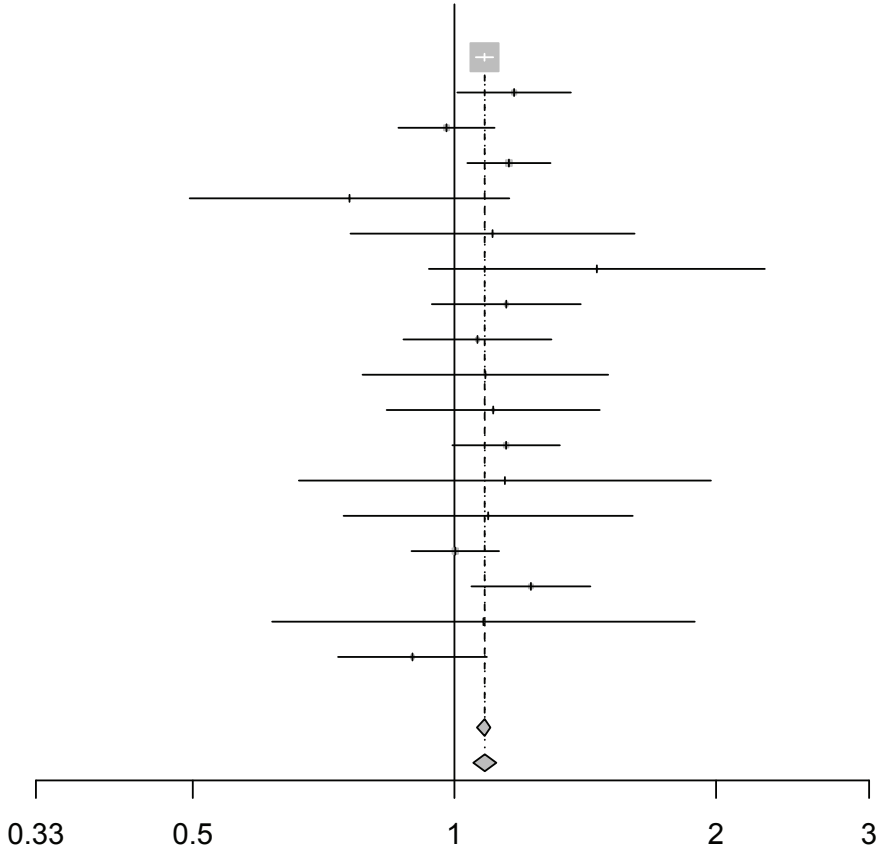
Heterogeneity: I^2 = 0%, p = 0.60



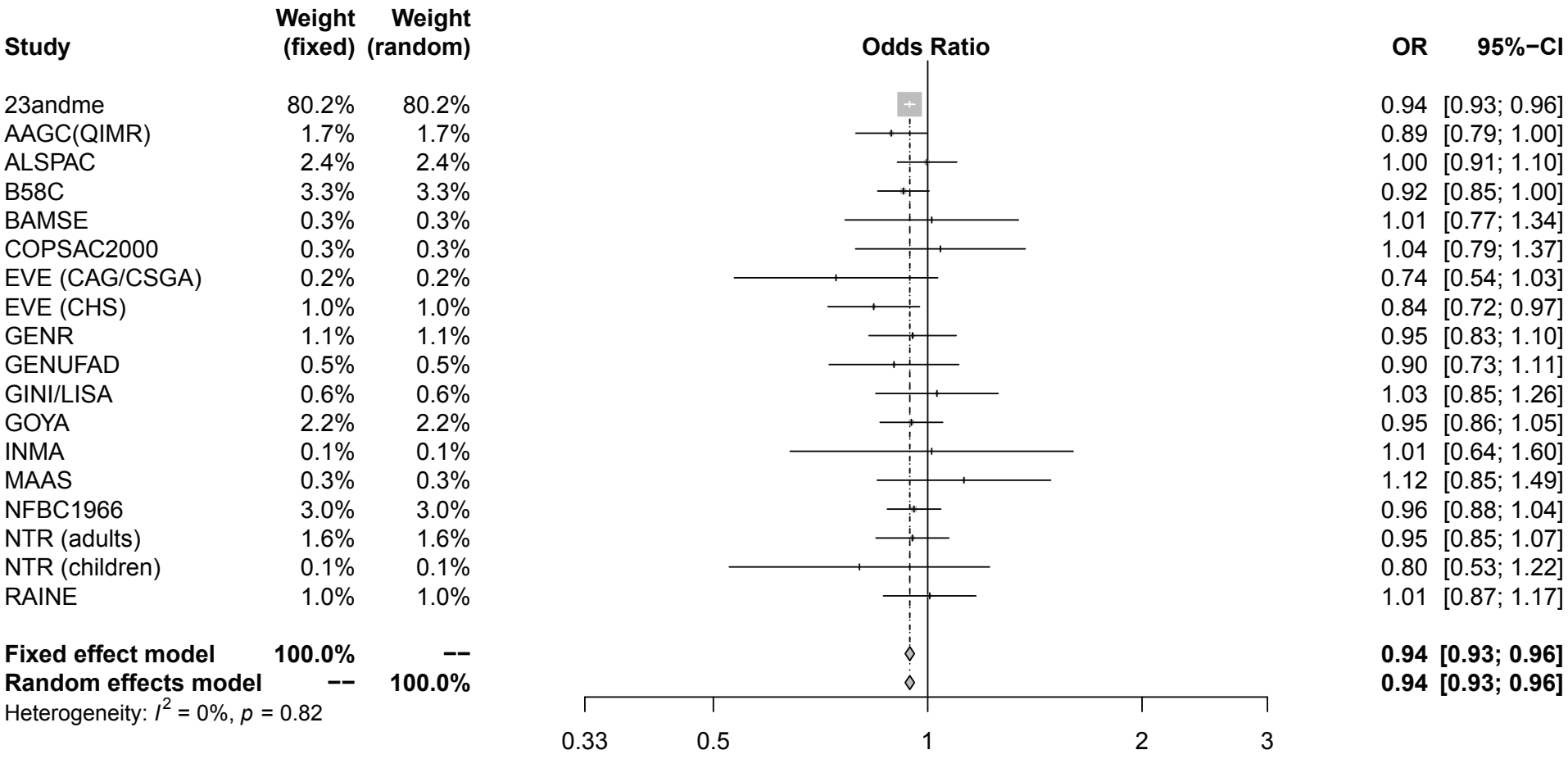
rs9775039[A]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	79.7%	57.3%	1.08	[1.06; 1.11]
AAGC(QIMR)	1.9%	3.9%	1.17	[1.01; 1.36]
ALSPAC	2.6%	5.4%	0.98	[0.86; 1.11]
B58C	3.5%	7.1%	1.16	[1.03; 1.29]
BAMSE	0.2%	0.5%	0.76	[0.50; 1.16]
COPSAC2000	0.3%	0.7%	1.11	[0.76; 1.61]
EVE (CAG/CSGA)	0.2%	0.5%	1.46	[0.93; 2.28]
EVE (CHS)	1.1%	2.3%	1.15	[0.94; 1.40]
GENR	1.1%	2.4%	1.06	[0.87; 1.29]
GENUFAD	0.4%	0.9%	1.09	[0.78; 1.50]
GINI/LISA	0.5%	1.2%	1.11	[0.84; 1.47]
GOYA	2.1%	4.4%	1.15	[0.99; 1.32]
INMA	0.1%	0.3%	1.14	[0.66; 1.97]
MAAS	0.3%	0.6%	1.09	[0.75; 1.60]
NFBC1966	3.1%	6.4%	1.00	[0.89; 1.13]
NTR (adults)	1.7%	3.6%	1.22	[1.05; 1.43]
NTR (children)	0.1%	0.3%	1.08	[0.62; 1.89]
RAINE	1.1%	2.3%	0.89	[0.73; 1.09]
Fixed effect model	**100.0%**	**−−**	**1.08**	**[1.06; 1.10]**
Random effects model	**−−**	**100.0%**	**1.08**	**[1.05; 1.12]**

Heterogeneity: I^2 = 5%, p = 0.39



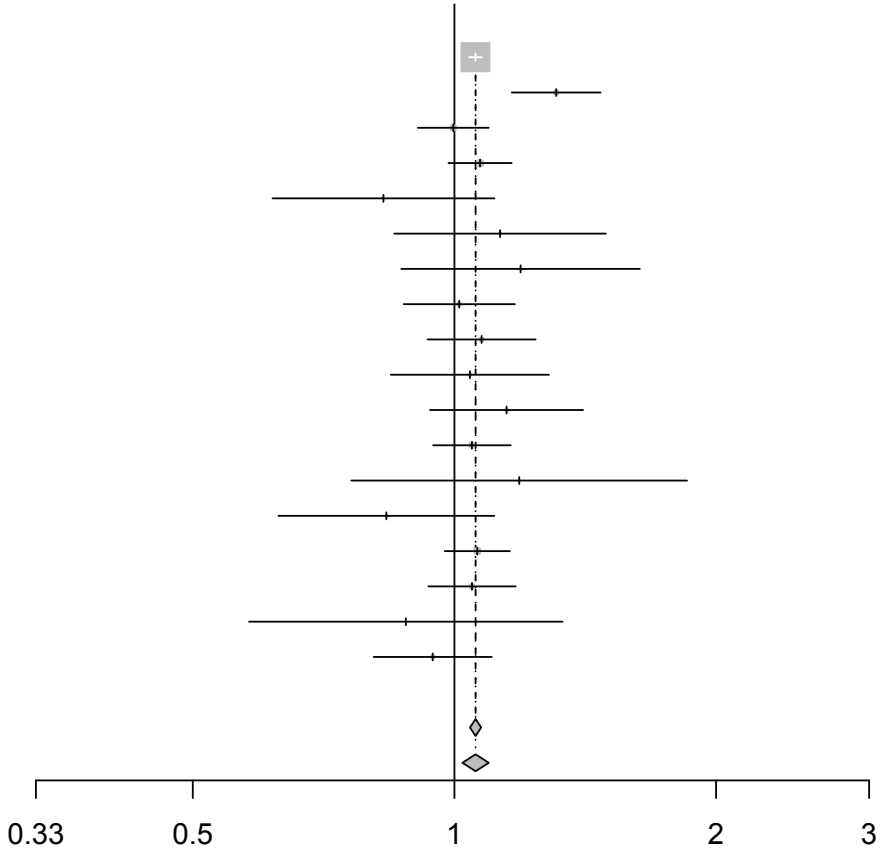
rs2164068[A]


Study
Weight (fixed)
Weight (random)
Odds Ratio
OR
95%−CI
23andme 80.2% 80.2% 0.94 [0.93; 0.96]
AAGC(QIMR) 1.7% 1.7% 0.89 [0.79; 1.00]
ALSPAC 2.4% 2.4% 1.00 [0.91; 1.10]
B58C 3.3% 3.3% 0.92 [0.85; 1.00]
BAMSE 0.3% 0.3% 1.01 [0.77; 1.34]
COPSAC2000 0.3% 0.3% 1.04 [0.79; 1.37]
EVE (CAG/CSGA) 0.2% 0.2% 0.74 [0.54; 1.03]
EVE (CHS) 1.0% 1.0% 0.84 [0.72; 0.97]
GENR 1.1% 1.1% 0.95 [0.83; 1.10]
GENUFAD 0.5% 0.5% 0.90 [0.73; 1.11]
GINI/LISA 0.6% 0.6% 1.03 [0.85; 1.26]
GOYA 2.2% 2.2% 0.95 [0.86; 1.05]
INMA 0.1% 0.1% 1.01 [0.64; 1.60]
MAAS 0.3% 0.3% 1.12 [0.85; 1.49]
NFBC1966 3.0% 3.0% 0.96 [0.88; 1.04]
NTR (adults) 1.6% 1.6% 0.95 [0.85; 1.07]
NTR (children) 0.1% 0.1% 0.80 [0.53; 1.22]
RAINE 1.0% 1.0% 1.01 [0.87; 1.17]
Fixed effect model 100.0% −− 0.94 [0.93; 0.96]
Random effects model −− 100.0% 0.94 [0.93; 0.96]
Heterogeneity: I^2 = 0%, p = 0.82
0.33
0.5
1
2
3

rs2030519[G]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	80.6%	24.1%	1.06	[1.04; 1.08]
AAGC(QIMR)	1.6%	6.6%	1.31	[1.16; 1.47]
ALSPAC	2.5%	8.9%	1.00	[0.91; 1.10]
B58C	3.2%	10.3%	1.07	[0.98; 1.16]
BAMSE	0.3%	1.3%	0.83	[0.62; 1.11]
COPSAC2000	0.3%	1.5%	1.13	[0.85; 1.49]
EVE (CAG/CSGA)	0.2%	1.2%	1.19	[0.87; 1.63]
EVE (CHS)	1.0%	4.6%	1.01	[0.87; 1.17]
GENR	1.1%	4.8%	1.07	[0.93; 1.24]
GENUFAD	0.5%	2.5%	1.04	[0.84; 1.28]
GINI/LISA	0.5%	2.7%	1.15	[0.94; 1.41]
GOYA	2.1%	7.9%	1.05	[0.95; 1.16]
INMA	0.1%	0.6%	1.19	[0.76; 1.85]
MAAS	0.3%	1.4%	0.84	[0.63; 1.11]
NFBC1966	3.0%	10.0%	1.06	[0.97; 1.16]
NTR (adults)	1.7%	6.7%	1.05	[0.93; 1.18]
NTR (children)	0.1%	0.7%	0.88	[0.58; 1.33]
RAINE	0.9%	4.2%	0.94	[0.81; 1.10]
Fixed effect model	**100.0%**	**−−**	**1.06**	**[1.04; 1.07]**
Random effects model	**−−**	**100.0%**	**1.06**	**[1.02; 1.10]**

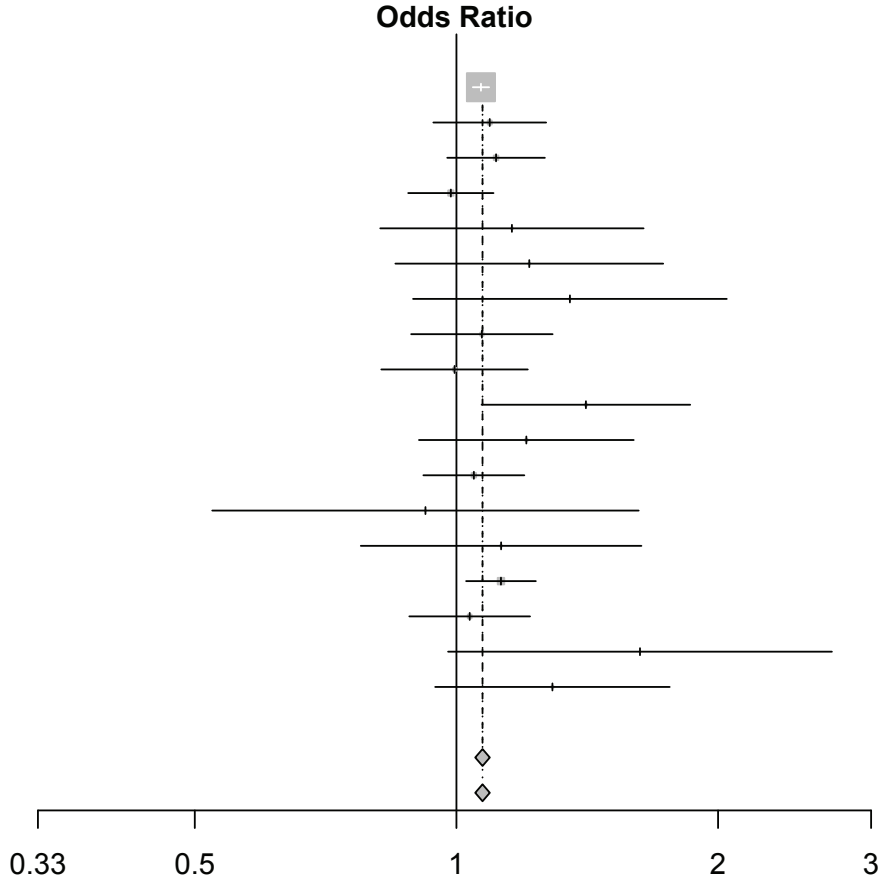
Heterogeneity: I^2 = 30%, p = 0.11



rs11256017[T]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	80.5%	80.5%	1.07	[1.04; 1.09]
AAGC(QIMR)	1.7%	1.7%	1.09	[0.94; 1.27]
ALSPAC	2.2%	2.2%	1.11	[0.98; 1.26]
B58C	2.9%	2.9%	0.99	[0.88; 1.10]
BAMSE	0.3%	0.3%	1.16	[0.82; 1.64]
COPSAC2000	0.3%	0.3%	1.21	[0.85; 1.73]
EVE (CAG/CSGA)	0.2%	0.2%	1.35	[0.89; 2.05]
EVE (CHS)	1.1%	1.1%	1.07	[0.89; 1.29]
GENR	1.0%	1.0%	1.00	[0.82; 1.21]
GENUFAD	0.5%	0.5%	1.41	[1.07; 1.86]
GINI/LISA	0.5%	0.5%	1.20	[0.91; 1.60]
GOYA	2.1%	2.1%	1.05	[0.92; 1.20]
INMA	0.1%	0.1%	0.92	[0.52; 1.62]
MAAS	0.3%	0.3%	1.13	[0.78; 1.63]
NFBC1966	4.4%	4.4%	1.13	[1.03; 1.23]
NTR (adults)	1.5%	1.5%	1.04	[0.88; 1.22]
NTR (children)	0.1%	0.1%	1.63	[0.98; 2.70]
RAINE	0.4%	0.4%	1.29	[0.95; 1.76]
Fixed effect model	**100.0%**	**−−**	**1.07**	**[1.05; 1.09]**
Random effects model	**−−**	**100.0%**	**1.07**	**[1.05; 1.09]**

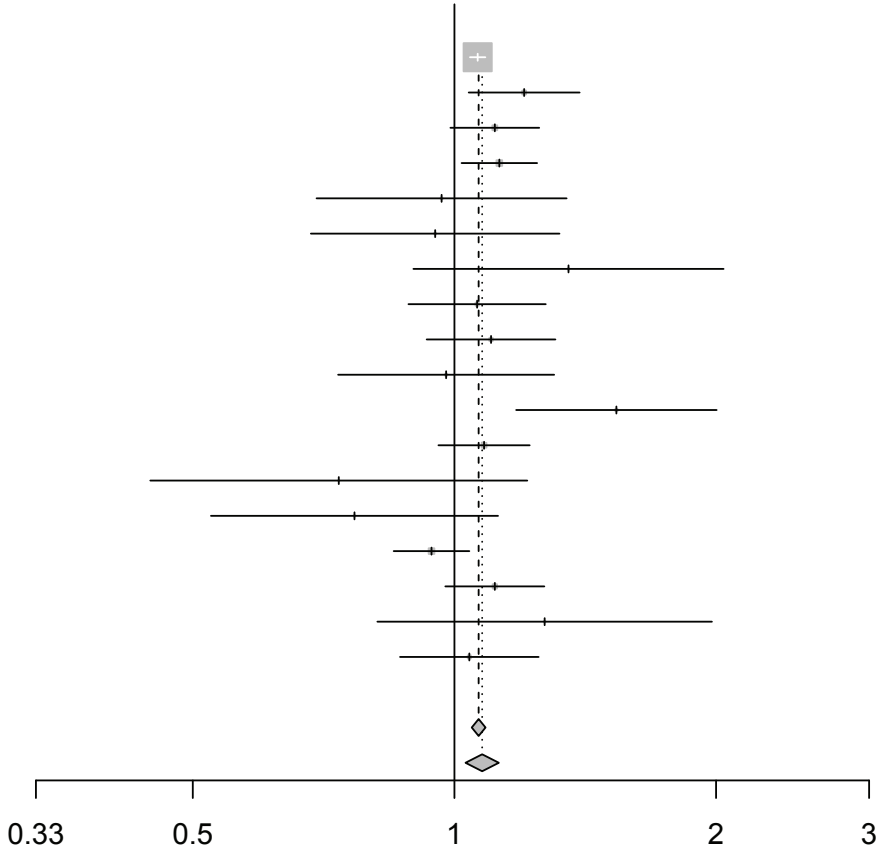
Heterogeneity: I^2 = 0%, p = 0.59



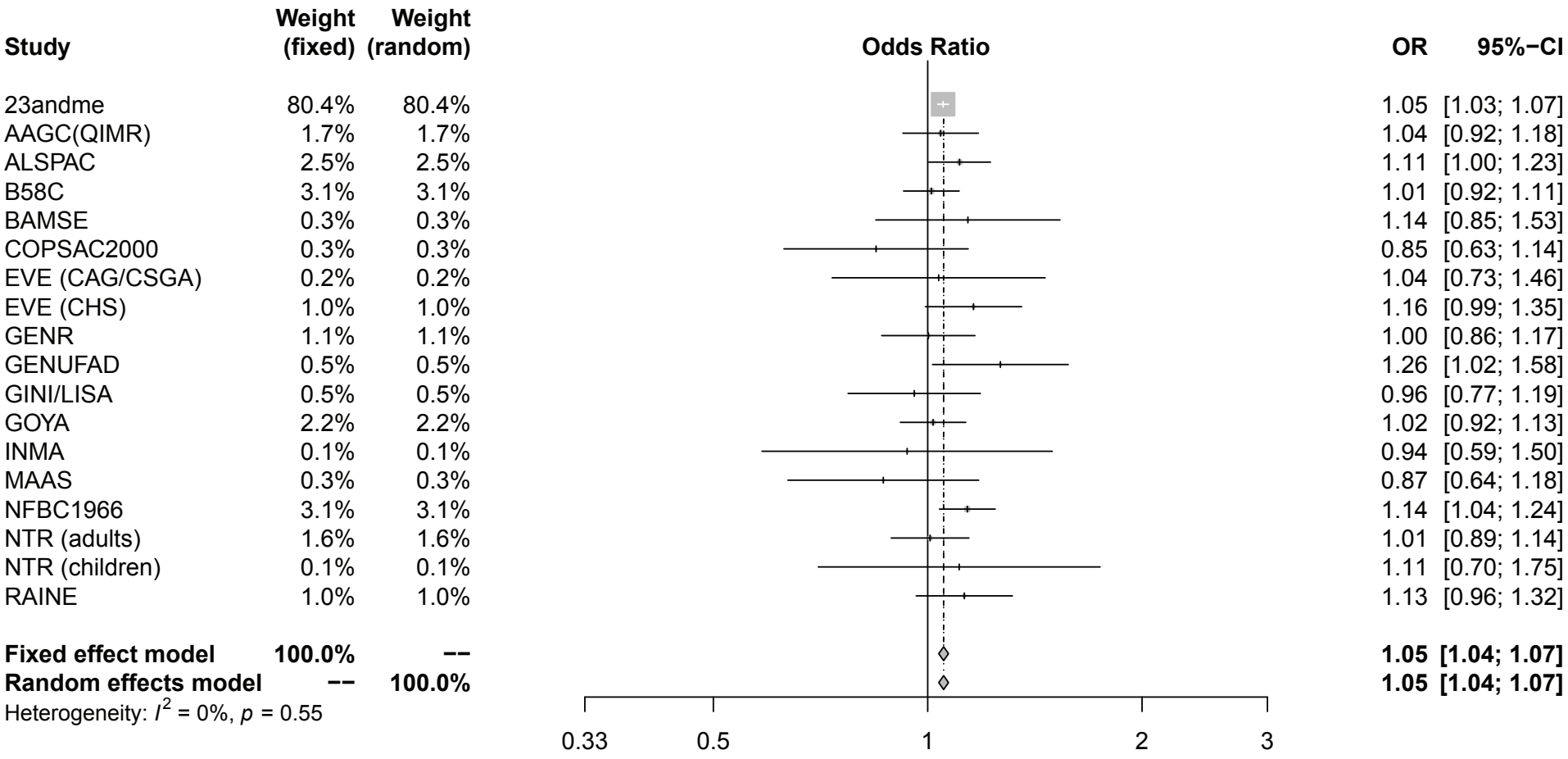
rs17294280[G]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	79.8%	21.6%	1.06	[1.04; 1.09]
AAGC(QIMR)	1.5%	6.4%	1.20	[1.04; 1.39]
ALSPAC	2.4%	8.7%	1.11	[0.99; 1.25]
B58C	3.3%	10.5%	1.13	[1.02; 1.24]
BAMSE	0.3%	1.6%	0.97	[0.69; 1.35]
COPSAC2000	0.3%	1.7%	0.95	[0.68; 1.32]
EVE (CAG/CSGA)	0.2%	1.1%	1.35	[0.90; 2.04]
EVE (CHS)	1.0%	4.7%	1.06	[0.89; 1.27]
GENR	1.1%	5.2%	1.10	[0.93; 1.31]
GENUFAD	0.4%	2.2%	0.98	[0.73; 1.30]
GINI/LISA	0.5%	2.5%	1.54	[1.18; 2.00]
GOYA	2.3%	8.4%	1.08	[0.96; 1.22]
INMA	0.1%	0.8%	0.74	[0.45; 1.21]
MAAS	0.2%	1.3%	0.77	[0.52; 1.12]
NFBC1966	3.3%	10.4%	0.94	[0.85; 1.04]
NTR (adults)	1.9%	7.5%	1.11	[0.98; 1.27]
NTR (children)	0.2%	0.9%	1.27	[0.82; 1.98]
RAINE	1.0%	4.6%	1.04	[0.87; 1.25]
Fixed effect model	**100.0%**	**−−**	**1.07**	**[1.05; 1.09]**
Random effects model	**−−**	**100.0%**	**1.08**	**[1.03; 1.12]**

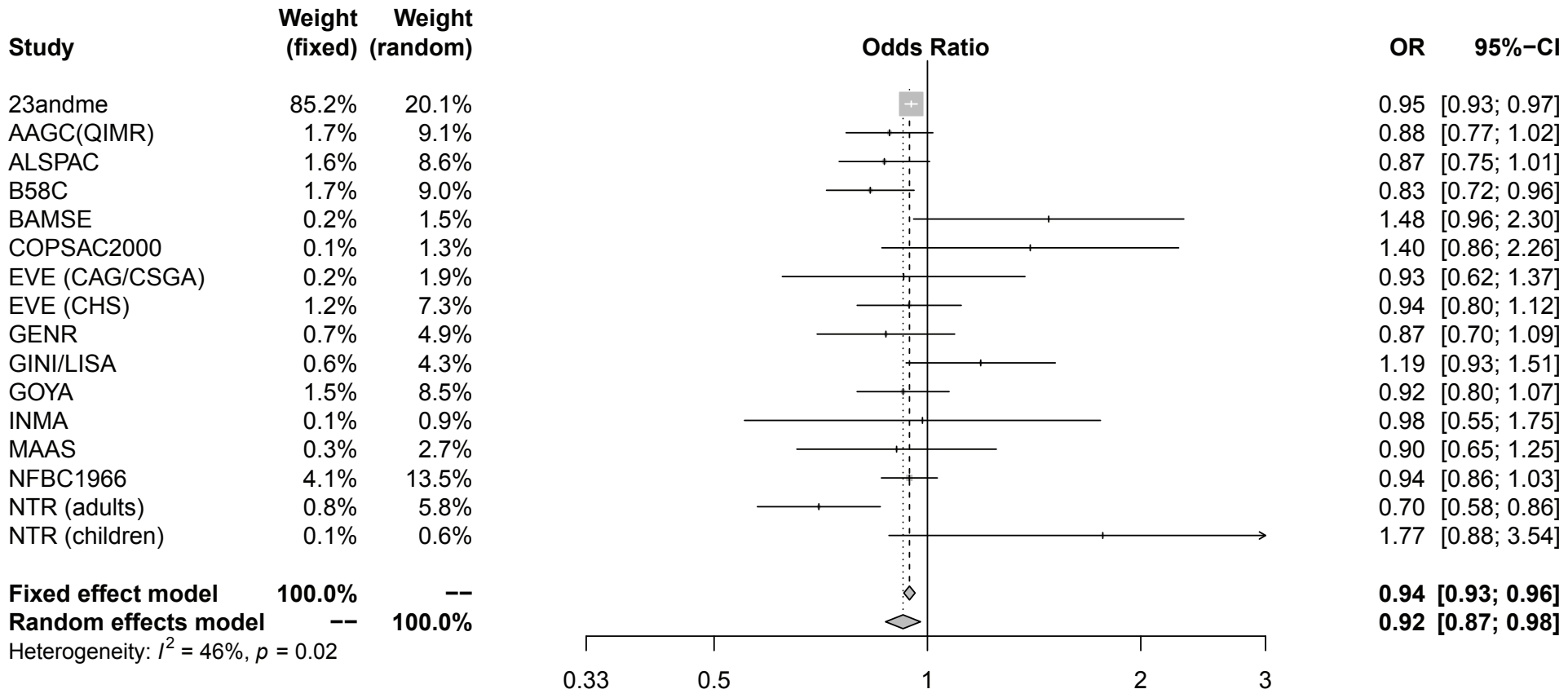
Heterogeneity: I^2 = 35%, p = 0.07



rs7824993[A]


Study
Weight (fixed)
Weight (random)
Odds Ratio
OR
95%−CI
23andme 80.4% 80.4% 1.05 [1.03; 1.07]
AAGC(QIMR) 1.7% 1.7% 1.04 [0.92; 1.18]
ALSPAC 2.5% 2.5% 1.11 [1.00; 1.23]
B58C 3.1% 3.1% 1.01 [0.92; 1.11]
BAMSE 0.3% 0.3% 1.14 [0.85; 1.53]
COPSAC2000 0.3% 0.3% 0.85 [0.63; 1.14]
EVE (CAG/CSGA) 0.2% 0.2% 1.04 [0.73; 1.46]
EVE (CHS) 1.0% 1.0% 1.16 [0.99; 1.35]
GENR 1.1% 1.1% 1.00 [0.86; 1.17]
GENUFAD 0.5% 0.5% 1.26 [1.02; 1.58]
GINI/LISA 0.5% 0.5% 0.96 [0.77; 1.19]
GOYA 2.2% 2.2% 1.02 [0.92; 1.13]
INMA 0.1% 0.1% 0.94 [0.59; 1.50]
MAAS 0.3% 0.3% 0.87 [0.64; 1.18]
NFBC1966 3.1% 3.1% 1.14 [1.04; 1.24]
NTR (adults) 1.6% 1.6% 1.01 [0.89; 1.14]
NTR (children) 0.1% 0.1% 1.11 [0.70; 1.75]
RAINE 1.0% 1.0% 1.13 [0.96; 1.32]
Fixed effect model 100.0% −− 1.05 [1.04; 1.07]
Random effects model −− 100.0% 1.05 [1.04; 1.07]
Heterogeneity: I^2 = 0%, p = 0.55
0.33
0.5
1
2
3

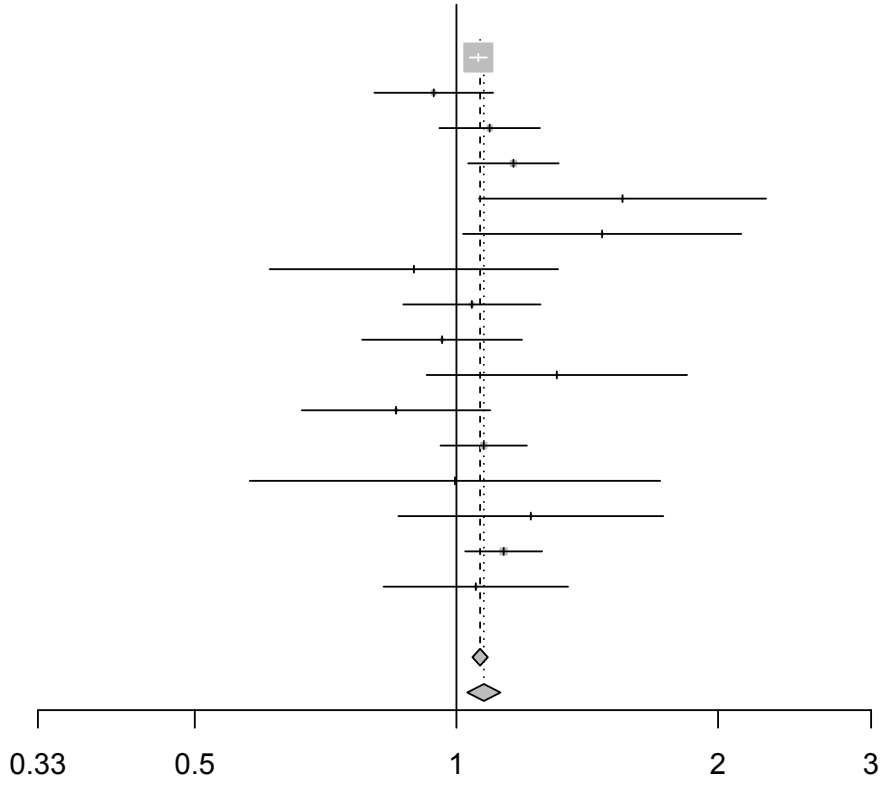
rs9282864[C]


Study
Weight (fixed)
Weight (random)
Odds Ratio
OR
95%−CI
23andme 85.2% 20.1% 0.95 [0.93; 0.97]
AAGC(QIMR) 1.7% 9.1% 0.88 [0.77; 1.02]
ALSPAC 1.6% 8.6% 0.87 [0.75; 1.01]
B58C 1.7% 9.0% 0.83 [0.72; 0.96]
BAMSE 0.2% 1.5% 1.48 [0.96; 2.30]
COPSAC2000 0.1% 1.3% 1.40 [0.86; 2.26]
EVE (CAG/CSGA) 0.2% 1.9% 0.93 [0.62; 1.37]
EVE (CHS) 1.2% 7.3% 0.94 [0.80; 1.12]
GENR 0.7% 4.9% 0.87 [0.70; 1.09]
GINI/LISA 0.6% 4.3% 1.19 [0.93; 1.51]
GOYA 1.5% 8.5% 0.92 [0.80; 1.07]
INMA 0.1% 0.9% 0.98 [0.55; 1.75]
MAAS 0.3% 2.7% 0.90 [0.65; 1.25]
NFBC1966 4.1% 13.5% 0.94 [0.86; 1.03]
NTR (adults) 0.8% 5.8% 0.70 [0.58; 0.86]
NTR (children) 0.1% 0.6% 1.77 [0.88; 3.54]
Fixed effect model 100.0% −− 0.94 [0.93; 0.96]
Random effects model −− 100.0% 0.92 [0.87; 0.98]
Heterogeneity: I^2 = 46%, p = 0.02
0.33
0.5
1
2
3

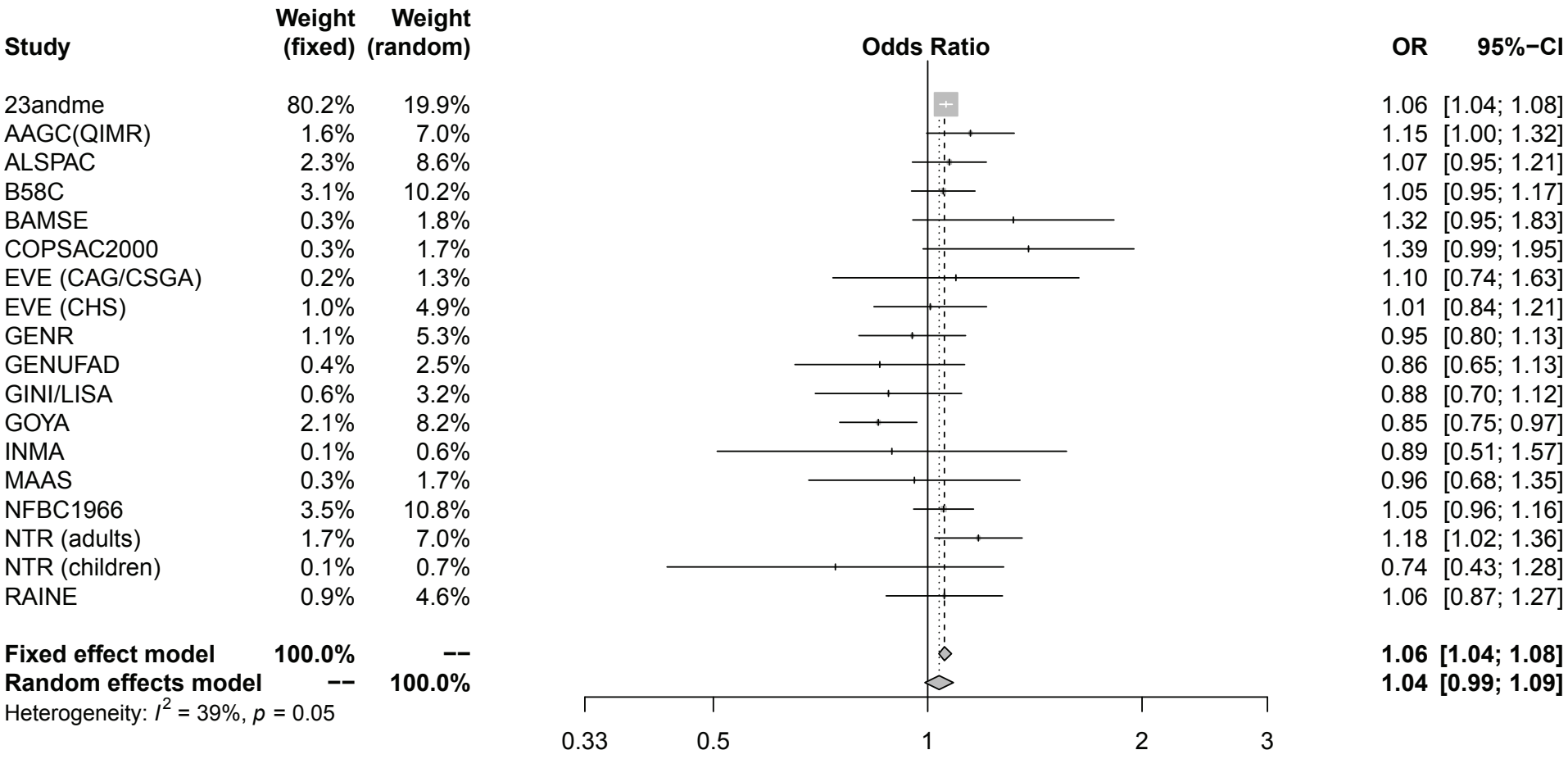
rs9687749[T]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	80.6%	31.3%	1.06	[1.04; 1.08]
AAGC(QIMR)	1.7%	6.3%	0.94	[0.80; 1.10]
ALSPAC	2.3%	8.2%	1.09	[0.96; 1.25]
B58C	2.9%	9.6%	1.16	[1.03; 1.31]
BAMSE	0.3%	1.3%	1.55	[1.06; 2.27]
COPSAC2000	0.3%	1.4%	1.47	[1.02; 2.13]
EVE (CAG/CSGA)	0.3%	1.3%	0.89	[0.61; 1.31]
EVE (CHS)	1.3%	5.0%	1.04	[0.87; 1.25]
GENR	0.9%	3.8%	0.96	[0.78; 1.19]
GENUFAD	0.3%	1.5%	1.30	[0.92; 1.84]
GINI/LISA	0.7%	2.8%	0.85	[0.66; 1.09]
GOYA	3.2%	10.3%	1.08	[0.96; 1.21]
INMA	0.1%	0.6%	1.00	[0.58; 1.72]
MAAS	0.3%	1.5%	1.22	[0.86; 1.73]
NFBC1966	4.0%	12.0%	1.13	[1.02; 1.25]
RAINE	0.7%	3.0%	1.05	[0.82; 1.34]
Fixed effect model	**100.0%**	**−−**	**1.06**	**[1.04; 1.09]**
Random effects model	**−−**	**100.0%**	**1.08**	**[1.03; 1.12]**

Heterogeneity: I^2 = 24%, p = 0.18



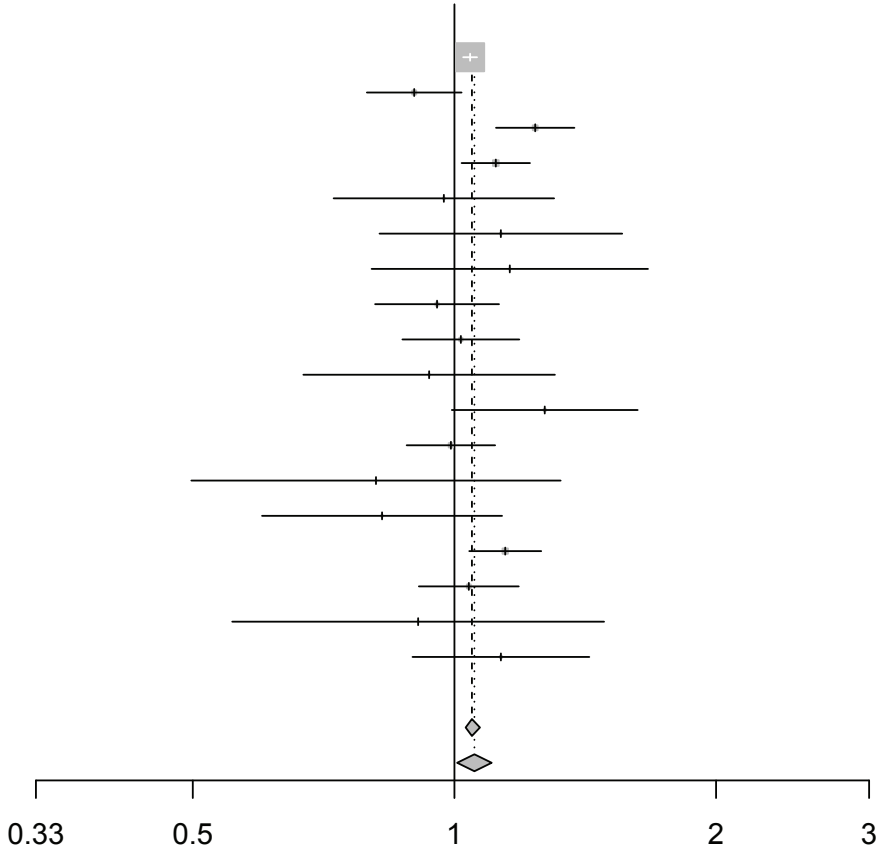
rs61977073[G]



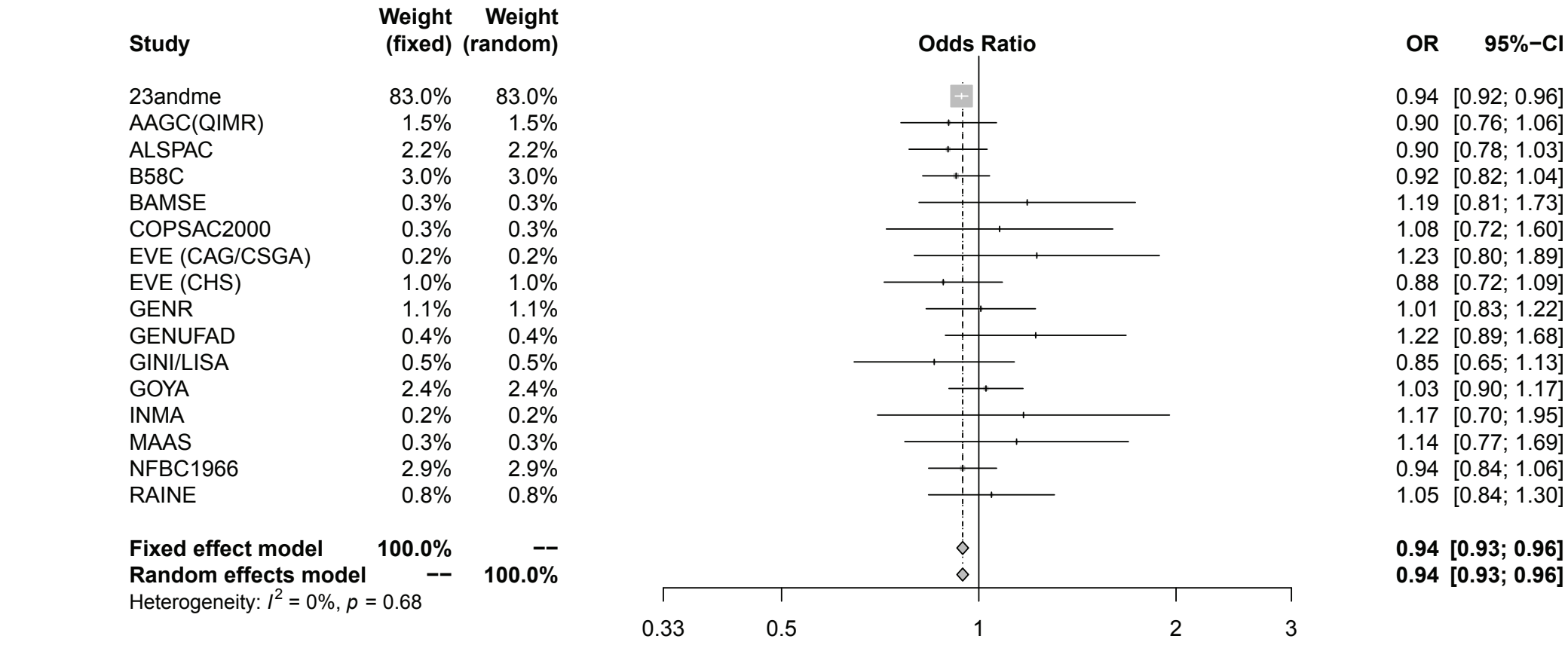
rs6470578[T]

Study	Weight (fixed)	Weight (random)	OR	95%−CI
23andme	81.1%	18.0%	1.04	[1.02; 1.06]
AAGC(QIMR)	1.7%	7.7%	0.90	[0.79; 1.02]
ALSPAC	2.5%	9.5%	1.24	[1.12; 1.37]
B58C	3.4%	10.7%	1.12	[1.02; 1.22]
BAMSE	0.3%	2.2%	0.97	[0.73; 1.30]
COPSAC2000	0.3%	1.8%	1.13	[0.82; 1.56]
EVE (CAG/CSGA)	0.2%	1.4%	1.16	[0.80; 1.67]
EVE (CHS)	1.0%	5.5%	0.96	[0.81; 1.13]
GENR	1.1%	5.9%	1.02	[0.87; 1.19]
GENUFAD	0.2%	1.7%	0.94	[0.67; 1.30]
GINI/LISA	0.5%	2.9%	1.27	[0.99; 1.62]
GOYA	2.0%	8.4%	0.99	[0.88; 1.11]
INMA	0.1%	0.8%	0.81	[0.50; 1.32]
MAAS	0.3%	1.9%	0.83	[0.60; 1.13]
NFBC1966	3.0%	10.3%	1.14	[1.04; 1.26]
NTR (adults)	1.6%	7.3%	1.04	[0.91; 1.19]
NTR (children)	0.1%	0.8%	0.91	[0.56; 1.49]
RAINE	0.5%	3.2%	1.13	[0.89; 1.43]
Fixed effect model	**100.0%**	**−−**	**1.05**	**[1.03; 1.07]**
Random effects model	**−−**	**100.0%**	**1.05**	**[1.01; 1.10]**

Heterogeneity: I^2 = 45%, p = 0.02



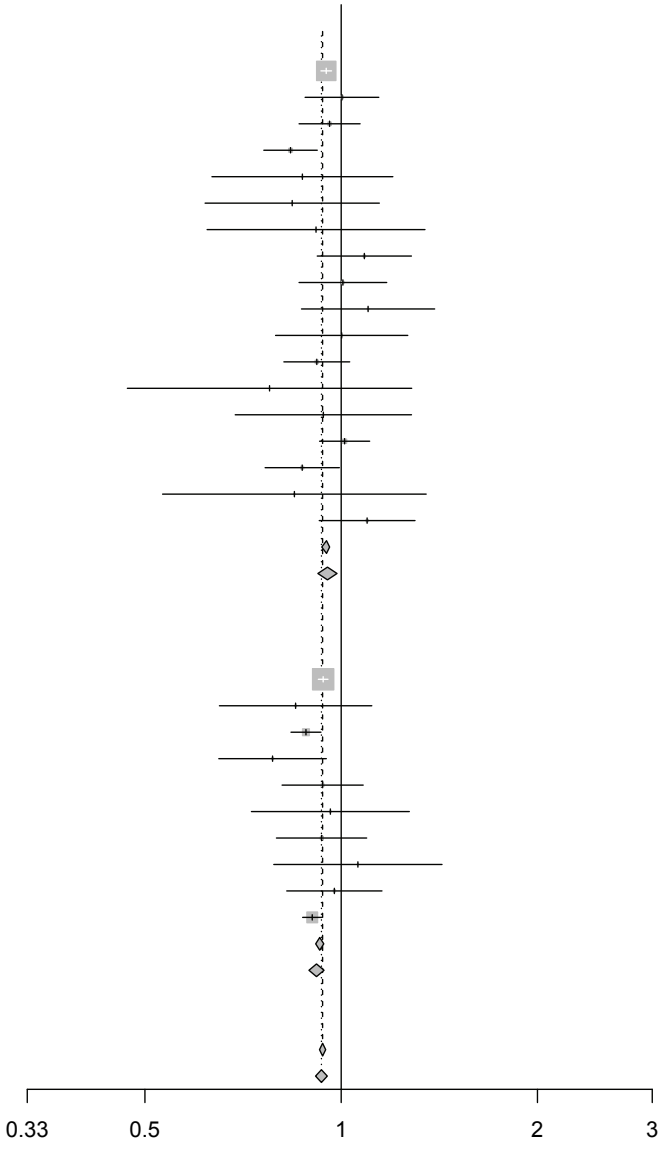
rs3787184[G]


Study
Weight (fixed)
Weight (random)
Odds Ratio
OR
95%−CI
23andme 83.0% 83.0% 0.94 [0.92; 0.96]
AAGC(QIMR) 1.5% 1.5% 0.90 [0.76; 1.06]
ALSPAC 2.2% 2.2% 0.90 [0.78; 1.03]
B58C 3.0% 3.0% 0.92 [0.82; 1.04]
BAMSE 0.3% 0.3% 1.19 [0.81; 1.73]
COPSAC2000 0.3% 0.3% 1.08 [0.72; 1.60]
EVE (CAG/CSGA) 0.2% 0.2% 1.23 [0.80; 1.89]
EVE (CHS) 1.0% 1.0% 0.88 [0.72; 1.09]
GENR 1.1% 1.1% 1.01 [0.83; 1.22]
GENUFAD 0.4% 0.4% 1.22 [0.89; 1.68]
GINI/LISA 0.5% 0.5% 0.85 [0.65; 1.13]
GOYA 2.4% 2.4% 1.03 [0.90; 1.17]
INMA 0.2% 0.2% 1.17 [0.70; 1.95]
MAAS 0.3% 0.3% 1.14 [0.77; 1.69]
NFBC1966 2.9% 2.9% 0.94 [0.84; 1.06]
RAINE 0.8% 0.8% 1.05 [0.84; 1.30]
Fixed effect model 100.0% −− 0.94 [0.93; 0.96]
Random effects model −− 100.0% 0.94 [0.93; 0.96]
Heterogeneity: I^2 = 0%, p = 0.68
0.33
0.5
1
2
3

Supplementary Figure 3

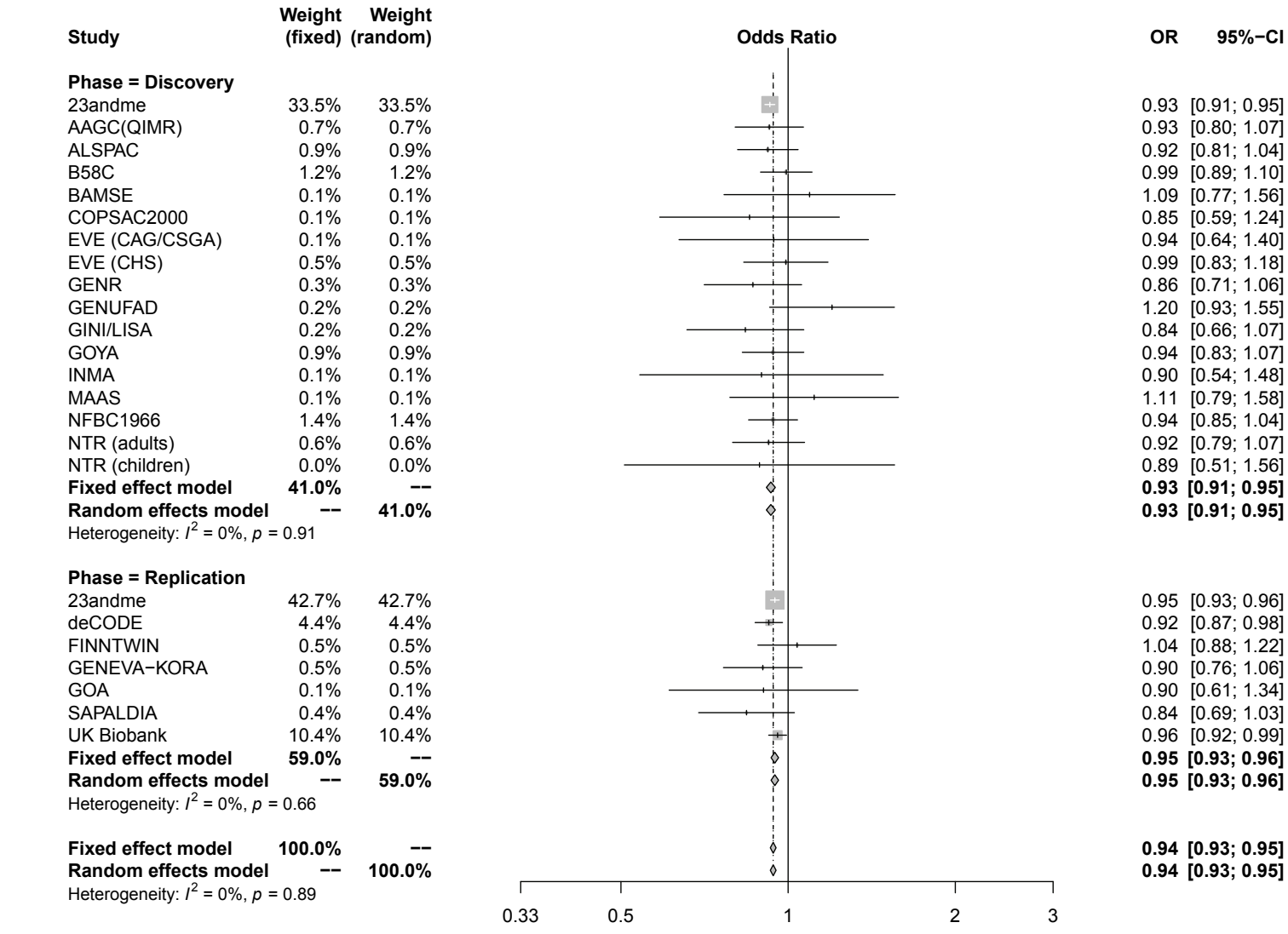
Forest plots from the discovery and replication phase for markers not previously associated to allergic rhinitis, including weights, estimates and 95% confidence intervals for inverse variance weighted fixed and random effects meta-analyses from genetic marker association of 16,531,985 genetic markers to allergic rhinitis from the discovery phase, including 212,120 individuals.

rs7717955[T]

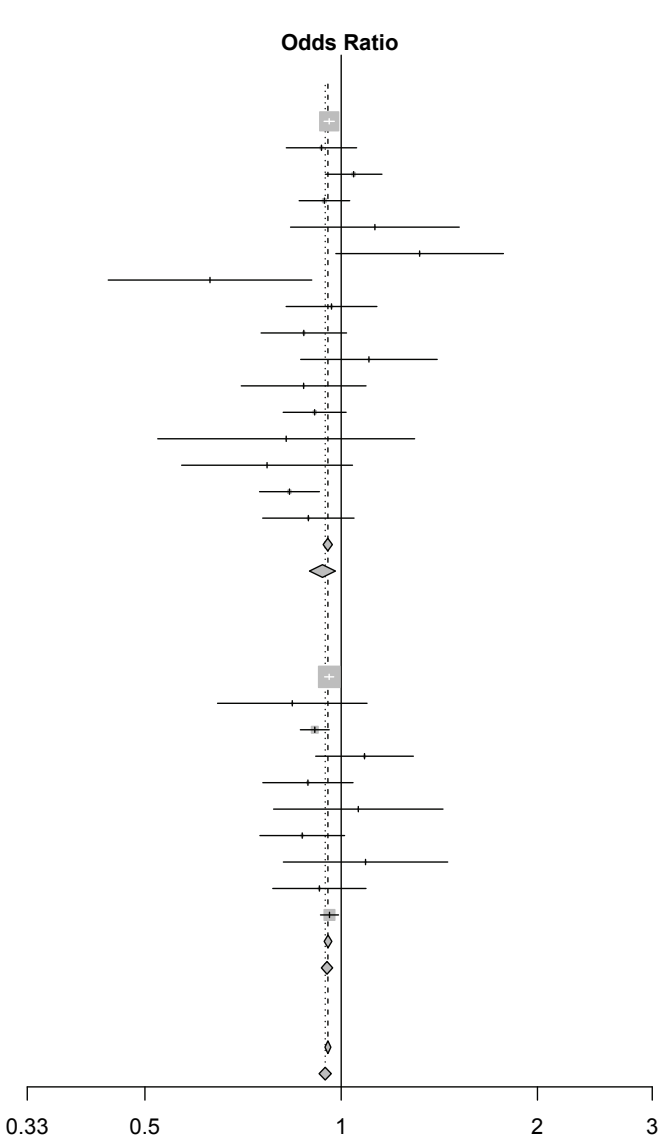


Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	33.8%	20.2%	0.95	[0.93; 0.97]
AAGC(QIMR)	0.7%	2.3%	1.00	[0.88; 1.14]
ALSPAC	1.0%	3.2%	0.96	[0.86; 1.07]
B58C	1.3%	4.0%	0.84	[0.76; 0.92]
BAMSE	0.1%	0.4%	0.87	[0.63; 1.20]
COPSAC2000	0.1%	0.4%	0.84	[0.62; 1.14]
EVE (CAG/CSGA)	0.1%	0.3%	0.91	[0.62; 1.34]
EVE (CHS)	0.4%	1.5%	1.09	[0.92; 1.28]
GENR	0.5%	1.7%	1.01	[0.86; 1.17]
GENUFAD	0.2%	0.8%	1.10	[0.87; 1.39]
GINI/LISA	0.2%	0.8%	1.00	[0.79; 1.27]
GOYA	0.9%	2.8%	0.92	[0.82; 1.03]
INMA	0.0%	0.2%	0.78	[0.47; 1.28]
MAAS	0.1%	0.4%	0.94	[0.69; 1.28]
NFBC1966	1.5%	4.5%	1.01	[0.93; 1.11]
NTR (adults)	0.7%	2.3%	0.87	[0.76; 0.99]
NTR (children)	0.1%	0.2%	0.85	[0.53; 1.35]
RAINE	0.4%	1.4%	1.10	[0.93; 1.30]
Fixed effect model	**42.2%**	**−−**	**0.95**	**[0.93; 0.96]**
Random effects model	**−−**	**47.3%**	**0.95**	**[0.92; 0.99]**
Heterogeneity: I^2 = 19%, p = 0.23				
Phase = Replication				
23andme	40.9%	20.8%	0.94	[0.92; 0.95]
CAMP	0.2%	0.6%	0.85	[0.65; 1.11]
deCODE	4.2%	9.4%	0.88	[0.84; 0.93]
ECRHS	0.3%	1.1%	0.78	[0.65; 0.95]
FINNTWIN	0.6%	1.9%	0.94	[0.81; 1.08]
GABRIELA	0.2%	0.5%	0.96	[0.73; 1.27]
GENEVA−KORA	0.5%	1.6%	0.93	[0.80; 1.09]
GOA	0.1%	0.5%	1.06	[0.79; 1.43]
SAPALDIA	0.4%	1.4%	0.98	[0.83; 1.15]
UK Biobank	10.4%	14.7%	0.90	[0.87; 0.93]
Fixed effect model	**57.8%**	**−−**	**0.93**	**[0.91; 0.94]**
Random effects model	**−−**	**52.7%**	**0.92**	**[0.89; 0.94]**
Heterogeneity: I^2 = 26%, p = 0.21				
Fixed effect model	**100.0%**	**−−**	**0.94**	**[0.93; 0.95]**
Random effects model	**−−**	**100.0%**	**0.93**	**[0.91; 0.95]**
Heterogeneity: I^2 = 28%, p = 0.09				

rs63406760[G]

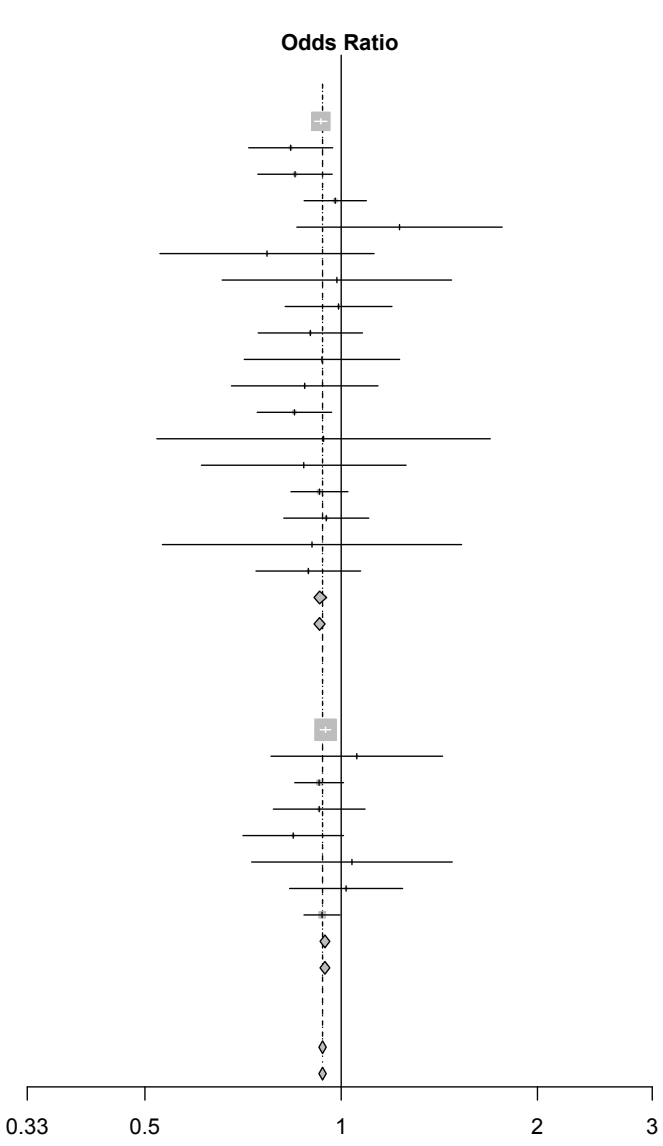


rs1504215[A]



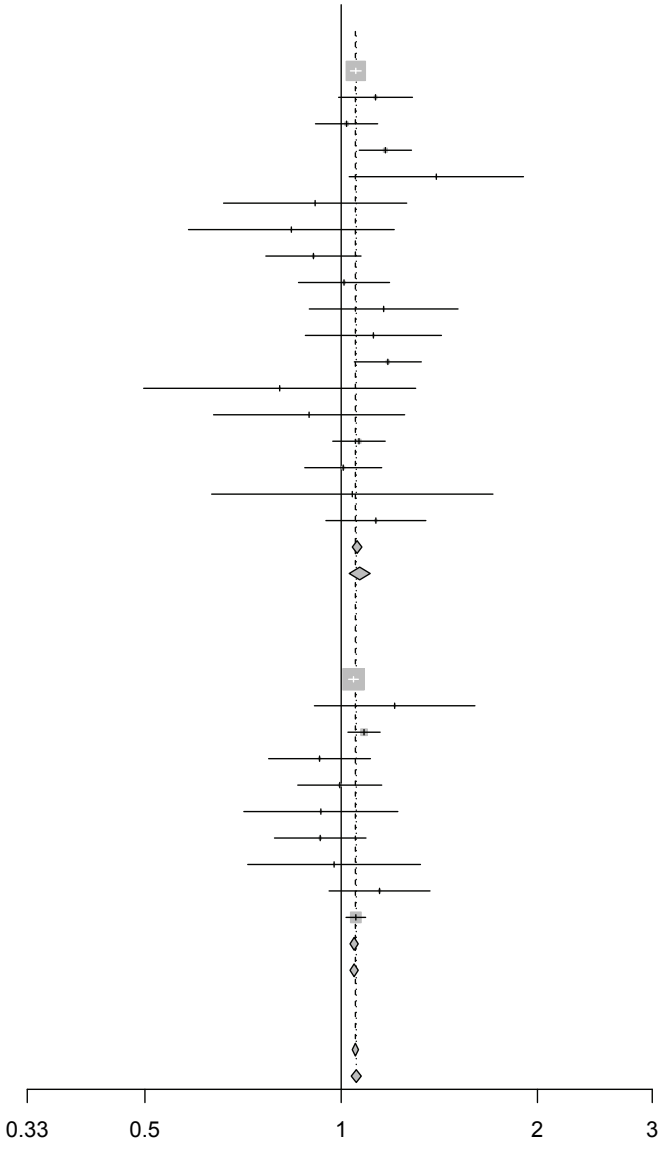
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	33.5%	19.5%	0.96	[0.94; 0.98]
AAGC(QIMR)	0.7%	2.7%	0.93	[0.82; 1.06]
ALSPAC	1.1%	3.9%	1.04	[0.95; 1.15]
B58C	1.4%	4.6%	0.94	[0.86; 1.03]
BAMSE	0.1%	0.5%	1.13	[0.84; 1.52]
COPSAC2000	0.1%	0.5%	1.32	[0.98; 1.77]
EVE (CAG/CSGA)	0.1%	0.4%	0.63	[0.44; 0.90]
EVE (CHS)	0.4%	1.7%	0.97	[0.82; 1.13]
GENR	0.5%	1.9%	0.88	[0.75; 1.02]
GENUFAD	0.2%	0.8%	1.10	[0.87; 1.40]
GINI/LISA	0.2%	0.9%	0.88	[0.70; 1.09]
GOYA	0.9%	3.2%	0.91	[0.81; 1.02]
INMA	0.1%	0.2%	0.82	[0.52; 1.30]
MAAS	0.1%	0.5%	0.77	[0.57; 1.04]
NFBC1966	1.0%	3.5%	0.83	[0.75; 0.93]
RAINE	0.4%	1.7%	0.89	[0.76; 1.05]
Fixed effect model	**40.8%**	**−−**	**0.95**	**[0.94; 0.97]**
Random effects model	**−−**	**46.4%**	**0.94**	**[0.89; 0.98]**
Heterogeneity: I^2 = 46%, p = 0.02				
Phase = Replication				
23andme	42.1%	20.1%	0.96	[0.94; 0.97]
CAMP	0.2%	0.6%	0.84	[0.65; 1.10]
deCODE	4.2%	10.0%	0.91	[0.87; 0.96]
ECRHS	0.4%	1.5%	1.09	[0.91; 1.29]
FINNTWIN	0.4%	1.7%	0.89	[0.76; 1.04]
GABRIELA	0.1%	0.5%	1.06	[0.79; 1.43]
GENEVA−KORA	0.5%	1.9%	0.87	[0.75; 1.01]
GOA	0.1%	0.5%	1.09	[0.82; 1.46]
SAPALDIA	0.4%	1.6%	0.93	[0.79; 1.09]
UK Biobank	10.8%	15.2%	0.96	[0.93; 0.99]
Fixed effect model	**59.2%**	**−−**	**0.95**	**[0.94; 0.97]**
Random effects model	**−−**	**53.6%**	**0.95**	**[0.93; 0.97]**
Heterogeneity: I^2 = 12%, p = 0.33				
Fixed effect model	**100.0%**	**−−**	**0.95**	**[0.94; 0.96]**
Random effects model	**−−**	**100.0%**	**0.95**	**[0.92; 0.97]**
Heterogeneity: I^2 = 34%, p = 0.05				

rs28361986[A]



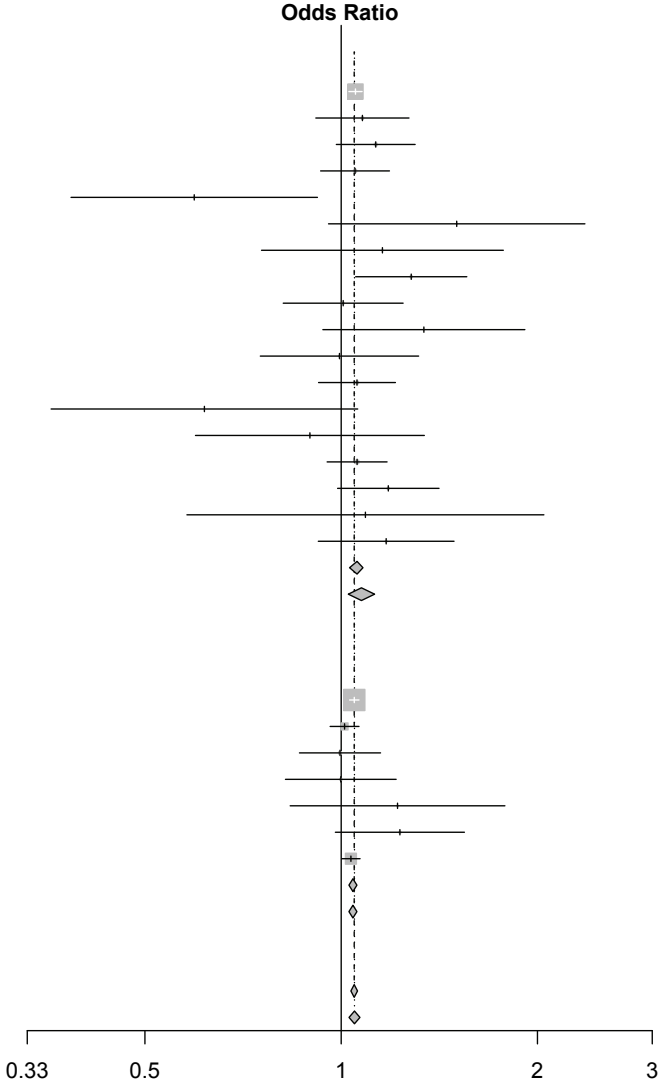
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	35.1%	35.1%	0.93	[0.91; 0.95]
AAGC(QIMR)	0.8%	0.8%	0.84	[0.72; 0.97]
ALSPAC	1.0%	1.0%	0.85	[0.75; 0.97]
B58C	1.4%	1.4%	0.98	[0.88; 1.09]
BAMSE	0.1%	0.1%	1.23	[0.86; 1.77]
COPSAC2000	0.1%	0.1%	0.77	[0.53; 1.12]
EVE (CAG/CSGA)	0.1%	0.1%	0.98	[0.66; 1.48]
EVE (CHS)	0.5%	0.5%	0.99	[0.82; 1.20]
GENR	0.5%	0.5%	0.90	[0.75; 1.08]
GENUFAD	0.2%	0.2%	0.93	[0.71; 1.23]
GINI/LISA	0.3%	0.3%	0.88	[0.68; 1.14]
GOYA	1.0%	1.0%	0.85	[0.74; 0.97]
INMA	0.0%	0.0%	0.94	[0.52; 1.69]
MAAS	0.1%	0.1%	0.88	[0.61; 1.26]
NFBC1966	1.7%	1.7%	0.93	[0.84; 1.02]
NTR (adults)	0.7%	0.7%	0.95	[0.82; 1.10]
NTR (children)	0.1%	0.1%	0.90	[0.53; 1.53]
RAINE	0.5%	0.5%	0.89	[0.74; 1.07]
Fixed effect model	**44.0%**	**−−**	**0.93**	**[0.91; 0.95]**
Random effects model	**−−**	**44.0%**	**0.93**	**[0.91; 0.94]**
Heterogeneity: I^2 = 0%, p = 0.86				
Phase = Replication				
23andme	47.7%	47.7%	0.95	[0.93; 0.96]
CAMP	0.2%	0.2%	1.06	[0.78; 1.43]
deCODE	2.3%	2.3%	0.93	[0.85; 1.01]
FINNTWIN	0.6%	0.6%	0.93	[0.79; 1.09]
GENEVA−KORA	0.5%	0.5%	0.84	[0.71; 1.01]
GOA	0.1%	0.1%	1.04	[0.73; 1.48]
SAPALDIA	0.4%	0.4%	1.02	[0.83; 1.24]
UK Biobank	4.1%	4.1%	0.93	[0.88; 1.00]
Fixed effect model	**56.0%**	**−−**	**0.94**	**[0.93; 0.96]**
Random effects model	**−−**	**56.0%**	**0.94**	**[0.93; 0.96]**
Heterogeneity: I^2 = 0%, p = 0.86				
Fixed effect model	**100.0%**	**−−**	**0.94**	**[0.92; 0.95]**
Random effects model	**−−**	**100.0%**	**0.94**	**[0.92; 0.95]**
Heterogeneity: I^2 = 0%, p = 0.91				

rs2070902[T]



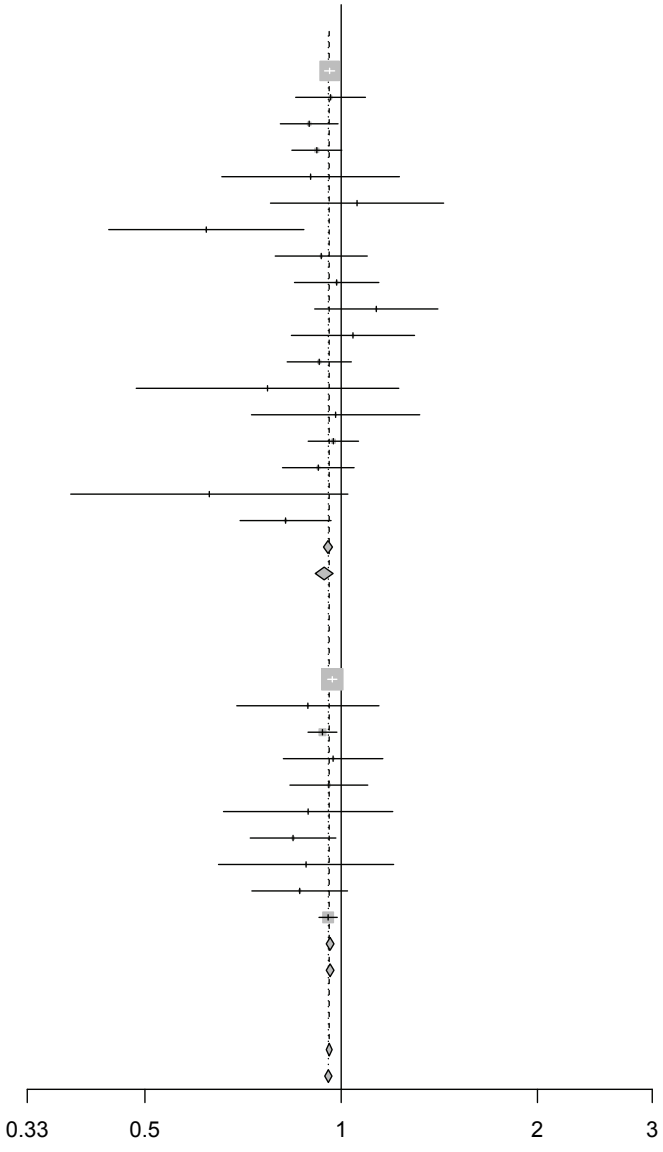
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	33.9%	25.0%	1.05	[1.03; 1.07]
AAGC(QIMR)	0.7%	1.8%	1.13	[0.99; 1.29]
ALSPAC	1.0%	2.4%	1.02	[0.91; 1.14]
B58C	1.4%	3.4%	1.17	[1.07; 1.28]
BAMSE	0.1%	0.3%	1.40	[1.03; 1.90]
COPSAC2000	0.1%	0.3%	0.91	[0.66; 1.26]
EVE (CAG/CSGA)	0.1%	0.2%	0.84	[0.58; 1.21]
EVE (CHS)	0.4%	1.1%	0.91	[0.77; 1.07]
GENR	0.5%	1.2%	1.01	[0.86; 1.19]
GENUFAD	0.2%	0.5%	1.16	[0.89; 1.51]
GINI/LISA	0.2%	0.5%	1.12	[0.88; 1.42]
GOYA	0.9%	2.1%	1.18	[1.05; 1.33]
INMA	0.1%	0.1%	0.80	[0.50; 1.30]
MAAS	0.1%	0.3%	0.89	[0.64; 1.25]
NFBC1966	1.4%	3.3%	1.06	[0.97; 1.17]
NTR (adults)	0.7%	1.6%	1.01	[0.88; 1.15]
NTR (children)	0.0%	0.1%	1.04	[0.63; 1.71]
RAINE	0.4%	1.0%	1.13	[0.95; 1.35]
Fixed effect model	**42.2%**	**−−**	**1.06**	**[1.04; 1.08]**
Random effects model	**−−**	**45.3%**	**1.07**	**[1.03; 1.11]**
Heterogeneity: I^2 = 24%, p = 0.17				
Phase = Replication				
23andme	41.6%	26.4%	1.04	[1.03; 1.06]
CAMP	0.2%	0.4%	1.21	[0.91; 1.60]
deCODE	3.8%	7.6%	1.08	[1.02; 1.15]
ECRHS	0.4%	0.9%	0.93	[0.77; 1.11]
FINNTWIN	0.6%	1.4%	0.99	[0.86; 1.15]
GABRIELA	0.2%	0.4%	0.93	[0.71; 1.22]
GENEVA−KORA	0.5%	1.2%	0.93	[0.79; 1.09]
GOA	0.1%	0.3%	0.98	[0.72; 1.32]
SAPALDIA	0.4%	1.0%	1.15	[0.96; 1.37]
UK Biobank	10.2%	15.0%	1.05	[1.02; 1.09]
Fixed effect model	**57.8%**	**−−**	**1.05**	**[1.03; 1.06]**
Random effects model	**−−**	**54.7%**	**1.05**	**[1.03; 1.06]**
Heterogeneity: I^2 = 0%, p = 0.45				
Fixed effect model	**100.0%**	**−−**	**1.05**	**[1.04; 1.06]**
Random effects model	**−−**	**100.0%**	**1.05**	**[1.04; 1.07]**
Heterogeneity: I^2 = 16%, p = 0.23				

rs111371454[G]



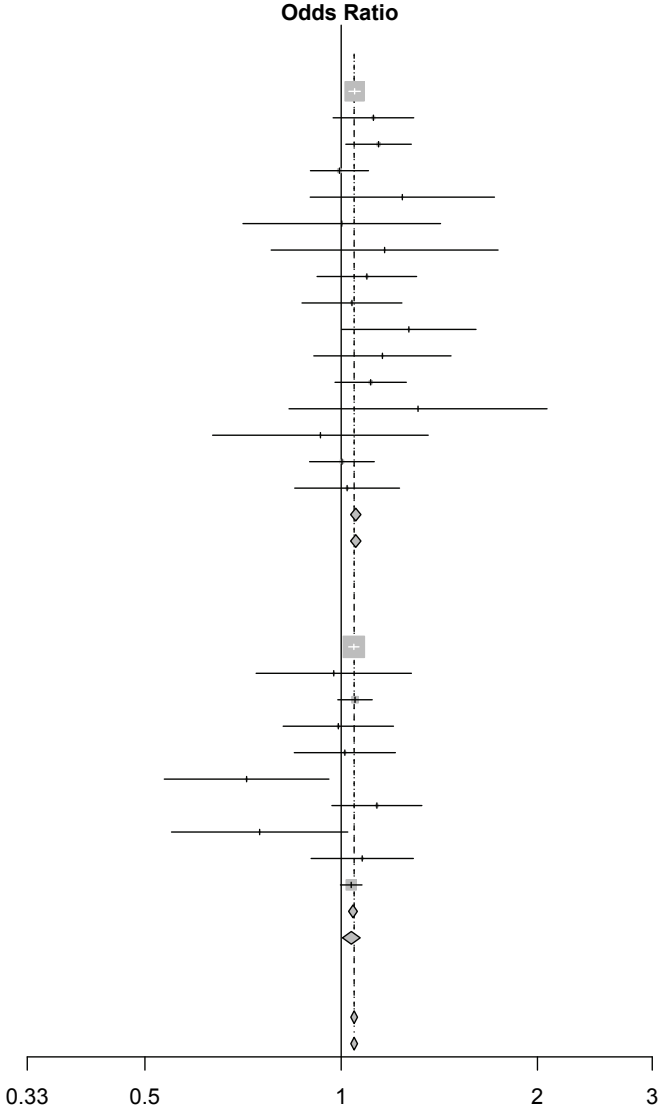
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	25.2%	23.8%	1.05	[1.03; 1.08]
AAGC(QIMR)	0.5%	1.3%	1.08	[0.91; 1.27]
ALSPAC	0.7%	1.8%	1.13	[0.98; 1.30]
B58C	0.9%	2.4%	1.05	[0.93; 1.19]
BAMSE	0.1%	0.2%	0.60	[0.39; 0.92]
COPSAC2000	0.1%	0.2%	1.50	[0.96; 2.36]
EVE (CAG/CSGA)	0.1%	0.2%	1.16	[0.75; 1.77]
EVE (CHS)	0.4%	0.9%	1.28	[1.05; 1.56]
GENR	0.3%	0.8%	1.01	[0.81; 1.24]
GENUFAD	0.1%	0.3%	1.34	[0.94; 1.91]
GINI/LISA	0.2%	0.5%	0.99	[0.75; 1.31]
GOYA	0.7%	1.9%	1.06	[0.92; 1.21]
INMA	0.0%	0.1%	0.62	[0.36; 1.06]
MAAS	0.1%	0.2%	0.90	[0.60; 1.34]
NFBC1966	1.2%	3.0%	1.06	[0.95; 1.18]
NTR (adults)	0.4%	1.1%	1.18	[0.99; 1.41]
NTR (children)	0.0%	0.1%	1.09	[0.58; 2.05]
RAINE	0.2%	0.6%	1.17	[0.92; 1.49]
Fixed effect model	**31.2%**	**−−**	**1.06**	**[1.03; 1.08]**
Random effects model	**−−**	**39.5%**	**1.07**	**[1.03; 1.13]**
Heterogeneity: I^2 = 25%, p = 0.16				
Phase = Replication				
23andme	49.0%	28.6%	1.05	[1.03; 1.06]
deCODE	5.3%	10.3%	1.01	[0.96; 1.06]
FINNTWIN	0.7%	1.7%	1.00	[0.86; 1.15]
GENEVA−KORA	0.4%	0.9%	1.00	[0.82; 1.21]
GOA	0.1%	0.3%	1.22	[0.84; 1.78]
SAPALDIA	0.3%	0.7%	1.23	[0.98; 1.55]
UK Biobank	13.1%	18.0%	1.03	[1.00; 1.07]
Fixed effect model	**68.8%**	**−−**	**1.04**	**[1.03; 1.06]**
Random effects model	**−−**	**60.5%**	**1.04**	**[1.03; 1.06]**
Heterogeneity: I^2 = 0%, p = 0.54				
Fixed effect model	**100.0%**	**−−**	**1.05**	**[1.03; 1.06]**
Random effects model	**−−**	**100.0%**	**1.05**	**[1.03; 1.07]**
Heterogeneity: I^2 = 17%, p = 0.22				

rs12509403[T]



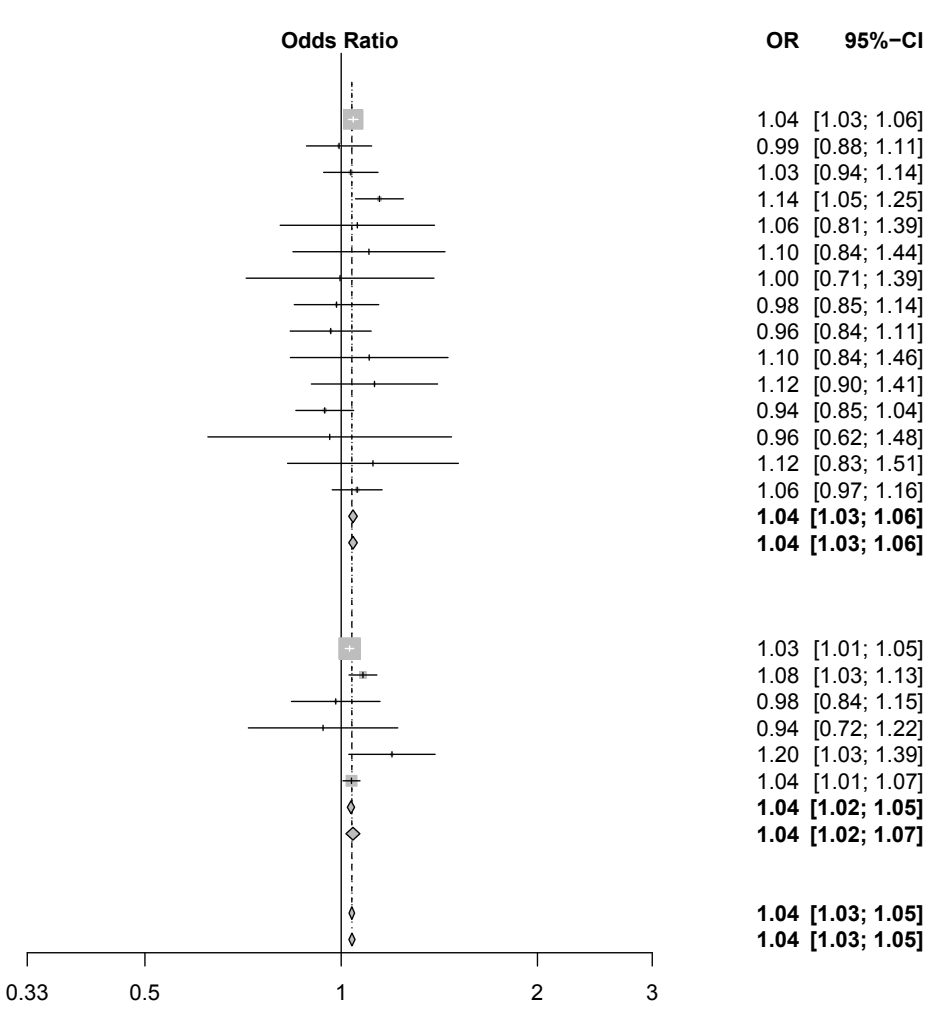
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	34.0%	31.0%	0.96	[0.94; 0.98]
AAGC(QIMR)	0.7%	1.1%	0.96	[0.85; 1.09]
ALSPAC	1.0%	1.5%	0.89	[0.81; 0.99]
B58C	1.4%	2.0%	0.92	[0.84; 1.00]
BAMSE	0.1%	0.2%	0.90	[0.66; 1.23]
COPSAC2000	0.1%	0.2%	1.06	[0.78; 1.44]
EVE (CAG/CSGA)	0.1%	0.1%	0.62	[0.44; 0.88]
EVE (CHS)	0.4%	0.6%	0.93	[0.79; 1.10]
GENR	0.5%	0.7%	0.98	[0.85; 1.14]
GENUFAD	0.2%	0.3%	1.13	[0.91; 1.41]
GINI/LISA	0.2%	0.3%	1.04	[0.84; 1.30]
GOYA	0.8%	1.2%	0.93	[0.83; 1.04]
INMA	0.0%	0.1%	0.77	[0.48; 1.23]
MAAS	0.1%	0.2%	0.98	[0.73; 1.32]
NFBC1966	1.4%	2.0%	0.97	[0.89; 1.06]
NTR (adults)	0.7%	1.0%	0.92	[0.81; 1.05]
NTR (children)	0.0%	0.1%	0.63	[0.38; 1.02]
RAINE	0.4%	0.6%	0.82	[0.70; 0.96]
Fixed effect model	**42.3%**	**−−**	**0.95**	**[0.94; 0.97]**
Random effects model	**−−**	**43.3%**	**0.94**	**[0.91; 0.97]**
Heterogeneity: I^2 = 16%, p = 0.26				
Phase = Replication				
23andme	41.1%	34.6%	0.97	[0.95; 0.98]
CAMP	0.2%	0.3%	0.89	[0.69; 1.14]
deCODE	4.1%	5.8%	0.94	[0.89; 0.98]
ECRHS	0.3%	0.5%	0.97	[0.82; 1.16]
FINNTWIN	0.6%	0.9%	0.96	[0.83; 1.10]
GABRIELA	0.1%	0.2%	0.89	[0.66; 1.20]
GENEVA−KORA	0.5%	0.7%	0.84	[0.73; 0.98]
GOA	0.1%	0.2%	0.88	[0.65; 1.20]
SAPALDIA	0.4%	0.6%	0.86	[0.73; 1.02]
UK Biobank	10.3%	13.1%	0.95	[0.92; 0.99]
Fixed effect model	**57.7%**	**−−**	**0.96**	**[0.95; 0.97]**
Random effects model	**−−**	**56.7%**	**0.96**	**[0.95; 0.97]**
Heterogeneity: I^2 = 0%, p = 0.58				
Fixed effect model	**100.0%**	**−−**	**0.96**	**[0.95; 0.97]**
Random effects model	**−−**	**100.0%**	**0.96**	**[0.94; 0.97]**
Heterogeneity: I^2 = 5%, p = 0.39				

rs9648346[G]

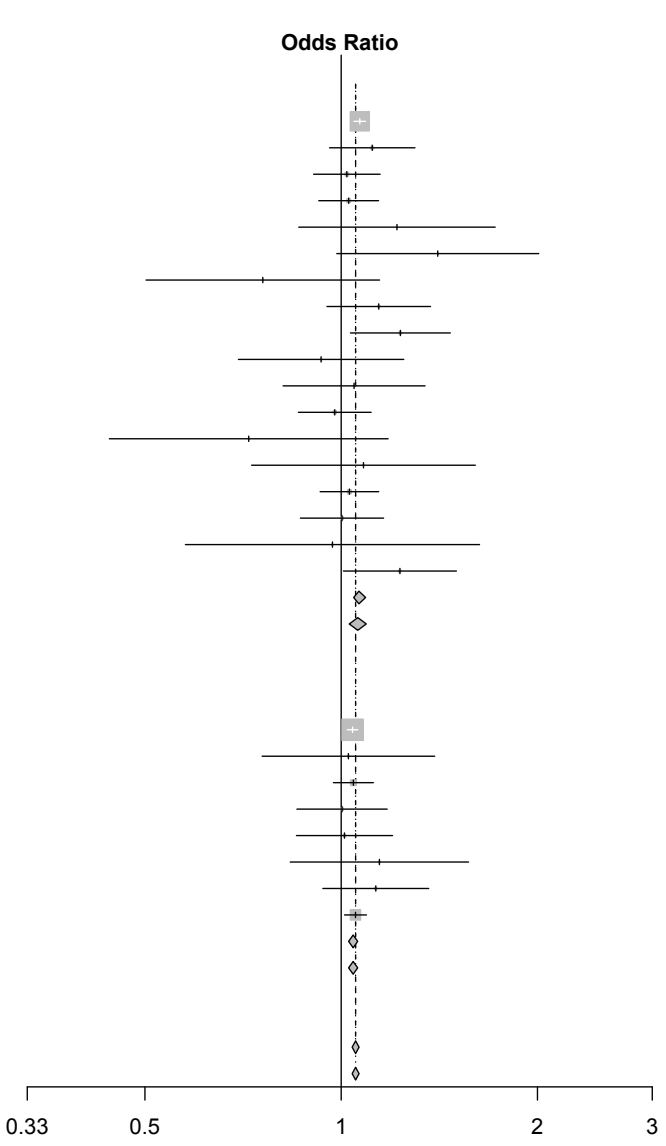


Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	34.6%	34.6%	1.05	[1.03; 1.07]
AAGC(QIMR)	0.7%	0.7%	1.12	[0.97; 1.29]
ALSPAC	1.0%	1.0%	1.14	[1.02; 1.28]
B58C	1.3%	1.3%	0.99	[0.90; 1.10]
BAMSE	0.1%	0.1%	1.24	[0.90; 1.72]
COPSAC2000	0.1%	0.1%	1.00	[0.71; 1.42]
EVE (CAG/CSGA)	0.1%	0.1%	1.17	[0.78; 1.74]
EVE (CHS)	0.5%	0.5%	1.09	[0.92; 1.31]
GENR	0.4%	0.4%	1.04	[0.87; 1.24]
GENUFAD	0.2%	0.2%	1.27	[1.00; 1.61]
GINI/LISA	0.2%	0.2%	1.16	[0.91; 1.47]
GOYA	0.9%	0.9%	1.11	[0.98; 1.26]
INMA	0.1%	0.1%	1.31	[0.83; 2.07]
MAAS	0.1%	0.1%	0.93	[0.64; 1.36]
NFBC1966	1.1%	1.1%	1.00	[0.89; 1.12]
RAINE	0.4%	0.4%	1.02	[0.85; 1.23]
Fixed effect model	**41.9%**	**--**	**1.05**	**[1.03; 1.07]**
Random effects model	**--**	**41.9%**	**1.05**	**[1.03; 1.07]**
Heterogeneity: I^2 = 0%, p = 0.73				
Phase = Replication				
23andme	42.4%	42.4%	1.05	[1.03; 1.06]
CAMP	0.2%	0.2%	0.97	[0.74; 1.28]
deCODE	3.8%	3.8%	1.05	[0.99; 1.12]
ECRHS	0.4%	0.4%	0.99	[0.81; 1.20]
FINNTWIN	0.4%	0.4%	1.01	[0.85; 1.21]
GABRIELA	0.2%	0.2%	0.72	[0.54; 0.96]
GENEVA−KORA	0.6%	0.6%	1.13	[0.97; 1.33]
GOA	0.1%	0.1%	0.75	[0.55; 1.02]
SAPALDIA	0.4%	0.4%	1.08	[0.90; 1.29]
UK Biobank	9.7%	9.7%	1.04	[1.00; 1.08]
Fixed effect model	**58.1%**	**--**	**1.04**	**[1.03; 1.06]**
Random effects model	**--**	**58.1%**	**1.04**	**[1.00; 1.07]**
Heterogeneity: I^2 = 30%, p = 0.17				
Fixed effect model	**100.0%**	**--**	**1.05**	**[1.03; 1.06]**
Random effects model	**--**	**100.0%**	**1.05**	**[1.03; 1.06]**
Heterogeneity: I^2 = 0%, p = 0.48				

rs35350651[C]

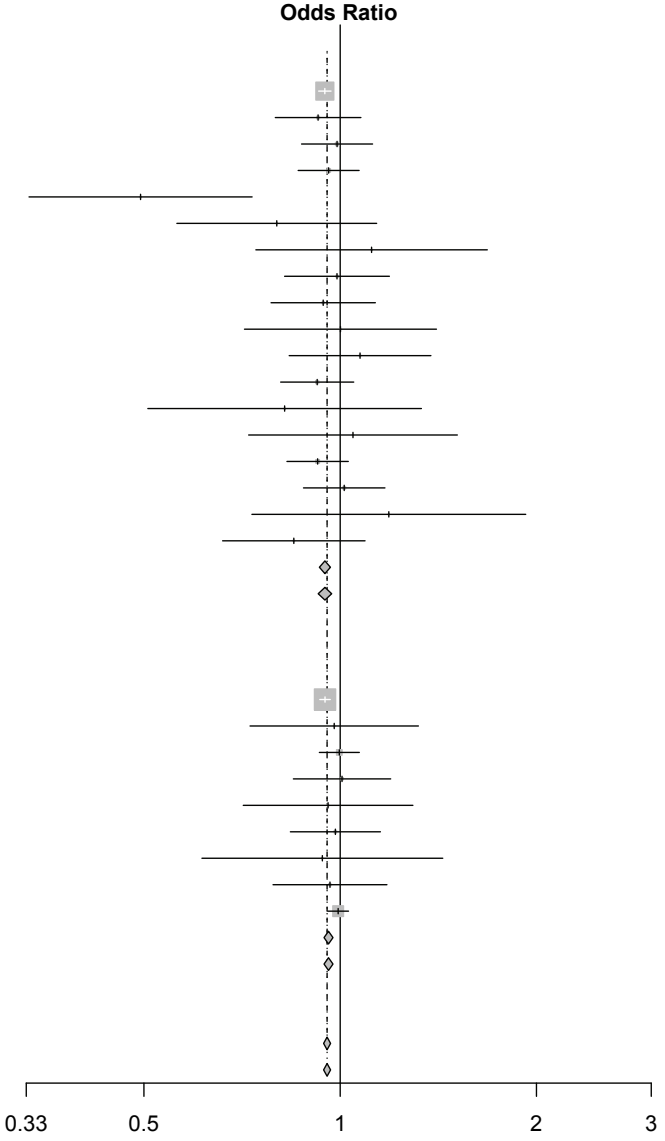


rs2519093[T]



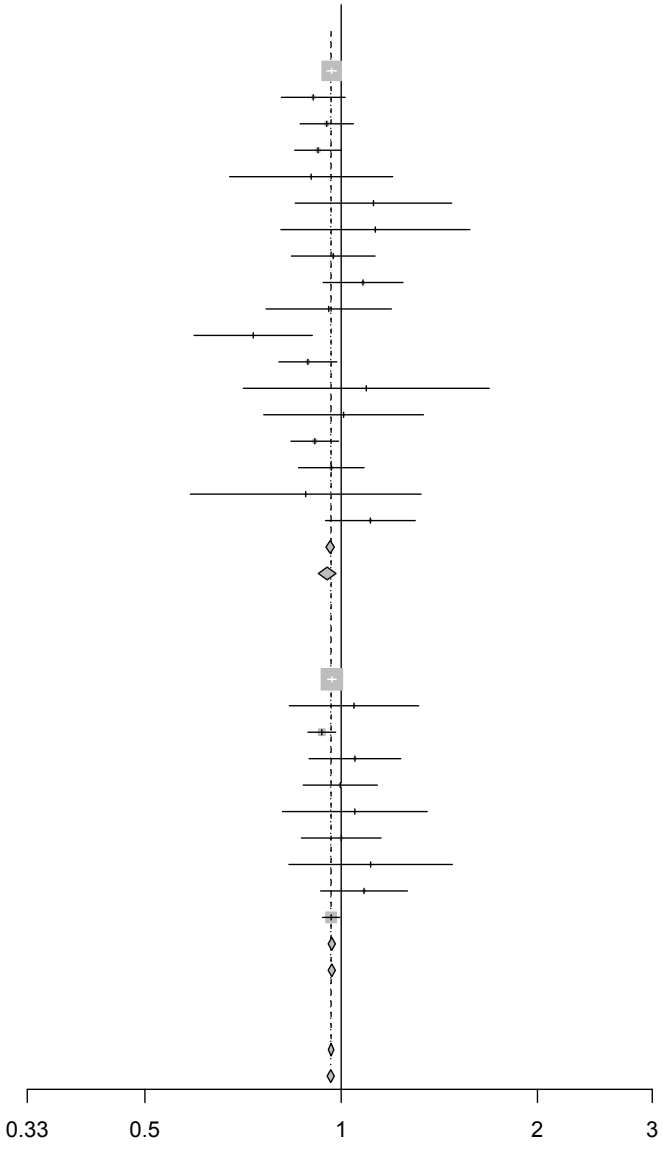
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	34.5%	34.5%	1.07	[1.05; 1.09]
AAGC(QIMR)	0.7%	0.7%	1.12	[0.96; 1.30]
ALSPAC	1.1%	1.1%	1.02	[0.91; 1.15]
B58C	1.3%	1.3%	1.03	[0.92; 1.14]
BAMSE	0.1%	0.1%	1.22	[0.86; 1.72]
COPSAC2000	0.1%	0.1%	1.41	[0.98; 2.01]
EVE (CAG/CSGA)	0.1%	0.1%	0.76	[0.50; 1.15]
EVE (CHS)	0.4%	0.4%	1.14	[0.95; 1.37]
GENR	0.5%	0.5%	1.23	[1.03; 1.47]
GENUFAD	0.2%	0.2%	0.93	[0.70; 1.25]
GINI/LISA	0.2%	0.2%	1.05	[0.81; 1.34]
GOYA	0.9%	0.9%	0.98	[0.86; 1.11]
INMA	0.1%	0.1%	0.72	[0.44; 1.18]
MAAS	0.1%	0.1%	1.08	[0.73; 1.61]
NFBC1966	1.4%	1.4%	1.03	[0.93; 1.14]
NTR (adults)	0.7%	0.7%	1.00	[0.87; 1.16]
NTR (children)	0.1%	0.1%	0.97	[0.58; 1.63]
RAINE	0.4%	0.4%	1.23	[1.01; 1.50]
Fixed effect model	**42.8%**	**−−**	**1.06**	**[1.04; 1.09]**
Random effects model	**−−**	**42.8%**	**1.06**	**[1.03; 1.09]**
Heterogeneity: I^2 = 7%, p = 0.37				
Phase = Replication				
23andme	42.4%	42.4%	1.04	[1.02; 1.06]
CAMP	0.2%	0.2%	1.03	[0.76; 1.39]
deCODE	3.0%	3.0%	1.04	[0.97; 1.12]
FINNTWIN	0.6%	0.6%	1.00	[0.86; 1.18]
GENEVA−KORA	0.5%	0.5%	1.01	[0.85; 1.20]
GOA	0.2%	0.2%	1.14	[0.84; 1.57]
SAPALDIA	0.4%	0.4%	1.13	[0.94; 1.36]
UK Biobank	9.9%	9.9%	1.05	[1.01; 1.09]
Fixed effect model	**57.2%**	**−−**	**1.04**	**[1.03; 1.06]**
Random effects model	**−−**	**57.2%**	**1.04**	**[1.03; 1.06]**
Heterogeneity: I^2 = 0%, p = 0.98				
Fixed effect model	**100.0%**	**−−**	**1.05**	**[1.04; 1.07]**
Random effects model	**−−**	**100.0%**	**1.05**	**[1.04; 1.07]**
Heterogeneity: I^2 = 0%, p = 0.61				

rs62257549[A]



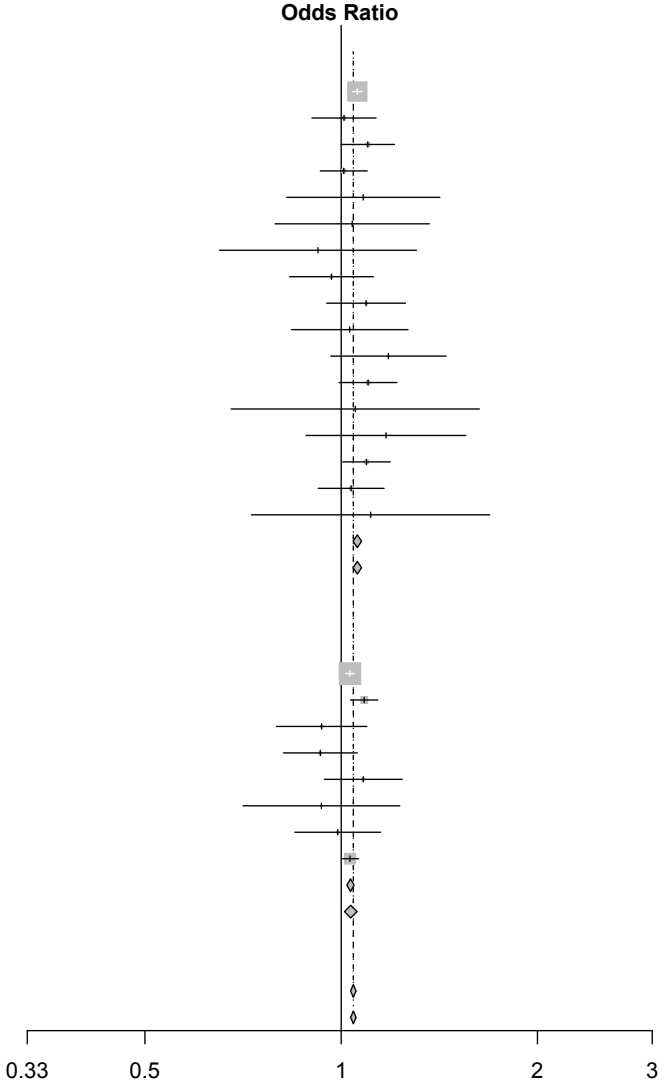
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	31.6%	31.6%	0.95	[0.93; 0.97]
AAGC(QIMR)	0.7%	0.7%	0.92	[0.80; 1.08]
ALSPAC	1.0%	1.0%	0.99	[0.87; 1.12]
B58C	1.3%	1.3%	0.96	[0.86; 1.07]
BAMSE	0.1%	0.1%	0.49	[0.33; 0.73]
COPSAC2000	0.1%	0.1%	0.80	[0.56; 1.14]
EVE (CAG/CSGA)	0.1%	0.1%	1.12	[0.74; 1.68]
EVE (CHS)	0.4%	0.4%	0.99	[0.82; 1.19]
GENR	0.4%	0.4%	0.94	[0.78; 1.13]
GENUFAD	0.1%	0.1%	1.00	[0.71; 1.41]
GINI/LISA	0.2%	0.2%	1.07	[0.84; 1.38]
GOYA	0.9%	0.9%	0.92	[0.81; 1.05]
INMA	0.1%	0.1%	0.82	[0.51; 1.33]
MAAS	0.1%	0.1%	1.05	[0.72; 1.51]
NFBC1966	1.3%	1.3%	0.92	[0.83; 1.03]
NTR (adults)	0.7%	0.7%	1.01	[0.88; 1.17]
NTR (children)	0.1%	0.1%	1.19	[0.73; 1.93]
RAINE	0.2%	0.2%	0.85	[0.66; 1.09]
Fixed effect model	**39.5%**	**−−**	**0.95**	**[0.93; 0.97]**
Random effects model	**−−**	**39.5%**	**0.95**	**[0.93; 0.97]**
Heterogeneity: I^2 = 2%, p = 0.43				
Phase = Replication				
23andme	44.2%	44.2%	0.95	[0.93; 0.97]
CAMP	0.2%	0.2%	0.98	[0.73; 1.32]
deCODE	3.0%	3.0%	1.00	[0.93; 1.07]
FINNTWIN	0.5%	0.5%	1.01	[0.85; 1.20]
GABRIELA	0.2%	0.2%	0.96	[0.71; 1.29]
GENEVA−KORA	0.6%	0.6%	0.98	[0.84; 1.15]
GOA	0.1%	0.1%	0.94	[0.61; 1.44]
SAPALDIA	0.4%	0.4%	0.96	[0.79; 1.18]
UK Biobank	11.3%	11.3%	0.99	[0.96; 1.03]
Fixed effect model	**60.5%**	**−−**	**0.96**	**[0.94; 0.97]**
Random effects model	**−−**	**60.5%**	**0.96**	**[0.94; 0.97]**
Heterogeneity: I^2 = 0%, p = 0.58				
Fixed effect model	**100.0%**	**−−**	**0.95**	**[0.94; 0.97]**
Random effects model	**−−**	**100.0%**	**0.95**	**[0.94; 0.97]**
Heterogeneity: I^2 = 0%, p = 0.53				

rs11677002[C]



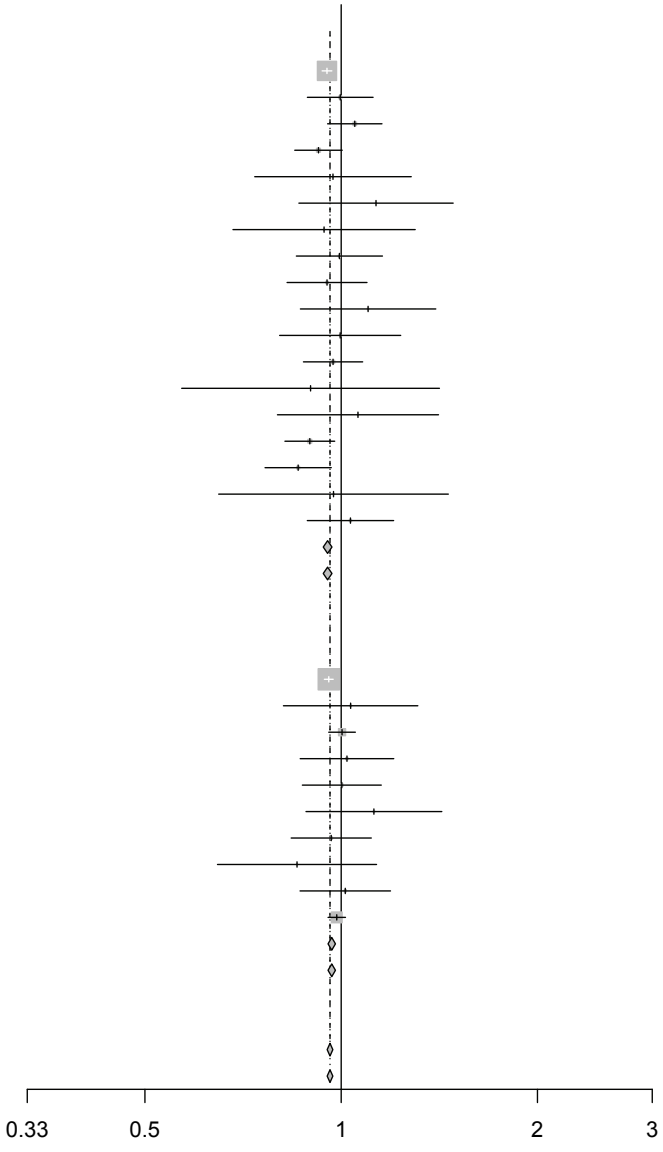
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	33.8%	29.5%	0.97	[0.95; 0.98]
AAGC(QIMR)	0.7%	1.3%	0.91	[0.81; 1.01]
ALSPAC	1.1%	1.8%	0.95	[0.86; 1.04]
B58C	1.4%	2.3%	0.92	[0.85; 1.00]
BAMSE	0.1%	0.2%	0.90	[0.67; 1.20]
COPSAC2000	0.1%	0.2%	1.12	[0.85; 1.48]
EVE (CAG/CSGA)	0.1%	0.1%	1.13	[0.81; 1.58]
EVE (CHS)	0.4%	0.7%	0.97	[0.84; 1.13]
GENR	0.5%	0.8%	1.08	[0.94; 1.24]
GENUFAD	0.2%	0.3%	0.96	[0.77; 1.19]
GINI/LISA	0.2%	0.4%	0.73	[0.59; 0.90]
GOYA	0.9%	1.5%	0.89	[0.80; 0.98]
INMA	0.1%	0.1%	1.09	[0.71; 1.69]
MAAS	0.1%	0.2%	1.01	[0.76; 1.34]
NFBC1966	1.3%	2.2%	0.91	[0.84; 0.99]
NTR (adults)	0.7%	1.2%	0.97	[0.86; 1.08]
NTR (children)	0.1%	0.1%	0.88	[0.59; 1.33]
RAINE	0.4%	0.7%	1.11	[0.95; 1.30]
Fixed effect model	**42.1%**	**−−**	**0.96**	**[0.95; 0.98]**
Random effects model	**−−**	**43.8%**	**0.95**	**[0.92; 0.98]**
Heterogeneity: I^2 = 21%, p = 0.20				
Phase = Replication				
23andme	41.4%	32.5%	0.97	[0.95; 0.98]
CAMP	0.2%	0.3%	1.05	[0.83; 1.31]
deCODE	3.9%	6.2%	0.93	[0.89; 0.98]
ECRHS	0.4%	0.6%	1.05	[0.89; 1.23]
FINNTWIN	0.6%	0.9%	1.00	[0.87; 1.14]
GABRIELA	0.1%	0.3%	1.05	[0.81; 1.35]
GENEVA−KORA	0.5%	0.8%	1.00	[0.87; 1.15]
GOA	0.1%	0.2%	1.11	[0.83; 1.48]
SAPALDIA	0.4%	0.7%	1.08	[0.93; 1.26]
UK Biobank	10.2%	13.7%	0.97	[0.94; 0.99]
Fixed effect model	**57.9%**	**−−**	**0.97**	**[0.96; 0.98]**
Random effects model	**−−**	**56.2%**	**0.97**	**[0.96; 0.98]**
Heterogeneity: I^2 = 0%, p = 0.61				
Fixed effect model	**100.0%**	**−−**	**0.97**	**[0.96; 0.97]**
Random effects model	**−−**	**100.0%**	**0.96**	**[0.95; 0.98]**
Heterogeneity: I^2 = 7%, p = 0.36				

rs35597970[−]



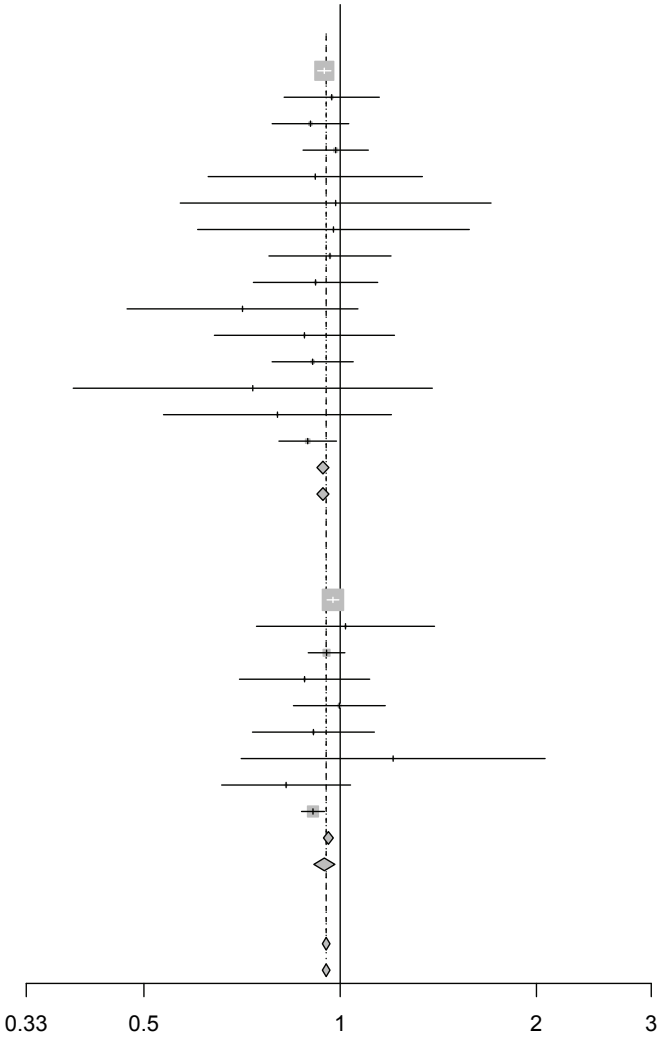
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	33.7%	33.7%	1.06	[1.04; 1.08]
AAGC(QIMR)	0.7%	0.7%	1.01	[0.90; 1.13]
ALSPAC	1.1%	1.1%	1.10	[1.00; 1.21]
B58C	1.4%	1.4%	1.01	[0.93; 1.10]
BAMSE	0.1%	0.1%	1.08	[0.82; 1.42]
COPSAC2000	0.1%	0.1%	1.04	[0.79; 1.37]
EVE (CAG/CSGA)	0.1%	0.1%	0.92	[0.65; 1.30]
EVE (CHS)	0.4%	0.4%	0.97	[0.83; 1.12]
GENR	0.5%	0.5%	1.09	[0.95; 1.26]
GENUFAD	0.2%	0.2%	1.03	[0.84; 1.27]
GINI/LISA	0.2%	0.2%	1.18	[0.96; 1.45]
GOYA	0.9%	0.9%	1.10	[0.99; 1.22]
INMA	0.0%	0.0%	1.05	[0.68; 1.63]
MAAS	0.1%	0.1%	1.17	[0.88; 1.55]
NFBC1966	1.3%	1.3%	1.09	[1.01; 1.19]
NTR (adults)	0.7%	0.7%	1.04	[0.92; 1.16]
NTR (children)	0.1%	0.1%	1.11	[0.73; 1.69]
Fixed effect model	**41.8%**	**−−**	**1.06**	**[1.04; 1.07]**
Random effects model	**−−**	**41.8%**	**1.06**	**[1.04; 1.07]**
Heterogeneity: I^2 = 0%, p = 0.95				
Phase = Replication				
23andme	41.8%	41.8%	1.03	[1.02; 1.05]
deCODE	4.1%	4.1%	1.08	[1.03; 1.14]
ECRHS	0.4%	0.4%	0.93	[0.80; 1.09]
FINNTWIN	0.6%	0.6%	0.93	[0.82; 1.06]
GENEVA−KORA	0.5%	0.5%	1.08	[0.94; 1.24]
GOA	0.1%	0.1%	0.93	[0.71; 1.23]
SAPALDIA	0.4%	0.4%	0.99	[0.85; 1.15]
UK Biobank	10.4%	10.4%	1.03	[1.00; 1.06]
Fixed effect model	**58.2%**	**−−**	**1.03**	**[1.02; 1.05]**
Random effects model	**−−**	**58.2%**	**1.03**	**[1.01; 1.06]**
Heterogeneity: I^2 = 25%, p = 0.23				
Fixed effect model	**100.0%**	**−−**	**1.04**	**[1.03; 1.05]**
Random effects model	**−−**	**100.0%**	**1.04**	**[1.03; 1.05]**
Heterogeneity: I^2 = 0%, p = 0.53				

rs2815765[T]



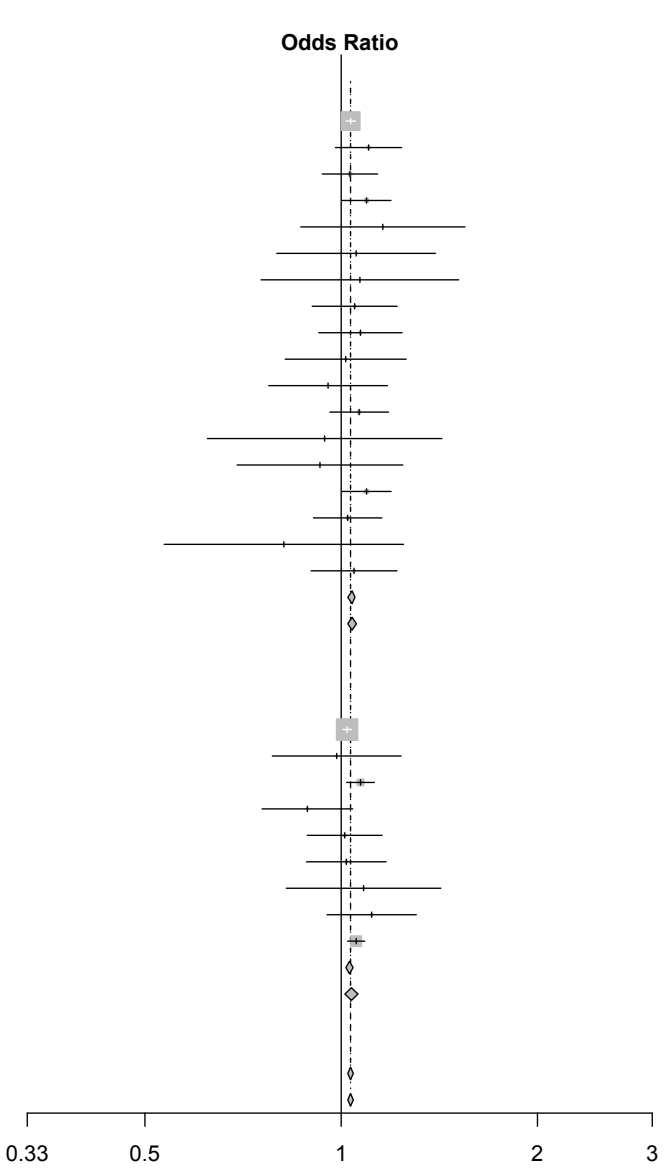
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	33.4%	33.4%	0.95	[0.93; 0.97]
AAGC(QIMR)	0.7%	0.7%	1.00	[0.89; 1.12]
ALSPAC	1.1%	1.1%	1.05	[0.95; 1.15]
B58C	1.4%	1.4%	0.92	[0.85; 1.00]
BAMSE	0.1%	0.1%	0.97	[0.74; 1.28]
COPSAC2000	0.1%	0.1%	1.13	[0.86; 1.49]
EVE (CAG/CSGA)	0.1%	0.1%	0.94	[0.68; 1.30]
EVE (CHS)	0.4%	0.4%	0.99	[0.85; 1.16]
GENR	0.5%	0.5%	0.95	[0.83; 1.10]
GENUFAD	0.2%	0.2%	1.10	[0.87; 1.40]
GINI/LISA	0.2%	0.2%	1.00	[0.80; 1.23]
GOYA	0.9%	0.9%	0.97	[0.88; 1.08]
INMA	0.0%	0.0%	0.90	[0.57; 1.41]
MAAS	0.1%	0.1%	1.06	[0.80; 1.41]
NFBC1966	1.3%	1.3%	0.90	[0.82; 0.98]
NTR (adults)	0.7%	0.7%	0.86	[0.76; 0.97]
NTR (children)	0.1%	0.1%	0.97	[0.65; 1.46]
RAINE	0.4%	0.4%	1.03	[0.89; 1.20]
Fixed effect model	**41.9%**	**−−**	**0.95**	**[0.94; 0.97]**
Random effects model	**−−**	**41.9%**	**0.95**	**[0.94; 0.97]**
Heterogeneity: I^2 = 0%, p = 0.58				
Phase = Replication				
23andme	40.8%	40.8%	0.96	[0.94; 0.97]
CAMP	0.2%	0.2%	1.03	[0.82; 1.31]
deCODE	4.5%	4.5%	1.00	[0.96; 1.05]
ECRHS	0.4%	0.4%	1.02	[0.87; 1.20]
FINNTWIN	0.5%	0.5%	1.00	[0.87; 1.15]
GABRIELA	0.2%	0.2%	1.12	[0.88; 1.43]
GENEVA−KORA	0.5%	0.5%	0.97	[0.84; 1.11]
GOA	0.1%	0.1%	0.86	[0.65; 1.13]
SAPALDIA	0.4%	0.4%	1.01	[0.86; 1.19]
UK Biobank	10.5%	10.5%	0.98	[0.95; 1.02]
Fixed effect model	**58.1%**	**−−**	**0.97**	**[0.95; 0.98]**
Random effects model	**−−**	**58.1%**	**0.97**	**[0.95; 0.98]**
Heterogeneity: I^2 = 0%, p = 0.45				
Fixed effect model	**100.0%**	**−−**	**0.96**	**[0.95; 0.97]**
Random effects model	**−−**	**100.0%**	**0.96**	**[0.95; 0.97]**
Heterogeneity: I^2 = 0%, p = 0.52				

rs11671925[A]



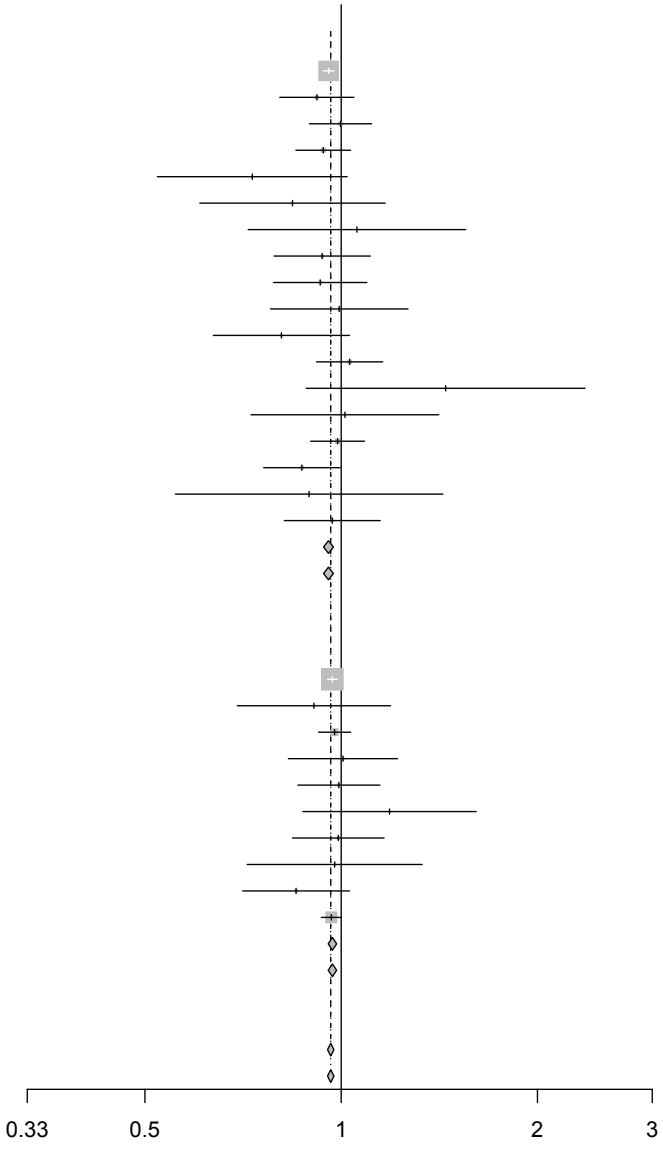
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	33.4%	33.4%	0.95	[0.92; 0.97]
AAGC(QIMR)	0.6%	0.6%	0.97	[0.82; 1.15]
ALSPAC	1.0%	1.0%	0.90	[0.79; 1.03]
B58C	1.4%	1.4%	0.98	[0.88; 1.10]
BAMSE	0.1%	0.1%	0.92	[0.63; 1.34]
COPSAC2000	0.1%	0.1%	0.98	[0.57; 1.70]
EVE (CAG/CSGA)	0.1%	0.1%	0.98	[0.60; 1.58]
EVE (CHS)	0.4%	0.4%	0.96	[0.78; 1.20]
GENR	0.4%	0.4%	0.92	[0.74; 1.14]
GENUFAD	0.1%	0.1%	0.71	[0.47; 1.06]
GINI/LISA	0.2%	0.2%	0.88	[0.64; 1.21]
GOYA	0.9%	0.9%	0.91	[0.79; 1.05]
INMA	0.0%	0.0%	0.73	[0.39; 1.38]
MAAS	0.1%	0.1%	0.80	[0.54; 1.20]
NFBC1966	1.8%	1.8%	0.89	[0.81; 0.99]
Fixed effect model	**40.5%**	**--**	**0.94**	**[0.92; 0.96]**
Random effects model	**--**	**40.5%**	**0.94**	**[0.92; 0.96]**
Heterogeneity: I^2 = 0%, p = 0.97				
Phase = Replication				
23andme	42.1%	42.1%	0.97	[0.95; 1.00]
CAMP	0.2%	0.2%	1.02	[0.74; 1.40]
deCODE	4.3%	4.3%	0.95	[0.89; 1.02]
ECRHS	0.3%	0.3%	0.88	[0.70; 1.11]
FINNTWIN	0.7%	0.7%	1.00	[0.85; 1.17]
GENEVA−KORA	0.4%	0.4%	0.91	[0.73; 1.13]
GOA	0.1%	0.1%	1.21	[0.70; 2.06]
SAPALDIA	0.3%	0.3%	0.83	[0.66; 1.04]
UK Biobank	11.1%	11.1%	0.91	[0.87; 0.95]
Fixed effect model	**59.5%**	**--**	**0.96**	**[0.94; 0.98]**
Random effects model	**--**	**59.5%**	**0.95**	**[0.91; 0.98]**
Heterogeneity: I^2 = 37%, p = 0.12				
Fixed effect model	**100.0%**	**--**	**0.95**	**[0.94; 0.96]**
Random effects model	**--**	**100.0%**	**0.95**	**[0.94; 0.96]**
Heterogeneity: I^2 = 0%, p = 0.60				

rs2461475[C]



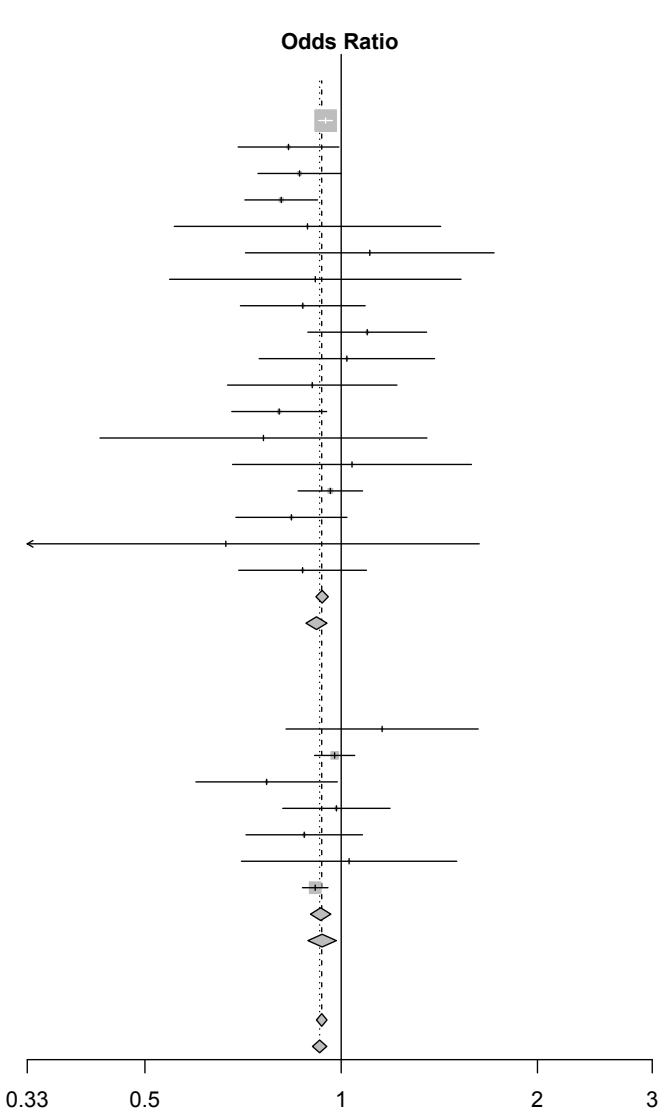
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	32.7%	32.7%	1.03	[1.02; 1.05]
AAGC(QIMR)	0.7%	0.7%	1.10	[0.98; 1.24]
ALSPAC	1.0%	1.0%	1.03	[0.94; 1.14]
B58C	1.3%	1.3%	1.09	[1.00; 1.19]
BAMSE	0.1%	0.1%	1.16	[0.87; 1.55]
COPSAC2000	0.1%	0.1%	1.05	[0.80; 1.40]
EVE (CAG/CSGA)	0.1%	0.1%	1.07	[0.75; 1.51]
EVE (CHS)	0.4%	0.4%	1.05	[0.90; 1.22]
GENR	0.4%	0.4%	1.07	[0.92; 1.24]
GENUFAD	0.2%	0.2%	1.02	[0.82; 1.26]
GINI/LISA	0.2%	0.2%	0.95	[0.77; 1.18]
GOYA	0.9%	0.9%	1.07	[0.96; 1.18]
INMA	0.1%	0.1%	0.94	[0.62; 1.43]
MAAS	0.1%	0.1%	0.93	[0.69; 1.24]
NFBC1966	1.3%	1.3%	1.09	[1.00; 1.19]
NTR (adults)	0.7%	0.7%	1.02	[0.91; 1.15]
NTR (children)	0.1%	0.1%	0.82	[0.54; 1.25]
RAINE	0.4%	0.4%	1.05	[0.90; 1.22]
Fixed effect model	**40.8%**	**−−**	**1.04**	**[1.02; 1.05]**
Random effects model	**−−**	**40.8%**	**1.04**	**[1.02; 1.06]**
Heterogeneity: I^2 = 0%, p = 0.97				
Phase = Replication				
23andme	42.7%	42.7%	1.02	[1.01; 1.04]
CAMP	0.2%	0.2%	0.98	[0.78; 1.24]
deCODE	4.0%	4.0%	1.07	[1.02; 1.12]
ECRHS	0.4%	0.4%	0.89	[0.76; 1.04]
FINNTWIN	0.5%	0.5%	1.01	[0.89; 1.16]
GENEVA−KORA	0.5%	0.5%	1.02	[0.88; 1.17]
GOA	0.1%	0.1%	1.08	[0.82; 1.42]
SAPALDIA	0.4%	0.4%	1.11	[0.95; 1.30]
UK Biobank	10.5%	10.5%	1.05	[1.02; 1.09]
Fixed effect model	**59.2%**	**−−**	**1.03**	**[1.02; 1.04]**
Random effects model	**−−**	**59.2%**	**1.04**	**[1.01; 1.06]**
Heterogeneity: I^2 = 25%, p = 0.22				
Fixed effect model	**100.0%**	**−−**	**1.03**	**[1.02; 1.04]**
Random effects model	**−−**	**100.0%**	**1.03**	**[1.02; 1.04]**
Heterogeneity: I^2 = 0%, p = 0.83				

rs6738964[G]



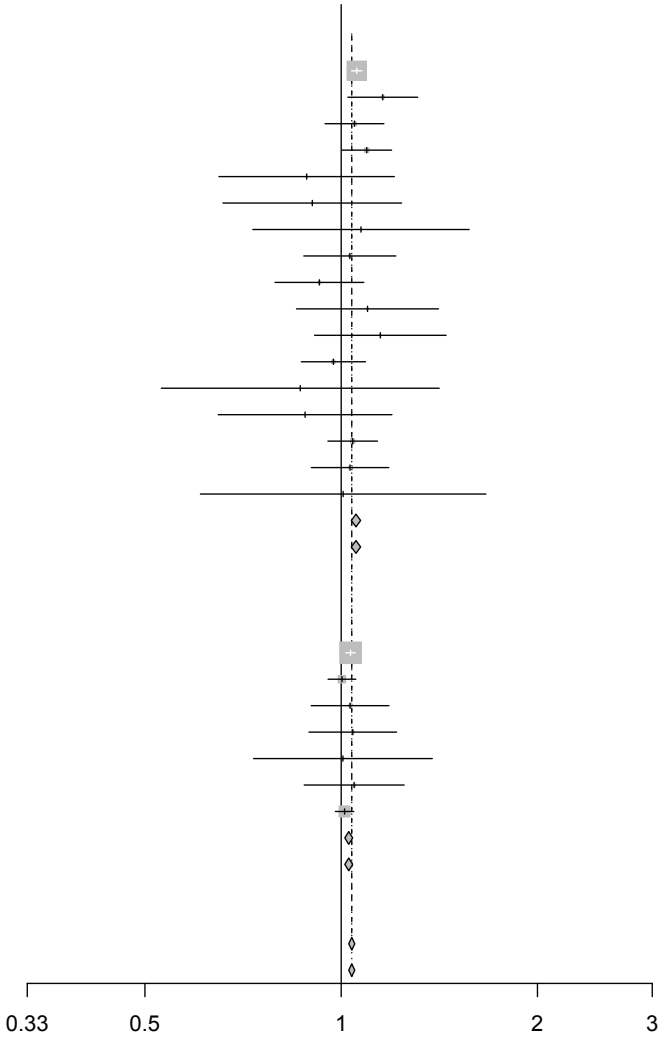
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	33.8%	33.8%	0.96	[0.94; 0.98]
AAGC(QIMR)	0.7%	0.7%	0.92	[0.81; 1.05]
ALSPAC	1.1%	1.1%	1.00	[0.89; 1.11]
B58C	1.4%	1.4%	0.94	[0.85; 1.03]
BAMSE	0.1%	0.1%	0.73	[0.52; 1.02]
COPSAC2000	0.1%	0.1%	0.84	[0.61; 1.17]
EVE (CAG/CSGA)	0.1%	0.1%	1.06	[0.72; 1.55]
EVE (CHS)	0.4%	0.4%	0.94	[0.79; 1.11]
GENR	0.5%	0.5%	0.93	[0.79; 1.09]
GENUFAD	0.2%	0.2%	0.99	[0.78; 1.27]
GINI/LISA	0.2%	0.2%	0.81	[0.64; 1.03]
GOYA	0.9%	0.9%	1.03	[0.92; 1.16]
INMA	0.1%	0.1%	1.45	[0.88; 2.37]
MAAS	0.1%	0.1%	1.01	[0.73; 1.41]
NFBC1966	1.4%	1.4%	0.99	[0.90; 1.09]
NTR (adults)	0.7%	0.7%	0.87	[0.76; 1.00]
NTR (children)	0.1%	0.1%	0.89	[0.56; 1.43]
RAINE	0.4%	0.4%	0.97	[0.82; 1.15]
Fixed effect model	**42.4%**	**−−**	**0.96**	**[0.94; 0.97]**
Random effects model	**−−**	**42.4%**	**0.96**	**[0.94; 0.97]**
Heterogeneity: I^2 = 0%, p = 0.71				
Phase = Replication				
23andme	41.0%	41.0%	0.97	[0.95; 0.99]
CAMP	0.2%	0.2%	0.91	[0.69; 1.19]
deCODE	4.0%	4.0%	0.98	[0.92; 1.03]
ECRHS	0.3%	0.3%	1.01	[0.83; 1.22]
FINNTWIN	0.6%	0.6%	0.99	[0.86; 1.15]
GABRIELA	0.1%	0.1%	1.19	[0.87; 1.61]
GENEVA−KORA	0.5%	0.5%	0.99	[0.84; 1.16]
GOA	0.1%	0.1%	0.98	[0.72; 1.33]
SAPALDIA	0.4%	0.4%	0.85	[0.71; 1.03]
UK Biobank	10.4%	10.4%	0.97	[0.93; 1.00]
Fixed effect model	**57.6%**	**−−**	**0.97**	**[0.96; 0.98]**
Random effects model	**−−**	**57.6%**	**0.97**	**[0.96; 0.98]**
Heterogeneity: I^2 = 0%, p = 0.91				
Fixed effect model	**100.0%**	**−−**	**0.96**	**[0.95; 0.97]**
Random effects model	**−−**	**100.0%**	**0.96**	**[0.95; 0.97]**
Heterogeneity: I^2 = 0%, p = 0.87				

rs10519067[A]



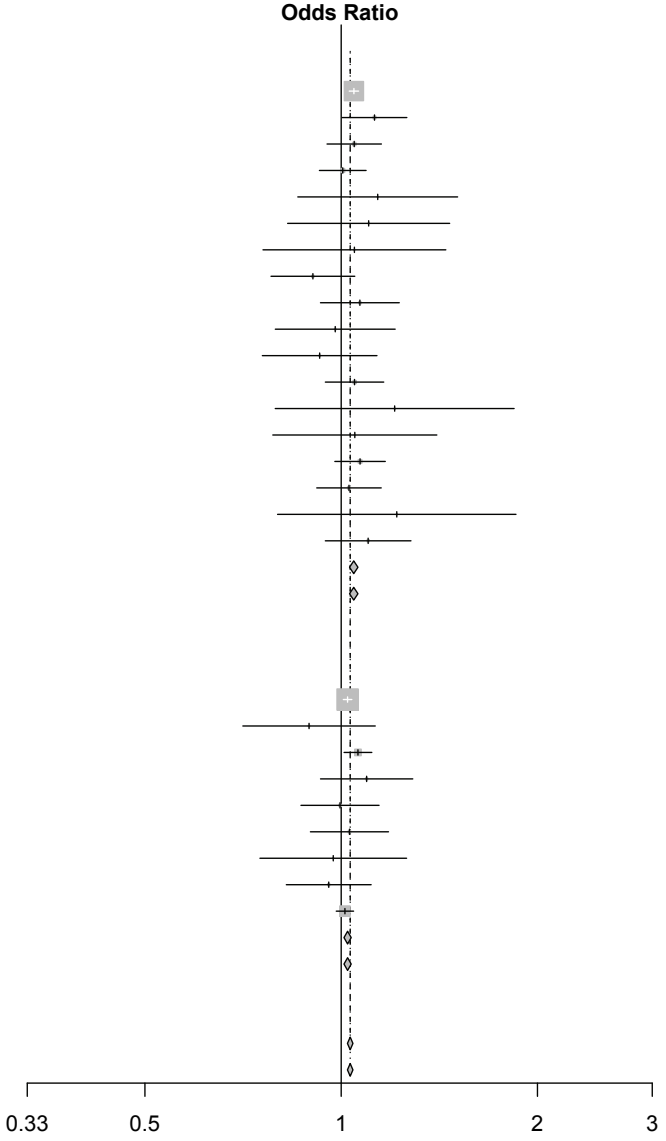
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	59.4%	39.7%	0.95	[0.92; 0.97]
AAGC(QIMR)	1.1%	2.0%	0.83	[0.69; 0.99]
ALSPAC	1.6%	2.8%	0.86	[0.75; 1.00]
B58C	2.1%	3.7%	0.81	[0.71; 0.92]
BAMSE	0.2%	0.3%	0.89	[0.55; 1.42]
COPSAC2000	0.2%	0.3%	1.11	[0.71; 1.72]
EVE (CAG/CSGA)	0.1%	0.2%	0.91	[0.55; 1.53]
EVE (CHS)	0.7%	1.3%	0.87	[0.70; 1.09]
GENR	0.8%	1.4%	1.10	[0.89; 1.35]
GENUFAD	0.4%	0.7%	1.02	[0.75; 1.39]
GINI/LISA	0.4%	0.7%	0.90	[0.67; 1.22]
GOYA	1.2%	2.2%	0.80	[0.68; 0.95]
INMA	0.1%	0.2%	0.76	[0.43; 1.35]
MAAS	0.2%	0.4%	1.04	[0.68; 1.58]
NFBC1966	2.7%	4.6%	0.96	[0.86; 1.08]
NTR (adults)	0.9%	1.6%	0.84	[0.69; 1.02]
NTR (children)	0.0%	0.1%	0.67	[0.27; 1.63]
RAINE	0.7%	1.2%	0.87	[0.70; 1.09]
Fixed effect model	**72.9%**	**−−**	**0.93**	**[0.91; 0.96]**
Random effects model	**−−**	**63.3%**	**0.92**	**[0.88; 0.95]**
Heterogeneity: I^2 = 8%, p = 0.36				
Phase = Replication				
CAMP	0.3%	0.6%	1.16	[0.82; 1.62]
deCODE	6.9%	10.5%	0.98	[0.91; 1.05]
ECRHS	0.6%	1.0%	0.77	[0.60; 0.99]
FINNTWIN	1.0%	1.7%	0.98	[0.81; 1.19]
GENEVA−KORA	0.8%	1.5%	0.88	[0.71; 1.08]
GOA	0.2%	0.4%	1.03	[0.70; 1.50]
UK Biobank	17.3%	21.0%	0.91	[0.87; 0.95]
Fixed effect model	**27.1%**	**−−**	**0.93**	**[0.90; 0.96]**
Random effects model	**−−**	**36.7%**	**0.93**	**[0.89; 0.98]**
Heterogeneity: I^2 = 18%, p = 0.30				
Fixed effect model	**100.0%**	**−−**	**0.93**	**[0.92; 0.95]**
Random effects model	**−−**	**100.0%**	**0.93**	**[0.90; 0.95]**
Heterogeneity: I^2 = 7%, p = 0.36				

rs138050288[−]



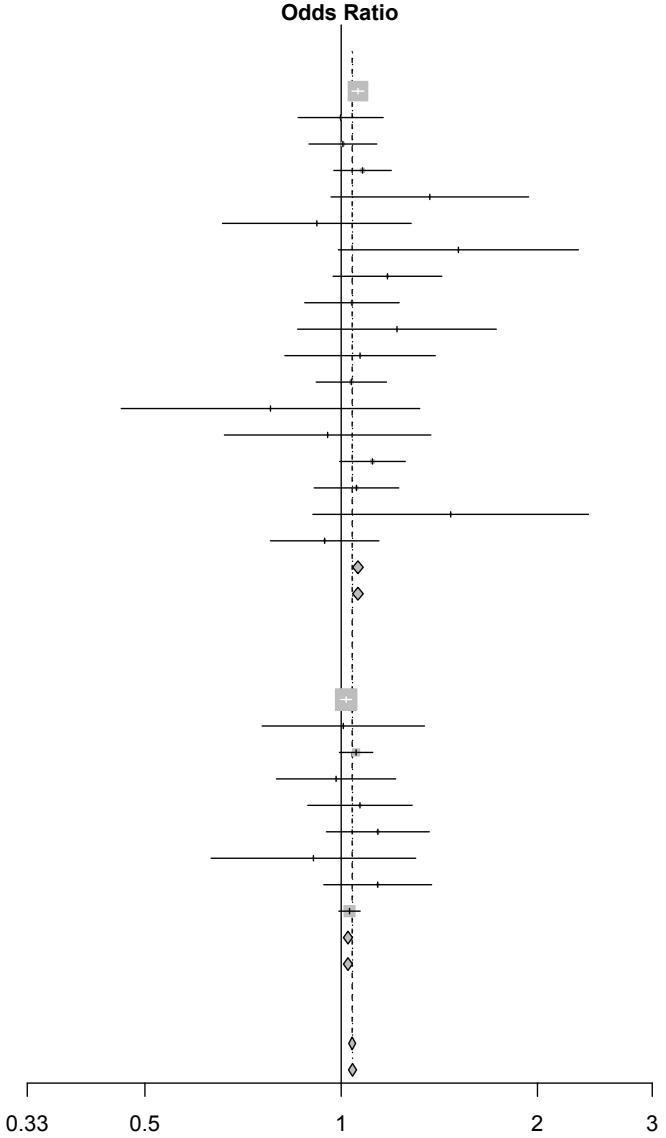
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	33.6%	33.6%	1.06	[1.04; 1.08]
AAGC(QIMR)	0.7%	0.7%	1.16	[1.02; 1.31]
ALSPAC	1.1%	1.1%	1.05	[0.94; 1.16]
B58C	1.5%	1.5%	1.09	[1.00; 1.19]
BAMSE	0.1%	0.1%	0.89	[0.65; 1.21]
COPSAC2000	0.1%	0.1%	0.90	[0.66; 1.24]
EVE (CAG/CSGA)	0.1%	0.1%	1.07	[0.73; 1.57]
EVE (CHS)	0.4%	0.4%	1.03	[0.88; 1.21]
GENR	0.5%	0.5%	0.93	[0.79; 1.08]
GENUFAD	0.2%	0.2%	1.10	[0.85; 1.41]
GINI/LISA	0.2%	0.2%	1.15	[0.91; 1.45]
GOYA	0.9%	0.9%	0.97	[0.87; 1.09]
INMA	0.0%	0.0%	0.87	[0.53; 1.41]
MAAS	0.1%	0.1%	0.88	[0.65; 1.20]
NFBC1966	1.5%	1.5%	1.04	[0.95; 1.14]
NTR (adults)	0.6%	0.6%	1.03	[0.90; 1.18]
NTR (children)	0.0%	0.0%	1.01	[0.61; 1.67]
Fixed effect model	**41.7%**	**−−**	**1.05**	**[1.04; 1.07]**
Random effects model	**−−**	**41.7%**	**1.05**	**[1.04; 1.07]**
Heterogeneity: I^2 = 0%, p = 0.71				
Phase = Replication				
23andme	41.4%	41.4%	1.03	[1.02; 1.05]
deCODE	4.7%	4.7%	1.00	[0.95; 1.05]
FINNTWIN	0.6%	0.6%	1.03	[0.90; 1.18]
GENEVA−KORA	0.5%	0.5%	1.04	[0.89; 1.22]
GOA	0.1%	0.1%	1.01	[0.73; 1.38]
SAPALDIA	0.4%	0.4%	1.05	[0.88; 1.25]
UK Biobank	10.7%	10.7%	1.01	[0.98; 1.05]
Fixed effect model	**58.3%**	**−−**	**1.03**	**[1.01; 1.04]**
Random effects model	**−−**	**58.3%**	**1.03**	**[1.01; 1.04]**
Heterogeneity: I^2 = 0%, p = 0.89				
Fixed effect model	**100.0%**	**−−**	**1.04**	**[1.03; 1.05]**
Random effects model	**−−**	**100.0%**	**1.04**	**[1.03; 1.05]**
Heterogeneity: I^2 = 0%, p = 0.63				

rs7328203[G]



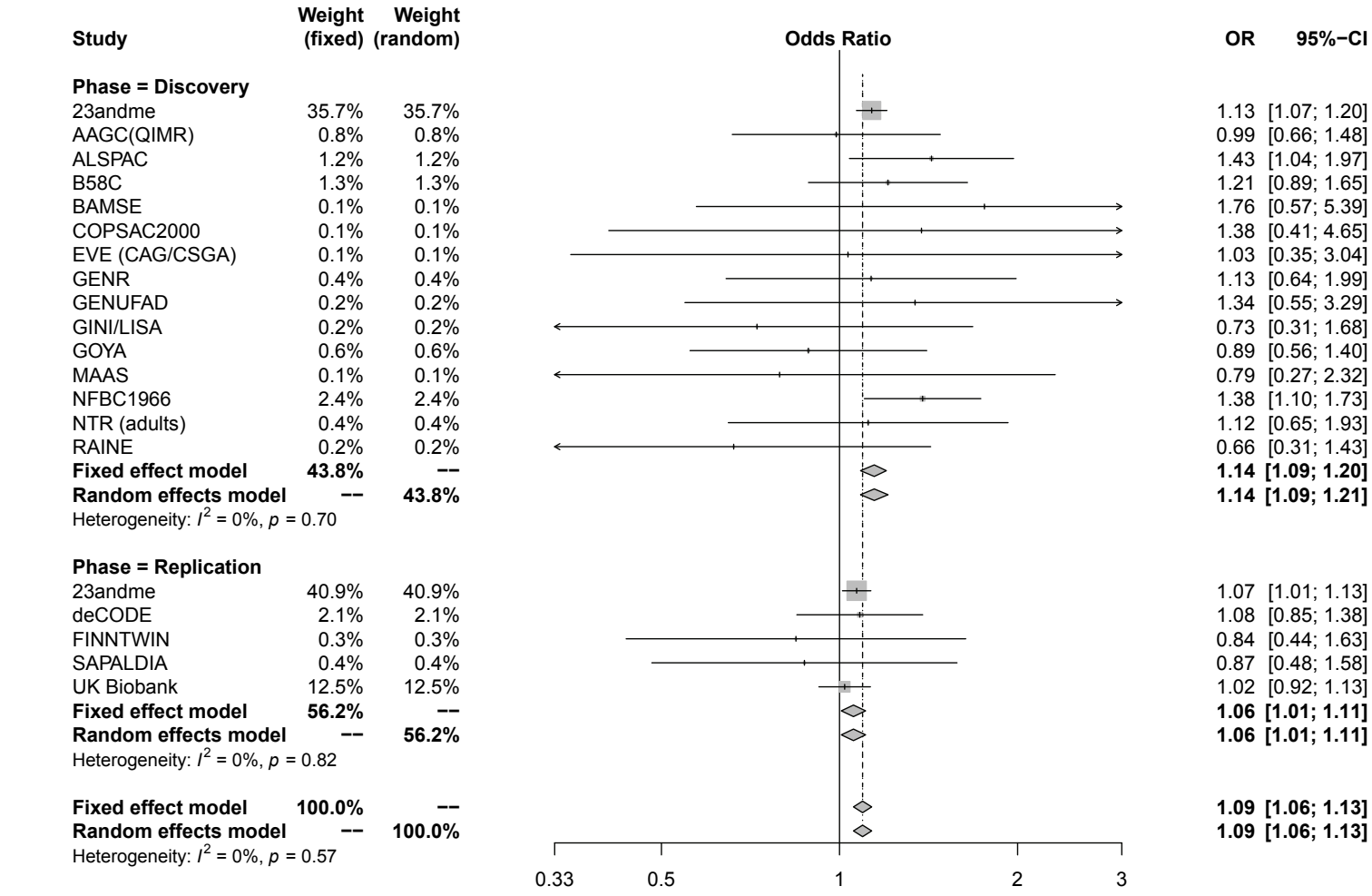
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	34.1%	34.1%	1.05	[1.03; 1.06]
AAGC(QIMR)	0.7%	0.7%	1.12	[1.00; 1.26]
ALSPAC	1.0%	1.0%	1.05	[0.95; 1.15]
B58C	1.4%	1.4%	1.01	[0.93; 1.09]
BAMSE	0.1%	0.1%	1.14	[0.86; 1.51]
COPSAC2000	0.1%	0.1%	1.10	[0.83; 1.47]
EVE (CAG/CSGA)	0.1%	0.1%	1.05	[0.76; 1.45]
EVE (CHS)	0.4%	0.4%	0.90	[0.78; 1.05]
GENR	0.5%	0.5%	1.07	[0.93; 1.23]
GENUFAD	0.2%	0.2%	0.98	[0.79; 1.21]
GINI/LISA	0.2%	0.2%	0.93	[0.76; 1.13]
GOYA	0.9%	0.9%	1.05	[0.95; 1.16]
INMA	0.1%	0.1%	1.21	[0.79; 1.84]
MAAS	0.1%	0.1%	1.05	[0.79; 1.40]
NFBC1966	1.2%	1.2%	1.07	[0.98; 1.17]
NTR (adults)	0.7%	0.7%	1.03	[0.92; 1.15]
NTR (children)	0.1%	0.1%	1.22	[0.80; 1.85]
RAINE	0.4%	0.4%	1.10	[0.95; 1.28]
Fixed effect model	**42.4%**	**−−**	**1.05**	**[1.03; 1.06]**
Random effects model	**−−**	**42.4%**	**1.05**	**[1.03; 1.06]**
Heterogeneity: I^2 = 0%, p = 0.90				
Phase = Replication				
23andme	41.6%	41.6%	1.02	[1.01; 1.04]
CAMP	0.2%	0.2%	0.89	[0.71; 1.13]
deCODE	3.9%	3.9%	1.06	[1.01; 1.11]
ECRHS	0.4%	0.4%	1.09	[0.93; 1.29]
FINNTWIN	0.5%	0.5%	1.00	[0.87; 1.14]
GENEVA−KORA	0.5%	0.5%	1.03	[0.90; 1.18]
GOA	0.1%	0.1%	0.97	[0.75; 1.26]
SAPALDIA	0.4%	0.4%	0.96	[0.82; 1.11]
UK Biobank	10.0%	10.0%	1.01	[0.98; 1.04]
Fixed effect model	**57.6%**	**−−**	**1.02**	**[1.01; 1.04]**
Random effects model	**−−**	**57.6%**	**1.02**	**[1.01; 1.04]**
Heterogeneity: I^2 = 0%, p = 0.70				
Fixed effect model	**100.0%**	**−−**	**1.03**	**[1.02; 1.04]**
Random effects model	**−−**	**100.0%**	**1.03**	**[1.02; 1.04]**
Heterogeneity: I^2 = 0%, p = 0.78				

rs11169225[A]

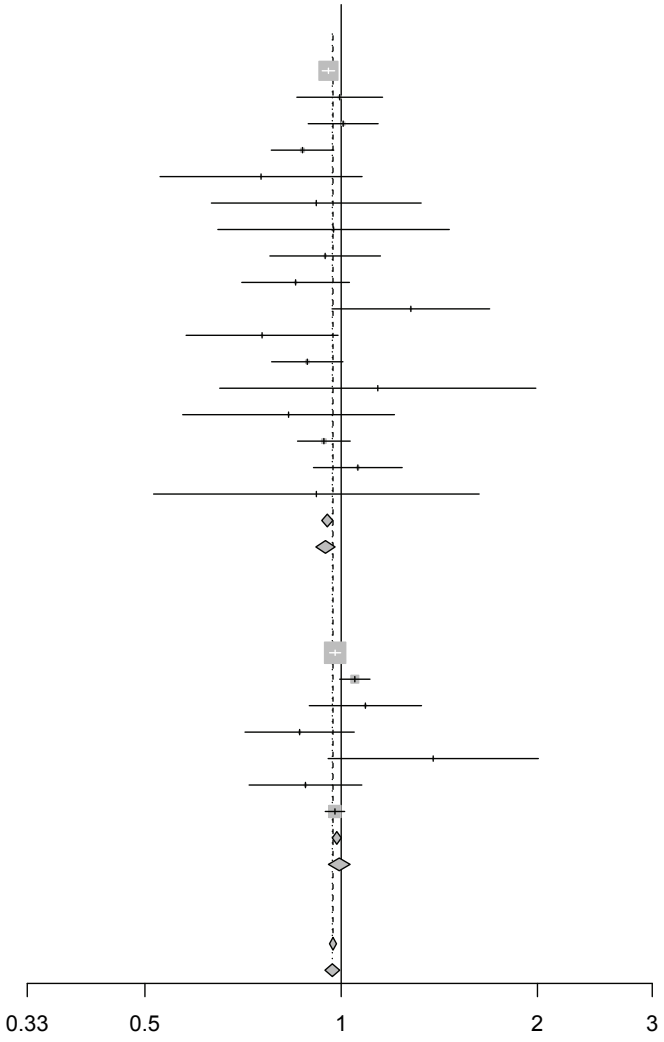


Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	33.9%	32.0%	1.06	[1.04; 1.08]
AAGC(QIMR)	0.7%	0.9%	1.00	[0.86; 1.16]
ALSPAC	1.0%	1.4%	1.01	[0.89; 1.13]
B58C	1.5%	1.9%	1.08	[0.97; 1.19]
BAMSE	0.1%	0.2%	1.37	[0.96; 1.94]
COPSAC2000	0.1%	0.2%	0.92	[0.66; 1.28]
EVE (CAG/CSGA)	0.1%	0.1%	1.51	[0.99; 2.31]
EVE (CHS)	0.4%	0.5%	1.18	[0.97; 1.43]
GENR	0.5%	0.7%	1.04	[0.88; 1.23]
GENUFAD	0.1%	0.2%	1.22	[0.86; 1.73]
GINI/LISA	0.2%	0.3%	1.07	[0.82; 1.39]
GOYA	1.0%	1.3%	1.04	[0.92; 1.17]
INMA	0.1%	0.1%	0.78	[0.46; 1.32]
MAAS	0.1%	0.2%	0.95	[0.66; 1.37]
NFBC1966	1.1%	1.5%	1.12	[0.99; 1.25]
NTR (adults)	0.7%	0.9%	1.06	[0.91; 1.23]
NTR (children)	0.1%	0.1%	1.47	[0.90; 2.39]
RAINE	0.4%	0.5%	0.94	[0.78; 1.14]
Fixed effect model	**42.1%**	**−−**	**1.06**	**[1.04; 1.08]**
Random effects model	**−−**	**42.9%**	**1.06**	**[1.04; 1.08]**
Heterogeneity: I^2 = 0%, p = 0.63				
Phase = Replication				
23andme	41.0%	36.4%	1.02	[1.00; 1.04]
CAMP	0.2%	0.2%	1.01	[0.76; 1.34]
deCODE	4.3%	5.5%	1.05	[0.99; 1.12]
ECRHS	0.3%	0.5%	0.98	[0.80; 1.21]
FINNTWIN	0.4%	0.6%	1.07	[0.89; 1.28]
GENEVA−KORA	0.5%	0.6%	1.14	[0.95; 1.37]
GOA	0.1%	0.2%	0.91	[0.63; 1.30]
SAPALDIA	0.4%	0.6%	1.14	[0.94; 1.38]
UK Biobank	10.6%	12.6%	1.03	[0.99; 1.07]
Fixed effect model	**57.9%**	**−−**	**1.02**	**[1.01; 1.04]**
Random effects model	**−−**	**57.1%**	**1.02**	**[1.01; 1.04]**
Heterogeneity: I^2 = 0%, p = 0.79				
Fixed effect model	**100.0%**	**−−**	**1.04**	**[1.03; 1.05]**
Random effects model	**−−**	**100.0%**	**1.04**	**[1.03; 1.06]**
Heterogeneity: I^2 = 3%, p = 0.42				

rs141023293[A]

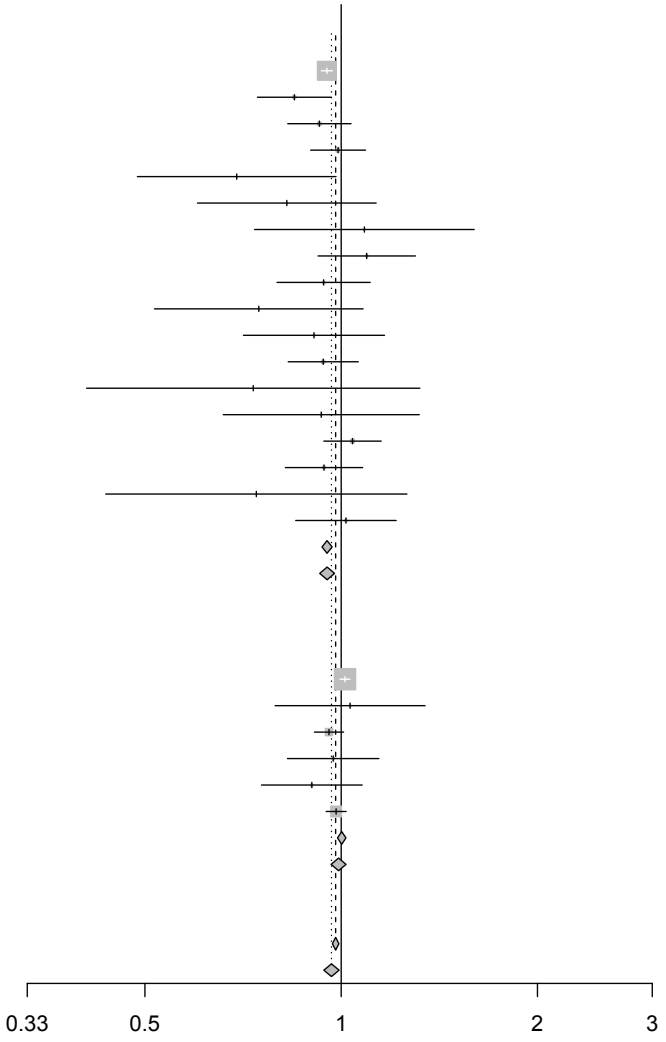


rs201428899[G]



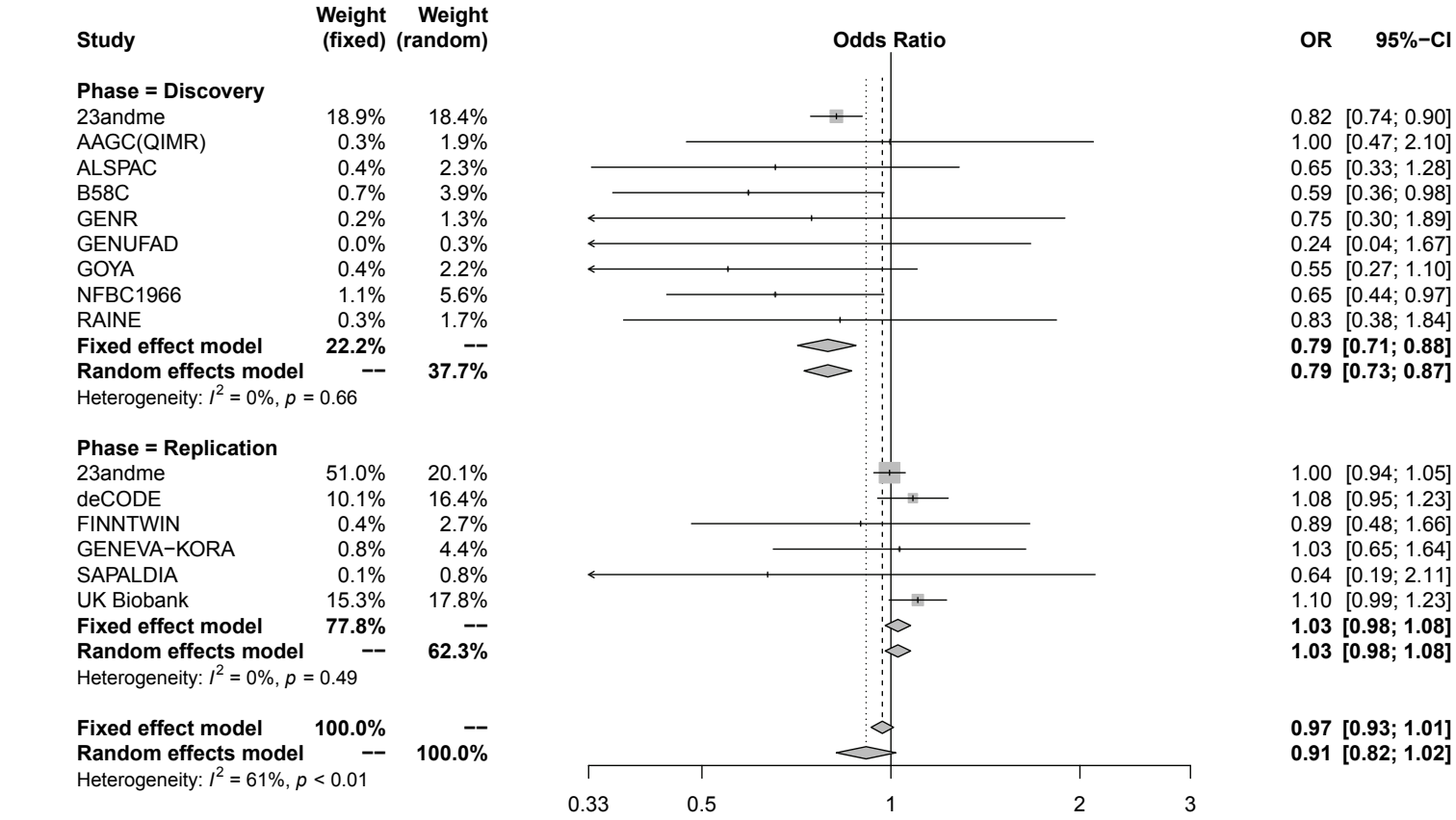
Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	32.0%	18.6%	0.96	[0.94; 0.98]
AAGC(QIMR)	0.7%	2.6%	0.99	[0.86; 1.16]
ALSPAC	1.0%	3.7%	1.01	[0.89; 1.14]
B58C	1.3%	4.4%	0.87	[0.78; 0.97]
BAMSE	0.1%	0.5%	0.75	[0.53; 1.08]
COPSAC2000	0.1%	0.5%	0.92	[0.63; 1.33]
EVE (CAG/CSGA)	0.1%	0.4%	0.97	[0.65; 1.46]
EVE (CHS)	0.4%	1.6%	0.94	[0.78; 1.15]
GENR	0.4%	1.7%	0.85	[0.70; 1.03]
GENUFAD	0.2%	0.8%	1.28	[0.97; 1.69]
GINI/LISA	0.2%	0.9%	0.76	[0.58; 0.99]
GOYA	1.0%	3.6%	0.89	[0.78; 1.01]
INMA	0.0%	0.2%	1.14	[0.65; 1.99]
MAAS	0.1%	0.5%	0.83	[0.57; 1.21]
NFBC1966	1.8%	5.8%	0.94	[0.86; 1.03]
NTR (adults)	0.6%	2.5%	1.06	[0.91; 1.24]
NTR (children)	0.0%	0.2%	0.92	[0.52; 1.63]
Fixed effect model	**40.0%**	**−−**	**0.95**	**[0.93; 0.97]**
Random effects model	**−−**	**48.6%**	**0.95**	**[0.91; 0.98]**
Heterogeneity: I^2 = 10%, p = 0.33				
Phase = Replication				
23andme	40.2%	19.1%	0.98	[0.96; 1.00]
deCODE	5.4%	11.3%	1.05	[0.99; 1.11]
ECRHS	0.4%	1.6%	1.09	[0.89; 1.33]
GENEVA−KORA	0.4%	1.7%	0.86	[0.71; 1.05]
GOA	0.1%	0.5%	1.38	[0.95; 2.01]
SAPALDIA	0.4%	1.6%	0.88	[0.72; 1.08]
UK Biobank	13.0%	15.6%	0.98	[0.95; 1.01]
Fixed effect model	**60.0%**	**−−**	**0.98**	**[0.97; 1.00]**
Random effects model	**−−**	**51.4%**	**0.99**	**[0.96; 1.03]**
Heterogeneity: I^2 = 54%, p = 0.04				
Fixed effect model	**100.0%**	**−−**	**0.97**	**[0.96; 0.98]**
Random effects model	**−−**	**100.0%**	**0.97**	**[0.94; 0.99]**
Heterogeneity: I^2 = 39%, p = 0.03				

rs6011016[C]

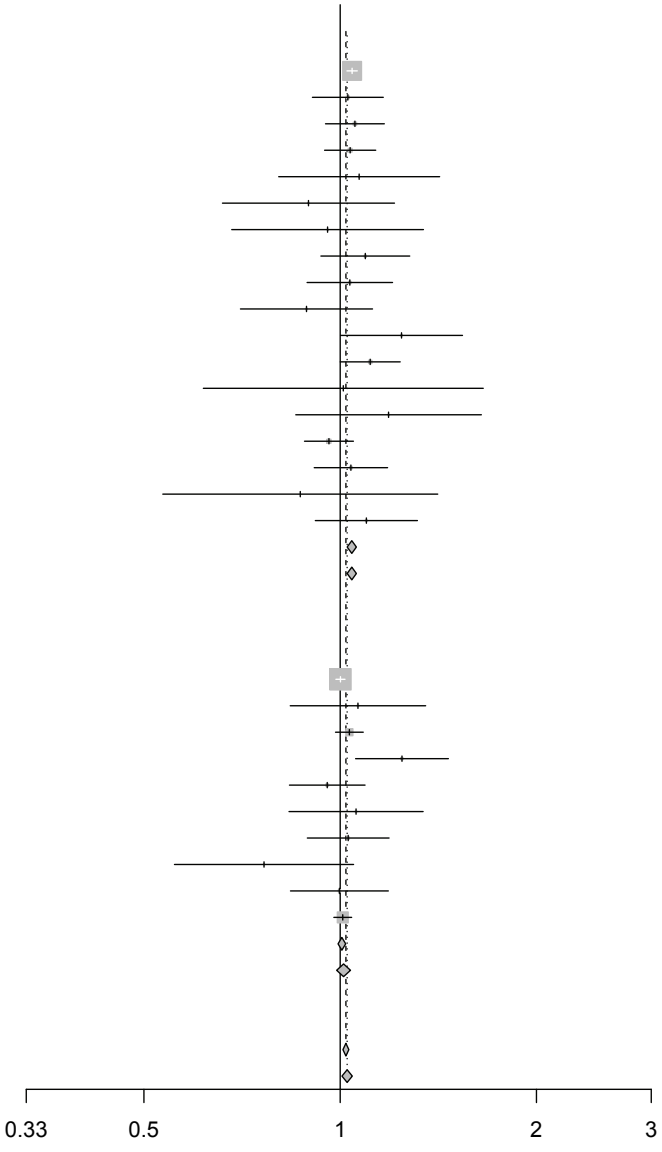


Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	33.2%	16.6%	0.95	[0.93; 0.97]
AAGC(QIMR)	0.8%	3.5%	0.85	[0.74; 0.97]
ALSPAC	1.1%	4.4%	0.93	[0.83; 1.04]
B58C	1.4%	5.4%	0.99	[0.90; 1.09]
BAMSE	0.1%	0.6%	0.69	[0.49; 0.98]
COPSAC2000	0.1%	0.7%	0.83	[0.60; 1.13]
EVE (CAG/CSGA)	0.1%	0.5%	1.09	[0.74; 1.60]
EVE (CHS)	0.4%	2.2%	1.09	[0.92; 1.30]
GENR	0.5%	2.3%	0.94	[0.80; 1.11]
GENUFAD	0.1%	0.5%	0.75	[0.52; 1.08]
GINI/LISA	0.2%	1.1%	0.91	[0.71; 1.17]
GOYA	0.9%	3.8%	0.94	[0.83; 1.06]
INMA	0.0%	0.2%	0.73	[0.41; 1.32]
MAAS	0.1%	0.6%	0.93	[0.66; 1.32]
NFBC1966	1.3%	5.1%	1.04	[0.94; 1.15]
NTR (adults)	0.7%	3.2%	0.94	[0.82; 1.08]
NTR (children)	0.0%	0.3%	0.74	[0.44; 1.26]
RAINE	0.4%	2.0%	1.02	[0.85; 1.21]
Fixed effect model	**41.4%**	**−−**	**0.95**	**[0.93; 0.97]**
Random effects model	**−−**	**52.9%**	**0.95**	**[0.93; 0.98]**
Heterogeneity: I^2 = 5%, p = 0.40				
Phase = Replication				
23andme	41.9%	16.9%	1.01	[1.00; 1.03]
CAMP	0.2%	1.0%	1.03	[0.79; 1.34]
deCODE	4.9%	10.9%	0.96	[0.91; 1.01]
FINNTWIN	0.5%	2.4%	0.97	[0.83; 1.14]
GENEVA−KORA	0.4%	2.0%	0.90	[0.75; 1.08]
UK Biobank	10.6%	13.9%	0.98	[0.95; 1.02]
Fixed effect model	**58.6%**	**−−**	**1.00**	**[0.99; 1.02]**
Random effects model	**−−**	**47.1%**	**0.99**	**[0.97; 1.02]**
Heterogeneity: I^2 = 32%, p = 0.20				
Fixed effect model	**100.0%**	**−−**	**0.98**	**[0.97; 0.99]**
Random effects model	**−−**	**100.0%**	**0.97**	**[0.94; 0.99]**
Heterogeneity: I^2 = 48%, p < 0.01				

rs149341190[C]



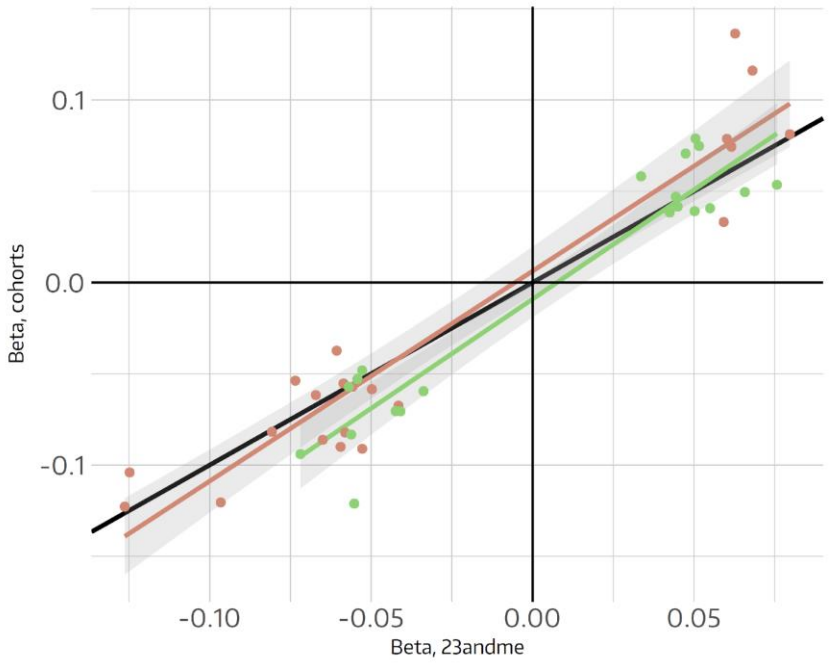
rs10760123[T]



Study	Weight (fixed)	Weight (random)	OR	95%−CI
Phase = Discovery				
23andme	32.5%	21.9%	1.04	[1.02; 1.06]
AAGC(QIMR)	0.7%	2.0%	1.03	[0.91; 1.16]
ALSPAC	1.0%	2.8%	1.05	[0.95; 1.17]
B58C	1.3%	3.6%	1.04	[0.95; 1.13]
BAMSE	0.1%	0.4%	1.07	[0.80; 1.42]
COPSAC2000	0.1%	0.4%	0.89	[0.66; 1.21]
EVE (CAG/CSGA)	0.1%	0.3%	0.96	[0.68; 1.34]
EVE (CHS)	0.4%	1.3%	1.09	[0.93; 1.28]
GENR	0.5%	1.4%	1.03	[0.89; 1.20]
GENUFAD	0.2%	0.6%	0.89	[0.70; 1.12]
GINI/LISA	0.2%	0.7%	1.24	[1.00; 1.54]
GOYA	1.0%	2.7%	1.11	[1.00; 1.24]
INMA	0.0%	0.1%	1.01	[0.62; 1.66]
MAAS	0.1%	0.3%	1.19	[0.85; 1.65]
NFBC1966	1.5%	3.9%	0.96	[0.88; 1.05]
NTR (adults)	0.6%	1.9%	1.04	[0.91; 1.18]
NTR (children)	0.0%	0.1%	0.87	[0.53; 1.41]
RAINE	0.3%	1.0%	1.10	[0.92; 1.31]
Fixed effect model	**40.8%**	**−−**	**1.04**	**[1.02; 1.06]**
Random effects model	**−−**	**45.3%**	**1.04**	**[1.02; 1.06]**
Heterogeneity: I^2 = 0%, p = 0.78				
Phase = Replication				
23andme	41.4%	23.0%	1.00	[0.98; 1.02]
CAMP	0.2%	0.6%	1.06	[0.84; 1.35]
deCODE	4.5%	9.3%	1.03	[0.98; 1.08]
ECRHS	0.4%	1.2%	1.24	[1.06; 1.47]
FINNTWIN	0.6%	1.8%	0.96	[0.84; 1.09]
GABRIELA	0.2%	0.6%	1.06	[0.83; 1.34]
GENEVA−KORA	0.5%	1.5%	1.03	[0.89; 1.19]
GOA	0.1%	0.3%	0.76	[0.56; 1.05]
SAPALDIA	0.4%	1.1%	1.00	[0.84; 1.19]
UK Biobank	10.9%	15.3%	1.01	[0.98; 1.04]
Fixed effect model	**59.2%**	**−−**	**1.01**	**[0.99; 1.02]**
Random effects model	**−−**	**54.7%**	**1.01**	**[0.99; 1.04]**
Heterogeneity: I^2 = 25%, p = 0.22				
Fixed effect model	**100.0%**	**−−**	**1.02**	**[1.01; 1.03]**
Random effects model	**−−**	**100.0%**	**1.02**	**[1.01; 1.04]**
Heterogeneity: I^2 = 22%, p = 0.14				

Supplementary Figure 4

Scatter plot of beta coefficients between 23andme and non-23andme cohorts from the discovery phase of the allergic rhinitis GWAS, from genetic marker association of 16,531,985 genetic markers to allergic rhinitis from the discovery phase, including 212,120 individuals. Only known and replicating novel markers are shown. Red indicates known loci, green indicates novel, replicating loci. The black line represents identity.

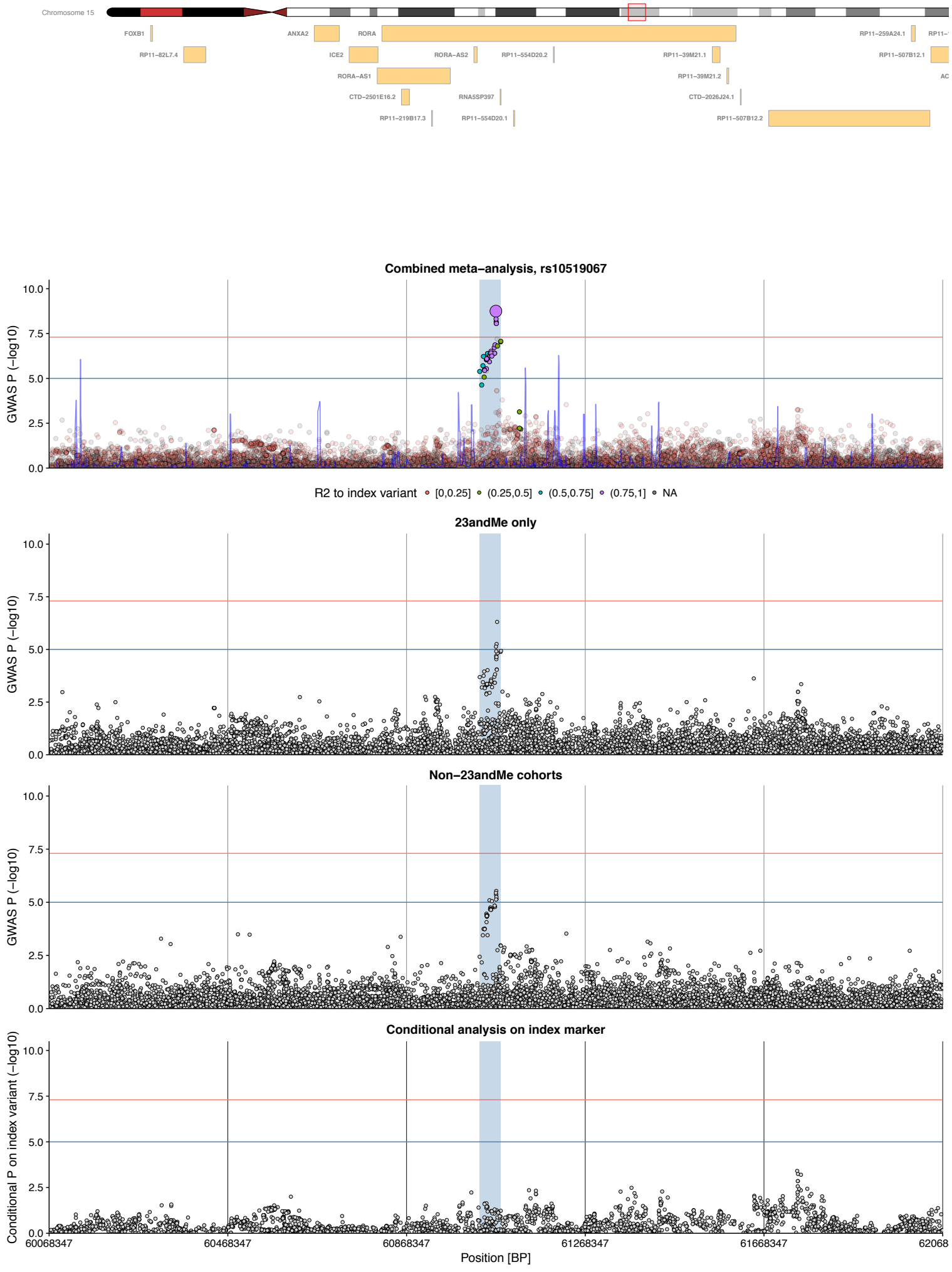


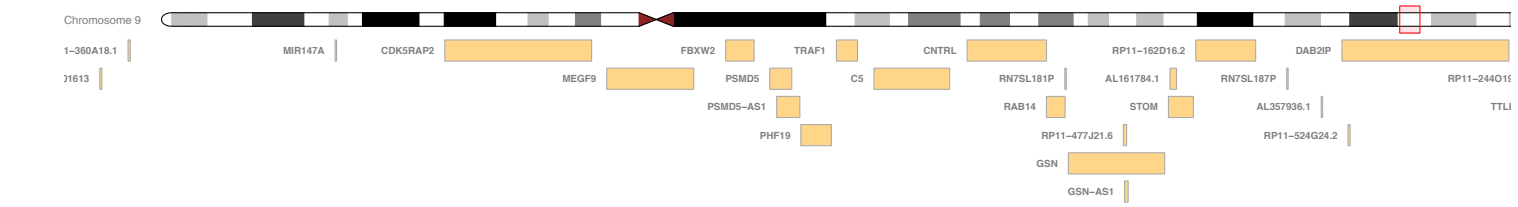
Supplementary Figure 5

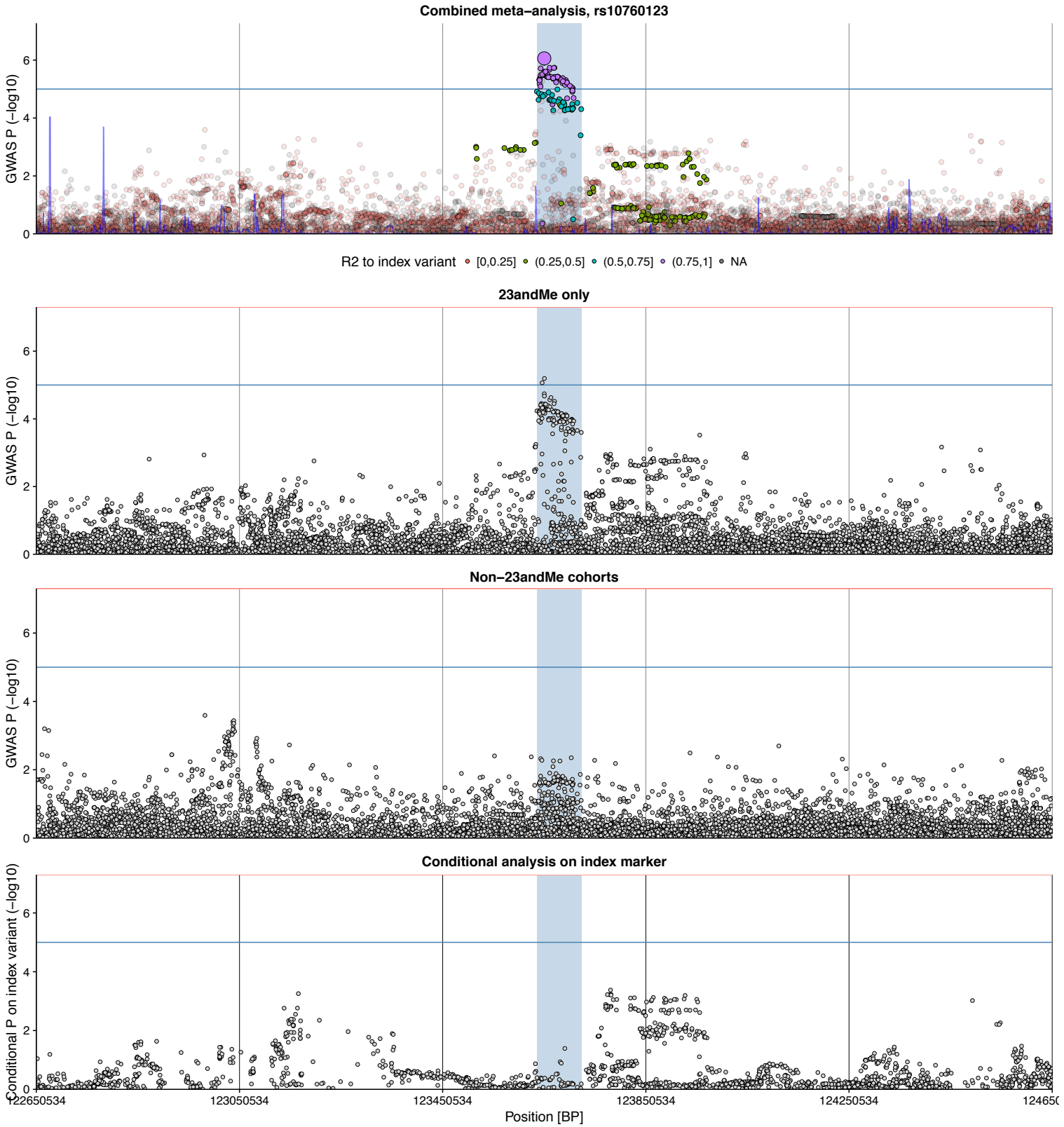
Regional locus plots of allergic rhinitis discovery phase markers. Top panel shows results from the discovery meta-analysis, color coded for 1000g R 2 values. AR discovery markers not present in 1000g are colored in grey ("NA"). Index marker has increased circle size. Locus definition (tag R2 > 0.5 to index) is indicated by blue region. Red line indicates p = 5e-8, blue line indicates p = 1e-5.

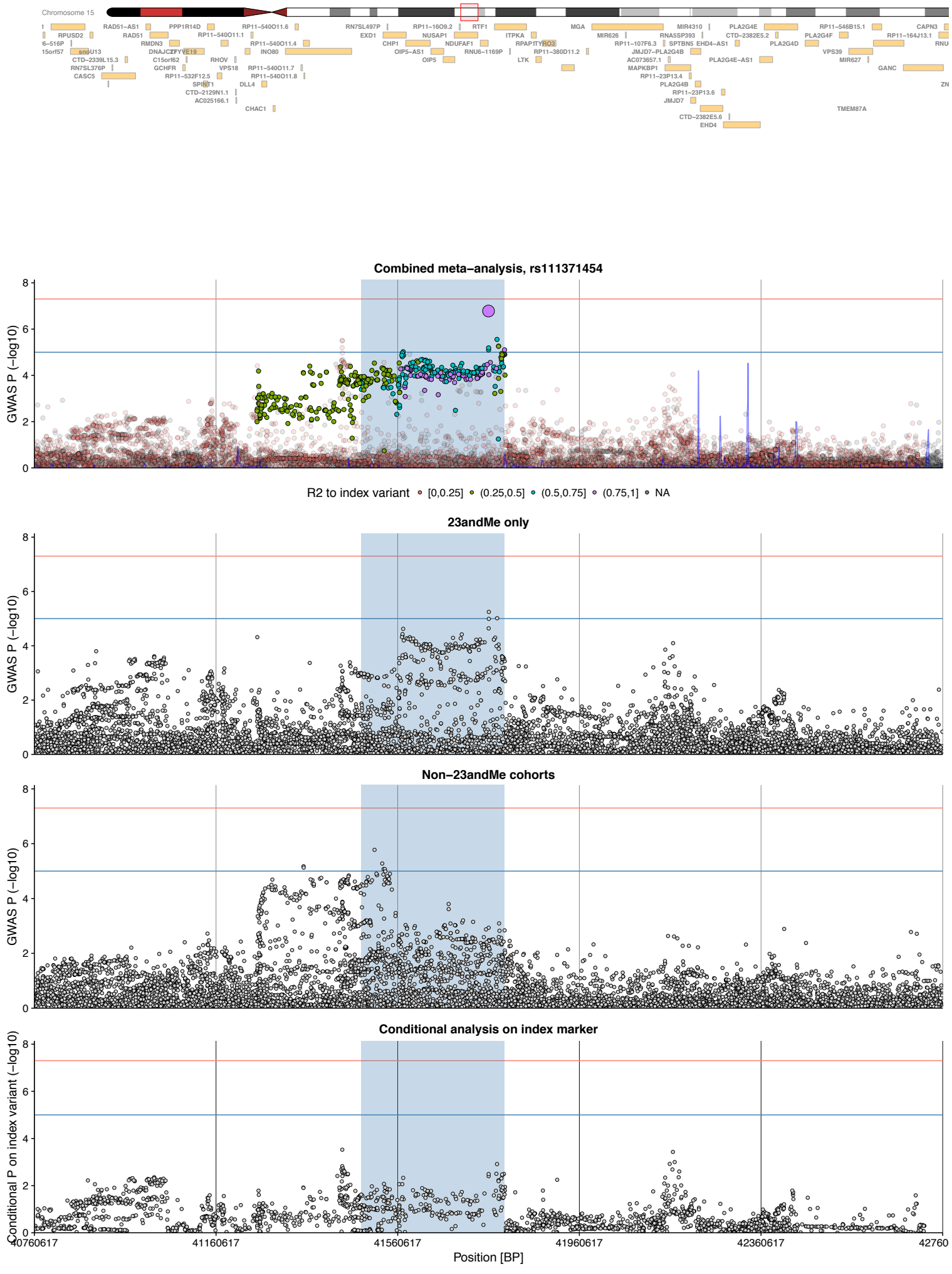
Middle upper panel shows association results for 23andme, and middle lower panel shows association values for non-23andme cohorts.

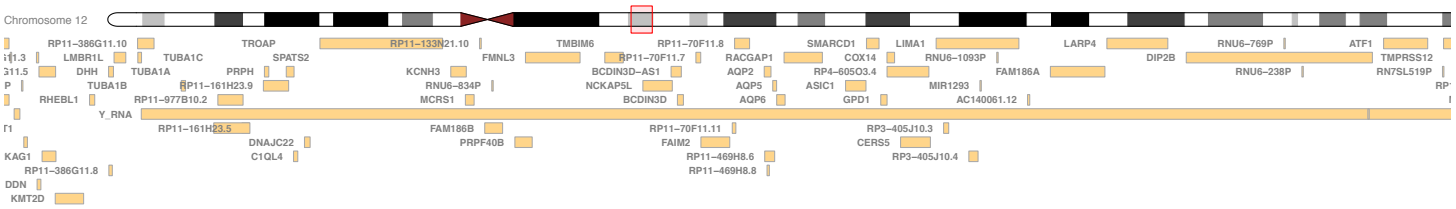
Bottom panel shows results for the conditional analysis on index marker.

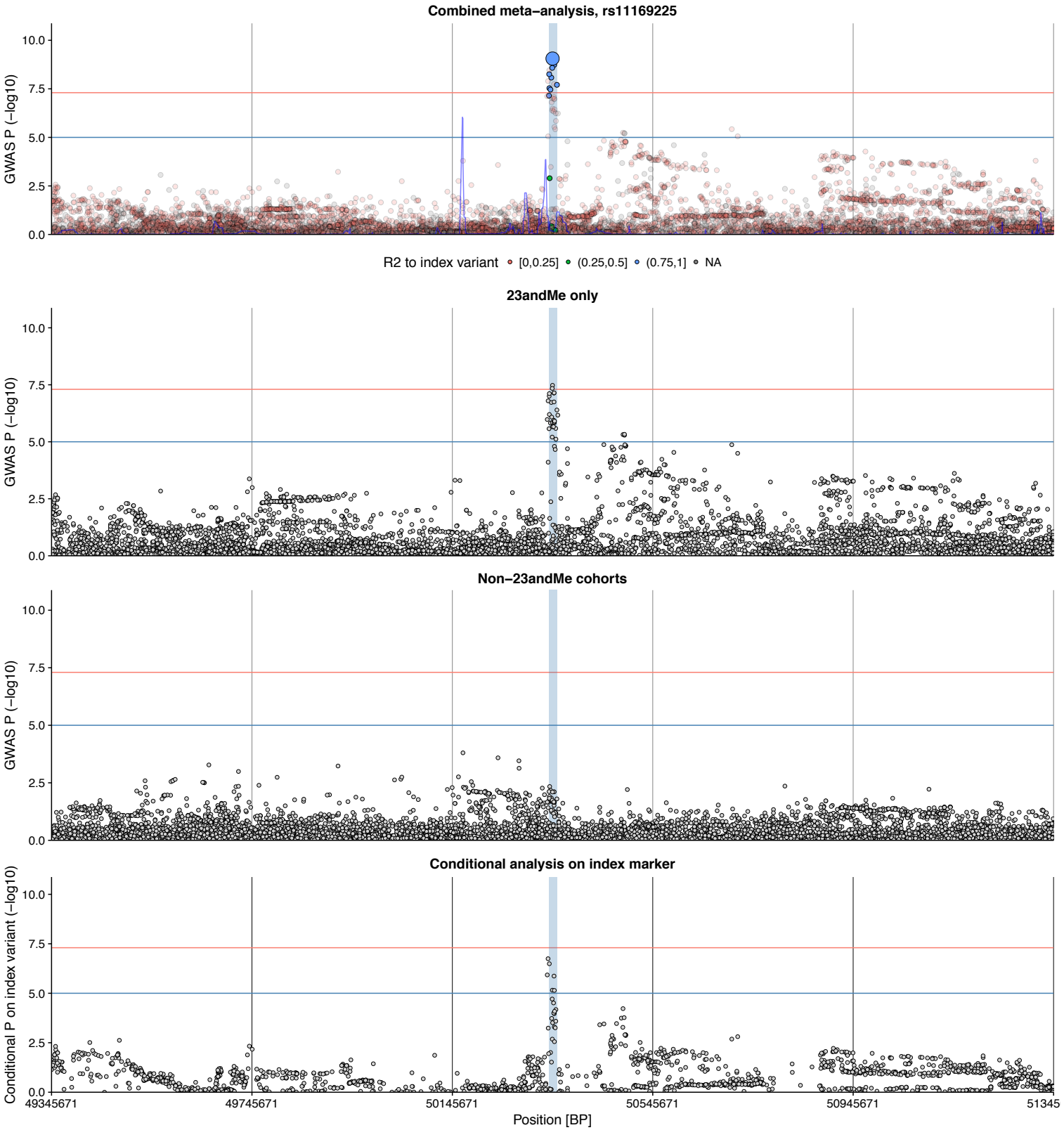

Chromosome 15
FOXB1
ANXA2
RORA
RP11−259A24.1
RP11−82L7.4
ICE2
RORA−AS2
RP11−554D20.2
RP11−39M21.1
RP11−507B12.1
RORA−AS1
RP11−39M21.2
CTD−2501E16.2
RNA5SP397
CTD−2026J24.1
RP11−219B17.3
RP11−554D20.1
RP11−507B12.2
Combined meta−analysis, rs10519067
GWAS P (−log10)
10.0
7.5
5.0
2.5
0.0
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
60068347
60468347
60868347
61268347
61668347
62068
Position [BP]

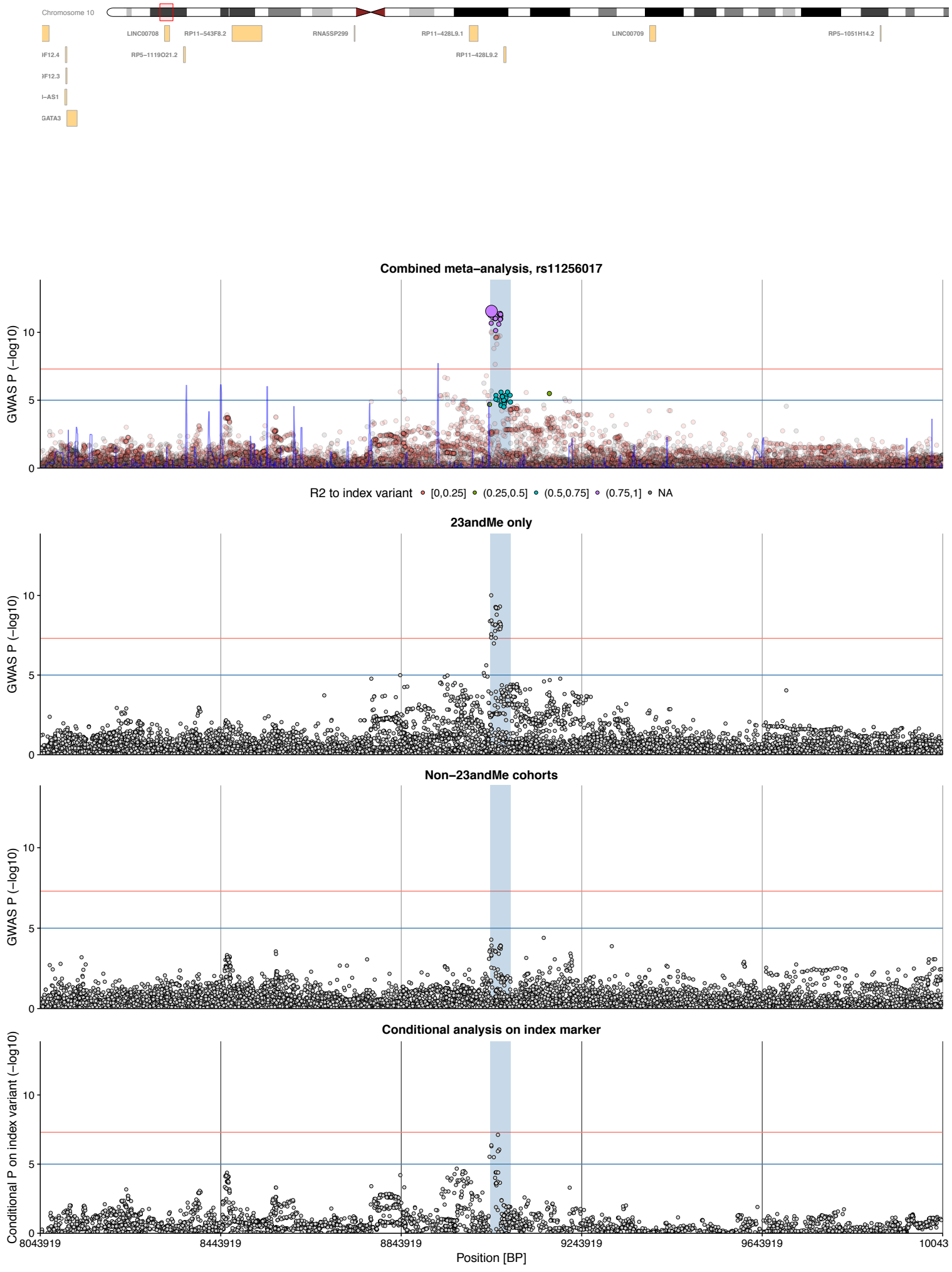

Chromosome 9
1–360A18.1
MIR147A
CDK5RAP2
FBXW2
TRAF1
CNTRL
RP11–162D16.2
DAB2IP
MEGF9
PSMD5
C5
RN7SL181P
AL161784.1
RN7SL187P
PSMD5–AS1
RAB14
STOM
AL357936.1
PHF19
RP11–477J21.6
RP11–524G24.2
GSN
GSN–AS1

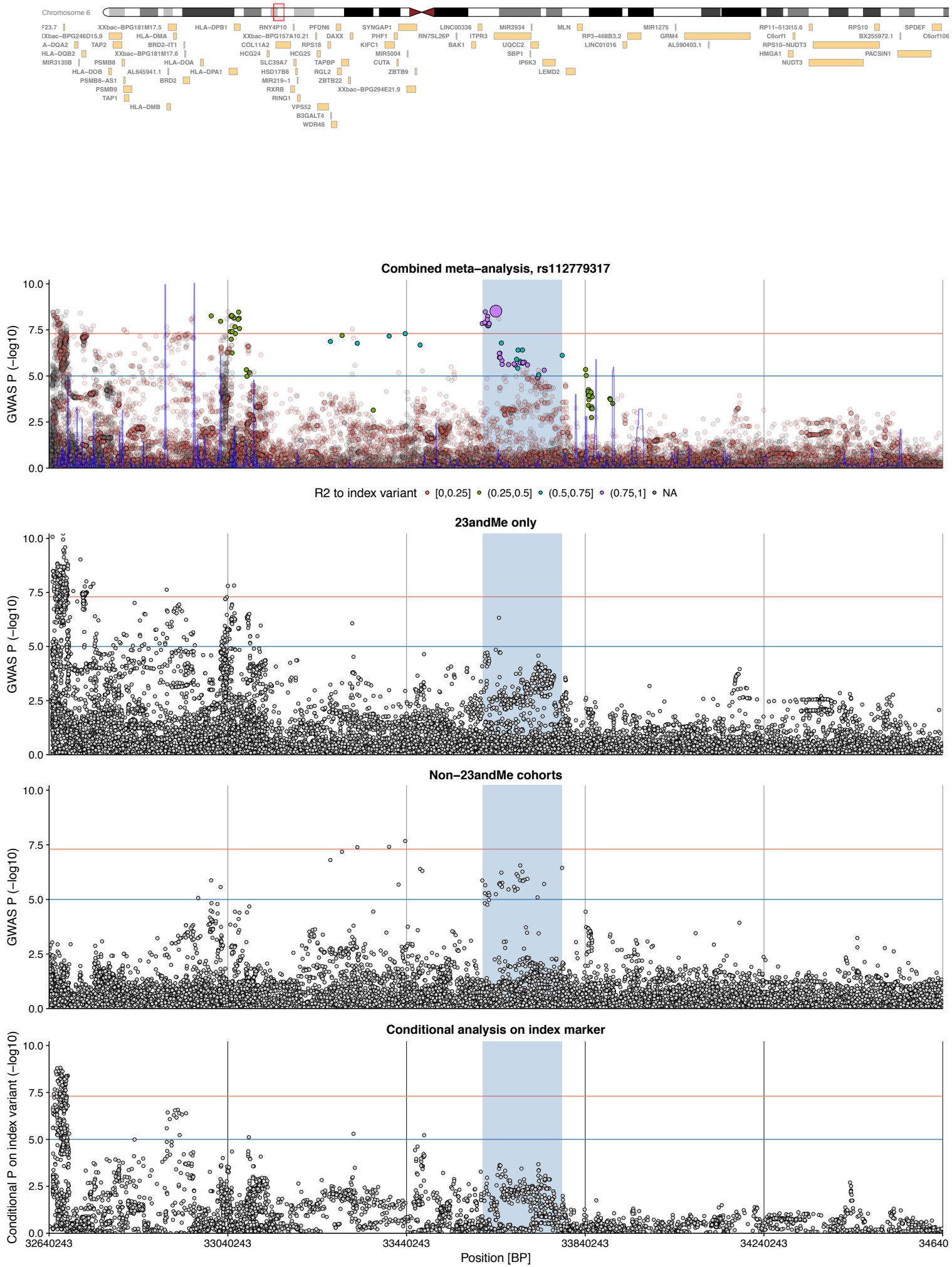

Combined meta–analysis, rs10760123
GWAS P (–log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non–23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (–log10)
122650534
123050534
123450534
123850534
124250534
Position [BP]

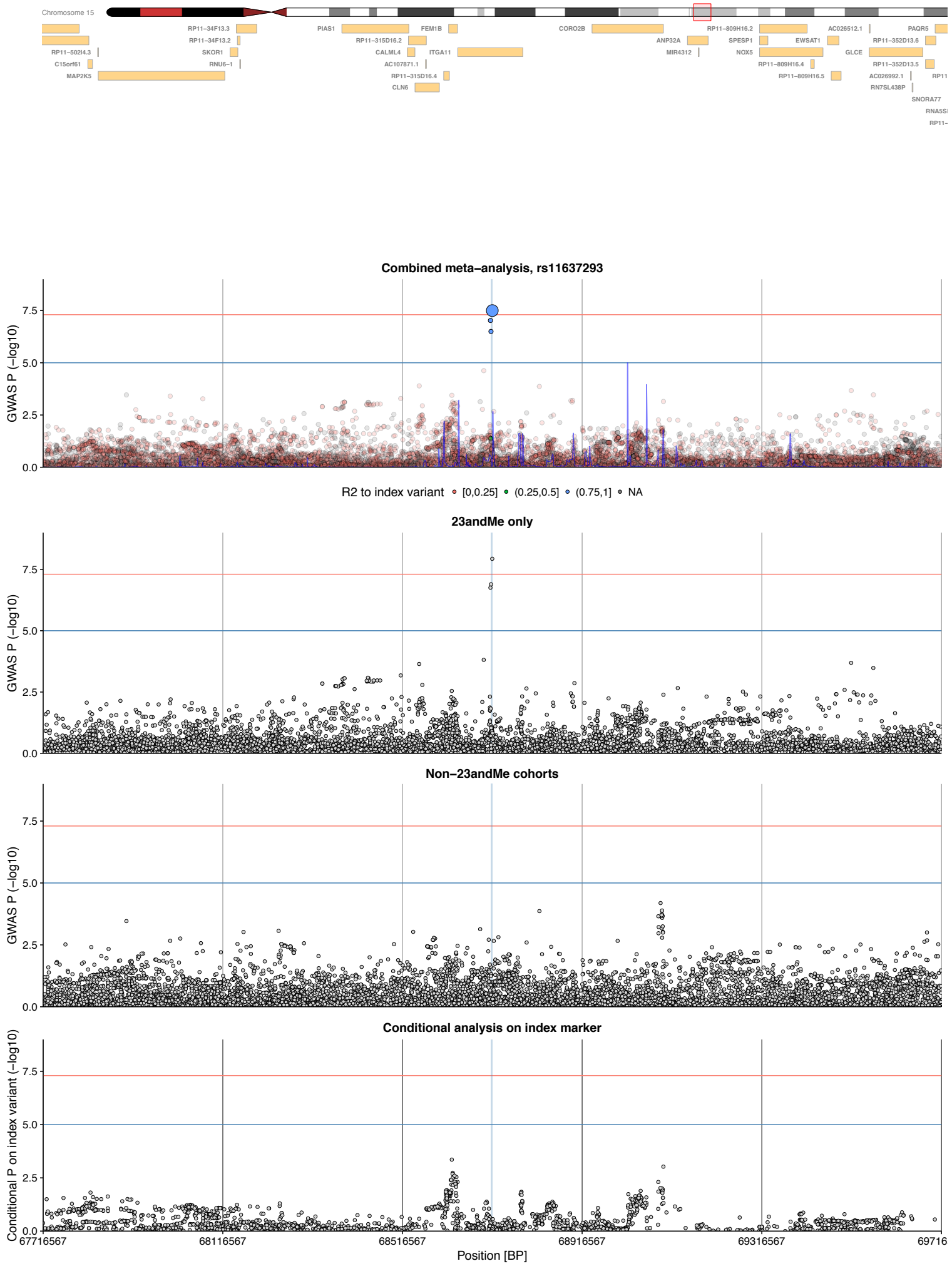

Chromosome 15
Combined meta−analysis, rs111371454
GWAS P (−log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
40760617
41160617
41560617
41960617
42360617
42760
Position [BP]

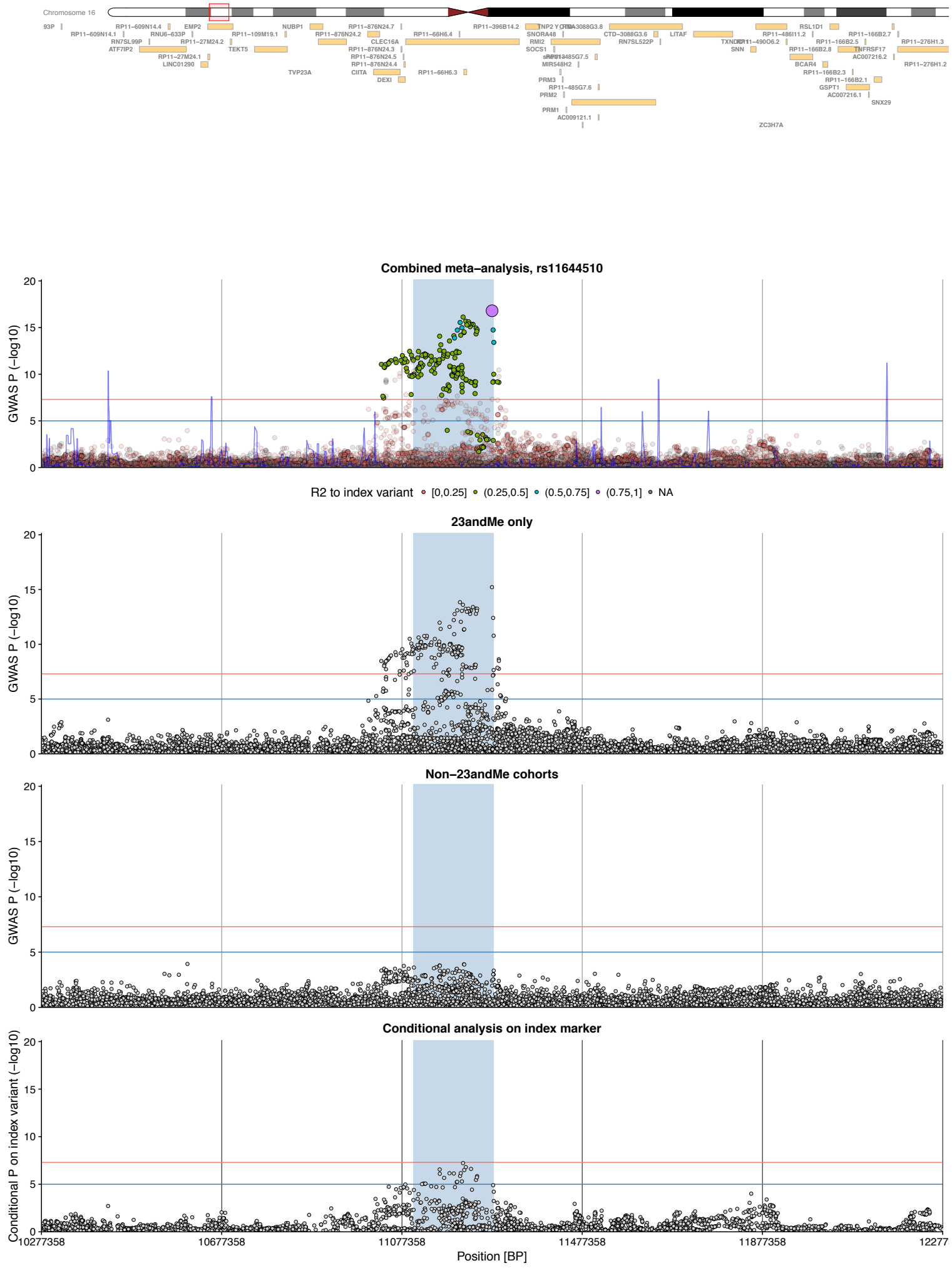

Chromosome 12
RP11−386G11.10
TROAP
RP11−133N21.10
TMBIM6
RP11−70F11.8
SMARCD1
LIMA1
LARP4
RNU6−769P
ATF1
LMBR1L
TUBA1C
SPATS2
FMNL3
RP11−70F11.7
RACGAP1
COX14
RNU6−1093P
DIP2B
TMPRSS12
DHH
TUBA1A
PRPH
KCNH3
BCDIN3D−AS1
AQP2
RP4−605O3.4
FAM186A
RNU6−238P
RN7SL519P
TUBA1B
RP11−161H23.9
RNU6−834P
NCKAP5L
AQP5
ASIC1
MIR1293
RHEBL1
RP11−977B10.2
MCRS1
BCDIN3D
AQP6
GPD1
AC140061.12
Y_RNA
RP11−161H23.5
FAM186B
RP11−70F11.11
RP3−405J10.3
DNAJC22
PRPF40B
FAIM2
CERS5
KAG1
C1QL4
RP11−469H8.6
RP3−405J10.4
RP11−386G11.8
RP11−469H8.8
DDN
KMT2D

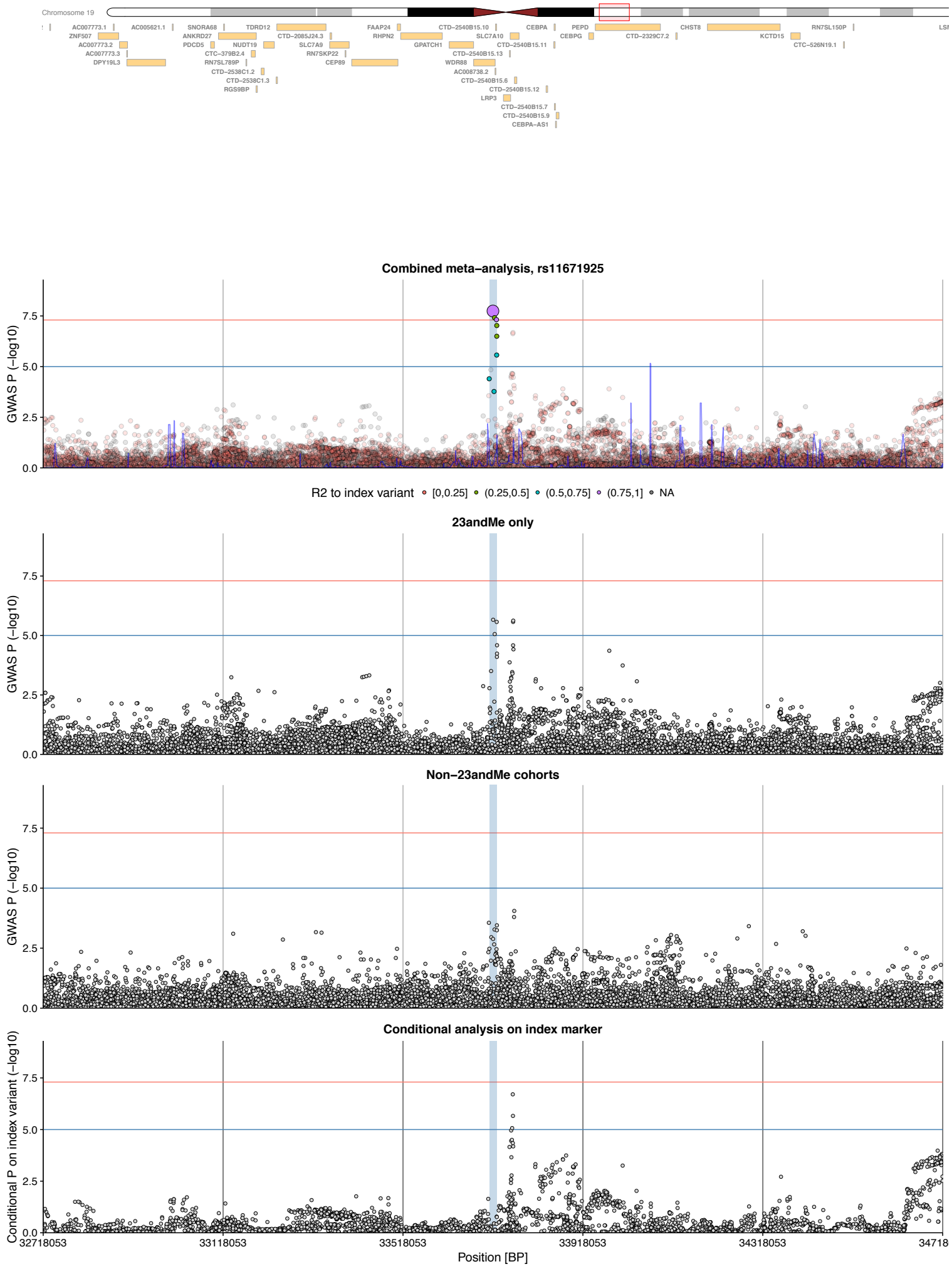

Combined meta−analysis, rs11169225
GWAS P (−log10)
10.0
7.5
5.0
2.5
0.0
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
49345671
49745671
50145671
50545671
50945671
51345
Position [BP]

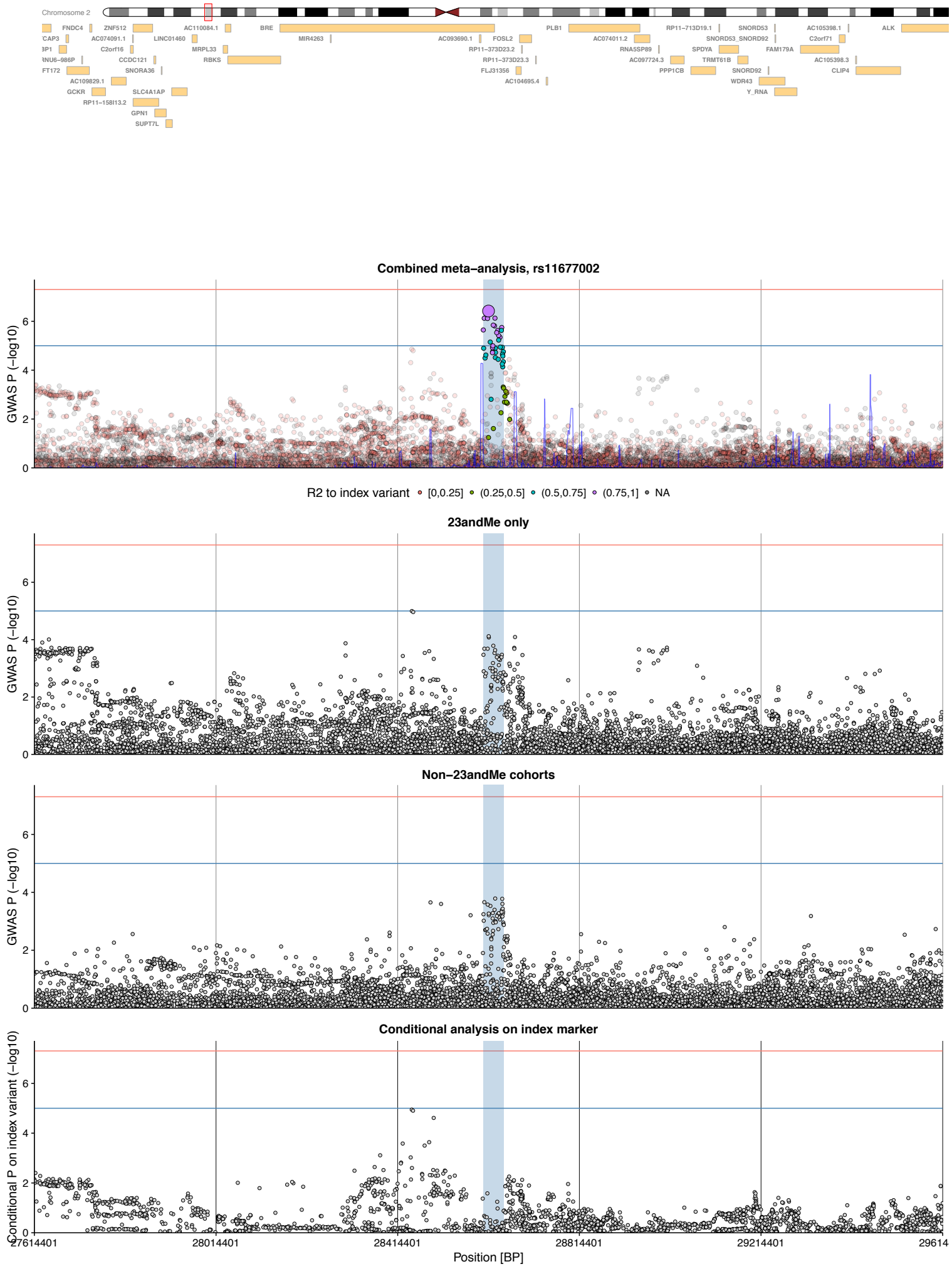

Chromosome 10
LINC00708
RP11−543F8.2
RNA5SP299
RP11−428L9.1
LINC00709
RP5−1051H14.2
RP5−1119O21.2
RP11−428L9.2
GATA3
Combined meta−analysis, rs11256017
GWAS P (−log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
8043919
8443919
8843919
9243919
9643919
10043
Position [BP]

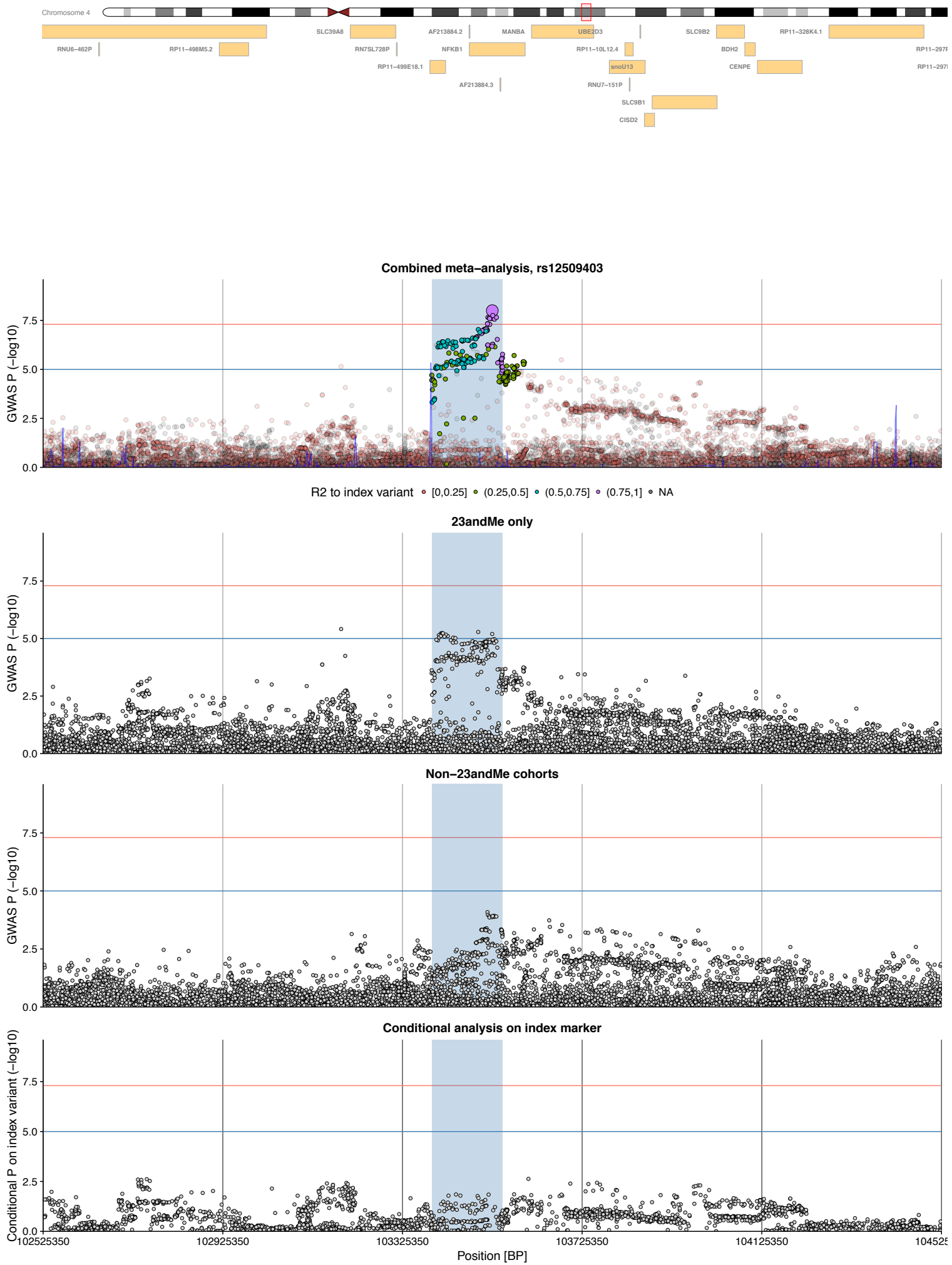

Chromosome 6
Combined meta–analysis, rs112779317
GWAS P (–log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non–23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (–log10)
Position [BP]

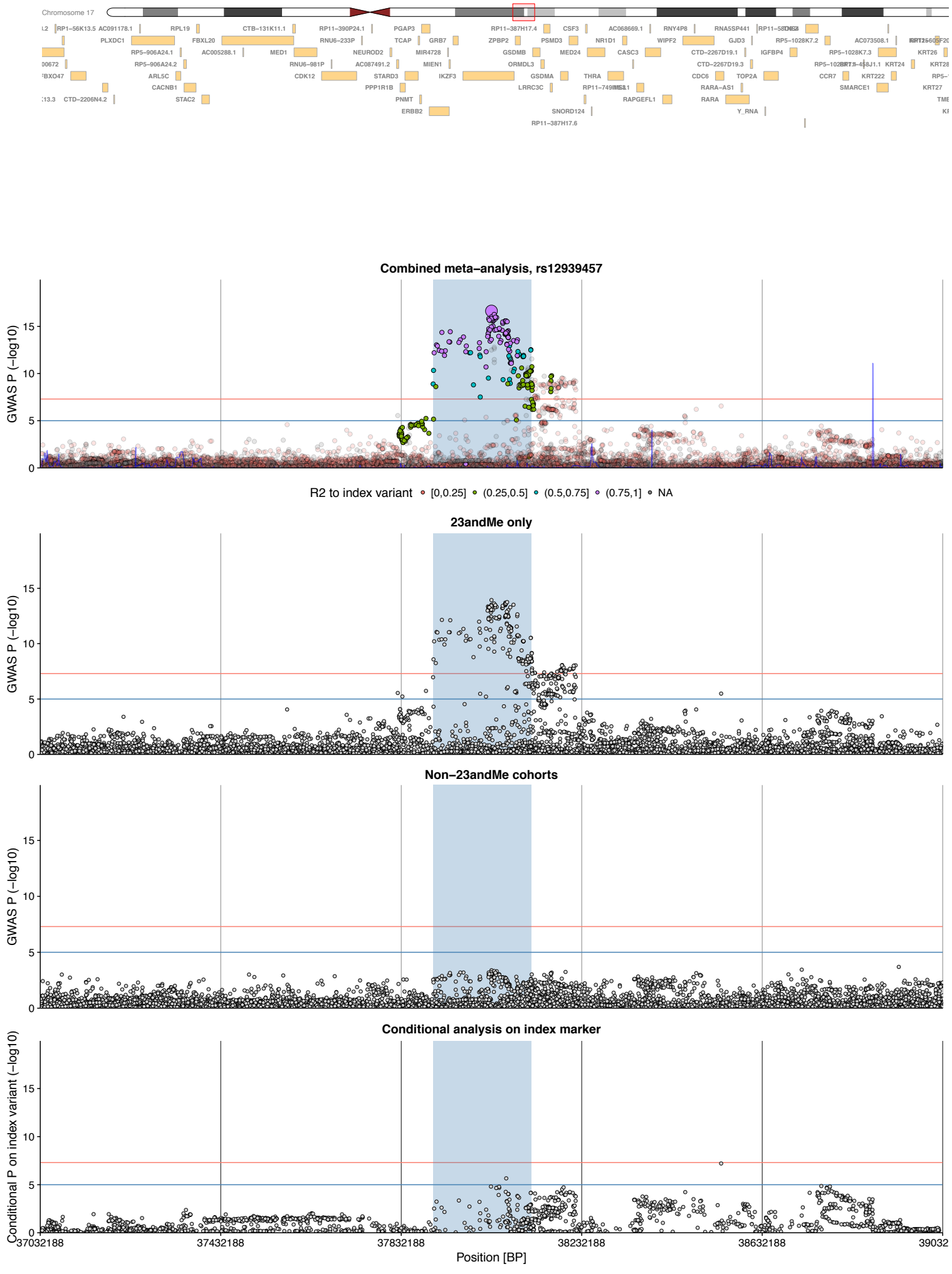

Chromosome 15
RP11−34F13.3
RP11−34F13.2
RP11−502I4.3
SKOR1
C15orf61
RNU6−1
MAP2K5
PIAS1
FEM1B
RP11−315D16.2
CALML4
ITGA11
AC107871.1
RP11−315D16.4
CLN6
CORO2B
ANP32A
MIR4312
RP11−809H16.2
SPESP1
EWSAT1
NOX5
RP11−809H16.4
RP11−809H16.5
AC026512.1
PAQR5
RP11−352D13.6
GLCE
RP11−352D13.5
AC026992.1
RN7SL438P
SNORA77
Combined meta−analysis, rs11637293
GWAS P (−log10)
7.5
5.0
2.5
0.0
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
67716567
68116567
68516567
68916567
69316567
Position [BP]

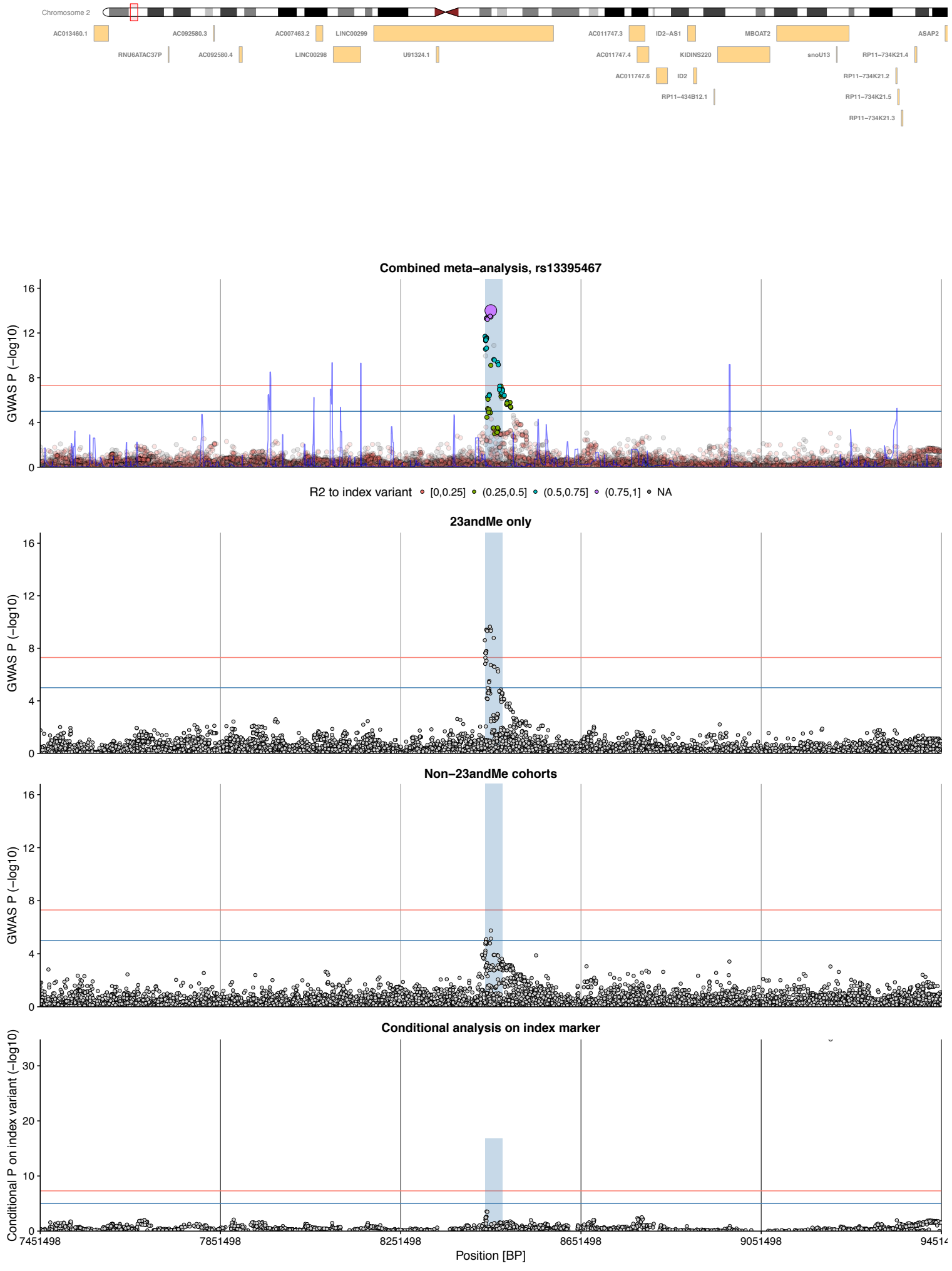

Chromosome 16
Combined meta–analysis, rs11644510
GWAS P (−log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
10277358
10677358
11077358
11477358
11877358
12277
Position [BP]

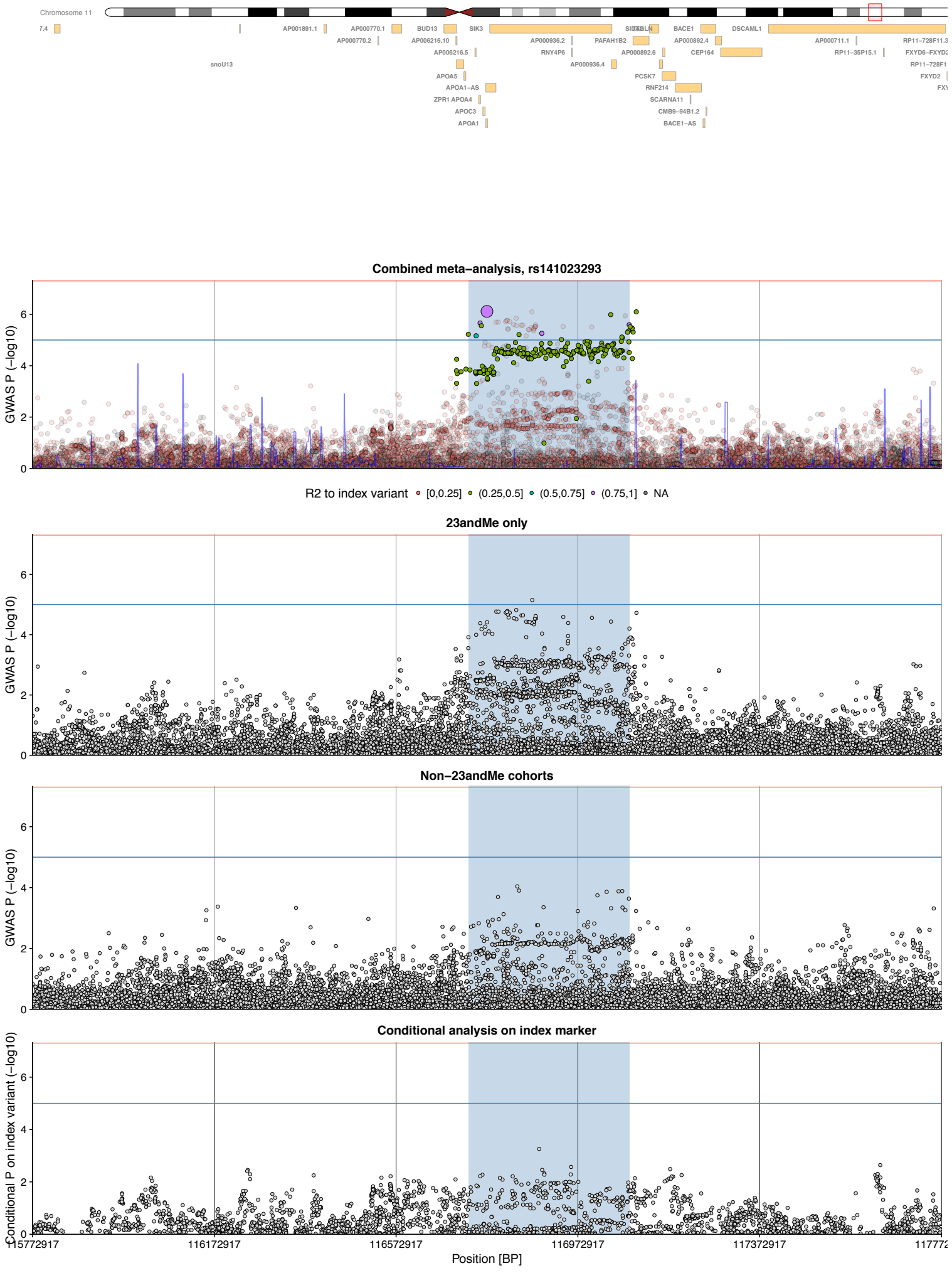

Chromosome 19
AC007773.1
AC005621.1
SNORA68
TDRD12
FAAP24
CTD−2540B15.10
CEBPA
PEPD
CHST8
RN7SL150P
ZNF507
ANKRD27
CTD−2085J24.3
RHPN2
SLC7A10
CEBPG
CTD−2329C7.2
KCTD15
AC007773.2
PDCD5
NUDT19
SLC7A9
GPATCH1
CTD−2540B15.11
CTC−526N19.1
AC007773.3
CTC−379B2.4
RN7SKP22
CTD−2540B15.13
DPY19L3
RN7SL789P
CEP89
WDR88
CTD−2538C1.2
AC008738.2
CTD−2538C1.3
CTD−2540B15.6
RGS9BP
CTD−2540B15.12
LRP3
CTD−2540B15.7
CTD−2540B15.9
CEBPA−AS1
Combined meta−analysis, rs11671925
GWAS P (−log10)
7.5
5.0
2.5
0.0
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
32718053
33118053
33518053
33918053
34318053
34718
Position [BP]

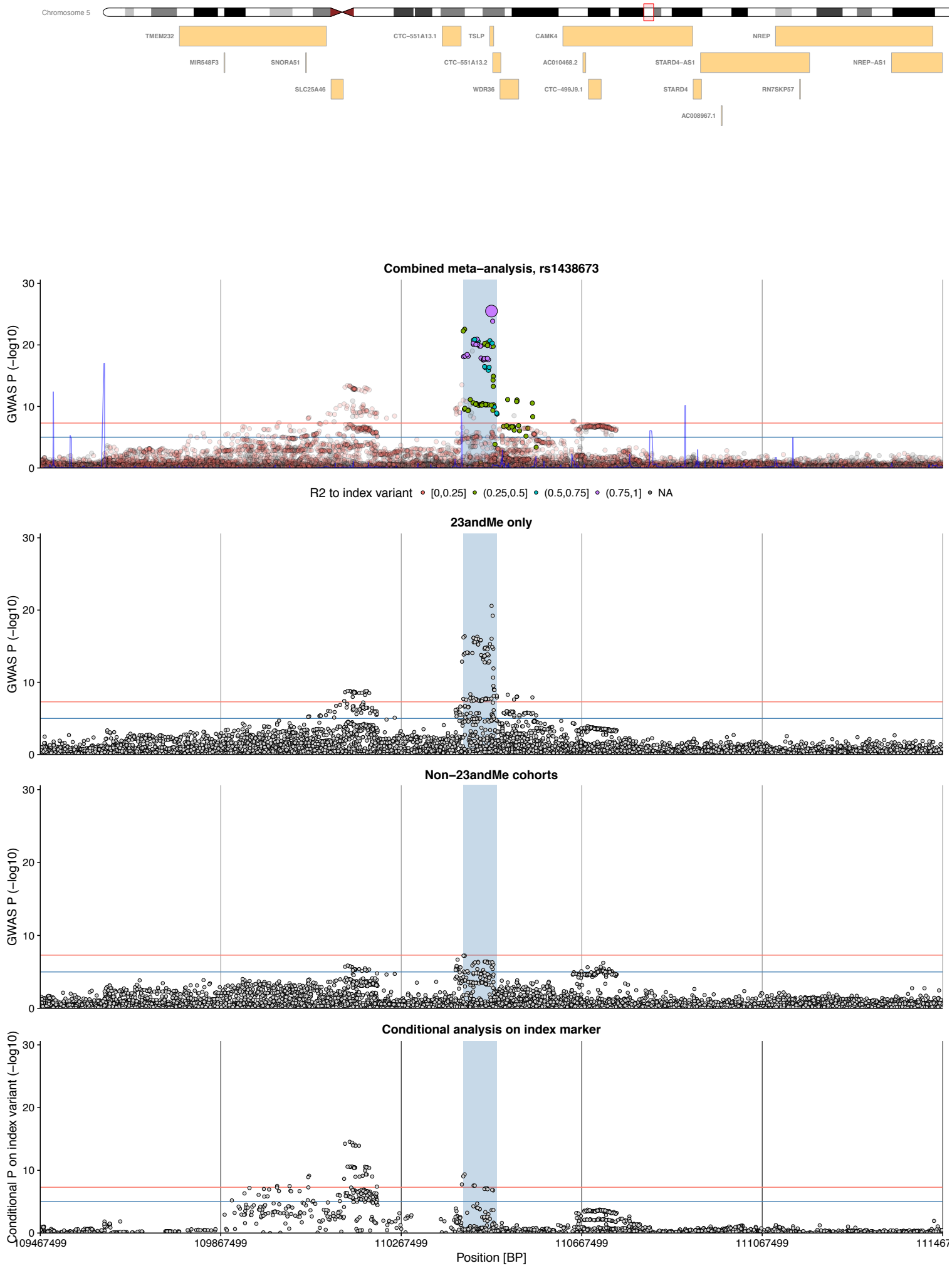

Chromosome 2
FNDC4
ZNF512
AC110084.1
BRE
PLB1
RP11−713D19.1
SNORD53
AC105398.1
ALK
AC074091.1
LINC01460
MIR4263
AC093690.1
FOSL2
AC074011.2
SNORD53_SNORD92
C2orf71
C2orf16
MRPL33
RP11−373D23.2
RNA5SP89
SPDYA
FAM179A
CCDC121
RBKS
RP11−373D23.3
AC097724.3
TRMT61B
AC105398.3
SNORA36
FLJ31356
PPP1CB
SNORD92
CLIP4
AC109829.1
AC104695.4
WDR43
GCKR
SLC4A1AP
Y_RNA
RP11−158I13.2
GPN1
SUPT7L
Combined meta−analysis, rs11677002
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
27614401
28014401
28414401
28814401
29214401
29614
Position [BP]

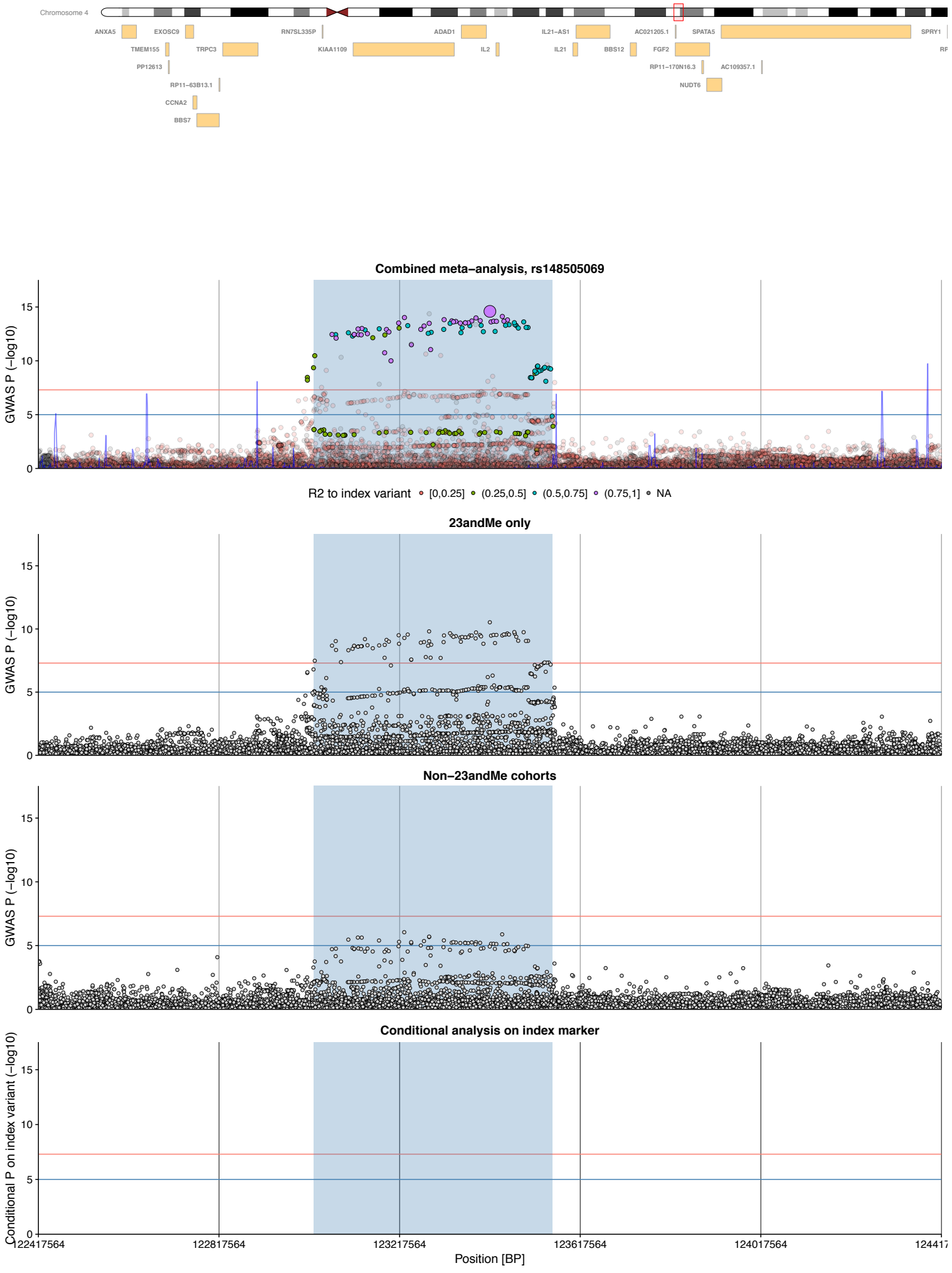

Chromosome 4
SLC39A8
AF213884.2
MANBA
UBE2D3
SLC9B2
RP11−328K4.1
RNU6−462P
RP11−498M5.2
RN7SL728P
NFKB1
RP11−10L12.4
BDH2
RP11−499E18.1
snoU13
CENPE
AF213884.3
RNU7−151P
SLC9B1
CISD2
Combined meta−analysis, rs12509403
GWAS P (−log10)
7.5
5.0
2.5
0.0
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
102525350
102925350
103325350
103725350
104125350
Position [BP]

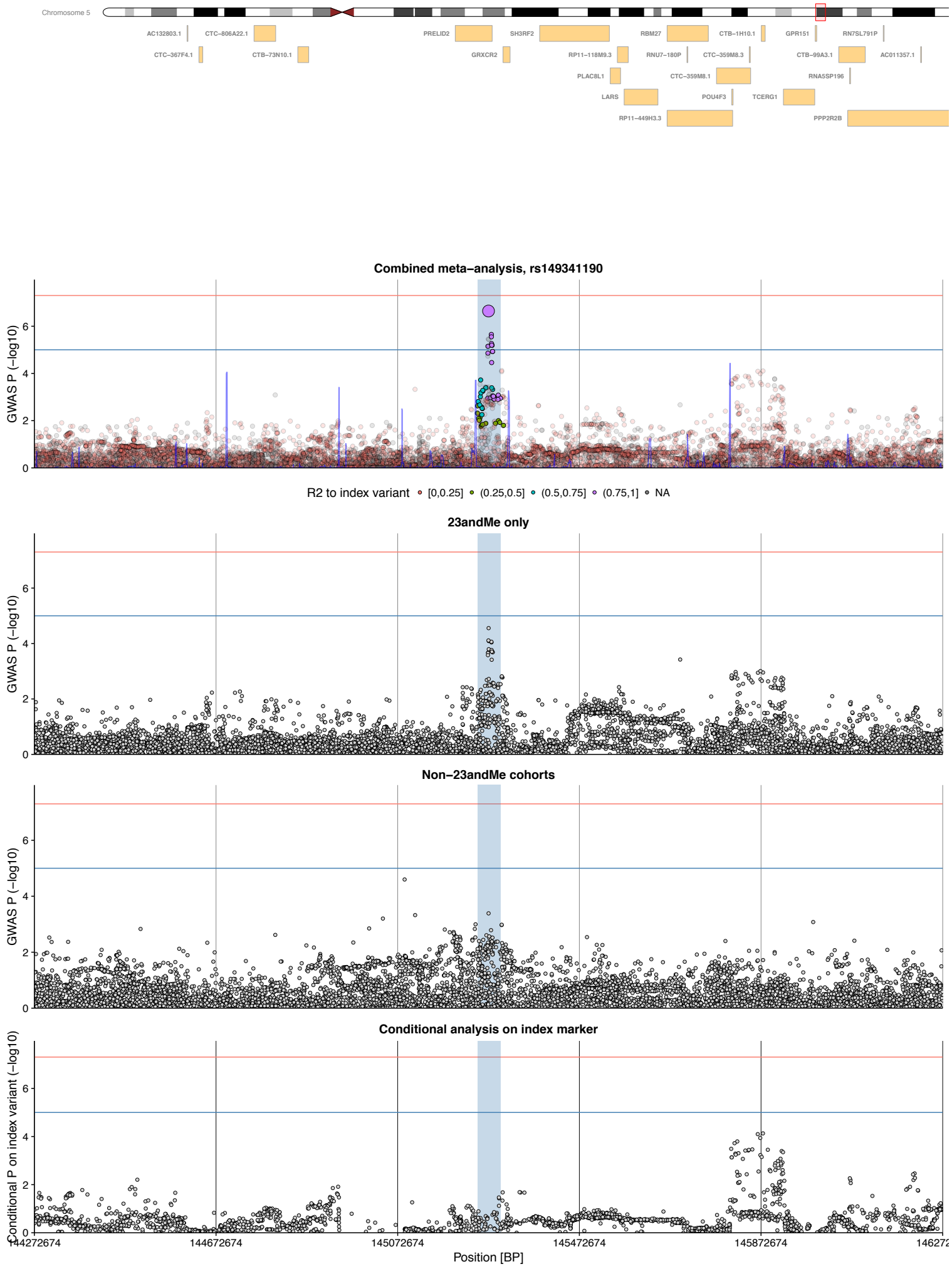

Chromosome 17
Combined meta–analysis, rs12939457
GWAS P (–log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non–23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (–log10)
37032188
37432188
37832188
38232188
38632188
39032
Position [BP]

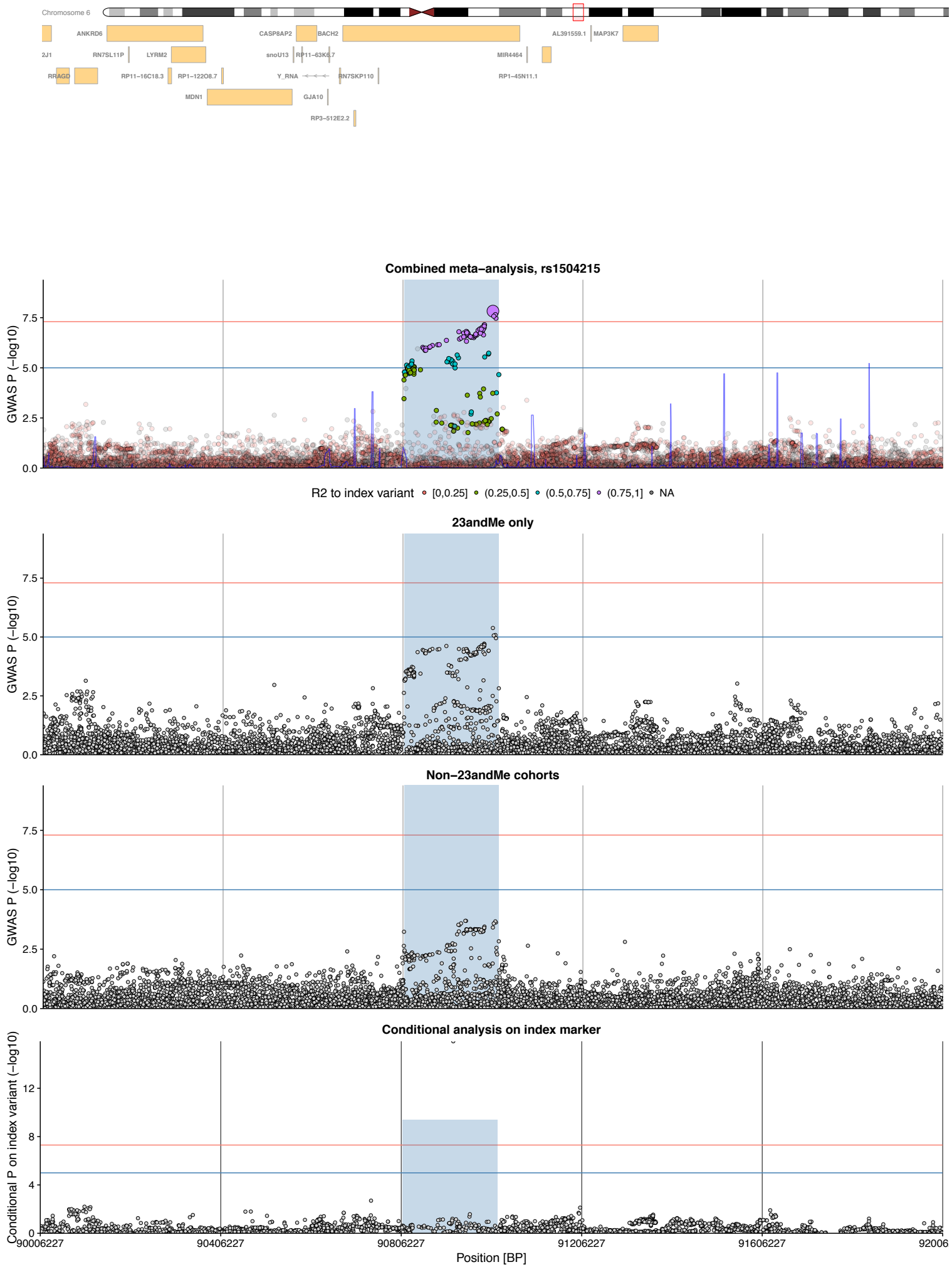

Chromosome 2
AC013460.1
AC092580.3
AC007463.2
LINC00299
AC011747.3
ID2−AS1
MBOAT2
ASAP2
RNU6ATAC37P
AC092580.4
LINC00298
U91324.1
AC011747.4
KIDINS220
snoU13
RP11−734K21.4
AC011747.6
ID2
RP11−734K21.2
RP11−434B12.1
RP11−734K21.5
RP11−734K21.3
Combined meta−analysis, rs13395467
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
7451498
7851498
8251498
8651498
9051498
Position [BP]

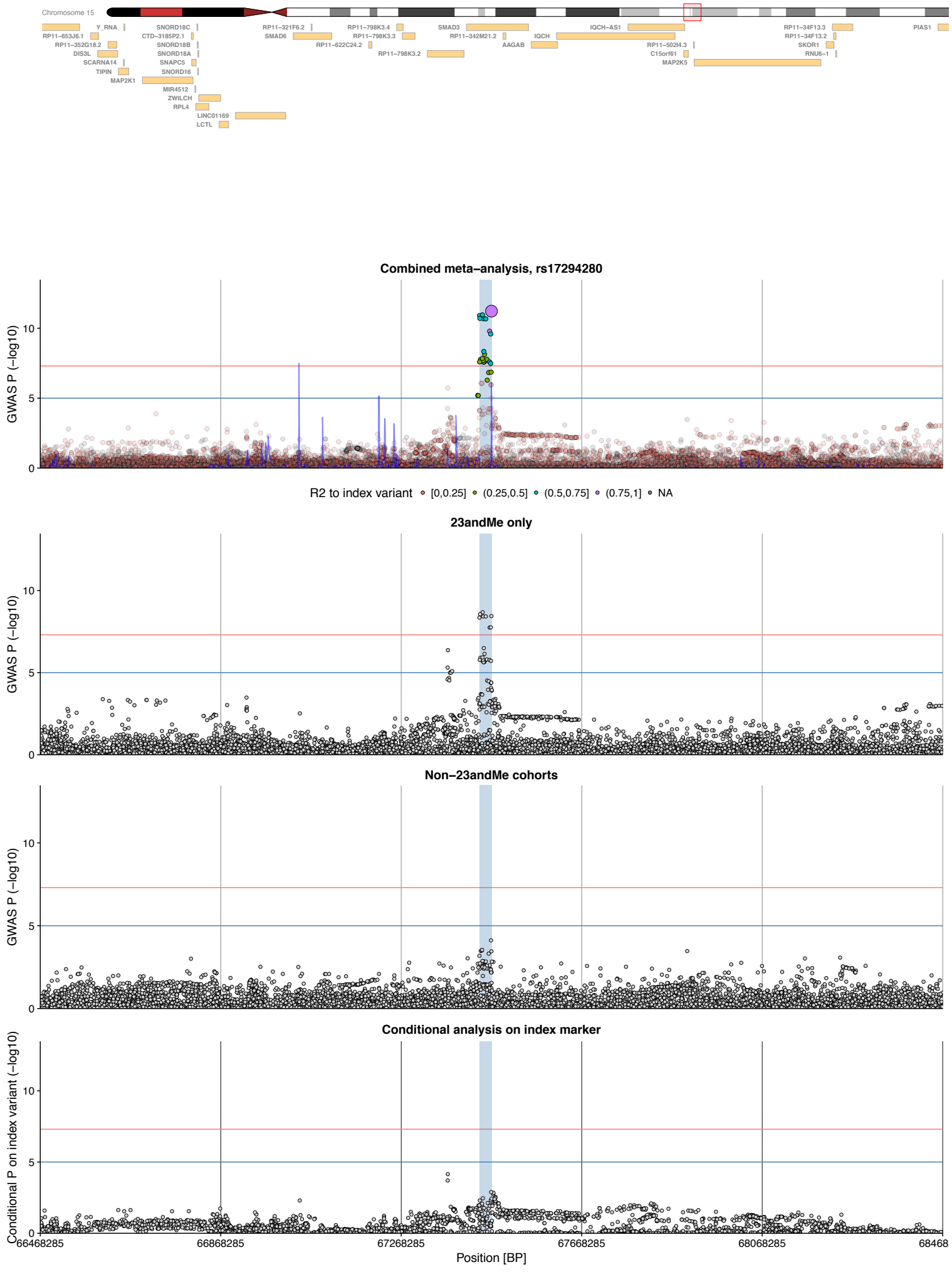

Chromosome 11
7.4
AP001891.1
AP000770.1
BUD13
SIK3
BACE1
DSCAML1
AP000770.2
AP006216.10
AP000936.2
PAFAH1B2
AP000892.4
AP000711.1
RP11−728F11.3
AP006216.5
RNY4P6
AP000892.6
CEP164
RP11−35P15.1
FXYD6−FXYD2
snoU13
AP000936.4
RP11−728F1
APOA5
PCSK7
FXYD2
APOA1−AS
RNF214
ZPR1 APOA4
SCARNA11
APOC3
CMB9−94B1.2
APOA1
BACE1−AS
Combined meta−analysis, rs141023293
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
115772917
116172917
116572917
116972917
117372917
Position [BP]

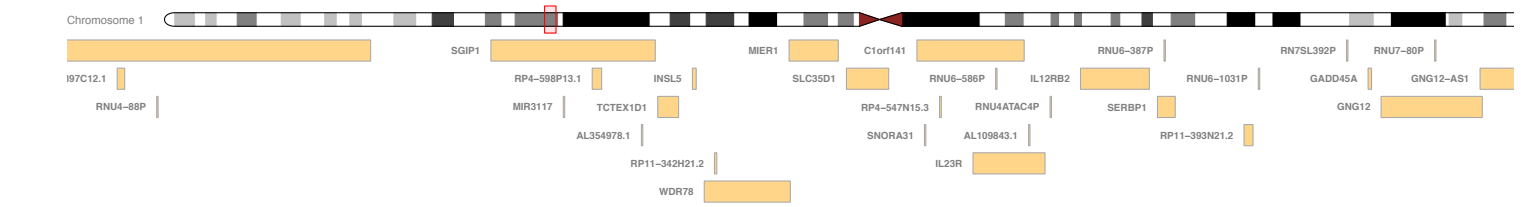

Chromosome 5
TMEM232
MIR548F3
SNORA51
SLC25A46
CTC−551A13.1
TSLP
CTC−551A13.2
WDR36
CAMK4
AC010468.2
CTC−499J9.1
NREP
STARD4−AS1
STARD4
AC008967.1
RN7SKP57
NREP−AS1
Combined meta−analysis, rs1438673
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
109467499
109867499
110267499
110667499
111067499
Position [BP]

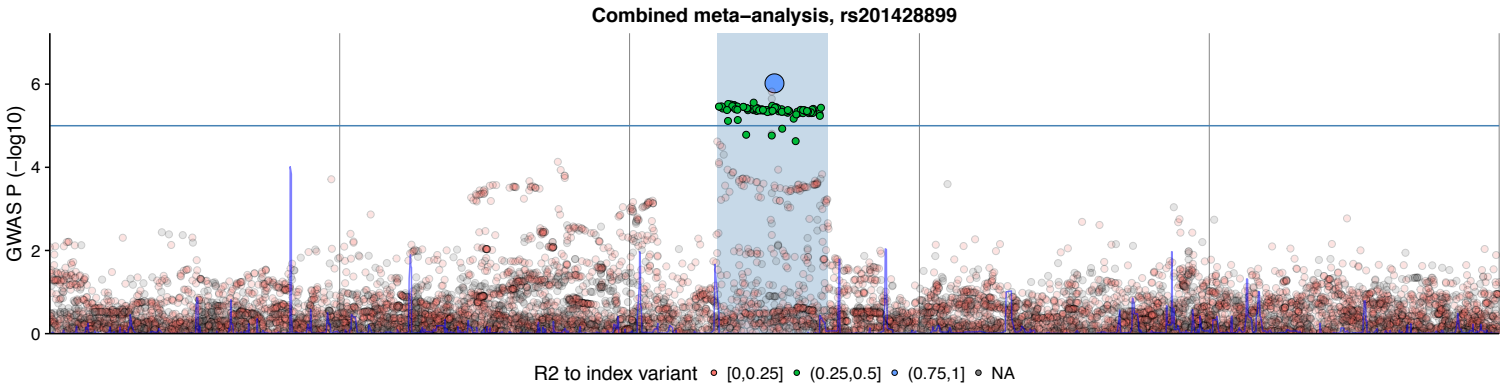

Chromosome 4
ANXA5
EXOSC9
RN7SL335P
ADAD1
IL21−AS1
AC021205.1
SPATA5
SPRY1
TMEM155
TRPC3
KIAA1109
IL2
IL21
BBS12
FGF2
PP12613
RP11−170N16.3
AC109357.1
RP11−63B13.1
NUDT6
CCNA2
BBS7
Combined meta−analysis, rs148505069
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
122417564
122817564
123217564
123617564
124017564
Position [BP]


Chromosome 5
AC132803.1
CTC−806A22.1
CTC−367F4.1
CTB−73N10.1
PRELID2
SH3RF2
GRXCR2
RBM27
CTB−1H10.1
RP11−118M9.3
RNU7−180P
CTC−359M8.3
GPR151
RN7SL791P
CTB−99A3.1
AC011357.1
PLAC8L1
CTC−359M8.1
RNA5SP196
LARS
POU4F3
TCERG1
RP11−449H3.3
PPP2R2B
Combined meta−analysis, rs149341190
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
144272674
144672674
145072674
145472674
145872674
Position [BP]

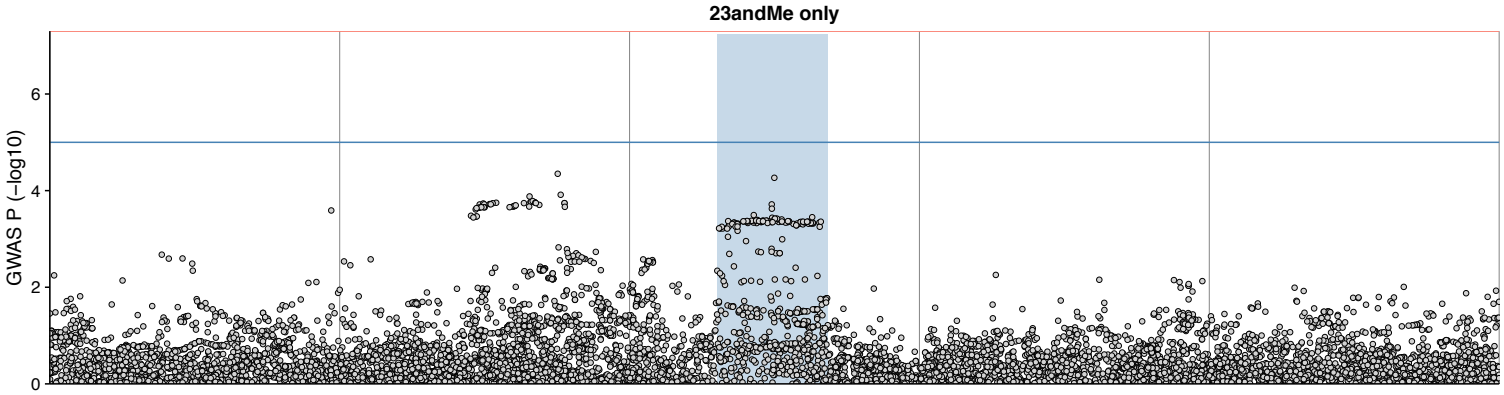

Chromosome 6
ANKRD6
CASP8AP2
BACH2
AL391559.1
MAP3K7
2J1
RN7SL11P
LYRM2
snoU13
RP11−63K6.7
MIR4464
RRAGD
RP11−16C18.3
RP1−122O8.7
Y_RNA
RN7SKP110
RP1−45N11.1
MDN1
GJA10
RP3−512E2.2
Combined meta−analysis, rs1504215
GWAS P (−log10)
7.5
5.0
2.5
0.0
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
12
8
4
0
90006227
90406227
90806227
91206227
91606227
92006
Position [BP]

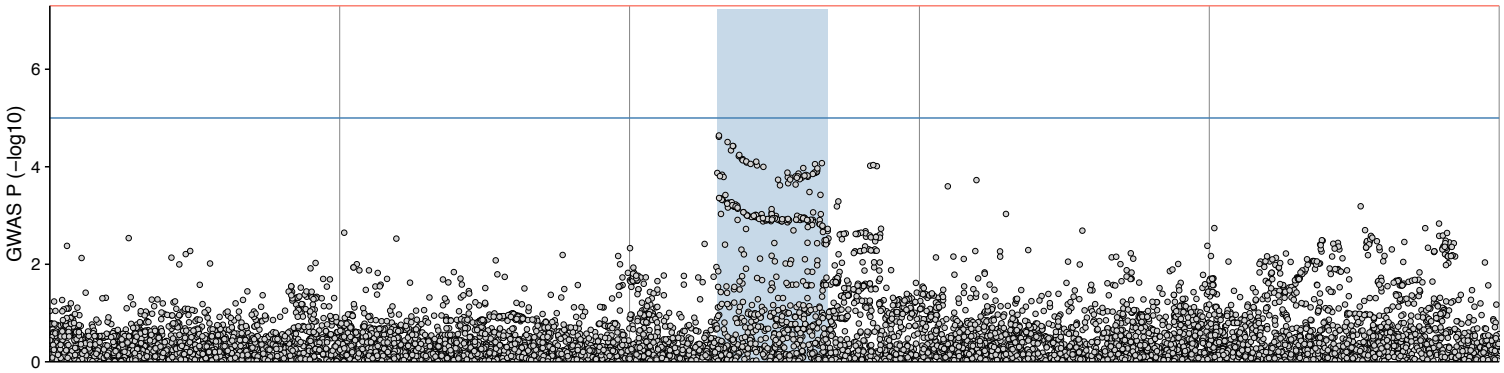

Chromosome 15
Y_RNA
SNORD18C
RP11−321F6.2
RP11−798K3.4
SMAD3
IQCH−AS1
RP11−34F13.3
PIAS1
RP11−653J6.1
CTD−3185P2.1
SMAD6
RP11−798K3.3
RP11−342M21.2
IQCH
RP11−34F13.2
RP11−352G18.2
SNORD18B
RP11−622C24.2
AAGAB
SKOR1
DIS3L
SNORD18A
RP11−798K3.2
RP11−502I4.3
RNU6−1
SCARNA14
SNAPC5
C15orf61
TIPIN
SNORD16
MAP2K5
MAP2K1
MIR4512
ZWILCH
RPL4
LINC01169
LCTL
Combined meta−analysis, rs17294280
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
66468285
66868285
67268285
67668285
68068285
68468
Position [BP]

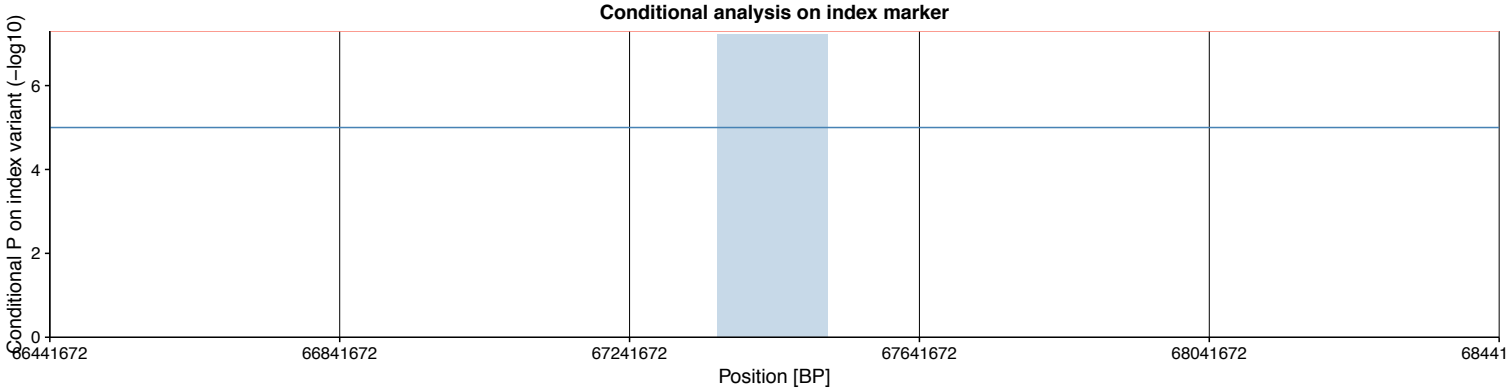


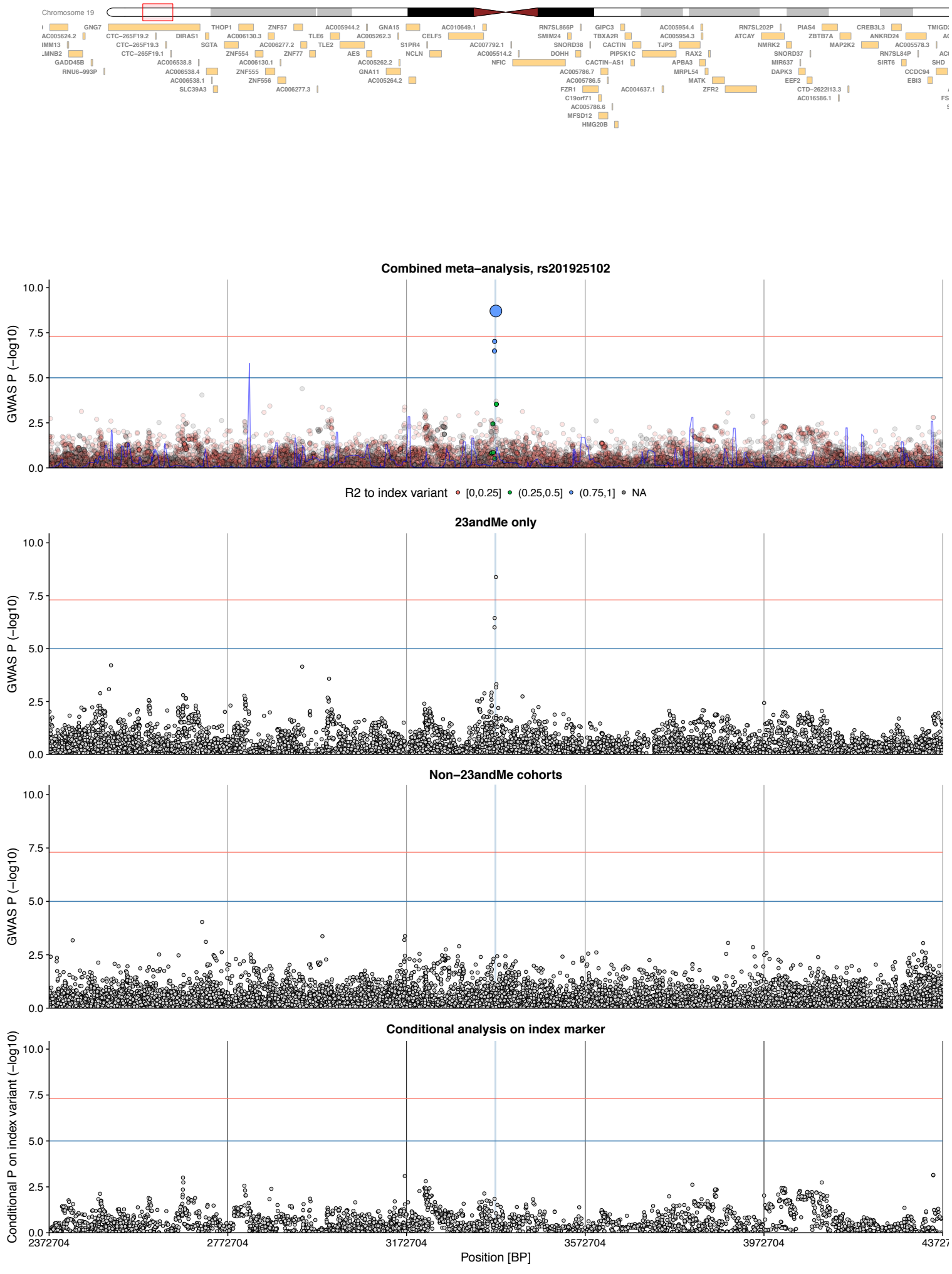


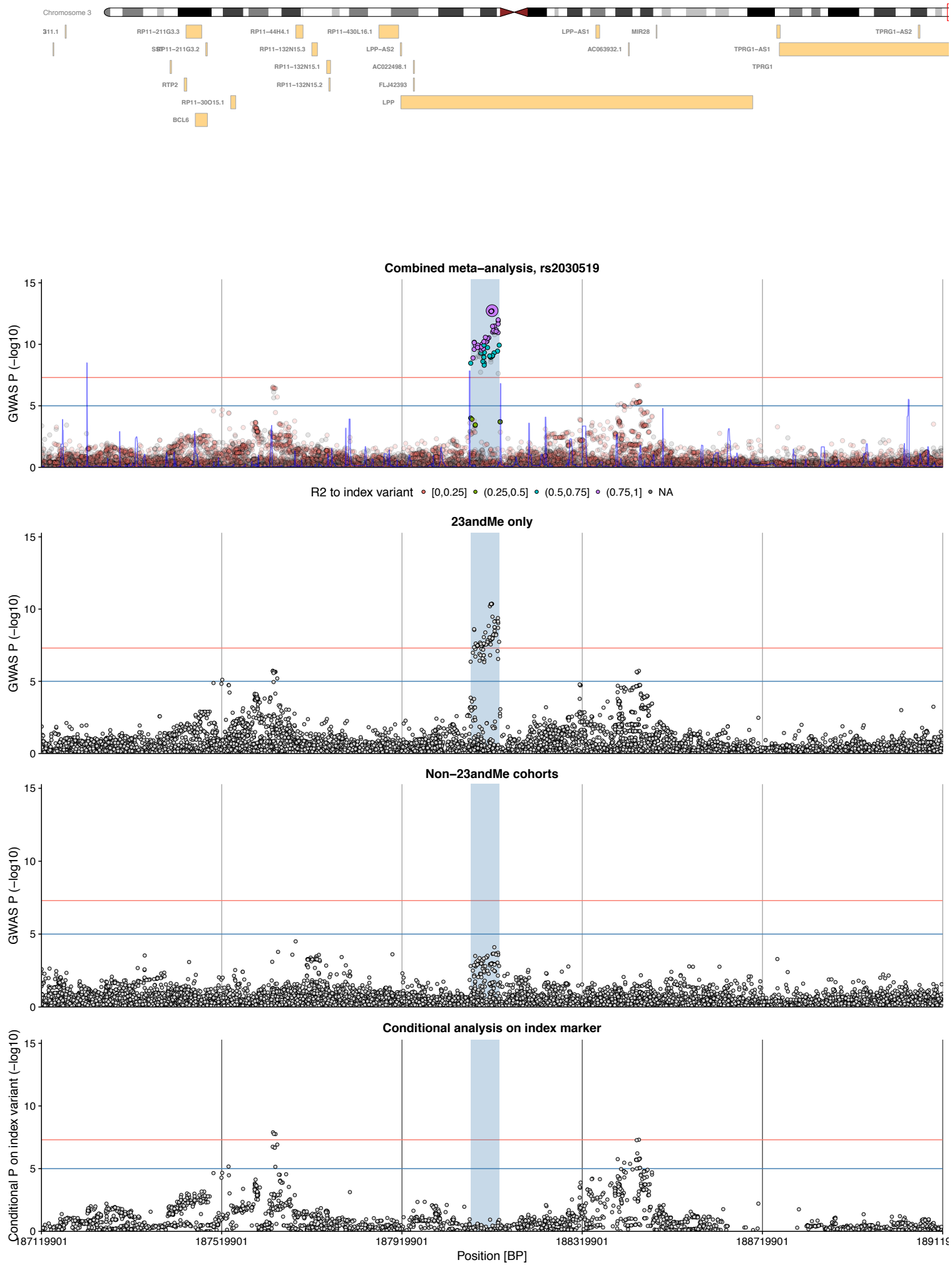
R2 to index variant [0,0.25] [0.25,0.5] (0.75,1] NA

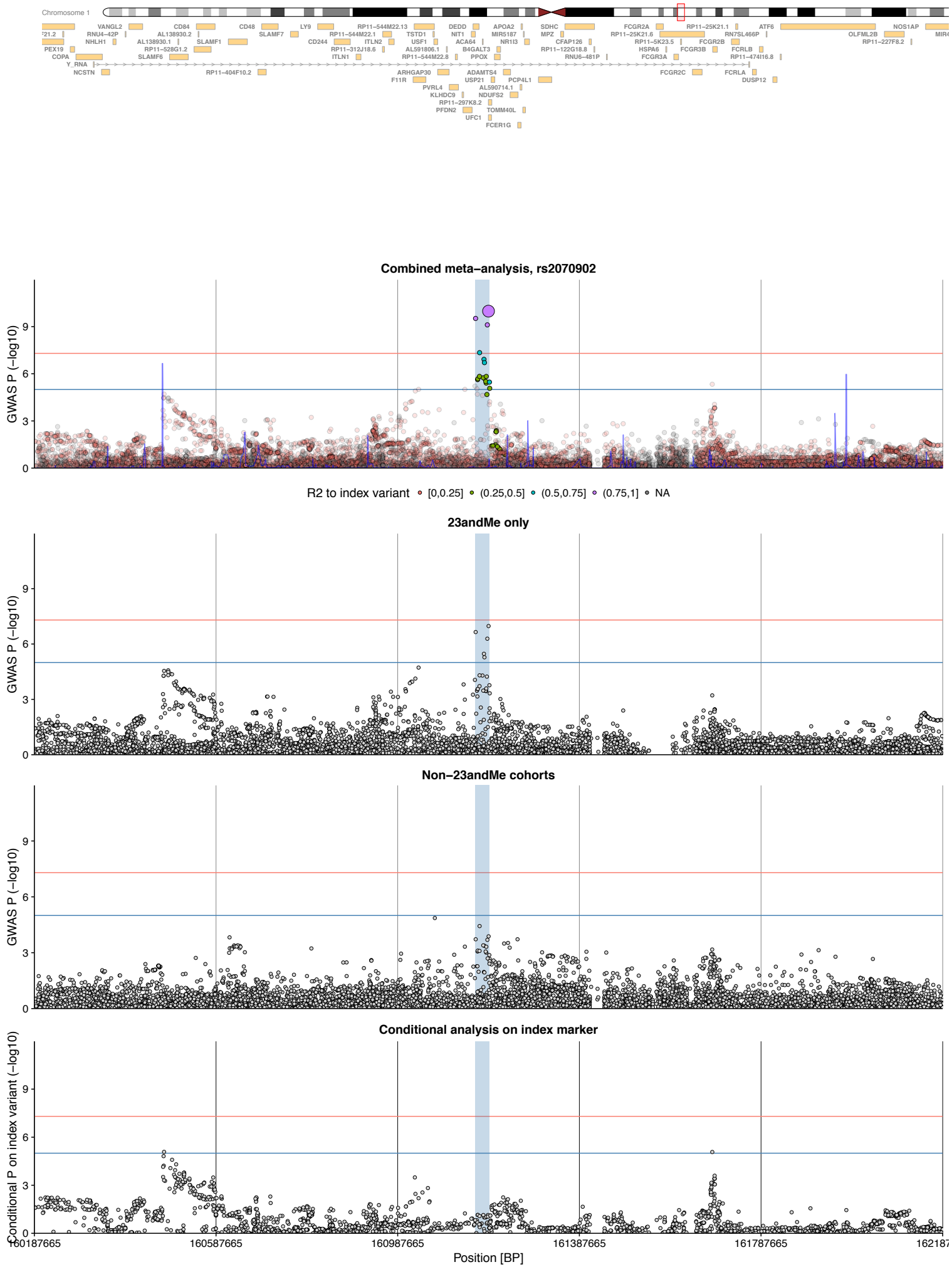


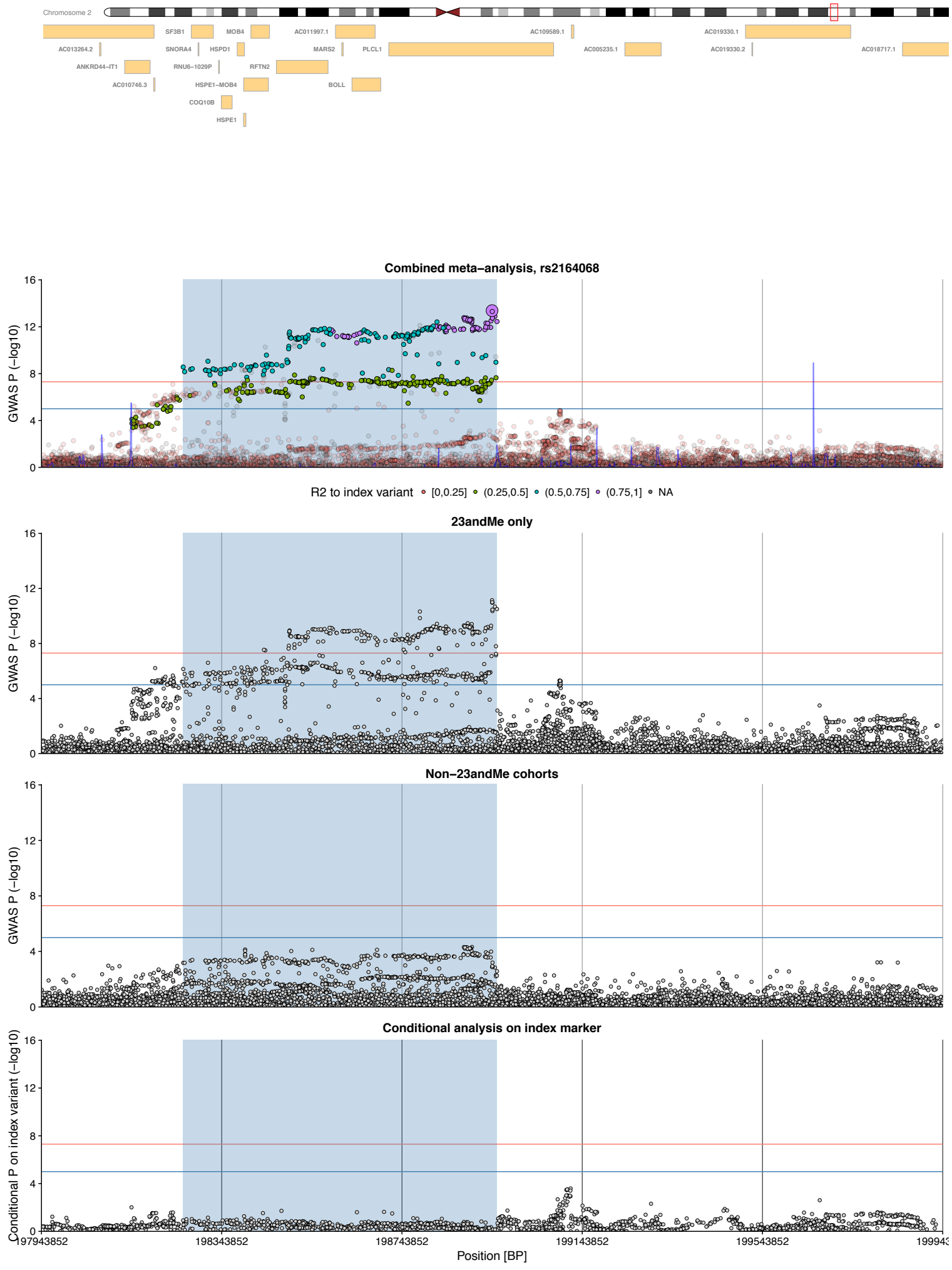


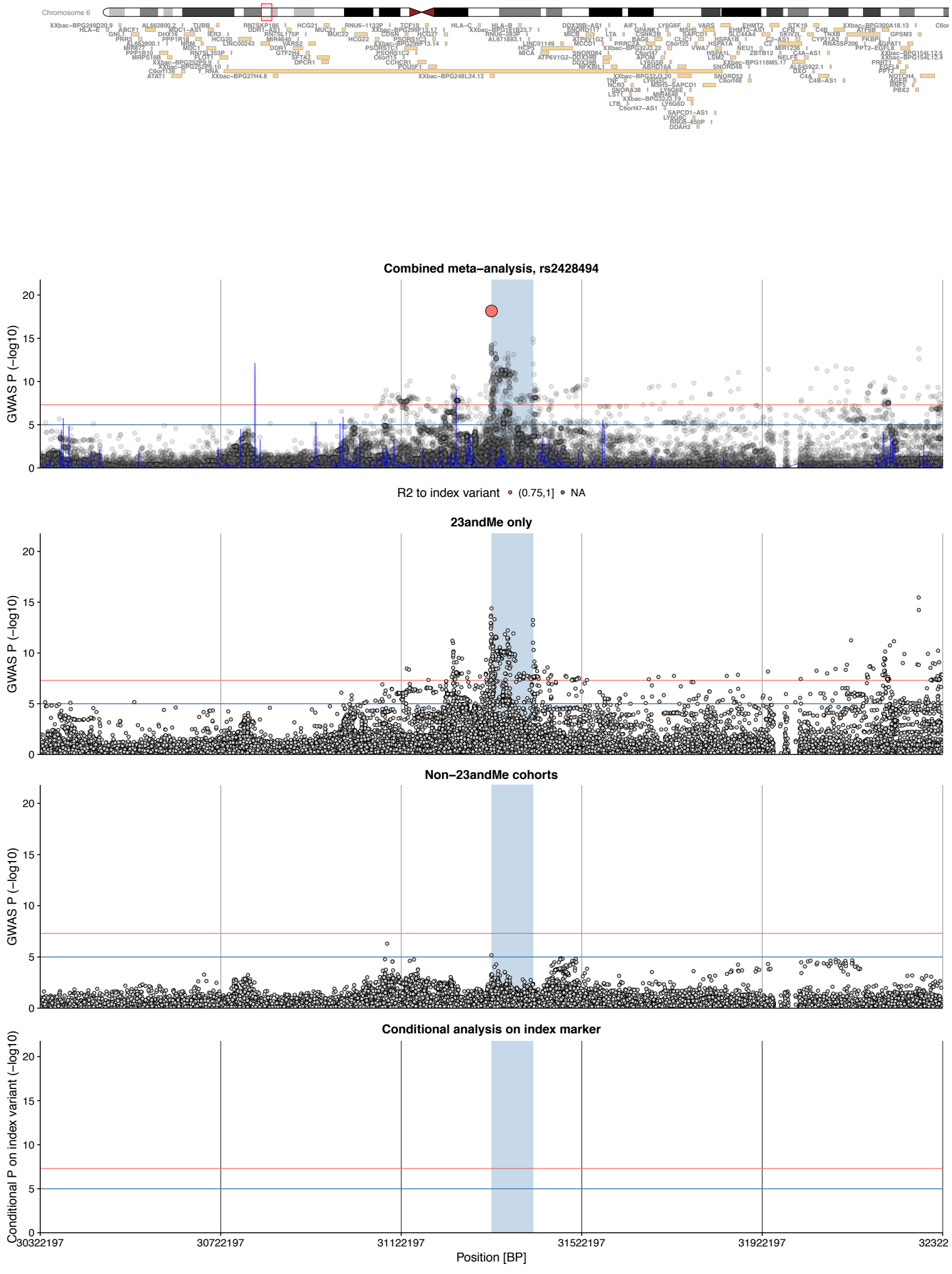


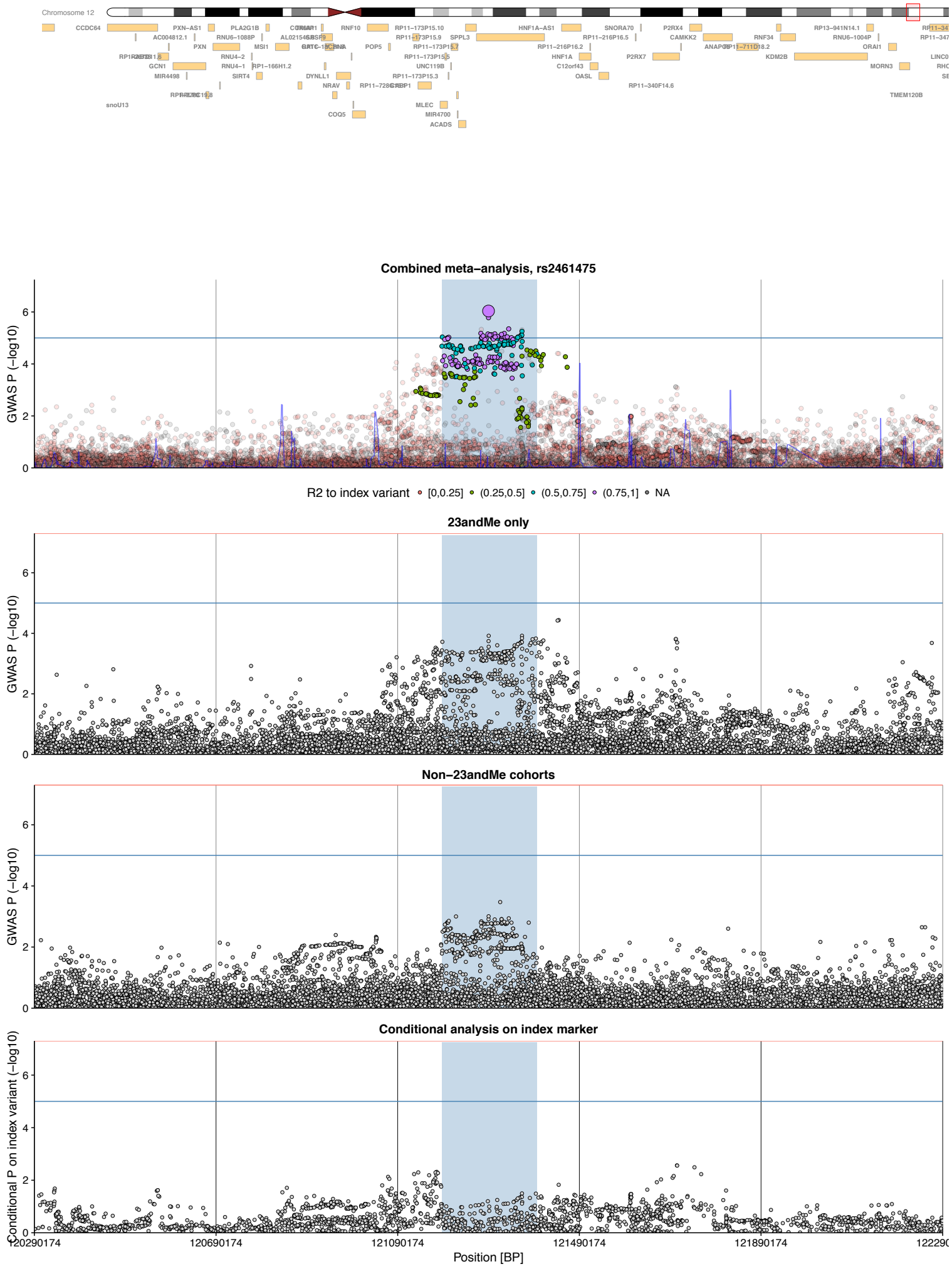

Chromosome 19
Combined meta−analysis, rs201925102
GWAS P (−log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.75,1] NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
2372704
2772704
3172704
3572704
3972704
Position [BP]

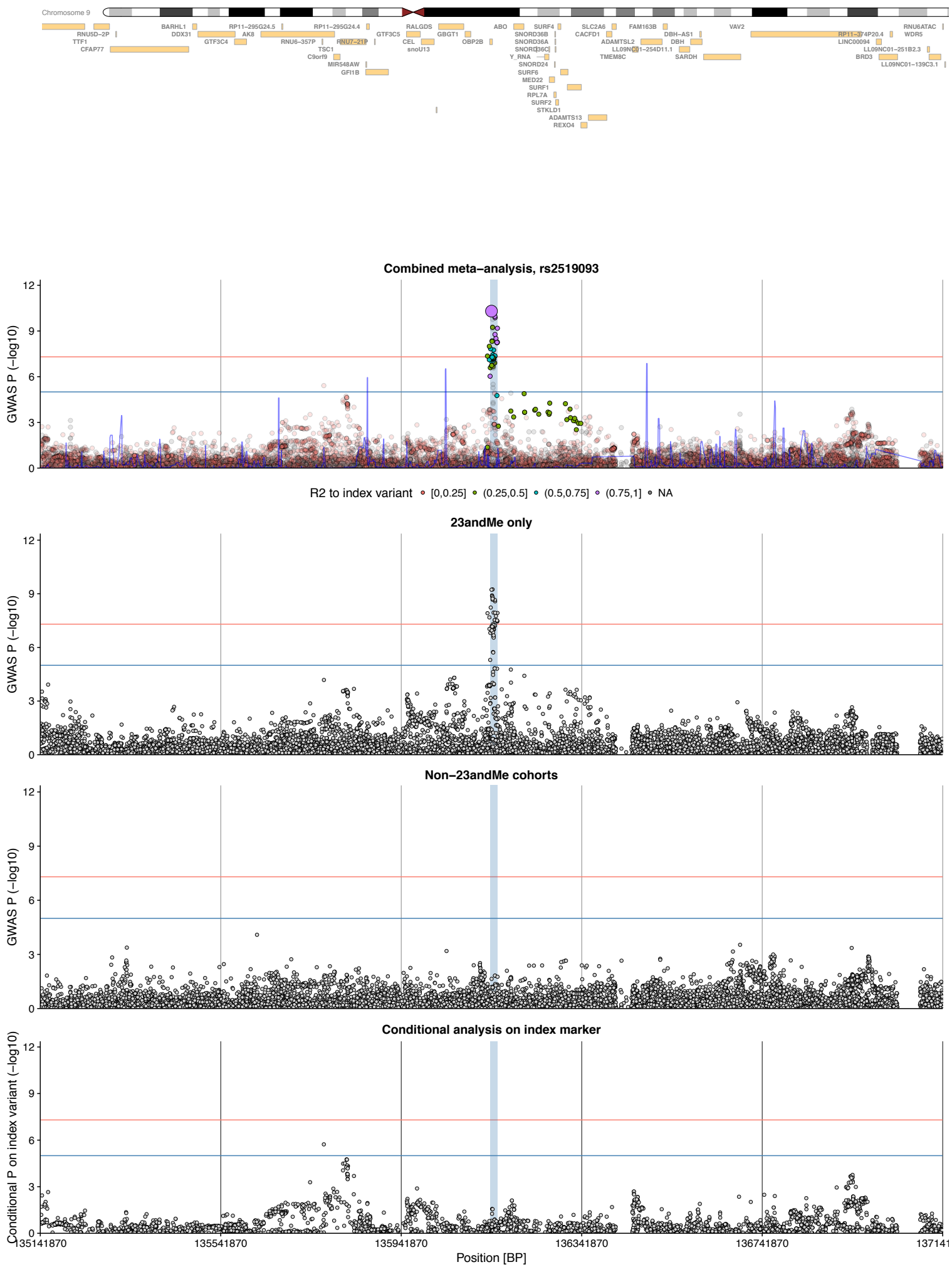

Chromosome 3
311.1
RP11−211G3.3
RP11−211G3.2
RP11−44H4.1
RP11−132N15.3
RP11−132N15.1
RP11−132N15.2
RP11−430L16.1
LPP−AS2
AC022498.1
FLJ42393
LPP
RTP2
RP11−30O15.1
BCL6
LPP−AS1
MIR28
AC063932.1
TPRG1−AS2
TPRG1−AS1
TPRG1
Combined meta−analysis, rs2030519
GWAS P (−log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
187119901
187519901
187919901
188319901
188719901
Position [BP]

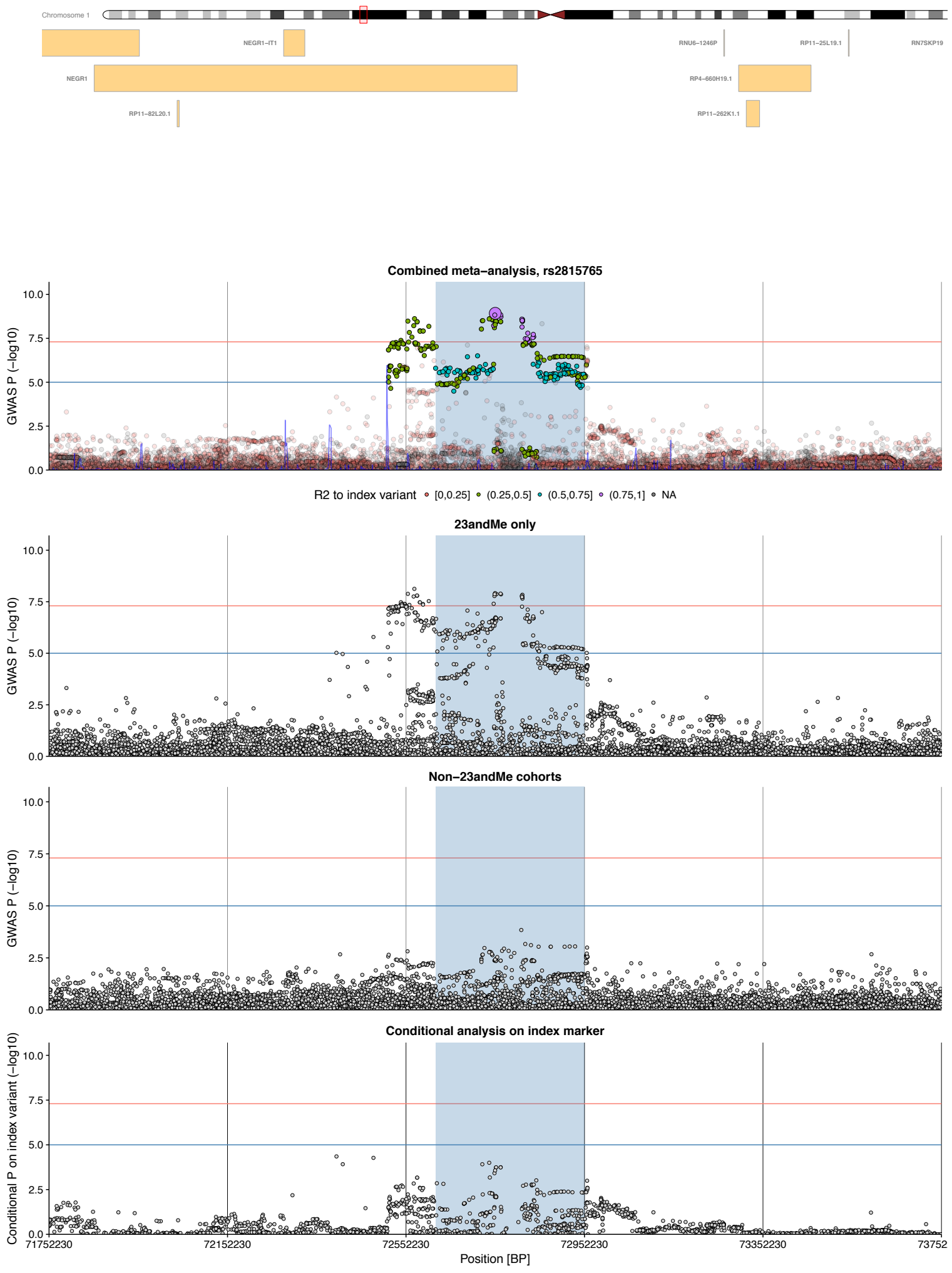

Chromosome 1
VANGL2
CD84
CD48
LY9
RP11−544M22.13
DEDD
APOA2
SDHC
FCGR2A
RP11−25K21.1
ATF6
NOS1AP
RNU4−42P
AL138930.2
SLAMF7
RP11−544M22.1
TSTD1
NIT1
MIR5187
MPZ
RP11−25K21.6
RN7SL466P
OLFML2B
NHLH1
AL138930.1
SLAMF1
CD244
ITLN2
USF1
ACA64
NR1I3
CFAP126
RP11−5K23.5
FCGR2B
RP11−227F8.2
PEX19
RP11−528G1.2
RP11−312J18.6
AL591806.1
B4GALT3
RP11−122G18.8
HSPA6
FCGR3B
FCRLB
COPA
SLAMF6
ITLN1
RP11−544M22.8
PPOX
RNU6−481P
FCGR3A
RP11−474H16.8
Y_RNA
NCSTN
RP11−404F10.2
ARHGAP30
ADAMTS4
FCGR2C
FCRLA
F11R
USP21
PCP4L1
DUSP12
PVRL4
AL590714.1
KLHDC9
NDUFS2
RP11−297K8.2
PFDN2
TOMM40L
UFC1
FCER1G
Combined meta−analysis, rs2070902
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
160187665
160587665
160987665
161387665
161787665
Position [BP]

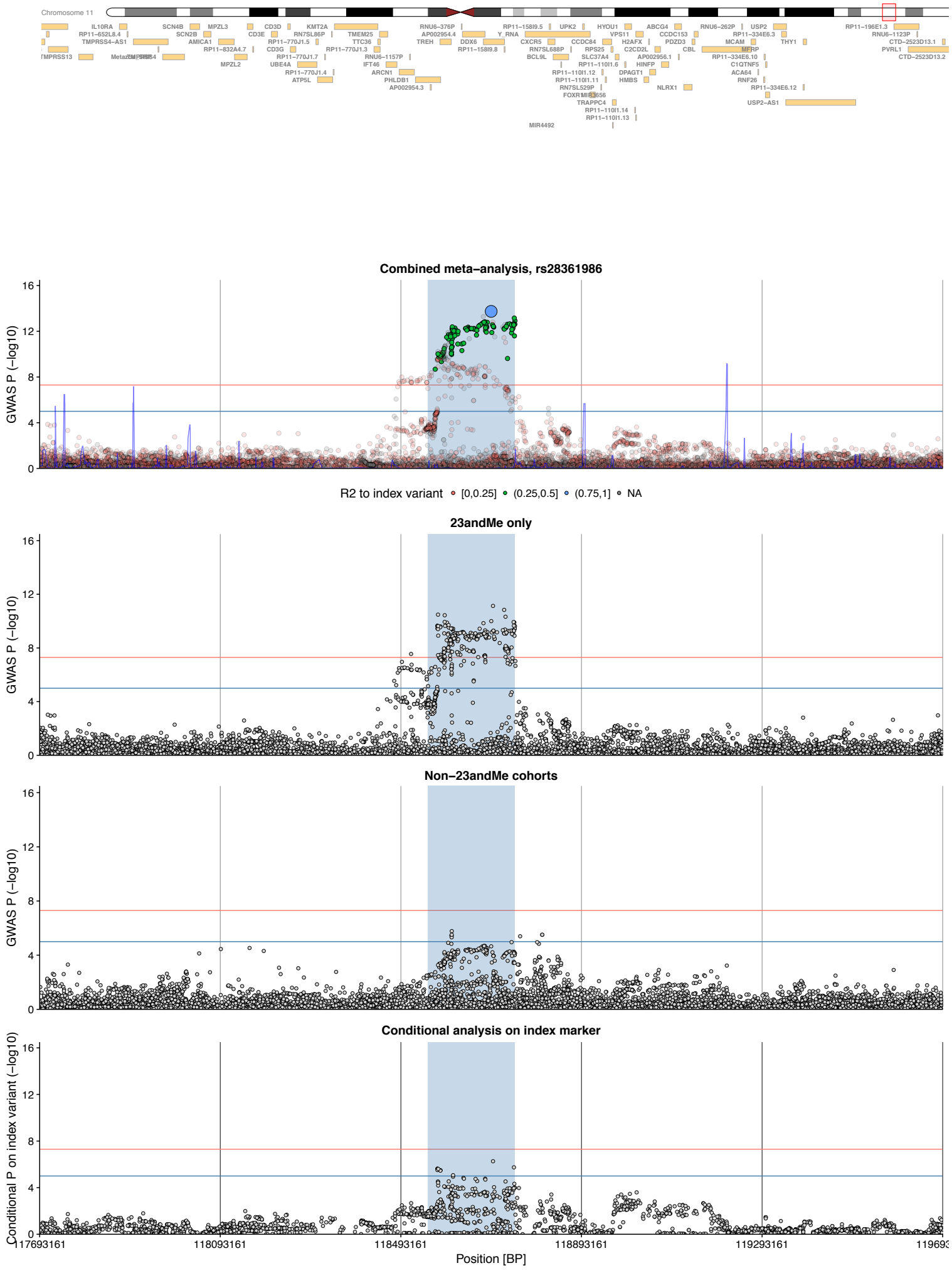

Chromosome 2
SF3B1
MOB4
AC011997.1
AC109589.1
AC019330.1
AC013264.2
SNORA4
HSPD1
MARS2
PLCL1
AC005235.1
AC019330.2
AC018717.1
ANKRD44−IT1
RNU6−1029P
RFTN2
AC010746.3
HSPE1−MOB4
BOLL
COQ10B
HSPE1
Combined meta−analysis, rs2164068
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
197943852
198343852
198743852
199143852
199543852
Position [BP]

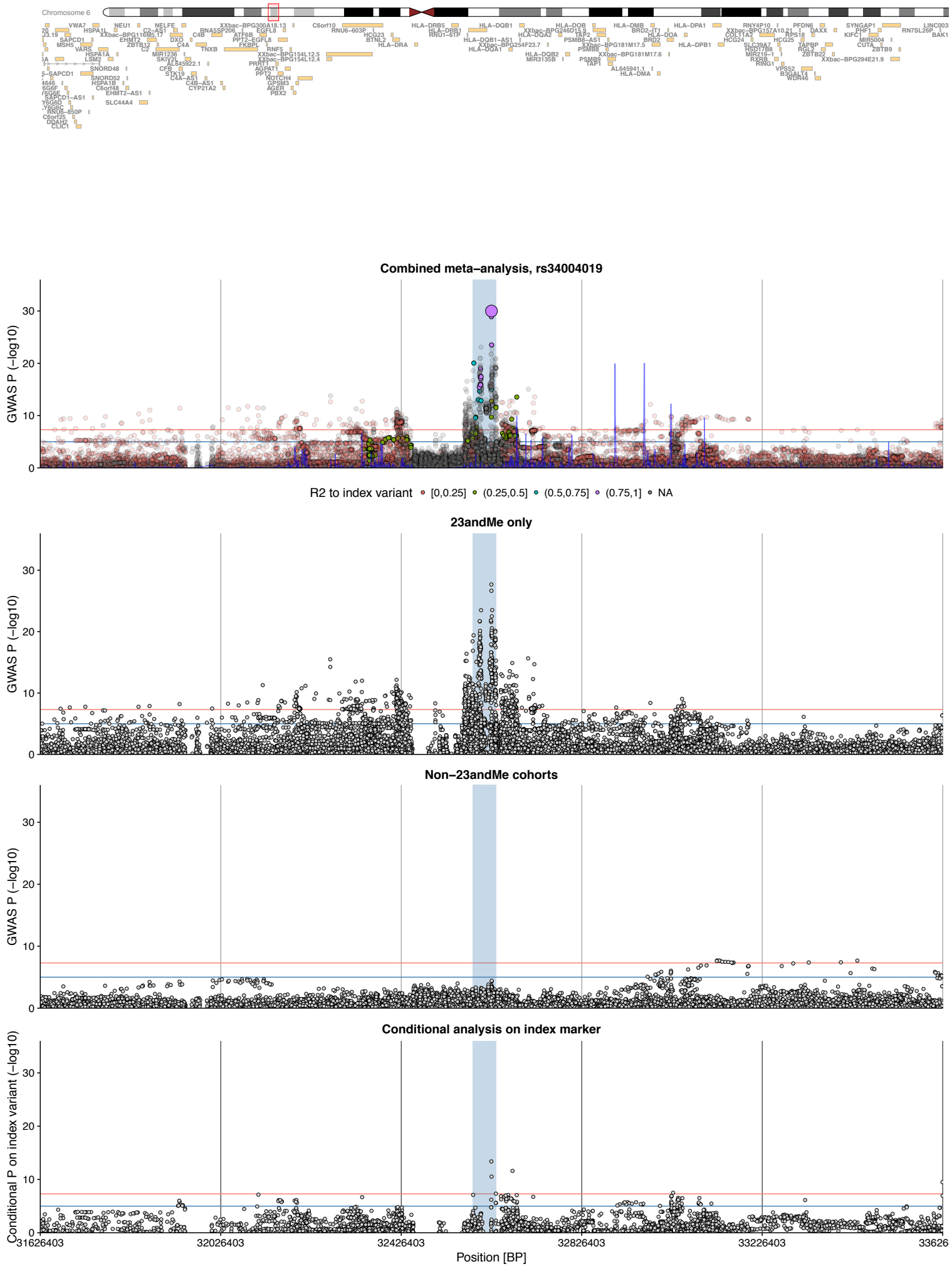

Chromosome 6
Combined meta−analysis, rs2428494
GWAS P (−log10)
R2 to index variant
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
30322197
30722197
31122197
31522197
31922197
32322
Position [BP]

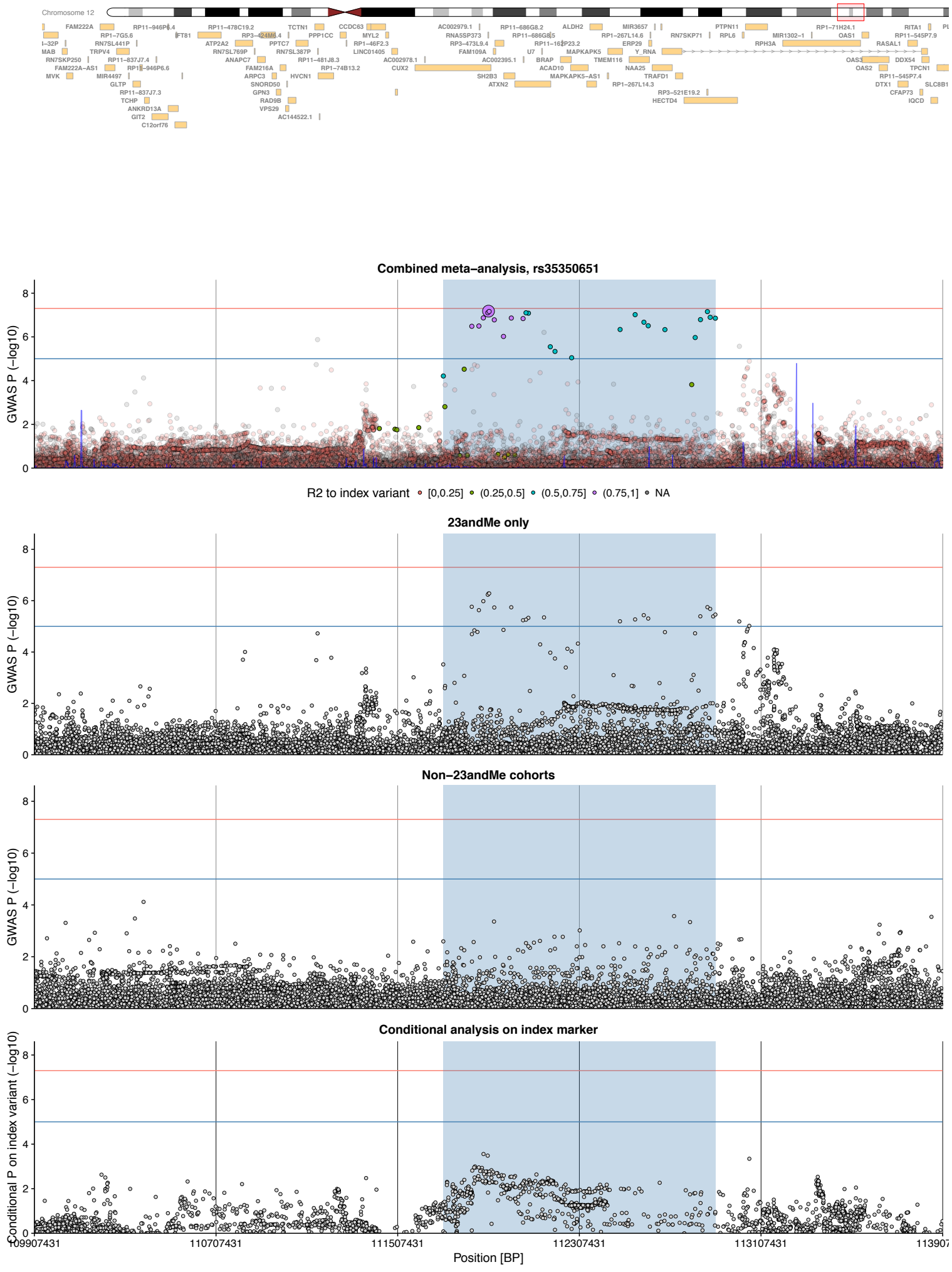

Chromosome 12
Combined meta−analysis, rs2461475
GWAS P (−log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
120290174
120690174
121090174
121490174
121890174
Position [BP]

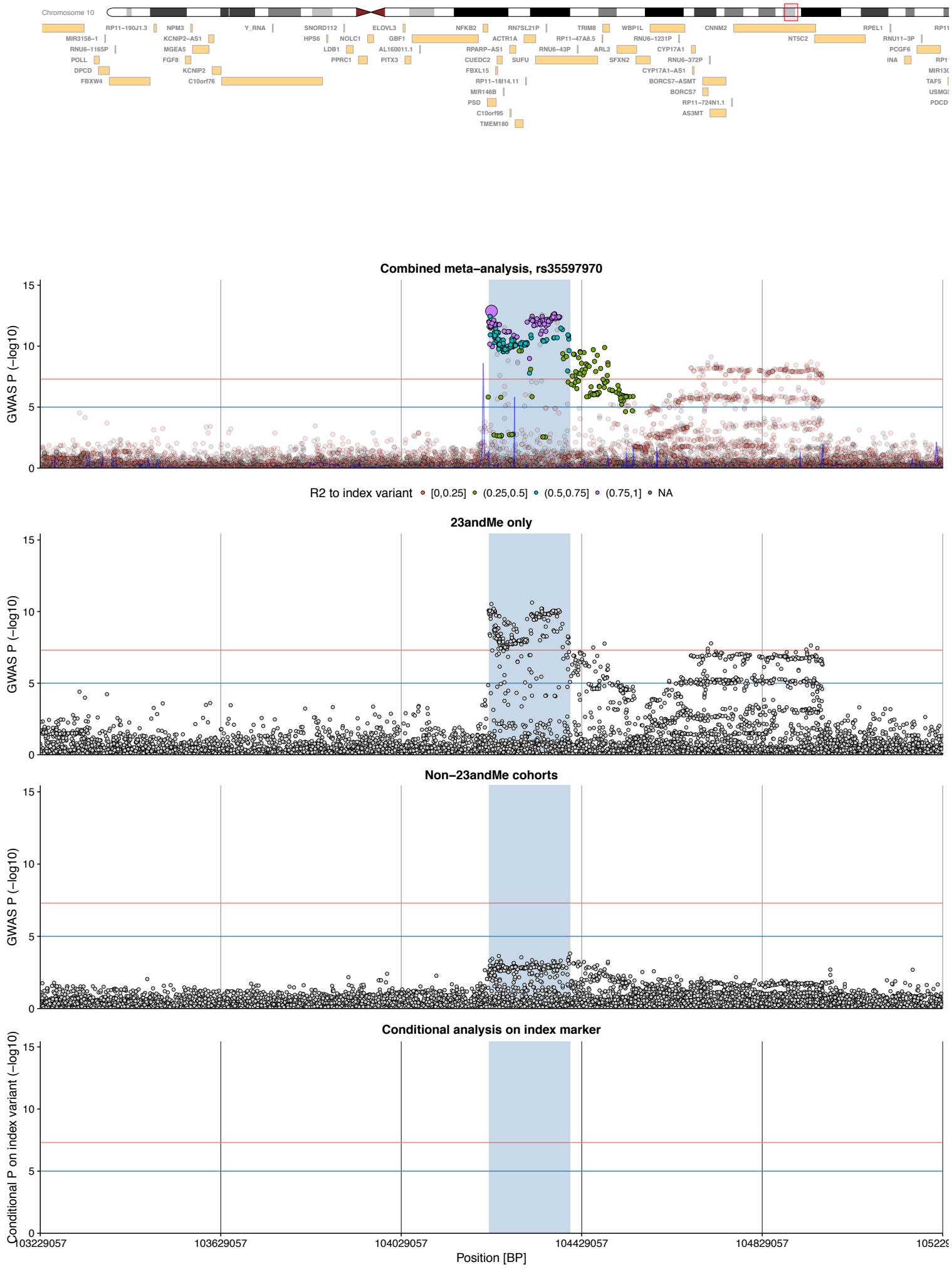

Chromosome 9
BARHL1
RP11−295G24.5
RP11−295G24.4
RALGDS
ABO
SURF4
SLC2A6
FAM163B
VAV2
RNU6ATAC
RNU5D−2P
DDX31
AK8
GTF3C5
GBGT1
SNORD36B
CACFD1
DBH−AS1
RP11−374P20.4
WDR5
TTF1
GTF3C4
RNU6−357P
RNU7−21P
CEL
OBP2B
SNORD36A
ADAMTSL2
DBH
LINC00094
CFAP77
TSC1
snoU13
SNORD36C
LL09NC01−254D11.1
LL09NC01−251B2.3
C9orf9
Y_RNA
TMEM8C
SARDH
BRD3
MIR548AW
SNORD24
LL09NC01−139C3.1
GFI1B
SURF6
MED22
SURF1
RPL7A
SURF2
STKLD1
ADAMTS13
REXO4
Combined meta−analysis, rs2519093
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
135141870
135541870
135941870
136341870
136741870
Position [BP]

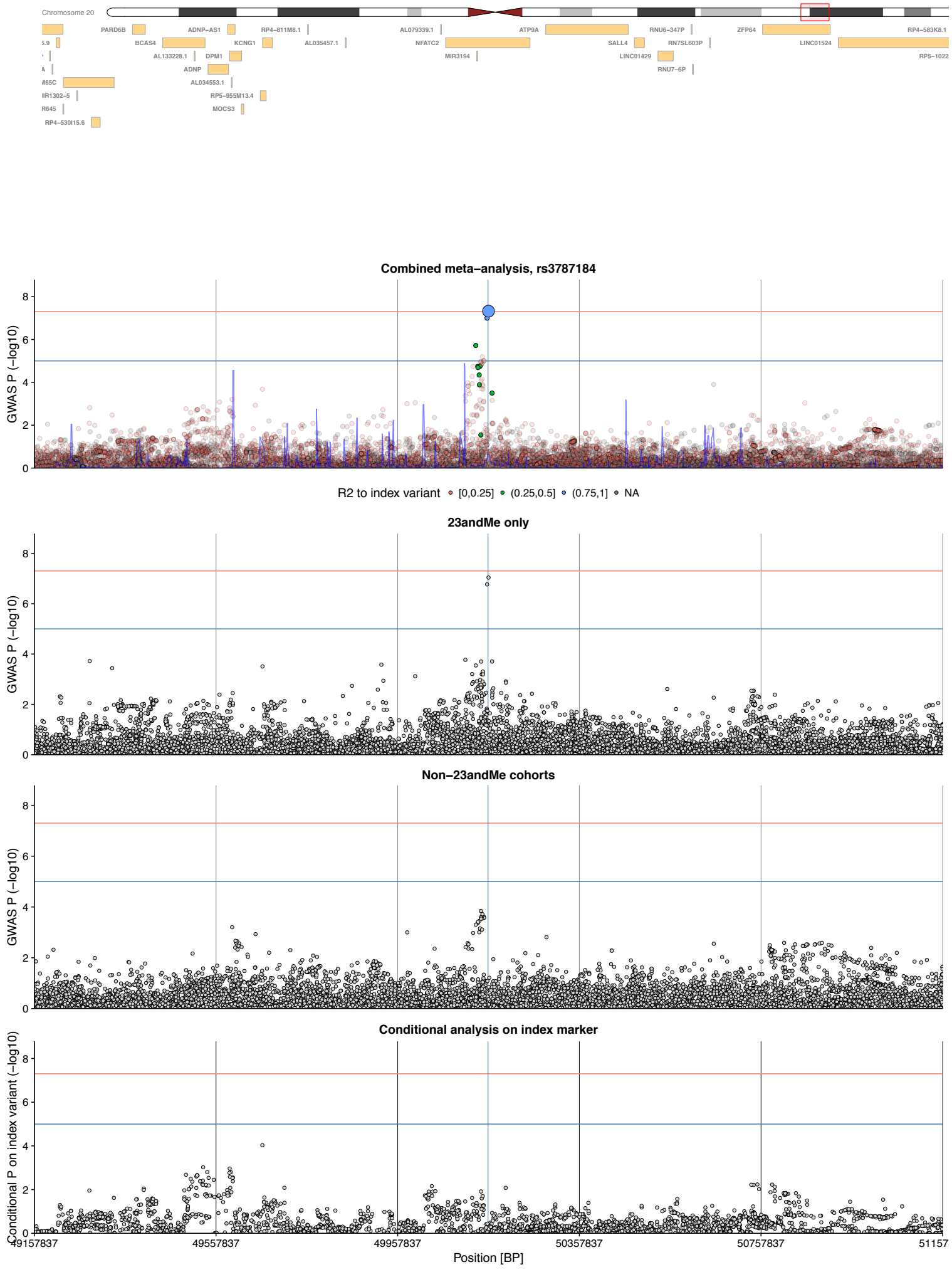

Chromosome 1
NEGR1–IT1
RNU6–1246P
RP11–25L19.1
RN7SKP19
NEGR1
RP4–660H19.1
RP11–82L20.1
RP11–262K1.1
Combined meta–analysis, rs2815765
GWAS P (–log10)
10.0
7.5
5.0
2.5
0.0
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non–23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (–log10)
71752230
72152230
72552230
72952230
73352230
73752
Position [BP]

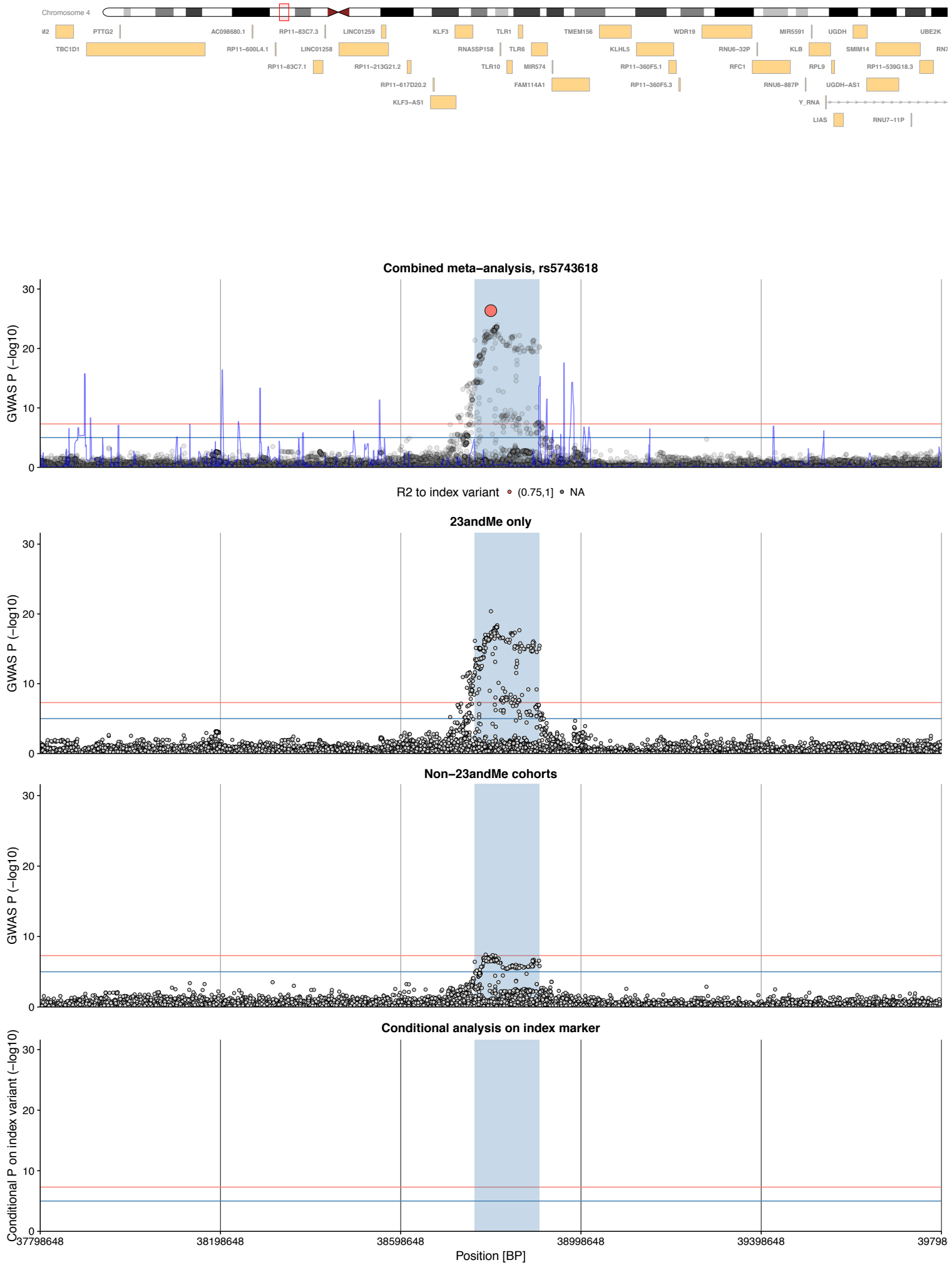

Chromosome 11
IL10RA
SCN4B
MPZL3
CD3D
KMT2A
RNU6−376P
RP11−158I9.5
UPK2
HYOU1
ABCG4
RNU6−262P
USP2
RP11−196E1.3
RP11−652L8.4
SCN2B
CD3E
RN7SL86P
TMEM25
AP002954.4
Y_RNA
VPS11
CCDC153
RP11−334E6.3
RNU6−1123P
TMPRSS4−AS1
AMICA1
RP11−770J1.5
TTC36
TREH
DDX6
CXCR5
CCDC84
H2AFX
PDZD3
MCAM
THY1
CTD−2523D13.1
RP11−832A4.7
CD3G
RP11−770J1.3
RP11−158I9.8
RN7SL688P
RPS25
C2CD2L
CBL
MFRP
PVRL1
TMPRSS13
RP11−770J1.7
RNU6−1157P
BCL9L
SLC37A4
AP002956.1
RP11−334E6.10
CTD−2523D13.2
MPZL2
UBE4A
IFT46
RP11−110I1.6
HINFP
C1QTNF5
RP11−770J1.4
ARCN1
RP11−110I1.12
DPAGT1
ACA64
ATP5L
PHLDB1
RP11−110I1.11
HMBS
RNF26
AP002954.3
RN7SL529P
NLRX1
RP11−334E6.12
TRAPPC4
USP2−AS1
RP11−110I1.14
RP11−110I1.13
MIR4492
Combined meta−analysis, rs28361986
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
Position [BP]

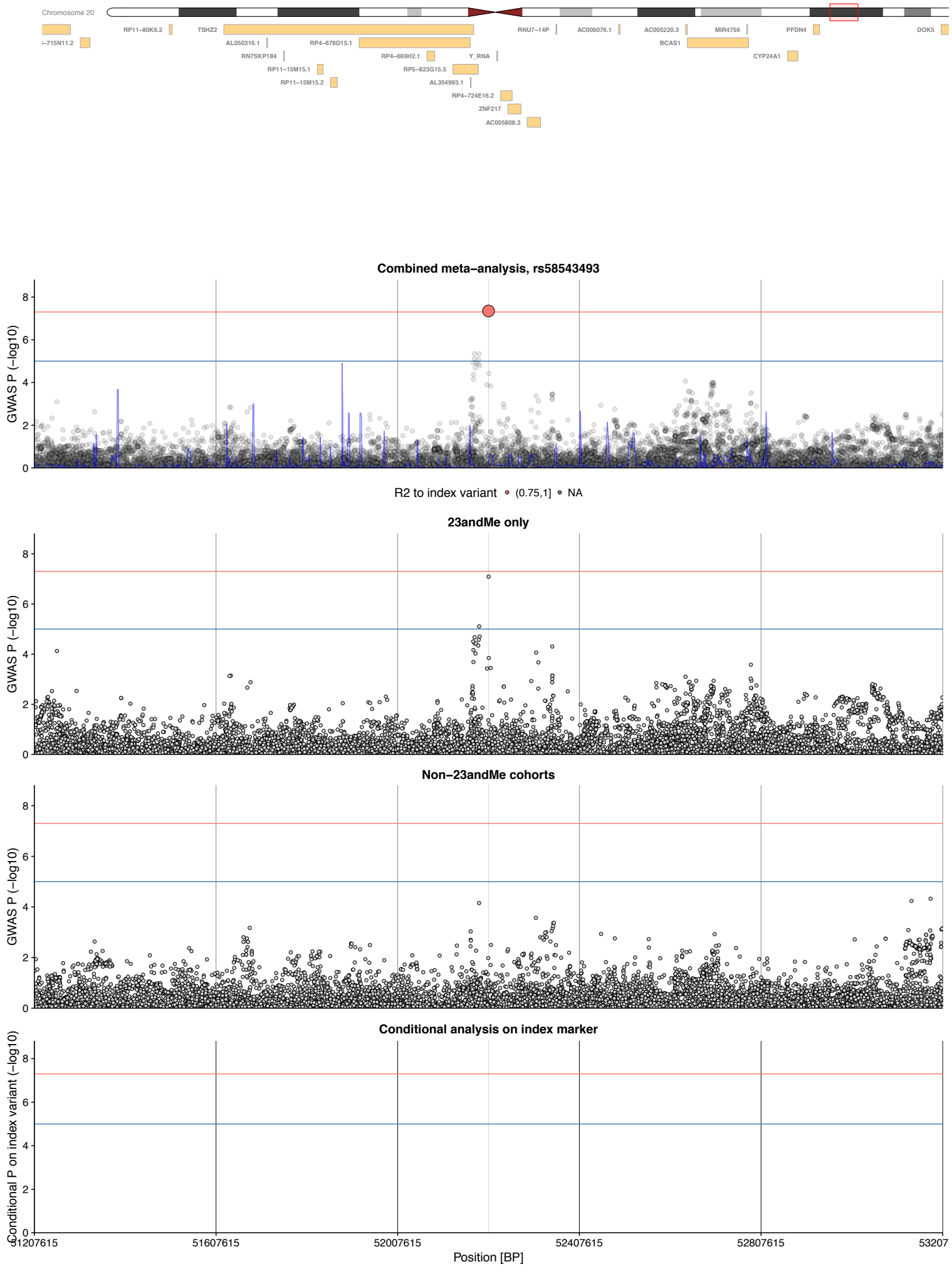

Chromosome 6
Combined meta−analysis, rs34004019
GWAS P (−log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
31626403
32026403
32426403
32826403
33226403
33626
Position [BP]

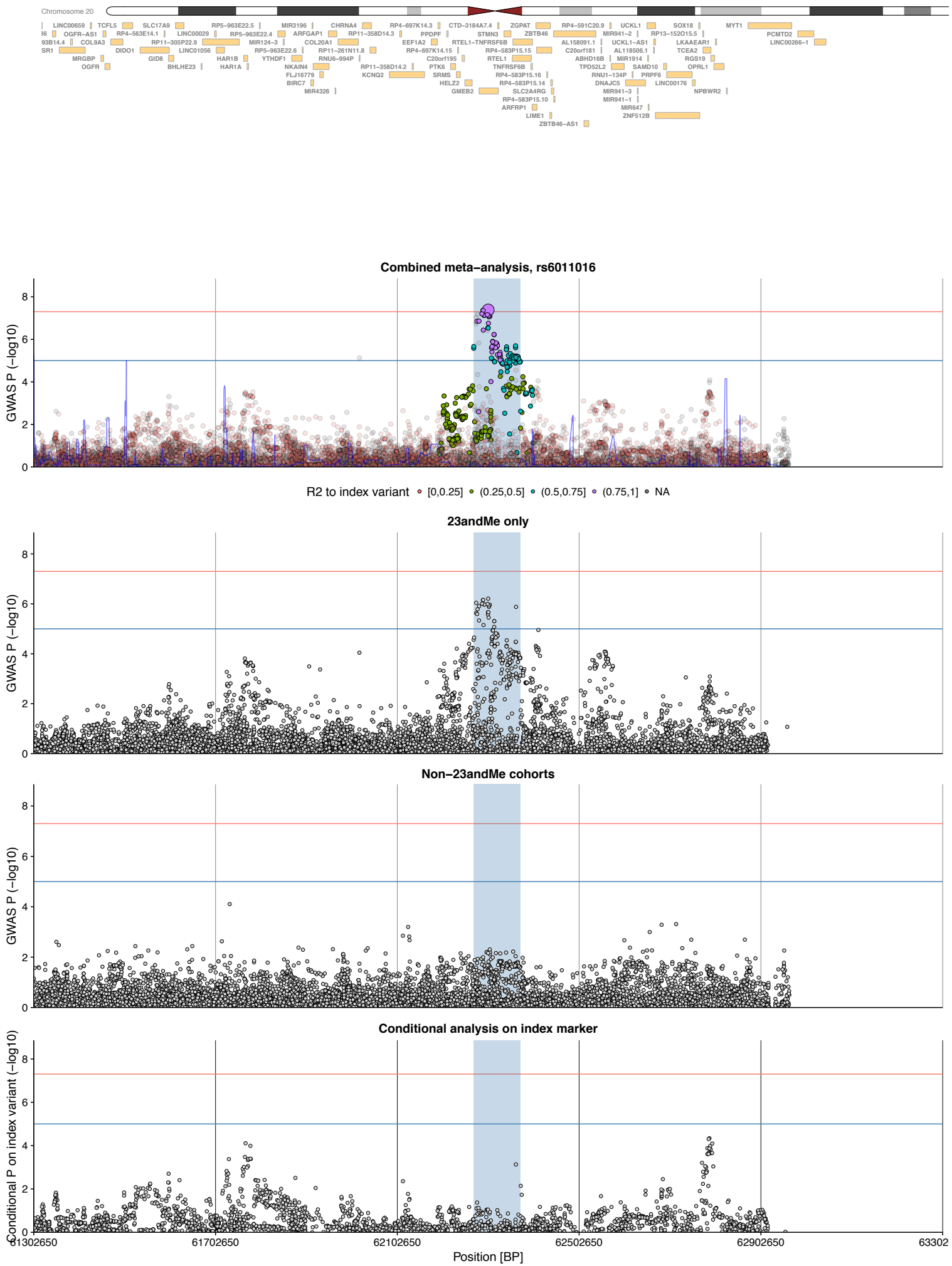

Chromosome 12
Combined meta−analysis, rs35350651
GWAS P (−log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
109907431
110707431
111507431
112307431
113107431
Position [BP]

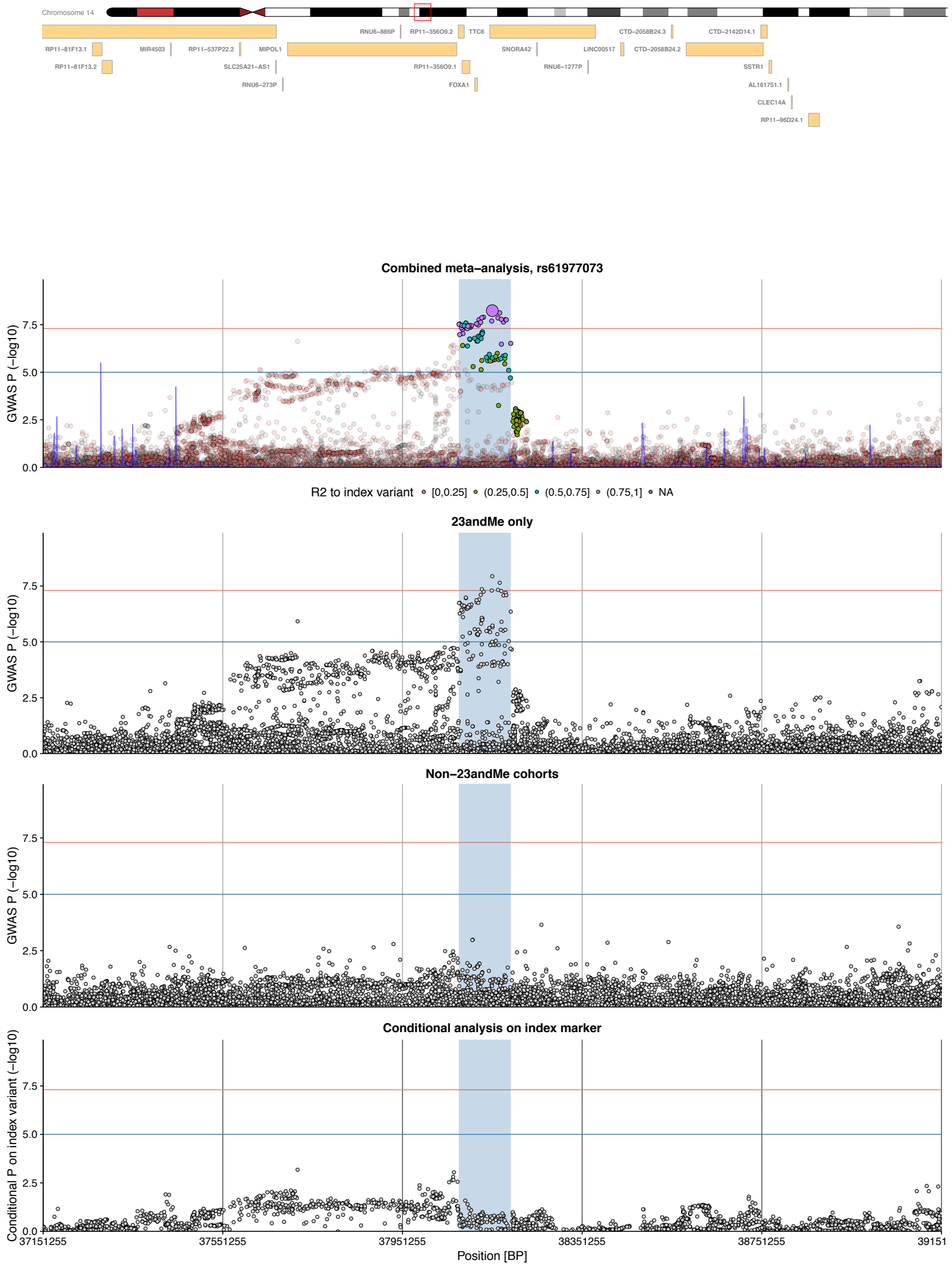

Chromosome 10
RP11−190J1.3
NPM3
Y_RNA
SNORD112
ELOVL3
NFKB2
RN7SL21P
TRIM8
WBP1L
CNNM2
RPEL1
MIR3158−1
KCNIP2−AS1
HPS6
NOLC1
GBF1
ACTR1A
RP11−47A8.5
RNU6−1231P
NT5C2
RNU11−3P
RNU6−1165P
MGEA5
LDB1
AL160011.1
RPARP−AS1
RNU6−43P
ARL3
CYP17A1
PCGF6
POLL
FGF8
PPRC1
PITX3
CUEDC2
SUFU
SFXN2
RNU6−372P
INA
DPCD
KCNIP2
FBXL15
CYP17A1−AS1
FBXW4
C10orf76
RP11−18I14.11
BORCS7−ASMT
TAF5
MIR146B
BORCS7
USMG5
PSD
RP11−724N1.1
PDCD7
C10orf95
AS3MT
TMEM180
Combined meta−analysis, rs35597970
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
103229057
103629057
104029057
104429057
104829057
Position [BP]

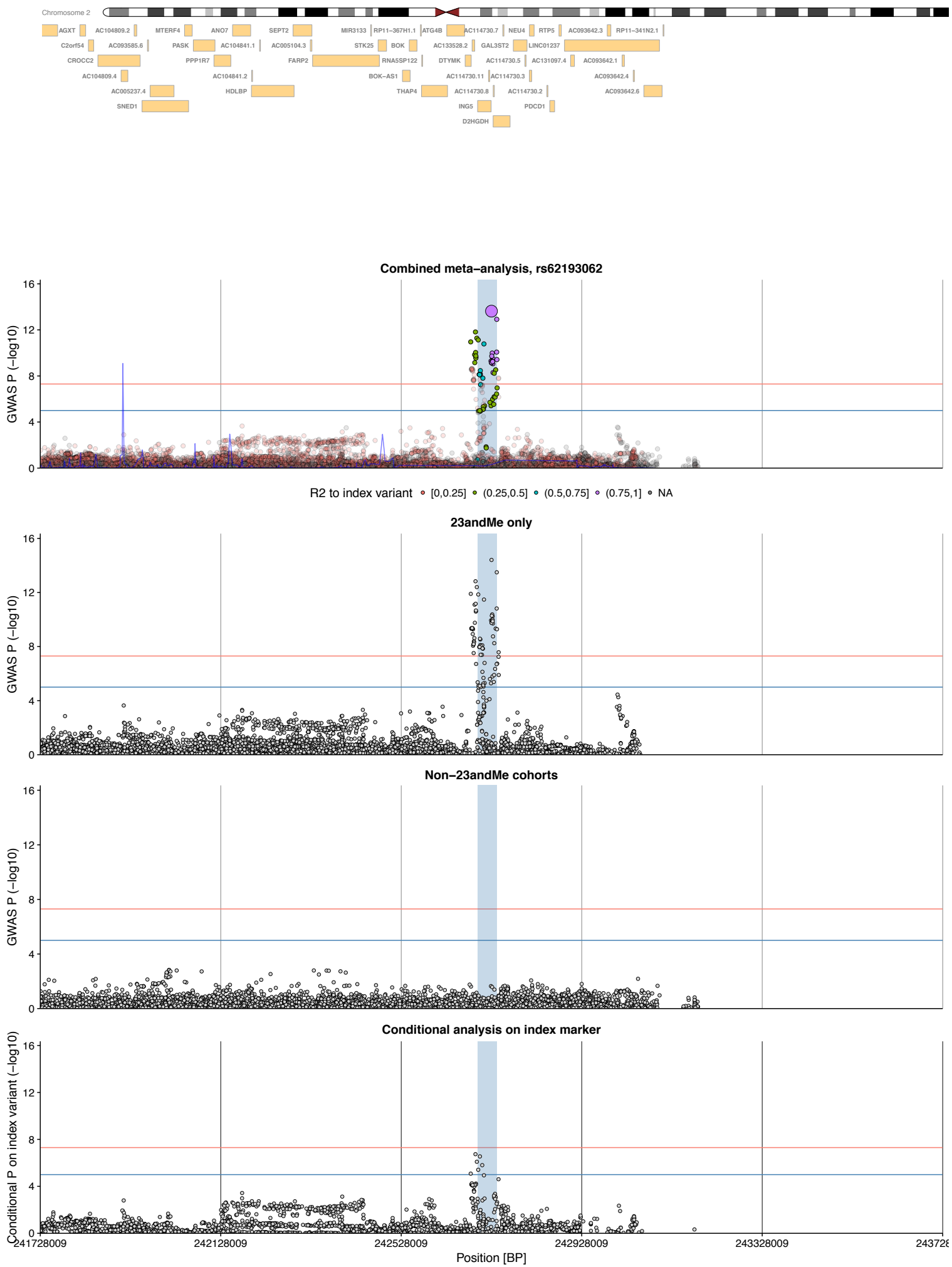

Chromosome 20
PARD6B
ADNP-AS1
RP4-811M8.1
AL079339.1
ATP9A
RNU6-347P
ZFP64
RP4-583K8.1
BCAS4
KCNG1
AL035457.1
NFATC2
SALL4
RN7SL603P
LINC01524
AL133228.1
DPM1
MIR3194
LINC01429
RP5-1022
ADNP
RNU7-6P
AL034553.1
RP5-955M13.4
MOCS3
RP4-530I15.6
Combined meta-analysis, rs3787184
GWAS P (-log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.75,1]
NA
23andMe only
Non-23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (-log10)
49157837
49557837
49957837
50357837
50757837
51157
Position [BP]

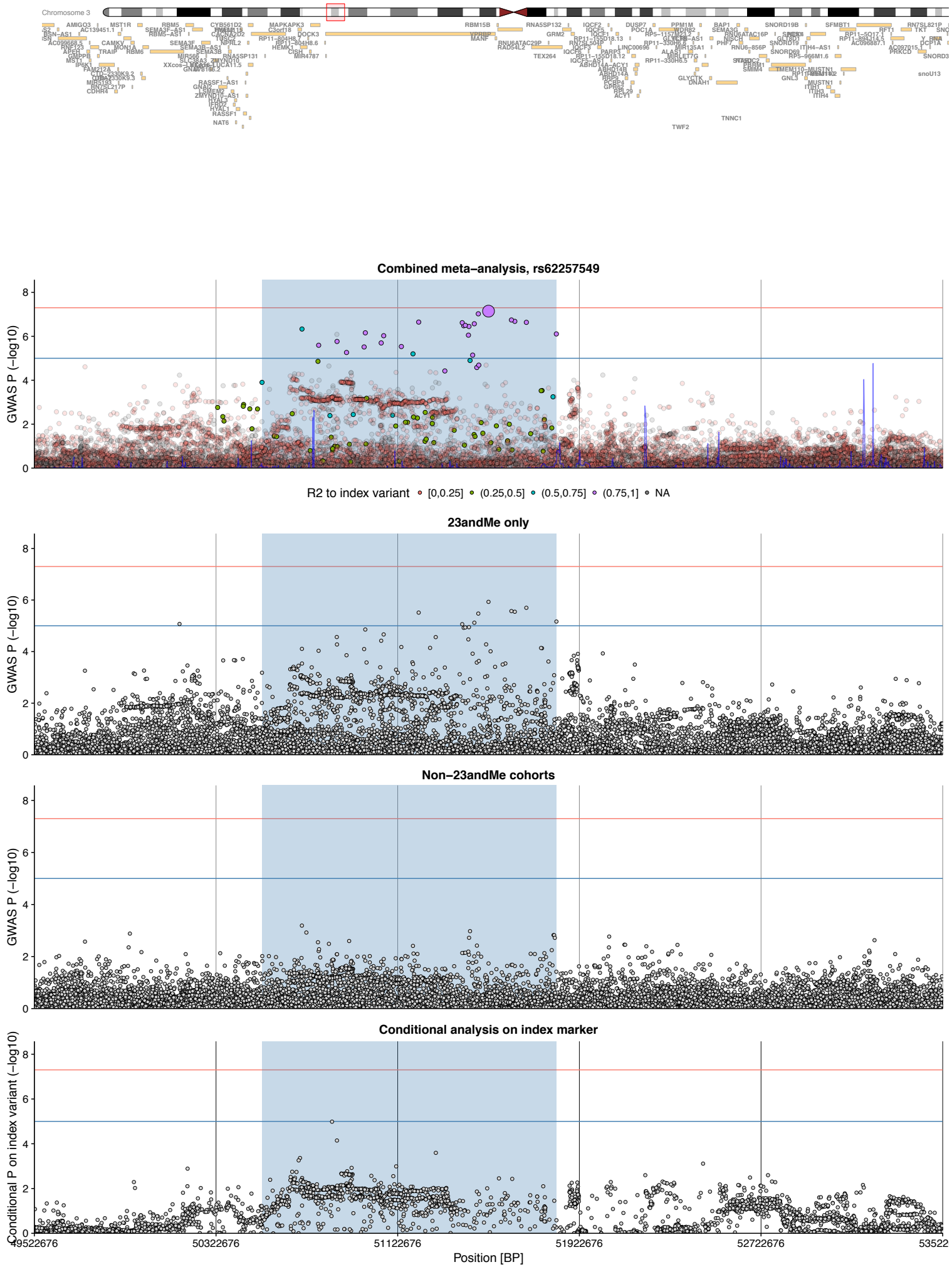

Chromosome 4
PTTG2
AC098680.1
RP11−83C7.3
LINC01259
KLF3
TLR1
TMEM156
WDR19
MIR5591
UGDH
UBE2K
TBC1D1
RP11−600L4.1
LINC01258
RNA5SP158
TLR6
KLHL5
RNU6−32P
KLB
SMIM14
RP11−83C7.1
RP11−213G21.2
TLR10
MIR574
RP11−360F5.1
RFC1
RPL9
RP11−539G18.3
RP11−617D20.2
FAM114A1
RP11−360F5.3
RNU6−887P
UGDH−AS1
KLF3−AS1
Y_RNA
LIAS
RNU7−11P
Combined meta−analysis, rs5743618
GWAS P (−log10)
R2 to index variant
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
37798648
38198648
38598648
38998648
39398648
39798
Position [BP]

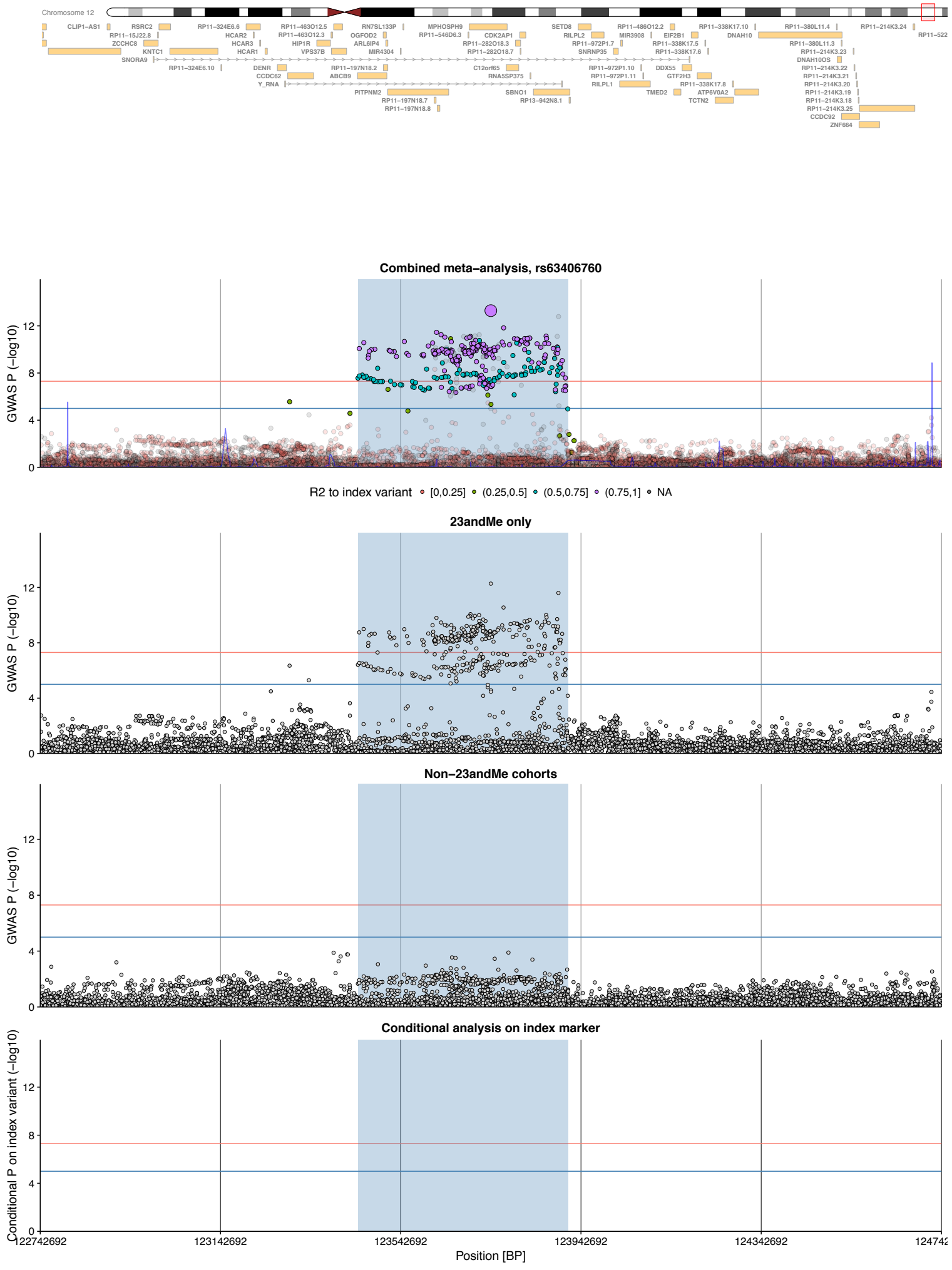

Chromosome 20
RP11−80K6.2
TSHZ2
RNU7−14P
AC006076.1
AC005220.3
MIR4756
PFDN4
DOK5
AL050316.1
RP4−678D15.1
BCAS1
RN7SKP184
RP4−669H2.1
Y_RNA
CYP24A1
RP11−15M15.1
RP5−823G15.5
RP11−15M15.2
AL354993.1
RP4−724E16.2
ZNF217
AC005808.3
Combined meta−analysis, rs58543493
GWAS P (−log10)
R2 to index variant
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
51207615
51607615
52007615
52407615
52807615
53207
Position [BP]

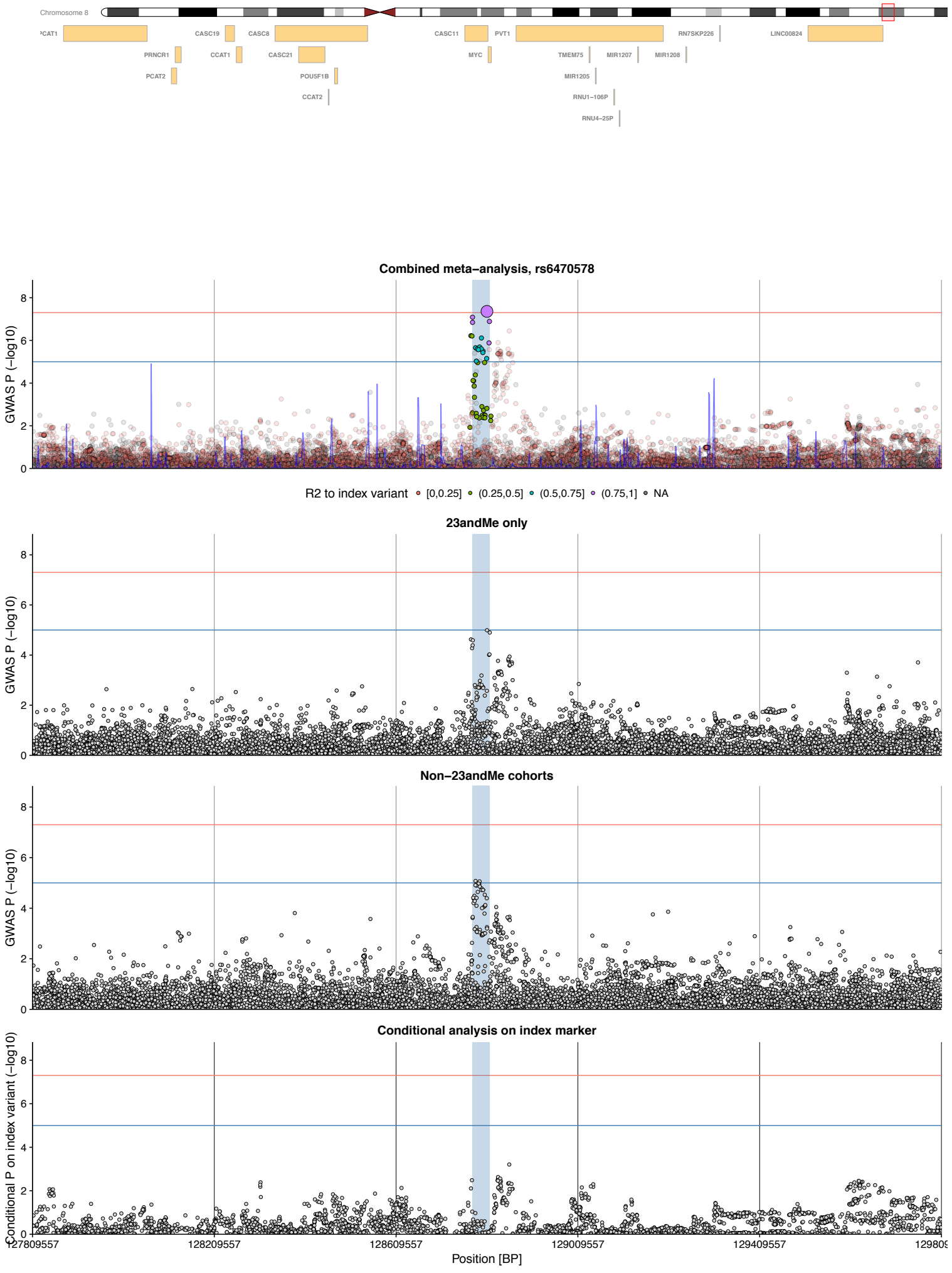

Chromosome 20
LINC00659 TCFL5 SLC17A9 RP5−963E22.5 MIR3196 CHRNA4 RP4−697K14.3 CTD−3184A7.4 ZGPAT RP4−591C20.9 UCKL1 SOX18 MYT1
OGFR−AS1 RP4−563E14.1 LINC00029 RP5−963E22.4 ARFGAP1 RP11−358D14.3 PPDPF STMN3 ZBTB46 MIR941−2 RP13−152O15.5 PCMTD2
93B14.4 COL9A3 RP11−305P22.9 MIR124−3 COL20A1 EEF1A2 RTEL1−TNFRSF6B AL158091.1 UCKL1−AS1 LKAAEAR1 LINC00266−1
SR1 DIDO1 LINC01056 RP5−963E22.6 RP11−261N11.8 RP4−697K14.15 RP4−583P15.15 C20orf181 AL118506.1 TCEA2
MRGBP GID8 HAR1B YTHDF1 RNU6−994P C20orf195 RTEL1 ABHD16B MIR1914 RGS19
OGFR BHLHE23 HAR1A NKAIN4 RP11−358D14.2 PTK6 TNFRSF6B TPD52L2 SAMD10 OPRL1
FLJ16779 KCNQ2 SRMS RP4−583P15.16 RNU1−134P PRPF6
BIRC7 HELZ2 RP4−583P15.14 DNAJC5 LINC00176
MIR4326 GMEB2 SLC2A4RG MIR941−3 NPBWR2
RP4−583P15.10 MIR941−1
ARFRP1 MIR647
LIME1 ZNF512B
ZBTB46−AS1
Combined meta−analysis, rs6011016
GWAS P (−log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
GWAS P (−log10)
Non−23andMe cohorts
GWAS P (−log10)
Conditional analysis on index marker
Conditional P on index variant (−log10)
61302650
61702650
62102650
62502650
62902650
63302
Position [BP]

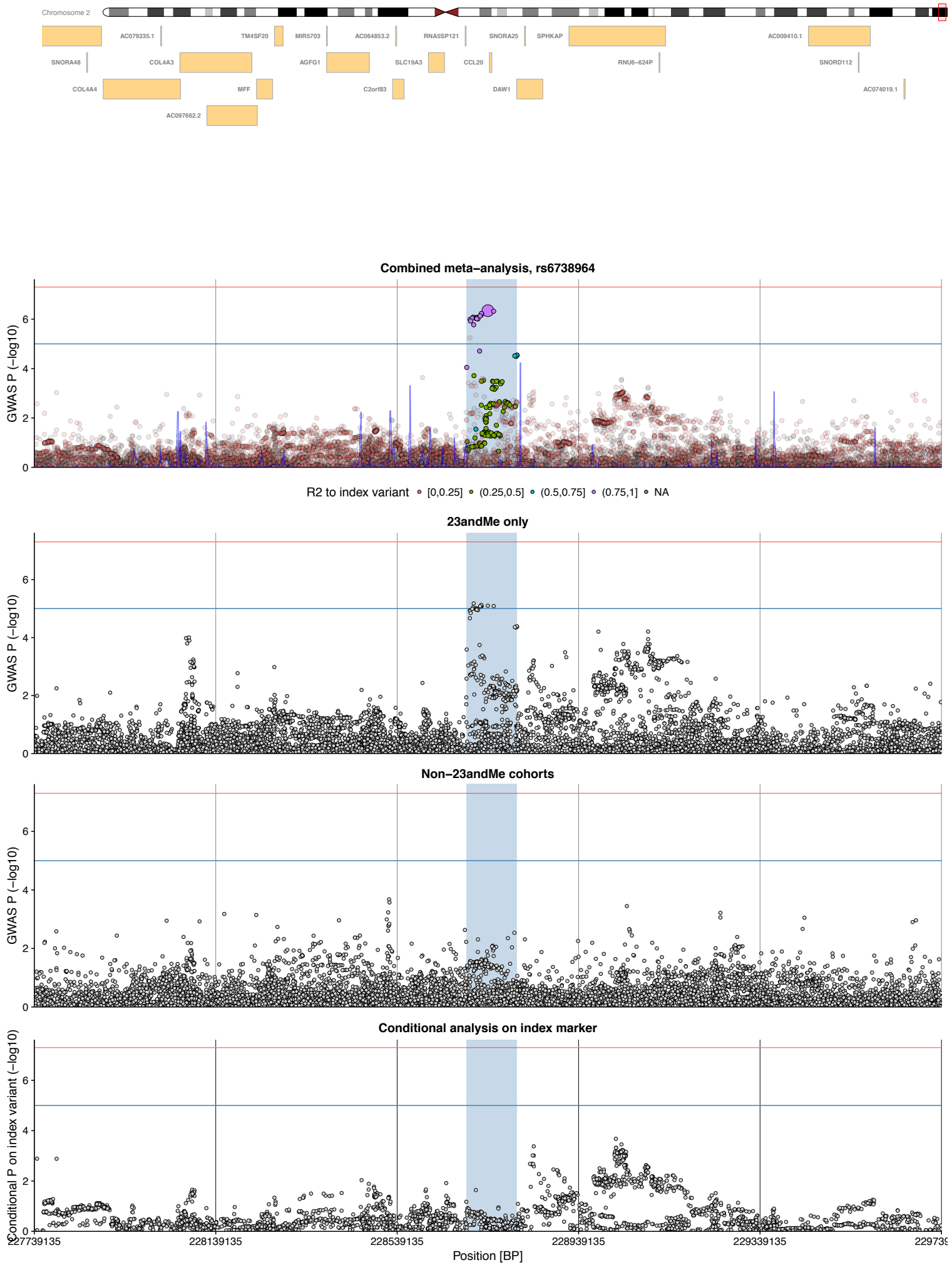

Chromosome 14
RNU6−886P
RP11−356O9.2
TTC6
CTD−2058B24.3
CTD−2142D14.1
RP11−81F13.1
MIR4503
RP11−537P22.2
MIPOL1
SNORA42
LINC00517
CTD−2058B24.2
RP11−81F13.2
SLC25A21−AS1
RP11−356O9.1
RNU6−1277P
SSTR1
RNU6−273P
FOXA1
AL161751.1
CLEC14A
RP11−96D24.1
Combined meta−analysis, rs61977073
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
37151255
37551255
37951255
38351255
38751255
39151
Position [BP]

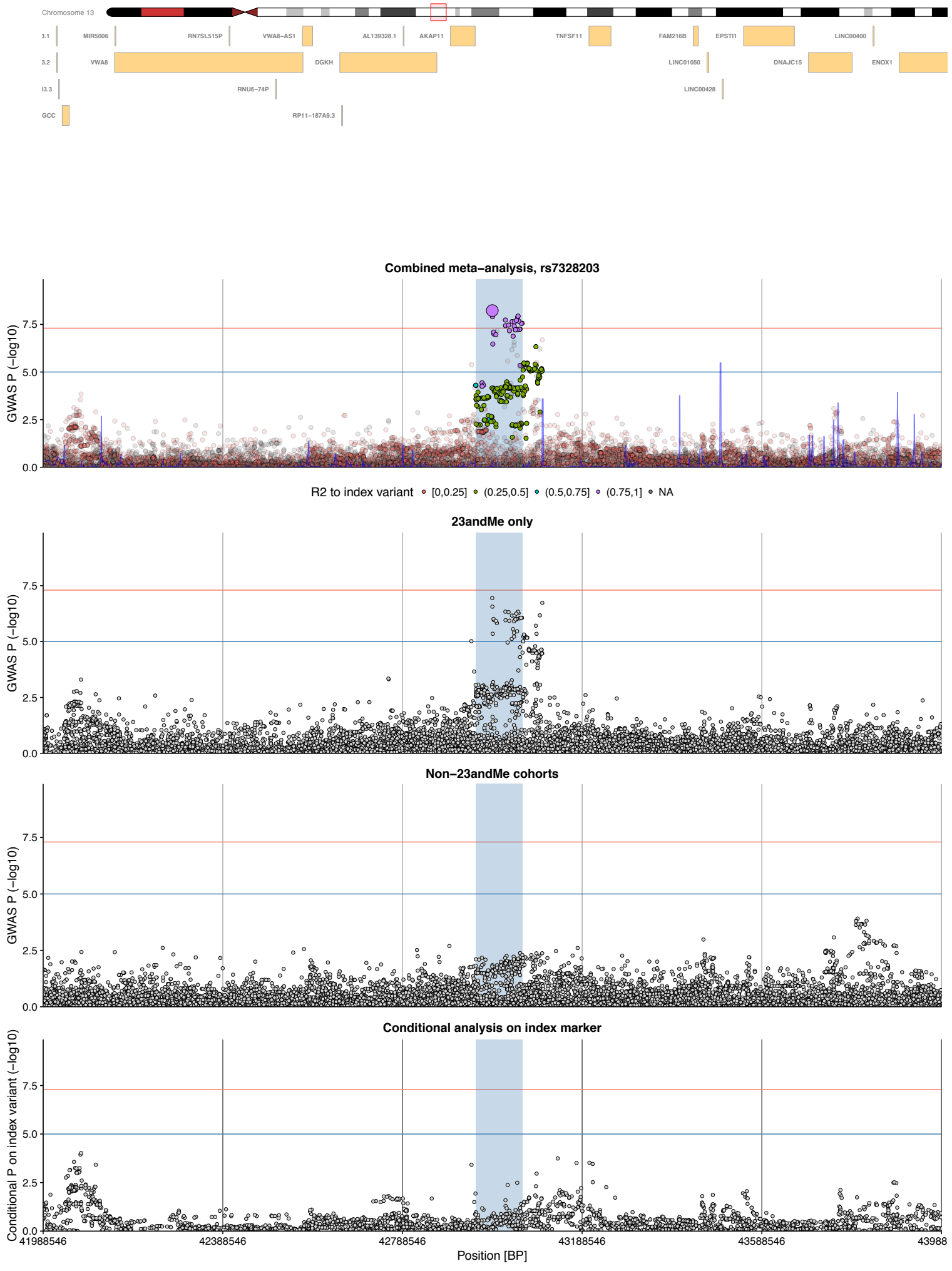

Chromosome 2
AGXT
AC104809.2
MTERF4
ANO7
SEPT2
MIR3133
RP11−367H1.1
ATG4B
AC114730.7
NEU4
RTP5
AC093642.3
RP11−341N2.1
C2orf54
AC093585.6
PASK
AC104841.1
AC005104.3
STK25
BOK
AC133528.2
GAL3ST2
LINC01237
CROCC2
PPP1R7
FARP2
RNA5SP122
DTYMK
AC114730.5
AC131097.4
AC093642.1
AC104809.4
AC104841.2
BOK−AS1
AC114730.11
AC114730.3
AC093642.4
AC005237.4
HDLBP
THAP4
AC114730.8
AC114730.2
AC093642.6
SNED1
ING5
PDCD1
D2HGDH
Combined meta−analysis, rs62193062
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
241728009
242128009
242528009
242928009
243328009
Position [BP]

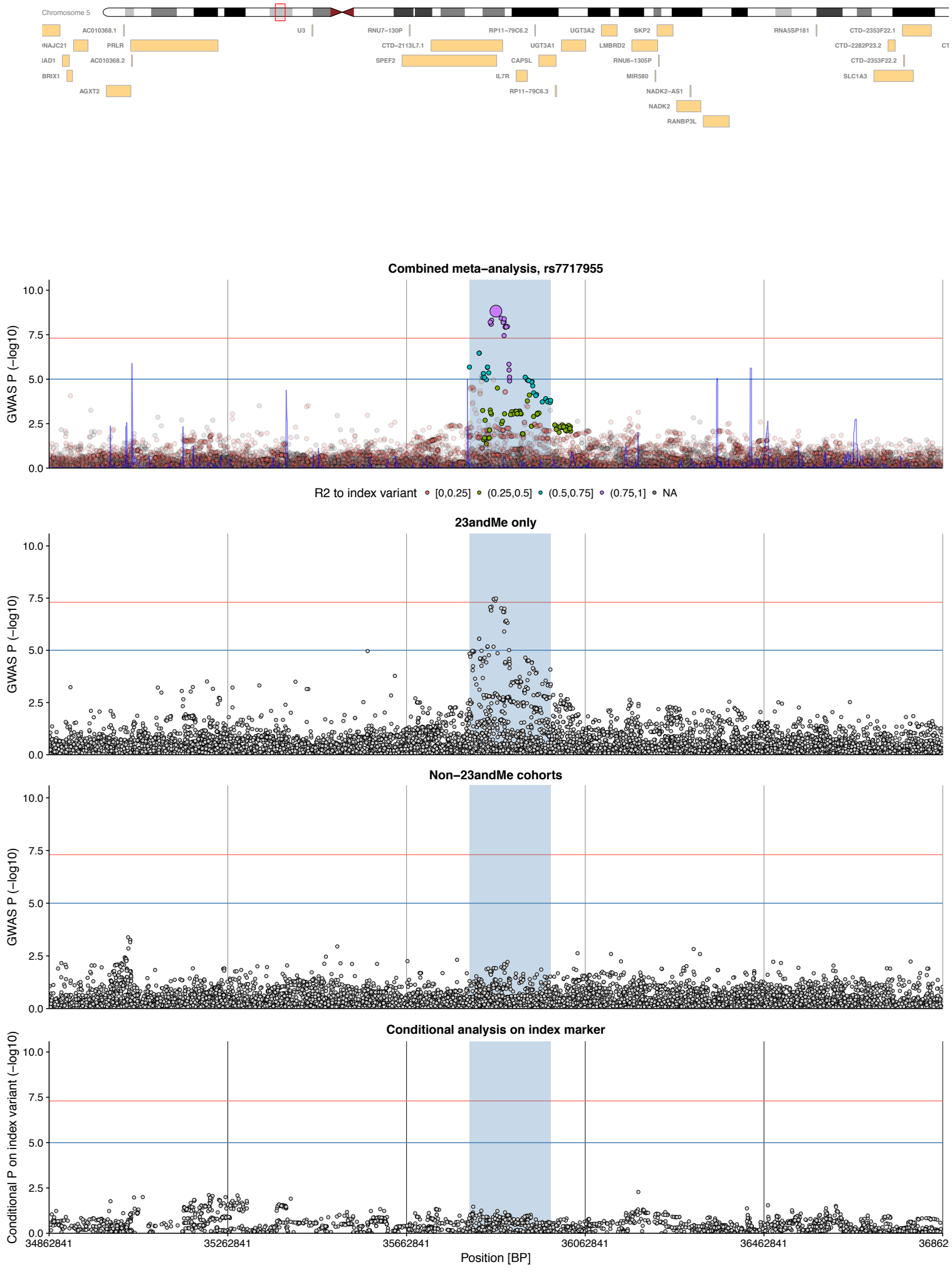

Chromosome 3
Combined meta−analysis, rs62257549
GWAS P (−log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
49522676
50322676
51122676
51922676
52722676
53522
Position [BP]

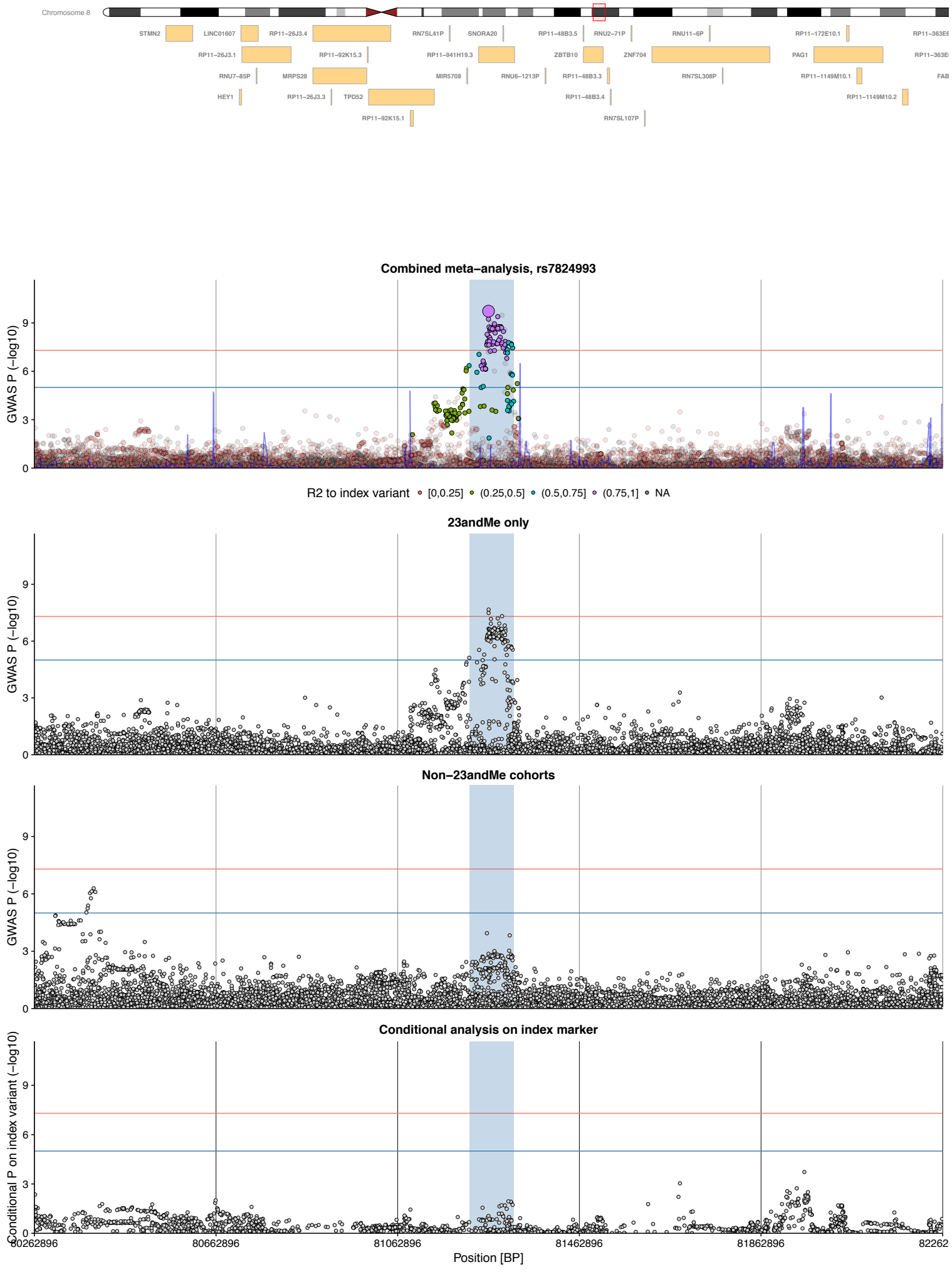

Chromosome 12
Combined meta−analysis, rs63406760
GWAS P (−log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
122742692
123142692
123542692
123942692
124342692
Position [BP]

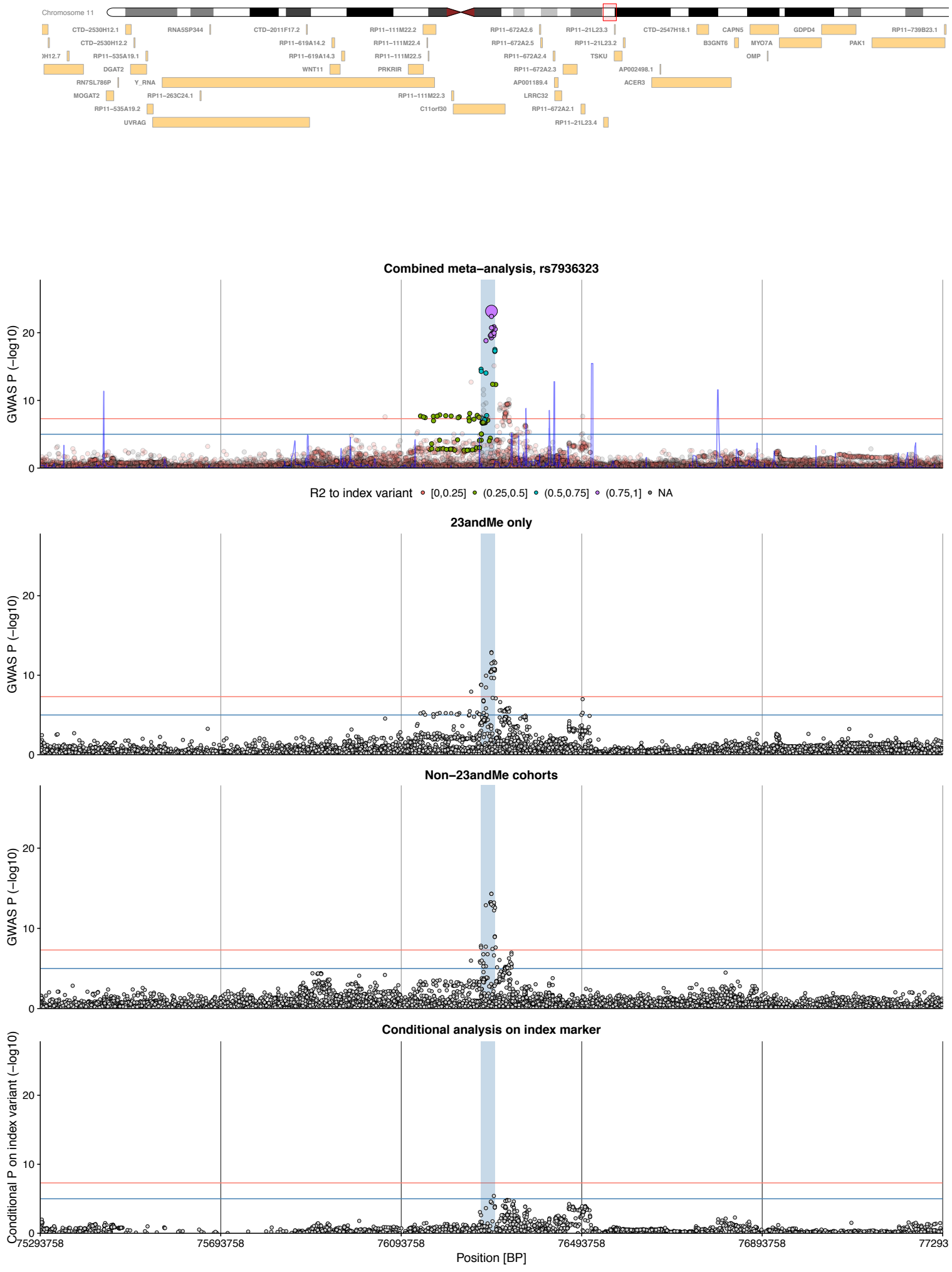

Chromosome 8
CASC19
CASC8
CASC11
PVT1
RN7SKP226
LINC00824
PRNCR1
CCAT1
CASC21
MYC
TMEM75
MIR1207
MIR1208
PCAT2
POU5F1B
MIR1205
CCAT2
RNU1–106P
RNU4–25P
Combined meta–analysis, rs6470578
GWAS P (–log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non–23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (–log10)
127809557
128209557
128609557
129009557
129409557
Position [BP]

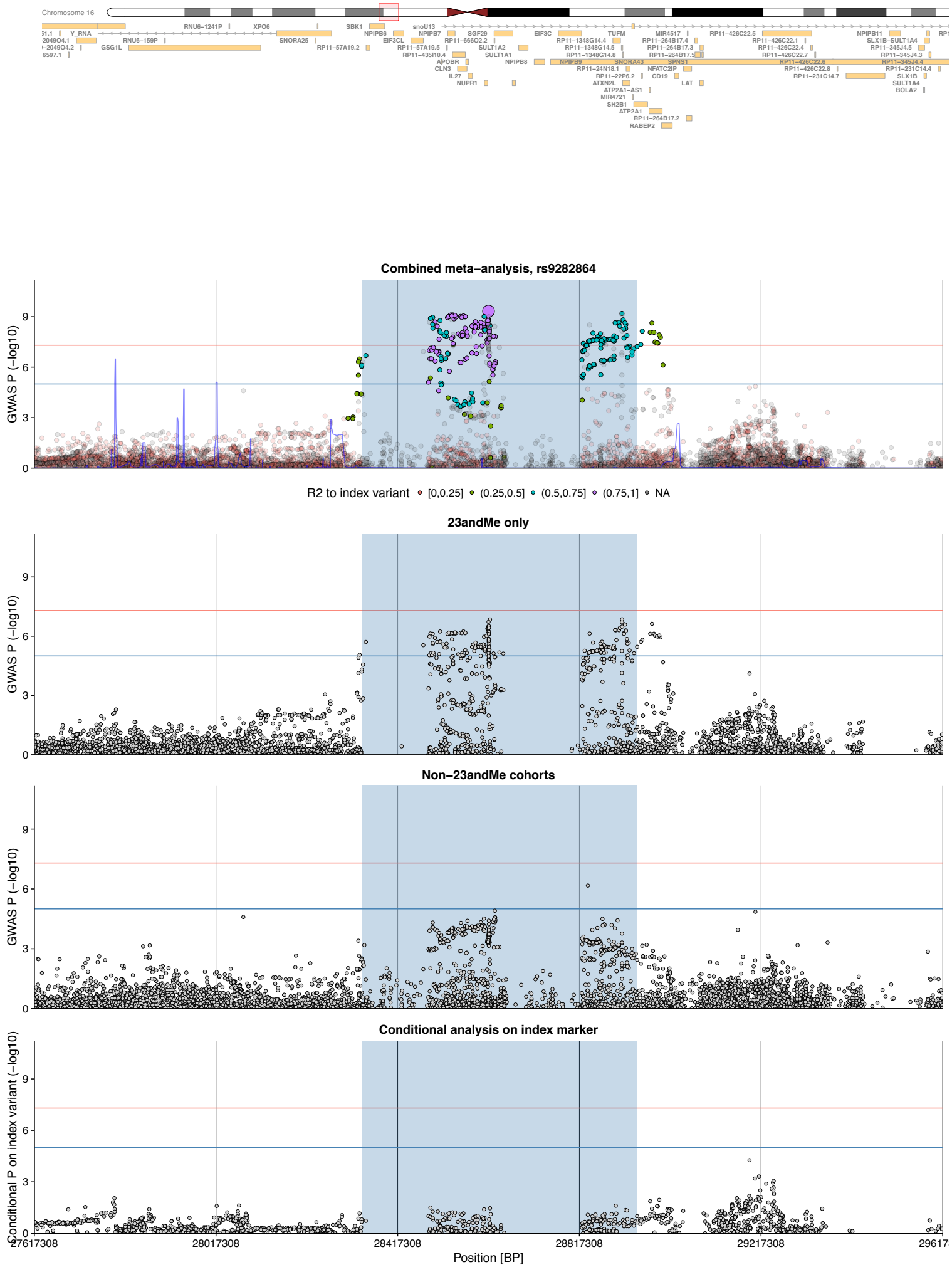

Chromosome 2
AC079235.1
TM4SF20
MIR5703
AC064853.2
RNA5SP121
SNORA25
SPHKAP
AC009410.1
SNORA48
COL4A3
AGFG1
SLC19A3
CCL20
RNU6−624P
SNORD112
COL4A4
MFF
C2orf83
DAW1
AC074019.1
AC097662.2
Combined meta−analysis, rs6738964
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
227739135
228139135
228539135
228939135
229339135
Position [BP]

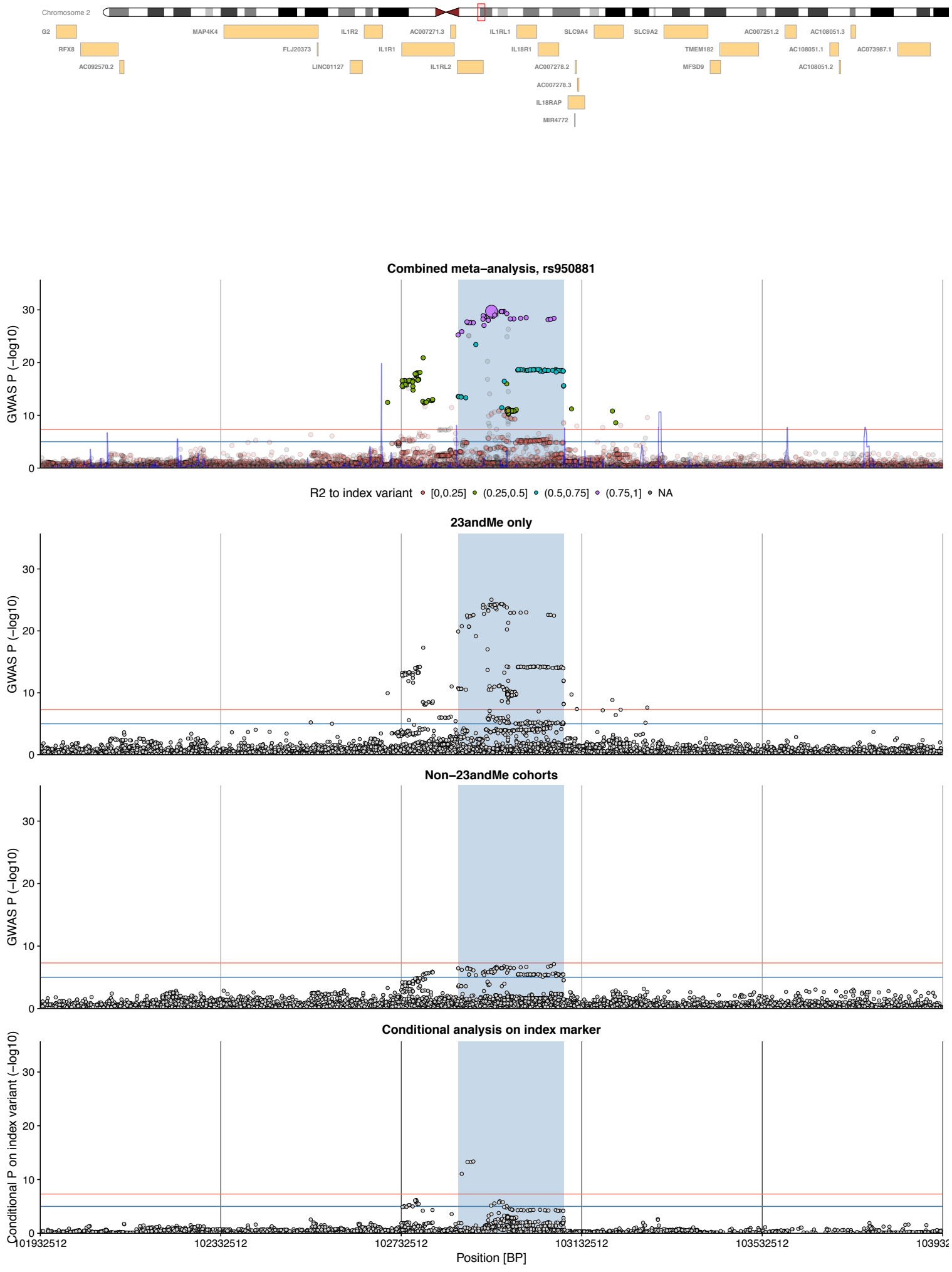

Chromosome 13
MIR5006
RN7SL515P
VWA8-AS1
AL139328.1
AKAP11
TNFSF11
FAM216B
EPSTI1
LINC00400
VWA8
DGKH
LINC01050
DNAJC15
ENOX1
RNU6-74P
LINC00428
RP11-187A9.3
Combined meta-analysis, rs7328203
GWAS P (-log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non-23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (-log10)
41988546
42388546
42788546
43188546
43588546
Position [BP]

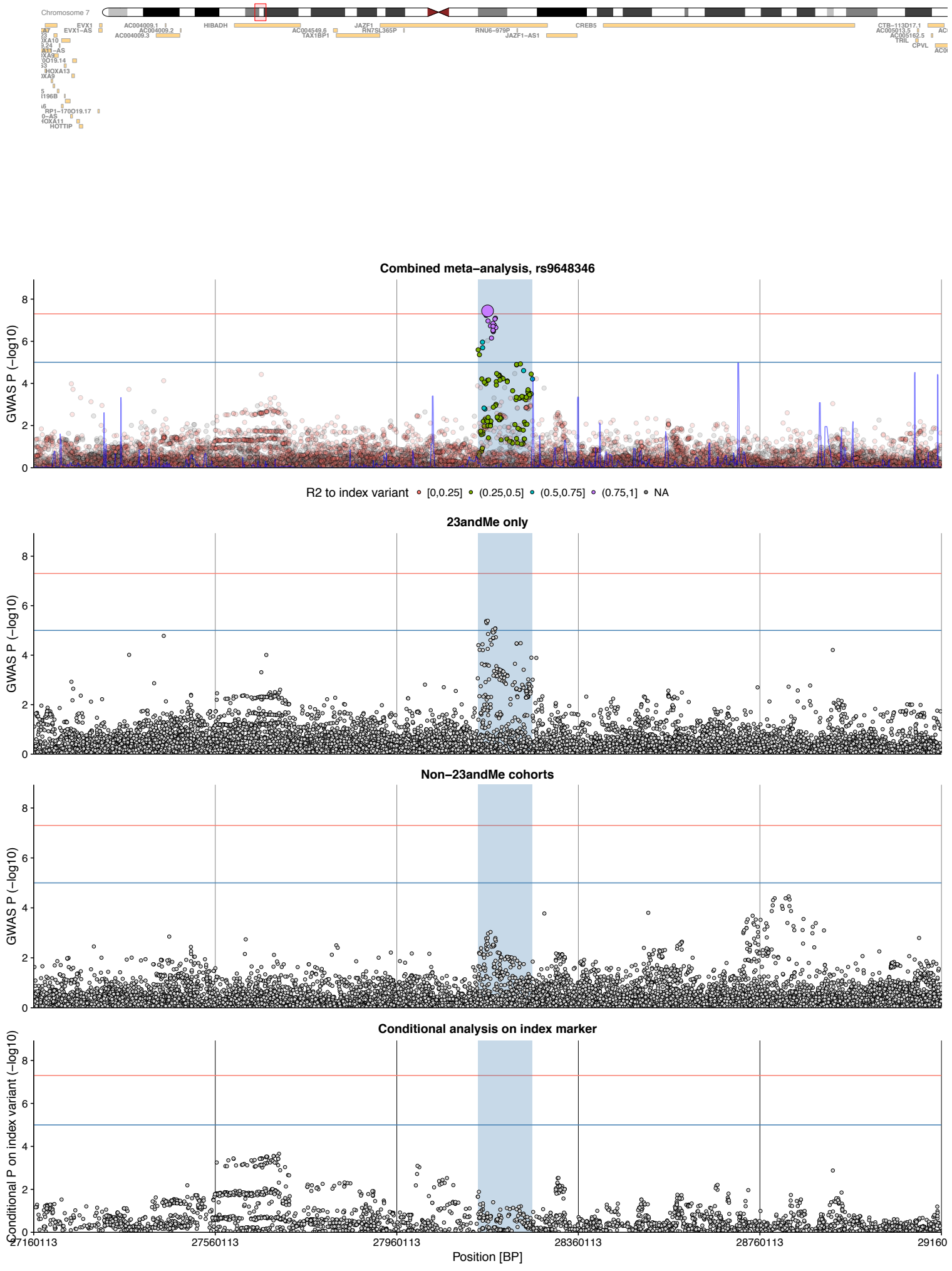

Chromosome 5
AC010368.1
U3
RNU7−130P
RP11−79C6.2
UGT3A2
SKP2
RNA5SP181
CTD−2353F22.1
PRLR
CTD−2113L7.1
UGT3A1
LMBRD2
CTD−2282P23.2
AC010368.2
SPEF2
CAPSL
RNU6−1305P
CTD−2353F22.2
BRIX1
IL7R
MIR580
SLC1A3
AGXT2
RP11−79C6.3
NADK2−AS1
NADK2
RANBP3L
Combined meta−analysis, rs7717955
GWAS P (−log10)
10.0
7.5
5.0
2.5
0.0
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
34862841
35262841
35662841
36062841
36462841
36862
Position [BP]

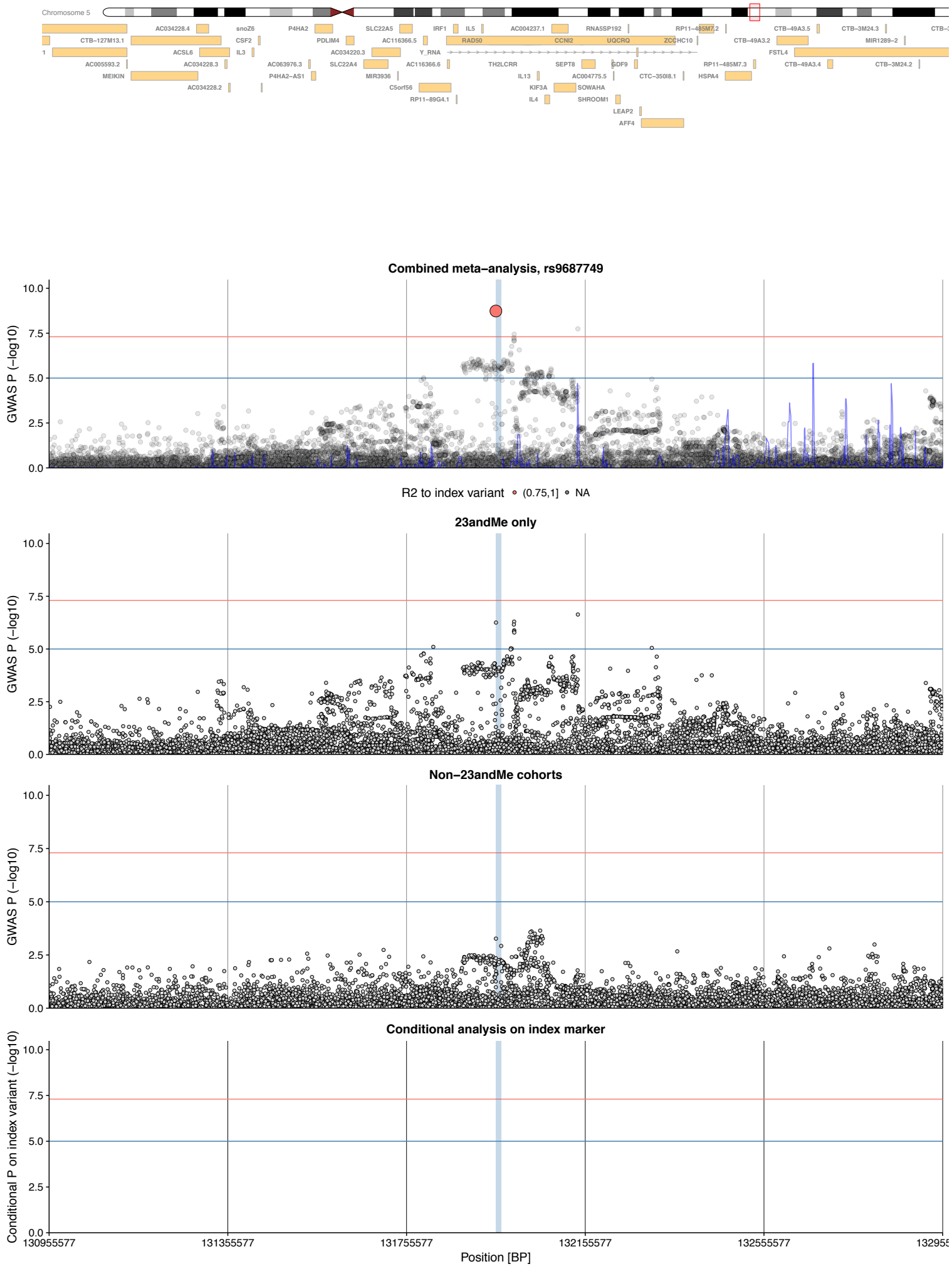

Chromosome 8
STMN2
LINC01607
RP11−26J3.4
RN7SL41P
SNORA20
RP11−48B3.5
RNU2−71P
RNU11−6P
RP11−172E10.1
RP11−26J3.1
RP11−92K15.3
RP11−941H19.3
ZBTB10
ZNF704
PAG1
RNU7−85P
MRPS28
MIR5708
RNU6−1213P
RP11−48B3.3
RN7SL308P
RP11−1149M10.1
HEY1
RP11−26J3.3
TPD52
RP11−48B3.4
RP11−1149M10.2
RP11−92K15.1
RN7SL107P
Combined meta−analysis, rs7824993
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
80262896
80662896
81062896
81462896
81862896
82262
Position [BP]

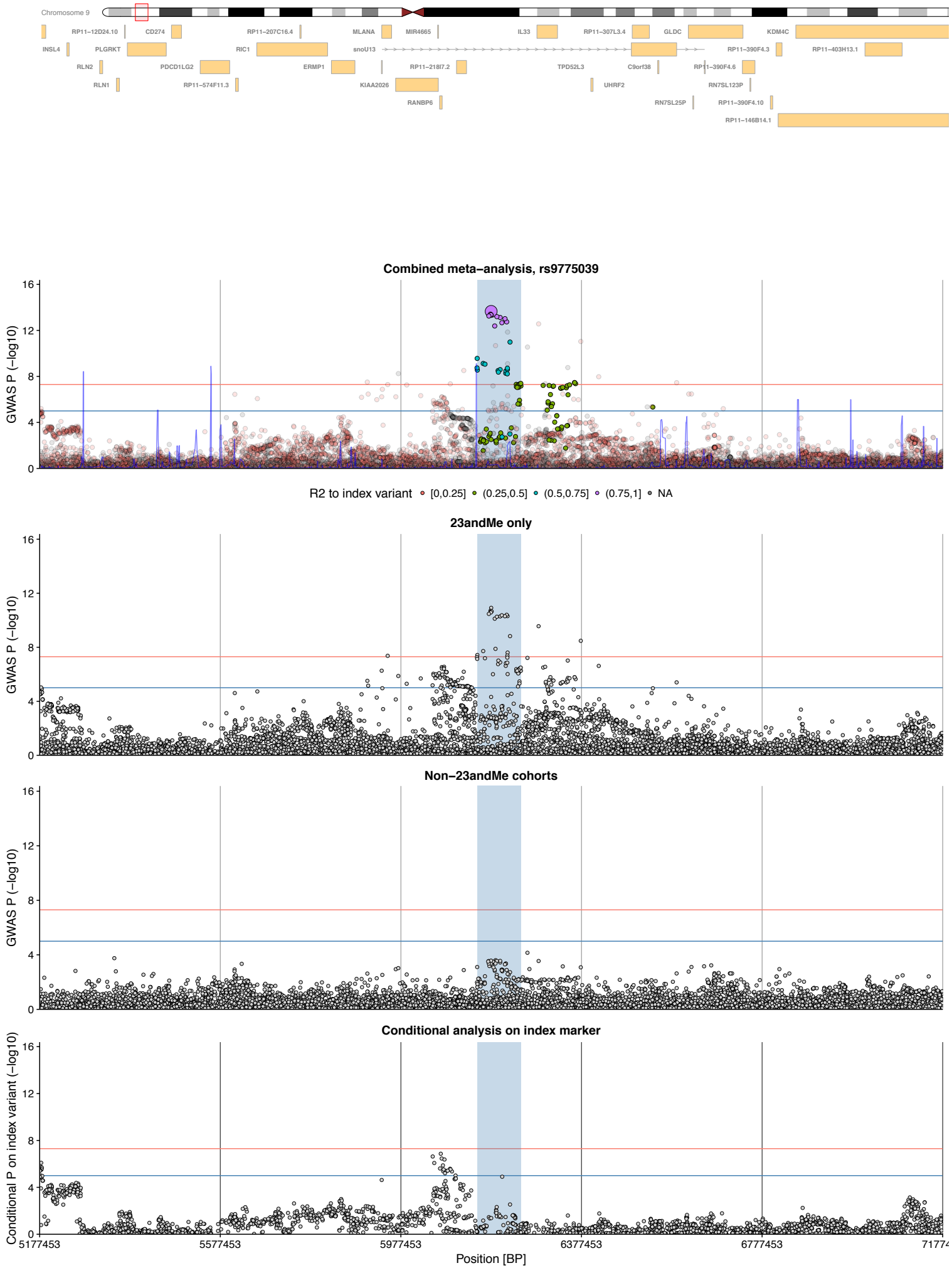

Chromosome 11
CTD−2530H12.1
RNA5SP344
CTD−2011F17.2
RP11−111M22.2
RP11−672A2.6
RP11−21L23.3
CTD−2547H18.1
CAPN5
GDPD4
RP11−739B23.1
CTD−2530H12.2
RP11−619A14.2
RP11−111M22.4
RP11−672A2.5
RP11−21L23.2
B3GNT6
MYO7A
PAK1
RP11−535A19.1
RP11−619A14.3
RP11−111M22.5
RP11−672A2.4
TSKU
OMP
DGAT2
WNT11
PRKRIR
RP11−672A2.3
AP002498.1
RN7SL786P
Y_RNA
AP001189.4
ACER3
MOGAT2
RP11−263C24.1
RP11−111M22.3
LRRC32
RP11−535A19.2
C11orf30
RP11−672A2.1
UVRAG
RP11−21L23.4
Combined meta−analysis, rs7936323
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
75293758
75693758
76093758
76493758
76893758
77293
Position [BP]


Chromosome 16
Combined meta−analysis, rs9282864
GWAS P (−log10)
R2 to index variant [0,0.25] (0.25,0.5] (0.5,0.75] (0.75,1] NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
27617308
28017308
28417308
28817308
29217308
Position [BP]


Chromosome 2
G2
RFX8
AC092570.2
MAP4K4
FLJ20373
LINC01127
IL1R2
IL1R1
IL1RL2
AC007271.3
AC007278.2
AC007278.3
IL18RAP
MIR4772
IL1RL1
IL18R1
SLC9A4
SLC9A2
TMEM182
MFSD9
AC007251.2
AC108051.3
AC108051.1
AC108051.2
AC073987.1
Combined meta−analysis, rs950881
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
101932512
102332512
102732512
103132512
103532512
Position [BP]

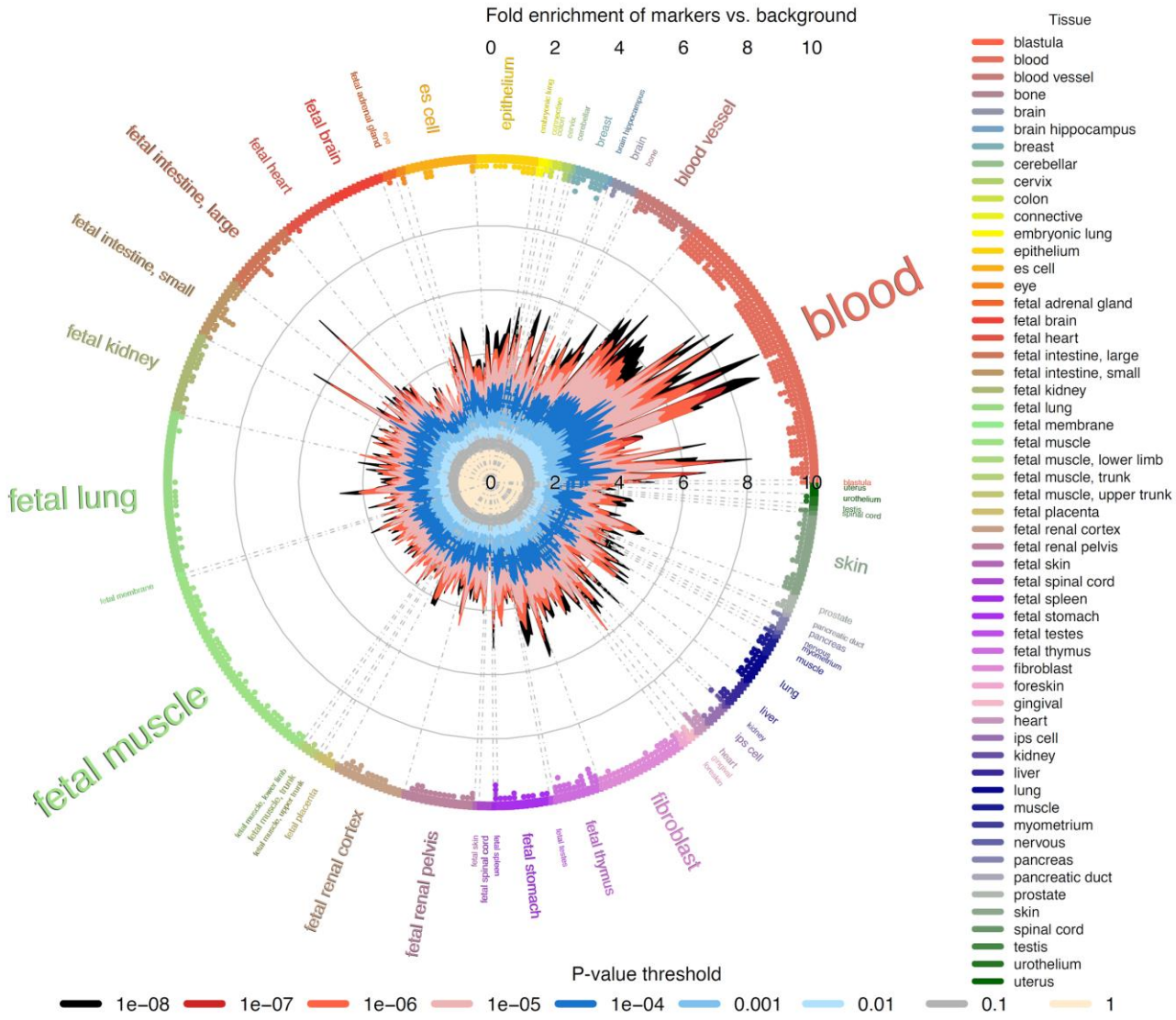

Chromosome 7
EVX1
EVX1−AS
AC004009.1
AC004009.2
AC004009.3
HIBADH
AC004549.6
TAX1BP1
JAZF1
RN7SL365P
RNU6−979P
JAZF1−AS1
CREB5
CTB−113D17.1
AC005013.5
AC005162.5
TRIL
CPVL
HOTTIP
Combined meta−analysis, rs9648346
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
27160113
27560113
27960113
28360113
28760113
29160
Position [BP]


Chromosome 5
AC034228.4
snoZ6
P4HA2
SLC22A5
IRF1
IL5
AC004237.1
RNA5SP192
RP11−485M7.2
CTB−49A3.5
CTB−3M24.3
CTB−127M13.1
CSF2
PDLIM4
AC116366.5
RAD50
CCNI2
UQCRQ
ZCCHC10
CTB−49A3.2
MIR1289−2
1
ACSL6
IL3
AC034220.3
Y_RNA
FSTL4
AC005593.2
AC034228.3
AC063976.3
SLC22A4
AC116366.6
TH2LCRR
SEPT8
GDF9
RP11−485M7.3
CTB−49A3.4
CTB−3M24.2
MEIKIN
P4HA2−AS1
MIR3936
IL13
AC004775.5
CTC−350I8.1
HSPA4
AC034228.2
C5orf56
KIF3A
SOWAHA
RP11−89G4.1
IL4
SHROOM1
LEAP2
AFF4
Combined meta−analysis, rs9687749
GWAS P (−log10)
10.0
7.5
5.0
2.5
0.0
R2 to index variant
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
130955577
131355577
131755577
132155577
132555577
Position [BP]


Chromosome 9
RP11−12D24.10
CD274
RP11−207C16.4
MLANA
MIR4665
IL33
RP11−307L3.4
GLDC
KDM4C
INSL4
PLGRKT
RIC1
snoU13
RP11−390F4.3
RP11−403H13.1
RLN2
PDCD1LG2
ERMP1
RP11−218I7.2
TPD52L3
C9orf38
RP11−390F4.6
RLN1
RP11−574F11.3
KIAA2026
UHRF2
RN7SL123P
RANBP6
RN7SL25P
RP11−390F4.10
RP11−146B14.1
Combined meta−analysis, rs9775039
GWAS P (−log10)
R2 to index variant
[0,0.25]
(0.25,0.5]
(0.5,0.75]
(0.75,1]
NA
23andMe only
Non−23andMe cohorts
Conditional analysis on index marker
Conditional P on index variant (−log10)
5177453
5577453
5977453
6377453
6777453
Position [BP]

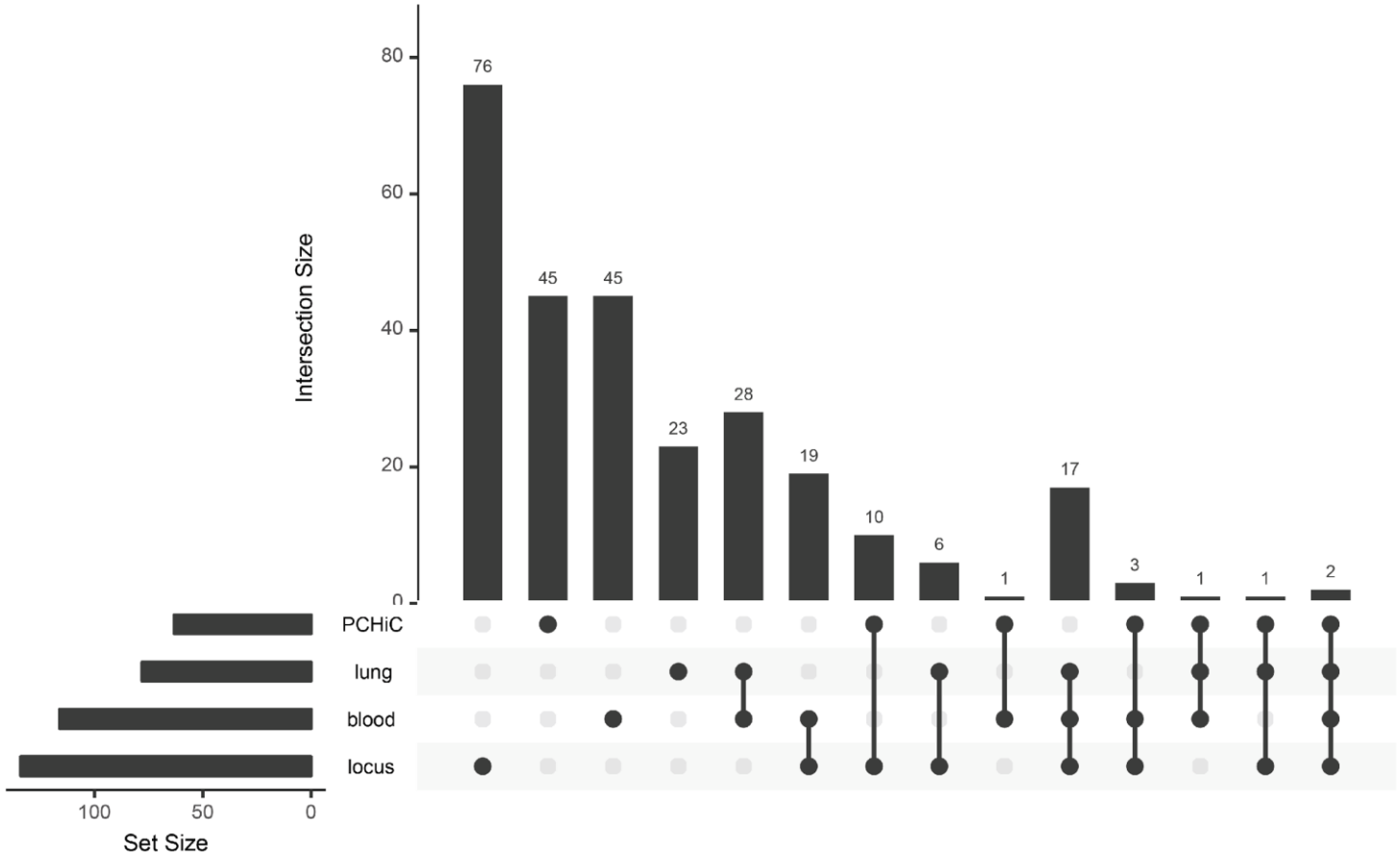
Supplementary Figure 6

Enrichment plot from the GARFIELD regulatory element enrichment pipeline, showing tissue specific enrichment of DNASE hypersensitivity hotspots. Each central layered densitogram color represents enrichment values for increasing marker p-value thresholds (1 through 1e-8). The GARFIELD analysis was run on p-values from genetic marker association of 16,531,985 genetic markers to allergic rhinitis from the discovery phase, including 212,120 individuals.



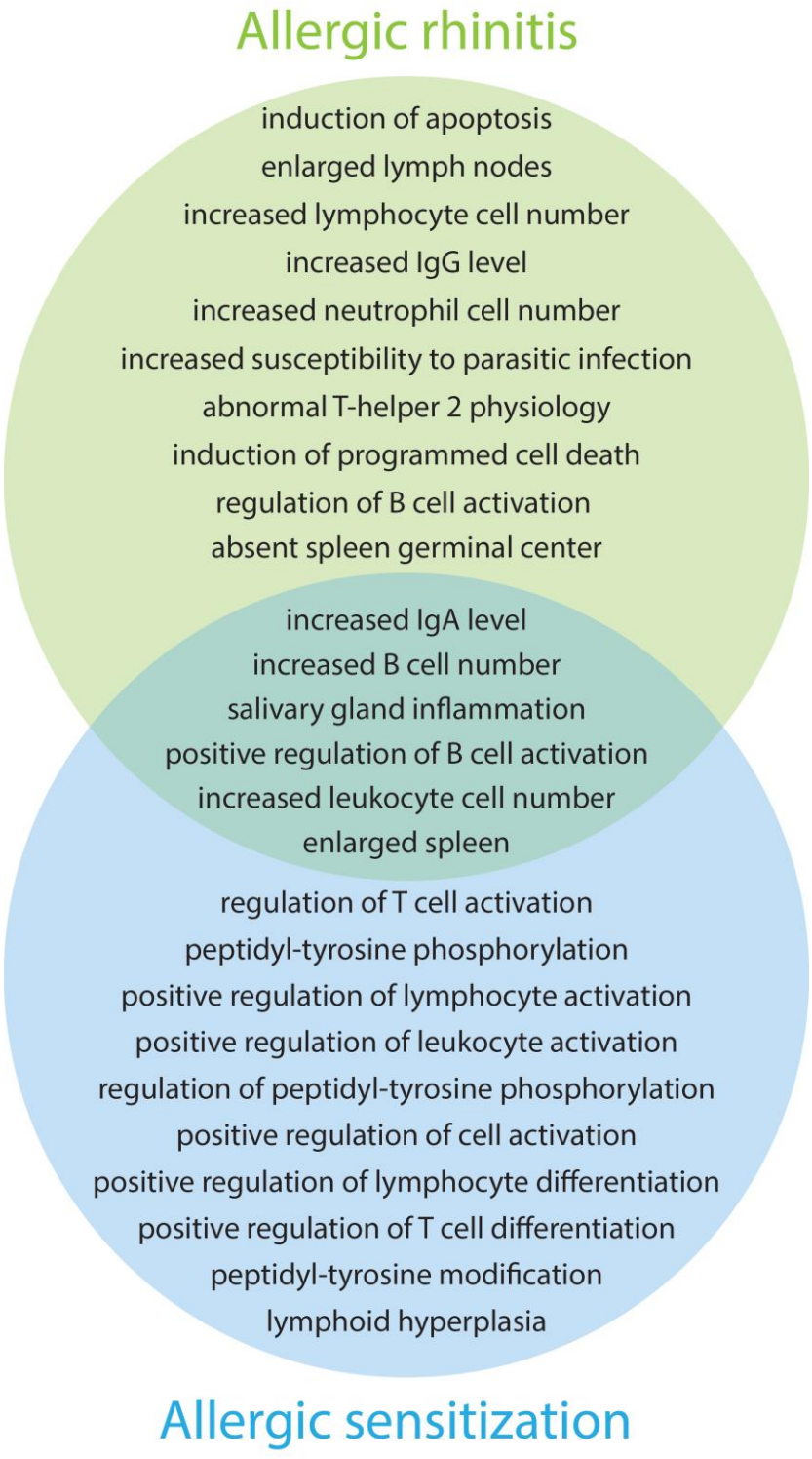
Supplementary Figure 7

"UpSet" set visualisation plot, indicating set intersections between e/meQTL blood, e/meQTL lung, PCHiC, and locus genes.



Supplementary Figure 8

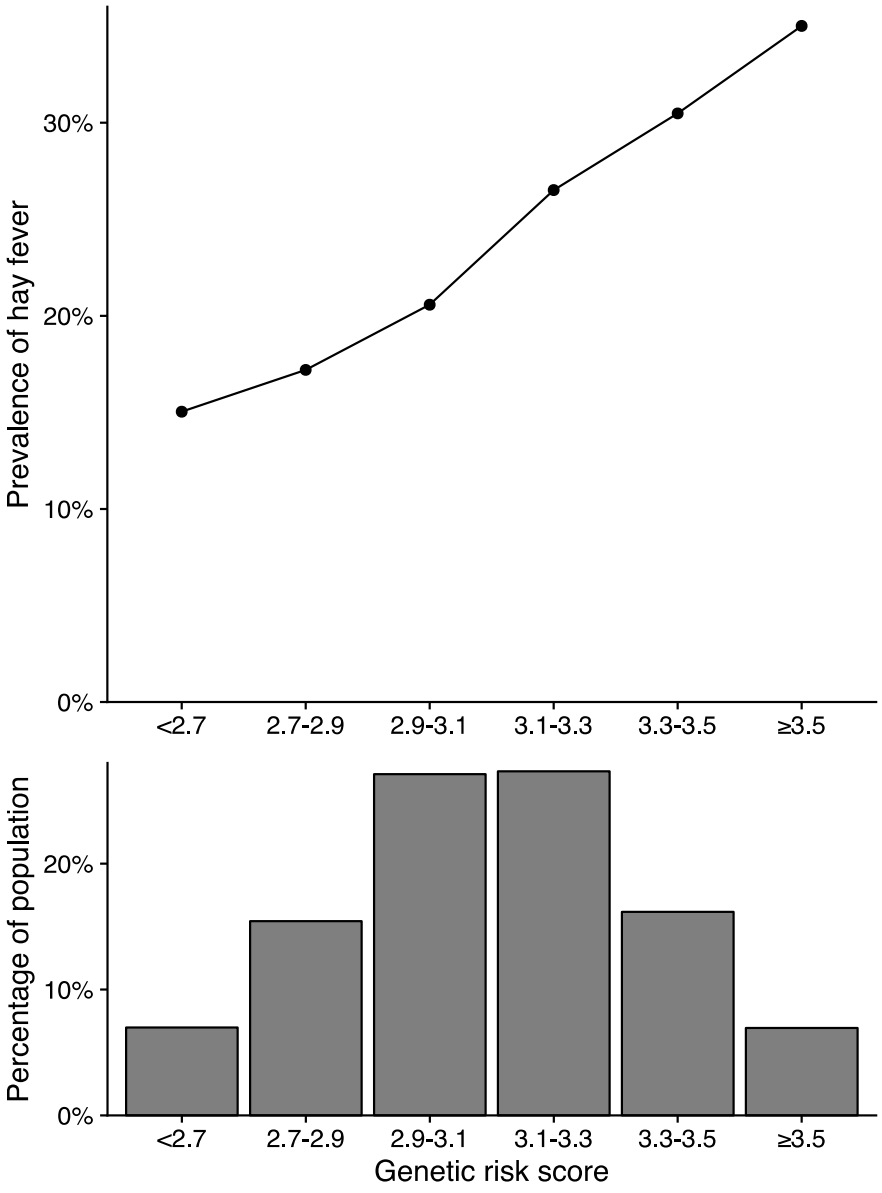
Overlap of enriched gene sets from the allergic rhinitis and allergic sensitization GWAS. All significant overlapping gene sets are shown together with the 10 gene sets with most statistically significant enrichment from each phenotype. Results for all gene sets are shown in Supplementary Tables 21 and 22. Enriched gene sets were calculated from an inverse variance weighted fixed-effect meta-analysis of genetic marker association of 16,531,985 genetic markers to allergic rhinitis from the discovery phase, including 212,120 individuals, and from inverse variance weighted fixed-effect meta-analysis of genetic marker association of 16,531,985 genetic markers to allergic sensitisation, including 24,481 individuals.



Supplementary Figure 9

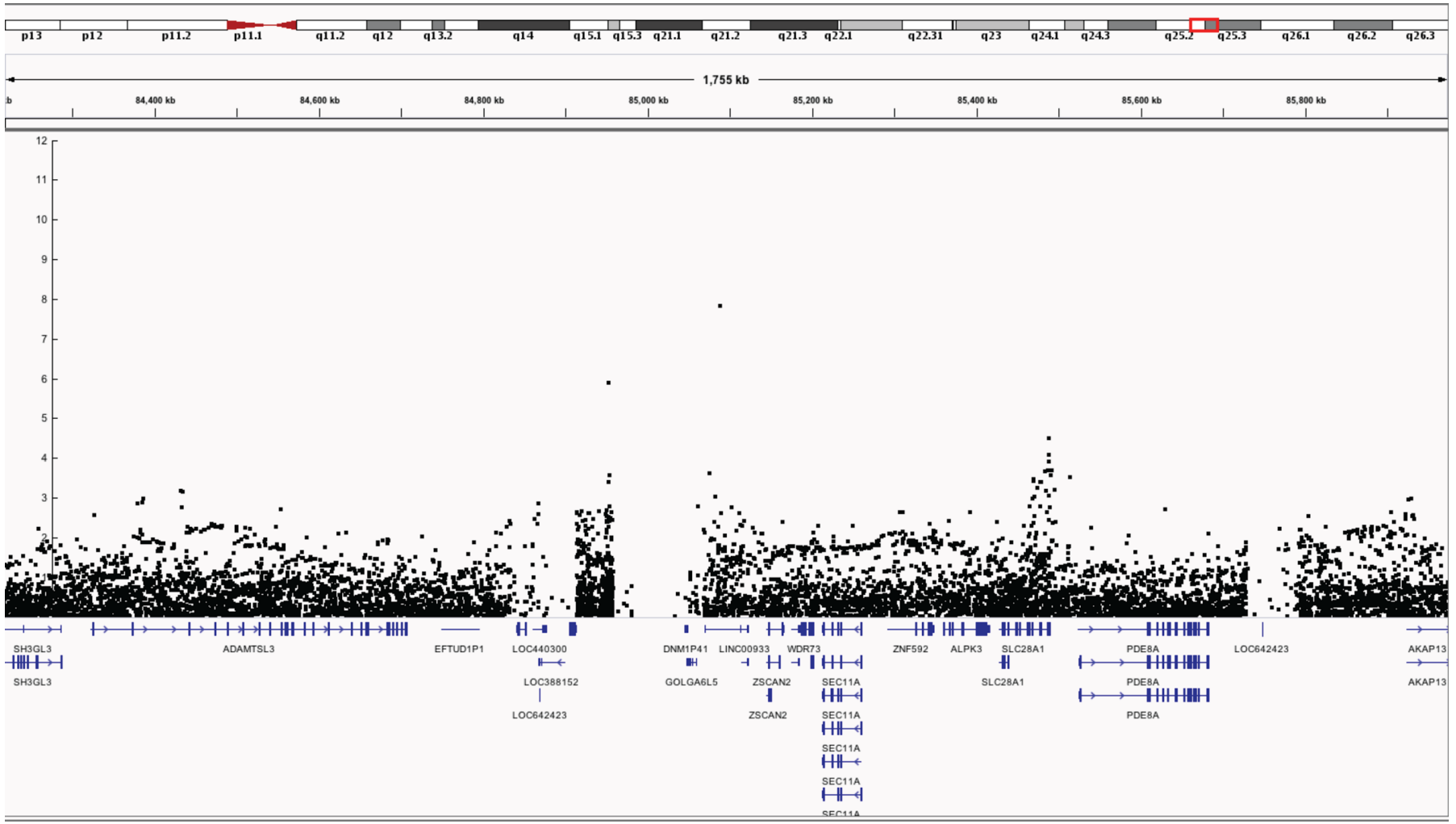
Relationship between allergic rhinitis genetic risk score and disease prevalence.

Combined impact of risk alleles from the 41 genome-wide significant and replicating loci on prevalence of allergic rhinitis (hay fever) in the population-based B58C study. For each individual, we calculated a genetic risk score by applying per-allele risk estimates from the replication sets to the number of higher-risk alleles at the 41 genome-wide significant and replicating loci (one SNP per locus). The risk score thus represents an index of the number of weighted risk alleles. Along the x axis, individuals in each risk score interval are shown (with the lower bound included and the upper bound excluded in each interval). The prevalence of hay fever in each interval is plotted in the upper panel, and the histogram in the bottom panel shows the percentage of the B58C population in each risk score interval.



Supplementary Figure 10

Regional locus plot of association values from the allergic rhinitis GWAS for marker rs193243426 (N=203,356). The locus has an index marker with p < 1e5-8 (p-value from genetic marker association from the discovery phase), but lacks a credible locus LD structure, and was excluded from the study.



Corresponding Author: Klaus Bønnelykke

Date: Apr 28, 2018

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work we publish. This form is published with all life science papers and is intended to promote consistency and transparency in reporting. All life sciences submissions use this form; while some list items might not apply to an individual manuscript, all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   We aimed to include all available data-sets on allergic rhinitis worldwide. In this study we achieved a sample size that was several times larger than previous GWAS- studies, which was sufficent for providing genome-wide significant findings.

2. Data exclusions

   Describe any data exclusions.

   Other than the preliminary analyses/experimental optimization, no data were excluded.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   No experimental studies were performed.

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   No experimental studies were performed. Cases and controls for genome-wide association were analyzed based upon predefined criteria.

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   No blinding was performed and not considered relevant in this meta-Genome-wide association study.

   Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

   For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly.
☒	☐	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. *p* values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A summary of the descriptive statistics, including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	This is described in details in the methods section. In short, for AR, AS, and NAR, meta-analysis for the discovery phase was conducted using GWAMA, while meta-analysis of replication candidates from the AR discovery phase was carried out using R version 3.4.0, and the meta package version 4.8-2 with an inverse variance weighted fixed-effect model. Additional software include EasyQC (v9.2), STRING (v10), ChEMBL Database (v22), GGraph (v1.0.0), iGraph (v1.0.1), TidyVerse (v1.1.1), LDHUB platform (v1.3.1), GENCODE (v19 and v24, Genome-wide Complex Trait Analysis (GCTA) (v. 1.26.0), Plink (v1.90b3.42), PyMOL (v1.8.2.1), LDSC (v.1.0), LDHUB platform (v1.3.1), GTEx (V6p), UpSetR package (v1.3.2), GWAS Analysis of Regulatory or Functional Information Enrichment with LD correction (GARFIELD) method (v1), Python (v2.7.12), HIBAG (v1.2.3).

For all studies, we encourage code deposition in a community repository (e.g. GitHub). Authors must make computer code available to editors and reviewers upon request. The *Nature Methods* guidance for providing algorithms and software for publication may be useful for any submission.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	No unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	No antibodies were used.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	No eukaryotic cell lines were used.
b. Describe the method of cell line authentication used.	No eukaryotic cell lines were used.
c. Report whether the cell lines were tested for mycoplasma contamination.	No eukaryotic cell lines were used.
d. If any of the cell lines used in the paper are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No eukaryotic cell lines were used.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	No animals were used.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	Relevant information on research participants in the individual studies is provided in the supplementary material.