

Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci

International Genetics of Ankylosing Spondylitis Consortium (IGAS)

Supplementary Online Material:

Supplementary Tables 1-11

Supplementary Figures 1-12

Supplementary Note

Supplementary Table 1 Association *P*-values for the combined analysis after correcting for genomic inflation factor ($\lambda_{1000} = 1.0285$).

SNP	Locus	Position	Combined <i>P</i> -value	Combined <i>P</i> _{corrected} -value
rs6600247	1p36	25,177,701	2.6×10^{-15}	6.48×10^{-15}
rs11209026	1p31	67,478,546	2.0×10^{-27}	1.10×10^{-26}
rs4129267	1q21	152,692,888	3.4×10^{-13}	7.54×10^{-13}
rs1801274	1q23	159,746,369	1.4×10^{-09}	2.36×10^{-09}
rs41299637	1q32	199,144,473	1.9×10^{-15}	4.72×10^{-15}
rs6759298	2p15	62,421,949	4.9×10^{-47}	1.03×10^{-45}
rs12615545	2q31	181,756,697	1.0×10^{-09}	1.80×10^{-09}
rs4676410	2q37	241,212,412	9.9×10^{-09}	1.62×10^{-08}
rs12186979	5p13	40,560,617	4.3×10^{-06}	5.87×10^{-06}
rs30187	5q15	96,150,086	4.4×10^{-45}	8.07×10^{-44}
rs6871626	5q33	158,759,370	3.1×10^{-08}	4.95×10^{-08}
rs17765610	6q15	90,722,494	5.3×10^{-08}	8.30×10^{-08}
rs1128905	9q34	138,373,660	7.0×10^{-09}	1.15×10^{-08}
rs1250550	10q22	80,730,323	1.5×10^{-09}	2.53×10^{-09}
rs11190133	10q24	101,268,715	4.9×10^{-14}	1.12×10^{-13}
rs1860545	12p13	6,317,038	2.8×10^{-10}	2.53×10^{-09}
rs11065898	12q24	110,346,958	4.7×10^{-08}	7.37×10^{-08}
rs11624293	14q31	87,558,574	1.5×10^{-10}	2.75×10^{-10}
imm_16_28525386	16p11	28,525,386	2.6×10^{-09}	4.41×10^{-09}
rs2531875	17q11	23,172,294	1.2×10^{-10}	2.25×10^{-10}

rs9901869	17q21	42,930,205	6.0×10^{-15}	1.49×10^{-14}
rs35164067	19p13	10,386,181	3.4×10^{-10}	6.17×10^{-10}
rs2836883	21q22	39,388,614	6.5×10^{-17}	1.81×10^{-16}
rs7282490	21q22	44,440,169	6.2×10^{-09}	1.03×10^{-08}

Supplementary Table 2 Suggestive associations ($5 \times 10^{-08} < P\text{-value} < 5 \times 10^{-07}$) with ankylosing spondylitis susceptibility. Locus plots for reported associations are in Supplementary Figure 5.

SNP	Locus	Position ^a	Nearby gene(s)	Combined P-value	Risk/Non-risk allele	Combined OR	European			East Asian		
							Risk Allele Frequency (Case/Control)	OR	P-value	Risk Allele Frequency (Case/Control)	OR	P-value
rs12758027	1q32	208,861,431	<i>HHAT</i>	3.7×10^{-07}	C/T	1.09	0.479/0.453	1.09	6.03×10^{-06}	0.558/0.518	1.22	9.61×10^{-04}
rs4851529	2q11	102,013,732	<i>IL1R2-IL1R1</i>	1.0×10^{-07}	G/A	1.10	0.630/0.609	1.10	1.02×10^{-06}	0.497/0.472	1.09	0.087
rs13093489	3p24	27,769,911	<i>EOMES</i>	8.8×10^{-08}	A/C	1.10	0.361/0.334	1.12	1.05×10^{-07}	0.441/0.424	1.07	0.262
rs11742270	5p13	35,917,200	<i>IL7R</i>	5.4×10^{-07}	G/A	1.11	0.750/0.728	1.11	1.89×10^{-06}	0.855/0.837	1.12	0.063
rs2402752	7q31	124,226,835	<i>GPR37</i>	4.8×10^{-07}	T/C	1.10	0.727/0.706	1.11	2.44×10^{-06}	0.736/0.724	1.08	0.161
rs2283790	22q11	20,286,653	<i>UBE2L3</i>	7.9×10^{-08}	G/A	1.12	0.216/0.196	1.12	1.15×10^{-06}	0.468/0.439	1.15	0.029

a) NCBI human genome build 36 coordinates.

Supplementary Table 3 TaqMan genotyping of 2,998 east Asian cases and 5,547 east Asian controls of the non-MHC SNPs reported to be associated to AS susceptibility in the Han Chinese populations (8).

SNP	Call Rate	Controls HWE test <i>P</i> -value	Genotypes		Association <i>P</i> -value
			Controls	Cases	
rs4552569	0.974	0.412	975/933/210	825/704/158	0.212
rs13210693	0.967	0.018	540/1071/491	479/806/406	0.098
rs17905830	0.978	0.243	1630/462/21	1336/359/28	0.178

Supplementary Table 4 HLA-B*27 interactions. Epistasis (interaction term significant at P -value $< 5 \times 10^{-5}$) was assessed via logistic regression as described in the methods. The term β_1 corresponds to the SNP effect and β_{12} to the interaction effect. See Supplementary Figure 4 for experiment wide HLA-B*27 correlations in controls and in cases and Q-Q plot of the interaction term for non-MHC SNPs.

SNP	LD (r) with HLA-B*27		β_1 (se)	P -value	β_{12} (se)	P -value
	Controls	Cases				
rs30187	-0.0052	0.0543	-0.0016 (0.0477)	0.9727	0.3897 (0.0586)	2.94×10^{-11}
rs10045403	0.0115	-0.0482	-0.0591 (0.0519)	0.2545	-0.2819 (0.0636)	9.29×10^{-6}

Supplementary Table 5 Lymphocyte counts and AS susceptibility. SNPs associated (P -value $< 5 \times 10^{-3}$) with at least one lymphocyte trait (CD4 T cells, CD8 T cells or the ratio CD4:CD8) and AS susceptibility are listed. Linkage disequilibrium between SNPs was computed using reference data from 1000 Genomes Project and paired SNPs had $D' > 0.6$. Coupled alleles correspond to the positively correlated alleles between two paired SNPs.

Locus	Gene	SNP	Allele	CD4		CD8		CD4:CD8		AS associated SNP	Risk/Non-risk Allele	LD (R ² /D')	Coupling
				Beta	P-value	Beta	P-value	Beta	P-value				
chr1p36	<i>RUNX3</i>	rs876109	C (T)	-0.022	0.5	-0.125	9.10×10^{-5}	0.118	2.50×10^{-4}	rs6600247	C/T	0.24/0.62	C,C
chr5p13	<i>IL7R</i>	rs991570	A (G)	0.056	0.091	-0.033	0.0027	-0.048	0.143	rs11742270	G/A	0.25/0.93	A,G
chr5p13	<i>PTGER4</i>	rs16869602	T (C)	-0.054	0.43	-0.235	5.6×10^{-4}	0.204	0.0029	rs62358232	C/G	0.01/0.70	T,C
chr5q15	<i>ERAP1</i>	rs2042381	G (A)	-0.059	0.088	0.065	0.055	-0.125	2.70×10^{-4}	rs30187	T/C	0.70/0.98	A,T
chr6q15	<i>BACH2</i>	rs2026637	T (G)	-0.303	4.60×10^{-4}	-0.153	0.075	-0.105	0.22	rs17765610	G/A	0.01/1.00	T,G
chr10q22	<i>ZMIZ1</i>	rs1782648	G (A)	-0.033	0.31	0.032	4.2×10^{-3}	0.074	0.021	rs1250550	G/T	0.93/0.98	G,G
chr12q24	<i>SH2B3</i>	rs11066054	T (G)	-0.171	3.80×10^{-5}	0.002	0.96	-0.162	9.40×10^{-5}	rs11065898	T/C	0.60/0.90	T,T
chr19p13	<i>TYK2</i>	rs10416653	C (T)	0.22	0.002	0.135	0.056	0.053	0.45	rs35164067	G/A	0.01/1.00	T,G

Supplementary Table 6 HLA-B*27 tagging performance by previously reported tag SNPs and rs116488202.

Supplementary Table 6a Validation datasets used to assess HLA-B*27 performance. Samples of British, Australia and China with PCR based HLA-B*27 typing.

	British & Australian		China	
	Cases	Controls	Cases	Controls
Number	542	754	104	5
HLA-B*27 positive	489 (90.2%)	67 (8.9%)	89 (85.6%)	0 (0.0%)

Supplementary Table 6b HLA-B*27 tagging performance. r^2 , correlation based on genotype allele counts.

SNP	Position ^a	Alleles (Minor/Major) ^b	British & Australian			China			
			r^2	Sensitivity	Specificity	MAF ^c	r^2	Sensitivity	Specificity
rs13202464	31,344,583	G/A	0.826	0.989	0.934	0.096	0.821	0.989	0.917
rs116488202	31,344,916	T/C	0.951	0.989	0.985	0.011	0.913	0.989	1.000
rs4349859	31,365,787	A/G	0.930	0.976	0.988	0.002	0.033	0.234	1.000

a) Chromosome 6 NCBI human genome build 37 coordinates.

b) Minor allele tags the HLA-B*27 allele.

c) Minor allele frequency from 1000 Genomes Project Phase I haplotypes (2010-11 data freeze) in European reference panel.

d) Minor allele frequency from 1000 Genomes Project Phase I haplotypes (2010-11 data freeze) in Asian reference panel.

Supplementary Table 6c HLA-B*27 tag SNP allele counts in cases and controls of European ancestry. Odds ratio of disease: $OR_{CT/CC} = 62$, $OR_{TT/CC} = 105$ and $OR_{TT/CT} = 1.7$.

rs116488202			
	CC	CT	TT
Cases	1,254	7,186	351
Controls	11,999	1,117	32

Supplementary Table 7 HLA-A*02 tagging performance and AS susceptibility association results in samples of European ancestry.

Supplementary Table 7a Correlation between rs2394250 and classical MHC alleles in 1,223 1958 Birth Cohort controls with PCR based genotypes in the *HLA-A* locus. MAF, minor allele frequency; PREV, allele prevalence; R2, Pearson correlation between *HLA-A* genotype and rs2394250 coded as 0, 1 or 2.

<i>HLA-A</i> allele	MAF	PREV	R2	Sensitivity	Specificity
2	0.283	0.482	0.411	0.669	0.934
1	0.182	0.321	0.179	0.224	0.462
3	0.130	0.245	0.112	0.173	0.594
11	0.061	0.117	0.063	0.169	1.000
32	0.042	0.083	0.055	0.118	0.997
24	0.070	0.135	0.042	0.102	0.792

Supplementary Table 7b Association results for HLA-A*0201 tag SNP rs2394250 in samples of European ancestry.

	All samples				HLA-B*27 negative samples ^a			
	MAF Controls	MAF Cases	OR	<i>P</i> -value	Controls	Cases	OR	<i>P</i> -value
rs2394250	0.421	0.53	1.541	8.00×10^{-51}	0.300	0.355	1.361	1.75×10^{-11}
<i>conditioning on rs116488202</i>			1.214	6.48×10^{-12}				

a) As tagged by rs116488202

Supplementary Table 8 Overlap with other immune mediated diseases or immune traits. Linkage disequilibrium between the AS SNP and the disease SNP was computed from the European reference panel from the 1000 Genomes Project (see Online Methods). Coupling refers to positively correlated alleles.

Locus	AS SNP	AS Risk/Non-risk Alleles	Coupling	Disease SNP	Disease Risk/Non-risk Alleles	Disease	Disease OR	LD (r ² /D')	PUBMED ID	Direction
<i>RUNX3</i>	rs6600247	C/T	C,G	rs10903122	G/A	Celiac disease	1.12	0.99/1.00	20190752	Concordant
<i>IL23R</i>	rs11209026	G/A	A,A	rs11209026	G/A	Ulcerative colitis	1.74	1.00/1.00	21297633	Concordant
			A,A	rs11209026	G/A	Crohn's disease	2.66	1.00/1.00	21102463	Concordant
<i>IL23R-IL12RB2</i>	rs12141575 ^b	T/C	A,A	rs11209026	G/A	Psoriasis	1.49	1.00/1.00	20953190	Concordant
			A,C	rs1495965	G/A	Behcet's disease	1.35	0.51/0.99	20622879	Concordant
<i>GPR25-KIF21B</i>	rs41299637	T/G	T,C	rs7554511	C/A	Ulcerative colitis	1.19	1.00/1.00	21297633	Concordant
			T,G	rs296547	G/A	Celiac disease	1.12	0.62/0.97	20190752	Concordant
			T,T	rs11584383	T/C	Crohn's disease	1.18	0.81/0.95	18587394	Concordant
<i>PTGER4</i>	rs12186979	C/T	A,T	rs11742570	C/T	Crohn's disease	1.33	0.43/0.12	21102463	Concordant
			A,G	rs6451493	T/G	Ulcerative colitis	1.08	0.43/0.12	21297633	Concordant
<i>ERAP1-ERAP2-LNPEP</i>	rs30187 rs2910686 ^b	T/C G/A	A,A	rs6896969	C/A	Multiple sclerosis	1.1	0.41/0.11	19525953	Concordant
			T,A	rs27524	A/G	Psoriasis	1.13	0.74/0.88	20953190	Concordant
<i>IL12B</i>	rs6871626 rs6556416 ^b	T/G G/T	T,T	rs2549794	C/T	Crohn's disease	1.05	0.91/1.00	21102463	Concordant
			A,A	rs6871626	A/C	Ulcerative colitis	1.17	1.00/1.00	21297633	Concordant
<i>CARD9</i>	rs1128905	C/T	A,A	rs10045431	C/A	Crohn's disease	1.11	0.89/1.00	18587394	Concordant
			T,G	rs10781499	A/G	Ulcerative colitis	1.12	0.55/0.85	21297633	Concordant
<i>LTBR-TNFRSF1A</i>	rs1860545	C/T	T,C	rs4077515	T/C	Crohn's disease	1.18	0.57/0.87	21102463	Concordant
			G,T	rs1800693	C/T	Multiple sclerosis	1.2	0.96/0.98	19525953	Discordant
<i>NPEPP5-TBKBP1-TBX21</i>	rs9901869	T/C	T,C	rs8070463	T/C	Multiple sclerosis	1.15	0.77/0.88	22190364	Discordant
			G,G	rs2836878	G/A	C-reactive protein	0.03 ^c	0.99/1.00	21300955	Concordant ^a
<i>IL6R</i>	rs4129267	G/A	G,G	rs2836878	G/A	Ulcerative colitis	1.25	0.99/1.00	21297633	Concordant
			C,C	rs4537545	T/C	C-reactive protein	-0.22 ^d	0.96/1.00	19567438	Concordant ^a

		G,G	rs4129267	T/C	Asthma	1.09	1.00/1.00	21907864	Discordant
<i>FCGR2A</i>	rs1801274	T,T	rs1801274	T/C	Ulcerative colitis	1.21	1.00/1.00	21297633	Concordant
<i>UBE2E3</i>	rs12615545	C,G	rs13010713	G/A	Celiac disease	1.13	0.81/0.91	20190752	Concordant
<i>GPR35</i>	rs4676410	G,G	rs4676406	T/G	Ulcerative colitis	1.14	0.16/0.85	21297633	Concordant
<i>NKX2-3</i>	rs11190133	C,T	rs4409764	T/G	Crohn's disease	1.22	0.32/0.90	21102463	Concordant
		C,T	rs6584283	T/C	Ulcerative colitis	1.21	0.32/0.90	21297633	Concordant
<i>ZMIZ1</i>	rs1250550	C,G	rs1250542	A/G	Multiple sclerosis	1.15	0.55/0.77	22190364	Discordant
		C,A	rs1250552	A/G	Celiac disease	1.12	0.40/0.83	20190752	Concordant
		C,C	rs1250550	G/T	Crohn's disease	1.19	1.00/1.00	21102463	Concordant
<i>SH2B3</i>	rs11065898	C,T	rs3184504	T/C	Rheumatoid arthritis	1.08	0.22/1.00	20453842	Discordant
		C,C	rs653178	G/A	Type 1 diabetes	1.33	0.22/1.00	19430480	Discordant
<i>GPR65</i>	rs11624293	T,C	rs8005161	T/C	Celiac disease	1.2	0.22/0.99	20190752	Discordant
<i>IL27-SULT1A1</i>	imm_16_28525386	G,A	rs151181	G/A	Crohn's disease	1.23	1.00/1.00	21102463	Concordant
		G,G	rs4788084	G/A	Crohn's disease	1.07	0.84/0.83	21102463	Concordant
<i>TYK2</i>	rs35164067	G,C	rs2304256	C/A	Type 1 diabetes	1.09	0.87/0.85	19430480	Discordant
<i>ICOSLG</i>	rs7282490	A,A	rs2838519	G/A	Type 1 diabetes	1.16	0.52/0.96	19966805	Concordant
<i>EOMES</i>	rs13093489	C,C	rs11129295	A/G	Ulcerative colitis	1.14	0.98/0.99	21297633	Concordant
<i>IL7R</i>	rs11742270	A,T	rs6897932	C/T	Crohn's disease	1.13	0.94/1.00	18587394	Concordant
<i>UBE2L3</i>	rs2283790	A,G	rs3194051	G/A	Multiple sclerosis	1.11	0.90/0.98	21833088	Concordant
		A,T	rs6897932	C/T	Multiple sclerosis	1.12	1.00/1.00	19525953	Concordant
		A,T	rs6897932	C/T	Type 1 diabetes	1.12	1.00/1.00	17554260	Concordant
		A,A	rs3194051	G/A	Ulcerative colitis	1.07	0.16/1.00	21297633	Concordant
		A,G	rs181359	T/C	Crohn's disease	1.1	1.00/1.00	21102463	Concordant
		A,G	rs5754217	T/G	Systemic lupus erythematosus	1.38	1.00/1.00	21408207	Concordant
		A,C	rs2298428	T/C	Celiac disease and Rheumatoid arthritis	NR	0.88/1.00	21383967	Concordant
		A,C	rs2298428	A/G	Celiac disease	1.13	0.88/1.00	20190752	Concordant

a) Increased susceptibility is associated with increased levels of C-reactive protein.

- b) Reported as a secondary signal in AS.
- c) Effect size of 1 unit change in the natural log transformed CPR (mg/L) per copy increment in the coded allele.
- d) Estimated absolute CRP (mg/L) effect size per copy of minor allele under an additive model adjusted for age, sex, and ethnicity.

Supplementary Table 9 Regions of association for the primary signal of association. For each locus five windows are reported: a 2 log drop in the p-value of association from the WTCCC2/TASC meta analysis⁶, the ImmunoChip European plus Asian analysis, and the European analysis, respectively; a window defined based on LD patterns ($r^2 > 0.9$) from ImmunoChip controls and from 1000 Genomes project European reference dataset, respectively.

LOCUS	Gene	Boundary Definition Method	Dataset	Lead SNP position	Region	Size
1p36	<i>RUNX3</i>	2 log(P-value) drop	WTCCC2/TASC	25,169,771	25,163,597-25,170,995	7,398
		2 log(P-value) drop	ImmunoChip combined cohorts	25,177,701	25,166,528-25,177,847	11,319
		2 log(P-value) drop	ImmunoChip European cohort	25,177,701	25,166,528-25,178,260	11,732
		R2>0.9	ImmunoChip combined controls	25,177,701	25,162,321-25,177,701	15,380
1p31	<i>IL23R</i>	R2>0.9	1000G European controls	25,177,701	25,162,321-25,177,759	15,438
		2 log(P-value) drop	WTCCC2/TASC	67,497,708	67,448,104-67,517,189	69,085
		2 log(P-value) drop	ImmunoChip combined cohorts	67,478,546	67,469,657-67,498,692	29,035
		2 log(P-value) drop	ImmunoChip European cohort	67,478,546	67,470,457-67,503,956	33,499
1q32	<i>KIF21B</i>	R2>0.9	ImmunoChip combined controls	67,478,546	67,478,546-67,498,692	20,146
		R2>0.9	1000G European controls	67,478,546	67,472,503-67,516,140	43,637
		2 log(P-value) drop	WTCCC2/TASC	199,152,956	199,132,391-199,247,619	115,228
		2 log(P-value) drop	ImmunoChip combined cohorts	199,144,473	199,129,678-199,265,154	135,476
2p15	<i>intergenic</i>	2 log(P-value) drop	ImmunoChip European cohort	199,144,473	199,129,678-199,156,456	26,778
		R2>0.9	ImmunoChip combined controls	199,144,473	199,141,718-199,145,350	3,632
		R2>0.9	1000G European controls	199,144,473	199,140,852-199,169,230	28,378
		2 log(P-value) drop	WTCCC2/TASC	62,413,836	62,402,788-62,423,747	20,959
2q11	<i>IL1R2-IL1R1</i>	2 log(P-value) drop	ImmunoChip combined cohorts	62,421,949	62,404,976-62,423,747	18,771
		2 log(P-value) drop	ImmunoChip European cohort	62,421,949	62,404,976-62,423,747	18,771
		R2>0.9	ImmunoChip combined controls	62,421,949	62,421,725-62,421,949	224
		R2>0.9	1000G European controls	62,421,949	62,421,725-62,424,077	2,352
2q11	<i>IL1R2-IL1R1</i>	2 log(P-value) drop	WTCCC2/TASC	102,013,732	101,977,074-102,042,561	65,487
		2 log(P-value) drop	ImmunoChip combined cohorts	102,013,732	102,013,264-102,016,332	3,068
		2 log(P-value) drop	ImmunoChip European cohort	102,013,732	101,846,430-102,029,252	182,822

5q15	ERAP1	R2>0.9	Immunochip combined controls	102,013,732	102,013,732-102,028,045	14,313
		R2>0.9	1000G European controls	102,013,732	102,013,732-102,028,045	14,313
5q15	ERAP1	2 log(P-value) drop	WTCCC2/TASC	96,147,750	96,144,608-96,150,203	5,595
		2 log(P-value) drop	Immunochip combined cohorts	96,152,064	96,150,086-96,173,489	23,403
		2 log(P-value) drop	Immunochip European cohort	96,152,064	96,135,967-96,223,958	87,991
		R2>0.9	Immunochip combined controls	96,152,064	96,147,750-96,152,064	4,314
12p13	LTBR-TNFRSF1A	R2>0.9	1000G European controls	96,152,064	96,147,471-96152,064	4,593
		2 log(P-value) drop	WTCCC2/TASC	6,373,003	6,358,711-6,386,116	27,405
		2 log(P-value) drop	Immunochip combined cohorts	6,361,386	6,357,422-6,369,794	12,372
		2 log(P-value) drop	Immunochip European cohort	6,360,642	6,357,422-6,369,794	12,372
21q22	intergenic	R2>0.9	Immunochip combined controls	6,361,386	6,310,270-6,317,246	6,976
		R2>0.9	1000G European controls	6,361,386	6,310,270-6,317,251	6,981
		2 log(P-value) drop	WTCCC2/TASC	39,382,729	39,376,918-39,388,148	11,230
		2 log(P-value) drop	Immunochip combined cohorts	39,388,614	39,381,402-39,389,293	7,891
21q22	intergenic	2 log(P-value) drop	Immunochip European cohort	39,385,153	39,382,729-39,387,048	4,319
		R2>0.9	Immunochip combined controls	39,388,614	39,385,153-39,390,708	5,555
		R2>0.9	1000G European controls	39,388,614	39,385,153-39,389,513	4,360

Supplementary Table 10 Heritability explained by susceptibility loci under an additive model. Allele frequency and odds ratio are as reported for samples of European ancestry. Disease prevalence 0.5%.

Locus	Gene	SNP	Risk Allele Frequency (Case/Control)	OR	Heritability explained (%)
1p36	<i>RUNX3</i>	rs6600247	0.540/0.501	1.164	0.138
1p31	<i>IL23R</i>	rs11209026	0.959/0.934	1.650	0.352
1p31	<i>IL23R</i>	rs12141575	0.370/0.330	1.147	0.100
1q21	<i>IL6R</i>	rs4129267	0.635/0.592	1.176	0.151
1q23	<i>FCGR2A</i>	rs1801274	0.487/0.476	1.123	0.080
1q23	<i>FCGR2A</i>	rs2039415	0.702/0.682	1.088	0.037
1q32	<i>GPR25-KIF21B</i>	rs41299637	0.757/0.715	1.201	0.162
1q32	<i>HHAT</i>	rs12758027	0.479/0.453	1.094	0.048
2p15	intergenic	rs6759298	0.447/0.378	1.308	0.407
2q11	<i>IL1R2-IL1R1</i>	rs4851529	0.630/0.609	1.103	0.055
2q12	<i>IL1R2-IL1R1</i>	rs2192752	0.239/0.222	1.112	0.047
2q31	<i>UBE2E3</i>	rs12615545	0.451/0.421	1.109	0.062
2q37	<i>GPR35</i>	rs4676410	0.232/0.209	1.131	0.060
3p24	<i>EOMES</i>	rs13093489	0.361/0.334	1.119	0.067
5p13	<i>PTGER4</i>	rs12186979	0.516/0.498	1.093	0.047
5p13	<i>IL7R</i>	rs11742270	0.750/0.728	1.113	0.054
5q15	<i>ERAP1</i>	rs30187	0.405/0.338	1.318	0.411
5q15	<i>ERAP1</i>	rs1065407	0.703/0.645	1.171	0.136
5q15	<i>ERAP2</i>	rs2910686	0.450/0.440	1.171	0.147
5q33	<i>IL12B</i>	rs6871626	0.360/0.337	1.117	0.066
5q33	<i>IL12B</i>	rs6556416	0.704/0.675	1.107	0.054
6p21	<i>HLA-B*27</i>	rs116488202	0.450/0.045	60	20.089
6p21	<i>HLA-A*02</i>	rs2394250	0.530/0.421	1.214	0.220
6q15	<i>BACH2</i>	rs17765610	0.131/0.118	1.172	0.064
6q15	<i>BACH2</i>	rs639575	0.624/0.609	1.081	0.034
7q31	<i>GPR37</i>	rs2402752	0.727/0.706	1.108	0.052
9q34	<i>CARD9</i>	rs1128905	0.529/0.503	1.124	0.082
10q22	<i>ZMIZ1</i>	rs1250550	0.678/0.652	1.110	0.059
10q24	<i>NKX2-3</i>	rs11190133	0.737/0.707	1.181	0.136
12p13	<i>LTBR-TNFRSF1A</i>	rs1860545	0.634/0.605	1.131	0.086
12p13	<i>LTBR-TNFRSF1A</i>	rs7954567	0.363/0.341	1.113	0.062
12q24	<i>SH2B3</i>	rs11065898	0.237/0.216	1.129	0.060
14q31	<i>GPR65</i>	rs11624293	0.106/0.087	1.234	0.086
16p11	<i>IL27-SULT1A1</i>	imm_16_28525386	0.421/0.393	1.112	0.064
16p11	<i>IL27-SULT1A1</i>	rs35448675	0.007/0.006	1.236	0.007
17q11	<i>NOS2</i>	rs2531875	0.396/0.367	1.122	0.074
17q11	<i>NOS2</i>	rs2297518	0.212/0.190	1.129	0.055

17q21	<i>NPEPPS-TBKBP1-TBX21</i>	rs9901869	0.548/0.516	1.146	0.111
19p13	<i>TYK2</i>	rs35164067	0.819/0.796	1.155	0.080
19p13	<i>TYK2</i>	rs6511701	0.220/0.218	1.098	0.036
21q22	intergenic	rs2836883	0.768/0.734	1.190	0.140
21q22	<i>ICOSLG</i>	rs7282490	0.411/0.390	1.100	0.052
22q11	<i>UBE2L3</i>	rs2283790	0.216/0.196	1.124	0.052

Supplementary Table 11 Association with low frequency variants.

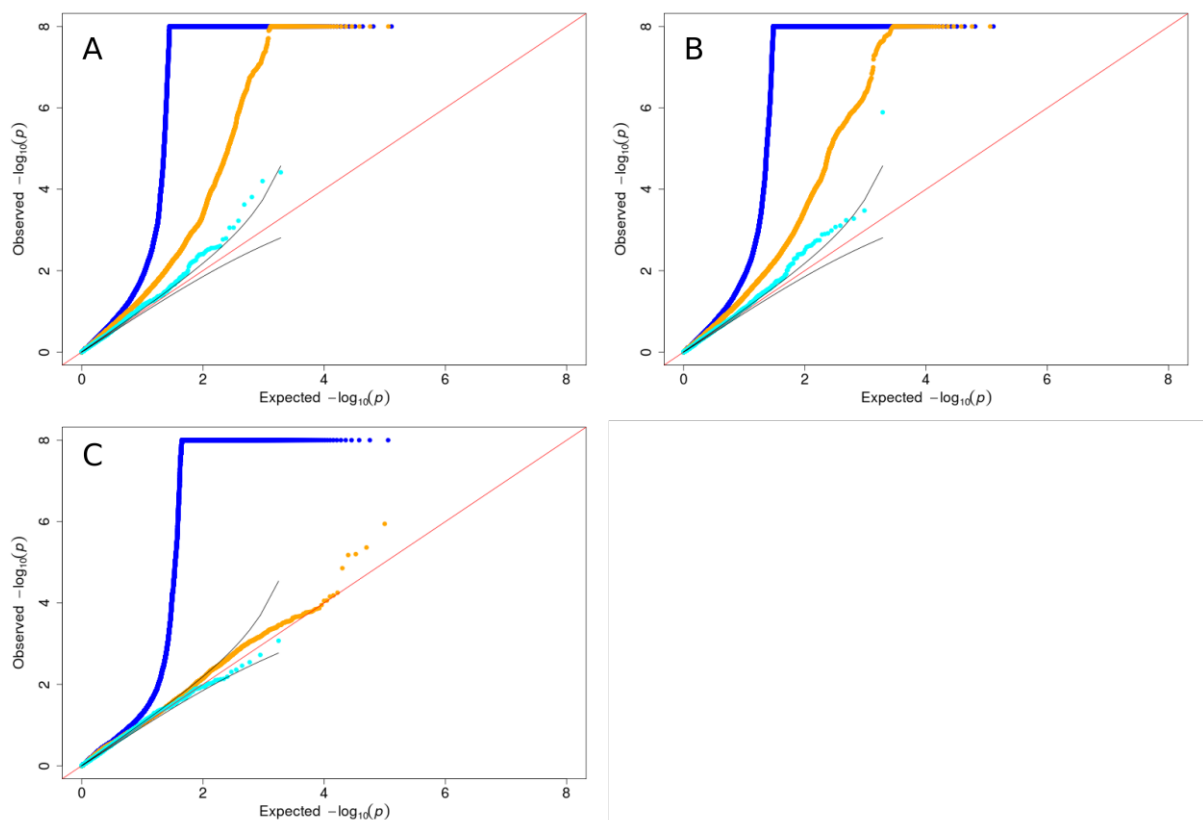
Supplementary Table 11a On the background of common variant association. Linkage disequilibrium (r^2 and D') is reported between the low frequency variants and the common variant associations in the locus reported on Table 1 and 3 used in the conditional analysis. Significance in the conditional analysis with logistic regression was also assessed by permutation, where phenotype labels were permuted 0.5M times. The reported permutation P -value is the proportion of times a test statistic was found to be greater than the observed.

SNP	CHR	Gene	Position	Change	MAF Controls	MAF Cases	Fisher's exact test P-value		Logistic regression P -value after conditioning on common variant associations	LD (r^2/D')	Permutation P -value	Locus-wide Permutation P -value
							OR	P -value				
vh_1_67407785	1	IL23R	67,407,785	N->K	0.0014	0.0028	2.20	1.47×10^{-3}	1.2×10^{-3}	0.00/0.74 (rs11209026) 0.00/0.08 (rs12141575)	2.50×10^{-3}	0.0526
rs76418789	1	IL23R	67,421,184	G->R	0.0058	0.0034	0.52	4.59×10^{-4}	1.9×10^{-3}	0.00/1.00 (rs11209026) 0.00/1.00 (rs12141575)	1.43×10^{-3}	0.0554
chr1_67407799	1	IL23R	67,407,799	R->Q	0.003	0.0017	0.66	9.08×10^{-2}	5.4×10^{-3}	0.00/0.92 (rs11209026) 0.00/0.90 (rs12141575)	2.49×10^{-2}	0.0579
rs61752351	5	LNPEP	96,341,076	I->M	0.019	0.023	1.23	4.40×10^{-3}	0.080	0.03/0.91 (rs30187) 0.01/0.85 (rs10045403) 0.02/0.99 (rs2910686)	3.87×10^{-3}	0.1284
rs1128905	9	CARD9	138,379,413	essential splice site	0.0053	0.0030	0.57	3.36×10^{-3}	9.2×10^{-4}	0.00/0.98 (rs1128905)	1.06×10^{-2}	0.2043
rs280518	19	TYK2	10,333,635	splice site	0.0001	0.0007	7.71	2.40×10^{-3}	0.019	0.00/1.00 (rs35164067) 0.00/0.34 (rs6511701)	7.37×10^{-3}	0.0332
rs35074907	19	KEAP1	10,461,418	snSNP	0.017	0.021	1.25	4.00×10^{-3}	2.96×10^{-3}	0.00/0.85 (rs35164067) 0.00/0.87 (rs6511701)	3.21×10^{-3}	0.0358

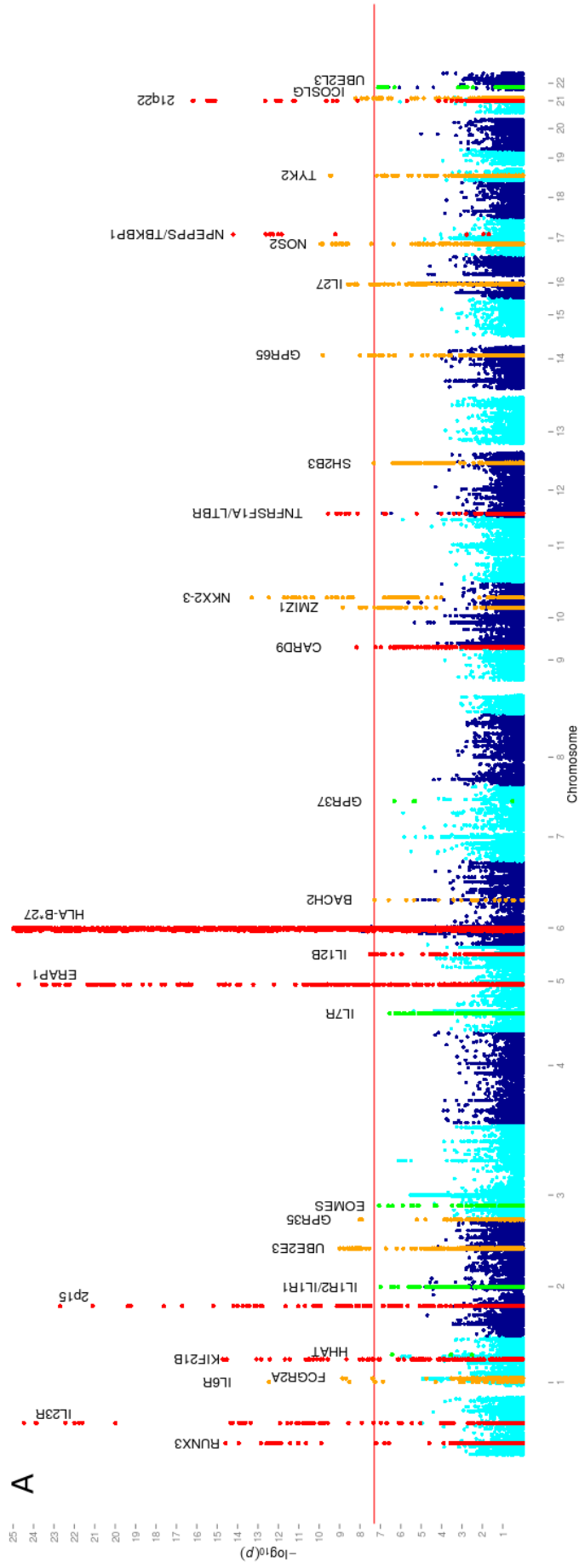
Supplementary Table 11b Low frequency exonic variants in the absence of common variant associations

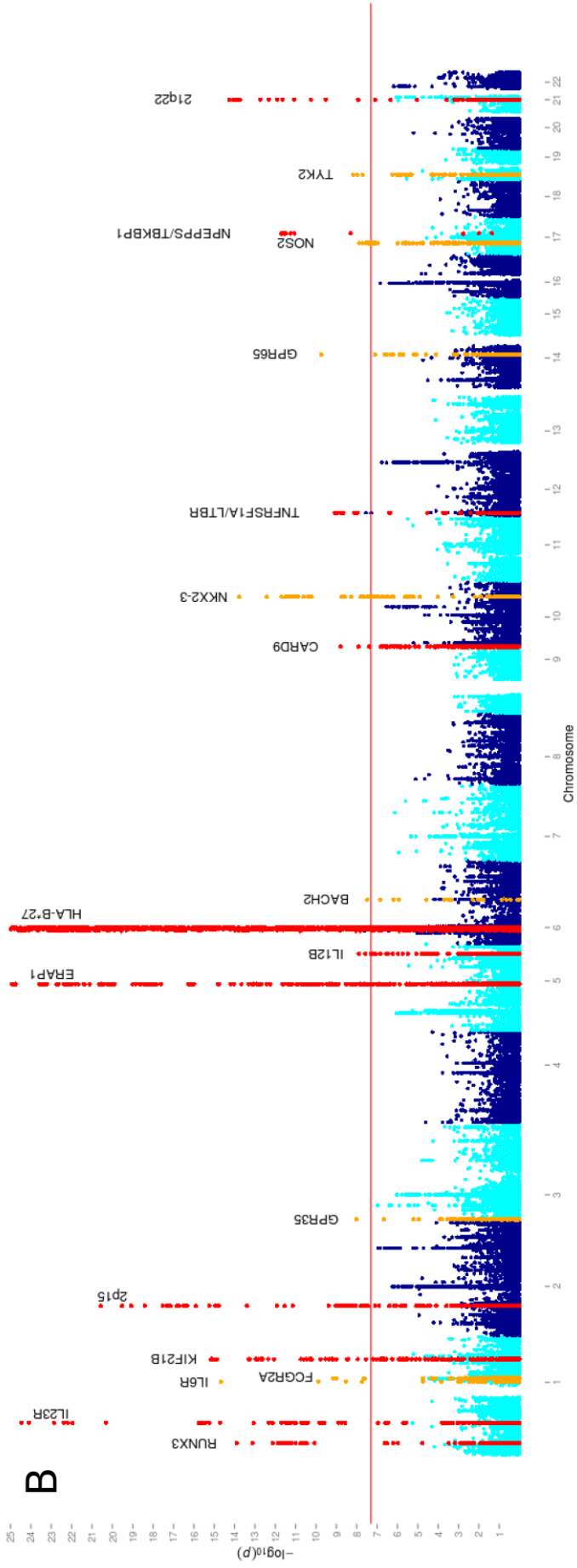
SNP	CHR	Position	gene	Type	A1	F_A	F_U	P -value	OR
rs61748604	4	187,396,399	KLKB1	splice site	B	0.016	0.0098	4.22×10^{-7}	1.61
rs114661226	5	132,006,650	RAD50	H->Y	A	0.0046	0.0022	7.20×10^{-5}	2.09
chr6_106660076	6	106,660,076	PRDM1	L->F	A	0.0028	0.0007	4.86×10^{-7}	4.06
rs35411407	12	4,579,169	DYRK4	synonymous	B	0.018	0.0244	7.41×10^{-6}	0.71
rs11848954	14	73,486,451	FAM161B	G->A	B	0.0055	0.0031	4.57×10^{-4}	1.78
rs12448775	16	31,292,149	ITGAX	synonymous	A	0.046	0.038	3.27×10^{-4}	1.21
imm_16_49259458	16	49,259,458	AC007608.2	splice site	A	0.042	0.035	1.82×10^{-4}	1.23

Supplementary Figure 1 Q-Q plots of association for the three analyses presented: A) all ethnicities, B) European ethnicity and C) east Asian ethnicity. Observed versus expected association for all SNPs tests are shown in blue, for all SNPs not on previously associated AS susceptibility loci are shown in orange and for a set of negative control SNPs in regions included on Immunochip for studies of reading and writing disabilities, psychosis and schizophrenia. 95% confidence intervals are shown for the expected association in the negative control SNP set. For each of the analyses, the negative control SNP lists contained SNPs that were unlinked ($r^2 < 0.3$) and common (MAF > 0.05) in the respective controls sample set. This resulted in a list of 1,922 SNPs for the all ethnicities analysis, 1,924 SNPs for the European ethnicity analysis and 1,767 SNPs for the east Asian ethnicity analysis. The genomic inflation factor in the negative control SNP list in the overall analysis was 1.047, in the white European analysis 1.110, and in the east Asian analysis 1.027.

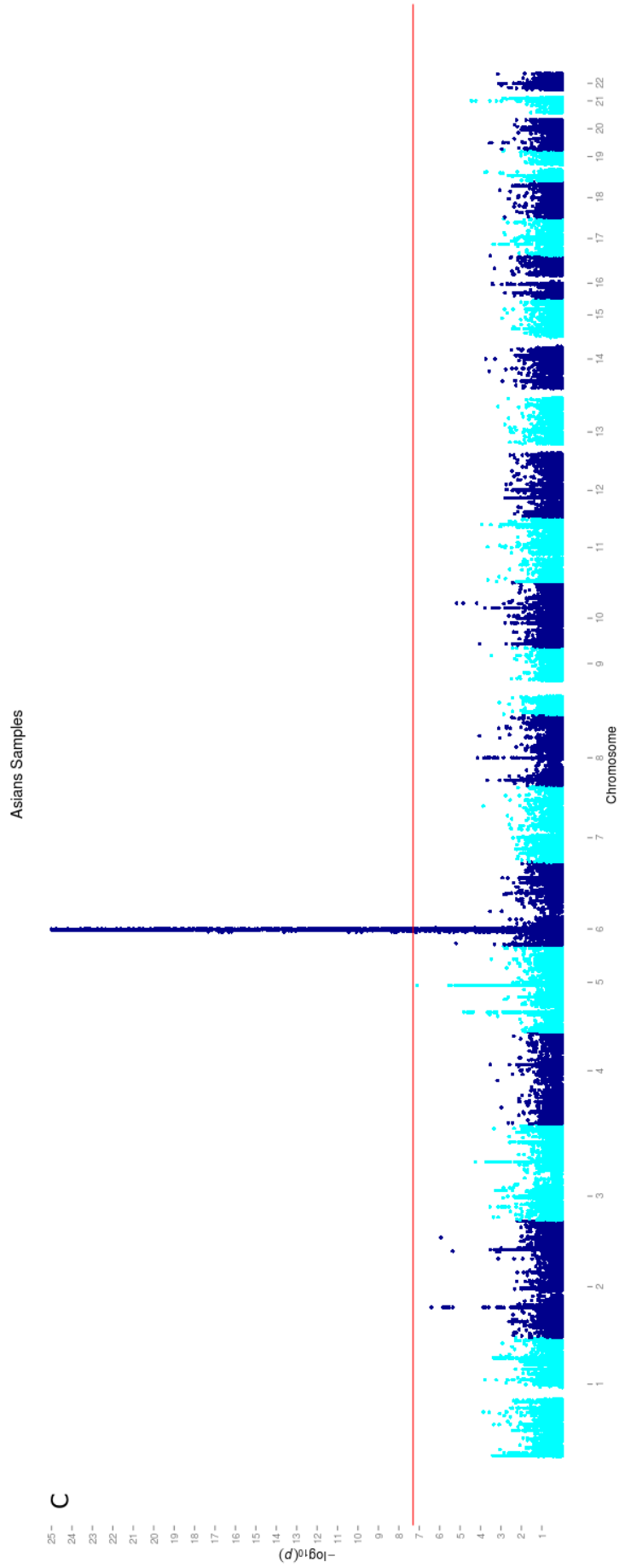


Supplementary Figure 2 Manhattan plot in European plus Asian cohort analysis (A), European cohort analysis (B) and Asian cohort analysis (C). Association peaks in European plus Asian cohort analysis are colour-coded in red if previously identified, orange if novel genome-wide significance loci and in green if loci was identified with suggestive association (P -value $< 5 \times 10^{-7}$).

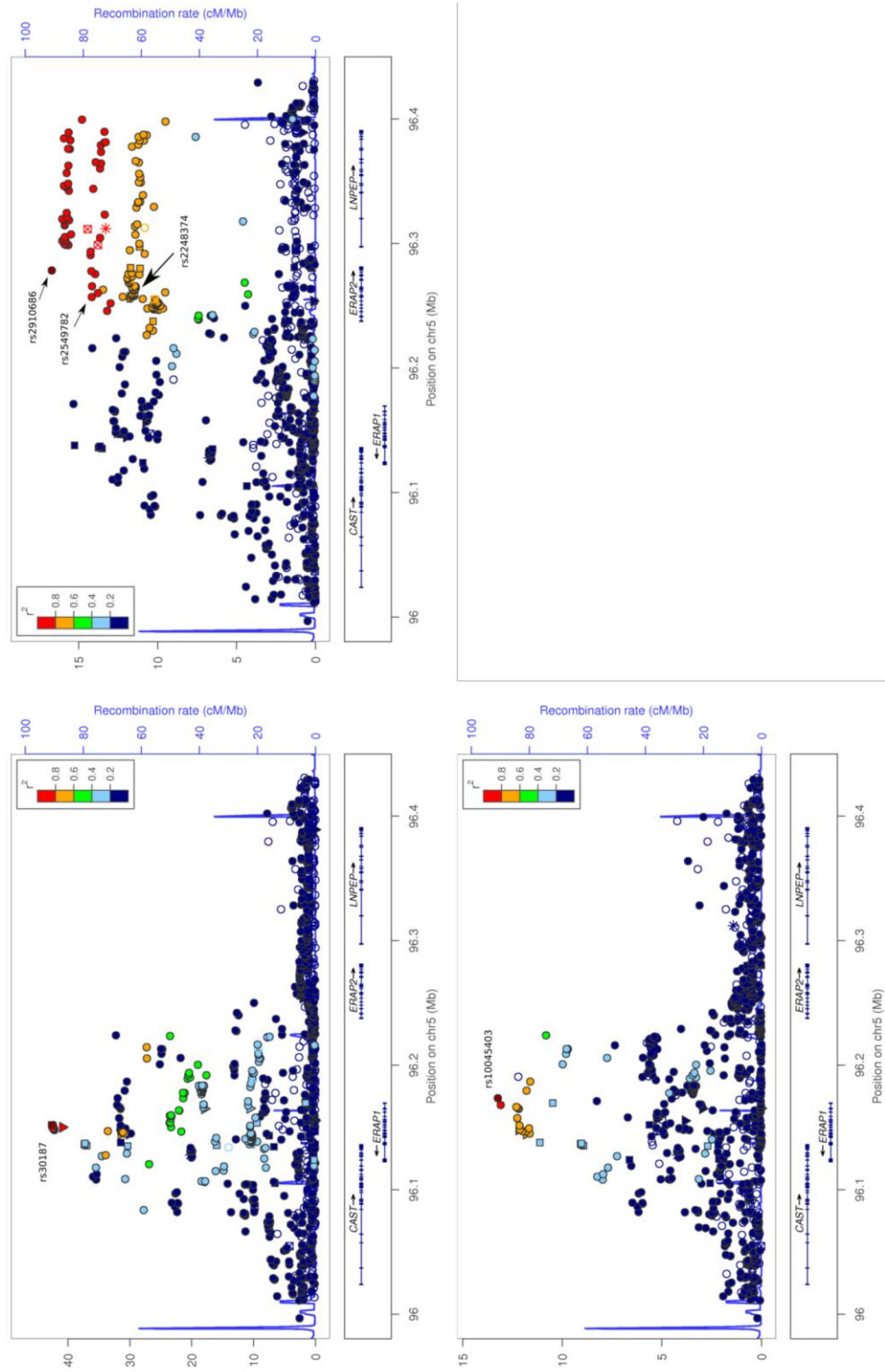




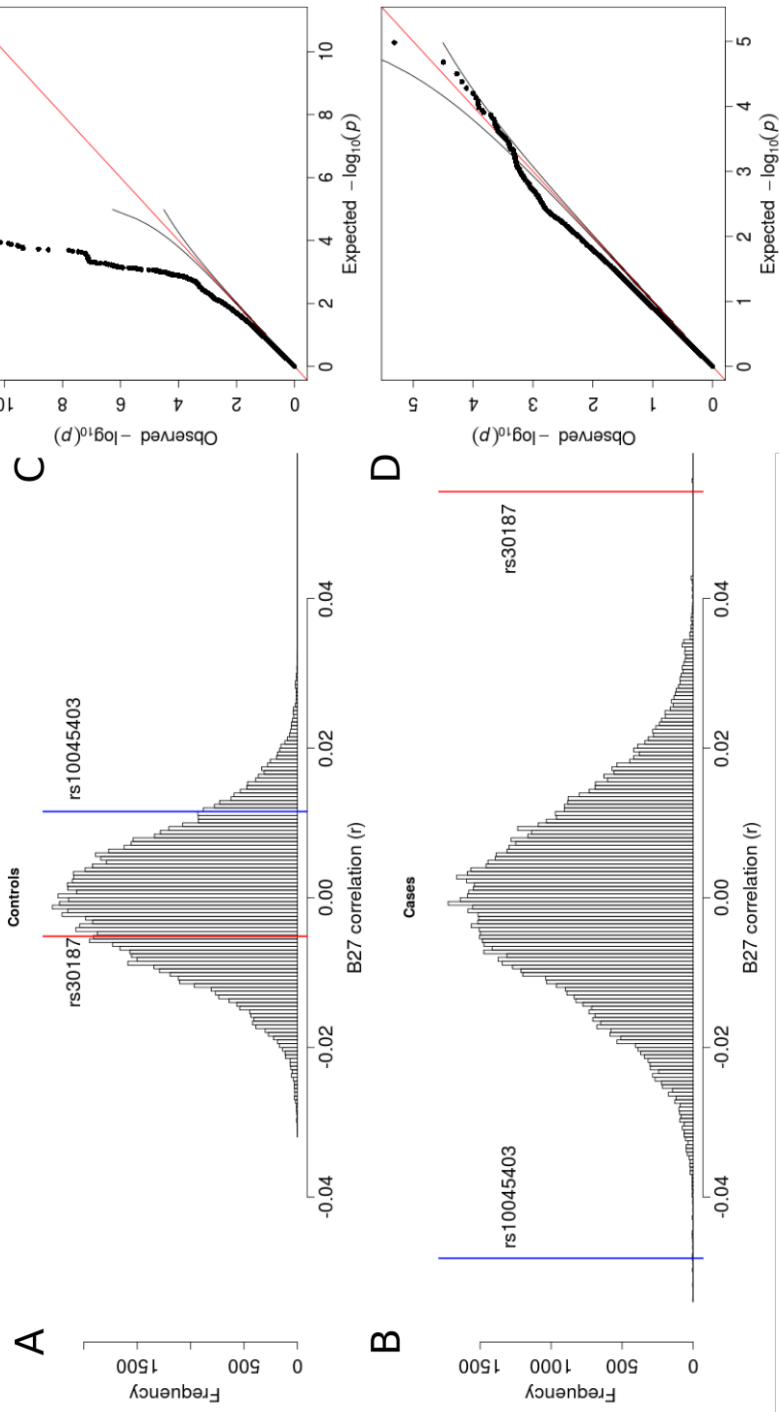
B



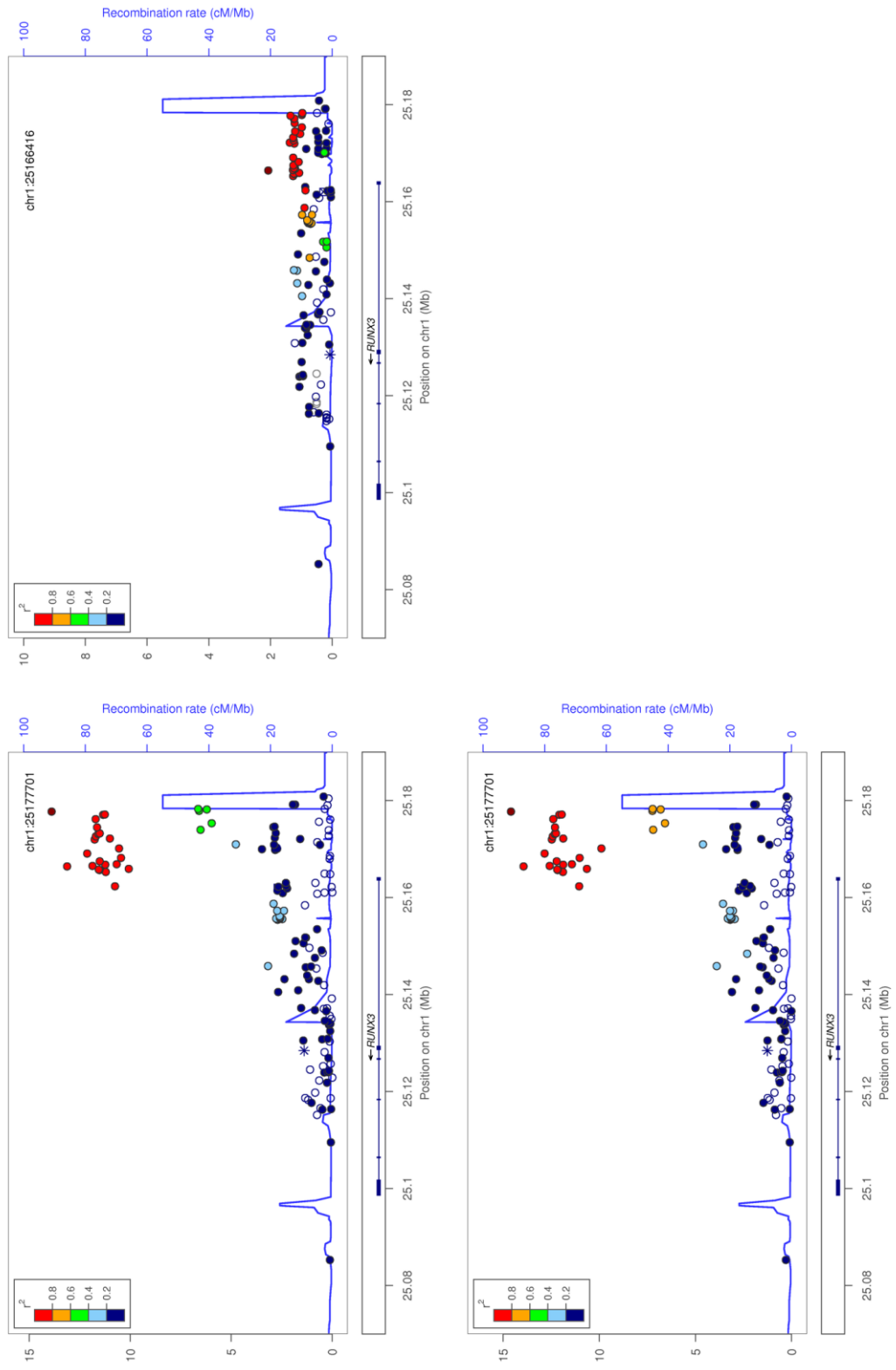
Supplementary Figure 3 European cohort AS susceptibility associations in the 5q15 locus containing the three aminopeptidases *ERAP1*, *ERAP2* and *LNPEP*.
 A) Unconditional association results tagged by the SNP rs30187. B) Associations over *ERAP2* and *LNPEP* after controlling for the SNP rs30187 association. C) Association results conditioning on rs30187 and rs2910686. Note the second signal over *ERAP1* remains significant (tagged by rs10045403) and no residual signal over *ERAP2/LNPEP* is seen.

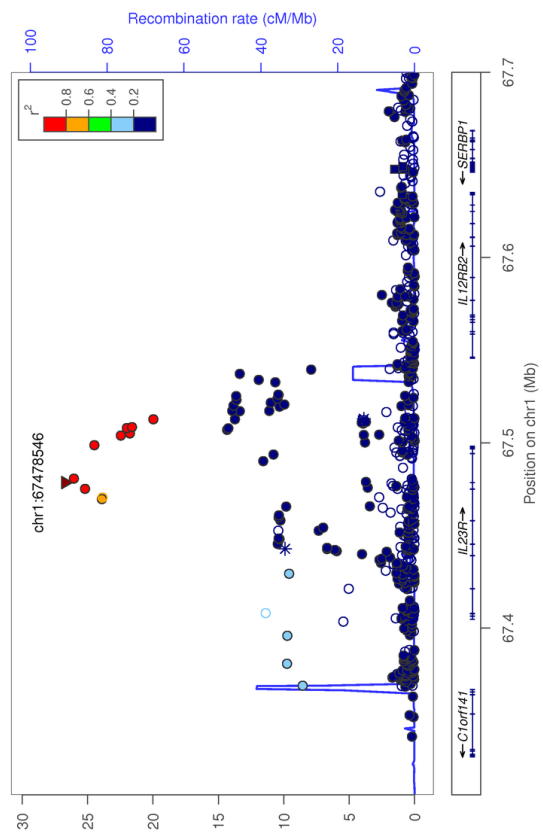
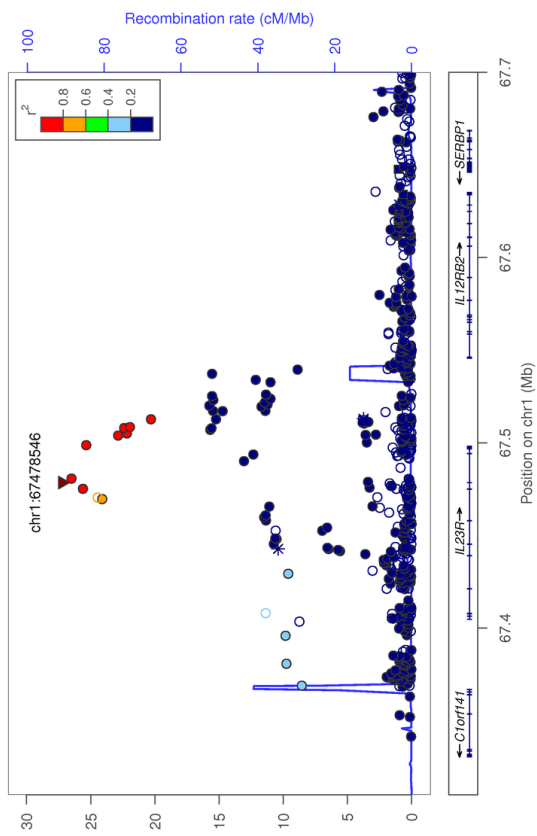
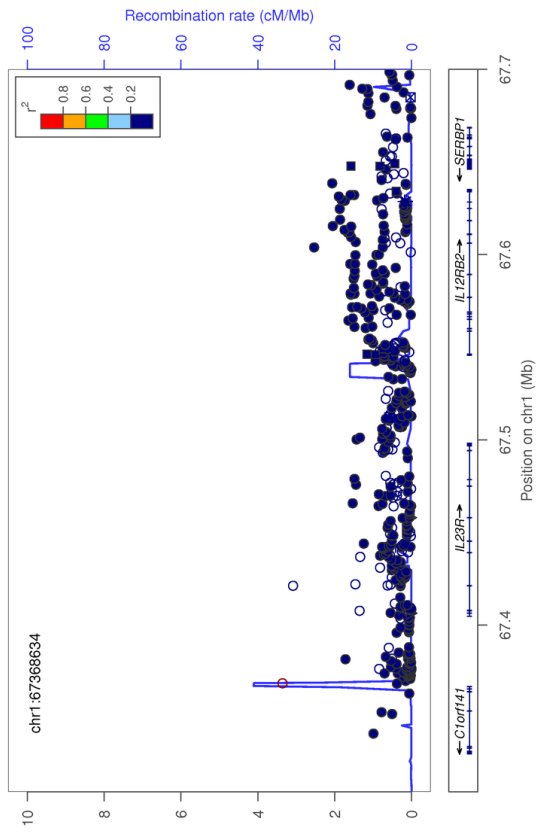


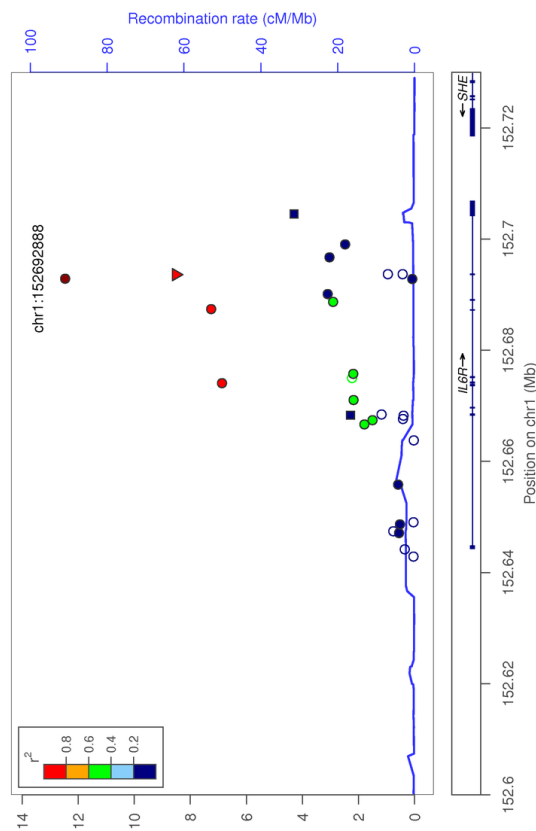
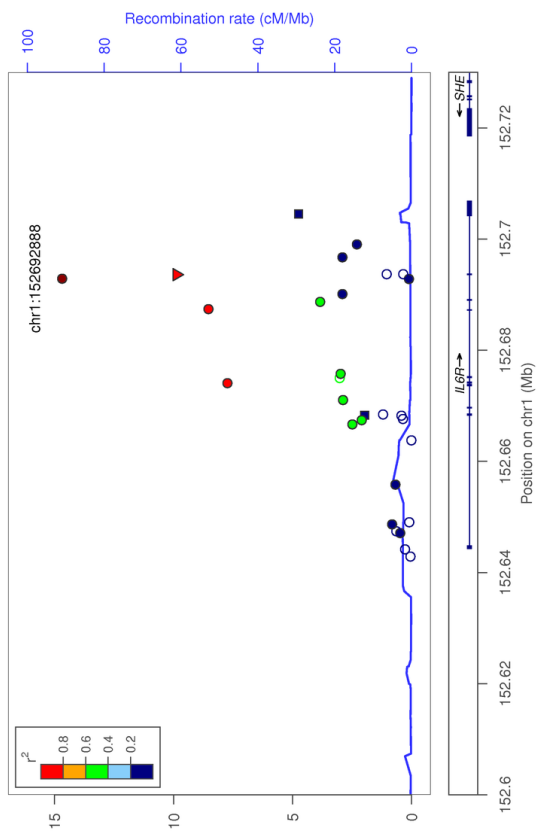
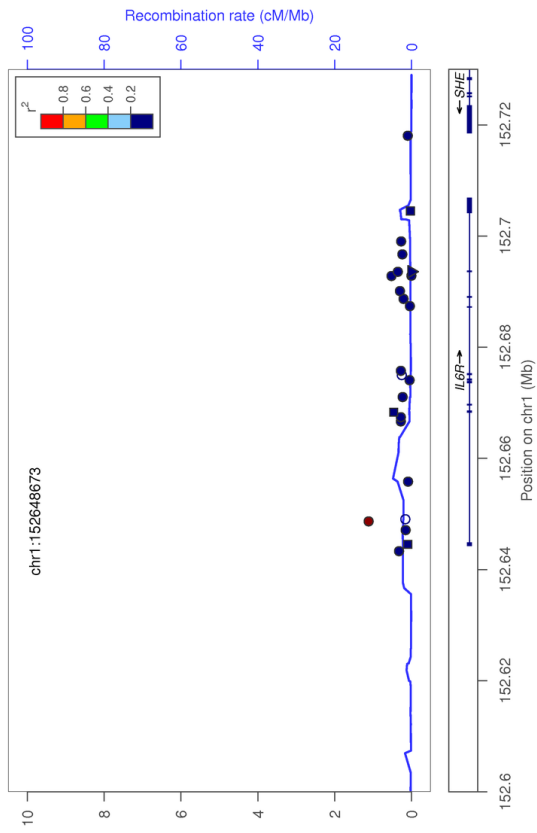
Supplementary Figure 4 ImmunoChip-wide HLA-B*27 interactions. Linkage disequilibrium (r) between the HLA-B*27 tagSNP rs116488202 and all SNPs assayed were computed in controls (A) and in cases (B). The blue and red vertical lines mark the LD observed between HLA-B*27 and the two identified independently associated *ERAP1* SNPs rs30187 and rs10045403. Epistasis between HLA-B*27 and all other non-MHC SNPs was assessed via logistic regression where we included an interaction term between the SNP tested and the tag SNP rs116488202, and population stratification was modeled with the first five principal components as covariates. Q-Q plots for the interaction term for all SNPs tested excluding the MHC (C) and for all SNPs tested excluding the MHC and those in the *ERAP1* locus (D).

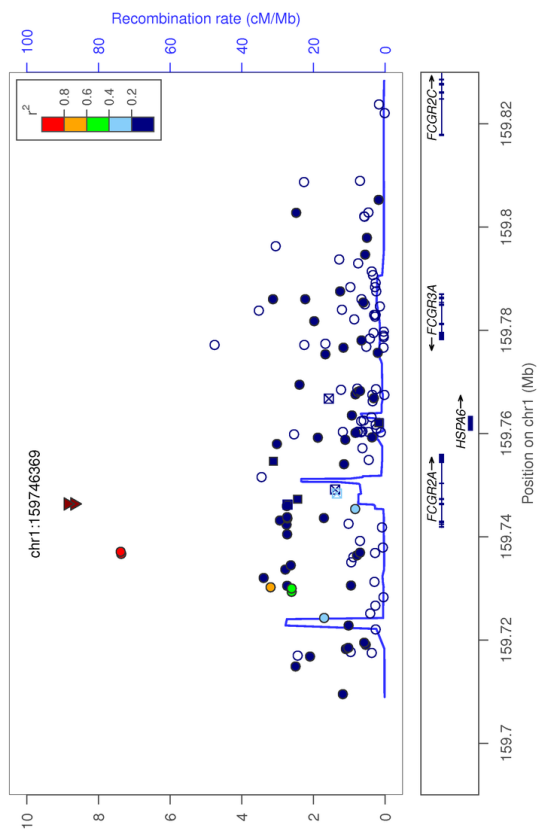
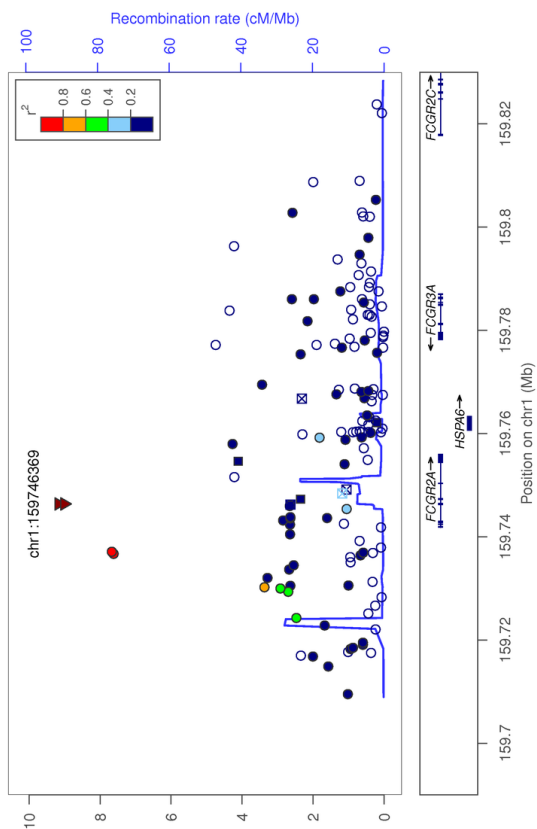
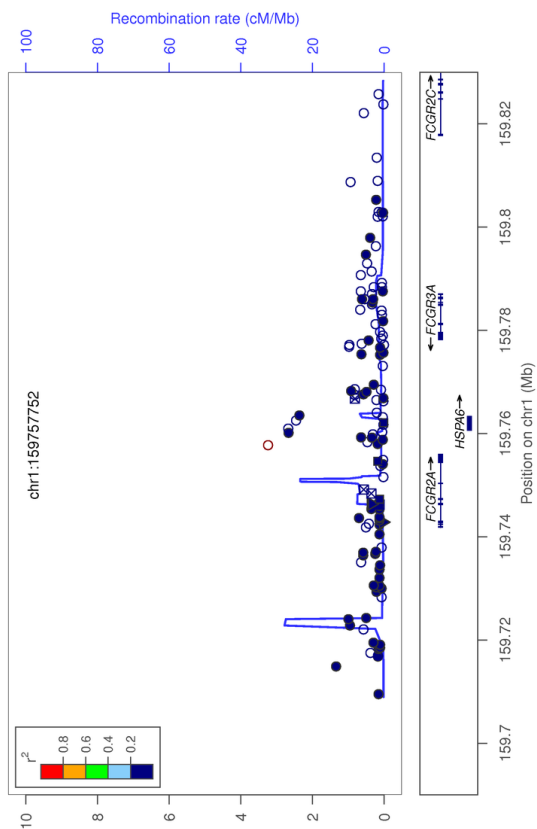


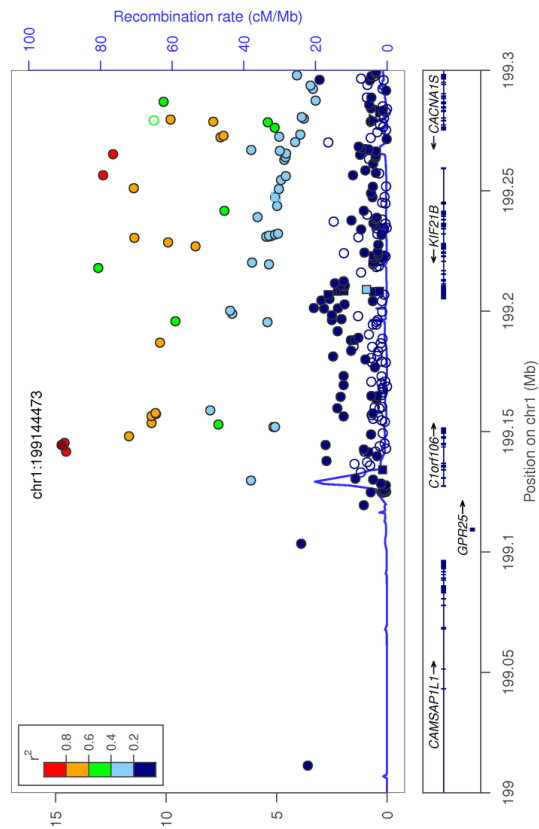
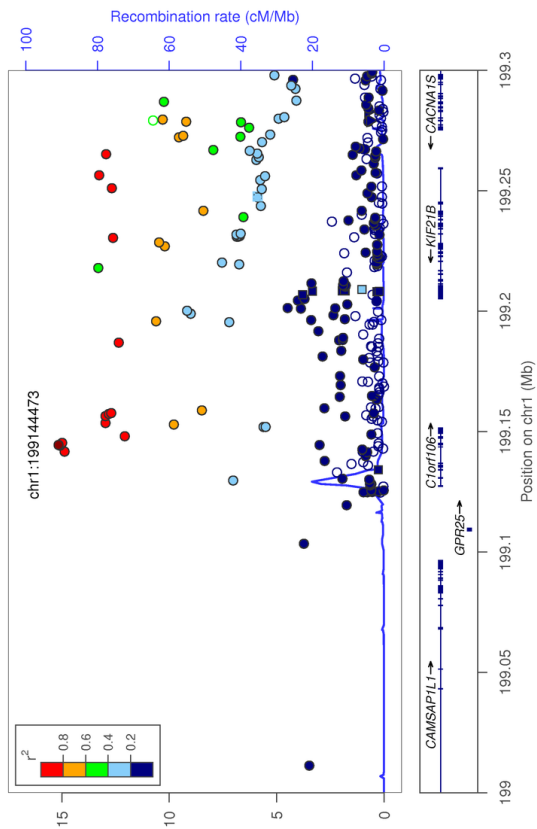
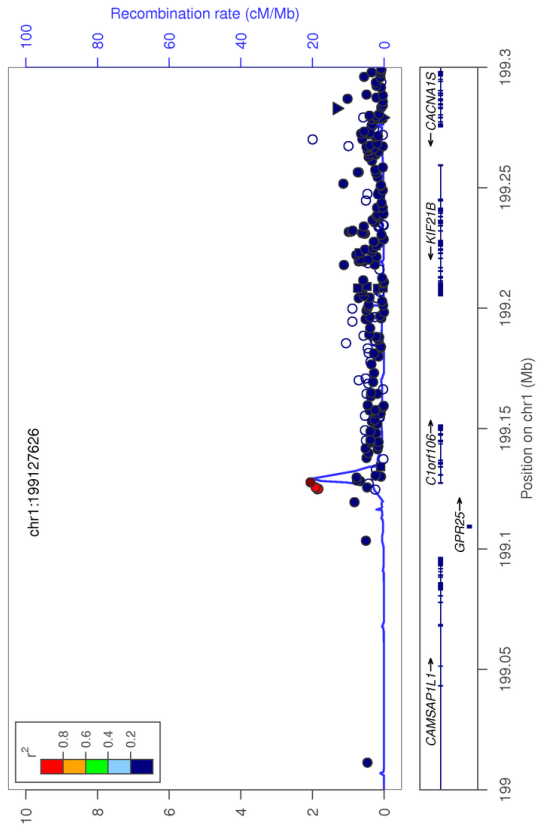
Supplementary Figure 5 Locus association plots and linkage disequilibrium patterns with lead SNP in each of the three analyses performed: European cohort (upper left), Asian cohort (upper right) and European plus Asian cohort (lower left).

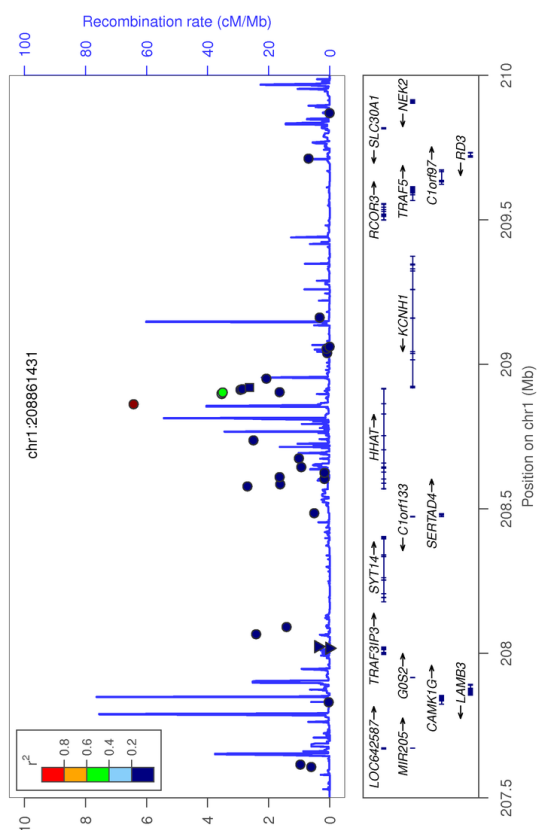
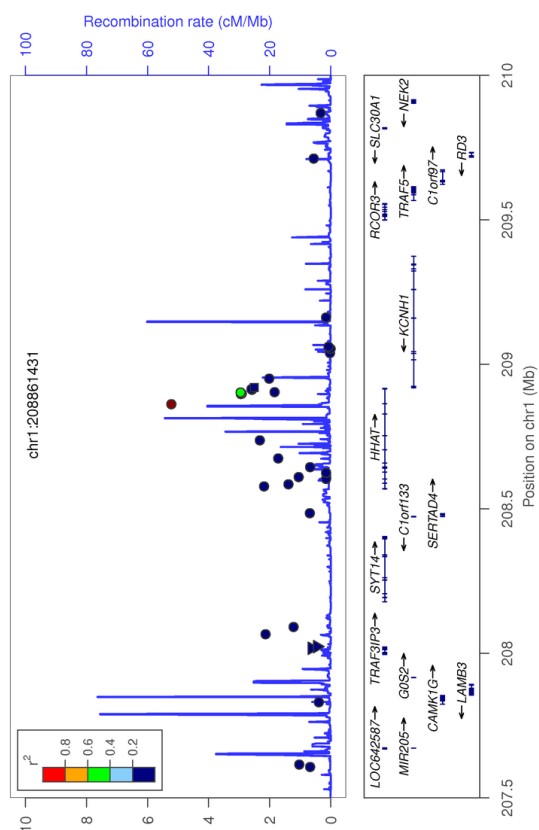
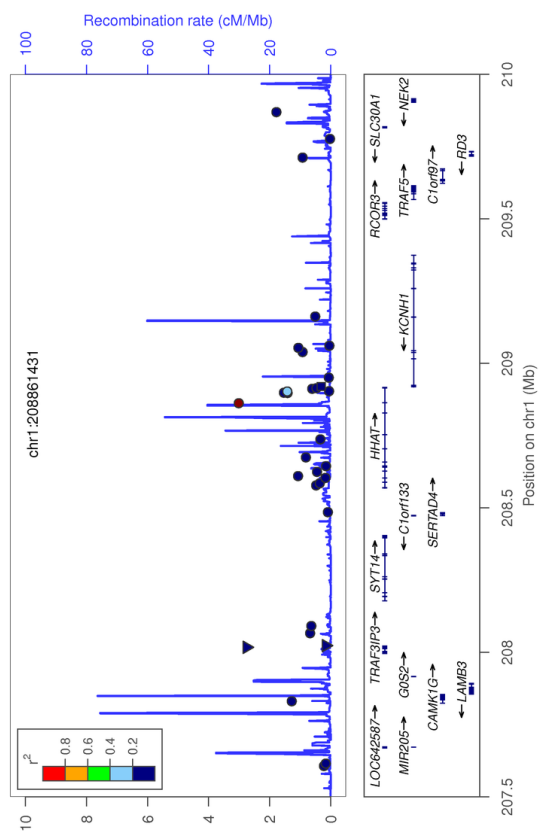


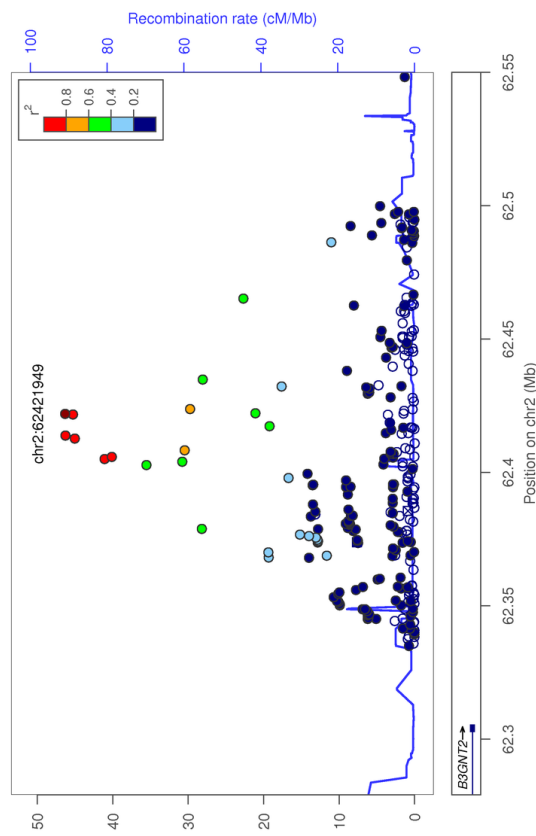
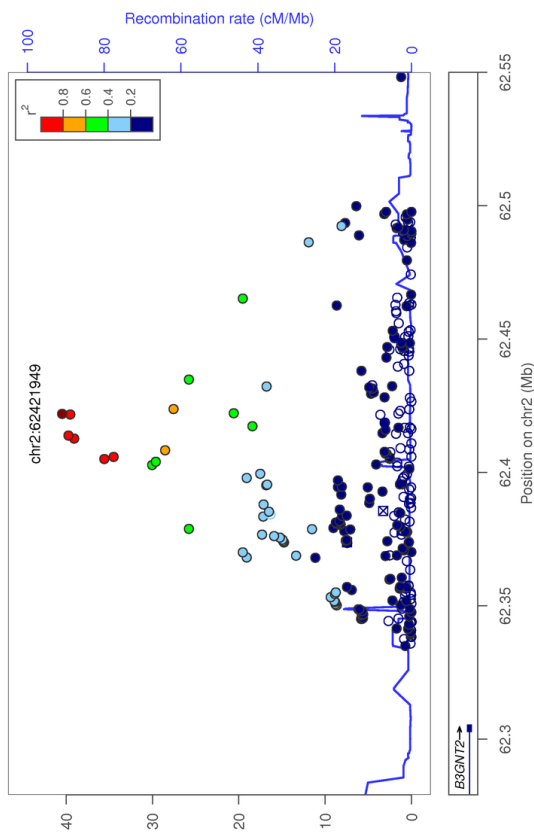
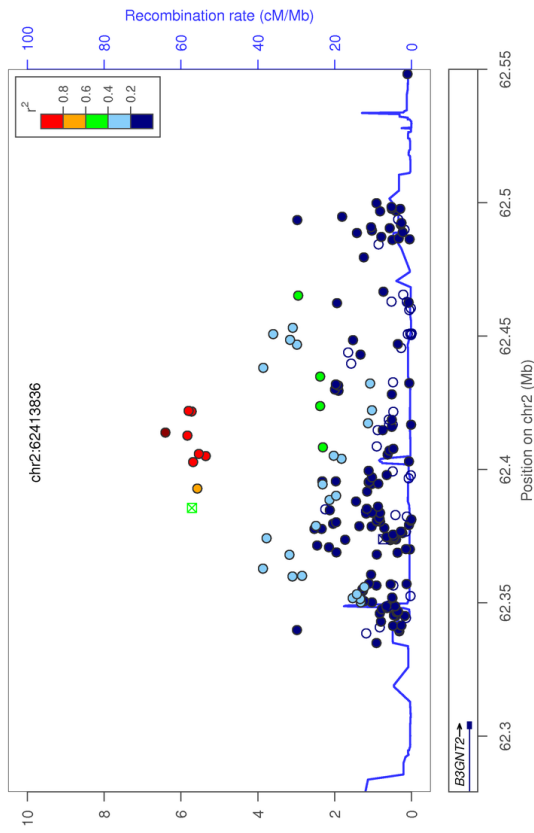


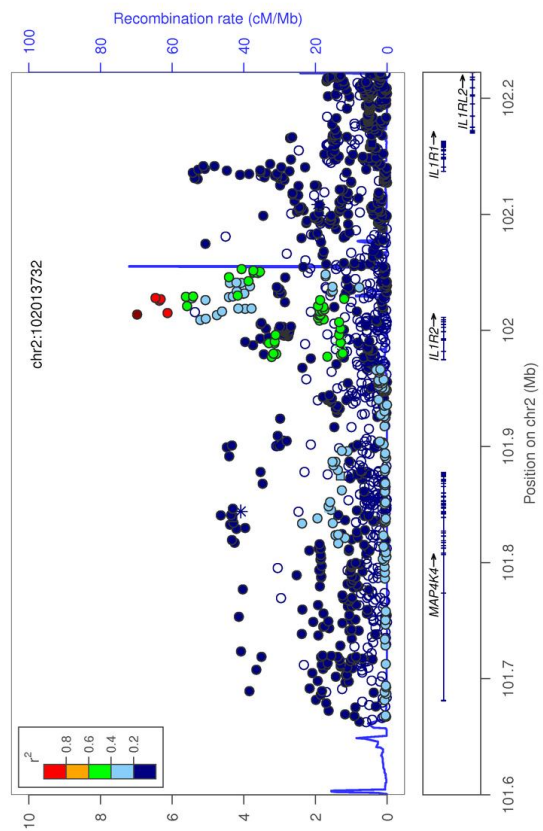
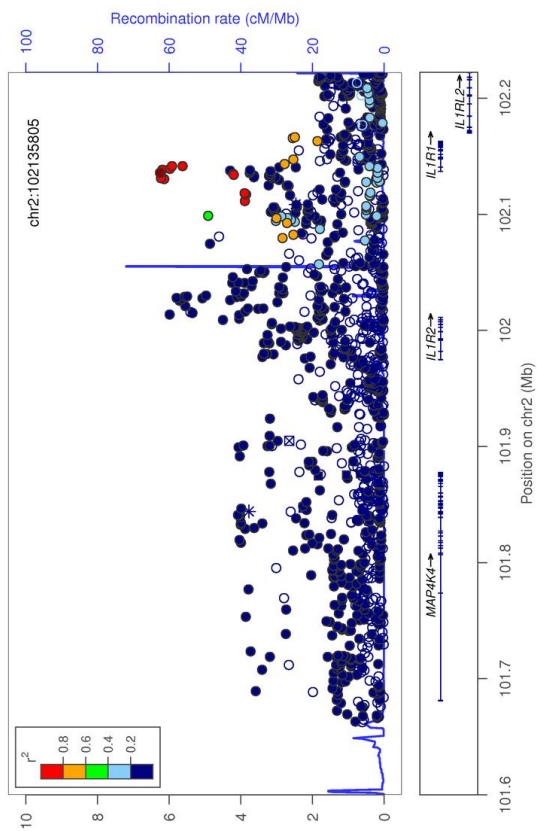
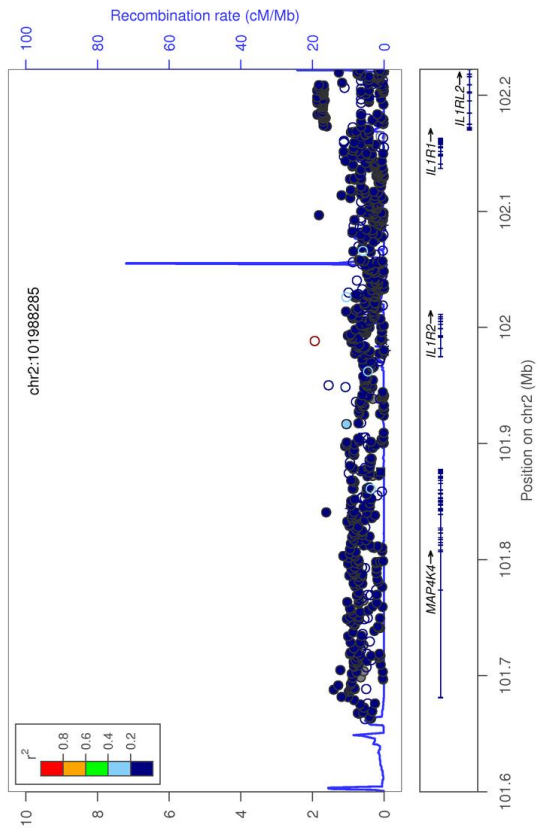


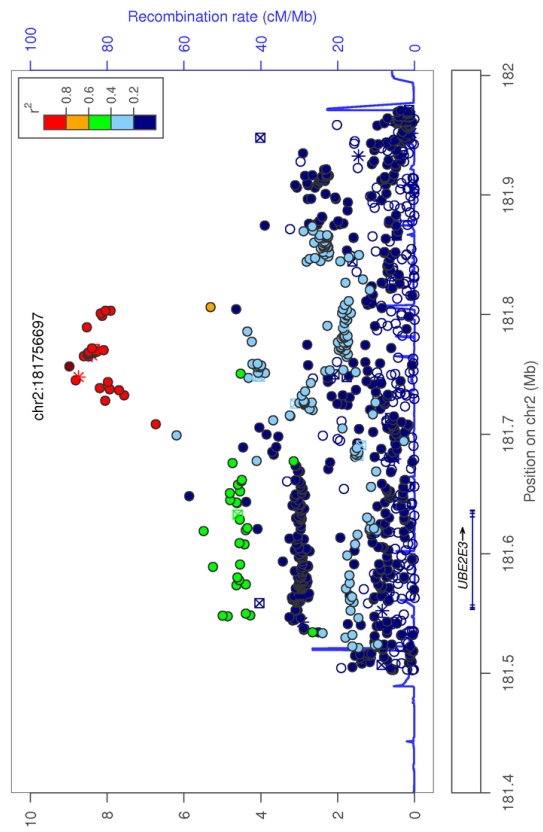
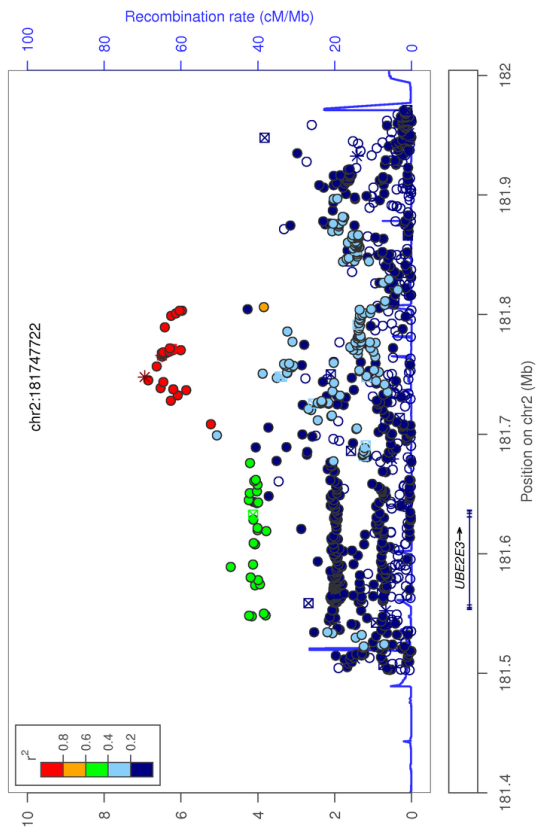
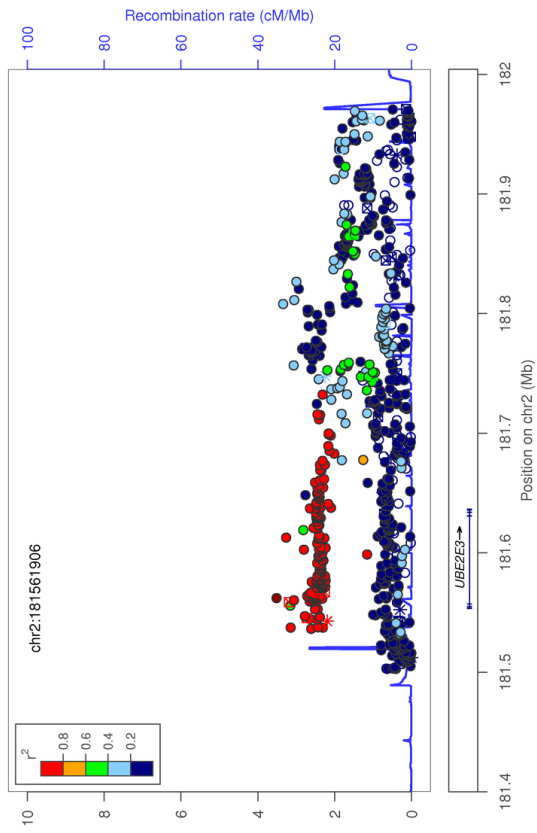


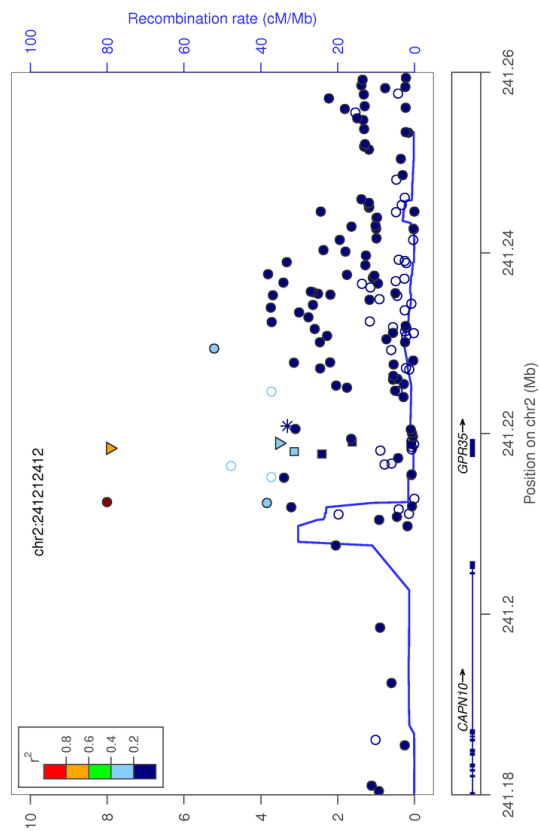
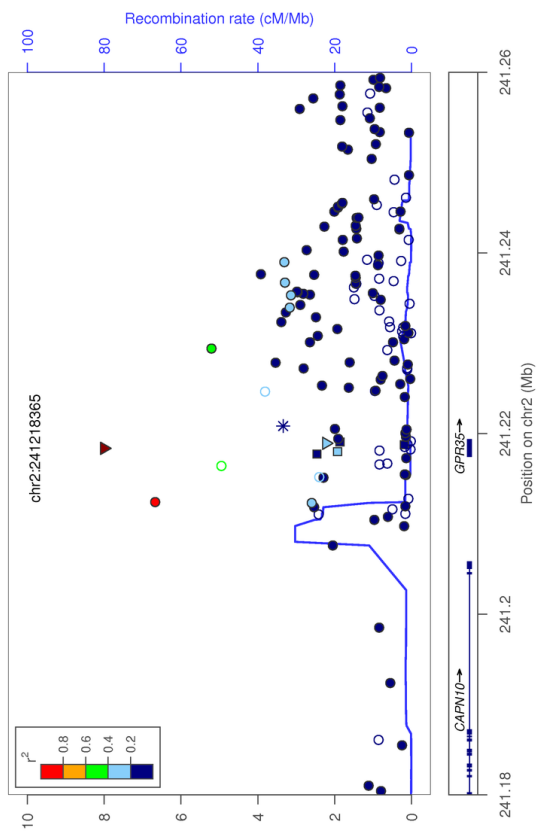
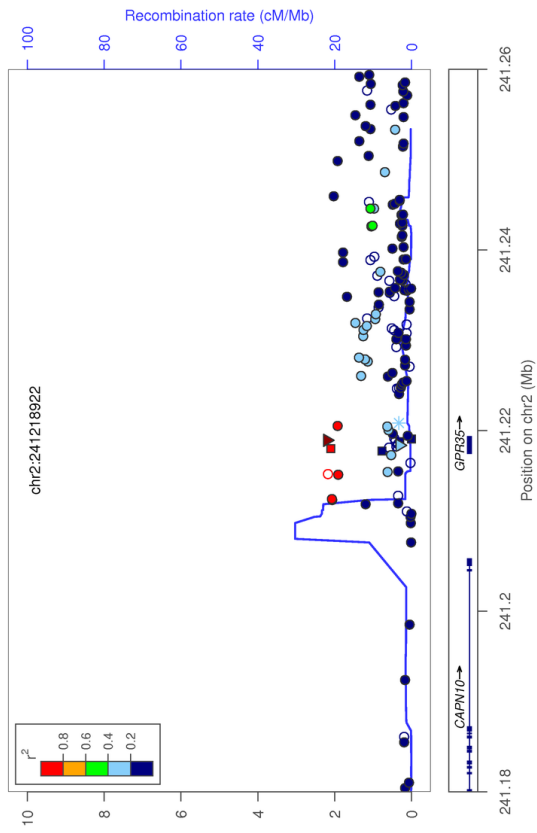


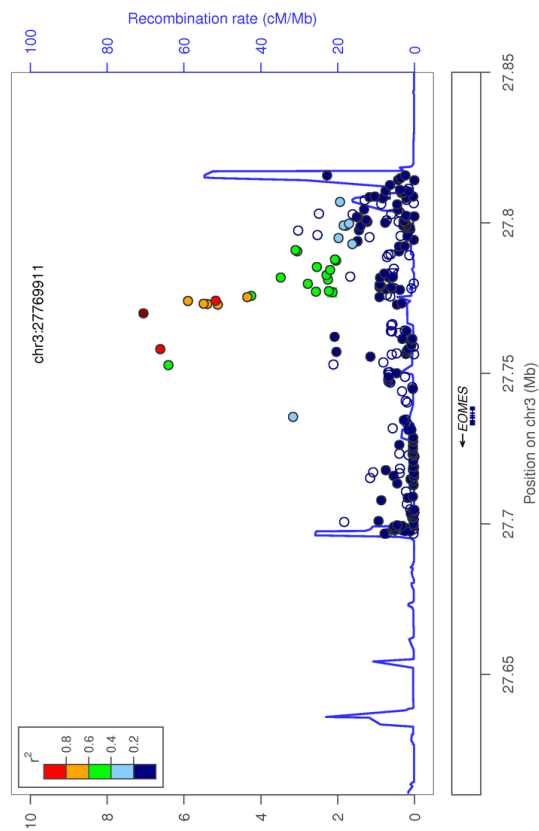
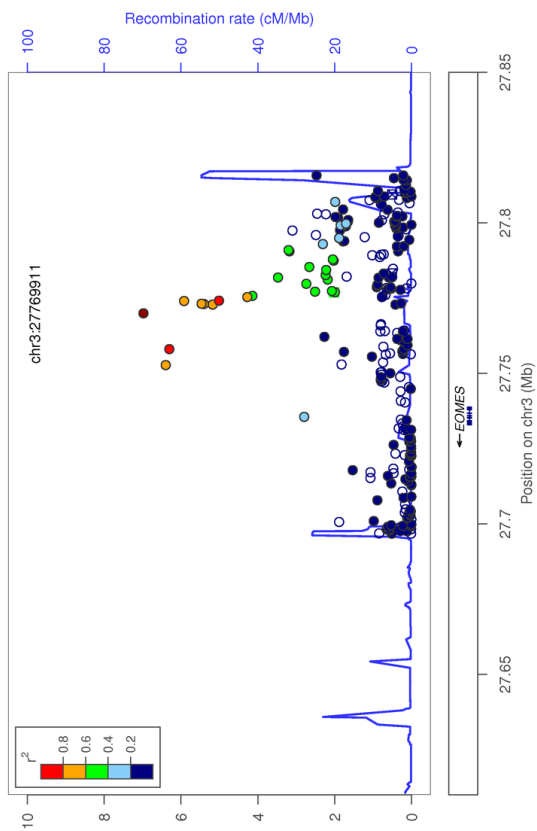
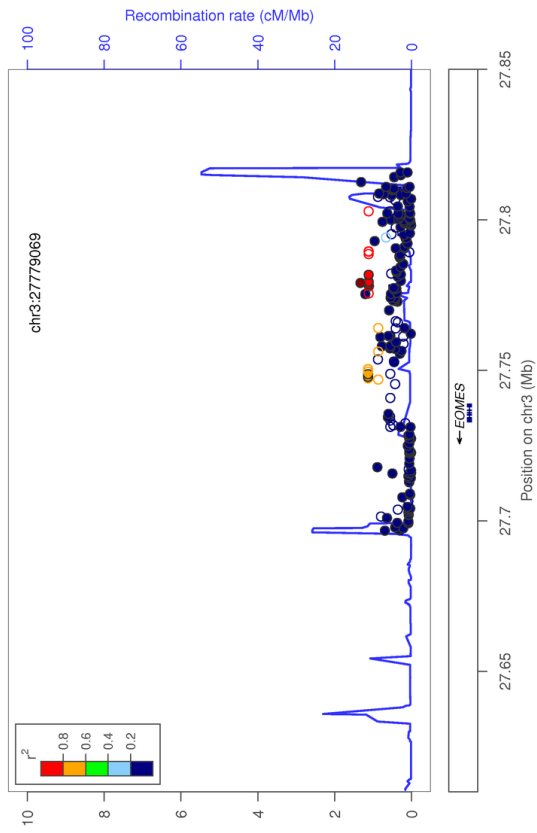


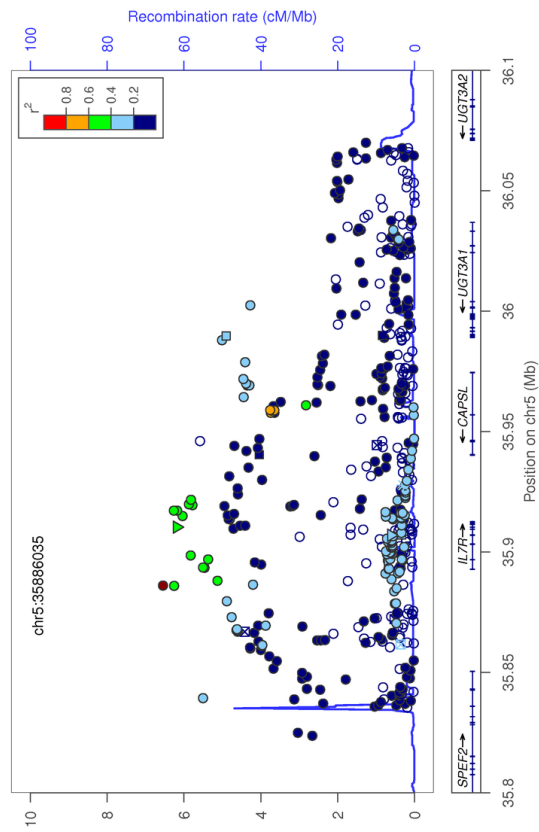
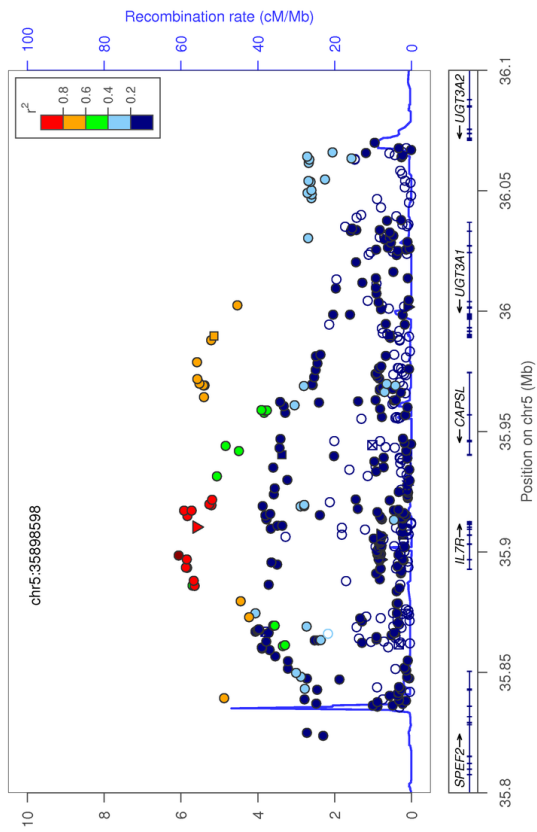
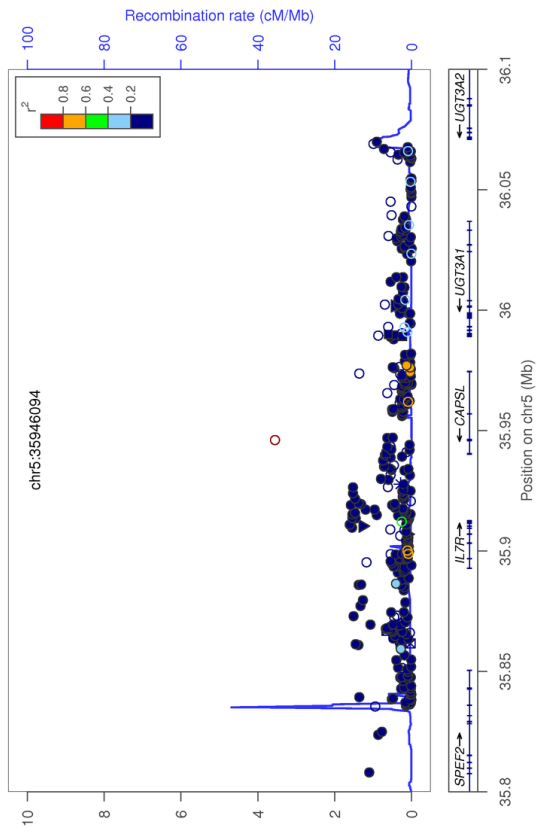


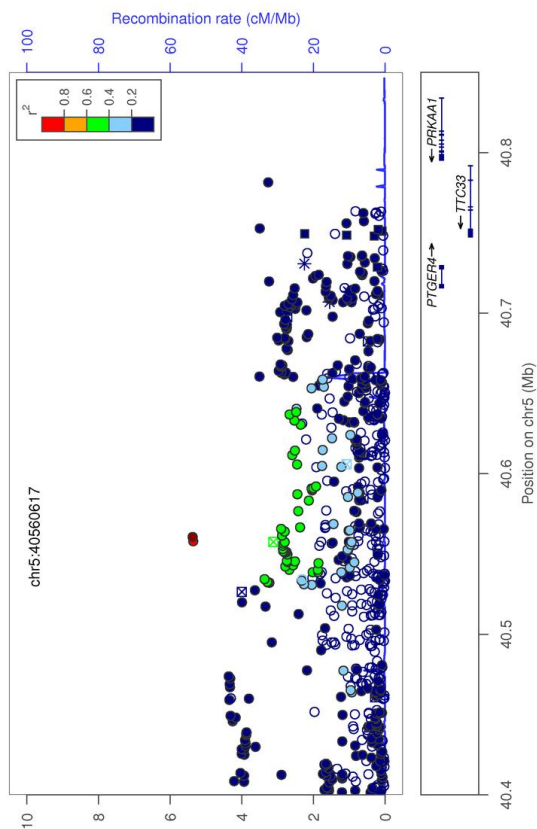
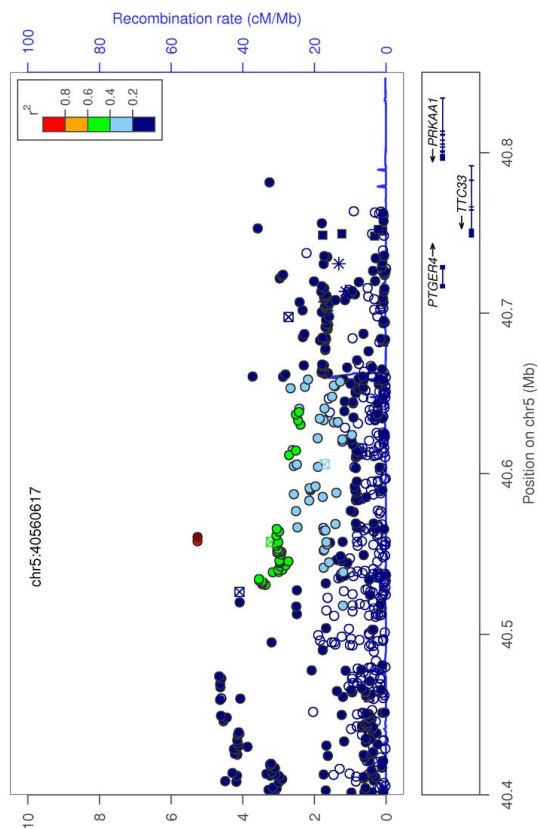
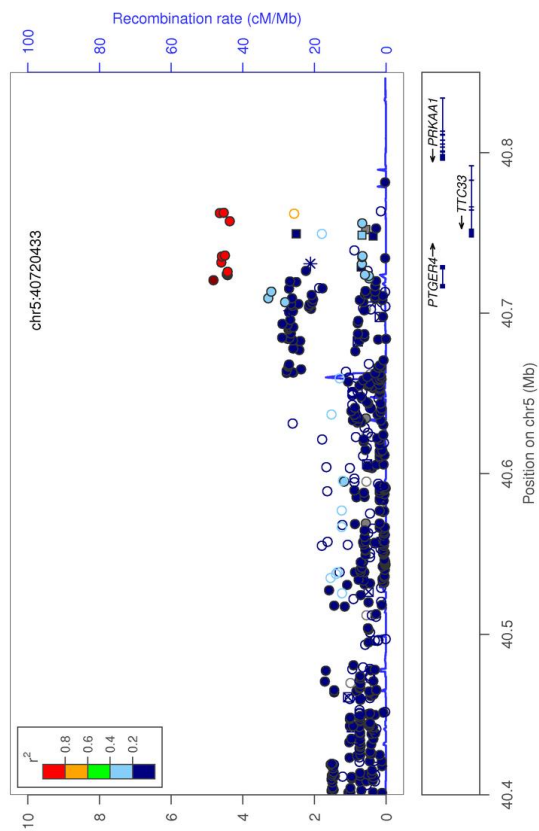


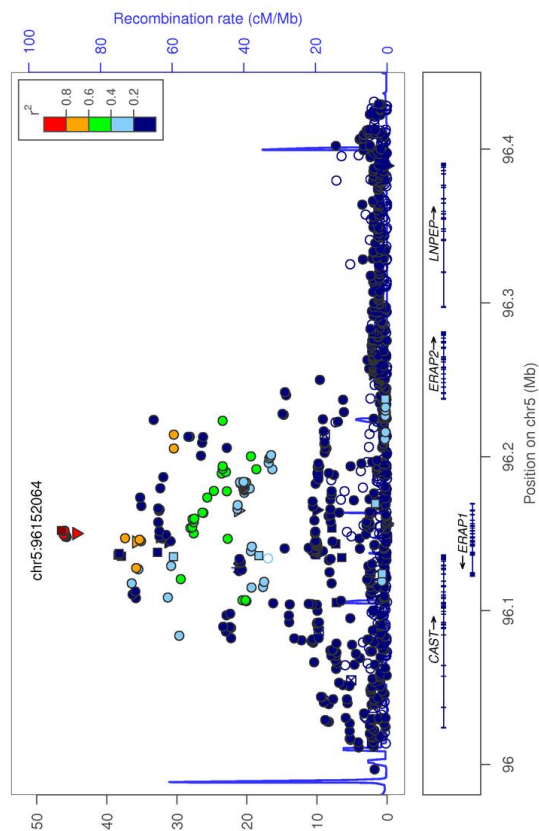
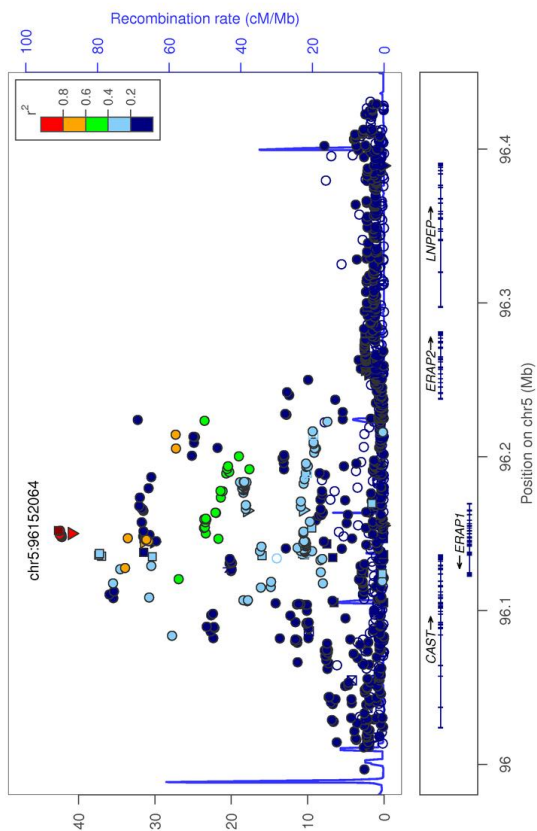
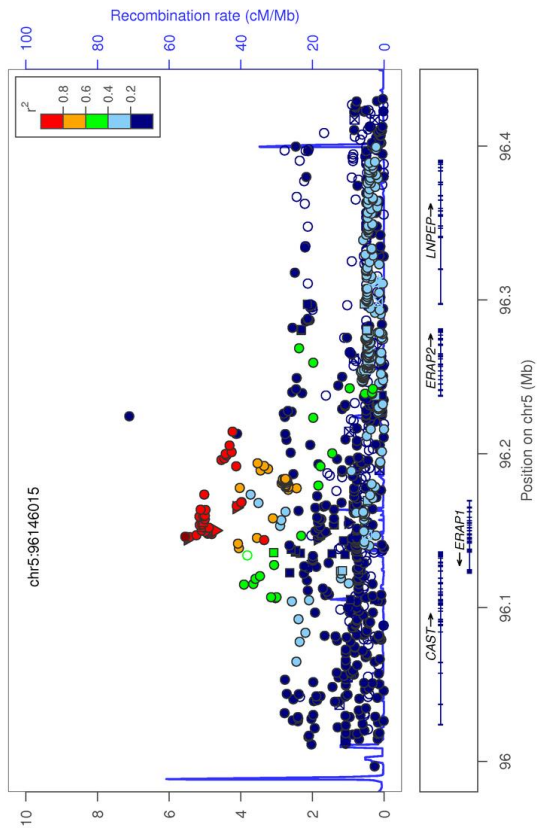


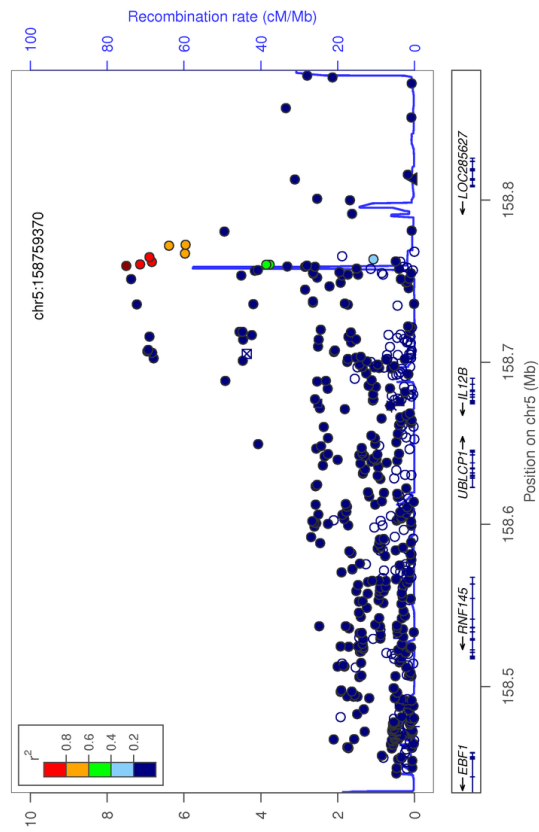
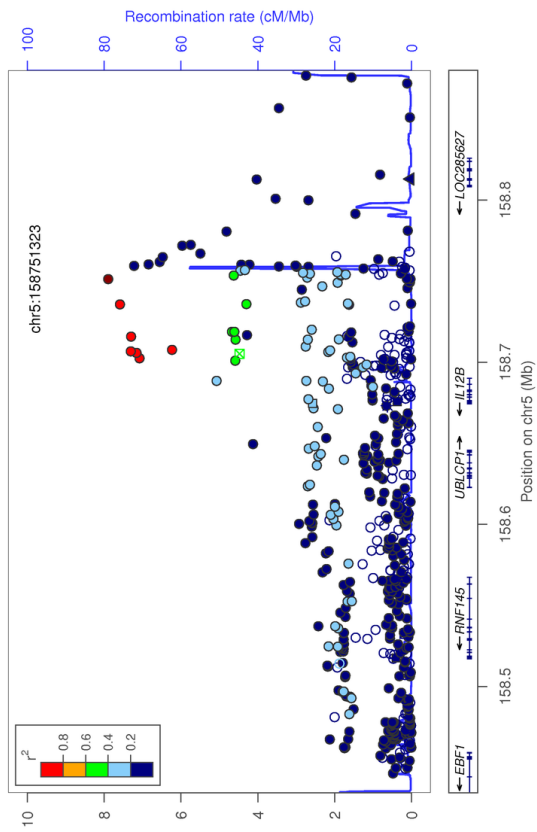
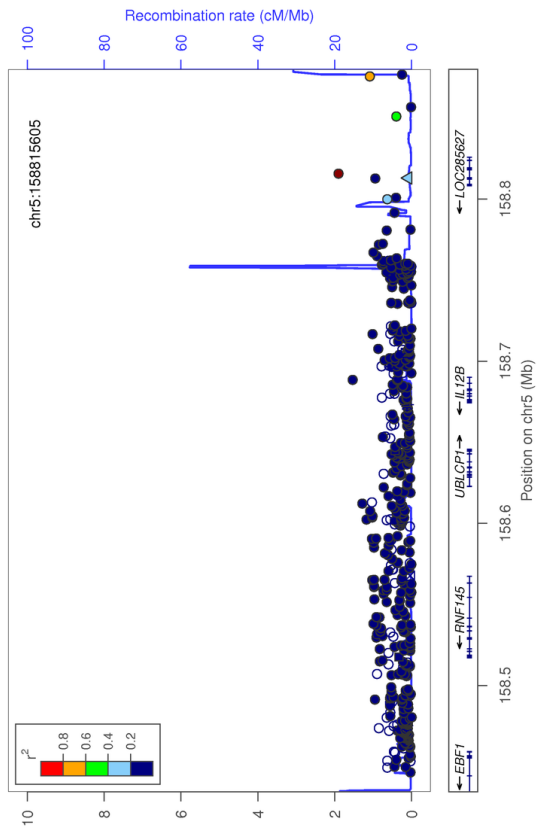


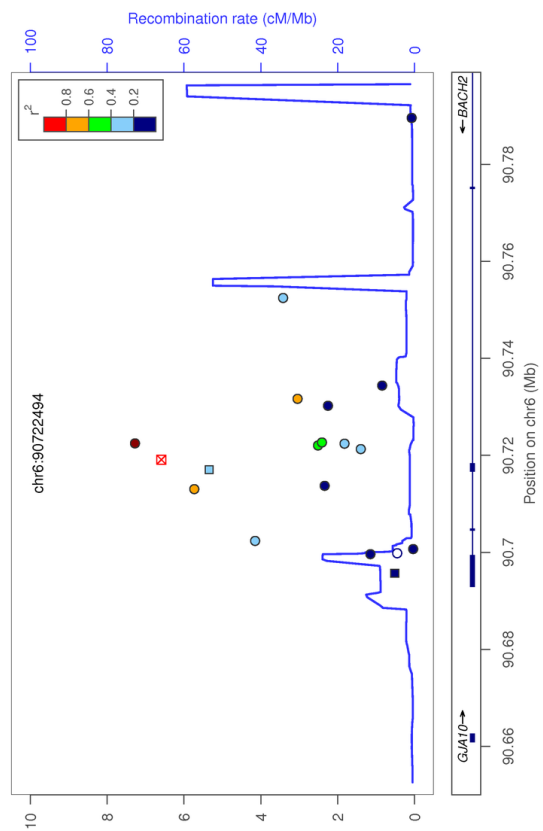
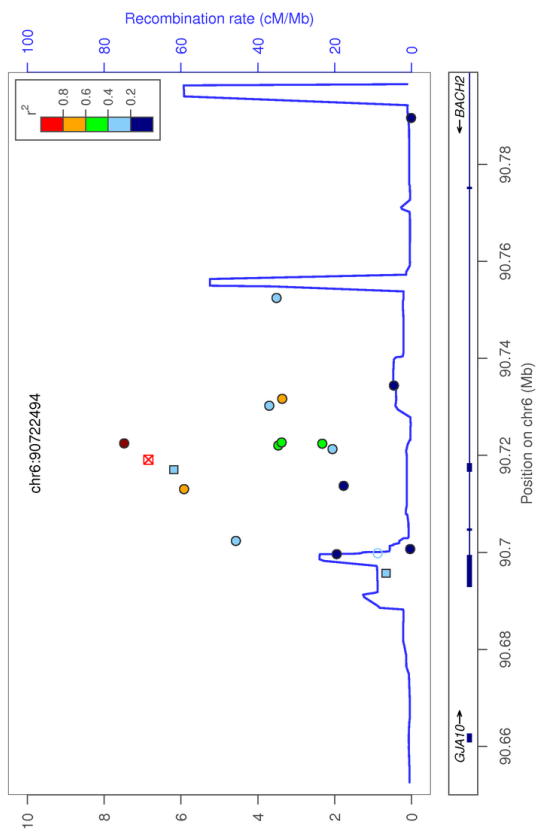
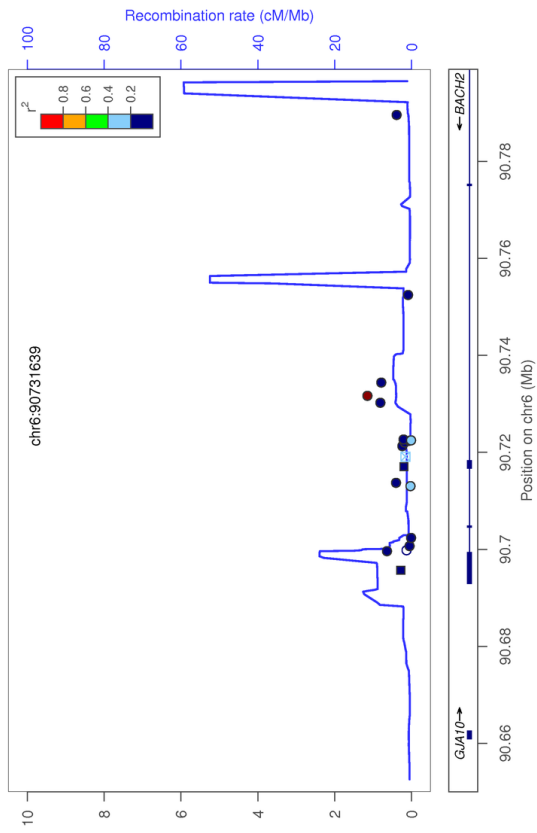


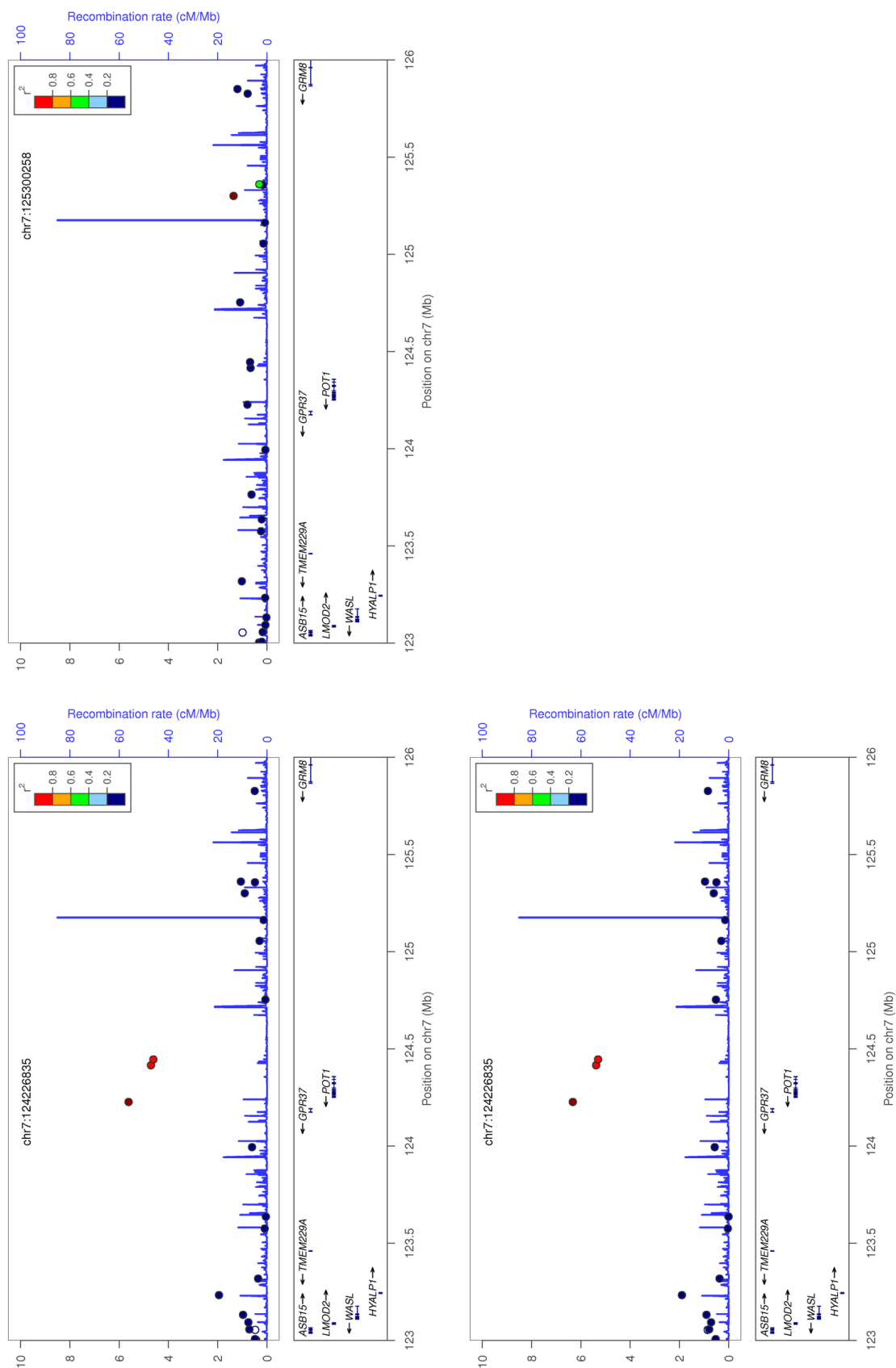


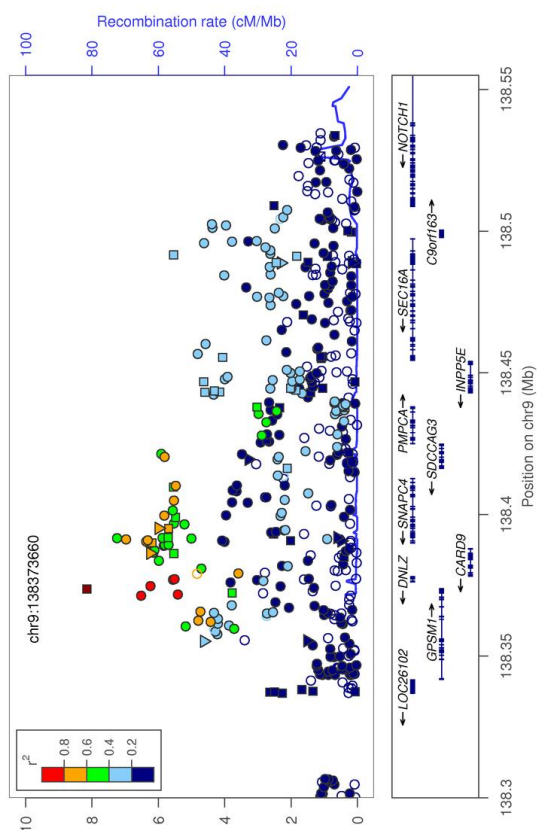
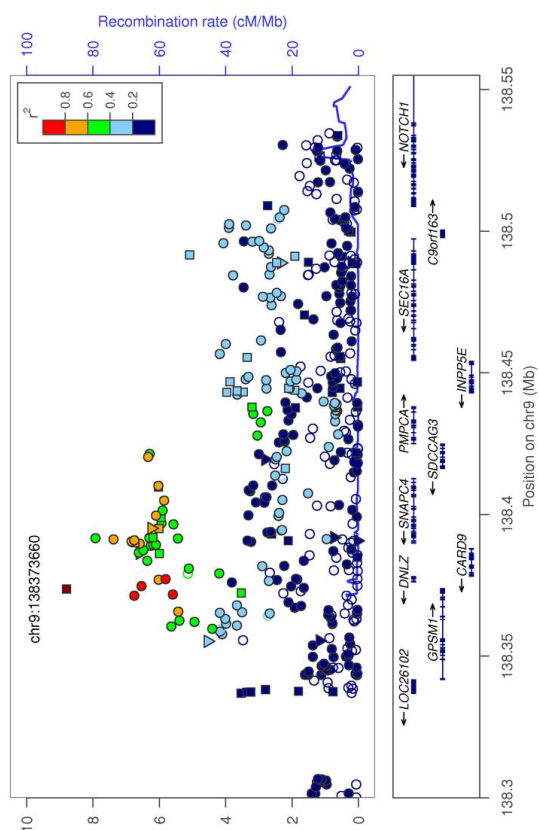
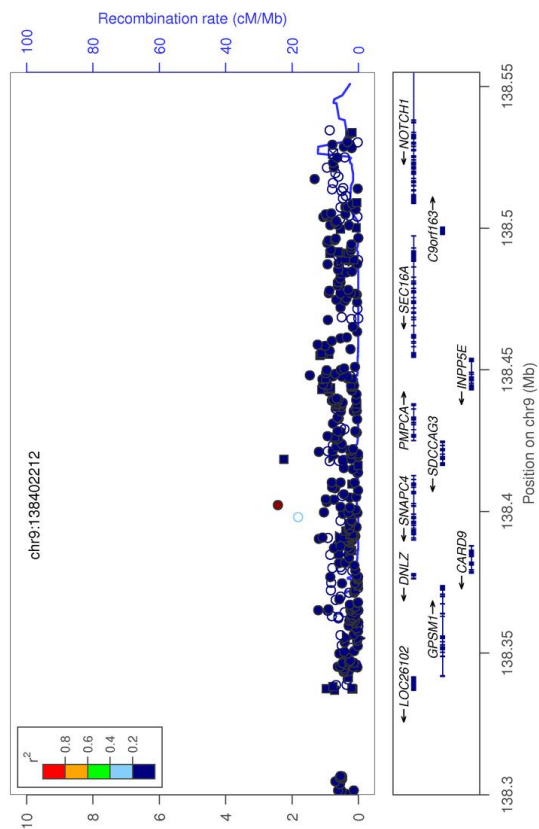


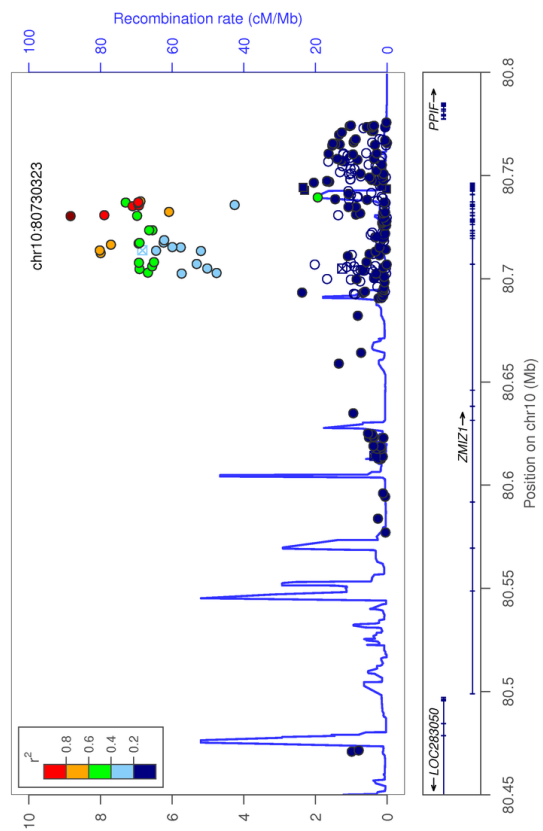
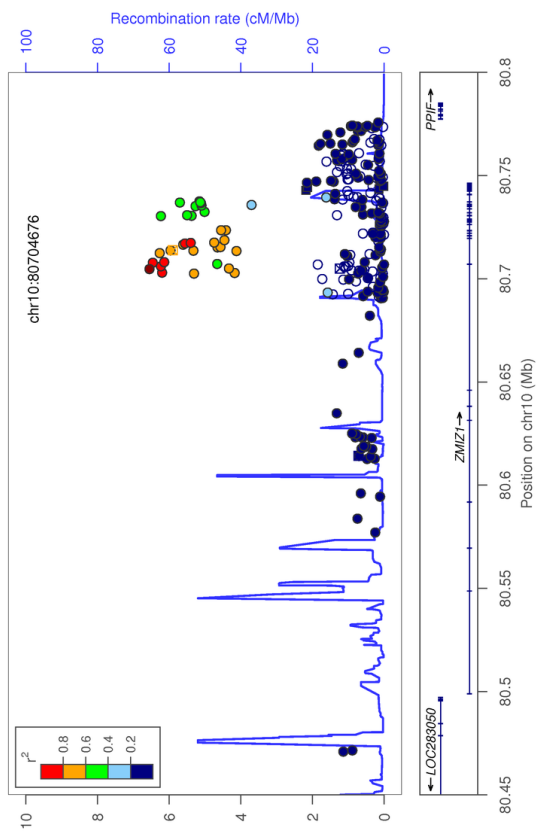
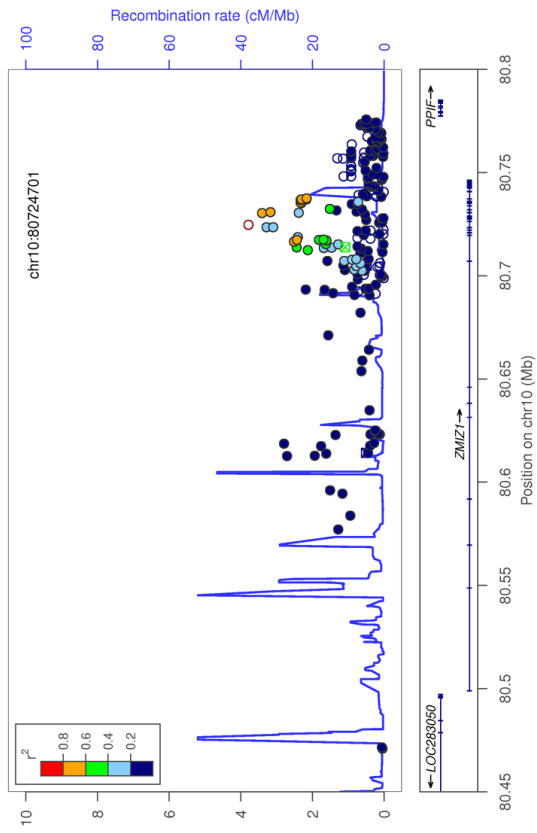


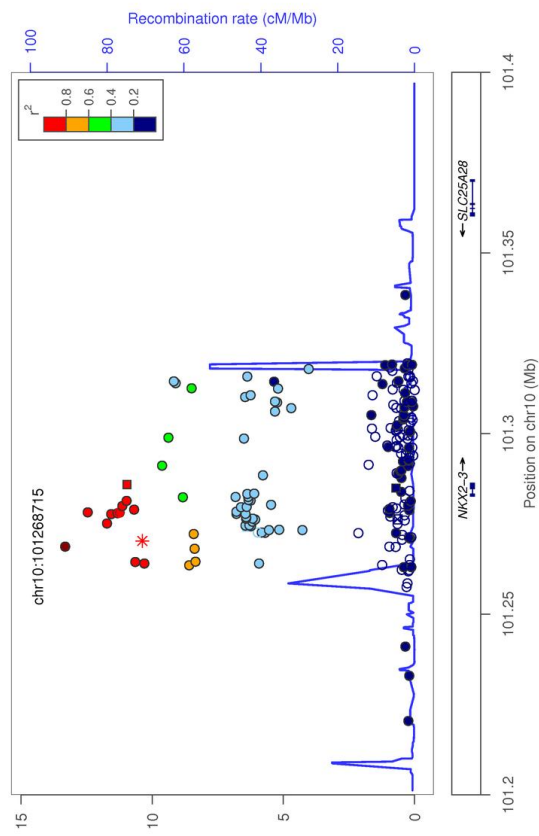
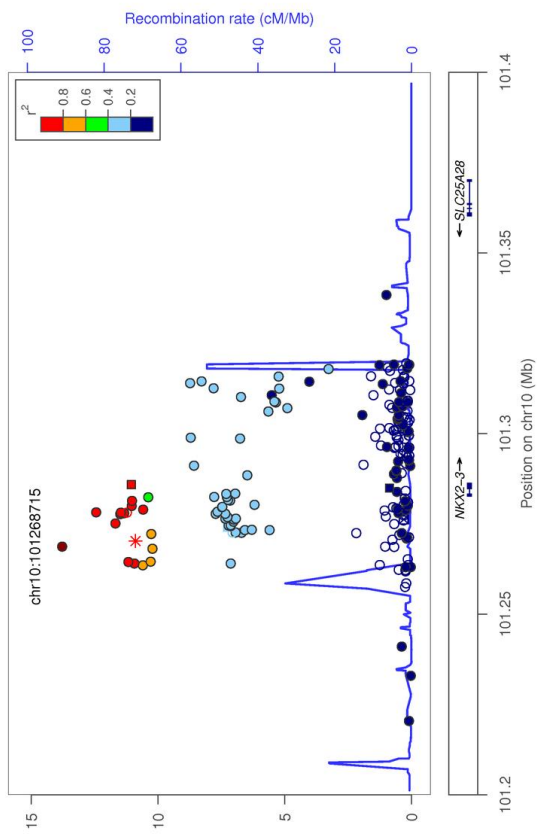
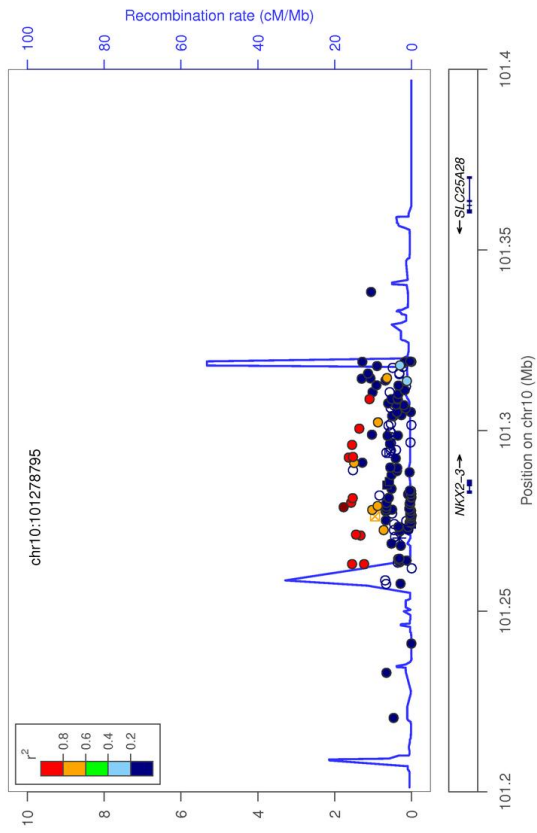


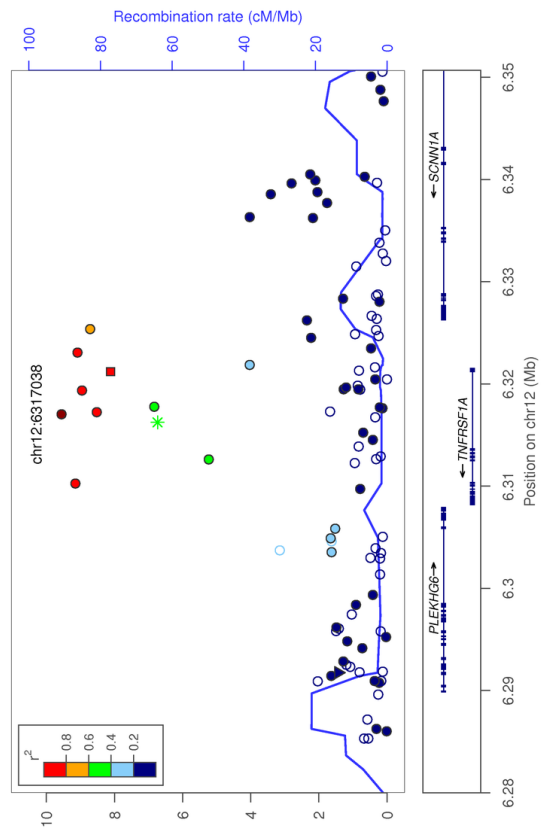
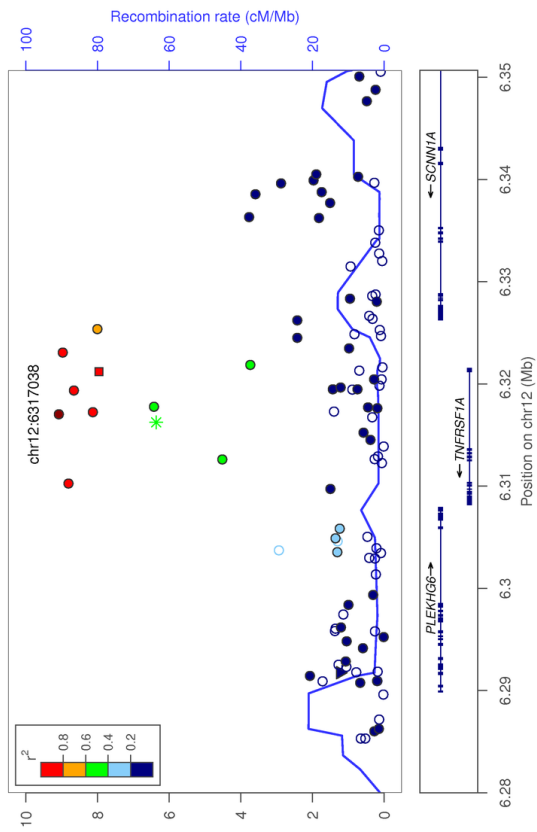
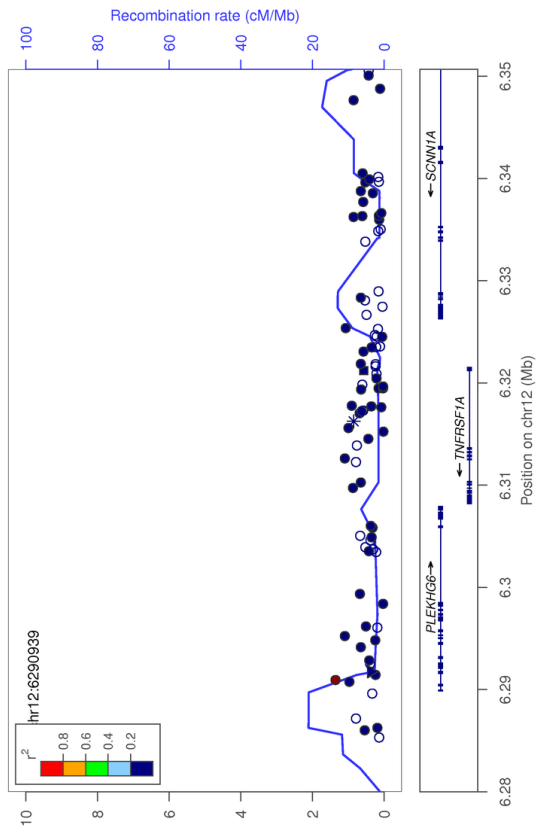


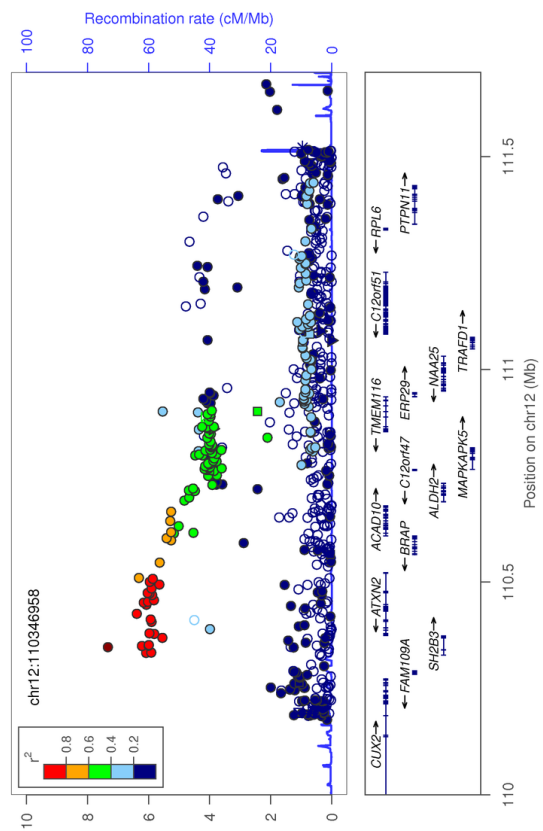
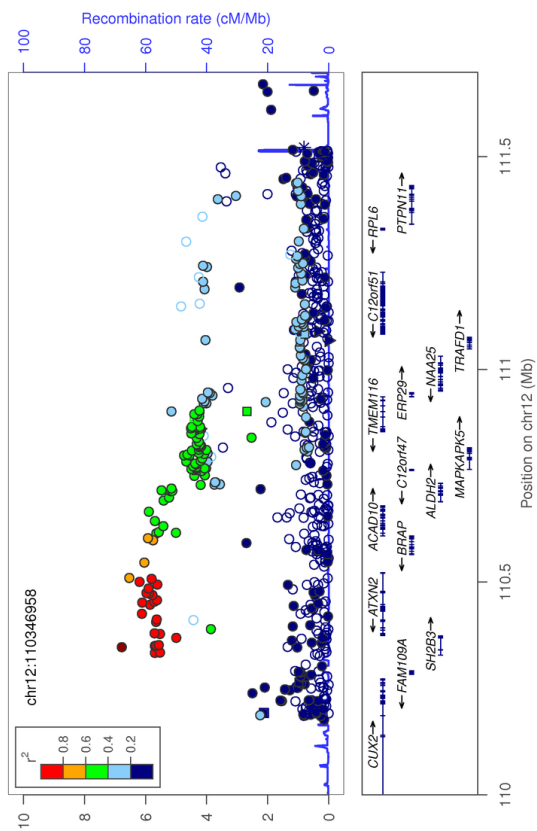
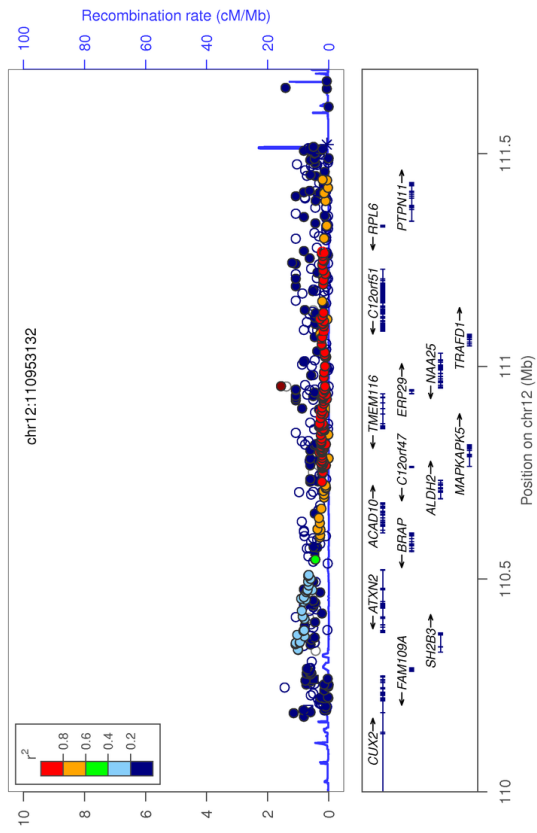


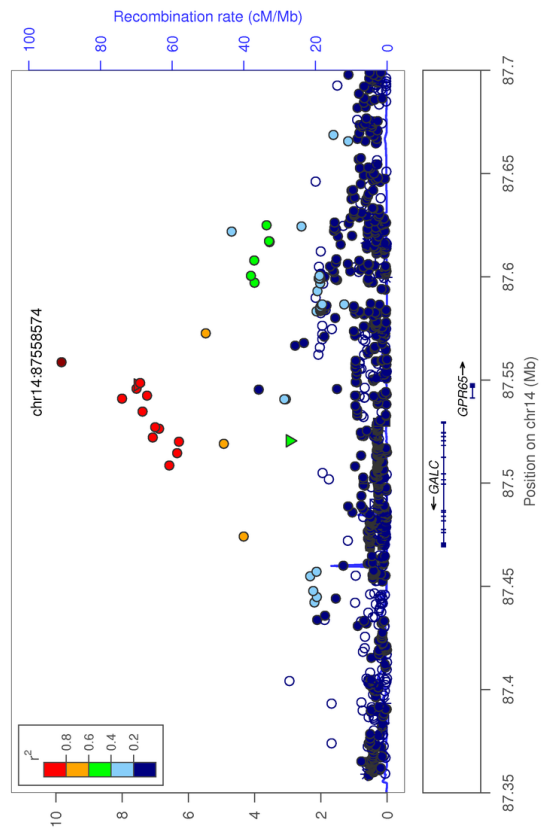
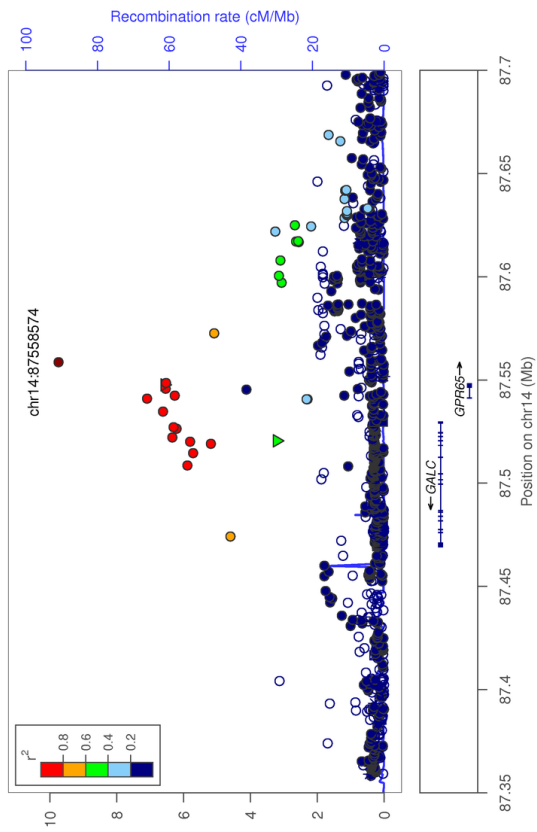
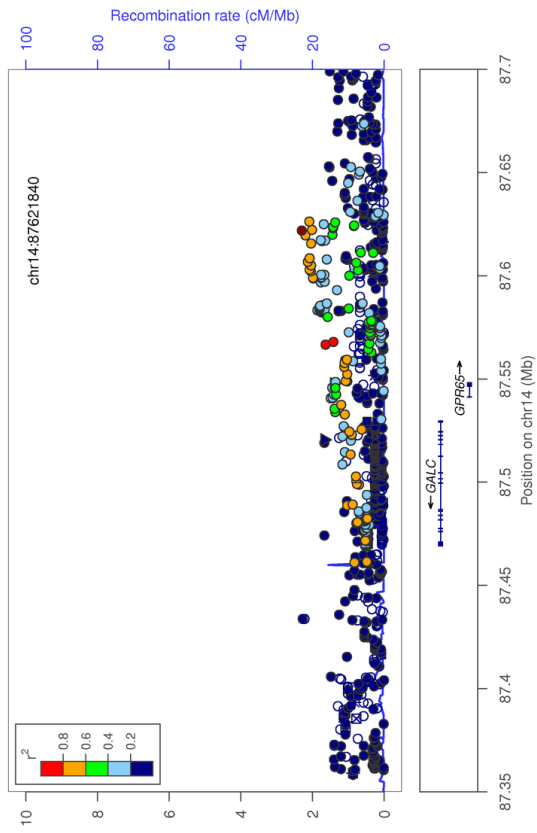


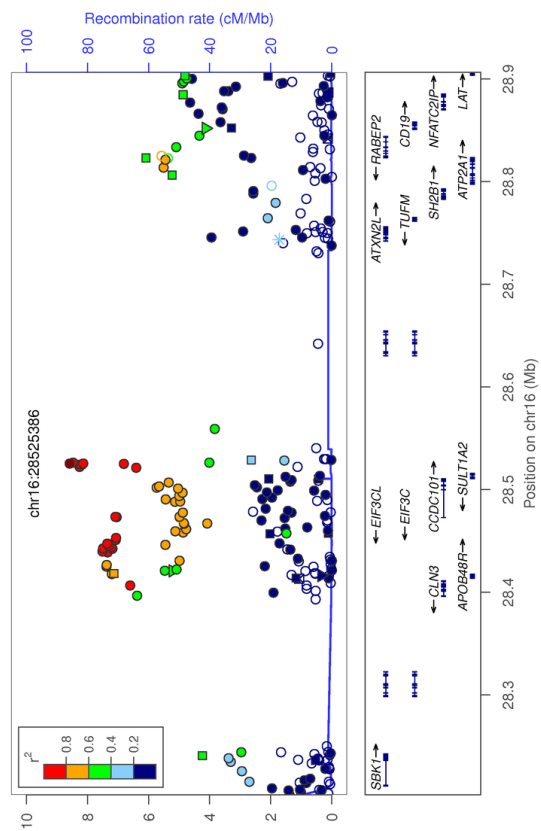
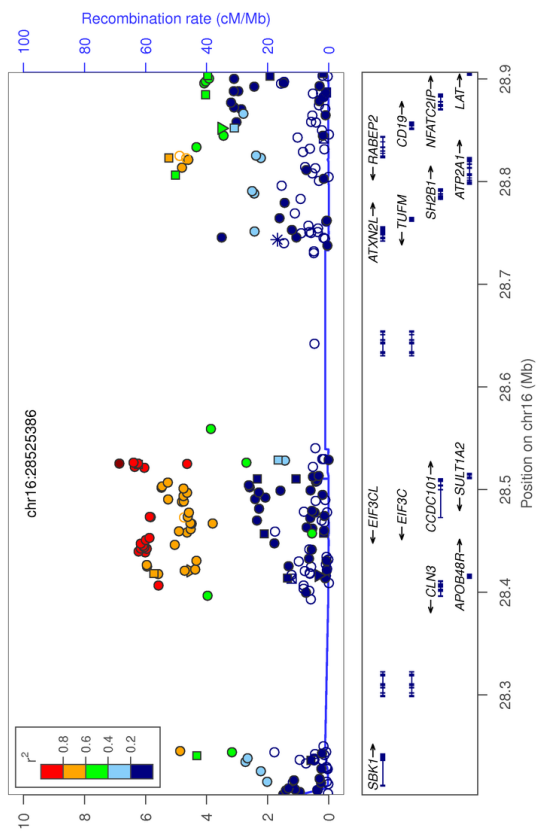
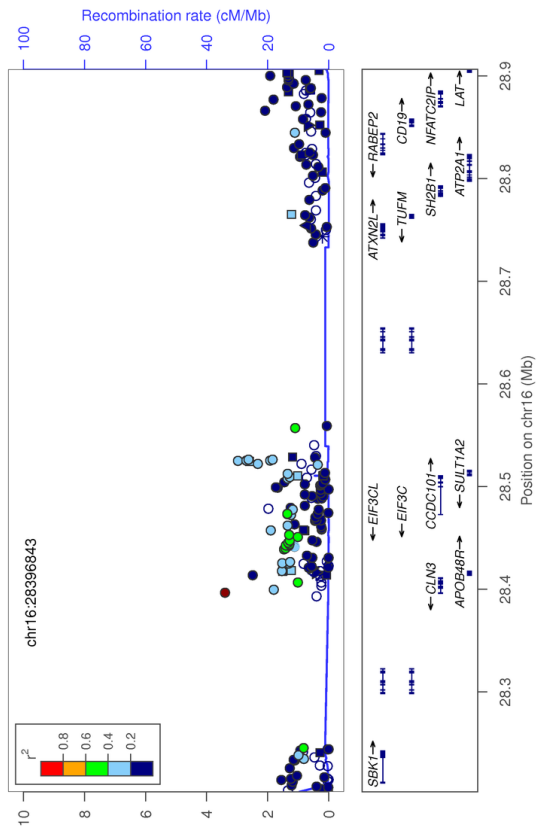


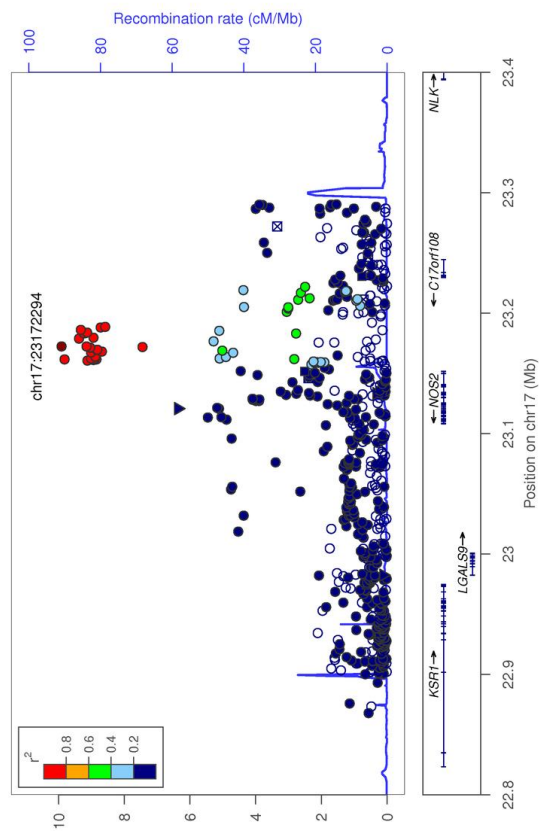
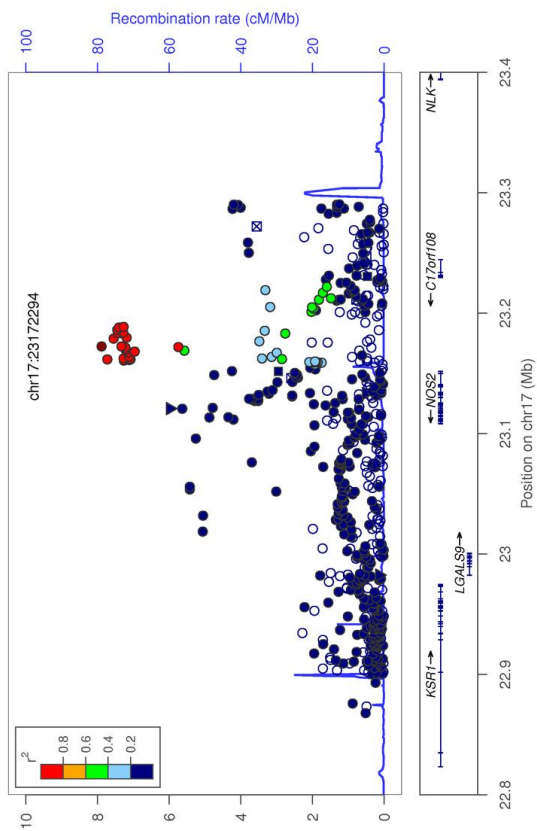
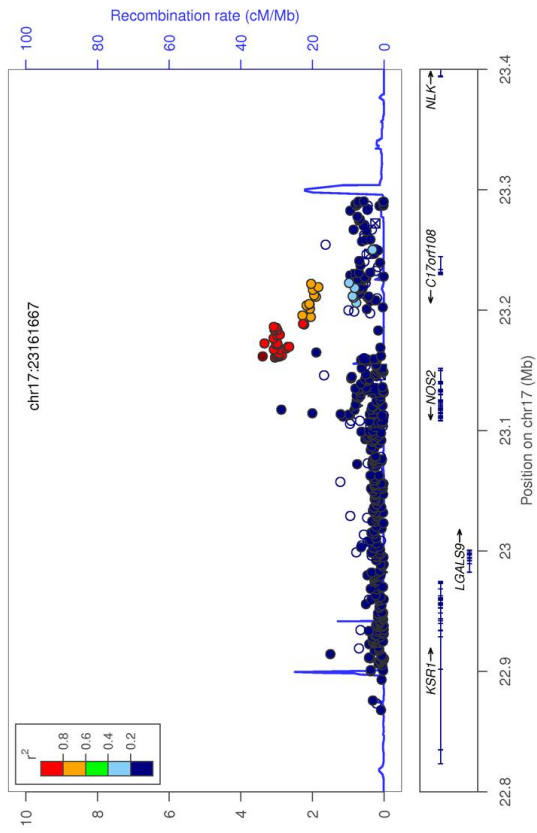


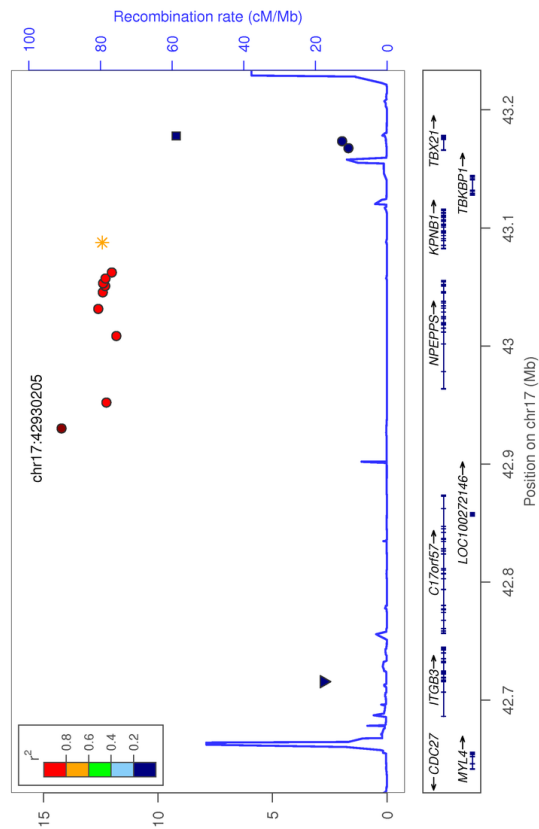
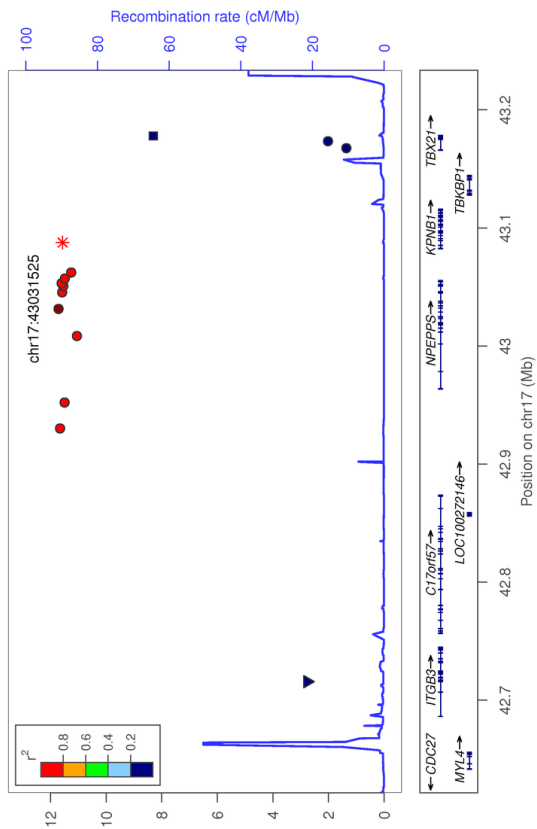
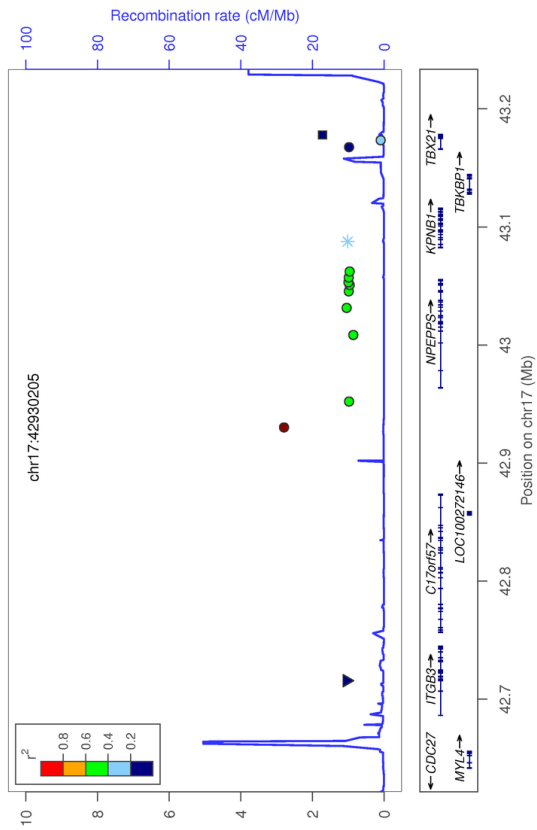


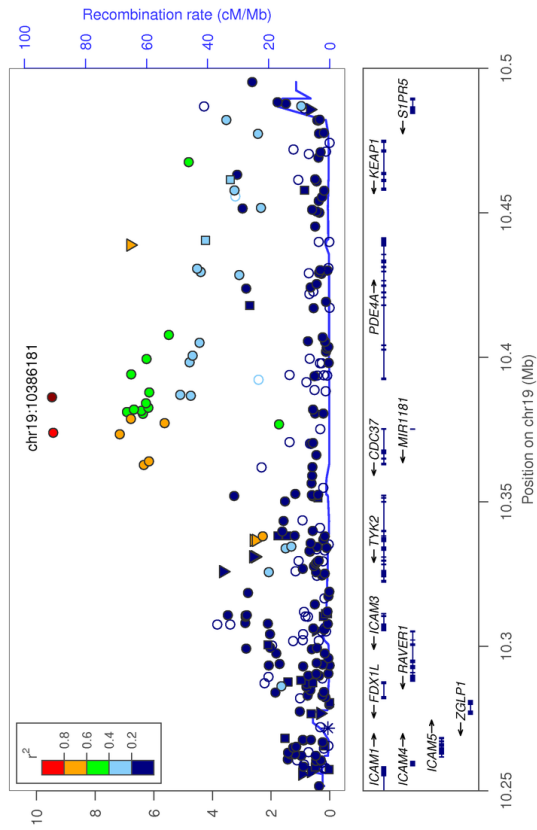
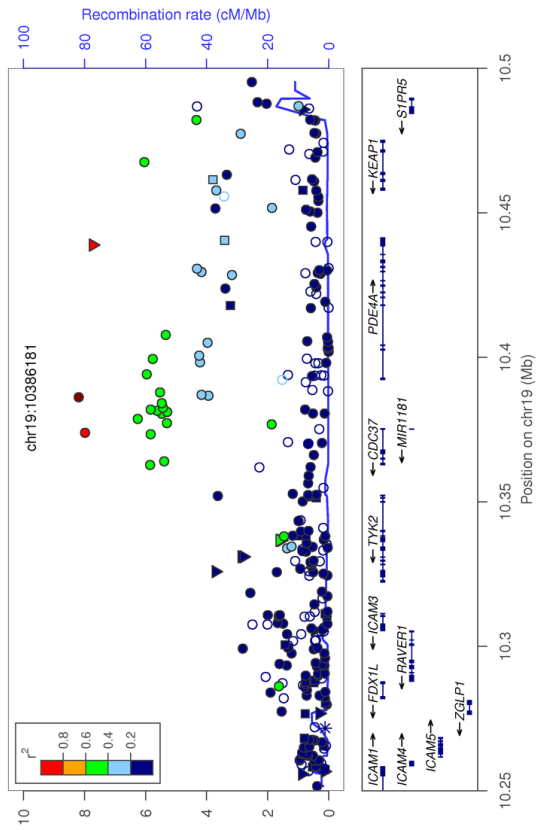
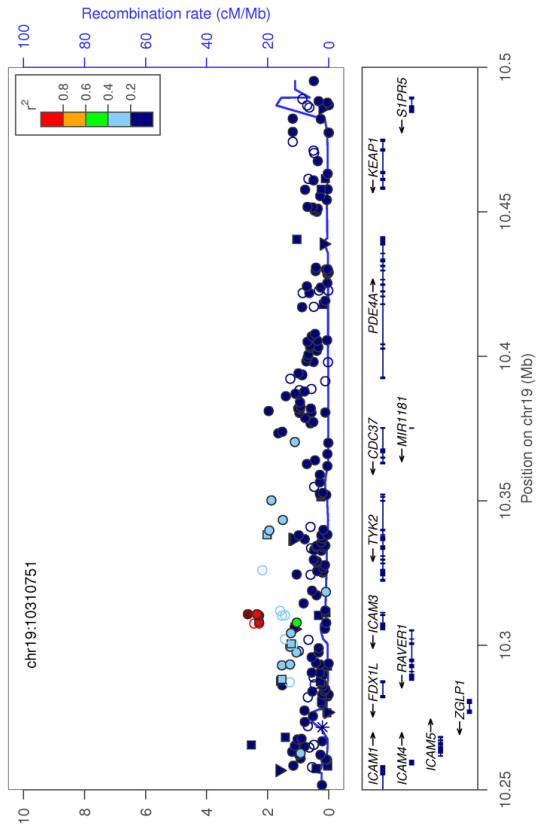


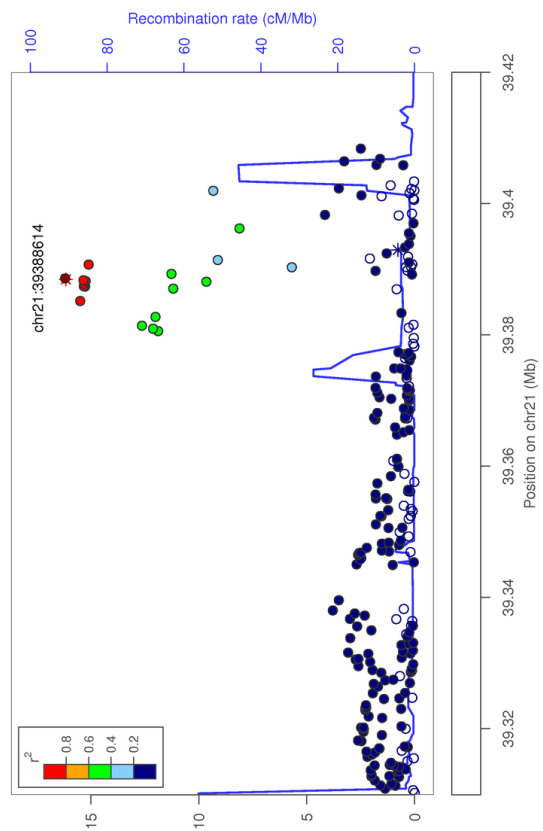
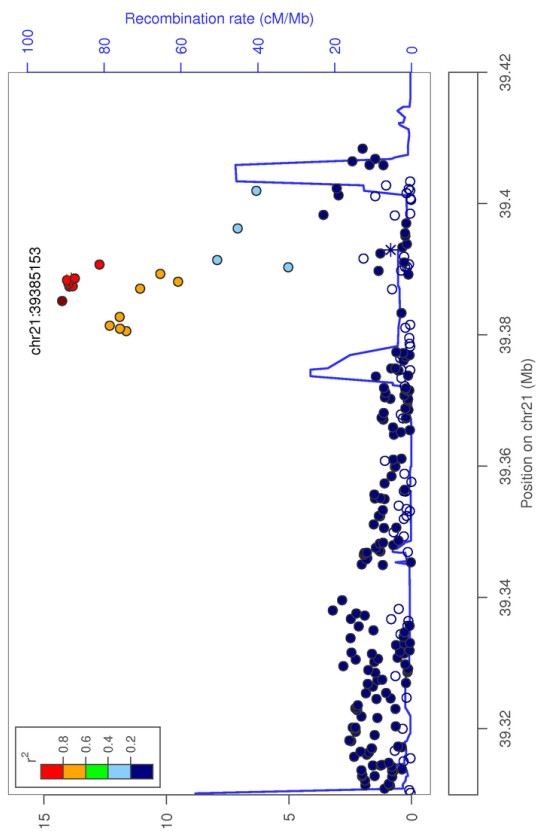
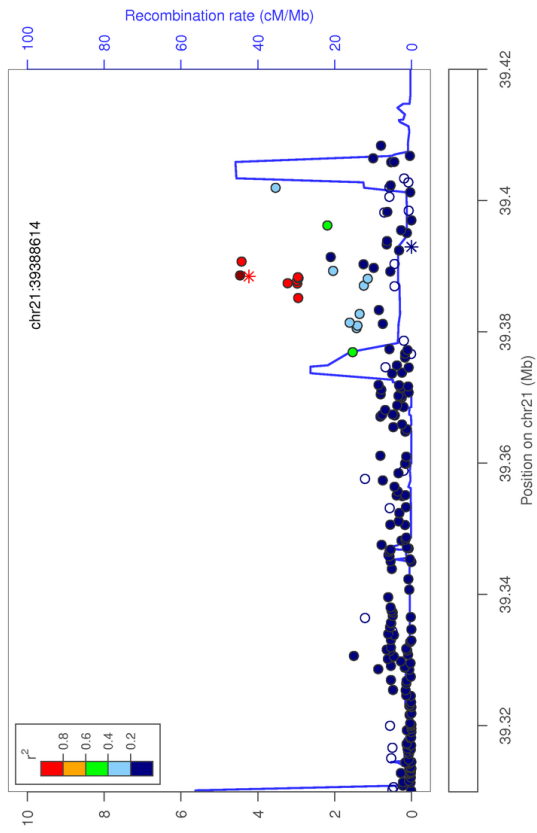


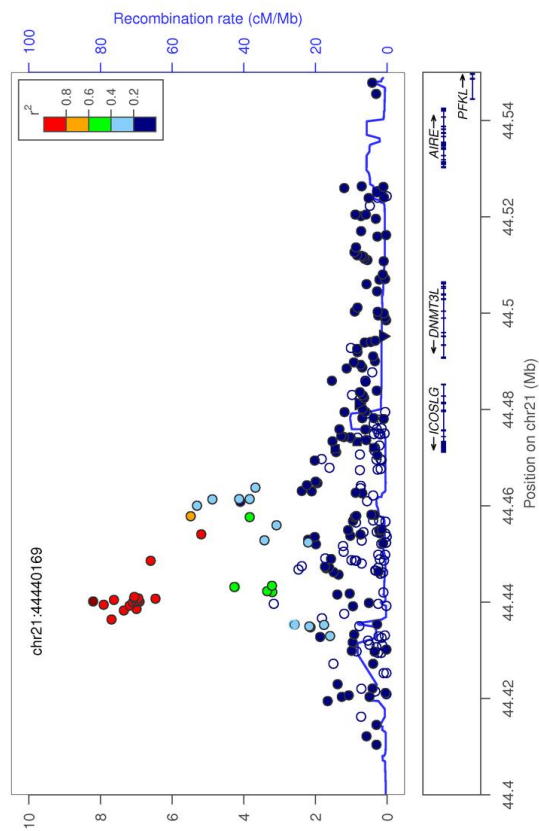
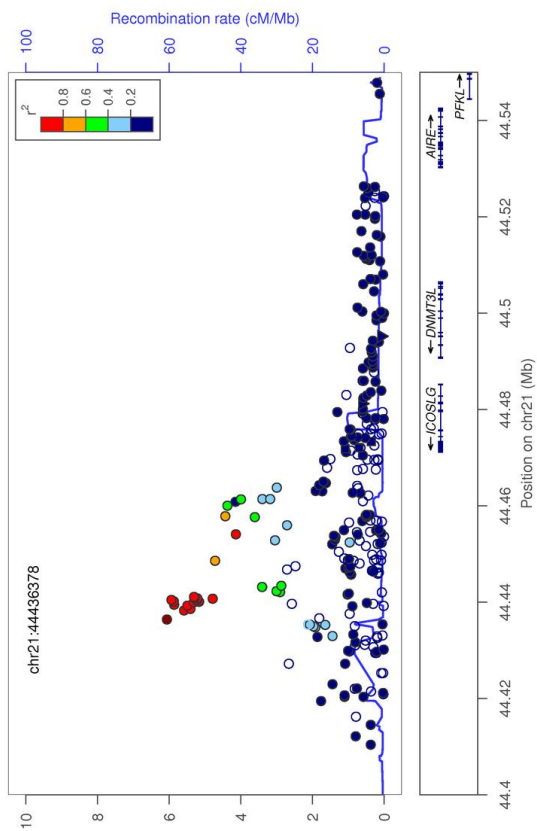
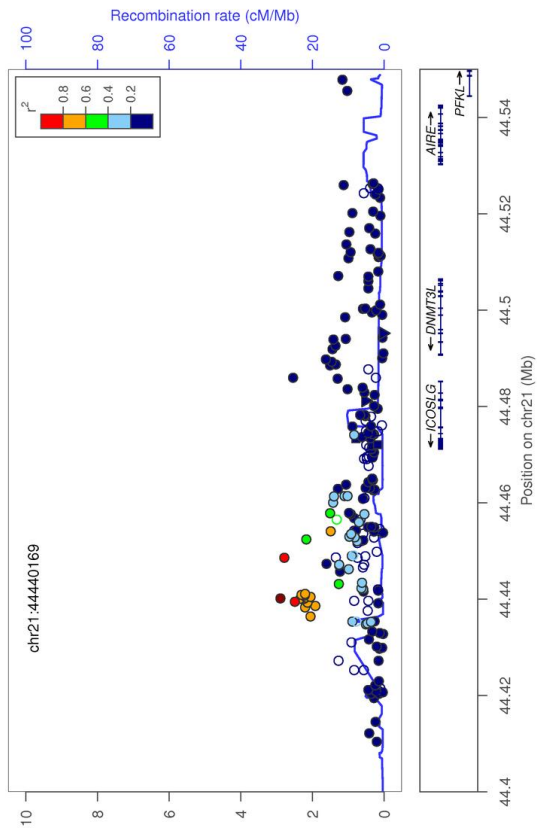


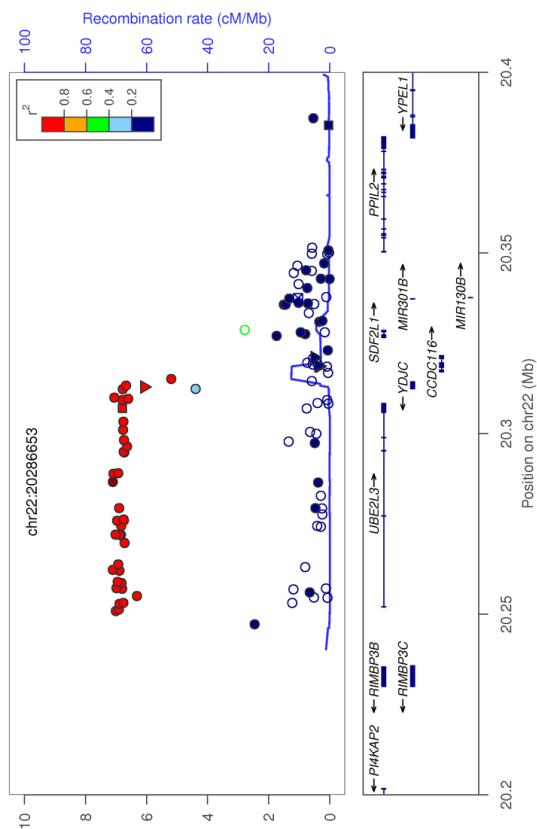
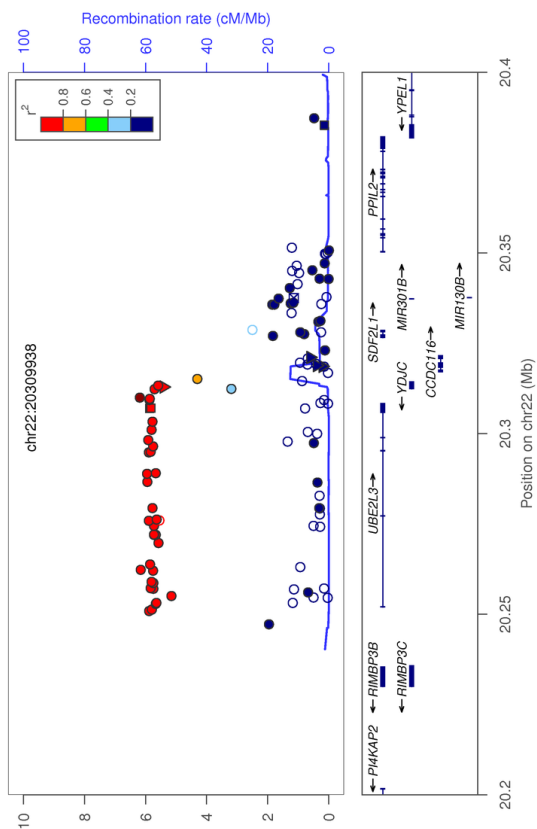
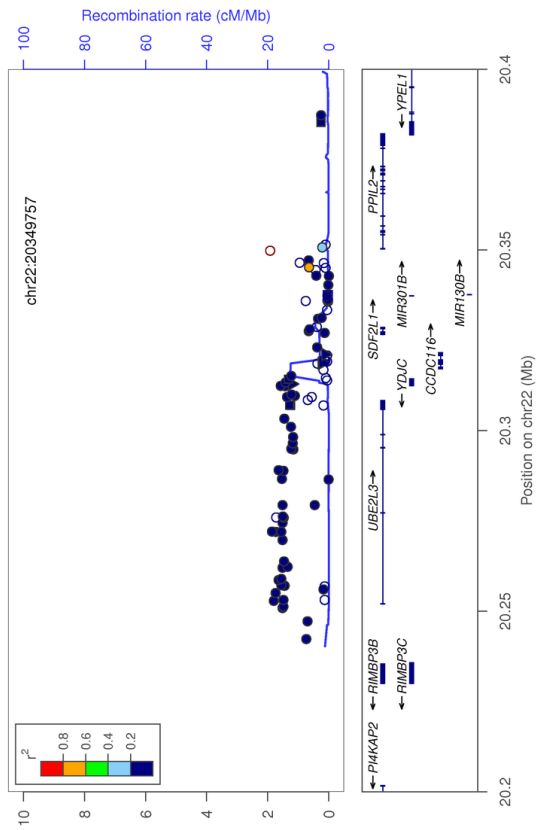




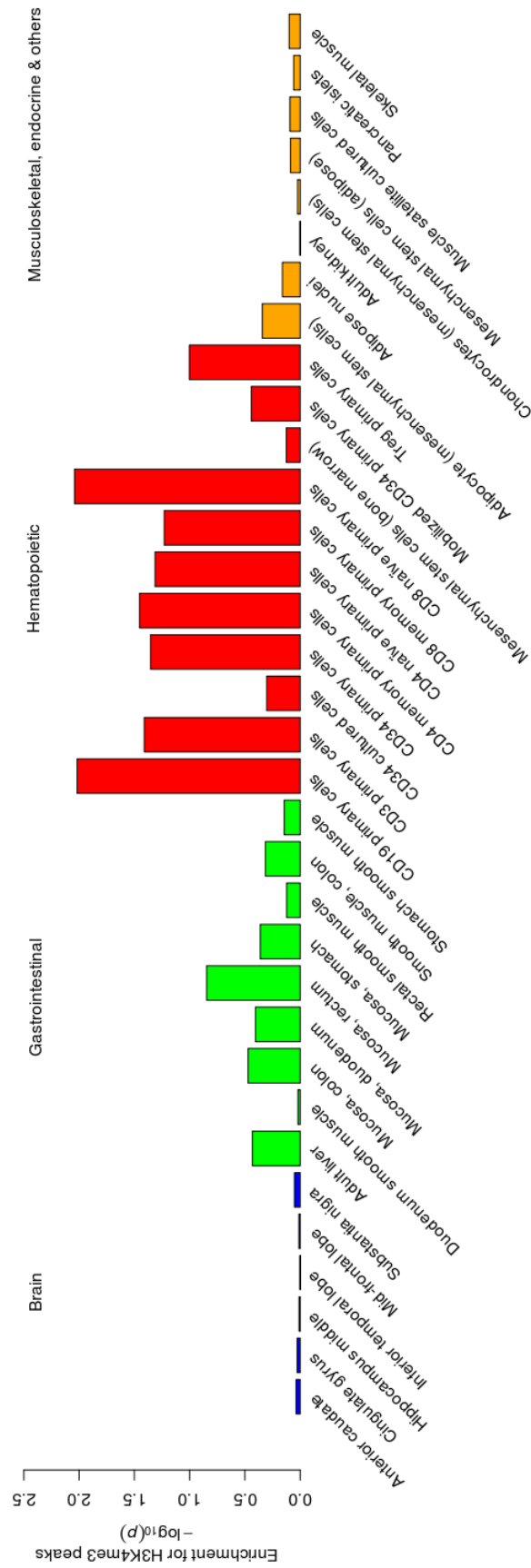




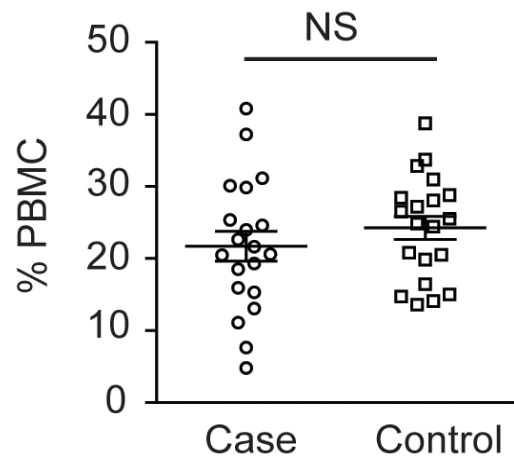




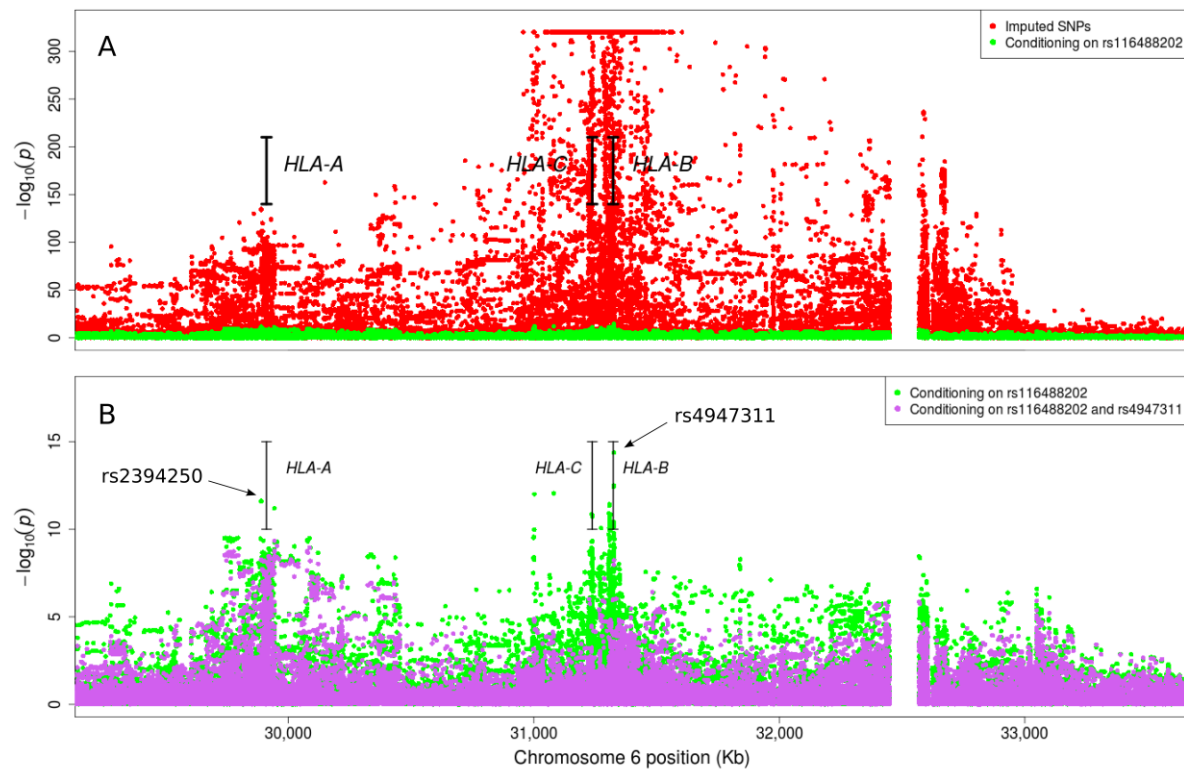
Supplementary Figure 6 Cell type-specificity of association of SNPs associated with AS with H3K4me3 chromatin marks using ENCODE cell line data. AS-associated SNPs are enriched for association specifically with immunological cell types. Analysis software and reference datasets can be obtained from the developer's web-site (<http://www.broadinstitute.org/mpg/epigwas/>).



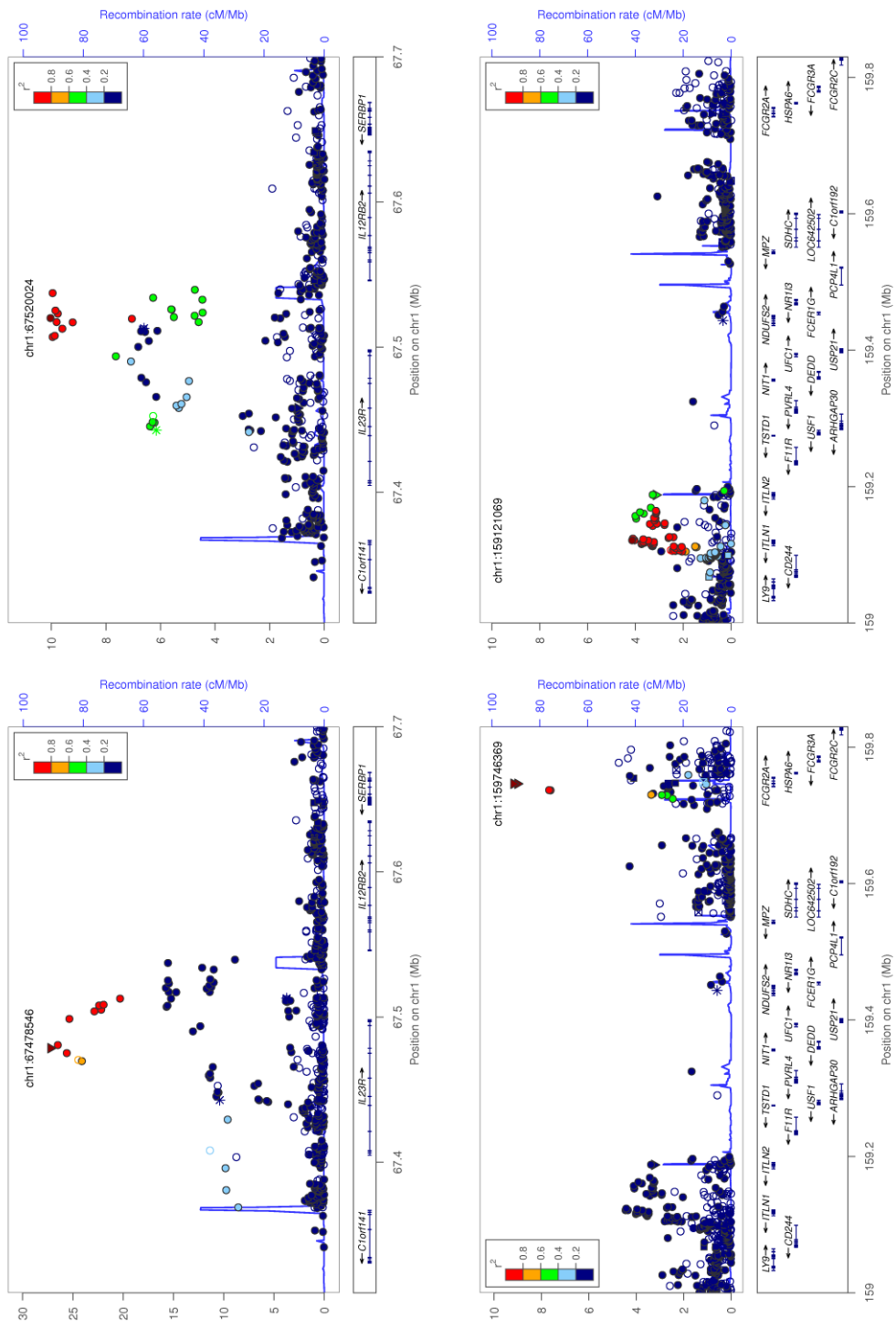
Supplementary Figure 7 CD4⁺ T cells are similar in the peripheral blood of AS patients and healthy controls. PBMC from patients with active AS and age-matched healthy controls (n = 20 for both groups) were stained for cell surface T cell markers and analysed by flow cytometry.

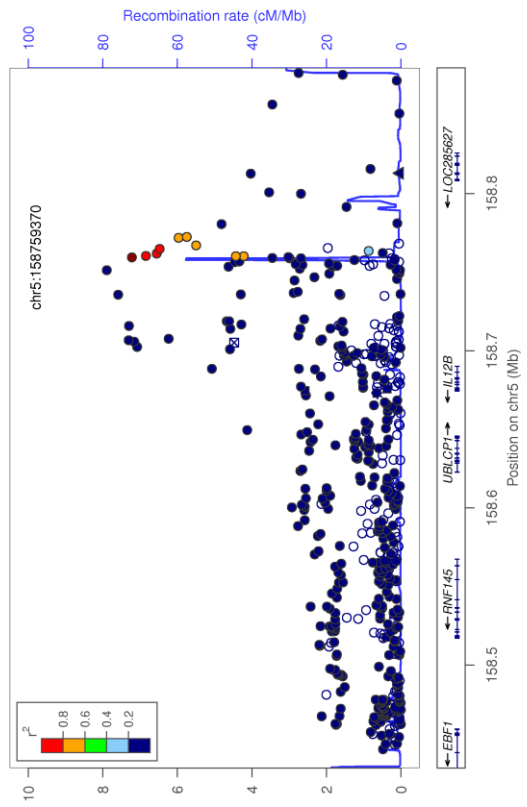
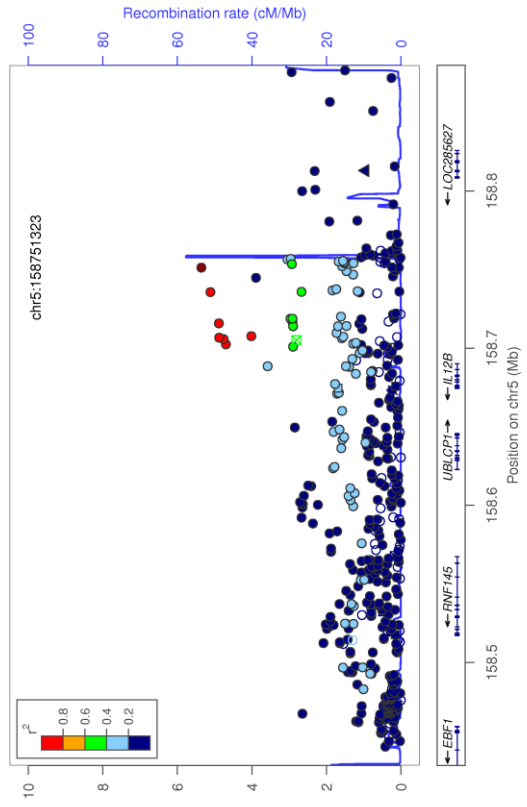
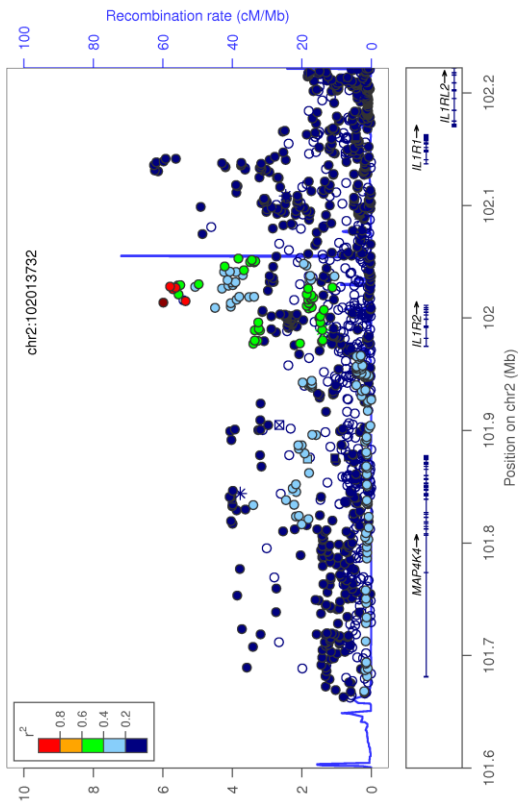
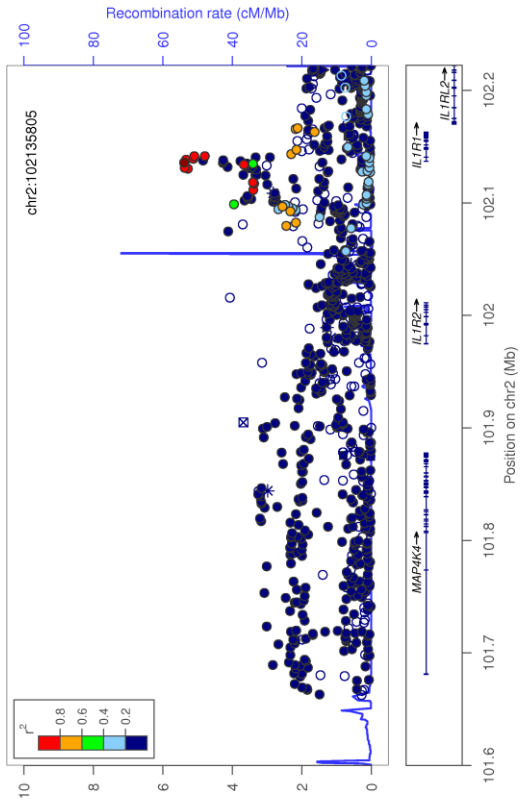


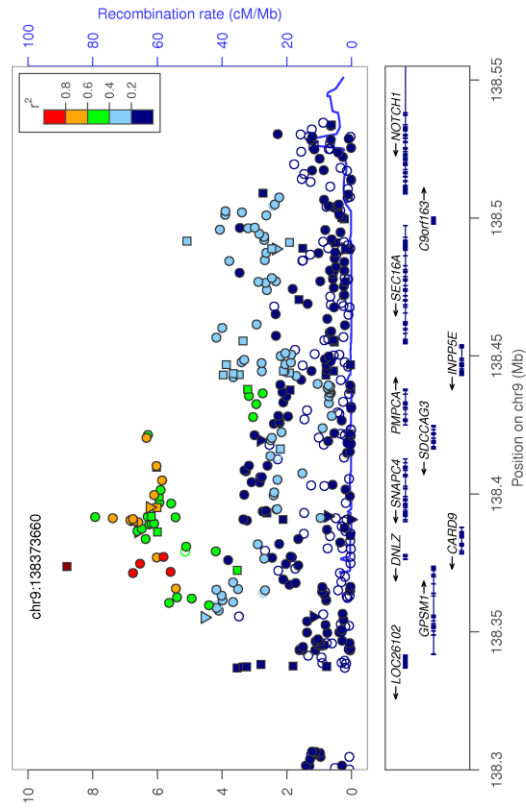
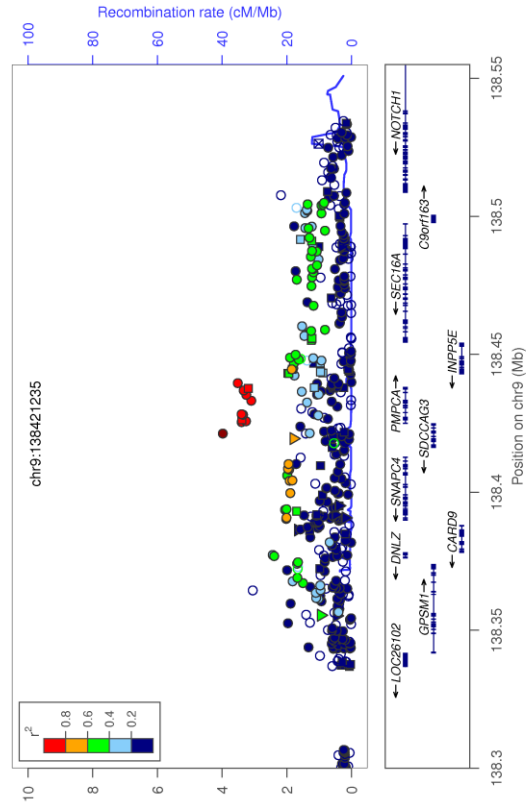
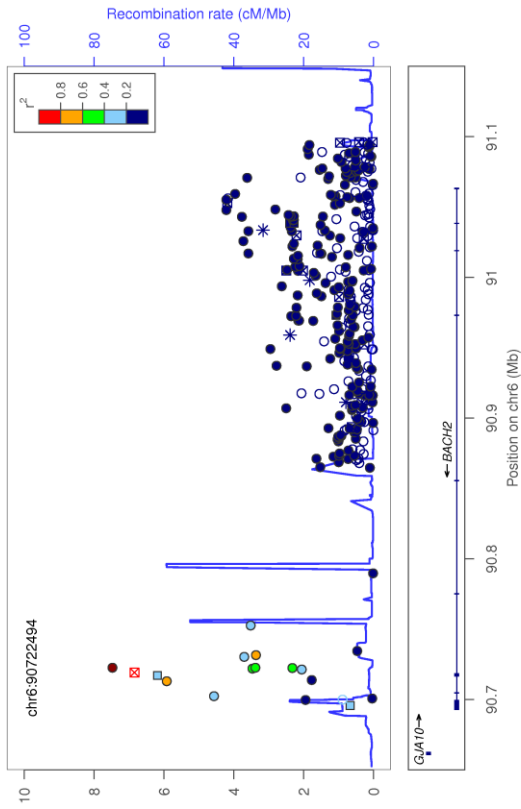
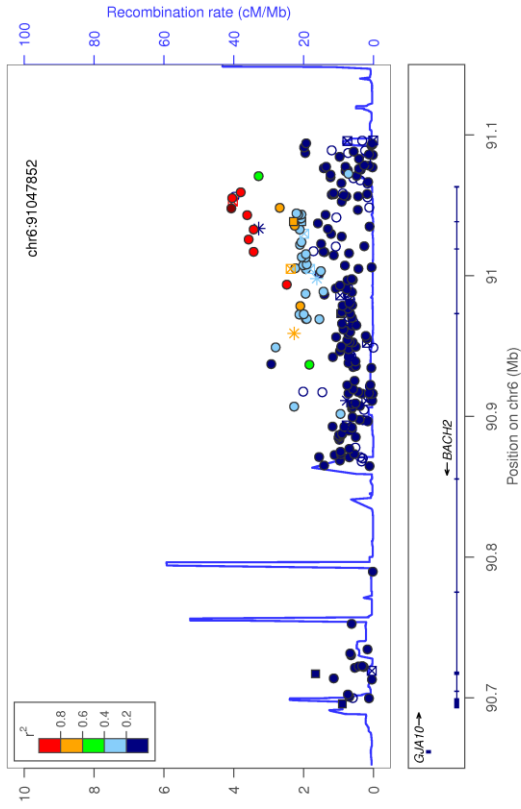
Supplementary Figure 8 Evidence of association to AS susceptibility in samples of European ancestry in the MHC region. A) Associations over the MHC correcting for ancestry with five principal components without and with conditioning on HLA-B*27 tagSNP rs116488202. SNS with P -value $< 10^{-320}$ are set P -value = 1×10^{-320} . B) Association at rs2394250, HLA-A*02 tagSNP, after further conditioning on HLA-B residual association signal rs4947311. The 85kb gap between positions 32,465kb and 32,550kb corresponds to an assembly correction between NCBI human genome build 36 and 37.

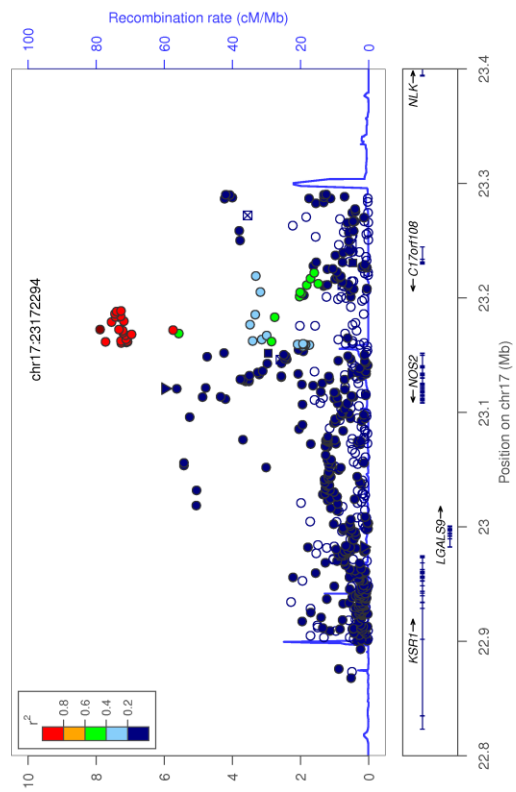
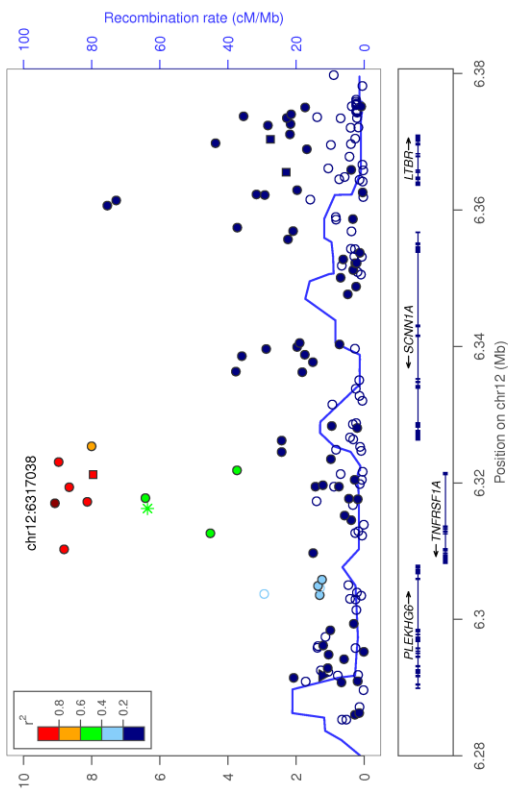
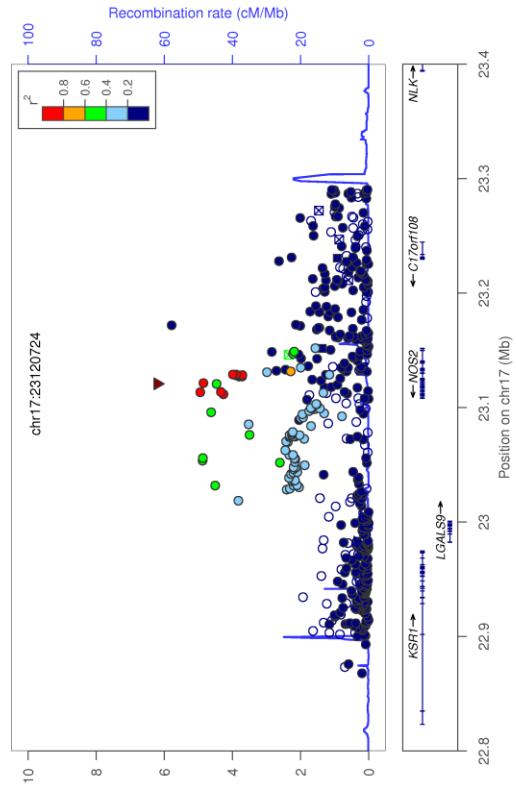
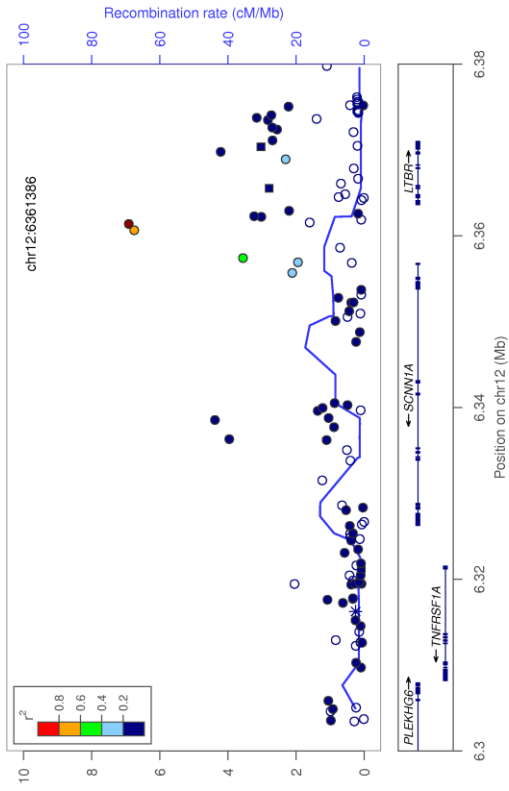


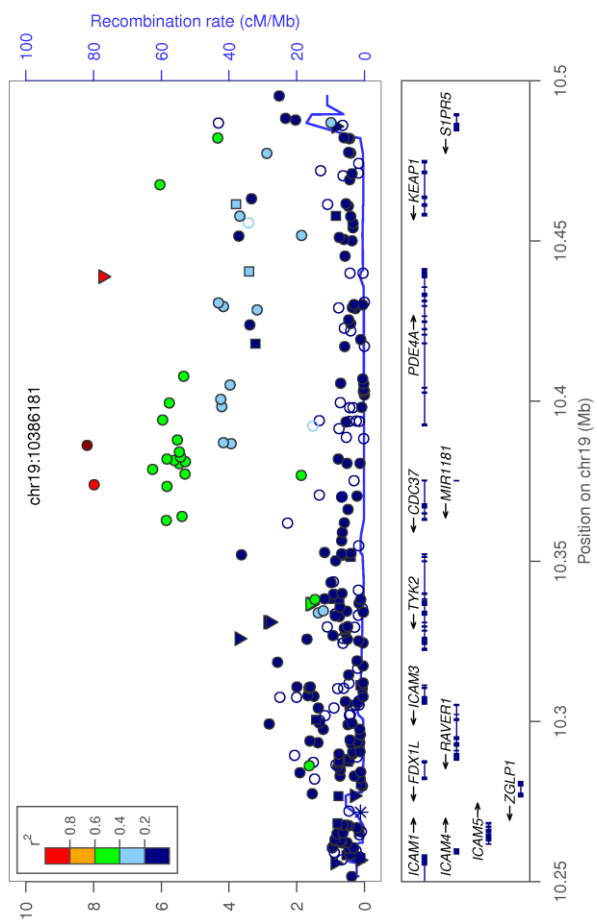
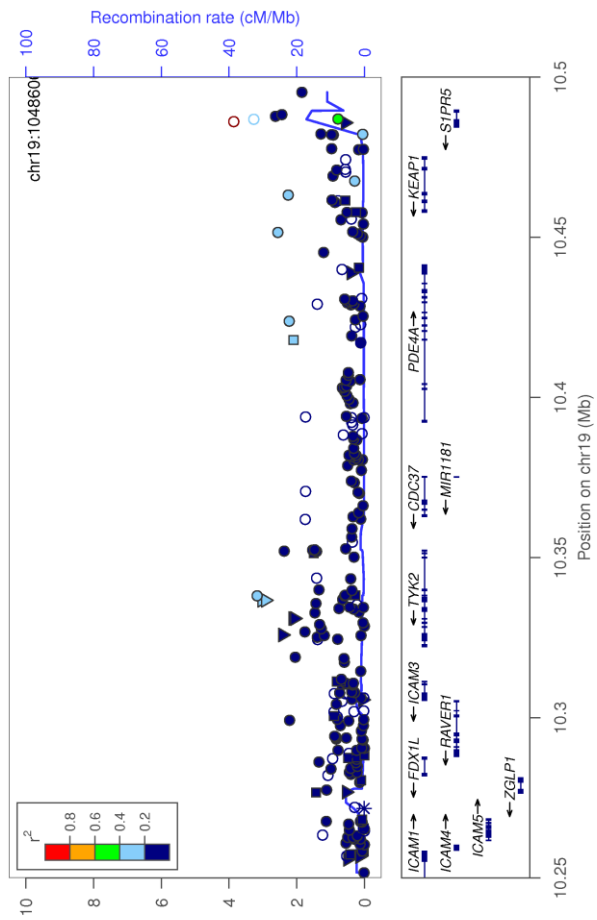
Supplementary Figure 9 Locus association plots and linkage disequilibrium patterns with lead SNP in unconditional and conditional analysis for each of the loci where an independent signal (P -value $< 5 \times 10^{-4}$) was observed for the European cohort analysis.



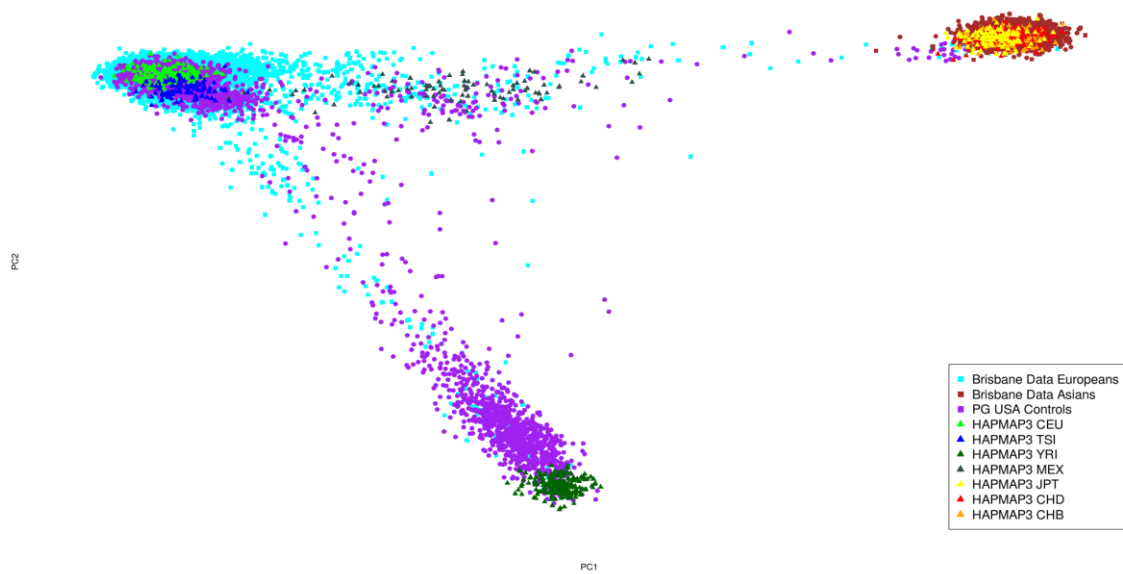




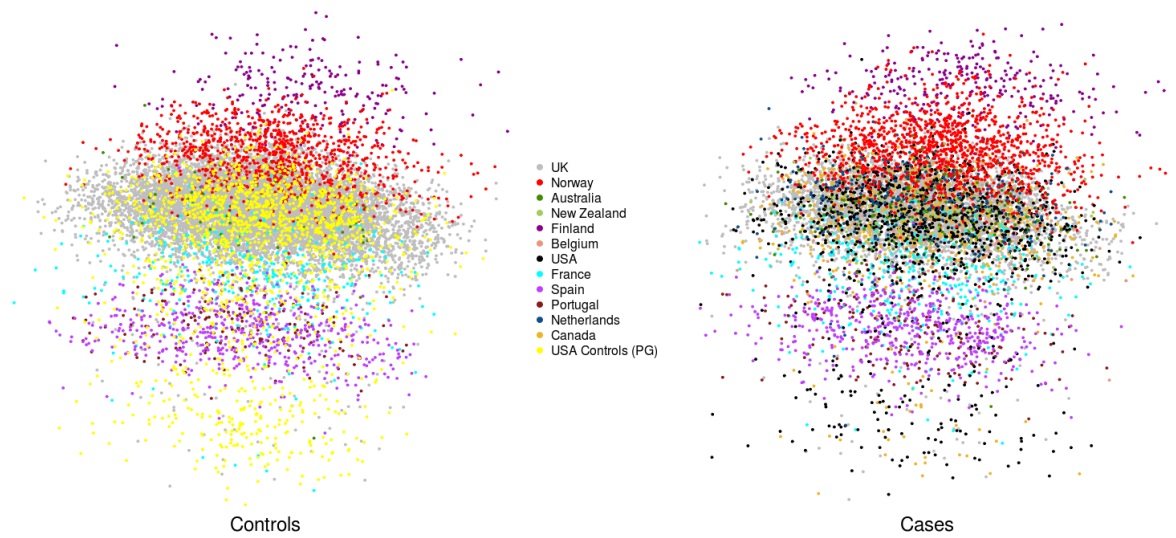




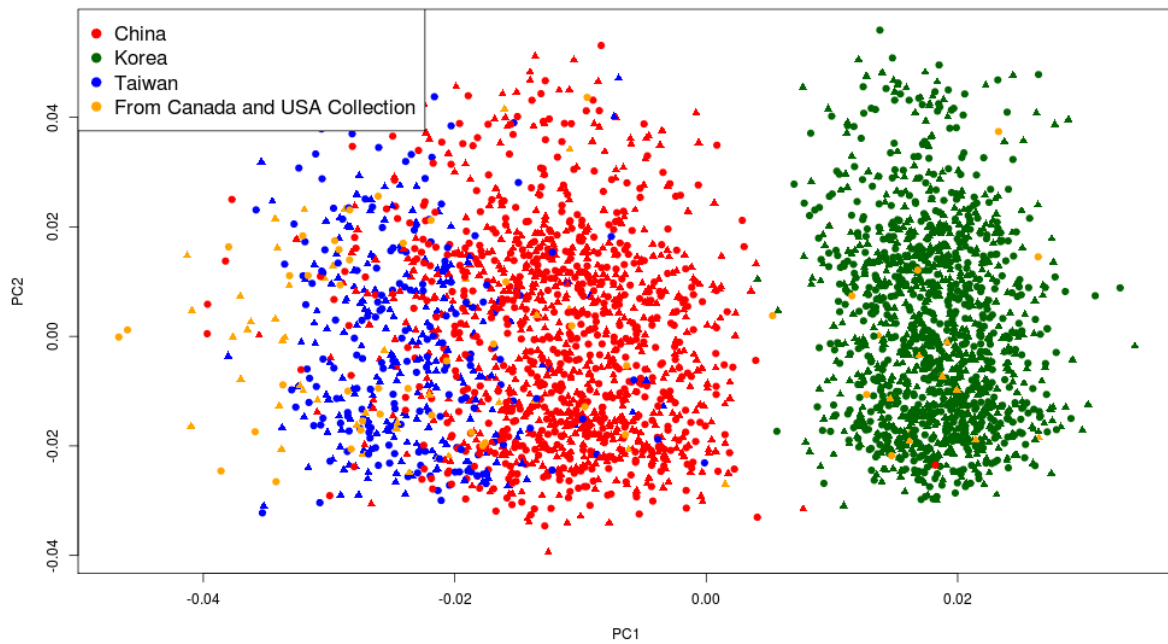
Supplementary Figure 10 Principal components analysis of ImmunoChip samples with 7 HapMap populations. Individuals are plotted for the first two principal components.



Supplementary Figure 11 Principal components analysis of ImmunoChip samples of European ancestry. Individuals are plotted for the first two principal components. Principal component one (plotted on the y-axis) correlates with the North-South gradient of European origin of individuals.



Supplementary Figure 12 Principal components analysis of ImmunoChip samples of East Asian ancestry.



Supplementary Note

Membership of Wellcome Trust Case Control Consortium 2 (WTCCC2)

Management Committee

Peter Donnelly (Chair)^{1,2}, Ines Barroso (Deputy Chair)³, Jenefer M Blackwell^{4,5}, Elvira Bramon⁶, Matthew A Brown⁷, Juan P Casas⁸, Aiden Corvin⁹, Panos Deloukas³, Audrey Duncanson¹⁰, Janusz Jankowski¹¹, Hugh S Markus¹², Christopher G Mathew¹³, Colin NA Palmer¹⁴, Robert Plomin¹⁵, Anna Rautanen¹, Stephen J Sawcer¹⁶, Richard C Trembath¹³, Ananth C Viswanathan¹⁷, Nicholas W Wood¹⁸

Data and Analysis Group

Chris C A Spencer¹, Gavin Band¹, Céline Bellenguez¹, Colin Freeman¹, Garrett Hellenthal¹, Eleni Giannoulatou¹, Matti Pirinen¹, Richard Pearson¹, Amy Strange¹, Zhan Su¹, Damjan Vukcevic¹, Peter Donnelly^{1,2}

DNA, Genotyping, Data QC and Informatics Group

Cordelia Langford³, Sarah E Hunt³, Sarah Edkins³, Rhian Gwilliam³, Hannah Blackburn³, Suzannah J Bumpstead³, Serge Dronov³, Matthew Gillman³, Emma Gray³, Naomi Hammond³, Alagurevathi Jayakumar³, Owen T McCann³, Jennifer Liddle³, Simon C Potter³, Radhi Ravindrarajah³, Michelle Ricketts³, Matthew Waller³, Paul Weston³, Sara Widaa³, Pamela Whittaker³, Ines Barroso³, Panos Deloukas³.

Publications Committee

Christopher G Mathew (Chair)¹³, Jenefer M Blackwell^{4,5}, Matthew A Brown⁷, Aiden Corvin⁹, Chris C A Spencer¹

1 Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK; 2 Dept Statistics, University of Oxford, Oxford OX1 3TG, UK; 3 Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; 4 Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, 100 Roberts Road, Subiaco, Western Australia 6008; 5 Cambridge Institute for Medical Research, University of Cambridge School of Clinical Medicine, Cambridge CB2 0XY, UK; 6 Department of Psychosis Studies, NIHR Biomedical Research Centre for Mental Health at the Institute of Psychiatry, King's College London and The South London and Maudsley NHS Foundation Trust, Denmark Hill, London SE5 8AF, UK; 7 University of Queensland Diamantina Institute, Brisbane, Queensland, Australia; 8 Dept Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT and Dept Epidemiology and Public Health, University College London WC1E 6BT, UK; 9 Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine, Trinity College Dublin, Dublin 2, Eire; 10 Molecular and Physiological Sciences, The Wellcome Trust, London NW1 2BE; 11 Department of Oncology, Old Road Campus, University of Oxford, Oxford OX3 7DQ, UK, Digestive Diseases Centre, Leicester Royal Infirmary, Leicester LE7 7HH, UK and Centre for Digestive Diseases, Queen Mary University of London, London E1 2AD, UK; 12 Clinical Neurosciences, St George's University of London, London SW17 0RE; 13 King's College London Dept Medical and Molecular Genetics, King's Health Partners, Guy's Hospital, London SE1 9RT, UK; 14 Biomedical Research Centre, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK; 15 King's College London Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Denmark Hill, London SE5 8AF, UK; 16 University of Cambridge Dept Clinical Neurosciences, Addenbrooke's

Hospital, Cambridge CB2 0QQ, UK; 17NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London EC1V 2PD, UK; 18Dept Molecular Neuroscience, Institute of Neurology, Queen Square, London WC1N 3BG, UK.

Membership of Spondyloarthritis Research Consortium of Canada (SPARCC)

Sample Collection

Dafna Gladman^{1,2,3}, Nigil Haroon⁴, Robert Inman⁴, Walter Maksymowych⁵, and Proton Rahman⁶.

1 Division of Rheumatology, University of Toronto, Toronto, Canada; 2 Toronto Western Research Institute, Toronto, Canada; 3 Psoriatic Arthritis Program, University Health Network; 4 Division of Rheumatology, Toronto Western Hospital, Toronto, Canada; 5 Department of Medicine, University of Alberta, Canada; 6 Memorial University of Newfoundland, Newfoundland, Canada.

Membership of Groupe Française d'Etude Génétique des Spondylarthrites (GFEGS)

Sample Collection

Maxime Breban^{1,2}, Gilles Chiochia^{1,2}, Henri-Jean Garchon^{1,2}, Roula Said-Nahal², Félicie Costantino², Ariane Leboime²

1 Institut Cochin, Université Paris-Descartes, CNRS (UMR 8104), INSERM U1016, France; 2 Ambroise Paré Hospital, Assistance Publique Hôpitaux de Paris, Versailles-Saint-Quentin en Yvelines University, Boulogne-Billancourt, France.

Membership of Australo-Anglo-American Spondyloarthritis Consortium (TASC)

Management Committee

John D. Reveille¹, Bryan Paul Wordsworth² and Matthew A. Brown³

Data and Analysis Group

Adrian Cortes³, Philip C. Robinson³, Paul Leo³, Tony J. Kenna³, Matthew A. Brown³

DNA, Genotyping and Data QC

Johanna Hadler³, Katie Cremin³, Karena Pryce³, Jessica Harris³

Sample Collection

Jenny P. Pointon², Tugce Karaderi², Michael Weisman⁴, Michael Ward⁵, Linda A. Bradbury³, Robert D. Inman⁶, John D. Reveille¹, Bryan Paul Wordsworth² and Matthew A. Brown³

1 Rheumatology and Clinical Immunogenetics, University of Texas Health Science Center at Houston, Houston, Texas, USA; 2 National Institute for Health Research (NIHR) Oxford Musculoskeletal Biomedical Research Unit, Nuffield Orthopaedic Centre, Headington, Oxford, United Kingdom; 3 University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Australia; 4 Department of Medicine/Rheumatology, Cedars-Sinai Medical Center, Los Angeles, California, USA; 5 National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland, USA; 6 Division of Rheumatology, Toronto Western Hospital, Toronto, Canada.

Membership of Nord-Trøndelag health study (HUNT)

Sample Collection

Vibeke Videm^{1,2}

1 Department of Laboratory Medicine, Children's and Women's Health, Norwegian University of Science and Technology, Trondheim, Norway; 2 Department of Immunology and Transfusion Medicine, Trondheim University Hospital, Norway.

SNP effect transformation from the linear model to the liability scale to calculate SNP disease odds ratios

SNP association analysis was performed with a linear mixed model with the package FaST-LMM. For each SNP, the output of FaST-LMM includes the effect size and its significance. In the linear mixed model the disease status is coded as y_{01} (controls 0; cases 1) and the genotype x is coded with the number of minor alleles an individual has (0, 1 or 2):

$$y_{01} = \mu + b_{01} \cdot x + e.$$

This model can be rewritten, solving for x , as:

$$x = \mu + a \cdot y_{01} + e,$$

which test the mean difference of x between cases and controls.

The SNP effect b_{01} is given by the formula:

$$b_{01} = \text{cov}(y_{01}, x) / \text{var}(x),$$

and equivalently:

$$a = \text{cov}(y_{01}, x) / \text{var}(y_{01}).$$

Substituting $\text{cov}(y_{01}, x)$ it follows that:

$$a = b_{01} \cdot \text{var}(x) / \text{var}(y_{01}).$$

If we let p_1 be the minor allele frequency in cases and p_0 in controls, by definition:

$E(x | y_{01} = 1) = \mu + a = 2 \cdot p_1$ and $E(x | y_{01} = 0) = \mu = 2 \cdot p_0$, so it follows that:

$$a = 2 \cdot (p_1 - p_0).$$

The odds ratio (OR) for disease is given by the formula:

$$\text{OR} = [p_1(1 - p_0)] / [p_0(1 - p_1)].$$

Rearranging terms and substituting $\text{var}(x) = p(1-p)$ and $\text{var}(y_{01}) = v(1-v)$, where v is the proportion of cases, the OR is given by:

$$\text{OR} = 1 + b_{01} \cdot p(1-p) / [v(1-v)p_2(1-p_1)].$$