



Published in final edited form as:

Alcohol Clin Exp Res. 2020 December ; 44(12): 2494–2518. doi:10.1111/acer.14492.

Identification of Functional Genetic Variants Associated with Alcohol Dependence and Related Phenotypes Using a High-Throughput Assay

Kriti S. Thapa, Ph.D.^{1,*}, Andy B Chen, Ph.D.², Dongbing Lai, Ph.D.², Xiaoling Xuei, Ph.D.², Leah Wetherill, Ph.D.², Jay A. Tischfield, Ph.D.³, Yunlong Liu, Ph.D.², Howard J. Edenberg, Ph.D.^{1,2}

¹Department of Biochemistry & Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana.

²Department of Medical & Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana.

³Department of Genetics, Rutgers University, Piscataway, NJ, 99999, USA

Abstract

Background: Genome Wide Association Studies (GWAS) of Alcohol Dependence (AD) and related phenotypes have identified multiple loci, but the functional variants underlying the loci have in most cases not been identified. Non-coding variants can influence phenotype by affecting gene expression; for example, variants in the 3' untranslated regions (3'UTR) can affect gene expression post-transcriptionally.

Methods: We adapted a high-throughput assay known as PASSPORT-seq (parallel assessment of polymorphisms in miRNA target-sites by sequencing) to identify among variants associated with AD and related phenotypes those that cause differential expression in neuronal cell lines. Based upon meta-analyses of alcohol-related traits in African American and European Americans in the Collaborative Study on the Genetics of Alcoholism, we tested 296 single nucleotide polymorphisms (SNPs with meta-analysis *p* values < 0.001) that were located in 3'UTRs.

Results: We identified 60 SNPs that affected gene expression (FDR < 0.05) in SH-SY5Y cells and 92 that affected expression in SK-N-BE(2) cells. Among these, 30 SNPs altered RNA levels in the same direction in both cell lines. Many of these SNPs reside in the binding sites of miRNAs and RNA-binding proteins and are expression quantitative trait loci (eQTLs) of genes including *KIF6*, *FRMD4A*, *CADM2*, *ADD2*, *PLK2*, and *GAS7*.

Conclusion: The SNPs identified in the PASSPORT-seq assay are functional variants that might affect the risk for AD and related phenotypes. Our study provides insights into gene regulation in

Address correspondence to: Howard J. Edenberg, Department of Biochemistry & Molecular Biology, Indiana University School of Medicine, 635 Barnhill Drive, MS4063, Indianapolis, Indiana 46202, **Telephone:** +1 317 274 2353, **Fax:** +1 317 274 4686, edenberg@iu.edu.

*Current address: Kriti S. Thapa, GENEWIZ, 2920 Fortune Circle West, Suite A, Indianapolis, Indiana 46241

The authors report no conflicts of interest.

AD and demonstrates the value of PASSPORT-seq as a tool to screen genetic variants in GWAS loci for one potential mechanism of action.

Keywords

Alcohol dependence; GWAS; gene expression; 3'Untranslated regions; neuronal cells

Introduction

Alcohol Dependence (AD) is a global health problem and one of the leading risk factors for death and disability (Collaborators, 2018, Rehm and Shield, 2019, Edenberg and Foroud, 2013). Alcohol dependence was defined in the fourth edition (revised) of the American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders (DSM–IV) (American Psychiatric Association, 2000) by the presence of 3 or more of 7 criteria in a 12-month period. Criteria are tolerance, withdrawal, drinking more than intended, desire to cut drinking, giving up activities, time spent drinking and excessive alcohol consumption despite problems. The fifth edition of the American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association, 2013) integrated AD and alcohol abuse into a single disorder, alcohol use disorder (AUD; for a comparison see NIH Publication No: 13–7999). There are 11 criteria in DSM-5 (that include those in DSM-IV), and an individual meeting of any two of the 11 criteria in a 12-month period is diagnosed with AUD.

AD and AUD are complex disorders in which both genetic differences and environmental factors affect the risk (Walters et al., 2018, Edenberg and Foroud, 2013, Edenberg and McClintick, 2018, Hart and Kranzler, 2015, Rietschel and Treutlein, 2013). Understanding the genetic differences associated with the risk of AD and AUD will provide insights into the biological mechanisms underlying the disease. They are heterogeneous disorders, and variations in many genes influence the risk. Individual variants make only small contributions to the risk (Edenberg and Foroud, 2013, Gelernter et al., 2014, Gelernter et al., 2019, Walters et al., 2018, Zhou et al., 2019), with the exception of the functional variants in two of the alcohol-metabolizing enzymes (*ADH1B* and *ALDH2*) (Edenberg and McClintick, 2018). Thus, identification of genetic variations that impact the risk of AD/AUD is challenging. With the idea that the heterogeneity may well reflect a range of genetic contributions to individual criteria and related traits, Genome Wide Association Studies (GWAS) have been carried out on AD and on individual criteria and related phenotypes. These have identified multiple loci (Walters et al., 2018, Lai et al., 2019a, Lai et al., 2019b, Wetherill et al., 2019, Gelernter et al., 2019, Zhou et al., 2019, Kranzler et al., 2019); however, the functional variants underlying those loci have not been characterized. The genome consists primarily of non-coding SNPs; therefore, it is not surprising that more than 90% of the variants identified by GWAS of complex traits reside in non-coding regions; however, the significant variants are enriched for potential regulatory elements (Maurano et al., 2012).

Variants in the 3' untranslated regions (3'UTR) may act as *cis*-regulators to modulate transcript levels. A very sensitive way to detect the effects of a variant in the 3'UTR is to

examine differential expression of each allele of a single nucleotide polymorphism (SNP) inserted into a reporter gene (allele-biased expression). We recently developed a high-throughput assay, PASSPORT-seq (parallel assessment of polymorphisms in miRNA target-sites by sequencing) (Ipe et al., 2018, Rao et al., 2019), that can simultaneously and efficiently detect whether hundreds of SNPs are *cis*-regulators. This assay combines pooled synthesis of oligonucleotides, bulk cloning into a reporter vector, transfection of the pool into mammalian cells, and next generation sequencing (NGS) to measure differential expression. In this study, we use this assay to test the potential effect of SNPs in loci associated with AD and related phenotypes (with meta-analysis *p* values < 0.001; (Lai et al., 2019a, Lai et al., 2019b)).

Materials and Methods

Selection of 3'UTR SNPs

The Collaborative study on the Genetics of Alcoholism (COGA) (Reich et al., 1998, Edenberg, 2002) is a multi-center collaboration study aimed at identifying genes/variants that affect the risk of AD and related phenotypes. AD probands and their family members were recruited from inpatient and outpatient AD treatment facilities in seven sites; community comparison families were recruited from a variety of sources in the same areas. This study was approved by institutional review boards from all sites and every participant provided informed consent or assent. The Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) and the child version of the SSAGA (Bucholz et al., 1994, Hesselbrock et al., 1999) were used for participants age 18 or over and younger than 18, respectively. COGA samples were genotyped on four genome wide genotyping arrays; genotyping data processing and QC were reported previously (Lai et al., 2019a, Lai et al., 2019b). GWAS of 11 phenotypes in the COGA African American (AA) and European American (EA) samples, as well as meta-analyses of them have been published (Lai et al., 2019a, Lai et al., 2019b, Wetherill et al., 2019). For the current study, we performed GWAS for another seven alcohol related phenotypes, making 18 GWAS in total (Supplemental Table S1). We then selected for PASSPORT-seq analysis 296 variants that are located in the 3' UTR of genes (defined by NCBI RefSeq annotation) and had *P*-values < 0.001 in any of those GWAS. An advantage of high-throughput assays is the ability to test many SNPs that are not genome-wide significant in current underpowered GWAS but might be functional and thereby potentially contribute to the risk for the disorder. The list of 296 SNPs used in this study is in Supplemental Table S2.

Screening for functional SNPs in the 3'UTR using PASSPORT-seq

The basic idea behind the PASSPORT assay is to determine whether the alternate alleles at a SNP are differentially expressed as RNA. This is done by transfecting a pool of expression vectors containing both alleles at all chosen SNPs, and examining biases in allelic expression by comparing at each SNP the ratio of alleles in the RNA to their ratio in the expression vectors transfected into the cell. The PASSPORT-seq assay was conducted as previously described in detail (Rao 2019). Briefly, the assay involves (a) synthesis of a pool of oligonucleotides that includes both alleles of 296 SNPs, each flanked by 25 nt of their genomic sequence (Oligomix®, LC Sciences, Houston, TX) (Supplementary Table S2), (b)

cloning them in bulk into reporter plasmid pIS-0 (12178, Addgene, Cambridge, MA) and preparing DNA, (c) transient transfection of the pooled DNA into cells, (d) extraction and purification of DNA and RNA from the cells, and (e) quantitation of the relative amounts of RNA and DNA for each oligonucleotide by next-generation sequencing. We tested expression in two neuroblastoma cell lines, SH-SY5Y (CRL-2266, ATCC, Manassas, VA) and SK-N-BE(2) (CRL-2271, ATCC). Both cell lines were cultured in a 1:1 mixture of EMEM (302003, ATCC) and F12K medium (10025-CV, Thermo Fisher Scientific, Waltham, MA) with 10% (vol/vol) fetal bovine serum (302020, ATCC) and 1% penicillin and streptomycin. Six independent biological replicates were conducted for each cell line. At 42 h post transfection, both plasmid DNA and RNA were extracted from the cells using the miRNeasy mini kit (217004, Qiagen, Germantown, MD) per manufacturer's instructions. The cDNAs were synthesized from total RNA using QuantiTech Reverse Transcription kit (205311, Qiagen) with primers complementary to a sequence in the reporter vector 3' of the inserted sequence. The target sequences from the cDNAs and also from the plasmid DNA extracted from the same cells were amplified using two-step PCR with primers containing sample barcodes to separately tag cDNA (representing expression from an allele) and input DNA from each replicate, unique molecular indices (UMI), and partial Illumina sequencing adapters (Illumina, San Diego, CA). The resulting PCR products of cDNAs and plasmid DNAs from each cell type were pooled to minimize potential batch effects in library preparation and sequencing. These PCR products were amplified using barcoded Illumina index adaptors and the two resulting libraries were pooled in equal molarity (again, to minimize batch effects) and sequenced (75 base paired-ends) on the Illumina HiSeq 4000 1 lane each at 2 different times.

Analyses

After sequencing, FASTQ files were demultiplexed based upon the barcodes (cell type, replicate, DNA or cDNA) with "cutadapt" (Martin, 2011). Each sequencing read was mapped to an oligonucleotide sequence using "bwa-mem" (Li, 2013) and the number of unique molecular identifiers (UMI) for each sample was counted using "umi_tools" (Smith et al., 2017). We only counted reads with exact matches (edit distance (NM tag) equal to 0) to the oligo sequences of either reference or alternative alleles.

The basic idea of the analysis is to compare the relative counts (unique UMI) of the RNA expressed from each allele of a given SNP to that of the input plasmid DNA of that SNP extracted from the same pool of cells (to control for potential bias in the plasmid pool). A generalized linear mixed effect model (GLMM) was used to model the sequencing reads based on allele type (reference or alternative), DNA source (DNA or RNA), batch number, and the interaction between allele type and cDNA or DNA. A random variable was used to account for data derived from the same biological sample, and a negative binomial distribution was applied. Thus, we modeled this as:

$$\log(\mu) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \beta_B X_B + r_{XS}$$

where μ is the expected number of sequencing reads, X_1 is the allele type (reference or alternative), X_2 is the DNA source (DNA or cDNA), X_B is the batch number (first or second

sequencing), and X_s is the sample replicate number. In this model, β_0 , β_1 , β_2 , and β_{12} are the coefficients of the fixed effects, β_B is the coefficient for the batch effect, and r is the coefficient for the random effect. For each SNP, we estimate the value of the coefficients.

In this study, we are interested in β_{12} , the coefficient for the effect of the interaction between allele type and DNA source that describes the SNP's alternative allele effect. Rejecting the null hypothesis ($\beta_{12} = 0$) suggests that the two alleles cause differences in RNA expression. Positive values for β_{12} indicate that the alternative allele increases the expression of the 3'UTR segment (i.e., the alternative allele is at higher frequency in RNA than in the DNA), while negative values suggest the opposite. The false discovery rate (FDR) was calculated according to Benjamini and Hochberg (Benjamini and Hochberg, 1995).

Annotation of the functional 3'UTR SNPs

For each SNP that showed differential expression, we examined the Genotype-Tissue Expression (GTEx) database (version V8)(GTEx Consortium, 2017) to determine whether it is a reported expression quantitative trait locus (eQTL) of the gene it resided in, and whether that gene is expressed in the brain. We annotated binding sites of RNA-binding proteins (RBP) using the ENCODE RNA Binding Protein track (Baroni et al., 2008) of the UCSC Genome Browser database (Haeussler et al., 2019). Using the Polymirns database (Bhattacharya et al., 2014) we identified miRNAs whose binding, as calculated by TargetScan (Agarwal et al., 2015), were predicted to be altered.

Results

To understand one potential mechanism of action underlying loci associated with AD and related phenotypes, we selected 296 SNPs that showed some evidence of association ($p < 0.001$) and were located in 3'UTRs, and simultaneously tested whether they affected gene expression using a high-throughput assay, PASSPORT-seq (Rao et al., 2019). Both alleles of 84% and 86% of the SNPs were detected in both DNA and RNA in every replicate in SH-SY5Y and SK-N-BE(2) cells, respectively (Supplementary Table S3). We identified 60 SNPs (FDR < 0.05) that affected gene expression in the SH-SY5Y and 92 in SK-N-BE(2) cells (Table 1). Among these, 30 SNPs altered the RNA levels relative to DNA levels in the same direction in both cell lines (Figure 1A; Table 2); this degree of overlap is significant (Fisher's exact test, $p = 4.91 \times 10^{-7}$). The alternative allele frequency of RNA vs DNA and the read depth for four representative SNPs with significant differential expression in both cell lines are shown in Figure 1B. Relaxing the FDR to < 0.2, there were 102 functional SNPs in SH-SY5Y and 128 in SK-N-BE(2), among which 57 SNPs showed significant differential expression in both cell lines, 51 in the same direction (Table 1; Table 2). Allele-biased expression results can be visualized on an R-shiny-based interactive website (https://yunlongliulab.shinyapps.io/coga_passport/).

There are many genes in which multiple SNPs in the 3'UTR affected expression. The very long 3' UTR of *KIF6* (kinesin family member 6) contained 13 SNPs that were tested, and of those, 7 showed significant allele-biased expression in both neuronal cell lines (FDR < 0.2), of which 2 (rs56370893 and rs16891940, both associated with the flushing phenotype) were significant at FDR < 0.05. *GAS7* (growth arrest specific 7) had 4 significant SNPs; it is

broadly expressed in brain, primarily in terminally differentiated cells, especially mature cerebellar Purkinje neurons, and plays a role in neuronal development. *NBN* (nibrin; also known as NBS1), with 2 significant SNPs, is involved in double-strand break repair. *PKDIL2* (polycystin 1 like 2) has 4 functional SNPs. *SIRPA* (signal regulatory protein alpha; 2 SNPs) encodes a transmembrane glycoprotein that regulates receptor tyrosine kinase signaling. *SIRPA* is expressed ubiquitously in brain at high levels, primarily in neurons, macrophages and dendritic cells. *ADD2* (adducin 2, expressed only in brain and hematopoietic tissues) had 3 significant SNPs. *CRKL* (CRK like proto-oncogene, adaptor protein) had 2 significant SNPs; it encodes a protein kinase that activates RAS and JUN kinase signaling and is expressed ubiquitously in brain. Other SNPs that affected gene expression were located in genes associated with neural phenotypes, including *CADM2*, *FRMD4A*, and *PLK2*.

Among the SNPs that showed differential expression significant at FDR = 0.20, 89 are in expression quantitative trait loci (eQTL) of their own genes and 116 were within the target site for an RNA-binding protein (RBP) (Table 2). The most prominent RBPs were PABPC1 (Poly(A) Binding Protein Cytoplasmic 1), with sites that overlap 88 SNPs, ELAVL1 (ELAV Like RNA Binding Protein 1) with sites that overlap 75 SNPs, SLBP (Stem-Loop Binding Protein) with sites that overlap 28 SNPs, CELF1 (CUGBP Elav-Like Family Member 1) with sites that overlap 20 SNPs, and IGF2BP1 (Insulin Like Growth Factor 2 mRNA Binding Protein 1) with sites that overlap 11 SNPs. Most of the functional SNPs (128 of the 173) were within predicted binding sites of microRNAs (miRNAs) (Table 2). In most cases, several miRNA were predicted to bind around the SNP; in total, 532 different miRNAs were predicted to bind these sites.

Discussion

Variations in many genes contribute to the risk for complex, heterogeneous disorders such as AD/AUD and traits related to them, and most have only a small effect (Edenberg and Foroud, 2013) (Gelernter et al., 2014) (Walters et al., 2018). More than 90% of the variants identified by GWAS lie in intergenic regions and do not affect the protein sequence (Hindorff et al., 2009, Maurano et al., 2012). Thus, most presumably act by altering gene expression. There are many variants within the linkage disequilibrium block that defines each locus, and the variants associated with a trait might not themselves be the functional variants. Therefore, to identify functional variants within the associated regions, highly sensitive and reliable high-throughput assays are needed.

One potential mechanism by which variants could affect gene expression is by affecting RNA stability. Here, we identified 296 SNPs that were weakly associated with alcohol dependence and related traits (P values = 0.001) in the meta-analyses of African American and European American samples in COGA (Lai et al., 2019a, Lai et al., 2019b, Wetherill et al., 2019) and that reside in the 3'UTR of a gene. We tested these in a high-throughput, parallel assay, PASSPORT-seq (Ipe et al., 2018, Rao et al., 2019). Most of the SNPs tested (58%) affected gene expression (FDR = 0.20), and most of those lie within binding sites of RNA binding proteins and microRNAs (67% and 74%, respectively; Table 2). Just over half

lie in known eQTLs for the genes they reside within. The overlap between differential expression and these known regulatory sites reinforces our findings.

Among the many genes in which multiple SNPs in the 3'UTR affected expression, some have relevant phenotypes. *KIF6* (kinesin family member 6), expressed in several regions of the brain, including amygdala, caudate, hippocampus and putamen, encodes a molecular motor involved in intracellular transport. It plays a role in ciliogenesis in vertebrate ependymal cell (specialized glial cells that forms the epithelial lining of the ventricular walls of the brain and spinal canal), and a frameshift mutation in *KIF6* was found in a child who displayed neurodevelopmental defects and intellectual disability (Konjikusic et al., 2018). A rare variant in *KIF6* was reported to play a role in some sudden unexpected deaths in people with epilepsy (Ge et al., 2020). *ADD2* (adducin 2) encodes a cytoskeleton associated protein involved in synaptic plasticity; alcohol exposure at early neurulation alters the expression of *ADD2* (Zhou et al., 2011). Polymorphisms in adducins affect cognitive functions in individuals with schizophrenia (Bosia et al., 2016). *NBN* (nibrin) affects neuronal proliferation and differentiation (Lee et al., 2007). Mutations in *NBN* cause Nijmegen breakage syndrome, characterized by progressive microcephaly, growth retardation, and predisposition to cancer (Varon et al., 1998). *PKDIL2* (polycystin 1 like 2) upregulation leads to a complex neuromuscular disease in mice (Mackenzie et al., 2009); it forms part of an ion channel in cilia of some cell types (DeCaen et al., 2013). Mutation of *SIRPA* (signal regulatory protein alpha) protected mice from oxidative stress and brain ischemia (Wang et al., 2012).

CADM2 (cell adhesion molecule 2) encodes a synaptic cell adhesion molecule expressed at very high levels in many brain regions. In a gene-based analysis, *CADM2* was associated with alcohol consumption in the UK Biobank, as are several SNPs in *CADM2* (rs13078384, rs1376935, rs67028245) (Clarke et al., 2017). *CADM2* is also associated with other psychological diseases and traits, including sensation seeking, drug experimentation, substance use involvement, and attention-deficit/hyperactivity disorder (Sanchez-Roige et al., 2019) and lifetime cannabis use (Pasman et al., 2018, Stringer et al., 2016). *FRMD4A* (FERM Domain Containing 4A), also expressed in multiple regions of the brain, is involved in regulating epithelial cell polarity. Polymorphism in *FRMD4A* is associated with nicotine dependence (Yoon et al., 2012) and as a risk locus for Alzheimer's disease (Lambert et al., 2013). *PLK2* (polo like kinase 2) encodes a member of the polo family of serine/threonine protein kinases and has a role in normal cell division. It is expressed in multiple regions of the brain (GTEx). Two studies have reported that *PLK2* expression is significantly upregulated in human Alzheimer's disease brain tissue (Mbefo et al., 2010, Lee et al., 2019). Polymorphisms in *PLK2* are associated with Alzheimer's disease (Bufill et al., 2015), and it has been suggested as an attractive novel target for Alzheimer's disease drug therapy (Lee et al., 2019, Lee et al., 2017).

While we identify functionally relevant SNPs that are associated with the risk of AD and associated phenotypes, and demonstrate the utility of a parallel assay for functional SNPs, our study has several limitations. One limitation is that the SNPs tested in this study were not themselves genome-wide significant. Another is that we only screened the variants in 3'UTR; obviously, variants in other regulatory regions likely play a role in influencing the

risk for AD and related phenotypes, and similar types of assays could be developed for those. Gene expression in any cell line depends both upon the unique genetics of that cell line and the exact culture conditions, which affect the expression of *trans*-acting factors. Many of the SNPs showed effects in the same direction, which is not unexpected since both cell lines are neuronal and presumably share many *trans*-acting factors. Those that differ in direction (or function in one cell line or the other) are presumably affected by a difference between the cell lines in the balance of *trans*-acting factors. Our use of 2 distinct cell lines provides a broader perspective than would a single one. Finally, any *in vitro* study, by definition, does not reproduce the environment in an intact brain.

In summary, our high-throughput assay detected 122 SNPs that significantly affect gene expression in either of two neuroblastoma cell lines (FDR < 0.05), 30 of which altered the RNA levels in the same direction in both. Although many of the SNPs had the same direction of effect in both cell types, some were cell-type specific, indicating that it is important to examine the effects of SNPs in multiple cell lines. These functional SNPs could contribute to the development of alcohol dependence and related phenotypes. Our study demonstrates one way of moving from GWAS association into biological mechanism, providing more insight into risk factors for AD/AUD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The Collaborative Study on the Genetics of Alcoholism (COGA), Principal Investigators B. Porjesz, V. Hesselbrock, T. Foroud; Scientific Director, A. Agrawal; Translational Director, D. Dick, includes eleven different centers: University of Connecticut (V. Hesselbrock); Indiana University (H.J. Edenberg, T. Foroud, J. Nurnberger Jr., Y. Liu); University of Iowa (S. Kuperman, J. Kramer); SUNY Downstate (B. Porjesz, J. Meyers, C. Kamarajan, A. Pandey); Washington University in St. Louis (L. Bierut, J. Rice, K. Bucholz, A. Agrawal); University of California at San Diego (M. Schuckit); Rutgers University (J. Tischfield, A. Brooks, R. Hart); The Children's Hospital of Philadelphia, University of Pennsylvania (L. Almasy); Virginia Commonwealth University (D. Dick, J. Salvatore); Icahn School of Medicine at Mount Sinai (A. Goate, M. Kapoor, P. Slesinger); and Howard University (D. Scott). Other COGA collaborators include: L. Bauer (University of Connecticut); L. Wetherill, D. Lai, S. O'Connor, M. Plawecki, (Indiana University); L. Acion (University of Iowa); G. Chan (University of Iowa; University of Connecticut); D.B. Chorlian, J. Zhang, S. Kinreich, G. Pandey (SUNY Downstate); ; M. Chao (Icahn School of Medicine at Mount Sinai); A. Anokhin, V. McCutcheon, S. Saccone (Washington University); F. Aliev, P. Barr (Virginia Commonwealth University); H. Chin and A. Parsian are the NIAAA Staff Collaborators. This national collaborative study is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA). Technical support was provided by Center for Medical Genomics facility at Indiana University School of Medicine.

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal (GTEx Analysis Release V8) on February 7, 2020.

This national collaborative study is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA). The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS.

References

- Agarwal V, Bell GW, Nam JW, Bartel DP (2015) Predicting effective microRNA target sites in mammalian mRNAs. *Elife* 4:e05005.
- American_Psychiatric_Association (2000) Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (Text Revision), Washington, DC.
- American_Psychiatric_Association (2013) Diagnostic and Statistical Manual of Mental Disorders: DSM-5, AMERICAN PSYCHIATRIC PUBLISHING.
- Baroni TE, Chittur SV, George AD, Tenenbaum SA (2008) Advances in RIP-chip analysis : RNA-binding protein immunoprecipitation-microarray profiling. *Methods Mol Biol* 419:93–108. [PubMed: 18369977]
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300.
- Bhattacharya A, Ziebarth JD, Cui Y (2014) PolymiRTS Database 3.0: linking polymorphisms in microRNAs and their target sites with human diseases and biological pathways. *Nucleic Acids Res* 42:D86–91. [PubMed: 24163105]
- Bosia M, Pigoni A, Zagato L, Merlino L, Casamassima N, Lorenzi C, Pirovano A, Smeraldi E, Manunta P, Cavallaro R (2016) ADDing a piece to the puzzle of cognition in schizophrenia. *Eur J Med Genet* 59:26–31. [PubMed: 26723519]
- Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JIJ, Reich T, Schmidt I, Schuckit MA (1994) A new semi-structured psychiatric interview for use in genetic linkage studies: A report of the reliability of the SSAGA. *J.Stud.Alcohol* 55:149–158. [PubMed: 8189735]
- Buñill E, Roura-Poch P, Sala-Matavera I, Anton S, Lleo A, Sanchez-Saudinos B, Tomas-Abadal L, Puig T, Abos J, Bernades S, Clarimon J, Blesa R (2015) Reelin signaling pathway genotypes and Alzheimer disease in a Spanish population. *Alzheimer Dis Assoc Disord* 29:169–172. [PubMed: 24384746]
- Clarke TK, Adams MJ, Davies G, Howard DM, Hall LS, Padmanabhan S, Murray AD, Smith BH, Campbell A, Hayward C, Porteous DJ, Deary IJ, McIntosh AM (2017) Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK Biobank (N=112 117). *Mol Psychiatry* 22:1376–1384. [PubMed: 28937693]
- Collaborators GAaDU (2018) The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Psychiatry* 5:987–1012. [PubMed: 30392731]
- DeCaen PG, Delling M, Vien TN, Clapham DE (2013) Direct recording and molecular identification of the calcium channel of primary cilia. *Nature* 504:315–318. [PubMed: 24336289]
- Edenberg HJ (2002) The collaborative study on the genetics of alcoholism: an update. *Alcohol Res Health* 26:214–218. [PubMed: 12875050]
- Edenberg HJ, Foroud T (2013) Genetics and alcoholism. *Nature reviews. Gastroenterology & hepatology* 10:487–494. [PubMed: 23712313]
- Edenberg HJ, McClintick JN (2018) Alcohol Dehydrogenases, Aldehyde Dehydrogenases, and Alcohol Use Disorders: A Critical Review. *Alcohol Clin Exp Res* 42:2281–2297. [PubMed: 30320893]
- Ge Y, Ding D, Zhu G, Kwan P, Wang W, Hong Z, Sander JW (2020) Genetic variants in incident SUDEP cases from a community-based prospective cohort with epilepsy. *J Neurol Neurosurg Psychiatry* 91:126–131. [PubMed: 31776209]
- Gelernter J, Kranzler HR, Sherva R, Almasy L, Koesterer R, Smith AH, Anton R, Preuss UW, Ridinger M, Rujescu D, Wodarz N, Zill P, Zhao H, Farrer LA (2014) Genome-wide association study of alcohol dependence: significant findings in African- and European-Americans including novel risk loci. *Mol Psychiatry* 19:41–49. [PubMed: 24166409]
- Gelernter J, Sun N, Polimanti R, Pietrzak RH, Levey DF, Lu Q, Hu Y, Li B, Radhakrishnan K, Aslan M, Cheung KH, Li Y, Rajeevan N, Sayward F, Harrington K, Chen Q, Cho K, Honerlaw J, Pyarajan S, Lencz T, Quaden R, Shi Y, Hunter-Zinck H, Gaziano JM, Kranzler HR, Concato J, Zhao H, Stein MB, Department of Veterans Affairs Cooperative Studies P, Million Veteran P

- (2019) Genome-wide Association Study of Maximum Habitual Alcohol Intake in >140,000 U.S. European and African American Veterans Yields Novel Risk Loci. *Biol Psychiatry* 86:365–376. [PubMed: 31151762]
- GTEEx_Consortium (2017) Genetic effects on gene expression across human tissues. *Nature* 550:204–213. [PubMed: 29022597]
- Haeussler M, Zweig AS, Tyner C, Speir ML, Rosenbloom KR, Raney BJ, Lee CM, Lee BT, Hinrichs AS, Gonzalez JN, Gibson D, Diekhans M, Clawson H, Casper J, Barber GP, Haussler D, Kuhn RM, Kent WJ (2019) The UCSC Genome Browser database: 2019 update. *Nucleic Acids Res* 47:D853–D858. [PubMed: 30407534]
- Hart AB, Kranzler HR (2015) Alcohol Dependence Genetics: Lessons Learned From Genome-Wide Association Studies (GWAS) and Post-GWAS Analyses. *Alcohol Clin Exp Res* 39:1312–1327. [PubMed: 26110981]
- Hesselbrock M, Easton C, Bucholz KK, Schuckit M, Hesselbrock V (1999) A validity study of the SSAGA—a comparison with the SCAN. *Addiction* 94:1361–1370. [PubMed: 10615721]
- Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 106:9362–9367. [PubMed: 19474294]
- Ipe J, Collins KS, Hao Y, Gao H, Bhatia P, Gaedigk A, Liu Y, Skaar TC (2018) PASSPORT-seq: A Novel High-Throughput Bioassay to Functionally Test Polymorphisms in Micro-RNA Target Sites. *Frontiers in genetics* 9:219. [PubMed: 29963077]
- Konjikusic MJ, Yeetong P, Boswell CW, Lee C, Roberson EC, Ittiwut R, Suphapeetiporn K, Ciruna B, Gurnett CA, Wallingford JB, Shotelersuk V, Gray RS (2018) Mutations in Kinesin family member 6 reveal specific role in ependymal cell ciliogenesis and human neurological development. *PLoS Genet* 14:e1007817. [PubMed: 30475797]
- Kranzler HR, Zhou H, Kember RL, Vickers Smith R, Justice AC, Damrauer S, Tsao PS, Klarin D, Baras A, Reid J, Overton J, Rader DJ, Cheng Z, Tate JP, Becker WC, Concato J, Xu K, Polimanti R, Zhao H, Gelernter J (2019) Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. *Nature communications* 10:1499.
- Lai D, Wetherill L, Bertelsen S, Carey CE, Kamarajan C, Kapoor M, Meyers JL, Anokhin AP, Bennett DA, Bucholz KK, Chang KK, De Jager PL, Dick DM, Hesselbrock V, Kramer J, Kuperman S, Nurnberger JI Jr., Raj T, Schuckit M, Scott DM, Taylor RE, Tischfield J, Hariri AR, Edenberg HJ, Agrawal A, Bogdan R, Porjesz B, Goate AM, Foroud T (2019a) Genome-wide association studies of alcohol dependence, DSM-IV criterion count and individual criteria. *Genes Brain Behav* 18:e12579. [PubMed: 31090166]
- Lai D, Wetherill L, Kapoor M, Johnson EC, Schwandt M, Ramchandani VA, Goldman D, Joslyn G, Rao X, Liu Y, Farris S, Mayfield RD, Dick D, Hesselbrock V, Kramer J, McCutcheon VV, Nurnberger J, Tischfield J, Goate A, Edenberg HJ, Porjesz B, Agrawal A, Foroud T, Schuckit M (2019b) Genome-wide association studies of the self-rating of effects of ethanol (SRE). *Addict Biol*:e12800. [PubMed: 31270906]
- Lambert JC, Grenier-Boley B, Harold D, Zelenika D, Chouraki V, Kamatani Y, Sleegers K, Ikram MA, Hiltunen M, Reitz C, Mateo I, Feulner T, Bullido M, Galimberti D, Concar L, Alvarez V, Sims R, Gerrish A, Chapman J, Deniz-Naranjo C, Solfrizzi V, Sorbi S, Arosio B, Spalletta G, Siciliano G, Epelbaum J, Hannequin D, Dartigues JF, Tzourio C, Berr C, Schrijvers EM, Rogers R, Tosto G, Pasquier F, Bettens K, Van Cauwenberghe C, Fratiglioni L, Graff C, Delepine M, Ferri R, Reynolds CA, Lannfelt L, Ingelsson M, Prince JA, Chillotti C, Pilotto A, Seripa D, Boland A, Mancuso M, Bossu P, Annoni G, Nacmias B, Bosco P, Panza F, Sanchez-Garcia F, Del Zompo M, Coto E, Owen M, O'Donovan M, Valdivieso F, Caffarra P, Scarpini E, Combarros O, Buee L, Campion D, Soininen H, Breteler M, Riemenschneider M, Van Broeckhoven C, Alperovitch A, Lathrop M, Tregouet DA, Williams J, Amouyel P (2013) Genome-wide haplotype association study identifies the FRMD4A gene as a risk locus for Alzheimer's disease. *Mol Psychiatry* 18:461–470. [PubMed: 22430674]
- Lee JS, Lee Y, Andre EA, Lee KJ, Nguyen T, Feng Y, Jia N, Harris BT, Burns MP, Pak DTS (2019) Inhibition of Polo-like kinase 2 ameliorates pathogenesis in Alzheimer's disease model mice. *PLoS One* 14:e0219691. [PubMed: 31306446]

- Lee WT, Chang WH, Huang CH, Wu KJ (2007) NBS1, the Nijmegen breakage syndrome gene product, regulates neuronal proliferation and differentiation. *J Neurochem* 102:141–152. [PubMed: 17442057]
- Lee Y, Lee JS, Lee KJ, Turner RS, Hoe HS, Pak DTS (2017) Polo-like kinase 2 phosphorylation of amyloid precursor protein regulates activity-dependent amyloidogenic processing. *Neuropharmacology* 117:387–400. [PubMed: 28257888]
- Li H Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. online]. 5–26-2013;arXiv:1303.3997.
- Mackenzie FE, Romero R, Williams D, Gillingwater T, Hilton H, Dick J, Riddoch-Contreras J, Wong F, Ireson L, Powles-Glover N, Riley G, Underhill P, Hough T, Arkell R, Greensmith L, Ribchester RR, Blanco G (2009) Upregulation of PKD1L2 provokes a complex neuromuscular disease in the mouse. *Human Molecular Genetics* 18:3553–3566. [PubMed: 19578180]
- Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 17:3.
- Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, Shafer A, Neri F, Lee K, Kutayavin T, Stehling-Sun S, Johnson AK, Canfield TK, Giste E, Diegel M, Bates D, Hansen RS, Neph S, Sabo PJ, Heimfeld S, Raubitschek A, Ziegler S, Cotsapas C, Sotoodehnia N, Glass I, Sunyaev SR, Kaul R, Stamatoyannopoulos JA (2012) Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337:1190–1195. [PubMed: 22955828]
- Mbefo MK, Paleologou KE, Boucharaba A, Oueslati A, Schell H, Fournier M, Olschewski D, Yin G, Zweckstetter M, Masliah E, Kahle PJ, Hirling H, Lashuel HA (2010) Phosphorylation of synucleins by members of the Polo-like kinase family. *J Biol Chem* 285:2807–2822. [PubMed: 19889641]
- Pasman JA, Verweij KJH, Gerring Z, Stringer S, Sanchez-Roige S, Treur JL, Abdellaoui A, Nivard MG, Baselmans BML, Ong JS, Ip HF, van der Zee MD, Bartels M, Day FR, Fontanillas P, Elson SL, andMe Research T, de Wit H, Davis LK, MacKillop J, Substance Use Disorders Working Group of the Psychiatric Genomics C, International Cannabis C, Derringer JL, Branje SJT, Hartman CA, Heath AC, van Lier PAC, Madden PAF, Magi R, Meeus W, Montgomery GW, Oldehinkel AJ, Pausova Z, Ramos-Quiroga JA, Paus T, Ribases M, Kaprio J, Boks MPM, Bell JT, Spector TD, Gelernter J, Boomsma DI, Martin NG, MacGregor S, Perry JRB, Palmer AA, Posthuma D, Munafo MR, Gillespie NA, Derks EM, Vink JM (2018) GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia. *Nat Neurosci* 21:1161–1170. [PubMed: 30150663]
- Rao X, Thapa KS, Chen AB, Lin H, Gao H, Reiter JL, Hargreaves KA, Ipe J, Lai D, Xuei X, Wang Y, Gu H, Kapoor M, Farris SP, Tischfield J, Foroud T, Goate AM, Skaar TC, Mayfield RD, Edenberg HJ, Liu Y (2019) Allele-specific expression and high-throughput reporter assay reveal functional genetic variants associated with alcohol use disorders. *Mol Psychiatry*.
- Rehm J, Shield KD (2019) Global Burden of Disease and the Impact of Mental and Addictive Disorders. *Current psychiatry reports* 21:10. [PubMed: 30729322]
- Reich T, Edenberg HJ, Goate A, Williams JT, Rice JP, Van Eerdewegh P, Foroud T, Hesselbrock V, Schuckit MA, Bucholz K, Porjesz B, Li TK, Conneally PM, Nurnberger JI Jr., Tischfield JA, Crowe RR, Cloninger CR, Wu W, Shears S, Carr K, Crose C, Willig C, Begleiter H (1998) Genome-wide search for genes affecting the risk for alcohol dependence. *Am J Med Genet* 81:207–215. [PubMed: 9603606]
- Rietschel M, Treutlein J (2013) The genetics of alcohol dependence. *Ann N Y Acad Sci* 1282:39–70. [PubMed: 23170934]
- Sanchez-Roige S, Fontanillas P, Elson SL, Gray JC, de Wit H, MacKillop J, Palmer AA (2019) Genome-Wide Association Studies of Impulsive Personality Traits (BIS-11 and UPPS-P) and Drug Experimentation in up to 22,861 Adult Research Participants Identify Loci in the CACNA1I and CADM2 genes. *J Neurosci* 39:2562–2572. [PubMed: 30718321]
- Smith T, Heger A, Sudbery I (2017) UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res* 27:491–499. [PubMed: 28100584]
- Stringer S, Minica CC, Verweij KJ, Mbarek H, Bernard M, Derringer J, van Eijk KR, Isen JD, Loukola A, Maciejewski DF, Mihailov E, van der Most PJ, Sanchez-Mora C, Roos L, Sherva R, Walters R,

Ware JJ, Abdellaoui A, Bigdeli TB, Branje SJ, Brown SA, Bruinenberg M, Casas M, Esko T, Garcia-Martinez I, Gordon SD, Harris JM, Hartman CA, Henders AK, Heath AC, Hickie IB, Hickman M, Hopfer CJ, Hottenga JJ, Huizink AC, Irons DE, Kahn RS, Korhonen T, Kranzler HR, Krauter K, van Lier PA, Lubke GH, Madden PA, Magi R, McGue MK, Medland SE, Meeus WH, Miller MB, Montgomery GW, Nivard MG, Nolte IM, Oldehinkel AJ, Pausova Z, Qaiser B, Quaye L, Ramos-Quiroga JA, Richarte V, Rose RJ, Shin J, Stallings MC, Stiby AI, Wall TL, Wright MJ, Koot HM, Paus T, Hewitt JK, Ribases M, Kaprio J, Boks MP, Snieder H, Spector T, Munafo MR, Metspalu A, Gelernter J, Boomsma DI, Iacono WG, Martin NG, Gillespie NA, Derks EM, Vink JM (2016) Genome-wide association study of lifetime cannabis use based on a large meta-analytic sample of 32 330 subjects from the International Cannabis Consortium. *Translational psychiatry* 6:e769. [PubMed: 27023175]

Varon R, Vissinga C, Platzer M, Cerosaletti KM, Chrzanowska KH, Saar K, Beckmann G, Seemanova E, Cooper PR, Nowak NJ, Stumm M, Weemaes CM, Gatti RA, Wilson RK, Digweed M, Rosenthal A, Sperling K, Concannon P, Reis A (1998) Nibrin, a novel DNA double-strand break repair protein, is mutated in Nijmegen breakage syndrome. *Cell* 93:467–476. [PubMed: 9590180]

Walters RK, Polimanti R, Johnson EC, McClintick JN, Adams MJ, Adkins AE, Aliev F, Bacanu SA, Batzler A, Bertelsen S, Biernacka JM, Bigdeli TB, Chen LS, Clarke TK, Chou YL, Degenhardt F, Docherty AR, Edwards AC, Fontanillas P, Foo JC, Fox L, Frank J, Giegling I, Gordon S, Hack LM, Hartmann AM, Hartz SM, Heilmann-Heimbach S, Herms S, Hodgkinson C, Hoffmann P, Jan Hottenga J, Kennedy MA, Alanne-Kinnunen M, Konte B, Lahti J, Lahti-Pulkkinen M, Lai D, Ligthart L, Loukola A, Maher BS, Mbarek H, McIntosh AM, McQueen MB, Meyers JL, Milanese Y, Palviainen T, Pearson JF, Peterson RE, Ripatti S, Ryu E, Saccone NL, Salvatore JE, Sanchez-Roige S, Schwandt M, Sherva R, Streit F, Strohmaier J, Thomas N, Wang JC, Webb BT, Wedow R, Wetherill L, Wills AG, andMe Research T, Boardman JD, Chen D, Choi DS, Copeland WE, Culverhouse RC, Dahmen N, Degenhardt L, Domingue BW, Elson SL, Frye MA, Gabel W, Hayward C, Ising M, Keyes M, Kiefer F, Kramer J, Kuperman S, Lucae S, Lynskey MT, Maier W, Mann K, Mannisto S, Muller-Myhsok B, Murray AD, Nurnberger JI, Palotie A, Preuss U, Raikonen K, Reynolds MD, Ridinger M, Scherbaum N, Schuckit MA, Soyka M, Treutlein J, Witt S, Wodarz N, Zill P, Adkins DE, Boden JM, Boomsma DI, Bierut LJ, Brown SA, Bucholz KK, Cichon S, Costello EJ, de Wit H, Diazgranados N, Dick DM, Eriksson JG, Farrer LA, Foroud TM, Gillespie NA, Goate AM, Goldman D, Gruzca RA, Hancock DB, Harris KM, Heath AC, Hesselbrock V, Hewitt JK, Hopfer CJ, Horwood J, Iacono W, Johnson EO, Kaprio JA, Karpyak VM, Kendler KS, Kranzler HR, Krauter K, Lichtenstein P, Lind PA, McGue M, MacKillop J, Madden PAF, Maes HH, Magnusson P, Martin NG, Medland SE, Montgomery GW, Nelson EC, Nothen MM, Palmer AA, Pedersen NL, Penninx B, Porjesz B, Rice JP, Rietschel M, Riley BP, Rose R, Rujescu D, Shen PH, Silberg J, Stallings MC, Tarter RE, Vanyukov MM, Vrieze S, Wall TL, Whitfield JB, Zhao H, Neale BM, Gelernter J, Edenberg HJ, Agrawal A (2018) Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci* 21:1656–1669. [PubMed: 30482948]

Wang L, Lu Y, Deng S, Zhang Y, Yang L, Guan Y, Matozaki T, Ohnishi H, Jiang H, Li H (2012) SHPS-1 deficiency induces robust neuroprotection against experimental stroke by attenuating oxidative stress. *Journal of Neurochemistry* 122:834–843. [PubMed: 22671569]

Wetherill L, Lai D, Johnson EC, Anokhin A, Bauer L, Bucholz KK, Dick DM, Hariri AR, Hesselbrock V, Kamarajan C, Kramer J, Kuperman S, Meyers JL, Nurnberger JI Jr., Schuckit M, Scott DM, Taylor RE, Tischfield J, Porjesz B, Goate AM, Edenberg HJ, Foroud T, Bogdan R, Agrawal A (2019) Genome-wide association study identifies loci associated with liability to alcohol and drug dependence that is associated with variability in reward-related ventral striatum activity in African- and European-Americans. *Genes Brain Behav* 18:e12580. [PubMed: 31099175]

Yoon D, Kim YJ, Cui WY, Van der Vaart A, Cho YS, Lee JY, Ma JZ, Payne TJ, Li MD, Park T (2012) Large-scale genome-wide association study of Asian population reveals genetic factors in FRMD4A and other loci influencing smoking initiation and nicotine dependence. *Hum Genet* 131:1009–1021. [PubMed: 22006218]

Zhou FC, Zhao Q, Liu Y, Goodlett CR, Liang T, McClintick JN, Edenberg HJ, Li L (2011) Alteration of gene expression by alcohol exposure at early neurulation. *BMC Genomics* 12:124. [PubMed: 21338521]

Zhou H, Sealock JM, Sanchez-Roige S, Clarke T-K, Levey D, Cheng Z, Li B, Polimanti R, Kember RL, Smith RV, Thygesen JH, Morgan MY, Atkinson SR, Thursz MR, Nyegaard M, Mattheisen M, Børglum AD, Johnson EC, Justice AC, Palmer AA, McQuillin A, Davis LK, Edenberg HJ, Agrawal A, Kranzler HR, Gelernter J (2020) Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. *Nat Neurosci.* 2020;23(7):809–18. [PubMed: 32451486]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

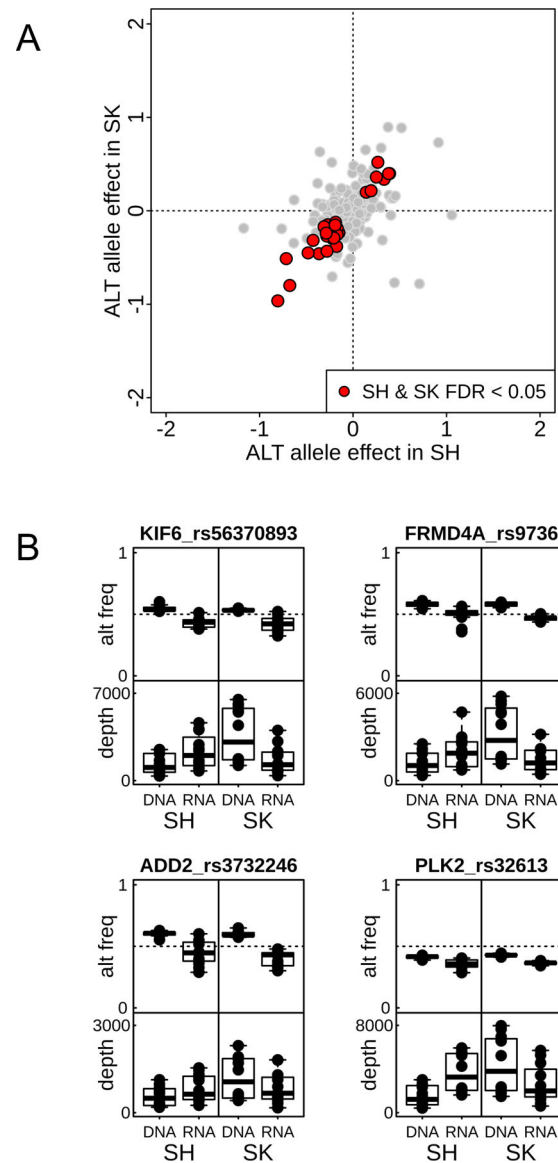


Figure 1.

PASSPORT-seq results in SH-SY5Y and SK-N-BE(2) cell lines. (A) Plot of the alternative allele effect (derived from the mixed effects generalized linear model) of the SNPs in SH-SY5Y [SH] and SK-N-BE(2) [SK] cell lines. SNPs with FDR < 0.05 in both cell lines are denoted in red. (B) Alternative allele frequency and total read depth (sum of reference and alternative alleles) for a few significant SNPs of interest. SH: SH-SY5Y; SK: SK-N-BE(2); ALT: alternate; FDR: false discovery rate; alt freq: alternative allele frequency.

Table 1.

Numbers of functional SNPs in SH-SY5Y and SK-N-BE(2) neuroblastoma cell lines (FDR < 0.05 or < 0.2) identified by PASSPORT-Seq. SNP: single nucleotide polymorphism; FDR: false discovery rate.

Cell type	#SNPs with FDR < 0.05	#SNPs with FDR < 0.20
SH-SY5Y	60	102
SK-N-BE(2)	92	128
Either SH-SY5Y or SK-N-BE(2)	122	173
Both SH-SY5Y and SK-N-BE(2)	30	57

Table 2.

SNPs significant ($FDR < 0.2$) in either SH-SY5Y or SK-N-BE(2) cells. SNP: single nucleotide polymorphism; ALT allele: alternate allele; ALT freq: alternative allele frequency; FDR: false discovery rate; eQTL (expression quantitative trait loci): whether the SNP is in eQTL for its gene in GTEx database; RBP binding: RNA-binding proteins whose target binding sites overlapped with the SNP; miRNA target: miRNAs whose binding potential may be altered by the variant allele; imputed (quality score 0.3) vs. genotyped.

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
SNP ^d	Gene	Description	Phenotype ^b	A	35%	-0.547	9.86E-02	-0.292	3.56E-01	yes		miR-6741-3p miR-6822-3p miR-4267 miR-5586-5p
rs3797615	ADAMTS2	ADAM metalloproteinase with thrombospondin type 1 motif 2	time_used									
rs35480486	ADD2	adducin 2	dsm5sev	T	20%	-0.099	3.86E-04	0.026	3.22E-01	yes	PABPC1	
rs3732246 ^d	ADD2	adducin 2	dsm5sev	T	20%	-0.411	1.35E-06	-0.573	4.13E-18	yes		miR-1243 miR-3654
rs17698193	ADD2	adducin 2	more_intend	G	21%	0.107	1.87E-02	0.039	6.42E-01	no		* miR-1250-5p miR-4462 miR-218-5p miR-5008-3p miR-636 miR-6737-3p miR-7157-3p
rs58092703 ^d	ARMC9	armadillo repeat containing 9	time_used	G	5%	0.057	8.34E-01	-0.216	4.03E-02	no	PABPC1, SLBP	
rs35498576 ^d	BAG5	BCL2 associated athanogene 5	dom_life	C	28%	-0.124	1.87E-02	-0.040	4.07E-01	yes	ELAVL1, PABPC1, CELF1	miR-138-2-3p
rs56738985	BMPRII	bone morphogenetic protein receptor type II	physical	A	3%	-0.201	7.01E-02	-0.027	7.48E-01	yes		miR-34c-3p miR-7112-3p
rs16854007	C1orf131	chromosome 1 open reading frame 131	more_intend	T	5%	-0.100	2.35E-01	-0.251	9.54E-04	no	PABPC1	miR-3152-3p miR-6867-5p

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
SNP ^a	Gene	Description	Phenotype ^b	C	35%	-5.354	1.88E-62	-5.574	3.53E-39	no	ELAVL1	miR-1302 miR-3122 miR-3913-5p miR-4298 * miR-635 miR-6774-5p miR-1343-3p miR-4793-5p miR-6742-3p miR-6783-3p miR-6852-5p miR-939-3p
rs77227433 ^d	CA12	carbonic anhydrase 12	dom_life	A	3%	-0.047		-0.039	8.73E-02	no	PABPC1	miR-214-5p miR-668-5p miR-6720-3p miR-500a-3p miR-6804-3p
rs77882042	CA12	carbonic anhydrase 12	dom_life	C	3%	-0.171	4.00E-03	0.071	5.81E-01	no	PABPC1	miR-223-5p miR-4302 miR-582-5p
rs2324980 ^d	CADM2	cell adhesion molecule 2	flush	T	10%	-0.105	1.04E-08	-0.081	1.19E-03	no		miR-3910 miR-510-3p
rs16925167	CAPRN1	cell cycle associated protein 1	tolerance	G	8%	-0.028	4.07E-01	-0.150	7.00E-06	yes		miR-676-3p miR-7-1-3p miR-7-2-3p miR-495-3p miR-5688
rs1918	CBL	Cbl proto-oncogene	more_intend	C	10%	-0.003	6.15E-01	0.076	5.38E-04	yes	ELAVL1, CELF1	miR-1253 miR-3123 miR-3925-5p miR-6832-5p miR-3125 miR-3150b-3p miR-3916 miR-4784 miR-6859-5p miR-877-5p
rs6932603	CCDC170	coiled-coil domain containing 170	more_intend	C	20%	0.066	4.79E-01	0.081	3.75E-02	yes	ELAVL1, SLBP	miR-6808-3p miR-7973
rs6932260	CCDC170	coiled-coil domain containing 170	more_intend	C	20%	0.484	4.56E-03	0.177	2.45E-01	yes	ELAVL1, SLBP	miR-16-2-3p miR-195-3p miR-656-3p

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
			Phenotype ^b									
			more_intend	T	21%	-0.052	7.79E-02	-0.021	6.37E-01	yes	ELAVL1, SLBP	miR-153-3p miR-20a-3p miR-544a miR-4520a-5p miR-4520b-5p
			flush	T	33%	0.238	1.55E-01	0.229	2.26E-02	yes	PABPC1, SLBP	miR-33a-3p miR-409-3p miR-656-3p
			dsm4sxct	A	2%	0.087	5.44E-03	0.152	1.47E-09	no	CELF1, ELAVL1, PABPC1	miR-132-3p miR-212-3p
			dsm4sxct	A	2%	-0.410	7.18E-02	0.334	6.99E-02	no	CELF1, ELAVL1, PABPC1	* miR-3936
			desire_cut	T	17%	0.028	6.61E-01	0.075	9.16E-03	no	ELAVL1, PABPC1	miR-1299 miR-6128 miR-875-3p miR-3688-3p miR-8076
			maxdrinks	G	34%	0.044	6.32E-01	0.136	3.00E-02	yes		miR-3173-3p * miR-6891-5p miR-6894-5p
			dsm5score	A	3%	-0.055	1.30E-02	-0.014		no	PABPC1, ELAVL1	miR-3688-3p miR-4528
			time_used	A	2%	0.204	3.09E-01	0.310	1.78E-02	no	ELAVL1	
			maxdrinks	C	24%	-0.246	5.79E-03	-0.088	2.85E-01	yes		
			crave	A	20%	-0.104	1.01E-01	0.090	1.13E-01	no	ELAVL1, PABPC1, SLBP	miR-4537 miR-3160-3p miR-3199 miR-365a-5p miR-365b-5p miR-4487 miR-556-5p miR-558 miR-8052
			dsm4dep	C	2%	-0.279		-0.368	1.17E-04	yes	ELAVL1	miR-5003-5p
			dsm4dep	T	5%	-0.076	1.31E-01	-0.004	9.28E-01	yes	ELAVL1	

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
				C	2%	-0.205	8.84E-02	-0.757		yes	ELAVL1	miR-142-3p miR-20b-3p miR-4687-3p miR-6073 miR-7974
				G	2%	-0.099	5.87E-08	-0.036	2.14E-01	yes	ELAVL1, PABPC1	miR-219a-2-3p miR-219b-3p miR-302f miR-1236-5p miR-3164 miR-6820-3p
				A	8%	-0.090	1.35E-01	0.029	8.48E-01	no		miR-1238-5p miR-4481 miR-4658 miR-4745-5p miR-4758-5p miR-6790-5p
rs4568				A	44%	0.125	1.11E-01	0.161	3.75E-03	yes	CELF1, PABPC1	miR-4436a * miR-497-3p miR-5000-3p miR-6715a-3p
rs4786494				A	28%	0.218	8.02E-02	0.400	3.11E-07	yes	PABPC1, CELF1, ELAVL1	miR-1193 miR-3690 miR-4254 miR-432-3p miR-4793-5p
rs2294532				T	30%	-0.104	3.57E-02	-0.169	4.00E-04	yes	IGF2BP1, PABPC1	miR-3909 miR-4421 miR-5699-3p miR-6748-3p miR-6852-3p miR-6881-3p miR-877-3p miR-1266-3p miR-2117 miR-26b-3p miR-4273 miR-4753-3p miR-6809-3p miR-7156-5p
rs1055131				T	3%	0.121	9.58E-06	0.148	1.47E-06	yes	PABPC1, ELAVL1	miR-548ao-3p miR-548u miR-7161-5p

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log ² ratio	FDR ^c	log ² ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
				G	27%	0.292	1.73E-09	0.030	6.75E-01	yes	ELAVL1	miR-424-3p miR-548ad
				G	26%	-0.052	1.17E-01	-0.113	1.39E-03	yes		miR-6793-3p
				C	8%	-0.169	4.38E-03	0.029	5.80E-01	yes	ELAVL1, SLBP	miR-4683 miR-6888-5p
				G	35%	0.134		0.234	1.39E-15	yes	IGF2BP1, PABPC1	
				C	9%	0.139	2.46E-01	0.109	1.20E-01	yes		
				G	15%	-0.035	2.46E-01	-0.064	5.91E-02	no	IGF2BP1, ELAVL1	miR-1270 miR-185-5p miR-4306 miR-432-5p miR-4531 miR-4644 miR-5192 miR-620
				G	6%	0.077	5.99E-01	0.170	3.25E-04	yes		
				A	20%	-0.302	1.46E-01	-0.091	4.37E-01	no		miR-4467 miR-6087 miR-566 miR-6134 miR-6789-3p miR-6815-5p miR-6865-5p
				G	39%	-0.087	2.65E-01	-0.264	1.86E-03	yes	ELAVL1, PABPC1, SLBP	*miR-3662 miR-580-5p miR-541-5p
				G	2%	-0.124		0.117	1.15E-01	no	IGF2BP1, PABPC1, SLBP	miR-548m miR-7848-3p miR-18a-5p miR-18b-5p miR-4290 miR-4735-3p
				A	7%	-0.228	2.29E-04	-0.312	3.93E-21	no	PABPC1, ELAVL1	miR-185-5p miR-4306 miR-4644 miR-5739 miR-139-5p miR-4797-5p

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
SNP ^d	Gene	Description	Phenotype ^b									
rs14384 ^d	<i>GADD45B</i>	growth arrest and DNA damage inducible beta	dom_life	C	12%	-0.127	4.55E-01	-0.585	1.95E-04	no		miR-136-5p miR-181a-3p miR-515-5p miR-519e-5p miR-3605-3p
rs3213494	<i>GALNT2</i>	polypeptide N-acetylgalactosaminyltransferase 2	withdrawal	T	10%	0.059	8.34E-01	-0.283	2.69E-03	no	CELF1	miR-181d-3p miR-4758-3p miR-5699-5p miR-6793-3p
rs3811484 ^d	<i>GALNT2</i>	polypeptide N-acetylgalactosaminyltransferase 2	withdrawal	G	11%	-0.271	3.91E-03	-0.271		no	CELF1	miR-548m miR-6881-5p miR-5090 miR-6775-5p miR-744-5p
rs638664	<i>GAREM1</i>	GRB2 associated regulator of MAPK1 subtype 1	withdrawal	C	19%	-0.016	4.35E-01	-0.083	3.90E-02	no		miR-4677-3p miR-4679 miR-7641 miR-192-5p miR-215-5p miR-3661 *miR-631
rs73974332	<i>GAS7</i>	growth arrest specific 7	desire_cut	T	5%	0.036	9.81E-01	0.453	2.30E-03	no	PABPC1	miR-451b miR-4696
rs73974331	<i>GAS7</i>	growth arrest specific 7	desire_cut	G	5%	0.045	5.82E-01	0.078	1.14E-01	no	PABPC1	miR-2110 miR-4271 miR-4471 miR-4725-3p miR-6780b-5p miR-8059 miR-1343-5p miR-3150a-3p miR-3160-3p miR-3175 miR-4300 miR-5591-5p miR-6726-5p miR-6763-5p miR-6825-5p miR-920 miR-939-5p
rs73974330	<i>GAS7</i>	growth arrest specific 7	desire_cut	A	5%	-0.143	3.67E-03	0.017	5.09E-01	no	PABPC1	
rs73974333	<i>GAS7</i>	growth arrest specific 7	desire_cut	T	5%	-0.077	5.45E-02	0.023	7.00E-01	no	PABPC1	miR-7108-3p

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
SNP ^d	Gene	Description	Phenotype ^b									
rs4963161	<i>GATD1</i>	glutamine amidotransferase like class 1 domain containing 1	dom_life	T	13%	-0.047	1.39E-01	-0.122	1.41E-05	no	IGF2BP1, ELAVL1, CELF1, PABPC1	miR-1204 miR-132-5p miR-1291 miR-146b-3p miR-663b miR-6775-3p miR-7108-5p
rs80071169 ^d	<i>GCN72</i>	glucosaminyl (N-acetyl) transferase 2 (I blood group)	dsm5dep	C	3%	-0.051	8.10E-02	-0.134	3.56E-05	no	ELAVL1, PABPC1, SLBP	
rs76447432 ^d	<i>GCN72</i>	glucosaminyl (N-acetyl) transferase 2 (I blood group)	dsm5sev	G	3%	-0.129	6.70E-04	-0.127		no	ELAVL1, PABPC1, SLBP	miR-181a-5p miR-181b-5p miR-181c-5p miR-181d-5p miR-3128 miR-4262 miR-4720-5p miR-4799-3p miR-5588-5p miR-6868-5p miR-3688-5p miR-4289 miR-4474-3p miR-7108-5p miR-891a-3p
rs1886176	<i>GJA3</i>	gap junction protein alpha 3	tolerance	A	27%	0.161	3.46E-10	0.168	6.77E-07	yes		miR-589-3p * miR-186-5p miR-3133 miR-4744
rs2073768	<i>GNB1L</i>	G protein subunit beta 1 like	desire_cut	A	21%	-0.418	1.46E-01	-0.143		yes	PABPC1, ELAVL1	
rs2891064 ^d	<i>GNG13</i>	G protein subunit gamma 13	tolerance	G	37%	0.042	7.56E-01	0.111	1.28E-02	no	ELAVL1	
rs199855774 ^d	<i>GOLGA8J</i>	golgin A8 family member J	time_used	C	8%	-0.139	1.39E-01	-0.469	3.56E-05	yes		
rs2911708 ^d	<i>GPATCH2</i>	G-patch domain containing 2	time_used	C	39%	-0.590	4.91E-05	-0.558	4.47E-06	no	ELAVL1, PABPC1, SLBP	miR-4456 miR-4738-3p miR-541-3p miR-582-3p miR-654-5p miR-6769a-5p miR-6769b-5p miR-92a-2-5p miR-6786-5p

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log ² ratio	FDR ^c	log ² ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
SNP ^d	Gene	Description	Phenotype ^b									
rs265123 ^d	GPATCH2	G-patch domain containing 2	time_used	A	39%	0.047	2.65E-01	0.052	1.71E-01	no	ELAVL1, PABPC1, SLBP	
rs166449	GPATCH2	G-patch domain containing 2	time_used	C	39%	0.235	2.88E-02	-0.067		no	ELAVL1, PABPC1, SLBP	
rs4669752 ^d	GREB1	growth regulating estrogen receptor binding 1	sre_total	C	5%	-0.440	1.84E-02	-0.040	7.13E-01	no	ELAVL1, PABPC1	miR-221-3p miR-222-3p
rs9469681	GRAM4	glutamate metabotropic receptor 4	more_intend	T	14%	-0.100	2.30E-03	-0.005	9.54E-01	no		
rs55762233 ^d	HAPLN4	hyaluronan and proteoglycan link protein 4	dsm4dep	G	17%	0.107	1.48E-01	0.074	3.21E-01	yes		miR-3912-5p
rs117169105 ^d	HAS3	hyaluronan synthase 3	desire_cut	T	1%	0.019	9.04E-01	0.068	1.69E-01	no	ELAVL1, PABPC1	miR-2114-5p miR-331-3p miR-5588-3p miR-6801-3p miR-6810-3p miR-10a-5p miR-10b-5p miR-624-5p
rs3734215	HDGFL1	HDGF like 1	dsm4dep	C	19%	-0.101	2.62E-01	-0.181	7.38E-02	no	ELAVL1	
rs3734214 ^d	HDGFL1	HDGF like 1	dsm5sev	G	25%	-0.104	7.79E-02	-0.271	1.57E-06	no	ELAVL1	
rs11751669	HDGFL1	HDGF like 1	dsm5sev	G	14%	-0.038	1.17E-01	-0.015	6.38E-01	no	ELAVL1	
rs7195	HLA-DRA	major histocompatibility complex, class II, DR alpha	withdrawal	G	3%	-0.084	2.89E-02	-0.229	1.52E-04	yes	ELAVL1, PABPC1	miR-6507-3p
rs7194	HLA-DRA	major histocompatibility complex, class II, DR alpha	withdrawal	A	3%	-0.162	2.65E-01	-0.715	2.62E-03	yes	ELAVL1, PABPC1	miR-4728-3p
rs7251209	HSH2D	hematopoietic SH2 domain containing	dsm5score	G	6%	0.099	4.89E-01	0.141	4.03E-02	no	SLBP, PABPC1	
rs7254531	HSH2D	hematopoietic SH2 domain containing	dsm5score	C	6%	-0.042	7.83E-02	0.038	7.68E-02	no	SLBP, PABPC1	
rs7250494	HSH2D	hematopoietic SH2 domain containing	dsm5score	T	6%	0.082	1.57E-01	-0.039	4.43E-01	no	SLBP, PABPC1	
rs11677686 ^d	ICA1L	islet cell autoantigen 1 like	crave	C	22%	0.033	8.45E-01	0.069	1.09E-01	no	PABPC1, SLBP	miR-2276-3p miR-4772-3p

SNP ^d	Gene	Description	Phenotype ^b	ALT allele	ALT freq	RNA alt allele freq vs DNA alt allele freq				eQTL	RBP binding	Potential miRNA targets
						SH-SY5Y		SK-N-BE(2)				
						log2 ratio	FDR ^c	log2 ratio	FDR ^c			
rs2240773	JMJD6	jumonji domain containing 6, arginine demethylase and lysine hydroxylase	sre_first5	A	14%	0.044	2.29E-01	0.131	8.52E-05	yes	PABPC1	miR-3160-3p miR-143-3p miR-4756-3p miR-4770 miR-6088
rs2276608 ^d	KCMF1	potassium channel modulatory factor 1	dsm4dep	C	8%	-0.204	5.41E-05	-0.120	1.33E-02	yes		miR-1183 miR-4773 miR-3190-3p miR-4318 miR-6833-5p miR-7114-5p
rs56370893 ^d	KIF6	kinesin family member 6	flush	A	7%	-0.328	2.03E-10	-0.339	3.81E-09	yes		
rs72858451 ^d	KIF6	kinesin family member 6	flush	G	7%	-0.002	5.89E-01	-0.092	5.64E-04	no		
rs16891940 ^d	KIF6	kinesin family member 6	flush	G	6%	-0.133	3.50E-03	-0.094	3.88E-02	yes		
rs11756084 ^d	KIF6	kinesin family member 6	flush	G	6%	-0.021	5.29E-01	-0.049	7.60E-02	yes		
rs72858444 ^d	KIF6	kinesin family member 6	flush	T	4%	0.107	2.08E-01	0.074	1.09E-01	yes		
rs78824997	KIF6	kinesin family member 6	flush	C	6%	0.121	2.65E-01	-0.206	1.63E-01	no		
rs11756848	KIF6	kinesin family member 6	flush	G	7%	-0.154	6.26E-03	-0.038	6.37E-01	no		
rs75700740 ^d	KPNA4	karyopherin subunit alpha 4	flush	C	4%	-0.592	5.61E-07	-0.181	2.54E-01	no	ELAVL1	miR-21-5p miR-216a-5p miR-216b-5p miR-5579-3p * miR-590-5p miR-452-3p
rs2665 ^d	KRTAP5-9	keratin associated protein 5-9	time_used	A	18%	-0.030	2.29E-01	-0.050	7.68E-02	no	ELAVL1, PABPC1, SLBP	
rs62137532 ^d	LHCGR	lutetizing hormone/choriogonadotropin receptor	dom_life	G	16%	0.353	6.32E-01	0.873	1.34E-03	yes	ELAVL1, PABPC1, SLBP, CELF1	miR-4639-3p miR-4432 miR-4509 miR-4744 miR-548u miR-7161-5p
rs1943488 ^d	LRRRC55	leucine rich repeat containing 55	flush	T	9%	0.000	7.19E-01	0.047	3.54E-02	yes		miR-8080 miR-3617-5p * miR-641

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
				A	10%	0.089		-0.223	3.88E-02	yes		
				A	27%	0.106		0.079	1.53E-01	yes	PABPC1, SLBP, ELAVL1	miR-6800-3p miR-6845-3p
				T	8%	0.257	1.92E-01	0.479	3.99E-02	yes	ELAVL1	
				A	31%	-0.010	4.62E-01	0.053	4.00E-04	yes		miR-891b
				C	27%	-0.090	5.47E-02	-0.228	1.44E-06	no	ELAVL1, PABPC1	miR-4719 miR-513b-5p miR-27a-3p *miR-27b-3p miR-3185 miR-513a-5p
				G	34%	0.180	7.05E-02	0.019	8.51E-01	no	ELAVL1, PABPC1	miR-377-5p miR-6086 miR-655-5p
				G	28%	-0.063	1.16E-04	-0.018		no	ELAVL1, PABPC1	miR-625-3p miR-139-5p miR-582-5p miR-935
				T	11%	-0.200	1.10E-06	-0.170	2.78E-02	no	PABPC1	miR-1469 miR-129-5p miR-4467 miR-4730 miR-6770-3p
				C	27%	0.166	5.29E-02	-0.020	8.18E-01	yes	PABPC1	miR-1178-5p miR-6823-5p
				G	35%	-0.073	1.38E-01	-0.180	6.29E-07	no	ELAVL1	miR-3613-3p
				C	8%	-0.197	9.66E-10	-0.135	3.45E-04	no	PABPC1, ELAVL1	miR-4643 miR-598-3p
				A	31%	0.141	3.95E-02	0.029	6.66E-01	yes	CELF1, PABPC1	miR-5003-5p
				G	34%	0.279	3.46E-10	0.244	1.03E-06	yes	ELAVL1, PABPC1, CELF1	miR-4652-3p miR-4743-3p miR-654-3p

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
SNP ^d	Gene	Description	Phenotype ^b									
rs9995	NBN	nibrin	dsm4sxc	G	34%	-0.114	3.09E-03	-0.152	2.78E-02	yes	ELAVL1, PABPC1, CELF1	miR-5580-3p miR-6082
rs484555	NFIB	nuclear factor 1B	dsm5modse _v	G	10%	-0.081	1.92E-01	-0.110	1.27E-01	no	IGF2BP1, ELAVL1	miR-3163 miR-3175 miR-577 miR-6799-5p miR-6825-5p
rs72629150 ^d	NLRP5	NLR family pyrin domain containing 5	time_used	A	17%	-0.026	2.08E-01	-0.108	1.34E-03	no	ELAVL1, PABPC1, SLBP	miR-431-5p miR-4635 miR-654-3p
rs12665231 ^d	NOX3	NADPH oxidase 3	tolerance	C	19%	-0.031		-0.262	4.99E-06	no	ELAVL1	miR-452-5p miR-4676-3p miR-892c-3p
rs2334868	NRIIP2	nuclear receptor interacting protein 2	physical	G	12%	-0.037	1.38E-01	-0.084	2.04E-02	yes	ELAVL1	miR-1294 miR-3174 miR-921
rs10046724	NUGGC	nuclear GTPase, germinal center associated	desire_cut	A	38%	0.018		0.093	9.16E-03	no	PABPC1, CELF1	*miR-186-5p
rs16930262	P4HAI	prolyl 4-hydroxylase subunit alpha 1	dsm5modse _v	T	18%	0.064	5.19E-02	0.135	5.29E-23	no	ELAVL1, PABPC1	
rs12610811 ^d	PGPEPI	pyroglutamyl-peptidase I	sre_first5	G	48%	-0.255	1.17E-01	0.460	7.10E-04	no	ELAVL1, CELF1, PABPC1	*miR-186-3p miR-4281 miR-7152-3p
rs78993050	PIK3CA	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha	flush	C	4%	0.024	9.63E-01	0.255	1.42E-08	no	ELAVL1	
rs12934430	PKDIL2	polycystin 1 like 2 (gene/pseudogene)	desire_cut	C	25%	-0.033	8.08E-01	-0.353	2.37E-05	yes	PABPC1	miR-6828-5p miR-665 miR-671-5p miR-6840-3p
rs11648241	PKDIL2	polycystin 1 like 2 (gene/pseudogene)	desire_cut	C	28%	-0.004	5.78E-01	-0.045	1.58E-01	yes	PABPC1	miR-4704-3p
rs7196623	PKDIL2	polycystin 1 like 2 (gene/pseudogene)	dsm5modse _v	T	28%	-0.109	1.03E-01	0.034	8.18E-01	yes	PABPC1	miR-101-3p miR-144-3p miR-3927-3p miR-4451 miR-6831-5p

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
						ALT allele	ALT freq	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
SNP ^d	Gene	Description	Phenotype ^b	ALT allele	ALT freq							
rs1901818	PKDIL2	polycystin 1 like 2 (gene/ pseudogene)	dsm5sev	G	33%			0.064	3.63E-02	yes	PABPC1	miR-6500-3p miR-494-3p
rs32613	PLK2	polo like kinase 2	time_used	C	2%			-0.224	4.91E-05	no	PABPC1, ELAVL1	miR-3130-5p miR-4482-5p miR-500b-3p
rs3741133	PPME1	protein phosphatase methyltransferase 1	sre_total	A	19%			0.050	6.61E-01	yes	CELF1	
rs6822927 ^d	PSAPL1	prosaposin like 1 (gene/ pseudogene)	dsm5sev	C	2%			-0.118	3.30E-02	no	PABPC1, ELAVL1	miR-1470 miR-331-3p miR-345-3p miR-6801-3p miR-6810-3p
rs73413219 ^d	PYGO1	pygopus family PHD finger 1	withdrawal	C	21%			-0.007	8.08E-01	no		
rs16976371	PYGO1	pygopus family PHD finger 1	withdrawal	C	21%			0.096	1.33E-01	no		
rs3811126 ^d	QSOX2	quiescin sulphydryl oxidase 2	dom_life	G	4%			-0.087	1.87E-02	no	IGF2BP1, PABPC1, SLBP, CELF1	miR-3190-3p miR-3650 miR-3911 miR-6833-5p
rs1128287	RAB24	RAB24, member RAS oncogene family	crave	T	17%			-0.335	3.33E-03	yes		miR-128-3p miR-216a-3p miR-27a-3p *miR-27b-3p miR-3681-3p miR-513a-5p miR-8070
rs226418 ^d	RASL10B	RAS like family 10 member B	flush	C	17%			-0.005	8.53E-01	yes		miR-4313 miR-6784-3p miR-6802-3p miR-6862-3p
rs226417	RASL10B	RAS like family 10 member B	flush	A	17%			-0.657	6.39E-02	yes		miR-4640-3p miR-6798-3p miR-4758-3p
rs75215933	RGS19	regulator of G protein signaling 19	dsm4dep	C	10%			0.025	9.71E-01	no	ELAVL1, IGF2BP1, PABPC1	miR-604 *miR-647 miR-6762-3p miR-6842-3p
rs1133926 ^d	RPL15	ribosomal protein L15	dsm5dep	G	13%			0.118	3.76E-02	no	PABPC1	miR-3679-3p miR-138-1-3p

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
SNP ^d	Gene	Description	Phenotype ^b									
rs974962	RPP21	ribonuclease P/MRP subunit p21	maxdrinks	G	20%	-0.018	4.62E-01	-0.043	4.27E-02	yes	IGF2BP1, PABPC1	miR-1273h-5p miR-30b-3p miR-30c-1-3p miR-30c-2-3p miR-3122 miR-3689a-3p miR-3689b-3p miR-3689c miR-3913-5p miR-450a-1-3p miR-5002-5p miR-6513-5p miR-6731-5p miR-6779-5p miR-6780a-5p miR-6788-5p miR-6878-5p miR-8085 miR-887-5p miR-425-3p miR-6746-5p miR-6771-5p
rs41292422 ^d	RTN4IP1	reticulon 4 interacting protein 1	dsm4dep	G	2%	-0.055	3.39E-01	-0.136	2.90E-02	yes	CELF1, ELAVL1, PABPC1	miR-3613-3p
rs3200575	SETD6	SET domain containing 6, protein lysine methyltransferase	tolerance	A	49%	0.105	1.03E-02	0.115	4.00E-04	yes		miR-1273h-3p miR-3158-5p miR-3166 miR-3173-5p miR-4771 miR-6799-3p miR-146b-3p miR-3684
rs4863654 ^d	SETD7	SET domain containing 7, histone lysine methyltransferase	physical	A	34%	0.034	5.89E-01	0.086	4.02E-02	no		
rs73071223	SIRPA	signal regulatory protein alpha	more_intend	A	4%	0.029	9.79E-01	-0.342	1.47E-06	no	PABPC1	miR-299-3p miR-4491 miR-4657 miR-147a miR-3650 miR-3911 *miR-3934-5p miR-644a miR-6867-5p

						RNA alt allele freq vs DNA alt allele freq							
						SH-SY5Y		SK-N-BE(2)					
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets	
SNP ^d	Gene	Description	Phenotype ^b	C	4%	-0.242	3.95E-02	0.016	7.13E-01	no	PABPC1	miR-299-3p miR-3151-5p miR-328-5p miR-4447 miR-4472 miR-491-5p miR-6742-5p miR-6795-5p miR-6796-5p miR-6885-5p miR-6887-5p miR-7109-5p miR-1275 miR-3147 miR-409-3p miR-4425 miR-4525 miR-4665-5p miR-4723-5p miR-4781-3p *miR-5010-5p miR-5698 miR-625-5p miR-6751-5p miR-6803-5p miR-6870-5p miR-7111-5p	
rs78695937 ^d	SLC25A4I	solute carrier family 25 member 41	withdrawal	T	3%	-0.037	6.26E-03	-0.025	1.13E-01	yes	ELAVL1, PABPC1	miR-197-3p miR-6511a-3p miR-6511b-3p miR-6777-3p	
rs2466295	SLC30A8	solute carrier family 30 member 8	dsm5score	T	3%	0.076		0.191	4.02E-02	no	SLBP, ELAVL1	* miR-3662 miR-586	
rs2919865 ^d	SMYDI	SET and MYND domain containing 1	more_intend	T	44%	-0.066	6.37E-04	-0.090	1.77E-03	yes			
rs2970889	SMYDI	SET and MYND domain containing 1	more_intend	T	44%	0.016	8.53E-01	-0.111	2.78E-02	yes		miR-145-5p miR-1825 miR-199a-5p *miR-199b-5p miR-203b-5p miR-5195-3p miR-6718-5p	

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
rs3815292	<i>SNED1</i>	sushi, nidogen and EGF like domains 1	crave	G	9%	0.187	1.47E-01	-0.167	7.50E-02	no	PABPC1, SLBP	miR-144-5p miR-34c-3p miR-3689a-5p miR-3689b-5p miR-3689e miR-3689f miR-1255b-2-3p
rs1060151 ^d	<i>SPARC</i>	secreted protein acidic and cysteine rich	dsm5sev	T	1%	-0.137	2.29E-01	0.238	5.17E-02	no	ELAVL1, PABPC1	*miR-1250-3p miR-153-5p miR-335-3p miR-5696 miR-579-3p miR-664b-3p
rs3087986	<i>SSR1</i>	signal sequence receptor subunit 1	dsm4sxct	T	38%	-0.033	3.06E-01	-0.261	2.48E-03	yes	ELAVL1, PABPC1	miR-4690-3p miR-5685 miR-3189-5p miR-4323
rs11071	<i>SSR1</i>	signal sequence receptor subunit 1	dsm4sxct	C	38%	-0.100	1.39E-01	-0.157	8.60E-02	yes	ELAVL1, PABPC1	miR-204-3p miR-4646-5p miR-582-3p miR-450a-1-3p
rs9505115^d	<i>SSR1</i>	signal sequence receptor subunit 1	dsm4sxct	C	37%	-0.107	4.38E-04	-0.194	1.62E-05	yes		*miR-3934-5p miR-552-3p miR-4438 miR-5095 miR-7151-3p
rs3772105	<i>STAC</i>	SH3 and cysteine rich domain	maxdrinks	G	16%	-0.041	2.62E-01	-0.102	1.81E-01	no	PABPC1	
rs2281789 ^d	<i>SUSD1</i>	sushi domain containing 1	flush	C	10%	0.239	3.31E-08	0.104	1.95E-01	no		*miR-1304-3p miR-217 miR-448 * miR-6807-3p miR-6890-3p miR-483-3p
rs6773595	<i>SUSD5</i>	sushi domain containing 5	dsm5sev	C	24%	-0.076	6.93E-04	0.027	1.99E-01	yes		miR-1910-5p miR-2114-3p miR-4264 miR-455-3p miR-6823-3p *miR-1225-3p miR-1233-3p miR-4323

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
SNP ^a	Gene	Description	Phenotype ^b									
rs59054390 ^d	<i>SVOP</i>	SV2 related protein	dsm5sev	A	10%	-0.084	4.91E-05	-0.051	7.02E-03	no		
rs274674	<i>TENT4A</i>	terminal nucleotidyltransferase 4A	dsm5sev	T	29%	0.082	8.84E-02	-0.017	6.62E-01	yes		miR-132-5p miR-4800-3p miR-7847-3p miR-4802-5p miR-7151-5p
rs2294532 ^e	<i>THAP3</i>	THAP domain containing 3	tolerance	T	30%	-0.096	8.59E-17	-0.164	3.53E-09	yes		miR-3909 miR-4421 miR-5699-3p miR-6748-3p miR-6852-3p miR-6881-3p miR-877-3p miR-1266-3p miR-2117 miR-266-3p miR-4273 miR-4753-3p miR-6809-3p miR-7156-5p
rs1062919	<i>TMEM143</i>	transmembrane protein 143	maxdrinks	A	34%	-0.206	6.48E-07	-0.192	2.21E-04	yes	IGF2BP1, PABPC1, SLBP	
rs3745720	<i>TMEM143</i>	transmembrane protein 143	maxdrinks	C	35%	-0.006	7.57E-01	0.076	1.47E-01	yes	IGF2BP1, PABPC1, SLBP	
rs4693446 ^d	<i>TMEM150_C</i>	transmembrane protein 150C	tolerance	A	10%	0.102		0.132	5.38E-04	no		miR-4645-5p miR-4673 miR-1911-3p miR-643 miR-6753-5p
rs7107522	<i>TMEM80</i>	transmembrane protein 80	dsm4sxct	G	21%	0.170	1.89E-01	0.323	7.34E-04	yes	PABPC1, ELAVL1	miR-199a-5p *miR-199b-5p miR-3151-5p miR-4302 miR-4447 miR-4472 miR-491-5p miR-6742-5p miR-6796-5p miR-6853-5p

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
SNP ^d	Gene	Description	Phenotype ^b									
rs1561743 ^d	<i>TOMM70</i>	translocase of outer mitochondrial membrane 70	sre_total	G	19%	-0.075	9.09E-02	-0.203	1.63E-01	yes	CELF1, PABPC1	miR-1273h-3p miR-2467-3p miR-3678-3p miR-6873-5p
rs17519439	<i>USP12</i>	ubiquitin specific peptidase 12	dom_life	A	5%	-0.016	4.66E-01	0.056	1.48E-01	yes	ELAVL1, CELF1, PABPC1	miR-1252-3p miR-1909-5p * miR-3662
rs11075337	<i>XYLT1</i>	xylosyltransferase 1	dsm5sev	T	6%	-0.065	3.33E-03	-0.024	3.47E-01	no	ELAVL1, PABPC1, SLBP	miR-4757-3p miR-1258
rs60382924 ^d	<i>ZC3H12D</i>	zinc finger CCCH-type containing 12D	dsm5sev	T	11%	0.066	1.68E-01	-0.023	5.30E-01	no	ELAVL1, PABPC1	miR-4474-3p miR-4687-3p miR-595 miR-644a miR-7974
rs7579561 ^d	<i>ZC3H6</i>	zinc finger CCCH-type containing 6	dsm4sxct	G	20%	0.040	8.45E-01	-0.088	3.54E-02	yes		miR-3120-5p miR-3664-5p miR-4522 miR-6888-3p miR-340-3p miR-6827-3p
rs7591439	<i>ZC3H6</i>	zinc finger CCCH-type containing 6	dsm4sxct	A	8%	-0.681	7.19E-05	0.102	6.42E-01	no		miR-197-3p miR-5096 miR-7154-5p miR-8060
rs17445167 ^d	<i>ZFX4</i>	zinc finger homeobox 4	dsm4sxct	A	1%	0.010	9.31E-01	-0.066	1.56E-02	yes		miR-494-3p
rs7206003	<i>ZFP1</i>	ZFP1 zinc finger protein	grave	G	44%	-0.293	5.02E-03	-0.477	1.47E-06	yes	ELAVL1, PABPC1	miR-7849-3p

						RNA alt allele freq vs DNA alt allele freq						
						SH-SY5Y		SK-N-BE(2)				
				ALT allele	ALT freq	log2 ratio	FDR ^c	log2 ratio	FDR ^c	eQTL	RBP binding	Potential miRNA targets
SNP ^a	Gene	Description	Phenotype ^b									
rs6443616 ^d	ZMAT3	zinc finger matrin-type 3	dsm4sxct	G	22%	0.009		0.166	1.59E-02	yes	ELAVL1, PABPC1	miR-147a miR-3151-5p miR-3911 miR-4447 miR-4472 miR-491-5p miR-644a miR-6742-5p miR-6751-5p miR-6752-5p miR-6796-5p miR-6803-5p miR-6835-5p miR-6842-5p miR-6867-5p miR-7110-5p miR-328-5p miR-6795-5p miR-6885-5p miR-6887-5p miR-7109-5p
rs729739	ZNF365	zinc finger protein 365	sre_total	A	13%	0.114	2.62E-01	0.336	1.03E-06	yes	ELAVL1, PABPC1	miR-135a-5p miR-135b-5p miR-299-5p miR-8069
rs1048192 ^d	ZNF469	zinc finger protein 469	dsm5modse _v	A	5%	0.015	8.16E-01	0.130	1.52E-04	no		
rs4806880	ZNF555	zinc finger protein 555	dom_life	T	26%	-0.123	7.95E-08	-0.111	2.62E-03	yes	PABPC1	miR-335-5p miR-4798-3p
rs2317033 ^d	ZNF555	zinc finger protein 555	dom_life	A	27%	-0.031	7.08E-01	-0.089	2.02E-02	yes		miR-4666a-5p
rs10418996	ZNF555	zinc finger protein 555	dom_life	G	30%	0.128	4.25E-01	0.239	2.78E-02	yes		miR-4455 miR-609 miR-6748-5p miR-6772-5p miR-562
rs6510709 ^d	ZNF555	zinc finger protein 555	dom_life	G	8%	0.052	4.70E-01	0.059	1.81E-01	no		miR-323a-5p miR-876-3p

SNP ^a	Gene	Description	Phenotype ^b	ALT allele	ALT freq	RNA alt allele freq vs DNA alt allele freq				eQTL	RBP binding	Potential miRNA targets
						SH-SY5Y		SK-N-BE(2)				
						log2 ratio	FDR ^c	log2 ratio	FDR ^c			
rs7256156 ^d	ZNF611	zinc finger protein 611	flush	C	15%	-0.065	1.35E-01	-0.158		yes	ELAVL1, PABPC1	miR-513a-5p miR-1281 miR-1915-3p miR-4726-3p miR-6764-5p miR-6778-3p miR-6840-3p
rs111703885 ^d	ZNF705E	zinc finger protein 705E	dsm4dep	G	6%	-0.081	3.01E-01	-0.161	1.91E-01	no		
rs2068630 ^d	ZNF765	zinc finger protein 765	flush	T	35%	-0.396	8.62E-03	-0.321	2.52E-01	yes	ELAVL1	*miR-3064-3p miR-4715-3p miR-4776-3p
rs1834950 ^d	ZNF99	zinc finger protein 99	sre_first5	C	12%	0.261	9.16E-05	0.245	3.39E-11	yes		*miR-25-5p miR-6087 miR-658 miR-1827 miR-1910-3p miR-2467-3p miR-3612 miR-4257 miR-650 miR-6511a-5p
rs71357930 ^d	ZNF99	zinc finger protein 99	sre_first5	T	12%	0.043	6.01E-01	0.218	1.47E-06	yes		miR-1227-3p miR-5186 miR-550b-2-5p miR-4727-5p miR-4769-3p miR-5001-3p miR-6738-3p miR-6817-5p
rs1834949 ^d	ZNF99	zinc finger protein 99	sre_first5	A	12%	-0.107	6.70E-04	-0.001		yes		miR-130a-5p miR-23a-3p miR-23b-3p *miR-23c miR-412-3p miR-5580-3p miR-6754-3p miR-6837-3p miR-5007-3p miR-8076

^aSNPs at FDR 0.05 in both cell lines are in **bold**.^bFor description of the phenotypes, see Supplementary Table S1

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

FDR false discovery rate; bold for p Imputed SNPs, imputation quality score 0.3
SNP located in 3'UTR of two genes on opposite strands; effect tested separately.
* miRNAs with median TPM > 0.5 in at least one brain tissue from GTEx v8 RNA-seq