Algorithm to Generate haploSNPs

The algorithm to generate haploSNPs proceed as described below:

results =()
active = ()
create 2 single SNP haploSNPs (one for each allele) for the first SNP and add to *active*();
while(*active* is not empty)
    current_haplo := remove the current element of active
    extend *current_haplo* until terminating *mismatch* or until max length reached
    add *current_haplo* to *results*
    if terminated due to *mismatch*
        create 2 new haploSNPs:
            current_haplo + *mismatch_allele1*
            current_haplo + *mismatch_allele2*
        add to *active*
    if *active* is empty and there are more SNPs
        create 2 single SNP haploSNPs for the current SNP
        add to *active*
return *results*

WTCCC2 MS Imputation using 1000 Genomes

In order to increase coverage at rare variants, we performed imputation the impute2 software package[1] and the provided 1000 Genomes phase 3 [2] reference panel. We used the exact set of 14,526 individuals that we analyzed as part of our analysis of genotyped SNPs. For SNP QC, we removed any SNP that had an info score below 0.5, missingness above 0.002, a deviation from Hardy-Weinberg equilibrium at a p-value below 0.01, or had differential missingness between cases and controls with a p-value below 0.05. This was identical to the QC imposed on genotyped SNPs, except that SNPs with MAF below 0.01 were not removed to allow for rare imputed variants to influence resulting heritability estimates. After QC, we analyzed 2.2 million imputed SNPs. We estimated h2g-imputed of 0.098 (s.e. 0.007). This is the result of the well-known phenomenon of LD-bias[3-5] influencing heritability estimates. To deal with this bias, we employed the LD-regression technique of [4] and obtained a revised h2g-imputed-ld of 0.25 (s.e. 0.014). We compare this with estimates from genotype data after performing LD-regression (h2g-ld) of 0.24 (s.e. 0.14), and conclude that imputed SNPs do not explain substantially more heritability than genotyped SNPs alone.

Table S1. Estimates of heritability from haplotypes using GCTA

Cohort ID	h_{hap}^2	S.E.
ajsz	0.61	0.17
boco	1.02	0.16
clm2	0.71	0.08
clo3	0.94	0.17
gras	0.80	0.30

irwt	1.01	0.24
mgs2	0.32	0.11
s234	0.53	0.16
swe5	0.41	0.17
swe6	-0.17	0.29
Meta-analysis	0.63	0.05

Table S1. We estimated the heritability explained by haplotypes in each cohort using GCTA, instead of PCGC regression. Results were similar to those observed with PCGC regression but slightly lower. This may be reflective of known biases in REML estimation of heritability.

Table S2. Detailed composition of the subsets used for cross cohort analysis

Cohort ID	Cases	Controls	Total
Swedish Subset			
ersw	258	277	535
swe5	1568	1529	3097
swe6	713	810	1523
umeb	199	224	423
umes	81	71	152
Total	2819	2911	5730
British Subset			
clm2	3466	3463	6929
clo3	2069	2055	4124
Total	5535	5518	11053

Table S2. This gives a detailed composition of the two subset of individuals used in cross-cohort analysis. We note that the Swedish subset included data from some cohorts with N < 1,000 (not in Table S1) to increase the number of individuals in the subset.

Table S3. Estimates from individual variance components for PGC2-SCZ

All Variance Components For PGC2-SCZ		
haploSNP MAF Bin	**0.001-0.005**	0.2
	0.005-0.01	0.07
	0.01-0.05	0.1
	0.05-0.1	-0.03
	0.1-0.25	-0.04
	0.25-0.5	-0.02
SNPs		0.39
Total		0.64

Table S3. This gives heritability estimates from each of 7 variance components. LD between SNPs and common haploSNPs results in an increase in SNP heritability with corresponding negative estimates of heritability explained by common haploSNPs. Overall, the results show a large component of heritability explained by common SNPs in the joint model. This suggests that both common and rare variants play a role in the genetic architecture of schizophrenia.

Table S4. GCTA Biases With Case Control Ascertainment: Multiple Sclerosis

REML Estimation (GCTA)								
	Case 20%		Cases 40 %		Cases 60 %		Cases 80 %	
Model Components	h_{hap}^2	**S.D.**	h_{hap}^2	**S.D.**	h_{hap}^2	**S.D.**	h_{hap}^2	**S.D.**
haplo > 0.001 + SNPs	1.05	0.06	0.44	0.03	0.26	0.01	-0.03	0.04
haplo > 0.005 + SNPs	0.52	0.06	0.33	0.02	0.29	0.01	0.30	0.03
haplo > 0.01 + SNPs	0.34	0.05	0.26	0.02	0.26	0.01	0.31	0.02
haplo > 0.05 + SNPs	0.22	0.03	0.18	0.01	0.18	0.00	0.22	0.02
SNPs Only	0.20	0.03	0.16	0.01	0.17	0.00	0.20	0.01
PCGC Regression								
	Case 20%		Cases 40 %		Cases 60 %		Cases 80 %	
Model Components	h_{hap}^2	**S.D.**	h_{hap}^2	**S.D.**	h_{hap}^2	**S.D.**	h_{hap}^2	**S.D.**
haplo > 0.001 + SNPs	0.80	0.10	0.73	0.06	0.68	0.02	0.61	0.06
haplo > 0.005 + SNPs	0.40	0.08	0.38	0.04	0.36	0.01	0.31	0.03
haplo > 0.01 + SNPs	0.30	0.07	0.29	0.03	0.28	0.01	0.25	0.02
haplo > 0.05 + SNPs	0.20	0.04	0.19	0.02	0.19	0.01	0.17	0.02
SNPs Only	0.19	0.04	0.19	0.02	0.19	0.00	0.17	0.01

Table S4. REML heritability estimation[6,7] assumes that the phenotype follows a multivariate normal distribution. However, under scenarios involving case ascertainment—where the prevalence of a disease is much higher in the study than in the population—this assumption is unlikely to hold, and GCTA has been shown to be subject to bias[8]. This scenario is of particular interest to us as Multiple Sclerosis is a rare disease with population prevalence of 0.001[4]. To assess whether GCTA was, in fact, subject to bias in these scenarios we subsampled our data to alter the case control ratio and re-estimated the heritability explained. All heritability estimates are on the liability scale. We compared the continuity of the subsampled estimates when using REML and when using PCGC. For example, REML produces estimates of h2haplo of 0.44 and 0.26 for 40% and 60% cases, respectively, when analyzing the rarest haploSNPs. Indeed, our results suggest that the bias is worst for the rarest component. The estimates were produced over 50 random subsamplings of the data and the standard deviation is estimated these 50 subsamplings. We caution that the subsamplings are not independent and the estimates of the standard deviation will be biased downward as a result.

Table S5. Additionally Stringent QC in MS Data

	SNPs remaining	h_{hap}^2	Random Drop h_{hap}^2	S.D
Level 0 QC	419962	0.74	-	-
Level 1 QC	374555	0.67	0.69	0.019
Level 2 QC	313514	0.63	0.61	0.019

Table S5. To avoid inflation in heritability estimates due to assay artifact, we performed increasingly stringent QC. Level 0 QC removed any SNPs that were below 0.01 minor allele frequency, above 0.002 missingness, had deviation from Hardy-Weinberg equilibrium at a p-value below 0.01, or had differential missingness between cases and controls with a p-value

below 0.05. Level 1 QC removed any SNPs that had minor allele frequency below 0.02, had missingness greater than 0.002, had deviation from Hardy-Weinberg equilibrium at a p-value below 0.05, or had differential missingness between cases and controls with a p-value below 0.05. Level 2 QC removed any SNPs that had minor allele frequency below 0.02, had missingness greater than 0.002, had deviation from Hardy-Weinberg equilibrium at a p-value below 0.1, or had differential missingness between cases and controls with a p-value below 0.1. For each level of QC we compared the estimate of heritability after QC, to the estimate after randomly dropping 10 sets of SNPs with equal size.

Table S6. Estimates from individual variance components for WTCCC2-MS

All Variance Components For WTCCC2-MS		
haploSNP MAF Bin	**0.001-0.005**	0.38
	0.005-0.01	0.04
	0.01-0.05	0.09
	0.05-0.1	-0.02
	0.1-0.25	-0.01
	0.25-0.5	-0.01
SNPs		0.2
Total		0.67

Table S6. This gives heritability estimates from each of 7 variance components. LD between SNPs and common haploSNPs results in an increase in SNP heritability with corresponding negative estimates of heritability explained by common haploSNPs. Overall, the results show a large component of heritability explained by common SNPs in the joint model. This suggests that both common and rare variants play a role in the genetic architecture of multiple sclerosis.

Table S7. FDR Controlled SNP Associations to MS

Index SNP	Chromosome	Position	P-Value
rs3748816	1	2516606	2.28E-10
rs6662618	1	92707999	1.93E-06
rs11581062	1	101180107	3.11E-05
rs2298116	1	110337726	7.12E-05
rs1335532	1	116902480	4.00E-08
rs506813	1	164145122	4.62E-05
rs6666839	1	206244148	2.72E-05
rs1151694	1	245795585	3.08E-06
rs990270	2	40435853	1.08E-05
rs12373588	2	112182736	5.06E-06
rs10186133	2	113553415	8.27E-05
rs11901513	2	117500792	7.71E-14
rs2286896	2	191243821	7.09E-06
rs4973263	2	230652673	3.22E-05
rs11128570	3	11738001	2.79E-05

rs12493245	3	27772814	2.26E-05
rs3732730	3	70642948	2.98E-05
rs6796183	3	97047528	6.42E-05
rs1132200	3	120633526	3.69E-05
rs6808500	3	122992571	4.14E-06
rs9290375	3	171048784	3.33E-05
rs2175525	3	185369639	4.44E-05
rs907314	4	37930848	2.11E-05
rs478454	4	111682726	5.13E-05
rs2246528	4	179703677	7.17E-05
rs6860349	5	2186412	5.29E-05
rs7716642	5	4808398	2.36E-05
rs2290616	5	6800104	6.21E-06
rs1993879	5	35987897	7.85E-06
rs4613763	5	40428485	8.93E-07
rs12654328	5	65121428	4.08E-05
rs9293440	5	72091136	4.30E-05
rs1862182	5	96025282	4.47E-05
rs447909	5	104055999	4.01E-06
rs10462946	5	164095625	6.01E-06
rs11759658	6	25467524	5.98E-06
rs13194491	6	27145059	3.00E-06
rs6903535	6	28525201	1.08E-08
rs3129941	6	32445664	4.71E-130
rs11154801	6	135781048	2.47E-08
rs1738074	6	159385965	7.46E-05
rs12718729	7	50638844	8.36E-05
rs6979448	7	89773742	4.11E-05
rs705352	7	90536300	5.61E-05
rs11166797	8	139290846	2.79E-05
rs2188229	9	20615875	2.69E-05
rs2104286	10	6139051	4.37E-10
rs10995250	10	64066927	2.31E-05
rs1250539	10	80707235	4.29E-05
rs10887926	10	90930338	1.22E-05
rs6421571	11	118248982	2.99E-05
rs1860545	12	6317038	5.40E-06
rs12819780	12	29659102	1.64E-05
rs4768412	12	41155407	5.40E-06
rs10783850	12	56515644	2.41E-06
rs11146946	12	131530313	2.68E-05

rs1931949	13	73881727	4.64E-05
rs1572432	13	91059626	5.92E-06
rs7322759	13	98707679	8.47E-08
rs10144160	14	75032462	4.28E-05
rs880569	14	78154458	5.72E-05
rs6575267	14	92117086	6.70E-06
rs8039451	15	34795656	4.84E-05
rs2083061	15	57945327	2.31E-05
rs8059610	16	7248898	1.07E-05
rs12928822	16	11311394	1.15E-08
rs17796129	16	22979606	6.27E-05
rs12926028	16	77669335	8.14E-05
rs386965	16	78210042	1.18E-05
rs4072683	16	80445042	5.45E-05
rs8078776	17	4724664	2.76E-06
rs11650810	17	41183305	2.43E-05
rs180515	17	55379057	2.61E-05
rs1108591	17	59759101	2.42E-05
rs1077667	19	6619972	6.92E-06
rs6512102	19	16315110	1.66E-05
rs1029804	19	46799047	7.61E-05
rs2303759	19	54560863	1.35E-05
rs1041606	20	14625788	3.79E-06
rs175284	20	15159056	4.22E-05
rs2835753	21	37750192	2.08E-05
rs473304	22	19297450	5.13E-05

Table S7. We list the FDR controlled SNP associations for MS. We corrected for 375k hypotheses.

Table S8. FDR Controlled haploSNP Associations to MS

Chromosome	Position	P-Value
1	2719300	3.85E-10
1	4903193	2.11E-06
1	21590058	1.78E-06
1	54758580	1.36E-05
1	57806709	1.14E-06
1	84506998	1.79E-05
1	92721204	7.34E-07
1	101165270	2.25E-05
1	107854649	1.52E-05

1	110286408	1.10E-05
1	116892906	6.16E-10
1	160594555	1.18E-05
1	199200640	1.93E-05
1	206244686	1.85E-05
1	223064272	4.81E-06
1	232478868	1.24E-05
1	233161680	2.07E-05
1	240756629	6.73E-06
1	245814858	4.73E-06
2	3088072	2.46E-05
2	8498044	1.46E-05
2	33682723	2.55E-05
2	40468962	1.05E-05
2	56229628	5.73E-06
2	60912800	3.13E-06
2	72231465	9.37E-08
2	79815717	1.66E-05
2	112182736	1.81E-05
2	117504882	3.01E-25
2	118588283	1.91E-05
2	148984193	1.20E-05
2	157454543	1.15E-05
2	170688986	1.15E-05
2	188198410	7.77E-07
2	191265119	2.50E-05
2	207613199	2.28E-05
2	226440109	1.07E-05
2	227752598	2.71E-06
2	236224252	1.24E-05
3	5463284	9.69E-06
3	23968156	1.06E-05
3	26714226	1.67E-05
3	27774888	8.68E-09
3	70555395	2.11E-05
3	103177328	1.32E-05
3	114404025	9.61E-06
3	120611734	2.00E-06
3	122959905	4.54E-06
3	127556267	1.32E-05
3	135199072	2.47E-08

3	159900180	2.07E-05
3	182935773	3.54E-06
3	189701320	2.10E-05
3	190570338	1.83E-05
4	7203397	1.06E-06
4	12098894	4.78E-06
4	30177894	1.90E-05
4	37966765	2.56E-05
4	39880111	1.05E-05
4	133962776	2.57E-05
4	156806535	2.04E-05
4	179700714	2.34E-06
5	3035219	3.07E-06
5	4814729	3.34E-06
5	5657190	1.96E-05
5	6795027	6.05E-07
5	10637412	2.26E-05
5	16661615	7.82E-06
5	34030996	2.18E-05
5	35959512	1.75E-06
5	40425837	8.09E-06
5	53979369	1.17E-05
5	61149871	1.02E-05
5	68350817	1.89E-06
5	72254014	8.88E-06
5	75722535	2.17E-05
5	95166538	2.36E-05
5	99331259	1.39E-05
5	104090427	1.53E-06
5	129490462	8.01E-06
5	164063346	5.84E-06
5	169417766	4.89E-06
6	1368862	2.12E-05
6	13259061	1.22E-05
6	23293998	1.90E-05
6	25756772	2.39E-06
6	26584037	1.44E-05
6	28563808	4.76E-10
6	32525043	1.00E-205
6	36614559	1.40E-05
6	41483143	1.09E-05

6	90037551	4.06E-06
6	130201584	1.78E-05
6	135891591	4.10E-08
6	138000832	6.30E-06
6	154331869	1.80E-05
6	169376662	1.59E-05
7	1911061	1.83E-05
7	10780669	9.56E-06
7	22470368	6.61E-06
7	35518350	6.50E-06
7	37357359	6.17E-07
7	49858202	2.00E-06
7	54664303	1.95E-05
7	77489985	4.81E-08
7	89773742	2.27E-06
7	90753974	1.65E-05
7	101139597	3.20E-08
7	111566139	5.99E-06
7	148799811	1.58E-05
8	1869368	1.53E-05
8	4264526	9.80E-07
8	17048433	1.67E-05
8	24267829	2.01E-06
8	27837682	8.85E-06
8	40840276	1.44E-05
8	59951849	1.25E-05
8	69327741	1.77E-05
8	125795092	1.73E-05
8	134739424	1.55E-05
8	139292523	1.69E-06
9	1210489	1.31E-05
9	4657779	6.05E-06
9	12475043	2.63E-06
9	14041407	1.95E-05
9	82716577	9.63E-06
9	91132600	2.26E-05
9	110995016	2.90E-06
9	121845129	2.57E-05
10	4595981	2.91E-06
10	5603845	9.66E-06
10	6148553	1.41E-09

10	8448016	2.29E-05
10	10807697	2.56E-06
10	18535453	8.05E-07
10	59113708	1.63E-05
10	64083506	6.02E-06
10	73870261	6.14E-07
10	90951097	1.74E-06
10	119224537	8.89E-06
10	120665729	2.40E-05
11	19377871	1.19E-05
11	19919235	9.52E-06
11	60571568	6.72E-06
11	61944338	1.03E-05
11	69262442	1.37E-05
11	117182349	1.33E-05
11	118206757	2.02E-06
11	122010118	1.52E-05
11	124497287	1.83E-05
12	2386039	2.51E-05
12	3940843	1.72E-05
12	6317142	1.32E-06
12	12106084	4.61E-06
12	29659102	1.56E-05
12	41140060	2.62E-06
12	56278383	3.01E-06
12	74343670	5.37E-06
12	91171775	3.61E-06
12	97661739	2.62E-05
12	113536025	2.00E-05
12	121240108	2.07E-05
12	128048998	2.88E-06
12	129789808	8.97E-06
12	131533266	2.37E-05
13	25958105	1.22E-05
13	27343216	2.61E-05
13	37651751	6.82E-06
13	46620697	1.65E-05
13	70425725	8.77E-06
13	71505252	2.03E-05
13	73886364	5.01E-06
13	87769054	2.44E-05

13	91073032	6.41E-07
13	92546367	4.11E-06
13	96254057	1.49E-05
13	98862510	9.59E-09
14	29900348	2.43E-05
14	33323490	6.49E-06
14	39266330	1.42E-05
14	51966340	1.82E-05
14	56024200	1.97E-05
14	62829727	8.30E-06
14	68663107	9.00E-07
14	75555069	4.43E-07
14	78267403	1.39E-05
14	79392412	1.63E-05
14	81134715	1.55E-05
14	92109123	2.67E-08
14	93320667	1.57E-05
14	98632991	1.96E-05
14	100381662	1.98E-05
15	20408038	2.02E-05
15	31787777	1.44E-05
15	34784817	1.34E-05
15	57945327	3.16E-06
15	65030768	2.62E-05
15	66414909	2.61E-05
15	70873784	8.36E-06
15	71876174	1.69E-05
15	77503085	1.97E-05
15	81412919	1.98E-05
15	84537769	9.97E-06
15	86651988	1.84E-05
15	90016200	1.33E-05
15	90722674	2.06E-05
15	96063588	1.85E-05
15	96675637	1.87E-06
15	99749783	1.16E-05
16	1009493	1.03E-05
16	7251801	2.36E-05
16	7801319	2.11E-05
16	11324901	1.82E-08
16	50584930	1.11E-06

16	56303417	9.98E-06
16	64075024	8.87E-07
16	77669690	7.11E-06
16	78210754	6.69E-08
16	79743832	2.10E-06
16	84549153	7.73E-06
16	86283428	1.13E-05
17	2166993	1.69E-06
17	4689933	1.73E-05
17	12175962	7.00E-06
17	21103513	1.58E-05
17	33845109	1.02E-05
17	41175986	7.67E-06
17	51196467	2.17E-05
17	59760566	8.64E-06
17	65832229	2.87E-06
17	75520041	1.20E-05
18	7484650	5.39E-06
18	8866229	6.95E-06
18	37383752	1.91E-06
18	41324120	1.37E-05
18	42414611	6.40E-06
18	54895303	1.76E-05
19	2421899	2.51E-05
19	6617444	1.46E-07
19	8234957	2.46E-05
19	11705091	6.78E-06
19	13402663	1.90E-05
19	17240740	2.17E-06
19	18260842	1.04E-05
19	46762332	2.61E-06
19	52404736	7.68E-06
19	54598698	8.33E-06
19	56960319	1.82E-05
19	58498311	6.06E-06
19	60131940	7.32E-06
20	1879232	1.30E-05
20	10995424	1.06E-05
20	14455455	6.86E-07
20	15166667	4.95E-06
21	20929660	1.08E-05

21	26817078	1.91E-05
21	37745214	2.19E-06
22	18117859	5.15E-06
22	19297450	1.14E-05
22	34092987	1.06E-05

Table S8. We list the FDR controlled haploSNP associations for MS. We corrected for 38.8 M hypotheses. The position corresponds to the center of the most highly associated haploSNP in the locus.

Table S9. GRAIL Genes of Interest

Gene	GRAIL P-Value
TNFSF14	1.18E-06
HLA-DQB2	3.55E-06
LTB	3.56E-06
CD70	4.84E-06
CD80	5.68E-06
*CD276**	2.90E-05
CXCR5**	3.55E-05
FLT3LG	3.91E-05
LTA	4.30E-05
NCR3	4.64E-05
MICB	7.90E-05
LTBR	0.000133877
HLA-DMA	0.000297397
IRF8	0.000377354
TAP1	0.000528402
LST1	0.000532375
EOMES	0.000670055
*CXCL16**	0.000705149
CD5	0.000845074
*NCR1**	0.000947797
MICA	0.001160222
*IL31**	0.001562778
IL7R	0.001890805
*TBX1**	0.002064077

*These genes are at novel loci

**This locus has been previously reported in Sawcer et al. 2011[9] but was not genome-wide significant in that analysis.

Table S9. This table lists the set of 24 genes showing a high degree of relatedness based on textual analysis using GRAIL[10]. These genes were chosen as significant at an FDR < 0.05 after

correcting for 821 (the number of genes) hypotheses tested. We note that we have removed 12 genes from this list that were all part of a family of small nucleolar RNAs. These genes showed a very high degree of text-based similarity to each other as the small number of abstracts that reference each of them is identical. Thus, we believe that the high degree of similarity of these genes is driven by the fact that little is known about each of them individually. We believe that augmenting text-based analyses to reduce the effect of such genes is an interesting avenue of future research.

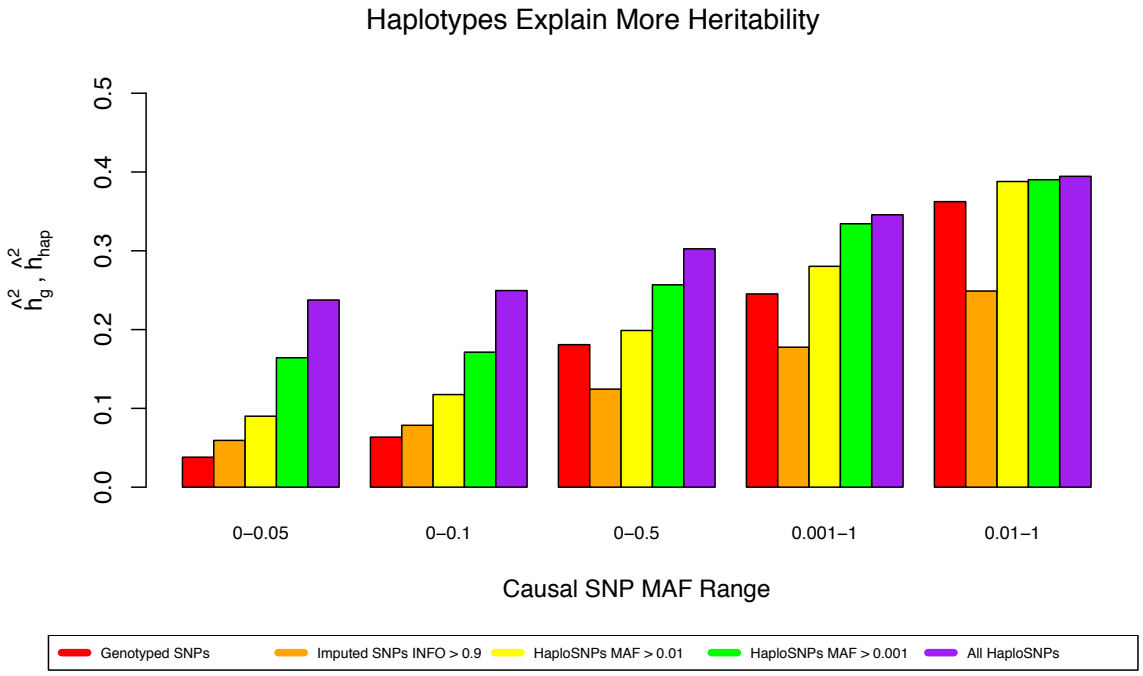

Figure S1. This is analogous to Figure 3 in the main text, including results from well-imputed (INFO > 0.9) SNPs. Our results show that well imputed SNPs do increase heritability explained for rare variant architectures. However, they are subject to known LD biases as genetic architectures become more common.

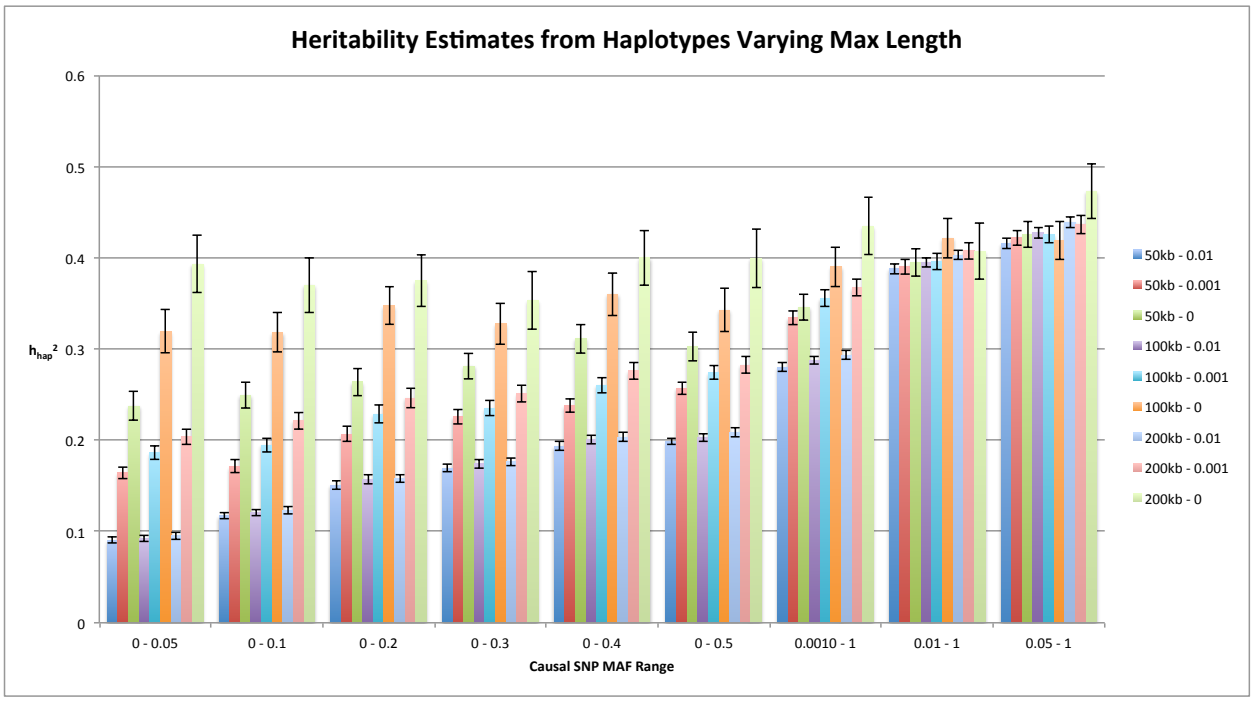

Figure S2. Assessing the heritability explained by haploSNP tags of variable lengths and with different MAF thresholds. The legend describes the maximum length and the minimum MAF of the haploSNPs. Given the large number of haploSNPs, and correspondingly large standard error, we chose to restrict to haploSNPs with MAF > 0.1% and length < 50kb in our real data analyses.

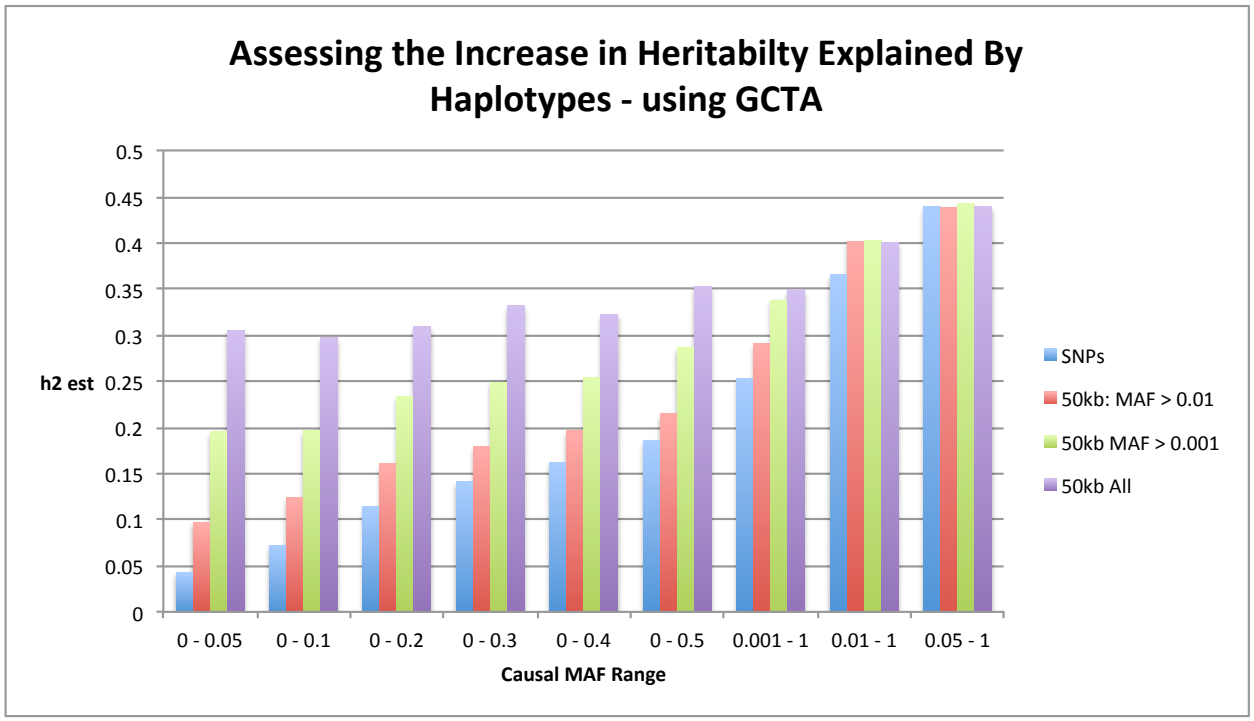

Figure S3. Assessing increases in tagging heritability in the UK10K data-set using haploSNPs using GCTA. Estimates are similar to those from PCGC regression as is expected in the absence of case control ascertainment.

References

1. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5,** e1000529 (2009).
2. 1000 Genomes Project Consortium *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491,** 56–65 (2012).
3. Speed, D., Hemani, G., Johnson, M. R. & Balding, D. J. Improved Heritability Estimation from Genome-wide SNPs. *The American Journal of Human Genetics* **91,** 1011–1021 (2012).
4. Gusev, A. *et al.* Quantifying Missing Heritability at Known GWAS Loci. *PLoS Genet* **9,** e1003993 (2013).
5. Lee, S. H. *et al.* Estimation of SNP heritability from dense genotype data. *Am. J. Hum. Genet.* **93,** 1151–1155 (2013).
6. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nature genetics* **42,** 565–569 (2010).
7. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88,** 76–82 (2011).
8. Golan, D., Lander, E. S. & Rosset, S. Measuring missing heritability: Inferring the contribution of common variants. *Proceedings of the National Academy of Sciences* **111,** E5272–81 (2014).
9. International Multiple Sclerosis Genetics Consortium *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476,** 214–219 (2011).
10. Raychaudhuri, S., Plenge, R. M., Rossin, E. J. & Ng, A. C. PLOS Genetics: Identifying Relationships among Genomic Disease Regions: Predicting Genes at Pathogenic SNP Associations and Rare Deletions. *PLoS Genet* (2009).