THE EUROPEAN SOCIETY OF MEDICINE
Medical Research Archives, Volume 12 Issue 9

RESEARCH ARTICLE

Digenic analysis confirms known and uncovers novel schizophrenia risk genes

Jian Ning, B.S. [1]; **Xingyu Wang, Ph.D.** [2]; **Jurg Ott, Ph.D.** [3]

[1] Center of Statistical Genetics, West Orange, NJ 07052, USA
[2] National Center for Human Genetic Resources, National Research Institute for Family Planning, Beijing, China 100012, and Rockefeller University, New York, NY 10065, USA
[3] Rockefeller University, New York, NY 10065, USA

OPEN ACCESS

PUBLISHED
30 September 2024

CITATION
Ning, J., Wang, X., et al., 2024. Digenic analysis confirms known and uncovers novel schizophrenia risk genes. Medical Research Archives, [online] 12(9).
https://doi.org/10.18103/mra.v12i9.5723

DOI
https://doi.org/10.18103/mra.v12i9.5723

ISSN
2375-1924

ABSTRACT

Based on an Ashkenazy Jewish case-control cohort for schizophrenia, we carried out (1) genetic association analysis for one variant at a time (GWAS) and (2) digenic analysis by comparing frequencies of genotype pairs between cases and controls. To control for genetic heterogeneity between sexes, we analyzed males and females separately. After pruning of variants in each of males and females, single-variant allelic analysis furnished 9 and 8 statistically significant variants in males and females, respectively, with 3 of these variants being significant in both males and females. Of the 14 distinct variants in males and females, 5 (36%) reside in genes. For digenic analysis, we evaluated all pairs of variants and, for a given variant pair, all nine genotype pairs. For each genotype pair, we applied the Fisher exact test to evaluate whether the given genotype pair was more frequent in cases than controls. We found 76 significant genotype pairs, comprising 36 distinct variants, 20 (56%) of which reside in genes, with many of which being known risk genes, thus lending credence to our approach.

Introduction

Schizophrenia is a rather common, highly genetic trait [1] that is also regulated epigenetically and environmentally [2]. The Human Gene Database lists 15,235 entries for schizophrenia. Many of its disease-associated variants are preserved across populations [3]. Because of its high genetic heterogeneity, we chose to look for disease-associated variants in a founder population [4,5] and downloaded an Ashkenazy Jewish case-control dataset from dbGaP.

Most investigations of schizophrenia-associated variants have been carried out in case-control studies [6,7] by assessing direct (main) effects of variants, that is, differences in allele or genotype frequencies between cases and controls, possibly combined over multiple variants in a gene [8] or in the form of polygenic risk scores over large numbers of variants [9]. To date, 287 loci with significant association to schizophrenia have been identified [9,10].

Mathematical models have shown that empirical segregation ratios of schizophrenia fit polygenic models better than single-locus models [11]. Thus, as outlined below, we proceeded to search for risk variants based on pairwise interactions between genotypes in addition to direct variant effects [12]. Such an exhaustive search for all pairs of variant genotypes has previously required enormous computing efforts [13], but modern workstations containing dozens of threads (central processing units, CPUs) have made this task more manageable [14]. Thus, our *Gpairs* program not only evaluates all possible pairs of variants, but, for a given variant pair, it tests each of the $3 \times 3 = 9$ pairs of genotypes whether the pair occurs in higher frequency in cases than controls [14].

Methods

DATA

A dataset entitled *Genetics of Schizophrenia in an Ashkenazi Jewish Case-Control Cohort* [15] was downloaded from *dbGaP* in the form of a binary *plink*-formatted [16] fileset. Customary quality-control measures reduced the original 1,016,422 genetic variants down to 892,850 variants, each genotyped in 3,096 individuals (1,044 cases and 2,052 controls; 2,164 males and 932 females).

Even though our data represent an ethnically homogeneous set of individuals, other sources of heterogeneity are of potential concern, notably differences in heritability between males and females from a genetic [17] and a biological [18] perspective. For example, for Parkinson Disease, a striking effect of sex on monocyte gene expression has been shown [19], with a note on the "importance of studies which examine the differential effects of sex on pathophysiology" of disease. Also, there are clear morphological differences in the brains of the two sexes [2], gene expression differences between male and female schizophrenics have been documented [20], association between troponin T levels and psychosis have only been found in women, not in men [21], a coronary heart disease genetic risk score predicted disease risk only in men, not women [22], sex differences in gene regulatory networks underlying lung cancer have been documented [23], and a sex difference exists in the association between cannabis use disorder and schizophrenia [24]. Allowing for sex in case-control studies may be accomplished through a logistic regression model. However, rather than imposing the constraints of such a model, we decided to analyze the data in sex-specific subgroups [24], which would eliminate any heterogeneity due to sex differences although at the cost of smaller numbers of individuals in each of the resulting two datasets. Analyzing a heterogeneous combined dataset would be expected to lead to false positive results. Thus, we proceeded to separately analyze the 2,164 males (660 cases, 1,504 controls) and 932 females (384 cases, 548 controls). In each of the two resulting data subsets, we imposed a minimum minor allele frequency of 0.01 and made variants relatively independent by applying the "indep 50 5 2" option in *plink*. These steps resulted in 179,104 variants in males and 179,898 variants in females.

Potential heterogeneity is often addressed by the use of principal components [25] as covariates in a logistic regression analysis. However, the steps outlined above did not seem to necessitate the use of principal components, and there are also concerns regarding their use [26].

GENETIC ANALYSIS

Initially, we carried out a standard GWAS in each of the two sexes and applied the Fisher exact test as implemented in *plink* (--assoc function). Empirical significance levels (*p*-values), corrected for multiple testing, were obtained in 100,000 permutations of phenotypes. Results were declared statistically significant for $p \leq 0.05$. For digenic analysis, we applied the *Gpairs* program as previously described [14] to evaluate, separately for males and females, all pairs of genotypes. For each genotype pair, we applied the Fisher exact test to see whether the given genotype pair was more frequent in cases than controls. Correction for multiple testing was carried out by the Bonferroni method.

Results

SINGLE-VARIANT GWAS

Table 1. Nine variants significant in single-variant association tests for males

chr	rsID	bp GRCh38	gene	function
1	rs7340057	73,269,583	—	—
1	rs6675786	248,066,288	OR2L13, LOC105373275	intron variant, genic upstream transcript variant
2	rs4387788	240,655,829	—	—
4	rs422548	68,745,554	—	—
8	rs7011530	125,838,349	—	—
11	rs1582781	12,561,634	—	—
13	rs2496577	38,226,298	—	—
22	rs8138145	22,146,185	—	—
22	rs5998848	33,288,529	LARGE1	genic downstream transcript variant, intron variant

In males, 9 variants were significant ($p \leq 0.05$), while 8 variants were significant in females, as shown in Tables 1 and 2. Of the 17 significant variants, 3 were shared between males and females, that is, single-variant analysis furnished a total of 14 unique significant variants, of which 5 (36%) reside in genes.

Table 2. Eight variants significant in single-locus association tests for females

chr	rsID	bp GRCh38	gene	function
1	rs7340057	73,269,583	--	--
2	rs6747270	13,446,506	--	--
4	rs422548	68,745,554	--	--
5	SNP5-70873072	*70,873,072*	BDP1	--
6	SNP6-26333526	*26,333,526*	H3C6	--
6	rs6915052	121,376,832	--	--
11	rs1582781	12,561,634	--	--
12	rs12296316	86,280,211	MGAT4C	intron variant, genic upstream transcript variant

Table 3. Thirty-six unique variants in males (m) or females (f) contributing to the 76 significant genotype pairs; *n* represents the number of other variants connected with the given variant.

ch	rsID	bp GRCh38	source	n	gene	function
1	rs7340057	73,269,583	m	3	—	—
1	rs2518417	100,539,811	m	0	GPR88	missense variant, coding sequence variant
1	rs7530249	199,369,333	m	0	LINC02789	intron variant
1	rs6675786	248,066,288	m	3	OR2L13, LOC105373275	intron variant, genic upstream transcript variant
2	rs6747270	13,446,506	mf	5	—	—
3	rs11924588	18,908,029	m	0	—	—
3	SNP3-186943722	*186,943,722*	m	0	IGF2BP2	—
4	rs422548	68,745,554	mf	6	—	—
5	rs10041314	38,584,216	m	1	LIFR,LIFR-AS1	intron variant, genic upstream transcript variant
5	rs16903025	88,245,670	m	9	TMEM161B	intron variant, genic upstream transcript variant
6	rs4134989	20,493,230	f	0	E2F3	genic downstream transcript variant, 3'UTR variant
6	SNP6-20737399	*20,737,399*	m	1	CDKAL1	—
6	SNP6-26333526	*26,333,526*	m	1	H3C6	—
6	rs6915052	121,376,832	m	2	—	—
7	rs28879680	91,748,465	m	1	—	—
8	SNP8-10507540	*10,507,540*	m	1	RP1L1	—
8	rs6982398	73,833,974	m	0	UBE2W	intron variant
8	rs11987092	122,014,250	m	1	—	—
8	rs7011530	125,838,349	m	3	—	—
9	rs7865153	126,799,253	m	3	—	—
10	rs2813274	50,148,357	m	1	—	—
11	rs1582781	12,561,634	mf	8	—	—
11	SNP11-24126646	*24,126,646*	m	0	LOC107984378	—
11	SNP11-64433849	*64,433,849*	m	1	ATG2A	—
11	rs597303	88,969,779	m	1	GRM5	intron variant
11	SNP11-93097180	*93,097,180*	f	3	CEP295	—
12	rs16933566	28,905,298	m	1	—	—
12	rs12296316	86,280,211	f	2	MGAT4C	intron variant, genic upstream transcript variant
12	rs12300389	90,558,317	m	2	—	—
12	rs11107087	93,545,255	m	1	—	—
13	rs2496577	38,226,298	m	6	—	—
13	rs1755808	66,304,158	m	10	PCDH9	3'UTR variant, genic downstream transcript variant
13	rs1044364	109,755,299	f	0	IRS2	3'UTR variant
18	rs3937015	69,932,280	m	0	CD226	intron variant, genic upstream transcript variant
22	rs8138145	22,146,185	m	21	—	—
22	rs5998848	33,288,529	m	19	LARGE1	intron variant, genic downstream transcript variant

DIGENIC ANALYSIS

In males, our analysis of genotype pairs (patterns) resulted in 69 patterns with significantly higher frequencies in cases than controls ($p \leq 0.05$) while in females, 7 genotype pairs were significant. None of the significant genotype pairs were the same in males and females although a few of the individual variants making up the pairs were shared (see below). Thus, we combined genotype pairs from males and females, resulting in 76 genotype pairs or, equivalently, 76 variant pairs leading to these genotype pairs. The combined 76 variant pairs are listed in Supplementary Table S1, along with their genic locations, if known.

For a given variant pair, the two component variants were either in two genes, only in one gene, or not in any gene, which occurred for 14 (18%), 44 (58%), and 18 (24%) variant pairs, respectively. The total of 76 variant pairs comprised 36 distinct variants, 20 (56%) of which reside in genes. Table 3 lists these variants, along with gene names where a variant is located, and the number *n* of other variants connected with the given variant.

Comparing Tables 1 (GWAS for males), 2 (GWAS for females), and 3 (variants furnishing significant genotype pairs in digenic analysis), we find the following overlaps, which add strength to the importance of these variants:

- Variant rs7340057 on chromosome 1 occurs in all three tables and, thus, is likely to be important for the etiology of schizophrenia even though it does not reside in a gene.
- Variant rs6675786 on chromosome 1 occurs in males (GWAS) and in digenic analysis; it resides in an intron of the OR2L13 gene and is a transcript variant.
- Variant rs6747270 on chromosome 2 occurs in females (GWAS) and in digenic analysis and does not reside in a gene.
- Variant rs422548 occurs in all three Tables and does not reside in a gene. Its function is unknown although it must be important for the etiology of schizophrenia.
- Based on flanking variants, we can assign variant SNP6-26333526 to gene H3C6 on chromosome 6. It is significant in GWAS for males and in digenic analysis, but its function is unknown.
- Variant rs6915052 on chromosome 6 occurs in GWAS for females and in digenic analysis but does not reside in a gene.
- Variant rs7011530 on chromosome 8 is significant in GWAS for males and in digenic analysis yet does not reside in a gene.
- Another variant that does not reside in a gene is rs1582781 on chromosome 11; it occurs in all three Tables and seems highly important for schizophrenia.
- Variant rs12296316 on chromosome 12 occurs in GWAS for females and in sigenic analysis. It is an upstream transcript variant in the MGAT4C gene.
- Variant rs2496577 on chromosome 13 occurs in GWAS for males and in digenic analysis but does not reside in a gene.
- Variant rs8138145 on chromosome 22 occurs in GWAS for males and in digenic analysis but does not reside in a gene.
- Finally, variant rs5998848 on chromosome 22 occurs in GWAS for males and in digenic analysis. It is a downstream transcript variant in the LARGE1 gene.

For the 14 variant pairs with both component variants of a pair located in genes, we constructed the interaction gene network shown in Figure 1.

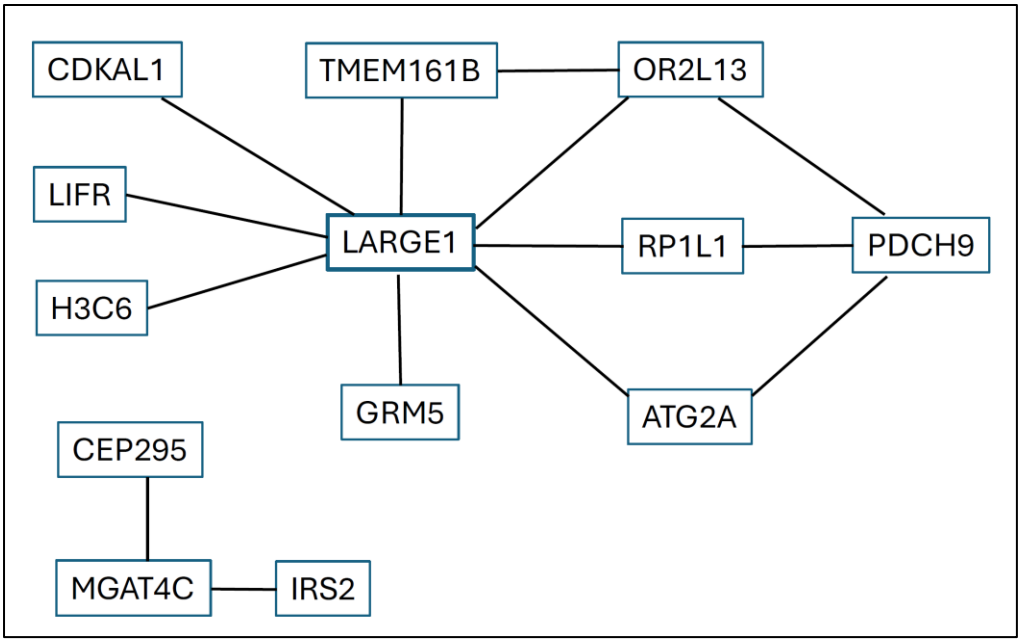

Figure 1. Connections among 13 genes in 14 significant gene pairs

Here we briefly discuss the 20 genes we have significantly identified as forming gene pairs associated with schizophrenia.

ATG2A, "autophagy-related 2A", also known as BLTP4A, has recently been mentioned as one of 11 autophagy-related [27] differentially expressed mRNA genes potentially involved in schizophrenia [28,29]. Various autophagy-related genes have previously been implicated as risk factors for schizophrenia and other psychiatric disorders [30].

CDKAL1 is a protein-coding gene without known function. Various reports have demonstrated an association

between CDKAL1 and gestational diabetes [31], and an older study showed association of CDKAL1 with type 2 diabetes [32]. These associations seem to be related to the known increased prevalence in schizophrenics of type 2 diabetes, which is associated with CDKAL1 [33].

CEP295, also known as KIAA1731 and SCKL11, is a protein coding gene that has just recently been shown to be associated with a Seckel-like syndrome involving intellectual disability and short stature [34], but we have been unable to find any references that associate the CEP295 gene with schizophrenia. Based on Supplementary Table S1, CEP295 is indirectly connected with several genes listed here via associations to variants outside of genes.

GRM5, also known as GPRC1E, MGLUR5, PPP1R86, and mGlu5, encodes a protein functioning as a metabatropic glutamate receptor, with a restricted expression toward the brain. It is an emerging target for the treatment of schizophrenia [35]. Recently, GRM2 and GRM3 have been implicated for their differential expressions in brains of schizophrenics versus controls [36]. Our results suggest that GRM5 may function in a similar manner.

H3C6, also known as H3.1, H3/d, H3C1, H3C2, H3C3, H3C4, H3C7, H3C8, H3FD, H3C10, H3C11, H3C12, and HIST1H3E, is a protein coding gene lacking introns. It encodes one of the histones responsible for the nucleosome structure. While we have not found any reports on a direct association of H3C6 with schizophrenia, post-translational modifications of histones have been suggested to play a role in the etiology of schizophrenia [37].

IRS2 has been listed as a schizophrenia-associated gene with a strong fold change in expression level compared with that in databases [20].

LARGE1, also known as LARGE, MDC1D, MDDGA6, and MDDGB6, is mostly expressed in brain and heart and, to a lesser degree, in various other tissues. We have not found much information about a direct association between this gene and schizophrenia, but a very recent report in *medRxiv* [38] lists LARGE1 as one of several genes involved in a pathway of neuroinflammatory response of the nervous system to various forms of damage, which may be connected to an inherent genetic predisposition to neurodegenerative aspects of schizophrenia.

LIFR (chr 5: 38,474,668..38,608,403 bp, complement), *Leukemia Inhibitary Factor Receptor*, is a protein coding gene, and **LIFR-AS1** (chr 5: 38,556,786..38,671,216 bp) is a noncoding RNA gene. LIFR is located at 22q12.1-q12.2, a hot spot for schizophrenia, and was associated more than ten years ago with schizophrenia [39,40]. Also, a large-scale transcriptomic meta-analysis of patient brain tissues with single-cell sequencing data of CNS neurons involving LIFR and other genes, was able to shed light on the well-known sexual dimorphism of schizophrenia [41].

MGAT4C is a protein coding gene with biased expression in thyroid, brain, and three other tissues. For several of the MGAT genes, but not for MGAT4C, post-translational protein modifications in schizophrenics have been demonstrated [42]. On the other hand, in a very recent schizophrenia case-control study of individuals of Chinese descent, recurrent somatic copy number variations were observed at several chromosomal regions including MGAT4C [43].

OR2L13, also known as OR2L14, is a protein coding gene on chr 1 (247,937,177..248,101,163 bp) and LOC105373275 is an uncharacterized non-coding RNA gene on chr 1 (248,047,705..248,095,542 bp, complement). A decrease in taste receptor expression in the brain has been reported for several genes but results for OR2L13 were not statistically significant [44]. Our results strengthen that earlier report.

PCDH9 is a protein coding gene with biased expression in brain, fat, and two other tissues. In a large GWAS for Major Depressive Disorder (MDD), PCDH9 was identified as a novel risk factor [45]. In that study, individuals with schizophrenia had been excluded, but our data strongly suggest that PCDH9 also plays a role in schizophrenia, if only through its connections with other genes. In other publications, however, PCDH9 has clearly been implicated in playing a role in familial schizophrenia [46,47].

RP1L1, also known as DCDC4B, OCMD, and RP88, is a protein coding gene and has in many publications been associated with photoreceptor diseases including macular dystrophy and retinitis pigmentosa [48]. Recently, genetic association analyses between cognitive impairment in schizophrenia showed results for large numbers of variants, including RP1L1, although the statistical significance for the involvement of RP1L1 was unclear [49].

TMEM161B, also known as FLB3342 and PRO1313, acts as a regulator of sonic hedgehog signaling and, in mouse models, plays a CNS-specific role [50]; it is also associated with defective formation of folds of the early brain development (polymicrogyria) [51]. In a case-control study of Chinese freshmen, major depressive disorder was associated with TMEM161B [52]. In that study, individuals with schizophrenia and bipolar disorder were excluded, but our analysis strongly suggests an involvement of TMEM161B in schizophrenia. Another transmembrane protein, TMEM204, has been listed as being differentially expressed in schizophrenics [53].

The following additional 7 schizophrenia-associated genes are not directly connected with other genes but with variants outside of genes:

CD226, also known as PTA1, DNAM1, DNAM-1, TLiSA1, encodes a glycoprotein on the surface of several cell types. We have not found evidence for direct association between CD226 and schizophrenia. In a mendelian randomization study [54], CD226 was one of five proteins with a causal relationship to psychiatric disorders.

E2F3 is a protein coding gene and encodes a transcription factor. In our analysis, E2F3 is significantly connected with three other genes (Figure 1). It is one of a large number of genes interacting with other genes in their relation to abnormal psychomotor behavior characteristics in schizophrenia and other severe mental disorders [55].

GPR88. The protein encoded by GPR88 is a G protein-coupled receptor with particularly robust expression in the brain [56,57]. It is emerging as a potential drug target for CNS-related diseases including schizophrenia [58].

GPR88 has been shown long ago to be a risk factor for psychiatric traits in three different populations [59]. Another G protein receptor gene, GPR56, has been published as being differentially expressed in schizophrenics [53].

IGF2BP2, also known as IMP-2, IMP2, and VICKZ2, is a protein coding gene. Significant associations between IGF2BP2 and type 2 diabetes as well as with schizophrenia have been found in Iran [60].

LINC02789 is a non-coding RNA gene. A recent study pointed out that plant-derived miRNAs can be found in the human body through eating and can then affect post-transcriptional gene regulation by binding to human mRNAs [61]. In that study, miRNAs were shown to bind to 33 human mRNAs associated with schizophrenia and other human traits. LINC02789 is one of many potential target genes of 84 wheat miRNAs identified in humans [61].

LOC107984378 is an uncharacterized non-coding RNA gene, located at chr 11p14.3 at bp 24,118,969..24,158,536. It contains an enhancer sequence, bp 24,155,823..24,156,117.

UBE2W, also known as UBC16 and UBC-16, is a protein coding gene with broad expression in brain, thyroid, and 25 other tissues. It encodes an enzyme, E2, in the ubiquitin proteasome system (UPS). In recent years, several publications reported association of schizophrenia with disruption of the UPS [62-65] although UBE2W is not generally mentioned specifically. Another ubiquitin conjugating enzyme, UBE2G1, has been implicated in schizophrenia [53].

Combining variants from our GWAS and digenic analysis, and eliminating duplicates, we wound up with 38 unique significant variants located outside of genes, shown in Supplementary Table S2.

Discussion

It has long been postulated that for common human traits, interactions among genes (and environmental effects) may be the norm rather than the exception [66]. Indeed, in our single-variant GWAS, only a relatively small number of variants were detected as being significantly associated with schizophrenia, but many more were significant based on pairs of genotypes involving different variants. Given that thousands of variants contribute to schizophrenia risk [67], it is gratifying to see that with powerful statistical methods, we can find 165 risk variants on the basis of only slightly more than 1,000 cases and 2,000 controls. Quite a few of these variants have previously been identified as being disease associated, which lends credence to our approach.

The statistical significance of our results appears highly reliable, particularly for our digenic analysis, where we had to rely on Bonferroni correction, which is known to be conservative. It is also immune to dependency among test items – genotype pairs in our situation, which are somewhat dependent as a given variant tends to occur in multiple genotype pairs. While many of our variants detected in digenic analysis are located in genes, many others were found outside of genes. The functions of these variants are unknown at this time but there can be no doubt that they are associated with schizophrenia.

A standard GWAS evaluates disease association for one variant at a time, which is most appropriate for monogenic traits. Polygenic traits like schizophrenia, however, should be addressed with methods allowing for the combined disease association of multiple variants. Early approaches in this direction considered family pedigree lod scores over multiple variants and their correlations [68], combination of *p*-values over multiple contiguous markers in the form of scan statistics [69,70], and sums of test statistics over large numbers of markers anywhere in the genome [71]. The current version of similar approaches for capturing the genetic liability to disease are polygenic risk scores (PRSs), several of which have recently been published for schizophrenia [72-74]. All these methods, including PRSs, represent aggregations of *main effects* while digenic analysis captures main and interaction effects although only over two variants at a time. A combination of multiple genotype pairs, perhaps over thousands of them, in the form of a polygenic risk score would presumably capture both main and interaction effects for large numbers of variants. We plan to develop such an approach.

Conclusion

Our digenic analysis has uncovered or confirmed 36 significant variants, quite a few more than the 14 significant variants found in standard GWAS. The value of our contribution is that it confirms previous tentative associations and points to new assignments not previously known, which are worth being followed up.

Conflicts of Interest Statement

The authors have no conflicts of interest to declare.

Acknowledgments

We are greatly indebted to Dr. Manoj Kandpal, Rockefeller University, for invaluable expertise in downloading and decrypting the data from dbGaP, supported in part by grant # UL1TR001866 from the NIH National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (NIH), Clinical and Translational Science Award (CTSA) program.

The datasets used for the analysis described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000448.v1.p1. Submission of the data, phs000448.v1.p1, to dbGaP was provided by Dr. Todd Lencz and on behalf of himself and his collaborator, Ariel Darvasi, Ph.D.. Support for the collection and analysis of the datasets was provided by RC2MH089964, R01MH084098, the North Shore - LIJ Health System Foundation, and the Hebrew University Genetic Resource.

Internet links

Gpairs program:
https://lab.rockefeller.edu/ott/programs/GPM,
https://github.com/jurgott/

Human Gene Database for schizophrenia:
https://www.genecards.org/Search/Keyword?queryString=Schizophrenia&sort=Score&sortdir=Descending&startPage=0&pageSize=-1

References

1. Tandon R, Nasrallah H, Akbarian S, et al. The schizophrenia syndrome, circa 2024: What we know and how that informs its nature. *Schizophr Res*. Dec 11 2023;264:1-28. doi:10.1016/j.schres.2023.11.015
2. R R, Devtalla H, Rana K, et al. A comprehensive update on genetic inheritance, epigenetic factors, associated pathology, and recent therapeutic intervention by gene therapy in schizophrenia. *Chem Biol Drug Des*. Jan 2024;103(1):e14374. doi:10.1111/cbdd.14374
3. Liu D, Meyer D, Fennessy B, et al. Schizophrenia risk conferred by rare protein-truncating variants is conserved across diverse human populations. *Nature Genetics*. 2023/03/01 2023;55(3):369-376. doi:10.1038/s41588-023-01305-1
4. McClellan JM, King M-C. A tipping point in neuropsychiatric genetics. *Neuron*. 2021/05/05/ 2021;109(9):1411-1413. doi:10.1016/j.neuron.2021.04.002
5. Lencz T, Yu J, Khan RR, et al. Novel ultra-rare exonic variants identified in a founder population implicate cadherins in schizophrenia. *Neuron*. 2021/05/05/ 2021;109(9):1465-1478.e4. doi:10.1016/j.neuron.2021.03.004
6. Schulze TG, McMahon FJ. Genetic association mapping at the crossroads: which test and why? Overview and practical guidelines. *Am J Med Genet*. Jan 8 2002;114(1):1-11. doi:10.1002/ajmg.10042
7. Uffelmann E, Huang QQ, Munung NS, et al. Genome-wide association studies. *Nature Reviews Methods Primers*. 2021/08/26 2021;1(1):59. doi:10.1038/s43586-021-00056-9
8. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am J Hum Genet*. Sep 2008;83(3):311-21. doi:10.1016/j.ajhg.2008.06.024
9. Trubetskoy V, Pardiñas AF, Qi T, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature*. 2022/04/01 2022;604(7906):502-508. doi:10.1038/s41586-022-04434-5
10. Nakamura T, Takata A. The molecular pathology of schizophrenia: an overview of existing knowledge and new directions for future research. *Molecular Psychiatry*. 2023/05/01 2023;28(5):1868-1889. doi:10.1038/s41380-023-02005-2
11. Ott J, Sun D. Multilocus association analysis under polygenic models. *International journal of data mining and bioinformatics*. 2012;6(5):482-9. doi:10.1504/IJDMB.2012.049302
12. Okazaki A, Ott J. Machine learning approaches to explore digenic inheritance. *Trends Genet*. Oct 2022;38(10):1013-1018. doi:10.1016/j.tig.2022.04.009
13. Zhang Q, Long Q, Ott J. AprioriGWAS, a new pattern mining strategy for detecting genetic variants associated with disease through interaction effects. Research Support, Non-U.S. Gov't. *PLoS Comput Biol*. Jun 2014;10(6):e1003627. doi:10.1371/journal.pcbi.1003627
14. Zhang Q, Bhatia M, Park T, Ott J. A multi-threaded approach to genotype pattern mining for detecting digenic disease genes. *Front Genet*. 2023;14:1222517. doi:10.3389/fgene.2023.1222517
15. Lencz T, Guha S, Liu C, et al. Genome-wide association study implicates NDST3 in schizophrenia and bipolar disorder. *Nature communications*. 2013/11/19 2013;4(1):2739. doi:10.1038/ncomms3739
16. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7. doi:10.1186/s13742-015-0047-8
17. Duncan LE, Ratanatharathorn A, Aiello AE, et al. Largest GWAS of PTSD (N=20,070) yields genetic overlap with schizophrenia and sex differences in heritability. *Mol Psychiatry*. Mar 2018;23(3):666-673. doi:10.1038/mp.2017.77
18. Judson O. *Dr. Tatiana's Sex Advice to All Creation: The Definitive Guide to the Evolutionary Biology of Sex*. Holt Paperbacks; 2003:320.
19. Carlisle SM, Qin H, Hendrickson RC, et al. Sex-based differences in the activation of peripheral blood monocytes in early Parkinson disease. *npj Parkinson's Disease*. 2021/04/13 2021;7(1):36. doi:10.1038/s41531-021-00180-z
20. Koch E, Kauppi K, Chen C-H. Candidates for drug repurposing to address the cognitive symptoms in schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*. 2023/01/10/ 2023;120:110637. doi:10.1016/j.pnpbp.2022.110637
21. Hukic DS, Lavebratt C, Olsson E, et al. Troponin T levels associated with genetic variants in NOTCH2 and MTNR1B in women with psychosis. *Psychiatry research*. 2017/04// 2017;250:217-220. doi:10.1016/j.psychres.2017.01.030
22. Hajek C, Guo X, Yao J, et al. Coronary Heart Disease Genetic Risk Score Predicts Cardiovascular Disease Risk in Men, Not Women. *Circulation: Genomic and Precision Medicine*. 2018/10/01 2018;11(10):e002324. doi:10.1161/CIRCGEN.118.002324
23. Saha E, Ben Guebila M, Fanfani V, et al. Gene regulatory networks reveal sex difference in lung adenocarcinoma. *Biology of Sex Differences*. 2024/08/06 2024;15(1):62. doi:10.1186/s13293-024-00634-y
24. Hjorthøj C, Compton W, Starzer M, et al. Association between cannabis use disorder and schizophrenia stronger in young males than in females. *Psychological Medicine*. 2023;53(15):7322-7328. doi:10.1017/S0033291723000880
25. Zhao H, Mitra N, Kanetsky PA, Nathanson KL, Rebbeck TR. A practical approach to adjusting for population stratification in genome-wide association studies: principal components and propensity scores (PCAPS). *Statistical Applications in Genetics and Molecular Biology*. 2018;17(6)doi:10.1515/sagmb-2017-0054
26. Elhaik E. Principal Component Analyses (PCA)-based findings in population genetic studies are highly biased and must be reevaluated. *Scientific reports*. 2022/08/29 2022;12(1):14683. doi:10.1038/s41598-022-14395-4
27. Grosjean I, Roméo B, Domdom M-A, et al. Autophagopathies: from autophagy gene

polymorphisms to precision medicine for human diseases. *Autophagy*. 2022/11/02 2022;18(11):2519-2536. doi:10.1080/15548627.2022.2039994
28. Li R, Wang Q, Qiu Y, et al. A Potential Autophagy-Related Competing Endogenous RNA Network and Corresponding Diagnostic Efficacy in Schizophrenia. *Front Psychiatry*. 2021;12:628361. doi:10.3389/fpsyt.2021.628361
29. Sepúlveda D, Grunenwald F, Vidal A, et al. Insulin-like growth factor 2 and autophagy gene expression alteration arise as potential biomarkers in Parkinson's disease. *Scientific reports*. 2022/02/07 2022;12(1):2038. doi:10.1038/s41598-022-05941-1
30. Tan Y, Zhu J, Hashimoto K. Autophagy-related gene model as a novel risk factor for schizophrenia. *Translational Psychiatry*. 2024/02/13 2024;14(1):94. doi:10.1038/s41398-024-02767-5
31. Huang C, Guo Y, Li W, et al. Association of the CDKAL1 gene polymorphism with gestational diabetes mellitus in Chinese women. *BMJ Open Diabetes Res Care*. Apr 2023;11(2). doi:10.1136/bmjdrc-2022-003164
32. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nature Genetics*. 2007/06/01 2007;39(6):770-775. doi:10.1038/ng2043
33. Yin XY, Chen P, Zhu HW, et al. The type 2 diabetes mellitus susceptibility gene CDKAL1 polymorphism is associated with depressive symptom in first-episode drug-naive schizophrenic patients. *Human Psychopharmacology: Clinical and Experimental*. 2021;36(5):e2790. doi:10.1002/hup.2790
34. Li N, Xu Y, Chen H, et al. Bi-allelic variants in CEP295 cause Seckel-like syndrome presenting with primary microcephaly, developmental delay, intellectual disability, short stature, craniofacial and digital abnormalities. *EBioMedicine*. Jan 2024;99:104940. doi:10.1016/j.ebiom.2023.104940
35. Dogra S, Conn PJ. Metabotropic glutamate receptors as emerging targets for the treatment of schizophreni. *Molecular Pharmacology*. 2022:MOLPHARM-MR-2021-000460. doi:10.1124/molpharm.121.000460
36. DelaCuesta-Barrutia J, Martínez-Peula O, Rivero G, et al. Effect of antipsychotic drugs on group II metabotropic glutamate receptor expression and epigenetic control in postmortem brains of schizophrenia subjects. *Transl Psychiatry*. Feb 23 2024;14(1):113. doi:10.1038/s41398-024-02832-z
37. Delphin N, Aust C, Griffiths L, Fernandez F. Epigenetic Regulation in Schizophrenia: Focus on Methylation and Histone Modifications in Human Studies. *Genes*. 2024;15(3):272. doi:10.3390/genes15030272
38. Rabe F, Smigielski L, Georgiadis F, et al. Genetic susceptibility to schizophrenia through neuroinflammatory pathways is associated with retinal thinning: Findings from the UK-Biobank. *medRxiv*. 2024:2024.04.05.24305387. doi:10.1101/2024.04.05.24305387
39. Okahisa Y, Ujike H, Kunugi H, et al. Leukemia inhibitory factor gene is associated with schizophrenia and working memory function. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*. 2010/02/01/ 2010;34(1):172-176. doi:10.1016/j.pnpbp.2009.10.020
40. Kang WS, Kim SK, Park JK, et al. Association between promoter polymorphisms of the LIFR gene and schizophrenia with persecutory delusion in a Korean population. *Mol Med Rep*. Jan 2012;5(1):270-4. doi:10.3892/mmr.2011.618
41. Lobentanzer S, Hanin G, Klein J, Soreq H. Integrative Transcriptomics Reveals Sexually Dimorphic Control of the Cholinergic/Neurokine Interface in Schizophrenia and Bipolar Disorder. *Cell Rep*. Oct 15 2019;29(3):764-777.e5. doi:10.1016/j.celrep.2019.09.017
42. Mueller TM, Meador-Woodruff JH. Post-translational protein modifications in schizophrenia. *NPJ Schizophr*. Mar 2 2020;6(1):5. doi:10.1038/s41537-020-0093-9
43. Chang K, Jian X, Wu C, et al. The contribution of mosaic chromosomal alterations to schizophrenia. *Biological Psychiatry*. 2024;doi:10.1016/j.biopsych.2024.06.015
44. Ansoleaga B, Garcia-Esparcia P, Pinacho R, Haro JM, Ramos B, Ferrer I. Decrease in olfactory and taste receptor expression in the dorsolateral prefrontal cortex in chronic schizophrenia. *J Psychiatr Res*. Jan 2015;60:109-16. doi:10.1016/j.jpsychires.2014.09.012
45. Xiao X, Zheng F, Chang H, et al. The Gene Encoding Protocadherin 9 (PCDH9), a Novel Risk Factor for Major Depressive Disorder. *Neuropsychopharmacology*. Apr 2018;43(5):1128-1137. doi:10.1038/npp.2017.241
46. Lee CC, Ye R, Tubbs JD, et al. Third-generation genome sequencing implicates medium-sized structural variants in chronic schizophrenia. *Front Neurosci*. 2022;16:1058359. doi:10.3389/fnins.2022.1058359
47. Yamamoto K, Kuriu T, Matsumura K, et al. Multiple alterations in glutamatergic transmission and dopamine D2 receptor splicing in induced pluripotent stem cell-derived neurons from patients with familial schizophrenia. *Translational Psychiatry*. 2021/10/25 2021;11(1):548. doi:10.1038/s41398-021-01676-1
48. Liu J, Hayden MR, Yang Y. Research progress of RP1L1 gene in disease. *Gene*. Jun 20 2024;912:148367. doi:10.1016/j.gene.2024.148367
49. Zhuo C, Tian H, Chen J, et al. Associations of cognitive impairment in patients with schizophrenia with genetic features and with schizophrenia-related structural and functional brain changes. *Front Genet*. 2022;13:880027. doi:10.3389/fgene.2022.880027
50. Akula SK, Marciano JH, Lim Y, et al. TMEM161B regulates cerebral cortical gyration, Sonic Hedgehog signaling, and ciliary structure in the developing central nervous system. *Proceedings of the National Academy of Sciences*. 2023/01/24 2023;120(4):e2209964120. doi:10.1073/pnas.2209964120
51. Wang L, Heffner C, Vong KI, et al. TMEM161B modulates radial glial scaffolding in neocortical development. *Proceedings of the National Academy of Sciences*. 2023/01/24 2023;120(4):e2209983120. doi:10.1073/pnas.2209983120

52. Yue S, Luo L, Feng Y, et al. Gene-gene interaction and new onset of major depressive disorder: Findings from a Chinese freshmen nested case-control study. *Journal of Affective Disorders*. 2022/03/01/ 2022;300:505-510. doi:10.1016/j.jad.2021.12.138
53. Merikangas AK, Shelly M, Knighton A, Kotler N, Tanenbaum N, Almasy L. What genes are differentially expressed in individuals with schizophrenia? A systematic review. *Molecular Psychiatry*. 2022/03/01 2022;27(3):1373-1383. doi:10.1038/s41380-021-01420-7
54. Chen Z, Wang X, Teng Z, et al. Modifiable lifestyle factors influencing psychiatric disorders mediated by plasma proteins: A systemic Mendelian randomization study. *Journal of Affective Disorders*. 2024/04/01/ 2024;350:582-589. doi:10.1016/j.jad.2024.01.169
55. Zhang Y, Zhao Y, Song X, et al. Modulation of Stem Cells as Therapeutics for Severe Mental Disorders and Cognitive Impairments. *Front Psychiatry*. 2020;11:80. doi:10.3389/fpsyt.2020.00080
56. Ye N, Li B, Mao Q, et al. Orphan Receptor GPR88 as an Emerging Neurotherapeutic Target. *ACS Chemical Neuroscience*. 2019/01/16 2019;10(1):190-200. doi:10.1021/acschemneuro.8b00572
57. Chen G, Xu J, Inoue A, et al. Activation and allosteric regulation of the orphan GPR88-Gi1 signaling complex. *Nature communications*. 2022/05/02 2022;13(1):2375. doi:10.1038/s41467-022-30081-5
58. Chen G, Xu J, Inoue A, et al. Activation and allosteric regulation of the orphan GPR88-Gi1 signaling complex. *Nature communications*. May 2 2022;13(1):2375. doi:10.1038/s41467-022-30081-5
59. Del Zompo M, Deleuze JF, Chillotti C, et al. Association study in three different populations between the GPR88 gene and major psychoses. *Mol Genet Genomic Med*. Mar 2014;2(2):152-9. doi:10.1002/mgg3.54
60. Sargazi S, Heidari Nia M, Saravani R, Jafari Shahroudi M, Jahantigh D, Shakiba M. IGF2BP2 polymorphisms as genetic biomarkers for either schizophrenia or type 2 diabetes mellitus: A case-control study. *Gene Reports*. 2020/09/01/ 2020;20:100680. doi:10.1016/j.genrep.2020.100680
61. Sánchez-Romo D, Hernández-Vásquez CI, Pereyra-Alférez B, García-García JH. Identification of potential target genes in Homo sapiens, by miRNA of Triticum aestivum: A cross kingdom computational approach. *Noncoding RNA Res*. Jun 2022;7(2):89-97. doi:10.1016/j.ncrna.2022.03.002
62. Bousman CA, Chana G, Glatt SJ, et al. Preliminary evidence of ubiquitin proteasome system dysregulation in schizophrenia and bipolar disorder: Convergent pathway analysis findings from two independent samples. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 2010/03/01 2010;153B(2):494-502. doi:10.1002/ajmg.b.31006
63. Andrews JL, Goodfellow FJ, Matosin N, et al. Alterations of ubiquitin related proteins in the pathology and development of schizophrenia: Evidence from human and animal studies. *Journal of Psychiatric Research*. 2017/07/01/ 2017;90:31-39. doi:10.1016/j.jpsychires.2017.01.009
64. Bousman CA, Luza S, Mancuso SG, et al. Elevated ubiquitinated proteins in brain and blood of individuals with schizophrenia. *Scientific reports*. Feb 19 2019;9(1):2307. doi:10.1038/s41598-019-38490-1
65. Luza S, Opazo CM, Bousman CA, Pantelis C, Bush AI, Everall IP. The ubiquitin proteasome system and schizophrenia. *The Lancet Psychiatry*. 2020/06/01/ 2020;7(6):528-537. doi:10.1016/S2215-0366(19)30520-6
66. McKinney BA, Reif DM, Ritchie MD, Moore JH. Machine Learning for Detecting Gene-Gene Interactions. *Applied Bioinformatics*. 2006/06/01 2006;5(2):77-88. doi:10.2165/00822942-200605020-00002
67. Purcell SM, Wray NR, Stone JL, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. Aug 6 2009;460(7256):748-52. doi:10.1038/nature08185
68. MacLean CJ, Sham PC, Kendler KS. Joint linkage of multiple loci for a complex disorder. *Am J Hum Genet*. Aug 1993;53(2):353-66.
69. Hoh J, Ott J. Scan statistics to scan markers for susceptibility genes. *Proc Natl Acad Sci USA*. 2000;97(17):9615-7. doi:10.1073/pnas.170179197
70. Hoh J, Ott J. Scan statistics in genome-wide scan for complex trait loci. In: Glaz J, Pozdnyakov V, Wallenstein S, eds. *Scan Statistics: Methods and Applications (Statistics for Industry and Technology)*. Birkhäuser Boston; 2009:197-204.
71. Hoh J, Wille A, Ott J. Trimming, weighting, and grouping SNPs in human case-control association studies. *Genome Res*. 2001;11(12):2115-9. doi:10.1101/gr.204001
72. Legge SE, Cardno AG, Allardyce J, et al. Associations Between Schizophrenia Polygenic Liability, Symptom Dimensions, and Cognitive Ability in Schizophrenia. *JAMA Psychiatry*. 2021;78(10):1143-1151. doi:10.1001/jamapsychiatry.2021.1961
73. Landi I, Kaji DA, Cotter L, et al. Prognostic value of polygenic risk scores for adults with psychosis. *Nature Medicine*. 2021/09/01 2021;27(9):1576-1581. doi:10.1038/s41591-021-01475-7
74. Barbu MC, Viejo-Romero M, Thng G, et al. Pathway-Based Polygenic Risk Scores for Schizophrenia and Associations With Reported Psychotic-like Experiences and Neuroimaging Phenotypes in the UK Biobank. *Biological Psychiatry Global Open Science*. 2023/10/01/ 2023;3(4):814-823. doi:10.1016/j.bpsgos.2023.03.004

Supplementary Table S1. List of 76 variant pairs in males and females furnishing significant genotype pairs.

ch1	rs1	bp1 GRCh38	ch2	rs2	bp2 GRCh38	sex	gene1	gene2
1	rs2518417	100,539,811	2	rs6747270	13,446,506	m	GPR88	—
1	rs6675786	248,066,288	5	rs16903025	88,245,670	m	OR2L13, LOC105373275	TMEM161B
1	rs7340057	73,269,583	5	rs16903025	88,245,670	m	—	TMEM161B
1	rs6675786	248,066,288	13	rs1755808	66,304,158	m	OR2L13, LOC105373275	PCDH9
1	rs7340057	73,269,583	13	rs1755808	66,304,158	m	—	PCDH9
1	rs7530249	199,369,333	13	rs2496577	38,226,298	m	LINC02789	—
1	rs6675786	248,066,288	22	rs5998848	33,288,529	m	OR2L13, LOC105373275	LARGE1
1	rs7340057	73,269,583	22	rs5998848	33,288,529	m	—	LARGE1
1	rs6675786	248,066,288	22	rs8138145	22,146,185	m	OR2L13, LOC105373275	—
1	rs7340057	73,269,583	22	rs8138145	22,146,185	m	—	—
2	rs6747270	13,446,506	5	rs16903025	88,245,670	m	—	TMEM161B
2	rs6747270	13,446,506	11	SNP11-93097180	*93,097,180*	f	—	CEP295
2	rs6747270	13,446,506	13	rs1755808	66,304,158	m	—	PCDH9
2	rs6747270	13,446,506	22	rs5998848	33,288,529	m	—	LARGE1
2	rs6747270	13,446,506	22	rs8138145	22,146,185	m	—	—
3	rs11924588	18,908,029	22	rs8138145	22,146,185	m	—	—
3	SNP3-186943722	*186,943,722*	22	rs8138145	22,146,185	m	IGF2BP2	—
4	rs422548	68,745,554	5	rs16903025	88,245,670	m	—	TMEM161B
4	rs422548	68,745,554	11	SNP11-93097180	*93,097,180*	f	—	CEP295
4	rs422548	68,745,554	12	rs12300389	90,558,317	m	—	—
4	rs422548	68,745,554	13	rs1755808	66,304,158	m	—	PCDH9
4	rs422548	68,745,554	18	rs3937015	69,932,280	m	—	CD226
4	rs422548	68,745,554	22	rs5998848	33,288,529	m	—	LARGE1
4	rs422548	68,745,554	22	rs8138145	22,146,185	m	—	—
5	rs16903025	88,245,670	8	rs7011530	125,838,349	m	TMEM161B	—
5	rs16903025	88,245,670	9	rs7865153	126,799,253	m	TMEM161B	—
5	rs16903025	88,245,670	11	rs1582781	12,561,634	m	TMEM161B	—
5	rs16903025	88,245,670	13	rs2496577	38,226,298	m	TMEM161B	—
5	rs10041314	38,584,216	22	rs5998848	33,288,529	m	LIFR,LIFR-AS1	LARGE1
5	rs16903025	88,245,670	22	rs5998848	33,288,529	m	TMEM161B	LARGE1
5	rs10041314	38,584,216	22	rs8138145	22,146,185	m	LIFR,LIFR-AS1	—
5	rs16903025	88,245,670	22	rs8138145	22,146,185	m	TMEM161B	—
6	rs4134989	20,493,230	11	rs1582781	12,561,634	f	E2F3	—
6	rs6915052	121,376,832	13	rs1755808	66,304,158	m	—	PCDH9
6	rs6915052	121,376,832	22	rs5998848	33,288,529	m	—	LARGE1
6	SNP6-20737399	*20,737,399*	22	rs5998848	33,288,529	m	CDKAL1	LARGE1
6	SNP6-26333526	*26,333,526*	22	rs5998848	33,288,529	m	H3C6	LARGE1
6	rs6915052	121,376,832	22	rs8138145	22,146,185	m	—	—
6	SNP6-20737399	*20,737,399*	22	rs8138145	22,146,185	m	CDKAL1	—
6	SNP6-26333526	*26,333,526*	22	rs8138145	22,146,185	m	H3C6	—
7	rs28879680	91,748,465	22	rs5998848	33,288,529	m	—	LARGE1

7	rs28879680	91,748,465	22	rs8138145	22,146,185	m	—	—
8	rs7011530	125,838,349	13	rs1755808	66,304,158	m	—	PCDH9
8	SNP8-10507540	*10,507,540*	13	rs1755808	66,304,158	m	RP1L1	PCDH9
8	rs11987092	122,014,250	22	rs5998848	33,288,529	m	—	LARGE1
8	rs7011530	125,838,349	22	rs5998848	33,288,529	m	—	LARGE1
8	SNP8-10507540	*10,507,540*	22	rs5998848	33,288,529	m	RP1L1	LARGE1
8	rs11987092	122,014,250	22	rs8138145	22,146,185	m	—	—
8	rs6982398	73,833,974	22	rs8138145	22,146,185	m	UBE2W	—
8	rs7011530	125,838,349	22	rs8138145	22,146,185	m	—	—
9	rs7865153	126,799,253	13	rs1755808	66,304,158	m	—	PCDH9
9	rs7865153	126,799,253	22	rs5998848	33,288,529	m	—	LARGE1
9	rs7865153	126,799,253	22	rs8138145	22,146,185	m	—	—
10	rs2813274	50,148,357	22	rs5998848	33,288,529	m	—	LARGE1
10	rs2813274	50,148,357	22	rs8138145	22,146,185	m	—	—
11	rs1582781	12,561,634	11	SNP11-93097180	*93,097,180*	f	—	CEP295
11	rs1582781	12,561,634	12	rs12296316	86,280,211	f	—	MGAT4C
11	SNP11-93097180	*93,097,180*	12	rs12296316	86,280,211	f	CEP295	MGAT4C
11	rs1582781	12,561,634	12	rs12300389	90,558,317	m	—	—
11	rs1582781	12,561,634	12	rs16933566	28,905,298	m	—	—
11	rs1582781	12,561,634	13	rs1755808	66,304,158	m	—	PCDH9
11	SNP11-64433849	*64,433,849*	13	rs1755808	66,304,158	m	ATG2A	PCDH9
11	rs1582781	12,561,634	22	rs5998848	33,288,529	m	—	LARGE1
11	rs597303	88,969,779	22	rs5998848	33,288,529	m	GRM5	LARGE1
11	SNP11-64433849	*64,433,849*	22	rs5998848	33,288,529	m	ATG2A	LARGE1
11	rs1582781	12,561,634	22	rs8138145	22,146,185	m	—	—
11	rs597303	88,969,779	22	rs8138145	22,146,185	m	GRM5	—
11	SNP11-24126646	*24,126,646*	22	rs8138145	22,146,185	m	LOC107984378	—
12	rs12296316	86,280,211	13	rs1044364	109,755,299	f	MGAT4C	IRS2
12	rs12300389	90,558,317	13	rs2496577	38,226,298	m	—	—
12	rs16933566	28,905,298	13	rs2496577	38,226,298	m	—	—
12	rs11107087	93,545,255	22	rs5998848	33,288,529	m	—	LARGE1
12	rs11107087	93,545,255	22	rs8138145	22,146,185	m	—	—
13	rs2496577	38,226,298	13	rs1755808	66,304,158	m	—	PCDH9
13	rs2496577	38,226,298	22	rs5998848	33,288,529	m	—	LARGE1
13	rs2496577	38,226,298	22	rs8138145	22,146,185	m	—	—

Supplementary Table S2. All 38 significant unique variants as obtained in single-variant and digenic analysis, where *n* indicates the number of other variants connected with the given variant.

ch	rsID	bp GRCh38	source	n	gene	function
1	rs7340057	73,269,583	m	3	—	—
1	rs2518417	100,539,811	m	0	GPR88	missense variant, coding sequence variant
1	rs7530249	199,369,333	m	0	LINC02789	intron variant
1	rs6675786	248,066,288	m	3	OR2L13, LOC105373275	intron variant, genic upstream transcript variant
2	rs6747270	13,446,506	mf	5	—	—
3	rs11924588	18,908,029	m	0	—	—
3	SNP3-186943722	*186,943,722*	m	0	IGF2BP2	—
4	rs422548	68,745,554	mf	6	—	—
5	rs10041314	38,584,216	m	1	LIFR,LIFR-AS1	intron variant, genic upstream transcript variant
5	rs16903025	88,245,670	m	9	TMEM161B	intron variant, genic upstream transcript variant
6	rs4134989	20,493,230	f	0	E2F3	genic downstream transcript variant, 3'UTR variant
6	SNP6-20737399	*20,737,399*	m	1	CDKAL1	—
6	SNP6-26333526	*26,333,526*	m	1	H3C6	—
6	rs6915052	121,376,832	m	2	—	—
7	rs28879680	91,748,465	m	1	—	—
8	SNP8-10507540	*10,507,540*	m	1	RP1L1	—
8	rs6982398	73,833,974	m	0	UBE2W	intron variant
8	rs11987092	122,014,250	m	1	—	—
8	rs7011530	125,838,349	m	3	—	—
9	rs7865153	126,799,253	m	3	—	—
10	rs2813274	50,148,357	m	1	—	—
11	rs1582781	12,561,634	mf	8	—	—
11	SNP11-24126646	*24,126,646*	m	0	LOC107984378	—
11	SNP11-64433849	*64,433,849*	m	1	ATG2A	—
11	rs597303	88,969,779	m	1	GRM5	intron variant
11	SNP11-93097180	*93,097,180*	f	3	CEP295	—
12	rs16933566	28,905,298	m	1	—	—
12	rs12296316	86,280,211	f	2	MGAT4C	intron variant, genic upstream transcript variant
12	rs12300389	90,558,317	m	2	—	—
12	rs11107087	93,545,255	m	1	—	—
13	rs2496577	38,226,298	m	6	—	—
13	rs1755808	66,304,158	m	10	PCDH9	3'UTR variant, genic downstream transcript variant
13	rs1044364	109,755,299	f	0	IRS2	3'UTR variant
18	rs3937015	69,932,280	m	0	CD226	intron variant, genic upstream transcript variant
22	rs8138145	22,146,185	m	21	—	—
22	rs5998848	33,288,529	m	19	LARGE1	intron variant, genic downstream transcript variant