Detection and replication of epistasis influencing transcription in humans

Gibran Hemani[1,2,*], Konstantin Shakhbazov[1,2], Harm-Jan Westra[3], Tonu Esko[4,5,6], Anjali K Henders[7], Allan F McRae[1,2], Jian Yang[1], Greg Gibson[8], Nicholas G Martin[7], Andres Metspalu[4], Lude Franke[3], Grant W Montgomery[7,+], Peter M Visscher[1,2,+], and Joseph E Powell[1,2,+]

[1]Queensland Brain Institute, University of Queensland, Brisbane, QLD, Australia. [2]University of Queensland Diamantina Institute, University of Queensland, Princess Alexandra Hospital, Brisbane, Queensland, Australia. [3]Department of Genetics, University Medical Center Groningen, University of Groningen, Hanzeplein 1, Groningen, the Netherlands. [4]Estonian Genome Center, University of Tartu, Tartu, 51010, Estonia. [5]Medical and Population Genetics, Broad Institute, Cambridge, MA, 02142, US. [6]Divisions of Endocrinology, Children's Hospital, Boston, MA, 02115, US. [7]Queensland Institute of Medical Research, Brisbane, Queensland, Australia. [8]School of Biology and Centre for Integrative Genomics, Georgia Institute of Technology, Atlanta, Georgia United States of America. [+]These authors contributed equally. [*]Corresponding author: g.hemani@uq.edu.au

Abstract

Epistasis is the phenomenon whereby one polymorphism's effect on a trait depends on other polymorphisms present in the genome. The extent to which epistasis influences complex traits[1] and contributes to their variation[2,3] is a fundamental question in evolution and human genetics. Though often demonstrated in artificial gene manipulation studies in model organisms[4,5], and some examples have been reported in other species[6], few examples exist for epistasis amongst natural polymorphisms in human traits[7,8]. Its absence from empirical findings may simply be due to low incidence in the genetic control of complex traits[2,3], but an alternative view is that it has previously been too technically challenging to detect due to statistical and computational issues[9]. Here we show that, using advanced computation[10] and a gene expression study design, many instances of epistasis are found between common single nucleotide polymorphisms (SNPs). In a cohort of 846 individuals with 7339 gene expression levels measured in peripheral blood, we found 501 significant pairwise interactions between common SNPs influencing the expression of 238 genes ($p < 2.91 \times 10^{-16}$). Replication of these interactions in two independent data sets[11,12] showed both concordance of direction of epistatic effects ($p = 5.56 \times 10^{-31}$) and enrichment of interaction p-values, with 30 being significant at a conservative threshold of $p < 0.05/501$. Forty-four of the genetic interactions are located within 2Mb of regions of known physical chromosome interactions[13] ($p = 1.8 \times 10^{-10}$). Epistatic networks of three SNPs or more influence the expression levels of 129 genes, whereby one *cis*-acting SNP is modulated by several *trans*-acting SNPs. For example MBNL1 is influenced by an additive effect at rs13069559 which itself is masked by *trans*-SNPs on 14 different chromosomes, with nearly identical genotype-phenotype (GP) maps for each *cis*-*trans* interaction. This study presents the first evidence for multiple instances of segregating common polymorphisms interacting to influence human traits.

Main text

In the genetic analysis of complex traits it is usual for SNP effects to be estimated using an additive model where they are assumed to contribute independently and cumulatively to the mean of a trait. This framework has been successful in identifying thousands of associations[14]. But to date, though its contribution to phenotypic variance is frequently the subject of debate[1–3], there is little empirical exploration of the role that epistasis plays in the architecture of complex traits in humans[7,8]. Beyond the prism of human association studies there is evidence for epistasis, not only at the molecular scale from artificially induced mutations[4] but also at the evolutionary scale in fitness adaptation[15] and speciation[16].

Methods are now available to overcome the computational problems involved in searching for epistasis, but its detection still remains problematic due to reduced statistical power. For example, increased dependence on linkage disequilibrium (LD) between causal SNPs and observed SNPs[17,18], increased model

complexity in fitting interaction terms[19], and more extreme significance thresholds to account for increased multiple testing[9] all make it more difficult to detect epistasis in comparison to additive effects. Thus, with small genetic effect sizes, as is expected in most complex traits of interest[14], the power to detect epistasis diminishes rapidly. There are two simple ways to overcome this problem. One is by using extremely large sample sizes[20]; another is by analysing traits that are likely to have large effect sizes among common variants. Because our focus was to ascertain the extent to which instances of epistasis arises from natural genetic variation we designed a study around the latter approach and searched for epistatic genetic effects that influence gene expression levels. Transcription levels can be measured for thousands of genes and like most complex diseases, these expression traits are typically heritable[21]. But unlike complex diseases, genetic associations with gene expression commonly have very large effect sizes that explain large proportions of the genetic variance[22], making them good candidates to search for epistasis, should it exist.

In our discovery dataset (Brisbane Systems Genetics Study, BSGS[23]) of 846 individuals genotyped at 528,509 SNPs, we used a two stage approach to identify genetic interactions. First, we exhaustively test every pair of SNPs for pairwise effects against each of 7339 expression traits in peripheral blood (1.03×10^{15} statistical tests, family-wise error rate of 5% corresponding to a significance threshold of $p < 2.91 \times 10^{-16}$, Methods). Second, we filtered the SNP pairs from stage 1 on LD and genotype class counts, and tested the remaining pairwise effects for significant interaction terms and used a Bonferroni correction for multiple testing (estimated type 1 error rate $0.05 \leq \alpha \leq 0.14$, Methods, Supplementary Figure S1). Using this design we identified 501 putative genetic interactions influencing the expression levels of 238 genes (Supplementary Table S1). We used strict quality control measures to avoid statistical associations being driven by technical artifacts (Methods). However it remains possible that unexplained technical artifacts may have led to the significant discovery interactions. Of the 501 discovery interactions, 434 had available data and passed filtering (Methods) in two independent replication datasets, Fehrmann[12] and the Estonian Genomics Centre University of Tartu (EGCUT)[11], in which we saw convincing evidence for replication. We used the summary statistics from the replication datasets to perform a meta analysis to obtain an independent p-value for the putative interactions, and 30 were significant after applying a Bonferroni correction for multiple testing (5% significance threshold $p < 0.05/501$, Table 1). To quantify the similarity of GP maps between the independent datasets (Figure 1) we decomposed the genetic effects of each of the SNP pairs into orthogonal additive, dominance and epistatic effects ($A1$, $A2$, $D1$, $D2$, $A1 \times A2$, $A1 \times D2$, $D1 \times A2$, $D1 \times D2$) and tested for concordance of the sign of the most signicant effect (Supplementary Table S3, Methods). Sign concordance between the discovery and both replication datasets was observed in 22 out of the 30 significantly replicated interactions (expected value = 7.5 under the null hypothesis of no interactions, $p = 3.76 \times 10^{-8}$).

In addition, using the meta analysis from the replication samples only, we observed that 316 of the remaining 404 discovery SNP pairs had replication

interaction p-values more extreme than the 2.5% confidence interval of the quantile-quantile plot against the null hypothesis of no interactions where p-values are assumed to be uniformly distributed ($p << 1.0 \times 10^{-16}$, Figure 2 and Supplementary Figure S2). Concordance of the direction of the effect of the largest variance component was also highly significant ($p = 5.71 \times 10^{-31}$, Supplementary Table S3). The congruence of the epistatic networks in discovery and replication datasets is shown in Figure 3, demonstrating that these complex genetic patterns are common even across independent datasets. A further replication was attempted using the Centre for Health Discovery and Wellbeing (CHDWB) dataset[24], but only 20 of the SNP pairs passed filtering because the sample size was small ($n = 139$), and likely due to insufficient power we found no evidence for replication (Supplementary Figure S6). It should be noted that although it is a necessary step to establish the veracity of the interactions from the discovery set, replication of epistatic effects in independent samples is difficult in practice due to LD (Methods).

Though seldom the focus of association studies, SNPs with known main effects are often tested for $A \times A$ genetic interactions[9], but our analysis suggests this is unlikely to be the best strategy for its detection. The majority of our discovery interactions comprised of one SNP that was significantly associated with the gene expression level in the discovery dataset, and one SNP that had no previous association[22] (439 out of 501, Methods). Only nine interactions were between SNPs that both had known main effects while 64 were between SNPs that had no known main effects. Additionally, we observed that the largest epistatic variance component for the 501 interactions was equally divided amongst $A \times A$, $A \times D$, $D \times A$ and $D \times D$ at the discovery stage ($p = 0.22$ for departure from expectation). This is not surprising because these patterns of epistasis used for statistical decomposition are simply convenient orthogonal parameterisations of a two locus model, and are not intended to model biological function[25].

Of the discovery interactions, 26 were *cis*-*cis* acting (within 1Mb of the transcription start site, mean distance between SNPs was 0.53Mb), 462 were *cis*-*trans*-acting, and 13 were *trans*-*trans*-acting. We observed a wide range of significant GP maps (Figure 1) but the most common pattern of epistasis that we detected involved a *trans*-SNP masking the effect of an additive *cis*-SNP. For example, MBNL1 (involved in RNA modification and regulation of splicing[26]) has a *cis* effect at rs13069559 which in turn is controlled by 13 *trans*-SNPs and one *cis*-SNP that each exhibit a masking pattern, such that when the *trans*-SNP is homozygous for the masking allele the decreasing allele of the *cis*-SNP no longer has an effect (Supplementary Figure S10). Each of these interactions has evidence for replication in at least one dataset and six are significantly replicated at the Bonferroni level (Supplementary Figure S3). We see similar epistatic networks involving multiple (eight or more) *trans*-acting SNPs for other gene expresson levels too, for example TMEM149 (Supplementary Figure S11), NAPRT1 (Supplementary Figure S12), TRAPPC5 (Supplementary Figure S13), and CAST (Supplementary Figure S14). We observed that from pedigree analysis these five gene expression phenotypes had non-additive variance component

estimates within the 95th percentile of the 17,994 gene expression phenotypes that were analysed previously[22] (Supplementary Table S2, Methods).

In total the 501 interactions comprised 781 unique SNPs, which we analysed for functional enrichment (Methods). We tested the SNPs for cell-type specific overlap with transcriptionally active chromatin regions, tagged by histone-3-lysine-4,tri-methylation (H3K4me3) chromatin marks, in 34 cell types[27] (Supplementary Figure S5). There was significant enrichment for *cis*-acting SNPs in haematopoietic cell types only ($p < 1 \times 10^{-4}$ for the three tissues with the strongest enrichment after adjusting for multiple testing). However *trans*-acting SNPs did not show any tissue specific enrichment ($p > 0.1$ for all tissues). This difference between *cis* and *trans* SNPs suggests different roles in epistatic interactions where tissue specificity is provided by the *cis* SNPs. There is also enrichment for *cis*-SNPs to be localised in regions with regulatory genomic features as measured by chromatin states[28] (Supplementary Figure S4).

We also demonstrate physical organisation of interacting loci within the cell, suggesting a mechanism by which biological function can lead to epistatic genetic variance. It has been shown that different chromosomal regions spatially colocalise in the cell through chromatin interactions[13]. We cross-referenced our epistatic SNPs with a map of chromosome interacting regions ($n = 96,139$) in K562 blood cell lines[29] (Methods) and found that 44 epistatic interactions mapped to within 5Mb ($p < 1.8 \times 10^{-10}$), (Supplementary Figure S15). Interaction of distant loci may occur through physical proximity in transcriptional factories that organise across different chromosome regions and can regulate transcription of related genes[30].

Quantifying the importance of epistasis in complex traits in humans remains an open question. Here we are able to identify 238 gene expression traits with at least one significant interaction given our experiment-wide threshold, where the minimum estimated variance explained by the epistatic effects of any interaction was 2.1% of phenotypic variance. Taking results from our previously published eQTL[23] we calculated that 1848 of the 7339 gene expression levels analysed were influenced by additive effects where the estimated additive variance of a locus was 2.1% or greater. Thus, we can infer that the number of instances of large additive effects is significantly greater than the number of instances of large epistatic effects.

In terms of their contribution to complex traits a more important metric might be the proportion of the variance that the epistatic loci explain[2]. Taking all additive effects detected in Powell *et al* (2012) that have additive variance explaining 2.1% or greater of phenotypic variance, we calculated that the proportion of total phenotypic variance of all 7339 gene expression levels explained by additive effects alone was 2.16%. By contrast, the estimated epistatic variance from the interacting SNPs detected in this study on average explain a total of 0.22% of phenotypic variance, approximately ten times lower than the estimated additive variance. There are several caveats to this comparison which we discuss in the Methods.

Overall, we have demonstrated that it is possible to identify and replicate epistasis in complex traits amongst common human variants, despite the rela-

tive contribution of pairwise epistasis to phenotypic variation being small. The bioinformatic analysis of the significant epistatic loci suggests that there are a large number of possible mechanisms that can lead to non-additive genetic variation. Further research into such epistatic effects may provide a useful framework for understanding molecular mechanisms and complex trait variation in greater detail. With computational techniques and data now widely available the search for epistasis in larger datasets for traits of broader interest is warranted.

Methods Summary

We searched for pairwise epistasis exhaustively in the BSGS discovery dataset[23], which comprises 846 individuals who are genotyped at 528,509 autosomal SNPs. Each individual had gene expression levels measured in peripheral blood at 7,339 probes representing 6,158 RefSeq genes (significant expression in $\geq 90\%$ of individuals). SNP pairs were modelled for full genetic effects, including marginal additive and dominance at both SNPs plus four interaction terms. We used permutation analysis to calculate an experiment-wide significance threshold of $T_e = 2.91 \times 10^{-16}$ at the 5% family-wise error rate (FWER). All SNP pairs with LD $r^2 > 0.1$ and $D'^2 > 0.1$ were removed to minimise the possibility of haplotype effects. All SNP pairs were required to have at least five data points in all nine genotype classes. If multiple SNP pairs were present on the same chromosomes for a particular expression trait then only the sentinel SNP pair was retained. Finally, a nested test contrasting the full genetic model against the marginal additive and dominance model was performed for each remaining SNP pair. The 501 significant SNP pairs were carried forward for replication in two independent datasets that used the same expression assays for analysing transcription in peripheral blood Fehrmann[12], $n = 1240$; EGCUT[11], ($n = 891$). A meta analysis on the interaction p-values from each replication dataset was performed to provide an overall replication statistic for each putative interaction.

Acknowledgements

We are grateful to the volunteers for their generous participation in these studies. We thank Bill Hill, Chris Haley and Lars Ronnegard for helpful discussions and comments.

This work could not have been completed without access to high performance GPGPU compute clusters. We acknowledge iVEC for the use of advanced computing resources located at iVEC@UWA (www.ivec.org), and the Multimodal Australian ScienceS Imaging and Visualisation Environment (MASSIVE) (www.massive.org.au). We also thank Jake Carroll and Irek Porebski from the Queensland Brain Institute Information Technology Group for HPC support.

The University of Queensland group is supported by the Australian National Health and Medical Research Council (NHMRC) grants 389892, 496667, 613601, 1010374 and 1046880, the Australian Research Council (ARC) grant

(DE130100691), and by National Institutes of Health (NIH) grants GM057091 and GM099568.

The QIMR researchers acknowledge funding from the Australian National Health and Medical Research Council (grants 241944, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 496688 and 552485), the and the National Institutes of Health (grants AA07535, AA10248, AA014041, AA13320, AA13321, AA13326 and DA12854). We thank Anthony Caracella and Lisa Bowdler for technical assistance with the micro-array hybridisations.

The CHDWB study funding support from the Georgia Institute of Technology Research Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript

The Fehrmann study was supported by grants from the Celiac Disease Consortium (an innovative cluster approved by the Netherlands Genomics Initiative and partly funded by the Dutch Government (grant BSIK03009), the Netherlands Organization for Scientific Research (NWO-VICI grant 918.66.620, NWO-VENI grant 916.10.135 to L.F.), the Dutch Digestive Disease Foundation (MLDS WO11-30), and a Horizon Breakthrough grant from the Netherlands Genomics Initiative (grant 92519031 to L.F.). This project was supported by the Prinses Beatrix Fonds, VSB fonds, H. Kersten and M. Kersten (Kersten Foundation), The Netherlands ALS Foundation, and J.R. van Dijk and the Adessium Foundation. The research leading to these results has received funding from the European Communitys Health Seventh Framework Programme (FP7/2007-2013) under grant agreement 259867.

The EGCUT study received targeted financing from Estonian Government SF0180142s08, Center of Excellence in Genomics (EXCEGEN) and University of Tartu (SP1GVARENG). We acknowledge EGCUT technical personnel, especially Mr V. Soo and S. Smit. Data analyzes were carried out in part in the High Performance Computing Center of University of Tartu.

Author contributions

G.H., J.E.P., P.M.V., and G.W.M. conceived and designed the study. G.H., J.E.P., K.S., H-J.W., and J.Y. performed the analysis. T.E. and A.M. provided the EGCUT data. A.K.H., A.F.M., G.W.M., N.G.M., and J.E.P. provided the BSGS data. G.G. provided the CHDWB data. H-J.W. and L.F. provided the Fehrmann data. G.H. and J.E.P. wrote the manuscript with the participation of all authors.

Author information

The authors declare no financial competing interests.

Tables

Table 1: Epistatic interactions significant at the Bonferroni level in two replication sets

	Gene (chr.)	SNP 1 (chr.)	SNP 2 (chr.)	BSGS[2]	Fehrmann[3]	EGCUT[3]	Meta[4]
1	ADK (10)	rs2395095 (10)	rs10824092 (10)	6.69[1]	18.33[1]	21.21[1]	39.82[1]
2	ATP13A1 (19)	rs4284750 (19)	rs873870 (19)	5.30	12.18	3.25	14.23
3	C21ORF57 (21)	rs9978658 (21)	rs11701361 (21)	9.42	6.08	16.36	21.67
4	CSTB (21)	rs9979356 (21)	rs3761385 (21)	11.99	25.20	16.72	42.27
5	CTSC (11)	rs7930237 (11)	rs556895 (11)	7.16	18.76	15.06	33.53
6	FN3KRP (17)	rs898095 (17)	rs9892064 (17)	16.16	28.24	29.39	59.95
7	GAA (17)	rs11150847 (17)	rs12602462 (17)	13.91	19.98	12.99	32.60
8	HNRPH1 (5)	rs6894268 (5)	rs4700810 (5)	15.38	8.55	3.01	10.37
9	LAX1 (1)	rs1891432 (1)	rs10900520 (1)	19.16	18.60	11.22	29.24
10	MBNL1 (3)	rs16864367 (3)	rs13079208 (3)	13.49	16.25	24.74	41.56
11	MBNL1 (3)	rs7710738 (5)	rs13069559 (3)	7.92	2.55	7.89	9.28
12	MBNL1 (3)	rs2030926 (6)	rs13069559 (3)	7.10	0.91	5.80	5.53
13	MBNL1 (3)	rs2614467 (14)	rs13069559 (3)	5.74	4.13	2.22	5.30
14	MBNL1 (3)	rs218671 (17)	rs13069559 (3)	7.63	0.62	5.82	5.23
15	MBNL1 (3)	rs11981513 (7)	rs13069559 (3)	7.71	0.43	5.36	4.58
16	MBP (18)	rs8092433 (18)	rs4890876 (18)	5.40	7.06	21.91	28.73
17	NAPRT1 (8)	rs2123758 (8)	rs3889129 (8)	8.45	15.12	16.08	30.77
18	NCL (2)	rs7563453 (2)	rs4973397 (2)	7.31	7.51	6.33	12.70
19	PRMT2 (21)	rs2839372 (21)	rs11701058 (21)	4.81	0.69	4.47	4.06
20	RPL13 (16)	rs352935 (16)	rs2965817 (16)	4.98	3.79	14.41	17.24
21	SNORD14A (11)	rs2634462 (11)	rs6486334 (11)	7.31	13.11	10.96	23.22
22	TMEM149 (19)	rs807491 (19)	rs7254601 (19)	12.16	81.55	45.78	145.78
23	TMEM149 (19)	rs8106959 (19)	rs6926382 (6)	5.80	3.06	8.80	10.72
24	TMEM149 (19)	rs8106959 (19)	rs914940 (1)	6.22	3.36	6.96	9.20
25	TMEM149 (19)	rs8106959 (19)	rs2351458 (4)	7.30	0.04	9.61	8.00
26	TMEM149 (19)	rs8106959 (19)	rs6718480 (2)	8.55	3.31	5.15	7.36
27	TMEM149 (19)	rs8106959 (19)	rs1843357 (8)	6.21	3.72	3.33	6.00
28	TMEM149 (19)	rs8106959 (19)	rs9509428 (13)	9.44	0.10	5.75	4.47
29	TRA2A (7)	rs7776572 (7)	rs11770192 (7)	8.23	3.19	1.89	4.09
30	VASP (19)	rs1264226 (19)	rs2276470 (19)	5.09	0.94	5.14	4.95

[1] $-\log_{10} p$-values for 4 *d.f.* interaction tests

[2] Discovery dataset

[3] Independent replication dataset

[4] Meta analysis of interaction terms between replication datasets only

Figures

Figure 1: **Replication of GP maps in two independent populations** The GP maps for each epistatic interaction that is significant at the Bonferroni level in both replication datasets are shown. Each GP map consists of nine tiles where each tile represents the expression level for that two-locus genotype class. Phenotypes are for gene transcript levels (dark coloured tiles = high expression, light coloured tiles = low expression). Columns of GP maps are for each independent dataset. Rows of GP maps are for each of 30 significantly replicated interactions at the Bonferroni level, corresponding to the rows in Table 1. There is a clear trend of the GP maps replicating across all three datasets.

Figure 2: **Q-Q plots of interaction p-values from replication datasets** The top panel shows all 434 discovery SNPs that were tested for interactions. Observed p-values (y-axis, $-\log_{10}$ scale) are plotted against the expected p-values (x-axis, $-\log_{10}$ scale). The multiple testing correction threshold for significance following Bonferroni correction is denoted by a dotted line. The bottom panel shows the same data as the top panel but excluding the 30 interactions that were significant at the Bonferroni level in the replication datasets. The shaded grey area represents the 5% confidence interval for the expected distribution of p-values. Dark blue points represent p-values that exceed the confidence interval, light blue are within the confidence interval.

Figure 3: **Discovery and replication of epistatic networks** All 434 putative genetic interactions (edges) with data common to discovery and replication sets is shown, where black nodes represent SNPs and red nodes represent traits (gene expression probes). Three hundred and forty-five interactions had p-values exceeding the 2.5% confidence interval following meta analysis of the replication data The remaining 89 interactions that did not replicate are depicted in grey. It is evident that a large proportion of the complex networks identified in the discovery set also exist in independent populations. An interactive version of this graph can be found here: `http://kn3in.github.io/detecting_epi/`

References

1. Carlborg, O. & Haley, C. S. Epistasis: too often neglected in complex trait studies? *Nature Reviews Genetics* **5,** 618–25. ISSN: 14710056 (Aug. 2004).
2. Hill, W. G., Goddard, M. E. & Visscher, P. M. Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits. *PLoS Genetics* **4.** doi:10.1371/journal.pgen.1000008 (Feb. 2008).
3. Crow, J. F. On epistasis: why it is unimportant in polygenic directional selection. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **365,** 1241–4. ISSN: 1471-2970 (Apr. 2010).
4. Costanzo, M. *et al.* The genetic landscape of a cell. *Science (New York, N.Y.)* **327,** 425–31. ISSN: 1095-9203 (Jan. 2010).
5. Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. o. & Kruglyak, L. Finding the sources of missing heritability in a yeast cross. *Nature,* 1–6. ISSN: 0028-0836 (Feb. 2013).
6. Carlborg, O., Jacobsson, L., Ahgren, P., Siegel, P. & Andersson, L. Epistasis and the release of genetic variation during long-term selection. *Nature Genetics* **38,** 418–420. ISSN: 1061-4036 (Apr. 2006).
7. Strange, A. *et al.* A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nature Genetics* **42,** 985–90. ISSN: 1546-1718 (Nov. 2010).
8. Evans, D. M. *et al.* Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nature Genetics* **43.** ISSN: 1061-4036. doi:10.1038/ng.873 (July 2011).
9. Cordell, H. J. Detecting gene-gene interactions that underlie human diseases. *Nature Reviews Genetics* **10,** 392–404. ISSN: 1471-0064 (June 2009).
10. Hemani, G., Theocharidis, A., Wei, W. & Haley, C. EpiGPU: exhaustive pairwise epistasis scans parallelized on consumer level graphics cards. *Bioinformatics (Oxford, England)* **27,** 1462–5. ISSN: 1367-4811 (June 2011).
11. Metspalu, A. The Estonian Genome Project. *Drug Development Research* **62,** 97–101. ISSN: 0272-4391 (June 2004).
12. Fehrmann, R. S. N. *et al.* Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. *PLoS genetics* **7,** e1002197. ISSN: 1553-7404 (Aug. 2011).
13. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (New York, N.Y.)* **326,** 289–93. ISSN: 1095-9203 (Oct. 2009).
14. Visscher, P. M., Brown, M. a., McCarthy, M. I. & Yang, J. Five years of GWAS discovery. *American journal of human genetics* **90,** 7–24. ISSN: 1537-6605 (Jan. 2012).

15. Weinreich, D. M., Delaney, N. F., Depristo, M. a. & Hartl, D. L. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science (New York, N.Y.)* **312,** 111–4. ISSN: 1095-9203 (Apr. 2006).

16. Breen, M. S., Kemena, C., Vlasov, P. K., Notredame, C. & Kondrashov, F. a. Epistasis as the primary factor in molecular evolution. *Nature* **490,** 535–538. ISSN: 1476-4687 (Oct. 2012).

17. Weir, B. S. Linkage disequilibrium and association mapping. *Annual review of genomics and human genetics* **9,** 129–42. ISSN: 1527-8204 (Jan. 2008).

18. Hemani, G., Knott, S. & Haley, C. An Evolutionary Perspective on Epistasis and the Missing Heritability. *PLoS Genetics* **9** (ed Mackay, T. F. C.) e1003295. ISSN: 1553-7404 (Feb. 2013).

19. Marchini, J., Donnelly, P. & Cardon, L. R. Genome-wide strategies for detecting multiple loci that influence complex diseases. *Nature Genetics* **37,** 413–417. ISSN: 1061-4036 (Apr. 2005).

20. Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467,** 832–8. ISSN: 1476-4687 (Oct. 2010).

21. Schadt, E. *et al.* Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422,** 297–302 (2003).

22. Powell, J. E. *et al.* Congruence of Additive and Non-Additive Effects on Gene Expression Estimated from Pedigree and SNP Data. *PLoS Genetics* **9** (ed Spector, T. D.) e1003502. ISSN: 1553-7404 (May 2013).

23. Powell, J. E. *et al.* The Brisbane Systems Genetics Study: genetical genomics meets complex trait genetics. *PloS one* **7,** e35430. ISSN: 1932-6203 (Jan. 2012).

24. Preininger, M. *et al.* Blood-informative transcripts define nine common axes of peripheral blood gene expression. *PLoS genetics* **9,** e1003362. ISSN: 1553-7404 (Mar. 2013).

25. Cockerham, C. C. An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics* **39,** 859–882 (Nov. 1954).

26. Ho, T. H. *et al.* Muscleblind proteins regulate alternative splicing. *The EMBO journal* **23,** 3103–12. ISSN: 0261-4189 (Aug. 2004).

27. Trynka, G. *et al.* Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nature genetics* **45,** 124–30. ISSN: 1546-1718 (Feb. 2013).

28. Hoffman, M., Buske, O., Wang, J. & Weng, Z. Unsupervised pattern discovery in human chromatin structure through genomic segmentation. *Nature Methods* **9,** 473–476 (2012).

29. Lan, X. *et al.* Integration of Hi-C and ChIP-seq data reveals distinct types of chromatin linkages. *Nucleic acids research* **40,** 7690–704. ISSN: 1362-4962 (Sept. 2012).

30. Rieder, D., Trajanoski, Z. & McNally, J. G. Transcription factories. *Frontiers in genetics* **3,** 221. ISSN: 1664-8021 (Jan. 2012).

Online methods

1 Discovery data

1.1 Data description

The Brisbane Systems Genetics Study (BSGS) comprises 846 individuals of European descent from 274 independent families[23]. DNA samples from each individual were genotyped on the Illumina 610-Quad Beadchip by the Scientific Services Division at deCODE Genetics Iceland. Full details of genotyping procedures are given in Medland et al.[31] Standard quality control (QC) filters were applied and the remaining 528,509 autosomal SNPs were carried forward for further analysis.

Gene expression profiles were generated from peripheral blood collected with PAXgene TM tubes (QIAGEN, Valencia, CA) using Illumina HT12-v4.0 bead arrays. The Illumina HT-12 v4.0 chip contains 47,323 probes, although some probes are not assigned to RefSeq genes. We removed any probes that did not match the following criteria: contained a SNP within the probe sequence with MAF > 0.05 within 1000 genomes data; did not map to a listed RefSeq gene; were not significantly expressed (based on a detection p-value < 0.05) in at least 90% of samples. After this stringent QC 7339 probes remained for 2D-eQTL mapping. These data are accessible through GEO Series accession number GSE53195.

1.2 Normalisation

Gene expression profiles were normalised and adjusted for batch and polygenic effects. Profiles were first adjusted for raw background expression in each sample. Expression levels were then adjusted using quantile and $\log_2$ transformation to standardise distributions between samples. Batch and polygenic effects were adjusted using the linear model

$$y = \mu + \beta_1 c + \beta_2 p + \beta_3 s + \beta_4 a + g + e \tag{1}$$

where μ is the population mean expression levels, c, p, s and a are vectors of chip, chip position, sex and generation respectively, fitted as fixed effects; and g is a random additive polygenic effect with a variance covariance matrix

$$G_{jk} = \begin{cases} \sigma_a^2 & j = k \\ 2\phi_{jk}\sigma_a^2 & j \neq k \end{cases} \tag{2}$$

The parameter σ_a^2 is the variance component for additive background genetic. Here, we are using family based pedigree information rather than SNP based IBD to account for relationships between individuals and so ϕ_{jk} is the kinship coefficient between individuals j and k. The residual, e, from equation 1 is assumed to follow a multivariate normal distribution with a mean of zero.

Residuals were normalised by rank transformation and used as the adjusted phenotype for the pairwise epistasis scan to remove any skewness and avoid results being driven by outliers. The GenABEL package for R was used to perform the normalisation[32].

2 Exhaustive 2D-eQTL analysis

2.1 Two stage search

We used epiGPU[10] software to perform an exhaustive scan for pairwise interactions, such that each SNP is tested against all other SNPs for statistical association with the expression values for each of the 7339 probes. This uses the massively parallel computational architecture of graphical processing units (GPUs) to speed up the exhaustive search. For each SNP pair there are 9 possible genotype classes. We treat each genotype class as a fixed effect and fit an 8 *d.f.* F-test to test the following hypotheses:

$$H_0 : \sum_{i=1}^{3}\sum_{j=1}^{3}(\bar{x}_{ij} - \mu)^2 = 0; \tag{3}$$

$$H_1 : \sum_{i=1}^{3}\sum_{j=1}^{3}(\bar{x}_{ij} - \mu)^2 > 0; \tag{4}$$

where μ is the mean expression level and x_{ij} is the pairwise genotype class mean for genotype i at SNP 1 and genotype j at SNP 2. This type of test does not parameterize for specific types of epistasis, rather it tests for the joint genetic effects at two loci. This has been demonstrated to be statistically more efficient when searching for a wide range of epistatic patterns, although will also include any marginal effects of SNPs which must be dealt with post-hoc[18].

2.1.1 Stage 1

The complete exhaustive scan for 7339 probes comprises 1.03×10^{15} F-tests. We used permutation analysis to estimate an appropriate significance threshold for the study. To do this we performed a further 1600 exhaustive 2D scans on permuted phenotypes to generate a null distribution of the extreme p-values expected to be obtained from this number of multiple tests given the correlation structure between the SNPs. We took the most extreme p-value from each of the 1600 scans and set the 5% FWER to be the 95% most extreme of these p-values, $T_* = 2.13 \times 10^{-12}$. The effective number of tests in one 2D scan being performed is therefore $N_* = 0.05/T_* \approx 2.33 \times 10^{10}$. To correct for the testing of multiple traits we established an experiment wide threshold of $T_e = 0.05/(N_* \times 7339) = 2.91 \times 10^{-16}$. This is likely to be conservative as it assumes independence between probes.

Filtering We used two approaches to filter SNPs from stage 1 to be tested for significant interaction effects in stage 2.

Filter 1 After keeping SNP pairs that surpassed the 2.91×10^{-16} threshold in stage 1 only SNP pairs with at least 5 data points in all 9 genotype classes were kept. We then calculated the LD between interacting SNPs (amongst unrelated individuals within the discovery sample and also from 1000 genomes data) and removed any pairs with $r^2 > 0.1$ or $D'^2 > 0.1$ to avoid the inclusion of haplotype effects and to increase the accuracy of genetic variance decomposition. If multiple SNP pairs were present on the same chromosomes for a particular expression trait then only the sentinel SNP pair was retained, *i.e.* if a probe had multiple SNP pairs that were on chromosomes one and two then only the SNP pair with the most significant p-value was retained. At this stage 6404 filtered SNP pairs remained.

Filter 2 We also performed a second filtering screen applied to the list of SNP pairs from stage 1 that was identical to filter 1 but an additional step was included where any SNPs that had previously been shown to have a significant additive or dominant effect ($p < 1.29 \times 10^{-11}$) were removed[22], creating a second set of 4751 unique filtered SNP pairs.

2.1.2 Stage 2

To ensure that interacting SNPs were driven by epistasis and not marginal effects we performed a nested ANOVA on each pair in the filtered set to test if the interaction terms were significant. We did this by contrasting the full genetic model (8 *d.f.*) against the reduced marginal effects model which included the additive and dominance terms at both SNPs (4 *d.f.*). Thus, a 4 *d.f.* F-test was performed on the residual genetic variation, representing the contribution of epistatic variance. Significance of epistasis was determined using a Bonferroni threshold of $0.05/(6404 + 4751) = 4.48 \times 10^{-6}$. This resulted in 406 and 95 SNP pairs with significant interaction terms from filters 1 and 2, respectively.

2.2 Type 1 error rate

Using a Bonferroni correction of 0.05 in the second stage of the two stage discovery scan implies a type 1 error rate of $\alpha = 0.05$. However, this could be underestimated because the number tests performed in the second stage depends on the number of tests in the first stage, and this depends on statistical power and model choice. We performed simulations to estimate the type 1 error rate of this study design.

We assumed a null model where there was one true additive effect and 7 other terms with no effect. To simulate a test statistic we simulated 8 z-scores, $z_1 \sim N(\sqrt{NCP}, 1)$ and $z_{2..8} \sim N(0, 1)$. Thus $z_{full} = \sum_{i=1}^{8} z_i \sim \chi^2_8$ (representing the 8 d.f. test) and $z_{int} = \sum_{i=5}^{8} z_i \sim \chi^2_4$ (representing the 4 d.f. test where the null hypothesis of no epistasis is true). For a particular value of NCP we simulated

100,000 z values, and calculated the p_{full}-value for the z_{full} test statistic. The n_{int} test statistics with $p_{full} < 2.31 \times 10^{-16}$ were kept for the second stage, where the type 1 error rate of stage 2 was calculated as the proportion of $p_{int} < 0.05/n_{int}$. The power at stage 1 was calculated as $n_{int}/100,000$. This procedure was performed for a range of NCP parameters that represented power ranging from ~ 0 to ~ 1.

2.3 Population stratification

We ruled out population stratification as a possible cause of inflated test statistics. To test for cryptic relatedness driving the interaction terms we tested for increased LD among the SNPs[33]. We calculated the mean of the off-diagonal elements of the correlation matrix of all unique SNPs from the 501 interactions (731 SNPs) using only unrelated individuals, $\bar{r^2} = 0.0039$. This is not significantly different from the null hypothesis of zero (sampling error $= 1/n_{\text{unrelated}} = 0.0039$).

2.4 Probe mapping

To avoid possibility that epistatic signals might arise due to expression probes hybridising in multiple locations we verified that probe sequences for genes with significant interactions mapped to only a single location. As an initial verification we performed a BLAST search of the full probe sequence against 1000 genomes phase 1 version 3 human genome reference and ensured that only one genomic location aligned significantly ($p < 0.05$). As a second step, to mitigate the possibility of weak hybridisation elsewhere in the genome we divided the probe sequence into three sections (1-25bp, 13-37bp, 26-50bp) and performed a BLAST search of these probe sequence fragments. No probe sequemces or probe sequence fragments mapped to positions other than the single expected genomic target ($p < 0.05$).

3 Replication

3.1 Data description

We attempted replication of the 501 significant interactions from the discovery set using three independent cohorts; Fehrmann, EGCUT, and CHDWB. It was required that LD $r^2 < 0.1$ and $D'^2 < 0.1$ between interacting SNPs (as measured in the replication sample directly), and all nine genotype classes had at least 5 individuals present in order to proceed with statistical testing for replication in both datasets. We also excluded any putative SNPs that had discordant allele frequencies in any of the datasets. Details of the cohorts are as follows.

Fehrmann: $n = 1240$ The Fehrmann dataset[12] consists of peripheral blood samples of 1240 unrelated individuals from the United Kingdom and the Netherlands. Some of these individuals are patients, while others are healthy controls.

Individuals were genotyped using the Illumina HumanHap300, Illumina HumanHap370CNV, and Illumina 610 Quad platforms. RNA levels were quantified using the Illumina HT-12 V3.0 platform. These data are accessible through GEO Series accession numbers GSE20332 and GSE20142.

EGCUT: $n = 891$ The Estonian Genome Center of the University of Tartu (EGCUT) study[11] consists of peripheral blood samples of 891 unrelated individuals from Estonia. They were genotyped using the Illumina HumanHap370CNV platform. RNA levels were quantified using the Illumina HT-12 V3.0 platform. These data are accessible through GEO Series accession number GSE48348.

CDHWB: $n = 139$ The Center for Health Discovery and Well Being (CDHWB) Study[24] is a population based cohort consisting of 139 individuals of European descent collected in Atlanta USA. Gene expression profiles were generated with Illumina HT-12 V3.0 arrays from peripheral blood collected from Tempus tubes that preserve RNA. Whole genome genotypes were measured using Illumina OmniQuad arrays. Due to the small sample size, most SNP pairs did not pass filtering in this dataset (20 SNP pairs remained) and so we have excluded it from the rest of the analysis.

3.2 Meta Analysis

The 4 *d.f.* interaction p-values for each independent replication dataset were calculated using the same statistical test as was performed in the discovery dataset. We then took the interaction p-values from EGCUT and Fehrmann and calculated a joint p-value using Fisher's method of combining p-values for a meta analysis as $-2\ln p_1 - 2\ln p_2 \sim \chi^2_{4d.f.}$. As in the discovery analysis, all gene expression levels were normalised using rank transformation to avoid skew or outliers in the distribution[34].

3.3 Concordance of direction of effects

We used four methods to calculate the concordance of the direction of effects between the discovery and replication datasets.

Test 1 Is the most significant epistatic effect in the discovery set in the same direction as the same epistatic effect in the replication sets? We decomposed the genetic variance into 8 orthogonal effects, four of which are epistatic ($A \times A$, $A \times D$, $D \times A$, $D \times D$). The sign of the epistatic effect that had the largest variance in the discovery was recorded, and then was compared to the same epistatic effect in the two replication datasets (regardless of whether or not the same epistatic effect was the largest in the replication datasets). The probability of the sign being the same in one dataset is 1/2. The probability of the sign being the same in two is 1/4.

Test 2 Is the most significant epistatic effect in the discovery the same as the largest epistatic effect in the replication set with the sign being concordant. As in Test 1, but this time we required that the largest effect was the same in the discovery and the replication, and that they had the same sign (*e.g.* if the largest effect in the discovery is $A \times A$, with a positive effect, then concordance is achieved if the same is true in the replication). The probability of one replication dataset being concordant by chance is 1/8, and concordance in both is 1/64.

Test 3 Do the epistatic effects that are significant at nominal $p < 0.05$ in the discovery have the same direction of effect as in the replication? Here we count all the epistatic variance components in the discovery that have $p < 0.05$ (1133 amongst the 434 discovery SNP pairs, *i.e.* each SNP pair has at least 1 and at most 4 significant epistatic variance components). Then we compare the direction of the effect in the replication dataset. The probability of the sign being the same in one dataset for any one significant effect is 1/2. The probability of the sign being the same in two is 1/4.

Test 4 If we count how many of the 4 epistatic effects are concordant between the discovery and replication data for each interaction then is this significant from what we expect by chance? There can be either 0, 1, 2, 3 or 4 concordant signs at each interaction, each with expectation of $p = 1/16, 4/16, 6/16, 4/16, 1/16$ under the null, respectively. Observed counts are multinomially distributed, and we tested if the observed proportions were statistically different from the expected proportions using an approximation of the multinomial test[35].

The probability of observing the number of concordant signs in tests 1-3 is calculated using a binomial test. All variance decompositions were calculated using the NOIA method[36].

4 Effects of LD on detection and replication

The power to detect genetic effects, when the observed markers are in LD with the causal variants, is proportional to r^x. For additive effects $x = 2$, but for non-additive effects x is larger, *i.e.* $x = 4$ for dominance or $A \times A$, $x = 6$ for $A \times D$ or $D \times A$, and $x = 8$ for $D \times D$. Many biologically realistic GP maps may be comprised of all 8 variance components[18].

This is important for both detection and for replication of epistasis. For detection, if the epistatic effect includes the $D \times D$ term then if the two causal variants are tagged by observed markers that are each in LD $r = 0.9$, then if the true variance is V_t then the observed variance V_o at the markers will be $0.9^8 V_t = 0.43 V_t$. Therefore, it is important to consider the sampling variation of $\hat{r}^x$ in a sample given some true population value of r.

4.1 Simulation 1

For some values of fixed population parameters, p_1 (minor allele frequency at observed marker), q_1 (minor allele frequency at causal variant), and r (LD between marker and causal variant), the expected haplotype frequencies are

$$h_{11} = r\sqrt{p_1 q_1 p_2 q_2} + p_1 q_1 \tag{5}$$
$$h_{12} = p_1 q_2 - r\sqrt{p_1 q_1 p_2 q_2} \tag{6}$$
$$h_{21} = p_2 q_1 - r\sqrt{p_1 q_1 p_2 q_2} \tag{7}$$
$$h_{22} = r\sqrt{p_1 q_1 p_2 q_2} + p_2 q_2 \tag{8}$$

where $p_2 = 1 - p_1$ and $q_2 = 1 - q_1$. For a range of population parameters we randomly sampled $2n$ haplotypes where the expected haplotype frequencies were $h_{11}, h_{12}, h_{21}, h_{22}$. From the sample haplotype frequencies we then calculated sample estimates of $\hat{r}$ where

$$\hat{r} = \frac{\hat{h}_{11} - \hat{p}_1 \hat{q}_1}{\sqrt{\hat{p}_1 \hat{q}_1 \hat{p}_2 \hat{q}_2}} \tag{9}$$

For each value of combination of the parameters p_1, q_1, r, n 1000 simulations were performed and the sampling mean and sampling standard deviation of $\hat{r}, \hat{r}^2, \hat{r}^4, \hat{r}^6, \hat{r}^8$ were recorded. It was observed that sampling variance increases for increasing x in $\hat{r}^x$.

4.2 Simulation 2

We assume that the discovery SNP pairs are ascertained (from a very large number of tests) have high $\hat{r}$ between observed SNPs and causal variants because otherwise power of detection would be low. We can hypothesis that the distribution of $\hat{r}$ in this ascertained sample will be a mixture of r that is high and r that is lower but with ascertained higher values from sampling. Therefore, we would expect those with truly high r to have a higher replication rate in independent datasets, and those with ascertained high $\hat{r}$ to have lower replication because resampling is unlikely to result in the same extreme ascertainment. To obtain empirical estimates of $\hat{r}$ in discovery and replication datasets we conducted the following simulation.

1. Using 1000 genomes data (phase 1, version 3, 379 European samples) we selected the 528,509 "markers" used in the original discovery analysis, plus 100,000 randomly chosen "causal variants" (CVs) with minor allele frequence > 0.05.

2. The 379 individuals were split into discovery (190) and replication (189) sets.

3. For each CV the marker with the maximum $\hat{r}^2_D$ from the marker panel was recorded in the discovery set. This marker was known as the "discovery marker" (DM).

4. The $\hat{r}^2_R$ for each CV/DM pair was then calculated in the replication set where the discovery LD was ascertained to be high, such that $\hat{r}^2_D > 0.9$.

We observed that there was an average decrease in $\hat{r}^x_R$ relative to $\hat{r}^x_D$, and that this decrease was larger with increasing x. We observed that $(\hat{r}^2_R - \hat{r}^2_D)/\hat{r}^2_D = 0.029$ whereas $(\hat{r}^8_R - \hat{r}^8_D)/\hat{r}^8_D = 0.092$. The average drop in in replication $\hat{r}^8$ was 3 times higher than the drop in $\hat{r}^2$.

4.3 Interpretation

Simulation 1 shows that sampling variance of r^x increases as x increases. Detection of epistatis is highly dependent upon high $\hat{r}$. Amongst the discovery SNPs there will be a mixture of interactions where observed SNPs are either in true high LD with causal variants, or will have highly inflated sample $\hat{r}^x$ compared to the population r^x. Simulation 2 shows that as x gets larger, the average decrease in $\hat{r}^x$ between discovery and replication becomes larger, likely to be a result of ascertained high $\hat{r}$ in the discovery and increased sampling variance with increasing x in the replication. These results demonstrate that if all else is equal, the impact of sampling variance of r alone will reduce the replication rate of epistatic effects compared to additive effects.

5 Additive and non-additive variance estimation

5.1 Fixed effects

To compare the relative contribution to the phenotypic variance of gene expression levels between additive and epistatic effects we are constrained by the problem that non-additive variance components for a phenotype cannot be calculated directly. Here, we only have SNP pairs that exceed a threshold of $p < 2.91 \times 10^{-16} = T_e$. A strong conclusion cannot be made about the genome-wide variance contribution, but we can compare the variance explained by SNP effects at this threshold for additive scans and epistatic scans.

In Powell *et al* 2012[23] an expression quantitative trait locus (eQTL) study was performed searching for additive effects in the same BSGS dataset as was used for the discovery here. Using the threshold T_e for the additive eQTL study, 453 of the 7339 probes analysed here had at least one significant additive effect. Assuming that the phenotypic variance for each of the probes is normalised to 1, the total phenotypic variance of all 7339 explained by the significant additive effects was 1.73%.

Following the same procedure, at the threshold T_e there were 238 gene expression probes with at least one significant pairwise epistatic interaction out of the 7339 tested. In total the proportion of the phenotypic variance explained by the epistatic effects at these SNP pairs was 0.25%.

5.2 Limitations of this type of comparison

Though it is useful to compare the relative variances of epistatic and additive effects, it must be stressed that our results here are approximations that are very limited by the study design. We estimate that additive effects explain approximately 10 times more variance than epistatic effects, but this could be an overestimate or an underestimate due to a number of different caveats. Firstly, the ratio of additive to epistatic variance may differ at different minimum variance thresholds, and our estimate is determined by the threshold used. Secondly, the power of a 1 *d.f.* test exceeds that of an 8 *d.f.* test. Thirdly, the non-additive variance at causal variants is expected to be underestimated by observed SNPs in comparison to estimates for additive variance. And forthly, the extent of winner's curse in estimation of effect sizes may differ between the two studies.

5.3 Pedigree estimates

The gene expression levels for MBNL1, TMEM149, NAPRT1, TRAPPC5 and CAST are influenced by large *cis-trans* epistatic networks (eight interactions or more). Though it is not possible to orthogonally estimate the non-additive genetic variance for non-clonal populations, an approximation of a component of non-additive variance can be estimated using pedigree information. The BSGS data is comprised of some related individuals and standard quantitative genetic analysis was used to calculate the additive and dominance variance components for each gene expression phenotype in Powell *et al* 2013[22]. The dominance effect is likely to capture additive $\times$ additive genetic variance plus some fraction of other epistatic variance components. We found that the aforementioned genes had dominance variance component estimates within the top 5% of all 17,994 gene expression probes that were analysed in Powell *et al* 2013.

6 Functional enrichment analysis

6.1 Tissue specific transcriptionally active regions

We employed a recently published method (`http://www.broadinstitute.org/mpg/epigwas/`)[27] that tests for cell-type-specific enrichment of active chromatin, measured through H3K4me3 chromatin marks[37] in regions surrounding the 731 SNPs that comprise the 501 discovery interactions. The exact method used to perform this analysis has been described previously[38]. Briefly, we tested the hypothesis that the 731 SNPs were more likely to be in transcriptionally active regions (as measured by chromatin marks) than a random set of SNPs selected from the same SNP chip. This hypothesis was tested for 34 cell types across four broad tissue types (haematopoietic, gastrointenstinal, musculoskeletal and endocrine, and brain).

6.2 Chromosome interactions

It has been shown[13] that different regions on different chromosomes or within chromosomes spatially colocalise within the cell. We shall refer to the colocalisation of two chromosome regions as a chromosome interaction. A map of pairwise chromosome interactions for K562 blood cell lines was recently produced[29], and we hypothesised that part of the underlying biological mechanism behind some of the 501 epistatic interactions may arise from chromosome interactions. We found that 44 of the putative epistatic interactions were amongst SNPs that were within 5Mb of known chromosome interactions. This means that SNP A was no more than 2.5Mb from the focal point of the chromosome interaction on chromosome A, and SNP B was no more than 2.5Mb from the focal point on chromosome B.

We performed simulations to test how extreme the observation of 44 epistatic interactions overlapping with chromosome interactions is compared to chance. Chromosome interactions fall within functional genomic regions[13,29], and the SNPs in our epistatic interactions are enriched for functional genomic regions. Therefore, we designed the simulations to ensure that the null distribution was of chromosome interactions between SNPs enriched for functional genomic regions but with no known epistatic interactions. To do this we used the 731 SNPs that form the 501 putative epistatic interactions and randomly shuffled them to create new sets of 501 pairs, disallowing any SNP combinations that were in the original set. Therefore, each new random set was enriched for functional regions but had no genetic interactions. We scanned the map of chromosome interactions for overlaps with the new sets and then repeated the random shuffling process. We performed 1,000 such permutations to generate a null distribution of chromosome interaction overlaps.

We repeated this process, searching for overlaps within 1Mb, 250kb, and 10kb.

6.3 SNP colocalisation with genomic features

We tested for enrichment of genomic features for the 687 IndexSNPs that comprise the 434 epistatic interactions with data present in discovery and replication datasets. For each of the 687 IndexSNPs we calculated LD with all regional SNPs within a radius of 0.5Mb and kept all regional SNPs with LD $r^2 > 0.8$. We then cross-referenced the remaining regional SNPs with the annotated chromatin structure reference[28]) querying whether the regional SNPs fell in Predicted promoter region including TSS (TSS), Predicted promoter flanking region (PF), Predicted enhancer (E), Predicted weak enhancer or open chromatin cis regulatory element (WE), CTCF enriched element (CTCF), Predicted transcribed region (T), or Predicted Repressed or Low Activity region (R) positions. Therefore a particular IndexSNP might cover multiple genomic features through LD.

We then performed the whole querying process for each of the 528,509 SNPs present in the SNP chip used in the scan, and used the results from this second

analysis to establish a null distribution for the expected proportion of SNPs for each genomic feature. We calculated p-values for enrichment of each of the seven genomic features independently, and for *cis*- and *trans*-SNPs separately, using a binomial test. For each genomic feature we used the expected proportion of SNPs as the expected probability of "success" (p). Here, a success is defined as an IndexSNP residing in a region that includes the genomic feature. The observed number of successes for each IndexSNP (k) out of the total count of IndexSNPs (n) was then modelled as $\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

6.4 Transcription factor enrichment

To test for enrichment of transcription factor binding sites (TFBS) we followed a procedure similar to that described in Section 6.3. For each of the 687 IndexSNPs we extracted regional SNPs as previously described. We then used the PWMEnrich package in Bioconductor (`http://www.bioconductor.org/packages/2.12/bioc/html/PWMEnrich.html`) to identify which TFBSs each of the regional SNPs for one IndexSNP falls in (within a radius of 250bp). Thus, the number of occurrences of a particular TFBS was counted for each IndexSNP. We used the "Threshold-free affinity" method for identifying TFBSs[39].

We constructed a null distribution of expected TFBS occurrences based on the same null hypothesis as described in Section 6.3 - the probability of an IndexSNP covering a particular TFBS is identical to any of the 528,509 SNPs in the discovery SNP chip. To do this, we performed the same procedure for each SNP in the discovery SNP chip as was performed for each IndexSNP to obtain an expected probability of covering a particular TFBS. We then tested the IndexSNPs for enrichment of each TFBS independently, and for *cis*- and *trans*-SNPs separately. p-values were obtained using Z-scores, calculated by using a normal approximation to the sum of binomial random variables representing motif hits along the sequence[40].

6.5 Defining previously identified SNP associations

The discovery dataset (BSGS) had previously been analysed for additive and dominant marginal effects for all gene expression levels[22,23]. To define SNPs that had been previously detected to have effects for a particular gene expression level we used a significance threshold accounting for multiple testing across SNPs and expression probes, $T_m = 0.05/(528509 \times 7339) = 1.29 \times 10^{-11}$. From this, we found that only nine of the 501 discovery interactions had known main effects, 64 were between SNPs that had no known marginal effects, and 439 were between a SNP with a known marginal effect and a SNP with no known marginal effect.

References

31. Medland, S. E. *et al.* Common variants in the trichohyalin gene are associated with straight hair in Europeans. *American journal of human genetics* **85,** 750–5. ISSN: 1537-6605 (Nov. 2009).

32. Aulchenko, Y. S., Ripke, S., Isaacs, A. & van Duijn, C. M. GenABEL: an R library for genome-wide association analysis. *Bioinformatics (Oxford, England)* **23,** 1294–6. ISSN: 14602059 (May 2007).

33. Yang, J. *et al.* Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics* **43,** 519–525. ISSN: 1061-4036 (May 2011).

34. Westra, H.-J. *et al.* MixupMapper: correcting sample mix-ups in genome-wide datasets increases power to detect small genetic effects. *Bioinformatics (Oxford, England)* **27,** 2104–11. ISSN: 1367-4811 (Aug. 2011).

35. Williams, D. A. Improved likelihood ratio tests for complete contingency tables. *Biometrika* **63,** 33–37. ISSN: 0006-3444 (1976).

36. Alvarez-Castro, J., Le Rouzic, A., Carlborg, O., Álvarez Castro, J. M. & Carlborg, O. How to perform meaningful estimates of genetic effects. *PLoS Genetics* **4,** e1000062 (May 2008).

37. Koch, C. M. *et al.* The landscape of histone modifications across 1% of the human genome in five human cell lines. *Genome research* **17,** 691–707. ISSN: 1088-9051 (June 2007).

38. Rietveld, C. A. *et al.* GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment. en. *Science.* ISSN: 0036-8075. doi:10.1126/science.1235488 (May 2013).

39. Stormo, G. DNA binding sites: representation and discovery. *Bioinformatics* **16,** 16–23 (2000).

40. Ho Sui, S. J. *et al.* oPOSSUM: identification of over-represented transcription factor binding sites in co-expressed genes. *Nucleic acids research* **33,** 3154–64. ISSN: 1362-4962 (Jan. 2005).

Supplementary Figures

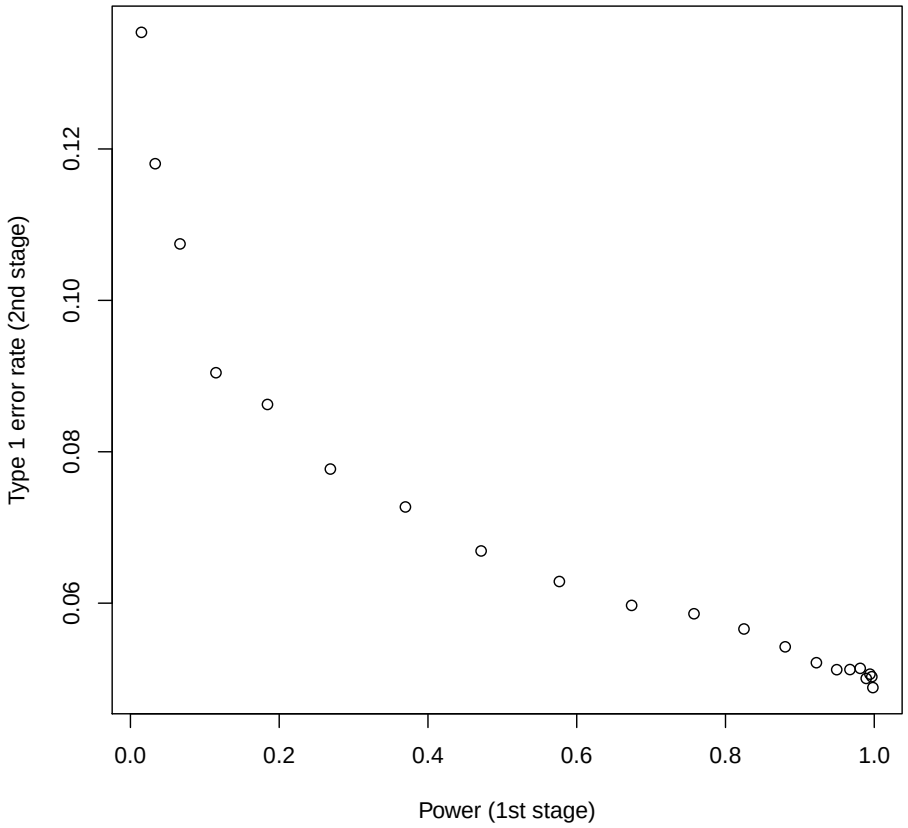

Figure S1: **Type 1 error rate of two stage design assuming a null model of one large additive effect and no epistasis** In stage 1 SNPs are tested for full genetic effects (8 d.f.) and those that surpass a threshold for multiple testing are then tested for significant interaction terms in stage 2. These interaction p-values are then adjusted (Bonferroni) for the total number of tests that passed stage 1. The type 1 error rate of this two stage design is dependent on the power, which is not known empirically.

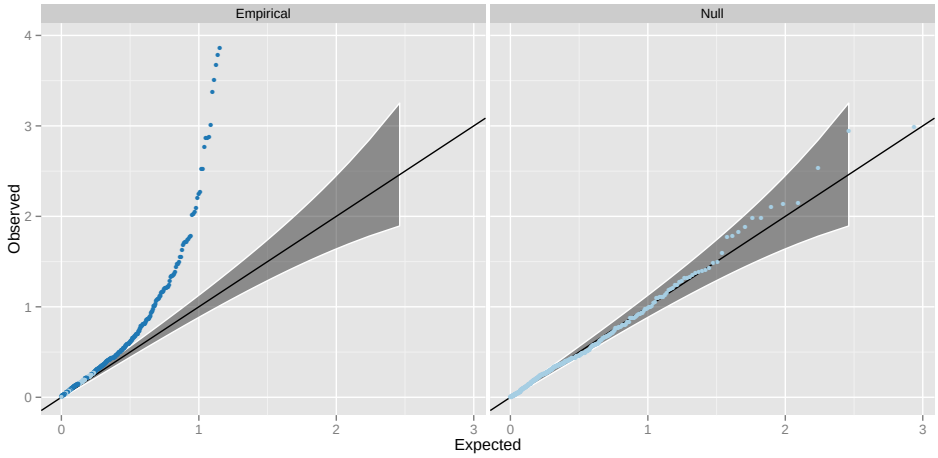

Figure S2: **Q-Q plots of interaction p-values from replication datasets, excluding the 30 points significant at the Bonferroni level** The right panel (Null) shows the interaction p-values from a meta analysis across two independent datasets on 434 SNP pairs where one SNP has a marginal effect. The left panel (Empirical) shows the interaction p-values from the 404 putative interactions that were not significant at the Bonferroni correction threshold. Dark blue points represent p-values that surpass the 2.5% FDR level, as in Figure 2.

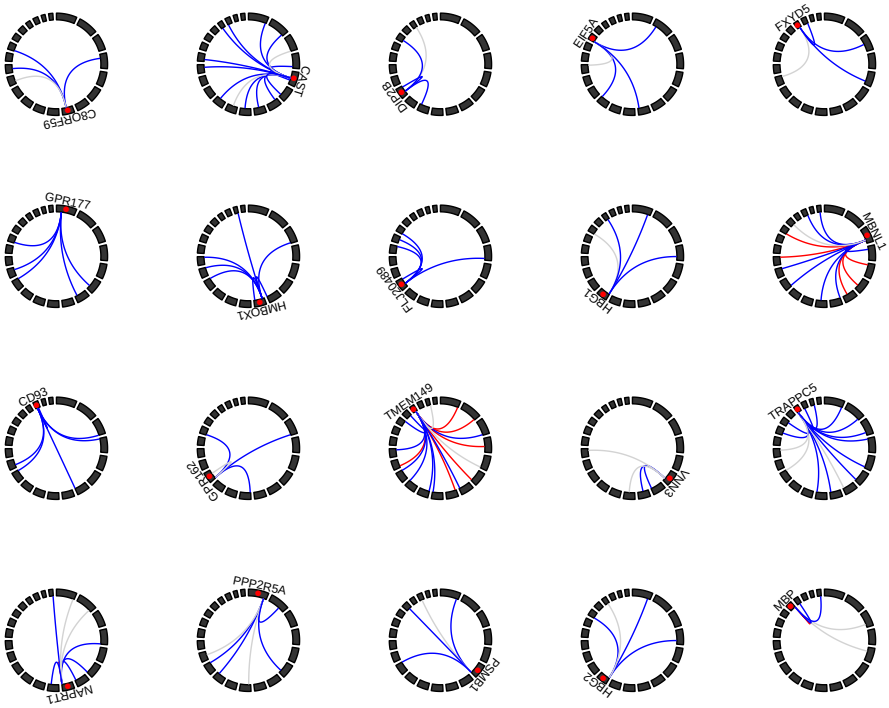

Figure S3: **Gene expression traits with four or more genetic interactions** Circle plots represent the genomic positions for SNPs (linking lines) and expression probes (red points). Chromosomes are represented by black blocks and ordered from 1 to 22 clockwise, starting from the top. Grey lines represent no evidence for replication, blue lines denote interactions that are outside the 97.5% confidence interval or the Q-Q plot (Figure 2), and red lines denote replication at the Bonferroni correction level. Most interactions are characterised as being *cis-trans* to the expression probe.

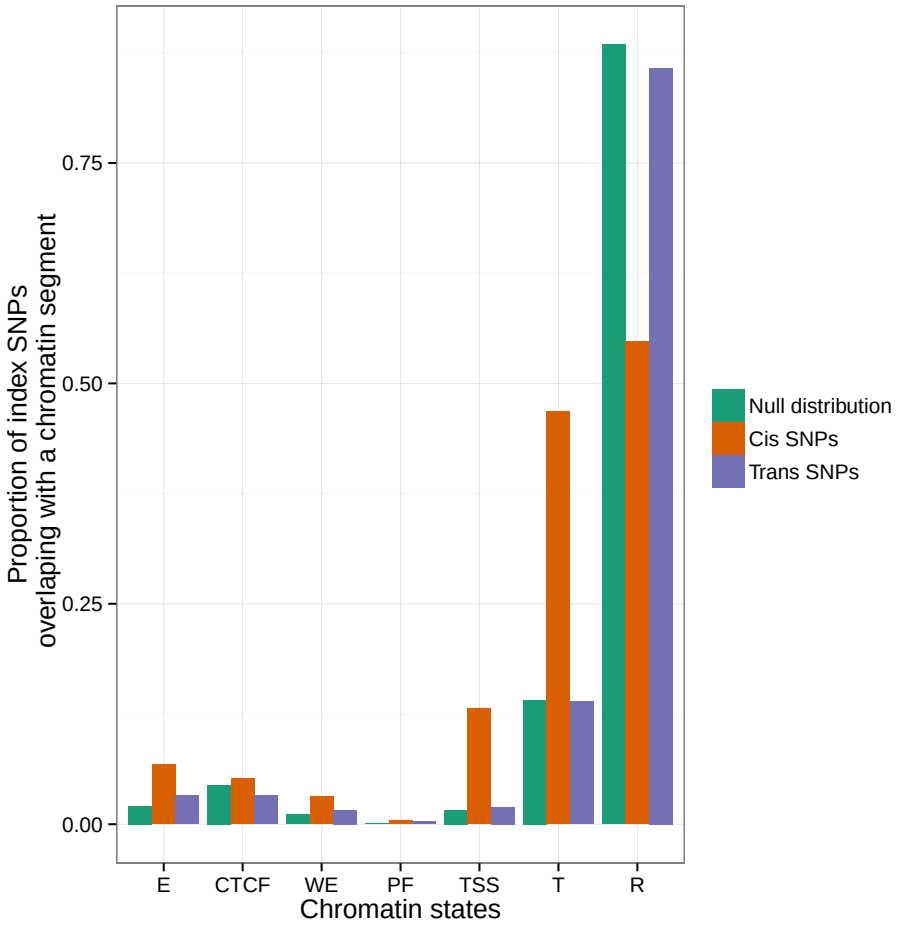

Figure S4: **Location of SNPs relative to genomic features** We used chromatin segmentation[28] as a method for labelling genomic features. All SNPs within 1Mb and $r^2 > 0.8$ of each *cis*- and *trans*-SNP were taken to find which genomic features (x-axis) were covered by the SNPs that compose the 501 significant interactions. Green bars represent the proportion (y-axis) of the 528,509 SNPs used in the analysis that fall within the range of the different genomic features. There is enrichment for *cis*-acting SNPs (red bars) in promotor regions, but *trans*-acting SNPs (blue bars) are not enriched for genomic features. The labels on the x-axis are as follows: E = Predicted enhancer, CTCF = CTCF enriched element, WE = Predicted weak enhancer or open chromatin cis regulatory element, PF = Predicted promoter flanking region, TSS = Predicted promoter region including transcriptional start site, T = Predicted transcribed region, R = Predicted Repressed or Low Activity region

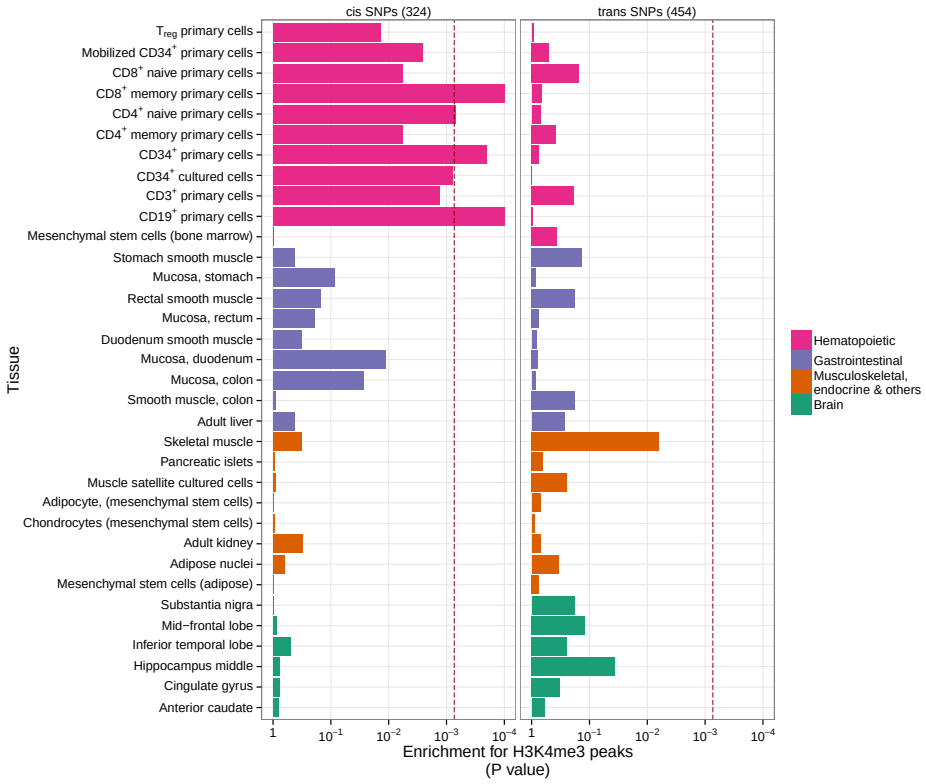

Figure S5: **Tissue specific enrichment of SNPs in transcriptionally active regions** The locations of transcriptional activity can be predicted by chromatin marks, assayed by H3K4me3[27]. Enrichment p-values are calculated using permutation analysis for 34 different cell types (y-axis) in four tissue types (Rows of boxes). The dotted red line denotes significance (Bonferroni correction for 34 cell types, x-axis). There is enrichment for *cis*-acting SNPs in Haematopoietic tissue types only. *Trans*-acting SNPs have no tissue specificity.

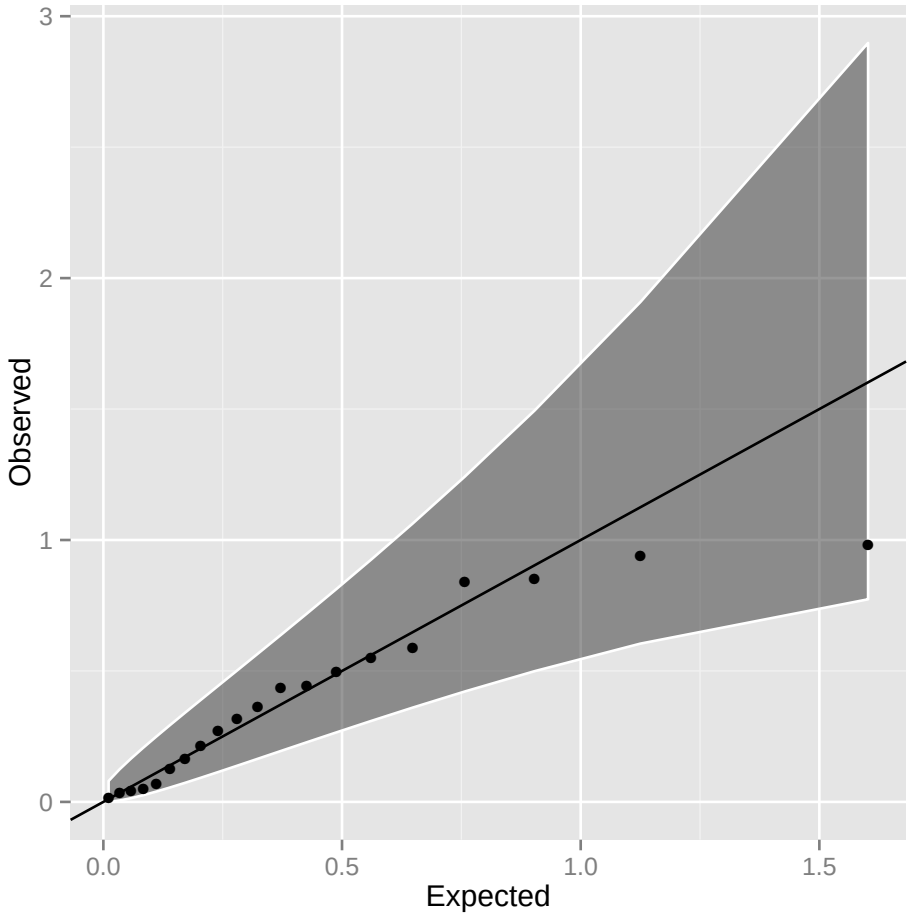

Figure S6: **Q-Q plot of interaction p-values in the CDHWB dataset** Twenty of the 501 discovery SNP pairs passed filtering in the CDHWB dataset (mainly due to small sample size). There is no evidence for enrichment of interaction terms, most likely due to insufficient power given the limited sample size.

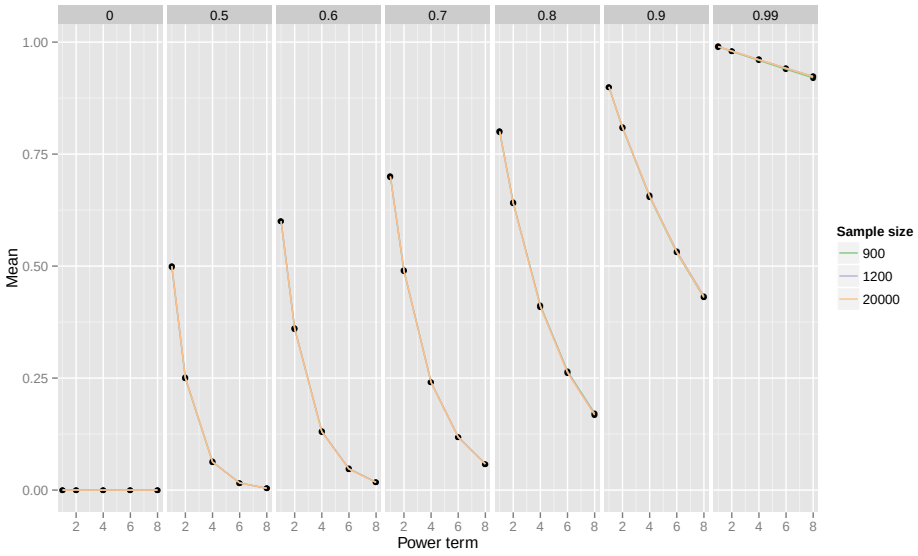

Figure S7: **Sampling mean for different power terms of population r values** Power of detection and replication of epistatic interactions depends not on r^2 between causal variants and observed SNPs, but on r^4, r^6, r^8. For a given population value of LD r (columns of plots), plotted is the sample mean (y-axis) of $\hat{r}$, $\hat{r}^2$ (additive), $\hat{r}^4$ (dominance, A×A), $\hat{r}^6$ (A×D), $\hat{r}^8$ (D×D) (x-axis) for different sample sizes (coloured lines). As true r reduces the statistical power to detect epistatic variants drops dramatically under the assumption that statistical power is proportional to higher moments of r.

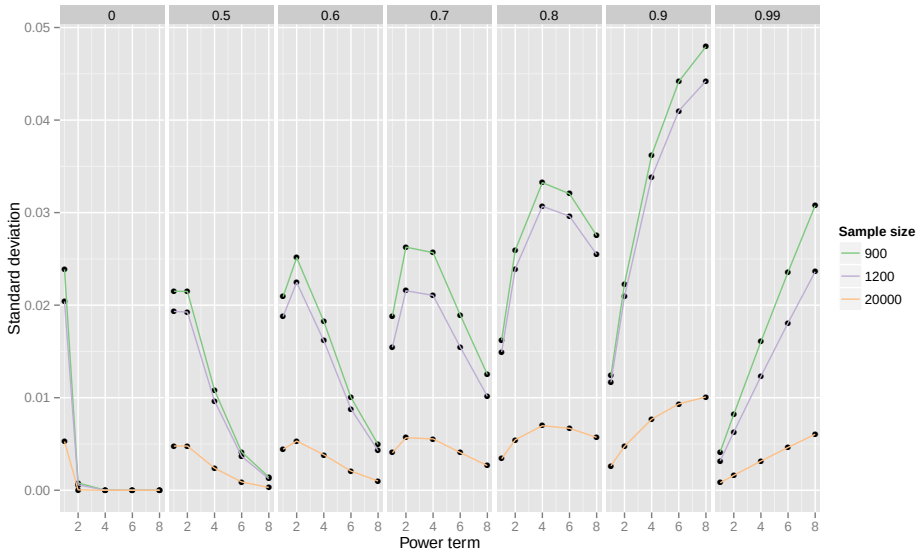

Figure S8: **Sampling standard deviation for different power terms of population r values** Power of detection and replication of epistatic interactions depends not on r^2 between causal variants and observed SNPs, but on r^4, r^6, r^8. For a given a population value of LD r (columns of plots), plotted is the sampling standard deviation (y-axis) of $\hat{r}$, $\hat{r}^2$ (additive), $\hat{r}^4$ (dominance, A×A), $\hat{r}^6$ (A×D), $\hat{r}^8$ (D×D) (x-axis) for different sample sizes (coloured lines). As the power term of r increases the sampling variance also increases. Supposing that there is sufficiently high r^x in the discovery sample for detection of epistasis, the replication sample is less likely to have similarly high r^x as x increases, leading to an expectation of reduced replication rates.

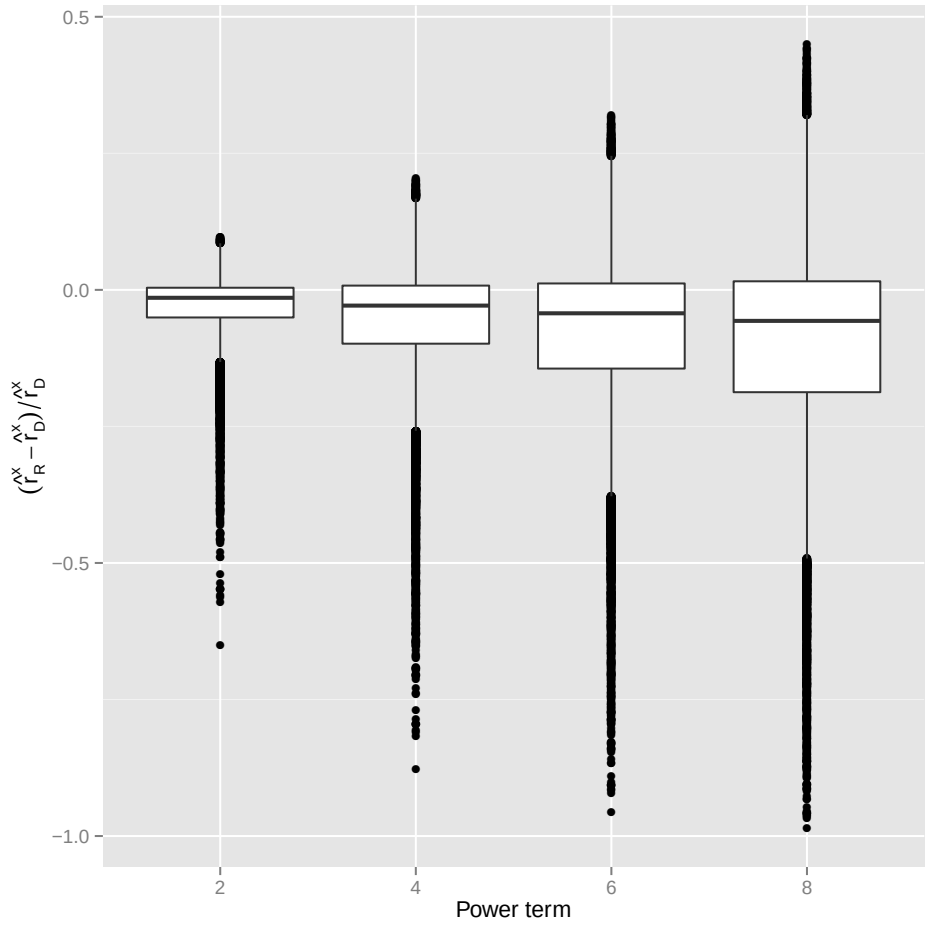

Figure S9: **Reduction in LD as estimated in replication data after ascertaining for high LD in discovery data** 100,000 "unobserved" causal variants (CVs) were tested for LD against a panel of 528,509 "observed" discovery markers (DMs). DM/CV pairs with LD $r^2 > 0.9$ were then tested in an independent sample. Simulation results of the proportional decrease between discovery and replication datasets in LD (y-axis) of $\hat{r}^2, \hat{r}^4, \hat{r}^6, \hat{r}^8$ (x-axis) are shown, where $\hat{r}_D^x$ and $\hat{r}_R^x$ are the sample LD measurements in the discovery and replication datasets, respectively. The average proportional decrease in the replication $\hat{r}_R^x$ was $2.8\%, 5.3\%, 7.4\%$ and 9.2% for $x = 2, 4, 6$ and 8, respectively.

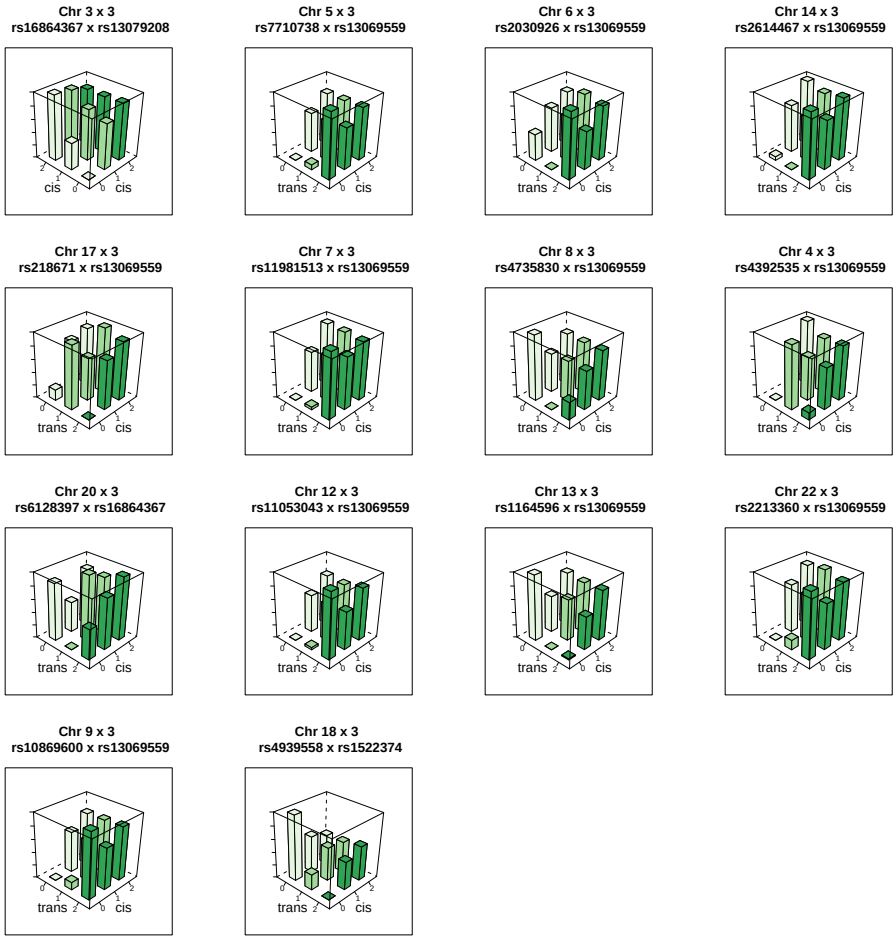

Figure S10: **Genotype-phenotype maps for 14 interactions influencing the expression of MBNL1** Each bar represents the mean phenotypic value for individuals in that genotype class. The rs13069559 SNP typically has a *cis*-additive decreasing effect on the expression of MBNL1, but in many of these interactions the *cis* effect is masked when the *trans* SNP is homozygous for the masking allele.

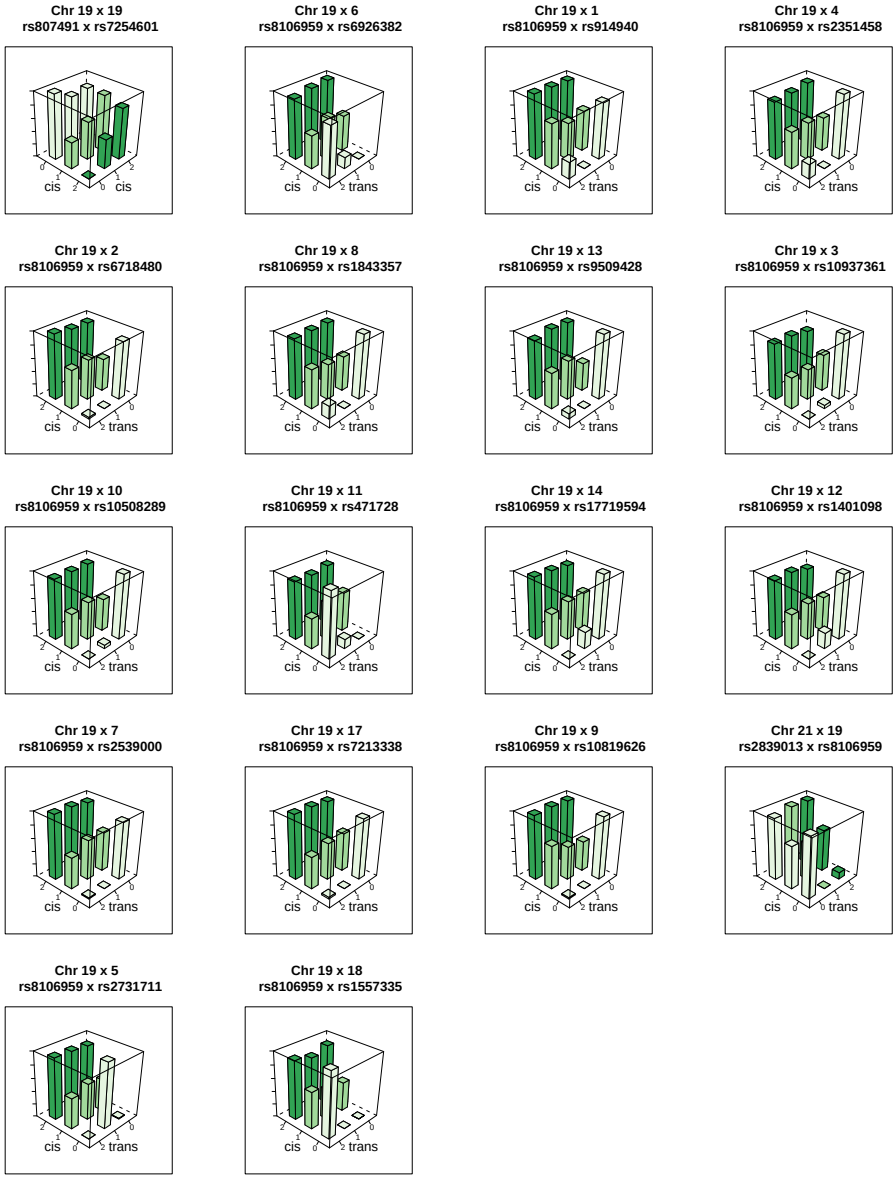

Figure S11: **Genotype-phenotype maps for 19 interactions influencing the expression of TMEM149** Each bar represents the mean phenotypic value for individuals in that genotype class. The rs13069559 SNP typically has a *cis*-additive decreasing effect on the expression of TMEM149, but in many of these interactions the *cis* effect is masked when the *trans* SNP is homozygous for the masking allele.

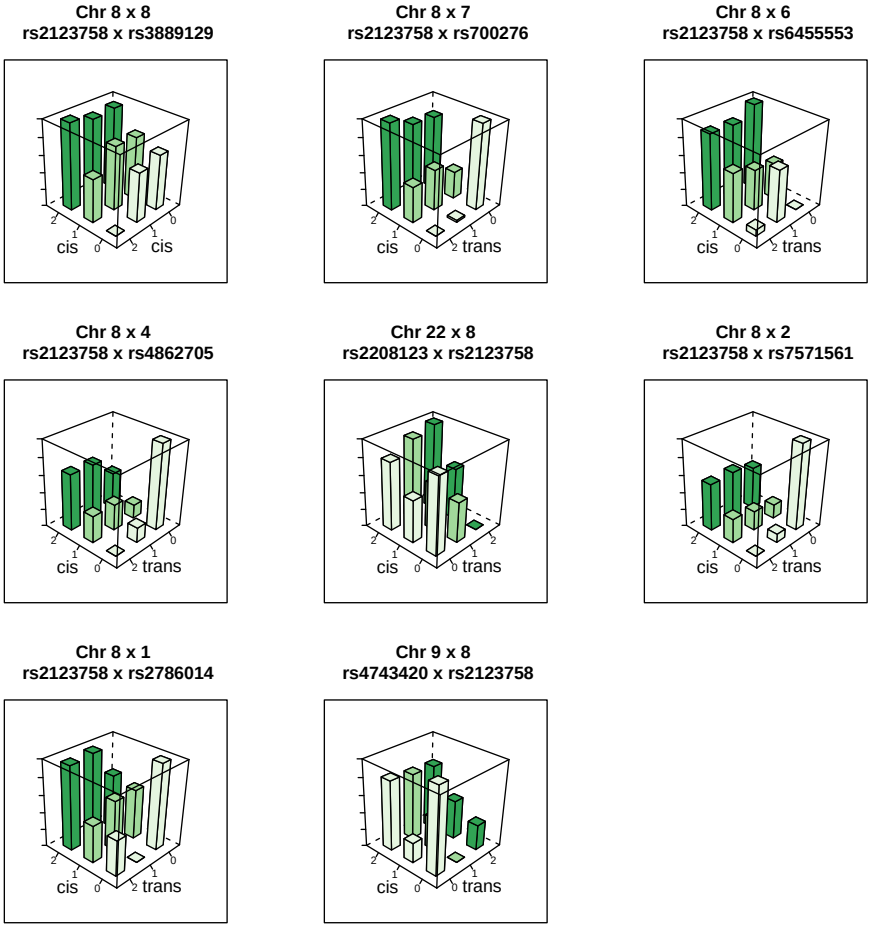

Figure S12: **Genotype-phenotype maps for 8 interactions influencing the expression of NAPRT1** Each bar represents the mean phenotypic value for individuals in that genotype class.

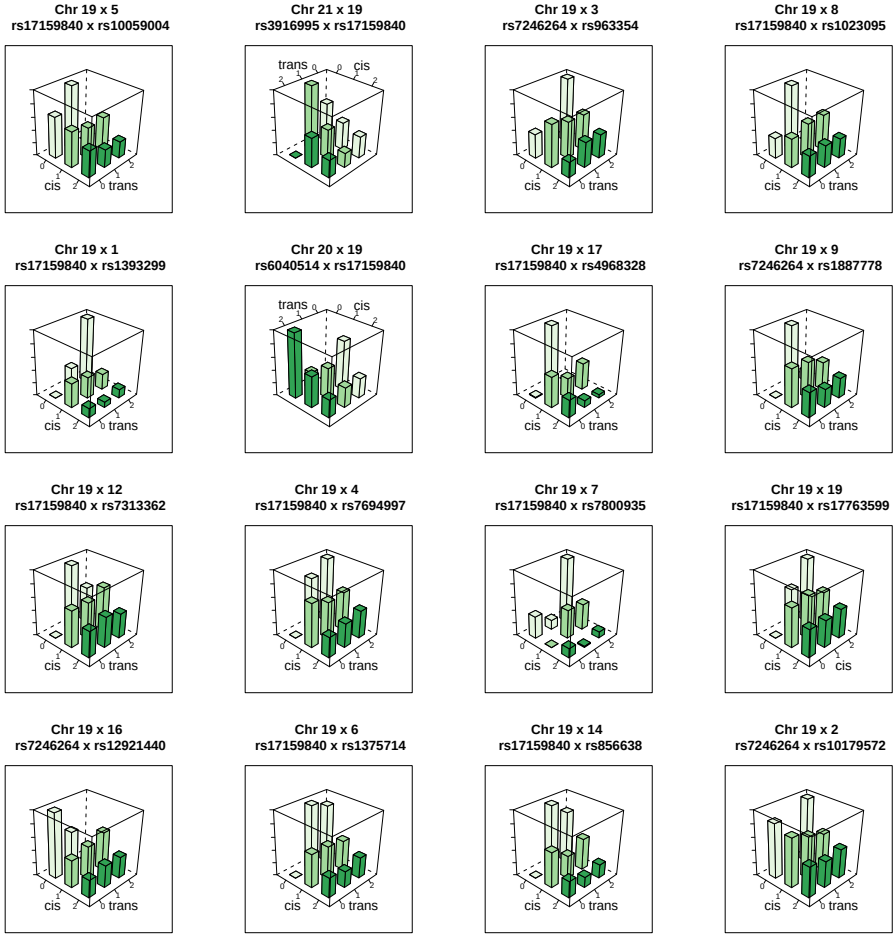

Figure S13: **Genotype-phenotype maps for 16 interactions influencing the expression of TRAPPC5** Each bar represents the mean phenotypic value for individuals in that genotype class.

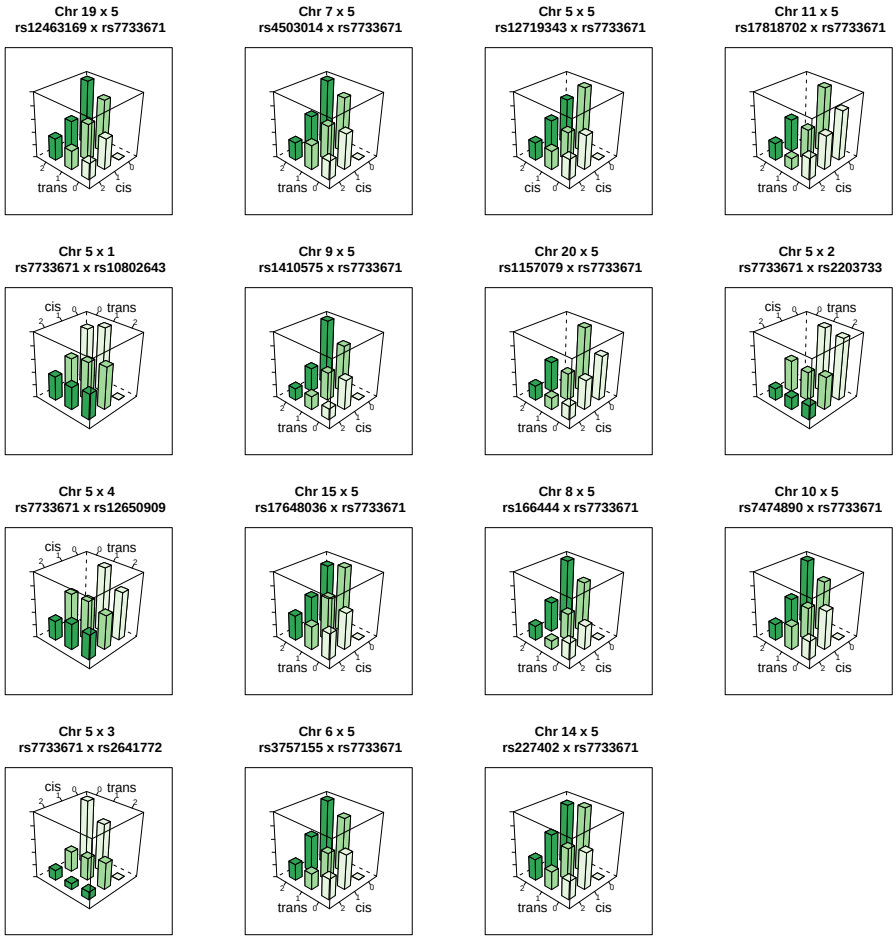

Figure S14: **Genotype-phenotype maps for 15 interactions influencing the expression of CAST** Each bar represents the mean phenotypic value for individuals in that genotype class.

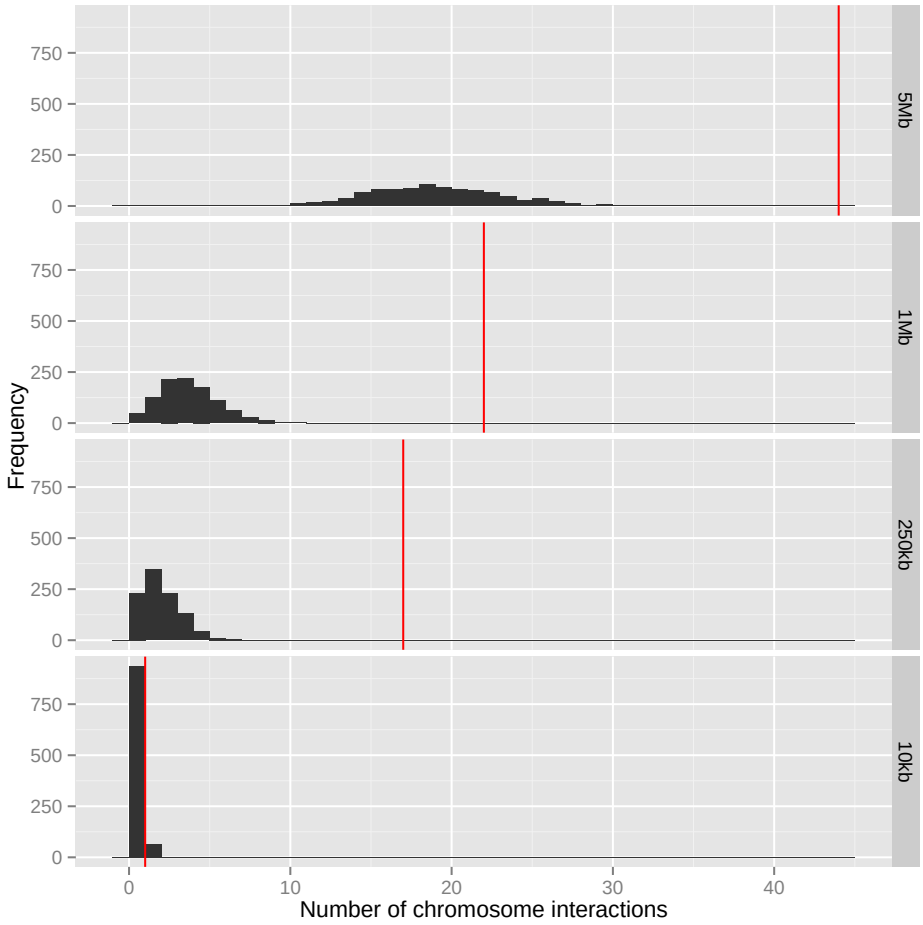

Figure S15: **Number of overlaps between chromosome interactions and epistatic interactions** Interacting chromosome regions may be a possible mechanism underlying epistatic interactions. The number of epistatic interactions within 20kb, 500kb, 2Mb and 10Mb of known chromosome interacting regions are shown by red vertical lines. The histograms represent the null distribution based on random sampling of 1,000 datasets for each window size.

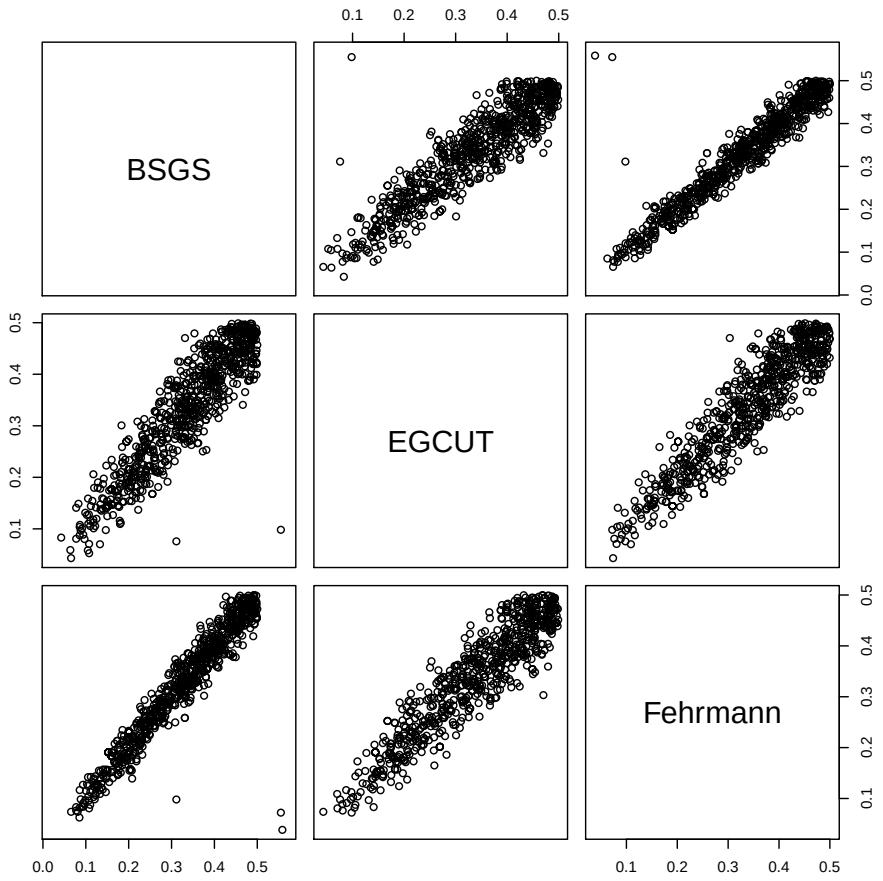

Figure S16: **Comparison of allele frequencies for 781 SNPs involved in genetic interactions across independent populations** Outliers were removed from the analysis as part of the filtering stage during replication.

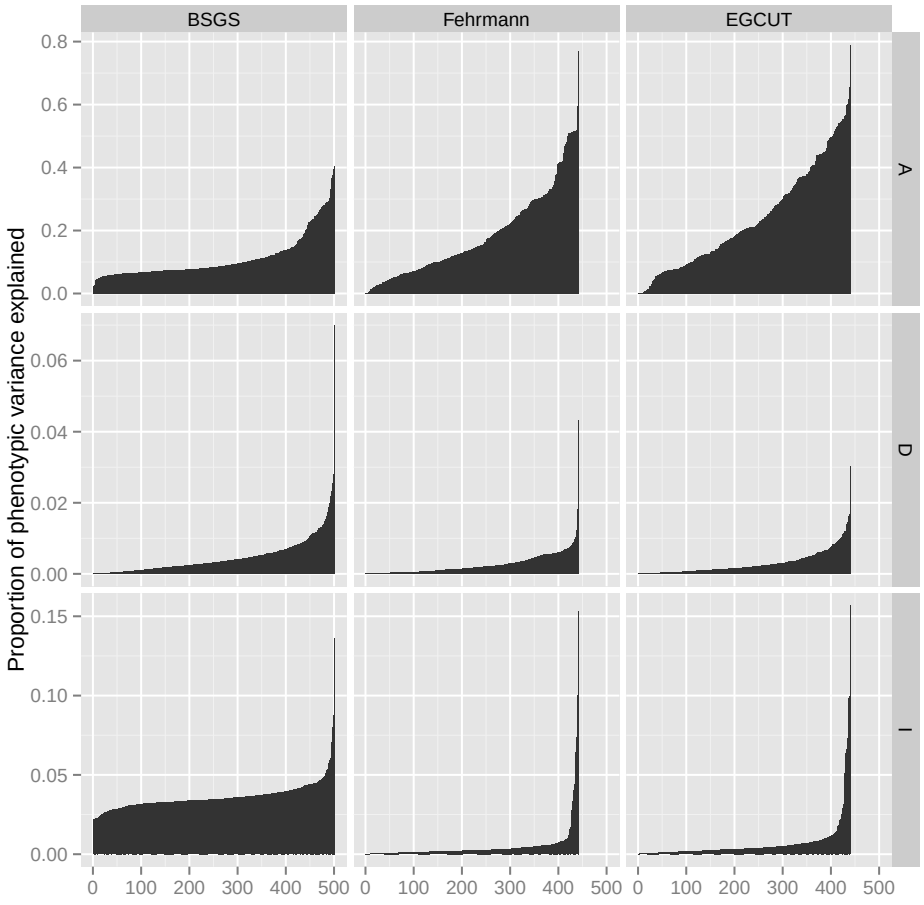

Figure S17: **Comparison of variance explained by additive, dominant and epistatic effects from different cohorts** How does the estimated variance decomposition change in different cohorts? The proportion of the phenotypic variance that is additive (A), dominant (D), or epistatic (I) for each putative interaction is shown on the y-axis (Note: different scales for each row). BSGS has 501 interactions whereas Fehrmann and EGCUT have 434 (x-axis). The variance estimates in each plot are ordered from lowest additive to highest. This is done independently for each cohort to depict the distribution of estimated effects.

Supplementary Tables

Table S1: **Details on 501 interactions discovered in BSGS dataset**

Expression trait			SNP 1				SNP 2				Interaction statistic / $-\log_{10}$ p-values				
Gene ID[a]	Probe ID[b]	Chr.	rs ID	Chr.	Pos/Mb[c]	Association[d]	rs ID	Chr.	Pos/Mb[c]	Association[d]	BSGS[e]	Fehrmann[f]	EGCUT[f]	Meta[g]	Distance / Mb[h]
ABCA7	ILMN_1743205	19	rs3752237	19	1047161	ABCA7	rs596183	6	158100199		5.82	0.38[i]	0.02[i]	0.09[j]	
ABCA7	ILMN_1743205	19	rs3752237	19	1047161	ABCA7	rs914737	9	139522101		5.50				
ABCC3	ILMN_1677814	17	rs9455	17	48771135	ABCC3	rs4732202	7	136057883		6.10	0.02	1.81	0.95	
ACAT1	ILMN_1800008	11	rs227064	11	108207393	ACAT1	rs4744894	9	72001517		6.59	1.04	1.78	2.02	
ADCK1	ILMN_1698777	14	rs12431896	14	78088813	ADCK1	rs4833241	4	122933691		5.59	0.36	1.14	0.87	
ADCK1	ILMN_1698777	14	rs8058066	16	88462550		rs12431896	14	78088813	ADCK1	6.58	2.04	0.83	2.05	
ADK	ILMN_2358626	10	rs2395095	10	76446305	ADK	rs10824092	10	75929517		6.69	18.33	21.21	39.82	0.517
AGAP6	ILMN_3239130	10	rs2611512	10	51515534	AGAP8	rs2547996	5	95174319		6.22				
AHSA2	ILMN_1798308	2	rs10881585	9	137112421		rs842647	2	61119471		7.15	1.83	1.93	2.88	
AHSA2	ILMN_1798308	2	rs2523971	6	29938258	HLA-G	rs1177303	2	61388355	AHSA2	5.45	0.92	0.64	0.94	
AKTIP	ILMN_1665982	16	rs2896940	16	57721127		rs13332406	16	53489705	AKTIP	6.91	0.16	0.99	0.57	4.231
AKTIP	ILMN_1665982	16	rs7189819	16	53536345	AKTIP	rs1362032	7	125543391		5.93	0.71	0.20	0.42	
AKTIP	ILMN_1665982	16	rs7189819	16	53536345	AKTIP	rs1473017	4	179323762		6.18	0.27	0.30	0.23	
ALDH3A2	ILMN_2401641	17	rs3760489	17	19581009	ALDH3A2	rs11720112	3	161996349		6.26	0.33	1.37	1.01	
ANG	ILMN_1760727	14	rs9322855	14	21153299	ANG	rs4866516	5	3032625		5.75	0.02	0.20	0.04	
ANPEP	ILMN_1763837	15	rs11073891	15	90363995	ANPEP	rs3823523	7	154511163		5.85	0.44	1.09	0.90	
ANPEP	ILMN_1763837	15	rs11073891	15	90363995	ANPEP	rs6846031	4	178019148		6.31	0.47	0.17	0.26	
AP3B1	ILMN_1768867	5	rs6453374	5	77508159	AP3B1	rs4684443	3	4818792		5.94	0.05			
APPL2	ILMN_1765076	12	rs935251	12	105580918	APPL2	rs2769594	9	87918528		5.60	0.80	1.02	1.16	
ARL17B	ILMN_3231952	17	rs12947580	17	75768225		rs8079215	17	44064851	ARL17B	5.96				31.703
ARL17B	ILMN_3231952	17	rs2834541	21	35932619		rs8079215	17	44064851	ARL17B	6.65				
ARL17B	ILMN_3231952	17	rs8079215	17	44064851	ARL17B	rs1950646	14	94722497		7.64				
ARL17B	ILMN_3231952	17	rs8079215	17	44064851	ARL17B	rs2197777	12	125831219		6.26				
ARL17B	ILMN_3231952	17	rs8079215	17	44064851	ARL17B	rs2684789	15	99492045		5.98				
ARL17B	ILMN_3231952	17	rs8079215	17	44064851	ARL17B	rs9834627	3	191203546		5.72				
ATP13A1	ILMN_2134224	19	rs4284750	19	19810050		rs873870	19	19738554		5.30	12.18	3.25	14.23	0.071
BID	ILMN_1763386	22	rs8919	22	18213057	BID	rs9804943	12	129906275		5.84	0.06	0.40	0.14	
BID	ILMN_2372413	22	rs181405	22	18233000		rs10888267	1	248059423		6.60	0.87	0.16	0.50	
C11ORF17	ILMN_1752988	11	rs2568061	11	8886260	C11ORF17	rs6553184	4	189150656		5.66	1.15	0.04	0.54	
C13ORF18	ILMN_2196550	13	rs2110603	16	6259852		rs674754	13	46913416	C13ORF18	6.66	0.28	0.28	0.22	
C13ORF18	ILMN_2196550	13	rs674754	13	46913416	C13ORF18	rs6857876	4	153610164		3.87	0.38	0.50	0.43	
C14ORF173	ILMN_2393450	14	rs11089825	22	37575398		rs4983382	14	105189504	C14ORF173	6.02	0.60	0.84	0.85	
C14ORF173	ILMN_2393450	14	rs3935344	15	92276674		rs4983382	14	105189504	C14ORF173	5.98	0.31	0.28	0.24	
C14ORF173	ILMN_2393450	14	rs4983382	14	105189504	C14ORF173	rs10754644	1	238724741		7.15	0.42	0.34	0.35	
C14ORF4	ILMN_1804396	14	rs1293455	18	13819673		rs2655991	14	77574438		4.87				
C14ORF4	ILMN_1804396	14	rs2655991	14	77574438		rs10972462	9	35427324		4.32				
C14ORF4	ILMN_1804396	14	rs2655991	14	77574438		rs6445340	3	63371601		4.40				
C14ORF4	ILMN_1804396	14	rs2655991	14	77574438		rs9787151	1	63179138		4.05				
C14ORF4	ILMN_1804396	14	rs4793445	17	70416307		rs2655991	14	77574438		3.85				
C14ORF4	ILMN_1804396	14	rs6010061	22	51151724		rs2655991	14	77574438		4.61				
C14ORF4	ILMN_1804396	14	rs7245800	19	52083552		rs2655991	14	77574438		4.69				
C17ORF60	ILMN_1747347	17	rs9907897	17	63502633		rs7405659	17		C17ORF60	6.79	0.53	0.05	0.19	
C1ORF86	ILMN_1726989	1	rs2334323	6	110577257		rs2257182	1	2082566	C1ORF86	5.90	0.01	0.50	0.13	
C1ORF86	ILMN_2097790	1	rs2279474	18	46384412		rs2460002	1	2119833	C1ORF86	5.65		0.03		
C1ORF86	ILMN_2097790	1	rs7188668	16	25711358		rs2460002	1	2119833	C1ORF86	5.59	0.29	0.50	0.37	
C21ORF57	ILMN_1795836	21	rs4819271	21	48052838		rs901964	12	48676038	ZNF641	4.91	0.65	0.08	0.28	
C21ORF57	ILMN_1795836	21	rs9978658	21	48027084		rs11701361	21	47764477		9.42	6.08	16.36	21.67	0.263
C5ORF4	ILMN_1728742	5	rs1122762	18	45866512		rs286595	5	154348552	C5ORF4	5.55	0.72	0.04	0.27	
C8ORF59	ILMN_1653205	8	rs12429804	13	36577930		rs2896452	8	86102223	C8ORF59	5.49	0.29	0.02	0.07	
C8ORF59	ILMN_1653205	8	rs12454561	18	31272238		rs2896452	8	86102223	C8ORF59	5.45	0.31			
C8ORF59	ILMN_1653205	8	rs2896452	8	86102223	C8ORF59	rs1004564	4	55242625		7.62	0.38	0.18	0.21	
C8ORF59	ILMN_1653205	8	rs7152284	14	52273663		rs2896452	8	86102223	C8ORF59	5.67	2.18	0.07	1.33	

Continued on next page

Table S1 – continued from previous page

Expression trait			SNP 1				SNP 2				Interaction statistic / $-\log_{10}$ p-values				
Gene ID[a]	Probe ID[b]	Chr.	rs ID	Chr.	Pos/Mb[c]	Association[d]	rs ID	Chr.	Pos/Mb[c]	Association[d]	BSGS[e]	Fehrmann[f]	EGCUT[f]	Meta[g]	Distance / Mb[h]
C8ORF59	ILMN_1653205	8	rs8051751	16	7188323		rs2896452	8	86102223	C8ORF59	5.79	1.39	0.18	0.87	
C9ORF72	ILMN_1741881	9	rs10122902	9	27556780	C9ORF72	rs2526698	1	242029101		6.36	0.96	0.01	0.37	
CABC1	ILMN_1731064	1	rs12765847	10	4353908		rs3738725	1	227174210	CABC1	6.36	0.94	0.00	0.34	
CARD9	ILMN_1712532	9	rs4266763	9	139289825	INPP5E	rs684040	1	82128660		5.81				
CARD9	ILMN_1712532	9	rs4573661	11	6026661		rs4077515	9	139266496	INPP5E	6.61	0.09	0.86	0.42	
CAST	ILMN_1717234	5	rs1157079	20	6778978		rs7733671	5	96000269	CAST	7.07	0.23	0.96	0.62	
CAST	ILMN_1717234	5	rs12463169	19	17321669		rs7733671	5	96000269	CAST	5.73	0.02	2.85	1.75	
CAST	ILMN_1717234	5	rs12599264	16	81840122		rs7733671	5	96000269	CAST	7.00				
CAST	ILMN_1717234	5	rs12719343	5	125369113		rs7733671	5	96000269	CAST	7.68	0.36	1.57	1.20	29.369
CAST	ILMN_1717234	5	rs1410575	9	78255630		rs7733671	5	96000269	CAST	6.55	0.13	1.34	0.78	
CAST	ILMN_1717234	5	rs166444	8	78392770		rs7733671	5	96000269	CAST	7.01	0.27	0.52	0.37	
CAST	ILMN_1717234	5	rs17648036	15	27311111		rs7733671	5	96000269	CAST	7.81	0.97	0.03	0.41	
CAST	ILMN_1717234	5	rs17818702	11	86107920		rs7733671	5	96000269	CAST	6.62	1.15	0.59	1.09	
CAST	ILMN_1717234	5	rs227402	14	70496867		rs7733671	5	96000269	CAST	6.12	0.11	0.01	0.01	
CAST	ILMN_1717234	5	rs2822124	21	15166804		rs7733671	5	96000269	CAST	6.87				
CAST	ILMN_1717234	5	rs3757155	6	136458593		rs7733671	5	96000269	CAST	7.24	0.07	0.33	0.12	
CAST	ILMN_1717234	5	rs4503014	7	31149140		rs7733671	5	96000269	CAST	5.88	0.92	1.56	1.72	
CAST	ILMN_1717234	5	rs7474890	10	59590078		rs7733671	5	96000269	CAST	6.74	0.49	0.12	0.23	
CAST	ILMN_1717234	5	rs7733671	5	96000269	CAST	rs10802643	1	238120177		7.42	0.75	0.78	0.93	
CAST	ILMN_1717234	5	rs7733671	5	96000269	CAST	rs12650909	4	170192890		7.42	0.23	0.78	0.50	
CAST	ILMN_1717234	5	rs7733671	5	96000269	CAST	rs2203733	2	224093101		6.07	0.22	0.87	0.54	
CAST	ILMN_1717234	5	rs7733671	5	96000269	CAST	rs2641772	3	195531841		6.93	0.19	0.26	0.15	
CAT	ILMN_1651705	11	rs872311	18	66175386		rs11032695	11	34447586	CAT	6.41	0.26	0.30	0.22	
CCDC88B	ILMN_1772208	11	rs2353203	19	17099980		rs541207	11	64125142	CCDC88B	5.68	0.33	0.37	0.31	
CCDC88B	ILMN_1772208	11	rs694739	11	64097233	CCDC88B	rs12771349	10	96998193		5.62	0.23	0.18	0.14	
CD36	ILMN_1784863	7	rs3211834	7	80280117		rs1254900	2	85816334	VAMP8	6.93	0.15	0.01	0.02	
CD55	ILMN_1800540	1	rs750801	11	76033374		rs6700168	1	207502534	CD55	5.09	0.08	0.03	0.02	
CD93	ILMN_1704730	20	rs1884655	20	23074375	CD93	rs10255470	7	157182040		6.06	1.74	0.24	1.20	
CD93	ILMN_1704730	20	rs1884655	20	23074375	CD93	rs4696726	4	7992632		5.71	0.13	0.80	0.42	
CD93	ILMN_1704730	20	rs1884655	20	23074375	CD93	rs7622580	3	196721395		5.56	0.04	0.27	0.08	
CD93	ILMN_1704730	20	rs1884655	20	23074375	CD93	rs838875	12	125145394		6.31	0.24	1.67	1.16	
CD93	ILMN_1704730	20	rs1884655	20	23074375	CD93	rs9576388	13	38434472		7.88	0.71	0.22	0.45	
CD93	ILMN_1704730	20	rs2868504	20	37771578		rs1884655	20	23074375	CD93	5.71	0.64	0.75	0.81	14.697
CD93	ILMN_1704730	20	rs4813479	20	23076914	CD93	rs10925747	1	238899903		7.43				
CD93	ILMN_1704730	20	rs4813479	20	23076914	CD93	rs2873420	8	136500554		7.02				
CD93	ILMN_1704730	20	rs4813479	20	23076914	CD93	rs4328531	18	74439542		6.13				
CD93	ILMN_1704730	20	rs4813479	20	23076914	CD93	rs4789981	17	77264482		6.08				
CDC16	ILMN_2339796	13	rs861544	14	104162263		rs7324744	13	115008038	CDC16	5.46	0.21	0.14	0.11	
CDK5R1	ILMN_1730928	17	rs9905940	17	46614102	HOXB2	rs11655031	17	30833162	CDK5R1	5.47	0.95	0.07	0.45	15.781
CEACAM21	ILMN_1745949	19	rs200609	20	51956250		rs4803481	19	42066556	CEACAM21	6.15	0.90	0.12	0.48	
CEACAM21	ILMN_1745949	19	rs4803481	19	42066556	CEACAM21	rs2421050	5	158943044		6.67	2.16	0.16	1.44	
CEP192	ILMN_1703754	18	rs6505780	18	13069782	CEP192	rs13132719	4	180265266		5.75	0.15	0.24	0.12	
CEP63	ILMN_1787808	3	rs3825569	14	101350298		rs13079012	3	134247706	ANAPC13	6.36	0.23	0.10	0.09	
CES1	ILMN_2359945	16	rs8192935	16	55861794	CES1	rs772788	2	235248562		5.65				
CHPT1	ILMN_2202940	12	rs591967	13	38838122		rs2695290	12	102087844	CHPT1	5.74	0.72	0.20	0.44	
CHPT1	ILMN_2202940	12	rs6539014	12	102277782		rs867578	11	81937002		4.75	0.92	0.02	0.36	
CLEC12A	ILMN_1663142	12	rs429790	16	84471642		rs7313235	12	10132283	CLEC12A	5.55	0.07	1.28	0.67	
CLEC12A	ILMN_2403228	12	rs7305054	12	10156646		rs3903088	10	134236688		7.54	0.95	0.36	0.73	
CLTB	ILMN_1674609	5	rs17129799	11	96929337		rs6863172	5	175595960	CLTB	5.55		0.27		
CNN2	ILMN_1770290	19	rs3752237	19	1047161	ABCA7	rs169130	16	63121080		7.56	0.07	0.07	0.02	
CNN2	ILMN_1770290	19	rs3752237	19	1047161	ABCA7	rs7336017	13	67713633		6.33	1.92	0.28	1.39	
CPSF1	ILMN_1654545	8	rs4333645	8	145569535		rs1455268	4	61738094		6.34	0.10	0.01	0.01	
CPVL	ILMN_1682928	7	rs12596791	16	26115562		rs245884	7	29188475	CPVL	5.74	0.06	0.57	0.23	

Continued on next page

Table S1 – continued from previous page

Expression trait			SNP 1				SNP 2				Interaction statistic / $-\log_{10}$ p-values				
Gene ID[a]	Probe ID[b]	Chr.	rs ID	Chr.	Pos/Mb[c]	Association[d]	rs ID	Chr.	Pos/Mb[c]	Association[d]	BSGS[e]	Fehrmann[f]	EGCUT[f]	Meta[g]	Distance / Mb[h]
CPVL	ILMN_1682928	7	rs2835998	21	39202070		rs245884	7	29188475	CPVL	5.55	0.19	0.03	0.04	
CRIPT	ILMN_1813256	2	rs2131290	4	188859908		rs1531133	2	46843631	CRIPT	5.47	0.28	0.10	0.12	
CRLS1	ILMN_1737685	20	rs6139887	20	5986234	CRLS1	rs1473927	5	62406408		6.18	0.10	0.36	0.15	
CSTB	ILMN_1761797	21	rs9979356	21	45230974		rs3761385	21	45198355		11.99	25.20	16.72	42.27	0.033
CTNNA1	ILMN_1804854	5	rs924943	18	69500505		rs176382	5	138226767	CTNNA1	5.74	0.02	0.41	0.11	
CTSC	ILMN_1696347	11	rs2457684	11	88139983	CTSC	rs7079264	10	108679892		5.67	0.92	0.74	1.03	
CTSC	ILMN_1696347	11	rs5752236	22	26250645		rs7128352	11	88087357	CTSC	5.84	0.49	0.80	0.73	
CTSC	ILMN_2242463	11	rs7930237	11	88117962		rs556895	11	88077479		7.16	18.76	15.06	33.53	0.040
CWF19L1	ILMN_1651886	10	rs7108734	11	11456027		rs12784396	10	102027407	CWF19L1	5.42	0.21	0.01	0.03	
CYBRD1	ILMN_1712305	2	rs2592948	4	129994690		rs888427	2	172368120	CYBRD1	5.89	0.23	0.53	0.34	
CYBRD1	ILMN_1712305	2	rs7852475	9	140698856		rs888427	2	172368120	CYBRD1	5.68	0.20	0.02	0.04	
CYBRD1	ILMN_2087692	2	rs11257679	10	12318284		rs888427	2	172368120	CYBRD1	5.81	0.39	1.87	1.47	
CYBRD1	ILMN_2087692	2	rs6137908	20	23344590		rs888427	2	172368120	CYBRD1	5.53	0.05	0.83	0.36	
CYBRD1	ILMN_2087692	2	rs888427	2	172368120	CYBRD1	rs7591849	2	160112881		5.85	0.87	0.10	0.44	12.255
CYP27A1	ILMN_1704985	2	rs6021982	20	36571928		rs933994	2	219650616	CYP27A1	5.42	0.29	0.86	0.60	
DAB2	ILMN_2128428	5	rs7778910	7	110451383		rs835223	5	39381357	DAB2	5.44	0.48	0.41	0.44	
DCAKD	ILMN_1811648	17	rs9900173	17	43111688		rs1343244	6	82076988		9.12	0.00	0.58	0.14	
DDT	ILMN_1690982	22	rs5760102	22	24248761	DDT	rs2378341	3	187475208		5.62	0.64	0.25	0.42	
DDX58	ILMN_1797001	9	rs4937097	11	125962645		rs7042042	9	32451144		5.31	0.61	0.29	0.44	
DEM1	ILMN_1783996	1	rs10120023	9	137810259	COQ10A	rs2519515	7	88204888		5.47	0.08	0.41	0.16	
DEM1	ILMN_1783996	1	rs12363827	11	106703727		rs10120023	9	137810259	COQ10A	6.39	0.77	0.02	0.29	
DHRS9	ILMN_1733998	2	rs1519956	12	89468283		rs7566044	2	169960422	DHRS9	6.00	0.06	1.17	0.58	
DHRS9	ILMN_1733998	2	rs1528529	7	147132505		rs7566044	2	169960422	DHRS9	6.48	0.37	0.34	0.32	
DHRS9	ILMN_2384181	2	rs2831914	21	29959453		rs2161037	2	169893419	DHRS9	5.51	0.88	0.04	0.37	
DHRS9	ILMN_2384181	2	rs7661304	4	187776431		rs2161037	2	169893419	DHRS9	7.64	0.05	0.11	0.03	
DIP2B	ILMN_1755589	12	rs11080134	17	29161503		rs11169322	12	50610976	LASS5	4.65	0.32	0.05	0.10	
DIP2B	ILMN_1755589	12	rs11169335	12	50636364	LASS5	rs2872008	7	153134888		4.87		0.58		
DIP2B	ILMN_1755589	12	rs338585	19	41711815		rs7134595	12	50730458	LASS5	5.31	0.30	0.22	0.19	
DIP2B	ILMN_1755589	12	rs7134595	12	50730458	LASS5	rs1808634	8	61971140		4.40	0.37			
DIP2B	ILMN_1755589	12	rs7312252	12	50744171	LASS5	rs4532958	10	115214154		5.03	0.09	0.02	0.01	
DIP2B	ILMN_1755589	12	rs871257	12	117994348		rs12427378	12	51074199	LASS5	5.92	0.48	0.00	0.11	66.920
DNAJB6	ILMN_1793770	7	rs2286842	7	157216093		rs3779589	7	157163614	DNAJB6	5.79	0.23	1.45	0.97	0.052
DPH3	ILMN_2349610	3	rs12232308	15	93409054		rs1566972	3	16320360	DPH3	6.17	1.58	0.27	1.12	
ECGF1	ILMN_2109708	22	rs140522	22	50971266	ECGF1	rs4891884	18	64004670		4.81	0.15	1.18	0.70	
ECHDC2	ILMN_1671568	1	rs4234091	2	241911027		rs11206043	1	53402552	ECHDC2	6.19	0.22	0.35	0.22	
ECHDC2	ILMN_1671568	1	rs5992637	22	17675900		rs11206043	1	53402552	ECHDC2	5.58	0.64	0.16	0.35	
EHD4	ILMN_1720083	15	rs10403312	19	53244938		rs1048166	15	42192040	EHD4	6.98	0.90	0.47	0.79	
EIF2B2	ILMN_1713380	14	rs6567288	18	60218334		rs175450	14	75590340	EIF2B2	5.56	0.23	0.11	0.10	
EIF5A	ILMN_1794522	17	rs7216490	17	7221707	EIF5A	rs1269096	14	99603119		5.44	0.56	0.08	0.24	
EIF5A	ILMN_1794522	17	rs7216490	17	7221707	EIF5A	rs1553474	2	49359676		5.55	0.28	0.59	0.41	
EIF5A	ILMN_1794522	17	rs7216490	17	7221707	EIF5A	rs2197210	8	129624067		6.36	0.08	0.05	0.02	
EIF5A	ILMN_1794522	17	rs7216490	17	7221707	EIF5A	rs4471434	11	126387391		5.52	0.05	1.12	0.53	
EMR2	ILMN_2353633	19	rs2827076	21	23196249		rs9305048	19	14879034	EMR2	6.51	0.36	0.04	0.11	
EMR2	ILMN_2353633	19	rs6132112	20	18761714		rs9305048	19	14879034	EMR2	5.56	0.45	0.40	0.41	
EMR2	ILMN_2353633	19	rs9305048	19	14879034	EMR2	rs3007765	13	102480759		6.03				
EPHX2	ILMN_1709237	8	rs1107764	11	127909396		rs13269963	8	27400604	EPHX2	5.70	0.20	0.58	0.35	
ERICH1	ILMN_1731001	8	rs10894861	11	134611176		rs12115088	8	578742	ERICH1	5.43	0.25	1.20	0.81	
ERICH1	ILMN_1731001	8	rs5766218	22	45337329		rs4735900	8	607161	ERICH1	6.11	0.20	0.11	0.09	
ERICH1	ILMN_1731001	8	rs726145	18	31187910		rs12115088	8	578742	ERICH1	5.65	0.29	0.04	0.08	
ERICH1	ILMN_2104696	8	rs4735895	8	600729	ERICH1	rs1517297	4	182786760		5.63	0.67	1.03	1.06	
EXOC3	ILMN_1789419	5	rs187076	10	55228462		rs12188164	5	428236	EXOC3	6.83	0.74	0.19	0.44	
FAHD1	ILMN_2246661	16	rs1560104	16	12708208		rs344363	16	1972548	FAHD1	5.61		1.38		10.736
FCN1	ILMN_1668063	9	rs12580388	12	129591144		rs10120023	9	137810259	COQ10A	6.33	0.27	0.30	0.23	

Continued on next page

Table S1 – continued from previous page

Expression trait			SNP 1				SNP 2				Interaction statistic / $-\log_{10}$ p-values				
Gene ID[a]	Probe ID[b]	Chr.	rs ID	Chr.	Pos/Mb[c]	Association[d]	rs ID	Chr.	Pos/Mb[c]	Association[d]	BSGS[e]	Fehrmann[f]	EGCUT[f]	Meta[g]	Distance / Mb[h]
FEZ2	ILMN_1739586	2	rs2356400	19	44321776		rs13406184	2	36791226	FEZ2	5.78	0.14	0.33	0.16	
FEZ2	ILMN_1739586	2	rs969010	4	159963132		rs11691600	2	36810133	FEZ2	6.59	0.14	0.28	0.14	
FGD2	ILMN_2115005	6	rs4803848	19	46205050		rs831486	6	37001267	FGD2	5.69	0.12	0.25	0.11	
FGD2	ILMN_2115005	6	rs902634	10	133943951		rs831489	6	36999682	FGD2	5.49	1.20	0.11	0.66	
FLJ20489	ILMN_1778144	12	rs17615703	12	117036766		rs3782908	12	48169526	FLJ20489	5.81	0.06	0.70	0.29	68.867
FLJ20489	ILMN_1778144	12	rs3782908	12	48169526	FLJ20489	rs897511	4	167695661		5.53	0.03	0.11	0.02	
FLJ20489	ILMN_1778144	12	rs4792199	17	7992118		rs3782908	12	48169526	FLJ20489	5.74	0.19	0.02	0.04	
FLJ20489	ILMN_1778144	12	rs4984440	15	97033129		rs3782908	12	48169526	FLJ20489	6.49	0.31	0.47	0.36	
FLJ20489	ILMN_1778144	12	rs7204135	16	50626195		rs3782908	12	48169526	FLJ20489	6.90	0.38	0.17	0.21	
FLJ20718	ILMN_1763663	16	rs9325634	21	43818790		rs2287197	16	50106594	FLJ20718	6.04	0.14	0.95	0.53	
FLJ43093	ILMN_2123450	6	rs17112712	14	107276627		rs6906101	6	36667610	FLJ43093	5.48	0.39	0.06	0.13	
FLJ43093	ILMN_2123450	6	rs6906101	6	36667610	FLJ43093	rs13214069	6	32705248		5.44	0.00	0.64	0.18	3.962
FN3KRP	ILMN_1652333	17	rs898095	17	80890638		rs9892064	17	80827903		16.16	28.24	29.39	59.95	0.063
FUCA1	ILMN_1752728	1	rs4971478	2	1346063		rs12744386	1	24168019	FUCA1	6.41	0.01	0.30	0.06	
FXYD5	ILMN_2309848	19	rs1633921	19	35695200		rs788178	13	98328559		3.70	0.09	0.41	0.17	
FXYD5	ILMN_2309848	19	rs17398183	20	55609148		rs2285515	19	35660450	FXYD5	6.58	0.03	0.48	0.15	
FXYD5	ILMN_2309848	19	rs2285515	19	35660450	FXYD5	rs11739594	5	141709563		5.70	0.07	0.17	0.05	
FXYD5	ILMN_2309848	19	rs2285515	19	35660450	FXYD5	rs13067700	3	95331048		6.00	0.09	0.51	0.22	
FXYD5	ILMN_2309848	19	rs2285515	19	35660450	FXYD5	rs17036504	2	47567329		6.10	0.28			
G3BP2	ILMN_2381758	4	rs10230232	7	29390239		rs1553985	4	76554604		5.19	0.08	0.37	0.14	
GAA	ILMN_2410783	17	rs11150847	17	78153130		rs12602462	17	78146016		13.91	19.98	12.99	32.60	0.007
GAA	ILMN_2410783	17	rs8068856	17	78100731	GAA	rs10902506	12	132678089		5.65	0.11	0.39	0.17	
GAPT	ILMN_1675191	5	rs10070522	5	57786110	GAPT	rs7605821	2	235695228		5.85	0.01	0.78	0.28	
GAPT	ILMN_1675191	5	rs7082031	10	128038717		rs10070522	5	57786110	GAPT	5.72	0.26	0.11	0.11	
GATS	ILMN_1699631	7	rs1147447	14	66460742		rs2950520	7	99827148	GATS	5.47	0.83	0.63	0.87	
GATS	ILMN_1699631	7	rs2425256	20	35056572		rs2950520	7	99827148	GATS	6.22		0.42		
GDPD3	ILMN_1774901	16	rs3809624	16	30102802	GDPD3	rs2197465	14	48572632		6.57	0.38	0.35	0.33	
GDPD3	ILMN_1774901	16	rs7204270	16	30156963	GDPD3	rs1015111	4	128972357		5.86	0.55	0.09	0.24	
GNLY	ILMN_1790692	2	rs4145072	13	110899955		rs7577293	2	85935282	GNLY	5.78	0.02	0.45	0.13	
GPN3	ILMN_3239426	12	rs7198646	16	26084476		rs7960552	12	111164237	GPN3	5.72				
GPR162	ILMN_1730816	12	rs1860563	16	6478898		rs2707210	12	6902002	GPR162	5.49	0.36	0.46	0.39	
GPR162	ILMN_1730816	12	rs2272500	12	79685913		rs2707210	12	6902002	GPR162	5.07	0.25	0.03	0.06	72.784
GPR162	ILMN_1730816	12	rs2707210	12	6902002	GPR162	rs4740848	9	6554558		5.47	0.25	0.06	0.07	
GPR162	ILMN_1730816	12	rs2707210	12	6902002	GPR162	rs9827054	3	188880113		6.21	0.96	0.06	0.44	
GPR177	ILMN_1660549	1	rs11057383	12	124369421		rs12065581	1	68732819	GPR177	5.45	0.72	0.67	0.81	
GPR177	ILMN_1660549	1	rs12527241	6	120468039		rs12065581	1	68732819	GPR177	5.76	0.17	0.40	0.22	
GPR177	ILMN_1660549	1	rs12532999	7	127939793		rs12065581	1	68732819	GPR177	6.50	0.79	1.43	1.50	
GPR177	ILMN_1660549	1	rs725613	16	11169683		rs12065581	1	68732819	GPR177	5.43	0.31	0.11	0.13	
GPR177	ILMN_1660549	1	rs9575097	13	82986268		rs12065581	1	68732819	GPR177	6.04	0.95	0.21	0.60	
GPR177	ILMN_2283325	1	rs6566669	18	70506011		rs12065581	1	68732819	GPR177	5.86	0.24	0.34	0.23	
GPR177	ILMN_2283325	1	rs9290426	3	171399321		rs12065581	1	68732819	GPR177	6.50	0.01	0.24	0.04	
GSDMB	ILMN_2347193	17	rs11557467	17	38028634	GSDMB	rs4965745	15	101508261		5.88	0.68	0.20	0.41	
GSTM1	ILMN_2391861	1	rs12248673	10	53192833		rs11101992	1	110266754	GSTM1	6.11	0.27	0.19	0.16	
GSTM1	ILMN_2391861	1	rs1547574	13	85344527		rs11101992	1	110266754	GSTM1	5.91	0.27	1.14	0.79	
GSTM2	ILMN_2201580	1	rs6492807	13	96159560		rs3754446	1	110253241	GSTM1	6.77				
H1F0	ILMN_1757467	22	rs139898	22	38399979		rs4853333	2	77919015		6.36	0.52	0.66	0.65	
H1F0	ILMN_1757467	22	rs139898	22	38399979		rs6497007	15	85877017		6.52	0.27	0.31	0.23	
H1F0	ILMN_1757467	22	rs139898	22	38399979		rs9983949	21	19532546		5.70	0.25	0.48	0.32	
HBG1	ILMN_1796678	11	rs11078523	17	4523167		rs2855039	11	5271671	HBG2	5.47	0.00	0.66	0.19	
HBG1	ILMN_1796678	11	rs12975066	19	35723501		rs2855039	11	5271671	HBG2	5.98	0.15	0.24	0.12	
HBG1	ILMN_1796678	11	rs2855039	11	5271671	HBG2	rs12042181	1	213088494	LQK1	6.78	0.08	0.52	0.21	
HBG1	ILMN_1796678	11	rs2855039	11	5271671	HBG2	rs12503379	4	141533832		6.42	0.01	0.46	0.11	
HBG2	ILMN_2084825	11	rs11078523	17	4523167		rs16912979	11	5309695	HBG2	6.06	0.01	0.41	0.10	

Continued on next page

Table S1 – continued from previous page

Expression trait			SNP 1				SNP 2				Interaction statistic / $-\log_{10}$ p-values				
Gene ID[a]	Probe ID[b]	Chr.	rs ID	Chr.	Pos/Mb[c]	Association[d]	rs ID	Chr.	Pos/Mb[c]	Association[d]	BSGS[e]	Fehrmann[f]	EGCUT[f]	Meta[g]	Distance / Mb[h]
HBG2	ILMN_2084825	11	rs12975066	19	35723501		rs2855039	11	5271671	HBG2	5.77	0.08	0.13	0.05	
HBG2	ILMN_2084825	11	rs2855039	11	5271671	HBG2	rs12042181	1	213088494	LQK1	6.84	0.06	0.54	0.21	
HBG2	ILMN_2084825	11	rs2855039	11	5271671	HBG2	rs12503379	4	141533832		5.98	0.00	0.46	0.10	
HDAC7	ILMN_3266186	12	rs2109029	16	6036851		rs4760636	12	48173352	HDAC7	5.75				
HEBP1	ILMN_1802557	12	rs3782567	12	13145613	HEBP1	rs17686635	8	135220622		5.98	0.15	0.59	0.32	
HEXDC	ILMN_1741180	17	rs1942719	18	71237270		rs7213057	17	80378939	HEXDC	5.81	1.61	0.34	1.22	
HLA-DRA	ILMN_2157441	6	rs4899635	14	77532672		rs7192	6	32411646	HLA-DRB6	5.94	0.90	0.16	0.52	
HLA-F	ILMN_1762861	6	rs11660982	18	75467313		rs2523404	6	29695713	HLA-H	5.69	1.00	0.47	0.86	
HMBOX1	ILMN_1720059	8	rs12435486	14	98670849		rs7837237	8	28876221	HMBOX1	6.54	0.92	1.11	1.34	
HMBOX1	ILMN_1720059	8	rs2837803	21	42112794		rs4732890	8	28751381	HMBOX1	6.62	0.05	1.01	0.46	
HMBOX1	ILMN_1720059	8	rs4765451	12	127237464		rs8180944	8	28904086	HMBOX1	5.80	0.39	3.13	2.52	
HMBOX1	ILMN_1720059	8	rs587639	8	132725731		rs7837237	8	28876221	HMBOX1	6.58	0.55	0.34	0.44	103.850
HMBOX1	ILMN_1720059	8	rs8180944	8	28904086	HMBOX1	rs4553956	3	189533772		6.88	3.38	0.03	2.20	
HMBOX1	ILMN_1720059	8	rs8180944	8	28904086	HMBOX1	rs7810884	7	158276926		6.12	0.34	0.66	0.52	
HMBOX1	ILMN_1720059	8	rs9521666	13	110897444		rs8180944	8	28904086	HMBOX1	5.45	0.67	0.26	0.45	
HNRPH1	ILMN_2101920	5	rs6894268	5	179032488		rs4700810	5	178991794		15.38	8.55	3.01	10.37	0.041
HSPC157	ILMN_3194087	1	rs555812	16	88882257		rs4654783	1	22439520	HSPC157	5.51				
HSPC157	ILMN_3194087	1	rs6063164	20	46486900		rs4654783	1	22439520	HSPC157	6.51				
HSPC157	ILMN_3194087	1	rs662739	12	121229893		rs4654783	1	22439520	HSPC157	6.61				
HSPC157	ILMN_3194087	1	rs7088558	10	101884937	CWF19L1	rs4654783	1	22439520	HSPC157	6.48				
IL32	ILMN_1778010	16	rs1554999	16	3115628	IL32	rs4759890	12	131757163		6.90	0.19	0.50	0.29	
IL32	ILMN_2368530	16	rs765044	19	2560423		rs1554999	16	3115628	IL32	5.53	0.69	0.23	0.44	
INPP5E	ILMN_1811301	9	rs8044524	16	81603771		rs1127152	9	139335599	INPP5E	5.58	1.46	0.84	1.55	
JAZF1	ILMN_1682727	7	rs757355	12	47970693		rs849341	7	28288174		8.16	0.02	0.26	0.05	
KCNJ15	ILMN_1675756	21	rs2186344	21	39606769	KCNJ15	rs424299	11	5570771		5.64	0.65	0.13	0.33	
KIR2DS5	ILMN_1691803	19	rs649216	19	55324635	KIR2DL1	rs6419960	4	189055298		4.74	0.46	0.89	0.77	
KTELC1	ILMN_1811104	3	rs4349034	13	84597119		rs727905	3	119119433	KTELC1	5.53	0.08	0.80	0.37	
KTELC1	ILMN_1811104	3	rs6815953	4	183109012		rs6414283	3	119195913	KTELC1	5.45	0.64	0.08	0.28	
L3MBTL2	ILMN_2336109	22	rs4822006	22	41519362	L3MBTL2	rs1294338	1	233438952		5.88	0.33	0.04	0.09	
LAP3	ILMN_1683792	4	rs7042087	9	132602868		rs7658240	4	17588950	LAP3	5.72	0.24	0.47	0.31	
LAX1	ILMN_1769782	1	rs1891432	1	203877662		rs10900520	1	203780591		19.16	18.60	11.22	29.24	0.097
LDLRAP1	ILMN_1809040	1	rs1552032	15	59971635		rs6687605	1	25889632	LDLRAP1	6.00				
LGALS9	ILMN_2412214	17	rs12450521	17	26083392		rs11749727	5	179608360		5.16	0.35	0.40	0.34	
LILRA5	ILMN_2357419	19	rs3859532	19	54827248	LILRA5	rs714789	18	71561497		6.13	0.23	0.03	0.05	
LINS1	ILMN_2338197	15	rs11247226	15	101120963	LINS1	rs1278387	10	127804531		5.89		0.13		
LRRC25	ILMN_2150196	19	rs6009951	22	51151350		rs8101804	19	18496107	LRRC25	5.68	0.11	0.35	0.15	
LY86	ILMN_1807825	6	rs977785	6	6588881	LY86	rs1543675	1	78946879		5.61	0.13	0.15	0.07	
LYZ	ILMN_1815205	12	rs2168029	12	69734641	LYZ	rs11981725	7	154137150		5.95	0.15	0.03	0.03	
LYZ	ILMN_2162972	12	rs177820	18	77276964		rs2168029	12	69734641	LYZ	5.71	0.49	0.03	0.16	
LYZ	ILMN_2162972	12	rs2168029	12	69734641	LYZ	rs2253135	9	130319560		6.31	0.61	0.36	0.49	
MAD1L1	ILMN_2358069	7	rs7783715	7	1923385	MAD1L1	rs6414306	3	127011798		5.62	0.25	0.88	0.59	
MAD2L1BP	ILMN_1694711	6	rs7983718	13	103203146		rs1096699	6	43528441	MAD2L1BP	5.93	0.63	1.11	1.09	
MAP1LC3A	ILMN_1776188	20	rs974607	21	29435869		rs6060034	20	33351864	MAP1LC3A	5.78	1.18			
MBNL1	ILMN_2313158	3	rs10869600	9	78225815		rs13069559	3	152187431	MBNL1	7.96	0.79	0.27	0.54	
MBNL1	ILMN_2313158	3	rs11053043	12	9932070		rs13069559	3	152187431	MBNL1	6.70	0.08	2.21	1.37	
MBNL1	ILMN_2313158	3	rs1164596	13	97100681		rs13069559	3	152187431	MBNL1	7.38	1.43	0.63	1.34	
MBNL1	ILMN_2313158	3	rs11981513	7	94648239		rs13069559	3	152187431	MBNL1	7.71	0.43	5.36	4.58	
MBNL1	ILMN_2313158	3	rs16864367	3	152234166		rs13079208	3	152116652		13.49	16.25	24.74	41.56	0.118
MBNL1	ILMN_2313158	3	rs2030926	6	114067127		rs13069559	3	152187431	MBNL1	7.10	0.91	5.80	5.53	
MBNL1	ILMN_2313158	3	rs218671	17	6604708		rs13069559	3	152187431	MBNL1	7.63	0.62	5.82	5.23	
MBNL1	ILMN_2313158	3	rs2213360	22	34291750		rs13069559	3	152187431	MBNL1	6.05	0.52	0.72	0.70	
MBNL1	ILMN_2313158	3	rs2305802	19	16038535		rs13069559	3	152187431	MBNL1	6.94	1.67			
MBNL1	ILMN_2313158	3	rs2614467	14	99770138		rs13069559	3	152187431	MBNL1	5.74	4.13	2.22	5.30	

Continued on next page

Table S1 – continued from previous page

Expression trait			SNP 1				SNP 2				Interaction statistic / $-\log_{10}$ p-values				
Gene ID[a]	Probe ID[b]	Chr.	rs ID	Chr.	Pos/Mb[c]	Association[d]	rs ID	Chr.	Pos/Mb[c]	Association[d]	BSGS[e]	Fehrmann[f]	EGCUT[f]	Meta[g]	Distance / Mb[h]
MBNL1	ILMN_2313158	3	rs4392535	4	41513423		rs13069559	3	152187431	MBNL1	8.39	0.02	4.33	3.02	
MBNL1	ILMN_2313158	3	rs4735830	8	895841		rs13069559	3	152187431	MBNL1	6.74	0.32	4.21	3.38	
MBNL1	ILMN_2313158	3	rs4939558	18	46278591		rs1522374	3	152235530		7.72	0.03	0.27	0.07	
MBNL1	ILMN_2313158	3	rs6128397	20	57253132		rs16864367	3	152234166		7.22	1.34	1.15	1.73	
MBNL1	ILMN_2313158	3	rs7710738	5	22101322		rs13069559	3	152187431	MBNL1	7.92	2.55	7.89	9.28	
MBP	ILMN_2331544	18	rs6079849	20	15462611		rs2051344	18	74715653	MBP	6.26	0.10	0.03	0.02	
MBP	ILMN_2398939	18	rs139568	22	42210985		rs2051344	18	74715653	MBP	5.56	0.03	0.23	0.05	
MBP	ILMN_2398939	18	rs2051344	18	74715653	MBP	rs1125539	3	155204939		5.79	0.02	0.76	0.27	
MBP	ILMN_2398939	18	rs2051344	18	74715653	MBP	rs2619046	5	55097534		6.03	0.15	0.50	0.26	
MBP	ILMN_2398939	18	rs4805021	19	33436367		rs2051344	18	74715653	MBP	5.82	0.03	0.47	0.14	
MBP	ILMN_2398939	18	rs8092433	18	74747424		rs4890876	18	74732087		5.40	7.06	21.91	28.73	0.015
MEGF9	ILMN_2290118	9	rs13039689	20	51922071		rs966396	9	123453281	MEGF9	4.63	1.13	1.33	1.71	
MFN2	ILMN_1651385	1	rs7989895	13	109401737		rs4846085	1	12050634	MFN2	5.76	0.61	0.25	0.41	
MGC13057	ILMN_1787526	2	rs12718598	7	50428445	MGC13057	rs11725347	4	171860973		5.81	0.13	0.30	0.14	
MGC13057	ILMN_1787526	2	rs674608	18	69070772		rs12718598	7	50428445	MGC13057	5.57	0.07	1.03	0.50	
MGC13057	ILMN_1787526	2	rs8058318	16	82628245		rs12718598	7	50428445	MGC13057	7.05	0.11	0.12	0.05	
MGC72104	ILMN_1688318	20	rs845787	20	26197931	MGC72104	rs2660665	8	137526799		4.17	0.05	0.08	0.02	
MGST3	ILMN_1751956	1	rs740441	17	55779644		rs4147592	1	165600146	MGST3	5.45	0.57	0.27	0.40	
MPZL2	ILMN_1752932	11	rs1805	11	118076069	MPZL2	rs11771552	7	154708716		5.90	0.01	0.23	0.04	
MPZL2	ILMN_1752932	11	rs7316716	12	19953193		rs1805	11	118076069	MPZL2	5.64	0.97	1.08	1.35	
MRPL36	ILMN_1800197	5	rs17469061	10	8436432		rs750495	5	1782046	MRPL36	6.89	0.34	0.18	0.19	
MRPL43	ILMN_2258774	10	rs6564769	16	80641040		rs2863095	10	102746503	MRPL43	5.71	0.26			
MRPL52	ILMN_1713966	14	rs1950857	14	26710271		rs3811188	14		MRPL52	6.56	0.14	0.44	0.22	
MRPS10	ILMN_1663664	6	rs10955512	8	110202230		rs722269	6	42194916	MRPS10	7.48	0.46	0.70	0.64	
MRPS10	ILMN_1663664	6	rs11698155	20	15063214		rs2395803	6	42158596	MRPS10	6.85	0.31	0.63	0.46	
MRPS10	ILMN_1663664	6	rs1420537	16	52453567		rs13217993	6	42164401	MRPS10	6.21	0.41	0.25	0.28	
MTMR15	ILMN_2152178	15	rs7178375	15	31215935	MTMR10	rs12431444	14	42068689		5.18	1.87	1.87	2.86	
MX1	ILMN_1662358	21	rs459498	21	42795027		rs11160227	14	95514596		6.31	0.46	0.52	0.50	
MX1	ILMN_1662358	21	rs459498	21	42795027		rs4973801	3	26706382		5.83	0.11	0.50	0.23	
MX1	ILMN_1662358	21	rs459498	21	42795027		rs8130120	21	29363604		6.78	0.29	0.92	0.65	13.431
MYBPC3	ILMN_1781184	11	rs10134030	14	61593110		rs1317149	11	47486885	MYBPC3	5.56	0.13	0.46	0.23	
MYBPC3	ILMN_1781184	11	rs7322768	13	109550561		rs7124681	11	47529947	MYBPC3	5.70	0.04	0.08	0.02	
MYOM1	ILMN_1680344	18	rs4798075	18	3247256	MYOM1	rs2737422	8	134485237		6.02	0.74	0.15	0.40	
N4BP1	ILMN_2201966	16	rs12444224	16	87580855		rs11649236	16	48632478	N4BP1	5.54	2.00	0.59	1.77	38.948
NAAA	ILMN_1668605	4	rs2707575	7	147638723		rs6826085	4	76870229	NAAA	5.65	0.20	0.03	0.04	
NAAA	ILMN_2391512	4	rs2071856	22	37770630		rs6826085	4	76870229	NAAA	5.46	0.27	0.43	0.30	
NAPRT1	ILMN_1710752	8	rs2123758	8	144663661	NAPRT1	rs2786014	1	234897243		6.08	0.07	0.48	0.18	
NAPRT1	ILMN_1710752	8	rs2123758	8	144663661	NAPRT1	rs3889129	8	144613680		8.45	15.12	16.08	30.77	0.050
NAPRT1	ILMN_1710752	8	rs2123758	8	144663661	NAPRT1	rs4862705	4	187445552		5.62	1.27	0.19	0.81	
NAPRT1	ILMN_1710752	8	rs2123758	8	144663661	NAPRT1	rs6455553	6	167811764		6.12	0.87	0.76	1.01	
NAPRT1	ILMN_1710752	8	rs2123758	8	144663661	NAPRT1	rs700276	7	146189057		6.86	1.10	2.58	2.77	
NAPRT1	ILMN_1710752	8	rs2123758	8	144663661	NAPRT1	rs7571561	2	213386267		6.03	0.13	0.47	0.23	
NAPRT1	ILMN_1710752	8	rs2208123	22	48214812		rs2123758	8	144663661	NAPRT1	6.60	0.29	0.88	0.63	
NAPRT1	ILMN_1710752	8	rs4743420	9	103488089		rs2123758	8	144663661	NAPRT1	5.50	0.12	0.17	0.08	
NAPSA	ILMN_1784040	19	rs1405655	19	50882619	NAPSB	rs930280	9	98391111		5.58	0.82	0.10	0.40	
NAPSB	ILMN_2109416	19	rs1405655	19	50882619	NAPSB	rs10882406	10	95976932		5.58	0.67	1.10	1.12	
NAPSB	ILMN_2109416	19	rs1405655	19	50882619	NAPSB	rs7577137	2	234721287		5.58	2.11	0.44	1.71	
NCL	ILMN_2121437	2	rs7563453	2	232301670		rs4973397	2	232291471		7.31	7.51	6.33	12.70	0.010
NDUFA12	ILMN_1737738	12	rs2746971	22	37101890		rs11107847	12	95386791	NDUFA12	3.88	0.39	0.18	0.22	
NMT2	ILMN_1656378	10	rs10906857	10	15239498	NMT2	rs12490878	3	183114008		6.84	0.42	0.34	0.35	
NOD2	ILMN_1762594	16	rs2967636	19	7067773		rs9302752	16	50719103	NOD2	5.90	0.24	0.04	0.06	
NRBF2	ILMN_3237385	10	rs11063498	12	5209048		rs7923609	10	65133822	NRBF2	5.45				
NRBF2	ILMN_3237385	10	rs2375269	11	69876894		rs7923609	10	65133822	NRBF2	5.53				

Continued on next page

Table S1 – continued from previous page

Expression trait			SNP 1				SNP 2				Interaction statistic / $-\log_{10}$ p-values				
Gene ID[a]	Probe ID[b]	Chr.	rs ID	Chr.	Pos/Mb[c]	Association[d]	rs ID	Chr.	Pos/Mb[c]	Association[d]	BSGS[e]	Fehrmann[f]	EGCUT[f]	Meta[g]	Distance / Mb[h]
NRBF2	ILMN_3237385	10	rs6025645	20	56157341		rs7923609	10	65133822	NRBF2	5.45				
NRBF2	ILMN_3237385	10	rs6517815	21	19819016		rs7923609	10	65133822	NRBF2	6.11				
NRD1	ILMN_1800897	1	rs4852124	2	240680022		rs6588415	1	52334047		6.13	0.47	0.05	0.17	
NUDT18	ILMN_1787885	8	rs5017351	11	25453482		rs1005901	8	21964378	NUDT18	5.44	0.03	0.46	0.15	
OAS1	ILMN_1658247	12	rs11613438	12	113480510		rs1047944	6	163997467		8.59	1.27	1.55	2.03	
OAS1	ILMN_1658247	12	rs13311	12	113448652		rs2072133	12	113409260		4.13	4.12	0.81	3.86	0.039
OAS1	ILMN_1675640	12	rs2892233	19	49160255		rs3741981	12		OAS1	4.38	0.87	0.46	0.76	
OPTN	ILMN_2381899	10	rs7192613	16	74286646		rs17512962	10	13169066	OPTN	5.64	0.42	0.06	0.14	
OSBPL5	ILMN_2307032	11	rs2829679	21	26662543		rs998639	11	3149249	OSBPL5	5.00	0.36	0.00	0.07	
OSTF1	ILMN_1742456	9	rs17780195	17	70624189		rs2273770	9	77755469	OSTF1	5.42	0.16	0.87	0.49	
OSTF1	ILMN_1742456	9	rs2273770	9	77755469	OSTF1	rs7718088	5	179590952		5.42	1.20	0.08	0.62	
OVGP1	ILMN_1734542	1	rs10802822	1	240132968		rs1264898	1	111992823	OVGP1	5.43	0.13	1.48	0.88	128.140
OVGP1	ILMN_1734542	1	rs347331	3	140148107		rs1264894	1	111969719	OVGP1	6.04	0.25	1.21	0.82	
PAM	ILMN_2313901	5	rs28092	5	102149795	PAM	rs784600	1	40139553	HPCAL4	5.59	0.66	0.44	0.59	
PCYOX1L	ILMN_1815951	5	rs2438490	5	148726162	PCYOX1L	rs2731939	3	21395989		6.20	0.19	0.26	0.16	
PEX5	ILMN_1660232	12	rs10444467	12	128052636		rs4329748	12	7364442	PEX5	5.85	0.09	0.71	0.32	120.688
PEX5	ILMN_1660232	12	rs7495797	15	27246462		rs4329748	12	7364442	PEX5	5.74	0.34	0.09	0.13	
PFAAP5	ILMN_1797893	13	rs131969	22	49151303		rs7328733	13	33126737	PFAAP5	5.64	0.87	0.36	0.67	
PGLYRP1	ILMN_1704870	19	rs12982353	19	46529456	PGLYRP1	rs1263806	14	21982957		6.51	0.03	0.65	0.24	
PHCA	ILMN_1812552	11	rs493642	11	123097386		rs10736812	11	76708086	PHCA	5.51	0.36	0.90	0.70	46.389
PIK3IP1	ILMN_1719986	22	rs4141404	22	31675185	PIK3IP1	rs2065841	1	61728597		5.60	0.20	0.01	0.03	
PISD	ILMN_1793934	22	rs470072	22	32263131	PISD	rs10498313	14	30398876		5.23	0.02	0.87	0.33	
PISD	ILMN_1793934	22	rs6518752	22	31999127	PISD	rs954627	1	18236681		7.11	0.00	1.19	0.48	
PISD	ILMN_1793934	22	rs715572	22	33234931		rs6518754	22	32097775	PISD	4.12	0.05	0.42	0.15	1.137
PNKD	ILMN_1774604	2	rs6869411	5	158781604		rs4672884	2	219182481	PNKD	6.35	0.16	0.04	0.04	
PNPLA7	ILMN_1662587	9	rs11639998	16	4527109		rs928046	9	140487108	PNPLA7	5.15	0.31	0.78	0.56	
PPFIBP2	ILMN_1675656	11	rs911019	20	49668255		rs4758001	11	7559930	PPFIBP2	4.44	0.29	0.33	0.26	
PPP2R3C	ILMN_1662617	14	rs12914603	15	58350896		rs11156875	14	35619816	PPP2R3C	5.81	0.12	0.42	0.19	
PPP2R5A	ILMN_1738784	1	rs10930170	2	166399467		rs12120009	1	212447167	PPP2R5A	5.63	0.72	0.48	0.66	
PPP2R5A	ILMN_1738784	1	rs12423255	12	123595064		rs12120009	1	212447167	PPP2R5A	5.72	0.08	0.95	0.46	
PPP2R5A	ILMN_1738784	1	rs1889083	13	66222691		rs12120009	1	212447167	PPP2R5A	5.61	0.36	0.13	0.17	
PPP2R5A	ILMN_1738784	1	rs682334	11	107417238		rs12120009	1	212447167	PPP2R5A	5.65	1.69	0.28	1.21	
PPP2R5A	ILMN_1738784	1	rs7757871	6	135030045		rs12120009	1	212447167	PPP2R5A	5.95	0.37	0.06	0.12	
PPP2R5A	ILMN_1738784	1	rs7871178	9	27148475		rs12120009	1	212447167	PPP2R5A	5.72	0.16	0.30	0.16	
PRDX5	ILMN_1711606	11	rs8019823	14	95040482		rs11600990	11	64082807	PRDX5	6.43	0.81	0.14	0.44	
PRKCB1	ILMN_1713603	16	rs2188355	16	23867776		rs10492793	16	12639800		7.34	0.53	0.11	0.25	11.228
PRMT2	ILMN_1675038	21	rs1029231	21	47931653	C21ORF57	rs958127	18	31497346		5.60	0.19	0.03	0.04	
PRMT2	ILMN_1675038	21	rs2839372	21	48063862		rs11701058	21	47776382	C21ORF57	4.81	0.69	4.47	4.06	0.287
PSMB1	ILMN_1789176	6	rs3862607	11	121774705		rs13207114	6	170877444	PSMB1	5.79		0.44		
PSMB1	ILMN_1789176	6	rs4890648	18	43983954		rs6928843	6	170890384	PSMB1	5.14	0.00	0.26	0.04	
PSMB1	ILMN_1789176	6	rs6060930	20	30347832		rs9295415	6	170823379	PSMB1	5.44	0.44	0.21	0.27	
PSMB1	ILMN_1789176	6	rs6928843	6	170890384	PSMB1	rs2769689	1	225797957		4.58	1.95	0.64	1.78	
PSMB1	ILMN_1789176	6	rs7299749	12	131727816		rs13207114	6	170877444	PSMB1	5.42	1.18	0.32	0.86	
PWP1	ILMN_1743049	12	rs2353567	14	95478823		rs11036212	11	5221825	PTDSS1	5.00	0.03	0.48	0.15	
PWP1	ILMN_1743049	12	rs4969205	17	76598123		rs11036212	11	5221825	PTDSS1	5.90	0.80	0.08	0.38	
PWP1	ILMN_1743049	12	rs631562	11	126852438		rs11036212	11	5221825	PTDSS1	5.70	0.02	0.40	0.11	
QDPR	ILMN_1672443	4	rs4946705	6	106348246		rs10020773	4	17526682	QDPR	5.75	1.03	1.25	1.55	
RAB3IP	ILMN_1803197	12	rs241730	22	33375704		rs7305307	12	70235726		6.55	0.25	0.08	0.09	
RABAC1	ILMN_2207363	19	rs1075728	19	42462788	RABAC1	rs7951628	11	120161117		6.42	0.28	0.84	0.59	
RBL2	ILMN_1756999	16	rs9931702	16	53526551	AKTIP	rs1863464	15	26938488		6.38	0.03	0.31	0.08	
RCN1	ILMN_1800276	11	rs10879131	12	41147155		rs4922579	11	32136436	RCN1	5.23	0.58	0.37	0.47	
RCN1	ILMN_1800276	11	rs4922579	11	32136436	RCN1	rs11166957	8	141177468		4.32	0.41	0.09	0.17	
RCN1	ILMN_1800276	11	rs4922579	11	32136436	RCN1	rs1341899	1	102740645		5.40	0.04	0.26	0.07	

Continued on next page

Table S1 – continued from previous page

Expression trait			SNP 1				SNP 2				Interaction statistic / $-\log_{10}$ p-values				
Gene ID[a]	Probe ID[b]	Chr.	rs ID	Chr.	Pos/Mb[c]	Association[d]	rs ID	Chr.	Pos/Mb[c]	Association[d]	BSGS[e]	Fehrmann[f]	EGCUT[f]	Meta[g]	Distance / Mb[h]
RERE	ILMN_1802380	1	rs4982958	14	24987865		rs301819	1	8501786	RERE	5.66	0.61	1.23	1.17	
RERE	ILMN_1802380	1	rs7697290	4	135248366		rs301819	1	8501786	RERE	5.74	0.14	0.10	0.06	
RERE	ILMN_2327795	1	rs11085829	19	13174312		rs301819	1	8501786	RERE	5.12	0.21	0.33	0.21	
RERE	ILMN_2327795	1	rs3852011	3	112844086		rs301819	1	8501786	RERE	5.71	0.08	0.60	0.26	
RNASE6	ILMN_1780533	14	rs11628398	14	21182800	RNASE6	rs7324365	13	100601327		5.48	0.42	0.21	0.26	
RNASE6	ILMN_1780533	14	rs6603134	19	8106521		rs11628398	14	21182800	RNASE6	5.11	0.09	0.22	0.08	
RNF167	ILMN_1794726	17	rs238230	17	4875566		rs4884857	13	54668512		4.37				
RNF167	ILMN_1794726	17	rs400688	17	4839930	RNF167	rs11706900	3	36348968		5.59	0.71	0.46	0.64	
RNPEP	ILMN_1738347	1	rs1107121	21	46127549		rs2819365	1	201983242		6.27	0.11	0.30	0.13	
RNPEP	ILMN_1738347	1	rs8071611	17	67153386		rs2819365	1	201983242		4.32	1.48	0.52	1.28	
RPL13	ILMN_2413278	16	rs352935	16	89648580		rs2965817	16	89513234		4.98	3.79	14.41	17.24	0.135
RPL23AP7	ILMN_2222750	2	rs1401202	16	80320056		rs4849261	2	114450028	RPL23AP7	5.55	0.13	0.73	0.38	
RPL36AL	ILMN_2189933	14	rs3007033	14	50103816	RPL36AL	rs17495030	9	138038093		5.46	0.09	0.06	0.02	
RPL36AL	ILMN_2189936	14	rs4900928	14	50020817	RPL36AL	rs1502991	6	66137260		5.86	0.32	0.20	0.19	
RPL8	ILMN_1764721	8	rs2958482	8	145984615	RPL8	rs1619856	1	234585790		4.59	0.10	0.37	0.15	
RPL8	ILMN_1764721	8	rs4143674	20	4741304		rs2958482	8	145984615	RPL8	4.33	0.13	0.45	0.22	
SEC13	ILMN_3297880	3	rs4889214	16	80913946		rs696221	3	10342876	SEC13	6.48				
SEMA4A	ILMN_1702787	1	rs17085428	5	95388015		rs7695	1	156147326	SEMA4A	5.70	0.22	1.73	1.17	
SESN3	ILMN_1694027	11	rs12147460	14	104412137		rs684856	11	94906111	SESN3	5.50	0.02	0.51	0.15	
SESN3	ILMN_1694027	11	rs355391	15	46591793		rs684856	11	94906111	SESN3	5.67	0.31	0.06	0.10	
SESN3	ILMN_1694027	11	rs684856	11	94906111	SESN3	rs7004947	8	134606425		5.60	0.21	0.51	0.31	
SH3BGRL2	ILMN_1762764	6	rs10838191	11	43893658		rs1354034	3	56849749	PPBP	5.52	0.70	0.12	0.35	
SH3BGRL2	ILMN_1762764	6	rs2545385	5	66383979		rs1354034	3	56849749	PPBP	5.97	0.20	0.51	0.30	
SH3BGRL2	ILMN_1762764	6	rs6845304	4	88280502		rs1354034	3	56849749	PPBP	5.23	0.32	0.71	0.53	
SH3GLB2	ILMN_2158336	9	rs1034120	21	18196922		rs17455517	9	131785369	SH3GLB2	7.40	0.22	0.18	0.13	
SIRPG	ILMN_1771801	20	rs1535883	20	1612819	SIRPG	rs6842739	4	60489510		5.74	0.29	0.18	0.17	
SLC22A18	ILMN_2382505	11	rs11673260	19	52181798		rs367035	11	2923826	SLC22A18	5.47	0.09	0.24	0.09	
SLC22A18	ILMN_2382505	11	rs367035	11	2923826	SLC22A18	rs3110874	7	153224179		5.70	0.15	0.10	0.06	
SLC22A18	ILMN_2382505	11	rs367035	11	2923826	SLC22A18	rs3772054	2	241678528		6.15	0.39	0.13	0.19	
SLC41A3	ILMN_2356111	3	rs1912136	11	24616743		rs6771703	3	125801067	SLC41A3	5.88	1.10	0.82	1.24	
SLC45A4	ILMN_1745778	8	rs6985508	8	142337734	SLC45A4	rs7701916	5	174598073		5.95	0.86	0.07	0.40	
SLC46A3	ILMN_1658639	13	rs949805	17	55602091		rs7981190	13	29259349	SLC46A3	5.52	0.09	0.58	0.26	
SMG7	ILMN_1706553	1	rs8035259	15	97403923		rs10911353	1	183489203	SMG7	6.52	0.17	0.09	0.06	
SMOX	ILMN_1775380	20	rs8118315	20	4161500	SMOX	rs11677815	2	65800982		5.68	0.39	0.62	0.52	
SNHG8	ILMN_3309349	4	rs1105621	9	133050233		rs705837	4	119225940	SNHG8	6.11				
SNORD14A	ILMN_1799381	11	rs1520429	15	46259108		rs214097	11	17291499	SNORD14A	6.60	0.29	1.03	0.72	
SNORD14A	ILMN_1799381	11	rs2634462	11	17339127		rs6486334	11	17015557		7.31	13.11	10.96	23.22	0.324
SNORD89	ILMN_3238662	2	rs10445863	2	115929241		rs750783	2	101889306	SNORD89	6.08				14.040
SNORD89	ILMN_3238662	2	rs11605822	11	122986326		rs750783	2	101889306	SNORD89	5.96				
SNORD89	ILMN_3238662	2	rs2135064	5	26778066		rs750783	2	101889306	SNORD89	6.33				
SNUPN	ILMN_1733932	15	rs8134646	21	46376528	SNUPN	rs7185362	16	81888905		6.45	0.13	1.41	0.83	
SNUPN	ILMN_2364535	15	rs8134646	21	46376528	SNUPN	rs1472075	3	193706323		5.59	0.34	0.00	0.06	
SPATA5L1	ILMN_1729179	15	rs1131620	19	41117869		rs4774580	15	45652086	SPATA5L1	5.44				
STARD10	ILMN_1717052	11	rs2221406	13	90174526		rs1000620	11	72509713		5.65	0.67	0.12	0.33	
STYXL1	ILMN_2210729	7	rs4073164	14	104947517		rs17685	7	75616105	STYXL1	5.88	0.57	0.17	0.31	
SULF2	ILMN_2345142	20	rs11700063	20	46153148	SULF2	rs939294	4	180439236		5.51	0.46	0.24	0.30	
SULT1A4	ILMN_2336133	16	rs1463965	18	74332954		rs3785354	16	28550667	TUFM	7.05	0.01	0.05	0.00	
SULT1A4	ILMN_2336133	16	rs2836657	21	40119768		rs3785354	16	28550667	TUFM	5.83				
SURF6	ILMN_1778032	9	rs6099626	20	56013994		rs3118663	9	136281753	SURF6	6.14	0.26	0.16	0.14	
SYTL2	ILMN_2336609	11	rs1375719	13	103410782		rs485485	11	85495269	SYTL2	5.47	0.28	0.31	0.24	
THBS3	ILMN_1804663	1	rs1939875	11	95422867		rs4072037	1	155162067	THBS3	5.55	0.03	0.15	0.03	
THBS3	ILMN_1804663	1	rs8014956	14	20687978		rs2049805	1	155194980	THBS3	5.65	0.31	0.76	0.55	
TIPRL	ILMN_1781457	1	rs2823245	21	16745523		rs1320993	1	168154599	TIPRL	5.22	0.07	0.40	0.15	

Continued on next page

Table S1 – continued from previous page

Expression trait			SNP 1				SNP 2				Interaction statistic / $-\log_{10}$ p-values				
Gene ID[a]	Probe ID[b]	Chr.	rs ID	Chr.	Pos/Mb[c]	Association[d]	rs ID	Chr.	Pos/Mb[c]	Association[d]	BSGS[e]	Fehrmann[f]	EGCUT[f]	Meta[g]	Distance / Mb[h]
TMED4	ILMN_1804148	7	rs1940400	11	132389627		rs17725246	7	44581986	TMED4	5.70	0.06	1.34	0.70	
TMEM149	ILMN_1786426	19	rs2839013	21	47248981		rs8106959	19	36219525	TMEM149	8.11	0.16	0.48	0.26	
TMEM149	ILMN_1786426	19	rs5762235	22	27925288		rs8106959	19	36219525	TMEM149	6.79				
TMEM149	ILMN_1786426	19	rs6090518	20	45207005		rs8106959	19	36219525	TMEM149	11.09	0.76			
TMEM149	ILMN_1786426	19	rs807491	19	36268923	SNX26	rs7254601	19	36147315	TMEM149	12.16	81.55	45.78	145.78	0.122
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs10508289	10	4799159		8.12	1.55	3.09	3.67	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs10819626	9	133025756		8.02	0.40	0.99	0.80	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs10937361	3	188359436		8.39	3.61	1.18	3.78	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs1401098	12	128884559		7.37	2.41	1.00	2.52	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs1557335	18	64268976		6.95	0.08	0.07	0.03	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs17719594	14	90932598		6.93	3.06	0.77	2.87	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs1843357	8	13822381		6.21	3.72	3.33	6.00	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs2351458	4	113317583		7.30	0.04	9.61	8.00	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs2539000	7	147619772		6.70	1.57	1.52	2.27	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs2731711	5	171792273		5.92	0.19	0.33	0.19	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs471728	11	129595460		8.89	0.90	3.62	3.51	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs6718480	2	233879066		8.55	3.31	5.15	7.36	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs6926382	6	161683974		5.80	3.06	8.80	10.72	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs7213338	17	80357420		5.49	0.07	3.14	2.10	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs914940	1	242889492		6.22	3.36	6.96	9.20	
TMEM149	ILMN_1786426	19	rs8106959	19	36219525	TMEM149	rs9509428	13	21473952		9.44	0.10	5.75	4.47	
TMEM63A	ILMN_1719649	1	rs1254086	13	72890603		rs4149226	1	226027323	TMEM63A	5.60				
TMEM80	ILMN_1708482	11	rs1548475	19	58058246		rs4963126	11	656845	TMEM80	5.79	0.64	0.12	0.32	
TNPO3	ILMN_1683811	7	rs1537146	9	4859303		rs10488630	7	128593948	IRF5	5.61	0.11	0.15	0.07	
TNPO3	ILMN_1683811	7	rs199793	20	22287303		rs10488630	7	128593948	IRF5	5.52	1.03	0.17	0.62	
TRA2A	ILMN_1731043	7	rs7776572	7	23528927		rs11770192	7	23498358		8.23	3.19	1.89	4.09	0.031
TRAPPC4	ILMN_1814650	11	rs1278760	13	113531675		rs3916581	11	118887887	TRAPPC4	5.61	0.28	0.40	0.29	
TRAPPC4	ILMN_1814650	11	rs1793823	11	131018917		rs3916581	11	118887887	TRAPPC4	5.52	0.93	0.01	0.36	12.131
TRAPPC5	ILMN_2372639	19	rs17159840	19	7758194	TRAPPC5	rs10059004	5	166970604		5.97	0.21	1.60	1.07	
TRAPPC5	ILMN_2372639	19	rs17159840	19	7758194	TRAPPC5	rs1023095	8	132022957		6.92	0.37	0.87	0.68	
TRAPPC5	ILMN_2372639	19	rs17159840	19	7758194	TRAPPC5	rs1375714	6	156404902		7.79	0.12	0.18	0.08	
TRAPPC5	ILMN_2372639	19	rs17159840	19	7758194	TRAPPC5	rs1393299	1	242329791		6.43	0.63	0.47	0.59	
TRAPPC5	ILMN_2372639	19	rs17159840	19	7758194	TRAPPC5	rs17763599	19	2369415		6.38	0.21	0.24	0.16	5.389
TRAPPC5	ILMN_2372639	19	rs17159840	19	7758194	TRAPPC5	rs4968328	17	57495457		6.51	0.50	0.38	0.44	
TRAPPC5	ILMN_2372639	19	rs17159840	19	7758194	TRAPPC5	rs7313362	12	129644342		7.08	0.04	0.65	0.25	
TRAPPC5	ILMN_2372639	19	rs17159840	19	7758194	TRAPPC5	rs7694997	4	9947811		5.86	0.20	0.36	0.22	
TRAPPC5	ILMN_2372639	19	rs17159840	19	7758194	TRAPPC5	rs7800935	7	146690926		6.27	0.15	0.33	0.16	
TRAPPC5	ILMN_2372639	19	rs17159840	19	7758194	TRAPPC5	rs856638	14	85439550		6.73	0.24	0.07	0.08	
TRAPPC5	ILMN_2372639	19	rs380708	22	22740855		rs17159840	19	7758194	TRAPPC5	7.58				
TRAPPC5	ILMN_2372639	19	rs3916995	21	45128454		rs17159840	19	7758194	TRAPPC5	7.73	0.85	0.78	1.01	
TRAPPC5	ILMN_2372639	19	rs6040514	20	11272861		rs17159840	19	7758194	TRAPPC5	8.10	0.51	0.55	0.56	
TRAPPC5	ILMN_2372639	19	rs7246264	19	7762978		rs10179572	2	228504503		6.71	0.14	0.02	0.02	
TRAPPC5	ILMN_2372639	19	rs7246264	19	7762978		rs12921440	16	30408765		7.34	0.14	0.26	0.13	
TRAPPC5	ILMN_2372639	19	rs7246264	19	7762978		rs1887778	9	134635088	RAPGEF1	7.05	0.08	0.86	0.40	
TRAPPC5	ILMN_2372639	19	rs7246264	19	7762978		rs963354	3	157393770		7.41	0.36	0.90	0.69	
TREM1	ILMN_1688231	6	rs10862975	12	85749398		rs2395771	6	41264577	TREM1	5.42	0.11	0.25	0.11	
TREM1	ILMN_1688231	6	rs12412964	10	108256422		rs2395771	6	41264577	TREM1	5.92	1.20	1.23	1.69	
TRIM38	ILMN_1697971	6	rs2527180	7	158808416		rs2032447	6	26044369	TRIM38	6.46	0.04	0.91	0.39	
TSPAN14	ILMN_1785060	10	rs968726	17	27194634		rs10748526	10	82273079	TSPAN14	6.00	0.07	0.18	0.06	
TSPAN32	ILMN_1718621	11	rs10838738	11	47663049	MYBPC3	rs12800998	11	2317951	TSPAN32	5.01				45.345
TSPAN32	ILMN_2389970	11	rs12800998	11	2317951	TSPAN32	rs620607	6	137947208		5.51				
TYMP	ILMN_3223126	22	rs140522	22	50971266	ECGF1	rs1198819	2	238746880		6.34				
TYMP	ILMN_3223126	22	rs470119	22	50966914	ECGF1	rs4783126	16	85147633		6.13				

Continued on next page

Table S1 – continued from previous page

Expression trait			SNP 1				SNP 2				Interaction statistic / $-\log_{10}$ p-values				
Gene ID[a]	Probe ID[b]	Chr.	rs ID	Chr.	Pos/Mb[c]	Association[d]	rs ID	Chr.	Pos/Mb[c]	Association[d]	BSGS[e]	Fehrmann[f]	EGCUT[f]	Meta[g]	Distance / Mb[h]
UBASH3A	ILMN_2338348	21	rs1893592	21	43855067	UBASH3A	rs7201194	16	83600397		5.91	0.59	0.42	0.52	
UBASH3A	ILMN_2338348	21	rs1893592	21	43855067	UBASH3A	rs7512594	1	214514361		6.01	0.48	1.29	1.10	
USP36	ILMN_1697227	17	rs2279308	17	76794981	USP36	rs7225546	17	75151717		5.71	0.03	0.14	0.03	1.643
VASP	ILMN_1743646	19	rs1264226	19	46063167		rs2276470	19	45974668		5.09	0.94	5.14	4.95	0.088
VNN2	ILMN_1678939	6	rs10435352	7	105252718		rs1883613	6	133077063	VNN2	5.64	0.84	0.15	0.46	
VNN2	ILMN_1678939	6	rs13044386	20	9116155		rs1883617	6	133072650	VNN2	5.44	0.39	0.69	0.57	
VNN2	ILMN_1678939	6	rs134447	22	49927332		rs1883617	6	133072650	VNN2	5.72				
VNN2	ILMN_1678939	6	rs216495	11	16834510		rs1883617	6	133072650	VNN2	5.77	0.33	0.19	0.19	
VNN3	ILMN_1804935	6	rs10278073	7	151662184		rs2267952	6	133067782	VNN3	6.44	0.16	0.74	0.41	
VNN3	ILMN_1804935	6	rs1443946	8	73006453		rs2267952	6	133067782	VNN3	5.74	0.23	0.48	0.31	
VNN3	ILMN_1804935	6	rs348462	9	75547169		rs2267952	6	133067782	VNN3	6.44	0.31	0.17	0.17	
VNN3	ILMN_1804935	6	rs7157055	14	83262064		rs2267952	6	133067782	VNN3	5.82	0.03	0.19	0.04	
VNN3	ILMN_2387680	6	rs2823165	21	16594253		rs2267952	6	133067782	VNN3	6.12	0.73	1.15	1.21	
VNN3	ILMN_2387680	6	rs9596457	13	51692548		rs2267952	6	133067782	VNN3	4.83	0.46	0.05	0.16	
VSTM1	ILMN_1763455	19	rs10500316	19	54553697	VSTM1	rs4552100	18	71024750		5.60	0.53	0.54	0.57	
VSTM1	ILMN_1763455	19	rs10500316	19	54553697	VSTM1	rs7895870	10	123098249		5.71	0.48	0.17	0.26	
VSTM1	ILMN_1763455	19	rs9625870	22	30261219		rs10500316	19	54553697	VSTM1	5.88	0.81	1.38	1.47	
WDR48	ILMN_1762103	3	rs1388935	4	188927822		rs6778963	3	39091812	WDR48	5.88	0.19	0.13	0.09	
WDR48	ILMN_1762103	3	rs1887778	9	134635088	RAPGEF1	rs883349	3	39067925	WDR48	6.34	0.57	1.35	1.22	
WDR48	ILMN_1762103	3	rs9554833	13	102624790		rs7619193	3	39044116	WDR48	5.85	0.18	0.61	0.35	
WDR6	ILMN_1669484	3	rs12362253	11	123371708		rs11715581	3	49194331	WDR6	4.86	1.64	1.43	2.25	
XAF1	ILMN_2370573	17	rs1533031	17	6673170	XAF1	rs12591171	15	93119799		5.48	2.38	0.17	1.63	
ZFP90	ILMN_1684628	16	rs909446	21	37040648		rs1182968	16	68573945	ZFP90	5.79	0.09	0.36	0.15	
ZNF500	ILMN_1700238	16	rs4823723	22	48283177		rs2290560	16	4799041	ZNF500	5.29	0.67	0.27	0.46	
ZYX	ILMN_1701875	7	rs6056281	20	8935312		rs2242601	7	143093824	ZYX	6.04	0.26	0.01	0.05	

[a] Phenotypes are expression levels of RefSeq Genes

[b] Illumina probe ID used to measure gene expression

[c] Physical SNP position in base pairs (HG19)

[d] RefSeq Gene ID of gene expression level that is influenced by the SNP (BSGS discovery dataset, significance threshold = 1.29×10^{-11})

[e] Interaction $-\log_{10}$ p-value from discovery dataset

[f] Interaction $-\log_{10}$ p-value from replication dataset

[g] Interaction $-\log_{10}$ p-value from meta analysis of replication datasets only

[h] Distance in Mb between interacting SNPs for *cis-cis* acting SNP pairs

[i] p-values are absent if the interaction did not pass the QC filtering in the replication dataset

[j] Meta analysis p-values are absent if the interaction did not pass the QC filtering in either replication dataset

Table S2: **Estimation of additive and non-additive variance components from pedigree information** Taken from previous analysis in Powell et al 2013[22]

		Additive		Non-additive	
Gene	Probe	Variance	s.e.	Variance	s.e.
NAPRT1	ILMN_1710752	0.37	0.03	0.14	0.05
TMEM149	ILMN_1786426	0.41	0.04	0.09	0.04
MBNL1	ILMN_2313158	0.18	0.03	0.11	0.04
TRAPPC5	ILMN_2372639	0.32	0.04	0.13	0.05
CAST	ILMN_1717234	0.31	0.03	0.10	0.04

Table S3: **Concordance of sign of epistatic variance components between discovery and replication datasets**

Test	Interactions[a]	Dataset	n[b]	Expected[c]	Observed[d]	p-value
1[e]	All	EGCUT	434	217.00	306	6.69×10^{-18}
		Fehrmann	434	217.00	278	5.04×10^{-9}
		Both	434	108.50	221	5.56×10^{-31}
	Significant	EGCUT	30	15.00	25	3.25×10^{-4}
		Fehrmann	30	15.00	24	1.43×10^{-3}
		Both	30	7.50	22	3.76×10^{-8}
2[f]	All	EGCUT	434	54.25	92	4.22×10^{-7}
		Fehrmann	434	54.25	79	6.18×10^{-4}
		Both	434	6.78	30	2.55×10^{-11}
	Significant	EGCUT	30	3.75	19	9.46×10^{-11}
		Fehrmann	30	3.75	19	9.46×10^{-11}
		Both	30	0.47	18	2.23×10^{-25}
3[g]	All	EGCUT	1133	566.50	775	7.10×10^{-36}
		Fehrmann	1133	566.50	726	1.90×10^{-21}
		Both	1133	283.25	562	1.39×10^{-70}
	Significant	EGCUT	73	36.50	55	1.69×10^{-5}
		Fehrmann	73	36.50	55	1.69×10^{-5}
		Both	73	18.25	46	7.86×10^{-12}

[a] "All" denotes 434 discovery interactions and "Significant" denotes 30 interactions with significant replication p-values

[b] Number of tests for concordance

[c] Expected number of concordant cases under the null hypothesis of no interactions

[d] Observed number of concordant cases

[e] The sign of the most significant epistatic variance component in discovery is the same as the corresponding variance component in the replication data.

[f] The largest epistatic variance component in the discovery is the same as in the replication with the same sign in both.

[g] The sign of all epistatic variance components in the discovery with $p < 0.05$ are the same as the corresponding variance components in the replication data.

Table S4: **Concordance of sign of epistatic variance components between discovery and replication datasets using test 4**

Interactions[a]	Dataset	n[b]	0[c]	1[c]	2[c]	3[c]	4[c]	p
Expected[d]	-	-	0.06	0.25	0.38	0.25	0.06	-
All	EGCUT	434	0.06	0.22	0.41	0.23	0.08	0.194
All	Fehrmann	434	0.07	0.22	0.39	0.24	0.08	0.385
All	Combined	868	0.07	0.22	0.40	0.23	0.08	0.0448
Significant	EGCUT	30	0.07	0.03	0.30	0.33	0.27	4.72×10^{-4}
Significant	Fehrmann	30	0.03	0.07	0.33	0.27	0.30	6.69×10^{-4}
Significant	Combined	60	0.05	0.05	0.32	0.30	0.28	5.49×10^{-8}

[a] "All" denotes 434 discovery interactions and "Significant" denotes 30 interactions with significant replication p-values.

[b] Number of tests for concordance.

[c] Proportion of tests that have 0, 1, 2, 3 or 4 concordant signs between discovery and replication.

[d] Expected proportion of concordant signs under the null hypothesis of no epistasis.

Table S5: **Details on linkage disequilibrium and relative positions of all discovery interactions with SNPs on the same chromosome**

Chr	Gene	SNP 1	SNP 2	Position 1	Position 2	Distance / Mb	R^2	D'
19	TMEM149	rs807491	rs7254601	36268923	36147315	0.122	0.000	0.001
17	FN3KRP	rs898095	rs9892064	80890638	80827903	0.063	0.063	0.088
21	CSTB	rs9979356	rs3761385	45230974	45198355	0.033	0.041	0.066
3	MBNL1	rs16864367	rs13079208	152234166	152116652	0.118	0.041	0.117
10	ADK	rs2395095	rs10824092	76446305	75929517	0.517	0.013	0.020
11	CTSC	rs7930237	rs556895	88117962	88077479	0.040	0.012	0.045
17	GAA	rs11150847	rs12602462	78153130	78146016	0.007	0.000	0.001
8	NAPRT1	rs2123758	rs3889129	144663661	144613680	0.050	0.053	0.060
1	LAX1	rs1891432	rs10900520	203877662	203780591	0.097	0.065	0.106
18	MBP	rs8092433	rs4890876	74747424	74732087	0.015	0.035	0.053
11	SNORD14A	rs2634462	rs6486334	17339127	17015557	0.324	0.008	0.012
21	C21ORF57	rs9978658	rs11701361	48027084	47764477	0.263	0.032	0.065
16	RPL13	rs352935	rs2965817	89648580	89513234	0.135	0.054	0.060
19	ATP13A1	rs4284750	rs873870	19810050	19738554	0.071	0.008	0.015
2	NCL	rs7563453	rs4973397	232301670	232291471	0.010	0.027	0.029
5	HNRPH1	rs6894268	rs4700810	179032488	178991794	0.041	0.000	0.001
19	VASP	rs1264226	rs2276470	46063167	45974668	0.088	0.018	0.022
7	TRA2A	rs7776572	rs11770192	23528927	23498358	0.031	0.064	0.064
21	PRMT2	rs2839372	rs11701058	48063862	47776382	0.287	0.100	0.122
12	OAS1	rs13311	rs2072133	113448652	113409260	0.039	0.002	0.016
16	N4BP1	rs12444224	rs11649236	87580855	48632478	38.948	0.007	0.021
5	CAST	rs12719343	rs7733671	125369113	96000269	29.369	0.001	0.001
7	DNAJB6	rs2286842	rs3779589	157216093	157163614	0.052	0.005	0.006
1	OVGP1	rs10802822	rs1264898	240132968	111992823	128.140	0.008	0.030
20	CD93	rs2868504	rs1884655	37771578	23074375	14.697	0.000	0.002
11	PHCA	rs493642	rs10736812	123097386	76708086	46.389	0.002	0.008
21	MX1	rs459498	rs8130120	42795027	29363604	13.431	0.000	0.000
16	AKTIP	rs2896940	rs13332406	57721127	53489705	4.231	0.000	0.001
17	CDK5R1	rs9905940	rs11655031	46614102	30833162	15.781	0.000	0.000
2	CYBRD1	rs888427	rs7591849	172368120	160112881	12.255	0.000	0.000
8	HMBOX1	rs587639	rs7837237	132725731	28876221	103.850	0.001	0.001
11	TRAPPC4	rs1793823	rs3916581	131018917	118887887	12.131	0.001	0.002
12	PEX5	rs10444467	rs4329748	128052636	7364442	120.688	0.000	0.000
12	FLJ20489	rs17615703	rs3782908	117036766	48169526	68.867	0.001	0.002
16	PRKCB1	rs2188355	rs10492793	23867776	12639800	11.228	0.000	0.000
14	MRPL52	rs1950857	rs3811188	26710271	23299135	3.411	0.002	0.004
17	C17ORF60	rs9907897	rs7405659	63502633	59874129	3.629	0.004	0.011
6	FLJ43093	rs6906101	rs13214069	36667610	32705248	3.962	0.000	0.000
19	TRAPPC5	rs17159840	rs17763599	7758194	2369415	5.389	0.000	0.000
22	PISD	rs715572	rs6518754	33234931	32097775	1.137	0.001	0.003
12	DIP2B	rs871257	rs12427378	117994348	51074199	66.920	0.001	0.001
12	GPR162	rs2272500	rs2707210	79685913	6902002	72.784	0.003	0.005
17	USP36	rs2279308	rs7225546	76794981	75151717	1.643	0.000	0.000