Common polygenic variation contributes to risk of schizophrenia that overlaps with bipolar disorder: Supplementary information

International Schizophrenia Consortium (ISC), 23rd April 2009

Single SNP association analysis

Multilocus analysis of common variation

Notation

In this document, we use lower-case r^2 to represent coefficients of linkage disequilibrium (LD) and upper case R^2 to represent measures of variance explained from logistic regression (specifically, Nagelkerke's pseudo-R^2 measure).

1) Study samples, genotyping, quality control and population stratification

Sample ascertainment and collection procedures have been described elsewhere (ISC, 2008). Briefly, cases were diagnosed based on either the DSM-IV (American Psychiatric Association, 2000), ICD-10 (Janca et al, 1993) or through hospital records (Lichtenstein et al, 2006). Controls were selected from the general population at each site. Controls from Aberdeen and The Portuguese Island Collection were screened for psychiatric disorders (ISC, 2008).

Genotyping was performed using the Affymetrix Mapping 500K Array and the Genome-wide Human SNP Array 5.0 or 6.0 as described elsewhere (ISC, 2008, Sklar et al, 2008; Ferreira et al, 2008). Genotypes were called from raw intensity data using the Birdseed component of the Birdsuite algorithm (McCarroll et al, 2008; Korn et al, 2008). Genotyping was performed by the Genetic Analysis Platform of the Broad Institute of Harvard and MIT.

As shown in Table S1, there are seven distinct sample collection sites; subsets of the Swedish sample were genotyped on either the 5.0 or 6.0 array. Therefore, for the primary association analyses, we conditioned on eight strata (i.e. seven sites, with two Swedish strata). For the multilocus analyses (which were restricted to high quality SNPs present on all array types) the two Swedish groups were pooled to form a single stratum, resulting in a total of seven strata.

Table S1: *Number of individuals (N=6,909) and SNPs (N=739,995) in the final (post-QC) dataset, stratified by sample collection and genotyping array.*

Sample	Ancestry	Cases (N)	Controls (N)	Genotyping Array	SNP (N)
University of Aberdeen	Scottish	720	702[a]	5.0	380,572
University College London	British	523	505	5.0/500K[b]	302,491
Portuguese Island Collection	Portuguese	347	216[a]	5.0	379,468
Karolinska Institutet	Swedish	170	170	5.0	380,573
Karolinska Institutet	Swedish	390	230	6.0	725,282
Cardiff University	Bulgarian	528	611	6.0	724,458
Trinity College Dublin	Irish	275	866	6.0	724,132
University of Edinburgh	Scottish	369	287	6.0	727,092

The number of cases and controls from each site included in the final analysis are listed. Genotyping platform indicates either Affymetrix Mapping 500K Array (500K) or Affymetrix Genome-wide Human SNP Array type (5.0 or 6.0). SNP refers to the number of SNPs for a given site that passed QC.

[a] Controls were screened for psychiatric disorders.

[b] Controls were genotyped on the 500K array and cases were genotyped on the 5.0 array.

All data quality control (QC) and analysis were performed using the PLINK software package (Purcell et al, 2007; http://pngu.mgh.harvard.edu/purcell/plink).

Step 1: Remove bad SNPs and individual samples
Prior to per-individual QC, we removed grossly-failing or uninformative SNPs that:

- had a call rate of less than 50%
- mapped to multiple locations
- were monomophic
- had shown consistently poor performance in previous studies (data not shown)

The initial raw dataset consisted of 3,798 cases and 3,998 controls. From this, we removed individual DNA samples with call rates of less than 95% or clearly incompatible levels of X chromosome heterozygosity.

Step 2: Detect and resolve array technical artefacts
We next performed a series of steps to detect and resolve potential technical differences in SNP performance or annotation between data generated on the 500K, 5.0 and 6.0 arrays:

- A comparison of allele frequencies between samples, within phenotype class and array type. A SNP was removed from a sample if it showed marked differences in allele frequency when comparing all cases in that sample to all cases from the other seven samples ($P < 1x10^{-6}$) *but did not* show a similar difference for controls ($P > 0.05$), or vice versa.
- A comparison of allele frequencies for UCL controls (genotyped on 500K) against Aberdeen controls (genotyped on 5.0); SNPs were removed from UCL only for tests with $P < 1x10^{-6}$

The UCL/Aberdeen control comparison was particularly important, as UCL is the only sample in which cases and controls were not genotyped on the identical platform at this same point in time. We therefore took a slightly conservative approach, additionally removing SNPs based on this extra QC step (Table S1) effectively removing a number of clearly spurious, highly significant associations that were specific to the UCL sample. These likely represent SNPs that have different performance characteristics between array types.

We also identified a subset of alleles that appeared to have undergone strand flips between arrays, based either on allele frequency, e.g. a ~20% versus an ~80% allele, or patterns of linkage disequilibrium, using the approach implemented in PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/dataman.shtml#flipscan). These were flipped to be consistent. The remaining SNPs with extreme differences between arrays that were not consistent with a strand flip (and so likely represent technical array-specific artefact) were removed (N=19).

Step 3: Quality control for individuals
The next set of QC steps were performed separately for each of the eight samples in Table S1. Specifically, in total we removed 625 samples based on the following criteria:

- outliers with respect to estimated heterozygosity (visual inspection)

- nearest neighbour (based on genome-wide identity-by-state (IBS) similarity) was more than 3 standard deviations from the sample mean nearest neighbour statistic
- 2nd degree relative or closer in the sample, based on estimated genome-wide identity-by-descent (IBD), estimated within each sample (Purcell et al, 2008)
- evidence of distant relatedness with multiple other individuals in the sample, which is often indicative of sample contamination
- appeared to be duplicated

Step 4: Quality control for SNPs

We next performed a SNP-level set of QC steps, removing SNPs with:

- Hardy-Weinberg $P < 1\times10^{-6}$ in controls
- minor allele frequency less than 0.01
- significantly different call rate between cases and controls ($P < 1\times10^{-6}$)
- non-random genotyping failure, as inferred by the flanking haplotypic background (PLINK mishap test, $P < 1\times10^{-10}$)
- evidence of gross non-random plate failure, based on a comparison of allele frequency of each plate to all others (if $P<1\times10^{-6}$; only genotypes for this plate, rather than the whole SNP, were removed from the final dataset).

Overall, these steps resulted in a dataset with 739,995 SNPs (from 862,921) for which there was data for at least one of the samples.

Step 5: Population stratification analysis

We next calculated identity-by-state (IBS) statistics for each pair of remaining samples and performed a multi-dimensional scaling (MDS) analysis. Based on the preliminary MDS analysis, we removed a further 259 individuals that appeared to be outliers with respect to their known population and a small numbers of individuals that formed small clusters. As illustrated in Figure S1, in the final dataset individuals cluster according to their known site; the four samples from the British Isles are largely overlapping, based on the first two MDS components. The final dataset consisted of 3,322 cases and 3,587 controls (Table S1).

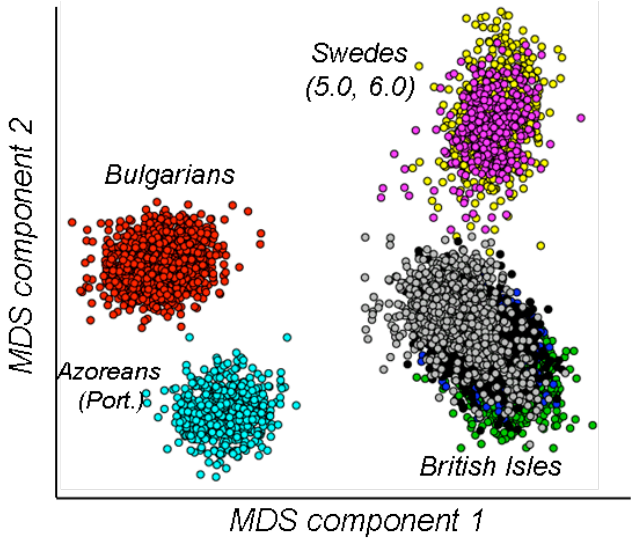


Figure S1: *Multidimensional scaling (MDS) plot for the individuals in the final post-QC dataset (both cases and controls). Known study samples are indicated by colour; the distinct clusters are labeled with the exception of the four British Isles samples (from Scotland, Ireland and England) that show near complete overlap on the first two dimensions.*

2) Primary association analysis of genotyped SNPs

The primary analysis of genotyped SNPs was a Cochran-Mantel-Haenszel (CMH) statistic as implemented in PLINK (Purcell et al, 2007), conditioning on the eight strata defined above. We also performed a Breslow-Day test for each SNP to assess the degree of heterogeneity of association across the eight analysis groups. Figure S2 shows the Q-Q plot for all SNPs. The genomic inflation factor (Devlin & Roeder, 1999) was 1.09.

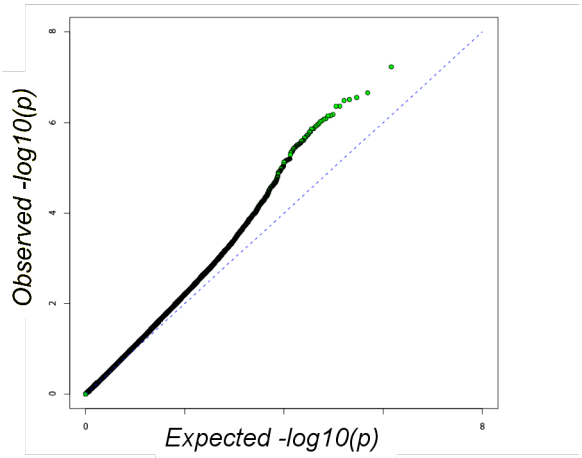


Figure S2: *Q-Q plot for primary single SNP statistics, based on the CMH test conditioning on eight strata as described above.*

The deviation from the diagonal in Figure S2 mostly reflects a large number of SNPs on chromosome 6p that are in a region of extensive linkage equilibrium and also associated with schizophrenia. When based on a subset of SNPs selected to be in approximate linkage equilibrium (as used in the score analyses, see sections S8-S17 below), the Q-Q plot does not show this trend to such a large extent (Figure S7, below). Additionally, given our score analyses that point to a highly polygenic model for disease risk, we would not expect the lambda to approach precisely 1.00 in the absence of any confounding, as explained below (Section S16).

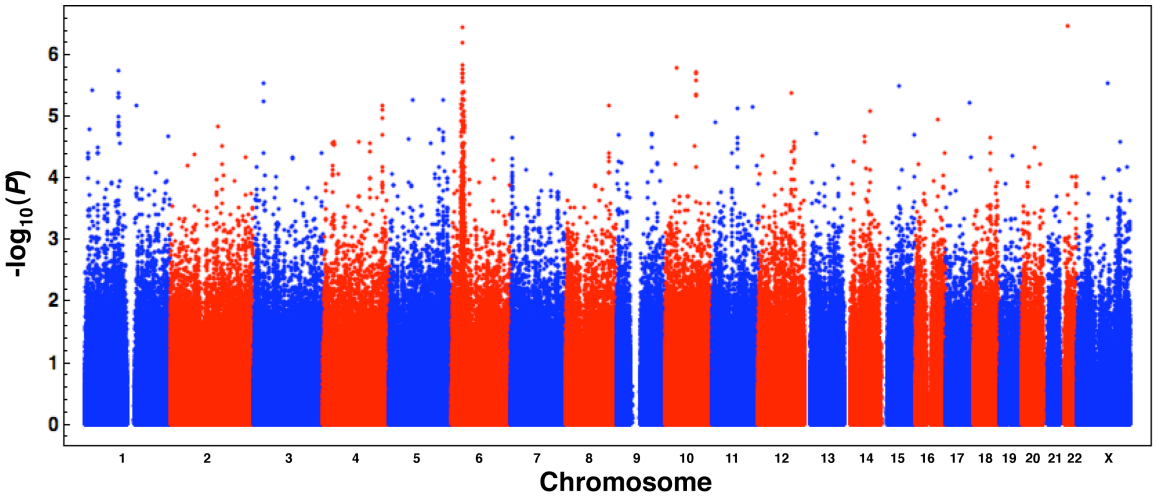


Figure S3: *Manhattan plot of single SNP Cochran-Mantel-Haenszel (CMH) test statistics, conditioning on the eight strata described above.*

Compared to the primary CMH tests, the results did not significantly vary when analyses were performed using a more finely matched IBS clustering within each site, or when including the first four components from the MDS analysis in a logistic regression of disease on genotype and covariates (data not shown).

Table S2 contains the associated regions, $P < 10^{-4}$. To present the single SNP results in terms of associated "clumps" rather than a list of SNPs, we used an LD-based clumping approach as implemented in PLINK. Specifically, the clumping procedure takes all SNPs that are significant at $P < 10^{-4}$ that have not already been clumped (denoting these as *index SNPs*) and forms clumps of all other SNPs that are within 10Mb of the index SNP and that are in linkage disequilibrium with the index SNP, based on an $r^2>0.1$, and significant at $P<0.01$. The large 10Mb threshold was used to because of the MHC association. The approach groups SNPs in LD-space rather than physical distance: as such clumps could overlap completely or partially and still represent independent associations. Although the MHC region contains multiple clumps, subsequent analyses that explicitly address the issue of independent effects (see below), suggest for the most part this represents only an artefact of the clumping procedure. That is, the clumping procedure doesn't show optimal performance when confronted with a complex region of very extensive LD (as clumps must be defined based on a fixed r^2 threshold with a single index SNP). Therefore, the six clumps in the MHC region should not be interpreted as providing evidence for six independent association signals.

With the exception of the MHC, we did not observe strong evidence for between-site heterogeneity or for sex-specific effects. We did not find support for reports from several smaller GWAS of schizophrenia, although our data were consistent with an association of rs1344706 in *ZNF804A* as reported in O'Donovan et al. 2008 (see Section S5 below). Of the three deletion regions we previously reported in this sample, we observed association at 22q11 (region-wide corrected $P = 0.023$; Figure S5 and Table S4 in Section S6 below), including SNPs in phosphatidylinositol 4-kinase alpha (*PI4KA*) consistent with previous reports (Vorstman et al. 2008; Jungerius et al. 2008).

Supplementary Table 2. ***Association analysis for genotyped SNPs in regions with SNP P < 1×10^{-4}***

Rank	Chr	SNP	BP	N	Position	Size (kb)	Genes within 20 kb	MA	MAF	P(CMH)	OR(CMH)	P-BD
1	22	rs5761163	24469378	3	chr22:24464026..24493996	30.0	MYO18B,ADRBK2	G	0.257	3.44E-07	0.800	0.069
2	6	rs3130375	30429711	301	chr6:25527073..32821845	7294.8	*MHC region	A	0.118	3.66E-07	0.733	0.012
3	10	rs1187102	33304038	8	chr10:33051235..33353003	301.8	ITGB1,C10orf68	G	0.392	1.64E-06	1.279	0.082
4	1	rs11165690	96886319	25	chr1:96705324..97055621	350.3	PTBP2	C	0.191	1.78E-06	1.237	0.320
5	10	rs11201716	87399953	10	chr10:87280738..87624513	343.8	GRID1	C	0.043	1.88E-06	0.663	0.360
6	3	rs6779328	24992222	6	chr3:24947402..25022232	74.8		C	0.468	2.92E-06	0.761	0.462
7	15	rs10162662	52963194	2	chr15:52963194..53053822	90.6		C	0.035	3.17E-06	0.520	0.252
8	1	rs2473277	22234432	9	chr1:22234432..22404140	169.7	WNT4,CDC42	G	0.476	3.75E-06	0.850	0.182
9	12	rs6538780	96034099	18	chr12:95982407..96119518	137.1		T	0.345	4.13E-06	0.844	0.830
10	5	rs34691	66323639	5	chr5:66316763..66376854	60.1	MAST4	A	0.408	5.43E-06	1.213	0.517
11	5	rs984078	153457658	35	chr5:153218516..153544302	325.8	MFAP3,GALNT10,FAM114A2	A	0.186	5.51E-06	0.744	0.754
12	17	rs6501685	69559670	1	chr17:69559670..69559670	0.0		T	0.104	6.12E-06	1.570	0.404
13	1	rs12139286	148529079	8	chr1:148268523..148687063	418.5	VPS45,PRPF3,PLEKHO1,OTUD7B,MRPS21,KIAA0460,CA14,C1orf54,C1orf51,APH1A, ANP32E	A	0.126	6.53E-06	0.786	0.439
14	8	rs17360690	125335988	12	chr8:125332999..125502081	169.1	TMEM65	A	0.298	6.67E-06	0.841	0.913
15	4	rs10017932	170474235	23	chr4:170466582..170863127	396.5	NEK1,CLCN3	G	0.497	6.69E-06	1.171	0.186
16	11	rs4938268	115777549	1	chr11:115777549..115777754	0.0		G	0.442	6.92E-06	0.753	0.399
17	11	rs12099027	72111568	9	chr11:72068288..72485279	417.0	STARD10,PDE2A,FCHSD2,CENTD2,ATG16L2	C	0.332	7.54E-06	0.846	0.261
18	14	rs2058919	74380565	15	chr14:74178043..74434815	256.8	YLPM1,RPS6KL1,PROX2,KIAA0317,FCF1,DLST	G	0.373	8.25E-06	0.851	0.639
19	6	rs9271850	32703038	7	chr6:32423632..32703038	279.4	HLA-DRB1,HLA-DRB5,HLA-DQA1,HLA-DRA,C6orf10,BTNL2	G	0.315	8.34E-06	0.723	0.453
20	6	rs370520	28650499	19	chr6:28525201..29505758	980.6	ZSCAN23,ZNF311,TRIM27,SCAND3,OR5V1,OR2W1,OR2J2,OR2J3,OR2B3P,OR14J1, OR12D2,OR12D3,OR11A1, OR10C1,LOC651503,GPX6,GPX5	T	0.443	9.09E-06	0.854	0.590
21	16	rs7198295	65034445	3	chr16:65034445..65041307	6.9		T	0.497	1.11E-05	0.857	0.004
22	6	rs3734536	26473325	6	chr6:26453120..27165497	712.4	ZNF322A,HMGN4,BTN2A3,BTN3A2,BTN3A3,BTN3A1,BTN2A1,BTN2A2,BTN1A1,ABT1	C	0.369	1.18E-05	0.853	0.190
23	11	rs17403795	10537793	3	chr11:10483198..10558098	74.9	RNF141,MRVI1,LYVE1,AMPD3	A	0.116	1.22E-05	1.273	0.398
24	2	rs12614381	136807739	4	chr2:136792895..136808052	15.2		C	0.180	1.48E-05	0.820	0.918
25	5	rs3822398	142368154	17	chr5:142249012..142382403	133.4	ARHGAP26	C	0.194	1.59E-05	0.825	0.964
26	1	rs2878677	11975851	10	chr1:11964848..12005513	40.7	PLOD1,MFN2,IIP45	T	0.340	1.61E-05	1.257	0.904
27	9	rs2779562	100317502	13	chr9:100307748..100360403	52.7	GABBR2	C	0.488	1.86E-05	0.861	0.643
28	13	rs8000946	36445260	6	chr13:36445260..36683910	238.7	FAM48A,EXOSC8,CSNK1A1L,ALG5	T	0.394	1.90E-05	1.167	0.600
29	9	rs10125618	6545311	2	chr9:6545311..6545417	0.1	GLDC	A	0.373	2.04E-05	1.247	0.836
30	15	rs12443391	96295482	5	chr15:96269059..96298976	29.9	ARRDC4	T	0.034	2.04E-05	0.534	0.411
31	14	rs10150328	58450585	9	chr14:58253028..58501856	248.8		C	0.148	2.08E-05	0.736	0.700
32	1	rs10926906	241116497	2	chr1:241116497..241121325	4.8		G	0.400	2.09E-05	0.804	0.310
33	7	rs10226475	2192688	29	chr7:1887352..2283378	396.0	SNX8,NUDT1,MAD1L1,FTSJ2	G	0.381	2.18E-05	0.857	0.639
34	18	rs1811441	49446590	19	chr18:49446590..49855111	408.5		A	0.274	2.24E-05	1.182	0.134
35	XY	rs553369	154616633	3	chrXY:154616633..15464382	27.2		T	0.486	2.28E-05	1.303	0.191
36	5	rs854041	57130457	3	chr5:57130457..57190980	60.5		T	0.013	2.33E-05	2.747	0.123
37	4	rs871061	102898360	9	chr4:102866586..102920255	53.7	BANK1	A	0.406	2.52E-05	1.179	0.963
38	4	rs7691359	29778591	6	chr4:29777078..29783327	6.2		G	0.276	2.56E-05	0.848	0.787
39	12	rs2468083	103704275	10	chr12:103669570..103775966	106.4	SLC41A2,CHST11	A	0.296	2.62E-05	1.176	0.413
40	23	rs5932307	126993095	20	chrX:126365396..127252479	887.1	ACTRT1	A	0.063	2.62E-05	1.715	0.714
41	5	rs1594002	120810165	3	chr5:120794627..120813081	18.5		C	0.024	2.68E-05	2.011	0.632
42	6	rs7748270	32556577	3	chr6:32556107..32716055	159.9	HLA-DRB1,HLA-DRB5,HLA-DQA1,HLA-DQB1	T	0.423	2.74E-05	1.283	0.486
43	4	rs1991976	133706306	13	chr4:133675794..133767613	91.8		G	0.164	2.74E-05	1.330	0.960
44	1	rs11164661	103313919	12	chr1:103192526..103503288	310.8	COL11A1	A	0.169	2.76E-05	1.329	0.504
45	4	rs4586917	26050597	9	chr4:25866797..26050597	183.8	RBPJ	T	0.085	2.77E-05	1.332	0.621
46	10	rs1113145	85608304	1	chr10:85608304..85608304	0.0		C	0.219	3.00E-05	1.213	0.596
47	2	rs4667369	149600091	3	chr2:149558655..149614988	56.3	LOC130576,KIF5C	C	0.373	3.05E-05	0.860	0.409
48	1	rs1002655	36965264	20	chr1:36919053..37012625	93.6		C	0.315	3.20E-05	1.170	0.408
49	20	rs211863	37887581	1	chr20:37887581..37887581	0.0		A	0.098	3.28E-05	1.419	0.396
50	3	rs4441603	195395782	3	chr3:195395499..195396246	0.7		T	0.420	3.92E-05	0.809	0.157
51	11	rs11570190	57317028	17	chr11:57237199..57465722	228.5	ZDHHC5,TXNDC14,MED19, CTNND1,C11orf31	C	0.344	3.93E-05	1.164	0.941
52	16	rs4843177	85186877	3	chr16:85184943..85189794	4.9	FOXL1	A	0.444	4.00E-05	0.812	0.638
53	1	rs172531	8418177	8	chr1:8346097..8891806	545.7	SLC45A1,RERE,ENO1	G	0.338	4.03E-05	1.165	0.343
54	2	rs741326	70912343	5	chr2:70897719..70914616	16.9	CLEC4F,CD207	G	0.442	4.14E-05	0.865	0.152
55	19	rs2041728	38881477	1	chr19:38881477..38881477	0.0	CHST8	A	0.160	4.35E-05	1.220	0.588
56	12	rs762721	11843691	5	chr12:11842382..11860837	18.5	ETV6	A	0.424	4.42E-05	0.812	0.481
57	6	rs10807124	33512042	11	chr6:33343733..33539547	195.8	ZBTB9,ZBTB22,WDR46,VPS52, TAPBP,SYNGAP1,RPS18,RGL2,PHF1,PFDN6,KIFC1,DAXX,CUTA, B3GALT4	A	0.278	4.49E-05	1.192	0.885
58	3	rs1164067	110948709	7	chr3:110627233..110966992	339.8		G	0.269	4.57E-05	0.851	0.632
59	17	rs750844	70632523	3	chr17:70632523..70697252	64.7	SUMO2,SLC16A5,NUP85,NT5C,HN1,ARMC7	A	0.290	4.58E-05	0.854	0.907
60	2	rs4621152	217617230	11	chr2:217571726..217644369	72.6		T	0.375	4.58E-05	0.863	0.119
61	6	rs7765368	119720155	16	chr6:119517635..119720155	202.5	MAN1A1,C6orf60	G	0.103	5.11E-05	0.790	0.733
62	9	rs10815532	7154061	15	chr9:7127500..7190905	63.4	JMJD2C	C	0.413	5.38E-05	0.866	0.979
63	14	*rs8018224	29894621	1	chr14:29894621..29894621	0.0		A	0.025	5.47E-05	1.775	4.59E-07
64	9	rs756624	118032202	4	chr9:118027711..118036771	9.1	PAPPA	C	0.394	5.66E-05	0.865	0.640
65	9	rs10738397	15441866	5	chr9:15419539..15459733	40.2	SNAPC3,PSIP1	A	0.187	5.80E-05	1.201	0.914
66	16	rs17144183	7571320	1	chr16:7571320..7571320	0.0	A2BP1	G	0.080	5.83E-05	1.436	0.002
67	10	rs2776632	30252761	3	chr10:30252761..30256536	3.8		A	0.441	5.85E-05	0.868	0.306
68	20	rs6127702	54265637	2	chr20:54264744..54265637	0.9	MC3R	A	0.119	5.91E-05	0.788	0.305
69	13	rs12584499	80651315	1	chr13:80651315..80651315	0.0		G	0.045	6.17E-05	1.839	0.480
70	7	rs7810949	4051242	12	chr7:4036402..4088371	52.0	SDK1	T	0.165	6.20E-05	0.828	0.609
71	11	rs7128882	129124689	1	chr11:129124689..129124689	0.0		T	0.095	6.23E-05	1.401	0.735
72	2	rs3850333	50856433	5	chr2:50752459..50894306	141.8	NRXN1	A	0.410	6.36E-05	0.867	0.007
73	3	rs358989	8385210	4	chr3:8378739..8396547	17.8		T	0.092	6.57E-05	1.410	0.778
74	20	rs7269093	20194569	3	chr20:20194569..20207855	13.3	C20orf26	C	0.430	6.68E-05	0.858	0.277
75	23	rs4827689	144991675	2	chrX:144940582..144991675	51.1		G	0.302	6.79E-05	0.822	0.359
76	14	rs7148664	65629502	3	chr14:65592567..65932496	339.9		T	0.029	6.97E-05	0.629	0.710
77	7	rs217426	44525384	2	chr7:44420900..44525384	104.5	NUDCD3,NPC1L1,LOC644907	C	0.036	7.20E-05	1.708	0.591
78	23	rs11796987	120853395	62	chrX:120552763..121043470	490.7		A	0.367	7.23E-05	0.778	0.131
79	14	rs4899649	76829350	6	chr14:76810128..76833519	23.4	TMEM63C,POMT2,NGB	T	0.198	7.31E-05	0.839	0.504
80	15	rs1194922	60582976	14	chr15:60550260..60651422	101.2		A	0.355	7.52E-05	0.812	0.226
81	X	rs6608182	123224785	2	chrX:123224785..123236187	11.4		C	0.383	7.54E-05	1.185	0.682
82	9	rs10738843	3050291	3	chr9:3050291..3055110	4.8		A	0.236	7.59E-05	0.791	0.785
83	6	rs12526001	8036524	8	chr6:8036524..8175770	139.2	EEF1E1	G	0.089	7.81E-05	0.783	0.238
84	5	rs6556290	157234667	6	chr5:157167055..157234760	67.7	CLINT1	C	0.193	7.96E-05	1.286	0.747
85	12	rs7955374	46166416	2	chr12:46166416..46174763	8.3		T	0.122	8.26E-05	1.239	0.922
86	1	rs12042196	203424382	11	chr1:203238075..203488282	250.2	TMEM81,TMCC2,RIPK5,RBBP5, NFASC,CNTN2	T	0.219	8.36E-05	1.198	0.294
87	4	rs16853762	41621125	1	chr4:41621125..41621125	0.0	TMEM33	C	0.020	8.52E-05	2.087	0.444
88	7	rs1918910	115746441	10	chr7:115643838..115776877	133.0	TES	C	0.443	8.56E-05	0.820	0.230
89	12	rs2238090	2553593	5	chr12:2517732..2553593	35.9	CACNA1C	A	0.317	8.58E-05	0.862	0.158
90	5	rs11242025	129845818	2	chr5:129777682..129845818	68.1		A	0.016	8.67E-05	2.326	0.201
91	5	rs924127	6887441	1	chr5:6887441..6887441	0.0		C	0.057	8.68E-05	1.390	0.773
92	2	rs7567863	153991620	5	chr2:153847699..153991620	143.9		C	0.247	9.00E-05	0.796	0.606
93	16	rs2316576	79327790	9	chr16:79304514..79345753	41.2	CDYL2	T	0.404	9.13E-05	0.818	0.023
94	3	rs7644703	30246650	9	chr3:30246650..30300551	53.9		C	0.244	9.24E-05	1.253	0.178
95	22	rs6520066	48680992	11	chr22:48561093..48700813	139.7	ZBED4,CRELD2,BRD1,ALG12	T	0.203	9.36E-05	1.186	0.243
96	22	rs2413398	35060893	1	chr22:35060893..35060893	0.0	MYH9	T	0.347	9.49E-05	1.227	0.552
97	3	rs349171	62026751	11	chr3:61986334..62056302	70.0	PTPRG	T	0.121	9.58E-05	0.811	0.781
98	5	rs7708885	168328507	14	chr5:168304390..168414893	110.5	SLIT3	T	0.368	9.82E-05	0.868	0.576
99	13	rs12853561	97993091	20	chr13:97911472..98048469	137.0	STK24,FARP1	C	0.481	9.92E-05	1.146	0.975
100	5	rs9327649	132616256	1	chr5:132616256..132616256	0.0	FSTL4	T	0.072	9.93E-05	0.681	0.533

Abbreviations: Chr – chromosome; BP – base pair position of SNP; P – CMH P value conditioned on 8 strata; N – Number of SNPs with r2<0.1 within 10 Mb with P < 0.01.

P-BD: Breslow-Day *P* for 8 strata.

3) Imputation of common SNPs and classical HLA alleles

We used the imputation procedure described in Ferreira et al (2008) and implemented in PLINK. The reference panel contained 2,379,233 filtered SNPs (MAF>0.01 and genotyping rate >95%) genotyped on the 60 HapMap CEU founders from HapMap release 23. For each reference SNP to be imputed, we selected a set of neighboring SNPs ("proxies") that were genotyped in the ISC. We then phased the proxy and reference SNPs using a standard EM algorithm on all samples (ISC and HapMap) jointly, to impute the unobserved genotype data.

For common reference SNPs (MAF>0.1), we selected up to 5 proxies from up to 15 flanking SNPs within 250 kb either side of the reference SNP. Proxies were selected to have at least 99% genotyping rate (i.e. present on both array types and in all samples). Based on prior simulation and experience with real data (dropping and re-imputing SNPs) we used the following heuristic. We selected proxies with the highest r^2 with the reference SNP but $r^2 < 0.5$ with an existing proxy; if two proxies were selected, any new proxy was required to have an $r^2 > 0.25$ with the reference. For less frequent reference SNPs (MAF<0.1 which are harder to impute), this last parameter was set to 0.01 instead of 0.25; additionally we searched up to 30 SNPs and 500kb, entering a maximum of 10 proxies. For each imputed SNP, the ratio of the observed to theoretical variance in dosage (based on allele frequency and assuming HWE) was used as an information score, to indicate the quality of imputation. Using this approach we also inferred genotypes for SNPs present on the 6.0 array, but not on the 5.0 array. Overall, a total of 1,639,653 HapMap SNPS were imputed with high confidence (information score $\geq$ 0.8). Additional description and evaluation of this method is presented in Ferreira et al (2008).

For imputed SNPs, the association analysis was a logistic regression of disease state on the expected fractional allele dosage, with seven dummy variables representing the eight strata entered as covariates. The Wald statistic for the dosage coefficient was the primary test statistic.

Classical HLA microsatellite alleles of the 60 CEU HapMap founders, described by de Bakker et al (2006), were obtained from the URL:
http://www.inflammgen.org/inflammgen/files/data/CEU+_HLAtypes.txt.
This dataset contains 96 alleles at six HLA genes: *A*, *B*, *C*, *DRB*, *DQA* and *DQB*. We recoded these alleles as 96 biallelic markers and merged this dataset with the SNP data for the HapMap CEU and ISC individuals. Using the LD relationships in the HapMap between the HLA alleles and the flanking HapMap SNPs that are also genotyped in the ISC, we imputed the HLA alleles for all ISC individuals, using the approach described above. Given the extended LD in this region, we increased the search space for proxies to span a greater distance.

Association of imputed classical HLA alleles within the full ISC sample was performed using logistic regression as described above. Of the 96 HLA alleles, 48 were imputed with high confidence (information score above 0.8) and an estimated allele frequency of 0.01 or above.

For the haplotype analysis of HLA alleles, we first imputed the most likely genotype for each of the loci. Genotypes with a maximum posterior probability of less than 0.8 were set to missing. We then used a standard E-M phasing approach to estimate haplotype frequencies in cases and controls. The association tests of HLA haplotype and disease were framed as a logistic regression on the expected number of HLA haplotypes, including sample as a dummy-coded covariate.

4) Conditional analyses of the MHC region and HLA alleles

The profile of LD (r^2) around the best genotyped SNP in the MHC (rs3130375) is shown in Figure S4 for the ISC overall, and also for each of the seven samples in which rs3130375 was genotyped. This SNP shows considerable LD (r^2>0.5) over multiple megabases of the MHC, although less so in the Portuguese samples. Consistent with the conditional analyses performed for SNPs and haplotypes in this region (see below), this suggests a single underlying locus drives the association in this region, as opposed to multiple, independent signals.

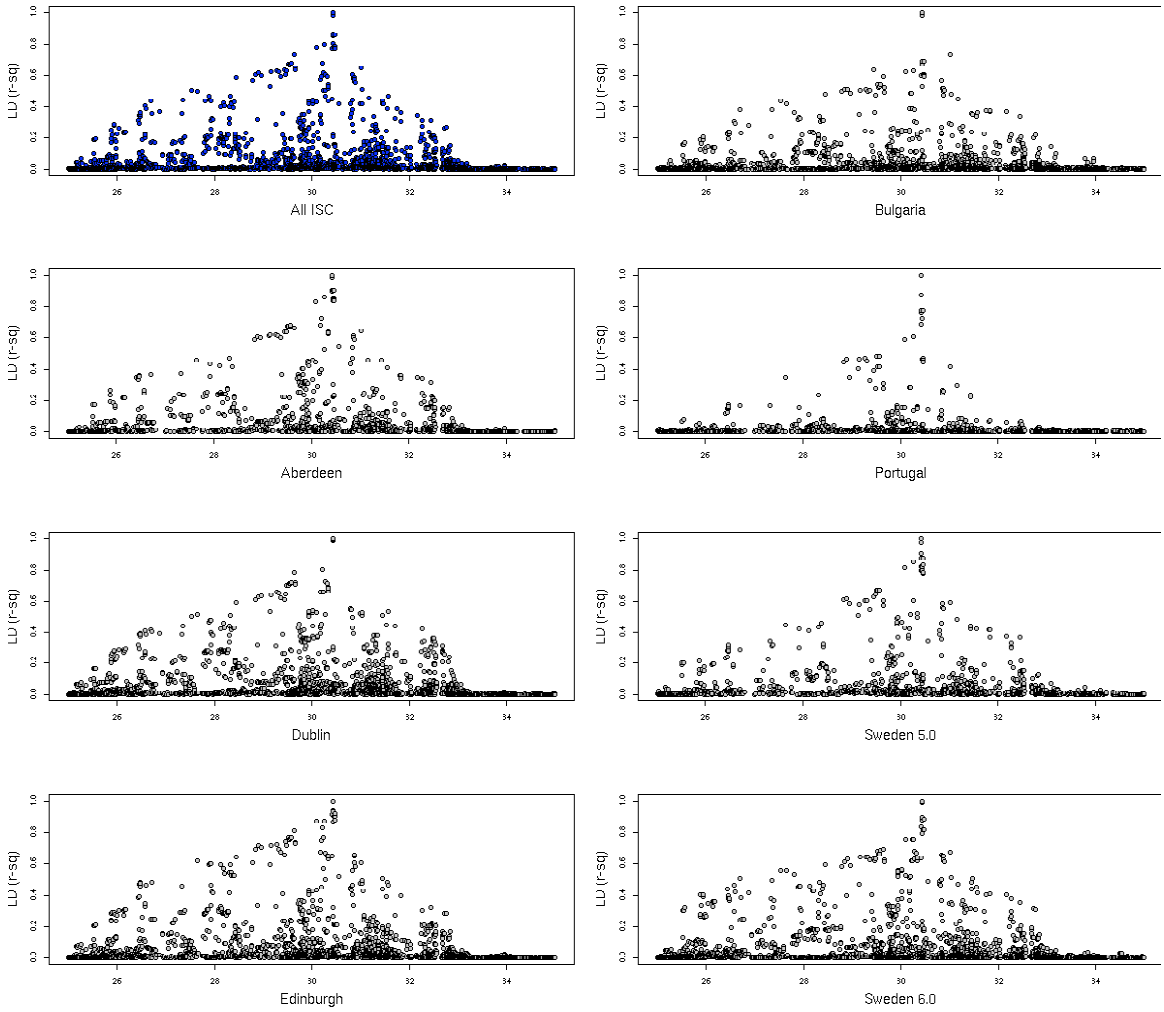


Figure S4: Linkage disequilibrium (r^2) around rs3130375 (chromosome 6).

We performed a logistic regression analysis of the MHC region (25-35Mb), controlling for genotype at rs3130375. The analysis is a regression of disease on the test SNP including as covariates genotype at rs3130375 (coded as 0,1 or 2 alleles) and also four components from the multidimensional scaling analysis to control for ancestry. Individuals missing a genotype at rs3130375 were excluded from analysis. This analysis yields only 4 SNPs at $P < 1x10^{-3}$. In contrast, removing rs3130375 as a covariate yields 190 significant SNPs at this threshold. At the $P < 1x10^{-4}$ threshold, there were no significant SNPs in the conditional analysis; in contrast, the unconditional analysis yielded 56 SNPs with $P < 10^{-4}$. This suggests that the signals observed many megabases away, e.g. at the histone cluster (~27Mb), cannot be statistically distinguished from those at rs3130375.

We also performed haplotype and conditional haplotype analyses of the HLA alleles, divided into class I (A, B and C) and class II (DRB, DQA and DQB) alleles. As shown in Table S3, we observed significant association with schizophrenia for haplotypes in both classes (class I, omnibus test $P = 4x10^{-5}$, df=4; class II, $P = 1x10^{-5}$; df=3). It was not possible to differentiate between the two signals by conditional analysis: here we tested for case/control frequency differences in class I haplotypes conditional on class II haplotypic background, and vice versa (class I, $P = 0.45$; df=5; class I, $P = 0.099$, df=4).

The two conditional tests compare a model in which the full (6 allele) haplotypes are compared against submodels in which haplotypes are grouped according the conditiong haplotypic background. See the URL
http://pnguweb.mgh.harvard.edu/~purcell/plink/whap.shtml
that describes the implementation of this test in PLINK for more details.

Table S3: Simple and conditional haplotype tests for class I and class II HLA alleles; omnibus test results for the four models are given in the main text.

3a) *Haplotype analysis of HLA haplotypes.*

HLA Haplotype	***Frequency***	***Odds ratio* (ref=AAAAAA)**	***Odds ratio* (ref=other haplotypes)**	***P value***
AAAAAA	0.05	(-ref-)	0.70	3×10^{-5}
*AAAAA	0.02	1.40	0.94	0.68
***AAA	0.02	1.01	0.73	0.032
*****A	0.04	1.23	0.84	0.11
****A*	0.10	1.36	1.02	0.78
AAA***	0.01	1.26	0.88	0.51
*A****	0.01	1.37	1.02	0.92
A*****	0.02	1.31	0.97	0.85
******	0.69	1.39	1.18	8×10^{-6}

HLA haplotype alleles are coded “A” to represent the six alleles, in the following order: HLA-A*0101, HLA-C*0701, HLA-B*0801, HLA-DRB*0301, HLA-DQB*0201 and HLA-DQA*0501. The symbol “*” indicates an alternative allele within the haplotype at that position. The test is a logistic regression of disease on

expected haplotype dosage; dummy variables for the eight samples were included as covariates. Two odds ratios (OR) are given: the first uses the first haplotype (AAAAAA) as the reference for all odds ratios; the second, for any specific haplotype, uses all other haplotypes as the reference. The p-value is based on the test of the second odds ratio.

3b) *Separation by class I and class II haplotypes and conditional tests*

Haplotype	F	OR	OR_{SUB}	P
Class I haplotypes (HLA-A*0101, HLA-C*0701, HLA-B*0801)				
AAA	0.062	(-ref-)		9.9E-06
*AA	0.023	1.23		0.25
A	0.021	1.28		0.68
A**	0.023	1.33		0.82
***	0.858	1.40		1.0E-05
Class II haplotypes (HLA-DRB*0301, HLA-DQB*0201, HLA-DQA*0501)				
AAA	0.099	(-ref-)		3.2E-06
**A	0.042	1.12		0.06
A	0.103	1.24		0.91
***	0.734	1.29		2.1E-05
Conditional test: class I haplotypes on class II background*				
AAA AAA	0.048	(-ref-)	(-ref-)	0.13
*AA AAA	0.019	1.40	|	
*** AAA	0.018	1.01	|	
*** **A	0.038	1.23	1.17	n/a
*** *A*	0.095	1.36	1.29	n/a
AAA ***	0.011	1.26	1.32	0.94
A ***	0.013	1.37	|	
A** ***	0.019	1.31	|	
*** ***	0.686	1.39	|	
Conditional test: class II haplotypes on class I background*				
AAA AAA	0.048	(-ref-)	(-ref-)	0.27
AAA ***	0.011	1.26	|	
*AA AAA	0.019	1.40	1.36	n/a
*** AAA	0.018	1.01	1.35	0.10
*** **A	0.038	1.23	|	
*** *A*	0.095	1.36	|	
*** ***	0.686	1.39	|	
A ***	0.013	1.37	1.37	n/a
A** ***	0.019	1.31	1.30	n/a

F = haplotype frequency; Haplotype: alleles are coded “A” to represent the allele on the ancestral haplotype; “*” indicates an alternative allele. OR = odds ratio under full model; OR_{SUB}= odds ratio under sub-model (a “|” indicates the haplotype is fixed to have the same odds ratio as the haplotype above it). P = for the unconditional models, simple haplotype specific p-values (that haplotype versus all others); for the conditional models, these p-value represent haplotype-specific tests performed within the haplotypic background defined by the other loci (i.e. class II for testing class I alleles, and vice versa). The omnibus p-values are given in the main text. * = For both conditional models, the six alleles correspond to *HLA-A*0101, HLA-C*0701*, *HLA-B*0801*, *HLA-DRB*0301*, *HLA-DQB*0201*, *HLA-DQA*0501* in that order.

5) ISC results for schizophrenia risk alleles from prior GWAS

O'Donovan et al (2008) recently reported an association with rs1344706 in the zinc finger protein 804A gene (*ZNF804A)*. Removing the Bulgarian and Dublin samples to avoid overlap with O'Donovan et al. replication samples, rs1344706 showed $P = 0.029$ (one-tailed). The odds ratio for the plus-strand *A* allele was 1.08 (1.12 in O'Donovan et al.).

We did not observe any evidence for the association reported in *reelin* (rs7341475) by Shifman et al (2008) (neither for our primary analyses nor the sex-specific genotypic models reported by Shifman et al; data not shown). Lencz et al (2007) reported association with rs4129148 in the pseudo-autosomal region of the X chromosome: this SNP was not genotyped, nor was it possible to impute it in our data. The nearest genotyped SNP, rs5988574, is ~2kb away, had $P = 0.056$ in the ISC, although it is not LD with rs4129148 (HapMap CEU $r^2 = 0.005$).

We recently performed a meta-analysis of bipolar disorder GWAS studies, that showed associations for SNPs in ankyrin 3, *ANK3*, and the calcium channel, voltage-dependent, L type gene, *CACNA1C*, reported in Ferreira et al (2008). For the top SNPs in these two genes, we did not observe any evidence for association in the ISC sample (after removing the samples that contained overlapping controls, UCL and Dublin). Of note, however, the region ranked 23rd in the non-overlapping ISC dataset is located in *CACNA1C* (rs2238090, $P = 7.7 \times 10^{-6}$) towards the 3' end of the gene, although this is not in LD with the Ferreira et al. SNP rs1006737. Given the evidence for a shared genetic component for schizophrenia and bipolar disorder, this and other calcium channel genes clearly warrant further investigation in both disorders.

6) Association analysis for CNV regions 22q11, 15q13 and 1q21

We recently reported association of large deletions at 22q11.2, 15q13.3 and 1q21.1 in this dataset (International Schizophrenia Consortium, 2008). We evaluated the evidence for association with common SNPs in these regions. We used a set-based permutation approach to evaluate the region-wide distribution of single SNP statistics from the primary CMH analysis within each locus. This approach implicitly allows for more than a single effect within each of these large regions that contain many genes. The set-based test implemented in PLINK works as follows, with *R* set to 0.1, *P* set to 0.05 and *N* set to 50:

1. For each set (i.e. each of the three deletion regions), for each SNP determine which other SNPs are in LD, above a certain r-squared threshold *R*
2. Perform standard single SNP analysis.
3. For each set, select up to *N* "independent" SNPs (as defined in step 1) with p-values below *P*. The best SNP is selected first; subsequent SNPs are selected in order of decreasing statistical significance, after removing SNPs in LD with previously selected SNPs.
4. From these subsets of SNPs, the statistic for each set is calculated as the mean of the single SNP statistics

5. Permute the dataset 10,000 times, keeping LD between SNPs constant (i.e. permute phenotype labels, within sample site)
6. For each permuted dataset, repeat steps 2 to 4 above.
7. Empirical p-value for the set is the number of times the permuted set-statistic exceeds the original one for that set.

We observed evidence for associated variants at 22q11 (corrected empirical *P* = 0.023) but not 15q13 (*P* = 0.26) or 1q21 (*P* = 0.27). The region-wide association at 22q11 represents at least two independent signals, centered on 18.3Mb and 19.5Mb, which correspond to the typical 1.5Mb and the atypical deleted regions respectively. In the first region, the strongest SNP is rs17210001 (*P* = 0.003), which is located in thioredoxin reductase 2 (*TXNRD2*), a gene that partially overlaps with catechol-O-methyltransferase (*COMT*). In the second region, the best SNP, rs165872 (*P* = 0.007), is within the genes phosphatidylinositol 4-kinase type 3 alpha (*PI4KA*) and serpin peptidase, clade D (heparin cofactor) member 1 (*SERPIND1)*. Figure S5 illustrates the association signals in this region.

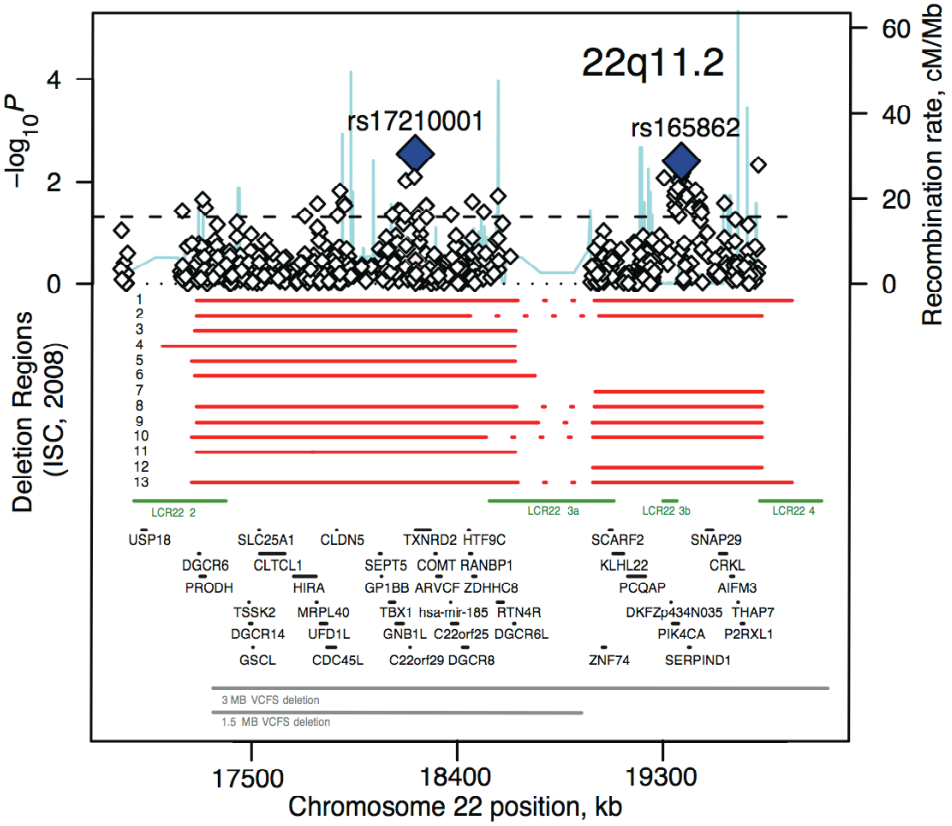


Figure S5: *Single SNP association statistics in the 22q11.2DS region. Note that this plot does not use the colour scheme used in Figure 1 (main text) to indicate LD. That is, all SNPs in this region are coloured white, but this does not imply they are in linkage equilibrium with the two most highly associated SNPs.*

Associations in the gene *PI4KA* have recently been reported for schizophrenia (Jungerius et al, 2008; Vorstman et al, 2008), including reports for association with

rs165862. The Vorstman et al. sample consists of individuals with the 22q11.2 deletion: the test for association is of 22q11.2DS patients with versus without schizophrenia. The direction of effect is consistent across the three studies, as shown in Table S4. The two other SNPs, rs2072513 and rs165793, are also reported as associated by Jungerius et al. In the ISC, rs2072513 shows a consistent effect; the SNP rs165793 was not in our dataset and could not be imputed with high confidence.

Table S4: *Comparison with previous association reports for PI4KA.*

			Jungerius et al.		Vorstman et al.		ISC	
SNP	**Position (hg18)**	**Allele (ref/*)**	***P***	**OR**	***P***	**OR**	***P***	**OR**
rs2072513	19429297	C/T	$6x10^{-6}$	1.54	0.14	2.02	0.029	1.091
rs165862	19465346	G/T	$6x10^{-5}$	1.46	0.054	2.50	0.0067	1.1
rs165793	19477293	G/A	$1x10^{-5}$	1.8	0.002	9.47	n/a	n/a

Alleles are coded for the positive stand, hg18. Odds ratios are relative to the first allele listed for each SNP.

7) Combined analysis with MGS-EA and SGENE

We exchanged summary data with two large case/control schizophrenia GWAS consortia (see companion manuscripts). The Molecular Genetics of Schizophrenia consortium (MGS) has 2,687 cases and 2,656 controls, all of European descent, genotyped on the Affymetrix 6.0 array. The SGENE consortium has 2,005 cases and 12,837 controls from multiple, but predominantly European, populations genotyped on the Illumina HumanHap300 and 550. The Aberdeen sample features in both the ISC and SGENE: for the comparison reported here, the Aberdeen samples have been removed from SGENE (and this is reflected in the SGENE sample N's given above).

For all ISC SNPs with a $P < 10^{-3}$, we received from the two other studies *P*-values, odds ratios for either the genotyped SNP, or the best proxy or imputed SNP. For regions of interest, we also combined summary statistics for all imputed SNPs. We combined results across m=3 studies by summing weighted Z-scores, such that the combined test is based on

$$Z = \frac{\eta_1 Z_1 + \cdots \eta_m Z_m}{[\eta_1^2 + \cdots + \eta_m^2]^{1/2}}$$

where the weights are given by

$$\eta_k = \left[2n \frac{n_{k,+1}}{n_k} \frac{n_{k,-1}}{n_k} \right]^{1/2}$$

where n_k is the sample size for the study k; $n_{k,+1}$ is the number of cases; $n_{k,-1}$ is the number of controls.

Table S5 lists three further regions containing SNPs with combined $P < 10^{-6}$. including the gene polypyrimidine tract binding protein 2 (*PTBP2)*, which was the 4th most associated region in the ISC. *PTBP2* promotes neuron-specific alternative splicing of transcripts (Makeyev et al, 2007) including neurexin1 (Resnick et al, 2008), a gene that

has been implicated by CNV analysis in schizophrenia (ISC 2008; Kirov et al 2008; Walsh et al 2008) and other neurodevelopmental disorders. The two other regions include the genes fragile X mental retardation, autosomal homolog 1 (*FXR1*) and transcription factor 4 (*TCF4*).

Table S5. Combined analysis of ISC with MGS-EA and SGENE datasets.

CHR	SNP	G/I (ISC)	MAF (ISC)	BP	A1/A2	P (COMB)	P (ISC)	P (MGS-EA)	P (SGENE)	OR (ISC)	OR (MGS-EA)	OR (SGENE)	GENES (+/-100kb)
1	rs7544736	I	0.184	96936931	G/A	5.7E-07	2.3E-06	6.7E-04	5.3E-01	1.25	1.19	1.04	PTBP2
3	rs7640601	I	0.289	182007458	G/C	7.3E-07	1.4E-02	5.5E-03	8.0E-04	0.90	0.87	0.85	FXR1
3	rs9869882	I	0.356	182014635	A/G	3.8E-07	4.4E-03	1.4E-02	5.4E-04	0.90	0.89	0.86	FXR1
3	rs9838229	I	0.264	182015945	C/A	2.1E-07	1.6E-02	1.6E-03	5.4E-04	0.91	0.86	0.84	FXR1
3	rs4488266	I	0.356	182018099	C/T	3.8E-07	4.4E-03	1.5E-02	5.4E-04	0.90	0.90	0.86	FXR1
3	rs2139551	I	0.264	182019059	A/G	1.8E-07	1.6E-02	1.5E-03	4.9E-04	0.91	0.86	0.84	FXR1
3	rs10937048	I	0.356	182022086	A/G	3.9E-07	4.4E-03	1.5E-02	5.4E-04	0.90	0.90	0.86	FXR1
3	rs6782299	G	0.268	182033396	G/T	1.4E-07	1.5E-02	1.2E-03	4.9E-04	0.91	0.87	0.84	FXR1
3	rs1879248	I	0.264	182033908	C/T	1.5E-07	1.6E-02	1.2E-03	4.9E-04	0.91	0.87	0.84	FXR1
3	rs1915104	I	0.264	182039598	C/A	1.5E-07	1.6E-02	1.1E-03	4.9E-04	0.91	0.87	0.84	FXR1
3	rs6767560	I	0.356	182048915	G/T	4.0E-07	4.4E-03	1.6E-02	5.0E-04	0.90	0.91	0.86	FXR1
3	rs2337743	I	0.271	182059367	T/A	1.6E-07	1.6E-02	2.0E-03	3.2E-04	0.91	0.87	0.84	FXR1
3	rs1607678	I	0.270	182067412	C/T	1.7E-07	1.6E-02	2.0E-03	3.2E-04	0.91	0.87	0.84	FXR1
3	rs1010471	I	0.329	182173786	A/G	6.5E-07	7.7E-03	9.9E-03	8.1E-04	0.90	0.90	0.86	FXR1
6	rs6904071	I	0.142	27155235	A/G	1.8E-08	3.0E-04	1.2E-02	3.7E-04	0.82	0.88	0.80	MHC (Figure 1)
6	rs926300	I	0.142	27167422	T/A	1.1E-08	3.0E-04	1.2E-02	2.1E-04	0.82	0.88	0.79	MHC (Figure 1)
6	rs6913660	I	0.142	27199404	A/C	2.4E-08	3.0E-04	1.7E-02	3.4E-04	0.82	0.88	0.80	MHC (Figure 1)
6	rs13219181	I	0.142	27244204	G/A	1.3E-08	3.0E-04	1.5E-02	2.1E-04	0.82	0.88	0.79	MHC (Figure 1)
6	rs13194053	I	0.142	27251862	C/T	9.5E-09	3.0E-04	1.5E-02	1.5E-04	0.82	0.88	0.78	MHC (Figure 1)
6	rs3800307	G	0.199	27293771	A/T	4.4E-08	3.4E-03	1.3E-02	6.1E-05	0.88	0.89	0.79	MHC (Figure 1)
6	rs6932590	I	0.238	27356910	C/T	7.1E-08	2.2E-03	3.4E-03	8.5E-04	0.88	0.87	0.83	MHC (Figure 1)
6	rs3800316	I	0.251	27364081	C/A	3.8E-08	3.5E-03	7.2E-04	1.1E-03	0.89	0.86	0.83	MHC (Figure 1)
6	rs7746199	G	0.180	27369303	T/C	5.0E-08	8.8E-04	6.8E-04	5.7E-03	0.86	0.84	0.84	MHC (Figure 1)
6	rs3800318	I	0.180	27371620	T/A	6.4E-08	8.8E-04	2.8E-03	2.3E-03	0.86	0.85	0.82	MHC (Figure 1)
6	rs9272219	G	0.270	32710247	T/G	6.9E-08	2.2E-05	1.3E-02	1.0E-02	0.85	0.90	0.88	MHC (Figure 1)
6	rs9272535	G	0.289	32714734	A/G	8.9E-08	2.5E-05	1.6E-02	9.9E-03	0.85	0.90	0.88	MHC (Figure 1)
18	rs9646596	G	0.032	51200210	A/G	2.6E-07	6.0E-04	1.7E-02	2.4E-03	1.42	1.29	1.44	TCF4
18	rs17594526	G	0.031	51209236	T/C	1.3E-07	3.5E-04	7.9E-03	4.0E-03	1.44	1.33	1.41	TCF4
18	rs17594665	G	0.030	51214717	A/G	1.6E-07	2.9E-04	8.6E-03	5.0E-03	1.45	1.33	1.40	TCF4
18	rs17594721	I	0.024	51216890	G/A	3.7E-07	4.9E-04	9.9E-03	6.8E-03	1.50	1.32	1.38	TCF4
18	rs11152369	G	0.030	51217326	C/A	2.0E-07	2.7E-04	9.8E-03	5.9E-03	1.45	1.32	1.39	TCF4
18	rs17509991	G	0.028	51218182	A/G	2.9E-07	2.7E-04	1.4E-02	5.9E-03	1.45	1.30	1.39	TCF4
18	rs17510124	I	0.027	51219678	C/T	3.8E-07	5.0E-04	1.1E-02	6.0E-03	1.47	1.31	1.39	TCF4
18	rs7228846	I	0.027	51225616	G/A	2.8E-07	3.5E-04	1.1E-02	6.0E-03	1.49	1.31	1.39	TCF4
18	rs8089309	I	0.027	51226290	C/A	3.8E-07	5.0E-04	1.1E-02	6.0E-03	1.47	1.31	1.39	TCF4
18	rs1371832	I	0.027	51233234	T/C	4.0E-07	5.0E-04	1.1E-02	6.2E-03	1.47	1.31	1.38	TCF4

Table S5 This table contains all SNPs with a combined p < 1e-6 across ISC, MGS-EA and SGENE studies. This table includes the Aberdeen samples in ISC and excludes them from SGENE.

CHR Chromosome
SNP RefSeq ID for SNP
G/I(ISC) Directly genotyped or imputed in ISC
MAF(ISC) Minor allele frequency in entire ISC
BP Physical position, base-pairs, hg18
A1/A2 Allele codes; A1 is the tested allele upon which OR is based
P(COMB) Combined p-value for ISC, MGS-EA and SGENE

P(ISC) Association p-value in ISC
P(MGS-EA) Association p-value in MGS-EA
P(SGENE) Association p-value in SGENE

OR(ISC) Odds ratio for A1 allele in ISC
OR(MGS-EA) Odds ratio for A1 allele in MGS-EA
OR(SGENE) Odds ratio for A1 allele in SGENE

GENES RefSeq genes within 100kb of the SNP
For the MHC region that contains many genes in a large region of extended LD, see Figure 1

8) Power to detect individual common variants in genome-wide association studies

In this section we outline the rationale behind our approach of creating scores to reflect the aggregate effects of common variants, in the context of a polygenic model for schizophrenia disease risk.

Our data are consistent with a model in which there are very few or no common variants of moderate to large effect that influence schizophrenia risk, as we were well-powered to detect at least some of them. Here we define a "moderate" effect as a heterozygote genotypic relative risk (GRR) of 1.5 (assuming a multiplicative model of gene action) for a variant with a minor allele frequency (MAF) of 0.20, which would explain ~0.7% of the total variance, assuming a liability-threshold model.

If directly and accurately genotyped, there is complete power (100%) even at a strict genome-wide significance level of 5×10^{-8} to detect such an effect. If the variant was imperfectly measured by its best proxy on the array, power of standard single SNP association would still remain high in many circumstances: if r^2 = 0.5 then power is 96%, if r^2 = 0.25 then power is 36%. Even at 36% power, if there were at least 10 such loci throughout the genome (collectively explaining ~7% of the total phenotype variance in liability), there is a 99% probability that at least one would have been detected.

Given that modern whole genome genotyping platforms (combined with imputation techniques) can be expected to measure the majority of common variation in samples of European descent, it is reasonable to assume that there are either no or very few common variants of such large, universal effect. Common variants of moderate to large effect might still exist in specific groups (e.g. defined by environmental exposure or ancestry or phenotypic homogeneity) for example, that lead to a 50% increase in risk per allele. Such heterogeneity would tend to reduce the apparent effect size in the combined population though.

If we take a GRR of 1.05 to represent a common variant of "very small effect", power will be extremely low in our sample, even before correction for multiple testing. For a 20% MAF variant, power is 8% for a nominal type I error rate α=0.01; it is almost 0% for a 5×10^{-8} error rate. It is therefore practically impossible that any one such variant would be identified by the current study as unambiguously associated with disease. For a 20% MAF variant, a GRR of 1.05 implies an allele frequency of 0.2079 in cases and 0.1999 in controls. Even with 10,000 cases and 10,000 controls, power at the relatively low threshold of 1×10^{-6} is only 0.2%

However, at order-of-magnitude less stringent significance thresholds, we would expect to see an appreciable enrichment of true positives of very small effect. For example, power is 46% for α=0.2 and 72% for α=0.5. (As the true effect size approaches the null, power approaches the type I error, i.e. 20% or 50% in this case.)

The implication is that the majority of the time (72%), such a variant would be in the top half of the distribution of all SNPs ($P < 0.5$) as opposed to the lower half ($P > 0.5$). Similarly, almost half (46%) of these variants would be in the top fifth of all results ($P < 0.2$). Consider, for example, if of 100,000 independent SNPs, 1% (1000) are in fact risk alleles of this type and the remaining 99,000 are not associated with disease. We would then expect a ~2.3-fold enrichment of true risk variants at the $P < 0.2$ threshold, i.e. $(1000 \times 0.46) / (1000 \times 0.46 + 99000 \times 0.2)$.

9) Creating the target sample scores

The scoring procedure aims to measure indirectly the collective effect of many weakly associated alleles that tend to show only very small allele frequency differences between cases and controls, but will nonetheless have higher average association test statistics and lower p-values than null loci.

First, we selected autosomal SNPs with a total sample MAF of 2% or greater and a genotyping rate threshold of 99% or greater (thereby effectively limiting the analysis to SNPs present on all versions of the Affymetrix arrays). We next pruned the SNP panel to remove SNPs in strong linkage disequilibrium with other SNPs (based on a pairwise r^2 threshold of 0.25, within a 200-SNP sliding window). Focusing these analyses on a subset of SNPs in approximate linkage equilibrium (LE) has several advantages and makes interpretation more straightforward, for example, the calibration of results with simulated data and the comparison of results across the frequency spectrum, but most importantly to ensure the score represents the aggregate effect of a large number of independent SNPs (see Section S13). As described in section S13, the results do not fundamentally change if we use the full set of QC+ SNPs. Focusing on autosomal SNPs avoids the issue of how to score haploid and diploid genotypes in males and females without creating artificial mean differences between the sexes.

In general, we have adopted a conservative approach, to demonstrate robustly the principle that many common variants of small effect influence risk for schizophrenia. In subsequent studies that aim to quantify this component of variance more precisely, it will be important to account for all measurable common variation by inclusion of correlated and sex-chromosome SNPs.

The pruned SNP set (N=74,062) contains only 10% of all QC+ SNPs (N=739,995) but still yields between ~30 to ~70% of the coverage relative to the full set (depending on the r^2 threshold, 0.8 to 0.2). To evaluate coverage, we used 2,486,974 SNPs from the Phase 2 CEU HapMap sample (filtered for 1% MAF and 90% genotyping rate, N=60 founders) including X chromosome SNPs that were obligatorily excluded from the pruned set. We defined coverage as the proportion of all HapMap SNPs with an above-threshold level of LD with at least one SNP within 50kb in the set under consideration, as shown in Table S6. The results indicate that SNPs in the 74K subset weakly tag ($r^2 > 0.2$) approximately half of all HapMap SNPs considered here, which includes the fact that no X chromosome SNPs will be tagged. When interpreting the results of the score analyses, it should therefore be borne in mind that the pruned SNP set weakly tags a

substantial proportion of all common variation, although coverage (for variants above 1% MAF) is far from complete.

Table S6: *Coverage for unpruned and pruned SNP sets, based on release 23 Phase 2 CEU HapMap.*

	# tagged (of 2,486,974)		% tagged		
R^2 threshold	Unpruned (N=739,995)	Pruned (N=74,062)	Unpruned	Pruned	Pruned / unpruned ratio
0.2	1,674,853	1,166,968	0.67	0.47	0.70
0.5	1,524,647	664,507	0.61	0.27	0.44
0.8	1,305,626	378,582	0.52	0.15	0.29

We formed independent discovery and target samples within the ISC in a number of ways. In each instance, the 74K SNPs were tested for association in the discovery sample, using a Cochran-Mantel-Haenszel statistic conditioning on site, mirroring the approach used in the primary association analysis reported above. In contrast to the primary analysis, however, the score analyses treated all Swedish individuals (genotyped on 5.0 and 6.0 arrays) as a single cluster, given that the analysis is restricted to SNPs on both platforms. The primary MDS analysis did not reveal any systematic differences between the two Swedish samples.

Based on the discovery sample association statistics, large sets of nominally-associated alleles were selected as "score alleles", for different significance thresholds (p_T). Specifically, we selected 11 overlapping sets: all SNPs with $p_T < 0.01$, $p_T < 0.05$, $p_T < 0.1$, $p_T < 0.2$, $p_T < 0.3$, $p_T < 0.4$ and $p_T < 0.5$; for some analyses we also considered sets excluding the most highly associated SNPs, $0.01 < p_T < 0.2$, $0.05 < p_T < 0.2$, $0.05 < p_T < 0.5$ and $0.2 < p_T < 0.5$.

In the target sample, we calculated the total score for each individual as the number of score alleles weighted by the log of the odds ratio from the discovery sample. Scores are additive across SNPs on the log odds scale and therefore multiplicative on the odds of disease scale. If an individual in the target sample is missing that genotype, the mean score is imputed for that genotype, based on the target sample allele frequency. The score is expressed as the mean score per SNP in the set; the number of non-missing genotypes used to calculate each score is also recorded per individual, for use as a covariate in subsequent target sample analysis.

We used PLINK's --score function to calculate scores, described at this URL http://pngu.mgh.harvard.edu/~purcell/plink/profile.shtml

10) Testing for association between score and disease in the target sample

The primary target sample test uses a logistic regression of disease state on score. Critical covariates include the number of non-missing genotypes of all SNPs used to calculate the score, to control for potential differences in genotyping rate between cases and controls. Study sample is taken into account by inclusion of dummy-coded covariates to represent the seven strata (e.g. Aberdeen, Dublin, etc). Study sample is also accounted for in all discovery sample analyses, by use of a Cochran-Mantel-Haenszel stratified analysis to calculate the common odds ratios.

In the target sample logistic regression analysis, we estimate the variance explained in disease state by the score as the difference in the Nagelkerke pseudo *R*-squared (Nagelkerke, 1991) from a model including the score and covariates versus a model including only the covariates. Below, the symbol R^2 represents this measure; the lower-case r^2 represents a measure of LD between two SNPs. All tests reported are two-sided. All reported significant associations are in the expected direction, such that a higher score is associated with an increased risk of disease.

11) Summary statistics for the SNP sets used in the score analysis

For the 11 p-value thresholds described above (labeled p_T), Table S7 shows the number of significant SNPs from the 74K SNP subset. These results are based on Cochran-Mantel-Haenszel tests calculated in the entire ISC sample. Proportionally, the lower p-value thresholds show a greater enrichment of significant results compared to chance (e.g. a 1.31-fold increase at the $p_T<0.01$ level). Because these SNPs have been selected to be in approximate linkage equilibrium, the comparison between O(N) and E(N) is informative. Although at lower p_T thresholds the ratio O(N) / E(N) approaches 1.00, the absolute enrichment difference O(N) – E(N) is greater.

Table S7: *Observed and expected counts of significant SNPs from the entire ISC within each range of p_T thresholds.*

Threshold (p_T)	O(N)	Proportion	E(N)		Ratio	Difference
$p_T < 0.01$	971	0.01	740.6		1.31	230.4
$p_T < 0.05$	4488	0.05	3703.1		1.21	784.9
$p_T < 0.1$	8537	0.10	7406.2		1.15	1130.8
$p_T < 0.2$	16124	0.20	14812.4		1.09	1311.6
$p_T < 0.3$	23563	0.30	22218.6		1.06	1344.4
$p_T < 0.4$	31029	0.40	29624.8		1.05	1404.2
$p_T < 0.5$	38274	0.50	37031.0		1.03	1243.0
$0.01 < p_T < 0.2$	15153	0.19	14071.8		1.08	1081.2
$0.05 < p_T < 0.2$	11636	0.15	11109.3		1.05	526.7
$0.05 < p_T < 0.5$	33786	0.45	33327.9		1.01	458.1
$0.2 < p_T < 0.5$	22139	0.30	22218.6		1.00	-79.6

Results based on the 74K SNP subset. The observed and expected number of significant results are labeled O(N) and E(N) respectively. The column "Proportion" indicates the approximate proportion of all SNPs included in this range (assuming a null distribution). The columns "Ratio" and "Difference" compare the O(N) and E(N) values as stated.

12) Results from the male/female score analysis

Table S8 shows the full set of p-values and pseudo-R^2 statistics for the ISC male/female score analyses, at the eleven different thresholds of p_T (i.e. the thresholds used to select the SNPs for scoring from the discovery sample analyses).

The test in the target samples is a logistic regression of disease state on score and the number of non-missing genotypes, considering all SNPs used to calculate that score. The measure of genotyping rate is added as a covariate to control for possible differences in non-random genotyping failure between cases and controls. The target sample p-value is from a Wald test of the coefficient for the score. As mentioned above, the R^2 measure is based on Nagelkerke's pseudo-R^2. Specifically, it is the difference between the R^2 for a model including terms for the intercept, genotyping rate and score, versus a model containing only the intercept and genotyping rate.

Of note, both the statistical significance and the variance explained increase as a greater number of SNPs with decreasing strength of association in the discovery sample are added. As discussed and modeled by simulation below, this is consistent with a highly polygenic model of disease risk, in which there are a large number of common variants of small individual effect.

Table S8: *Results of the ISC male / female score analyses.*

	Discovery: females **Target:** males			**Discovery:** males **Target:** females	
Discovery sample score threshold	Target p-value	Target R^2		Target p-value	Target R^2
$p_T < 0.01$	9×10^{-4}	0.004		2×10^{-5}	0.007
$p_T < 0.05$	3×10^{-8}	0.011		2×10^{-7}	0.011
$p_T < 0.1$	6×10^{-12}	0.016		2×10^{-11}	0.018
$p_T < 0.2$	7×10^{-14}	0.019		7×10^{-16}	0.027
$p_T < 0.3$	1×10^{-15}	0.022		1×10^{-16}	0.028
$p_T < 0.4$	3×10^{-17}	0.025		9×10^{-18}	0.030
$p_T < 0.5$	1×10^{-17}	0.025		9×10^{-19}	0.032
$0.01 < p_T < 0.2$	3×10^{-12}	0.017		2×10^{-13}	0.022
$0.05 < p_T < 0.2$	8×10^{-9}	0.011		1×10^{-12}	0.021
$0.05 < p_T < 0.5$	3×10^{-13}	0.018		8×10^{-16}	0.026
$0.2 < p_T < 0.5$	2×10^{-7}	0.009		3×10^{-6}	0.009

See main text for details.

The last four rows of Table S8 show scenarios in which we excluded the most associated SNPs from the score: in all cases, we were able to create scores that were significantly correlated with disease state in independent target samples. For variants of

very small effect, power is expected to be relatively low even at nominal significance thresholds such as 0.01 or 0.05; as such, under a highly polygenic we would not expect the majority of true positives to be represented in the top 1% or even 5% of SNPs, as detailed in the section above.

The effects are slightly stronger for the male/female discovery/target pairing than for the female/male pairing. We note that the sample size of both the discovery and target samples impact on significance: a larger discovery sample implies better estimates of effect sizes and therefore a more discriminating set of score alleles, whereas a larger target set yields higher significance for the same variance explained. There are 2176/1642 male cases/controls and 1146/1945 female.

13) Addressing population stratification and other possible confounders

Section Overview

In part (*a*) below, we note that population stratification within one or more samples could produce a qualitatively similar profile of results compared to a polygenic model. If the score indexed alleles that were particularly subject to stratification and more frequent in the ancestral backgrounds of the sampled cases, this could lead to many spuriously associated SNPs in both discovery and target samples. We therefore selected discovery and target samples from different populations, to reduce the chance of overlapping substructure. As described below, for stratification alone to explain the results of these additional analyses would require that the same ancestral strata are over-represented in individuals with schizophrenia in every sample, an explanation that is much less plausible than the alternative hypothesis of polygenic inheritance. We also report analyses that include the first four MDS components as covariates to control for stratification.

In this section, we also consider possible confounders other than population stratification. We recently reported an increased rate of rare copy number variants (CNVs) in this sample (International Schizophrenia Consortium, 2008). As described below, repeating the primary score analyses (males/females) including CNV burden as a covariate did not change results (part *b*). Similarly, controlling for individuals' estimated heterozygosity did not change the results (part *c*).

In the final part (*d*), we report various other technical and analytic issues. In particular, using a different set of SNPs with complete genotyping did not change our conclusions. The results were also essentially unchanged if all 739,995 SNPs were used to create scores; scores based on unweighted allele counts also gave similar results.

a) Population stratification

As shown by multidimensional scaling (MDS, Figure S1 above), the seven geographically distinct samples cluster into four groups: those samples from the British Isles (Aberdeen, Edinburgh, Dublin and UCL) cluster closely together; the three other sites (Sweden, Portugal/Azores and Bulgaria) form three separate clusters. Note that extensive QC analysis had already been performed on these samples and SNPs prior to all score analyses, thus the relatively tight clustering within each sample.

For stratification to give false positives in this context, both discovery and target samples must necessarily have the same population strata over-represented in the same phenotypic class. By selecting the British Isles (BI, 1887 cases, 2360 controls) as the discovery sample and other samples (non-BI, 1435 cases, 1227 controls) as the target sample (and vice versa), we reduced this possibility. This does not assume the BI samples to be homogeneous: sample site is still explicitly controlled for in both discovery and target analyses. We observed a similar pattern of results, with score alleles designated in the BI samples showing significantly higher rates in non-BI cases (Table S9; the $p_T < 0.5$ threshold result is $P = 2.6 \times 10^{-6}$) and vice versa. The effects are

slightly less pronounced than in the initial comparison. This might reflect some residual degree of unaccounted for within-site stratification, although it could also represent the aggregate effect of the many subtle differences in allele frequency and patterns of linkage disequilibrium between BI and non-BI sites.

Table S9: *Results of the British Isles / non-British Isles score analyses.*

	Discovery: BI samples **Target:** non-BI samples			**Discovery:** non-BI samples **Target:** BI samples	
Discovery sample score threshold	Target p-value	Target R^2		Target p-value	Target R^2
$p_T < 0.01$	9×10^{-1}	0.001		7×10^{-3}	0.002
$p_T < 0.05$	5×10^{-2}	0.002		7×10^{-4}	0.003
$p_T < 0.1$	3×10^{-4}	0.006		2×10^{-5}	0.005
$p_T < 0.2$	6×10^{-5}	0.008		1×10^{-7}	0.008
$p_T < 0.3$	2×10^{-5}	0.009		4×10^{-8}	0.009
$p_T < 0.4$	1×10^{-6}	0.012		2×10^{-7}	0.008
$p_T < 0.5$	3×10^{-6}	0.011		1×10^{-6}	0.007
$0.01 < p_T < 0.2$	1×10^{-4}	0.007		1×10^{-6}	0.007
$0.05 < p_T < 0.2$	9×10^{-5}	0.008		6×10^{-6}	0.006
$0.05 < p_T < 0.5$	3×10^{-6}	0.011		8×10^{-5}	0.004
$0.2 < p_T < 0.5$	4×10^{-3}	0.004		3×10^{-1}	0.000

We also repeated the score analyses with each single sample as the target sample and the other six as the discovery sample. We observed similar significant results each time, indicating that no single sample drives the results (Table S10). The *P*-values in Table S10 tend to be less significant than for the entire sample: the sample size of both the discovery and target samples impact on significance and R^2, as noted above.

Table S10: *Results of the "all but sample X" / "sample X" score analyses.*

Discovery: remaining six samples **Target:** this sample	Target p-value for $p_T < 0.5$ score
Aberdeen	8×10^{-7}
Bulgaria	2×10^{-6}
Dublin	2×10^{-5}
Edinburgh	3×10^{-7}
Portugal	1×10^{-3}
Sweden	6×10^{-5}
UCL	9×10^{-4}

We also performed the primary male/female analysis with the first four components from the MDS analysis (performed on the 74K SNP subset) included as continuous covariates, as well as binary dummy variables coded for sample. The same pattern of results emerged from this analysis (Table S11). This analysis was performed both for all

seven samples, as well as an analysis restricted to the four British Isles samples. For reference, Table S12 shows the case and control counts by sex for each site.

Table S11: *Including quantitative estimates of ancestry from multidimensional scaling analysis (first four components) and sample site as covariates in the male versus females score analysis.*

	BI samples only (N=4)			All samples (N=7)	
	Disc: females **Target:** males	**Disc:** males **Target:** females		**Disc:** females **Target:** males	**Disc:** males **Target:** females
Discovery sample score threshold	Target p-value	Target p-value		Target p-value	Target p-value
$p_T < 0.01$	1×10^{-2}	1×10^{-5}		1×10^{-3}	5×10^{-5}
$p_T < 0.05$	2×10^{-4}	5×10^{-5}		6×10^{-8}	1×10^{-6}
$p_T < 0.1$	5×10^{-6}	5×10^{-10}		2×10^{-11}	1×10^{-10}
$p_T < 0.2$	9×10^{-6}	4×10^{-14}		4×10^{-13}	3×10^{-15}
$p_T < 0.3$	2×10^{-7}	1×10^{-13}		7×10^{-15}	7×10^{-16}
$p_T < 0.4$	3×10^{-8}	1×10^{-14}		2×10^{-16}	4×10^{-17}
$p_T < 0.5$	2×10^{-8}	5×10^{-15}		6×10^{-17}	3×10^{-18}
$0.01 < p_T < 0.2$	8×10^{-5}	2×10^{-11}		1×10^{-11}	7×10^{-13}
$0.05 < p_T < 0.2$	3×10^{-3}	4×10^{-13}		3×10^{-8}	3×10^{-12}
$0.05 < p_T < 0.5$	2×10^{-6}	2×10^{-13}		2×10^{-12}	3×10^{-15}
$0.2 < p_T < 0.5$	3×10^{-5}	5×10^{-4}		4×10^{-7}	6×10^{-6}

Table S12: *Sample size stratified by sample site, sex and phenotype.*

	Males		Females	
	Control	Case	Control	Case
Aberdeen	449	536	253	184
Dublin	260	189	606	86
Edinburgh	148	268	139	101
UCL	213	372	292	151
Bulgarian	291	271	320	257
Portugal	80	214	136	133
Sweden	201	326	199	234

The four samples from the British Isles are listed first, followed by the three samples from outside of the British Isles.

For the results presented in Tables S9, S10 and S11 purely to reflect stratification would require that the same groups or clines of ancestry exist within each separate sample. Furthermore, it is necessary to assume that those minority groups or clines are also systematically over-represented, or systematically under-represented, in the cases from each sample (either through ascertainment or true between-group differences in disease prevalence).

To give a concrete example: if population stratification accounted for our entire result, this would imply that the scores indexed alleles that are more frequent in the ancestral backgrounds of British, Scottish and Irish patients with schizophrenia compared to controls (i.e. as could result if cases and controls are not well-matched). However, for the scores to also predict disease in target samples outside of the British Isles, then these same alleles must also be more frequent in the ancestral backgrounds of Swedish, Bulgarian and Portuguese patients compared to the (badly-matched) controls in those samples. From the ascertainment protocols and the MDS analysis, we can rule out major trans-national migration, effectively excluding the possibility that many individuals with schizophrenia originated from a one or a small number of ethnic groups and were more likely than their unaffected compatriots to migrate recently to all of the countries studied here.

b) Analysis incorporating individuals' genome-wide CNV burden

We recently reported that cases in this sample had significantly higher rates of rare copy number variants (CNVs) compared to controls (International Schizophrenia Consortium, 2008). Since the same data were used to identify structural variants as well as to genotype SNPs, we wanted to exclude the possibility that our score analyses indirectly represented this established difference in the rate of rare CNVs, due to technical bias in genotype calling for example.

We repeated the primary analysis with the CNV burden reported in the previous manuscript as a covariate: this did not change the results, as shown in Table S13. In addition, the covariate of CNV burden (number of rare deletions and duplications per individual) was significant in both analyses; other covariates were genotyping rate for the SNPs used to form the score and dummy variables to represent sample site. UCL was excluded, as there were no CNV data available for controls.

Table S13: *Including individuals burden of rare copy number variants (CNVs) in the primary male/female score analyses.*

	All samples except UCL (N=6)	
	Disc: females **Target:** males	**Disc:** males **Target:** females
Discovery sample score threshold	Target p-value	Target p-value
$p_T < 0.01$	2×10^{-3}	3×10^{-4}
$p_T < 0.05$	3×10^{-9}	3×10^{-5}
$p_T < 0.1$	5×10^{-13}	3×10^{-7}
$p_T < 0.2$	7×10^{-16}	4×10^{-11}
$p_T < 0.3$	9×10^{-18}	2×10^{-12}
$p_T < 0.4$	2×10^{-18}	3×10^{-13}
$p_T < 0.5$	2×10^{-18}	5×10^{-14}
$0.01 < p_T < 0.2$	1×10^{-14}	3×10^{-9}
$0.05 < p_T < 0.2$	2×10^{-10}	3×10^{-9}
$0.05 < p_T < 0.5$	2×10^{-13}	2×10^{-12}
$0.2 < p_T < 0.5$	3×10^{-6}	8×10^{-6}

c) Controlling for per-individual heterozygosity (inbreeding coefficient)

We tested for differential levels of inbreeding coefficients between cases and controls within each ISC site, in terms of Wrights F_{ST} statistic, as calculated by PLINK (Purcell et al, 2007). These estimates compare the observed rate of heterozygosity per individual to the expected rate given allele frequency and assuming Hardy-Weinberg equilibrium. The estimates were not constrained to be positive (negative values indicate that fewer homozygotes than expected were observed).

A linear regression of estimated F_{ST} (based on the 74K subset of SNPs, calculated using sample-specific allele frequencies) on phenotype, controlling for site as a dummy-coded factor, showed no association (P = 0.67). There were strong between-site effects, notably with the Portuguese individuals showing higher F_{ST} levels, consistent with this population being more isolated.

Stratifying by site and testing for case/control differences in heterozygosity showed no strong associations; the UCL sample showed a nominally significant association (P = 0.03), as shown in Table S14. As UCL cases and controls were genotyped on different versions of the Affymetrix platform ("500K" controls versus "5.0" cases), this might reflect some proportion of SNPs showing subtle technical differences in genotyping.

Table S14: *Analysis of per-individual heterozygosity by disease state, within sample collection site, based on Wright's inbreeding coefficient, F_{ST}.*

Site	Cases (mean F_{ST})	Controls (mean F_{ST})	p-value (t-test)
Aberdeen	0.000	0.000	0.36
Bulgaria	0.001	0.001	0.81
Dublin	-0.001	0.000	0.53
Edinburgh	-0.001	-0.002	0.10
Portugal/Azores	0.004	0.003	0.36
Sweden	-0.001	-0.001	0.78
UCL	-0.001	0.000	0.03

Adding estimated F_{ST} value as a covariate did not impact the primary female/male score analysis (full data not shown). For example: the $p_T < 0.5$ score target sample P-values were 1×10^{-17} and 9×10^{-19} for females-to-males (discovery-to-target) and males-to-females respectively (unchanged from the original analysis).

d) Other analytic, QC and interpretative issues involved in the scoring procedure

We performed a set of further analytic controls. Specifically, we:

- investigated possible confounding effects of missing genotype data
- repeated the analysis with complete (100%) genotyping SNPs
- used all 740K SNPs as well as the LD-pruned 74K subset
- confirmed that the results were not driven by a small number of SNPs of very large effect, or a small number of genomic sites (as opposed to a true polygenic effect involving many sites)

- repeated the analysis using unweighted scores
- checked for cryptic relatedness and sample duplication

For the 74K subset of SNPs the ISC sample, there was no correlation between the standard quality control indices and level of genotype-phenotype association, as indexed by either the signed Z-score per SNP (with direction of effect based on minor allele frequency) or chi-square statistic. Note that all genotypes were previously subjected to standard QC procedures, outlined above. Performed separately for cases and controls, there was no correlation between the association statistics and genotyping rate or Hardy-Weinberg p-value; also, there was no correlation with a measure of differential genotyping rate between cases and controls (data not shown). There were also no significant associations between primary association statistics and the specific strand or allele queried by the Affymetrix platform (data not shown). In addition, we obtained similar results (Table S15) when using a set of SNPs with complete (100%) genotyping in the entire ISC. We selected these SNPs prior to the LD-pruning stage; after LD-pruning and filtering on a 2% MAF, this yielded a set of 25,215 SNPs, of which 15,590 were in the original 74K set (i.e. 9,624 unique to this set).

Table S15: *The male/female and British Isles/non-British Isles score analyses on a set of SNPs with 100% genotyping*

	Disc: females **Target:** males	**Disc:** males **Target:** females		**Disc:** BI **Target:** non-BI	**Disc:** non-BI **Target:** BI
Discovery sample score threshold	Target p-value	Target p-value		Target p-value	Target p-value
$p_T < 0.01$	5×10^{-4}	2×10^{-6}		2×10^{-2}	8×10^{-7}
$p_T < 0.05$	2×10^{-9}	1×10^{-8}		2×10^{-4}	8×10^{-5}
$p_T < 0.1$	3×10^{-12}	5×10^{-9}		2×10^{-7}	3×10^{-4}
$p_T < 0.2$	6×10^{-14}	2×10^{-12}		3×10^{-7}	4×10^{-5}
$p_T < 0.3$	1×10^{-14}	1×10^{-12}		2×10^{-7}	2×10^{-7}
$p_T < 0.4$	8×10^{-16}	9×10^{-14}		5×10^{-8}	3×10^{-9}
$p_T < 0.5$	5×10^{-15}	1×10^{-14}		1×10^{-9}	2×10^{-9}
$0.01 < p_T < 0.2$	4×10^{-12}	4×10^{-9}		3×10^{-6}	2×10^{-2}
$0.05 < p_T < 0.2$	4×10^{-7}	2×10^{-6}		2×10^{-4}	4×10^{-2}
$0.05 < p_T < 0.5$	7×10^{-9}	5×10^{-9}		5×10^{-7}	2×10^{-6}
$0.2 < p_T < 0.5$	9×10^{-4}	1×10^{-4}		5×10^{-4}	6×10^{-7}

We confirmed that the scores indeed represented variation across many sites, or more specifically, did not over-represent variation from a small number of regions. This could occur if a few SNPs had extreme odds ratios. We can discount this, as there were no very strong effects for common SNPs in this sample. Furthermore, an unweighted analysis, in which 0, 1 or 2 alleles at each nominally-associated SNP are simply summed, gave similar results to the weighted analysis for p_T thresholds up to 0.5 (data not shown).

Alternatively, a relatively small number of SNPs could be over-represented if a region contained many SNPs in linkage disequilibrium that were nominally associated and so

that region would effectively be represented multiple times in the overall score. Given that the MHC region shows an association with schizophrenia in our sample and that this region has hundreds of SNPs in LD that, to a large extent, all index the same association signal, this would be a concern. By focusing on a set of SNPs in LE, we can discount this second possibility. When based on the LE-pruned 74K subset, there are no individual SNPs in the full 740K ISC set that correlate with the any of the scores with a $P < 1\times10^{-10}$ and ordinary least squares (OLS) $R^2 > 0.01$. In contrast, there are 356 SNPs that correlate with the $p_T < 0.01$ score when it is based on all 740K SNPs instead of the pruned set; even for the $p_T < 0.5$ score based on all SNPs, there are 5 SNPs that correlate with it at this level of significance. In fact, for the $p_T < 0.01$ score based on all 740K SNPs, there are 19 SNPs that have an OLS $R^2 > 0.1$ with the score. Any correlation between the score and disease in target samples could therefore easily represent relatively small number of SNPs. For the LD-pruned 74K set, we see no such effects, thus the score represents the aggregate of a large number of SNPs, each of very small effect.

Even if individual SNPs were in fact over-represented in the score, this would not bias subsequent tests in independent target samples per se. However, the *interpretation* of a significant target test (as representing a polygenic effect from many sites genome-wide) would be harder to justify.

We also note that using a set of quasi-independent SNPs makes several steps of downstream analysis and interpretation easier (for example, the calibration of results with simulated data and the comparison of results across the frequency spectrum). However, our main results are fundamentally unchanged when based on all SNPs rather than the 74K subset (data not shown).

Lastly, we calculated pairwise identity-by-state (IBS) values for all pairs of individuals across the ISC and the other schizophrenia and bipolar samples used in the extended score analyses (described in Section S15, below). If a substantial number of individuals were present in both discovery and target samples, this could induce a correlation between target sample phenotype and score. No individual samples were duplicated within or across studies. (As discussed below, where necessary we removed samples that would have otherwise created a known overlap: in particular, we only used the independent STEP-BD component from the Sklar et al (2008) dataset, excluding the UCL cases and controls, as the UCL controls were also used in the ISC.)

14) Score analyses stratified by SNP location relative to nearest gene

Scores based on SNPs within genes showed stronger effects than those based on SNPs from “gene-deserts” (defined here as more than 500kb from the nearest RefSeq known gene). Of the 74K SNP set, 5,316 SNPs (~7%) are in gene-deserts, whereas 27,950 SNPs are within genes (~38%). In the ISC male/female discovery/target samples, the scores based on the gene-desert SNPs showed only marginal association with disease, particular when using males for the discovery sample (all $P > 0.01$; Table S16).

For a fair comparison with scores based on genic SNPs, we randomly sampled 5,316 of the 27,950 genic SNPs and used these to form scores; we repeated this procedure 500 times.

Based on the median *P*-values, scores based on genic SNPs tended to show greater differences between cases and controls in the target samples (Table S16). In addition, the proportion of the 500 replicates in which the target sample analysis was more significant than the equivalent gene-desert result was generally well above 50%. For the male/female discovery/target analysis, the results were particularly clear – this trend was seen in most score analyses and likely reflects the larger number of male cases leading to a more powerful discovery analysis.

Table S16: *Stratifying the score analyses by genic versus non-genic SNPs.*

	Discovery: females **Target:** males			**Discovery:** males **Target:** females		
Discovery sample score threshold	Target p-value (gene desert)	Median target p-value (genic)	Proportion genic < gene desert p-value	Target p-value (gene desert)	Median target p-value (genic)	Proportion genic < gene desert p-value
$p_T < 0.01$	0.31	0.198	0.62	0.84	0.054	0.98
$p_T < 0.05$	0.11	0.082	0.55	0.44	0.089	0.85
$p_T < 0.1$	0.01	0.031	0.37	0.34	0.020	0.94
$p_T < 0.2$	0.11	0.013	0.81	0.24	0.002	0.98
$p_T < 0.3$	0.11	0.012	0.81	0.17	0.002	0.97
$p_T < 0.4$	0.04	0.009	0.74	0.18	0.002	0.96
$p_T < 0.5$	0.03	0.005	0.73	0.08	0.002	0.92
$0.01 < p_T < 0.2$	0.17	0.030	0.80	0.14	0.012	0.88
$0.05 < p_T < 0.2$	0.46	0.085	0.86	0.33	0.009	0.96
$0.05 < p_T < 0.5$	0.08	0.030	0.67	0.09	0.010	0.84
$0.2 < p_T < 0.5$	0.04	0.182	0.24	0.11	0.327	0.22

Although this analysis is based on a very limited set of markers (just over 5000 SNPs to cover all genic common variation) it suggests the biologically plausible result that risk alleles tend to reside in genes more often than in gene-deserts. This also makes it less likely that stratification drives these results. If scores are based on all ~28K genic SNPs, the results are consistently order-of-magnitude stronger than the non-genic 5K SNPs (data not shown).

15) External replication/target samples used in the score analyses

We asked whether scores based on the entire ISC as a discovery sample were associated with disease in independent target samples. The target samples selected are independent whole-genome association studies of schizophrenia, bipolar disorder and other non-psychiatric disease. Each study used a similar genotyping platform (Affymetrix arrays, with at least 500K SNPs).

Schizophrenia case/control studies:

1) Molecular Genetics of Schizophrenia (MGS) consortium European-American sample (MGS-EA). These individuals were genotyped on the Affymetrix 6.0 array; there is no overlap with any individuals in the ISC. Following recommendations from MGS investigators (DL, PG), we included five principal component scores, calculated by MGS investigators, as covariates in the target sample analyses to account for within-sample stratification. The sample collection and GWAS QC procedures are described in Sanders et al (2008) and the companion manuscript describing the MGS-EA sample. This is the same sample as the MGS sample described above in the MGS and SGENE combined analyses of single SNP results.

2) Molecular Genetics of Schizophrenia (MGS) consortium African-American sample (MGS-AA). These individuals were genotyped on the Affymetrix 6.0 array; there is no overlap with any individuals in the ISC. Following recommendations from MGS investigators (DL, PG), we included two principal component scores, calculated by MGS investigators, as covariates in the target sample analyses to account for within-sample stratification. The sample collection and GWAS QC procedures are described in Sanders et al (2008) and the companion manuscript describing the MGS-AA sample.

3) *O'Donovan et al (2008).* Genotyping was performed on the Affymetrix 500K two-chip array; the controls are from the WTCCC control sample (effectively complete overlap). There is no overlap with any of the samples in the ISC. The O'Donovan et al manuscript reported results on a handful of SNPs from the ISC Bulgarian and Dublin samples; however, these individuals were not part of their whole-genome study, and are therefore not included in this dataset. The sample collection and GWAS QC procedures are described in O'Donovan et al (2008).

Bipolar disorder case/control studies:

1) *Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD).* These individuals were genotyped on the Affymetrix 500K two-chip array. The sample collection and QC procedures are described in Sklar et al (2008). That report also includes GWAS data on additional individuals collected in the UK (UCL sample, Hugh Gurling). Because the UCL controls described in Sklar et al overlap with controls used in the ISC, we removed all UCL individuals (cases and controls) from the Sklar et al dataset, leaving the STEP-BD individuals only. Of the samples included here, there is no overlap with the ISC. These samples were also reported in the meta-analysis by Ferreira et al (2008).

2) *Wellcome Trust Case Control Consortium (WTCCC) bipolar disorder study.* These individuals were genotyped on Affymetrix 500K two-chip arrays. Sample collection and GWAS QC details are reported in WTCCC (2007). These samples were also reported in the meta-analysis by Ferreira et al (2008). There is no overlap with any ISC samples; the controls overlap with the O'Donovan et al dataset and also with the other WTCCC non-psychiatric diseases.

Non-psychiatric diseases:

1) *Wellcome Trust Case Control Consortium (WTCCC) study.* Excluding bipolar disorder, six diseases were included: coronary artery disease (CAD), Crohn's disease (CD), hypertension (HT), rheumatoid arthritis (RA), type 1 diabetes (T1D) and type 2 diabetes (T2D). These individuals were genotyped on Affymetrix 500K two-chip arrays. Sample collection and GWAS QC details are reported in WTCCC (2007). As mentioned above, the same WTCCC control sample is used for the O'Donovan et al, the WTCCC bipolar comparisons, as well as these six diseases. As indicated, all studies used similar genotyping platforms containing the ~74K SNP subset; naturally, in particular external datasets some small proportion of the 74K SNPs will be missing, removed in their QC stages.

Table S17 gives the results for in terms of target sample p-values and pseudo-R^2 values for all comparisons performed. We restricted our attention to the p_T>0.1 threshold values, as these were the more highly significant comparisons within the ISC. The data in Table S17b is also represented in Figure 2 (main text).

Table S17a,b: *Extension of the score analyses to external samples:*

a)		**Target sample p-value**				
	$p_T < X$	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
Schizophrenia	MGS-EA	1×10^{-21}	3×10^{-22}	1×10^{-24}	5×10^{-27}	2×10^{-28}
	MGS-AA	3×10^{-2}	3×10^{-2}	3×10^{-2}	9×10^{-3}	8×10^{-3}
	O'Donovan	8×10^{-7}	6×10^{-9}	7×10^{-10}	3×10^{-11}	5×10^{-11}
Bipolar disorder	STEP-BD	5×10^{-6}	1×10^{-8}	2×10^{-8}	2×10^{-9}	7×10^{-9}
	WTCCC-BP	4×10^{-9}	6×10^{-11}	6×10^{-11}	1×10^{-11}	1×10^{-12}
Non-psychiatric	WTCCC-CAD	0.28	0.60	0.62	0.64	0.71
	WTCCC-CD	0.05	0.13	0.05	0.06	0.05
	WTCCC-HT	0.69	0.44	0.36	0.38	0.30
	WTCCC-RA	0.13	0.42	0.65	0.54	0.65
	WTCCC-T1D	0.06	0.04	0.15	0.18	0.23
	WTCCC-T2D	0.34	0.09	0.07	0.12	0.06

b)		**Target sample R^2 (x100)**				
	$p_T < X$	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
Schizophrenia	MGS-EA	2.24	2.41	2.76	3.00	3.18
	MGS-AA	0.27	0.24	0.23	0.32	0.36
	O'Donovan	1.30	1.80	2.02	2.34	2.30
Bipolar disorder	STEP-BD	1.16	1.81	1.75	2.05	1.88
	WTCCC-BP	0.99	1.23	1.22	1.31	1.45
Non-psychiatric	WTCCC-CAD	0.03	0.01	0.01	0.01	0.00
	WTCCC-CD	0.11	0.07	0.11	0.10	0.11
	WTCCC-HT	0.00	0.02	0.02	0.02	0.03
	WTCCC-RA	0.07	0.02	0.01	0.01	0.01
	WTCCC-T1D	0.10	0.11	0.06	0.05	0.04
	WTCCC-T2D	0.02	0.08	0.09	0.07	0.10

The ISC is the discovery sample, used to generated scores (indexing schizophrenia risk) that are applied in a range of independent target samples, including schizophrenia, bipolar disorder and non-psychiatric disease. The top panel gives the target sample p-value; panel b gives the variance explained (difference in pseudo-R^2).

Figure S6 shows the mean score values in target sample cases and controls, based on the $p_T < 0.5$ score and using the entire ISC as the discovery sample. Specifically, the y-axis represents the residual score after regressing score on genotyping rate; it is rescaled so that the observed minimum and maximums are set at 0.0 and 1.0 respectively. The results reported in Figure 2 therefore represent direct tests of the within-study case/control differences shown here. As evident in Figure 2, there was a clear pattern in which all studies of psychotic mental illness (schizophrenia and bipolar disorder) showed higher rates of ISC-derived schizophrenia score alleles than their matched controls. In contrast, there were no differences for the $p_T<0.5$ scores in the six non-psychiatric diseases. Figure S6 also shows that the African-American sample (MGS-AA) had a markedly lower rate of all these alleles. This is consistent with the results from the analyses stratified by score allele frequency (see below) that reported a tendency for less frequent SNPs to show stronger effects in the score analyses. Such differences are largely expected and represent between-study and between-country differences in ancestral composition.

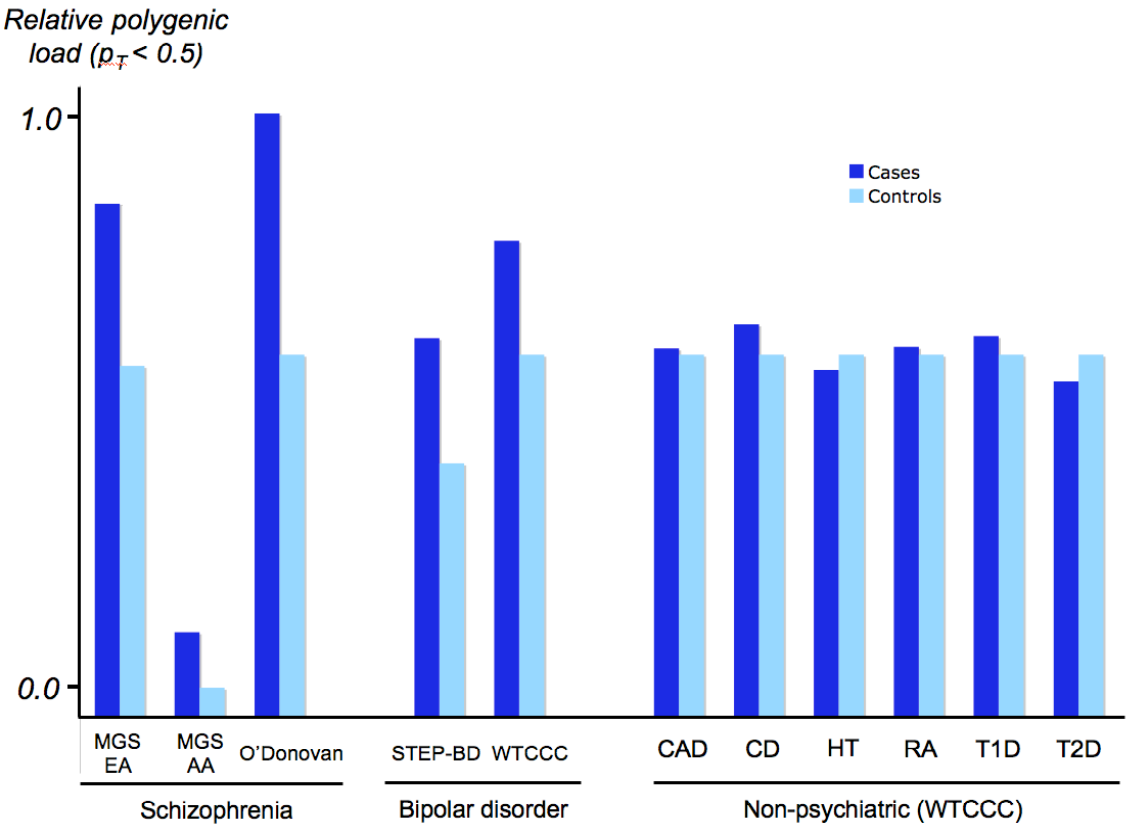


Figure S6: *Mean target score for cases and controls in the extended score analyses. Scores are based on the entire ISC as the discovery sample, for the $p_T<0.5$ threshold. The y-axis represents the residual score after regressing the score on genotyping rate and rescaled so that the observed minimum and maximum score means are set to 0.0 and 1.0 respectively.*

16) Simulation study to estimate the variance explained by common variants

To gain insight into the types of genetic models that are consistent with our data, we adopted a simulation-based approach. In our primary male/female analysis, and also in the ISC/MGS-EA analysis (i.e. utilizing the largest pair of single studies to form discovery and target), the polygenic scores captured up to ~3% of the variance in disease risk. Of the SNPs included in the score, only a small proportion will be true risk alleles and the remainder will be null loci (i.e. false positives at the p_T threshold). Therefore, questions we might want to ask are:

Q1. What is the proportion of all independent SNPs that are true risk variants?

Q2. Of true risk alleles, what distributions of effect sizes (genotypic relative risks) are consistent with our results?

Q3. Of the true risk alleles, what distributions of allele frequencies are consistent with our results?

Q4. What is the total variance explained by the subset of true risk alleles (that is diluted in our scoring approach because many false positives are also included)?

However, there will be many possible configurations of the number and frequency/effect size distributions of true risk alleles that are consistent with our data. A further confounding factor is the average extent of LD, as our GWAS comprises randomly selected markers and associations are very unlikely to reflect the actual causal variant – instead, the SNP will be in some degree of LD with the causal variant. Because of these factors, it will not be possible to obtain precise answers from our data alone. However, we can make some broad statements about a) the likely substantial total proportion of variance explained by these risk alleles and b) the inability of extreme models only invoking a relative small number of variants or multiple rare variants (MRV) to account for our findings.

Overview of approach

Our approach is to:

a) Simulate discovery and target datasets that are comparable to those used in this study, under a variety of genetic models
b) Repeat the score analyses across a range of p_T thresholds for each pair of simulated discovery/target datasets, in order to identify models that produce profiles of results similar to the real data in terms of variance explained by the score, R^2
c) Calculate the implied variance explained by the subset of true risk alleles from the selected model

Considering the number of variants and their average effect size, there will be many different combinations, all other things being equal, that lead to the same variance explained. For example, 5% of the additive genetic variance could be accounted for by:

- 5 loci that each explain, on average, 1% of the variance
- 50 loci that each explain, on average, 0.1% of the variance
- 500 loci that each explain, on average, 0.01% of the variance

Depending on sample size and therefore power, these three models would give quite different profiles in the score analyses, however, when considering the range of p_T thresholds used to select score alleles. In the first scenario, there would be greater power to detect these loci of large effect at highly significant *P*-values. As such, we would predict scores based on very low p_T thresholds to perform well; adding in further SNPs at less stringent p_T thresholds would only be adding noise, and so the target sample R^2 would drop. In contrast, the third scenario would be likely to show a different profile, in which the pseudo-R^2 continues to increase with less stringent p_T thresholds, as more of the 500 risk alleles are detected and included in the score. The second scenario will show an intermediate profile. Given our current sample size, we expect to be able to distinguish between some but not all possible models that lead to a similar total variance explained.

Range of models simulated

The primary set of simulations generated datasets with 1661 cases and 1793 controls for both the discovery and target samples, reflecting the average sample sizes for the male/female analyses reported above.

For each dataset, we generated genotype data for autosomal SNPs assuming linkage and Hardy-Weinberg equilibrium from the empirical allele frequency distribution observed in ISC controls. Each pair of discovery and target sample was drawn from the same, homogeneous population.

Even though the 74K SNP panel was pruned to remove SNPs in strong linkage disequilibrium, for the purpose of appropriately calibrating the simulations, we simulated datasets based on the estimated number of effectively independent SNPs rather than 74,062 SNPs exactly. Although the residual LD in the 74K panel is not an issue for the primary score analyses, as demonstrated above, ignoring it induces a minor bias (quantified below) as we are simulating SNPs under complete LE. We estimated the effective number of independent SNPs, $N_{LE,}$ as

$$74{,}062 \times (2 * 74{,}062 / \mathrm{Var}(S))$$

where *S* is, for each of 3,000 datasets simulated under the null, the sum of chi-square statistics for association tests of all 74,062 SNPs. Null datasets were generated by permuting phenotype labels within the seven sites for the ISC dataset, thereby preserving the LD structure between the 74K SNPs and any between-site variation.

The sum of *N* independent 1 df chi-squares has expected variance 2*N*. To the extent that SNPs are in LD, the variance of *S* will be larger. Based on the ratio of expected to observed variances, the equation above gives an estimate of the equivalent number of SNPs in LE. The estimated ratio was 0.7421, leading to N_{LE} = 54,963. (Note: this does

not imply that a subset of 55K actual SNPs in perfect LE exists, which we might otherwise have used in the initial score analyses instead of the 74K).

In the first instance, we directly simulated effects for the "observed" SNP markers. In each case, a certain proportion of markers were assigned non-null effects from one of three effect size distributions:

a) fixed GRRs, such that non-null markers all have the same population GRR irrespective of allele frequency
b) fixed variance explained, such that non-null markers with lower minor allele frequencies will have larger GRRs
c) exponential distribution of GRRs, such that non-null markers will have a population GRR drawn from an exponential distribution with a certain mean, irrespective of allele frequency

In reality, we expect most associated genotyped SNPs will not be causal but will be in LD with the causal allele. We extended the simulations to explicitly model effects at an unobserved causal variant, which is in LD with the marker. The effect observed at the marker is therefore attenuated, depending on the extent of LD. In this context, we specify that a certain proportion of the marker SNPs will be in LD with a causal variant. Causal variants are simulated from one of two frequency distributions: uniform or U-shaped, bounded to a minor allele frequency of 2% in the initial simulations, as described in Wray et al (2007). The same three effect distributions (fixed GRR, fixed variance explained and exponential GRR) were similarly applied to the causal SNPs. The allele frequency for the causal variant was simulated independent of the allele frequency of the marker. In each case, we fixed D'=1 in order to calculate the GRRs at the marker and the implied r^2 between marker and causal variant. We also performed simulations in which we set D'=1 but added the further constraint that two minor alleles should be in coupling phase. This effectively increases the mean r^2 generating the maximum r^2 given the allele frequencies and D' = 1; the primary results reported here are based on simulations with this constraint. Our specification of LD is necessarily arbitrary and unrealistic and we will not be able to estimate the true extent of LD between SNPs and the unobserved causal variants. Rather, we are focusing on two simple scenarios: complete LD (r^2=1) versus moderate/strong LD (D'=1, r^2<1).

The proportion of SNPs with a non-null effect, or that tag a causal variant, is set at different values depending on the particular model under consideration. In all, we performed a first round of simulations, with 560 different models selected to form a grid across a range of values for:

- Effect size distribution type and mean effect size
- Whether the effect is modeled at the marker, or at a causal variant in LD with the marker
- Frequency distribution of unobserved causal variants
- Proportion of independent non-null sites

For each simulated dataset, we calculated the variance explained by the score in the disease state (R^2), using the same analytic approach as for the real data, for the seven

p_T thresholds from 0.01 to 0.5. We also calculated the implied total additive genetic variance, both at the markers (V_M) and at the underlying causal variants (V_A) as well as the mean LD (r^2) between marker and causal variant. We calculated the variance explained using a liability-threshold model to combine the effects of multiple variants (and so estimate the variance explained in *liability*, rather than on disease state per se). If the GRR is small then an approximation of the variance explained on the liability scale for variant *l* is:

$$v_l = \text{Var(on liability scale)} = 2p_l(1-p_l)(GRR_l-1)^2 / i^2$$

with allele frequency p_l and $i = z / K$, with K as the disease prevalence and z the height of the curve at the truncation point (Falconer & Mackay, 1996). Taking K=0.01 implies i = 2.665. A variant with a MAF of 0.2 and GRR of 1.05 therefore accounts for ~0.011% of the variance in liability. The difference in variance explained in liability varies by an order of magnitude for a SNP of this GRR and MAF of 0.02 (~0.0014% of the variance) compared to a MAF of 0.5 (0.018% of the variance). The total additive variance is the sum $V = \Sigma_l v_l$.

Results of simulations

The majority of the models considered did not produce consistent profiles of results (data not shown): the target sample R^2 was typically far too small or great (e.g. 0% or >>10%), compared to the observed results, that ranged from ~0.5 to ~3%. Further, because we simulated across a grid of values, for convenience, not all models were biologically possible, e.g. implying a total genetic variance greater 100%; these were excluded from further consideration.

However, some models produced profiles of results that were broadly consistent with those observed in the ISC. To follow these up we performed 15 replicates across a finer grid in the parameter space around these models. Here we focus on seven consistent models, that illustrate a) how a number of different models are consistent with the data, but b) that all consistent models yield similar estimates of the total variance explained, V_M. Importantly, this list of seven models is not exhaustive: there will be a further spectrum of other models that fit the data, for example interpolating between the models presented here.

The first model selected (M_1) assumes a fixed GRR distribution, with a mean GRR of 1.05 modeled at the marker itself (i.e. r^2=1) at 6% of all independent sites. These parameters, along with those for the other six models considered here, are given in Table S18a. Averaging over replicates, this model produces a profile of pseudo R^2 values consistent with those in the ISC (Figure 3a and 3b; Table S18c). The mean V_M implied by this model is 34.8%; the mean V_A will necessarily be the same, as we assume r^2=1.

An equivalent overall fit can be obtained for similar models but with different balances of number and average size of effect, within a certain range. For example, model M_2 specifies more sites (25%) but with smaller average effects (GRR=1.025). Naturally,

many other combinations of these two parameters would also be consistent (but not all, as we show below). Model M_2 implies a total V_M of 34.7%, which is very similar to M_1.

We selected the five remaining models to illustrate how models with different effect size distributions and frequency/LD models can also yield consistent results for particular combinations of mean effect and proportion of sites. Model M_3 relaxes the assumption that the marker perfectly tags the causal variant (which is to say, we now simulate effects at an unobserved marker rather than directly at the marker). In this case, assuming a uniform frequency distribution for the causal variants, a model with 11% of sites having a mean GRR of 1.05 provides a good fit to the data. The implied V_M is very similar to models M_1 and M_2, however, at 34.1%. The estimate of V_A under this model is 80.5%, which reflects the average LD between marker and SNP, which is r^2=0.43 (Table S18b).

We next selected a model similar to M_3 but with a U-shaped allele frequency distribution for the causal variants. This model, M_4, again provides a good fit to the data and again provides a similar estimate of the V_M. Similarly, the next two models used different effect size distributions, either fixing the variance explained (M_5) or using an exponential distribution of GRRs (M_6). In both cases it was possible to find values of the mean effect and proportion of sites which lead to profiles of results similar to those observed in the ISC (Figure 2 main text; Table S18c below) but also broadly similar estimates of V_M compared to the other models. Finally, M_7 is another exponential model for which all sites show some non-null effect, with a small mean GRR of 1.012.

Table S18: *Results from polygenic simulations for seven consistent models.*
18a) Model descriptions.

Model	Effect dist.	LD model	Frequency model	Mean effect	Proportion of sites
M_1	Fixed GRR	$r^2 = 1$	Empirical (ISC)	1.05	6.25%
M_2	Fixed GRR	$r^2 = 1$	Empirical (ISC)	1.025	25%
M_3	Fixed GRR	$r^2 < 1$	Uniform	1.05	12%
M_4	Fixed GRR	$r^2 < 1$	U-shaped	1.04	32%
M_5	Fixed VE	$r^2 = 1$	Empirical (ISC)	0.00006	11%
M_6	Exponential	$r^2 < 1$	Uniform	1.025	25%
M_7	Exponential	$r^2 < 1$	Uniform	1.012	100%

18b) Mean estimates of variance explained and marker/causal variant LD.

Model	V_M	V_A	Mean r^2 (LD)
M_1	34.8	$=V_M$	=1
M_2	34.7	$=V_M$	=1
M_3	34.1	80.5	0.43
M_4	31.9	97.9	0.26
M_5	36.1	$=V_M$	=1
M_6	32.5	76.8	0.42
M_7	35.5	84.1	0.42

18c) Mean target sample pseudo R^2 (and in the observed data, Obs).

	Target sample 100 × R^2, based on p_T < X						
Model	**0.01**	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
Obs.	*0.55*	*1.10*	*1.70*	*2.30*	*2.50*	*2.75*	*2.85*
M_1	0.34	1.10	1.62	2.15	2.57	2.75	2.83
M_2	0.45	1.33	1.67	2.08	2.30	2.48	2.54
M_3	0.54	1.29	1.74	2.14	2.49	2.60	2.73
M_4	0.27	0.81	1.28	1.87	2.10	2.29	2.37
M_5	0.49	1.20	1.67	2.30	2.55	2.84	3.02
M_6	0.46	1.03	1.50	2.04	2.35	2.55	2.69
M_7	1.09	1.80	2.06	2.71	2.82	2.97	3.02

Although more models could be listed here, two points are important to re-iterate. First, of the total set of models considered, all models that produced a profile of results consistent with the ISC yielded similar estimates of V_M These results are consistent with the simulations reported in Wray et al (2007). The models listed in Table S18 shows values between 32 and 36%, with a mean of 34% for V_M. Second, not all models that explained ~34% of the variance yielded consistent patterns of results, however, as described below when considering models with either a smaller number of loci (100 or fewer) or only rare effects.

If the simulations were based on 74,062 SNPs instead of 54,963 (i.e. using the actual number of SNPs instead of the effective number of independent SNPs) a very similar set of results and models was obtained. The models selected tended to specify slightly fewer or slightly weaker effects, and the mean estimate of V_M was slightly larger, ~38%). In both cases, however, we should note that the estimates depend on the true disease prevalence as well as the simplifying assumptions of the genetic models used.

Here we illustrate how this type of polygenic model appears with respect to the standard genome-wide association study metrics such as the genomic control lambda and Q-Q plots. Figure S7 shows the Q-Q plots for the ISC observed data and the equivalent simulated data (3,322 cases and 3,587 controls for 74K SNPs), generated under a model in which 5% of variants have a GRR of 1.05. For the simulated data, the lambda was ~1.05. For the observed ISC data (74K subset), the lambda was slightly higher at 1.08. As the simulations indicated, we would not expect the lambda to converge to exactly 1.00 even in the absence of any confounding factors, if there were in fact a substantial polygenic component underlying disease risk. In both plots, we see a characteristic small departure from the diagonal, starting at a relatively liberal *P*-value threshold.

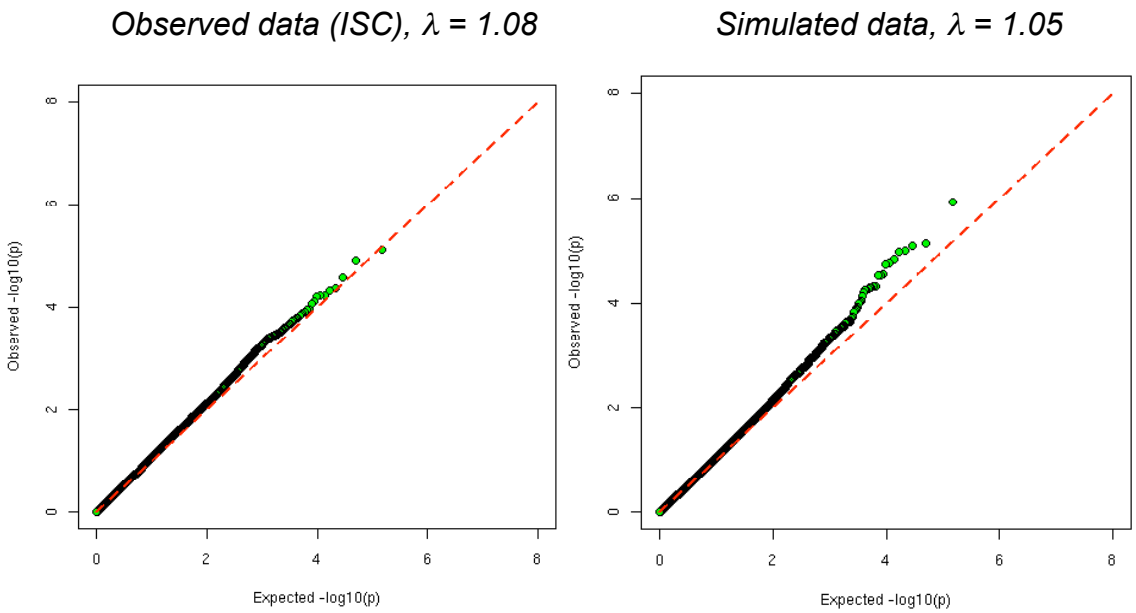


Figure S7: Q-Q plots of –log10(p) for the 74K SNP subset, in the ISC and in simulated data. The simulated data represents an equivalently-sized sample to the ISC, sampled from a single homogeneous population, under a polygenic model of disease risk (4.5% of SNPs with a GRR of 1.05 and uniform allele frequency distribution).

Inconsistent models

As described above, the simulations cannot distinguish between multiple possible models from those considered, although we obtain consistent estimates of the total variance explained. However, this is not to say that the simulations yield absolutely no information above and beyond the total variance explained. Here we consider models that predict a similar total variance explained by the marker genotypes (34%) from a relatively small number of causal variants, with relatively large effects. Specifically, we consider four models: 10, 20, 50 or 100 non-null marker loci, each explaining 3.4%, 1.7%, 0.68% or 0.34% of the variance. The implied GRRs for different minor allele frequencies is shown in Table S19.

Table S19: *Implied genotypic relative risks for models involving 10, 20, 50 or 100 loci, when the total variance explained is constrained to equal 34%.*

	Implied GRR, if MAF =		
Number of non-null markers	**0.05**	**0.2**	**0.5**
10	2.59	1.87	1.69
20	2.13	1.61	1.49
50	1.71	1.39	1.31
100	1.50	1.27	1.22

As shown in Figure 3c (main text), these models lead to qualitatively different profiles of results given the current sample sizes.

17) The risk allele frequency distribution and multiple rare variant (MRV) models

To gain insight into possible risk allele frequency spectra that might underlie the polygenic component of variance we detected, we repeated the score analyses stratified by allele frequency, i.e. only including certain SNPs in the discovery and target stages. In addition, we repeated the simulations described above adopting the same stratification procedure. We also performed a set of simulations to test whether a model involving only rare variants could account for our results.

Analysis stratified by score allele frequency

Based on the ISC controls, we obtained allele frequency quintiles with the following thresholds (truncated at 2 and 98%): 0.02, 0.136, 0.351, 0.65, 0.863 and 0.98. To assign the risk-increasing, or "scored" allele, we used a Cochran-Mantel-Haenszel analysis that conditions on sample strata (i.e. the same approach as used for the primary WGAS and primary discovery sample analyses).

Figure 4a (main text) illustrates the ISC/MGS discovery/target analysis, stratified by SAF as described above, and a similar pattern was observed from the male/female ISC analyses. The five PCs are included as covariates in the MGS-EA target analysis. We observed significant target analysis results across all quintiles, although the effects were particularly strong for the second and third quintiles, spanning score allele frequencies of ~14% to 65%. The inverted-U pattern across the frequency quintiles was expected, as it reflects the greater power to detect variants with a greater minor allele frequency (for a fixed GRR) in the discovery analysis and the greater variance they explain in the target analysis.

Stratified analysis of simulated data

We repeated the simulations for two illustrative models, stratifying by SAF quintile as described above (Figure 4b and 4c, main text). The data represented in Figure 4b were simulated under a "common variant" model, in which 20% of sites had a mean GRR of 1.05 and were in LD (D'=1) with the genotyped SNP. The central point is the relative profile of results by the SAF quintiles: similar profiles are observed for all models involving large numbers of common variants. Of note, the 2nd and 4th quintiles showed equivalent results for the simulated data (Figure 4b) but were clearly asymmetric in the ISC (Figure 4a). This is consistent with schizophrenia risk-increasing alleles being, on average, less frequent than randomly selected alleles at SNPs not associated with schizophrenia.

The common variant polygenic model implies that most affected individuals (and many unaffected individuals) will posses multiple risk alleles that are common in the population and of small effect. In contrast, a polygenic model that involves multiple rare variants (MRVs) describes a scenario in which there are many disease alleles *in the population*, but a given affected individual will possess only a small number (typically one) and unaffected individuals will typically possess none. This implies a large average

effect, for example involving GRRs of 10 or more. Here we ask whether our findings are consistent with a MRV model alone explaining the results.

We simulated data under MRV models. To allow for the possibility that multiple rare variants at the same locus happen, by chance, to reside on the same common haplotype background, we constrain causal allele frequencies to be less than 5% (based on a uniform frequency distribution between 1/5000 and 1/20). Under this model, we set 20% of sites to have a mean GRR of 1.5 at the causal locus. As above, common SNP markers (following the empirical allele frequency distribution of the ISC) were simulated to have D'=1 with the rarer causal variants. Although D' is high, because of the difference in allele frequencies, LD in terms of r^2 is necessarily very low: as such, these large effects at the rare causal locus emerge as very small apparent effects at the "genotyped" SNPs used in the score analyses, if at all. Rare variants will typically have different profiles of LD, such that a rare variant will have a D'=1 with multiple "independent" common SNPs across an extended region. Because of these complications, we do not attempt to calculate the total variance explained by the true subset of rare variants.

Crucially, though, we do not see the same profile of results across the range of score allele frequencies as we do for the ISC. Figure 4c (main text) shows results of this MRV model stratified by SAF: the skewing in favour of the less frequent SNPs is much greater than for the observed ISC results (Figure 4a). If the causal alleles are constrained to be much rarer (e.g. 1/1000 rather than 1/20), the skewing is even more marked. In contrast, in the ISC data we see the polygenic signal even for very common alleles (above 50% frequency), which would not capture very rare variation at other loci at all well. As such, we can with confidence reject extreme forms of the MRV model as being the sole explanation for the polygenic component underlying the score analysis results presented here.

18) Impact of discovery sample size on polygenic score estimation

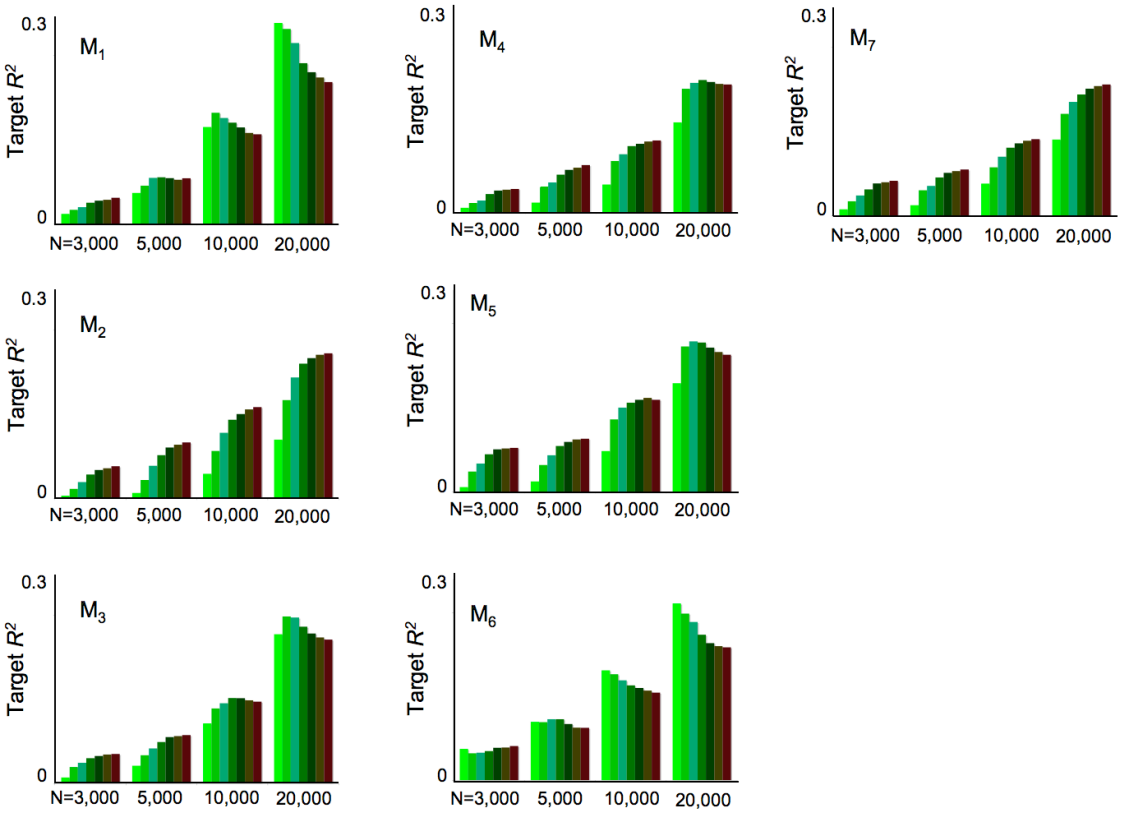
Given the current discovery sample sizes, the measured polygenic score accounts for only a small percentage of disease risk, up to ~3%. In this Section, we illustrate how increasing the discovery sample size, via large-scale meta-analytic efforts, for example, increases the variance explained by the measured score and also allows for discrimination between the multiple models that currently fit the data equally well.

We simulated data under the seven models (M_1-M_7) for discovery sample sizes of 3,000, 5,000, 10,000 and 20,000 case/control pairs. (We also set the target sample size to equal the discovery size in each case, although once the target sample size is above a certain minimum level, increasing it will have little effect on these simulations. That is, it is the discovery sample size that drives the effects shown here, and we would have observed similar results if the target sample sized was fixed to N=1,000 pairs in all cases, for instance.)

Figure S8 shows the pseudo R^2 metric for the seven models under these four sample size scenarios (plotted for the seven p_T thresholds as above). Of note, for the largest

sample size considered, the variance explained by the score is over 20%. Also, the seven models that show similar profiles with respect the p_T thresholds in the original simulations, and also for the N=3,000 condition, now show a range of profiles for the N=20,000 condition, suggesting that with larger samples we will be able to refine the underlying model. Of course, if larger sample collections introduce greater levels of heterogeneity and population differences, we might not realize these benefits. Nonetheless, these results suggest that with appropriately sized samples further exploration of the polygenic models proposed here is feasible. Under the models M_1 and M_6, for example, the most predictive score (from the seven p_T values considered here) is obtained with the p_T<0.01 threshold, as opposed to the p_T<0.5 threshold. If these models are true, genotyping panels of hundreds (rather than tens of thousands) of SNPs may have the potential for use in genetic risk prediction, as concluded by Wray et al (2007, 2008).

Figure S8: *Impact of increasing sample size on score analysis.*



The bars in each sample size bin (from light to dark / left to right) reflect the seven p_T thresholds: 0.01, 0.05, 0.1, 0.2, 0.3, 0.4 and 0.5. The Y-axis represents the target sample pseudo R^2 metric. The *N* in each panel represents the discovery and target sample size (in case/control pairs).

19) Supplementary References

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), 4th edn.*, (American Psychiatric Association, Washington, DC, 2000).

de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, Ke X, Monsuur AJ, Whittaker P, Delgado M, Morrison J, Richardson A, Walsh EC, Gao X, Galver L, Hart J, Hafler DA, Pericak-Vance M, Todd JA, Daly MJ, Trowsdale J, Wijmenga C, Vyse TJ, Beck S, Murray SS, Carrington M, Gregory S, Deloukas P, Rioux JD. (2006) A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. Nat Genet 38, 1166-72.

Devlin B & Roeder K (1999) Genomic Control for Association Studies. Biometrics, 55:997-1004.

Falconer DS & MacKay TFC (1996) Introduction to Quantitative Genetics, Ed 4. Longmans Green, Harlow, Essex, UK.

Ferreira MA, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, Fan J, Kirov G, Perlis RH, Green EK, Smoller JW, Grozeva D, Stone J, Nikolov I, Chambert K, Hamshere ML, Nimgaonkar VL, Moskvina V, Thase ME, Caesar S, Sachs GS, Franklin J, Gordon-Smith K, Ardlie KG, Gabriel SB, Fraser C, Blumenstiel B, Defelice M, Breen G, Gill M, Morris DW, Elkin A, Muir WJ, McGhee KA, Williamson R, Macintyre DJ, Maclean AW, St Clair D, Robinson M, Van Beck M, Pereira AC, Kandaswamy R, McQuillin A, Collier DA, Bass NJ, Young AH, Lawrence J, Nicol Ferrier I, Anjorin A, Farmer A, Curtis D, Scolnick EM, McGuffin P, Daly MJ, Corvin AP, Holmans PA, Blackwood DH; Wellcome Trust Case Control Consortium, Gurling HM, Owen MJ, Purcell SM, Sklar P, Craddock N. (2008) Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. Nat Genet, [Epub ahead of print] PMID: 18711365.

Friedman, A. & Perrimon, N. Genome-wide high-throughput screens in functional genomics. Curr Opin Genet Dev 14, 470-6 (2004).

International Schizophrenia Consortium (2008) Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature, 455, 237-41.

Janca, A., Ustun, T.B., Early, T.S. & Sartorius, N. The ICD-10 symptom checklist: a companion to the ICD-10 classification of mental and behavioural disorders. Soc Psychiatry Psychiatr Epidemiol 28, 239-42 (1993).

Jungerius BJ, Hoogendoorn MLC, Bakker SC, van't Slot R, Bardoel AF, Ophoff RA, Wijmenga C, Kahn RS & Sinke RJ (2008) An association screen of myelin-related genes implicates the chromosome 22q11 PIK4CA gene in schizophrenia. Molecular Psychiatry 13, 1060 – 1068.

Kirov G, Gumus D, Chen W, Norton N, Georgieva L, Sari M, O'Donovan MC, Erdogan F, Owen MJ, Ropers HH, Ullmann R. (2008) Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. Hum Mol Genet 17, 458-65.

Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, Cawley S, Hubbell E, Veitch J, Collins PJ, Darvishi K, Lee C, Nizzari MM, Gabriel SB, Purcell S, Daly MJ, Altshuler D. (2008) Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat Genet. 40(10):1253-60.

Lichtenstein P, Björk C, Hultman CM, Scolnick E, Sklar P, Sullivan PF. (2006) Recurrence risks for schizophrenia in a Swedish national cohort. Psychol Med 36, 1417-25.

Makeyev, E. V., Zhang, J., Carrasco, M. A. & Maniatis, T. (2007) The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. Mol Cell 27, 435-48.

McCarroll, S., Kuruvilla FG, Korn JM, Cawley S, Nemesh J, Wysoker A, Shapero MH, deBakker PW, Maller J, Kirby A, Elliott AE, Parkin M, Hubbell E, Webster T, Mei R, Handsaker R, Lincoln S, Nizzari M, Blume J, Jones K, Rava R, Daly MJ, Gabriel SB, and Altshuler D Integrated detection and population-genetic analysis of SNPs and copy number variation Nat Genet, in press (2008).

Nagelkerke NJD (1991) A note on a general definition of the coefficient of determination. *Biometrika* 78(3):691-692.

O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskvina V, Nikolov I, Hamshere M, Carroll L, Georgieva L, Dwyer S, Holmans P, Marchini JL, Spencer CC, Howie B, Leung HT, Hartmann AM, Möller HJ, Morris DW, Shi Y, Feng G, Hoffmann P, Propping P, Vasilescu C, Maier W, Rietschel M, Zammit S, Schumacher J, Quinn EM, Schulze TG, Williams NM, Giegling I, Iwata N, Ikeda M, Darvasi A, Shifman S, He L, Duan J, Sanders AR,

Levinson DF, Gejman PV; Molecular Genetics of Schizophrenia Collaboration, Gejman PV, Sanders AR, Duan J, Levinson DF, Buccola NG, Mowry BJ, Freedman R, Amin F, Black DW, Silverman JM, Byerley WF, Cloninger CR, Cichon S, Nöthen MM, Gill M, Corvin A, Rujescu D, Kirov G, Owen MJ (2008) Identification of loci associated with schizophrenia by genome-wide association and follow-up. Nat Genet, Jul 30. [Epub ahead of print] PMID: 18677311

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 81(3):559-75.

Reich, DE & Goldstein, DB (2001) Detecting association in a case-control study while correcting for population stratification. Genet Epidemiol 20, 4-16.

Resnick M, Segall A, G, GR, Lupowitz, Z, Zisapel, N. (2008) Alternative splicing of neurexins: a role for neuronal polypyrimidine tract binding protein. Neurosci Lett 439, 235-40.

Sanders AR, Duan J, Levinson DF, Shi J, He D, Hou C, Burrell GJ, Rice JP, Nertney DA, Olincy A, Rozic P, Vinogradov S, Buccola NG, Mowry BJ, Freedman R, Amin F, Black DW, Silverman JM, Byerley WF, Crowe RR, Cloninger CR, Martinez M, Gejman PV. (2008) No significant association of 14 candidate genes with schizophrenia in a large European ancestry sample: implications for psychiatric genetics. Am J Psychiatry, 165(4):497-506.

Shi et al. Full title and authorship to be determined. SGENE companion manuscript

Sklar P, Smoller JW, Fan J, Ferreira MA, Perlis RH, Chambert K, Nimgaonkar VL, McQueen MB, Faraone SV, Kirby A, de Bakker PI, Ogdie MN, Thase ME, Sachs GS, Todd-Brown K, Gabriel SB, Sougnez C, Gates C, Blumenstiel B, Defelice M, Ardlie KG, Franklin J, Muir WJ, McGhee KA, MacIntyre DJ, McLean A, VanBeck M, McQuillin A, Bass NJ, Robinson M, Lawrence J, Anjorin A, Curtis D, Scolnick EM, Daly MJ, Blackwood DH, Gurling HM, Purcell SM. (2008) Whole-genome association study of bipolar disorder. Mol Psychiatry, 13(6):558-69.

Stefansson H, Rujescu D, Cichon S, Pietiläinen OP, Ingason A, Steinberg S, Fossdal R, Sigurdsson E, Sigmundsson T, Buizer-Voskamp JE, Hansen T, Jakobsen KD, Muglia P, Francks C, Matthews PM, Gylfason A, Halldorsson BV, Gudbjartsson D, Thorgeirsson TE, Sigurdsson A, Jonasdottir A, Jonasdottir A, Bjornsson A, Mattiasdottir S, Blondal T, Haraldsson M, Magnusdottir BB, Giegling I, Möller HJ, Hartmann A, Shianna KV, Ge D, Need AC, Crombie C, Fraser G, Walker N, Lonnqvist J, Suvisaari J, Tuulio-Henriksson A, Paunio T, Toulopoulou T, Bramon E, Di Forti M, Murray R, Ruggeri M, Vassos E, Tosato S, Walshe M, Li T, Vasilescu C, Mühleisen TW, Wang AG, Ullum H, Djurovic S, Melle I, Olesen J, Kiemeney LA, Franke B, Sabatti C, Freimer NB, Gulcher JR, Thorsteinsdottir U, Kong A, Andreassen OA, Ophoff RA, Georgi A, Rietschel M, Werge T, Petursson H, Goldstein DB, Nöthen MM, Peltonen L, Collier DA, St Clair D, Stefansson K (2008) Large recurrent microdeletions associated with schizophrenia. Nature, 455:232-6.

Steffanson et al. MGS companion manuscript . Full title and authorship to be determined

Vorstman JA, Chow EW, Ophoff RA, van Engeland H, Beemer FA, Kahn RS, Sinke RJ, Bassett AS (2008) Association of the PIK4CA schizophrenia-susceptibility gene in adults with the 22q11.2 deletion syndrome. Am J Med Genet B Neuropsychiatr Genet. Jul 21. [Epub ahead of print]

Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, Nord AS, Kusenda M, Malhotra D, Bhandari A, Stray SM, Rippey CF, Roccanova P, Makarov V, Lakshmi B, Findling RL, Sikich L, Stromberg T, Merriman B, Gogtay N, Butler P, Eckstrand K, Noory L, Gochman P, Long R, Chen Z, Davis S, Baker C, Eichler EE, Meltzer PS, Nelson SF, Singleton AB, Lee MK, Rapoport JL, King MC, Sebat J. (2008) Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 320, 539-43.

Wray NR, Goddard ME, Visscher PM (2007) Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Research*, 17(10):1520-1528.

Wray NR, Goddard ME, Visscher PM (2008) Prediction of individual genetic risk of complex disease. *Current opinion in Genetics & Development*, 18:257-263.