

1 Detection and replication of epistasis influencing
2 transcription in humans

3 Gibran Hemani^{1,2,*}, Konstantin Shakhbazov^{1,2}, Harm-Jan Westra³,
4 Tonu Esko^{4,5,6}, Anjali K Henders⁷, Allan F McRae^{1,2}, Jian Yang²,
5 Greg Gibson⁸, Nicholas G Martin⁷, Andres Metspalu⁴, Lude
6 Franke³, Grant W Montgomery^{7,+}, Peter M Visscher^{1,2,+}, and
7 Joseph E Powell^{1,2,+}

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

Abstract

Epistasis is the phenomenon whereby a polymorphism's effect on a trait depends on other polymorphisms present in the genome. The extent to which epistasis influences complex traits¹ and contributes to their variation^{2,3} is a fundamental question in evolution and human genetics. Though epistasis has been demonstrated in artificial gene manipulation studies in model organisms,^{4,5} and examples have been reported in other species,⁶ few convincing examples with independent replication exist for epistasis amongst natural polymorphisms in human traits.^{7,8} Its absence from empirical findings may simply be due to its 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 power and computational issues.⁹ Here we show that, using advanced computation techniques¹⁰ 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 data on 7339 gene expression levels in peripheral blood, we found 501 significant pairwise epistatic interactions between common SNPs acting on the expression levels 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 $p < 0.05/501$. There was evidence of functional enrichment for the interacting SNPs, for instance 44 of the genetic interactions are located within 5Mb of regions of known physical chromosome interactions¹³ ($p = 1.8 \times 10^{-10}$). 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 mutations⁴ but also at the evolutionary scale in fitness adaptation¹⁵ 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 re-

duced 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 testing⁹ all make it more difficult to detect epistasis in comparison to additive effects. Thus, when combined 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, BSGS²³) 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 (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, Fehrmann¹² 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 ($A1$, $A2$, $D1$, $D2$, $A \times A$, $A \times D$, $D \times A$, $D \times D$) and tested for concordance of the sign of the most significant 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

113 observed that 316 of the remaining 404 discovery SNP pairs had replication
 114 interaction p -values more extreme than the 2.5% confidence interval of the
 115 quantile-quantile plot against the null hypothesis of no interactions where p -
 116 values are assumed to be uniformly distributed ($p \ll 1.0 \times 10^{-16}$, Figure 2 and
 117 Supplementary Figure S2). Concordance of the direction of the effect of the
 118 largest variance component was also highly significant ($p = 5.71 \times 10^{-31}$, Sup-
 119plementary Table S3). The congruence of the epistatic networks in discovery
 120 and replication datasets is shown in Figure 3, demonstrating that these com-
 121 plex genetic patterns are common even across independent datasets. A further
 122 replication was attempted using the Centre for Health Discovery and Wellbeing
 123 (CHDWB) dataset,²⁴ but only 20 of the SNP pairs passed filtering because the
 124 sample size was small ($n = 139$), and likely due to insufficient power we found
 125 no evidence for replication (Supplementary Figure S6).

126 It should be noted that although it is a necessary step to establish the
 127 veracity of the interactions from the discovery set, replication of epistasis is
 128 difficult in practice. For example, LD between causal variants and observed
 129 markers plays an important role. Not only is the dependence on LD much
 130 greater for epistatic effects than for additive effects (Supplementary Figure S7),
 131 but when estimating epistatic variance it is more sensitive to changes in LD
 132 between observed SNPs and causal variants between independent samples when
 133 compared to additive effects (Supplementary Figure S8). This has a direct effect
 134 on statistical power for replication. The sampling variance of LD r leads to the
 135 ascertainment of marker associations with higher sample r in the discovery stage
 136 in comparison to the replication stage. However, the average decrease in \hat{r}^x in
 137 replication samples becomes larger as x increases (Methods, Supplementary
 138 Figure S9). For example, the decrease in \hat{r}^8 (which is proportional to the power
 139 of detecting $D \times D$ effects), is on average three fold greater than the decrease in
 140 \hat{r}^2 (which is proportional to the power of detecting additive effects).

141 Though seldom the focus of association studies, SNPs with known main
 142 effects are often tested for additive \times additive genetic interactions,⁹ but our
 143 analysis shows that this is unlikely to be the most effective strategy for its
 144 detection. The majority of our discovery interactions comprised of one SNP
 145 that was significantly associated with the gene expression level in the discov-
 146 ery dataset, and one SNP that had no previous association²² (439 out of 501,
 147 Methods). Only nine interactions were between SNPs that both had known
 148 main effects while 64 were between SNPs that had no known main effects. Ad-
 149 ditionally, we observed that the largest epistatic variance component for the
 150 501 interactions was equally divided amongst additive \times additive, additive \times
 151 dominance, dominance \times additive and dominance \times dominance at the discovery
 152 stage ($p = 0.22$ for departure from expectation). This is not surprising because
 153 the patterns of epistasis used for statistical decomposition (*i.e.* $A \times A$, $A \times D$,
 154 $D \times A$, $D \times D$) are simply convenient orthogonal parameterisations of a two
 155 locus model, and are not intended to model biological function.²⁵

156 Of the discovery interactions, 26 were *cis-cis* acting (within 1Mb of the
 157 transcription start site, mean distance between SNPs was 0.53Mb), 462 were
 158 *cis-trans*-acting, and 13 were *trans-trans*-acting. We observed a wide range of

159 significant GP maps (Figure 1) but the most common pattern of epistasis that
 160 we detected involved a *trans*-SNP masking the effect of an additive *cis*-SNP. For
 161 example, MBNL1 (involved in RNA modification and regulation of splicing²⁶)
 162 has a *cis* effect at rs13069559 which in turn is controlled by 13 *trans*-SNPs and
 163 one *cis*-SNP that each exhibit a masking pattern, such that when the *trans*-
 164 SNP is homozygous for the masking allele the decreasing allele of the *cis*-SNP
 165 no longer has an effect (Supplementary Figure S10). Each of these interac-
 166 tions has evidence for replication in at least one dataset and six are significantly
 167 replicated at the Bonferroni level (Supplementary Figure S3). We see similar
 168 epistatic networks involving multiple (eight or more) *trans*-acting SNPs for other
 169 gene expression levels too, for example TMEM149 (Supplementary Figure S11),
 170 NAPRT1 (Supplementary Figure S12), TRAPPC5 (Supplementary Figure S13),
 171 and CAST (Supplementary Figure S14). We observed that from pedigree anal-
 172 ysis these five gene expression phenotypes had non-additive variance component
 173 estimates within the 95th percentile of the 17,994 gene expression phenotypes
 174 that were analysed previously²² (Supplementary Table S2, Methods).

175 In total the 501 interactions comprised 781 unique SNPs, which we analysed
 176 for functional enrichment (Methods). We tested the SNPs for cell-type specific
 177 overlap with transcriptionally active chromatin regions, tagged by histone-3-
 178 lysine-4,tri-methylation (H3K4me3) chromatin marks, in 34 cell types²⁷ (Sup-
 179 plementary Figure S5). There was significant enrichment for *cis*-acting SNPs
 180 in haematopoietic cell types only ($p < 1 \times 10^{-4}$ for the three tissues with the
 181 strongest enrichment after adjusting for multiple testing). However *trans*-acting
 182 SNPs did not show any tissue specific enrichment ($p > 0.1$ for all tissues). This
 183 difference between *cis* and *trans* SNPs suggests different roles in epistatic in-
 184 teractions where tissue specificity is provided by the *cis* SNPs. There is also
 185 enrichment for *cis*-SNPs to be localised in regions with regulatory genomic fea-
 186 tures as measured by chromatin states²⁸ (Supplementary Figure S4).

187 We also demonstrate physical organisation of interacting loci within the cell,
 188 suggesting a mechanism by which biological function can lead to epistatic ge-
 189 netic variance. It has been shown that different chromosomal regions spatially
 190 colocalise in the cell through chromatin interactions.¹³ We cross-referenced our
 191 epistatic SNPs with a map of chromosome interacting regions ($n = 96,139$)
 192 in K562 blood cell lines²⁹ (Methods) and found that 44 epistatic interactions
 193 mapped to within 5Mb ($p < 1.8 \times 10^{-10}$), (Supplementary Figure S15). Inter-
 194 action of distant loci may occur through physical proximity in transcriptional
 195 factories that organise across different chromosome regions and can regulate
 196 transcription of related genes.^{30,31}

197 Though we present many instances of epistasis, quantifying its relative im-
 198 portance to complex traits in humans remains an open question. In this study
 199 we are able to identify 238 gene expression traits with at least one significant
 200 interaction given our experiment-wide threshold, where the minimum estimated
 201 variance explained by the epistatic effects of any interaction was 2.1% of phe-
 202 notypic variance. Taking results from our previously published eQTL²³ we
 203 calculated that 1848 of the 7339 gene expression levels analysed were influenced
 204 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 Ideally one would approach this question from a whole genome perspective³² but this is intractable for non-additive variance components. Nevertheless, some inference can be made from the ascertained effects in these analyses and it is evident that estimated additive variance is overall a larger component than estimated epistatic variance, as has been argued previously.^{2,3} 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. 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. This is due to differences in the rate of decay of the estimate of the genetic variance of the causal SNPs as LD decreases with the observed SNPs. And forthly, the extent of winner’s curse in estimation of effect sizes may differ between the the two studies.

Overall, we have demonstrated that it is possible to identify and replicate epistasis in complex traits amongst common human variants, despite the relative 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 47,323 probes. Only the probes that passed quality control and had significant expression in $\geq 90\%$ of individuals were used in the analysis (7,339 probes representing 6,158 RefSeq genes). Recent hardware and software¹⁰ advances that use graphics processing units (GPUs) made it possible to perform the 1.03×10^{15} statistical tests to complete this analysis. We used permutation analysis³³ to calculate an experiment-wide significance threshold of $T_e = 2.91 \times 10^{-16}$ at the 5% family-wise error rate (FWER). SNP pairs were modelled for

249 full genetic effects, including marginal additive and dominance at both SNPs
 250 plus four interaction terms. Though we could have used a less complex model to
 251 improve statistical efficiency, we deemed it important to be agnostic about the
 252 type of epistasis that might exist, and therefore chose not to over-parameterise
 253 the test.^{18,19} Because there are many large marginal effects present in these data
 254 it was necessary to perform several filtering steps to exclude SNP pairs that were
 255 significant due to marginal effects alone. All SNP pairs with LD $r^2 > 0.1$ and
 256 $D'^2 > 0.1$ were removed to minimise the possibility of haplotype effects. All
 257 SNP pairs were required to have at least five data points in all nine genotype
 258 classes. If multiple SNP pairs were present on the same chromosomes for a
 259 particular expression trait then only the sentinel SNP pair was retained. Finally,
 260 a nested test contrasting the full genetic model against the marginal additive
 261 and dominance model was performed for each remaining SNP pair (Methods),
 262 resulting in 501 significant interactions after Bonferroni correction for multiple
 263 testing of the filtered SNPs. The 501 significant SNP pairs were carried forward
 264 for replication in two independent datasets that used the same expression assays
 265 for analysing transcription in peripheral blood, the Fehrmann dataset¹² ($n =$
 266 1240) and the Estonian Genome Centre University of the University of Tartu
 267 (EGCUT) dataset¹¹ ($n = 891$). Of these, 434 passed filtering in both replication
 268 datasets. A meta analysis on the interaction p -values from each replication
 269 dataset was performed to provide an overall replication statistic for each putative
 270 interaction.

271 Acknowledgements

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

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

281 The University of Queensland group is supported by the Australian National
 282 Health and Medical Research Council (NHMRC) grants 389892, 496667,
 283 613601, 1010374 and 1046880, the Australian Research Council (ARC) grant
 284 (DE130100691), and by National Institutes of Health (NIH) grants GM057091
 285 and GM099568.

286 The QIMR researchers acknowledge funding from the Australian National
 287 Health and Medical Research Council (grants 241944, 389875, 389891, 389892,
 288 389938, 442915, 442981, 496739, 496688 and 552485), the and the National In-
 289 stitutes of Health (grants AA07535, AA10248, AA014041, AA13320, AA13321,
 290 AA13326 and DA12854). We thank Anthony Caracella and Lisa Bowdler for
 291 technical assistance with the micro-array hybridisations.

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

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

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

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

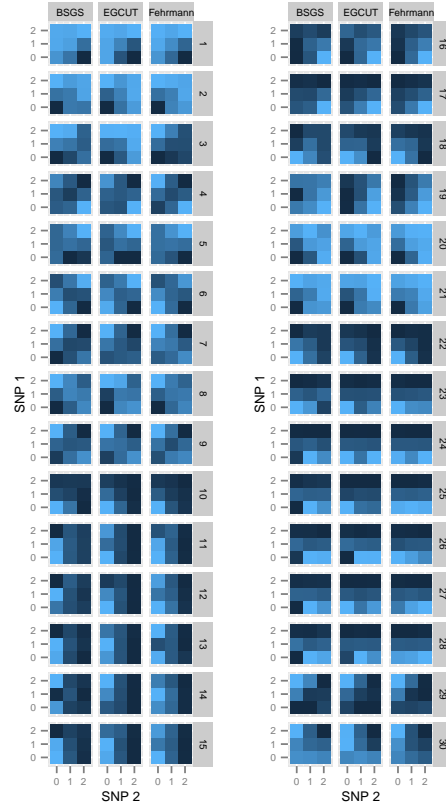


Figure 1: Replication of GP maps in two independent populations
The GP maps for each epistatic interaction that is significant at the Bonferroni level in both replication datasets are shown. Each GP map consists of nine tiles where each tile represents the expression level for that two-locus genotype 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.

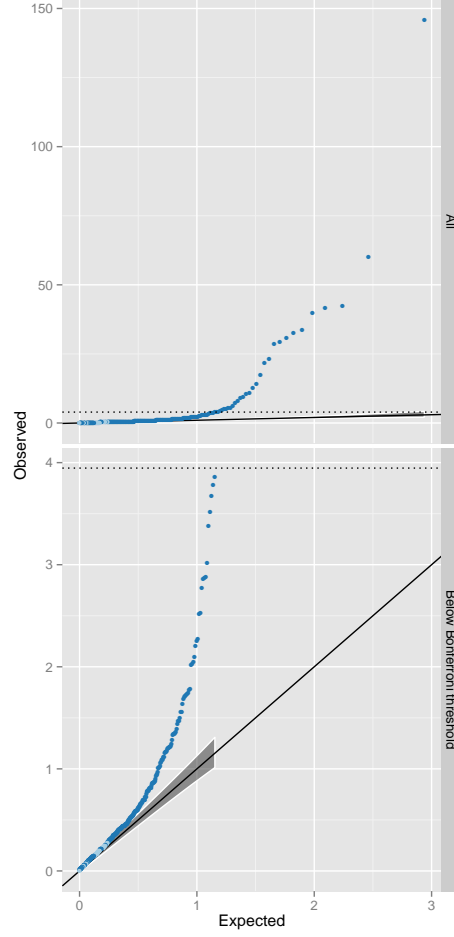


Figure 2: **Q-Q plots of interaction p -values from replication datasets** The top panel shows all 434 discovery SNPs that were tested for interactions. Observed p -values (y -axis, $-\log_{10}$ scale) are plotted against the expected p -values (x -axis, $-\log_{10}$ scale). The multiple testing correction 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.

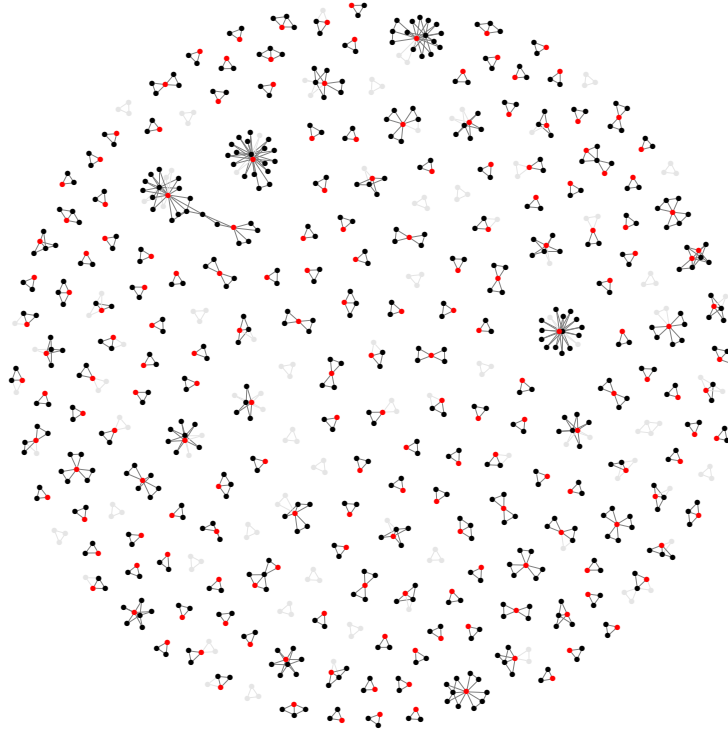


Figure 3: **Discovery and replication of epistatic networks** All 434 putative genetic interactions (edges) with data common to discovery and replication sets is shown, where black nodes represent SNPs and red nodes represent traits (gene expression probes). Three hundred and forty-five interactions 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 (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** (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 (2010).
- ⁴ Costanzo, M. *et al.* The genetic landscape of a cell. *Science (New York, N.Y.)* **327**, 425–31 (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 (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 (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 (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** (2011).
- ⁹ Cordell, H. J. Detecting gene-gene interactions that underlie human diseases. *Nature Reviews Genetics* **10**, 392–404 (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 (2011).
- ¹¹ Metspalu, A. The Estonian Genome Project. *Drug Development Research* **62**, 97–101 (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 (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 (2009).

- 350 ¹⁴ Visscher, P. M., Brown, M. a., McCarthy, M. I. & Yang, J. Five years of
351 GWAS discovery. *American journal of human genetics* **90**, 7–24 (2012).
- 352 ¹⁵ Weinreich, D. M., Delaney, N. F., Depristo, M. a. & Hartl, D. L. Darwinian
353 evolution can follow only very few mutational paths to fitter proteins. *Science*
354 (*New York, N.Y.*) **312**, 111–4 (2006).
- 355 ¹⁶ Breen, M. S., Kemena, C., Vlasov, P. K., Notredame, C. & Kondrashov, F. a.
356 Epistasis as the primary factor in molecular evolution. *Nature* **490**, 535–538
357 (2012).
- 358 ¹⁷ Weir, B. S. Linkage disequilibrium and association mapping. *Annual review*
359 *of genomics and human genetics* **9**, 129–42 (2008).
- 360 ¹⁸ Hemani, G., Knott, S. & Haley, C. An Evolutionary Perspective on Epistasis
361 and the Missing Heritability. *PLoS Genetics* **9**, e1003295 (2013).
- 362 ¹⁹ Marchini, J., Donnelly, P. & Cardon, L. R. Genome-wide strategies for de-
363 tecting multiple loci that influence complex diseases. *Nature Genetics* **37**,
364 413–417 (2005).
- 365 ²⁰ Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and
366 biological pathways affect human height. *Nature* **467**, 832–8 (2010).
- 367 ²¹ Schadt, E. *et al.* Genetics of gene expression surveyed in maize, mouse and
368 man. *Nature* **422**, 297–302 (2003).
- 369 ²² Powell, J. E. *et al.* Congruence of Additive and Non-Additive Effects on
370 Gene Expression Estimated from Pedigree and SNP Data. *PLoS Genetics* **9**,
371 e1003502 (2013).
- 372 ²³ Powell, J. E. *et al.* The Brisbane Systems Genetics Study: genetical genomics
373 meets complex trait genetics. *PloS one* **7**, e35430 (2012).
- 374 ²⁴ Preininger, M. *et al.* Blood-informative transcripts define nine common axes
375 of peripheral blood gene expression. *PLoS genetics* **9**, e1003362 (2013).
- 376 ²⁵ Cockerham, C. C. An extension of the concept of partitioning hereditary
377 variance for analysis of covariances among relatives when epistasis is present.
378 *Genetics* **39**, 859–882 (1954).
- 379 ²⁶ Ho, T. H. *et al.* Muscleblind proteins regulate alternative splicing. *The EMBO*
380 *journal* **23**, 3103–12 (2004).
- 381 ²⁷ Trynka, G. *et al.* Chromatin marks identify critical cell types for fine mapping
382 complex trait variants. *Nature genetics* **45**, 124–30 (2013).
- 383 ²⁸ Hoffman, M., Buske, O., Wang, J. & Weng, Z. Unsupervised pattern dis-
384 covery in human chromatin structure through genomic segmentation. *Nature*
385 *Methods* **9**, 473–476 (2012).

- 386 ²⁹ Lan, X. *et al.* Integration of Hi-C and ChIP-seq data reveals distinct types
387 of chromatin linkages. *Nucleic acids research* **40**, 7690–704 (2012).
- 388 ³⁰ Osborne, C. S. *et al.* Active genes dynamically colocalize to shared sites of
389 ongoing transcription. *Nature genetics* **36**, 1065–71 (2004).
- 390 ³¹ Rieder, D., Trajanoski, Z. & McNally, J. G. Transcription factories. *Frontiers*
391 *in genetics* **3**, 221 (2012).
- 392 ³² Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era—
393 concepts and misconceptions. *Nature Reviews Genetics* **9**, 255–66 (2008).
- 394 ³³ Churchill, G. A. & Doerge, R. W. Empirical threshold values for quantitative
395 trait mapping. *Genetics* **138**, 963–71 (1994).

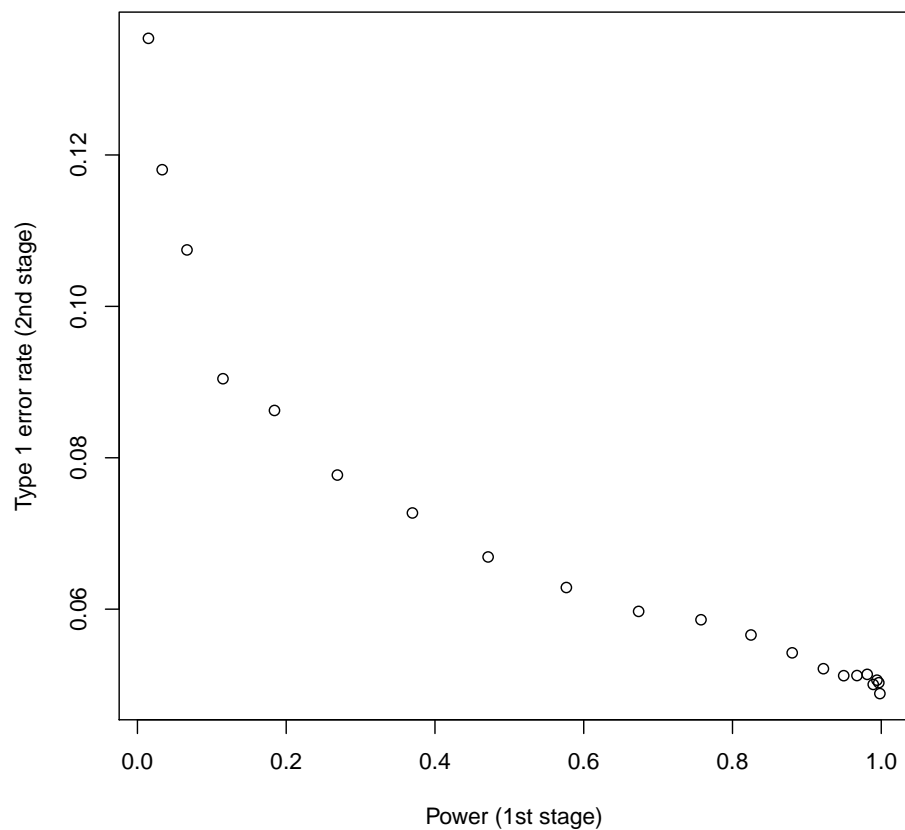


Figure S1: **Type 1 error rate of two stage design assuming a null model of one large additive effect and no epistasis** In stage 1 SNPs are tested for full genetic effects (8 d.f.) and those that surpass a threshold for multiple testing are then tested for significant interaction terms in stage 2. These 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.

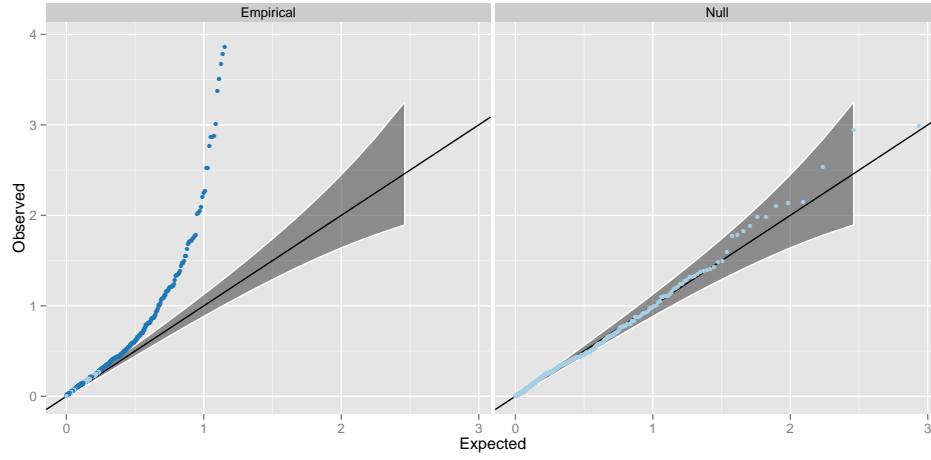


Figure S2: **Q-Q plots of interaction p -values from replication datasets, excluding the 30 points significant at the Bonferroni level** The right panel (Null) shows the interaction p -values from a meta analysis across two independent datasets on 434 randomly drawn SNP pairs. The left panel (Empirical) 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.

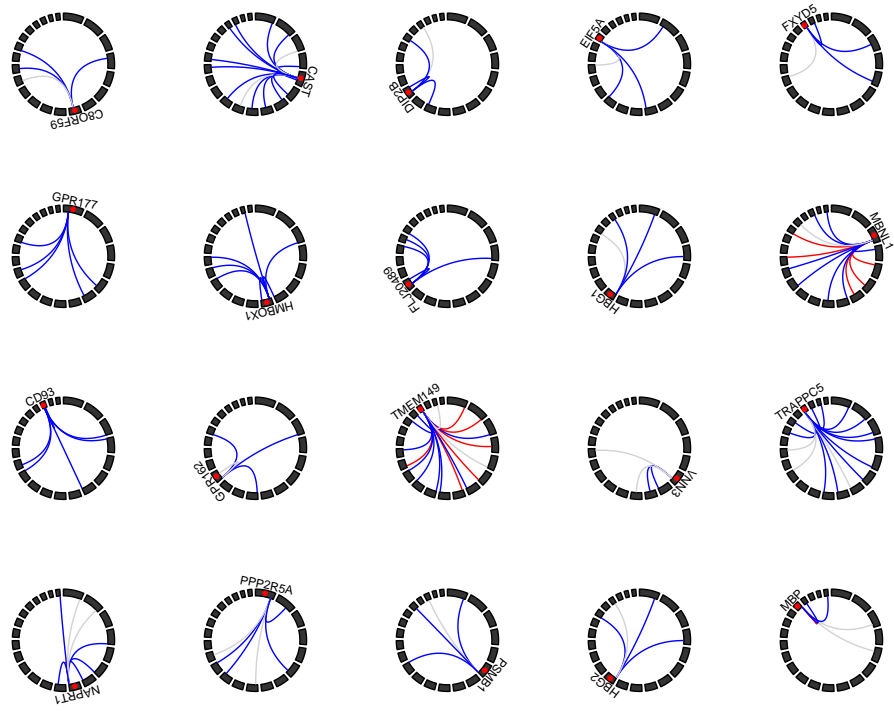


Figure S3: **Gene expression traits with four or more genetic interactions** Circle plots represent the genomic positions for SNPs (linking lines) and expression probes (red points). Chromosomes are represented by black blocks and ordered from 1 to 22 clockwise, starting from the top. Grey lines represent 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.

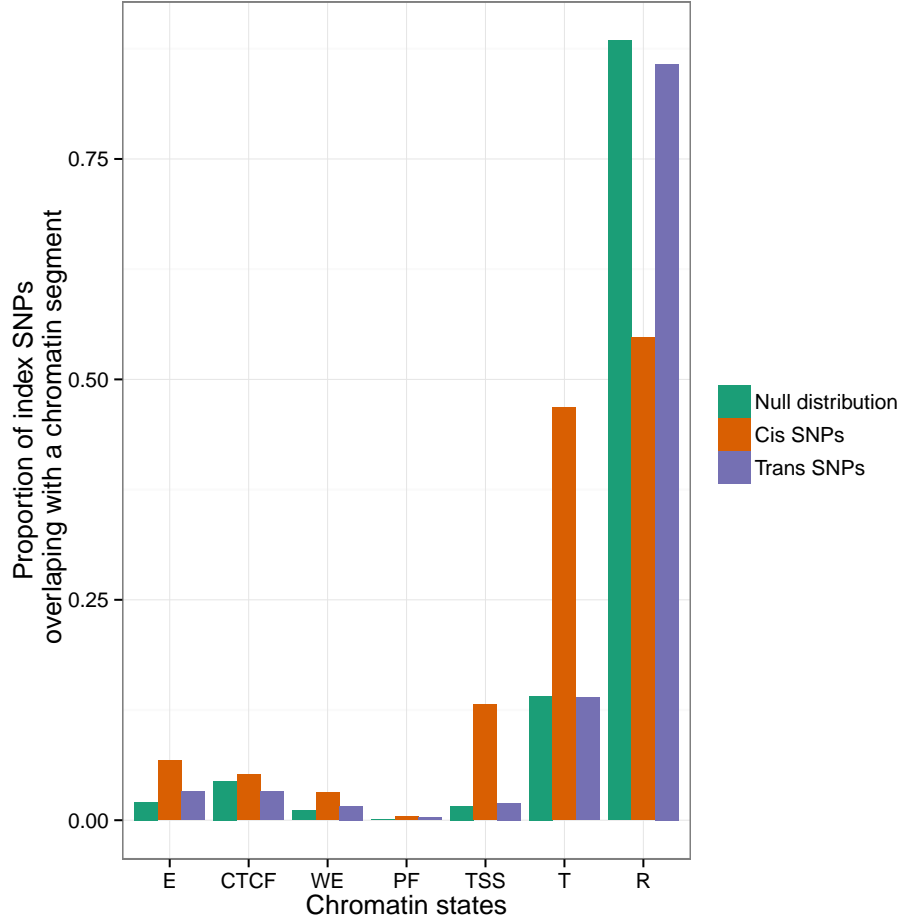


Figure S4: **Location of SNPs relative to genomic features** We used chromatin segmentation²⁸ as a method for labelling genomic features. All SNPs within 1Mb and $r^2 > 0.8$ of each *cis*- and *trans*-SNP were taken to find which genomic features (x -axis) were covered by the SNPs that compose the 501 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

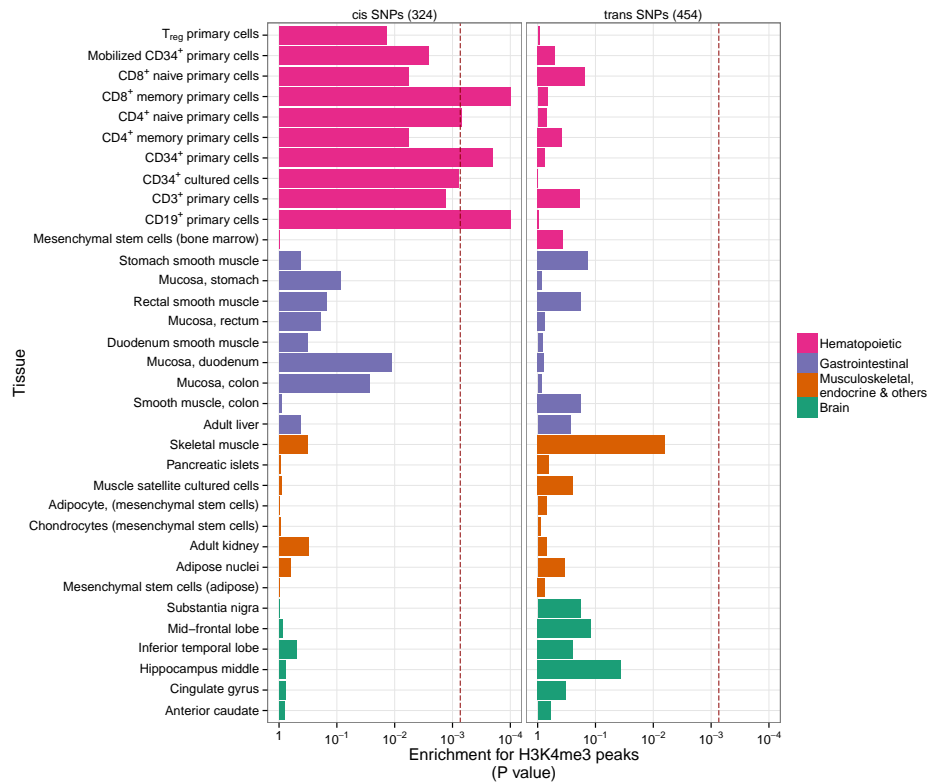


Figure S5: Tissue specific enrichment of SNPs in transcriptionally active regions The locations of transcriptional activity can be predicted by chromatin marks, assayed by H3K4me3.²⁷ Enrichment p -values are calculated using permutation analysis for 34 different cell types (y -axis) in four tissue types (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.

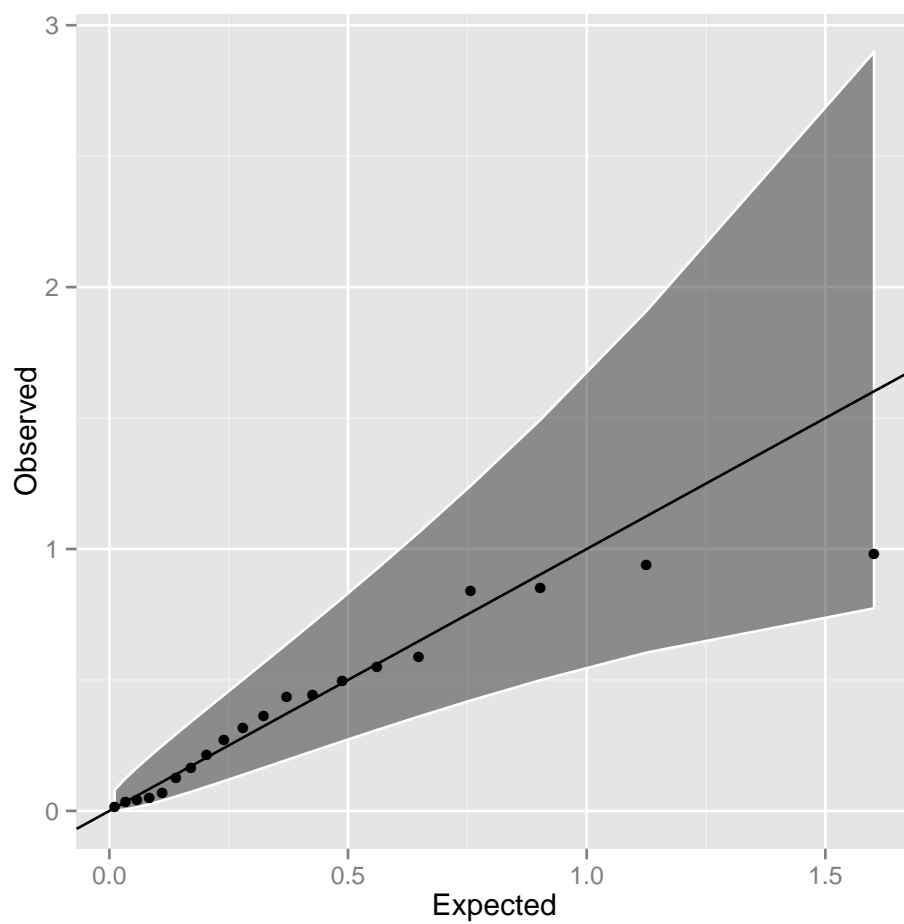


Figure S6: **Q-Q plot of interaction p -values in the CDHWB dataset**
 Twenty of the 501 discovery SNP pairs passed filtering in the CDHWB dataset (mainly due to small sample size). There is no evidence for enrichment of interaction terms, most likely due to insufficient power given the limited sample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

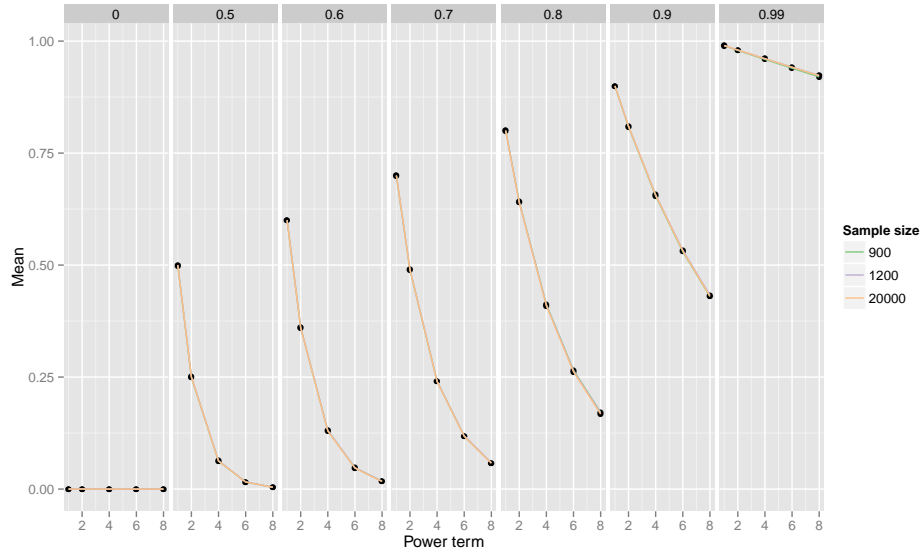


Figure S7: **Sampling mean for different power terms of population r values** Power of detection and replication of epistatic interactions depends not on r^2 between causal variants and observed SNPs, but on r^4, r^6, r^8 . For a given population value of LD r (columns of plots), plotted is the sample mean (y -axis) of \hat{r} , \hat{r}^2 (additive), \hat{r}^4 (dominance, $A \times A$), \hat{r}^6 ($A \times D$), \hat{r}^8 ($D \times D$) (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 .

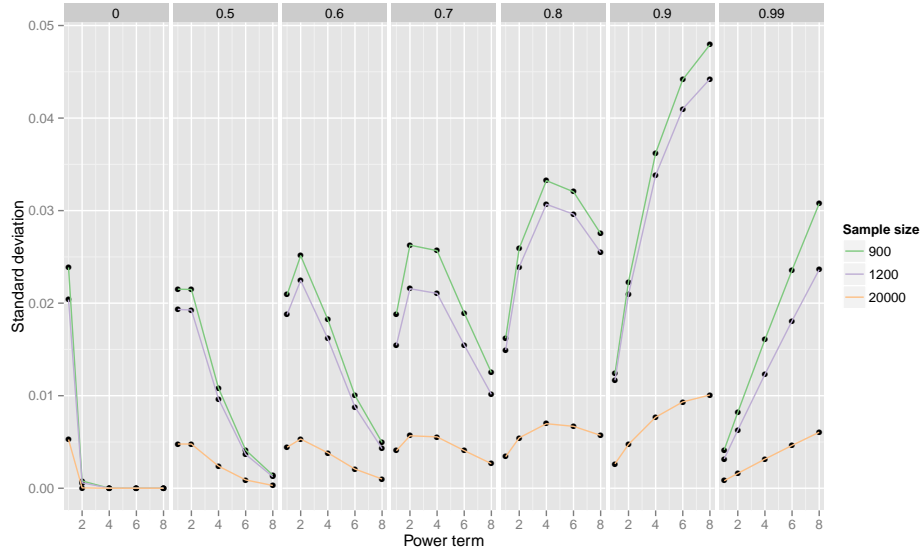


Figure S8: Sampling standard deviation for different power terms of population r values Power of detection and replication of epistatic interactions depends not on r^2 between causal variants and observed SNPs, but on r^4, r^6, r^8 . For a given a population value of LD r (columns of plots), plotted is the sampling standard deviation (y -axis) of \hat{r} , \hat{r}^2 (additive), \hat{r}^4 (dominance, $A \times A$), \hat{r}^6 ($A \times D$), \hat{r}^8 ($D \times D$) (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.

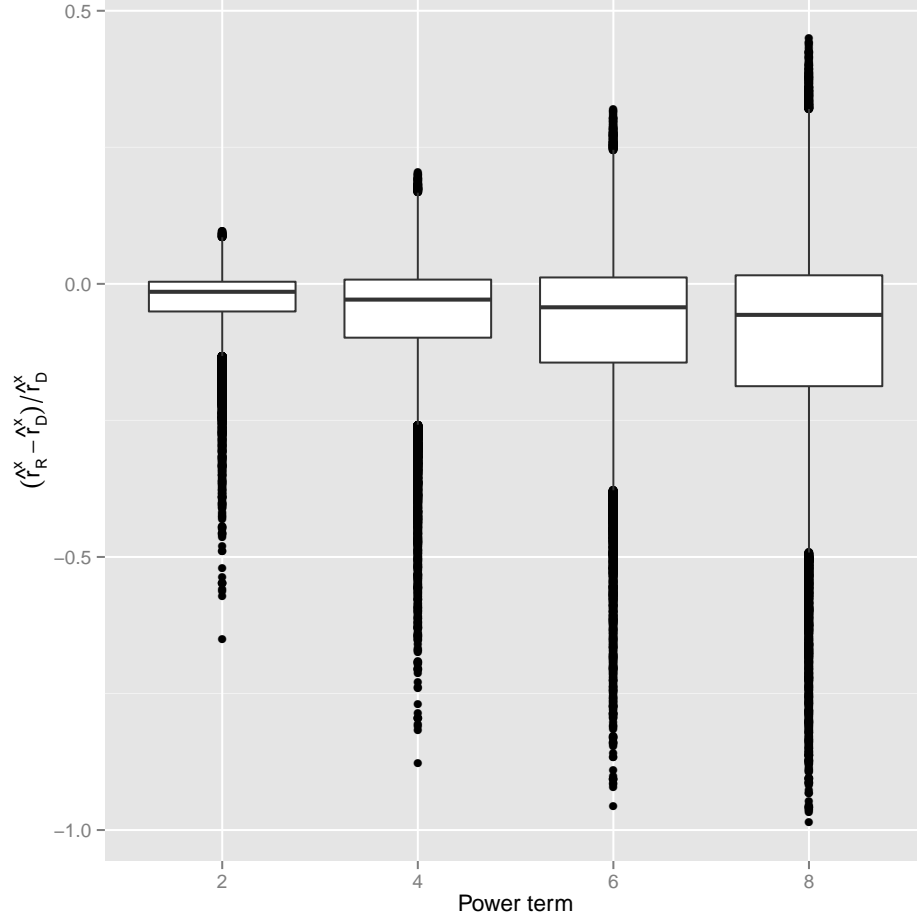


Figure S9: **Reduction in LD as estimated in replication data after ascertaining for high LD in discovery data** 100,000 “unobserved” causal variants (CVs) were tested for LD against a panel of 528,509 “observed” discovery markers (DMs). DM/CV pairs with LD $r^2 > 0.9$ were then tested in an independent 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.

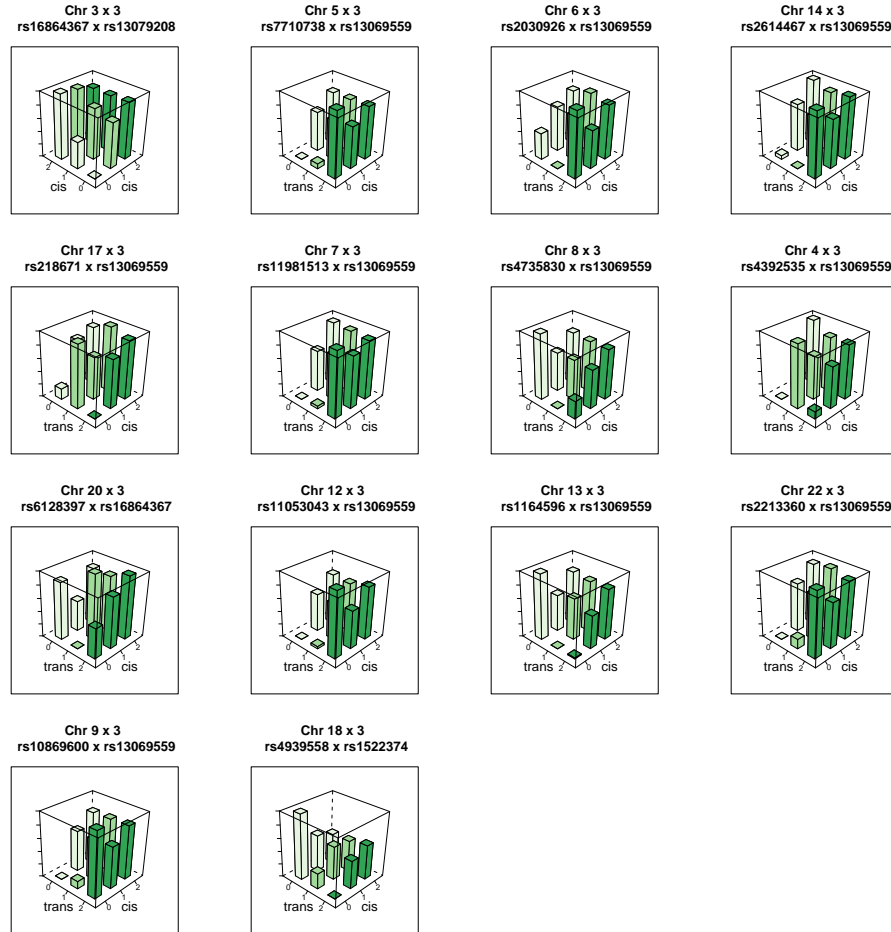


Figure S10: **Genotype-phenotype maps for 14 interactions influencing the expression of MBNL1** Each bar represents the mean phenotypic value for individuals in that genotype class. The rs13069559 SNP typically has a *cis*-additive decreasing effect on the expression of MBNL1, but in many of these interactions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

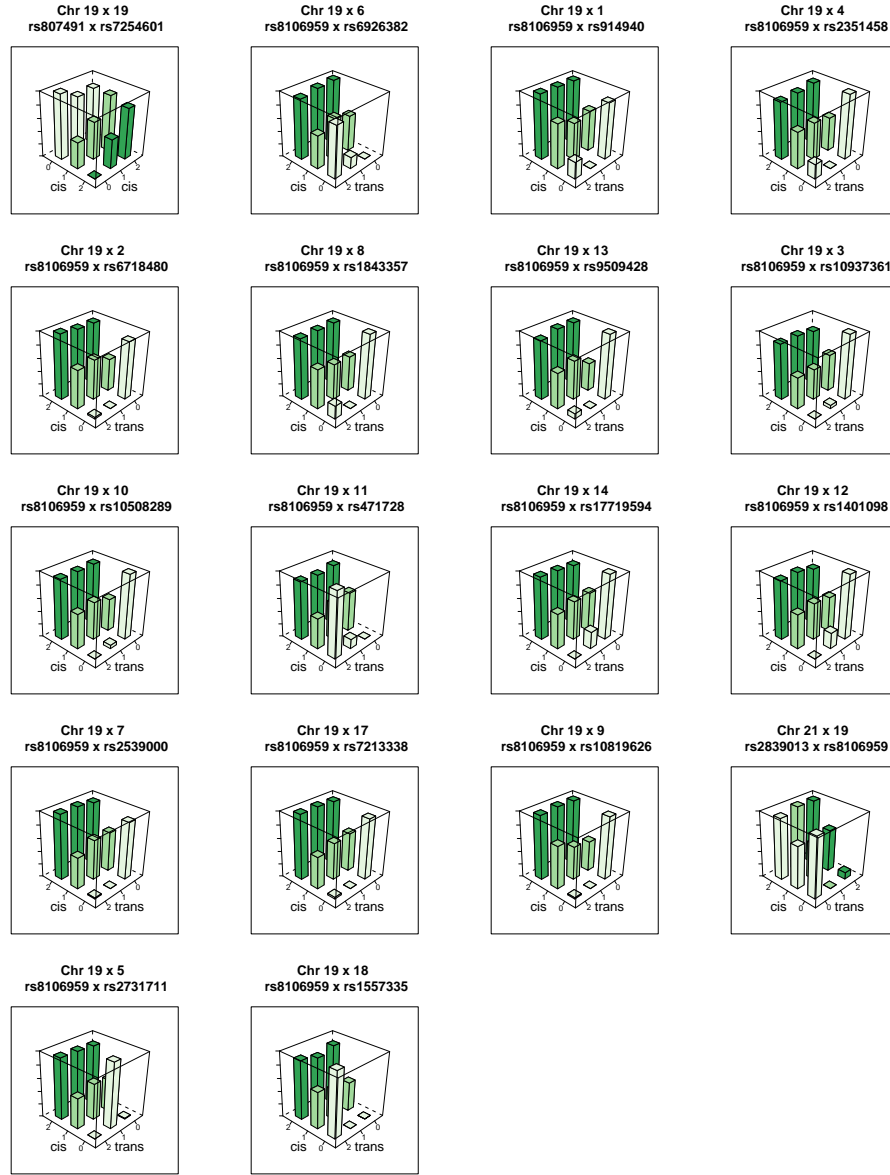


Figure S11: **Genotype-phenotype maps for 19 interactions influencing the expression of TMEM149** Each bar represents the mean phenotypic value for individuals in that genotype class. The rs13069559 SNP typically has a *cis*-additive decreasing effect on the expression of TMEM149, but in many of these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

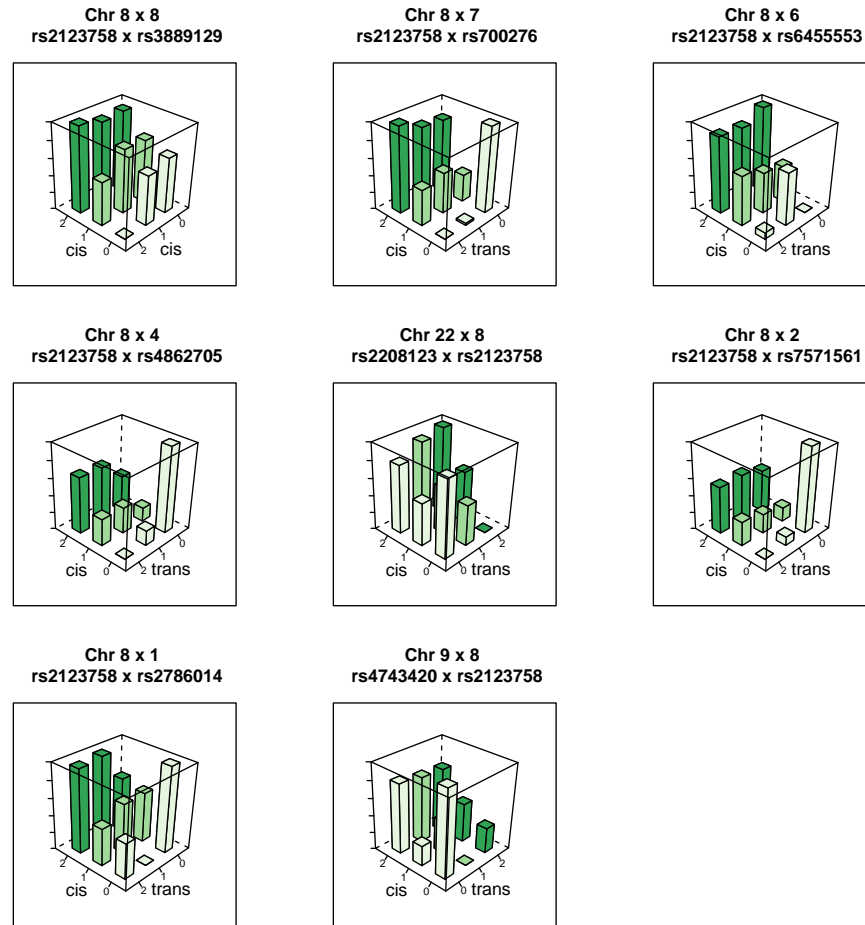


Figure S12: **Genotype-phenotype maps for 8 interactions influencing the expression of NAPRT1** Each bar represents the mean phenotypic value for individuals in that genotype class.

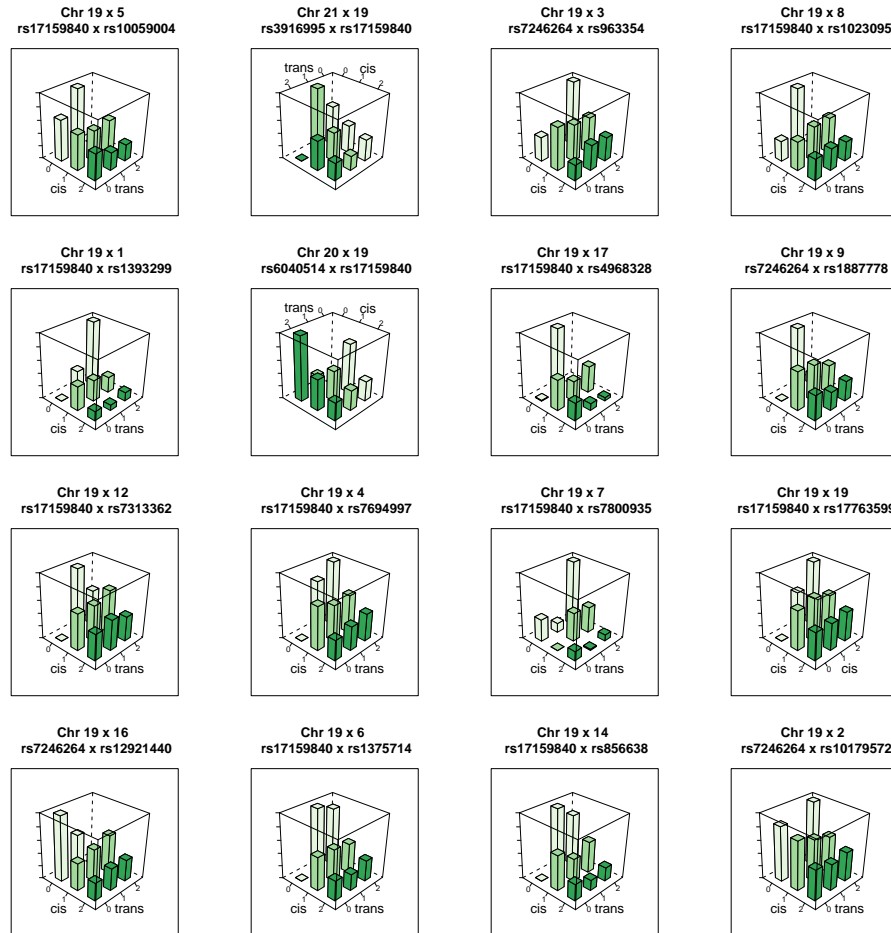


Figure S13: **Genotype-phenotype maps for 16 interactions influencing the expression of TRAPPC5** Each bar represents the mean phenotypic value for individuals in that genotype class.

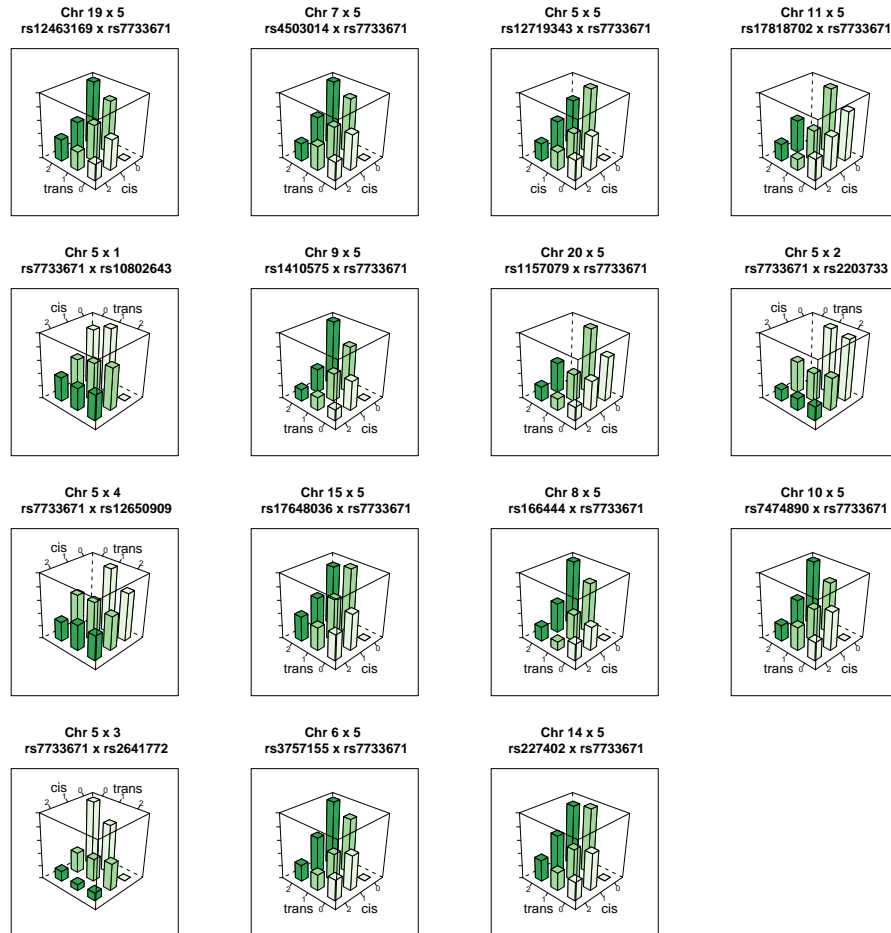


Figure S14: **Genotype-phenotype maps for 15 interactions influencing the expression of CAST** Each bar represents the mean phenotypic value for individuals in that genotype class.

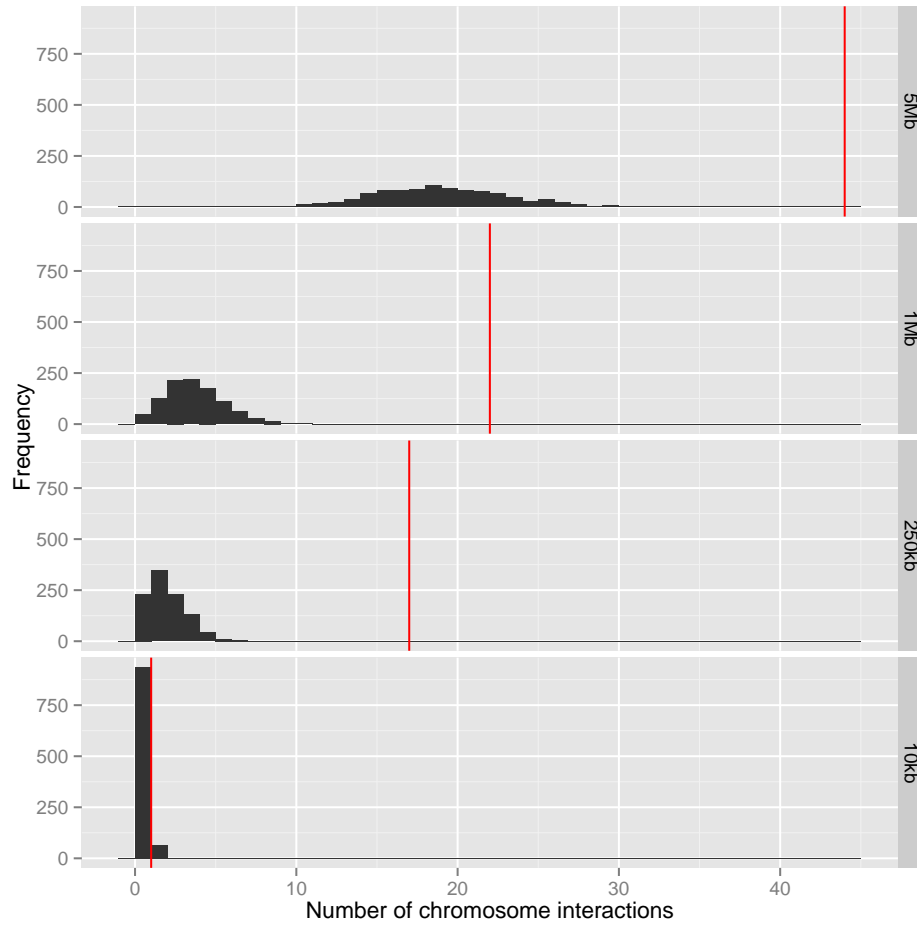


Figure S15: Number of overlaps between chromosome interactions and epistatic interactions Interacting chromosome regions may be a possible mechanism underlying epistatic interactions. The number of epistatic interactions within 20kb, 500kb, 2Mb and 10Mb of known chromosome interacting regions are shown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

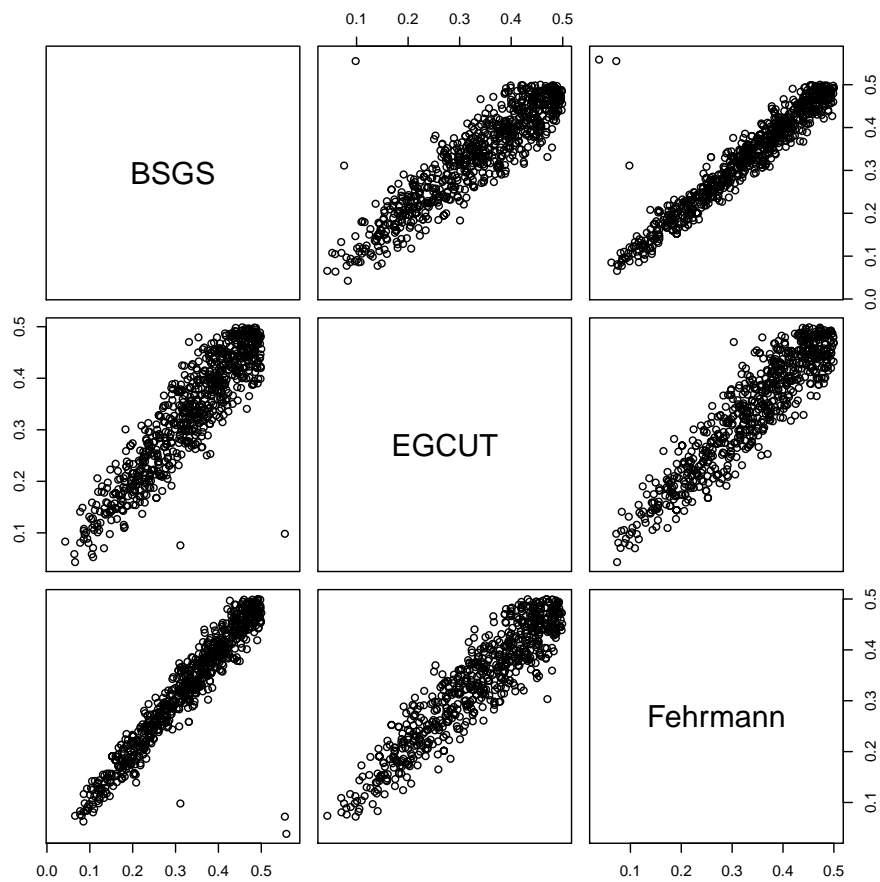


Figure S16: **Comparison of allele frequencies for 781 SNPs involved in genetic interactions across independent populations** Outliers were removed from the analysis as part of the filtering stage during replication.

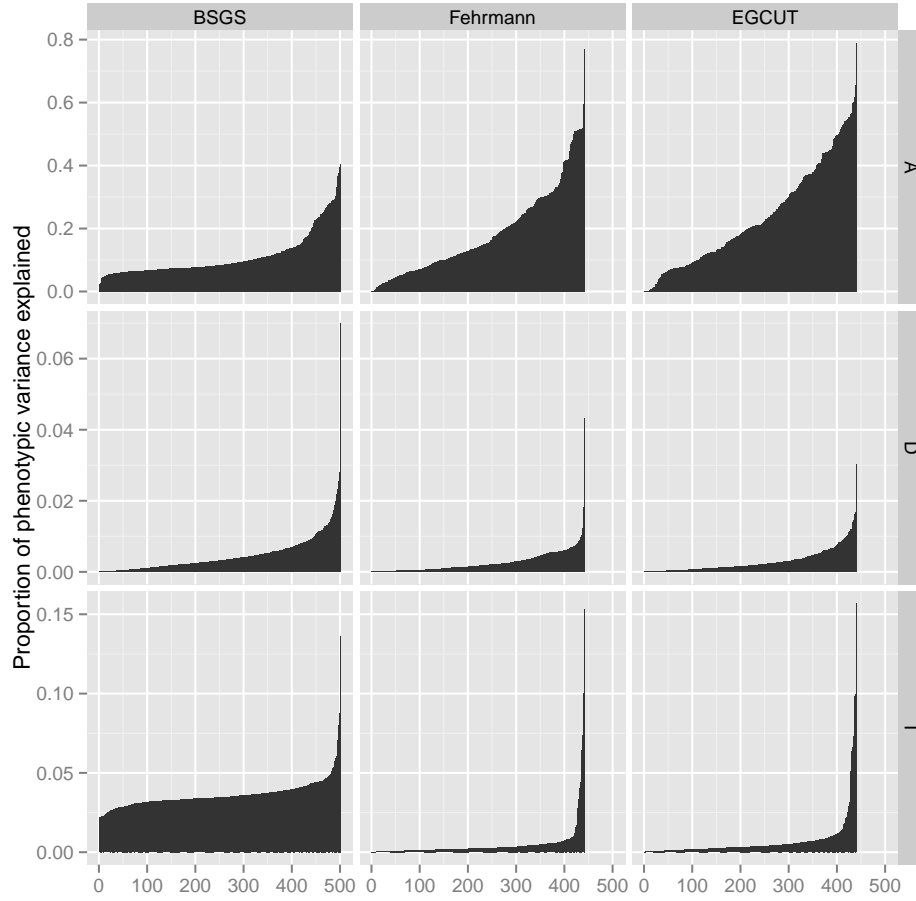


Figure S17: Comparison of variance explained by additive, dominant and epistatic effects from different cohorts How does the estimated variance decomposition change in different cohorts? The proportion of the phenotypic variance that is additive (A), dominant (D), or epistatic (I) for each putative interaction 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.

³⁹⁷ **Supplementary Tables**

Table S1 – continued from previous page

Gene ID ^a	Expression trait ^b	SNP 1			SNP 2			Interaction statistic / -log ₁₀ p-values		
		rs ID	Chr.	Pos/Mb ^c	Association ^d	rs ID	Chr.	Pos/Mb ^c	Association ^d	Distance / Mb ^h
CBORF69	ILMN_1653205	rs8051751	16	7188323		rs2890452	8	86102223	CBORF59	5.79
CBORF72	ILMN_1741881	rs10122902	9	27550780	C9ORF72	rs2420910	1	24209101	CABC1	0.18
CAC1	ILMN_1731064	rs12765847	10	4353908	INPP5E	rs3738725	1	22174210		0.96
CARD9	ILMN_1712532	rs4260763	9	139289825		rs654040	1	82128600		0.01
CARD9	ILMN_1712532	rs4573069	11	6026601		rs4077515	9	139266486	INPP5E	3.81
CAST	ILMN_1717234	rs112463169	20	6778978		rs7733671	5	96000269	CAST	0.61
CAST	ILMN_1717234	rs12463169	19	17321669		rs7733671	5	96000269	CAST	0.09
CAST	ILMN_1717234	rs12599264	16	81840122		rs7733671	5	96000269	CAST	0.23
CAST	ILMN_1717234	rs12719343	5	125369113		rs7733671	5	96000269	CAST	0.02
CAST	ILMN_1717234	rs1410575	9	78255630		rs7733671	5	96000269	CAST	0.62
CAST	ILMN_1717234	rs160444	8	27311111		rs7733671	5	96000269	CAST	1.75
CAST	ILMN_1717234	rs17048036	15	78392770		rs7733671	5	96000269	CAST	0.36
CAST	ILMN_1717234	rs17818702	11	86107920		rs7733671	5	96000269	CAST	0.13
CAST	ILMN_1717234	rs227402	14	70496867		rs7733671	5	96000269	CAST	0.27
CAST	ILMN_1717234	rs2822124	21	13160804		rs7733671	5	96000269	CAST	0.37
CAST	ILMN_1717234	rs3757155	6	136458593		rs7733671	5	96000269	CAST	0.41
CAST	ILMN_1717234	rs4003014	7	31149140		rs7733671	5	96000269	CAST	0.59
CAST	ILMN_1717234	rs4747890	10	39590078		rs7733671	5	96000269	CAST	0.01
CAST	ILMN_1717234	rs7733671	5	96000269	CAST	rs10802643	1	238120177		0.07
CAST	ILMN_1717234	rs7733671	5	96000269	CAST	rs12630909	1	17012880		0.33
CAST	ILMN_1717234	rs7733671	5	96000269	CAST	rs2203753	2	224095101		0.92
CAST	ILMN_1717234	rs7733671	5	96000269	CAST	rs2641772	3	195531841		1.72
CAST	ILMN_1717234	rs8723203	11	66175386		rs41152695	11	34115386	CAT	0.22
CCDC88B	ILMN_1692705	rs2352303	18	60973980		rs41152695	11	34115386	CCDC88B	0.15
CCDC88B	ILMN_1722268	rs694739	19	1709590		rs1271549	10	6415142		0.26
CD86	ILMN_1784863	rs5211834	7	80283117	CCDC88B	rs1254900	1	8498183	YAMP8	0.37
CD86	ILMN_1800940	rs7508015	11	76053374		rs670105	7	20755354	CD55	0.18
CD86	ILMN_1704730	rs1884655	20	23074375	CD93	rs4607740	7	15782340		0.13
CD93	ILMN_1704730	rs1884655	20	23074375	CD93	rs7623520	4	7692632		0.01
CD93	ILMN_1704730	rs1884655	20	23074375	CD93	rs8388750	3	196721305		0.92
CD93	ILMN_1704730	rs1884655	20	23074375	CD93	rs8388750	12	125145394		0.24
CD93	ILMN_1704730	rs1884655	20	23074375	CD93	rs8576388	13	38434372		0.82
CD93	ILMN_1704730	rs2889504	20	37771578		rs186858	20	38434372		0.27
CD93	ILMN_1704730	rs4813479	20	23076914		rs10925247	1	238890903		1.67
CD93	ILMN_1704730	rs4813479	20	23076914		rs2873430	8	138500554	CD93	0.71
CD93	ILMN_1704730	rs4813479	20	23076914		rs4295531	17	77264432		0.22
CD93	ILMN_1704730	rs4813479	20	23076914		rs7294744	13	115080398		0.64
CD93	ILMN_2230796	rs9015940	14	104162263		rs11655031	17	3083182	CDK16	0.21
CEACAM21	ILMN_1730928	rs200690	17	46614102	HOXB2	rs2421050	5	158043044	CEACAM21	0.15
CEACAM21	ILMN_1745949	rs4803481	19	42068556		rs4803481	19	42068556		0.07
CEACAM21	ILMN_1707554	rs6505780	18	13069792	CEACAM21	rs13132719	3	134247706		0.12
CEACAM21	ILMN_1707554	rs6505780	18	13069792	CEP102	rs13079012	3	180265266	ANAPC13	0.24
CEP63	ILMN_1787808	rs3825569	14	101350298		rs2695290	12	102087844		0.09
CEP63	ILMN_2259945	rs8199935	16	55861794	CES1	rs2695290	12	102087844		0.72
CHPT1	ILMN_2209240	rs501967	13	38838122		rs867578	11	81937002	CHPT1	0.44
CHPT1	ILMN_2209240	rs501967	13	38838122		rs7313235	12	10132283		0.36
CLEC12A	ILMN_1663142	rs429790	12	84471642		rs3903088	10	134236688	CLEC12A	0.02
CLEC12A	ILMN_2403228	rs7405054	11	96929337		rs6863172	5	175595960		0.73
CLTB	ILMN_1674609	rs17129799	11	96929337		rs169130	16	63121080	CLTB	0.27
CNN2	ILMN_1770290	rs3752237	19	1047161	ABCA7	rs169130	16	63121080		0.07
CNN2	ILMN_1770290	rs3752237	19	1047161	ABCA7	rs7713633	13	67713633		0.28
CP5F1	ILMN_1654545	rs4333645	8	145569535		rs1455268	4	61738094		1.39
CPVL	ILMN_1682928	rs12596791	16	26115562		rs2455884	7	29188475	CPVL	0.01
										0.57

Continued on next page

Table S1 – continued from previous page

Gene ID ^a		Expression trait		SNP 1		SNP 2		Interaction statistic / -log ₁₀ p-values		Meta ^g		Distance / Mb ^h			
CPVL	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 ^g	Meta ^g	Distance / Mb ^h
CPVL	ILMN-1682928	7	rs2835998	21	39202070		rs245884	7	29188475	CPVL	5.55	0.19	0.03	0.04	
	ILMN-1813256	2	rs2131290	4	188859908		rs1531133	7	46843631	CPVL	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	
	ILMN-1761797	21	rs6979356	21	43230974		rs3761385	21	45196355		11.99	25.20	16.72	42.27	0.033
CTNNA1	ILMN-1804854	5	rs924943	18	69000505		rs176382	5	138226707	CTNNA1	5.74	0.02	0.41	0.11	
	ILMN-1696347	11	rs2457684	11	88139983	CTSC	rs7079264	10	10679892		5.67	0.92	0.74	0.03	
CTSC	ILMN-2242463	11	rs75732236	22	26250645		rs1728352	11	88077357	CTSC	5.84	0.49	0.73	0.73	
	ILMN-2242463	11	rs7930237	11	88117962		rs12784396	10	102027407		7.16	18.76	15.06	33.53	0.040
CWFI9L1	ILMN-1651886	10	rs7108734	11	11450027		rs12784396	10	102027407	CWFI9L1	5.42	0.21	0.01	0.03	
	ILMN-1712305	4	rs2592948	4	129994690		rs888427	2	172366120	CYBRD1	5.89	0.23	0.53	0.34	
CYBRD1	ILMN-1712305	2	rs7852475	9	140698856		rs888427	2	172366120	CYBRD1	5.68	0.20	0.02	0.04	
	ILMN-2087692	2	rs11257679	10	12318284		rs888427	2	172366120	CYBRD1	5.81	0.39	1.87	1.47	
CYBRD1	ILMN-2087692	2	rs6137908	20	23344590		rs888427	2	172366120	CYBRD1	5.53	0.05	0.83	0.36	
	ILMN-2087692	2	rs888427	2	172366120	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	
	ILMN-2128428	5	rs7778910	17	110451383		rs835223	5	39381357	DAB2	5.44	0.48	0.41	0.44	
DAB2	ILMN-1811648	17	rs9900173	17	133111688		rs1343244	9	824705988		5.12	0.00	0.38	0.14	
	ILMN-1690982	22	rs9760102	22	24248761	DDT	rs275341	3	184747208		5.62	0.64	0.25	0.42	
DDT	ILMN-1797001	9	rs4857097	11	125962645		rs7042042	7	32451144		5.31	0.61	0.29	0.44	
	ILMN-1783996	1	rs10120023	9	137810259	COQ10A	rs2519515	7	88204888		5.47	0.08	0.41	0.16	
DEN1	ILMN-1733998	1	rs12363827	13	106703727		rs10120023	9	137810259	COQ10A	6.39	0.77	0.02	0.29	
	ILMN-1733998	2	rs1519956	12	89468283		rs7566044	2	169960422	DHRS9	6.00	0.06	1.17	0.58	
DHRS9	ILMN-2384181	2	rs1328529	12	177153535		rs7566044	2	169960422	DHRS9	6.48	0.37	0.34	0.32	
	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	
	ILMN-1755589	12	rs11080134	17	29161503	LASS5	rs11169322	12	50610976	LASS5	4.65	0.32	0.05	0.10	
DIP2B	ILMN-1755589	12	rs1166935	12	50636364		rs2872008	7	153134888	LASS5	4.87	0.58	0.22	0.19	
	ILMN-1755589	12	rs3383858	19	41711815		rs7345459	12	50730458		5.31	0.30	0.30	0.52	
DIP2B	ILMN-1755589	12	rs3383858	19	41711815		rs1808634	8	619711454		4.40	0.37	0.09	0.02	0.01
	ILMN-1755589	12	rs7134595	12	50730458	LASS5	rs4532958	10	115214154	LASS5	5.03	0.48	0.00	0.11	66.920
DIP2B	ILMN-1755589	12	rs7319252	12	50744171	LASS5	rs4532958	10	115214154	LASS5	5.92	0.23	1.45	0.97	0.052
	ILMN-1755589	12	rs871257	12	117994348		rs3779589	7	157163614	DNABJ6	4.17	1.58	0.27	1.12	
DNABJ6	ILMN-1793770	7	rs2286842	7	157216093		rs4891884	3	6320360	DPH3	6.17	1.58	0.18	0.70	
	ILMN-2349610	3	rs12232308	15	93409054		rs1566972	3	16320360	DPH3	4.81	0.15	1.18	0.70	
ECGF1	ILMN-2109708	22	rs140522	15	93409054		rs4891884	3	6320360	DPH3	4.81	0.15	1.18	0.70	
	ILMN-2109708	22	rs140522	15	93409054		rs4891884	3	6320360	DPH3	4.81	0.15	1.18	0.70	
ECHDC2	ILMN-1671568	1	rs4324091	22	241911027	ECGF1	rs1566972	3	16320360	DPH3	4.81	0.15	1.18	0.70	
	ILMN-1671568	1	rs4324091	22	241911027	ECGF1	rs1566972	3	16320360	DPH3	4.81	0.15	1.18	0.70	
ECHDC2	ILMN-1720083	15	rs5092637	22	17675900		rs11206043	1	53402552	ECHDC2	6.19	0.22	0.35	0.22	
	ILMN-1720083	15	rs5092637	22	17675900		rs11206043	1	53402552	ECHDC2	6.19	0.22	0.35	0.22	
EHD4	ILMN-1719380	14	rs6567288	18	53244938		rs1048166	15	42192040	EHD4	6.98	0.90	0.47	0.79	
	ILMN-1719380	14	rs6567288	18	53244938		rs1048166	15	42192040	EHD4	6.98	0.90	0.47	0.79	
EIF2B2	ILMN-1794522	17	rs7216490	17	7221707	EIF5A	rs1269096	14	99603119	EIF2B2	5.56	0.23	0.11	0.10	
	ILMN-1794522	17	rs7216490	17	7221707	EIF5A	rs1269096	14	99603119	EIF2B2	5.56	0.23	0.11	0.10	
EIF5A	ILMN-1794522	17	rs7216490	17	7221707	EIF5A	rs1269096	14	99603119	EIF2B2	5.56	0.23	0.11	0.10	
	ILMN-1794522	17	rs7216490	17	7221707	EIF5A	rs1269096	14	99603119	EIF2B2	5.56	0.23	0.11	0.10	
EIF5A	ILMN-1794522	17	rs7216490	17	7221707	EIF5A	rs1269096	14	99603119	EIF2B2	5.56	0.23	0.11	0.10	
	ILMN-1794522	17	rs7216490	17	7221707	EIF5A	rs1269096	14	99603119	EIF2B2	5.56	0.23	0.11	0.10	
EMR2	ILMN-2353633	19	rs2827076	21	132196249		rs1553474	2	49359676		5.54	0.56	0.08	0.24	
	ILMN-2353633	19	rs2827076	21	132196249		rs1553474	2	49359676		5.54	0.56	0.08	0.24	
EMR2	ILMN-2353633	19	rs6132112	20	18761714		rs9305048	19	14879034	EMR2	6.51	0.36	0.04	0.11	
	ILMN-2353633	19	rs6132112	20	18761714		rs9305048	19	14879034	EMR2	6.51	0.36	0.04	0.11	
EMR2	ILMN-2353633	19	rs9305048	19	14879034	EMR2	rs3007765	13	102480759	EMR2	5.70	0.20	0.58	0.35	
	ILMN-2353633	19	rs9305048	19	14879034	EMR2	rs3007765	13	102480759	EMR2	5.70	0.20	0.58	0.35	
EPHX2	ILMN-1709237	8	rs1107764	19	12790396		rs12115088	8	607161	ERCHI	6.11	0.25	1.20	0.81	
	ILMN-1709237	8	rs1107764	19	12790396		rs12115088	8	607161	ERCHI	6.11	0.25	1.20	0.81	
ERCHI	ILMN-1731001	8	rs10894861	11	12790396		rs4735900	8	578742	ERCHI	5.65	0.29	0.04	0.08	
	ILMN-1731001	8	rs10894861	11	12790396		rs4735900	8	578742	ERCHI	5.65	0.29	0.04	0.08	
ERCHI	ILMN-1731001	8	rs726145	18	31187910		rs1517297	4	182786760	EXOC3	5.63	0.67	1.03	1.06	
	ILMN-1731001	8	rs726145	18	31187910		rs1517297	4	182786760	EXOC3	5.63	0.67	1.03	1.06	
EXOC3	ILMN-2104696	5	rs4735895	8	600729		rs12188164	5	428236	EXOC3	6.83	0.74	1.19	0.44	
	ILMN-1789419	5	rs4735895	8	600729		rs12188164	5	428236	EXOC3	6.83	0.74	1.19	0.44	
FAHD1	ILMN-2246661	16	rs1560104	16	12708208		rs344363	5	4928236	FAHD1	5.61	0.37	0.38	0.23	10.736
	ILMN-2246661	16	rs1560104	16	12708208		rs344363	5	4928236	FAHD1	5.61	0.37	0.38	0.23	

Continued on next page

Table S1 – continued from previous page

Gene ID ^a	Probe ID ^b	Expression trait		SNP 1			SNP 2			Interaction statistic ^f / -log ₁₀ p-values			Distance / Mb ^h		
		rs ID	Chr.	Pos/Mb ^c	Association ^d	rs ID	Chr.	Pos/Mb ^c	Association ^d	BSGS ^e	Fehrmann ^f	EGCUT ^g			
FE22	ILMN-1739586	2	rs2356400	19	44321776			rs13406184	2	36791226	FE22	5.78	0.14	0.33	0.16
FE22	ILMN-1739586	2	rs969010	4	159963132			rs11691600	2	36810133	FE22	6.59	0.14	0.28	0.14
FGD2	ILMN-2115005	6	rs4803848	19	46203030			rs31486	6	37001267	FGD2	5.69	0.12	0.25	0.11
FLJ20489	ILMN-2115005	12	rs9024634	10	133943931			rs831489	6	36999652	FLJ20489	5.49	1.20	0.11	0.66
FLJ20489	ILMN-1778144	12	rs17615703	12	117036766	FLJ20489		rs3782908	12	48169526	FLJ20489	5.81	0.06	0.70	0.29
FLJ20489	ILMN-1778144	12	rs472408	12	4569326			rs769951	12	16769951	FLJ20489	5.73	0.03	0.10	0.02
FLJ20489	ILMN-1778144	12	rs472408	12	4569326			rs831489	12	48169526	FLJ20489	5.73	0.03	0.10	0.02
FLJ20489	ILMN-1778144	12	rs934440	17	97033126			rs3782908	12	48169526	FLJ20489	6.49	0.31	0.47	0.36
FLJ20489	ILMN-1778144	12	rs204135	6	50626195			rs3782908	12	48169526	FLJ20489	6.04	0.38	0.17	0.21
FLJ20489	ILMN-1763663	16	rs9325634	21	43818790			rs3782908	12	48169526	FLJ20489	6.04	0.38	0.17	0.21
FLJ43093	ILMN-2124350	6	rs17112712	14	107276627			rs6906101	6	50116594	FLJ43093	5.48	0.14	0.95	0.53
FLJ43093	ILMN-2124350	6	rs6906101	14	107276627	FLJ43093		rs13214069	6	36676710	FLJ43093	5.48	0.39	0.06	0.13
FN3KRP	ILMN-1652333	17	rs898095	17	80890638			rs9892064	17	32705248	FN3KRP	16.16	28.24	29.39	59.95
FUC1	ILMN-1752728	17	rs4971478	17	13406033			rs12744386	17	808927903	FUC1	6.41	0.01	0.30	0.06
FXD5	ILMN-2309848	19	rs1633921	19	35695290			rs788175	13	98338559	FXD5	3.70	0.09	0.41	0.17
FXD5	ILMN-2309848	19	rs17398183	20	55609148			rs2285515	13	98338559	FXD5	6.58	0.03	0.48	0.15
FXD5	ILMN-2309848	19	rs2285515	19	35660450	FXD5		rs11739594	5	141709563	FXD5	5.70	0.07	0.17	0.05
FXD5	ILMN-2309848	19	rs2285515	19	35660450	FXD5		rs13067700	3	95331048	FXD5	6.00	0.09	0.09	0.51
FXD5	ILMN-2309848	19	rs2285515	19	35660450	FXD5		rs17036504	4	47567329	FXD5	6.10	0.28	0.08	0.08
G3BP2	ILMN-2317558	4	rs10230232	7	29390239			rs1553985	4	7654604	G3BP2	5.19	0.08	0.37	0.14
GAA	ILMN-2410783	17	rs11150847	17	78153130			rs12602462	17	78146016	GAA	13.91	19.98	12.99	32.60
GAA	ILMN-2410783	17	rs8068856	17	78100731	GAA		rs10902506	12	1332678089	GAA	5.65	0.11	0.39	0.17
GAPT	ILMN-1675191	5	rs10070522	5	57786110	GAPT		rs7605821	2	2335655228	GAPT	5.85	0.01	0.78	0.28
GAPT	ILMN-1675191	5	rs20838717	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	5.47	0.83	0.63	0.87
GDPD3	ILMN-1774901	16	rs3809624	16	30102802	GDPD3		rs2197465	14	48572632	GDPD3	6.57	0.38	0.35	0.33
GDPD3	ILMN-1774901	16	rs7204270	16	30102802	GDPD3		rs1015111	4	128972357	GDPD3	5.86	0.55	0.09	0.24
GDPD3	ILMN-1774901	16	rs4145072	13	110899955	GDPD3		rs7577283	4	128972357	GDPD3	5.78	0.02	0.45	0.13
GNLY	ILMN-1790692	2	rs1986426	16	26084476			rs7960552	12	111164237	GNLY	5.72	0.02	0.45	0.13
GNLY	ILMN-3239426	12	rs1860563	16	6478898			rs2707210	12	6902002	GNLY	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
GPR162	ILMN-1730816	12	rs2707210	12	79685913			rs4740848	9	6554558	GPR162	5.47	0.25	0.06	0.07
GPR162	ILMN-1730816	12	rs2707210	12	6902002			rs827054	3	188880113	GPR162	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	5.43	0.79	1.43	1.50
GPR177	ILMN-1660549	1	rs225613	16	11169683			rs12065581	1	68732819	GPR177	5.43	0.11	0.13	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	171399932			rs12065581	1	68732819	GPR177	6.50	0.01	0.24	0.04
GSDMB	ILMN-2347193	17	rs11557467	17	38028634	GSDMB		rs4965745	15	101508261	GSDMB	5.88	0.68	0.20	0.41
GSTM1	ILMN-2391861	1	rs12248673	10	53192833			rs11101992	1	110266754	GSTM1	6.11	0.27	1.14	0.79
GSTM1	ILMN-2391861	1	rs1547574	13	85345257			rs11101992	1	110266754	GSTM1	5.91	0.27	1.14	0.79
GSTM2	ILMN-2201580	1	rs6432807	13	96159560			rs3754446	1	110253241	GSTM2	6.77	0.66	0.77	0.66
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs4533333	2	77919015	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs4947007	15	58677017	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs9883949	21	19532546	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039	11	5271671	H1FO	6.36	0.32	0.52	0.65
H1FO	ILMN-1757467	22	rs139898	22	38399979			rs2855039							

Continued on next page

Table S1 – continued from previous page

[illegible]

Continued on next page

Gene ID ^a	Expression trait		SNP 1		SNP 2		Interaction statistic ^c		Distance / Mb ^e	
	rs ID	Chr.	Pos/Mb/c	Association ^d	rs ID	Chr.	Pos/Mb/c	Association ^d	BSGG ^f	EGCU [†] - log ₁₀ p-values
NRRF2	r56025645	20	56157341	OSTF1	rs79236009	10	65133822	NRRF2	5.45	
	ILMN_3237385	10			rs6517815	20	19819016		6.11	
	ILMN_3237385	10			r56517815	20	19819016		6.11	
	ILMN_1800897	1			r56588415	1	52334047	NRRF2	6.13	0.05
	ILMN_1787885	8			r54057351	21	25453482	NUDT18	5.44	0.03
	ILMN_1658247	12			ILMN_1658247	12	1134480510		8.59	0.15
	ILMN_1658247	12			OAS1	12	1134480510		4.13	0.039
	ILMN_1658247	12			ILMN_1658247	12	1134480510		4.13	0.039
	ILMN_1658247	12			ILMN_1658247	12	1134480510		4.13	0.039
	ILMN_1658247	12			ILMN_1658247	12	1134480510		4.13	0.039
OSTF1	rs79236009	10	65133822	OSTF1	rs79236009	10	65133822	NRRF2	5.45	
	ILMN_3237385	10			rs6517815	20	19819016		6.11	
	ILMN_3237385	10			r56517815	20	19819016		6.11	
	ILMN_1800897	1			r56588415	1	52334047	NRRF2	6.13	0.05
	ILMN_1787885	8			r54057351	21	25453482	NUDT18	5.44	0.03
	ILMN_1658247	12			ILMN_1658247	12	1134480510		8.59	0.15
	ILMN_1658247	12			OAS1	12	1134480510		4.13	0.039
	ILMN_1658247	12			ILMN_1658247	12	1134480510		4.13	0.039
	ILMN_1658247	12			ILMN_1658247	12	1134480510		4.13	0.039
	ILMN_1658247	12			ILMN_1658247	12	1134480510		4.13	0.039
OSTF1	rs79236009	10	65133822	OSTF1	rs79236009	10	65133822	NRRF2	5.45	
	ILMN_3237385	10			rs6517815	20	19819016		6.11	
	ILMN_3237385	10			r56517815	20	19819016		6.11	
	ILMN_1800897	1			r56588415	1	52334047	NRRF2	6.13	0.05
	ILMN_1787885	8			r54057351	21	25453482	NUDT18	5.44	0.03
	ILMN_1658247	12			ILMN_1658247	12	1134480510		8.59	0.15
	ILMN_1658247	12			OAS1	12	1134480510		4.13	0.039
	ILMN_1658247	12			ILMN_1658247	12	1134480510		4.13	0.039
	ILMN_1658247	12			ILMN_1658247	12	1134480510		4.13	0.039
	ILMN_1658247	12			ILMN_1658247	12	1134480510		4.13	0.039
OSTF1	rs79236009	10	65133822	OSTF1	rs79236009	10	65133822	NRRF2	5.45	
	ILMN_3237385	10			rs6517815	20	19819016		6.11	
	ILMN_3237385	10			r56517815	20	19819016		6.11	
	ILMN_1800897	1			r56588415	1	52334047	NRRF2	6.13	0.05
	ILMN_1787885	8			r54057351	21	25453482	NUDT18	5.44	0.03
	ILMN_1658247	12			ILMN_1658247	12	1134480510		8.59	0.15
	ILMN_16582									

Continued on next page

Table S1 – continued from previous page

Gene ID ^a	Expression trait		SNP 1		SNP 2		Interaction statistic / -log ₁₀ p-values			
	Probe ID ^b	Chr.	rs ID	Chr.	Pos/Mb ^c	Association ^d	rs ID	Chr.	Pos/Mb ^c	Association ^d
REBE	ILMN-1802380	1	rs4982958	14	24987865		rs301819	1	8501786	REBE
REBE	ILMN-1802380	1	rs7697290	4	135248366		rs301819	1	8501786	REBE
REBE	ILMN-2327795	1	rs11085829	19	13174312		rs301819	1	8501786	REBE
REBE	ILMN-2327795	1	rs3852011	3	112844086	RNASE6	rs301819	1	8501786	REBE
RNASE6	ILMN-1780533	14	rs11628398	14	8106521		rs11628398	14	100601327	RNASE6
RNASE6	ILMN-1780533	14	rs6003134	19	8106521		rs11628398	14	100601327	RNASE6
RNASE6	ILMN-1794726	17	rs2382330	17	4875566		rs11706900	3	36348908	
RNF167	ILMN-1794726	17	rs400658	17	4839930	RNF167	rs11706900	3	36348908	
RNFEP	ILMN-1738347	1	rs1107121	17	46127349		rs2819365	1	201983242	
RNFEP	ILMN-1738347	1	rs8071611	17	67153586		rs2819365	1	201983242	
RPL13	ILMN-2413278	16	rs352935	16	89048580		rs2965817	16	89513234	
RPL23AP7	ILMN-2222730	12	rs1401202	16	80320056		rs4848261	2	114450028	RPL23AP7
RPL36AL	ILMN-2186933	14	rs3007033	14	50103816	RPL36AL	rs17450530	9	138035083	
RPL36AL	ILMN-2186933	14	rs4009028	14	50020817	RPL36AL	rs1502991	6	66137260	
RPL8	ILMN-1764721	8	rs2958482	8	143984615	RPL8	rs1619856	1	234585790	
RPL8	ILMN-1764721	8	rs4143674	20	4741304		rs2958482	8	143984615	
SEC13	ILMN-3297880	3	rs4889214	16	80913946		rs696221	3	10342876	SEC13
SEC13	ILMN-3297880	3	rs17085428	3	83388015		rs7695	1	136147326	SEC13
SES3	ILMN-1702787	11	rs12147460	14	104412137		rs684856	11	94906111	SES3
SES3	ILMN-1694027	11	rs355391	15	46391793	SES3	rs684856	11	94906111	SES3
SES3	ILMN-1694027	11	rs684856	15	46391793		rs7004947	8	134606425	PPBP
SH3BGL2	ILMN-1694027	11	rs10838191	11	43893658		rs1354034	3	56849749	PPBP
SH3BGL2	ILMN-1767664	6	rs2345385	5	46838939		rs1354034	3	56849749	PPBP
SH3BGL2	ILMN-1767664	6	rs6845364	4	88280592		rs1745517	9	131785369	SH3BGL2
SH3BGL2	ILMN-2158336	9	rs1034294	21	18196922	SIRPG	rs6842739	14	60429380	
SIRPG	ILMN-2158336	9	rs1355883	20	1512549		rs367035	17	15233826	SLC22A18
SIRPG	ILMN-2158336	9	rs1355883	20	1512549		rs367035	17	15233826	SLC22A18
SLC22A18	ILMN-2382605	11	rs11673260	19	5215198	SLC22A18	rs367035	17	15233826	SLC22A18
SLC22A18	ILMN-2382605	11	rs367035	19	5215198	SLC22A18	rs367035	17	15233826	SLC22A18
SLC22A18	ILMN-2382605	11	rs367035	19	5215198	SLC22A18	rs367035	17	15233826	SLC22A18
SLC41A3	ILMN-226111	3	rs1912136	11	2923674		rs771703	3	12587558	SLC41A3
SLC41A3	ILMN-226111	3	rs698508	8	14233774	SLC41A3	rs771703	3	12587558	SLC41A3
SLC46A3	ILMN-1658639	13	rs19805	17	5502091		rs7081100	13	28259349	SLC46A3
SLC46A3	ILMN-1658639	13	rs803259	15	97030923		rs10911353	1	18349203	SLC46A3
SMO7	ILMN-1775380	20	rs11677215	20	4161500	SMO7	rs11677215	2	65800982	SMO7
SMO7	ILMN-1775380	20	rs11677215	20	4161500	SMO7	rs11677215	2	65800982	SMO7
SNHG8	ILMN-3309380	4	rs1105621	9	133050233		rs214097	11	19225940	SNHG8
SNHG8	ILMN-3309380	4	rs1105621	9	133050233		rs214097	11	19225940	SNHG8
SNORD14A	ILMN-1709381	11	rs1504220	15	46250108		rs6148334	11	1701499	SNORD14A
SNORD14A	ILMN-1709381	11	rs2634462	11	17339197		rs6148334	11	1701499	SNORD14A
SNORD89	ILMN-3238662	2	rs1605863	2	115929241		rs750783	2	101889306	SNORD89
SNORD89	ILMN-3238662	2	rs1605863	2	115929241		rs750783	2	101889306	SNORD89
SNORD89	ILMN-3238662	2	rs1605863	2	115929241		rs750783	2	101889306	SNORD89
SNUPN	ILMN-1733932	15	rs2135064	5	26778066	SNUPN	rs7185362	16	81888905	SNUPN
SNUPN	ILMN-1733932	15	rs2135064	5	26778066	SNUPN	rs7185362	16	81888905	SNUPN
SPATA5L1	ILMN-2364535	15	rs1346466	21	46376528		rs1472075	3	193706323	SPATA5L1
SPATA5L1	ILMN-2364535	15	rs1346466	21	46376528		rs1472075	3	193706323	SPATA5L1
STARD10	ILMN-1729179	15	rs131620	19	41117869		rs1006620	15	45652086	STARD10
STARD10	ILMN-1729179	15	rs131620	19	41117869		rs1006620	15	45652086	STARD10
STYXL1	ILMN-2210752	11	rs2221406	13	90174526		rs17685	7	75616105	STYXL1
STYXL1	ILMN-2210752	11	rs2221406	13	90174526		rs17685	7	75616105	STYXL1
SULT1A4	ILMN-2345142	20	rs11700063	14	104947517	SULT1A4	rs392994	4	180439236	TUFM
SULT1A4	ILMN-2345142	20	rs11700063	14	104947517	SULT1A4	rs392994	4	180439236	TUFM
SULT1A4	ILMN-2345142	20	rs11700063	14	104947517	SULT1A4	rs392994	4	180439236	TUFM
SURF6	ILMN-2336133	16	rs1463965	18	74332954		rs3785354	16	28550667	SURF6
SURF6	ILMN-2336133	16	rs1463965	18	74332954		rs3785354	16	28550667	SURF6
SYTL2	ILMN-1778032	9	rs6099626	20	56013994		rs3118663	9	136281753	SYTL2
SYTL2	ILMN-1778032	9	rs6099626	20	56013994		rs3118663	9	136281753	SYTL2
THBS3	ILMN-2336609	11	rs1375719	13	103410782		rs485485	11	85495269	THBS3
THBS3	ILMN-2336609	11	rs1375719	13	103410782		rs485485	11	85495269	THBS3
THBS3	ILMN-1804663	1	rs1939875	11	95422867		rs4072037	1	155194980	THBS3
THBS3	ILMN-1804663	1	rs1939875	11	95422867		rs4072037	1	155194980	THBS3
TIPRL	ILMN-1781457	1	rs8014956	14	20687978		rs32049805	1	168154599	TIPRL
TIPRL	ILMN-1781457	1	rs2823245	21	16745523		rs32049805	1	168154599	TIPRL

Continued on next page

	SNP 2	Interaction statistic / —	og ₁₀ p-values

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 ^g	Meta ^g	Distance / Mb ^h
UBASH3A	LMN-2338348	21	rs1893592	21	43855067	UBASH3A	rs7201194	16	83600397		5.91	0.59	0.42	0.52	
UBASH3A	LMN-2338348	21	rs1893592	21	43855067	UBASH3A	rs7512594	1	214517361		6.01	0.48	1.29	1.10	
USP36	LMN-1697227	17	rs2279308	17	76794981	USP36	rs7225546	17	75151717		5.71	0.03	0.14	0.03	1.643
VASP	LMN-1743646	19	rs1264226	19	40603167		rs2276470	19	45974668	VNN2	5.09	0.94	5.14	4.95	0.088
VNN2	LMN-1678939	6	rs10435352	7	103252718		rs1883613	6	133077063	VNN2	5.04	0.84	0.15	0.46	
VNN2	LMN-1678939	6	rs13044386	20	9116155		rs1883617	6	133072650	VNN2	5.44	0.39	0.69	0.57	
VNN2	LMN-1678939	6	rs134447	22	49927332		rs1883617	6	133072650	VNN2	5.72				
VNN3	LMN-1678939	6	rs216495	11	16834510		rs1883617	6	133072650	VNN2	5.77	0.33	0.19	0.19	
VNN3	LMN-1678939	6	rs10278073	7	151662184		rs2267932	6	133067782	VNN3	6.44	0.16	0.74	0.41	
VNN3	LMN-1804935	6	rs1443946	8	73006453		rs2267932	6	133067782	VNN3	5.74	0.23	0.48	0.31	
VNN3	LMN-1804935	6	rs348462	9	75547169		rs2267952	6	133067782	VNN3	6.44	0.31	0.17	0.17	
VNN3	LMN-1804935	6	rs7157055	14	83262064		rs2267952	6	133067782	VNN3	5.82	0.03	0.19	0.04	
VNN3	LMN-2387680	6	rs2823165	21	5694253		rs2267952	6	133067782	VNN3	6.12	0.73	1.15	1.21	
VNN3	LMN-2387680	6	rs9596457	13	51692548		rs2267952	6	133067782	VNN3	4.83	0.46	0.05	0.16	
VSTM1	LMN-1763455	19	rs9596457	19	54553697	VSTM1	rs4552100	18	71024750		5.60	0.53	0.54	0.57	
VSTM1	LMN-1763455	19	rs10500316	19	54553697	VSTM1	rs7895870	10	123095249		5.71	0.48	0.17	0.26	
VSTM1	LMN-1763455	19	rs10500316	19	54553697	VSTM1	rs10500316	19	54553697	VSTM1	5.88	0.81	1.38	1.47	
WDR48	LMN-1762103	3	rs9628570	22	30261219		rs6778963	3	39091812	WDR48	5.88	0.09	0.33	0.09	
WDR48	LMN-1762103	3	rs1388935	4	18827822		rs833349	3	39067325	WDR48	6.34	0.57	1.35	1.22	
WDR48	LMN-1762103	3	rs1887778	9	134635088	RAPGEF1	rs7619193	3	39044116	WDR48	5.85	0.18	0.61	0.35	
WDR6	LMN-1669484	3	rs9554833	13	102624790		rs7619193	3	39044116	WDR6	5.86	1.64	1.43	2.25	
XAF1	LMN-2330573	17	rs12362253	11	123571708		rs17175581	15	93119799		4.86	2.38	0.17	1.63	
XAF1	LMN-2330573	17	rs1535031	21	6073170	XAF1	rs12591171	15	68173945		5.79	0.09	0.36	0.15	
ZEP00	LMN-1680928	16	rs906446	21	3906468		rs1829968	16	47573945	ZEP00	5.79	0.67	0.27	0.46	
ZNF500	LMN-1700228	16	rs4282723	22	48283177		rs2240500	16	4799041	ZNF500	5.29	0.67	0.27	0.46	
ZYX	LMN-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 – log₁₀ p -value from discovery dataset^f Interaction – log₁₀ p -value from replication dataset^g Interaction – log₁₀ p -value from meta analysis of replication datasets only^h Distance in Mb between interacting SNPs for *cis-cis* acting SNP pairsⁱ 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 2013²²

Gene	Probe	Additive		Non-additive	
		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	Fehrman	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	Fehrman	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