

Supplemental Online Material

Supplementary Table 1. Means and standard deviations of receptive language measures

Measure	Girls (<i>n</i> = 2434)	Boys (<i>n</i> = 1943)	<i>F</i>	<i>p</i>	μ^2
	M (SD)	M (SD)			
Vocabulary	0.01 (1.01)	0.00 (.99)	0.02	0.89	0.000
Listening grammar	0.03 (1.00)	-0.06 (.99)	8.94	0.003	0.002
Making inferences	0.06 (0.99)	-0.05 (1.02)	12.07	0.001	0.003
Figurative language	0.01 (0.99)	0.03 (0.99)	0.42	0.52	0.000

Note: Each measure was standardized to a mean of 0 and standard deviation of 1 based on the whole sample. *F*, *p* and μ^2 were derived from multivariate analyses of variance (MANOVAs) performed using one member of each twin pair. MANOVA data are as follows: intercept, $F(4, 4372) = 1.17, p < 0.01$; sex, $F(4, 4372) = 6.51, p < 0.01$.

Supplementary Table 2. Genomewide association results for the 114 SNPs with $p \leq 1 \times 10^{-4}$ in the discovery sample. Chromosome positions are based on NCBI genome assembly b36 (hg18) and alleles are given relative to the positive strand. Chr= chromosome; POS=position; FREQ= frequency of allele B; BETA= the unstandardised regression coefficient; SE = Standard error; P = p-value of the test of association under an additive model; HWE p = Hardy Weinburg p-value; Imputed: 1: imputed SNP; 0: genotyped SNP. Gene = If the SNP maps to within the genomic bounds of a gene annotated in both Ensembl and Entrz-gene databases, the gene name is given.

SNP	Chr	POS	N	Allele		FREQ	BETA	SE	P	HWE p	Imputed	Gene Name
				A	B							
rs12474600	2	179593131	2329	A	G	0.89	-0.239	0.047	4.57E-07	0.274	1	CCDC141
rs17362972	2	179617712	2329	C	T	0.11	0.238	0.047	5.43E-07	0.272	1	CCDC141
rs1326167	10	92091006	2329	C	T	0.77	0.174	0.035	9.59E-07	0.348	1	--
rs11897059	2	179648121	2329	C	T	0.89	-0.226	0.047	1.36E-06	0.329	0	--
rs17363057	2	179644417	2329	G	T	0.89	-0.212	0.046	5.22E-06	0.522	0	--
rs11885400	2	179644482	2329	G	T	0.11	0.212	0.047	5.61E-06	0.453	1	--
rs16866624	2	179645015	2329	C	G	0.89	-0.211	0.047	5.96E-06	0.452	0	--
rs17454527	2	179842263	2329	C	G	0.08	0.250	0.055	6.08E-06	0.662	0	--
rs11896002	2	179644132	2329	A	C	0.11	0.210	0.047	7.19E-06	0.456	1	--
rs2335212	19	14725032	2329	A	G	0.28	0.148	0.033	7.25E-06	0.534	1	EMR2
rs10403842	19	14718891	2327	A	G	0.28	0.148	0.033	7.27E-06	0.534	0	EMR2
rs2335213	19	14725228	2329	C	T	0.72	-0.148	0.033	7.36E-06	0.569	1	EMR2
rs11885227	2	179644092	2329	A	G	0.89	-0.210	0.047	7.38E-06	0.456	1	--
rs999967	19	14724522	2329	A	G	0.28	0.148	0.033	7.44E-06	0.568	1	EMR2
rs10497530	2	179632448	2329	A	G	0.11	0.208	0.047	8.31E-06	0.455	1	--
rs6709090	2	179631930	2329	C	T	0.11	0.208	0.047	8.45E-06	0.394	1	--
rs11893153	2	179642788	2327	A	G	0.11	0.205	0.046	9.48E-06	0.524	0	--
rs7596842	2	179643609	2329	C	T	0.89	-0.207	0.047	9.62E-06	0.522	1	--
rs7575047	2	179641155	2329	A	G	0.89	-0.206	0.047	9.85E-06	0.524	1	--
rs6433748	2	179641569	2329	A	G	0.89	-0.206	0.047	9.93E-06	0.524	1	--
rs6433749	2	179641616	2329	A	C	0.11	0.206	0.047	9.94E-06	0.524	1	--
rs7609150	2	179643550	2329	A	G	0.11	0.206	0.047	9.95E-06	0.522	1	--

rs6433750	2	179641763	2329	A	G	0.89	-0.206	0.047	9.97E-06	0.524	1	--
rs6433751	2	179641820	2329	C	T	0.89	-0.206	0.047	9.98E-06	0.524	1	--
rs6433752	2	179641887	2329	A	T	0.11	0.206	0.047	1.00E-05	0.524	1	--
rs7578521	2	179642060	2329	C	T	0.11	0.206	0.047	1.00E-05	0.524	1	--
rs7579263	2	179642519	2329	A	G	0.89	-0.206	0.047	1.02E-05	0.524	1	--
rs7582862	2	179643480	2329	C	G	0.89	-0.206	0.047	1.04E-05	0.523	1	--
rs7248991	19	14709467	2329	A	T	0.73	-0.146	0.033	1.05E-05	0.320	0	EMR2
rs10497531	2	179643369	2329	A	T	0.11	0.205	0.047	1.09E-05	0.523	1	--
rs11903793	2	179642828	2329	C	G	0.11	0.204	0.046	1.09E-05	0.524	0	--
rs4808387	19	14700160	2329	C	T	0.72	-0.146	0.033	1.16E-05	0.147	0	--
rs11186126	10	92077633	2329	A	G	0.72	0.140	0.032	1.17E-05	0.070	0	--
rs10497532	2	179643401	2328	A	G	0.11	0.203	0.046	1.23E-05	0.524	0	--
rs7099231	10	82379506	2329	G	T	0.04	0.323	0.074	1.39E-05	0.184	0	SH2D4B
rs12613938	2	29940969	2329	C	T	0.29	0.141	0.033	1.91E-05	0.063	1	ALK
rs1326162	10	92072230	2329	C	T	0.27	-0.137	0.032	2.20E-05	0.068	1	--
rs7924398	11	67849769	2329	C	T	0.06	-0.278	0.066	2.38E-05	0.574	1	LRP5
rs4797788	18	13580752	2326	A	G	0.48	-0.122	0.029	2.54E-05	0.361	0	C18orf1
rs16908342	8	139223221	2329	C	T	0.87	0.183	0.043	2.64E-05	0.350	1	FAM135B
rs12616992	2	29934322	2329	C	T	0.28	0.139	0.033	2.67E-05	0.113	1	ALK
rs6445445	3	65012525	2329	G	T	0.63	0.126	0.030	2.73E-05	0.287	0	--
rs2397150	6	53942338	2329	A	C	0.35	-0.130	0.031	2.91E-05	0.437	1	C6orf142
rs4338148	8	139228210	2329	C	T	0.86	0.176	0.042	3.04E-05	0.481	0	FAM135B
rs9367546	6	53939693	2329	C	T	0.65	0.129	0.031	3.12E-05	0.410	1	C6orf142
rs9357779	6	53939788	2329	A	G	0.65	0.129	0.031	3.23E-05	0.410	0	C6orf142
rs12620548	2	29943845	2329	C	T	0.70	-0.136	0.033	3.23E-05	0.038	0	ALK
rs8067574	17	76006986	2329	C	T	0.57	0.121	0.029	3.28E-05	0.138	0	--
rs2206574	6	53939420	2328	A	G	0.35	-0.128	0.031	3.39E-05	0.385	0	C6orf142
rs10165099	2	3381598	2329	C	T	0.41	0.124	0.030	3.43E-05	1.000	1	TTC15
rs17072468	3	65127206	2329	A	G	0.75	0.137	0.033	3.53E-05	0.077	0	--
rs742525	6	53931538	2329	C	T	0.35	-0.128	0.031	3.87E-05	0.465	1	C6orf142
rs7555255	1	215417006	2329	C	T	0.84	-0.162	0.039	3.88E-05	0.162	0	--
rs11826539	11	134185495	2326	A	G	0.97	-0.362	0.088	4.16E-05	0.262	0	--

rs6510896	19	6353648	2329	C	T	0.34	-0.126	0.031	4.22E-05	0.578	0	--
rs621527	18	64236042	2329	G	T	0.86	0.175	0.043	4.33E-05	0.932	1	--
rs2570548	10	79901113	2329	C	T	0.84	-0.161	0.039	4.33E-05	0.650	0	--
rs17006175	2	70795321	2329	C	G	0.22	0.142	0.035	4.33E-05	0.070	0	ADD2
rs9395894	6	53936350	2328	C	T	0.65	0.126	0.031	4.48E-05	0.410	0	C6orf142
rs4524263	3	1648354	2329	A	G	0.17	-0.157	0.039	4.60E-05	0.560	0	--
rs4627605	2	3390930	2329	A	G	0.41	0.121	0.030	4.73E-05	0.797	1	TTC15
rs16941631	18	21447583	2329	A	G	0.96	-0.321	0.079	4.77E-05	0.766	0	--
rs715338	15	55003159	2329	A	G	0.42	0.122	0.030	4.96E-05	0.348	1	TCF12
rs11826391	11	134181630	2329	C	T	0.97	-0.355	0.088	5.07E-05	0.265	0	--
rs16974087	15	84971701	2329	C	T	0.91	0.206	0.051	5.15E-05	0.521	0	AGBL1
rs3755369	2	70796472	2329	A	G	0.78	-0.141	0.035	5.23E-05	0.183	1	ADD2
rs10865541	2	3371082	2329	C	T	0.41	0.122	0.030	5.23E-05	0.932	1	TTC15
rs9913006	17	76005526	2329	G	T	0.43	-0.118	0.029	5.40E-05	0.099	1	--
rs11537721	6	55727976	2329	A	G	0.77	0.138	0.034	5.40E-05	0.452	1	BMP5
rs12528988	6	165734163	2329	C	T	0.93	-0.226	0.056	5.46E-05	0.450	1	PDE10A
rs6910005	6	55725321	2329	A	G	0.23	-0.138	0.034	5.51E-05	0.418	1	--
rs1661110	5	3214133	2329	C	T	0.89	0.194	0.048	5.60E-05	0.376	0	--
rs4971505	2	3411557	2329	G	T	0.39	0.122	0.030	5.62E-05	0.931	1	TTC15
rs228140	6	55736303	2329	C	T	0.23	-0.138	0.034	5.76E-05	0.486	0	BMP5
rs11002554	10	79873831	2329	A	G	0.85	-0.165	0.041	5.77E-05	0.744	1	--
rs467651	5	3210365	2329	A	G	0.90	0.198	0.049	5.85E-05	0.359	0	--
rs7579868	2	3410164	2329	G	T	0.39	0.121	0.030	5.85E-05	0.931	1	TTC15
rs4397868	11	133072195	2329	C	T	0.81	0.147	0.037	6.10E-05	0.642	0	--
rs11002556	10	79874158	2329	A	C	0.15	0.165	0.041	6.15E-05	0.744	1	--
rs468700	5	3210181	2329	A	T	0.10	-0.195	0.049	6.32E-05	0.910	1	--
rs10206964	2	3404955	2329	C	T	0.39	0.120	0.030	6.63E-05	0.931	1	TTC15
rs11002559	10	79877749	2328	A	C	0.85	-0.163	0.041	6.64E-05	0.870	0	--
rs11186132	10	92085667	2329	A	G	0.37	-0.120	0.030	6.66E-05	0.722	0	--
rs7558487	2	3404521	2329	C	G	0.39	0.120	0.030	6.67E-05	0.931	1	TTC15
rs4821670	22	36300812	2328	C	T	0.39	-0.120	0.030	6.68E-05	0.965	0	LGALS2
rs11631541	15	55148968	2329	C	G	0.51	-0.121	0.030	6.83E-05	0.772	1	TCF12

rs461401	5	3209508	2329	C	T	0.90	0.194	0.049	6.97E-05	0.910	1	--
rs13020433	2	3396584	2329	A	G	0.39	0.120	0.030	7.06E-05	0.965	1	TTC15
rs7804190	7	1960425	2329	A	C	0.65	0.121	0.030	7.12E-05	0.650	0	MAD1L1
rs6855056	4	5741184	2329	C	T	0.67	-0.123	0.031	7.13E-05	0.574	0	EVC2
rs666410	5	3213404	2325	A	G	0.90	0.192	0.048	7.17E-05	0.570	0	--
rs2361072	16	8366215	2329	A	G	0.05	-0.262	0.066	7.32E-05	0.684	1	--
rs11002558	10	79875419	2329	A	G	0.15	0.163	0.041	7.39E-05	0.744	1	--
rs12606330	18	55333319	2329	A	G	0.97	0.324	0.082	7.45E-05	0.736	0	CCBE1
rs30472	5	5035643	2329	A	G	0.68	-0.123	0.031	7.54E-05	0.214	1	--
rs3011217	1	44075853	2329	A	G	0.72	0.127	0.032	7.64E-05	0.151	0	ST3GAL3
rs3803455	15	54997589	2329	A	G	0.45	0.118	0.030	7.76E-05	0.900	1	--
rs11016422	10	130348359	2329	A	G	0.95	0.254	0.064	7.97E-05	0.538	0	--
rs11002555	10	79873906	2328	G	T	0.15	0.161	0.041	8.04E-05	0.743	0	--
rs467974	5	3213194	2329	C	G	0.90	0.191	0.048	8.06E-05	0.653	0	--
rs467873	5	3212357	2329	A	C	0.90	0.191	0.048	8.06E-05	0.571	0	--
rs1691133	5	3214367	2329	G	T	0.10	-0.192	0.049	8.06E-05	0.734	1	--
rs10824609	10	79876909	2329	A	G	0.85	-0.161	0.041	8.07E-05	0.742	1	--
rs1691132	5	3214324	2329	A	G	0.90	0.191	0.048	8.10E-05	0.571	0	--
rs13145438	4	13061363	2329	A	G	0.74	-0.139	0.035	8.44E-05	0.520	1	RAB28
rs6080100	20	15951406	2329	C	T	0.67	0.121	0.031	8.44E-05	0.512	0	MACROD2
rs17660394	5	2696974	2329	G	T	0.64	-0.123	0.031	8.86E-05	0.500	1	--
rs12753507	1	213446975	2327	A	G	0.39	0.123	0.031	8.91E-05	0.000	0	KCNK2
rs10889889	1	71022940	2329	A	C	0.38	0.120	0.031	9.08E-05	0.125	1	--
rs1409982	1	71040479	2329	C	T	0.39	0.120	0.031	9.33E-05	0.073	1	--
rs4971502	2	3393036	2329	A	G	0.61	-0.117	0.030	9.40E-05	0.794	0	TTC15
rs9379751	6	25400193	2329	A	G	0.48	-0.115	0.030	9.48E-05	0.618	1	LRRC16A
rs16192	7	24246951	2329	A	C	0.09	0.202	0.052	9.48E-05	0.142	1	--
rs2174146	11	13883488	2328	A	G	0.81	0.146	0.037	9.88E-05	0.736	0	--

Genotyping and quality control protocol

DNA extracted from buccal cheek swabs for the discovery sample of 4444 children was sent to Affymetrix, Santa Clara, California, USA. In total, 3,665 samples were successfully hybridized to AffymetrixGeneChip 6.0 SNP genotyping arrays

(http://www.affymetrix.com/support/technical/datasheets/genomewide_snp6_datasheet.pdf)

using experimental protocols recommended by the manufacturer (Affymetrix Inc., Santa Clara, CA). The raw image data from the arrays were normalized and pre-processed at the Wellcome Trust Sanger Institute, Hinxton, UK for genotyping as part of the Wellcome Trust Case Control Consortium 2 (WTCCC) (<https://www.wtccc.org.uk/ccc2/>) according to the manufacturer's guidelines. Genotypes for the Affymetrix arrays were called using CHIAMO (<http://www.stats.ox.ac.uk/marchini/software/gwas/chiamo.html>). Where there was a sufficient quantity of DNA, samples were also re-genotyped on a panel of 30 SNPs (including 26 autosomal SNPs present on the Affymetrix array, and 4 SNPs on the X chromosome to verify gender) using the Sequenom iPlex Gold assay (Sequenom Inc., San Diego, CA).

Quality control of the genotyping data was performed as part of the WTCCC2 (Barrett et al., 2009). We identified and removed samples whose genome-wide patterns of diversity differed from those of the collection at large, interpreting these differences as likely to be due to biases or artifacts. To obtain a set of putatively unrelated individuals we used a hidden Markov model (HMM) to infer identify by descent along the genome between pairs of individuals. Among pairs of closely related individuals, we excluded the member of the pair with the lowest call rate, iteratively repeating this procedure to obtain a set of individuals with pairwise identity by decent less than 5% (Barrett et al., 2009).

Outlying individuals were identified on the basis of call rate, heterozygosity, relatedness and ancestry using a Bayesian clustering approach (Barrett et al., 2009). Specifically, of the individuals genotyped, 377 samples were excluded because of low call rate or heterozygosity outliers, 9 due to unusual hybridization intensity, 59 due to atypical population ancestry, 83 due to sample duplication or relatedness to other sample members, and 13 due to gender mismatches (13). In addition, 54 samples were excluded because less than 90% of genotypes were called identically on the genome-wide array and Sequenom panel. The remaining samples were consistent with previous genotyping. In total, 513 samples were excluded by these quality control criteria. The remaining sample of 3,152 individuals included 1,446 males and 1,706 females.

A measure of information for the allele frequency at each of 932,533 called SNPs was calculated using SNPTEST version 2.1.1 (Marchini et al., 2007). Autosomal SNPs were excluded if this information measure was below 0.975, if the minor allele frequency was less than 1%, if greater than 2% of genotype data were missing, or if the Hardy Weinberg *p*-value was lower than 10^{-20} . Association between the SNP and the plate on which samples were genotyped was calculated; SNPs with a plate effect *p*-value less than 10^{-6} were also excluded. In addition, SNPs were manually filtered for call quality by visual inspection of the hybridization intensity plots using EVOKER software (<http://sourceforge.net/projects/evoker/>).

Statistical analysis

Imputation was carried out using the IMPUTE version 2 software (Howie, Donnelly & Marchini, 2009) on the genotype data after application of quality control procedures, using a two-stage approach with both a haploid reference panel and a diploid reference panel. For the haploid reference panel we used HapMap phase II and III SNP data on the 120 unrelated CEU trios. 5,175 WTCCC2 controls were genotyped on both Affymetrix 6.0 and Illumina Human1.2M-Duo arrays (Illumina Inc., La Jolla, CA), and these were used for the diploid reference panel. SNPs were retained for analysis if they were genotyped using the Affymetrix 6.0 array, if they were genotyped using the Illumina Human1.2M-Duo array and obtained an information score ≥ 0.90 , or if they were imputed and obtained an information score ≥ 0.98 . Using these criteria, 1,724,317 SNPs were retained for the association analyses.

We performed Principal Component Analysis on a subset of 105,556 autosomal SNPs remaining after applying our quality control criteria, after pruning to remove SNPs in high linkage disequilibrium ($r^2 > 0.2$) and excluding high linkage disequilibrium genomic regions so as to ensure that only genome-wide effects were detected (Fellay et al., 2007). Application of the Tracy-Widom test (Patterson, Price, & Reich, 2006) indicated that eight principal components were significant using a threshold of $p < 0.05$. Phenotype scores were normalized by transforming the ranked data to the quantiles of a standard normal distribution using the van der Waerden transformation (van der Waerden, 1952).

SOM References

UK IBD Genetics Consortium, Barrett. J. C., Lee, J. C., Lees, C. W., Prescott, N. J., Anderson, C. A....Strachan, D. P. (2009). Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nature Genetics*, 41, 1330-1334. doi: 10.1038/ng.483.

Fellay, J., Shianna, K. V., Ge, D., Colombo, S., Ledergerber, B., Weale, M... Goldstein, D. B. (2007). A whole-genome association study of major determinants for host control of HIV-1. *Science*, 317(5840), 944-947. doi: 10.1126/science.1143767.

Howie, B. N., Donnelly, P., & Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *Plos Genetics*, 5(6), e1000529.

Marcini, J., Howie, B., Myers, S., McVean, G., & Donnelly, P. (2007). A new multipoint method for genome-wide association studies via imputation of genotypes. *Nature Genetics*, 38, 906-913. doi:10.1038/ng2088.

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. [Research Support, N.I.H., External Research Support, Non-U.S. Gov't]. *Plos Genetics*, 2(12), e190.

van der Waerden, B.L. (1952). "Order tests for the two-sample problem and their power", *Indagationes Mathematicae*, 14, 453–458.