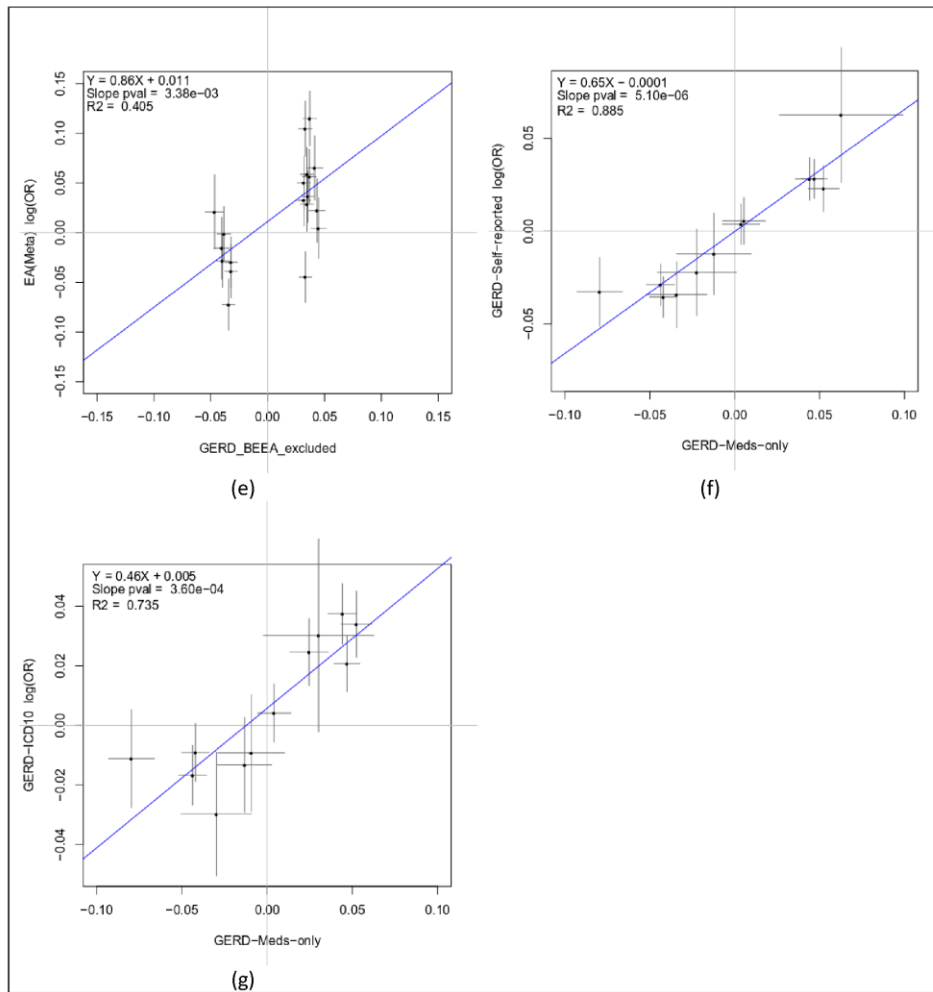
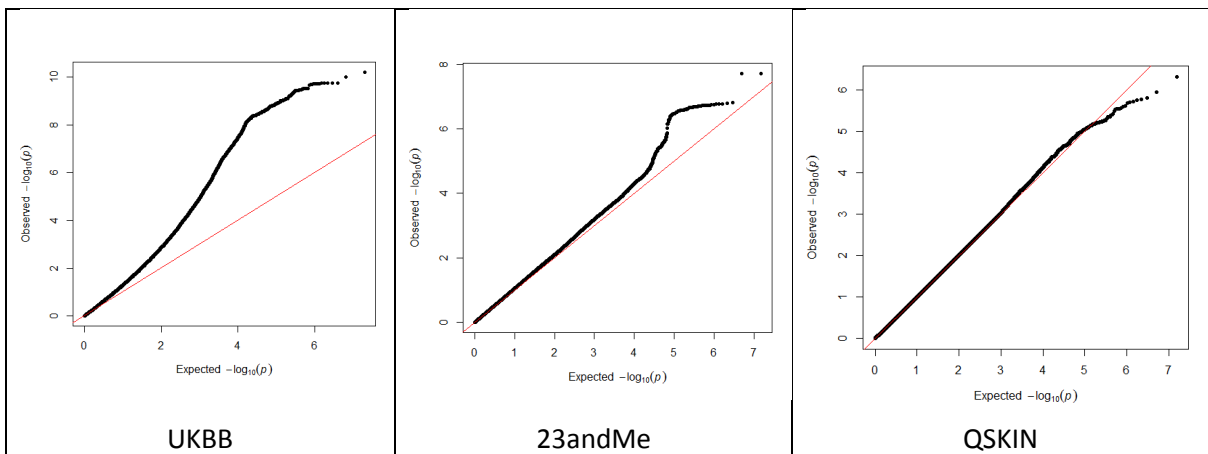
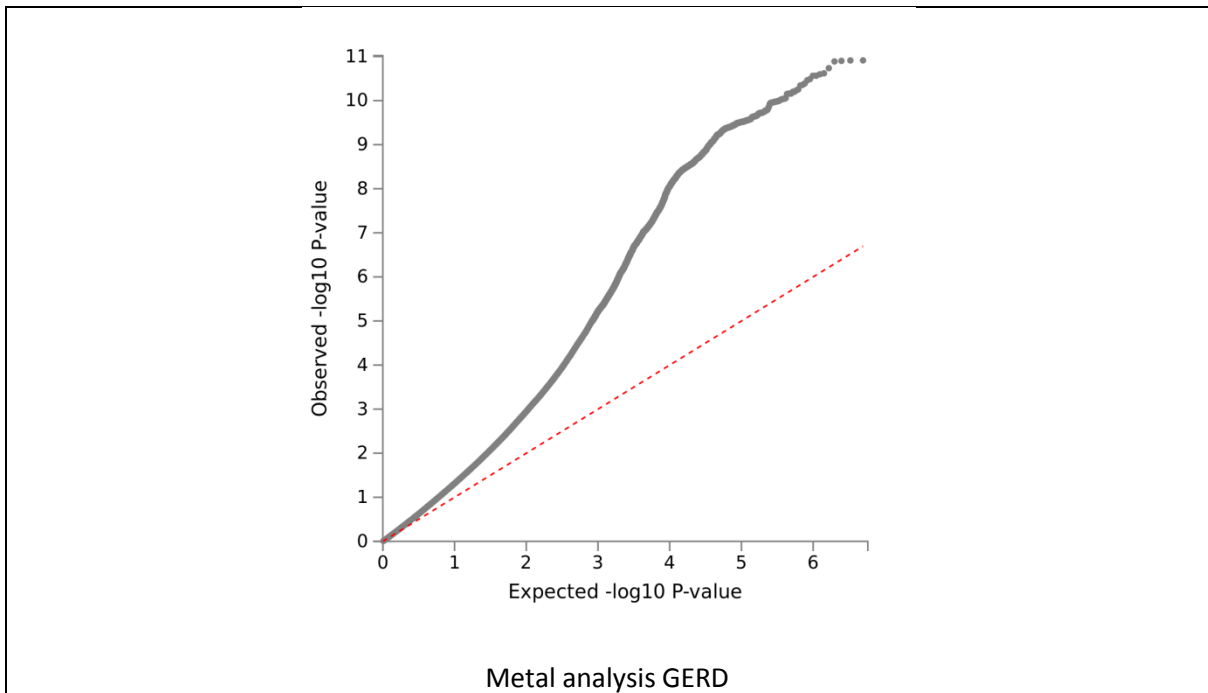


An *et al.* Gastroesophageal reflux GWAS identifies first risk loci that also associate with subsequent severe esophageal diseases

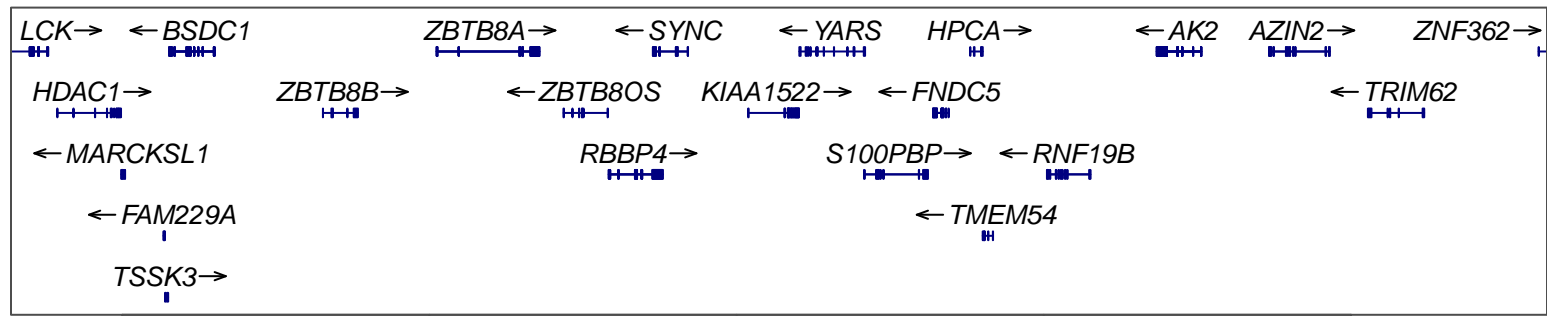
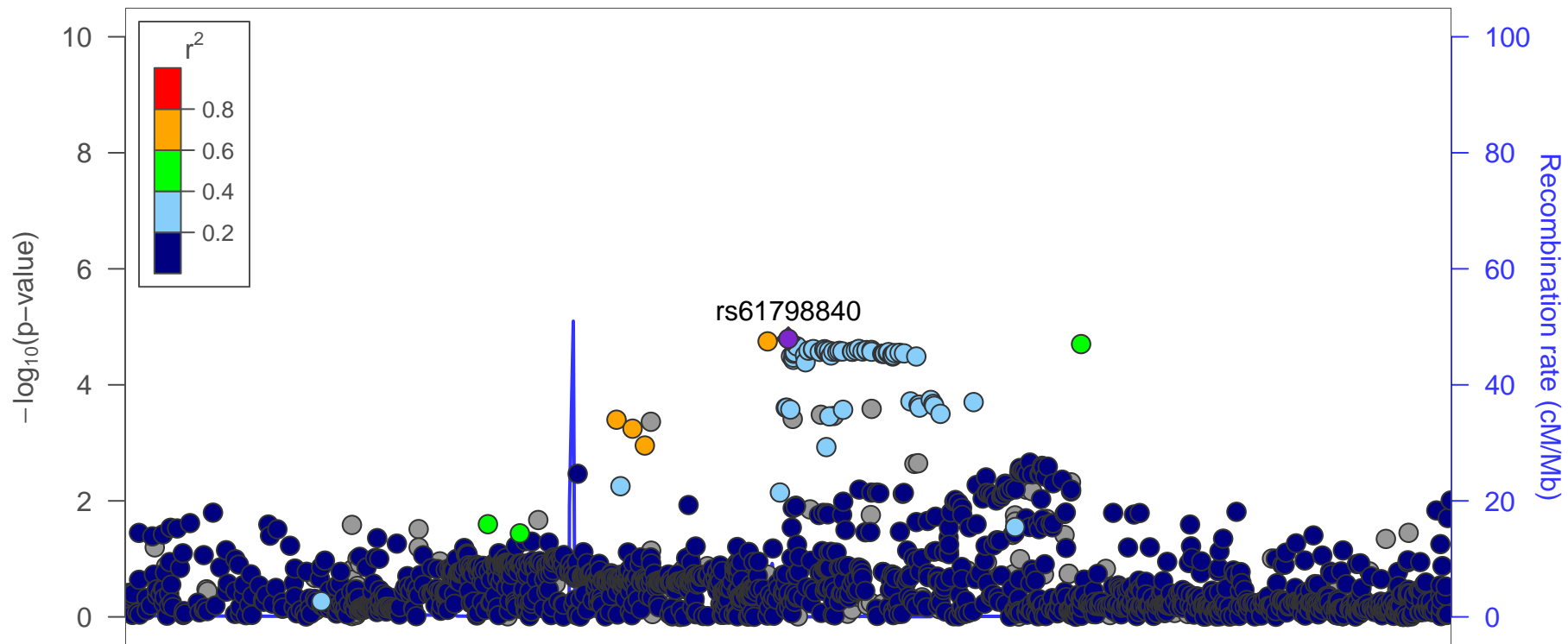


Supplementary Figure 1. (a) The correlation of UKBB GERD vs 23andMe. (b) The correlation of UKBB GERD vs QSKIN GERD. (c) The correlation of Meta GERD (BE/EA samples excluded) vs Meta BEEA. (d) The correlation of Meta GERD (BE/EA samples excluded) vs Meta BE. (e) The correlation of Meta GERD (BE/EA samples excluded) vs EA. Log of significant independent SNPs from GCTA COJO are represented on the x-axis. The corresponding logOR of SNPs from 23andMe and QSKIN are shown on the y-axis. The lengths of the lines for each SNP represent the standard error for the respective logOR. Linear regression was undertaken using the weight of the MAF variance ($MAF * (1-MAF)$). (f) The correlation of GERD from medicine pheno type only (self-report GERD and ICD10 medical record samples excluded) vs GERD from self-reported pheno type only. (g) The correlation of GERD from medicine pheno type only (self-report GERD and ICD10 medical record samples excluded) vs GERD from ICD10 medical record pheno type only.



Supplementary Figure 2. QQ plot of GERD GWAS using summary statistics.

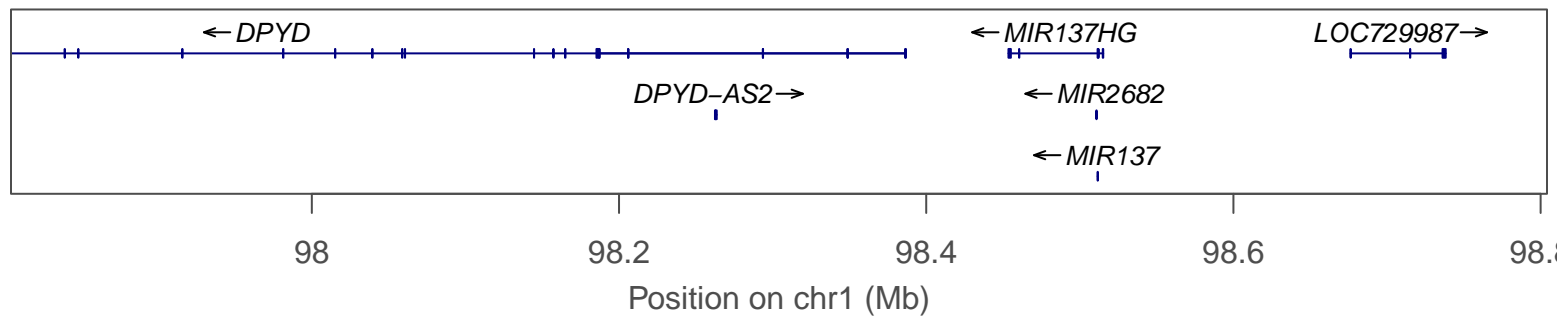
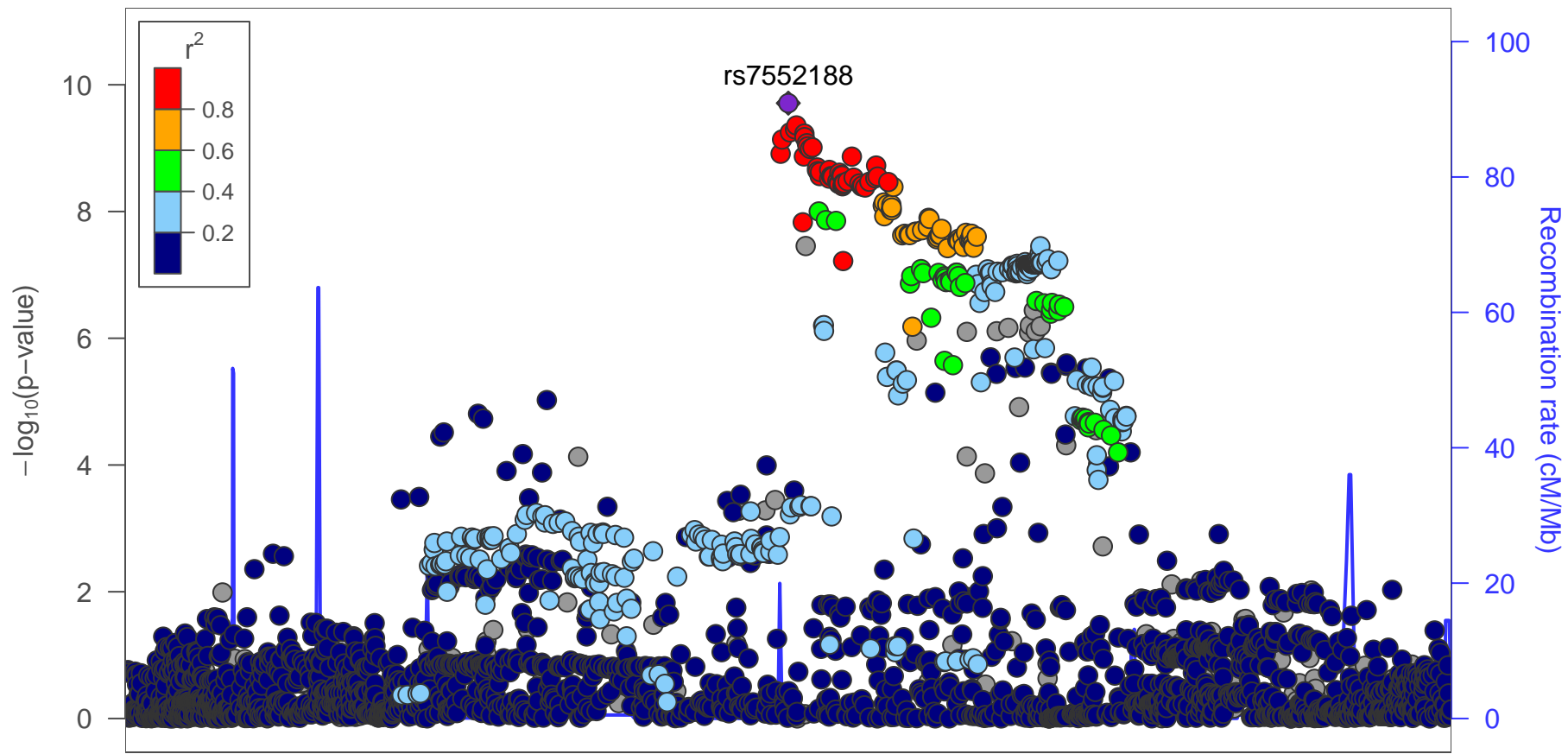
Plotted SNPs



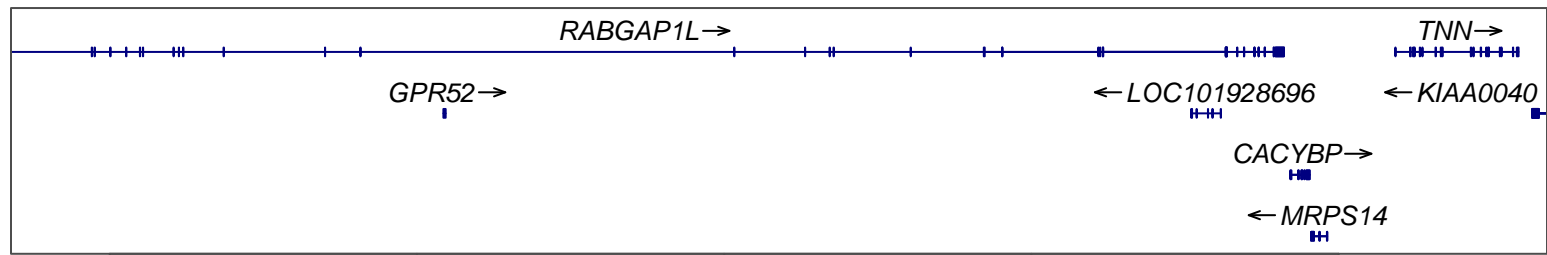
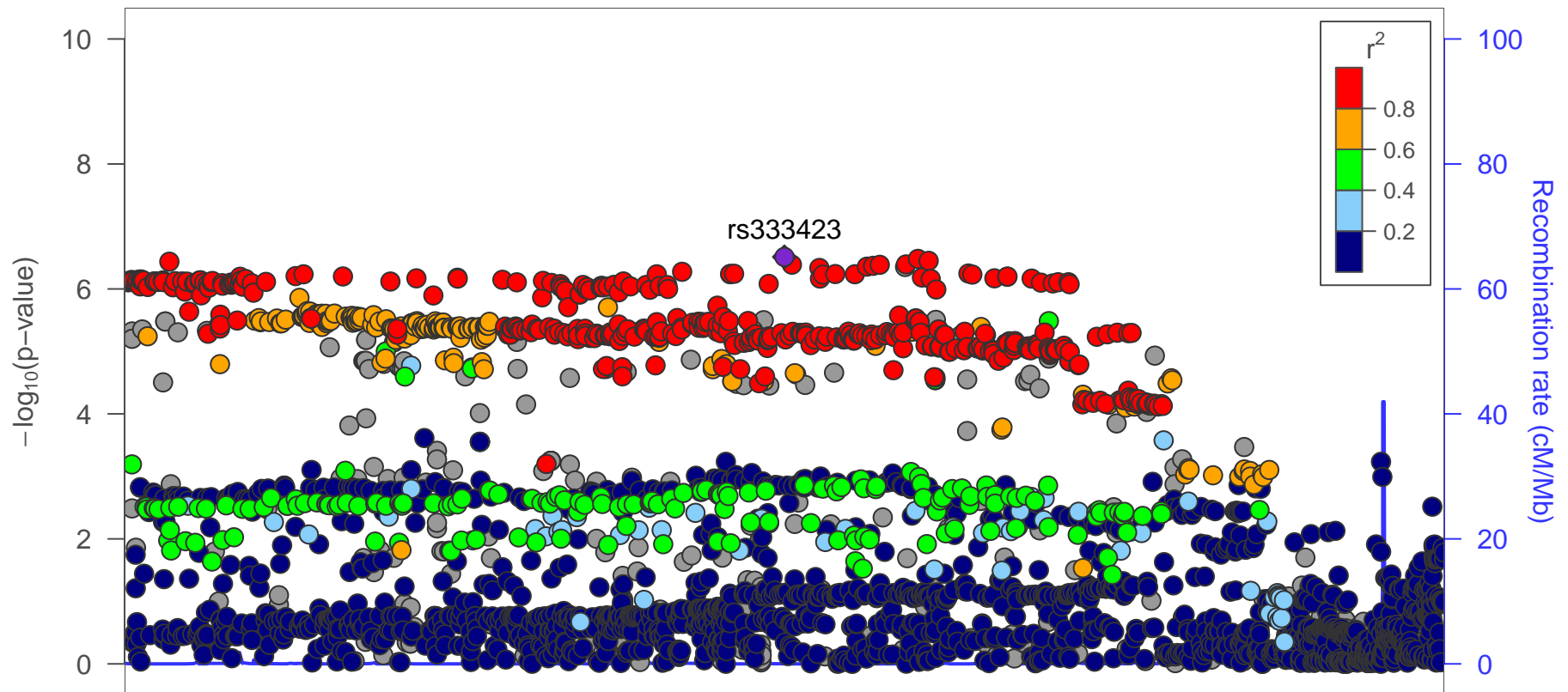
32.8 33 33.2 33.4 33.6

Position on chr1 (Mb)

Plotted SNPs



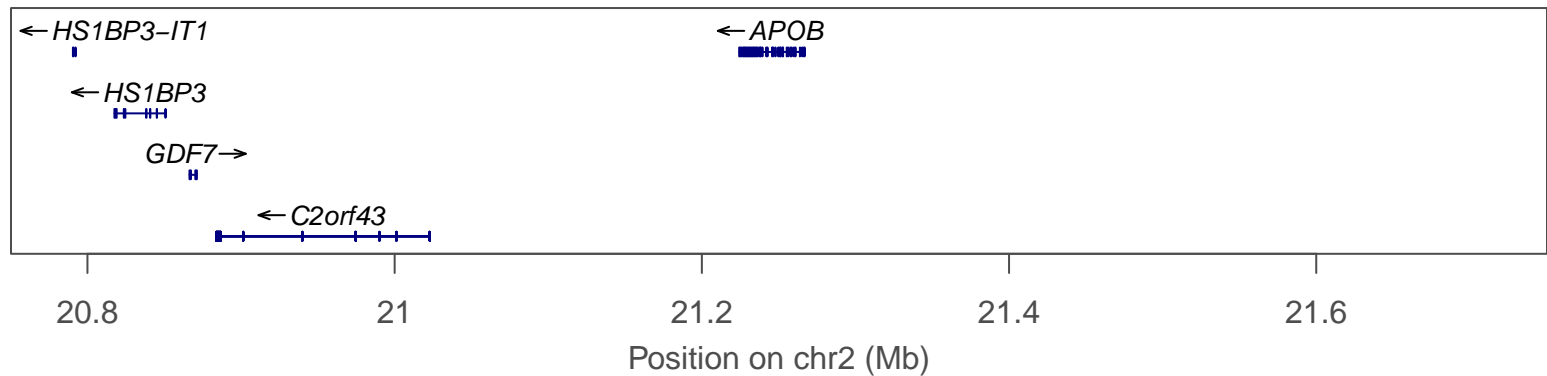
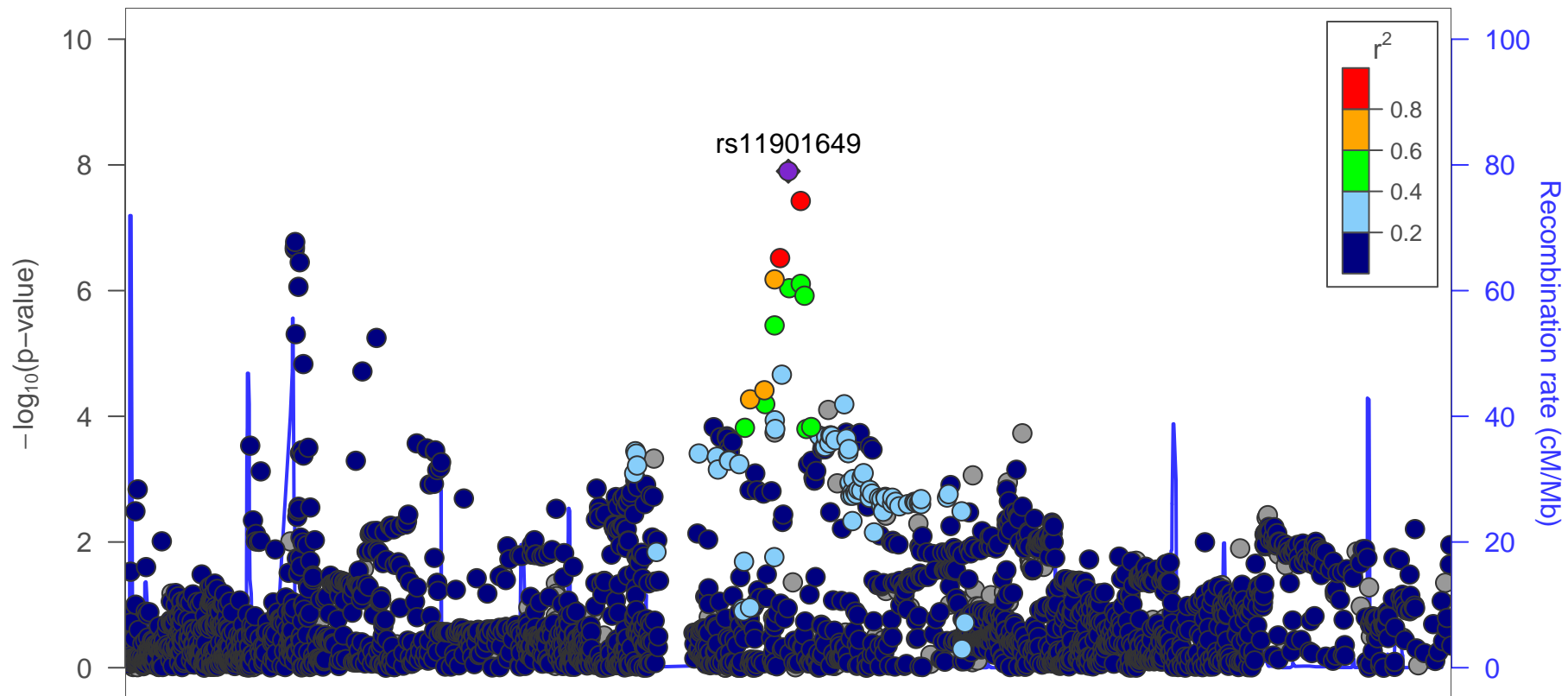
Plotted SNPs



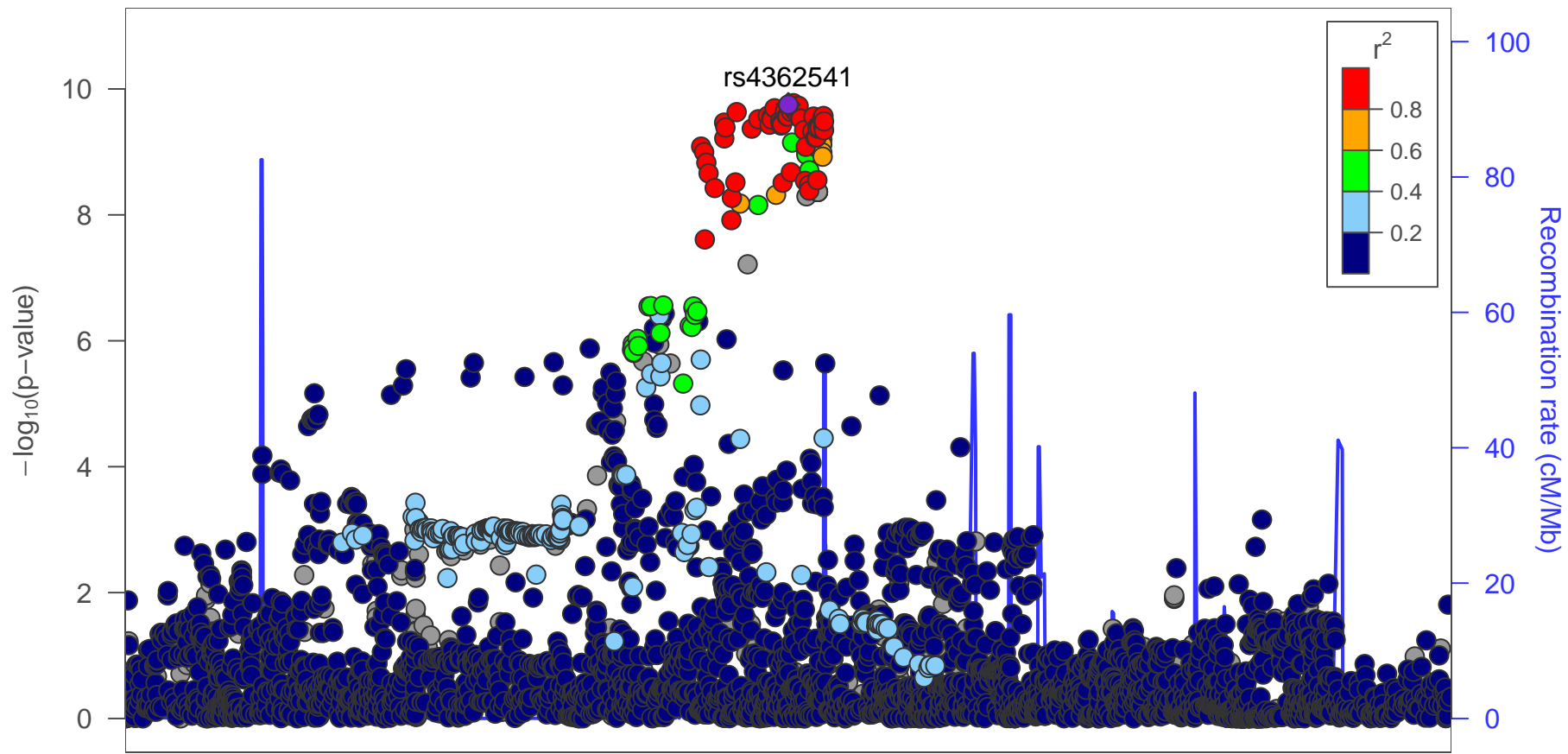
174.2 174.4 174.6 174.8 175

Position on chr1 (Mb)

Plotted SNPs



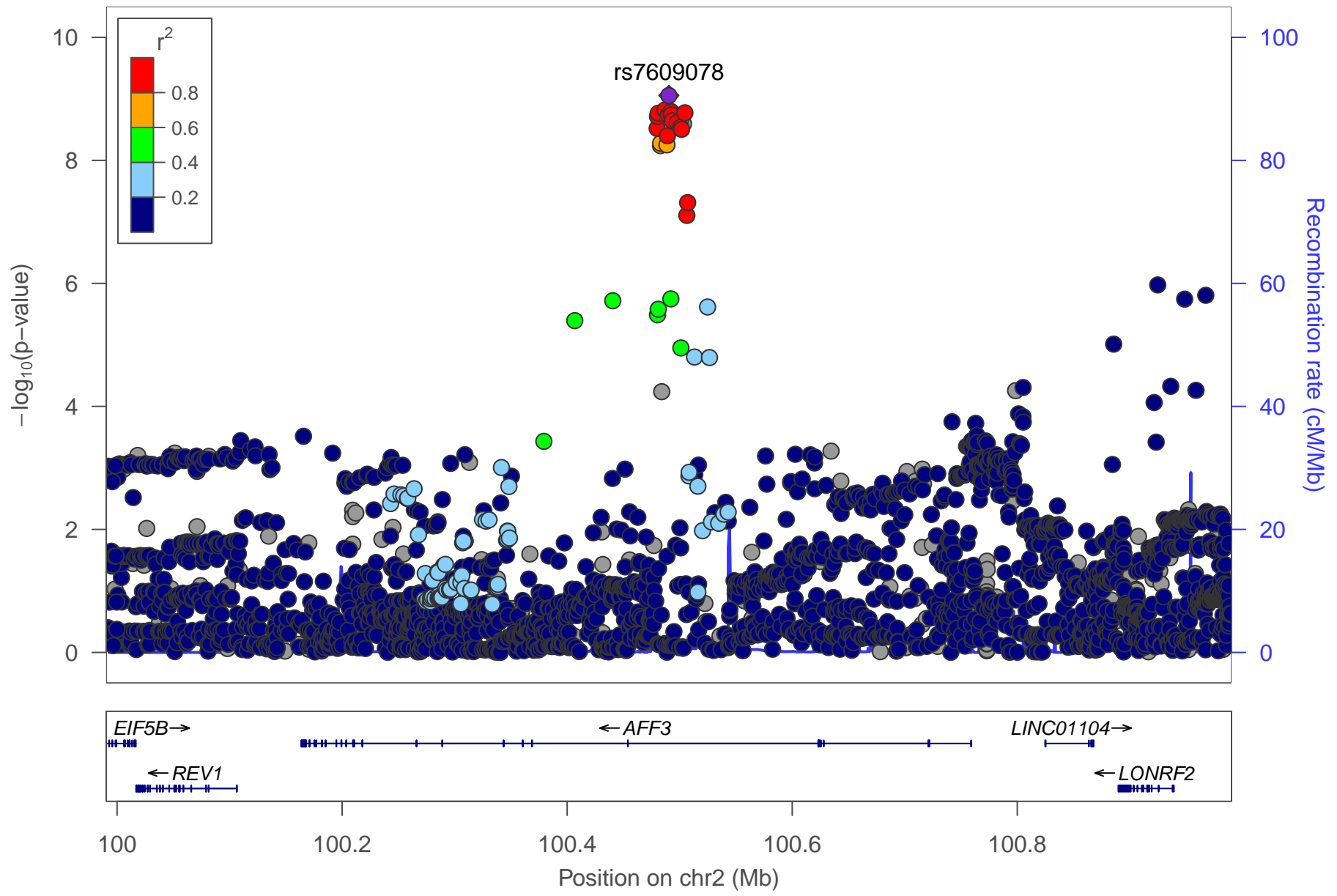
Plotted SNPs



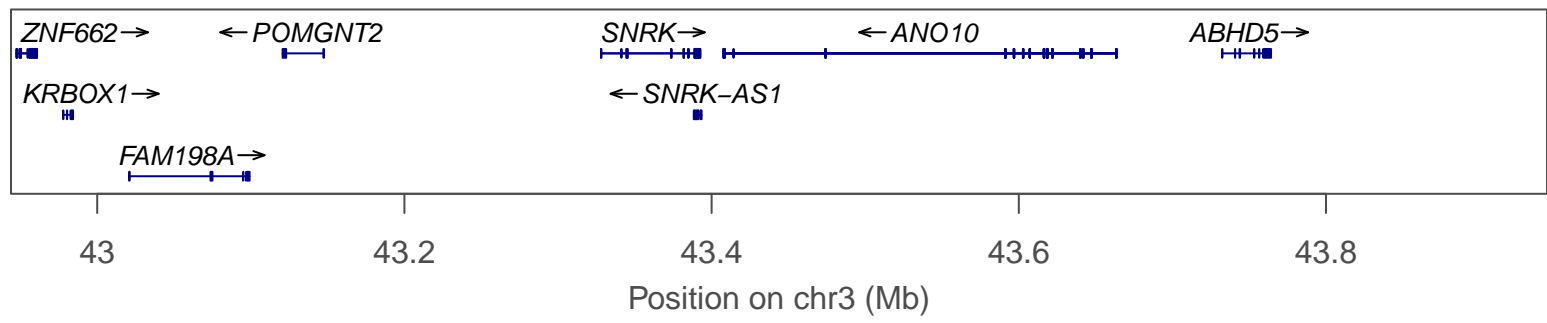
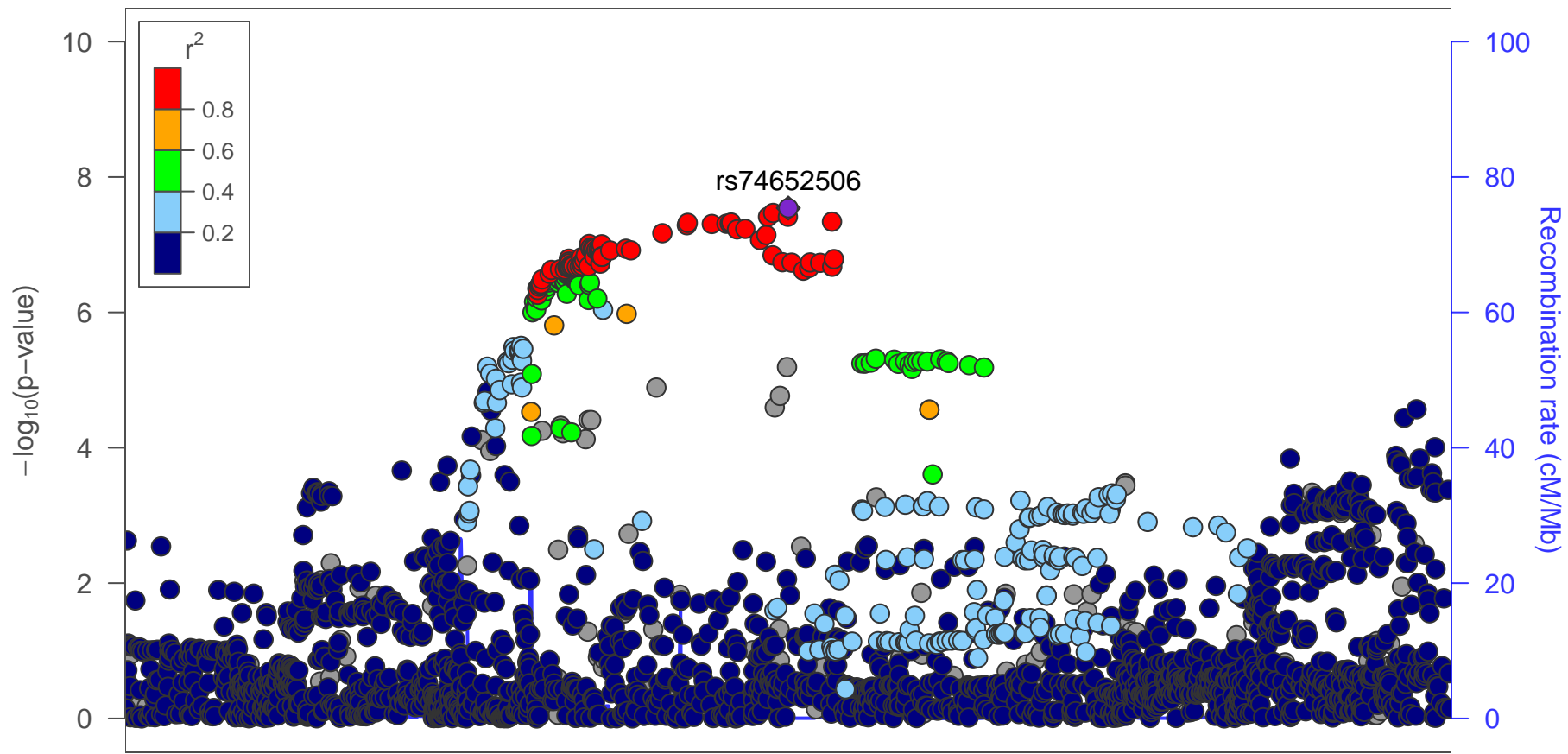
67.4 67.6 67.8 68 68.2

Position on chr2 (Mb)

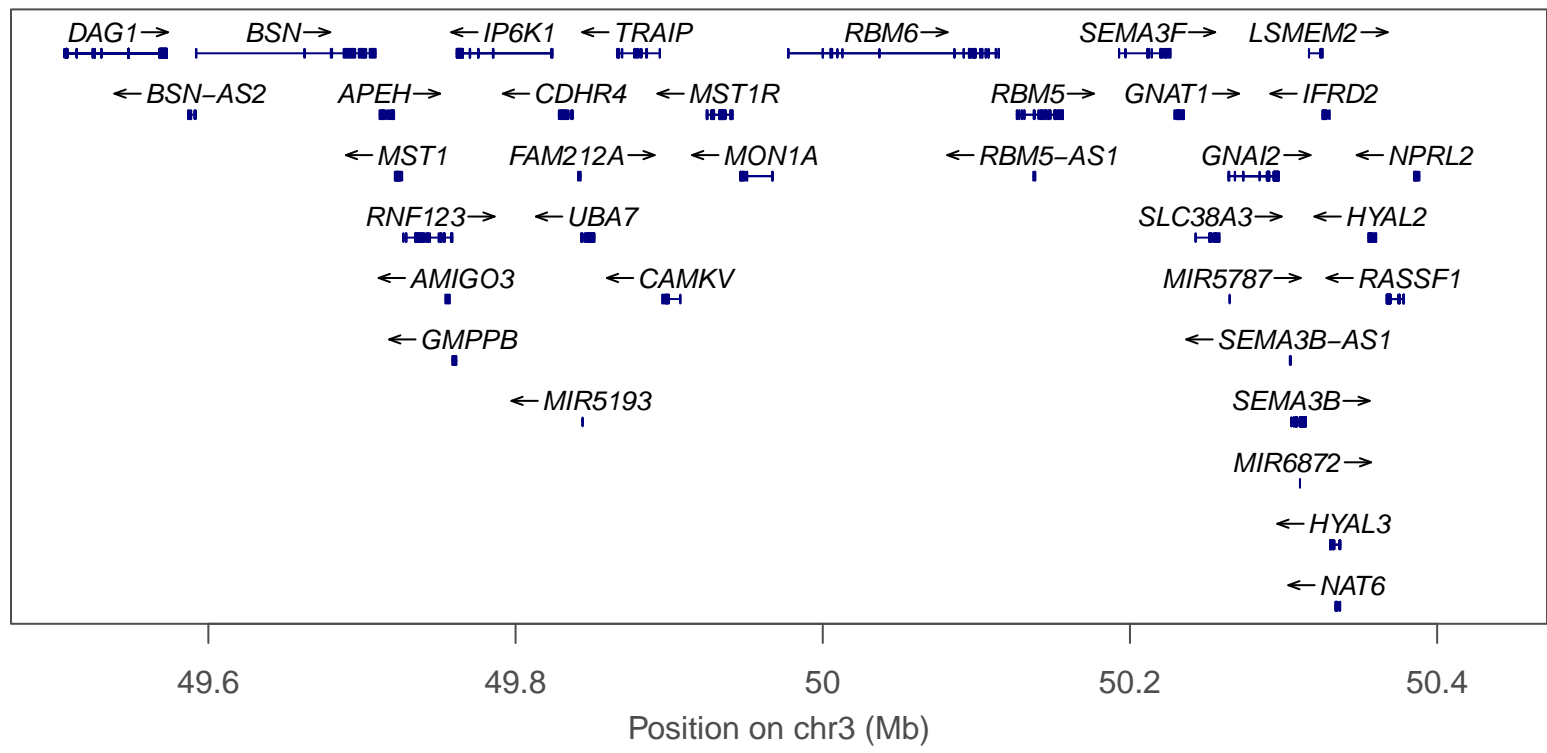
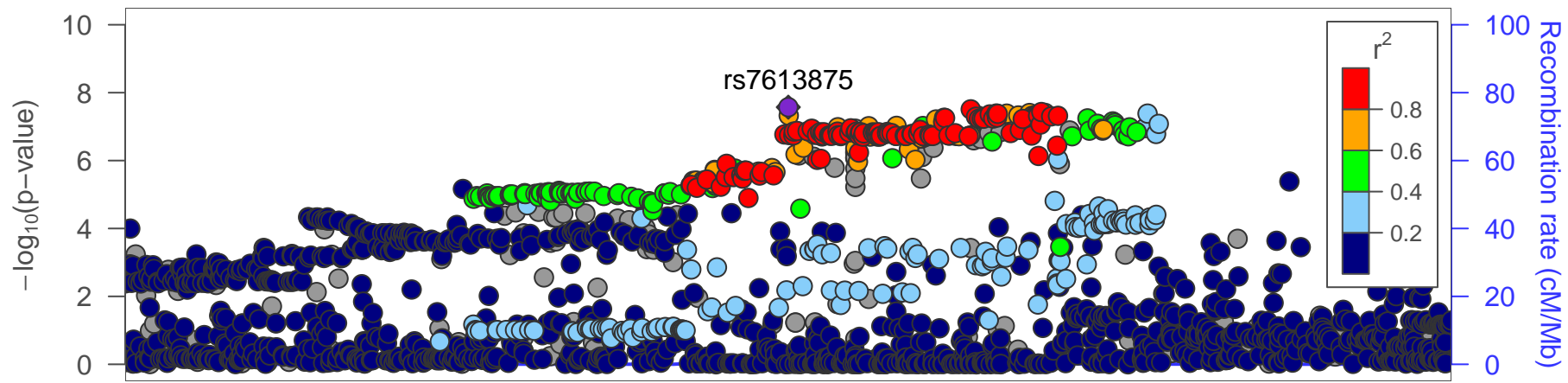
Plotted SNPs



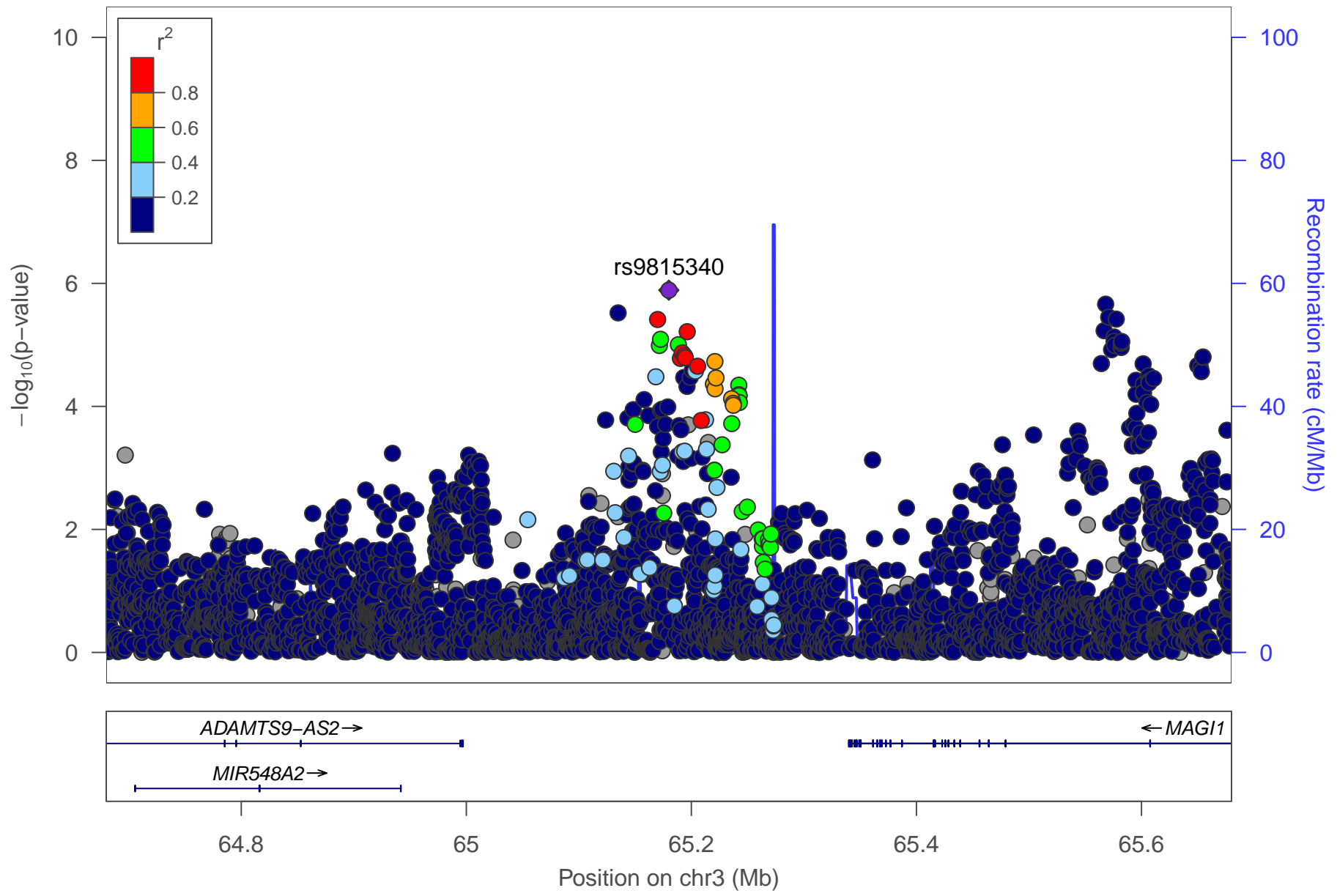
Plotted SNPs



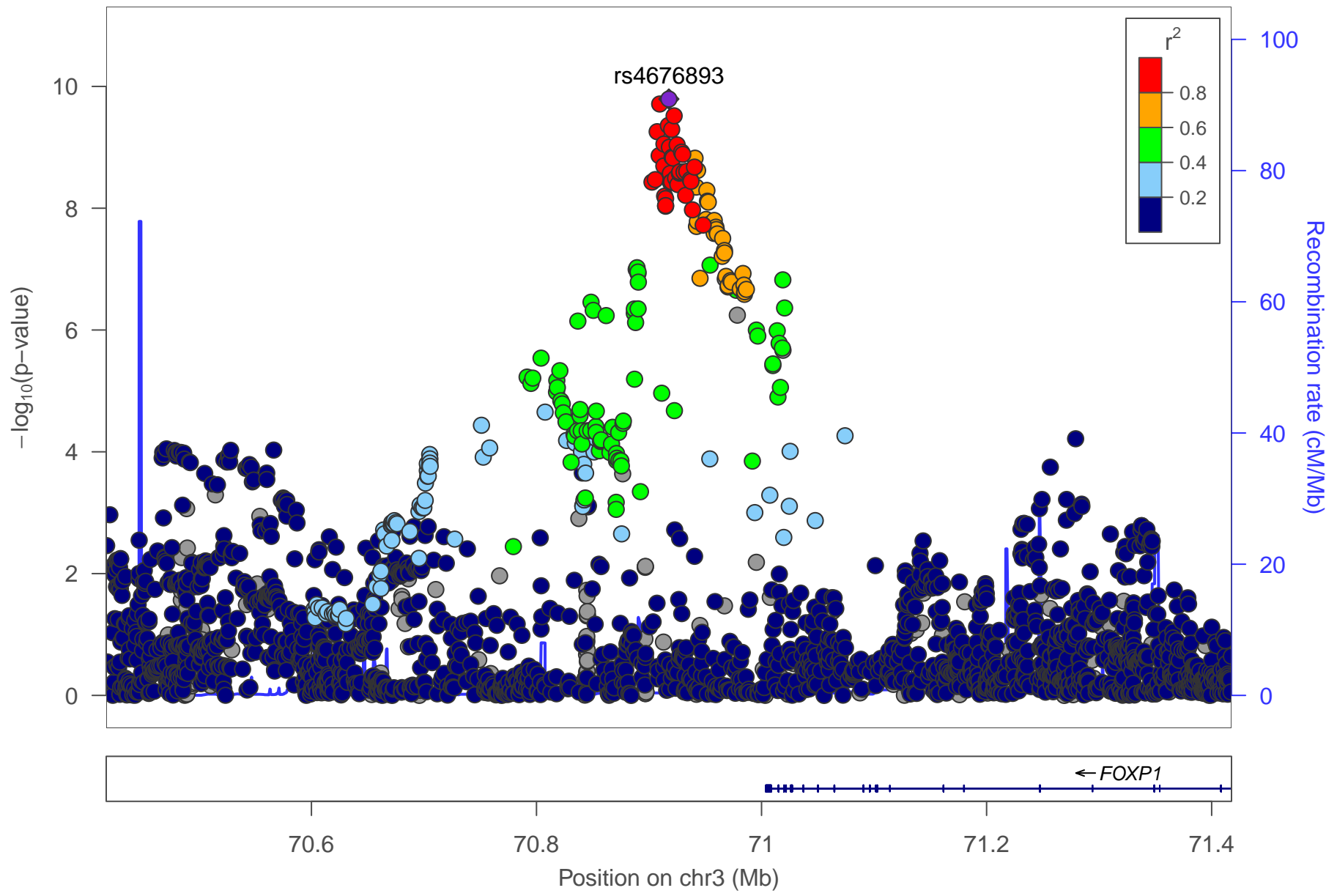
Plotted SNPs



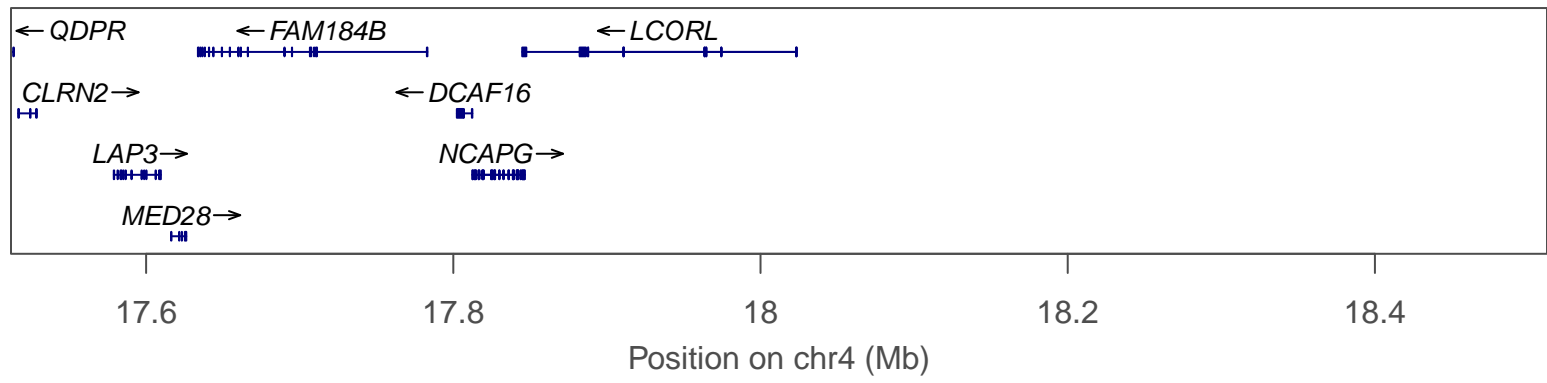
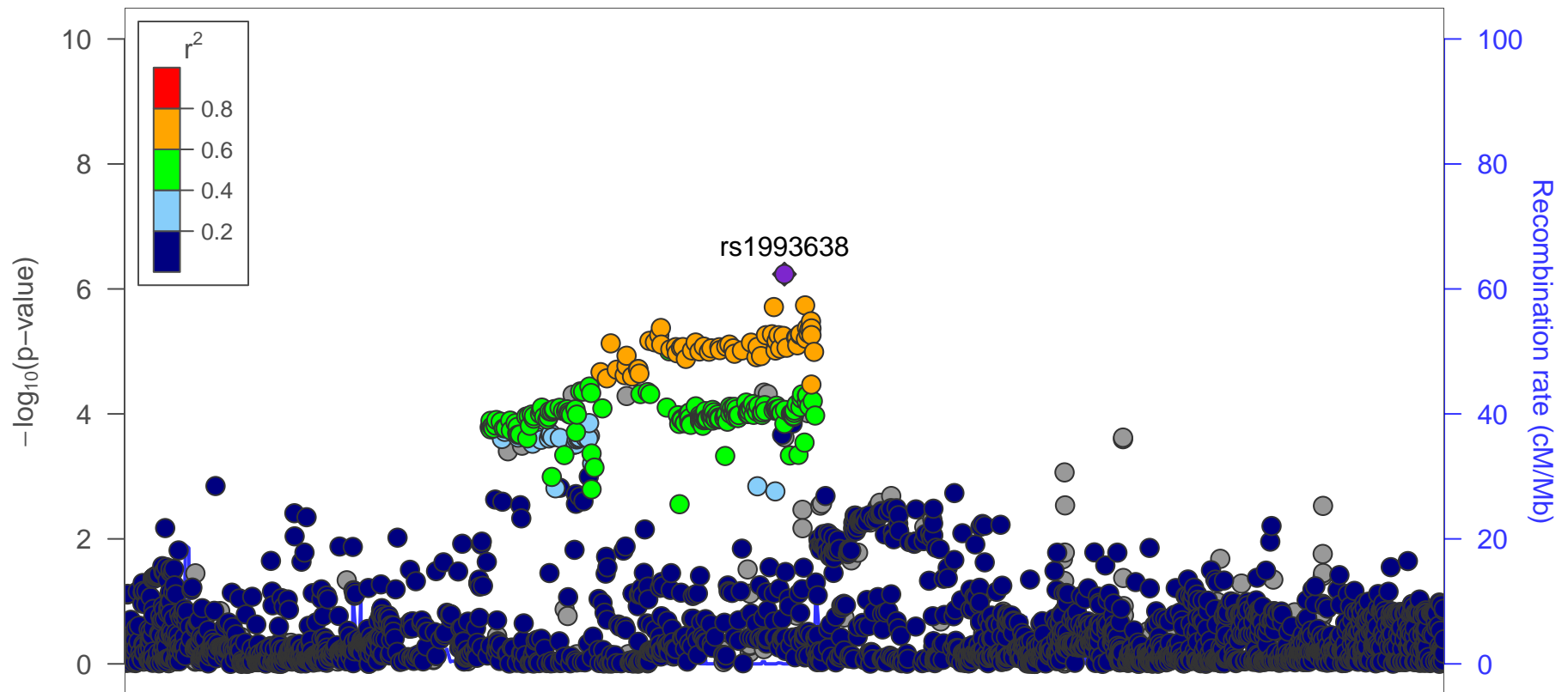
Plotted SNPs



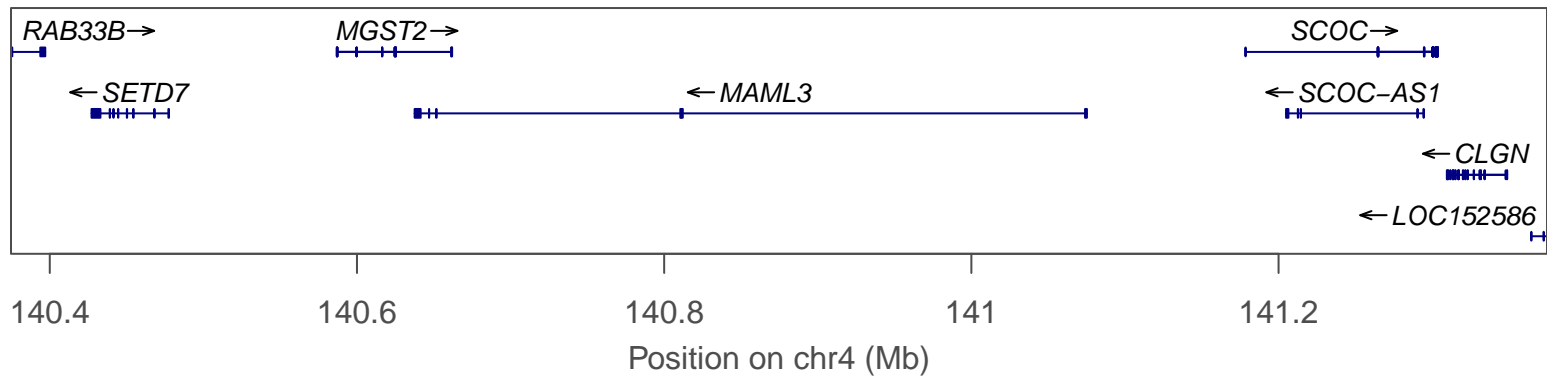
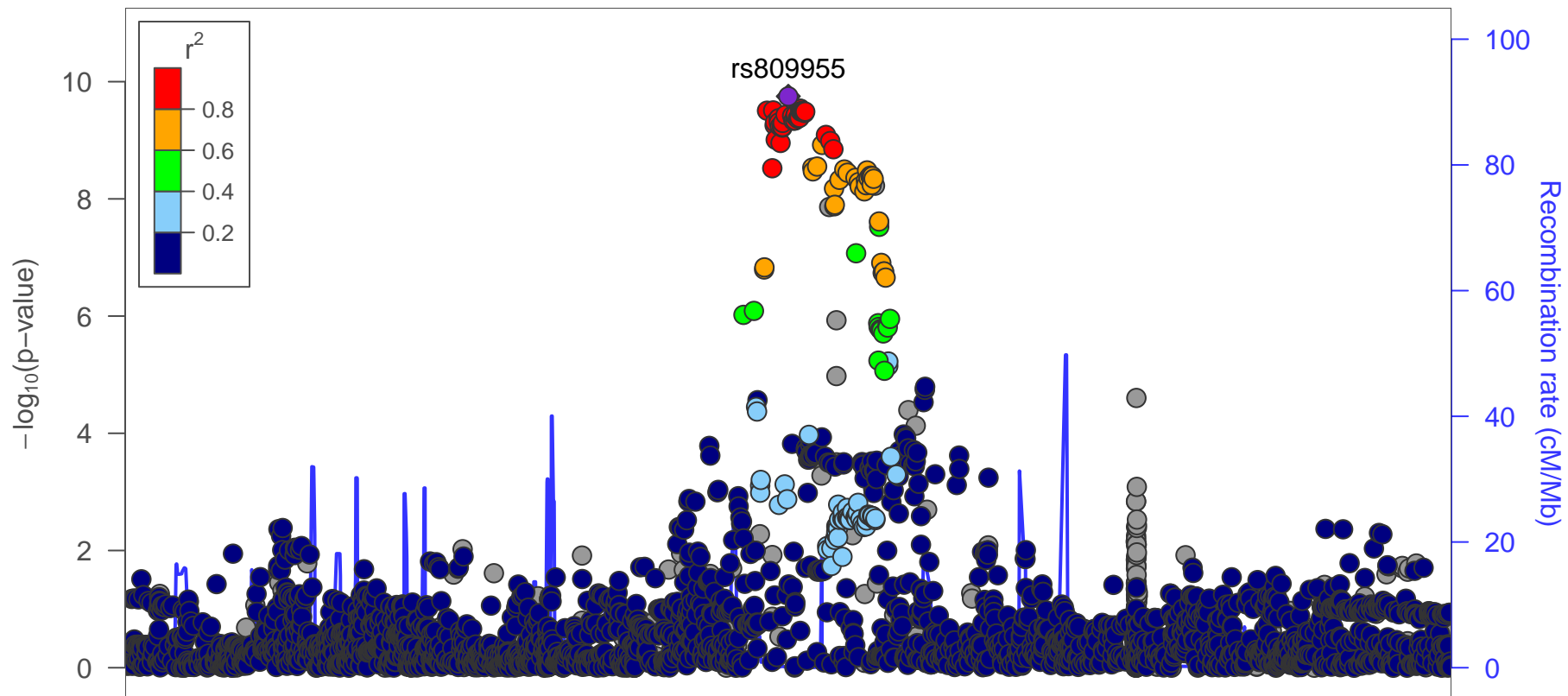
Plotted SNPs



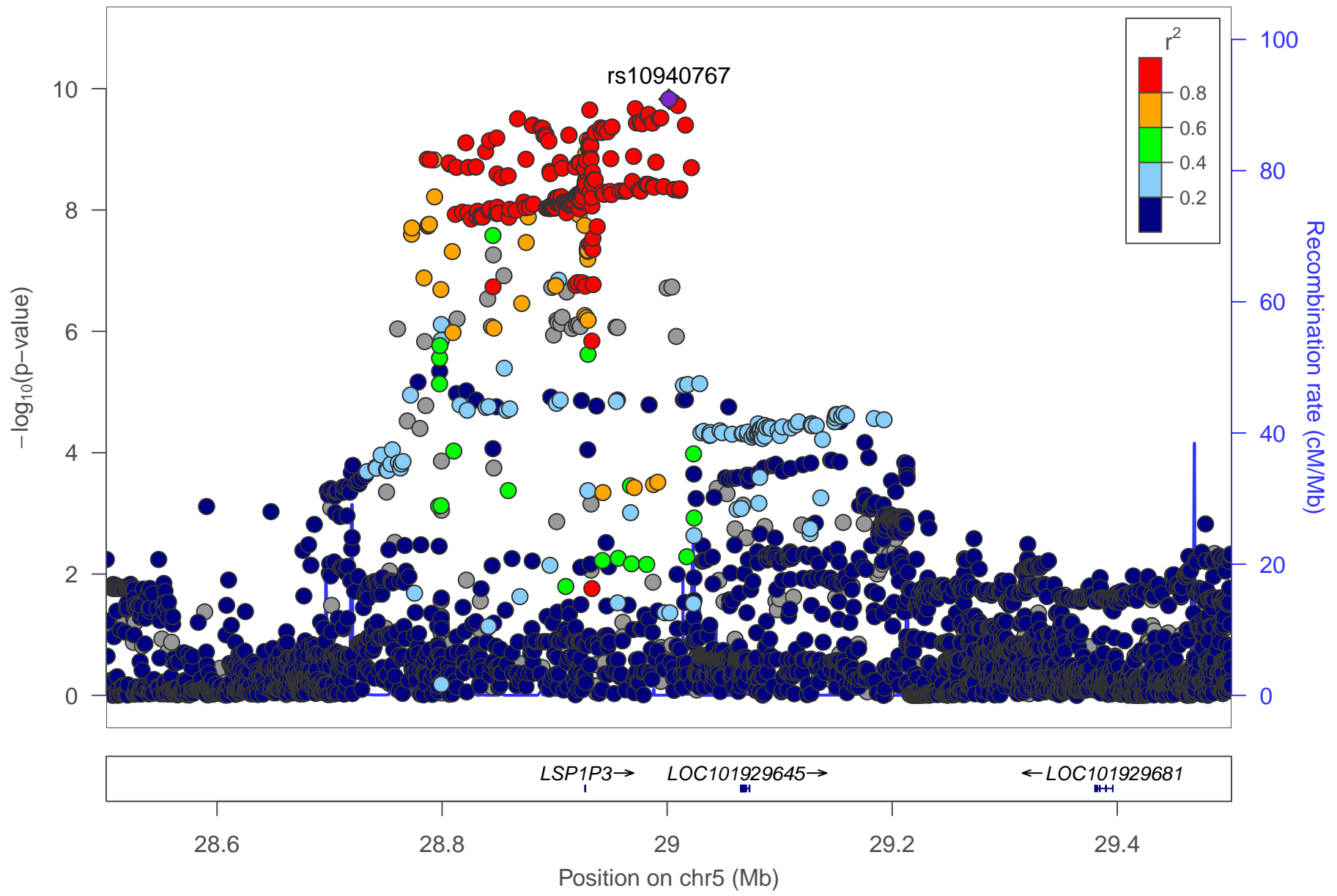
Plotted SNPs



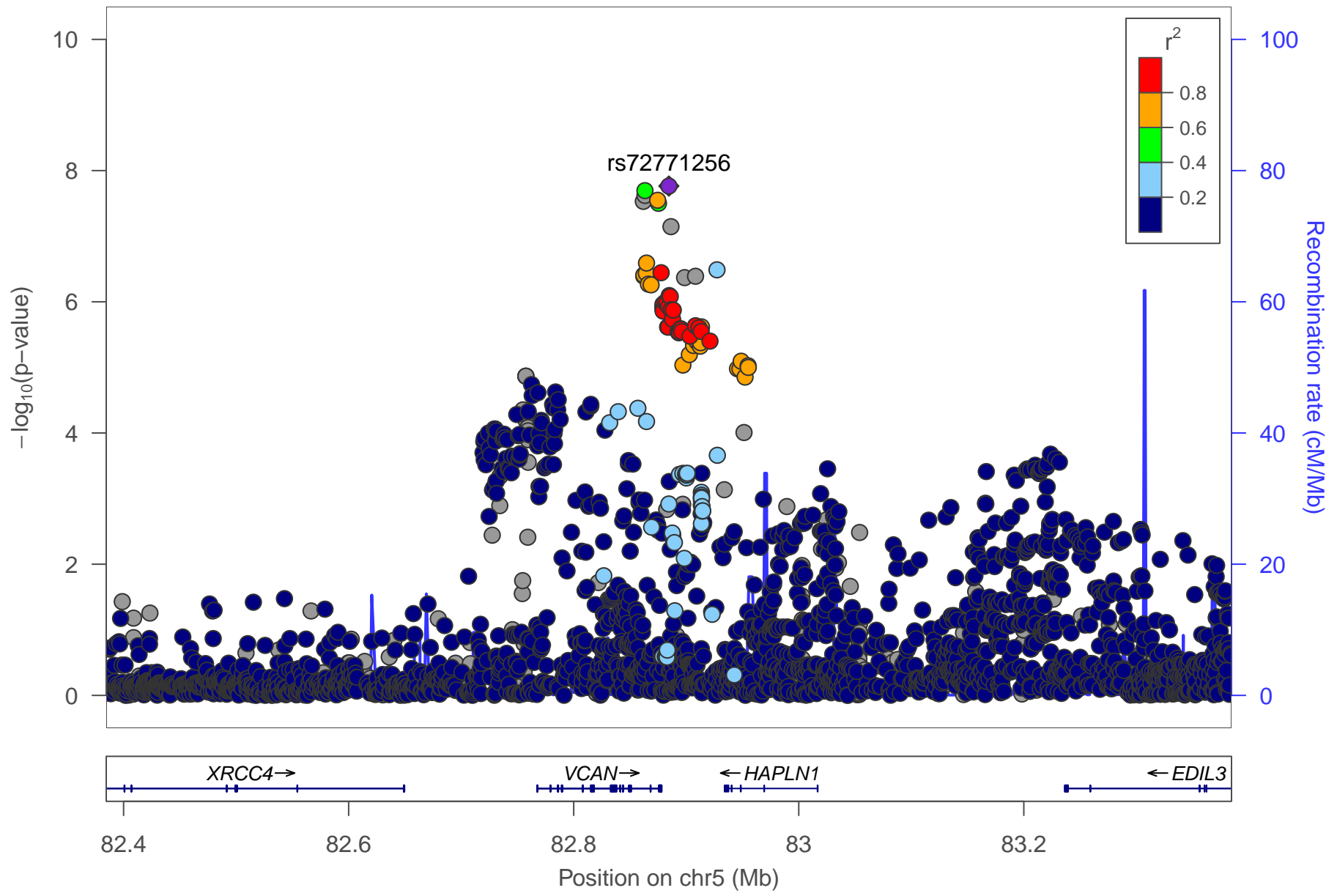
Plotted SNPs



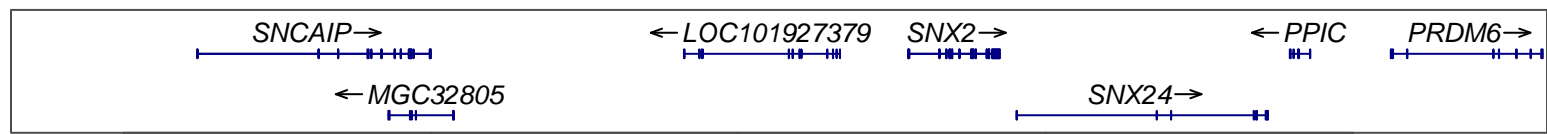
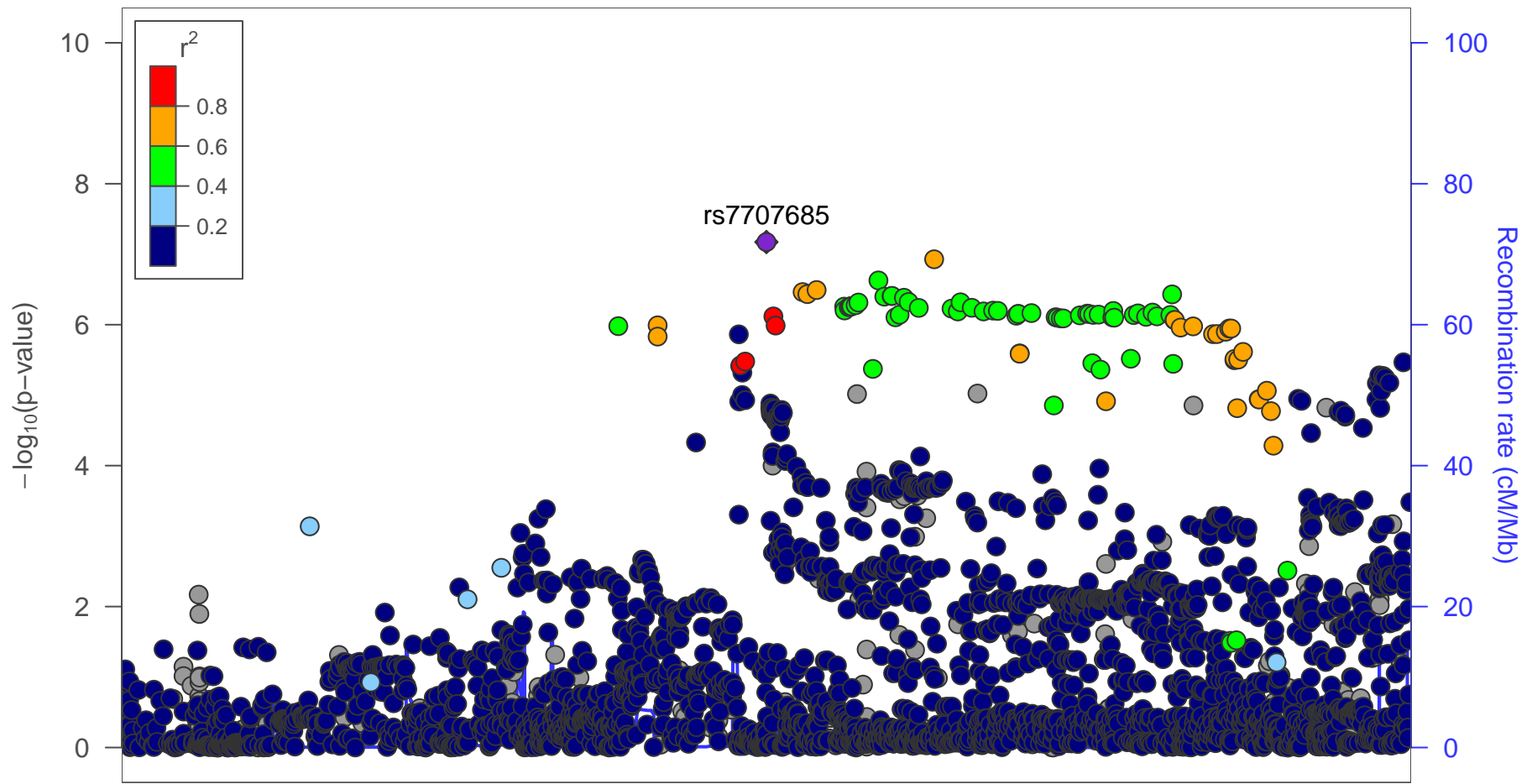
Plotted SNPs



Plotted SNPs



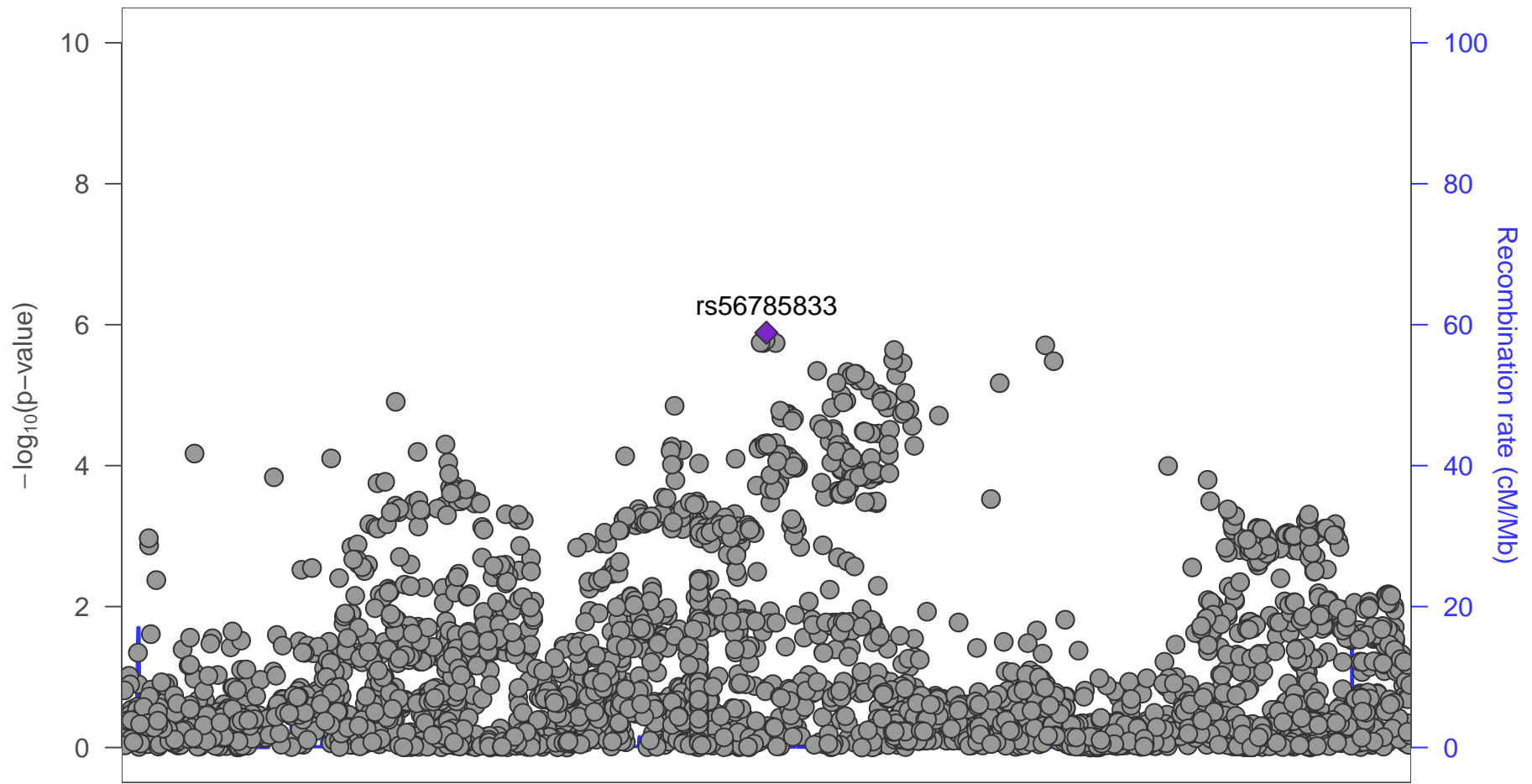
Plotted SNPs



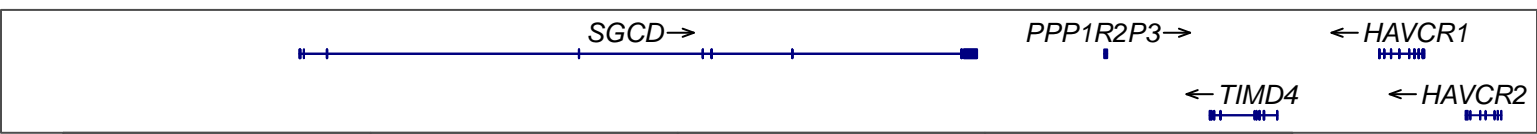
121.6 121.8 122 122.2 122.4

Position on chr5 (Mb)

Plotted SNPs



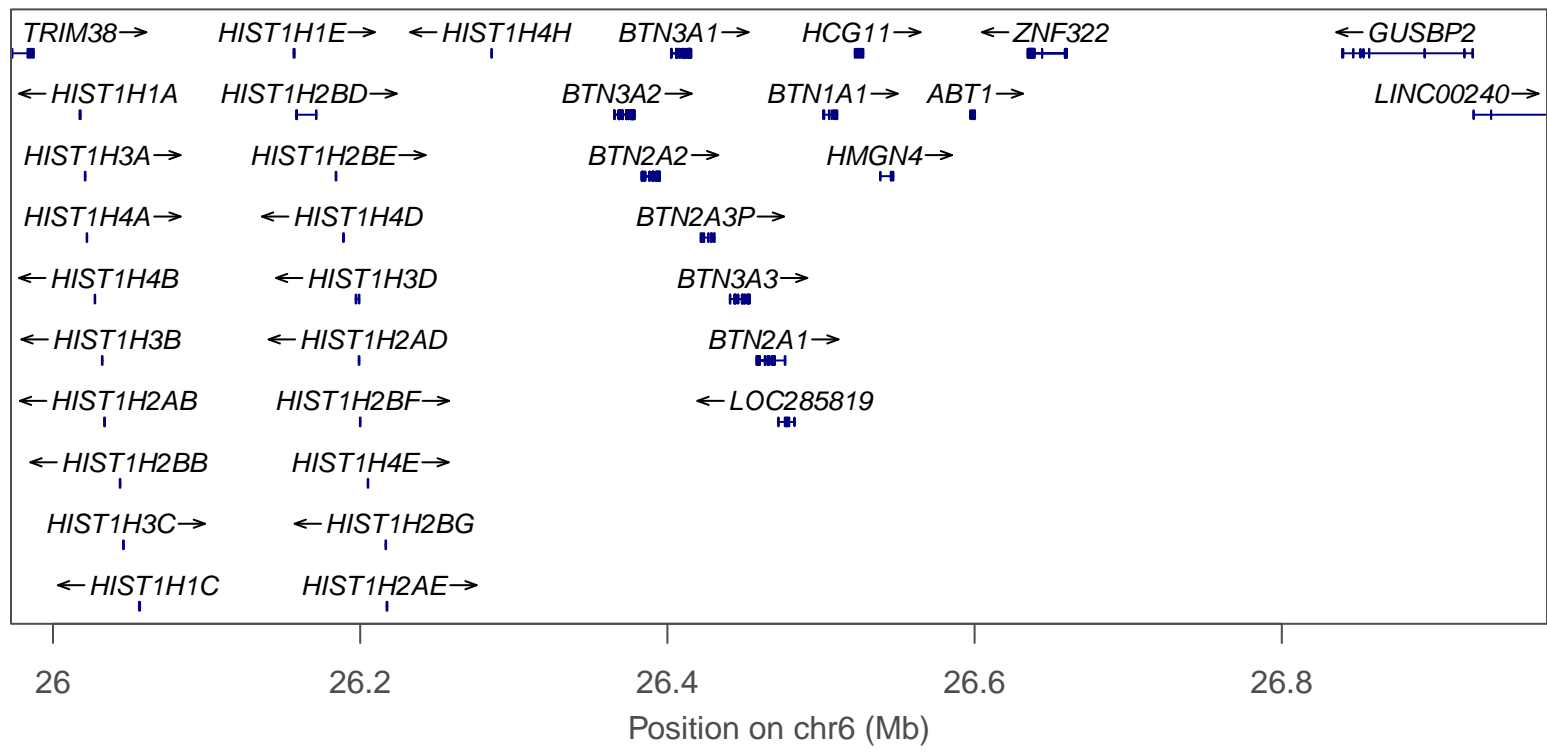
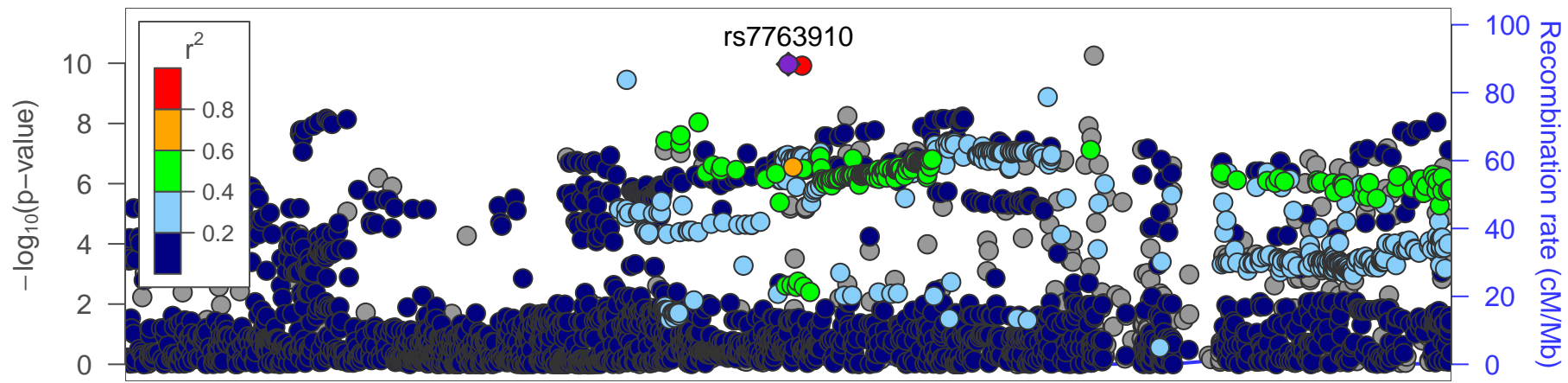
rs56785833



155.6 155.8 156 156.2 156.4

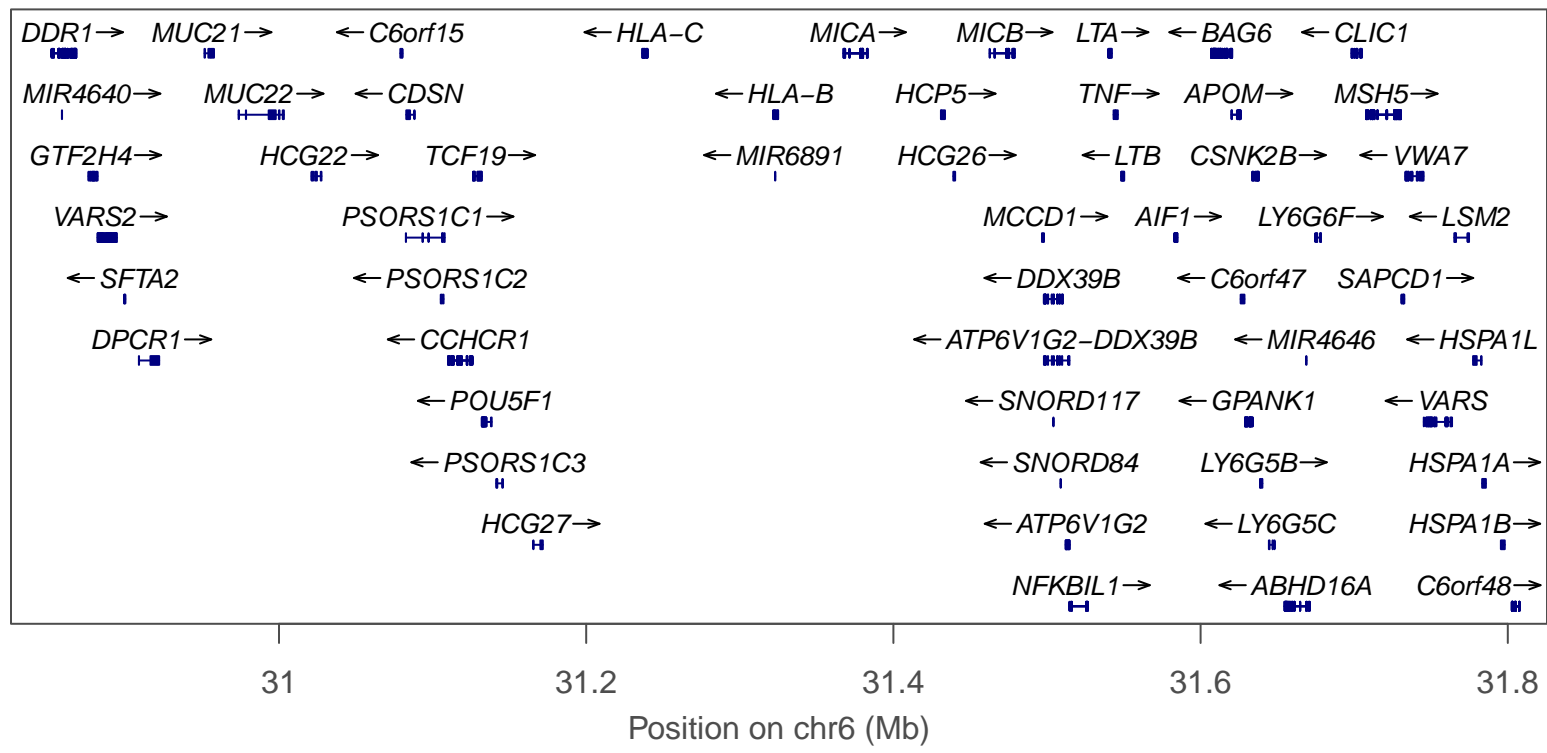
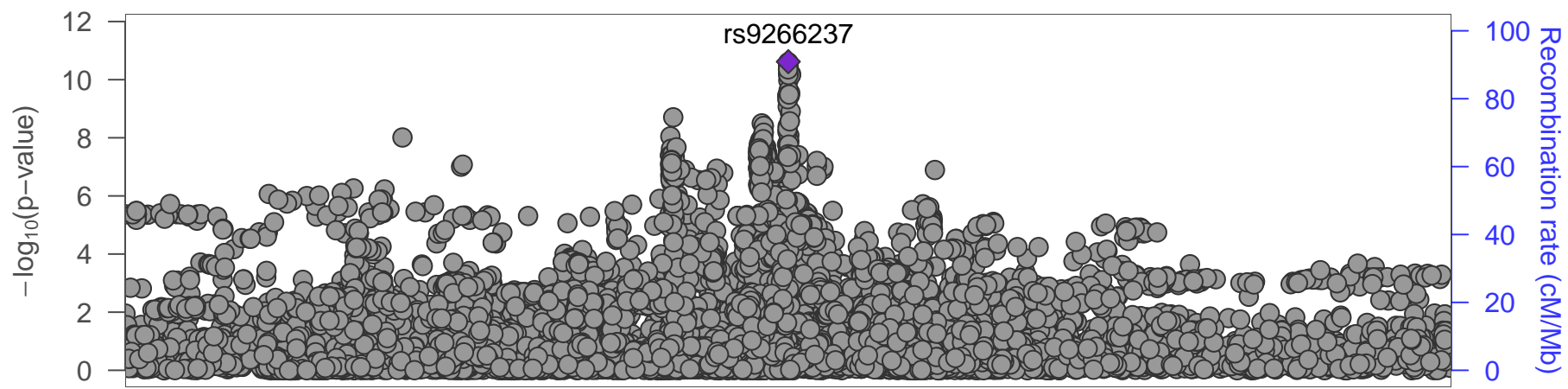
Position on chr5 (Mb)

Plotted SNPs

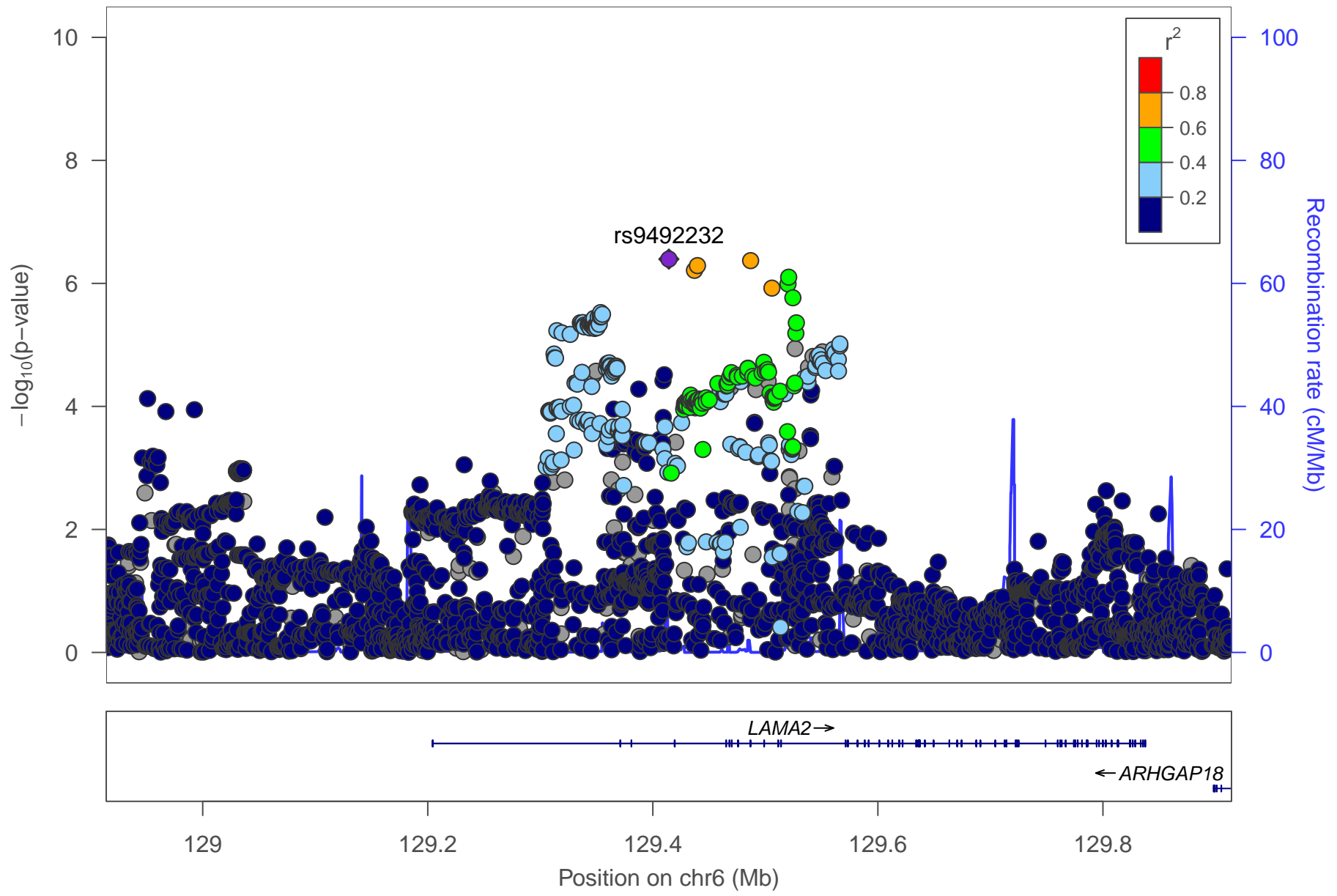


13 genes omitted

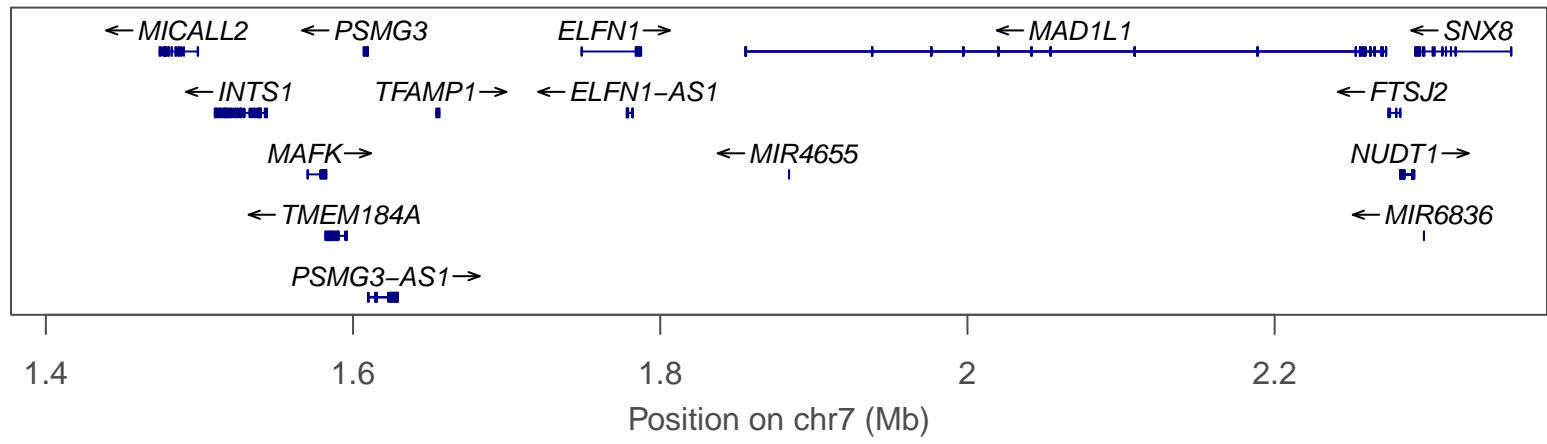
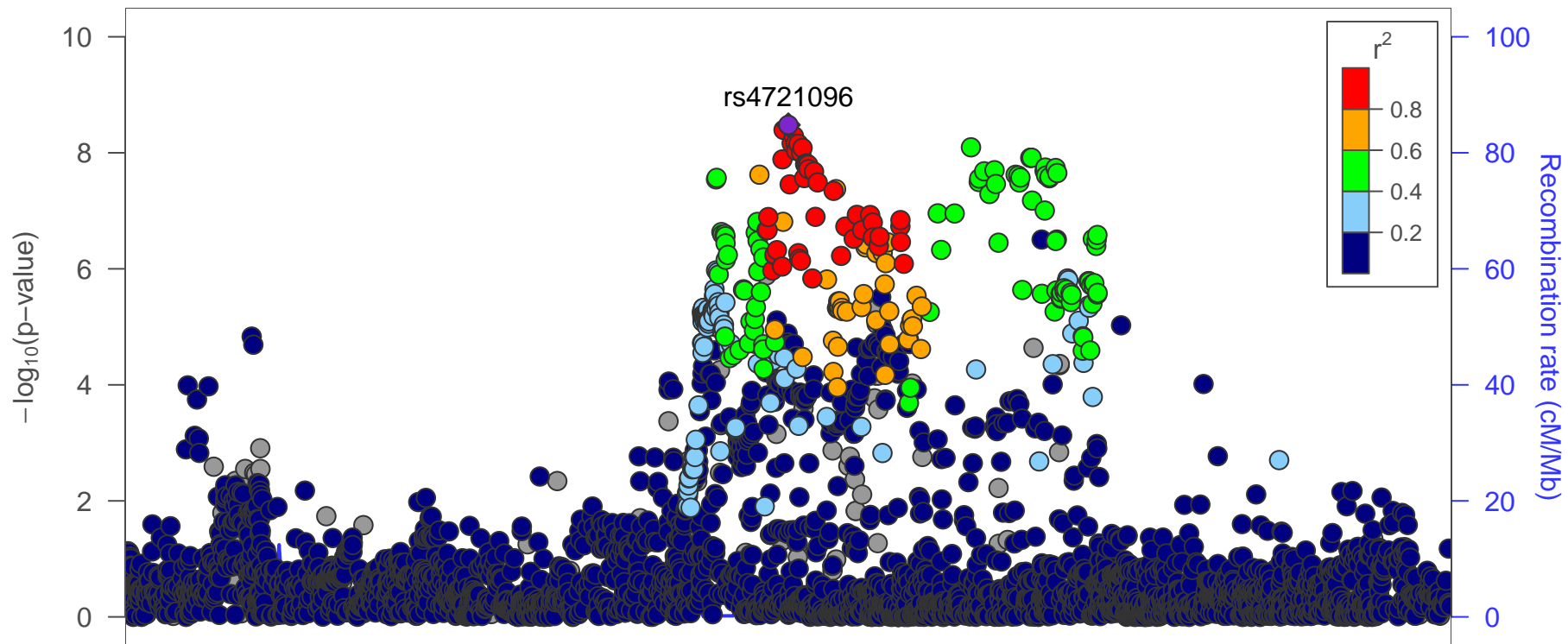
Plotted SNPs

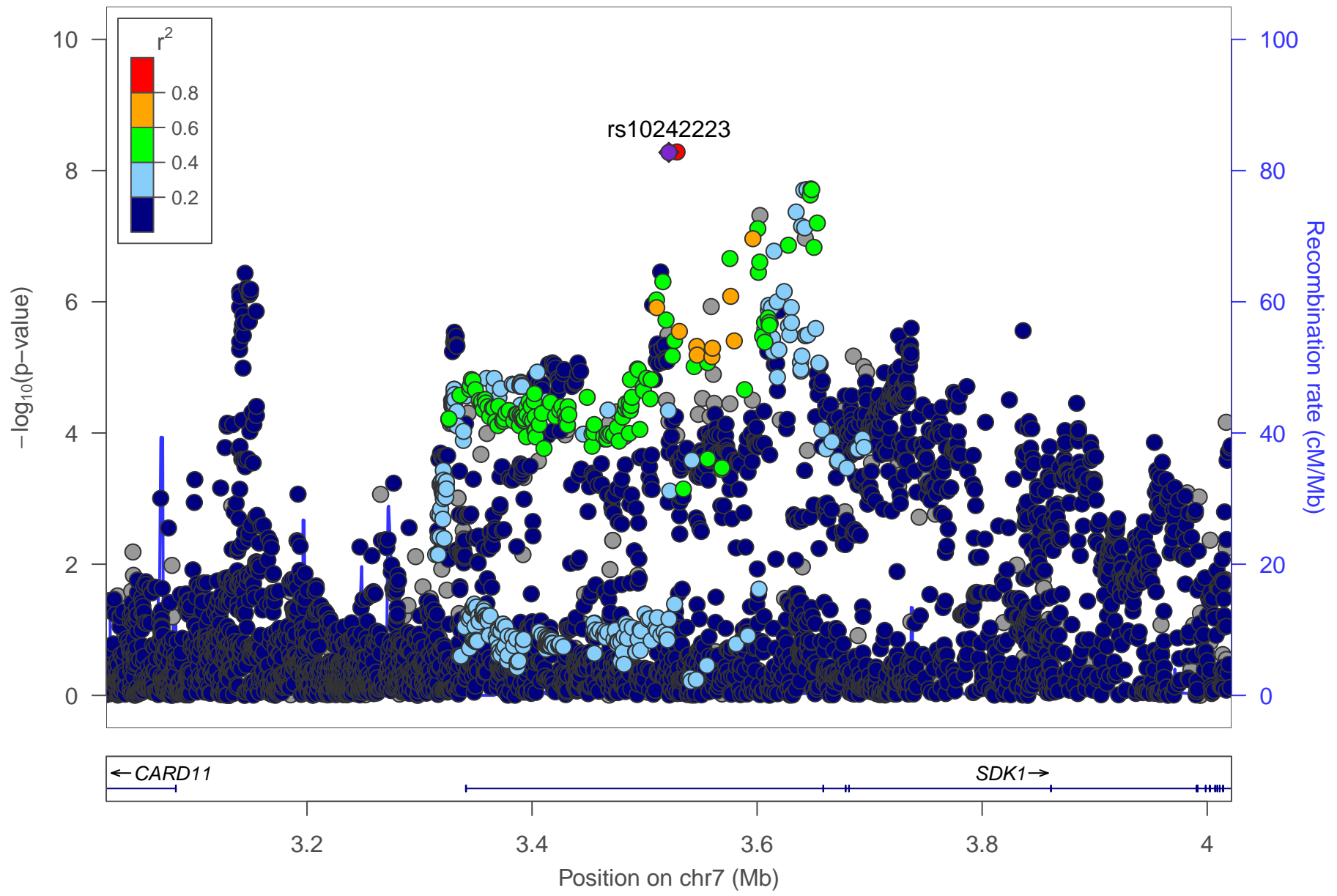


Plotted SNPs

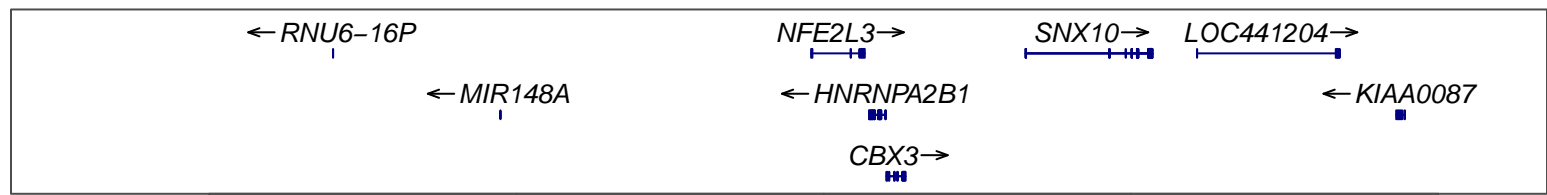
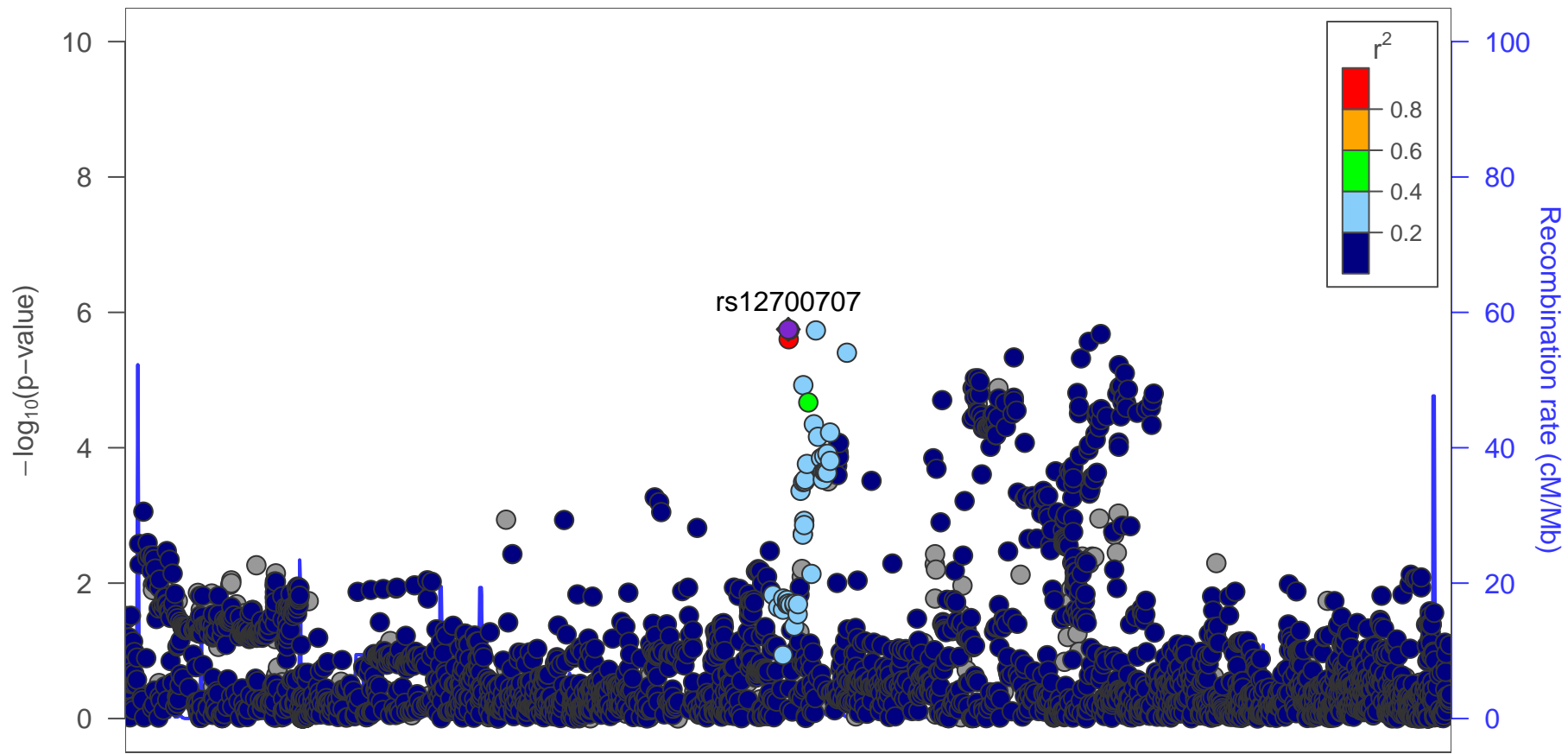


Plotted SNPs





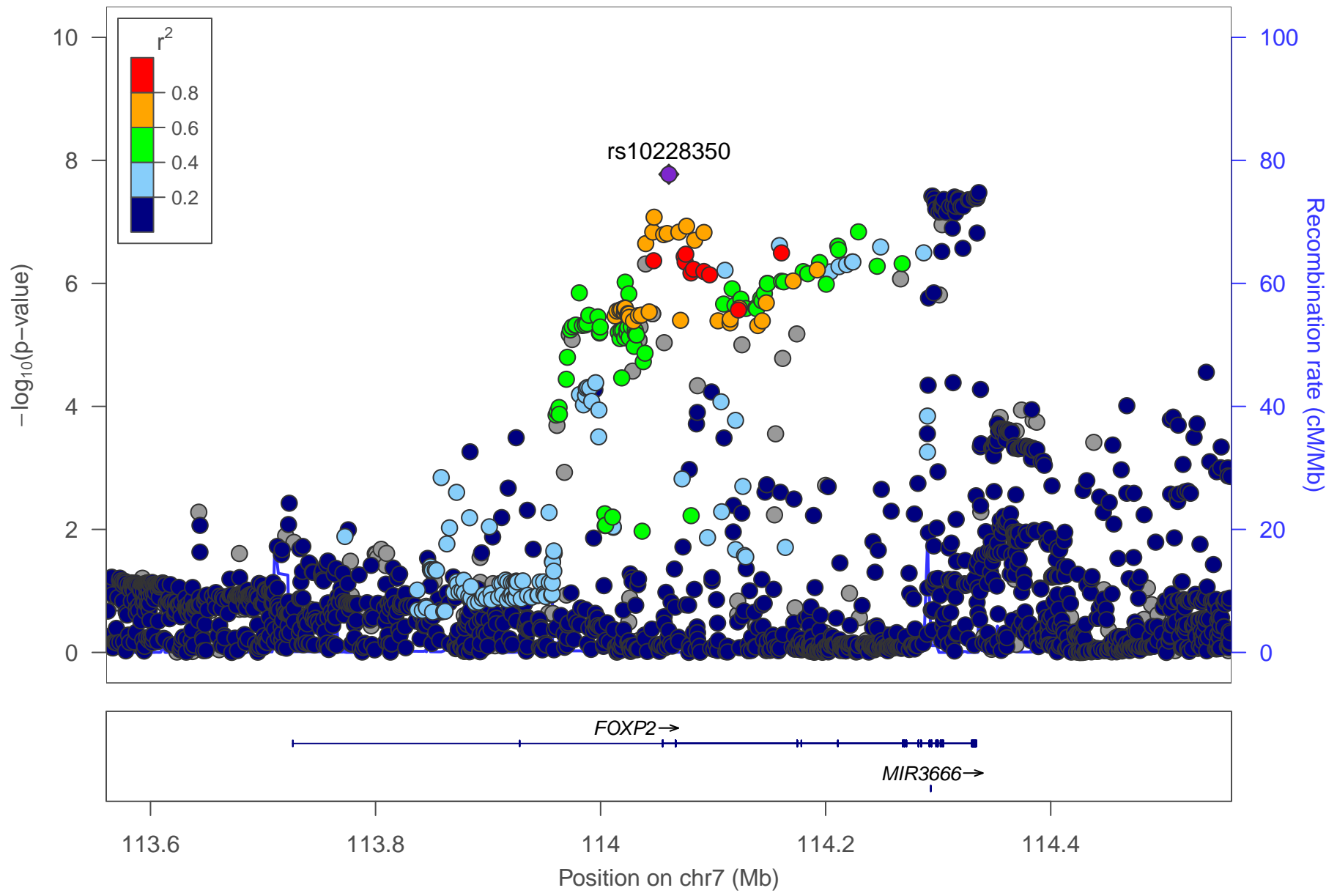
Plotted SNPs



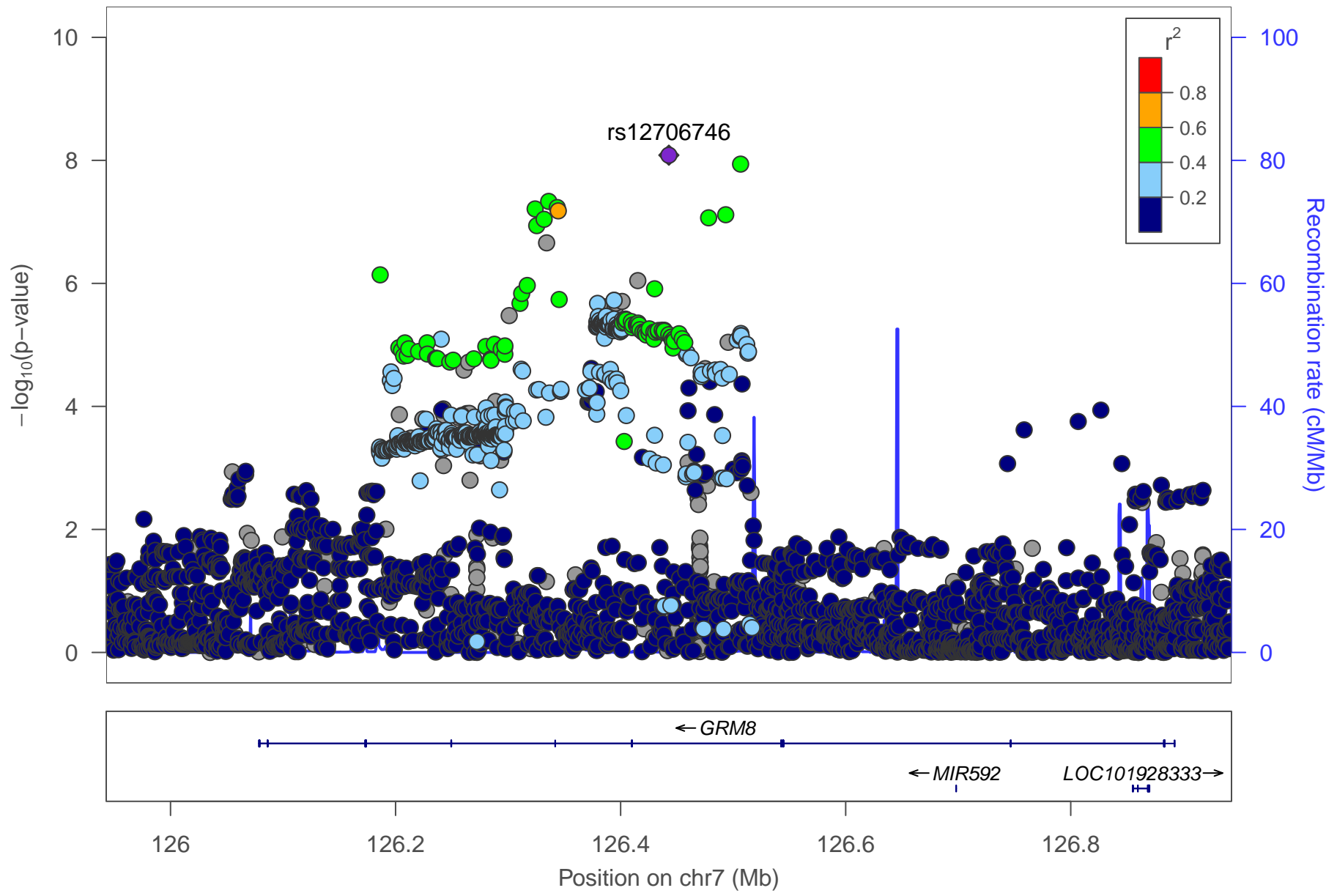
25.8 26 26.2 26.4 26.6

Position on chr7 (Mb)

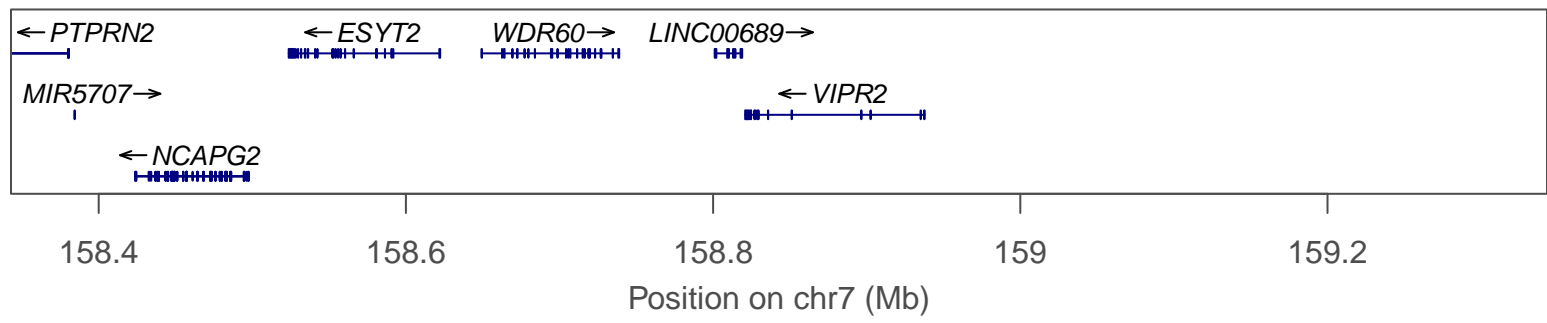
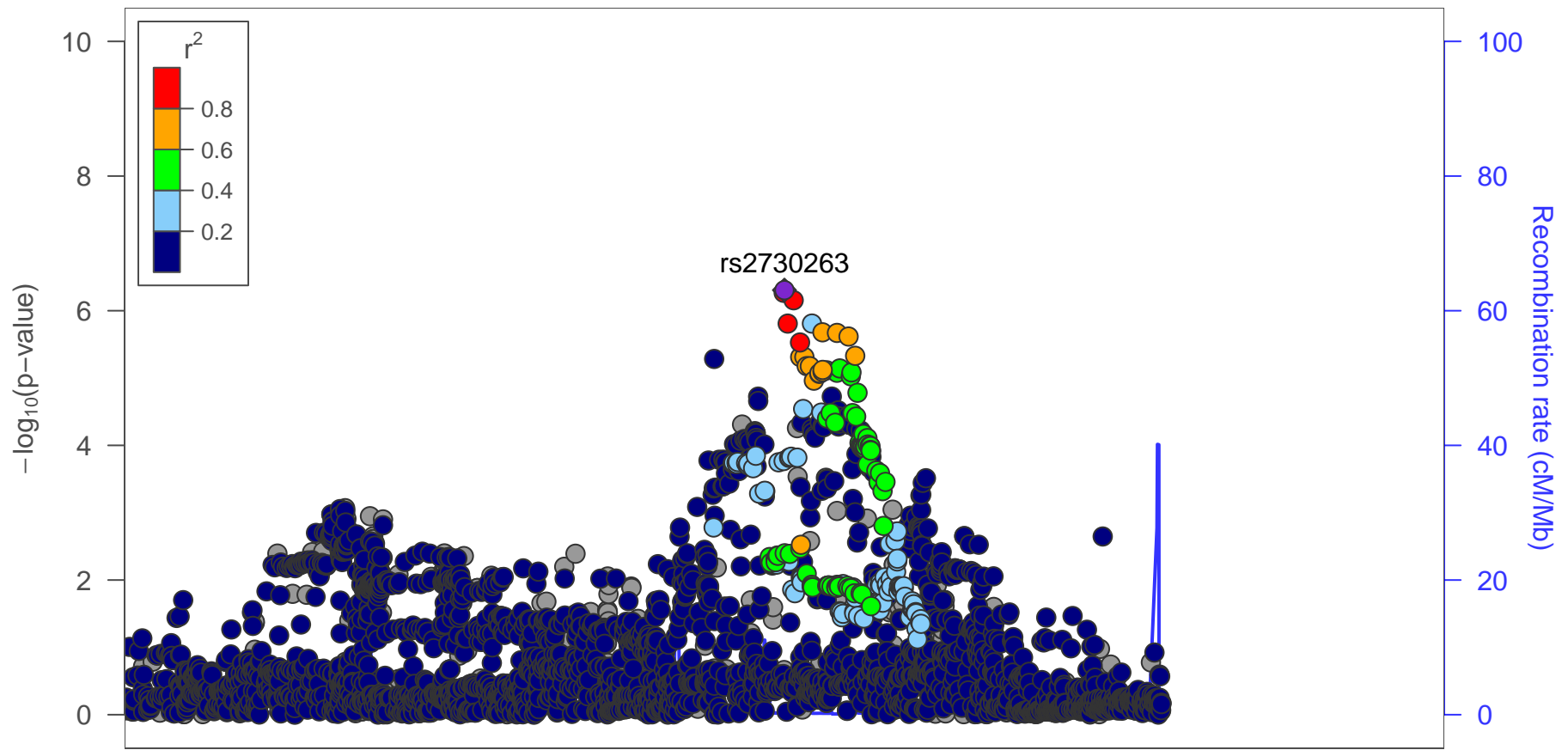
Plotted SNPs



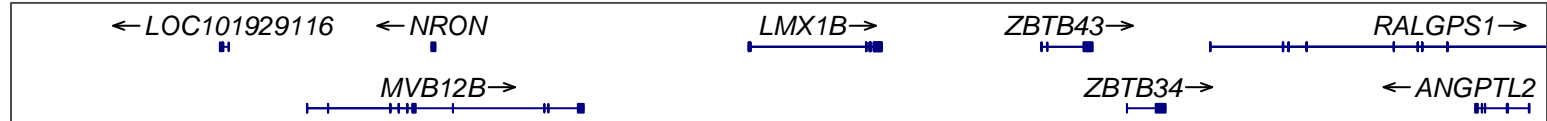
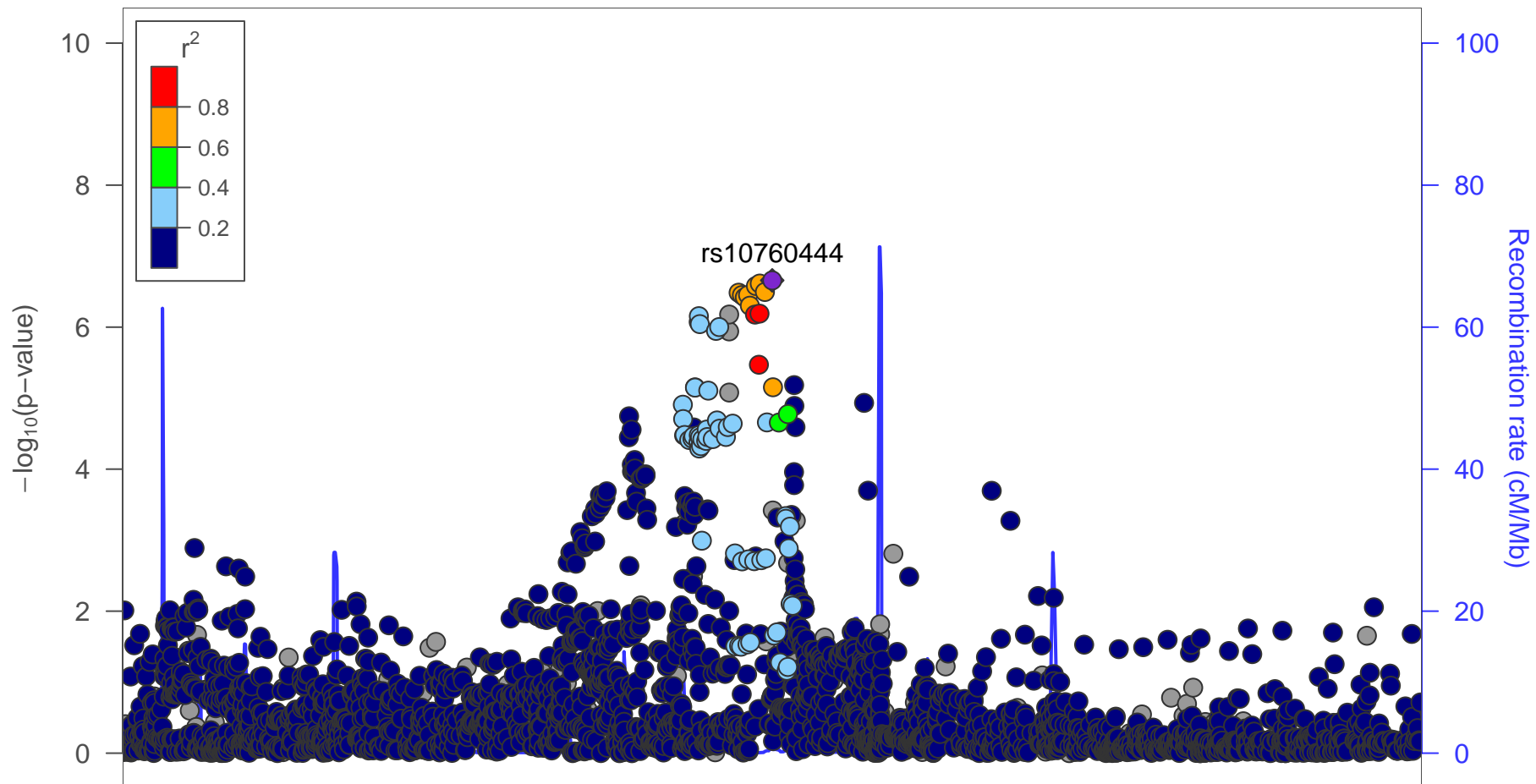
Plotted SNPs



Plotted SNPs



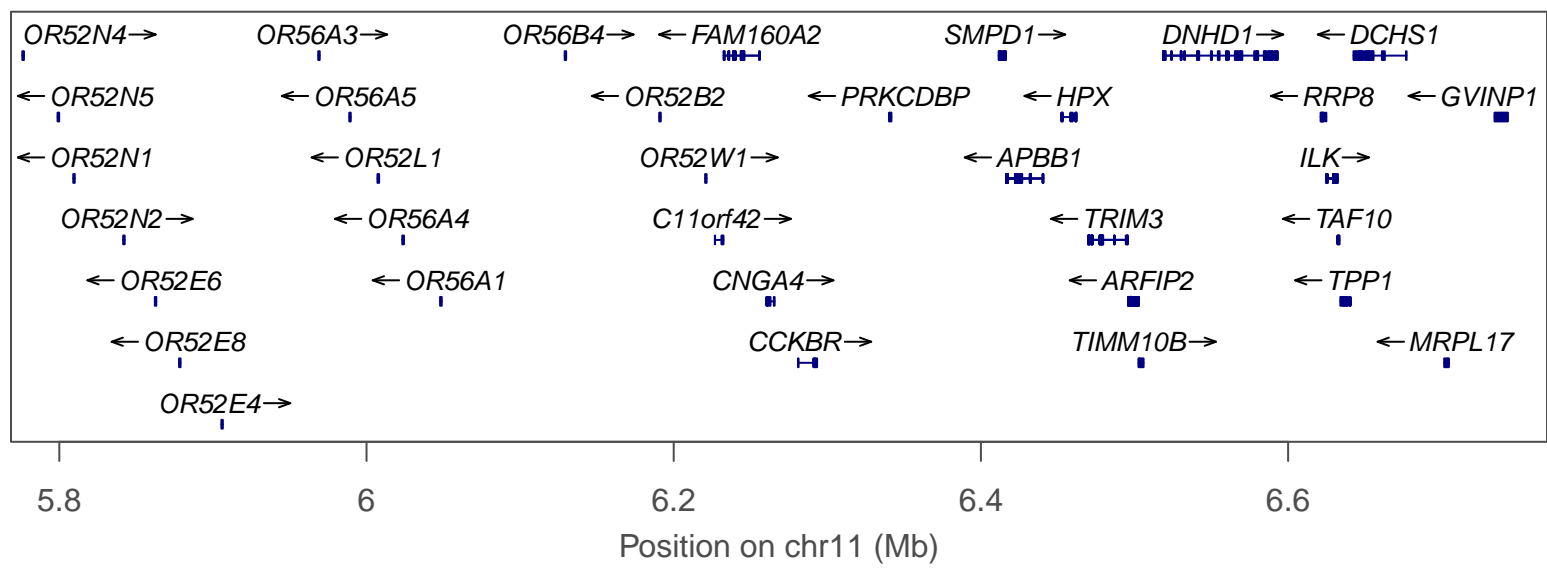
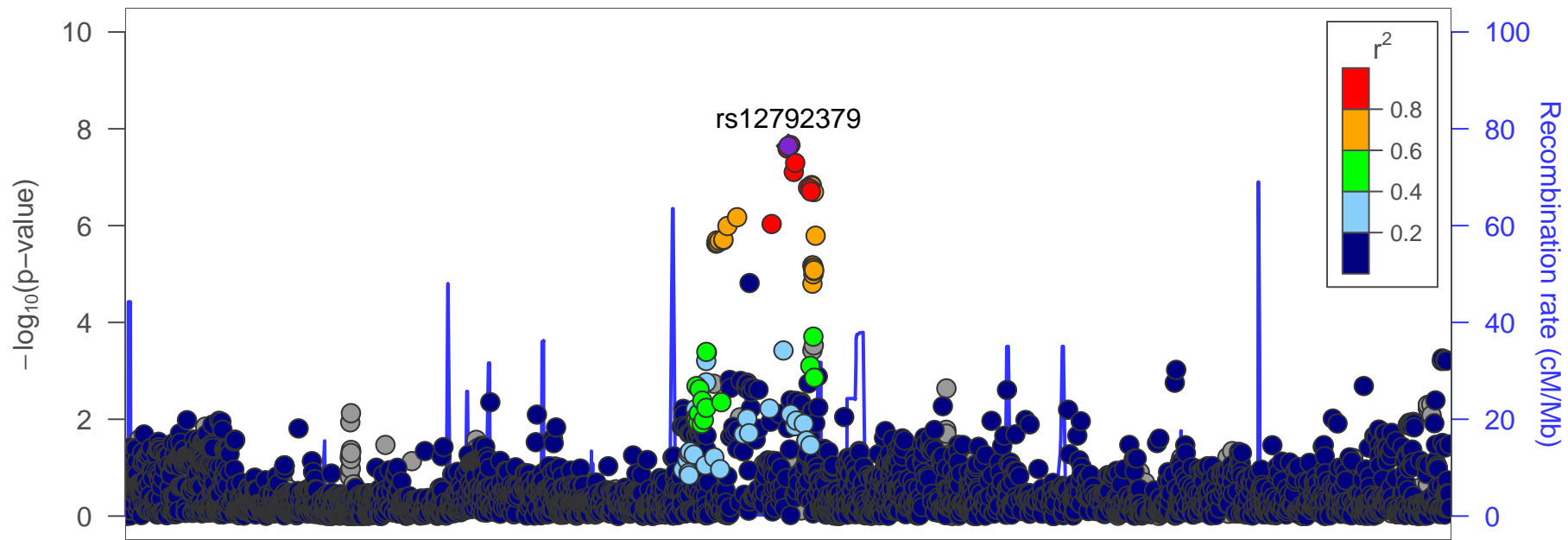
Plotted SNPs



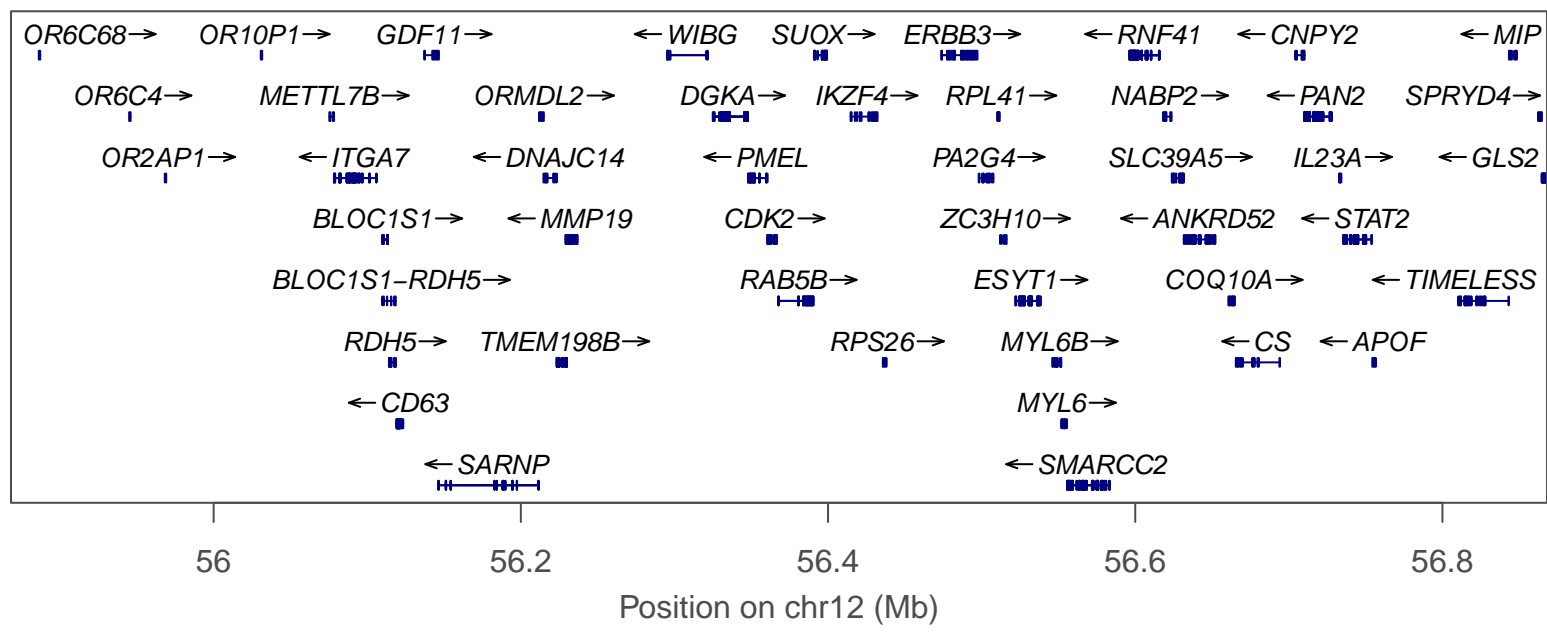
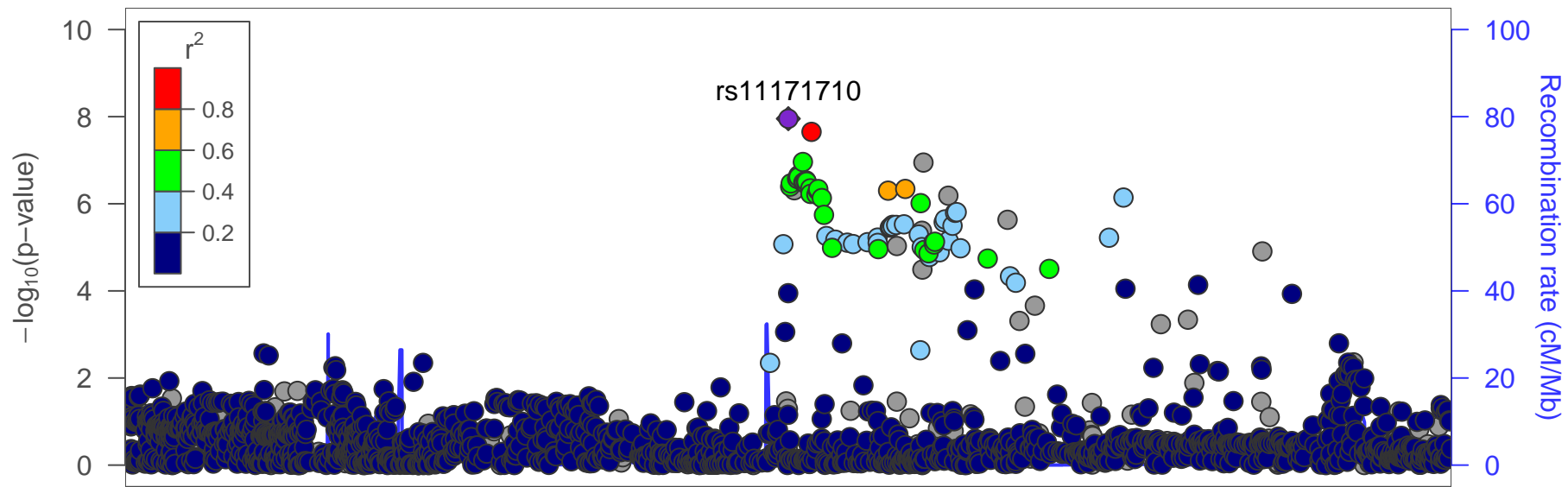
129 129.2 129.4 129.6 129.8

Position on chr9 (Mb)

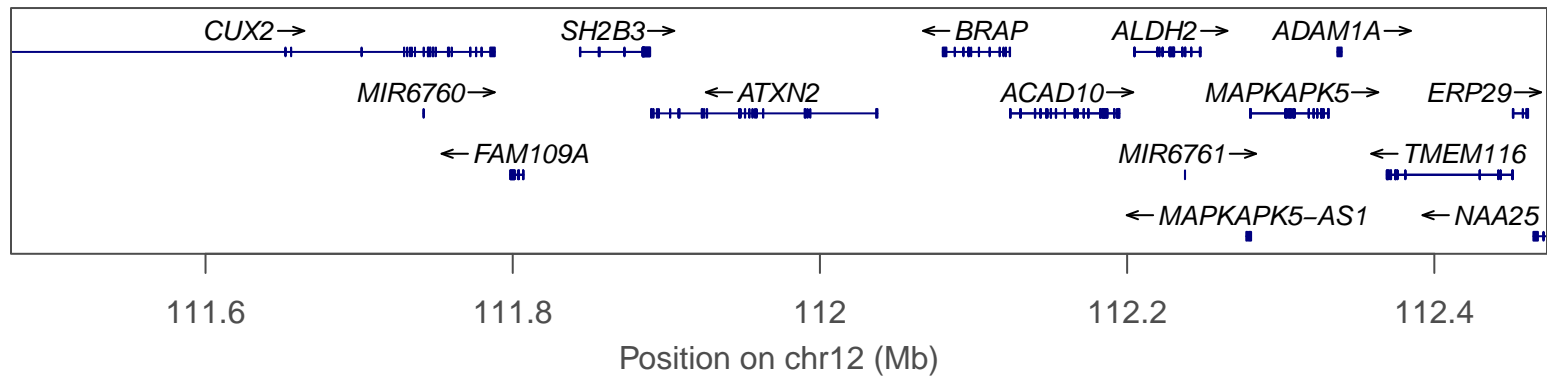
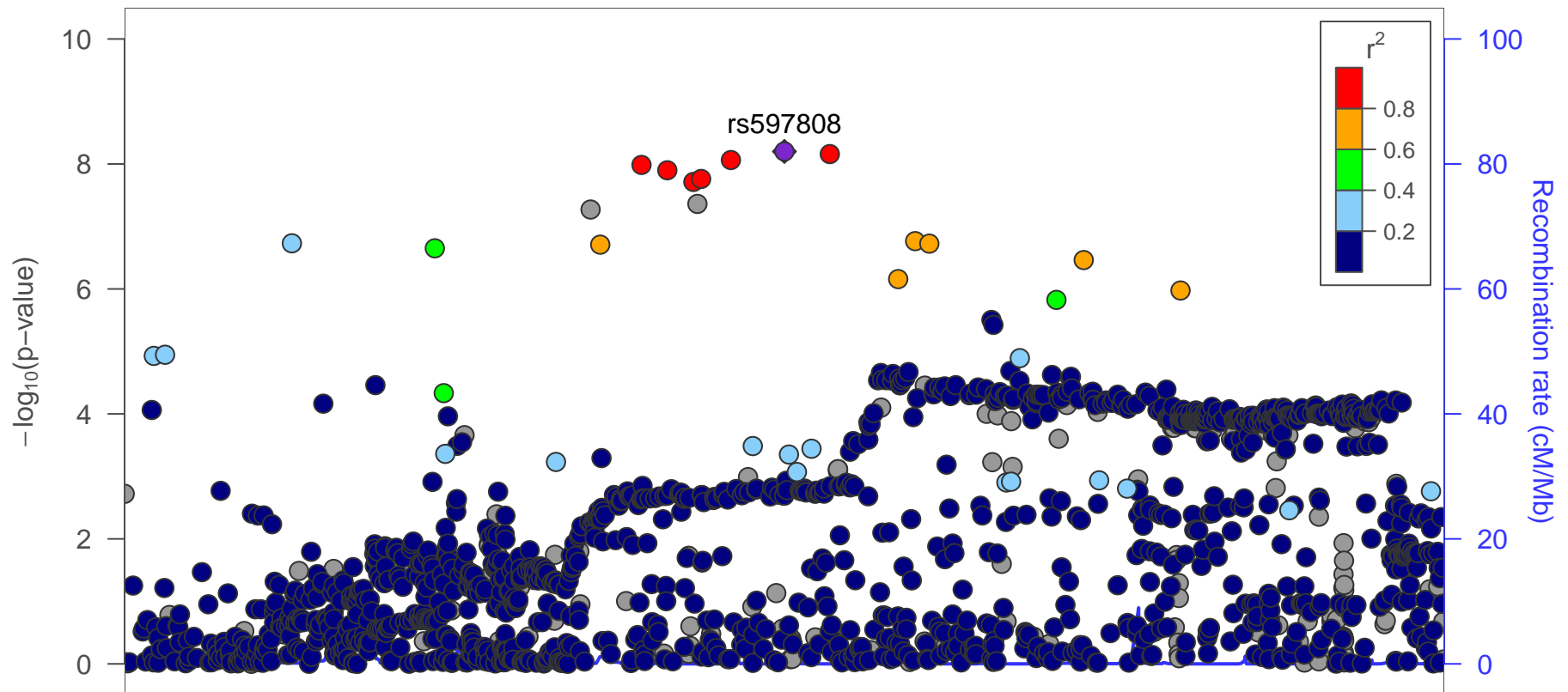
Plotted SNPs



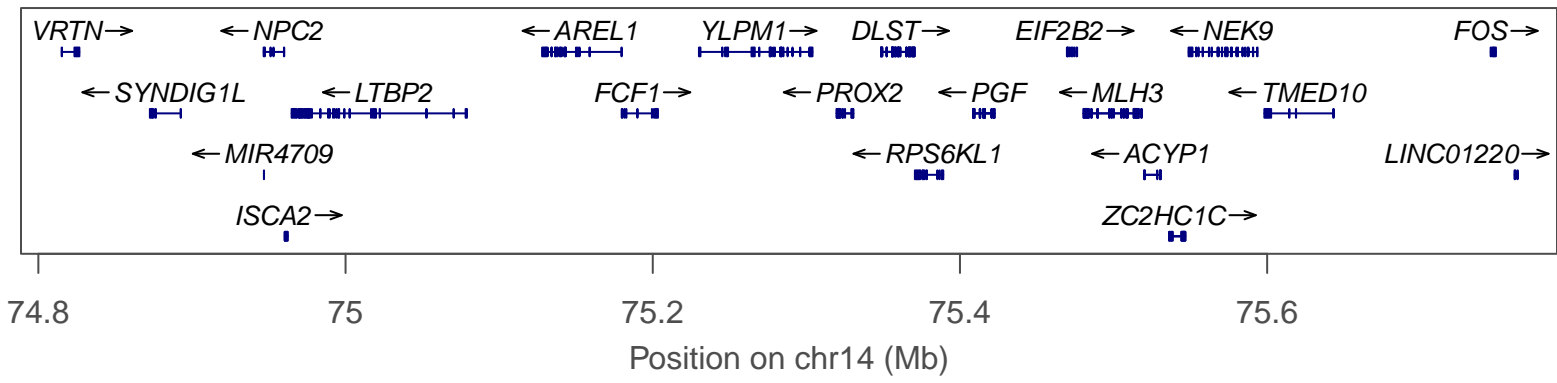
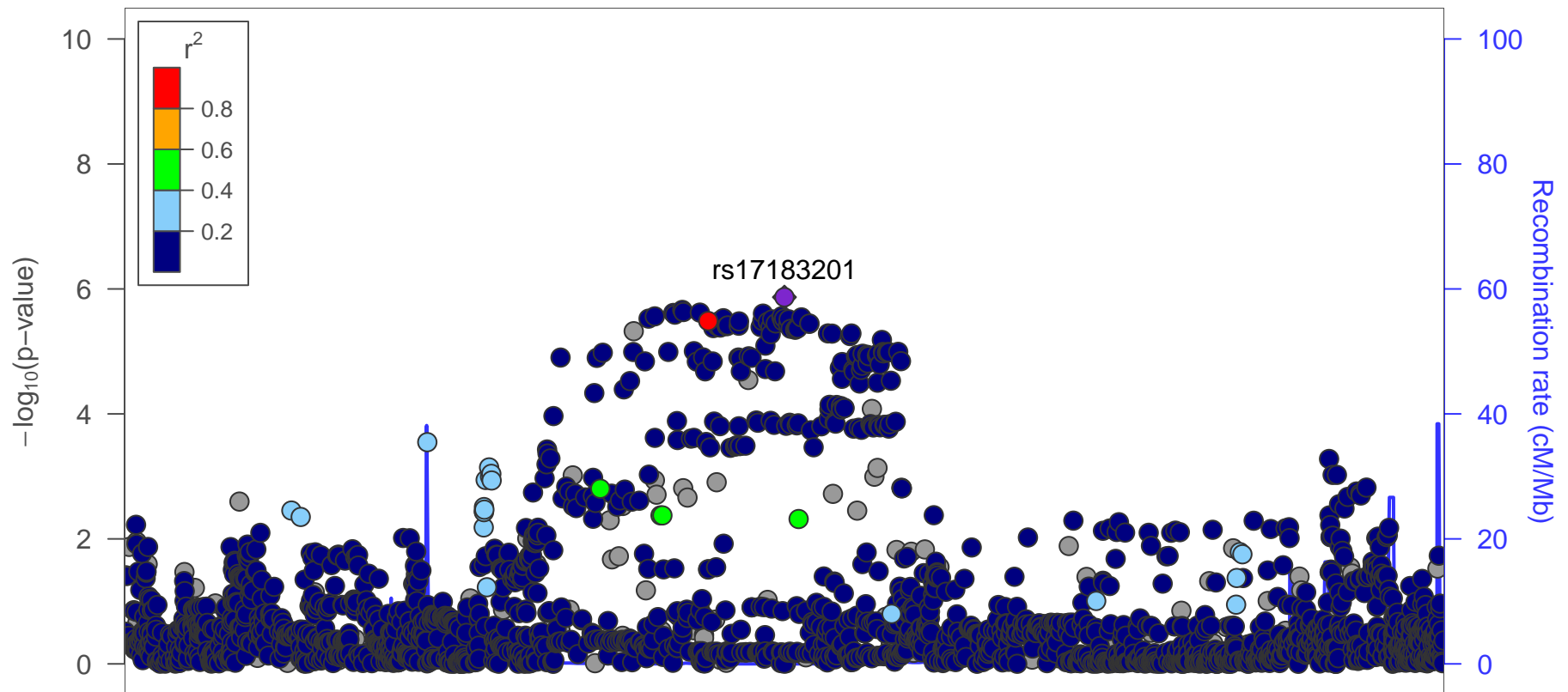
Plotted SNPs



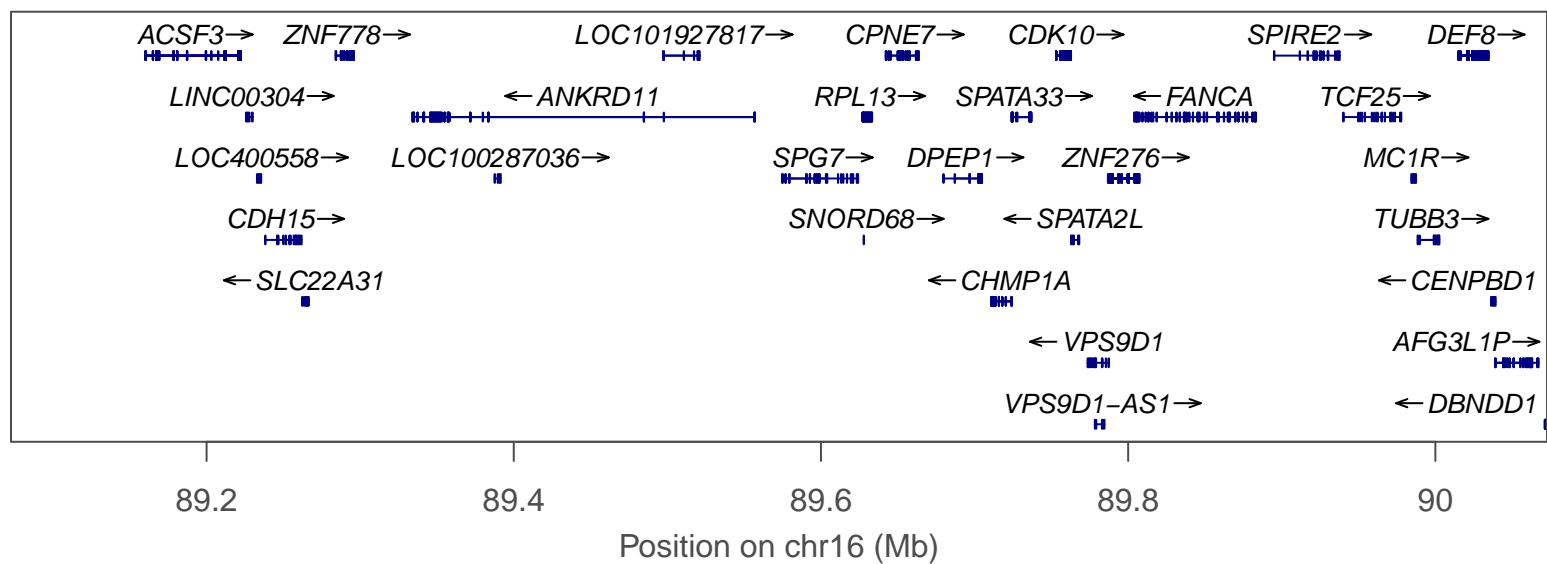
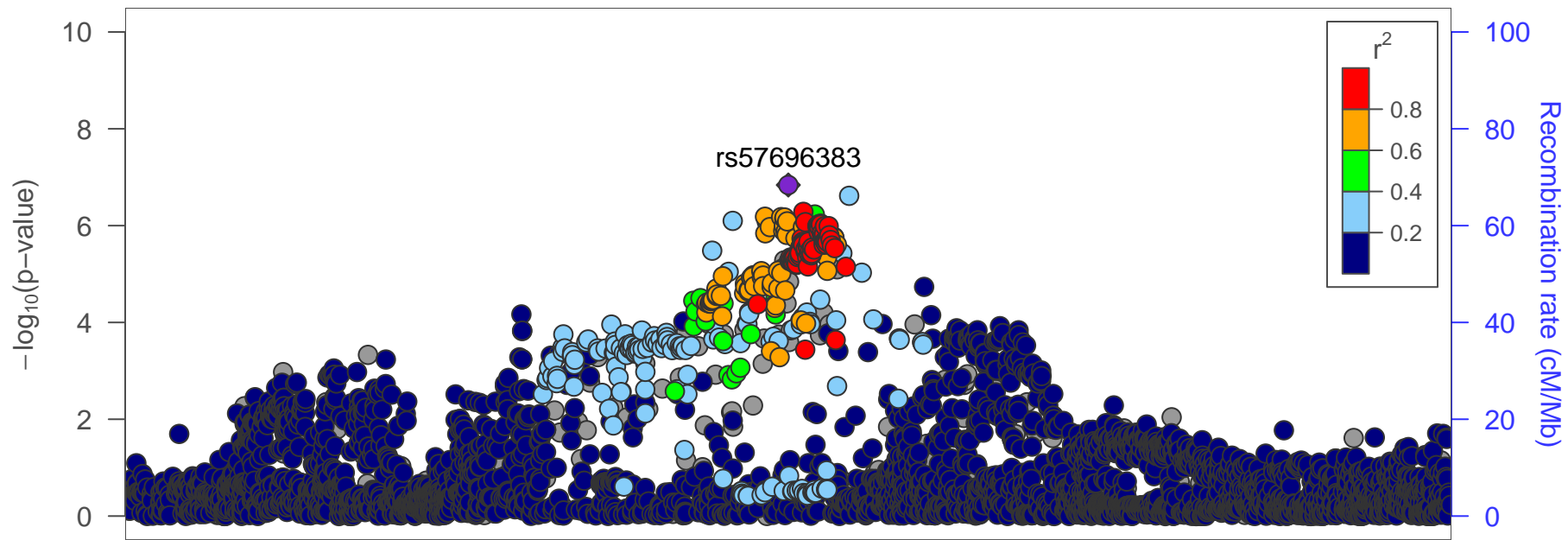
Plotted SNPs



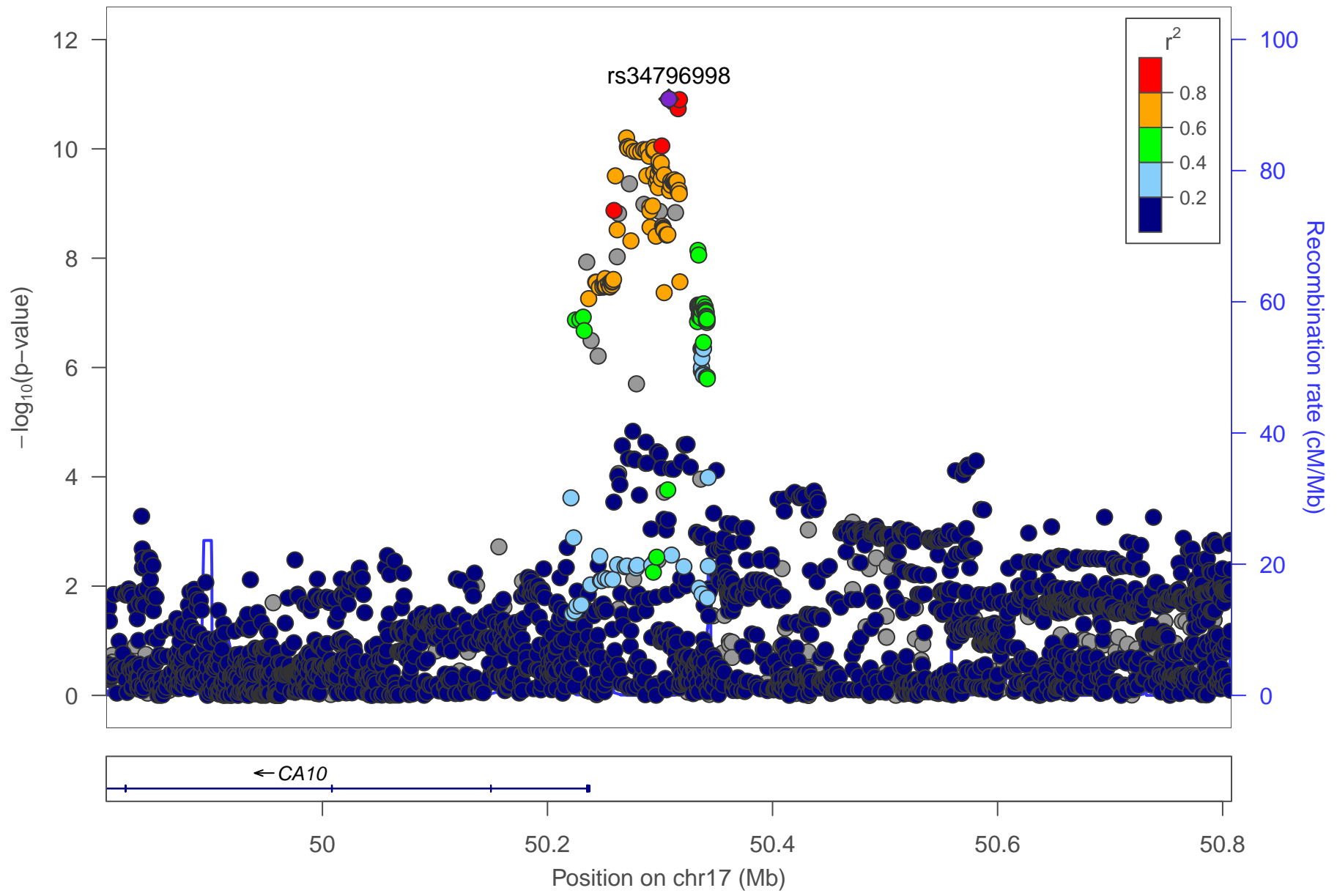
Plotted SNPs

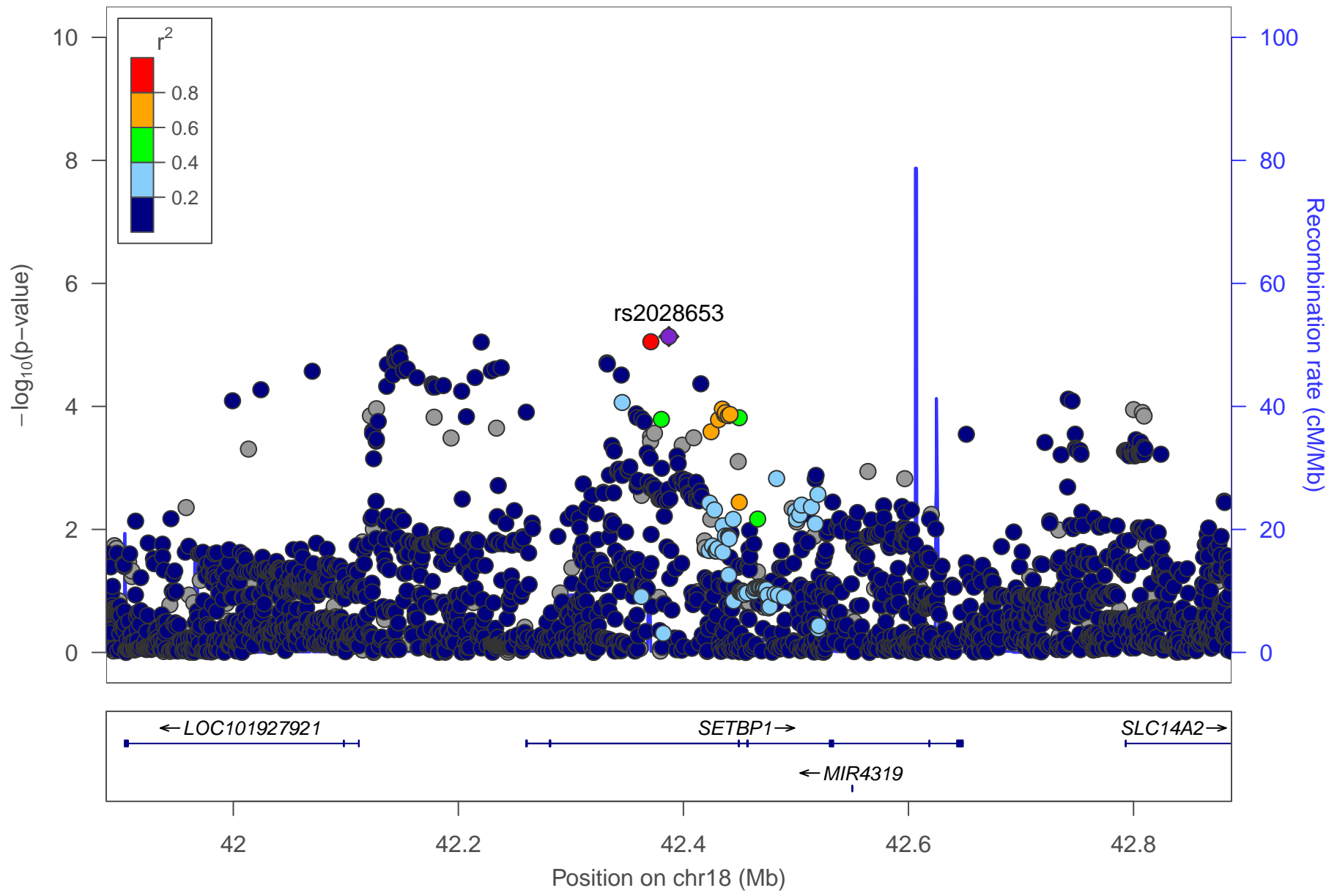


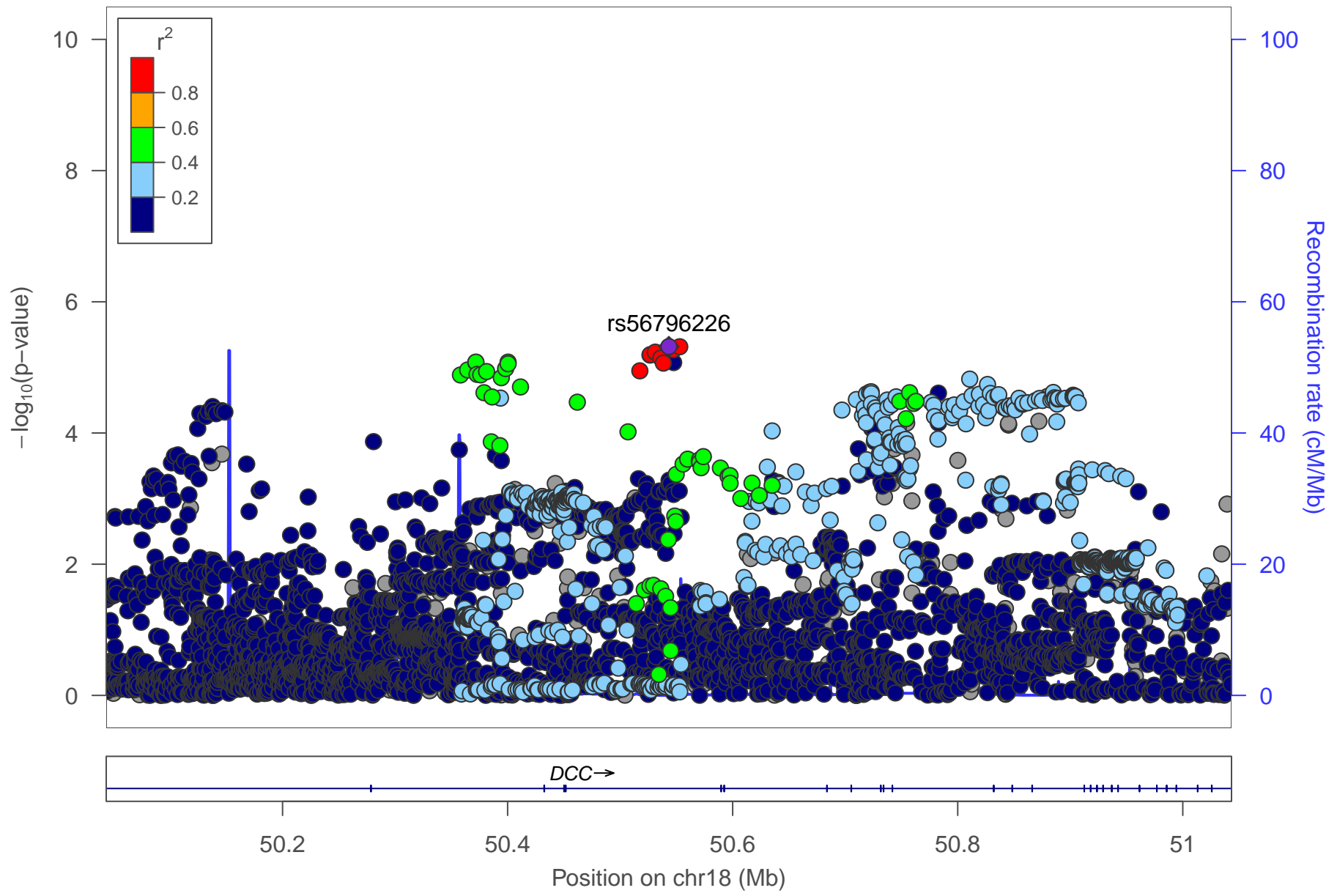
Plotted SNPs



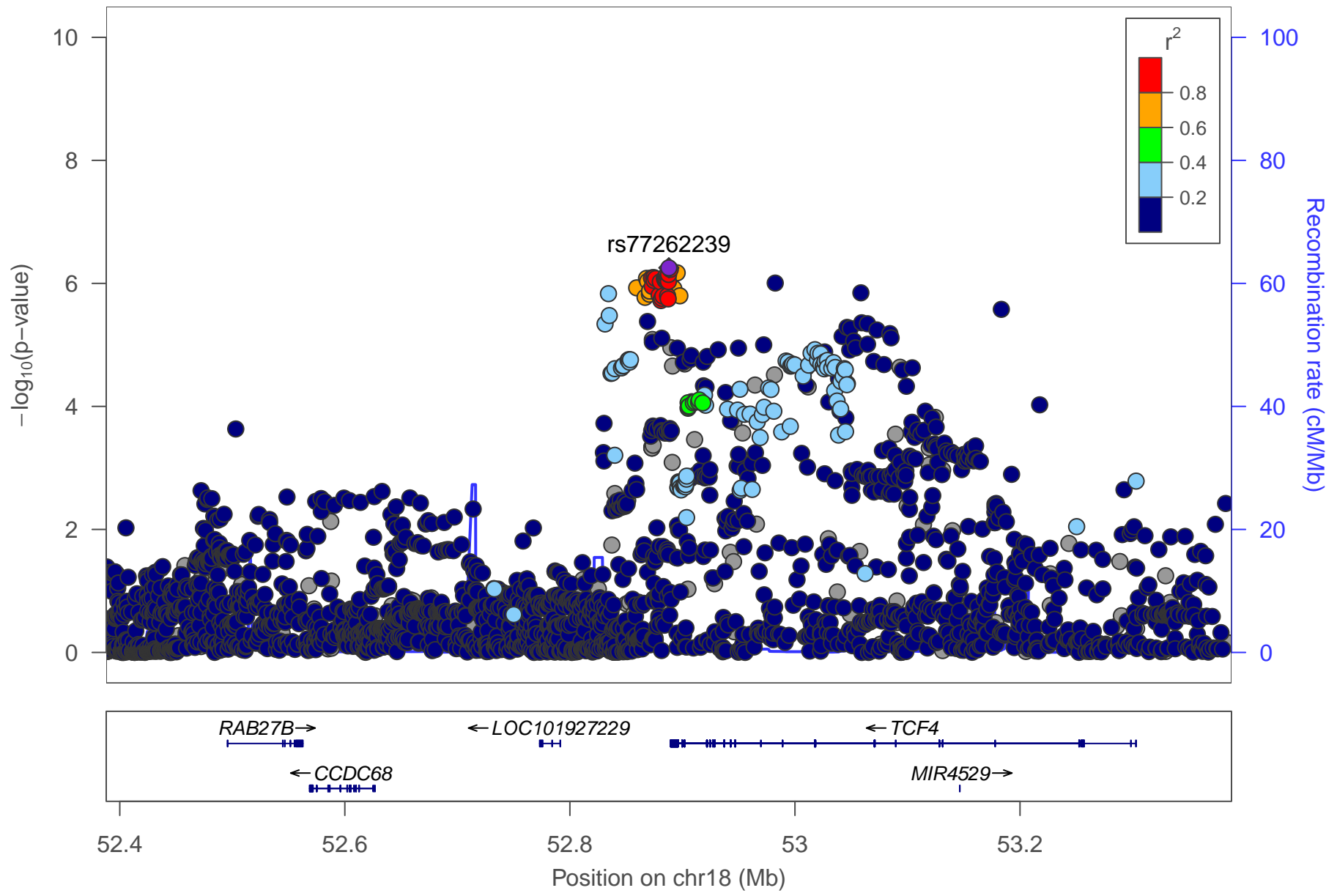
Plotted SNPs



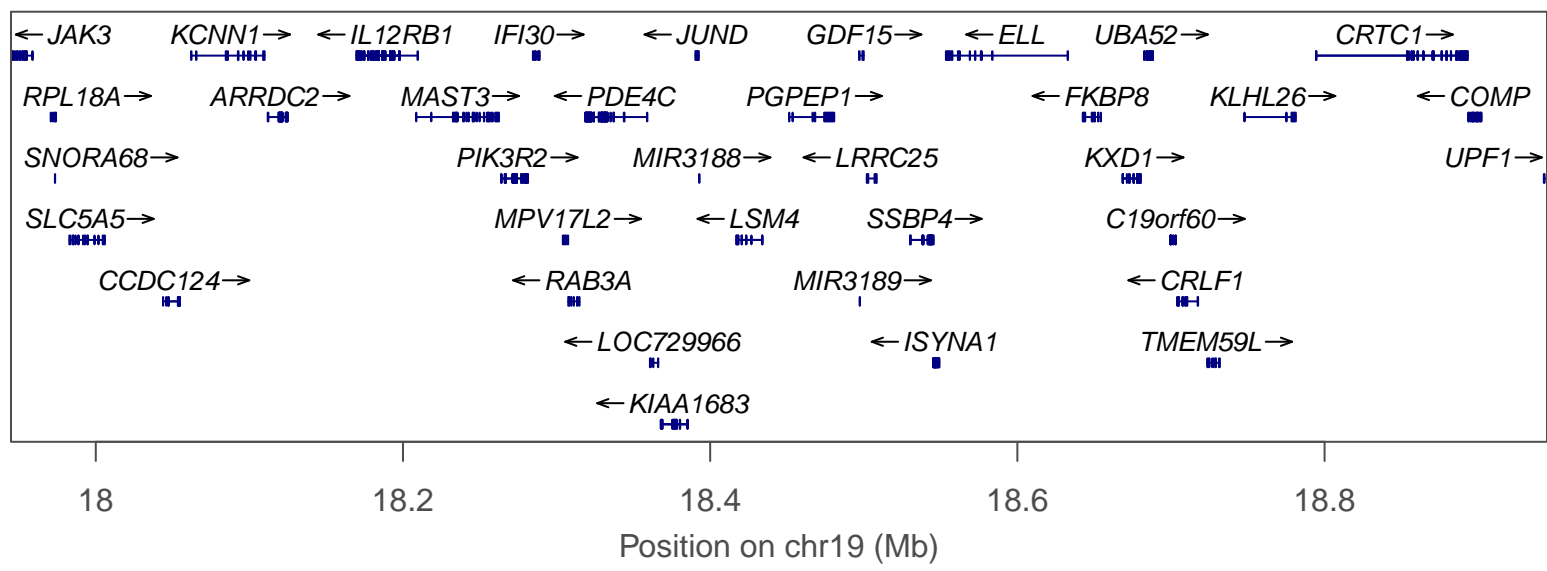
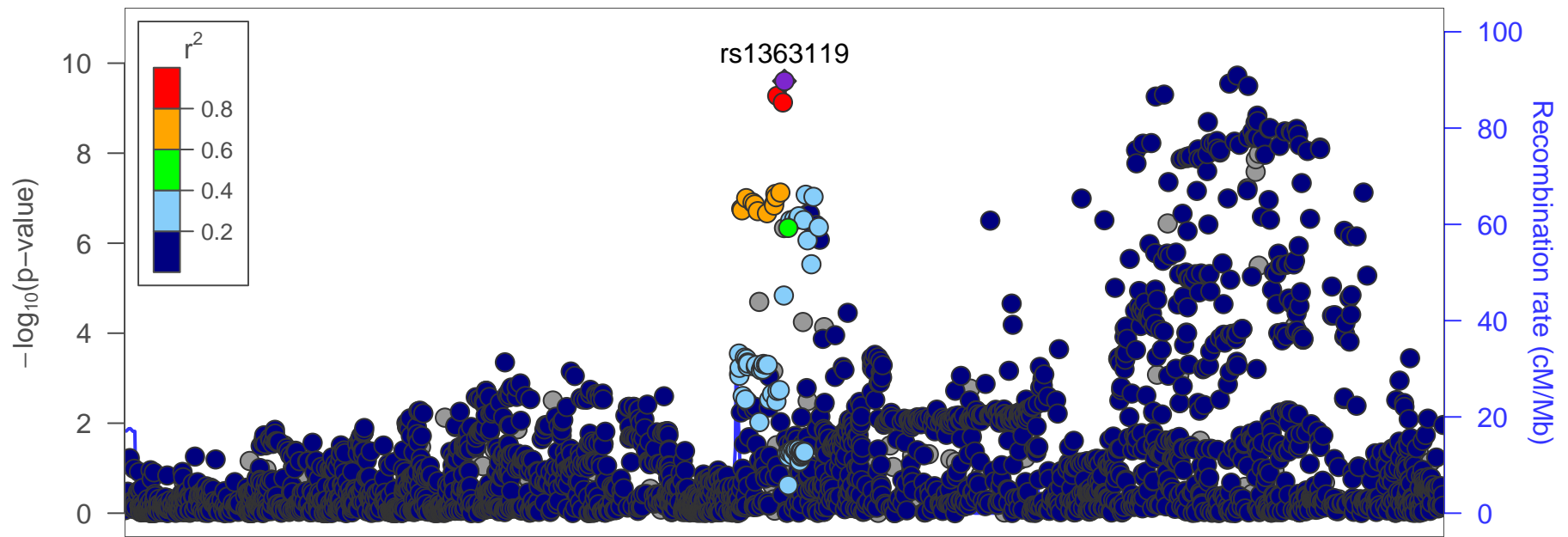




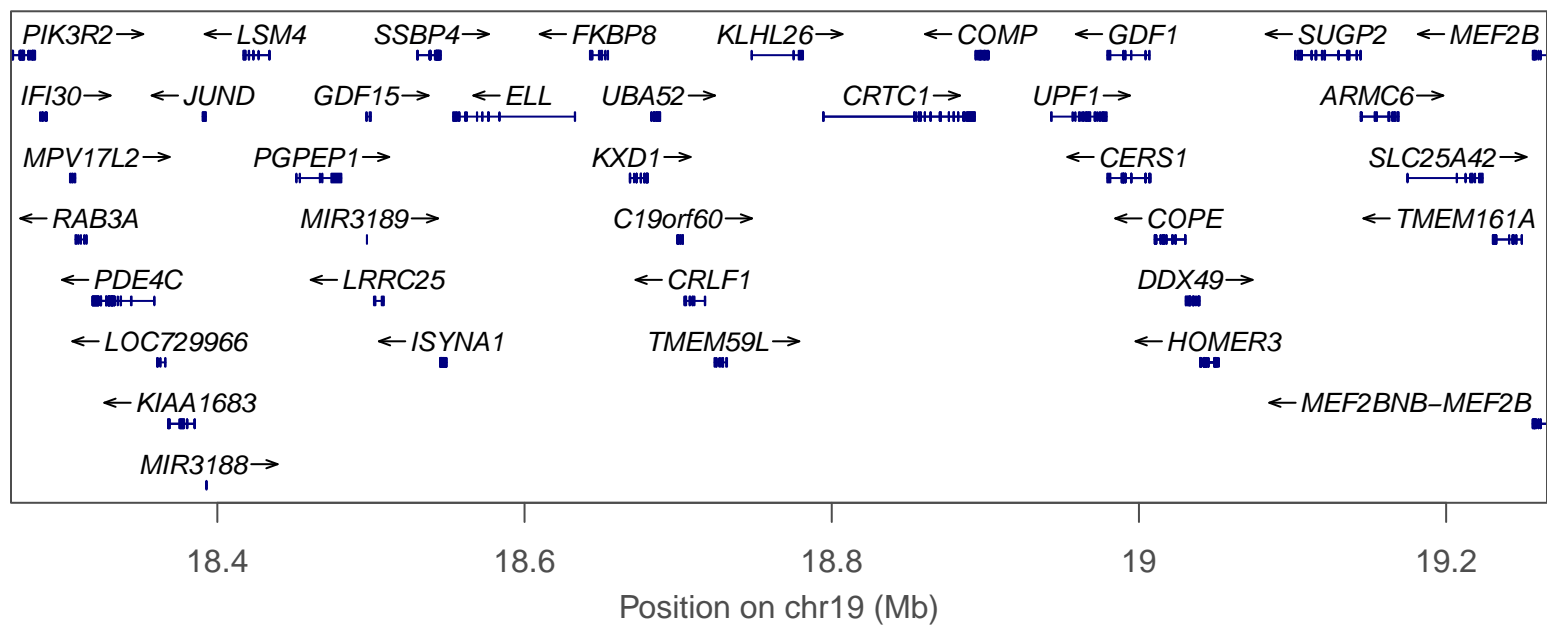
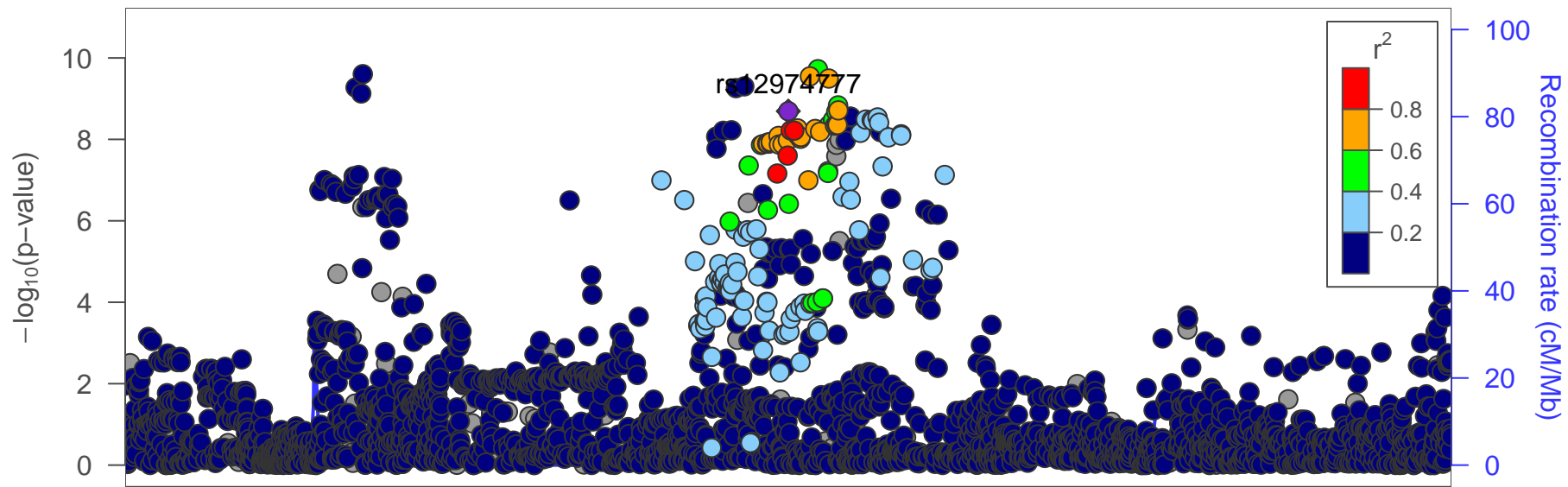
Plotted SNPs



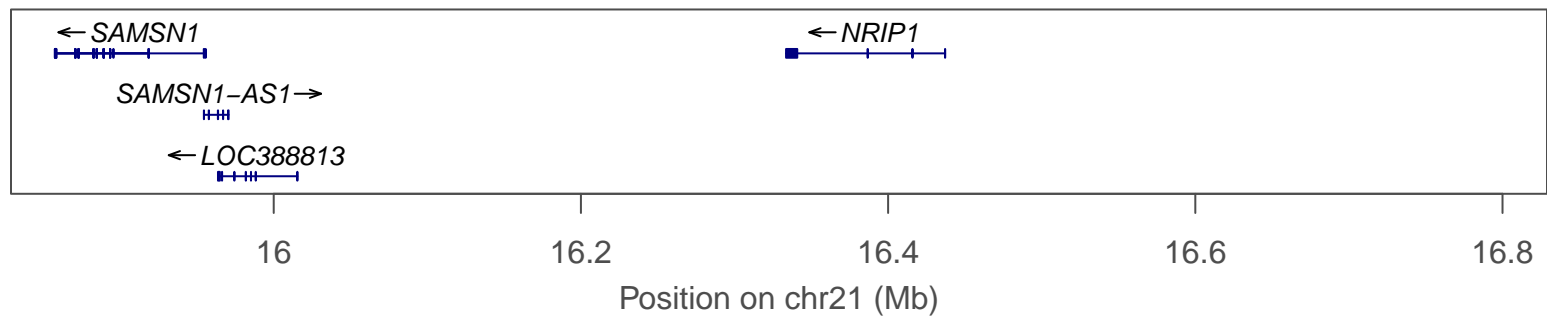
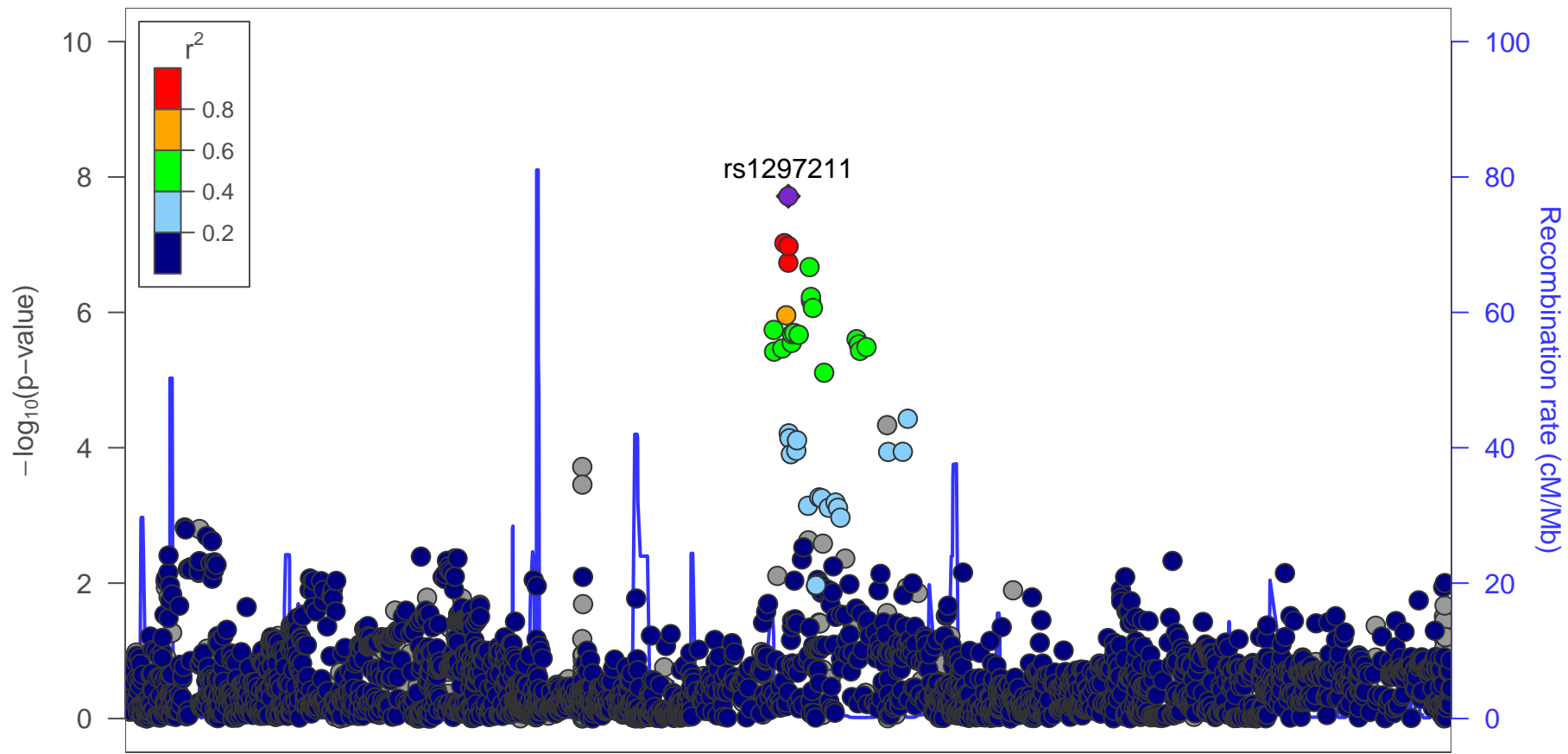
Plotted SNPs



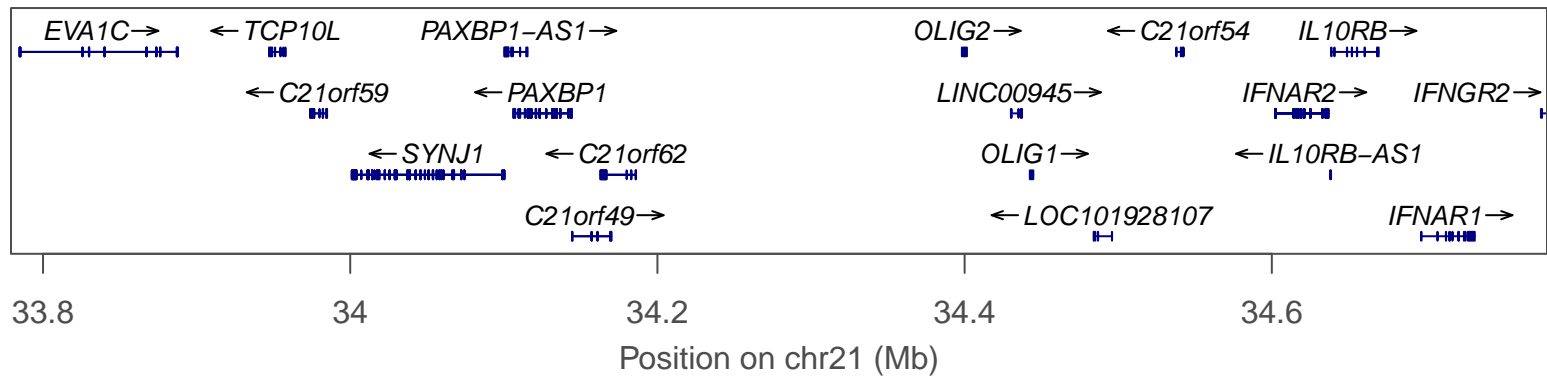
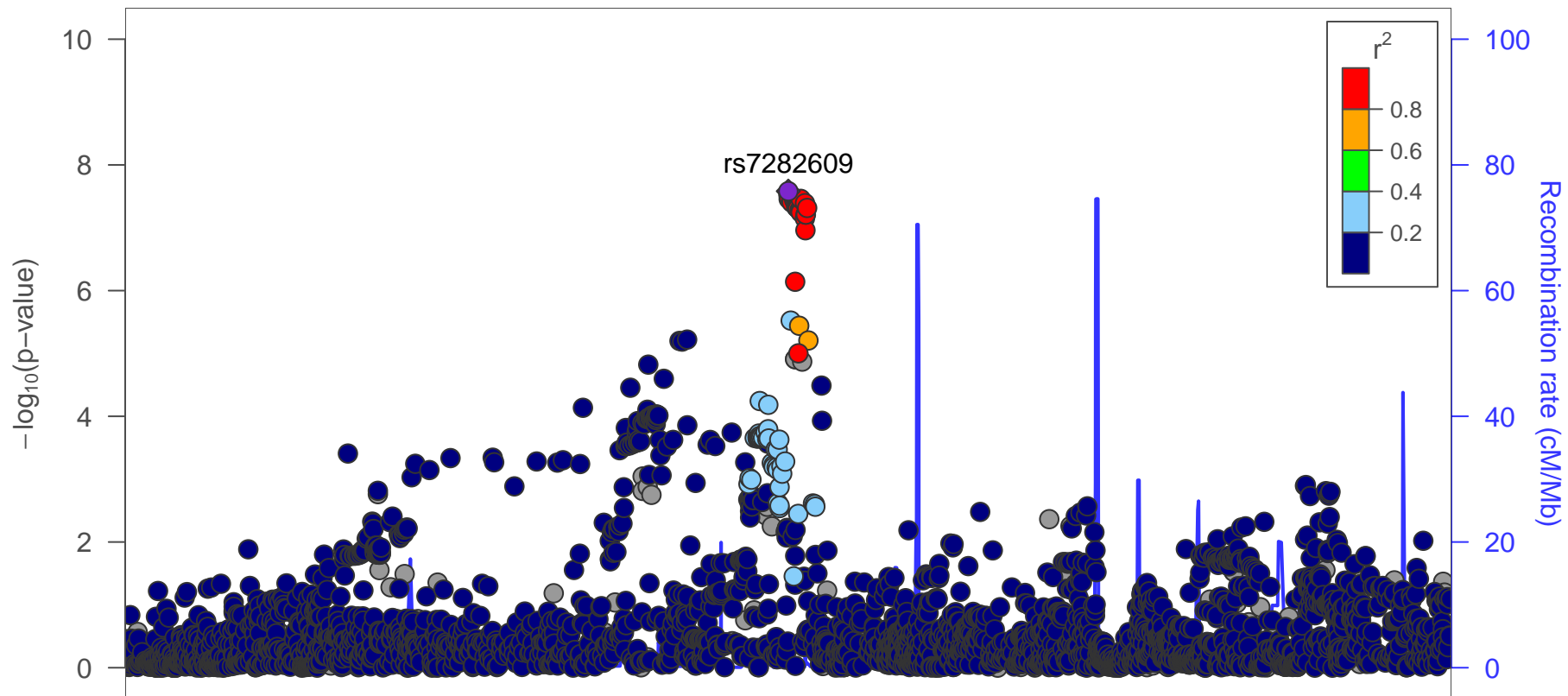
Plotted SNPs



Plotted SNPs



Plotted SNPs



Supplementary Figure 3. Locuszoom plots for each genome-wide significant locus in the GERD GWAS.

Supplimentary Table 1. GERD phenotype data extraction

source	Field ID	Description	Category	code	count*
ICD10	41202	Diagnoses - main ICD10	Summary Information	K21.0/K21.9	16,420
	41204	Diagnosis - secondary ICD10	Summary Information (diagnoses)		15,419
Self-report	20002	Non-cancer illness code, self-reported	Medical conditions	1138	23,774
ICD9	41203	Diagnoses - main ICD9	Summary Information	53010/53011	17
	41205	Diagnosis - secondary ICD9	Summary Information (diagnoses)		3
OPCS	41200	Operative procedures - main OPCS	Summary Information	G24/G25	1000
	41210	Operative procedures - secondary OPCS	Summary Information (operations)		205

*count: the occurrence of the corresponding code.

Supplementary Table 2. Phenotypic code exclude in control.

source	Exclude code	Data code
ICD10	K00-K93	19
Self-report	1134-1143	6
ICD9	5200-5799	87
OPCS	G-J	240

Supplementary Table 3. Extract GERD phenotype from Treatment/medication

Medications-verbal interview)	medicine	count
Reduce acid production	cimetidine	4
As above	tagamet HB	15
As above	pepcid AC	13
As above	nizatidine	103
As above	ranitidine	8,959
As above	ranitil-150mg	7
As above	Zantac	271
Block acid production and heal the esophagus	esomeprazole	1,763
As above	Nexium	882
As above	lansoprazole	19,353
As above	omeprazole	33,233
As above	pantoprazole	984
As above	gaviscon liquid	2,876

Antacids that neutralize stomach acid: Mylanta, Roloids and Tums (none of them can be found in UKB)

Supplementary Table 4. Number of individuals in UKBB with GERD phenotype

phenotype	case	control	total
union (meds+self_ICD10+etc.)	75,356	279,997	355,353
union (meds+self_ICD10+etc.)	75,356	279,997	355,353
self-ICD10	43,532	279,997	323,529
ICD10-only	27,163	279,997	307,160
self_only	21,981	279,997	301,978
meds_only	58,170	279,997	338,167
CD9_only	19	279,997	280,016
operative procedure	1,101	279,997	281,098
self_ICD10	43,659	279,997	323,656

Supplementary Table 5. BE(Barrett's esophagus) phenotype data extraction

source	Field ID	Description	Category	code	count*
ICD10	41202	Diagnoses - main ICD10	Summary Information (diagnoses)	K22.7	2,400
	41204	Diagnosis - secondary ICD10	Summary Information (diagnoses)		1,211

Supplimentary Table 6. Esophageal Adenocarcinoma phenotype data extraction

source	Field ID	Description	Category	code	count*
ICD10	40006	Type of Cancer: ICD10	Cancer register - Health-related outcomes	C15	1,536
Tumour	40011	Histology of cancer tumour	Cancer register - Health-related outcomes	8140	38360
				8141	28
				8144	271
				8147	14
				8160	301
				8190	2
				8210	306
				8211	1481
				8220	11
				8221	5
				8245	14
				8250	72
				8260	736
				8262	23
				8263	955
				8270	63
				8290	95
				8310	1301
				8313	3
				8323	32
				8330	281
				8331	17
				8332	7
				8333	2
				8400	10
				8401	38
				8408	2
				8410	18
				8480	1451
				8481	246
				8500	30746
				8503	138
				8570	21
8571	1				
8573	6				
9014	6				
9015	4				

items(8154;8161;8215;8251;8255;8322;8382;8283;8384;8413;8420;8525;8572;8576) have zero count

Supplementary Table 7. PBS code in QSKIN data

reduce acid production:	
Code & Prescriber	medical product
1158Y	cimetidine
2487X,2488Y	pepcid AC
1504E,1505F	nizatidine
1977C,1978D,1937Y,8162N	ranitidine
block acid production	
Code & Prescriber	medical product
3401B,8601Q,10295W,10343J,8600P,8886Q,10330Q,10331R	esomeprazole
9477T,9478W,8198L,2240X,2241Y,9331D	lansoprazole
1326T,1327W,8332M,8331L,8333N,9109K,9110L	omeprazole
8399C,9423Y,9424B,8007K,8008L	pantoprazole
8507R, 8508T, 8509W	RABEPRAZOLE

Supplementary Table 8. Number of controls in BE and EA cohorts

	UKBB	BEACON	Cambridge	Oxford	Bonn	total
BE	2,831	2,406	873	1,851	1,037	8,998
EA	568	1,508	995	N/A	1,609	4,680
#control	2,088	6,718	3,408	3,496	3,537	19,247

Supplementary Table 9. Top SNPs in BEEA in GERD GWAS after excluding BE/EA cases

SNP	Chr	Position	Tested allele	Other allele	Nearest gene or region	effect	se	pvalue
rs7255	2	20878820	T	C	GDF7 and LDAH	0.0236	0.0056	2.73E-05
rs1396065	2	200045039	T	C	SATB2	0.0154	0.0059	9.71E-03
rs2687202	3	70929983	T	C	FOXP1	0.0341	0.0061	2.61E-08
rs9918259	5	663092	T	C	TPPP and CEP72	-0.0311	0.007	9.80E-06
rs9257809	6	29356331	A	G	MHC region	0.0464	0.0085	5.00E-08
rs6242317	6	62195368	A	G	KHDRBS2 and MTRNR2L9	-0.0071	0.0076	3.50E-01
rs1745175	7	117256712	A	G	CFTR	0.0129	0.0082	1.17E-01
rs1774915	8	10068073	A	G	MSRA	-0.0023	0.0076	7.67E-01
rs1010851	8	11435516	T	C	LINC00208and BLK	0.0053	0.0056	3.42E-01
rs7852462	9	100310501	T	C	TMOD1	0.0066	0.0058	2.54E-01
rs1247942	12	114673723	C	G	LOC105369996and TBX5	0.0149	0.0057	8.75E-03
rs2464469	15	58362025	A	G	ALDH1A2	0.0034	0.0057	5.47E-01
rs1979654	16	86396835	C	G	LOC732275and FOXF1	0.0001	0.0057	9.98E-01
rs1996205	19	18804294	T	TG	CRTC1	0.0254	0.0063	5.19E-05