Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis: supplementary tables and figures

George F. Mells*, James A.B. Floyd*, Katherine I. Morley*,
Heather J. Cordell, Christopher S. Franklin, So-Youn Shin,
Michael A. Heneghan, James M. Neuberger, Peter T. Donaldson,
Darren B. Day, Samantha J. Ducker, Agnes W. Muriithi,
Elizabeth F. Wheater, Christopher J. Hammond, Muhammad F. Dawwas,
The UK PBC Consortium, The Wellcome Trust Case Control Consortium 3,
David E. Jones, Leena Peltonen, Graeme J. Alexander,
Richard N. Sandford# and Carl A. Anderson#

Contents

Supplementary Tables **2**

Supplementary Figures **9**

Consortium Membership **33**
The UK PBC Consortium 33
The WTCCC3 Consortium 35

List of Tables

1 Individuals failing quality control thresholds 2
2 SNPs failing quality control thresholds 2
3 Significance and odds ratio estimates for all SNPs taken forward for replication. 3
4 Comparison of most significant genotyped and imputed SNPs 4
5 Interaction between the most significant SNP for each locus and HLA region 5
6 Results of conditional analysis for the *SCHIP1-IL12A* locus at 3q25 5
7 Overlap with loci associated with other autoimmune diseases 6
8 GRAIL results for all novel PBC loci 7
9 Non-synonymous SNP analysis using 1,000 Genomes data 8
10 Identification of cis-eQTL within all associated loci 8

List of Figures

1 Quantile-quantile plot of observed *versus* expected p-values 9
2 Manhattan plot of association results from imputed data 10
3 Linkage disequilibrium for *SCHIP1-IL12A* locus at 3q25 11
4 Regional association plots of all genome-wide significant loci 12
5 Plot of first two principal components from the GWAS samples and HapMap data 23
6 Heterozygosity and missingness for analysed SNPs 24
7 Cluster plots for all SNPs taken forward for replication 25

Supplementary Tables

Table 1: Individuals failing quality control thresholds. WTCCC2 indicates sample failed QC criteria for the Wellcome Trust Case-Control Consortium 2 study.

Sample	Criteria						Total[a]
	Heterozygosity	Call rate	Gender	Non-European	Relatedness	WTCCC2	
PBC cases	28	15	5	37	21	N/A	81
NBS controls	42	54	12	13	51	236	249
58C controls	68	84	5	44	23	255	231

[a] Some individuals failed quality control for multiple reasons.

Table 2: SNPs failing quality control thresholds. HWE indicates Hardy-Weinberg Equilibrium test; MAF indicates minor allele frequency; NMCAR indicates not missing completely at random. WTCCC2 indicates sample failed QC criteria for the Wellcome Trust Case-Control Consortium 2 study.

Sample	Criteria					Total[a]
	Call rate	HWE	Low MAF	NMCAR	WTCCC2	
PBC cases	9,022	10,164	67,387	5,217	N/A	79,941
NBS controls	16,958	15,167	195,422	11,545	214,848	237,455
58C controls	15,317	14,721	196,294	12,209	215,732	239,454

[a] Some SNPs failed quality control for multiple reasons.

Table 3: Significance and odds ratio estimates for all SNPs taken forward for replication.

SNP	CHR	Position (bp)	P-value	OR	95% C.I.
rs11586136	1	101008725	8.35×10^{-6}	1.27	(1.14-1.41)
rs1539414	1	196010129	2.32×10^{-9}	1.31	(1.20-1.43)
rs12134279	1	196047821	1.07×10^{-9}	1.32	(1.21-1.44)
rs4952108	2	30283239	1.16×10^{-7}	1.32	(1.19-1.46)
rs3738863	2	30311065	1.82×10^{-6}	1.23	(1.13-1.35)
rs11692152	2	153429754	2.12×10^{-6}	1.29	(1.16-1.44)
rs16831902	2	153441617	3.21×10^{-6}	1.29	(1.16-1.43)
rs2286896	2	191243821	1.09×10^{-11}	1.44	(1.29-1.59)
rs10931468	2	191246807	2.55×10^{-12}	1.46	(1.31-1.62)
rs322684	3	25352227	8.71×10^{-6}	1.22	(1.12-1.32)
rs1483831	3	25358670	8.34×10^{-6}	1.22	(1.12-1.33)
rs12494314	3	120605510	2.08×10^{-10}	1.40	(1.26-1.56)
rs2293370	3	120702624	7.70×10^{-11}	1.41	(1.27-1.56)
rs7665090	4	103770651	5.33×10^{-7}	1.21	(1.13-1.31)
rs2866413	4	103776125	5.40×10^{-7}	1.21	(1.13-1.31)
rs6897932	5	35910332	3.16×10^{-10}	1.33	(1.22-1.45)
rs860413	5	35978799	3.09×10^{-10}	1.33	(1.22-1.45)
rs26232	5	102624619	5.87×10^{-6}	1.21	(1.11-1.31)
rs674726	5	102705263	8.15×10^{-6}	1.20	(1.11-1.31)
rs6974491	7	37341035	3.39×10^{-6}	1.25	(1.14-1.38)
rs7045605	9	8252469	1.28×10^{-6}	1.25	(1.14-1.37)
rs192705	9	85781524	3.25×10^{-6}	1.20	(1.11-1.30)
rs4743150	9	99779945	9.23×10^{-6}	1.23	(1.12-1.35)
rs874610	9	99822516	9.34×10^{-6}	1.23	(1.12-1.35)
rs4938573	11	118247052	5.02×10^{-10}	1.39	(1.25-1.54)
rs6421571	11	118248982	3.53×10^{-10}	1.40	(1.26-1.55)
rs1800693	12	6310270	5.51×10^{-8}	1.23	(1.14-1.33)
rs4149581	12	6317246	8.27×10^{-7}	1.21	(1.12-1.31)
rs3213989	12	41150068	6.64×10^{-8}	1.24	(1.14-1.34)
rs4768412	12	41155407	3.70×10^{-8}	1.24	(1.15-1.34)
rs11066188	12	111095097	3.19×10^{-6}	1.20	(1.11-1.29)
rs9594738	13	41850145	6.86×10^{-6}	1.19	(1.10-1.28)
rs9533108	13	41922710	2.97×10^{-6}	1.20	(1.11-1.29)
rs2208397	14	67823040	5.71×10^{-9}	1.30	(1.19-1.42)
rs911263	14	67823346	1.68×10^{-9}	1.31	(1.20-1.43)
rs8017161	14	102632948	4.71×10^{-6}	1.20	(1.11-1.29)
rs725613	16	11077184	3.13×10^{-10}	1.30	(1.20-1.41)
rs12924729	16	11095284	7.68×10^{-11}	1.32	(1.21-1.44)
rs11117432	16	84576772	2.00×10^{-6}	1.26	(1.15-1.39)
rs8070723	17	41436901	5.96×10^{-6}	1.24	(1.13-1.35)
rs618671	18	352199	7.04×10^{-6}	1.19	(1.10-1.28)
rs963168	20	1586578	1.05×10^{-6}	1.27	(1.15-1.39)
rs6043722	20	1609245	2.39×10^{-6}	1.38	(1.20-1.57)
rs2831525	21	28427528	8.31×10^{-6}	1.19	(1.10-1.28)
rs968451	22	38000797	4.31×10^{-7}	1.27	(1.16-1.39)
rs1003643	22	38006440	5.83×10^{-7}	1.24	(1.14-1.35)

Table 4: Comparison of most significant genotyped and imputed SNPs for each region reaching genome-wide significance in the combined analysis.

CHR	Region (Mb)	Genotyped data			Imputed data		
		SNP	Position (bp)	P-value	SNP	Position (bp)	P-value
1	2.39 - 2.78	rs10752747	2514775	2.65×10^{-3}	rs2843404	2520418	2.45×10^{-3}
1	67.53 – 67.71	rs17129789	67563186	9.48×10^{-20}	rs11209050	67564324	9.60×10^{-19}
1	195.58 - 196.21	rs12134279	196047821	1.07×10^{-9}	rs16841904	195968615	2.07×10^{-9}
2	190.77 - 191.61	rs10931468	191246807	2.55×10^{-12}	rs3771317	191252207	1.61×10^{-12}
3	16.82 - 17.13	rs1372072	16930263	1.38×10^{-4}	rs11928330	16933371	4.07×10^{-5}
3	120.58 - 120.79	rs2293370	120702624	7.70×10^{-11}	rs3732421	120632779	2.37×10^{-10}
3	160.96 - 161.3	rs485499	161228557	2.23×10^{-16}	rs564799	161211681	8.19×10^{-16}
4	103.61 - 104.24	rs7665090	103770651	5.33×10^{-7}	rs1054037	103771757	4.91×10^{-7}
5	35.74 - 36.08	rs860413	35978799	3.09×10^{-10}	rs7717955	35898598	3.23×10^{-10}
6	26.21 - 33.74	rs7774434	32765556	3.86×10^{-34}	rs3128966	33163924	3.02×10^{-30}
7	37.32 - 37.41	rs6974491	37341035	3.39×10^{-6}	rs1962401	37343697	1.62×10^{-6}
7	128.33 - 128.57	rs12531711	128404702	8.90×10^{-17}	rs12539476	128444719	1.47×10^{-16}
11	63.60 - 64.04	rs538147	63886298	1.01×10^{-5}	rs510372	63871713	1.11×10^{-6}
11	117.82 - 118.30	rs6421571	118248982	3.53×10^{-10}	rs10892294	118172567	5.61×10^{-10}
12	6.29 - 6.33	rs1800693	6310270	5.51×10^{-8}	rs11064145	6325359	2.37×10^{-6}
14	67.34 - 67.98	rs911263	67823346	1.68×10^{-9}	rs3784099	67819680	2.37×10^{-9}
14	102.54 - 102.68	rs8017161	102632948	4.71×10^{-6}	rs2297067	102636538	3.74×10^{-12}
16	84.55 - 84.58	rs11117432	84576772	1.20×10^{-6}	rs1568387	84582380	1.98×10^{-4}
16	10.92 - 11.22	rs12924729	11095284	7.68×10^{-11}	rs12935413	11117948	1.85×10^{-10}
17	34.61 - 35.49	rs7208487	34796975	7.89×10^{-7}	rs2879258	34652905	2.79×10^{-7}
19	55.52 - 55.73	rs3745516	55618554	1.63×10^{-13}	rs1726773	55619382	2.73×10^{-7}
22	37.87 - 38.19	rs968451	38000797	4.31×10^{-7}	rs1003644	38006690	3.20×10^{-7}

Table 5: Results for analysis of interaction between the most significant SNP for each genome-wide significant locus and HLA region. P-values are the reported significance for each respective regression coefficient as estimated by PLINK for the full logistic regression model incorporating the SNP in question (SNP), HLA score (HLA) and interaction between the two ($SNP \times HLA$).

SNP	CHR	Position (bb)	P_{SNP}	P_{HLA}	$P_{SNP \times HLA}$
rs10752747	1	2514775	2.6×10^{-3}	8.06×10^{-50}	0.43
rs17129789	1	67563186	8.95×10^{-18}	3.06×10^{-69}	0.41
rs12134279	1	196047821	4.74×10^{-9}	3.45×10^{-76}	0.04
rs10931468	2	191246807	9.10×10^{-10}	6.01×10^{-76}	0.24
rs1372072	3	16930263	2.29×10^{-5}	1.44×10^{-55}	0.25
rs2293370	3	120702624	1.94×10^{-9}	7.87×10^{-77}	0.56
rs485499	3	161228557	8.13×10^{-18}	2.74×10^{-44}	0.38
rs7665090	4	103770651	2.80×10^{-7}	1.89×10^{-44}	0.53
rs860413	5	35978799	7.76×10^{-10}	1.15×10^{-62}	0.67
rs6974491	7	37341035	1.1×10^{-5}	1.15×10^{-79}	0.26
rs12531711	7	128404702	1.46×10^{-15}	3.18×10^{-81}	0.34
rs538147	11	63886298	1.19×10^{-5}	8.83×10^{-51}	0.74
rs6421571	11	118248982	1.72×10^{-9}	4.04×10^{-78}	0.82
rs1800693	12	6310270	1.69×10^{-7}	7.42×10^{-51}	0.15
rs911263	14	67823346	2.93×10^{-8}	7.52×10^{-68}	0.57
rs8017161	14	102632948	8.68×10^{-6}	1.79×10^{-47}	0.53
rs11117432	16	84576772	2.439×10^{-6}	1.28×10^{-73}	0.59
rs12924729	16	11095284	3.78×10^{-9}	2.38×10^{-64}	0.24
rs7208487	17	34796975	1.55×10^{-5}	7.30×10^{-85}	0.47
rs3745516	19	55618554	1.18×10^{-12}	2.06×10^{-67}	0.42
rs968451	22	38000797	2.17×10^{-5}	1.39×10^{-65}	0.37

Table 6: Results of conditional analysis for the *SCHIP1-IL12A* locus at 3q25. The most significant SNP at this locus is rs485499; analyses of SNPs in this region were performed including this SNP as covariate in a logistic regression model. For each SNP reaching genome-wide significance in the conditional analysis, the significance of the regression coefficient for each SNP is shown for the models with and without rs485499 included (P_{GWAS} and $P_{Conditional}$ respectively). Linkage disequilibrium between rs485499 and each SNP is shown in the r^2 column; see Supplementary Figure 3 for a graphical representation of the LD between these SNPs.

SNP	Position (bp)	P_{GWAS}	$P_{Conditional}$	r^2	Type
rs4525910	161129306	2.18×10^{-9}	2.73×10^{-8}	0.001	Imputed
rs7610160	161169039	2.61×10^{-9}	1.73×10^{-8}	0.006	Genotyped
rs17811014	161170723	6.98×10^{-8}	1.10×10^{-11}	0.016	Genotyped
rs13064168	161174900	3.31×10^{-7}	5.73×10^{-10}	0.011	Imputed
rs2366408	161178793	3.64×10^{-11}	1.14×10^{-8}	0.024	Genotyped

Table 7: Overlap with loci associated with other autoimmune diseases for all novel PBC loci with $P_{Combined} \leq 5 \times 10^{-8}$. Acronyms are as follows: MS, multiple sclerosis; CD, Crohn's disease; SLE, systemic lupus erythematosus; RA, rheumatoid arthritis; Coeliac, Coeliac disease; T1D, type 1 diabetes; UC, ulcerative colitis; PS, psoriasis; AS, ankylosing spondylitis.

CHR	PBC			Other AI Diseases						Reference
	SNP	Region		SNP	Region		Gene	Disease	P-value	
1	rs10752747	2394116	2775531	rs3748816	2396747	2775531	?	Coeliac	3.28×10^{-9}	[6]
1	rs17129789	67533953	67713755	rs2201841	67367147	67540581	*IL23R*	PS	3.00×10^{-8}	[14]
								CD	6.66×10^{-63}	[4]
								UC	1.30×10^{-8}	[15]
				rs11209032	67367147	67541181	*IL23R*	AS	7.50×10^{-9}	[5]
2	rs10931468	190765686	191605082	rs7574865	191581798	191715979	*STAT4*	SLE	5.17×10^{-42}	[8]
3	rs2293370	120581993	120788069	rs11712165	120587671	120783345	*CD80/KTELC1*	Coeliac	8.03×10^{-9}	[6]
							TMEM39A	MS	3.09×10^{-8}	[10]
3	rs485499	160955049	161303447	rs17810546	161068993	161234305	*IL12A*	Coeliac	3.98×10^{-28}	[6]
				rs4680534	161086908	161237201	*IL12A*	MS	3.08×10^{-8}	[10]
5	rs860413	35743657	36078801	rs6897932	35835053	36070623	*IL7R*	MS	1.21×10^{-17}	[9]
7	rs12531711	128326964	128572244	rs2070197	128336804	128564756	*IRF5*	SLE	5.82×10^{-24}	[7]
								RA	4.00×10^{-9}	[16]
11	rs6421571	117820059	118296360	rs4639966	117847131	118270810	Multiple genes	SLE	1.25×10^{-16}	[8]
12	rs1800693	6286255	6334123	rs1800693	6286255	6334123	*TNFRSF1A*	MS	1.59×10^{-11}	[11]
16	rs12924729	10924559	11223140	rs12708716	10924559	11214525	*CLEC16A*	MS	1.60×10^{-16}	[9]
								T1D	2.20×10^{-16}	[2]
								Coeliac	3.12×10^{-8}	[6]
16	rs11117432	84554782	84584334	rs17445836	84551081	84581605	*IRF8*	MS	3.73×10^{-9}	[11]
17	rs7208487	34610560	35493742	rs2872507	34636200	35493742	*ORMDL3*	CD	5.00×10^{-9}	[4]
								T1D	5.50×10^{-13}	[2]

Table 8: GRAIL results for all novel PBC loci with $P_{Combined} \leq 5 \times 10^{-8}$.

GWAS Results			GRAIL Results		
CHR	SNP	Region (Mb)	Gene	P_{text}	Putative function
1	rs12134279	195.58 - 196.21	*DENND1B*	0.01	Associated with asthma.
2	rs10931468	190.77 - 191.61	*STAT4*	1.08×10^{-5}	Mediates responses to IL12 in lymphocytes; regulates T-cell differentiation.
3	rs1372072	16.82 - 17.13	*PLCL2*	0.05	Expressed in skeletal muscle.
3	rs2293370	120.58 - 120.79	*CD80*	1.6×10^{-3}	Provides regulatory signals for T lymphocytes.
4	rs7665090	103.61 - 104.24	*NHEDC2*	0.32	Sodium hydrogen antiporter.
5	rs860413	35.74 - 36.08	*IL7R*	1.5×10^{-3}	Receptor for IL7; involved in activation of T-cells and SCID.
7	rs6974491	37.32 - 37.41	*ELMO1*	1	Promotes phagocytosis.
11	rs538147	63.60 - 64.04	*MACROD1*	0.41	Role in invasion, metastasis of gastric cancer.
11	rs6421571	117.82 - 118.30	*DDX6*	0.10	RNA helicase found in P-bodies and stress granules.
12	rs1800693	6.29 - 6.33	*TNFRSF1A*	8.8×10^{-3}	Activates NF-kappaB, mediates apoptosis, regulator of inflammation.
14	rs911263	67.34 - 67.98	*RAD51L1*	0.24	Involved in DNA repair and recombination.
14	rs8017161	102.54 - 102.68	*TNFAIP2*	0.40	Expression induced by tumor necrosis factor alpha
16	rs12924729	10.92 - 11.22	*CIITA*	4.5×10^{-4}	Positive regulator of class II MHC genes.
16	rs11117432	84.55 - 84.58	*IRF8*	2.0×10^{-3}	Transcription factor of interferon regulatory factors.
22	rs968451	37.87 - 38.19	*SNORD43*	0.76	Noncoding small nucleolar RNA.

Table 9: Results of non-synonymous SNP analysis using 1,000 Genomes data [1] for those SNPs reaching genome-wide significance in the combined analyses. Proxy SNP indicates the non-synonymous SNP used as a proxy for the genome-wide significant SNP identified in the GWA analysis, with r^2 providing an estimate of the linkage disequilibrium between the two SNPs.

CHR	SNP	SNP Position	Proxy SNP	Distance to hit SNP (bp)	r^2	Gene
4q24	rs7665090	103770651	rs2866413	5,474	0.97	*MANBA*
5p13	rs860413	35978799	rs6897932	68,467	0.80	*IL7R*
19q13	rs3745516	55618554	rs11546996	477	1.0	*SPIB*

Table 10: Identification of cis-eQTL within all known associated PBC loci that are in LD ($r^2 > 0.8$) with the most associated SNP at that locus, using data from the Gene Expression Analysis dataset at the University of Michigan Center for Statistical Genetics (`http://www.sph.umich.edu/csg/liang/imputation`). Results where a SNP has $p < 1 \times 10^{-5}$ for an effect on gene expression are reported in the table.

Chromosome	SNP	Gene	eQTL p-value
6p21	rs7774434	*HLA-DRB4*	6.4×10^{-12}
6p21	rs7774434	*HLA-DQA1*	1.7×10^{-10}
6p21	rs7774434	*HLA-DRB1*	2.1×10^{-10}
17q12	rs7208487	*CRKRS*	2.4×10^{-7}
17q12	rs2879258	*CRKRS*	3.2×10^{-7}
17q12	rs11655972	*CRKRS*	3.8×10^{-7}
17q12	rs8073907	*CRKRS*	3.5×10^{-7}
17q12	rs590051	*CRKRS*	3.5×10^{-7}
17q12	rs2302073	*CRKRS*	2.4×10^{-7}
17q12	rs9908131	*CRKRS*	2.4×10^{-7}
17q12	rs9906612	*CRKRS*	2.4×10^{-7}
17q12	rs6503513	*CRKRS*	2.3×10^{-8}
17q12	rs9646419	*CRKRS*	1.2×10^{-7}
17q12	rs12449852	*CRKRS*	2.2×10^{-7}
17q12	rs8069074	*CRKRS*	1.1×10^{-7}
17q12	rs7503377	*CRKRS*	1.9×10^{-7}
22q13	rs968451	*SYNGR1*	1.9×10^{-9}
22q13	rs5757611	*SYNGR1*	1.1×10^{-12}
22q13	rs2014842	*SYNGR1*	4.3×10^{-10}
22q13	rs2076125	*SYNGR1*	1.1×10^{-12}
22q13	rs12627970	*SYNGR1*	2.8×10^{-13}

Supplementary Figures

Figure 1: Quantile-quantile plot of observed *versus* expected p-values. Blue points indicate the results of the complete GWAS. Yellow points indicate the results after all known loci (including the HLA region) are removed. Red points indicate the results after all known loci and the genome-wide significant loci identified in this study are removed. $\hat{\lambda}_{GC}$ was estimated from the complete GWAS data.

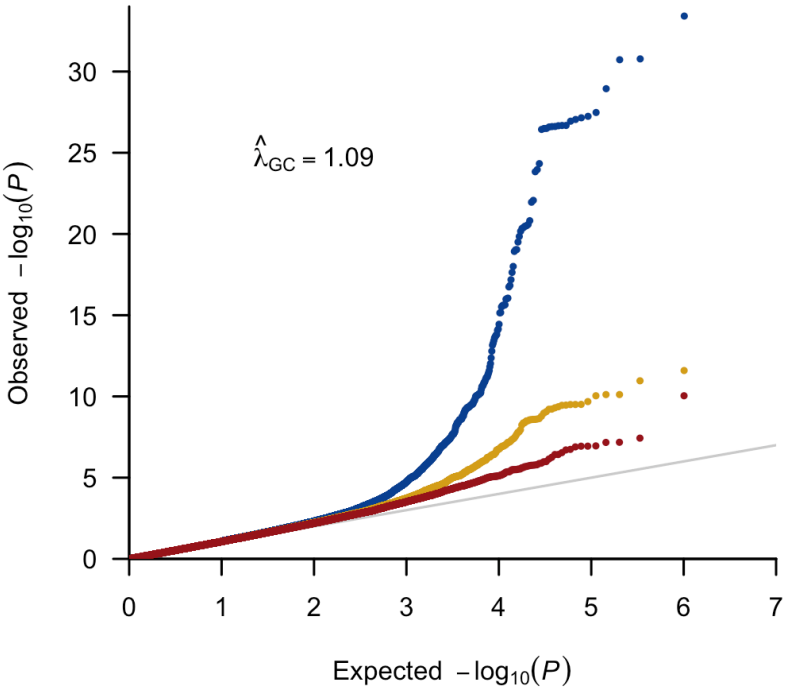


Figure 2: Manhattan plot of association results from data imputed using the HapMap3 Indo-European dataset. Imputation was performed using only those markers that passed QC and were genotyped in cases and controls. Analysis of the imputed data was conducted for those SNPs with a quality score (predicted dosage r^2) > 0.4 in PLINK version 1.07 (Purcell et al., 2007). Gold points indicate SNPs with a suggestive ($\leq 1 \times 10^{-5}$) level of association, red points indicate SNPs with a significant ($\leq 5 \times 10^{-8}$) level of association. Filled dots indicate SNPs for which only missing data were imputed, empty dots indicate completely imputed SNPs.

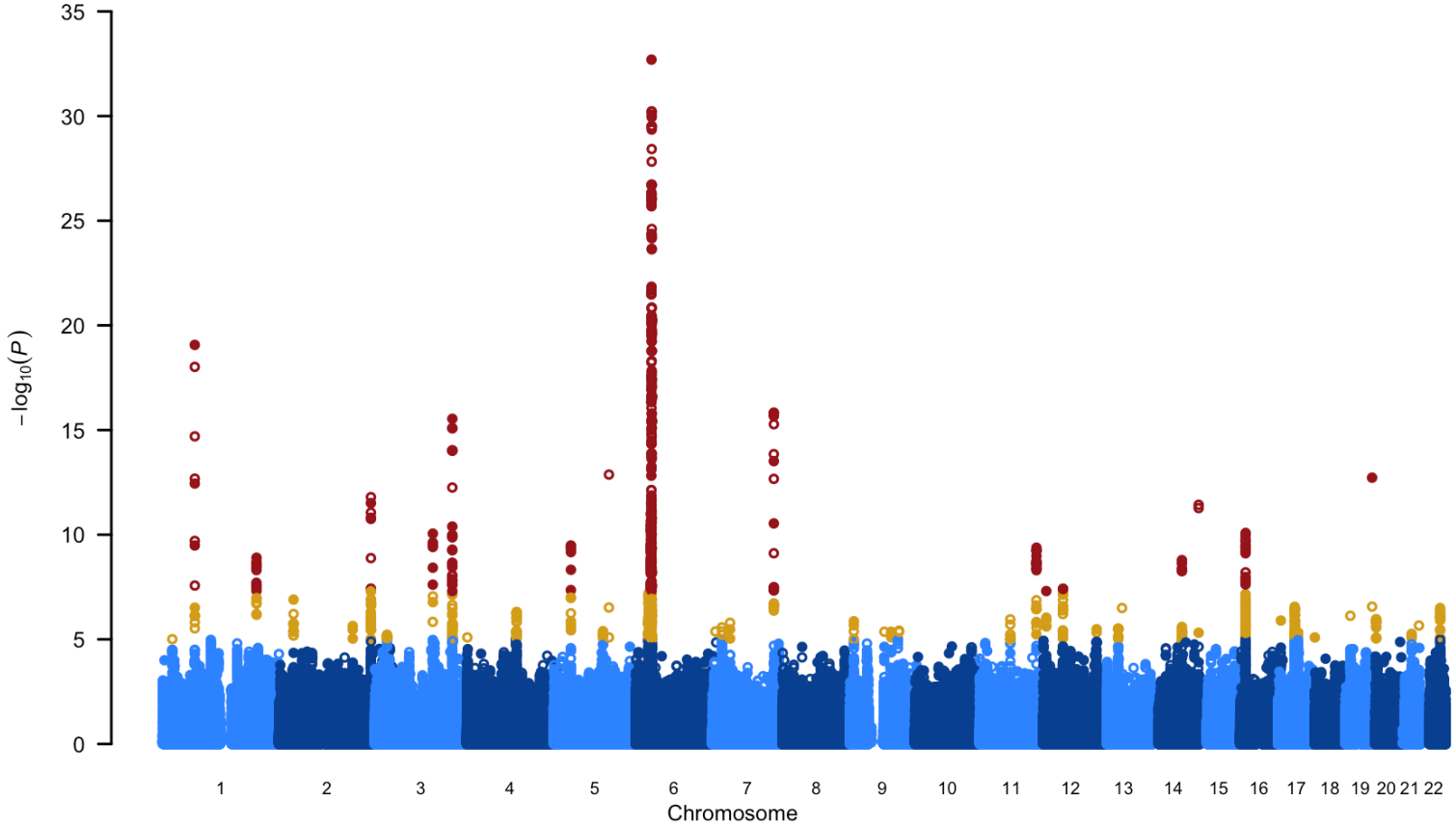


Figure 3: Linkage disequilibrium (r^2) between genome-wide significant SNPs at the *SCHIP1-IL12A* locus at 3q25. The most significant genotyped SNP in the region was rs485499. The remaining SNPs shown were still genome-wide significant ($P \leq 5 \times 10^{-8}$) in an analysis conditioning on rs485499. Three of these SNPs were genotyped (rs7610160, rs17811014, rs2366408), the remaining two were imputed using HapMap3 data. See Supplementary Table 6 for association results for each SNP. Plot was created using Haploview [3].

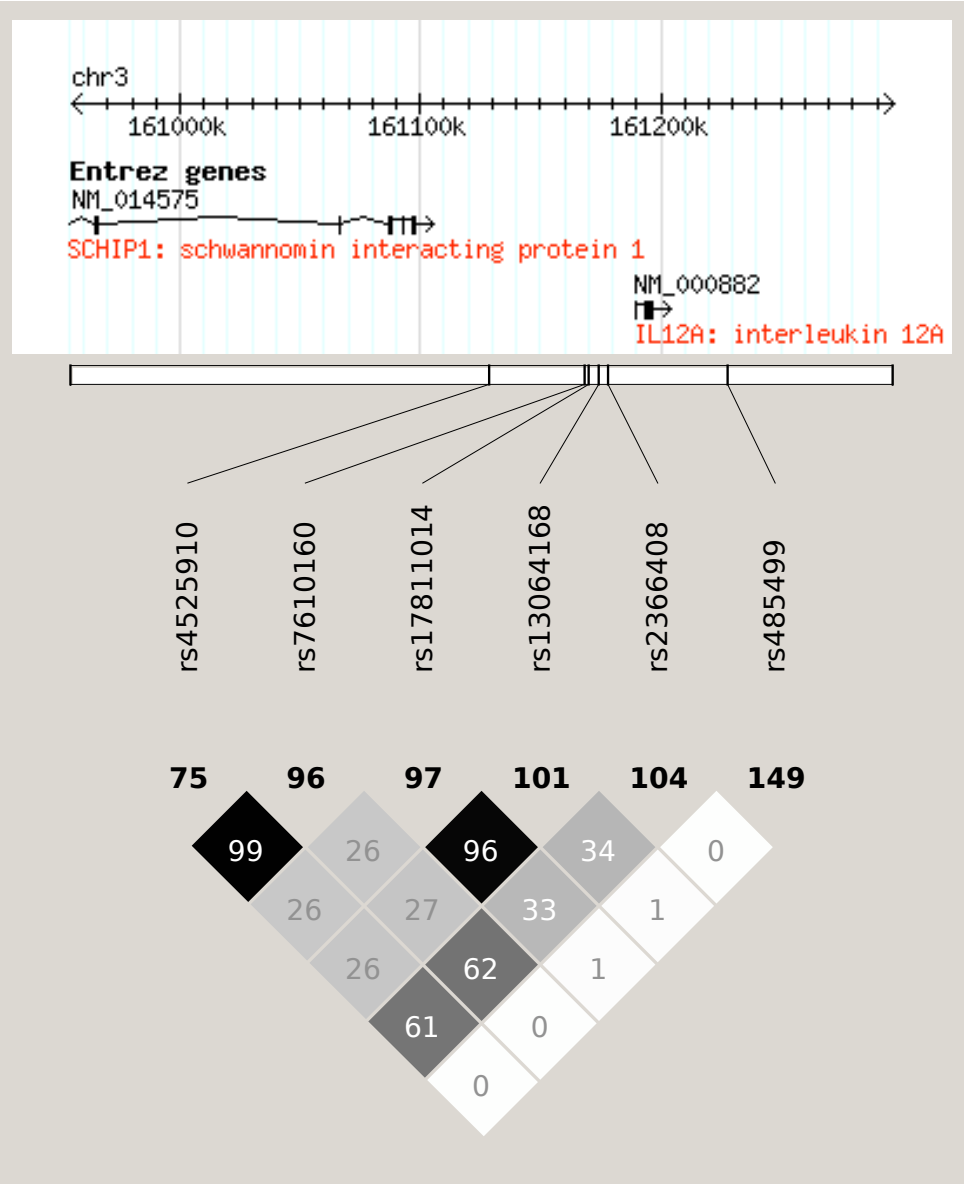
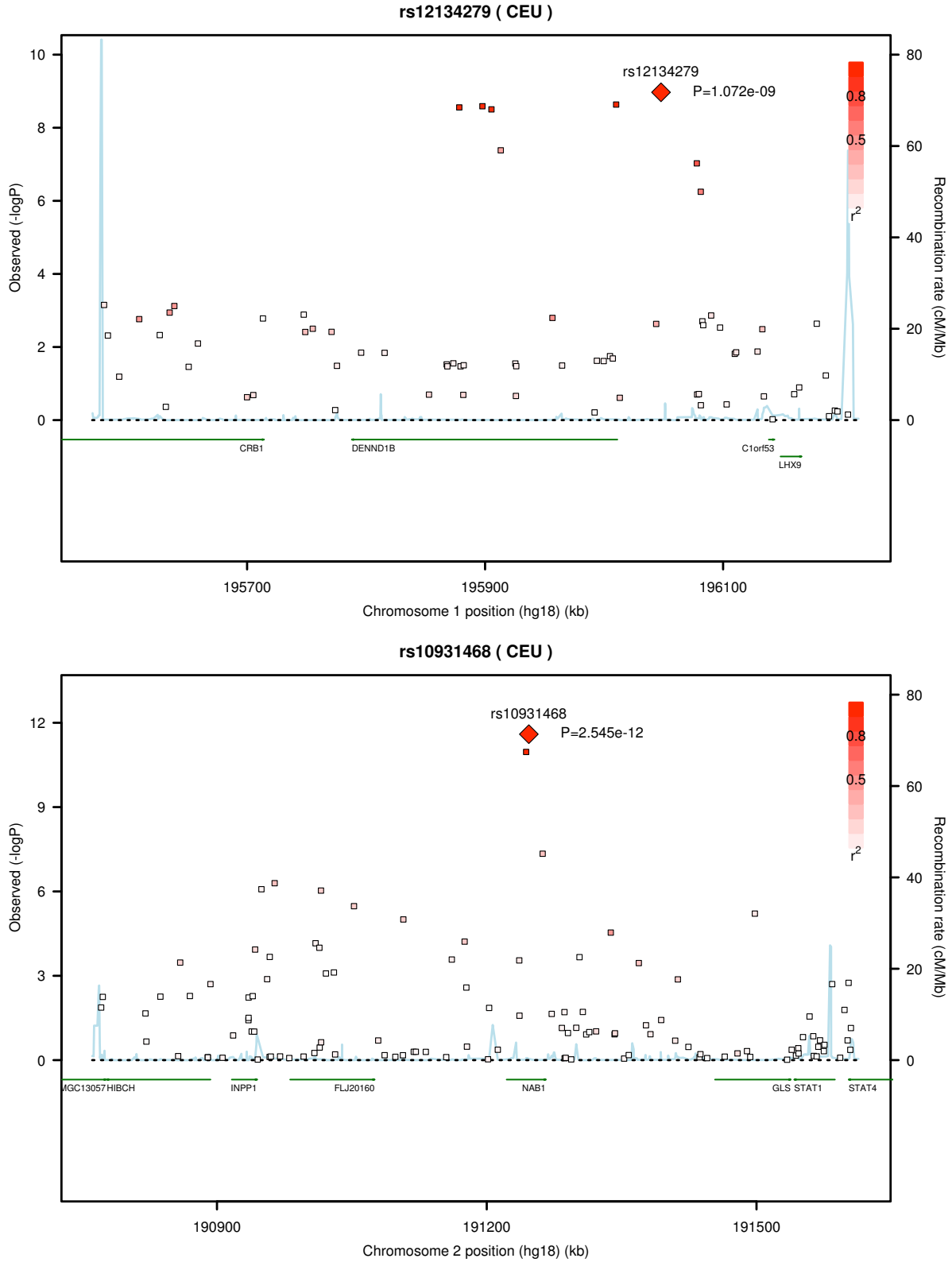
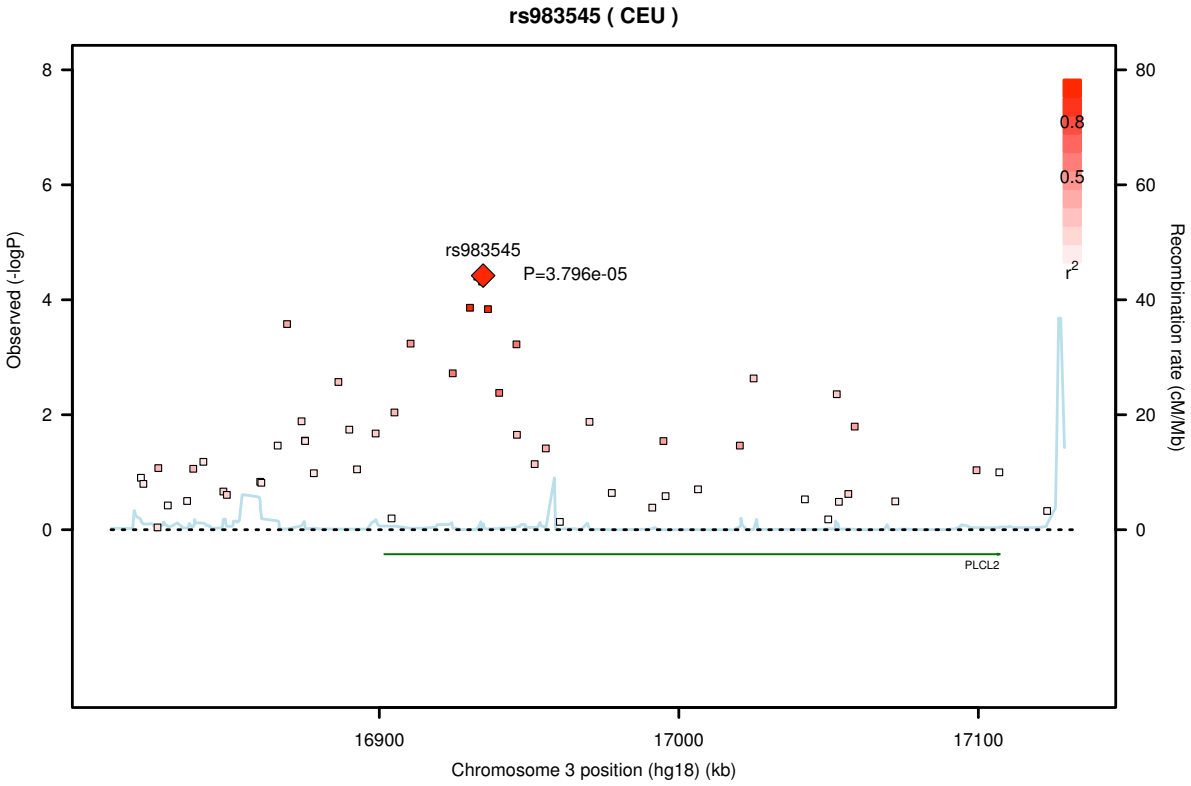
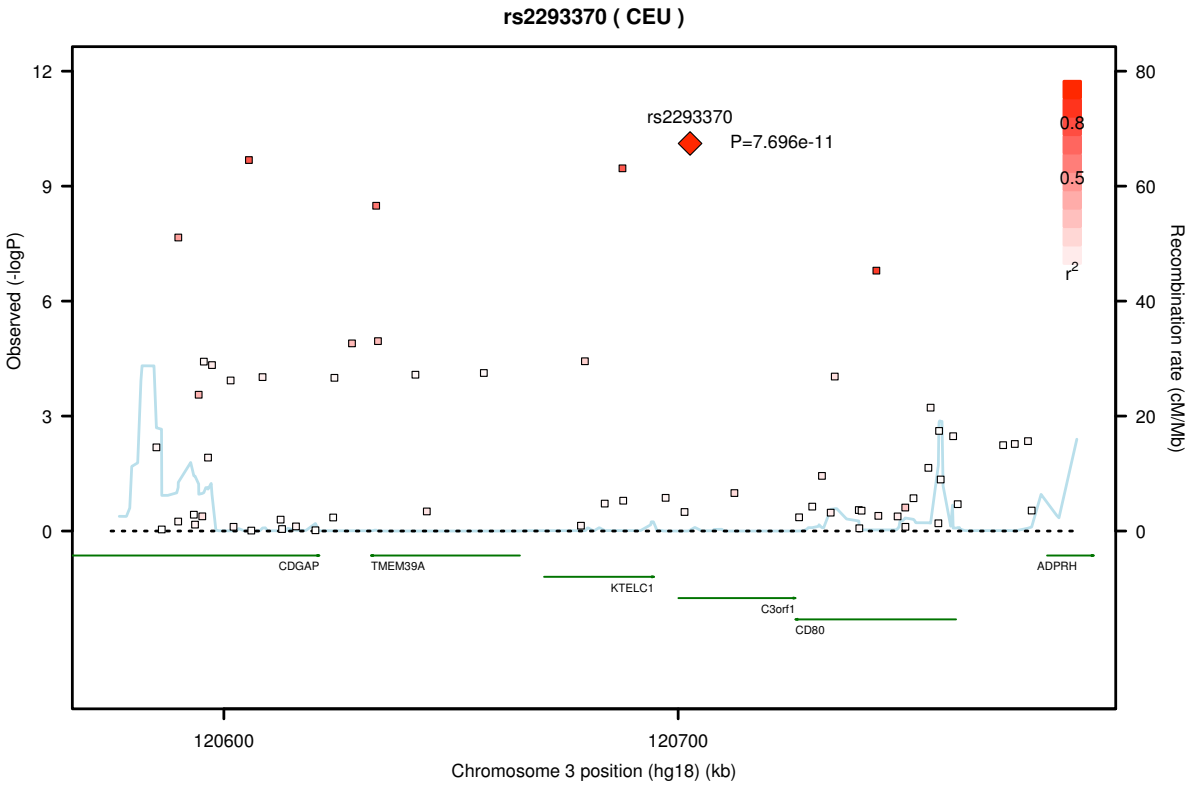
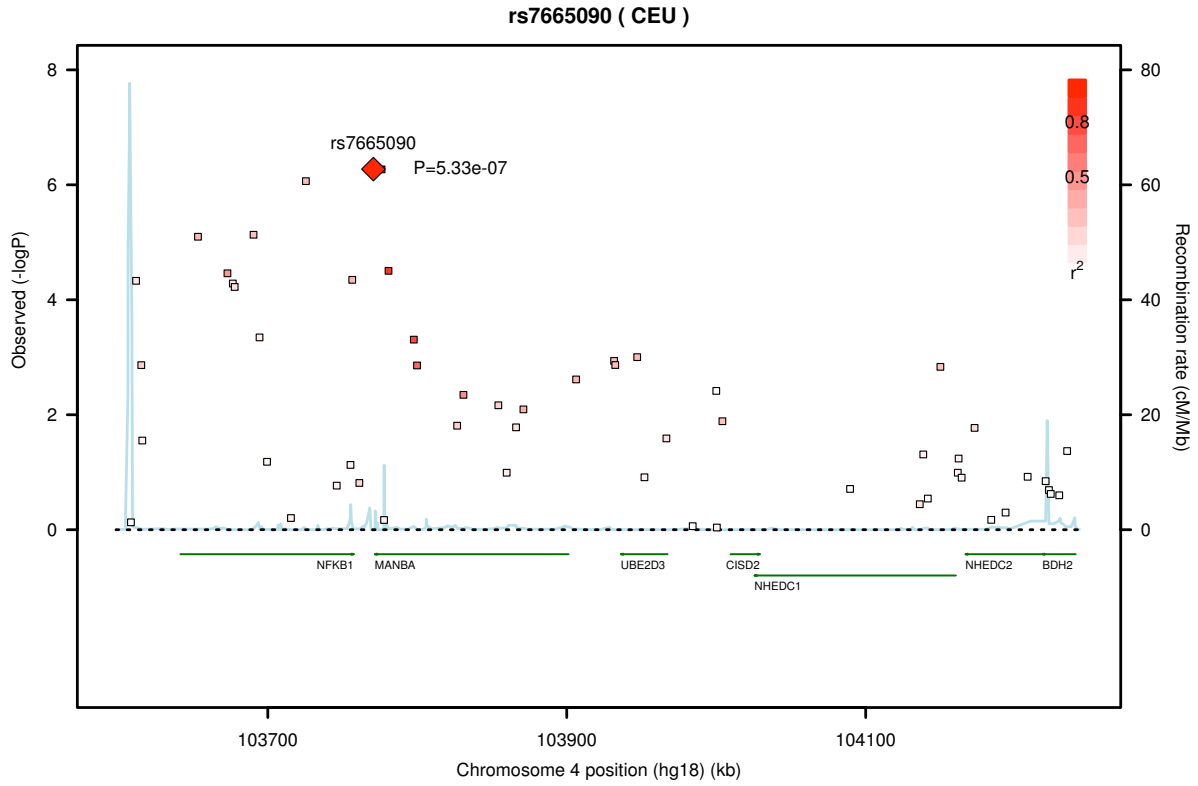


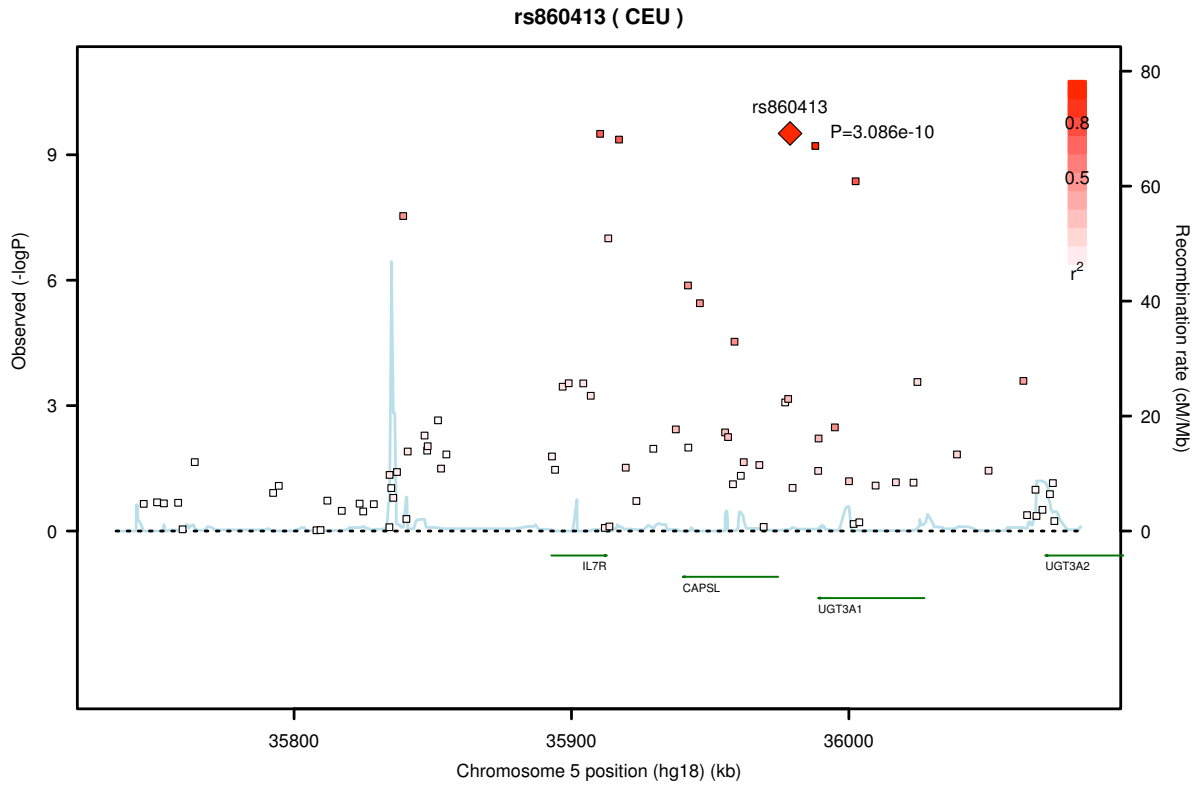
Figure 4: Regional association plots of all genome-wide significant loci, created using SNAP [12].

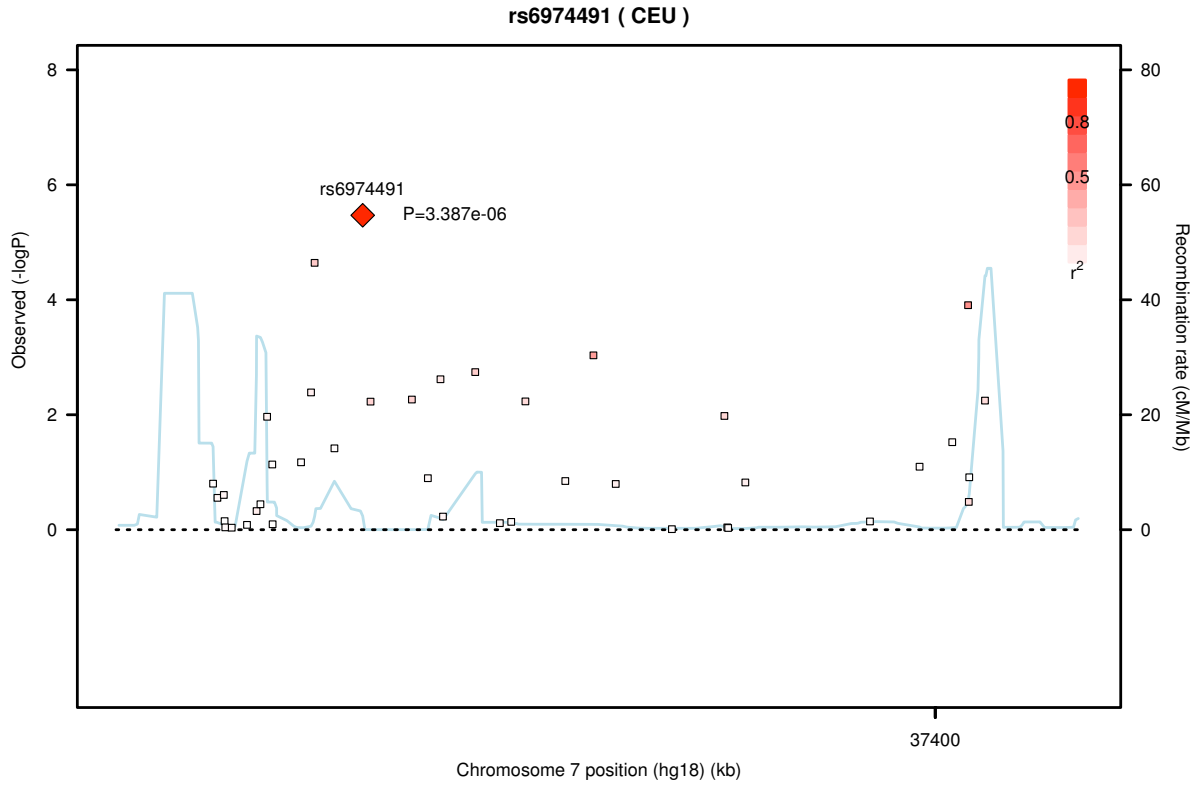


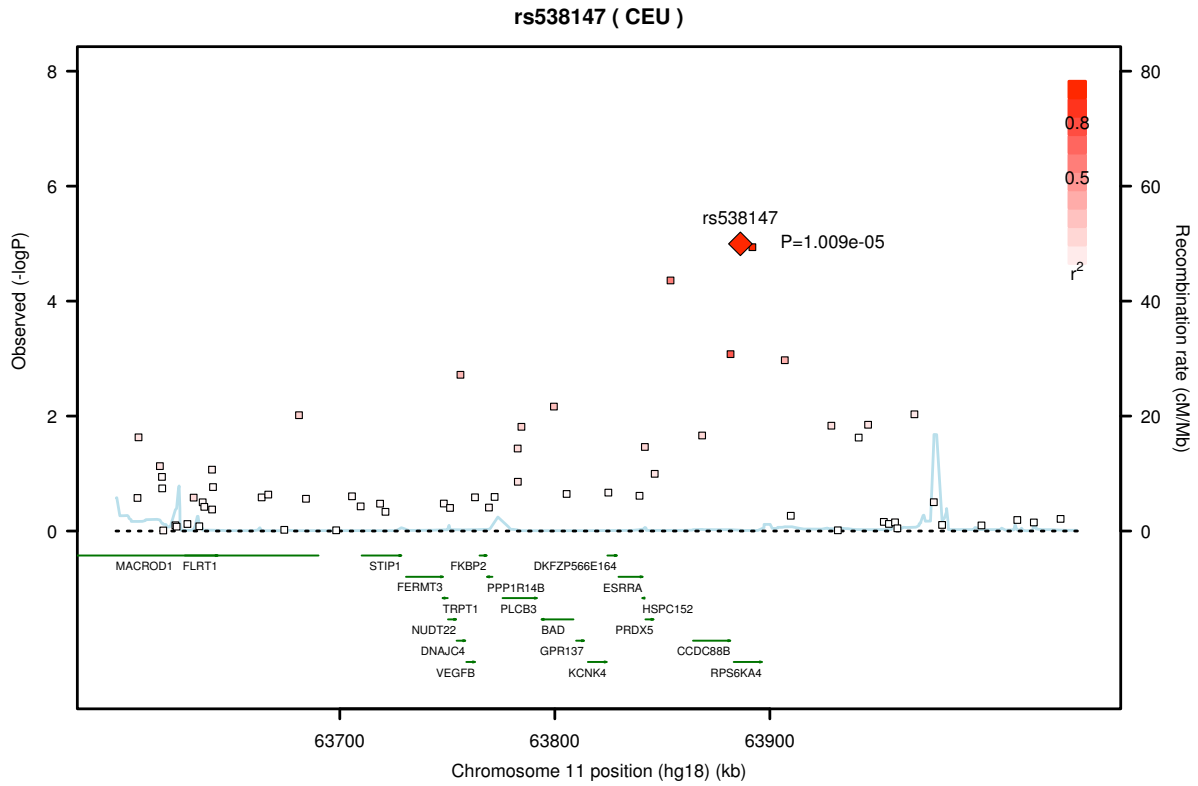

rs983545 (CEU)
Observed (-logP)
Recombination rate (cM/Mb)
rs983545
P=3.796e-05
0.8
0.5
r2
PLCL2
16900
17000
17100
Chromosome 3 position (hg18) (kb)

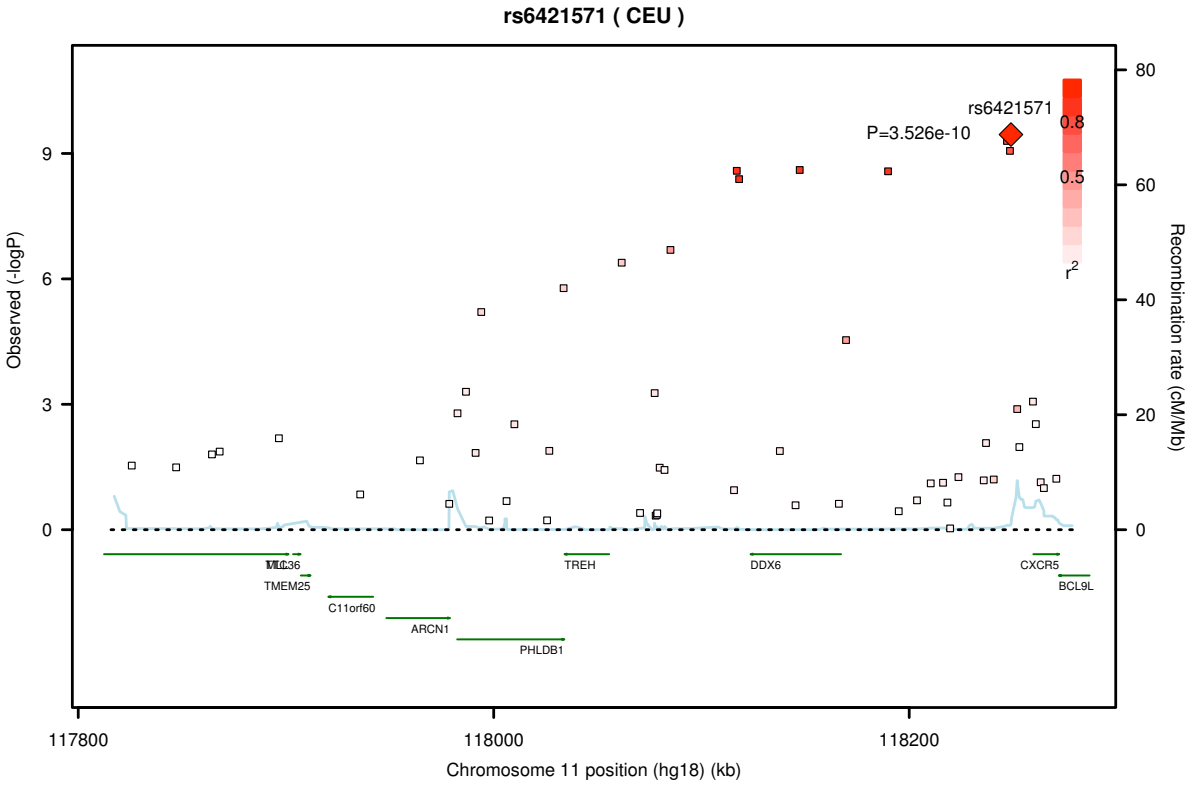

rs2293370 (CEU)
Observed (-logP)
Recombination rate (cM/Mb)
rs2293370
P=7.696e-11
0.8
0.5
r2
CDGAP
TMEM39A
KTELC1
C3orf1
CD80
ADPRH
120600
120700
Chromosome 3 position (hg18) (kb)

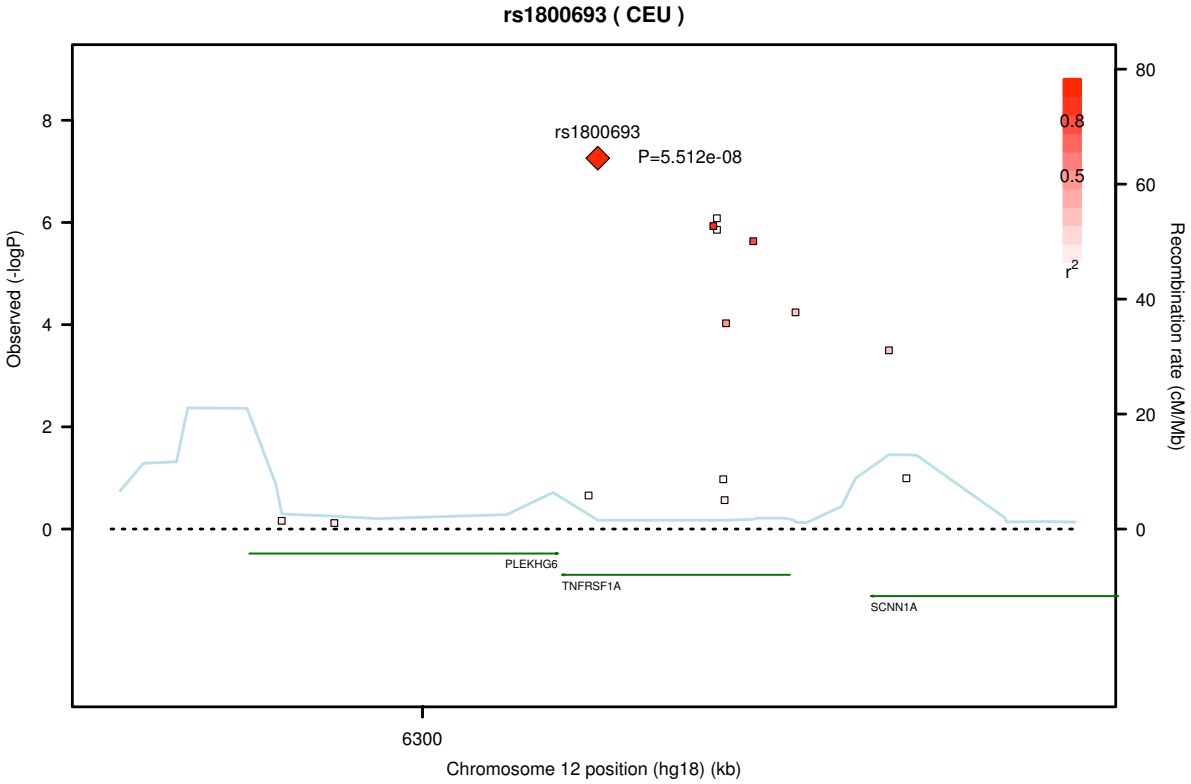

rs7665090 (CEU)
rs7665090
P=5.33e-07
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
NFKB1
MANBA
UBE2D3
CISD2
NHEDC1
NHEDC2
BDH2
103700
103900
104100
Chromosome 4 position (hg18) (kb)

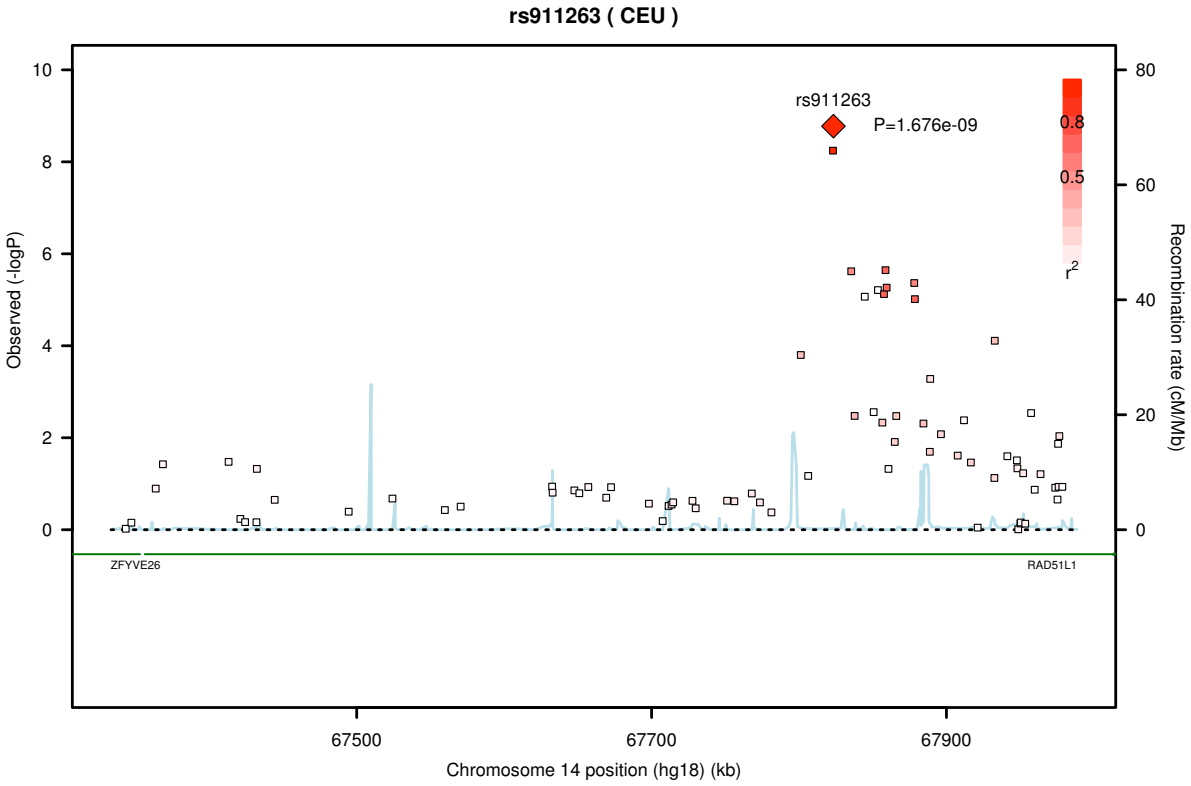

rs860413 (CEU)
rs860413
P=3.086e-10
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
IL7R
CAPSL
UGT3A1
UGT3A2
35800
35900
36000
Chromosome 5 position (hg18) (kb)

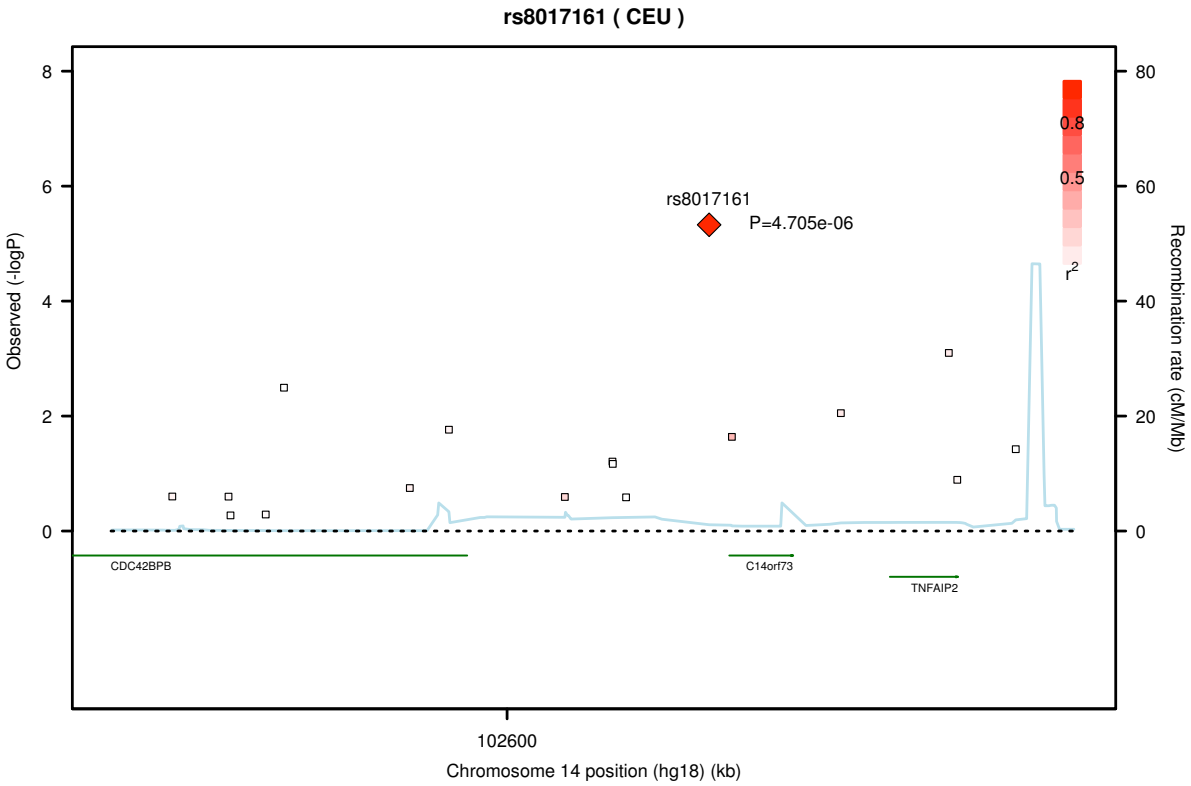

rs6974491 (CEU)
rs6974491
P=3.387e-06
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
37400
Chromosome 7 position (hg18) (kb)

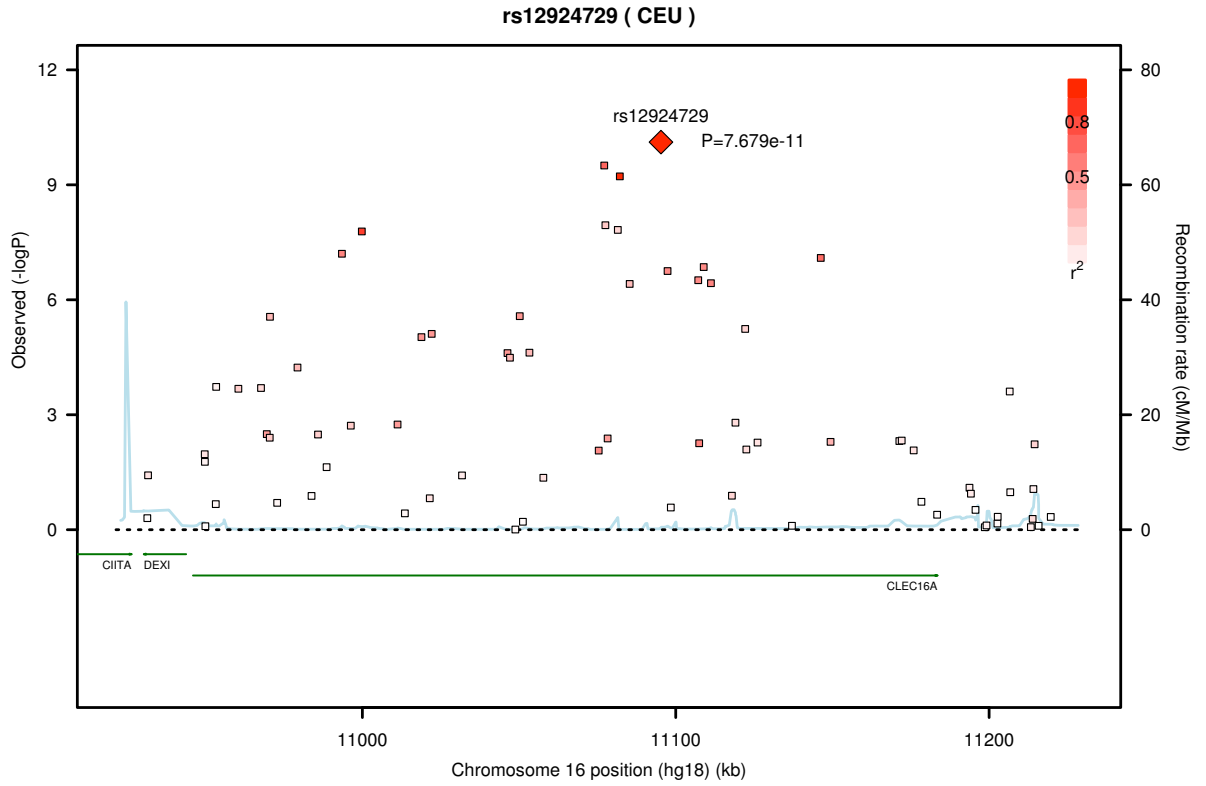

rs538147 (CEU)
rs538147
P=1.009e-05
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
MACROD1
FLRT1
STIP1
FKBP2
DKFZP566E164
FERMT3
PPP1R14B
ESRRA
TRPT1
PLCB3
HSPC152
NUDT22
BAD
PRDX5
DNAJC4
GPR137
CCDC88B
VEGFB
KCNK4
RPS6KA4
63700
63800
63900
Chromosome 11 position (hg18) (kb)

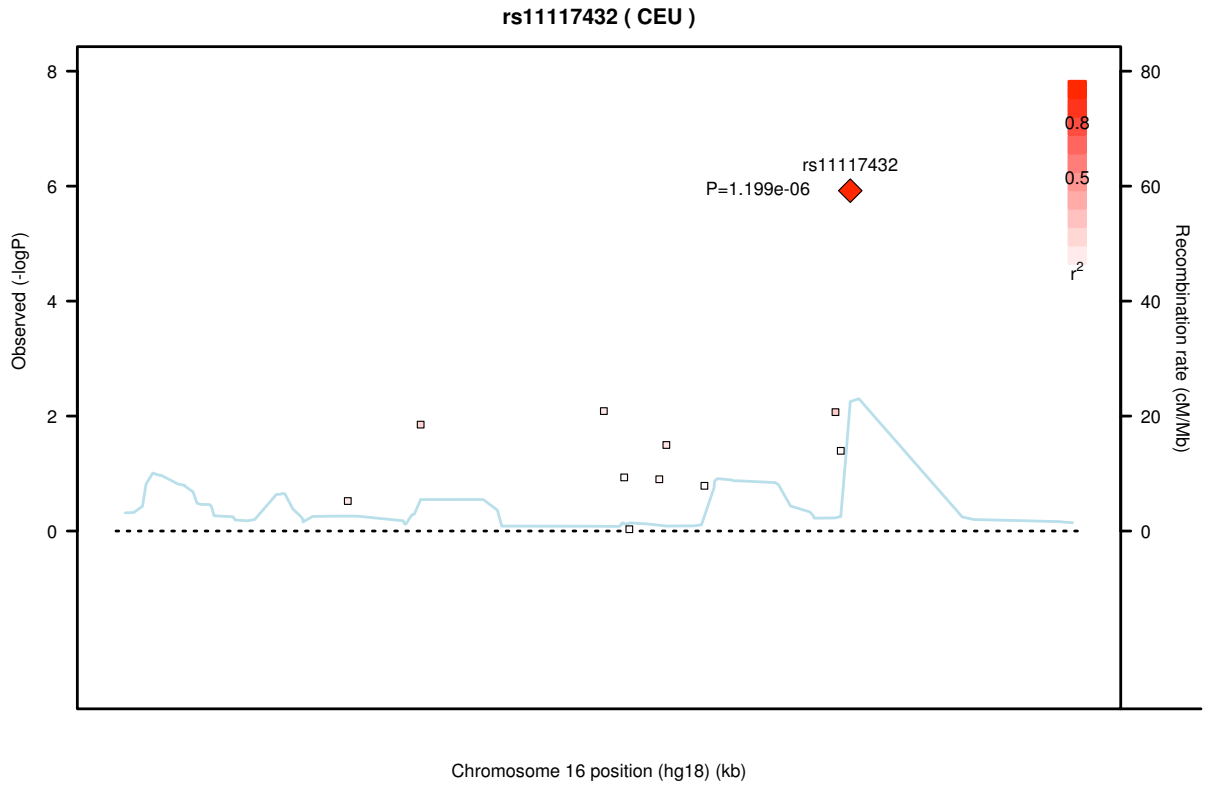

rs6421571 (CEU)
rs6421571
P=3.526e-10
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
TMEM25
C11orf60
ARCN1
PHLDB1
TREH
DDX6
CXCR5
BCL9L
117800
118000
118200
Chromosome 11 position (hg18) (kb)

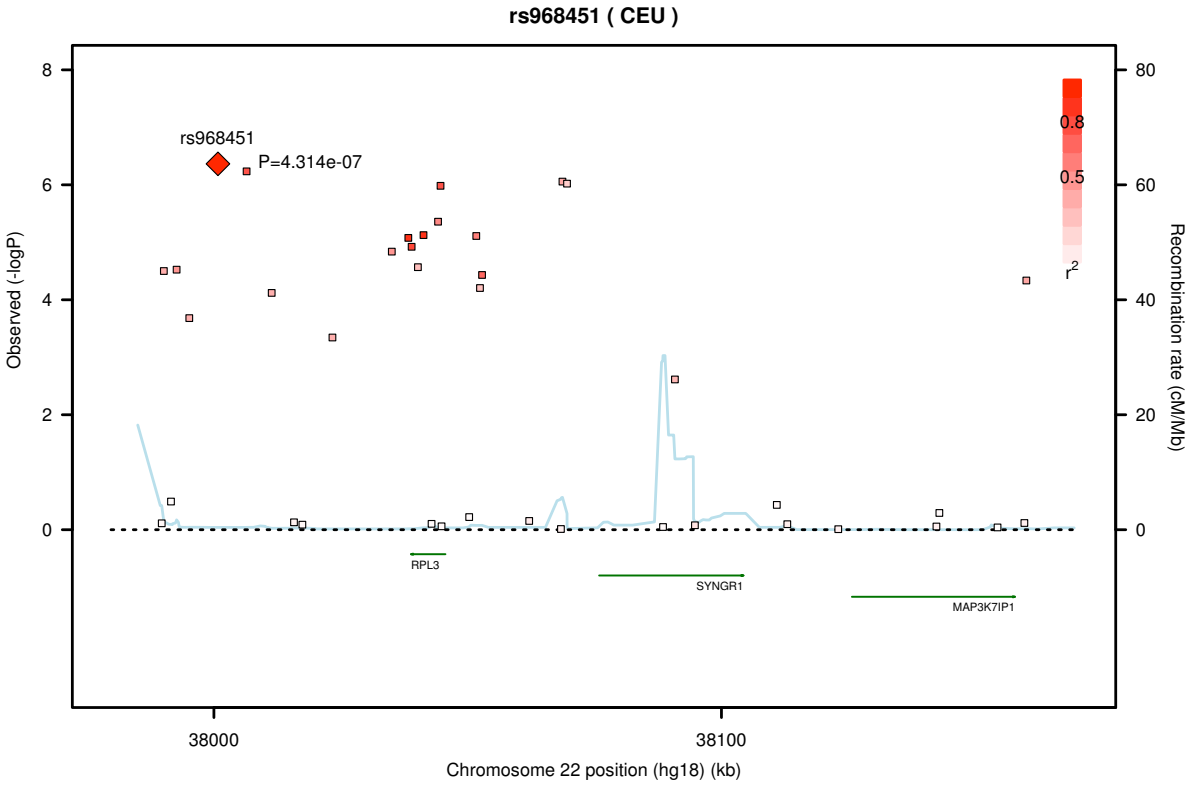

rs1800693 (CEU)
rs1800693
P=5.512e-08
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
PLEKHG6
TNFRSF1A
SCNN1A
6300
Chromosome 12 position (hg18) (kb)

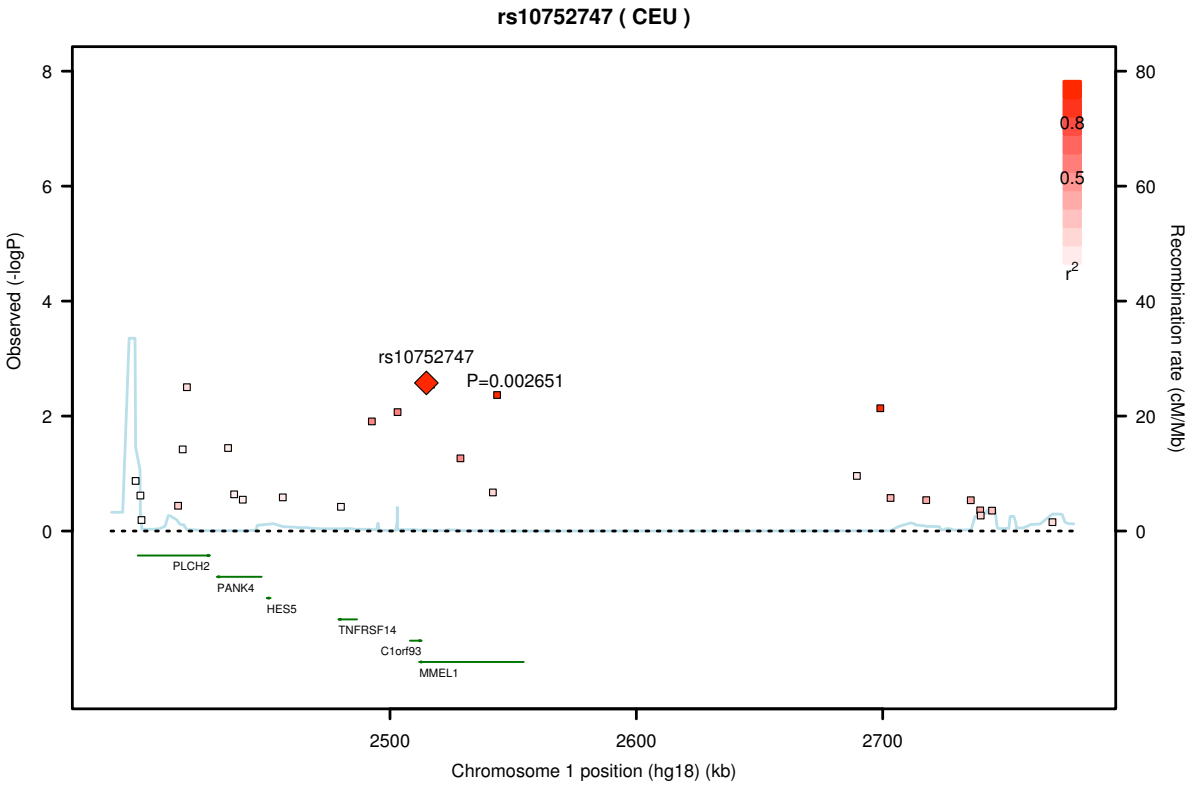

rs911263 (CEU)
rs911263
P=1.676e-09
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
ZFYVE26
RAD51L1
67500
67700
67900
Chromosome 14 position (hg18) (kb)

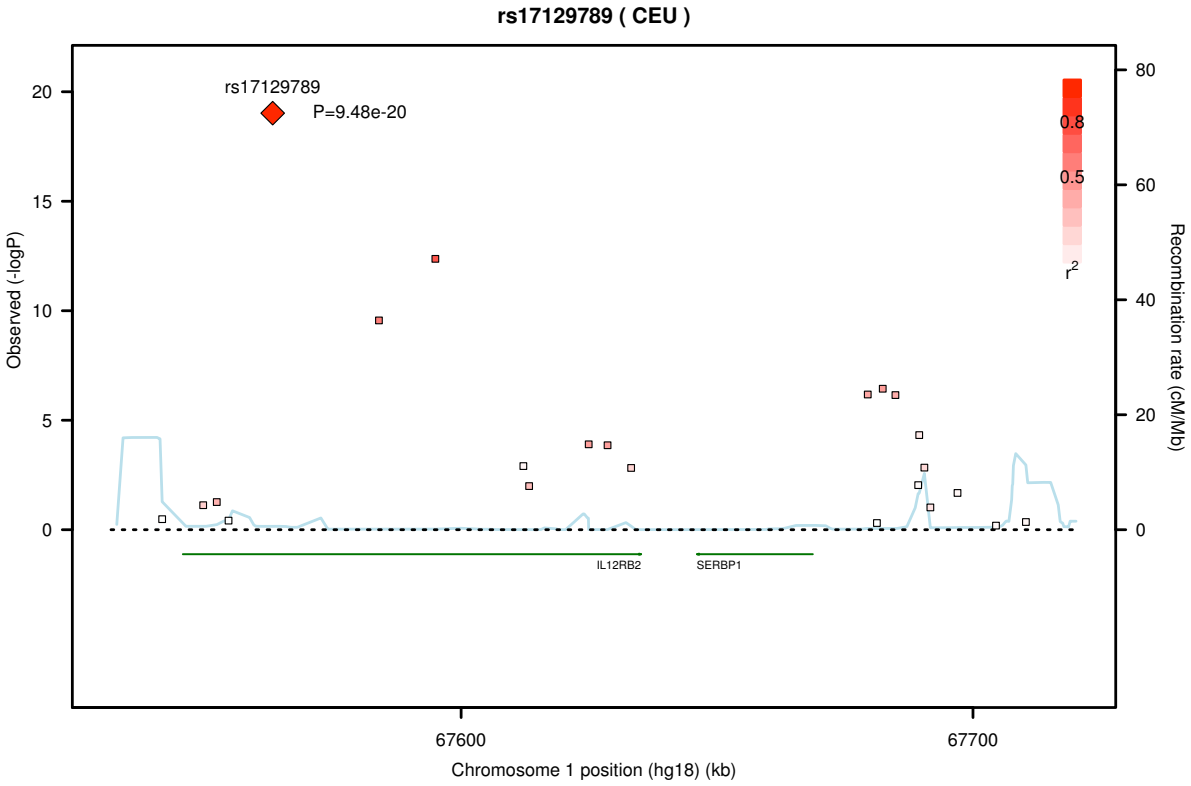

rs8017161 (CEU)
rs8017161
P=4.705e-06
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
CDC42BPB
C14orf73
TNFAIP2
102600
Chromosome 14 position (hg18) (kb)

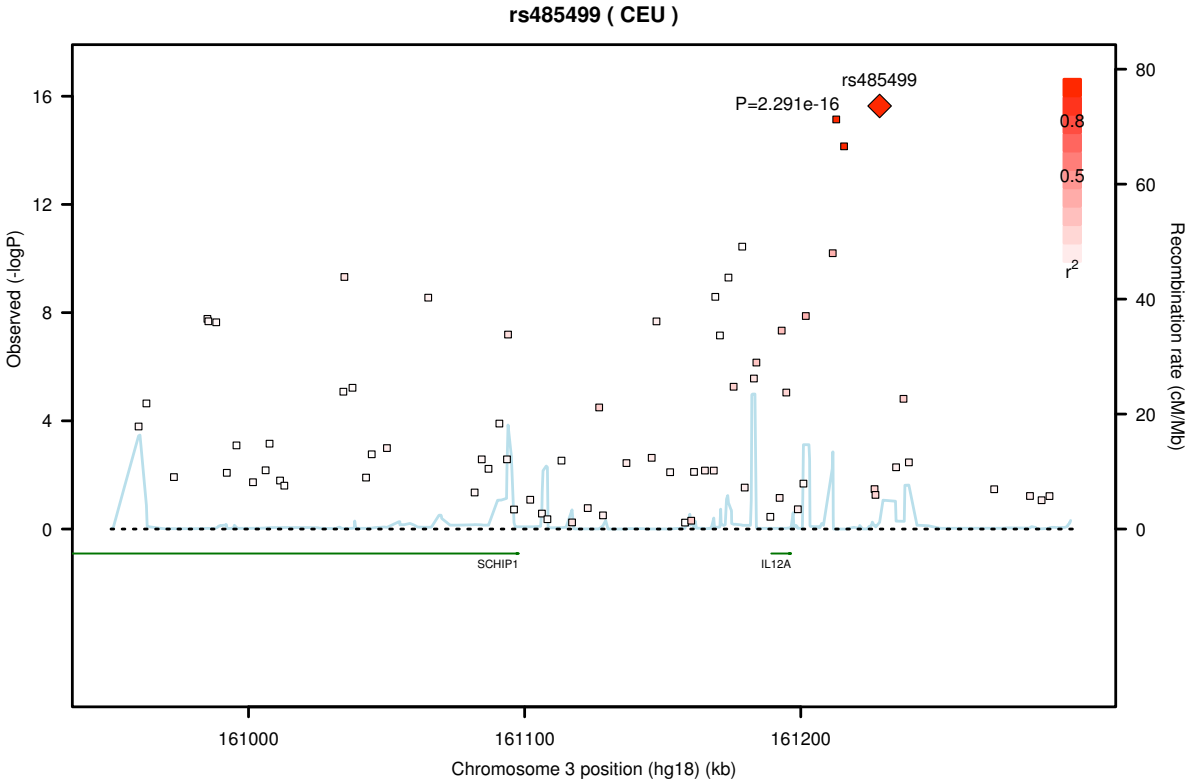

rs12924729 (CEU)
rs12924729
P=7.679e-11
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
CIITA
DEXI
CLEC16A
11000
11100
11200
Chromosome 16 position (hg18) (kb)

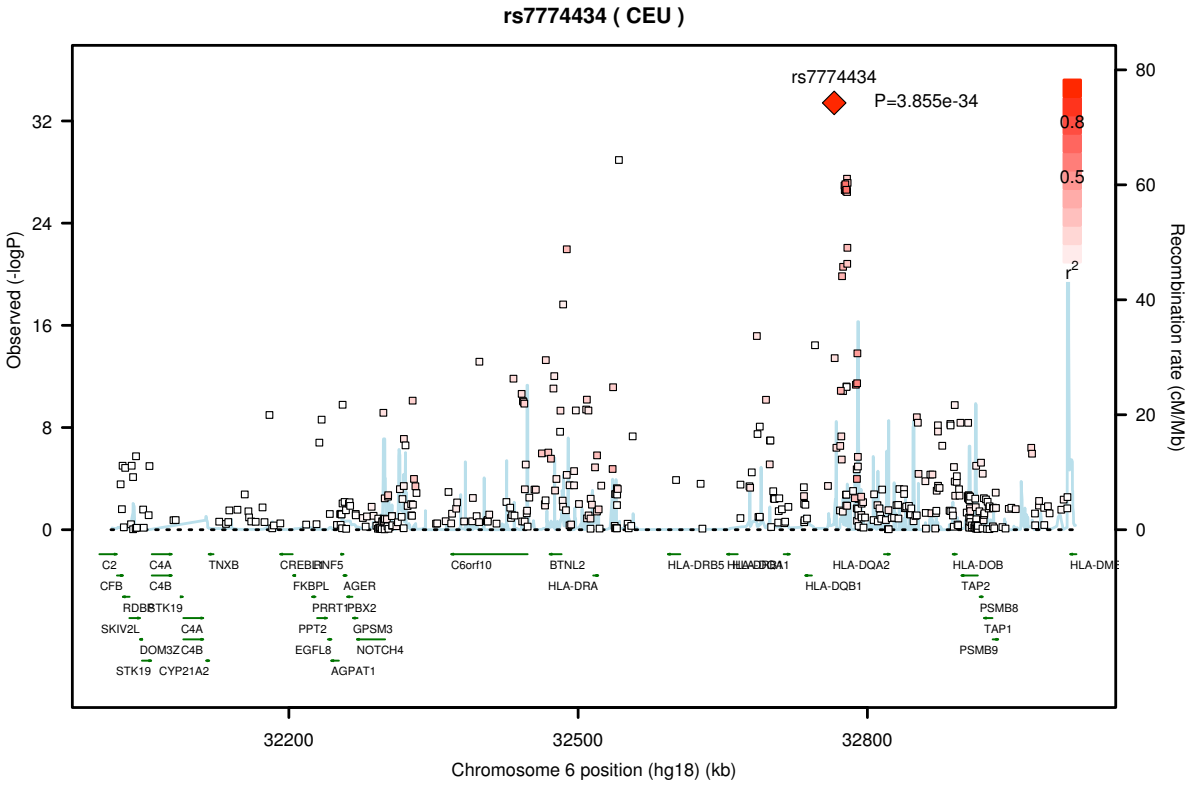

rs11117432 (CEU)
rs11117432
P=1.199e-06
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
Chromosome 16 position (hg18) (kb)

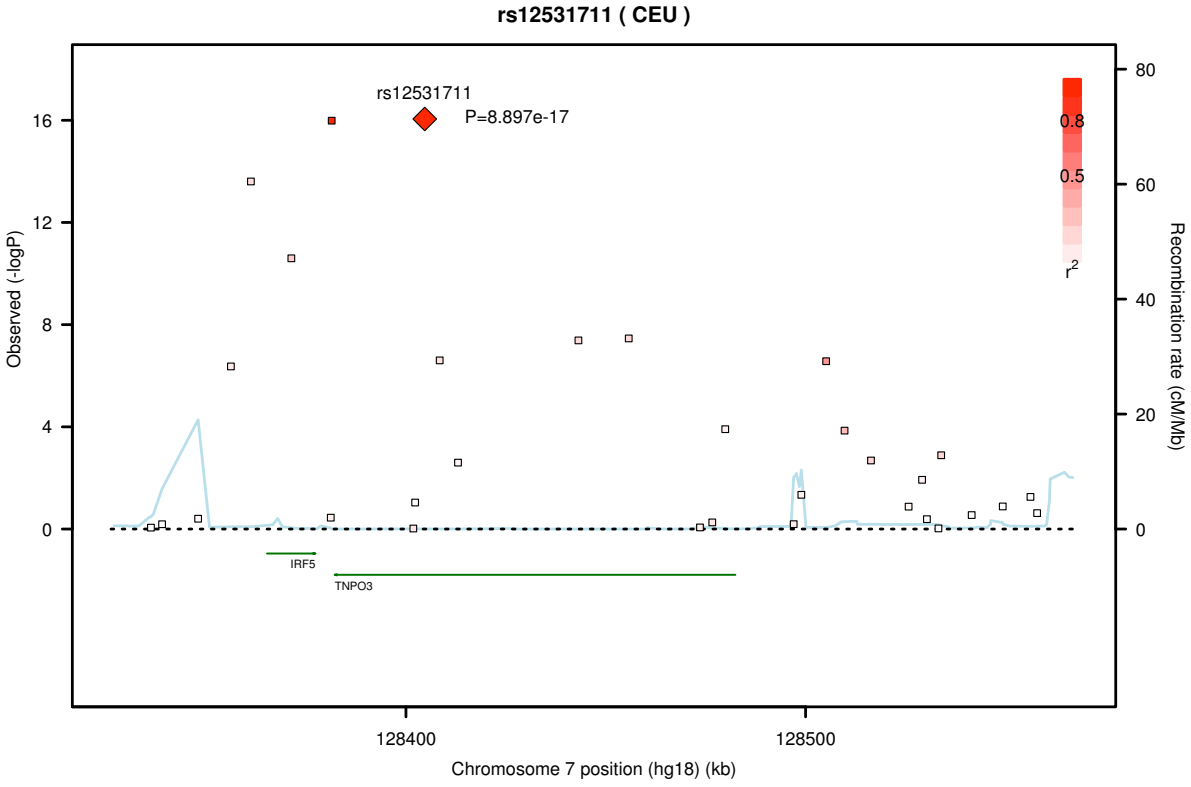

rs968451 (CEU)
rs968451
P=4.314e-07
Observed (-logP)
Recombination rate (cM/Mb)
r2
0.8
0.5
RPL3
SYNGR1
MAP3K7IP1
38000
38100
Chromosome 22 position (hg18) (kb)

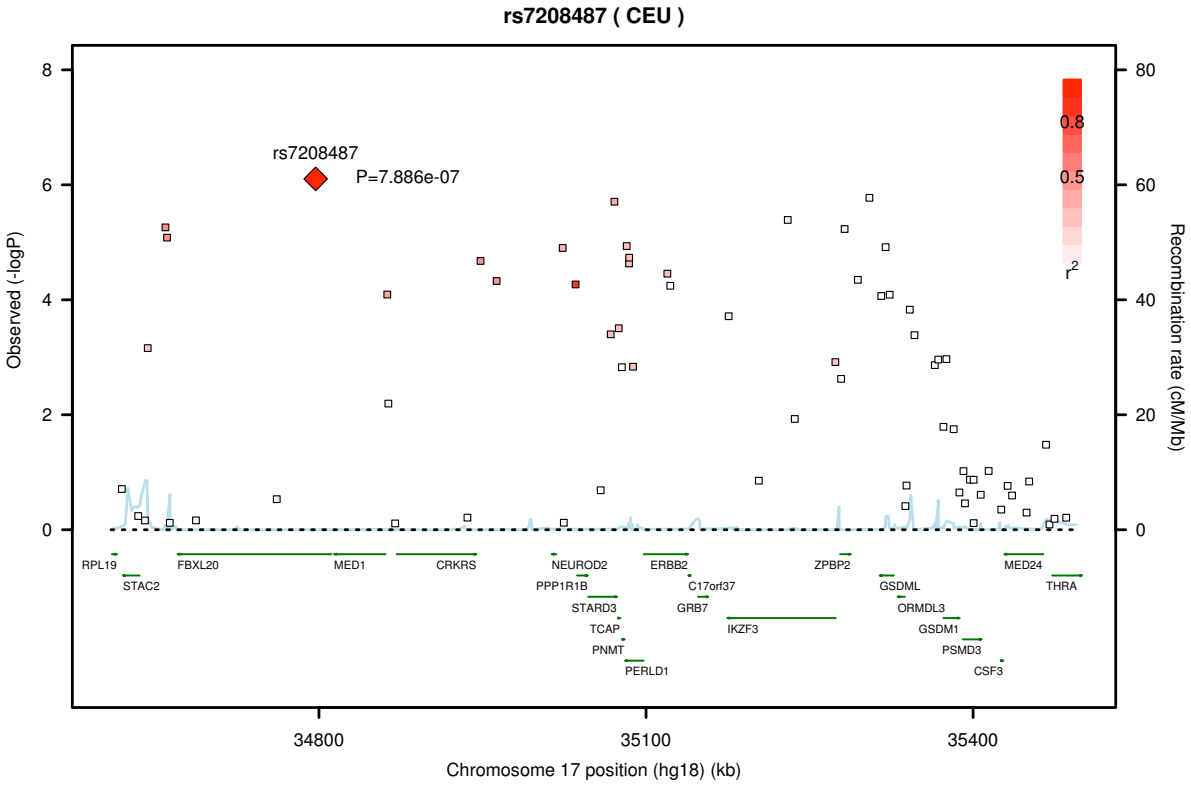

rs10752747 (CEU)
rs10752747
P=0.002651
Observed (-logP)
Recombination rate (cM/Mb)
r2
0.8
0.5
PLCH2
PANK4
HES5
TNFRSF14
C1orf93
MMEL1
2500
2600
2700
Chromosome 1 position (hg18) (kb)


rs17129789 (CEU)
rs17129789
P=9.48e-20
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
IL12RB2
SERBP1
67600
67700
Chromosome 1 position (hg18) (kb)


rs485499 (CEU)
rs485499
P=2.291e-16
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
SCHIP1
IL12A
161000
161100
161200
Chromosome 3 position (hg18) (kb)


rs7774434 (CEU)
rs7774434
P=3.855e-34
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
32
24
16
8
0
80
60
40
20
0
C2
CFB
C4A
C4B
TNXB
RDBP
TK19
SKIV2L
C4A
DOM3Z
C4B
STK19
CYP21A2
CREBL1
IRF5
FKBPL
AGER
PRRT1
PBX2
PPT2
GPSM3
EGFL8
NOTCH4
AGPAT1
C6orf10
BTNL2
HLA-DRA
HLA-DRB5
HLA-DQA1
HLA-DQA2
HLA-DQB1
HLA-DOB
TAP2
PSMB8
TAP1
PSMB9
HLA-DM
32200
32500
32800
Chromosome 6 position (hg18) (kb)


rs12531711 (CEU)
rs12531711
P=8.897e-17
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
16
12
8
4
0
80
60
40
20
0
IRF5
TNPO3
128400
128500
Chromosome 7 position (hg18) (kb)


rs7208487 (CEU)
rs7208487
P=7.886e-07
Observed (-logP)
Recombination rate (cM/Mb)
r2
0.8
0.5
RPL19
STAC2
FBXL20
MED1
CRKRS
NEUROD2
PPP1R1B
STARD3
TCAP
PNMT
PERLD1
ERBB2
C17orf37
GRB7
IKZF3
ZPBP2
GSDML
ORMDL3
GSDM1
PSMD3
CSF3
MED24
THRA
34800
35100
35400
Chromosome 17 position (hg18) (kb)

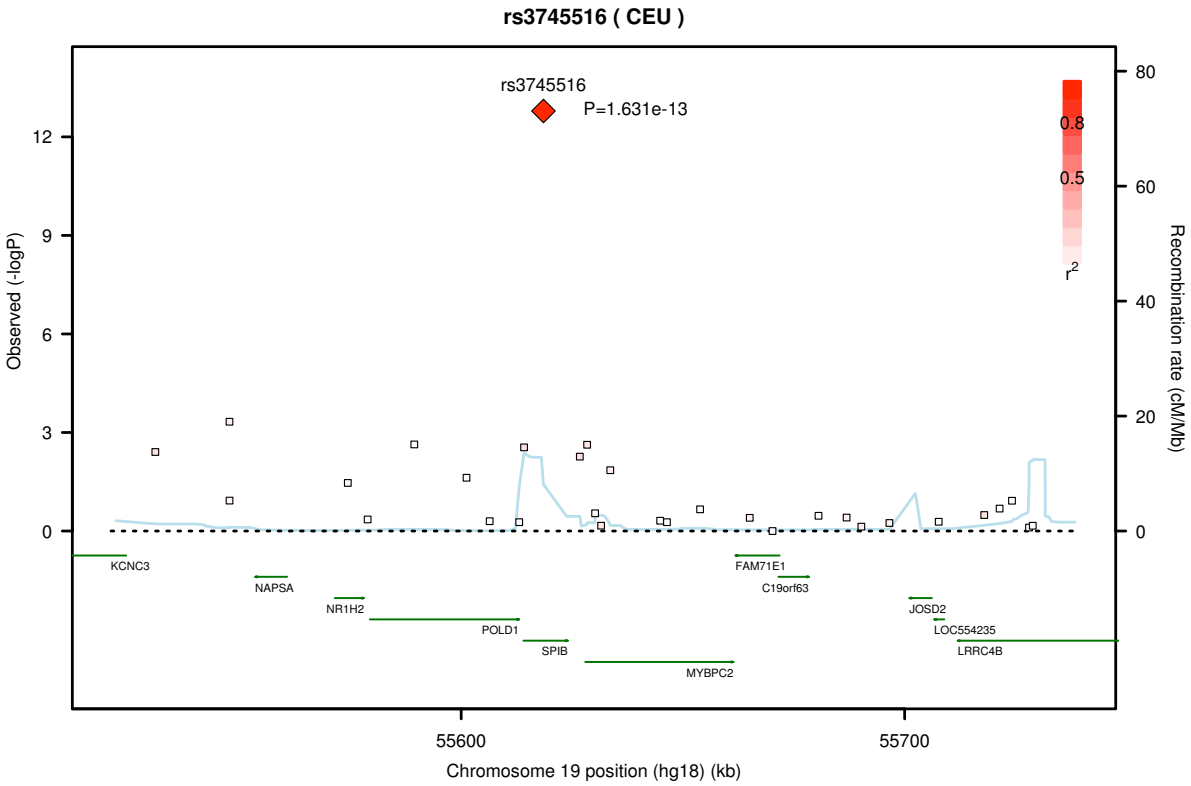

rs3745516 (CEU)
rs3745516
P=1.631e-13
Observed (-logP)
Recombination rate (cM/Mb)
r2
0.8
0.5
KCNC3
NAPSA
NR1H2
POLD1
SPIB
MYBPC2
FAM71E1
C19orf63
JOSD2
LOC554235
LRRC4B
55600
55700
Chromosome 19 position (hg18) (kb)

Figure 5: Plot of first two principal components from the GWAS samples and HapMap data. Eigenvalues for 61,863 SNPs were calculated for the CEU, CHB, JPT and YRI HapMap samples, and these were then applied to the GWAS samples. Red crosses represent individuals from the GWAS showing non-European or mixed ancestry.

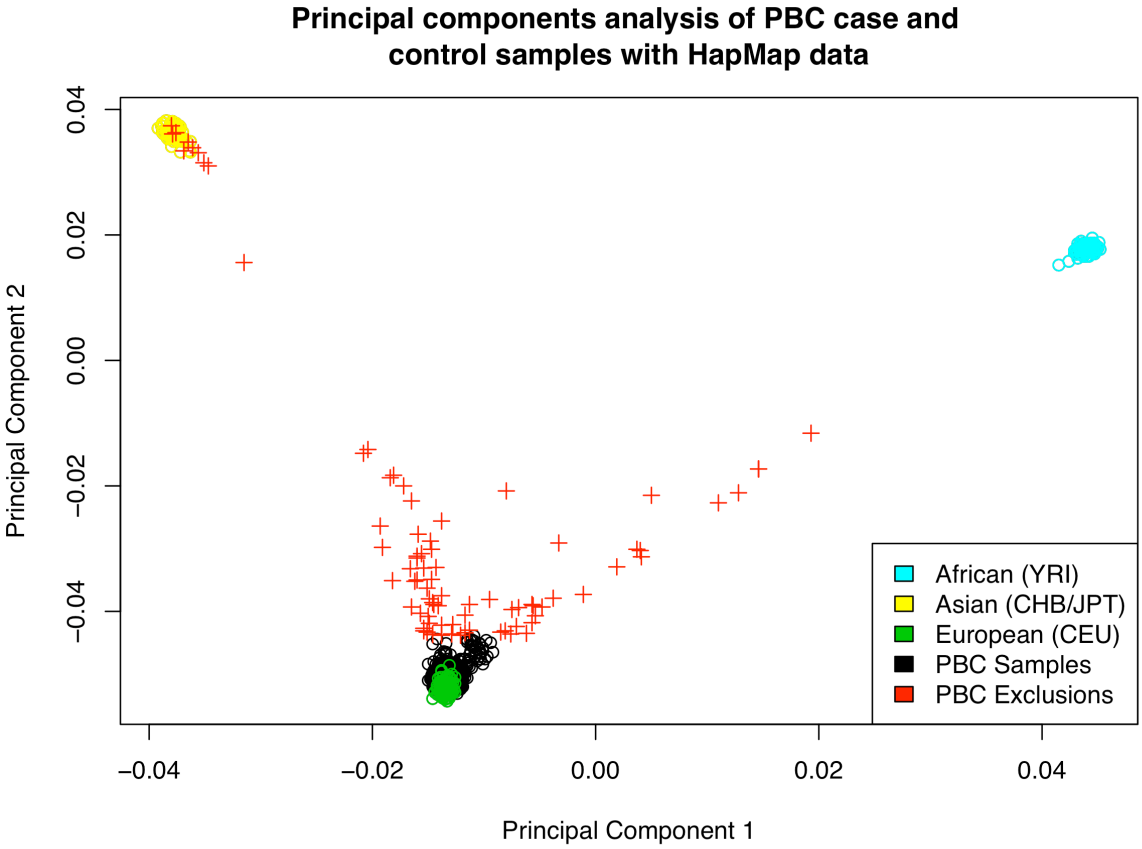


Figure 6: Heterozygosity and missingness are plotted for all GWAS samples based on the SNPs analysed in this study. The red dashed horizontal lines indicate ±3 standard deviations from the mean, and the red dashed vertical line indicates the call rate threshold of 0.02.

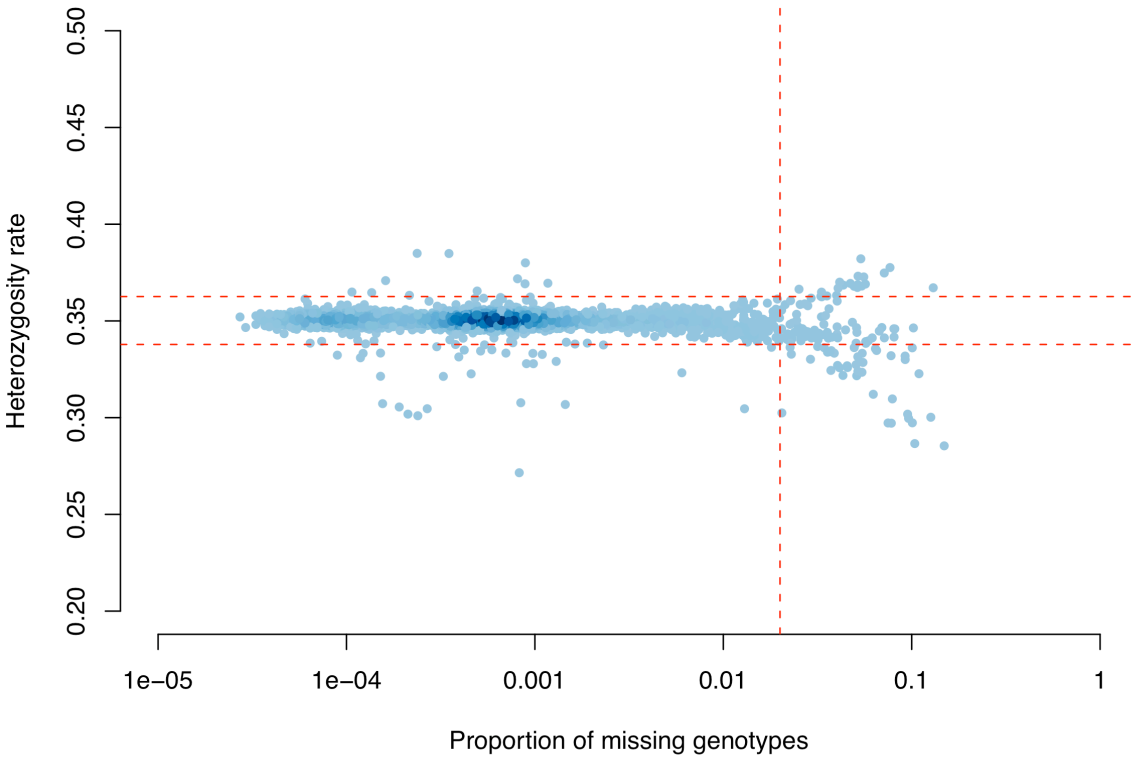
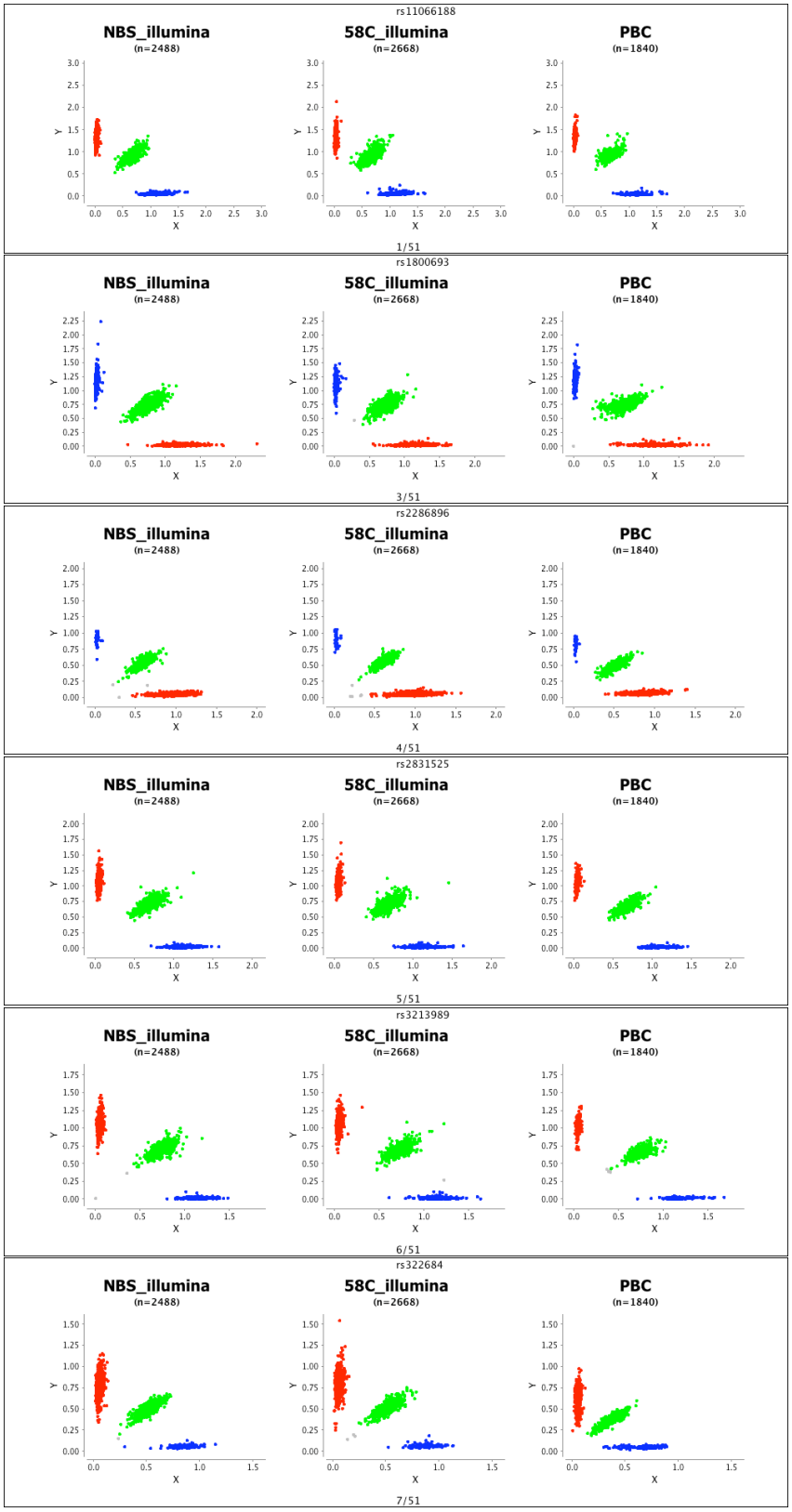
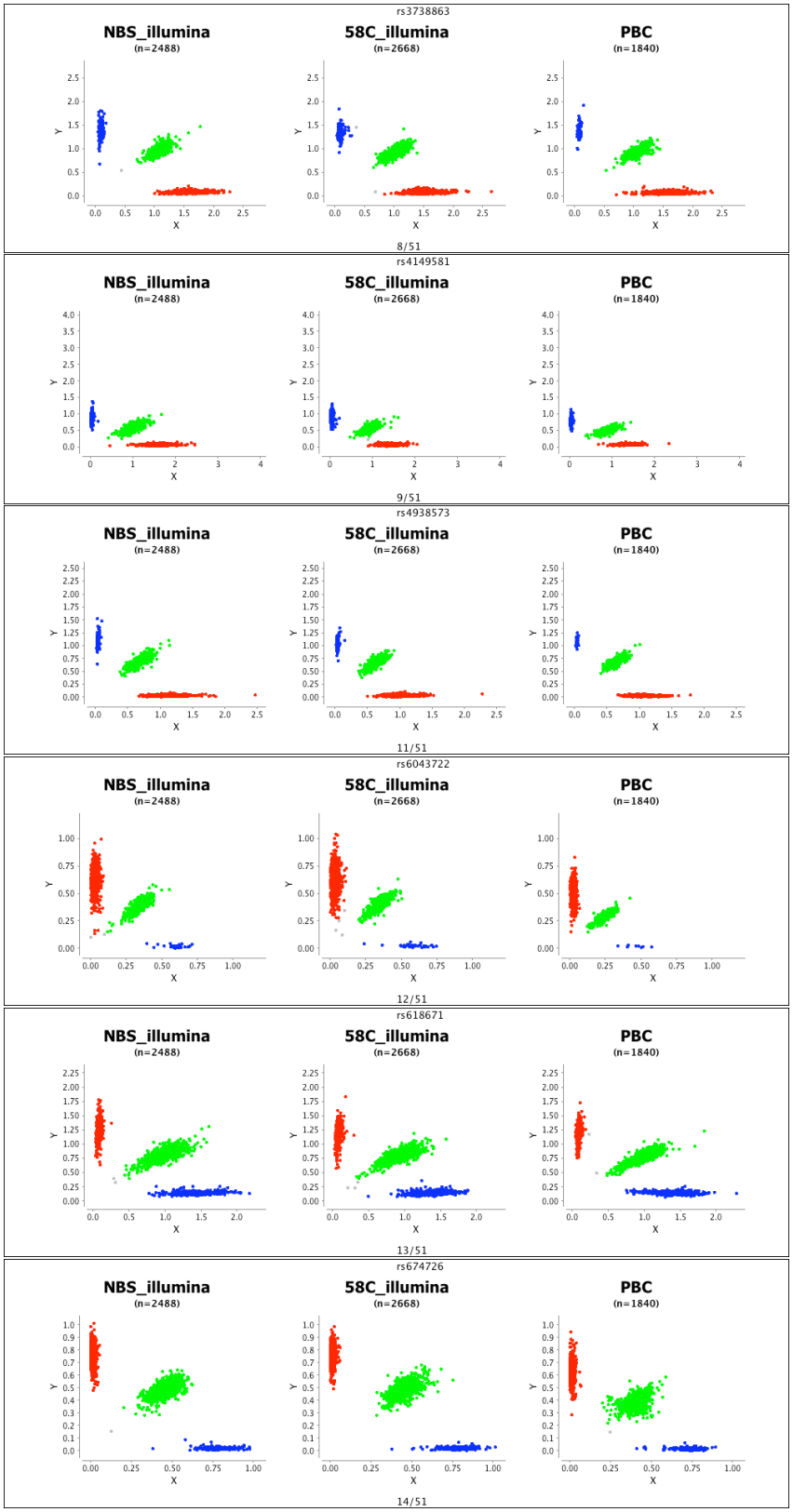
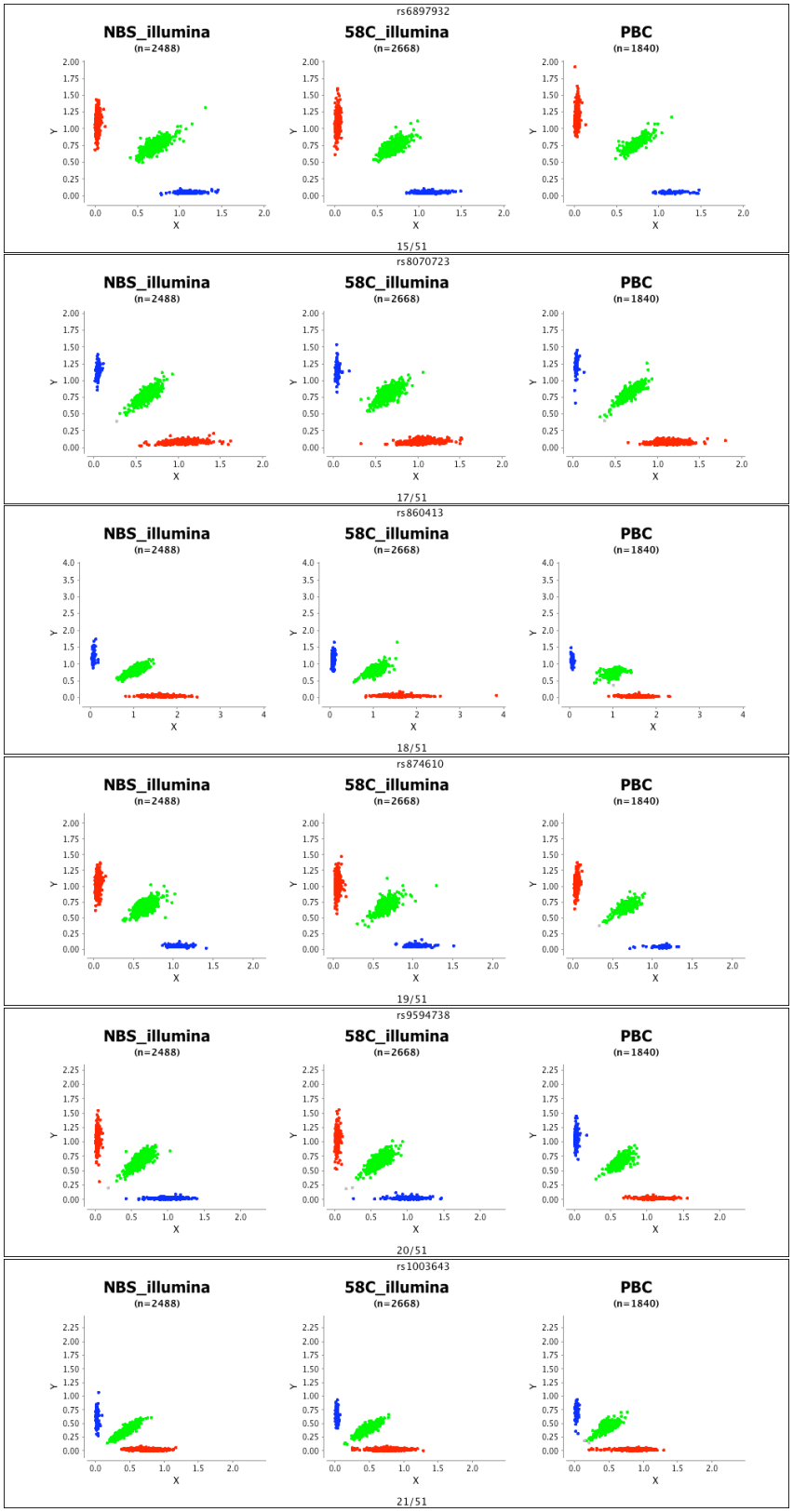
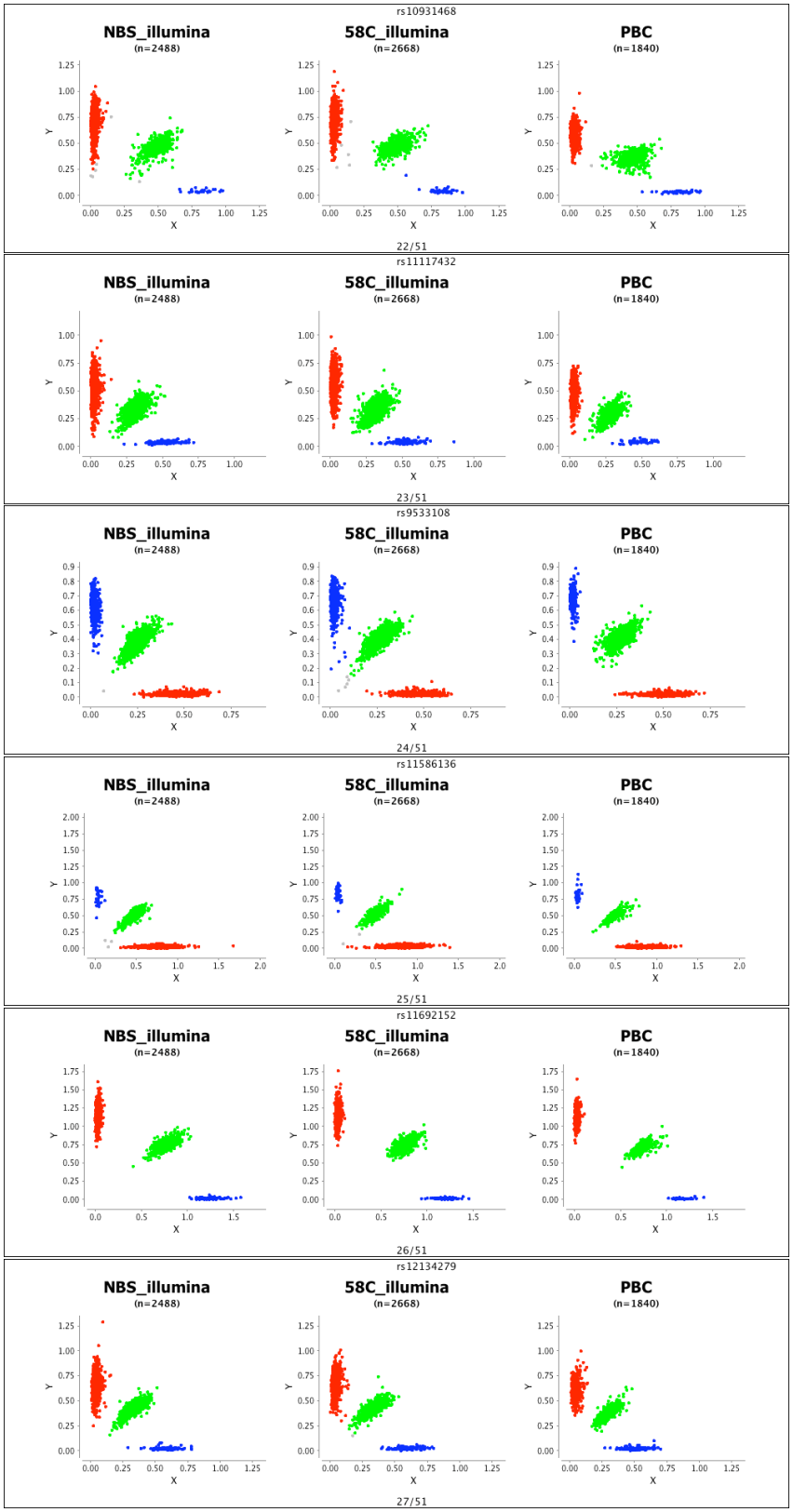


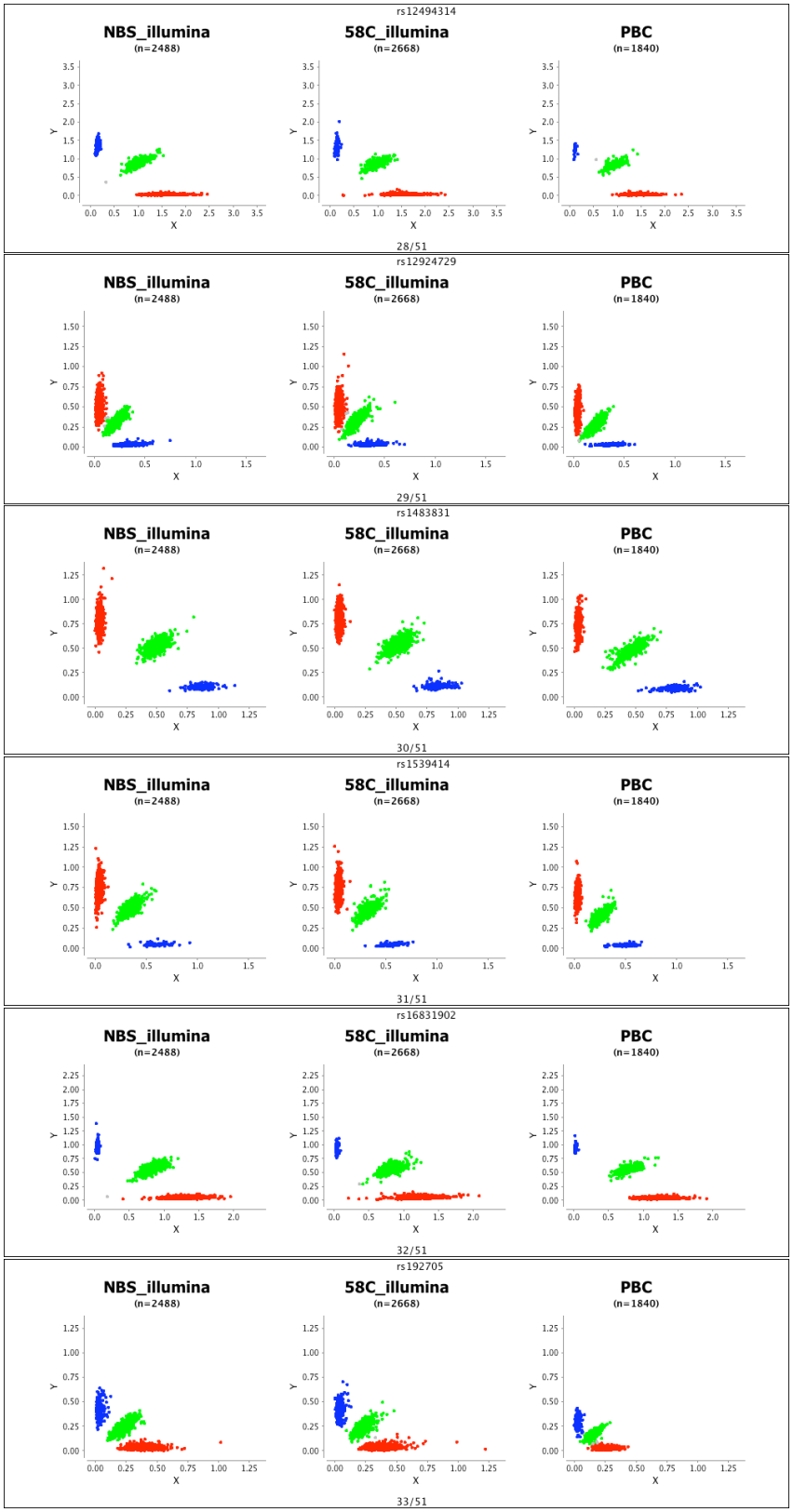
Figure 7: Cluster plots for all SNPs taken forward for replication, created using Evoker [13].

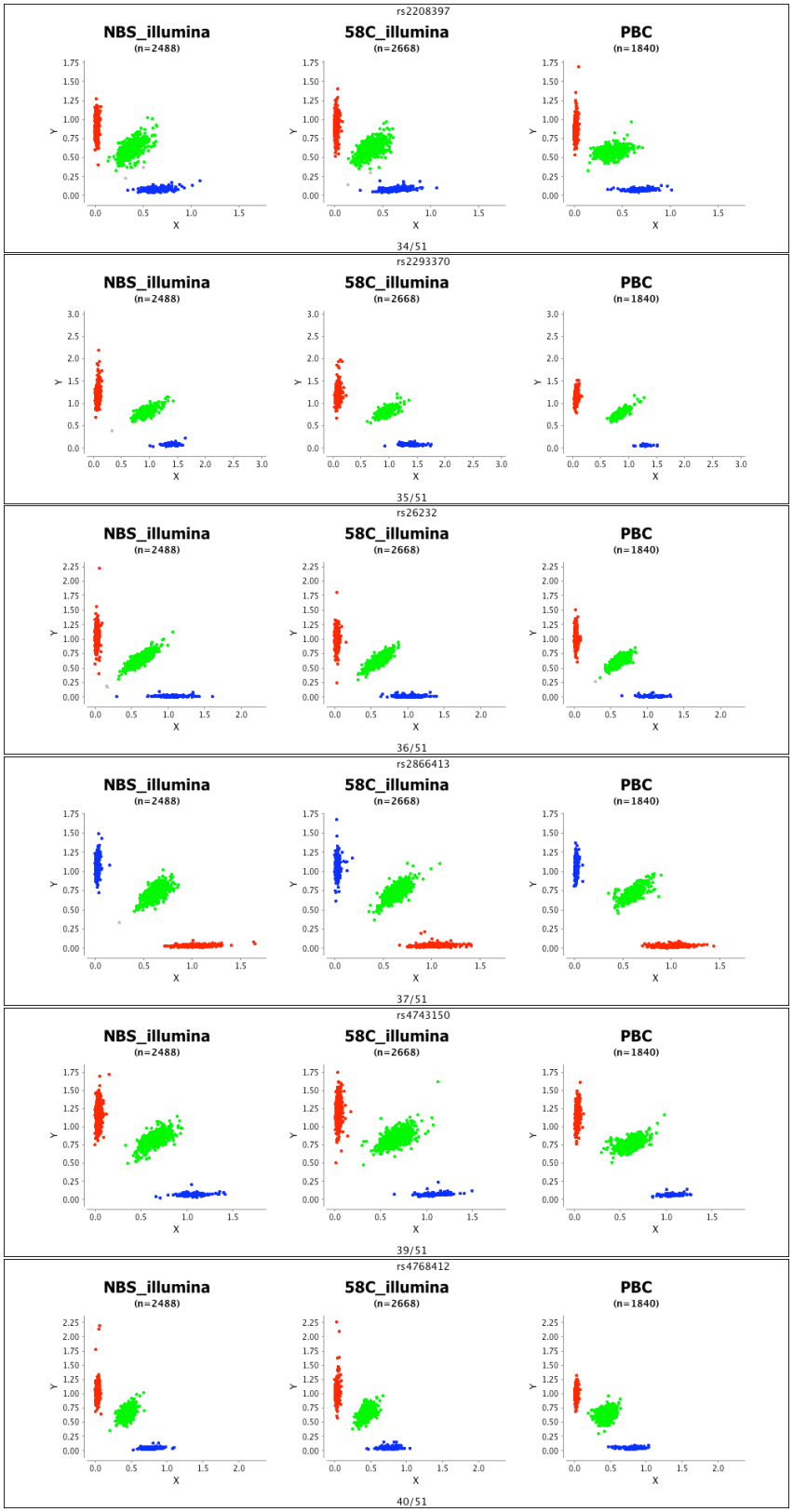


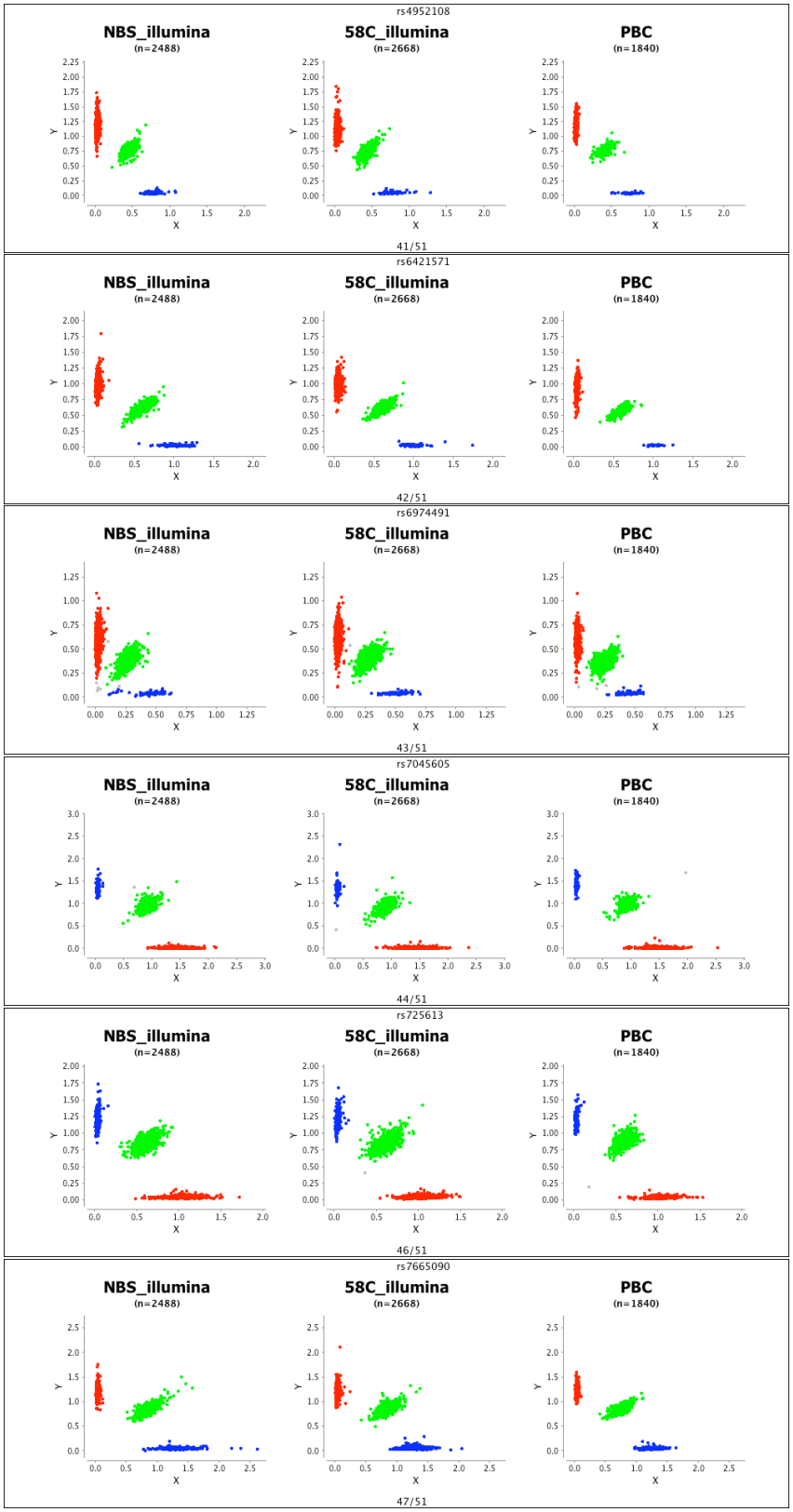

rs3738863
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
8/51
rs4149581
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
9/51
rs4938573
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
11/51
rs6043722
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
12/51
rs618671
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
13/51
rs674726
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
14/51

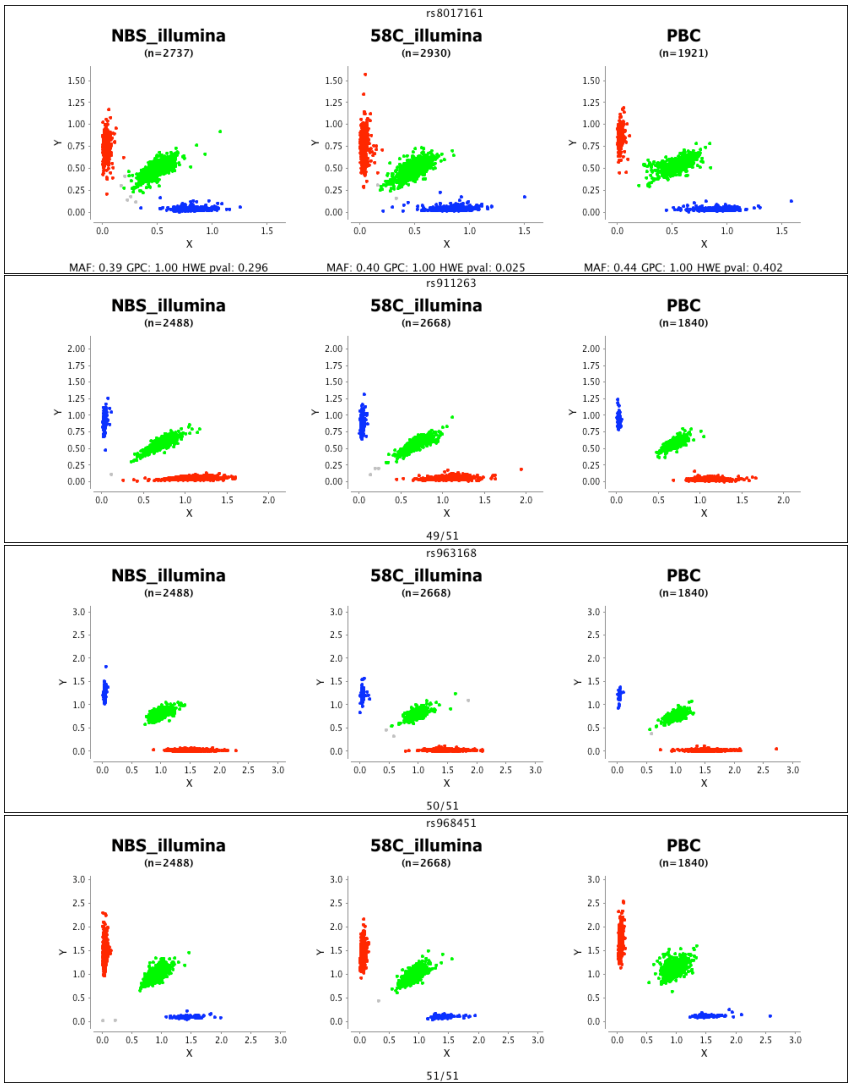

rs6897932
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
15/51
rs8070723
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
17/51
rs860413
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
18/51
rs874610
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
19/51
rs9594738
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
20/51
rs1003643
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
21/51


rs10931468
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
22/51
rs11117432
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
23/51
rs9533108
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
24/51
rs11586136
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
25/51
rs11692152
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
26/51
rs12134279
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
27/51


rs12494314
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
28/51
rs12924729
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
29/51
rs1483831
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
30/51
rs1539414
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
31/51
rs16831902
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
32/51
rs192705
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
33/51


rs2208397
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
34/51
rs2293370
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
35/51
rs26232
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
36/51
rs2866413
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
37/51
rs4743150
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
39/51
rs4768412
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
40/51


rs4952108
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
41/51
rs6421571
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
42/51
rs6974491
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
43/51
rs7045605
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
44/51
rs725613
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
46/51
rs7665090
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
X
Y
47/51


rs8017161
NBS_illumina
(n=2737)
58C_illumina
(n=2930)
PBC
(n=1921)
MAF: 0.39 GPC: 1.00 HWE pval: 0.296
MAF: 0.40 GPC: 1.00 HWE pval: 0.025
MAF: 0.44 GPC: 1.00 HWE pval: 0.402
rs911263
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
49/51
rs963168
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
50/51
rs968451
NBS_illumina
(n=2488)
58C_illumina
(n=2668)
PBC
(n=1840)
51/51

Consortium Membership

The UK PBC Consortium

Management committee: Graeme Alexander[1], Heather Cordell[2], Peter Donaldson[3], Michael Heneghan[4], David Jones[3], George Mells[1,5], James Neuberger[6], Collette Thain[7], Richard Sandford[5] (chair).

1. Department of Hepatology, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ
2. Institute of Human Genetics, Newcastle University, Newcastle upon Tyne NE1 3BZ
3. Institute of Cellular Medicine, Medical School, Newcastle University, Newcastle upon Tyne NE2 4HH
4. Institute of Liver Studies, King's College Hospital NHS Foundation Trust, Denmark Hill, London SE5 9RS
5. Academic Department of Medical Genetics, Cambridge University, Cambridge CB2 0QQ
6. The Liver Unit, Queen Elizabeth Hospital, Birmingham B15 2TH
7. The PBC Foundation, 54 Queen Street, Edinburgh EH2 3NS

Collaborating centres in the UK PBC Consortium: Ashford and St Peter's Hospitals NHS Trust (**Dr John Thornton**), Barnet and Chase Farm Hospitals NHS Trust (**Dr Stephen Mann**), Barts and The London NHS Trust (**Dr Richard Marley**), Basingstoke and North Hampshire NHS Foundation Trust (**Dr John Ramage**), Bedford Hospitals NHS Trust (**Dr Rory Harvey**), Brighton and Sussex University Hospitals NHS Trust (**Dr Jeremy Tibble**), Cambridge University Hospitals NHS Foundation Trust (**Dr George Mells, Dr Muhammad Dawwas, Dr Graeme Alexander**), Colchester Hospital University NHS Foundation Trust (**Dr Ian Gooding**), Dartford And Gravesham NHS Trust (**Dr Roland Ede**), Derby Hospitals NHS Foundation Trust (**Dr Andrew Austin**), East and North Hertfordshire NHS Trust (**Dr Martyn Carter, Dr Peter McIntyre**), East Sussex NHS Trust (**Dr David Neal**), Epsom and St Helier University Hospitals NHS Trust (**Dr Guan Lim**), Frimley Park NHS Foundation Trust (**Dr Aftab Ala**), Guy's and St Thomas' NHS Trust (**Dr Mark Wilkinson**), Heatherwood & Wexham Park Hospitals NHS Trust (**Dr Sass Levi**), Heart of England NHS Foundation Trust (**Dr Theodore Ngatchu**), Hinchingbrooke Health Care NHS Trust (**Dr Richard Dickinson**), Homerton University Hospital NHS Foundation Trust (**Dr Ray Shidrawi**), Imperial College Healthcare NHS Trust (**Dr Ashley Brown, Professor Salim Khakoo**), Ipswich Hospital NHS Trust (**Dr Simon Williams**), James Paget University Hospitals NHS Foundation Trust (**Dr Matthew Williams**), Kettering General Hospital NHS Foundation Trust (**Dr Andrew Chilton**), Kings College Hospital NHS Foundation Trust (**Dr Rachel Westbrook, Dr Michael Heneghan**), Kingston Hospital NHS Trust (**Dr Chris Rodrigues**), Luton and Dunstable Hospital NHS Foundation Trust (**Dr Sambit Sen**), Maidstone and Tunbridge Wells NHS Trust (**Dr George Bird**), Medway NHS Foundation Trust (**Dr Gray Smith-Laing**), Milton Keynes Hospital NHS Foundation Trust (**Dr George MacFaul**), Newham University Hospital NHS Trust (**Dr Beate Mengelkoch**), Norfolk and Norwich University Hospitals NHS Foundation Trust (**Dr Hugh Kennedy**), Northampton General Hospital NHS Trust (**Dr Udi Shmueli**), Nottingham University Hospitals NHS Trust (**Dr Steve Ryder**), Peterborough Stamford Hospitals NHS Foundation Trust (**Dr Mary Ninkovic**), Royal Free Hampstead NHS Trust (**Professor Andrew Burroughs**), South London Healthcare NHS Trust (**Dr Howard Curtis, Dr Alastair McNair**), Southampton University Hospitals NHS Trust (**Dr Mark Wright**), Southend University Hospital NHS Foundation Trust (**Dr Gary Bray**), The Hillingdon Hospital NHS Trust (**Dr Sarah Lean**), The North West London Hospitals NHS Trust (**Dr Maxton Pitcher**), The Princess Alexandra Hospital NHS Trust (**Dr David Preston**), The Queen Elizabeth Hospital King's Lynn NHS Trust (**Dr Andrew Douds**), The Whittington Hospital NHS Trust (**Dr Voi Shim Wong**), United Lincolnshire Hospitals NHS Trust (**Dr Aditya Mandal**), University College London Hospitals NHS Foundation Trust (**Dr Steve Pereira**), University Hospitals Bristol NHS Foundation Trust (**Dr Fiona Gordon**), University Hospitals of Leicester NHS Trust (**Dr Allister Grant**), West Suffolk Hospitals NHS Trust (**Dr Simon Whalley**), Western Sussex Hospitals NHS Trust (**Dr Andy Li**), Whipps Cross University Hospital NHS Trust (**Dr Af Sawyerr**), Airedale NHS Trust (**Dr Chris Healey**), Barnsley Hospital NHS Foundation Trust (**Dr Kapil Kapur**), Blackpool, Fylde and Wyre Hospitals NHS Foundation Trusts (**Dr CJ Shorrock**), Bradford Teaching Hospitals NHS Trust (**Dr Paul Southern**), Calderdale And Huddersfield NHS Trust (**Dr Sue Jones**), City Hospitals Sunderland NHS Foundation Trust (**Dr Harriet Mitchison**), County Durham and Darlington NHS Foundation Trust (**Dr Tony Macklon**), East Lancashire Hospitals NHS Trust (**Dr Charles Grimley**), Gateshead Health NHS Foundation Trust (**Dr Athar Saeed**), Harrogate and District NHS Foundation Trust (**Dr Jo Ridpath**), Hull And East Yorkshire Hospitals NHS Trust (**Dr George Abouda**), Lancashire Teaching Hospitals NHS Foundation Trust (**Dr IM Drake**), Leeds Teaching Hospitals NHS Trust (**Dr Mervyn Davies,**

Dr Charles Millson), NHS Ayrshire & Arran (Dr Amir Shah), NHS Borders (**Dr Chris Evans**), NHS Dumfries & Galloway (**Dr Subrata Saha**), NHS Fife (**Dr Church**), NHS Forth Valley (**Dr Peter Bramley**), NHS Grampian (**Dr Andrew Fraser**), NHS Greater Glasgow and Clyde (**Dr Peter Mills**), NHS Highland (**Dr Tim Shallcross**), NHS Lanarkshire (**Dr Richard Crofton**), NHS Lothian (**Dr Andrew Bathgate**), North Cumbria University Hospitals NHS Trust (**Dr Chris McDonald**), North Tees And Hartlepool NHS Foundation Trust (**Dr Jane Metcalf**), Northumbria Healthcare NHS Trust (**Dr Mark Welfare**), Rotherham NHS Foundation Trust (**Dr Barbara Hoeroldt**), Scarborough And North East Yorkshire Health Care NHS Trust (**Dr Sathish Babu**), Sheffield Teaching Hospitals NHS Foundation Trust (**Dr Dermot Gleeson**), South Tees Hospitals NHS Trust (**Dr Andrew Douglas**), South Tyneside NHS Foundation Trust (**Dr Simon Panter**), The Newcastle upon Tyne Hospitals NHS Foundation Trust (**Professor David Jones**), University Hospitals of Morecambe Bay NHS Trust (**Dr Andrew Higham**), York Hospitals NHS Foundation Trust (**Dr Alastair Turnbull**), Abertawe Bro Morgannwg University NHS Trust (**Dr Chin Lye Ch'ng, Dr Lai, Dr Tom Yapp**), Aintree University Hospitals NHS Foundation Trust (**Dr Richard Sturgess**), Aneurin Bevan Health Board (**Dr Marek Czajkowski**), Belfast Health and Social Care Trust (**Dr Neil MacDougall**), Buckinghamshire NHS Trust (**Dr David Gorard**), Burton Hospitals NHS Trust (**Dr Altaf Palegwala**), Cardiff and Vale NHS Trust (**Dr Sunil Dolwani**), Central Manchester and Manchester Children's University Hospitals NHS Trust (**Dr Martin Prince**), Countess of Chester Hospital NHS Foundation Trust (**Dr Mazn Karmo**), Cwm Taf Health Board (**Dr Minesh Patel**), Doncaster and Bassetlaw Hospitals NHS Foundation Trust (**Dr Joanne Sayer**), Dorset County Hospitals NHS Foundation Trust (**Dr Chris Hovell**), Dudley Group of Hospitals NHS Trust (**Dr Neil Fisher**), East Cheshire NHS Trust (**Dr Konrad Koss**), George Eliot Hospital NHS Trust (**Dr Gordon Wood**), Gloucestershire Hospitals NHS Foundation Trust (**Professor Jonathan Brown**), Hereford Hospitals NHS Trust (**Dr Rupert Ransford**), Hywel Dda Health Board (**Dr Ian Rees, Dr Imroz Salam**), Mid Cheshire Hospitals NHS Foundation Trust (**Dr Kevin Yoong**), Mid Staffordshire General Hospitals NHS Trust (**Dr Ray Mathew**), NHS Isle Of Wight (**Dr Christopher Sheen**), North Bristol NHS Trust (**Dr Robert Przemioslo**), North Wales NHS Trust (**Dr Thiriloganathan Mathialahan**), North Wales NHS Trust (**Dr David Ramanaden, Dr Richard Evans**), North West Wales NHS Trust (**Dr Gasem**), Northern Devon Healthcare NHS Trust (**Dr Andrew Davis**), Oxford Radcliffe Hospitals NHS Trust (**Dr Jane Collier, Dr Roger Chapman**), Plymouth Hospitals NHS Trust (**Dr Matthew Cramp**), Portsmouth Hospitals NHS Trust (**Dr Patrick Goggin**), Royal Bolton Hospital NHS Foundation Trust (**Dr George Lipscomb**), Royal Bournemouth And Christchurch Hospitals NHS Foundation Trust (**Dr Earl Williams**), Royal Cornwall Hospitals NHS Trust (**Dr Hyder Hussaini**), Royal Devon and Exeter NHS Foundation Trust (**Dr Reuben Ayres**), Royal Liverpool And Broadgreen University Hospitals NHS Trust (**Dr Martin Lombard**), Royal United Hospital Bath NHS Trust (**Dr Duncan Robertson, Dr Mark Farrant**), Salisbury NHS Foundation Trust (**Dr Andrew Tanner**), Sandwell and West Birmingham Hospitals NHS Trust (**Dr Saket Singhal**), Shrewsbury and Telford Hospital NHS Trust (**Dr Jeff Butterworth**), South Devon Healthcare NHS Trust (**Dr Keith George**), South Warwickshire General Hospitals NHS Trust (**Dr Jeremy Shearman**), Southport And Ormskirk Hospital NHS Trust (**Dr Graham Butcher**), St Helens And Knowsley Hospitals NHS Trust (**Dr John Mclindon**), Stockport NHS Foundation Trust (**Dr Debashis Das**), The Royal Wolverhampton Hospitals NHS Trust (**Dr Matthew Brookes**), University Hospital Birmingham NHS Foundation Trust (**Professor James Neuberger**), University Hospital of North Staffordshire NHS Trust (**Dr Alison Brind**), University Hospital of South Manchester NHS Foundation Trust (**Dr Gill Watts**), University Hospitals Coventry and Warwickshire NHS Trust (**Dr Esther Unitt**), Walsall Hospitals NHS Trust (**Dr Mark Cox**), Weston Area Health NHS Trust (**Dr Andrew Bell**), Winchester And Eastleigh Healthcare NHS Trust (**Dr Harriet Gordon**), Wirral University Teaching Hospital NHS Foundation Trust (**Dr Amit Singhal**), Worcestershire Acute Hospitals NHS Trust (**Dr Ian Gee**), Wrightington, Wigan And Leigh NHS Trust (**Dr Yeng Ang**), Yeovil District Hospital NHS Foundation Trust (**Dr James Gotto**).

The WTCCC3 Consortium

DNA, Genotyping, and Informatics Group: Hannah Blackburn[1], Sarah Edkins[1], Mathew Gillman[1], Emma Gray[1], Sarah E. Hunt[1], Cordelia Langford[1], Simon Potter[1], Douglas Simpkin[1], Pamela Whittaker[1].

Data Analysis Group: Carl A. Anderson[1], Jeffrey C. Barrett[1], James A.B. Floyd[1], Christopher S. Franklin[1], Ralph McGinnis[1], Nicole Soranzo[1], Eleftheria Zeggini[1].

UK Blood Services Controls: Jennifer Sambrook[2], Jonathan Stephens[2], Willem H. Ouwehand[2].

1958 Birth Cohort Controls: Wendy L. McArdle[3], Susan M. Ring[3], David P. Strachan[4].

Management Committee: Graeme Alexander[5], Jeffrey C. Barrett[1], Cynthia M. Bulik[6], David Collier[7], Peter J. Conlon[8], Anna Dominiczak[9], Audrey Duncanson[10], Adrian Hill[11], Cordelia Langford[1], Graham Lord[12], Alexander P. Maxwell[13], Linda Morgan[14], Richard N. Sandford[15], Neil Sheerin[12], Nicole Soranzo[1], Fredrik O. Vannberg[11], Leena Peltonen[1] (chair).

1. The Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA
2. Division of Transfusion Medicine, Department of Haematology, University of Cambridge, NHSBT Cambridge Centre, Long Road, Cambridge, CB2 0PT
3. Department of Social Medicine, University of Bristol, Bristol BS8 2BN
4. St. George's University, Division of Community Health Sciences, London SW19 0RE
5. Department of Hepatology, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ
6. Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina
7. Institute of Psychiatry, King's College London, London SE5 8AF
8. Department of Nephrology, Beaumont Hospital, Dublin, Ireland and Royal College of Surgeons Dublin, Ireland
9. BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow G12 8TA
10. Gibbs Building, 215 Euston Road, London NW1 2BE
11. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX1 2JA
12. MRC Centre for Transplantation, King's College London, London SE1 9RT
13. Belfast City Hospital, Lisburn Road, Belfast BT9 7AB
14. School of Molecular Medical Sciences, University of Nottingham, Nottingham NG7 2UH
15. Academic Department of Medical Genetics, Cambridge University, Cambridge CB2 0QQ

References

[1] 1000 Genomes Project Consortium, R M Durbin, G R Abecasis, D L Altshuler, A Auton, L D Brooks LD, R A Gibbs, M E Hurles, and G A McVean. A map of human genome variation from population-scale sequencing. *Nature*, 467(7319):1061–1073, 2010.

[2] Jeffrey C Barrett, David G Clayton, P Concannon, B Akolkar, J Cooper, H Erlich, C Julier, G Morahan, J Nerup, C Nierras, V Plagnol, F Pociot, H Schuilenburg, D Smyth, H Stevens, J Todd, N Walker, S Rich, and The Type 1 Diabetes Genetics Consortium. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet*, May 2009.

[3] Jeffrey C Barrett, B Fry, Julian Maller, and Mark J Daly. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, 21(2):263–5, Jan 2005.

[4] Jeffrey C Barrett, Sarah Hansoul, Dan L Nicolae, Judy H Cho, Richard H Duerr, John D Rioux, Steven R Brant, Mark S Silverberg, Kent D Taylor, M Michael Barmada, Alain Bitton, Themistocles Dassopoulos, Lisa Wu Datta, Todd Green, Anne M Griffiths, Emily O Kistner, Michael T Murtha, Miguel D Regueiro, Jerome I Rotter, L Philip Schumm, A Hillary Steinhart, Stephan R Targan, Ramnik J Xavier, NIDDK IBD Genetics Consortium, Cécile Libioulle, Cynthia Sandor, Mark Lathrop, Jacques Belaiche, Olivier Dewit, Ivo Gut, Simon Heath, Debby Laukens, Myriam Mni, Paul Rutgeerts, André Van Gossum, Diana Zelenika, Denis Franchimont, Jean-Pierre Hugot, Martine de Vos, Severine Vermeire, Edouard Louis, Belgian-French IBD Consortium, Wellcome Trust Case Control Consortium, Lon R Cardon, Carl A Anderson, Hazel Drummond, Elaine R Nimmo, Tariq Ahmad, Natalie J Prescott, Clive M Onnie, Sheila A Fisher, Jonathan L Marchini, Jilur Ghori, Suzannah J Bumpstead, Rhian Gwilliam, Mark Tremelling, Panos Deloukas, John C Mansfield, Derek P Jewell, Jack Satsangi, Christopher G Matthew, Miles Parkes, Michel Georges, and Mark J Daly. Genome-wide association defines more than 30 distinct susceptibility loci for crohn's disease. *Nat Genet*, 40(8):955–62, Aug 2008.

[5] Wellcome Trust Case Control Consortium, Australo-Anglo-American Spondylitis Consortium (TASC), Paul R Burton, David G Clayton, Lon R Cardon, Nick Craddock, Panos Deloukas, Audrey Duncanson, Dominic P Kwiatkowski, Mark I McCarthy, Willem H Ouwehand, Nilesh J Samani, John A Todd, Peter Donnelly, Jeffrey C Barrett, Dan Davison, Douglas F Easton, David M Evans, Hin-Tak Leung, Jonathan L Marchini, Andrew P Morris, Chris C A Spencer, Martin D Tobin, Antony P Attwood, James P Boorman, Barbara Cant, Ursula Everson, Judith M Hussey, Jennifer D Jolley, Alexandra S Knight, Kerstin Koch, Elizabeth Meech, Sarah Nutland, Christopher V Prowse, Helen E Stevens, Niall C Taylor, Graham R Walters, Neil M Walker, Nicholas A Watkins, Thilo Winzer, Richard W Jones, Wendy L McArdle, Susan M Ring, David P Strachan, Marcus Pembrey, Gerome Breen, David St Clair, Sian Caesar, Katharine Gordon-Smith, Lisa Jones, Christine Fraser, Elaine K Green, Detelina Grozeva, Marian L Hamshere, Peter A Holmans, Ian R Jones, George Kirov, Valentina Moskivina, Ivan Nikolov, Michael C O'Donovan, Michael J Owen, David A Collier, Amanda Elkin, Anne Farmer, Richard Williamson, Peter McGuffin, Allan H Young, I Nicol Ferrier, Stephen G Ball, Anthony J Balmforth, Jennifer H Barrett, Timothy D Bishop, Mark M Iles, Azhar Maqbool, Nadira Yuldasheva, Alistair S Hall, Peter S Braund, Richard J Dixon, Massimo Mangino, Suzanne Stevens, John R Thompson, Francesca Bredin, Mark Tremelling, Miles Parkes, Hazel Drummond, Charles W Lees, Elaine R Nimmo, Jack Satsangi, Sheila A Fisher, Alastair Forbes, Cathryn M Lewis, Clive M Onnie, Natalie J Prescott, Jeremy Sanderson, Christopher G Matthew, Jamie Barbour, M Khalid Mohiuddin, Catherine E Todhunter, John C Mansfield, Tariq Ahmad, Fraser R Cummings, Derek P Jewell, John Webster, Morris J Brown, Mark G Lathrop, John Connell, Anna Dominiczak, Carolina A Braga Marcano, Beverley Burke, Richard Dobson, Johannie Gungadoo, Kate L Lee, Patricia B Munroe, Stephen J Newhouse, Abiodun Onipinla, Chris Wallace, Mingzhan Xue, Mark Caulfield, Martin Farrall, Anne Barton, Biologics in RA Genetics, Genomics Study Syndicate (BRAGGS) Steering Committee, Ian N Bruce, Hannah Donovan, Steve Eyre, Paul D Gilbert, Samantha L Hilder, Anne M Hinks, Sally L John, Catherine Potter, Alan J Silman, Deborah P M Symmons, Wendy Thomson, Jane Worthington, David B Dunger, Barry Widmer, Timothy M Frayling, Rachel M Freathy, Hana Lango, John R B Perry, Beverley M Shields, Michael N Weedon, Andrew T Hattersley, Graham A Hitman, Mark Walker, Kate S Elliott, Christopher J Groves, Cecilia M Lindgren, Nigel W Rayner, Nicolas J Timpson, Eleftheria Zeggini, Melanie Newport, Giorgio Sirugo, Emily Lyons, Fredrik Vannberg, Linda A Bradbury, Claire Farrar, Jennifer J Pointon, Paul Wordsworth, Jayne A Franklyn, Joanne M Heward, Matthew J Simmonds, Stephen C L Gough, Sheila Seal, Breast Cancer Susceptibility Collaboration (UK), Michael R Stratton, Nazneen Rahman, Maria Ban, An Goris, Stephen J Sawcer, Alastair Compston, David Conway, Muminatou Jallow, Melanie Newport, Giorgio Sirugo, Kirk A Rockett, Suzannah J Bumpstead, Amy Chaney, Kate Downes, Mohammed J R Ghori, Rhian Gwilliam, Sarah E Hunt, Michael Inouye, Andrew Keniry, Emma King, Ralph McGinnis, Simon Potter, Rathi Ravindrarajah, Pamela Whittaker, Claire Widden, David Withers, Niall J Cardin, Dan Davison, Teresa Ferreira, Joanne Pereira-Gale, Ingeleif B Hallgrimsdo'ttir, Bryan N Howie, Zhan Su, Yik Ying Teo, Damjan Vukcevic, David Bentley, Alastair Compston, Martin Farrall, Alistair S Hall,

Andrew T Hattersley, Adrian V S Hill, Miles Parkes, Marcus Pembrey, Michael R Stratton, Sarah L Mitchell, Paul R Newby, Oliver J Brand, Jackie Carr-Smith, Simon H S Pearce, A Keniry, John D Reveille, Xiaodong Zhou, Anne-Marie Sims, Alison Dowling, Jacqueline Taylor, Tracy Doan, John C Davis, Laurie Savage, Michael M Ward, Thomas L Learch, Michael H Weisman, and Matthew A Brown. Association scan of 14,500 nonsynonymous snps in four diseases identifies autoimmunity variants. *Nat Genet*, 39(11):1329–37, Nov 2007.

[6] Patrick C A Dubois, Gosia Trynka, Lude Franke, Karen A Hunt, Jihane Romanos, Alessandra Curtotti, Alexandra Zhernakova, Graham A R Heap, Róza Ádány, Arpo Aromaa, Maria Teresa Bardella, Leonard H van den Berg, Nicholas A Bockett, Emilio G de la Concha, Bárbara Dema, Rudolf S N Fehrmann, Miguel Fernández-Arquero, Szilvia Fiatal, Elvira Grandone, Peter M Green, Harry J M Groen, Rhian Gwilliam, Roderick H J Houwen, Sarah E Hunt, Katri Kaukinen, Dermot Kelleher, Ilma Korponay-Szabo, Kalle Kurppa, Padraic Macmathuna, Markku Mäki, Maria Cristina Mazzilli, Owen T McCann, M Luisa Mearin, Charles A Mein, Muddassar M Mirza, Vanisha Mistry, Barbara Mora, Katherine I Morley, Chris J J Mulder, Joseph A Murray, Concepción Núñez, Elvira Oosterom, Roel A Ophoff, Isabel Polanco, Leena Peltonen, Mathieu Platteel, Anna Rybak, Veikko Salomaa, Joachim J Schweizer, Maria Pia Sperandeo, Greetje J Tack, Graham Turner, Jan H Veldink, Wieke H M Verbeek, Rinse K Weersma, Victorien M Wolters, Elena Urcelay, Bozena Cukrowska, Luigi Greco, Susan L Neuhausen, Ross McManus, Donatella Barisani, Panos Deloukas, Jeffrey C Barrett, Paivi Saavalainen, Cisca Wijmenga, and David A van Heel. Multiple common variants for celiac disease influencing immune gene expression. *Nat Genet*, 42(4):295–302, Jan 2010.

[7] Vesela Gateva, Johanna K Sandling, Geoff Hom, Kimberly E Taylor, Sharon A Chung, Xin Sun, Ward Ortmann, Roman Kosoy, Ricardo C Ferreira, Gunnel Nordmark, Iva Gunnarsson, Elisabet Svenungsson, Leonid Padyukov, Gunnar Sturfelt, Andreas Jönsen, Anders A Bengtsson, Solbritt Rantapää-Dahlqvist, Emily C Baechler, Elizabeth E Brown, Graciela S Alarcón, Jeffrey C Edberg, Rosalind Ramsey-Goldman, Gerald McGwin, John D Reveille, Luis M Vilá, Robert P Kimberly, Susan Manzi, Michelle A Petri, Annette Lee, Peter K Gregersen, Michael F Seldin, Lars Rönnblom, Lindsey A Criswell, Ann-Christine Syvänen, Timothy W Behrens, and Robert R Graham. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. *Nat Genet*, 41(11):1228–33, Nov 2009.

[8] Jian-Wen Han, Hou-Feng Zheng, Yong Cui, Liang-Dan Sun, Dong-Qing Ye, Zhi Hu, Jin-Hua Xu, Zhi-Ming Cai, Wei Huang, Guo-Ping Zhao, Hong-Fu Xie, Hong Fang, Qian-Jin Lu, Jian-Hua Xu, Xiang-Pei Li, Yun-Feng Pan, Dan-Qi Deng, Fan-Qin Zeng, Zhi-Zhong Ye, Xiao-Yan Zhang, Qing-Wen Wang, Fei Hao, Li Ma, Xian-Bo Zuo, Fu-Sheng Zhou, Wen-Hui Du, Yi-Lin Cheng, Jian-Qiang Yang, Song-Ke Shen, Jian Li, Yu-Jun Sheng, Xiao-Xia Zuo, Wei-Fang Zhu, Fei Gao, Pei-Lian Zhang, Qing Guo, Bo Li, Min Gao, Feng-Li Xiao, Cheng Quan, Chi Zhang, Zheng Zhang, Kun-Ju Zhu, Yang Li, Da-Yan Hu, Wen-Sheng Lu, Jian-Lin Huang, Sheng-Xiu Liu, Hui Li, Yun-Qing Ren, Zai-Xing Wang, Chun-Jun Yang, Pei-Guang Wang, Wen-Ming Zhou, Yong-Mei Lv, An-Ping Zhang, Sheng-Quan Zhang, Da Lin, Yi Li, Hui Qi Low, Min Shen, Zhi-Fang Zhai, Ying Wang, Feng-Yu Zhang, Sen Yang, Jian-Jun Liu, and Xue-Jun Zhang. Genome-wide association study in a chinese han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat Genet*, 41(11):1234–7, Nov 2009.

[9] International Multiple Sclerosis Genetics Consortium (IMSGC). Refining genetic associations in multiple sclerosis. *Lancet Neurology*, 7(7):567–9, Jul 2008.

[10] International Multiple Sclerosis Genetics Consortium (IMSGC). Comprehensive follow-up of the first genome-wide association study of multiple sclerosis identifies KIF21B and TMEM39A as susceptibility loci. *Hum Mol Genet*, 19(5):953–62, Mar 2010.

[11] Philip L De Jager, Xiaoming Jia, Joanne Wang, Paul I W de Bakker, Linda Ottoboni, Neelum T Aggarwal, Laura Piccio, Soumya Raychaudhuri, Dong Tran, Cristin Aubin, Rebeccah Briskin, Susan Romano, International MS Genetics Consortium, Sergio E Baranzini, Jacob L Mccauley, Margaret A Pericak-Vance, Jonathan L Haines, Rachel A Gibson, Yvonne Naeglin, Bernard Uitdehaag, Paul M Matthews, Ludwig Kappos, Chris Polman, Wendy L McArdle, David P Strachan, Denis Evans, Anne H Cross, Mark J Daly, Alastair Compston, Stephen J Sawcer, Howard L Weiner, Stephen L Hauser, David A Hafler, and Jorge R Oksenberg. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet*, 41(7):776–82, Jul 2009.

[12] Andrew D Johnson, Robert E Handsaker, Sara L Pulit, Marcia M Nizzari, Christopher J O'Donnell, and Paul I W de Bakker. SNAP: a web-based tool for identification and annotation of proxy snps using hapmap. *Bioinformatics*, 24(24):2938–9, Dec 2008.

[13] James A Morris, Joshua C Randall, Julian B Maller, and Jeffrey C Barrett. Evoker: a visualization tool for genotype intensity data. *Bioinformatics*, 26(14):1786–7, Jul 2010.

[14] Rajan P Nair, Kristina Callis Duffin, Cynthia Helms, Jun Ding, Philip E Stuart, David E Goldgar, Johann E Gudjonsson, Yun Li, Trilokraj Tejasvi, Bing-Jian Feng, Andreas Ruether, Stefan Schreiber, Michael Weichenthal, Dafna Gladman, Proton Rahman, Steven J Schrodi, Sampath Prahalad, Stephen L Guthery, Judith Fischer, Wilson Liao, Pui-Yan Kwok, Alan Menter, G Mark Lathrop, Carol A Wise, Ann B Begovich, John J Voorhees, James T Elder, Gerald G Krueger, Anne M Bowcock, Gonçalo R Abecasis, and Collaborative Association Study of Psoriasis. Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. *Nat Genet*, 41(2):199–204, Feb 2009.

[15] Mark S Silverberg, Judy H Cho, John D Rioux, Dermot P.B Mcgovern, Jing Wu, Vito Annese, Jean-Paul Achkar, Philippe Goyette, Regan Scott, Wei Xu, M Michael Barmada, Lambertus Klei, Mark J Daly, Clara Abraham, Theodore M Bayless, Fabrizio Bossa, Anne M Griffiths, Andrew F Ippoliti, Raymond G Lahaie, Anna Latiano, Pierre Paré, Deborah D Proctor, Miguel D Regueiro, A Hillary Steinhart, Stephan R Targan, L Philip Schumm, Emily O Kistner, Annette T Lee, Peter K Gregersen, Jerome I Rotter, Steven R Brant, Kent D Taylor, Kathryn Roeder, and Richard H Duerr. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nat Genet*, 41(2):216–20, Feb 2009.

[16] Eli A Stahl, Soumya Raychaudhuri, Elaine F Remmers, Gang Xie, Stephen Eyre, Brian P Thomson, Yonghong Li, Fina A S Kurreeman, Alexandra Zhernakova, Anne Hinks, Candace Guiducci, Robert Chen, Lars Alfredsson, Christopher I Amos, Kristin G Ardlie, BIRAC Consortium, Anne Barton, John Bowes, Elisabeth Brouwer, Noel P Burtt, Joseph J Catanese, Jonathan Coblyn, Marieke J H Coenen, Karen H Costenbader, Lindsey A Criswell, J Bart A Crusius, Jing Cui, Paul I W de Bakker, Philip L De Jager, Bo Ding, Paul Emery, Edward Flynn, Pille Harrison, Lynne J Hocking, Tom W J Huizinga, Daniel L Kastner, Xiayi Ke, Annette T Lee, Xiangdong Liu, Paul Martin, Ann W Morgan, Leonid Padyukov, Marcel D Posthumus, Timothy R D J Radstake, David M Reid, Mark Seielstad, Michael F Seldin, Nancy A Shadick, Sophia Steer, Paul P Tak, Wendy Thomson, Annette H M van der Helm-van Mil, Irene E van der Horst-Bruinsma, C Ellen van der Schoot, Piet L C M van Riel, Michael E Weinblatt, Anthony G Wilson, Gert Jan Wolbink, B Paul Wordsworth, YEAR Consortium, Cisca Wijmenga, Elizabeth W Karlson, Rene E M Toes, Niek de Vries, Ann B Begovich, Jane Worthington, Katherine A Siminovitch, Peter K Gregersen, Lars Klareskog, and Robert M Plenge. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet*, 42(6):508–14, Jun 2010.