OPEN

Citation: Transl Psychiatry (2014) **4**, e391; doi:10.1038/tp.2014.29

www.nature.com/tp

ORIGINAL ARTICLE

Genetic risk prediction and neurobiological understanding of alcoholism

DF Levey[1], H Le-Niculescu[1], J Frank[2], M Ayalew[1], N Jain[1], B Kirlin[1], R Learman[1], E Winiger[1], Z Rodd[1], A Shekhar[1], N Schork[3], F Kiefe[4], N Wodarz[5], B Müller-Myhsok[6], N Dahmen[7], GESGA Consortium, M Nöthen[8], R Sherva[9], L Farrer[9], AH Smith[10], HR Kranzler[11], M Rietschel[2], J Gelernter[10] and AB Niculescu[1,12]

We have used a translational Convergent Functional Genomics (CFG) approach to discover genes involved in alcoholism, by gene-level integration of genome-wide association study (GWAS) data from a German alcohol dependence cohort with other genetic and gene expression data, from human and animal model studies, similar to our previous work in bipolar disorder and schizophrenia. A panel of all the nominally significant *P*-value single-nucleotide length polymorphisms (SNPs) in the top candidate genes discovered by CFG ($n = 135$ genes, 713 SNPs) was used to generate a genetic risk prediction score (GRPS), which showed a trend towards significance ($P = 0.053$) in separating alcohol dependent individuals from controls in an independent German test cohort. We then validated and prioritized our top findings from this discovery work, and subsequently tested them in three independent cohorts, from two continents. In order to validate and prioritize the key genes that drive behavior without some of the pleiotropic environmental confounds present in humans, we used a stress-reactive animal model of alcoholism developed by our group, the D-box binding protein (DBP) knockout mouse, consistent with the surfeit of stress theory of addiction proposed by Koob and colleagues. A much smaller panel ($n = 11$ genes, 66 SNPs) of the top CFG-discovered genes for alcoholism, cross-validated and prioritized by this stress-reactive animal model showed better predictive ability in the independent German test cohort ($P = 0.041$). The top CFG scoring gene for alcoholism from the initial discovery step, synuclein alpha (SNCA) remained the top gene after the stress-reactive animal model cross-validation. We also tested this small panel of genes in two other independent test cohorts from the United States, one with alcohol dependence ($P = 0.00012$) and one with alcohol abuse (a less severe form of alcoholism; $P = 0.0094$). SNCA by itself was able to separate alcoholics from controls in the alcohol-dependent cohort ($P = 0.000013$) and the alcohol abuse cohort ($P = 0.023$). So did eight other genes from the panel of 11 genes taken individually, albeit to a lesser extent and/or less broadly across cohorts. SNCA, GRM3 and MBP survived strict Bonferroni correction for multiple comparisons. Taken together, these results suggest that our stress-reactive DBP animal model helped to validate and prioritize from the CFG-discovered genes some of the key behaviorally relevant genes for alcoholism. These genes fall into a series of biological pathways involved in signal transduction, transmission of nerve impulse (including myelination) and cocaine addiction. Overall, our work provides leads towards a better understanding of illness, diagnostics and therapeutics, including treatment with omega-3 fatty acids. We also examined the overlap between the top candidate genes for alcoholism from this work and the top candidate genes for bipolar disorder, schizophrenia, anxiety from previous CFG analyses conducted by us, as well as cross-tested genetic risk predictions. This revealed the significant genetic overlap with other major psychiatric disorder domains, providing a basis for comorbidity and dual diagnosis, and placing alcohol use in the broader context of modulating the mental landscape.

Translational Psychiatry (2014) **4,** e391; doi:10.1038/tp.2014.29; published online 20 May 2014

INTRODUCTION

'Drunkenness is nothing but voluntary madness.'

—Seneca

Alcohol use and overuse (alcoholism) have deep historical and cultural roots, as well as important medical and societal consequences.[1] Whereas there is evidence for roles for both

[1]Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN, USA; [2]Central Institute of Mental Health, Mannheim, Germany; [3]Department of Human Biology, The J. Craig Venter Institute, La Jolla, CA, USA; [4]Department of Addictive Behavior and Addiction Medicine, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Heidelberg, Germany; [5]Department of Psychiatry, University Medical Center Regensburg, University of Regensburg, Regensburg, Germany; [6]Department of Statistical Genetics, Max-Planck-Institute of Psychiatry, Munich, Germany; [7]Department of Psychiatry, University of Mainz, Mainz, Germany; [8]Department of Genomics, Life & Brain Center, Institute of Human Genetics, University of Bonn, Bonn, Germany; [9]Department of Medicine (Biomedical Genetics), Boston University School of Medicine, Boston, MA, USA; [10]Division of Human Genetics, Department of Psychiatry, Yale University School of Medicine, VA CT Healthcare Center, New Haven, CT, USA; [11]Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, and Philadelphia VAMC, Philadelphia, PA, USA and [12]Indianapolis VA Medical Center, Indianapolis, IN, USA. Correspondence: Professor AB Niculescu, Department of Psychiatry, Indiana University School of Medicine, Neuroscience Research Building, 320 W. 15th Street, Indianapolis, IN 46202, USA.
E-mail: anicules@iupui.edu
Received 23 January 2014; accepted 18 March 2014

genes and environment in alcoholism, a comprehensive biological understanding of the disorder has been elusive so far, despite extensive work in the field. Most notably, there has been until recently insufficient translational integration across functional and genetic studies, and across human and animal model studies, resulting in missed opportunities for a comprehensive understanding.

As part of a translational Convergent Functional Genomics (CFG) approach, developed by us over the last 15 years,[2] and expanding upon our earlier work on identifying genes for alcoholism,[3–5] we set out to comprehensively identify candidate genes, pathways and mechanisms for alcoholism, integrating the available evidence in the field to date. We have used data from a published German genome-wide association study for alcoholism.[6] We integrated those data in a Bayesian-like manner with other human genetic data (association or linkage) for alcoholism, as well as human gene expression data, post-mortem brain gene expression data and peripheral (blood and cell culture) gene expression data. We also used relevant animal model genetic data (transgenic and quantitative trait loci (QTL)), as well as animal model gene expression data (brain and blood) generated by our group and others (Figures 1 and 2). Human data provide specificity for the illness, and animal model data provide sensitivity of detection. Together, they helped to identify and prioritize candidate genes for the illness using a polyevidence CFG score, resulting in essence in a *de facto* field-wide integration putting together all the available lines of evidence to date. Once that is done, biological pathway analyses can be conducted and mechanistic models can be constructed.

An obvious next step is developing a way of applying that knowledge to genetic testing of individuals to determine risk for the disorder. On the basis of our comprehensive identification of top candidate genes described in this paper, we have chosen all the nominally significant *P*-value SNPs corresponding to each of those 135 genes from the GWAS data set used for discovery (top candidate genes prioritized by CFG with the score of 8 and above ($\geq$ 50% maximum possible CFG score of 16) and assembled a Genetic Risk Prediction panel out of those 713 SNPs. We then developed a Genetic Risk Prediction Score (GRPS) for alcoholism based on the presence or absence of the alleles of the SNPs associated with the illness from the discovery GWAS, and tested the GRPS in an independent German cohort,[51] to see whether it can differentiate alcohol-dependent subjects from controls, observing a trend towards significance.

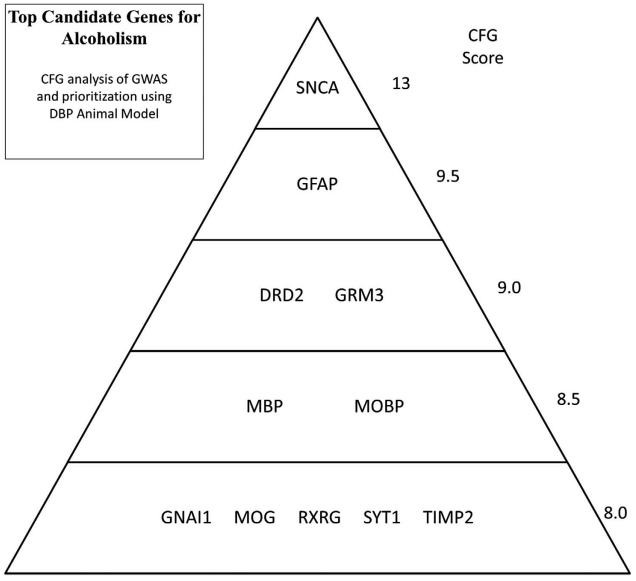


Figure 2. Top candidate genes for alcoholism.

In order to validate and prioritize genes in this panel using a behavioral prism, we then looked at the overlap between our panel of 135 top candidate genes and genes changed in expression in a stress-reactive animal model for alcoholism developed by our group, the DBP knockout mouse.[4,5] We used this overlap to reduce our panel to 11 genes (66 SNPs).

This small panel of 11 genes was subsequently tested and shown to be able to differentiate between alcoholics and controls in the three independent test cohorts, one German[51] and two US-based,[52] suggesting that the animal model served in essence as a filter to identify from the larger list of CFG-prioritized genes the key behaviorally relevant genes. Our results indicate that panels of SNPs in top genes identified and prioritized by CFG analysis and by a behaviorally relevant animal model can differentiate between alcoholics and controls at a population level (Figure 3), although at

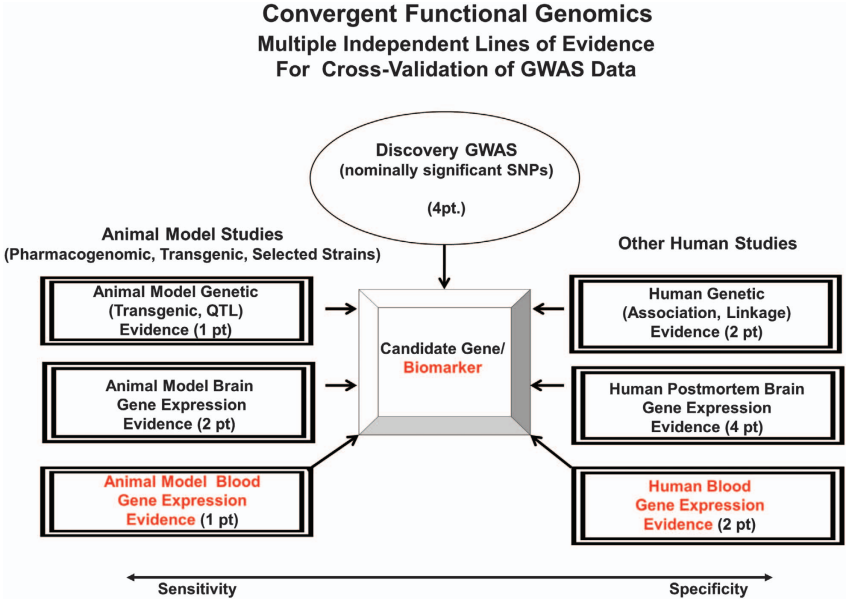
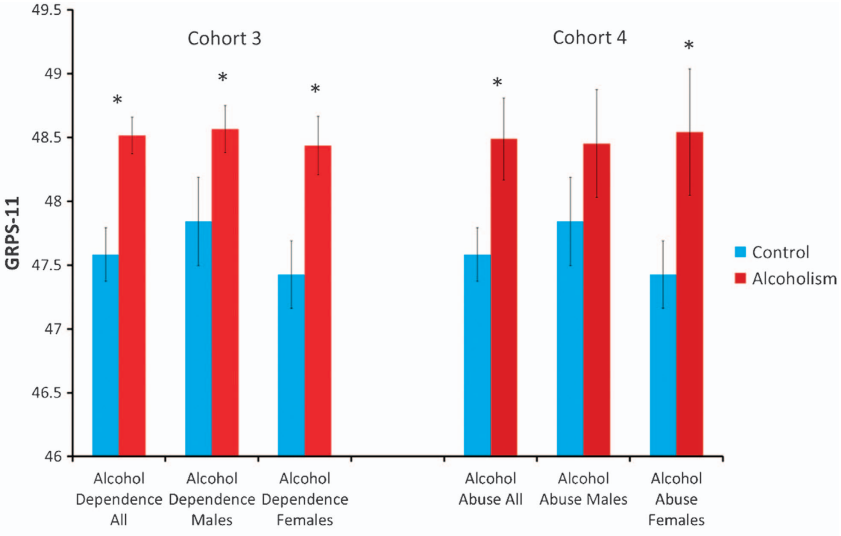


Figure 1. Convergent Functional Genomics.



Comparison	p-value
Alcohol Dependence All (2768) vs. Control (1261)	**0.00012486**
Alcohol Dependence Females (1081) vs. Control Females (786)	**0.00196973**
Alcohol Abuse All (600) vs. Control (1261)	**0.00939608**
Alcohol Abuse Females (234) vs. Control Females (786)	**0.02399935**
Alcohol Dependence Males (1687) vs. Control Males (475)	**0.03215217**
Alcohol Abuse Males (366) vs. Control Males (475)	0.13134934

Figure 3. Genetic Risk Prediction using a panel of top candidate genes for alcoholism (GRPS-11). Testing in independent cohorts 3 and 4.

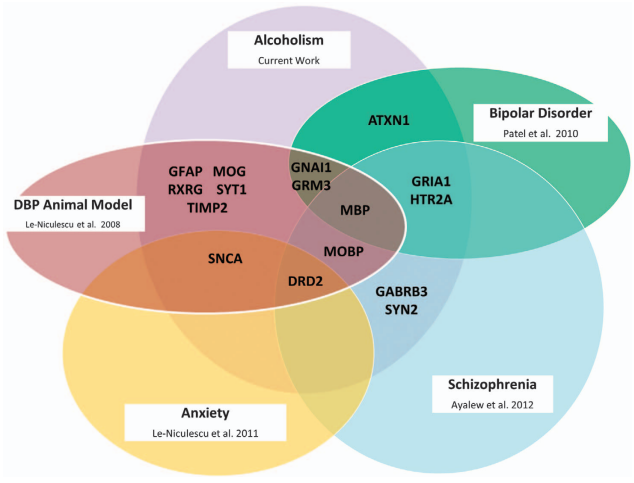


Figure 4. Overlap of alcoholism versus other major psychiatric disorders. Top candidate genes for alcoholism identified by CFG ($n = 135$) in the current study versus top candidate genes for other psychiatric disorders and a stress-driven animal model of alcoholism (DBP knockout mouse) from our previous work.

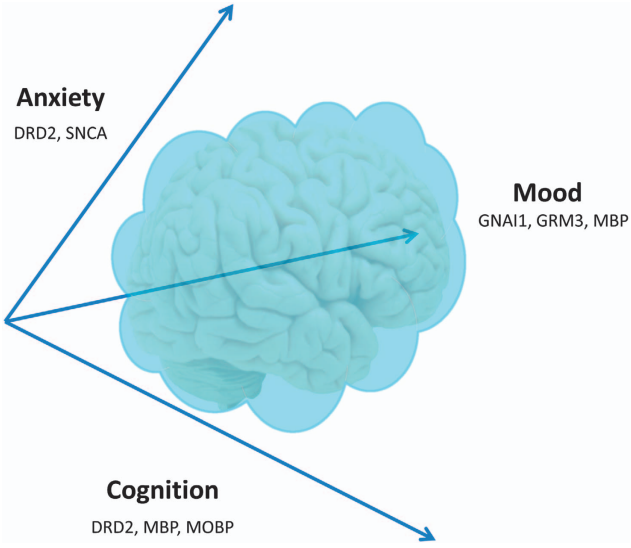


Figure 5. Mindscape (mental landscape)-dimensional view of genes that may be involved in alcoholism and other major psychiatric disorders.

an individual level the margin may be small (Supplementary Figure S2). The latter point suggests that, similar to bipolar disorder[53] and schizophrenia,[54] the contextual cumulative combinatorics of common gene variants and environment[55] has a major role in risk for illness.

Lastly, we have looked at overlap with genes for other major psychiatric disorder domains (bipolar disorders, anxiety disorders, schizophrenias) from our previous studies and provide evidence for shared genes (Figures 4 and 5) as well as shared genetic risk (Figure 6).

Overall, this work sheds light on the genetic architecture and pathophysiology of alcoholism, provides mechanistic targets for therapeutic intervention and has implications for genetic testing to assess risk for illness before the illness manifests itself clinically, opening the door for enhanced prevention strategies at a young age. As alcoholism is a disease that does not exist if the exogenous

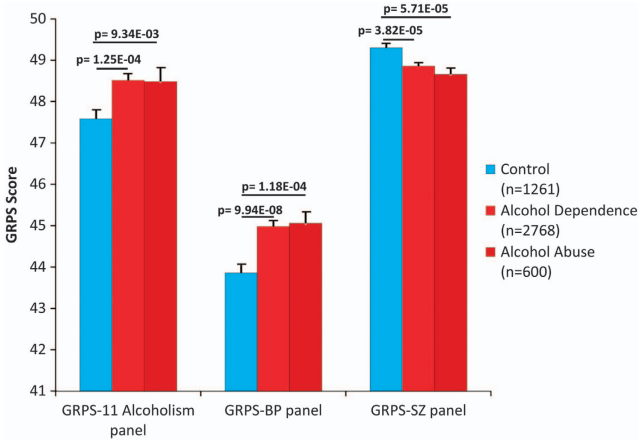


Figure 6. Genetic load for bipolar disorder and schizophrenia in alcoholism. A total of 34 out of 66 SNPs in our alcohol GRPS-11 panel (current work; in $n=10$ genes), 42 out of 224 SNPs in our bipolar GRPS[53] (in $n=34$ genes) and 151 out of 542 SNPs in our schizophrenia GRPS[54] (in $n=35$ genes) were present and tested in the alcohol cohorts 3 and 4. See also Supplementary Table S7.

agent (alcohol) is not consumed, the use of genetic information to inform lifestyle choices could be quite powerful.

MATERIALS AND METHODS

Human subject cohorts

Discovery cohort (cohort 1): GWAS for alcohol dependence from Germany. Data for the discovery CFG work (Cohort 1) were obtained from a GWA study of self-reported German descent subjects, consisting of 411 alcohol-dependent male subjects and 1307 population-based controls (663 male and 644 female subjects).[6] Individuals were genotyped using HumanHap 550 BeadChips (Illumina Inc, San Diego, CA, USA). SNPs with a nominal allelic *P*-value <0.05 were selected for analysis. No Bonferroni correction was performed.

Test cohort 2 (alcohol dependence, Germany). An independent test cohort of German descent[51] consisting of 740 alcohol-dependent male subjects and 861 controls (276 male and 585 female subjects) was used for testing the results of the discovery analyses. Individuals were genotyped using Illumina Human610Quad or Illumina Human660w Quad BeadChips (Illumina Inc). The controls were genotyped using Illumina HumanHap550 Bead Chips.

Test cohort 3 (alcohol dependence, United States) and test cohort 4 (alcohol abuse, United States). The sample consisted of small nuclear families originally collected for linkage studies, and unrelated individuals, Caucasians and African-American, male and female subjects. The subjects were recruited at five US clinical sites: Yale University School of Medicine (APT Foundation; New Haven, CT, USA), the University of Connecticut Health Center (Farmington, CT, USA), the University of Pennsylvania Perelman School of Medicine (Philadelphia, PA, USA), the Medical University of South Carolina (Charleston, SC, USA) and McLean Hospital (Belmont, MA, USA). All subjects were interviewed using the Semi-Structured Assessment for Drug Dependence and Alcoholism to derive diagnoses for lifetime alcohol dependence, alcohol abuse and other major psychiatric traits according to the DSM-IV criteria. There were 1687 male subjects with alcohol dependence, 366 male subjects with alcohol abuse and 475 male controls. There were 1081 female subjects with alcohol dependence, 234 female subjects with alcohol abuse and 786 female controls (Table 1). Individuals were genotyped on the Illumina HumanOmni1-Quad v1.0 microarray (988 306 autosomal SNPs). GWAS genotyping was conducted at the Yale Center for Genome Analysis and the Center for Inherited Disease Research. Genotypes were called using the

Table 1. Discovery and test cohorts

	Alcohol dependence		*Control*
Discovery cohort 1 GWAS Germany			
Male	411		663
Female	0		644
Ethnicity	All Caucasians		All Caucasians
Test cohort 2 Germany			
Male	740		276
Female	0		585
Ethnicity	All Caucasians		All Caucasians
	Alcohol dependence	*Alcohol abuse*	*Control*
Test cohorts 3 and 4, United States			
Male	1687	366	475
Female	1081	234	786
Male ethnicity (Caucasian/African-American)	802/885	201/165	168/307
Female ethnicity (Caucasian/African-American)	471/610	123/111	220/566

Abbreviation: GWAS, genome-wide association study.

GenomeStudio software V2011.1 and genotyping module version 1.8.4 (Illumina Inc).[52]

Gene identification in discovery cohort 1

Quality control. Genotype data had been filtered using stringent quality-control criteria as described earlier[51] and accounted for call rate, population substructure, cryptic relatedness, minor allele frequency and batch effects.

Association test in discovery sample. Association testing was performed using PLINK 1.07 (http://pngu.mgh.harvard.edu/~purcell)[56] software package. A logistic regression modelling approach was applied to correct for population stratification. Therefore, principal component analysis was conducted considering only independent autosomal SNPs with minor allele frequency >0.05 and pairwise $R^2 < 0.05$ within a 200-SNP window. LD filtering resulted in a set of 28 505 SNPs used for principal component analysis, which was carried out using GCTA 1.04 (http://www.complex-traitgenomics.com/software/gcta/).[57] The first two principal components resulting from this analysis were included as covariates in the logistic regression model.

Assignment of SNPs to genes. Genes corresponding to SNPs were identified initially using the annotation file from the Illumina website (http://www.illumina.com, HumanHAP550v3_Gene_Annotation). Next, genes were cross-checked with GeneCards (http://www.genecards.org) to ensure that each gene symbol was current. Any gene symbol that matched to a different gene symbol in Gene Cards was checked to verify chromosome number and location match with the original gene, and was replaced with the current GeneCards gene symbol. SNPs from the original annotation files that had no gene matches in the annotation file and UCSC Genome Browser (that is, not falling within an exon or intron of a known gene) were assumed to regulate and thus implicate the gene closest to the SNP location, using the refSNP database from NCBI (http://www.ncbi.nlm.nih.gov/snp/?SITE=NcbiHome&submit=Go).

Convergent functional genomic analyses

Databases. We have established in our laboratory (Laboratory of Neurophenomics, Indiana University School of Medicine, www.neurophenomics.info) manually curated databases of all the human gene expression (post-mortem brain, blood and cell cultures), human genetic (association, copy number variants (CNVs) and linkage), animal model genetic and animal model gene expression studies published to date on psychiatric disorders. Only the findings deemed significant in the primary publication, by the study authors, using their particular experimental design and thresholds, are included in our databases. Our databases include only primary literature data and do not include review papers or other secondary data integration analyses to avoid redundancy and circularity. These large and constantly updated databases have been used in our CFG cross-validation and prioritization (Figure 1).

Human post-mortem brain, blood and other peripheral tissue gene expression evidence. Information about genes was obtained and imported in our databases searching the primary literature with PubMed (http://ncbi.nlm.nih.gov/PubMed), using various combinations of keywords. For this work, the keywords were as follows: alcohol, alcoholism, human, brain, postmortem, lymphocytes, blood, cells and gene expression.

Human genetic evidence (association, linkage). To designate convergence for a particular gene, the gene had to have independent published evidence of association or linkage for alcoholism. We sought to avoid using any association studies that included subjects who were also included in our discovery or test cohorts. For linkage, the location of each gene was obtained through GeneCards (http://www.genecards.org), and the sex-averaged cM location of the start of the gene was then obtained through http://compgen.rutgers.edu/old/map-interpolator/. For linkage convergence, per our previously published criteria, the start of the gene had to map within 5 cM of the location of a marker linked to the disorder with a lod score of ≥ 2.

Animal model brain and blood gene expression evidence. For animal model brain and blood gene expression evidence, we have used our own rat model data sets,[3] as well as published reports from the literature curated in our databases.

The rat animal model experimental work from our group was previously described.[3] The experimental approaches used to produce the animal model data for CFG analysis were carried out in two rat lines selectively bred for divergent alcohol preference: inbred alcohol-preferring (iP) versus inbred alcohol-non-preferring (iNP) rats. Following five brain regions were chosen for gene expression studies in these rat lines: the frontal cortex, amygdala, caudate–putamen, nucleus accumbens and hippocampus. Animal studies, as well as human imaging and post-mortem analyses, had previously provided evidence that these regions are implicated in alcoholism.

Data for the analysis came from studies of three experimental paradigms. Paradigm 1 examined basal level of gene expression in the brains of the alcohol-naive iP and iNP lines of rats. This basal comparison was performed to determine innate differences between these two lines with a marked divergence in the willingness to consume alcohol. We hypothesized that the innate differences in gene expression between the iP and iNP would involve some of the genes associated with an increased susceptibility for alcohol dependence. Paradigm 2 examined the effects of chronic 24-h free-choice alcohol consumption on gene expression in iP rats compared with alcohol-naive iP rats. This paradigm looked for gene expression changes in the brain associated with the direct influence of peripherally self-administered alcohol in the genetically susceptible rats. In Paradigm 3, iP rats were allowed to self-infuse alcohol directly into the posterior ventral tegmental area, the originating area of the mesolimbic dopamine system. The advantage of this latter procedure is that it isolates the neurocircuitry involved in alcohol reinforcement, and eliminates the peripheral effects of alcohol. Following the establishment of alcohol self-administration into the posterior VTA, gene expression levels in target brain areas were measured and compared with P rats that received artificial cerebral spinal fluid infusions into the posterior VTA.

Animal model genetic evidence. To search for mouse genetic evidence (transgenic and QTL) for our candidate genes, we utilized PubMed as well as the Mouse Genome Informatics (http://www.informatics.jax.org; Jackson Laboratory, Bar Harbor, ME, USA) database, and used the search 'Genes and Markers' form to find transgenic in categories for abnormal alcohol consumption, alcohol preference, alcohol aversion, impaired behavioral response to alcohol, hyperactivity elicited by ethanol administration and enhanced behavioral response to alcohol. For QTL convergence, the start of the gene had to map within 5 cM of the location of these markers.

CFG scoring. We used a nominal *P*-value threshold (having at least one SNP with $P < 0.05$) for including genes from the discovery GWAS in the CFG analysis. No Bonferroni correction was performed.

Internal score: For each of these genes implicated by SNPs, we calculated the percent of SNPs that were nominally significant (ratio of number of nominally significant SNPs over total number of SNPs tested for that gene, multiplied by 100), obtaining a distribution of values. The genes in the top 0.1% of the distribution were given an internal score of 4 points, those in the top 5% of the distribution were given 3 points and the remaining genes all received 2 points. The internal score provides a prioritization of genes based on GWAS results and might prioritize genes that have higher biological relevance and heterogeneity.

External score: Human and animal model data, genetic and gene expression were integrated and tabulated, resulting in a polyevidence CFG score. All six cross-validating lines of evidence (human data and animal model data) were weighted such that evidence from human studies was prioritized 2x over evidence from animal models, gene expression evidence was prioritized 2x over genetic evidence and brain evidence was prioritized 2x over peripheral tissue evidence (Figure 1). For human genetic evidence, 2 points were assigned if it was from association and 1 point if it was from linkage studies. For animal model genetic evidence, 2 points if it was from transgenic and 1 point if it was from QTL. The maximum possible external score for each gene is 12.

We have capped (one positive study scores maximum points) the hypothesis-driven candidate gene genetic association evidence and animal model genetic (transgenic) lines of evidence, regardless of how many other such studies support that gene, to avoid potential 'popularity' biases, where some genes are more studied than others. For discovery-driven gene expression studies, we have capped (one positive study scores maximum points) the human post-mortem brain work because of the paucity of brain collections and the fact that such studies often use the same brain bank sources. However, we have not similarly capped the

animal model brain and blood gene expression evidence, as such studies are not only discovery-based, but use independent cohorts of animals. These were scored differentially, based on the number of studies showing evidence for a given gene: three or more different studies received full maximum points, two studies 0.75 of maximum points and one study 0.5 of the maximum points. Our group generated data sets for three independent animal studies for this analysis (see above).

The more lines of evidence for a gene—that is, the more times a gene shows up as a positive finding across independent studies, platforms, methodologies and species—the higher its external CFG score (Figure 1). This is similar conceptually to the Google PageRank algorithm, in which the more links to a page, the higher it comes up on the search prioritization list. It has not escaped our attention that other ways of weighing the lines of evidence may give slightly different results in terms of prioritization, if not in terms of the list of genes *per se*. Nevertheless, we think this simple scoring system provides a good separation of genes, with specificity provided by human data and sensitivity provided by animal model data.

Prioritizing top alcoholism candidate genes that overlap with a stress-reactive animal model of alcoholism. Stress has been proposed as a driver of alcoholism, notably by Koob and colleagues,[58,59] as well as by Heilig and colleagues.[60] We have previously identified the circadian clock gene *DBP* as a candidate gene for bipolar disorder,[61] as well as for alcoholism,[3] using a CFG approach. In follow-up work, we established mice with a homozygous deletion of DBP (DBP KO) as a stress-reactive genetic animal model of bipolar disorder and alcoholism.[4] We reported that DBP KO mice have lower locomotor activity, blunted responses to stimulants and gain less weight over time. In response to a stress paradigm that translationally mimics what can happen in humans (chronic stress-isolation housing for 4 weeks, with acute stress, on top of that- experimental handling in week 3), the mice exhibit a diametric switch in these phenotypes. DBP KO mice are also activated by sleep deprivation, similar to bipolar patients, and that activation is prevented by treatment with the mood stabilizer drug valproate. Moreover, these mice show increased alcohol intake following exposure to stress. Microarray studies of brain and blood revealed a pattern of gene expression changes that may explain the observed phenotypes. CFG analysis of the gene expression changes identified a series of candidate genes and blood biomarkers for bipolar disorder, alcoholism and stress reactivity. Subsequent studies by us showed that treatment with the omega-3 fatty acid docosahexaenoic acid (DHA) normalized the gene expression (brain and blood) and behavioral phenotypes of this mouse model, including reducing alcohol consumption.[5]

We examined the overlap between the top candidate genes for alcoholism from the current analysis and the top candidate genes from the DBP KO stress mice, thus reducing the list from 135 to 11 (Figure 4).

Pathway analyses. IPA 9.0 (Ingenuity Systems, www.ingenuity.com, Redwood City, CA, USA) was used to analyze the biological roles, including top canonical pathways and diseases, of the candidate genes resulting from our work (Table 2 and Supplementary Table S2), as well as used to identify genes in our data sets that are the targets of existing drugs (Supplementary Table S3). Pathways were identified from the IPA library of canonical pathways that were most significantly associated with genes in our data set. The significance of the association between the data set and the canonical pathway was measured in 2 ways: (1) a ratio of the number of molecules from the data set that map to the pathway divided by the total number of molecules that map to the canonical pathway is displayed; (2) Fisher's exact test was used to calculate a *P*-value determining the probability that the association between the genes in the data set and the canonical pathway is explained by chance alone. We also conducted a KEGG pathway analysis through the Partek Genomic Suites 6.6 software package, Partek Inc, Saint Louis, MO, USA), and GeneGo MetaCore from Thomson Reuters, New York, NY, USA) pathway analyses (https://portal.genego.com/).

Epistasis testing. The test cohort 2 data were used to test for epistatic interactions among the best *P*-value SNPs in the 11 top candidate genes from our work. SNP–SNP allelic epistasis was tested for each distinct pair of SNPs between genes, using the PLINK software package (Supplementary Table S5).

Genetic risk prediction

The software package PLINK 1.07 (http://pngu.mgh.harvard.edu/~purcell)[56] was used to extract individual genotype information for each subject from the test cohorts 2, 3 and 4 data files.

As we had previously performed for bipolar disorder and schizophrenia, we developed a polygenic GRPS for alcoholism based on the presence or absence of the alleles of the SNPs associated with illness in the discovery GWAS cohort 1, and tested the GRPS in three independent cohorts, from different geographic areas, ethnicities and different types of alcoholism. We tested two panels: a larger panel containing all the nominally significant SNPs in top CFG scoring candidate genes ($n = 135$) from the discovery GWAS1 in the top CFG-prioritized genes (Supplementary Tables S1 and S4) and a smaller one ($n = 11$) containing genes out of the larger panel that were cross-validated using an animal model of alcoholism.

Of note, our genes, SNP panels and choice of affected alleles were based solely on analysis of the discovery GWAS1, which is our discovery cohort, completely independently from the test cohorts. Each SNP has two alleles (represented by base letters at that position). One of them is associated with the illness (affected), the other not (non-affected), based on the odds ratios from the discovery GWAS1. We assigned the affected allele a score of 1 and the non-affected allele a score of 0. A two-dimensional matrix of subjects by GRP panel alleles is generated, with the cells populated by 0 or 1. A SNP in a particular individual subject can have any permutation of 1 and 0 (1 and 1, 0 and 1, 0 and 0). By adding these numbers, the minimum score for a SNP in an individual subject is 0, and the maximum score is 2. By adding the scores for all the alleles in the panel, averaging that and multiplying by 100, we generated for each subject an average score corresponding to a genetic loading for disease, which we call Genetic Risk Predictive Score.[53,54]

To test for significance, a one-tailed *t*-test with unequal variance was performed between the alcoholic subjects and the control subjects, looking at differences in GRPS.

Receiver operating characteristic curves. Receiver operating characteristic curves were plotted using IBM SPSS Statistics 21. Diagnosis was converted to a binary call of 0 (control) or 1 (alcohol-dependent or abuser) and entered as the state variable, with calculated GRPS entered as the test variable (Supplementary Figure S2).

RESULTS

Top candidate genes

To minimize false-negatives, we initially cast a wide net, using as a filter a minimal requirement for a gene to have both some GWAS evidence and some additional independent evidence. Thus, out of the 6085 genes with at least a SNP at $P < 0.05$ in the discovery GWAS cohort 1, we generated a list of 3142 genes that also had some additional line of evidence (human or animal model data), implicating them in alcoholism (CFG score ≥ 2.5 (≥ 2 internal) +(≥ 0.5 external)). This suggests, using these minimal thresholds and requirements, that the repertoire of genes potentially involved directly or indirectly in alcohol consumption and alcoholism may be quite large, similar to what we have previously seen for bipolar disorder[62] and schizophrenia.[54] To minimize false-positives, we used an internal score based on percent of SNPs in a gene that were nominally significant, with 4 points for those in the top 0.1% of the distribution ($n = 77$), 3 points for those in the top 5% of the distribution ($n = 561$) and 2 points for the rest of the nominally significant SNPs ($n = 5447$). We then used the CFG analysis and scoring integrating multiple lines of evidence to prioritize this list of genes (Figure 1) and focused our subsequent analyses on only the top CFG scoring candidate genes. Overall, 135 genes had a CFG score of 8 and above ($\geq 50\%$ of maximum possible score of 16).

Of note, there was no correlation between CFG prioritization and gene size, thus excluding a gene-size effect for the observed enrichment (Supplementary Figure S1).

Table 2. Pathway analyses: (A) biological pathways, (B) disease and disorders

	Ingenuity pathways				*KEGG pathways*			*GeneGO pathways*		
(A)	*Numbers*	*Top canonical pathways*	P-*value*	*Ratio*	*Pathway name*	*Enrichment score*	*Enrichment* P-*value*	*Networks*	P-*value*	*Ratio*
GRPS-11, top DBP KO genes out of CFG score ≥ 8.0 genes ($n = 11$ genes)	1	Gαl signaling	4.68E − 05	3/135 (0.022)	Cocaine addiction	11.9589	6.40226E − 06	Neurophysiological process_Transmission of nerve impulse	6.010E − 06	6/212
	2	cAMP-mediated signaling	2.64E − 04	3/226 (0.013)	Gap junction	5.71209	0.00330577	Development_Neurogenesis in general	9.163E − 04	4/192
	3	G-protein-coupled receptor signaling	4.37E − 04	3/276 (0.011)	Glutamatergic synapse	5.21128	0.00545466	Reproduction_GnRH signaling pathway	6.603E − 03	3/166
	4	14-3-3-mediated signaling	2.14E − 03	2/121 (0.017)	Dopaminergic synapse	5.0126	0.00665355	Transport_Synaptic vesicle exocytosis	7.762E − 03	3/176
	5	Synaptic long-term depression	3.33E − 03	2/161 (0.012)	Neuroactive ligand–receptor interaction	3.62169	0.0267375	Development_Neurogenesis_Synaptogenesis	8.258E − 03	3/180

	Ingenuity				*GeneGO*		
(B)	*Number*	*Diseases and disorders*	P-*value*	*Molecules*	*Diseases*	P-*value*	*Ratio*
GRPS-11, top DBP KO genes out of CFG score ≥ 8.0 genes ($n = 11$ genes)	1	Hereditary disorder	1.66E − 08 to 1.29E − 02	9	Schizophrenia	1.476E − 10	11/1033
	2	Neurological disease	1.66E − 08 to 1.64E − 02	10	Suicide	1.076E − 09	6/151
	3	Psychological disorders	1.66E − 08 to 1.53-02	9	Depressive disorder, major	2.106E − 09	10/966
	4	Cancer	1.90E − 08 to 1.82E − 02	10	Bipolar disorder	2.983E − 09	9/706
	5	Skeletal and muscular disorders	1.50E − 07 to 1.34E − 02	6	Alcoholism	3.468E − 09	7/291

Abbreviations: CFG, Convergent Functional Genomics; DBP, D-box-binding protein. Pathway analyses of top candidate genes.

Biological pathways and drug targets

Pathway analyses were carried out on the top candidate genes (Table 2). Notably, Gαi signaling, cocaine addiction and transmission of nerve impulses were the top biological pathways in alcoholism, which may be informative for treatments and drug discovery efforts by pharmaceutical companies. Of note, these top candidate genes were identified and prioritized only for evidence for alcoholism before pathway analyses; therefore, the overlap with cocaine addiction is a completely independent result, suggesting a shared drive and neurobiology. Consistent with that, two of our 135 top candidate genes for alcoholism (CPE and VWF) had SNPs with $P < 10^{-5}$ in a recent GWAS of cocaine addiction.[63]

Some of the top alcohol candidate genes have prior evidence of being modulated by the omega-3 fatty acid DHA in our DBP mouse animal model (Table 3 and Supplementary Table S1). That is of particular interest, as we have previously shown that treatment with the omega-3 fatty acid DHA decreased alcohol consumption in that animal model, as well as in another independent animal model, the alcohol-preferring P rats.[5] Omega-3 fatty acids, particularly DHA, have been described to have alcoholism, mood, psychosis and suicide-modulating properties, in preclinical models as well as some human clinical trials and epidemiological studies. For example, deficits in omega-3 fatty acids have been linked to increased depression and aggression in animal models[64,65] and humans.[66,67] DHA prevents ethanol damage *in vitro* in rat hippocampal slices.[68] Omega-3 supplementation can prevent oxidative damage caused by prenatal alcohol exposure in rats.[69] Of note, deficits in DHA have been reported in erythrocytes[70] and in the post-mortem orbitofrontal cortex of patients with bipolar disorder, and were greater in those that had high versus those that had low alcohol abuse.[71] Low DHA levels may be a risk factor for suicide.[72,73] Omega-3 fatty acids have been reported to be clinically useful in the treatment of both mood[74–77] and psychotic disorders.[78–80]

Other existing pharmacological drugs that modulate alcohol candidate genes identified by us include, besides benzodiazepines, dopaminergic agents, glutamatergic agents, serotonergic agents, as well as statins (Supplementary Table S3).

Genetic risk prediction score

Once the genes involved in a disorder are identified, and prioritized for likelihood of involvement, then an obvious next step is developing a way of applying that knowledge to genetic testing of individuals to determine risk for the disorder. On the basis of our identification of top candidate genes described above using CFG, we pursued a polygenic panel approach, with digitized binary scoring for presence or absence, similar to the one we have devised and used in the past for biomarkers testing[53,81] and for genetic testing in bipolar disorder[53] and schizophrenia.[54] Somewhat similar approaches but without CFG prioritization, attempted by other groups, have been either unsuccessful[82] or have required very large panels of markers.[83]

We chose all the nominally significant *P*-value SNPs ($P < 0.05$) in each of our top CFG-prioritized genes ($n = 135$ with CFG score ≥ 8; Supplementary Table S1) in the GWAS1 data set used for discovery, and assembled a GRPS-135 panel out of those SNPs (Table 4). We then tested the GRPS-135 in the independent German test cohort 2, based on the presence or absence of the alleles of the SNPs associated with the illness, comparing the alcoholic subjects to controls (Table 4), and showed that, although there was a trend, we were not able to distinguish alcoholics from controls in both independent test cohorts.

We then prioritized a smaller panel of 11 genes (Table 3) out of this larger panel, by using as a cross-validator the top genes from a stress-reactive mouse animal model for alcoholism, the DBP knockout mouse[4] (Figure 4). The small panel (GRPS-11) showed more robust results than the larger panel (Table 4), suggesting that it captures the key behaviorally relevant genes.

DISCUSSION

Our CFG approach helped to prioritize a very rich-in-signal and biologically interesting set of genes (Table 3 and Supplementary Table S1). Some, such as SNCA, CPE, DRD2 and GRM3, have weaker evidence based on the GWAS data but strong independent evidence in terms of gene expression studies and other prior human or animal genetic work. Conversely, some of the top previous genetic findings in the field,[84] such as ADH1C[85] (CFG score of 9), GABRA2[86] (CFG score of 8), as well as AUTS2 (CFG score of 7), CHRM2 and KCNJ6 (CFG scores of 4) have fewer different independent lines of evidence, and thus received a lower CFG prioritization score in our analysis (Supplementary Table S1), although they are clearly involved in alcoholism-related processes. Whereas we cannot exclude that more recently discovered genes have had less hypothesis-driven work performed and thus might score lower on CFG, it is to be noted that the CFG approach integrates predominantly non-hypothesis-driven, discovery-type data sets, such as GWAS data, linkage, quantitative traits loci and, particularly, gene expression. We also cap each line of evidence from an experimental approach (Figure 1), to minimize any 'popularity' bias, whereas multiple studies of the same kind are conducted on better-established genes. In the end, it is gene-level reproducibility across multiple approaches and platforms that is built into the approach and gets prioritized most by CFG scoring during the discovery process. Our top results subsequently show good reproducibility and predictive ability in independent cohort testing, the litmus test for any such work.

At the very top of our list of candidate genes for alcoholism, with a CFG score of 13, we have SNCA, a pre-synaptic chaperone that has been reported to be involved in modulating brain plasticity and neurogenesis, as well as neurotransmission, primarily as a brake.[87,88] On the pathological side, low levels of SNCA might offer less protection against oxidative stress,[89] whereas high levels of SNCA may have a role in neurodegenerative diseases, including in Parkinson disease. SNCA has been identified as a susceptibility gene for alcohol cravings[7] and response to alcohol cues.[90] The evidence provided by our data and other previous human genetic association studies suggest a genetic rather than purely environmental (alcohol consumption and stress) basis for its alteration in disease, and its potential utility as trait rather than purely state marker.

Alcoholics carry a genetic variant that leads to reduced baseline expression of *SNCA*.[8] SNCA is also downregulated in expression in the frontal cortex and caudate–putamen of inbred alcohol-preferring rats,[17] as well as in the brain (amygdala) and blood of our stress-reactive DBP animal model of alcoholism, before exposure to any alcohol. SNCA is upregulated in expression in blood in human alcoholism,[12,13] as well as in the blood of monkeys consuming alcohol, and in rats after alcohol administration.[3] Thus, it may serve as a blood biomarker. Overall, we may infer that, whereas low levels of SNCA may predispose to cravings for alcohol and consequent alcoholism, possibly mediated through increased neurobiological activity and drive (*the SNCA deficit hypothesis*), excessive alcohol consumption then increases SNCA expression beyond that seen in non-alcohol-consuming controls, potentially compounding risk for neurodegenerative diseases in individuals that have mutations that lead to its aggregation. This observation is also biologically consistent with the fact that dementia is often observed late in the course of alcohol dependence.

GFAP (glial fibrillary acidic protein), a top candidate gene with a CFG score of 9.5, is an astrocyte intermediate filament-type protein involved in neuron–astrocyte interactions, cell adhesion, process formation and cell–cell communication. It is decreased in

Table 3. Top candidate genes for alcoholism

Gene symbol/name	*Discovery GWAS1 best P-value SNP*	*Internal score nominally significant SNPs/total SNPs tested (% significant)*	*Human genetic evidence*	*Human post-mortem brain expression evidence*	*Human peripheral expression evidence*	*Animal model genetic evidence*	*Animal model brain expression evidence*	*Animal model peripheral (blood) expression evidence*	*CFG score*	*DBP stress*[4]	*DBP stress DHA*[5]
SNCA, synuclein alpha (non A4 component of amyloid precursor)	0.02363 rs17015982	2 4/68 (5.89%)	Association[7,8]	(D) PFC[8–10] (I) FC[11]	(I) Blood[12] (I) Blood[13] (I) Serum[14] DNA hypermethylation[15]	QTL[16]	(D) FC, CP[17] (I) HIP P1 AMY P3[3]	(I) Blood male adult cynomolgus monkeys[18]	13	(D) AMY, blood	(I) Blood
GFAP, glial fibrillary acidic protein	0.01052 rs744281	3 4/12 (33.34%)	Linkage[19]	(D) FC[9,11,20]			(I) FC[21] HIP, NAC, PFC P1[3]		9.5	(I) AMY	
DRD2, dopamine receptor D2	0.03652 rs4938019	2 2/46 (4.35%)	Association[22–26]	(D) FC, CP[27] (D)[28]		(Transgenic) alcohol aversion decreased alcohol consumption[29]			9	(D) PFC	(I) AMY
GRM3, glutamate receptor, metabotropic 3	0.001126 rs41440	2 15/133 (11.28%)	Linkage[30,31]	(D) HIP[32]			(I) NAC[33] (D) FC[34] (D) CP P2[3]		9	(I) AMY	
MBP, myelin basic protein	0.006503 rs1124941	2 16/109 (14.68%)		(D) FC[11]			(I) PFC[35] (D) cerebellum[36] delayed expression[37]	(D) CG-4 glial cells rat brain[38]	8.5	(D) PFC (I) AMY	(D) Blood (I) HIP
MOBP, myelin-associated oligodendrocyte basic protein	0.01231 rs562545	2 2/43 (4.66%)		(D) FC[11]		QTL[39]	(I) PFC[35] (D) NAC[40] (I) whole brain[41] (I) PFC P1[3]		8.5	(D) PFC (I) AMY	(D) Blood (I) HIP
GNAI1, guanine nucleotide binding protein (G protein), alpha-inhibiting activity polypeptide 1	**0.00059** rs12706724	2 11/66 (16.67%)	Linkage[42,43]	(D) HIP[32]			(D) NAC P3[3]		8	(D) PFC	

Table. 3. (Continued)

Gene symbol/name	*Discovery GWAS1 best P-value SNP*	*Internal score nominally significant SNPs/total SNPs tested (% significant)*	*Human genetic evidence*	*Human post-mortem brain expression evidence*	*Human peripheral expression evidence*	*Animal model genetic evidence*	*Animal model brain expression evidence*	*Animal model peripheral (blood) expression evidence*	*CFG score*	*DBP stress*[4]	*DBP stress DHA*[5]
MOG, myelin oligodendrocyte glycoprotein	0.01429 rs3117292	2 3/19 (15.79%)		(D) FC, HIP[11,32]			(D) VTA[44] (I) PFC[35] (I) PFC P1[3]		8	(D) PFC	(I) HIP
RXRG, retinoid X receptor, gamma	0.005815 rs10800098	2 1/37 (2.71%)	Linkage[45]	(D) FC[11]			(I) CP P3[3]		8	(D) PFC	
SYT1, synaptotagmin I	0.04139 rs1245810	2 7/117 (5.99%)		(D) NAC[46]			(D) FC[34] (D) HIP[47]	(I) Cultured neurons[48] (I) Cortical neurons[49]	8	(D) PFC	(D) AMY
TIMP2, TIMP metallopeptidase inhibitor 2	0.04309 rs7502935	2 1/25 (4%)	Linkage[45]	(D) FC, HIP, NAC[32,46,50]			(I) VTA[44]		8	(D) PFC	(D) Blood

Abbreviations: AMY, amygdala; Association, association evidence; CFG, convergent functional genomics; CP, caudate–putamen; D, decreased in expression; DBP, D-box-binding protein; DHA, docosahexaenoic acid; GWAS, genome-wide association study; HIP, hippocampus; I, increased; Linkage, linkage evidence; NAC, nucleus accumbens; P1, paradigm 1; P2, Paradigm 2; P3, paradigm 3 in the Rodd *et al*;[3] PFC, prefrontal cortex; QTL, quantitative trait loci; SNP, single-nucleotide length polymorphism; TG, transgenic; VT, ventral tegmentum; VTA, ventral tegmental area. Top genes with a CFG score of 8 and above that overlapped with top genes from the stress-reactive animal model are shown ($n = 11$; Figure 4). Best *P*-value SNP within the gene or flanking regions is depicted. A more complete list of genes with CFG score of 8 and above ($n = 135$) is available in the Supplementary Information section (Supplementary Table S1). Underlined gene symbol represents means gene is a blood biomarker candidate. Bold *P*-values < 0.001.

Table 4. Genetic Risk Prediction Score (GRPS)-Panels from Discovery Cohort 1

	Test in cohort 2 *alcohol-dependent versus control*
GRPS-135, genes with CFG score of ≥ 8 all nominally significant SNPs in each gene ($n = 713$)	$P = 0.053$ (135 genes, 713 SNPs)
GRPS-11, top animal model (DBP mouse) prioritized genes out of genes with CFG score of ≥ 8 all nominally significant SNPs in each gene ($n = 66$)	$P = 0.041$ (11 genes, 66 SNPs)
	Test in cohort 3 *alcohol-dependent versus control*
GRPS-11, top animal model (DBP mouse) prioritized genes out of genes with CFG score of ≥ 8 all nominally significant SNPs in each gene ($n = 66$)	$P = 0.00012$ (10 genes, 34 SNPs present)
GRPS-SNCA, top CFG gene all nominally significant SNPs in it ($n = 4$)	$P = 0.000013$ (1 gene, 1 SNP *rs17015888* present)
	Test in cohort 4 *alcohol abuse versus control*
GRPS-11, top animal model (DBP mouse) prioritized genes out of genes with CFG score of ≥ 8 all nominally significant SNPs in each gene ($n = 66$)	$P = 0.0094$ (10 genes, 34 SNPs present)
GRPS-SNCA, top CFG gene all nominally significant SNPs in it ($n = 4$)	$P = 0.023$ (1 gene, 1 SNP *rs17015888* present)

Abbreviations: CFG, Convergent Functional Genomics; DBP, DNA-box-binding protein; SNCA, synuclein alpha; SNP, single-nucleotide length polymorphism. Differentiation between alcoholics and controls in three independent test cohorts using, GRPS-135, a panel composed of all the nominally significant SNPs from GWAS1 in the top candidate genes prioritized by CFG; GRPS-11, a panel additionally prioritized by a stress-reactive animal model for alcoholism, the DBP KO-stressed mouse; and GRPS-SNCA, the top candidate gene from our analyses. *P*-values depict one-tailed *t*-test results between alcoholics and controls.

expression in post-mortem brain of alcoholics, but increased in expression in brains of animal models of predisposition to alcoholism, before exposure to alcohol (Table 3). This is consistent with a model for increased physiological robustness in individuals predisposed to alcoholism,[3] as well as with the neurodegenerative consequences of protracted alcohol use.

DRD2 (dopamine receptor D2), another top candidate gene with a CFG score of 9, has prior human genetic association evidence. It is reduced in expression in the frontal cortex in the human brain from alcoholics, as well as in the DBP animal model before any exposure to alcohol. One possible interpretation would be that lower levels of dopamine receptors are associated with reduced dopaminergic signaling and anhedonia, leading individuals to overcompensate by alcohol and drug abuse. Another interpretation, consistent with the low SNCA and consequently higher neurotransmitter (including dopamine) levels, would be that these individuals are in fact in a compulsive, hyperdopaminergic state, which drives them to hedonic activities and leads to compensatory homeostatic downregulation of their DRD2 receptors. Consistent with this later scenario, mice that have a constitutive knockout of their DRD2 receptors, not because of a hyperdopaminergic state, in fact consume less alcohol,[29] unless they are exposed to stress.[91]

Another top candidate gene, *GRM3*, is also involved in neurotransmitter signaling. Prior evidence in the field had implicated another metabotropic glutamate receptor, GRM2.[92]

Other top candidate genes in the panel (*MOBP*, *MBP* and *MOG*) are involved in myelination (Table 3). They are decreased in expression in the prefrontal cortex of human alcoholics, as well as in our stress-reactive DBP animal model of alcoholism, before exposure to any alcohol. Decreased myelination may lead to decreased connectivity. Interestingly, MOBP and MBP are increased in expression in the amygdala in the DBP mice, opposite to the direction of change in the PFC, consistent with a frontal deactivation and a limbic hyperactivity, which could lead to impulsivity.

Epistasis testing of top candidate genes for alcoholism

For the top 11 candidate genes, best *P*-value SNPs from GWAS1 were used to test for gene–gene interactions in GWAS2 (Supplementary Table S5). Nominally significant interactions were found between SNPs in SNCA and RXRG, DRD2 and SYT1, MOBP and TIMP2. As a caveat, the *P*-value was not corrected for multiple comparisons. The corresponding genes merit future follow-up work to elucidate the biological and pathophysiological relevance of their interactions.

Pathways and mechanisms

Our pathway analysis (Table 2 and Supplementary Table S2) results are consistent with the accumulating evidence about the role of neuronal excitability and signaling in alcoholism.[83,93,94]

Overlap with other psychiatric disorders

Despite using lines of evidence for our CFG approach that have to do only with alcoholism, the list of genes identified has a notable overlap at a pathway analysis level (Table 2B and Supplementary Table S2B) and at a gene level (Figures 4 and 5) with other psychiatric disorders. This is a topic of major interest and debate in the field. We demonstrate an overlap between top candidate genes for alcoholism and top candidate genes for schizophrenia, anxiety and bipolar disorder, previously identified by us through CFG (Figure 4), thus providing a possible molecular basis for the frequently observed clinical comorbidity and interdependence between alcoholism and those other major psychiatric disorders, as well as cross-utility of pharmacological agents. Moreover, we tested in alcoholics genetic risk predictive panels for bipolar

Table 5. Individual top genes and genetic risk prediction in independent cohorts

Gene/SNPs	*CFG score*	*Mean GRPS*		t-*test*
		Control (n = *861*)	*Alcohol dependence cohort 2* (n = *740*)	
(A) Test cohort 2				
Panel of 11 top genes				
66 SNPs	≥8	53.98	54.61	*0.041*
SNCA				
rs7668883	13	93.93	92.84	0.086
rs17015888				
rs17015982				
rs6532183				
GFAP				
rs3744473	9.5	63.99	64.69	0.303
rs3169733				
rs736866				
rs744281				
DRD2				
rs4648317	9	13.07	15.51	*0.024*
rs4938019				
GRM3				
rs17160519	9	55.44	54.94	0.271
rs6944937				
rs13236080				
rs17315854				
rs12668989				
rs41440				
rs2373124				
rs13222675				
rs2708553				
rs12673599				
rs4236502				
rs1554888				
rs10499898				
rs1527769				
rs17161018				
MBP				
rs470131	8.5	44.92	47.07	***0.002***
rs2282566				
rs736421				
rs1789094				
rs9951586				
rs1667952				
rs1789105				
rs1789103				
rs1812680				
rs1789139				
rs4890912				
rs9947485				
rs1562771				
rs1015820				
rs1124941				
rs11877526				
MOBP				
rs562545	8.5	49.88	49.93	0.487
rs2233204				
GNAI1				
rs4731111	8	72.97	72.71	0.393
rs6466884				
rs7803811				
rs17802148				
rs7805663				
rs10486920				
rs2523189				
rs2886611				
rs2886609				
rs12706724				
rs4731302				
MOG				
rs3117292	8	34.53	34.56	0.493
rs2747442				
rs3117294				
RXRG				
rs10800098	8	6.04	5.27	0.174
SYT1				
rs1569033	8	39.16	40.89	0.113
rs10735416				
rs1245810				
rs1245819				
rs1268463				

Table. 5. (Continued)

Gene/SNPs	*CFG score*	*Mean GRPS*		t-*test*
		Control (n = *861*)	*Alcohol dependence cohort 2* (n = *740*)	
rs1245840				
rs10861755				
TIMP2				
rs7502935	8	67.65	70.61	*0.038*

Gene/SNPs	*CFG score*	*Mean GRPS*			t-*test*	
		Control (n = *1261*)	*Alcohol dependence cohort 3* (n = *2768*)	*Alcohol abuse cohort 4* (n = *600*)	*Alcohol dependence cohort 3*	*Alcohol abuse cohort 4*
(B) Test cohorts 3 and 4						
Panel of 10 top genes						
34 SNPs	≥8	47.58	48.51	48.49	***0.00012***	*0.0094*
SNCA						
rs17015888	13	72.28	76.96	75.58	***0.000013***	*0.023*
GFAP						
rs3169733	9.5	58.92	60.38	60.17	*0.042*	0.158
rs736866						
DRD2						
rs4648317	9	15.38	14.92	15.61	0.293	0.429
GRM3						
rs17160519	9	35.13	37.38	35.55	***0.000061***	0.309
rs6944937						
rs17315854						
rs4236502						
MBP						
rs470131	8.5	47.23	48.01	48.31	*0.0443*	0.059
rs2282566						
rs736421						
rs1789094						
rs9951586						
rs1789103						
rs4890912						
rs9947485						
rs1124941						
MOBP						
rs562545	8.5	50.28	50.80	50.75	0.233	0.337
rs2233204						
GNAI1						
rs4731111	8	61.58	63.03	62.87	*0.006435*	0.072
rs6466884						
rs17802148						
rs10486920						
rs2523189						
rs2886611						
rs2886609						
MOG						
rs3117292	8	48.78	46.44	46.08	0.020216	0.056
rs2747442						
rs3117294						
RXRG						
rs10800098	8	2.62	3.42	3.33	*0.024279*	0.126
SYT1						
rs1569033	8	41.56	42.48	44.25	0.11474	*0.0087*
rs1245819						
rs1268463						
rs10861755						

Abbreviations: GRPS, Genetic Risk Prediction Score; SNCA, synuclein alpha; SNP, single-nucleotide length polymorphism. Italic, nominally significant; bold italic, survived Bonferroni correction.

disorder[53] and for schizophrenia[54] generated in previous studies by us, and show that they are significantly different in alcoholics versus controls (Figure 6), beyond the overlap in genes with alcohol. There seems to be an increased genetic load for bipolar disorder, consistent with increased drive, and a decreased genetic load for schizophrenia, consistent with increased connectivity before alcohol use. These results led us to develop a heuristic, testable model of alcoholism (Figure 5). Some people may drink to

be calm, mitigating the effects of stress and anxiety, some people may drink to be happy, the common drive with bipolar disorder, and some people may drink to be drunk, to disconnect from reality and/or get unstuck from internal obsessions and ruminations.

Genetic risk prediction

Of note, our SNP panels and choice of affected alleles were based solely on analysis of the discovery GWAS, completely independently from the test cohorts. Our results show that a relatively limited and well-defined panel of SNPs identified based on our CFG analysis could differentiate between alcoholism subjects and controls in three independent cohorts. The fact that our genetic testing worked for both alcohol dependence and alcohol abuse suggests that these two diagnostic categories are actually overlapping, supporting the DSM-V reclassification of a single category of alcohol use disorders.

Reproducibility among studies

Our work provides striking evidence for the advantages, reproducibility and consistency of gene-level analyses of data, as opposed to SNP level analyses, pointing to the fundamental issue of genetic heterogeneity at a SNP level. In fact, it may be that the more biologically important a gene is for higher mental functions, the more heterogeneity it has at a SNP level and the more evolutionary divergence, for adaptive reasons. On top of that, CFG provides a way to prioritize genes based on disease relevance, not study-specific effects (that is, fit-to-disease as opposed to fit-to-cohort). Reproducibility of findings across different studies, experimental paradigms and technical platforms is deemed more important (and scored as such by CFG) than the strength of finding in an individual study (for example, *P*-value in a GWAS).

Potential limitations and confounds

The GWAS study (cohort 1) on which our discovery was based contained males as probands but contained males and females as controls. This was the case for the German test cohort (cohort 2) as well. It is possible that some of the nominally significant SNPs identified in the discovery GWAS have to do with gender differences rather than to alcoholism *per se*, or at least may have to do with male alcoholism. Stratification across gender and ethnicities may also be a factor in our test cohorts 3 and 4 (Table 1). The issue of possible ethnicity differences in alleles, genes and the consequent neurobiology may need to be explored more in the future, with larger sample sizes, and with environmental and cultural factors taken into account. However, the use of a CFG approach using evidence from other studies of alcoholism, including animal model studies, to prioritize the findings decreases the likelihood that our final top results are ethnicity- or gender-related. Of note, our GRPS predictions separate alcoholics from controls in independent test cohorts, in both genders, and in fact work even better at separating female alcoholics from female controls (Figure 3). Moreover, a series of individual genes from the panel, not just SNCA, separates alcoholics from controls in independent cohorts (Table 5).

The conversion from SNPs to genes as part of our discovery assumed the rule of proximity—that is, an intragenic SNP implicates the gene inside which it falls, or if it falls into an intergenic region, it implicates the most proximal gene to it. That may not be true in reality in all cases, generating potentially false-positives and false-negatives. However, the convergent approach and focus on the top CFG scoring genes reduce the likelihood of false-positives.

The only SNP for SNCA that was present/tested for in cohorts 3 and 4 (rs17015888) was relatively far away upstream (0.13 MB) from SNCA. However, no other known genes are present in that region, SNCA is the closest gene, and the distance is well within the range of known examples of regulatory regions (enhancers). In addition, the risk allele for this SNP (G/G) seems to be the major variant in the population (Supplementary Table S6), suggesting that this allele *per se* is evolutionarily advantageous, when not coupled with the exogenous ingestion of alcohol.

A relatively large list of genes ($n = 6085$) was implicated by nominally significant SNPs from the discovery GWAS. There is a risk that out of such a large list CFG will find something to prioritize. We have tried to mitigate that by developing an internal score for each gene based on the proportion of SNPs tested in a gene that were nominally significant. Moreover, in the end, we tested the reproducibility and predictive ability of our top findings in multiple independent cohorts, which is the ultimate litmus test for any genetic or biomarker study.

CONCLUSION

Overall, whereas multiple mechanistic entry points may contribute to alcoholism pathogenesis, it is likely at its core a disease of an exogenous agent (alcohol) modulating different mind domains/dimensions (anxiety, mood and cognition),[95] precipitated by environmental stress on a background of genetic vulnerability (Figure 5). The degree to which various mind domains/dimensions are affected in different individuals is a fertile area for future research into subtypes of alcoholism and lends itself to personalization of diagnosis and treatment, by integrating genetic data, blood gene expression biomarker data and clinical data. Lastly, it is important to note that individuals with a predisposition to alcoholism but no exposure to alcohol may in fact have a robust physiology and strong neurobiological drive that can be harnessed for other, more productive endeavors.

CONFLICT OF INTEREST

The authors declare no conflict of interest directly related to this work. Although not directly relevant to this work, HRK has been a consultant or advisory board member with Alkermes, Lilly, Lundbeck, Pfizer, and Roche. He has also received honoraria from the Alcohol Clinical Trials Initiative (ACTIVE) of the American Society of Clinical Psychopharmacology, which is supported by Lilly, Lundbeck, AbbVie and Pfizer.

ACKNOWLEDGMENTS

This work is, in essence, a field-wide collaboration. We would like to acknowledge our debt of gratitude for the efforts and results of the many other groups, cited in our paper, who have conducted and published studies (genetic and gene expression) in alcoholism. With their arduous and careful work, a convergent approach, such as ours, is possible. We would particularly like to thank the subjects who volunteered to participate in these studies. Without their generous contribution, such work to advance the understanding of alcoholism and help others would not be possible. This work was supported by an NIH Directors' New Innovator Award (1DP2OD007363) and a VA Merit Award (1I01CX000139-01) to ABN, as well as by NIH grants R01 DA12690, R01 DA12849, R01 AA11330, R01 and AA017535 to JG and collaborators, and by grant FKZ 01GS08152 from the National Genome Research Network of the German Federal Ministry of Education and Research to MR and collaborators.

AUTHOR CONTRIBUTIONS

ABN designed the study, with input from JG, MR, NS and AS, and wrote the manuscript. DFL, HL-N, JF and MA analyzed the data. NJ, BK, EW and RL performed database work. HL-N and ZR generated the animal model studies. JF, FK, NW, BM-M, ND, MN, MR and the GESGA consortium generated German GWAS cohorts 1 and 2. RS, LF, AHS, HRK and JG generated US cohorts 3 and 4. All authors discussed the results and commented on the manuscript.

REFERENCES

1 Schuckit MA. Alcohol-use disorders. *Lancet* 2009; **373**: 492–501.

2 Niculescu AB, Le-Niculescu H. Convergent Functional Genomics: what we have learned and can learn about genes, pathways, and mechanisms. *Neuropsychopharmacology* 2010; **35**: 355–356.
3 Rodd ZA, Bertsch BA, Strother WN, Le-Niculescu H, Balaraman Y, Hayden E *et al.* Candidate genes, pathways and mechanisms for alcoholism: an expanded convergent functional genomics approach. *Pharmacogenomics J* 2007; **7**: 222–256.
4 Le-Niculescu H, McFarland MJ, Ogden CA, Balaraman Y, Patel S, Tan J *et al.* Phenomic, convergent functional genomic, and biomarker studies in a stress-reactive genetic animal model of bipolar disorder and co-morbid alcoholism. *Am J Med Genet* 2008; **147B**: 134–166.
5 Le-Niculescu H, Case NJ, Hulvershorn L, Patel SD, Bowker D, Gupta J *et al.* Convergent functional genomic studies of omega-3 fatty acids in stress reactivity, bipolar disorder and alcoholism. *Transl Psychiatry* 2011; **1**: e4.
6 Treutlein J, Cichon S, Ridinger M, Wodarz N, Soyka M, Zill P *et al.* Genome-wide association study of alcohol dependence. *Arch Gen Psychiatry* 2009; **66**: 773–784.
7 Foroud T, Wetherill LF, Liang T, Dick DM, Hesselbrock V, Kramer J *et al.* Association of alcohol craving with alpha-synuclein (SNCA). *Alcohol Clin Exp Res* 2007; **31**: 537–545.
8 Janeczek P, Mackay RK, Lea RA, Dodd PR, Lewohl JM. Reduced expression of alpha-synuclein in alcoholic brain: influence of SNCA-Rep1 genotype. *Addict Biol* 2014; **19**: 509–515.
9 Mayfield RD, Lewohl JM, Dodd PR, Herlihy A, Liu J, Harris RA. Patterns of gene expression are altered in the frontal and motor cortices of human alcoholics. *J Neurochem* 2002; **81**: 802–813.
10 Lewohl JM, Van Dyk DD, Craft GE, Innes DJ, Mayfield RD, Cobon G *et al.* The application of proteomics to the human alcoholic brain. *Ann N Y Acad Sci* 2004; **1025**: 14–26.
11 Lewohl JM, Wang L, Miles MF, Zhang L, Dodd PR, Harris RA. Gene expression in human alcoholism: microarray analysis of frontal cortex. *Alcohol Clin Exp Res* 2000; **24**: 1873–1882.
12 Taraskina AE, Filimonov VA, Kozlovskaya YA, Morozova MN, Gaschin DV, Schwarzman AL. High level of alpha-synuclein mRNA in peripheral lymphocytes of patients with alcohol dependence syndrome. *Bull exp Biol Med* 2008; **146**: 609–611.
13 Bonsch D, Reulbach U, Bayerlein K, Hillemacher T, Kornhuber J, Bleich S. Elevated alpha synuclein mRNA levels are associated with craving in patients with alcoholism. *Biol Psychiatry* 2004; **56**: 984–986.
14 Bonsch D, Greifenberg V, Bayerlein K, Biermann T, Reulbach U, Hillemacher T *et al.* Alpha-synuclein protein levels are increased in alcoholic patients and are linked to craving. *Alcohol Clin Exp Res* 2005; **29**: 763–765.
15 Bonsch D, Lenz B, Kornhuber J, Bleich S. DNA hypermethylation of the alpha synuclein promoter in patients with alcoholism. *Neuroreport* 2005; **16**: 167–170.
16 Liang T, Spence J, Liu L, Strother WN, Chang HW, Ellison JA *et al.* alpha-Synuclein maps to a quantitative trait locus for alcohol preference and is differentially expressed in alcohol-preferring and -nonpreferring rats. *Proc Natl Acad Sci USA* 2003; **100**: 4690–4695.
17 Liang T, Kimpel MW, McClintick JN, Skillman AR, McCall K, Edenberg HJ *et al.* Candidate genes for alcohol preference identified by expression profiling in alcohol-preferring and -nonpreferring reciprocal congenic rats. *Genome Biol* 2010; **11**: R11.
18 Walker SJ, Grant KA. Peripheral blood alpha-synuclein mRNA levels are elevated in cynomolgus monkeys that chronically self-administer ethanol. *Alcohol* 2006; **38**: 1–4.
19 Dick DM, Aliev F, Bierut L, Goate A, Rice J, Hinrichs A *et al.* Linkage analyses of IQ in the collaborative study on the genetics of alcoholism (COGA) sample. *Behav Genet* 2006; **36**: 77–86.
20 Liu J, Lewohl JM, Dodd PR, Randall PK, Harris RA, Mayfield RD. Gene expression profiling of individual cases reveals consistent transcriptional changes in alcoholic human brain. *J Neurochem* 2004; **90**: 1050–1058.
21 Kimpel MW, Strother WN, McClintick JN, Carr LG, Liang T, Edenberg HJ *et al.* Functional gene expression differences between inbred alcohol-preferring and -non-preferring rats in five brain regions. *Alcohol* 2007; **41**: 95–132.
22 Ponce G, Perez-Gonzalez R, Aragues M, Palomo T, Rodriguez-Jimenez R, Jimenez-Arriero MA *et al.* The ANKK1 kinase gene and psychiatric disorders. *Neurotoxic Res* 2009; **16**: 50–59.
23 Dick DM, Wang JC, Plunkett J, Aliev F, Hinrichs A, Bertelsen S*et al.* Family-based association analyses of alcohol dependence phenotypes across DRD2 and neighboring gene ANKK1. *Alcohol Clin Exp Res* 2007; **31**: 1645–1653.
24 Kraschewski A, Reese J, Anghelescu I, Winterer G, Schmidt LG, Gallinat J *et al.* Association of the dopamine D2 receptor gene with alcohol dependence: haplotypes and subgroups of alcoholics as key factors for understanding receptor function. *Pharmacogenet Genomics* 2009; **19**: 513–527.
25 Munafo MR, Matheson IJ, Flint J. Association of the DRD2 gene Taq1A polymorphism and alcoholism: a meta-analysis of case-control studies and evidence of publication bias. *Mol Psychiatry* 2007; **12**: 454–461.
26 Joe KH, Kim DJ, Park BL, Yoon S, Lee HK, Kim TS *et al.* Genetic association of DRD2 polymorphisms with anxiety scores among alcohol-dependent patients. *Biochem Biophys Res Commun* 2008; **371**: 591–595.
27 Noble EP, Blum K, Ritchie T, Montgomery A, Sheridan PJ. Allelic association of the D2 dopamine receptor gene with receptor-binding characteristics in alcoholism. *Arch Gen Psychiatry* 1991; **48**: 648–654.
28 Volkow ND, Wang GJ, Fowler JS, Logan J, Hitzemann R, Ding YS *et al.* Decreases in dopamine receptors but not in dopamine transporters in alcoholics. *Alcohol Clin Exp Res* 1996; **20**: 1594–1598.
29 Phillips TJ, Brown KJ, Burkhart-Kasch S, Wenger CD, Kelly MA, Rubinstein M*et al.* Alcohol preference and sensitivity are markedly reduced in mice lacking dopamine D2 receptors. *Nat Neurosci* 1998; **1**: 610–615.
30 Zhu X, Cooper R, Kan D, Cao G, Wu X. A genome-wide linkage and association study using COGA data. *BMC Genet* 2005; **6**(Suppl 1): S128.
31 Wang S, Huang S, Liu N, Chen L, Oh C, Zhao H. Whole-genome linkage analysis in mapping alcoholism genes using single-nucleotide polymorphisms and microsatellites. *BMC Genet* 2005; **6**(Suppl 1): S28.
32 McClintick JN, Xuei X, Tischfield JA, Goate A, Foroud T, Wetherill L *et al.* Stress-response pathways are altered in the hippocampus of chronic alcoholics. *Alcohol* 2013; **47**: 505–515.
33 Sommer W, Arlinde C, Heilig M. The search for candidate genes of alcoholism: evidence from expression profiling studies. *Addict Biol* 2005; **10**: 71–79.
34 Worst TJ, Tan JC, Robertson DJ, Freeman WM, Hyytia P, Kiianmaa K *et al.* Transcriptome analysis of frontal cortex in alcohol-preferring and nonpreferring rats. *J Neurosci Res* 2005; **80**: 529–538.
35 Tapocik JD, Solomon M, Flanigan M, Meinhardt M, Barbier E, Schank JR*et al.* Coordinated dysregulation of mRNAs and microRNAs in the rat medial prefrontal cortex following a history of alcohol dependence. *Pharmacogenomics J* 2013; **13**: 286–296.
36 Zoeller RT, Butnariu OV, Fletcher DL, Riley EP. Limited postnatal ethanol exposure permanently alters the expression of mRNAS encoding myelin basic protein and myelin-associated glycoprotein in cerebellum. *Alcohol Clin Exp Res* 1994; **18**: 909–916.
37 Chiappelli F, Taylor AN, Espinosa de los Monteros A, de Vellis J. Fetal alcohol delays the developmental expression of myelin basic protein and transferrin in rat primary oligodendrocyte cultures. *Int J Dev Neurosci* 1991; **9**: 67–75.
38 Bichenkov E, Ellingson JS. Ethanol alters the expressions of c-Fos and myelin basic protein in differentiating oligodendrocytes. *Alcohol* 2009; **43**: 627–634.
39 Weng J, Symons MN, Singh SM. Ethanol-responsive genes (Crtam, Zbtb16, and Mobp) located in the alcohol-QTL region of chromosome 9 are associated with alcohol preference in mice. *Alcohol Clin Exp Res* 2009; **33**: 1409–1416.
40 Bell RL, Kimpel MW, McClintick JN, Strother WN, Carr LG, Liang T *et al.* Gene expression changes in the nucleus accumbens of alcohol-preferring rats following chronic ethanol consumption. *Pharmacol Biochem Behav* 2009; **94**: 131–147.
41 Treadwell JA, Singh SM. Microarray analysis of mouse brain gene expression following acute ethanol treatment. *Neurochem Res* 2004; **29**: 357–369.
42 Bergen AW, Yang XR, Bai Y, Beerman MB, Goldstein AM, Goldin LR. Genomic regions linked to alcohol consumption in the Framingham Heart Study. *BMC genet* 2003; **4**(Suppl 1): S101.
43 Foroud T, Edenberg HJ, Goate A, Rice J, Flury L, Koller DL *et al.* Alcoholism susceptibility loci: confirmation studies in a replicate sample and further mapping. *Alcohol Clin Exp Res* 2000; **24**: 933–945.
44 McBride WJ, Kimpel MW, McClintick JN, Ding ZM, Hauser SR, Edenberg HJ *et al.* Changes in gene expression within the ventral tegmental area following repeated excessive binge-like alcohol drinking by alcohol-preferring (P) rats. *Alcohol* 2013; **47**: 367–380.
45 Hill SY, Shen S, Zezza N, Hoffman EK, Perlin M, Allan W. A genome wide search for alcoholism susceptibility genes. *Am J Med Genet* 2004; **128**: 102–113.
46 Flatscher-Bader T, van der Brug M, Hwang JW, Gochee PA, Matsumoto I, Niwa S *et al.* Alcohol-responsive genes in the frontal cortex and nucleus accumbens of human alcoholics. *J Neurochem* 2005; **93**: 359–370.
47 Kabbaj M, Evans S, Watson SJ, Akil H. The search for the neurobiological basis of vulnerability to drug abuse: using microarrays to investigate the role of stress and individual differences. *Neuropharmacology* 2004; **47**(Suppl 1): 111–122.
48 Pignataro L, Miller AN, Ma L, Midha S, Protiva P, Herrera DG *et al.* Alcohol regulates gene expression in neurons via activation of heat shock factor 1. *J Neurosci* 2007; **27**: 12957–12966.
49 Varodayan FP, Pignataro L, Harrison NL. Alcohol induces synaptotagmin 1 expression in neurons via activation of heat shock factor 1. *Neuroscience* 2011; **193**: 63–71.
50 Liu J, Lewohl JM, Harris RA, Iyer VR, Dodd PR, Randall PK *et al.* Patterns of gene expression in the frontal cortex discriminate alcoholic from nonalcoholic individuals. *Neuropsychopharmacology* 2006; **31**: 1574–1582.
51 Frank J, Cichon S, Treutlein J, Ridinger M, Mattheisen M, Hoffmann P *et al.* Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster. *Addict Biol* 2012; **17**: 171–180.

52 Gelernter J, Kranzler HR, Sherva R, Almasy L, Koesterer R, Smith AH *et al.* Genome-wide association study of alcohol dependence:significant findings in African- and European-Americans including novel risk loci. *Mol Psychiatry* 2014; **19**: 41–49.
53 Patel SD, Le-Niculescu H, Koller DL, Green SD, Lahiri DK, McMahon FJ *et al.* Coming to grips with complex disorders: genetic risk prediction in bipolar disorder using panels of genes identified through convergent functional genomics. *Am J Med Genet* 2010; **153B**: 850–877.
54 Ayalew M, Le-Niculescu H, Levey DF, Jain N, Changala B, Patel SD *et al.* Convergent functional genomics of schizophrenia: from comprehensive understanding to genetic risk prediction. *Mol Psychiatry* 2012; **17**: 887–905.
55 Niculescu AB, Le-Niculescu H. The *P*-value illusion: how to improve (psychiatric) genetic studies. *Am J Med Genet* 2010; **153B**: 847–849.
56 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M, Bender D *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Med Genet* 2007; **81**: 559–575.
57 Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Med Genet* 2011; **88**: 76–82.
58 Edwards S, Baynes BB, Carmichael CY, Zamora-Martinez ER, Barrus M, Koob GF *et al.* Traumatic stress reactivity promotes excessive alcohol drinking and alters the balance of prefrontal cortex-amygdala activity. *Transl Psychiatry* 2013; **3**: e296.
59 Koob GF, Buck CL, Cohen A, Edwards S, Park PE, Schlosburg JE *et al.* Addiction as a stress surfeit disorder. *Neuropharmacology* 2014; **76 Pt B**: 370–382.
60 Schank JR, Ryabinin AE, Giardino WJ, Ciccocioppo R, Heilig M. Stress-related neuropeptides and addictive behaviors: beyond the usual suspects. *Neuron* 2012; **76**: 192–208.
61 Niculescu AB 3rd, Segal DS, Kuczenski R, Barrett T, Hauger RL, Kelsoe JR. Identifying a series of candidate genes for mania and psychosis: a convergent functional genomics approach. *Physiol Genomics* 2000; **4**: 83–91.
62 Le-Niculescu H, Patel SD, Bhat M, Kuczenski R, Faraone SV, Tsuang MT *et al.* Convergent functional genomics of genome-wide association data for bipolar disorder: comprehensive identification of candidate genes, pathways and mechanisms. *Am J Med Genet* 2009; **150B**: 155–181.
63 Gelernter J, Sherva R, Koesterer R, Almasy L, Zhao H, Kranzler HR *et al.* Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene. *Mol Psychiatry* 2014; **19**: 41–49.
64 DeMar JC Jr, Ma K, Bell JM, Igarashi M, Greenstein D, Rapoport SI. One generation of n-3 polyunsaturated fatty acid deprivation increases depression and aggression test scores in rats. *J Lipid Res* 2006; **47**: 172–180.
65 Rao JS, Ertley RN, Lee HJ, DeMar JC Jr, Arnold JT, Rapoport SI *et al.* n-3 polyunsaturated fatty acid deprivation in rats decreases frontal cortex BDNF via a p38 MAPK-dependent mechanism. *Mol Psychiatry* 2007; **12**: 36–46.
66 Zanarini MC, Frankenburg FR. omega-3 Fatty acid treatment of women with borderline personality disorder: a double-blind, placebo-controlled pilot study. *Am J Psychiatry* 2003; **160**: 167–169.
67 Lin PY, Huang SY, Su KP. A meta-analytic review of polyunsaturated fatty acid compositions in patients with depression. *Biol Psychiatry* 2010; **68**: 140–147.
68 Collins MA, Moon KH, Tajuddin N, Neafsey EJ, Kim HY. Docosahexaenoic acid (DHA) prevents binge ethanol-dependent aquaporin-4 elevations while inhibiting neurodegeneration: experiments in rat adult-age entorhino-hippocampal slice cultures. *Neurotox Res* 2013; **23**: 105–110.
69 Patten AR, Brocardo PS, Christie BR. Omega-3 supplementation can restore glutathione levels and prevent oxidative damage caused by prenatal ethanol exposure. *J Nutr Biochem* 2012; **24**: 760–769.
70 McNamara RK. DHA deficiency and prefrontal cortex neuropathology in recurrent affective disorders. *J Nutr* 2010; **140**: 864–868.
71 McNamara RK, Jandacek R, Rider T, Tso P, Stanford KE, Hahn CG *et al.* Deficits in docosahexaenoic acid and associated elevations in the metabolism of arachidonic acid and saturated fatty acids in the postmortem orbitofrontal cortex of patients with bipolar disorder. *Psychiatry Res* 2008; **160**: 285–299.
72 Sublette ME, Hibbeln JR, Galfalvy H, Oquendo MA, Mann JJ. Omega-3 polyunsaturated essential fatty acid status as a predictor of future suicide risk. *Am J Psychiatry* 2006; **163**: 1100–1102.
73 Lewis MD, Hibbeln JR, Johnson JE, Lin YH, Hyun DY, Loewke JD. Suicide deaths of active-duty US military and omega-3 fatty-acid status: a case-control comparison. *J Clin Psychiatry* 2011; **72**: 1585–1590.
74 Stoll AL, Severus WE, Freeman MP, Rueter S, Zboyan HA, Diamond E *et al.* Omega 3 fatty acids in bipolar disorder: a preliminary double-blind, placebo-controlled trial. *Arch Gen Psychiatry* 1999; **56**: 407–412.
75 Parker G, Gibson NA, Brotchie H, Heruc G, Rees AM, Hadzi-Pavlovic D. Omega-3 Fatty acids and mood disorders. *Am J Psychiatry* 2006; **163**: 969–978.
76 Osher Y, Belmaker RH. Omega-3 fatty acids in depression: a review of three studies. *CNS Neurosci Ther* 2009; **15**: 128–133.
77 Clayton EH, Hanstock TL, Hirneth SJ, Kable CJ, Garg ML, Hazell PL. Reduced mania and depression in juvenile bipolar disorder associated with long-chain omega-3 polyunsaturated fatty acid supplementation. *Eur J Clin Nutr* 2009; **63**: 1037–1040.
78 Peet M, Stokes C. Omega-3 fatty acids in the treatment of psychiatric disorders. *Drugs* 2005; **65**: 1051–1059.
79 Berger GE, Wood SJ, Wellard RM, Proffitt TM, McConchie M, Amminger GP *et al.* Ethyl-eicosapentaenoic acid in first-episode psychosis. A 1H-MRS study. *Neuropsychopharmacology* 2008; **33**: 2467–2473.
80 Amminger GP, Schafer MR, Papageorgiou K, Klier CM, Cotton SM, Harrigan SM *et al.* Long-chain omega-3 fatty acids for indicated prevention of psychotic disorders: a randomized, placebo-controlled trial. *Arch Gen Psychiatry* 2010; **67**: 146–154.
81 Kurian SM, Le-Niculescu H, Patel SD, Bertram D, Davis J, Dike C *et al.* Identification of blood biomarkers for psychosis using convergent functional genomics. *Mol Psychiatry* 2009; **16**: 37–58.
82 Yan J, Aliev F, Webb BT, Kendler KS, Williamson VS, Edenberg HJ *et al.* Using genetic information from candidate gene and genome-wide association studies in risk prediction for alcohol dependence. *Addict Biol* advance online publication, 30 January 2013; doi:10.1111/adb.12035 (e-pub ahead of print).
83 Kos MZ, Yan J, Dick DM, Agrawal A, Bucholz KK, Rice JP *et al.* Common biological networks underlie genetic risk for alcoholism in African- and European-American populations. *Genes Brain Behavior* 2013; **12**: 532–542.
84 Edenberg HJ, Foroud T. Genetics and alcoholism. *Nat Rev Gastroenterol Hepatol* 2013; **10**: 487–494.
85 Biernacka JM, Geske JR, Schneekloth TD, Frye MA, Cunningham JM, Choi DS *et al.* Replication of genome wide association studies of alcohol dependence: support for association with variation in ADH1C. *PLoS ONE* 2013; **8**: e58798.
86 Villafuerte S, Strumba V, Stoltenberg SF, Zucker RA, Burmeister M. Impulsiveness mediates the association between GABRA2 SNPs and lifetime alcohol problems. *Genes Brain Behavior* 2013; **12**: 525–531.
87 Winner B, Regensburger M, Schreglmann S, Boyer L, Prots I, Rockenstein E *et al.* Role of alpha-synuclein in adult neurogenesis and neuronal maturation in the dentate gyrus. *J Neurosci* 2012; **32**: 16906–16916.
88 Lundblad M, Decressac M, Mattsson B, Bjorklund A. Impaired neurotransmission caused by overexpression of alpha-synuclein in nigral dopamine neurons. *Proc Natl Acad Sci USA* 2012; **109**: 3213–3219.
89 Musgrove RE, King AE, Dickson TC. alpha-Synuclein protects neurons from apoptosis downstream of free-radical production through modulation of the MAPK signalling pathway. *Neurotox Res* 2013; **23**: 358–369.
90 Wilcox CE, Claus ED, Blaine SK, Morgan M, Hutchison KE. Genetic variation in the alpha synuclein gene (SNCA) is associated with BOLD response to alcohol cues. *J Stud Alcohol Drugs* 2013; **74**: 233–244.
91 Delis F, Thanos PK, Rombola C, Rosko L, Grandy D, Wang GJ *et al.* Chronic mild stress increases alcohol intake in mice with low dopamine D2 receptor levels. *Behav Neurosci* 2013; **127**: 95–105.
92 Zhou Z, Karlsson C, Liang T, Xiong W, Kimura M, Tapocik JD*et al.* Loss of metabotropic glutamate receptor 2 escalates alcohol consumption. *Proc Natl Acad Sci USA* 2013; **110**: 16963–16968.
93 Heilig M, Goldman D, Berrettini W, O'Brien CP. Pharmacogenetic approaches to the treatment of alcohol addiction. *Nat Rev Neurosci* 2011; **12**: 670–684.
94 Zhao Z, Guo AY, van den Oord EJ, Aliev F, Jia P, Edenberg HJ *et al.* Multi-species data integration and gene ranking enrich significant results in an alcoholism genome-wide association study. *BMC Genomics* 2012; **13**(Suppl 8): S16.
95 Niculescu AB 3rd, Schork NJ, Salomon DR. Mindscape: a convergent perspective on life, mind, consciousness and happiness. *J Affect Disord* 2010; **123**: 1–8.

Supplementary Information accompanies the paper on the Translational Psychiatry website (http://www.nature.com/tp)

Supplementary Information

Table S1.Top candidate genes for alcoholism - CFG analysis of GWAS data. Top genes with a CFG score of 8 and above (n= 135) are shown. I - increased; D – decreased in expression. I –increased; D – decreased. PFC- prefrontal cortex. AMY-amygdala. Gene symbols underlined are blood biomarker candidate genes (20 out of 135). Best p-value SNP within the gene or flanking regions are depicted. **P-values in bold are <0.001.**

Gene Symbol	GWAS 1 Best p-values SNP	P-Value	Internal Score	Human Genetic Evidence	Human Brain Expression Evidence	Human Peripheral Expression Evidence	Animal Model Genetic Evidence	Animal Model Brain Expression Evidence	Animal Model Peripheral Expression Evidence	Total CFG Score		DBP Stress	DBP Stress DHA
SNCA	rs17015982	0.02363	2	2	4	2	0.5	2	0.5	13		(D)DBP Stress AMY (D)DBP Stress Blood	(I)DBP-DHA Blood
CPE	rs17688688	0.03277	2	2	4	2	0	1.5	0	11.5			(D)DBP-DHA Blood
FCGRT	rs2335534	0.04521	3	0	4	2	0.5	1.5	0	11			
INSIG1	rs3923644	0.03586	2	1	4	2	0	1.5	0	10.5			
SPARC	rs6861486	0.02757	2	1	4	2	0	1.5	0	10.5			(D)DBP-DHA Blood
ACSL3	rs6436346	0.01857	2	1	4	2	0	1	0	10			
GABRG2	rs17060763	0.009148	2	2	4	0	0	2	0	10			
SYT11	rs822519	0.007779	4	0	4	0	0.5	1.5	0	10			
CXCL12	rs17156360	0.005029	2	0	4	2	0	1.5	0	9.5			
GABRB3	rs8027455	0.002829	2	2	4	0	0.5	1	0	9.5		(D)DBP Stress PFC ; (D)DBP Stress AMY	(D)DBP-DHA HIP
GFAP	rs744281	0.01052	3	1	4	0	0	1.5	0	9.5		(I)DBP Stress AMY	
NRXN3	rs10873325	0.001212	2	2	4	0	0	1.5	0	9.5			
PDYN	rs9679771	0.03805	2	2	4	0	0	1.5	0	9.5			
PSMA1	rs2575850	0.02751	2	0	4	2	0	1.5	0	9.5			
PTPRE	rs12415045	0.00429	2	0	4	2	0	1.5	0	9.5			
SCD	rs7068970	0.01188	2	0	4	2	0	1.5	0	9.5			(I)DBP-DHA HIP; (I)DBP-DHA PFC ; (I)DBP-DHA HIP
ADH1C	rs2173201	**0.000618**	3	2	4	0	0	0	0	9			
CAP2	rs16879781	0.02187	2	1	4	0	0	2	0	9			
DDX5	rs2075552	0.01493	3	2	4	0	0	0	0	9			
DRD2	rs4938019	0.03652	2	2	4	0	1	0	0	9		(D)DBP Stress PFC	(I)DBP-DHA AMY
DST	rs7760542	0.02887	2	1	4	2	0	0	0	9			(I)DBP-DHA NAC
FN1	rs6713637	0.002112	2	1	4	0	0	2	0	9			

GRM3	rs41440	0.001126	2	1	4	0	0	2	0	9		(I)DBP Stress AMY	
HMGCR	rs3761740	0.03334	2	1	4	0	0	2	0	9			
KDR	rs2305948	**0.000462**	2	1	4	0	0	2	0	9			
MAPT	rs4792891	0.009256	2	1	4	0	0	2	0	9			(I)DBP-DHA HIP
MPDZ	rs10514823	0.001443	2	2	4	0	0	1	0	9			(D)DBP-DHA HIP
NFIB	rs2382456	**0.00072**	2	0	4	2	0	1	0	9		(D)DBP Stress AMY	(D)DBP-DHA Blood
NTRK2	rs11140688	0.009416	2	1	4	0	0	2	0	9			
QDPR	rs2697705	0.02419	2	1	4	0	0	2	0	9			
SDC4	rs2267869	0.0212	2	1	4	0	0	2	0	9		(D)DBP Stress PFC ; (D)DBP Stress AMY	(D)DBP-DHA PFC; (D)DBP-DHA NAC
SOX9	rs8068789	0.001764	2	0	4	2	0	1	0	9			
STX1A	rs6951030	0.02611	3	0	4	0	0	2	0	9			
SYN2	rs17669026	0.03978	2	1	4	0	0	2	0	9			
TFAP2B	rs2635727	0.01621	2	2	4	0	0	1	0	9			
USP32	rs8079220	0.02085	2	1	4	2	0	0	0	9			
VLDLR	rs10967188	0.002602	2	2	4	0	0	1	0	9		(I)DBP Stress AMY	
YWHAZ	rs3134369	**0.000535**	2	0	4	0	1	2	0	9			(D)DBP-DHA Blood
ADAM10	rs12439189	0.007005	2	1	4	0	0	1.5	0	8.5			
ANXA2	rs1630535	0.0109	2	1	4	0	0	1.5	0	8.5			
ARIH1	rs1160608	0.01122	3	1	4	0	0.5	0	0	8.5			(I)DBP-DHA Blood
CAST	rs27772	0.00013	2	1	4	0	0	1.5	0	8.5			
CERS2	rs267738	0.02953	4	0	4	0	0.5	0	0	8.5			
CITED2	rs9389724	0.009349	2	1	4	0	0.5	1	0	8.5			
ENO1	rs6682376	0.04055	2	0	4	0	0.5	2	0	8.5			
ENPP2	rs7000665	0.007509	2	0	4	0	0.5	2	0	8.5			
FOXG1	rs1885147	0.01082	2	0	4	2	0.5	0	0	8.5			
IGF1	rs7306935	0.007657	2	1	4	0	0	1	0.5	8.5			
KLK6	rs1654537	0.007661	3	0	4	0	0.5	1	0	8.5			
MAP1B	rs1217785	0.01705	2	1	4	0	0	1.5	0	8.5		(D)DBP Stress PFC ; (I)DBP Stress AMY	(D)DBP-DHA HIP
MBP	rs1124941	0.006503	2	0	4	0	0	2	0.5	8.5		(D)DBP Stress PFC ; (I)DBP Stress AMY	(I)DBP-DHA HIP; (D)DBP-DHA Blood
MOBP	rs562545	0.01231	2	0	4	0	0	2	0	8.5			
MPZL2	rs2187557	0.02351	3	1	4	0	0.5	0	0	8.5			
MYO1B	rs16833762	0.02792	2	1	4	0	0	1.5	0	8.5			
SCG2	rs13402129	0.01329	2	1	4	0	0	1.5	0	8.5			

SESTD1	rs1968618	0.007605	2	2	4	0	0.5	0	0	8.5			
TCF12	rs10518890	0.03998	2	1	4	0	0	1.5	0	8.5			
VEZF1	rs9904523	0.03904	2	1	4	0	0	1.5	0	8.5			
ACACA	rs2542663	0.02582	2	1	4	0	0	1	0	8			
ACO1	rs3936927	**4.97E-05**	2	1	4	0	0	1	0	8			
ACOT12	rs4703516	0.04173	2	0	4	2	0	0	0	8			
AGMO	rs6962150	0.005631	2	0	4	2	0	0	0	8			
AGT	rs4762	0.01166	2	0	4	0	0	2	0	8			(I) DBP-DHA AMY
ANLN	rs11765557	0.01028	2	0	4	2	0	0	0	8			
ANXA5	rs11735972	0.02185	2	0	4	0	0	2	0	8			
ATAD5	rs3764421	0.02827	3	1	4	0	0	0	0	8			
ATXN1	rs1473730	**0.000537**	2	1	4	0	0	1	0	8		(D)DBP Stress PFC	(D)DBP-DHA Blood
B4GALT2	rs869896	0.04123	4	0	4	0	0	0	0	8			
CCNA1	rs1517897	**0.000617**	3	1	4	0	0	0	0	8			
CD74	rs2071368	**0.000424**	3	1	4	0	0	0	0	8			
CEBPD	rs13266791	0.03991	2	0	4	0	0.5	1.5	0	8			(D)DBP-DHA Blood
CLK1	rs7224	0.01583	3	1	4	0	0	0	0	8			
CNTNAP2	rs2538971	**0.00082**	2	2	4	0	0	0	0	8			
COTL1	rs1469479	0.004344	2	0	4	2	0	0	0	8			(D)DBP-DHA Blood (D)DBP-DHA HIP
CSNK1A1	rs7712431	0.005067	2	1	4	0	0	1	0	8			(I)DBP-DHA Blood
CUX2	rs3809277	0.00291	2	2	4	0	0	0	0	8			
CYB5R3	rs8190423	0.006456	2	0	4	0	0	2	0	8		(D)DBP Stress PFC	
CYTH4	rs9610713	0.01097	2	0	4	2	0	0	0	8			
DLC1	rs10503435	0.02245	2	1	4	0	0	1	0	8			
DPYSL2	rs17055687	0.03831	2	2	4	0	0	0	0	8			
ELL2	rs3815768	0.007039	3	1	4	0	0	0	0	8			
FAM3C	rs998057	0.04555	2	0	4	2	0	0	0	8			
FARP2	rs1476698	0.01722	2	2	4	0	0	0	0	8			
FSD1L	rs12335518	0.009382	2	2	4	0	0	0	0	8			
GABRA2	rs1025852	0.0307	2	2	4	0	0	0	0	8			
GABRB2	rs10037137	0.04398	2	2	4	0	0	0	0	8			
GBAP1	rs2049805	0.0468	4	0	4	0	0	0	0	8			
GNAI1	rs12706724	**0.00059**	2	1	4	0	0	1	0	8		(D)DBP Stress PFC	
GNG12	rs17531147	0.02658	2	1	4	0	0	1	0	8			(D)DBP-DHA Blood
GOT2	rs13339064	0.02053	2	0	4	0	0	2	0	8		(D)DBP Stress PFC	
GRHPR	rs309453	0.0319	2	1	4	0	0	1	0	8			

GRIA1	rs4530817	0.04338	2	0	4	0	0	2	0	8			
H2AFV	rs3801403	0.03082	3	1	4	0	0	0	0	8			
HNRNPA2B1	rs12672536	0.02132	3	0	4	0	0	1	0	8			
HSPH1	rs932715	0.01153	2	1	4	0	0	1	0	8			
HTR1B	rs2320160	0.001425	2	2	4	0	0	0	0	8			
HTR2A	rs1923885	0.003812	2	2	4	0	0	0	0	8			
IFITM3	rs11246074	0.03415	2	1	4	0	0	1	0	8			
IGF1R	rs3743264	0.00708	2	0	4	0	0	2	0	8			
IST1	rs4788449	0.01156	3	1	4	0	0	0	0	8			
JMJD8	rs6597	0.003844	4	0	4	0	0	0	0	8			
KDM4C	rs12553351	**0.000602**	2	2	4	0	0	0	0	8			
MAN2A1	rs2416227	0.01905	2	1	4	0	0	1	0	8			
MAT2B	rs1433779	0.005647	2	0	4	2	0	0	0	8			
METAP2	rs282323	**0.000198**	3	1	4	0	0	0	0	8			(I)DBP-DHA Blood
MKLN1	rs3800678	0.01405	2	2	4	0	0	0	0	8			
MOG	rs3117292	0.01429	2	0	4	0	0	2	0	8		(D)DBP Stress PFC	(I)DBP-DHA HIP
MYL2	rs933296	0.01066	2	2	4	0	0	0	0	8			
MYO1D	rs225184	**0.000219**	2	1	4	0	0	1	0	8			
NOL11	rs4496198	0.009119	2	0	4	2	0	0	0	8			
NSMAF	rs4737508	0.04528	2	0	4	2	0	0	0	8			
NTM	rs3099787	0.008411	2	1	4	0	0	1	0	8		(D)DBP Stress PFC ; (I)DBP Stress AMY	
PDK4	rs854061	0.03023	2	1	4	0	0	1	0	8			
PFDN6	rs456261	0.01101	4	0	4	0	0	0	0	8			
PGM2L1	rs3193507	0.01831	2	2	4	0	0	0	0	8			(D)DBP-DHA Blood
PLD1	rs3850188	0.001101	2	0	4	0	0	2	0	8			
PLLP	rs16969211	0.03152	2	0	4	0	0	2	0	8			
PPM1B	rs4952703	0.03325	2	0	4	0	0	2	0	8			
PPP1R12C	rs575144	0.03198	2	0	4	2	0	0	0	8			
PRKCZ	rs908742	0.04567	2	0	4	0	0.5	1.5	0	8			
PSMD13	rs6598055	0.02309	2	1	4	0	0	1	0	8			
RFTN1	rs9990208	0.009928	2	0	4	2	0	0	0	8			
RHOB	rs342056	**0.000614**	2	0	4	0	0.5	1.5	0	8			(D)DBP-DHA Blood
RXRG	rs10800098	0.005815	2	1	4	0	0	1	0	8		(D)DBP Stress PFC	
SETD1A	rs897986	0.01892	4	0	4	0	0	0	0	8			
SH2B3	rs10774623	0.01612	3	0	4	0	0	1	0	8			(D)DBP-DHA Blood
SLC11A1	rs2290708	0.005329	3	1	4	0	0	0	0	8			(D)DBP-DHA Blood

SLC12A2	rs36693	0.00193	2	0	4	0	0	2	0	8			(I)DBP-DHA HIP
SYT1	rs1245810	0.04139	2	0	4	0	0	1.5	0.75	8		(D)DBP Stress PFC	(D) DBP-DHA AMY
TIMP2	rs7502935	0.04309	2	1	4	0	0	1	0	8		(D)DBP Stress PFC	(D)DBP-DHA Blood
TIMP3	rs762886	0.01733	2	1	4	0	0	1	0	8			
TMED4	rs217361	0.01461	3	1	4	0	0	0	0	8			
TMEM109	rs555835	0.03234	4	0	4	0	0	0	0	8		(I)DBP Stress PFC	(D)DBP-DHA AMY
URI1	rs11671255	0.01216	2	0	4	2	0	0	0	8			
VWF	rs1063856	0.01716	2	1	4	0	0	1	0	8			

Table S2. Pathway Analyses of all top candidate genes for alcoholism CFG 8 and up (n=135) A. Biological Pathways. B. Disease and Disorders.

A.	Ingenuity Pathways				KEGG Pathways			GeneGO Pathways		
	#	Top Canonical Pathways	P-Value	Ratio	Pathway Name	Enrichment Score	Enrichment p-value	Networks	pValue	Ratio
GRPS-135 CFG score >=8.0 genes (N=135 genes)	1	Hepatic Fibrosis / Hepatic Stellate Cell Activation	3.78E-05	7/155 (0.045)	Nicotine addiction	10.3005	3.36E-05	Cell adhesion_Synaptic contact	1.927E-06	14/184
	2	RhoA Signaling	1.43E-04	6/123 (0.049)	Morphine addiction	8.97218	0.000127	**Neurophysiological process_Transmission of nerve impulse**	2.021E-06	15/212
	3	Signaling by Rho Family GTPases	1.57E-04	8/263 (0.03)	Retrograde endocannabinoid signaling	8.9528	0.000129	Development_Neurogenesis_Synaptogenesis	1.907E-04	11/180
	4	p70S6K Signaling	1.87E-04	6/132 (0.045)	GABAergic synapse	8.48991	0.000206	Reproduction_Progesterone signaling	2.081E-04	12/213
	5	IL-8 Signaling	2.75E-04	7/225 (0.031)	**Cocaine addiction**	6.55773	0.001419	Development_Regulation of angiogenesis	3.182E-04	12/223

B.	Ingenuity				GeneGO		
	#	Diseases and Disorders	P-Value	Molecules	Diseases	P-Value	Ratio
GRPS-135 CFG score >=8.0 genes (N=135 genes)	1	Hereditary Disorder	1.39E-17 - 3.92E-04	44	Substance-Related Disorders	1.531E-28	22/847
	2	Neurological Disease	1.39E-17 - 5.77E-04	72	Alcoholism	4.572E-28	34/291
	3	Psychological Disorders	1.39E-17 - 5.77E-04	54	Bipolar Disorder	9.395E-28	47/706
	4	Skeletal and Muscular Disorders	2.56E-15 - 4.60E-04	48	Depressive Disorder, Major	8.210E-23	48/966
	5	Metabolic Disease	1.17E-13 - 1.19E-04	41	Diabetus Mellitus, Type 2	1.038E-22	59/1529

Table S3. Ingenuity drug targets analysis. Repositioning of existing drugs to be potentially tested for treating alcoholism. Genes with CFG score of 8 and above (n= 135)

Gene Symbol/Gene Name	CFG Score	Location	Type(s)	Drug(s)
GABRG2 gamma-aminobutyric acid (GABA) A receptor, gamma 2	**10**	Plasma Membrane	ion channel	pagoclone,alphadolone,SEP174559,clobazam,nitrazepam,tracazolate,adinazolam,sevoflurane,isoflurane,gaboxadol,isoniazid,felbamate,etomidate,muscimol,halothane,estazolam,clorazepate,eszopiclone,quazepam,diazepam,temazepam,zolpidem,chlordiazepoxide,lorazepam,olanzapine,triazolam,flumazenil,clonazepam,flurazepam,midazolam,oxazepam,alprazolam ,pregnenolone
GABRB3 gamma-aminobutyric acid (GABA) A receptor, beta 3	**9.5**	Plasma Membrane	ion channel	pagoclone,ethchlorvynol,alphadolone,SEP174559,clobazam,nitrazepam,adinazolam,piperazine,acetaminophen/butalbital/caffeine,sevoflurane,isoflurane,gaboxadol,isoniazid,felbamate,etomidate,muscimol,halothane,fluoxetine/olanzapine,amobarbital,estazolam,atropine/hyoscyamine/phenobarbital/scopolamine,clorazepate,acetaminophen/butalbital,eszopiclone,quazepam,mephobarbital,hyoscyamine/phenobarbital,acetaminophen/butalbital/caffeine/codeine,butabarbital,diazepam,temazepam,zolpidem,chlordiazepoxide,lorazepam,olanzapine,triazolam,clonazepam,flurazepam,midazolam,oxazepam,alprazolam,zaleplon,secobarbital,butalbital,phenobarbital,pentobarbital,thiopental,propofol,ezogabine,desflurane,methoxyflurane,enflurane,pregnenolone
ADH1C alcohol dehydrogenase 1C (class I), gamma polypeptide	**9**	Cytoplasm	enzyme	fomepizole,ethanol
DRD2 dopamine receptor D2	**9**	Plasma Membrane	G-protein coupled receptor	paliperidone,risperidone,buspirone,bifeprunox,iloperidone,blonanserin,asenapine,pardoprunox,ocaperidone,abaperidone,methotrimeprazine,fluspirilene,SLV314,cariprazine,rotigotine,acetophenazine,sultopride,zuclopenthixol,thioproperazine,lurasidone,opipramol, pipothiazine,chloropromazine,domperidone,metoclopramide,sulpiride,meloxicam,amantadine,flupenthixol,chlorprothixene,trifluoperazine,fluphenazine,pimozide,clozapine,haloperidol ,fluphenazine decanoate,thiothixene, quetiapine,pramipexol,olanzapine,lisuride,sertindole,cabergoline,ziprasidone,mesoridazine,thioridazine,aripiprazole,ropinirole,dihydroergocryptine,dihydroergotamine,bromocriptine,apomorphine,pergolide,dopamine,droperidol,thiethylperazine,droperidol/fentanyl,L-dopa
GRM3 glutamate receptor, metabotropic 3	**9**	Plasma Membrane	G-protein coupled receptor	fasoracetam
HMGCR 3-hydroxy-3-methylglutaryl-CoA reductase	**9**	Cytoplasm	enzyme	aspirin/pravastatin,atorvastatin/ezetimibe,simvastatin/sitagliptin,pitavastatin,lovastatin/niacin,ezetimibe/simvastatin,amlodipine/atorvastatin,fluvastatin,cerivastatin,atorvastatin,pravastatin,simvastatin,lovastatin,rosuvastatin
KDR kinase insert domain receptor (a type III receptor tyrosine kinase)	**9**	Plasma Membrane	kinase	AEE 788,sunitinib,cediranib,pazopanib,axitinib,XL647,CEP 7055,BMS-582664,CHIR-265,tivozanib,OSI930,telatinib,cabozantinib,regorafenib,vatalanib,sorafenib,vandetanib
NTRK2 neurotrophic tyrosine kinase, receptor, type 2	**9**	Plasma Membrane	kinase	cabozantinib

GABRA2 gamma-aminobutyric acid (GABA) A receptor, alpha 2	**8**	Plasma Membrane	ion channel	methohexital,primidone,meprobamate,aspirin/butalbital/caffeine,aspirin/butalbital/caffeine/codeine,hexobarbital,pagoclone,ethchlorvynol,alphadolone,SEP 174559,zopiclone,clobazam,nitrazepam,adinazolam,butobarbital,acetaminophen/butalbital/caffeine,sevoflurane,isoflurane,gaboxadol,felbamate,etomidate,muscimol,halothane,fluoxetine/olanzapine,amobarbital,estazolam,atropine/hyoscyamine/phenobarbital/scopolamine,clorazepate,acetaminophen/butalbital,eszopiclone,quazepam,mephobarbital,hyoscyamine/phenobarbital,amitriptyline/chlordiazepoxide,acetaminophen/butalbital/caffeine/codeine,butabarbital,diazepam,temazepam,zolpidem,chlordiazepoxide,lorazepam,olanzapine,triazolam,clonazepam,flurazepam,midazolam,flunitrazepam,oxazepam,alprazolam,zaleplon,thiamylal,secobarbital,barbital,butalbital,phenobarbital,pentobarbital,thiopental,ezogabine,desflurane,methoxyflurane,enflurane,pregnenolone
GABRB2 gamma-aminobutyric acid (GABA) A receptor, beta 2	**8**	Plasma Membrane	ion channel	methohexital,aspirin/butalbital/caffeine,aspirin/butalbital/caffeine/codeine,fospropofol,pagoclone,ethchlorvynol,alphadolone,SEP 174559,clobazam,nitrazepam,adinazolam,acetaminophen/butalbital/caffeine,sevoflurane,isoflurane,gaboxadol,isoniazid,felbamate,etomidate,muscimol,halothane,fluoxetine/olanzapine,amobarbital,estazolam,atropine/hyoscyamine/phenobarbital/scopolamine,clorazepate,acetaminophen/butalbital,eszopiclone,mephobarbital,hyoscyamine/phenobarbital,acetaminophen/butalbital/caffeine/codeine,butabarbital,diazepam,temazepam,zolpidem,chlordiazepoxide,lorazepam,olanzapine,triazolam,clonazepam,flurazepam,midazolam,oxazepam,alprazolam,zaleplon,secobarbital,butalbital,phenobarbital,pentobarbital,thiopental,propofol,ezogabine,desflurane,methoxyflurane,enflurane,pregnenolone
GRIA1 glutamate receptor, ionotropic, AMPA 1	**8**	Plasma Membrane	ion channel	talampanel,Org24448,LY451395,tezampanel,perampanel,sevoflurane,isoflurane,\desflurane,methoxyflurane,enflurane
HTR1B 5-hydroxytryptamine (serotonin) receptor 1B, G protein-coupled	**8**	Plasma Membrane	G-protein coupled receptor	caffeine/ergotamine,asenapine,donitriptan,sumatriptan,eletriptan,frovatriptan,almotriptan,fenfluramine,dihydroergotamine,ergotamine,naratriptan,zolmitriptan,rizatriptan
HTR2A 5-hydroxytryptamine (serotonin) receptor 2A, G protein-coupled	**8**	Plasma Membrane	G-protein coupled receptor	paliperidone,risperidone,buspirone,caffeine/ergotamine,iloperidone,eplivanserin,blonanserin,flibanserin,asenapine,ocaperidone,abaperidone,psilocybine,APD125,thioproperazine,lurasidone,cinitapride,pipothiazine,chloropromazine,trazodone,flupenthixol,cisapride,chlorprothixene,clozapine,doxepin,desipramine,cyproheptadine,clomipramine,fluoxetine/olanzapine,thiothixene,loxapine,epinastine,fenfluramine,quetiapine,olanzapine,nefazodone,mirtazapine,amitriptyline,cyclobenzaprine,nortriptyline,lisuride,sertindole,ziprasidone,mesoridazine,thioridazine,aripiprazole,methysergide,dihydroergotamine,apomorphine,ergotamine,azatadine
IGF1R insulin-like growth factor 1 receptor	**8**	Plasma Membrane	transmembrane receptor	OSI-906,cixutumumab,ganitumab,AVE1642,BIIB022,IGF1
METAP2 methionyl aminopeptidase 2	**8**	Cytoplasm	peptidase	PPI-2458
PRKCZ protein kinase C, zeta	**8**	Cytoplasm	kinase	ingenol 3-angelate
RXRG retinoid X receptor, gamma	**8**	Nucleus	ligand-dependent nuclear receptor	etretinate,bexarotene,adapalene,acitretin,tretinoin,9-cis-retinoic acid
SLC12A2 solute carrier family 12 (sodium/potassium/chloride transporters), member 2	**8**	Plasma Membrane	transporter	bumetanide,quinethazone

Table S4: GRPS-135 large panel used for genetic risk prediction of alcoholism (n=135 genes).

Gene Symbol/Gene Name	HumanHap550v3_Gene_AnnotationSNP	Gene	Location	A1 (Cohort 1)	A2 (Cohort 1)	OR (Cohort 1)	P (Cohort 1)	A1 (Cohort 2)	A2 (Cohort 2)	A1 (Cohort 3,4)	A2 (Cohort 3,4)
ACACA acetyl-CoA carboxylase alpha	rs732770	NM_198838.1	flanking_3UTR	A	G	0.8171	0.02853	A	G	A	G
	rs9906543	NM_198837.1	intron	C	T	0.8384	0.04879	C	T	A	G
	rs2305097	NM_198837.1	intron	G	T	0.7933	0.04048	G	T		
	rs2542663	NM_198838.1	intron	C	T	0.7781	0.02582	C	T		
	rs2898659	NM_198838.1	intron	T	C	0.7874	0.03417	T	C		
	rs9330248	NM_198837.1	intron	A	G	0.7631	0.03123	A	G		
ACO1 aconitase 1, soluble	rs1033216	NM_002197.1	flanking_5UTR	T	G	1.234	0.02069	T	G	C	A
	rs10757998	NM_002197.1	flanking_5UTR	G	A	1.406	0.00929	G	A	G	A
	rs10813415	NM_002197.1	flanking_5UTR	G	A	1.68	0.000117	G	A	G	A
	rs10813420	NM_002197.1	flanking_5UTR	T	C	1.284	0.004285	T	C	A	G
	rs10969854	NM_002197.1	flanking_5UTR	T	C	1.194	0.02782	C	T	G	A
	rs12552052	NM_002197.1	flanking_5UTR	T	C	1.235	0.01145	T	C	G	A
	rs13302787	NM_002197.1	flanking_5UTR	T	G	0.7463	0.02173	T	G	A	C
	rs2888998	NM_002197.1	flanking_5UTR	G	A	1.407	0.005323	G	A	G	A
	rs2992117	NM_002197.1	flanking_5UTR	C	T	1.336	0.03935	C	T	G	A
	rs3936927	NM_002197.1	flanking_5UTR	T	C	1.688	4.97E-05	T	C	A	G
	rs7027645	NM_002197.1	flanking_5UTR	T	C	1.229	0.0487	T	C	A	G
	rs7035270	NM_002197.1	flanking_5UTR	G	A	0.7626	0.02332	G	A	G	A
	rs10970046	NM_002197.1	flanking_5UTR	G	A	1.33	0.001177	G	A		
	rs13285154	NM_002197.1	flanking_5UTR	T	C	0.666	0.02638	T	C		
	rs13295330	NM_002197.1	flanking_5UTR	C	T	0.6879	0.04114	C	T		
	rs1335137	NM_002197.1	flanking_5UTR	T	C	1.518	0.000433	T	C		
	rs17776734	NM_002197.1	flanking_5UTR	A	C	1.188	0.04649	A	C		
	rs2375104	NM_002197.1	flanking_5UTR	A	G	1.252	0.006665	G	A		
	rs7021230	NM_002197.1	flanking_5UTR	A	G	1.222	0.02708	A	G		
	rs7038348	NM_002197.1	flanking_5UTR	C	T	1.242	0.007778	C	T		
ACOT12 acyl-CoA thioesterase 12	rs4703516	NM_130767.1	intron	T	G	0.7476	0.04173	T	G	A	C
ACSL3 acyl-CoA synthetase long-chain family member 3	rs6436346	NM_004457.3	flanking_5UTR	G	A	0.8066	0.01857	G	A	G	A
	rs2395926	NM_004457.3	flanking_5UTR	A	C	0.7846	0.03489	A	C		
ADAM10 ADAM metallopeptidase domain 10	rs12439189	NM_001110.2	intron	C	A	1.308	0.007005	C	A	C	A
	rs1869135	NM_001110.2	flanking_3UTR	C	T	1.239	0.0305	C	T	G	A
	rs2081703	NM_001110.2	intron	G	A	1.237	0.04393	G	A	G	A
	rs4238331	NM_001110.2	intron	T	G	1.243	0.01568	T	G	A	C

ADH1C alcohol dehydrogenase 1C (class I), gamma polypeptide	rs11499823	NM_000669.3	flanking_5UTR	C	T	1.348	0.03112	C	T	G	A
	rs1789891	NM_000669.3	flanking_3UTR	A	C	1.377	0.00327	A	C	A	C
	rs1789924	NM_000669.3	flanking_5UTR	T	C	1.269	0.00476	T	C	A	G
	rs2173201	NM_000669.3	flanking_3UTR	A	C	0.6908	0.000618	A	C	A	C
	rs2851300	NM_000669.3	flanking_5UTR	T	C	1.276	0.003949	T	C	A	G
AGMO alkylglycerol monooxygenase	rs2191994	NM_001004320.1	flanking_3UTR	C	T	1.396	0.01513	C	T	G	A
	rs6962150	NM_001004320.1	flanking_3UTR	T	G	1.324	0.005631	T	G	A	C
	rs10243870	NM_001004320.1	flanking_3UTR	G	T	0.7897	0.04595	G	T		
	rs10274730	NM_001004320.1	flanking_3UTR	T	C	0.8234	0.02214	T	C		
	rs4422690	NM_001004320.1	flanking_3UTR	C	T	0.7909	0.04612	C	T		
	rs4599703	NM_001004320.1	intron	A	G	0.7904	0.04643	A	G		
AGT angiotensinogen (serpin peptidase inhibitor, clade A, member 8)	rs2478545	NM_000029.2	intron	T	C	1.252	0.01919	T	C	A	G
	rs2493137	NM_000029.2	flanking_5UTR	C	T	1.187	0.04627	C	T	A	G
	rs4762	NM_000029.2	coding	T	C	1.34	0.01166	T	C	A	G
	rs7536290	NM_000029.2	flanking_3UTR	G	A	1.263	0.02717	G	A	G	A
ANLN anillin, actin binding protein	rs11765557	NM_018685.2	flanking_3UTR	T	G	0.739	0.01028	T	G	A	C
	rs3801317	NM_018685.2	intron	T	C	1.295	0.04953	T	C		
	rs6954831	NM_018685.2	intron	A	G	1.532	0.01362	A	G		
ANXA2 annexin A2	rs12148854	NM_001002858.1	flanking_3UTR	G	A	1.258	0.04113	G	A	G	A
	rs1630535	NM_001002858.1	flanking_3UTR	A	G	1.338	0.0109	A	G	A	G
	rs11735972	NM_001154.2	flanking_3UTR	C	T	1.299	0.02185	C	T	G	A
ANXA5 annexin A5	rs1158024	NM_001154.2	flanking_3UTR	T	C	1.19	0.04948	T	C		
ARIH1 ariadne RBR E3 ubiquitin protein ligase 1	rs1562250	NM_005744.2	intron	T	C	1.376	0.02367	T	C	A	G
	rs17753089	NM_005744.2	flanking_3UTR	T	C	0.8062	0.01767	T	C	A	G
	rs17825866	NM_005744.2	flanking_3UTR	C	T	0.5841	0.04745	C	T	G	A
	rs1160608	NM_005744.2	intron	A	G	1.431	0.01122	A	G		
	rs2278694	NM_005744.2	intron	A	G	1.36	0.03474	A	G		
ATAD5 ATPase family, AAA domain containing 5	rs3816780	NM_024857.3	coding	T	C	1.33	0.03496	T	C	A	G
	rs3764421	NM_024857.3	coding	C	A	1.353	0.02827	C	A		
ATXN1 ataxin 1	rs1150639	NM_000332.2	intron	A	G	0.6902	0.002488	A	G	A	G
	rs12204969	NM_000332.2	intron	C	T	1.408	0.006857	C	T	G	A
	rs1473730	NM_000332.2	flanking_5UTR	G	A	1.558	0.000537	G	A	G	A
	rs16879127	NM_000332.2	flanking_5UTR	A	C	0.7814	0.007641	A	C	A	C
	rs179944	NM_000332.2	intron	A	G	1.506	0.004517	A	G	A	G
	rs3812199	NM_000332.2	intron	C	T	0.81	0.01391	C	T	G	A
	rs3812200	NM_000332.2	intron	T	C	1.236	0.0157	T	C	A	G
	rs3812206	NM_000332.2	intron	T	C	0.8291	0.04534	T	C	A	G
	rs4074661	NM_000332.2	flanking_5UTR	A	G	0.7534	0.001507	A	G	A	G
	rs6903850	NM_000332.2	intron	T	G	1.359	0.003059	T	G	A	C
	rs697739	NM_000332.2	intron	A	G	0.8375	0.04964	A	G	G	A

	rs9396669	NM_000332.2	intron	A	G	1.394	0.01705	A	G	A	G
	rs9396726	NM_000332.2	flanking_5UTR	T	G	1.245	0.04965	T	G	A	C
	rs1144699	NM_000332.2	intron	G	A	0.7146	0.002979	G	A		
	rs12196135	NM_000332.2	intron	A	G	1.415	0.02666	A	G		
	rs17595094	NM_000332.2	intron	A	G	0.6719	0.01244	A	G		
	rs17606174	NM_000332.2	intron	T	C	1.305	0.02656	T	C		
	rs2744413	NM_000332.2	intron	C	T	1.958	0.009804	C	T		
	rs3793118	NM_000332.2	intron	A	G	0.6505	0.001042	A	G		
	rs9350018	NM_000332.2	flanking_5UTR	C	A	0.75	0.002317	C	A		
B4GALT2 UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 2	rs869896	NM_001005417.1	intron	T	C	1.418	0.04123	T	C		
CAP2 CAP, adenylate cyclase-associated protein, 2 (yeast)	rs16879781	NM_006366.2	intron	T	C	1.405	0.02187	T	C	A	G
	rs13194542	NM_006366.2	intron	T	C	1.187	0.04138	T	C		
CAST calpastatin	rs10053056	NM_173060.1	intron	T	C	1.256	0.0057	T	C	G	A
	rs13362120	NM_001750.4	intron	C	T	0.7182	0.000324	C	T	G	A
	rs155051	NM_001750.4	flanking_5UTR	C	T	0.8386	0.03831	C	T	G	A
	rs155053	NM_001750.4	flanking_5UTR	A	G	1.193	0.03298	A	G	A	G
	rs25862	NM_173061.1	intron	A	G	0.7465	0.000433	A	G	A	G
	rs261227	NM_001750.4	flanking_5UTR	G	T	1.196	0.03497	G	T	C	A
	rs26505	NM_173061.1	coding	T	C	1.274	0.02555	T	C	A	G
	rs27991	NM_173060.1	intron	A	G	1.275	0.01524	A	G	A	G
	rs4400148	NM_001750.4	flanking_5UTR	T	C	0.7745	0.002181	T	C	A	G
	rs10037212	NM_173061.1	intron	A	C	0.7197	0.00031	A	C		
	rs155039	NM_001750.4	flanking_5UTR	T	C	0.8186	0.01787	T	C		
	rs155040	NM_001750.4	flanking_5UTR	A	C	0.8374	0.04581	A	C		
	rs17476752	NM_001750.4	flanking_5UTR	G	T	1.32	0.01299	G	T		
	rs1862609	NM_001750.4	intron	A	G	0.7115	0.000193	A	G		
	rs27772	NM_173061.1	intron	G	A	1.406	0.00013	G	A		
	rs469532	NM_173060.1	intron	C	A	1.397	0.000808	C	A		
CCNA1 cyclin A1	rs1517881	NM_003914.2	flanking_3UTR	A	G	0.7518	0.000797	A	G	A	G
	rs1517896	NM_003914.2	flanking_3UTR	T	C	0.7738	0.002482	T	C	A	G
	rs1517897	NM_003914.2	flanking_3UTR	T	C	0.7488	0.000617	T	C	A	G
	rs6563485	NM_003914.2	flanking_3UTR	C	T	0.7961	0.01624	C	T	G	A
	rs9315422	NM_003914.2	flanking_3UTR	T	G	0.7992	0.01593	T	G	A	C
	rs943725	NM_003914.2	flanking_3UTR	T	C	0.8045	0.01949	T	C	A	G
	rs9576081	NM_003914.2	flanking_3UTR	A	G	1.221	0.03475	A	G	A	G
	rs9603055	NM_003914.2	flanking_3UTR	T	C	0.739	0.00189	T	C	A	G
	rs4391925	NM_003914.2	flanking_3UTR	C	A	0.7889	0.01098	C	A		
	rs7330345	NM_003914.2	flanking_3UTR	C	T	1.263	0.004912	C	T		
CD74 CD74 molecule,	rs13175409	NM_004355.2	flanking_5UTR	T	C	1.218	0.02556	T	C	A	G

major histocompatibility complex, class II invariant chain	rs2071368	NM_004355.2	flanking_5UTR	C	T	0.5014	0.000424	C	T	G	A
	rs2288817	NM_004355.2	intron	G	A	0.6692	0.01734	G	A	G	A
CEBPD CCAAT/enhancer binding protein (C/EBP), delta	rs12541086	NM_005195.2	flanking_3UTR	T	C	0.8331	0.04242	T	C	G	A
	rs13266791	NM_005195.2	flanking_3UTR	C	T	1.44	0.03991	C	T		
CERS2 ceramide synthase 2	rs267738	NM_181746.1	coding	C	A	0.7938	0.02953	C	A	C	A
CITED2 Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2	rs2080112	NM_006079.3	flanking_5UTR	T	C	1.206	0.03502	T	C	A	G
	rs3010311	NM_006079.3	flanking_5UTR	G	A	1.196	0.03638	G	A	G	A
	rs4896491	NM_006079.3	flanking_5UTR	T	C	1.52	0.02062	T	C	A	G
	rs6903961	NM_006079.3	flanking_5UTR	A	G	1.246	0.03076	A	G	A	G
	rs6913402	NM_006079.3	flanking_3UTR	T	G	1.185	0.04384	T	G	C	A
	rs6925308	NM_006079.3	flanking_3UTR	A	G	0.7524	0.02677	A	G	A	G
	rs887780	NM_006079.3	flanking_5UTR	G	A	1.205	0.04075	G	A	G	A
	rs9389724	NM_006079.3	flanking_5UTR	A	G	1.262	0.009349	A	G	A	G
	rs9495538	NM_006079.3	flanking_5UTR	C	T	1.191	0.03641	C	T	A	G
	rs12525047	NM_006079.3	flanking_5UTR	T	C	0.7854	0.02464	T	C		
	rs13204837	NM_006079.3	flanking_5UTR	A	G	0.7854	0.02464	A	G		
	rs1587171	NM_006079.3	flanking_5UTR	C	T	0.8129	0.04821	C	T		
	rs17069288	NM_006079.3	flanking_5UTR	T	G	1.312	0.01671	T	G		
	rs7759402	NM_006079.3	flanking_5UTR	A	G	1.179	0.04878	A	G		
CLK1 CDC-like kinase 1	rs11903236	NM_001024646.1	intron	C	T	1.343	0.0251	C	T	G	A
	rs7224	NM_004071.2	3UTR	C	A	1.413	0.01583	C	A		
CNTNAP2 contactin associated protein-like 2	rs10238572	NM_014141.3	intron	C	T	0.8393	0.03688	C	T	A	G
	rs10243377	NM_014141.3	intron	T	C	0.7018	0.03093	T	C	A	G
	rs10250570	NM_014141.3	intron	A	G	1.212	0.01834	A	G	A	G
	rs10251377	NM_014141.3	intron	G	A	1.241	0.02137	G	A	G	A
	rs10268483	NM_014141.3	intron	G	A	1.244	0.03362	G	A	G	A
	rs10276770	NM_014141.3	intron	T	C	1.249	0.03233	T	C	A	G
	rs10279409	NM_014141.3	intron	A	G	0.7819	0.04541	A	G	A	G
	rs12112963	NM_014141.3	intron	C	A	1.296	0.02161	C	A	C	A
	rs12531630	NM_014141.3	intron	A	G	0.7384	0.02622	A	G	A	G
	rs1404732	NM_014141.3	intron	C	T	1.508	0.0102	C	T	G	A
	rs1881726	NM_014141.3	intron	C	T	0.8147	0.03531	C	T	G	A
	rs2011815	NM_014141.3	intron	G	T	1.33	0.007536	G	T	C	A
	rs2204465	NM_014141.3	flanking_5UTR	C	T	1.231	0.0166	C	T	G	A
	rs2533096	NM_014141.3	intron	G	A	0.6446	0.02503	G	A	G	A
	rs2538971	NM_014141.3	intron	C	T	0.6712	0.00082	C	T	G	A
	rs4421277	NM_014141.3	intron	A	G	0.7802	0.02921	A	G	A	G
	rs6943628	NM_014141.3	intron	A	G	1.433	0.00644	A	G	A	G
	rs6970519	NM_014141.3	flanking_3UTR	T	G	1.744	0.007191	T	G	A	C
	rs700303	NM_014141.3	intron	A	G	0.7948	0.02234	A	G	A	G

	rs7785839	NM_014141.3	intron	T	C	1.519	0.008994	T	C	A	G
	rs7790550	NM_014141.3	intron	T	C	1.25	0.02599	T	C	A	G
	rs7808274	NM_014141.3	intron	A	G	1.226	0.01209	A	G	G	A
	rs826810	NM_014141.3	intron	A	G	0.8236	0.03838	A	G	A	G
	rs10272638	NM_014141.3	intron	A	G	1.217	0.04147	A	G		
	rs11762071	NM_014141.3	intron	G	A	1.221	0.01416	G	A		
	rs1177007	NM_014141.3	intron	G	A	0.8014	0.0141	G	A		
	rs12538810	NM_014141.3	flanking_3UTR	A	C	1.742	0.007308	A	C		
	rs13225365	NM_014141.3	intron	A	G	1.409	0.02894	A	G		
	rs1351312	NM_014141.3	intron	T	C	0.7076	0.003337	T	C		
	rs1517770	NM_014141.3	flanking_5UTR	T	C	0.8073	0.02458	T	C		
	rs1534702	NM_014141.3	intron	T	C	0.8058	0.02598	T	C		
	rs1718101	NM_014141.3	intron	A	G	1.673	0.001685	A	G		
	rs2074711	NM_014141.3	intron	G	A	0.7835	0.03526	G	A		
	rs7779225	NM_014141.3	intron	G	A	1.256	0.02287	G	A		
	rs7784277	NM_014141.3	intron	T	G	1.279	0.01752	T	G		
	rs7805359	NM_014141.3	intron	C	T	0.8381	0.03182	C	T		
	rs7806237	NM_014141.3	intron	A	C	1.497	0.003595	A	C		
	rs826823	NM_014141.3	intron	C	T	0.641	0.02532	C	T		
	rs851715	NM_014141.3	intron	G	A	0.8255	0.03185	G	A		
	rs851826	NM_014141.3	intron	C	T	0.8379	0.04741	C	T		
COTL1 coactosin-like 1 (Dictyostelium)	rs1429264	NM_021149.2	intron	C	T	0.7973	0.01439	C	T		
	rs1469479	NM_021149.2	intron	G	T	1.341	0.004344	G	T		
	rs7188531	NM_021149.2	intron	T	C	0.823	0.03581	T	C		
CPE carboxypeptidase E	rs17688688	NM_001873.1	flanking_3UTR	A	G	1.443	0.03277	A	G	A	G
	rs4691213	NM_001873.1	flanking_3UTR	A	G	1.206	0.03841	A	G		
CSNK1A1 casein kinase 1, alpha 1	rs7712431	NM_001025105.1	intron	C	T	0.7117	0.005067	C	T	G	A
	rs6580605	NM_001025105.1	flanking_3UTR	C	T	0.6479	0.04637	C	T		
	rs9325142	NM_001025105.1	flanking_3UTR	A	G	1.437	0.0141	A	G		
CUX2 cut-like homeobox 2	rs7300860	NM_015267.1	intron	C	T	1.254	0.02541	C	T	G	A
	rs3809277	NM_015267.1	intron	A	C	1.493	0.00291	A	C		
	rs3809278	NM_015267.1	intron	T	G	1.281	0.03329	T	G		
CXCL12 chemokine (C-X-C motif) ligand 12	rs1002588	NM_000609.4	flanking_5UTR	T	C	1.204	0.042	T	C	A	G
	rs17156360	NM_000609.4	flanking_5UTR	T	C	1.387	0.005029	T	C	A	G
	rs17156715	NM_000609.4	flanking_5UTR	A	C	0.725	0.04005	A	C	A	C
	rs17406863	NM_000609.4	flanking_5UTR	C	T	1.309	0.01828	C	T	G	A
	rs2297630	NM_001033886.1	intron	A	G	0.8034	0.02376	A	G	A	G
	rs11591534	NM_000609.4	flanking_5UTR	A	G	1.279	0.03067	A	G		
	rs12412154	NM_000609.4	flanking_5UTR	T	C	0.7984	0.03291	T	C		
	rs7075674	NM_000609.4	flanking_5UTR	T	C	1.203	0.04371	T	C		

	rs994179	NM_000609.4	flanking_5UTR	C	T	0.7659	0.01034	C	T		
CYB5R3 cytochrome b5 reductase 3	rs2285141	NM_007326.2	intron	T	G	0.6722	0.006635	T	G		
	rs8190423	NM_007326.2	intron	T	C	0.6713	0.006456	T	C		
CYTH4 cytohesin 4	rs4821628	NM_013385.2	flanking_3UTR	G	A	1.18	0.04695	G	A	G	A
	rs9610713	NM_013385.2	flanking_3UTR	C	T	0.8062	0.01097	C	T	G	A
DDX5 DEAD (Asp-Glu-Ala-Asp) box helicase 5	rs2075552	NM_004396.2	intron	A	G	1.433	0.01493	A	G		
DLC1 deleted in liver cancer 1	rs10503435	NM_006094.3	intron	C	T	1.686	0.02245	C	T	G	A
DPYSL2 dihydropyrimidinase-like 2	rs11775967	NM_001386.4	intron	A	C	1.239	0.03957	A	C	A	C
	rs11780026	NM_001386.4	intron	G	A	1.236	0.04146	G	A		
	rs17055687	NM_001386.4	flanking_3UTR	T	G	1.488	0.03831	T	G		
	rs17088251	NM_001386.4	flanking_3UTR	A	G	1.488	0.03831	A	G		
DRD2 dopamine receptor D2	rs4648317	NM_016574.2	intron	T	C	1.258	0.04409	T	C	A	G
	rs4938019	NM_016574.2	intron	C	T	1.268	0.03652	C	T		
DST dystonin	rs1024196	NM_020388.2	coding	G	A	0.5529	0.03412	G	A	G	A
	rs11966986	NM_183380.1	intron	T	C	0.5529	0.03412	T	C		
	rs7760542	NM_183380.1	intron	G	A	0.5434	0.02887	G	A		
ELL2 elongation factor, RNA polymerase II, 2	rs147295	NM_012081.3	flanking_3UTR	T	C	1.199	0.03721	T	C	G	A
	rs3777179	NM_012081.3	intron	T	C	1.266	0.01821	T	C	A	G
	rs3777185	NM_012081.3	intron	A	G	1.272	0.009015	A	G	A	G
	rs7705279	NM_012081.3	flanking_5UTR	A	G	1.256	0.02091	A	G	A	G
	rs7717348	NM_012081.3	flanking_5UTR	T	G	1.197	0.0496	T	G	A	C
	rs918629	NM_012081.3	flanking_5UTR	A	G	1.24	0.0239	A	G	A	G
	rs10035477	NM_012081.3	intron	C	T	1.249	0.01618	C	T		
	rs3815768	NM_012081.3	coding	A	G	1.282	0.007039	A	G		
	rs4538631	NM_012081.3	flanking_5UTR	T	C	1.206	0.04001	T	C		
	rs6556898	NM_012081.3	flanking_5UTR	G	A	1.224	0.03632	G	A		
ENO1 enolase 1, (alpha)	rs6682376	NM_001428.2	flanking_5UTR	C	T	1.192	0.04055	C	T	A	G
	rs12124851	NM_001428.2	flanking_3UTR	G	A	1.194	0.04723	G	A		
ENPP2 ectonucleotide pyrophosphatase/phosphodiesterase 2	rs1058913	NM_006209.2	coding	T	C	1.214	0.04919	T	C	A	G
	rs1992721	NM_006209.2	intron	G	A	1.22	0.02332	G	A	G	A
	rs2289887	NM_006209.2	intron	A	G	1.57	0.008782	A	G	A	G
	rs6993464	NM_006209.2	flanking_3UTR	T	C	1.203	0.02435	T	C	A	G
	rs10505364	NM_006209.2	flanking_3UTR	G	A	1.218	0.01604	G	A		
	rs11782176	NM_006209.2	flanking_3UTR	G	A	1.197	0.02883	G	A		
	rs4520196	NM_006209.2	intron	G	A	1.175	0.05	G	A		
	rs7000665	NM_006209.2	flanking_3UTR	C	T	1.245	0.007509	C	T		
	rs7846200	NM_006209.2	flanking_3UTR	A	G	1.205	0.02289	A	G		
FAM3C family with sequence similarity 3, member C	rs2697177	NM_014888.1	flanking_5UTR	G	A	1.611	0.04748	G	A	G	A
	rs998057	NM_014888.1	flanking_5UTR	A	G	0.683	0.04555	A	G		

FARP2 FERM, RhoGEF and pleckstrin domain protein 2	rs1476698	NM_014808.1	intron	C	T	1.223	0.01722	C	T	G	A
	rs3771570	NM_014808.1	intron	T	C	1.264	0.03552	T	C		
FCGRT Fc fragment of IgG, receptor, transporter, alpha	rs2335534	NM_004107.3	flanking_5UTR	A	G	0.8057	0.04521	A	G	A	G
FN1 fibronectin 1	rs1250085	NM_212482.1	flanking_5UTR	G	A	0.814	0.02947	G	A	A	G
	rs1250064	NM_212482.1	flanking_5UTR	T	G	0.8317	0.03772	T	G		
	rs1250079	NM_212482.1	flanking_5UTR	T	C	0.7999	0.009874	T	C		
	rs4411688	NM_212482.1	flanking_5UTR	G	A	1.218	0.02491	G	A		
	rs6713637	NM_212482.1	flanking_5UTR	G	A	1.35	0.002112	G	A		
FOXG1 forkhead box G1	rs10141935	NM_005249.3	flanking_5UTR	G	A	1.227	0.01131	G	A	G	A
	rs1885147	NM_005249.3	flanking_5UTR	A	G	0.811	0.01082	A	G	G	A
	rs7143611	NM_005249.3	flanking_5UTR	C	T	1.478	0.0179	C	T	G	A
	rs730195	NM_005249.3	flanking_5UTR	T	C	1.34	0.01415	T	C	A	G
	rs10132568	NM_005249.3	flanking_5UTR	A	G	1.224	0.01205	A	G		
	rs10483345	NM_005249.3	flanking_3UTR	A	G	1.213	0.02237	A	G		
	rs176338	NM_005249.3	flanking_5UTR	A	G	0.843	0.03431	A	G		
	rs2224437	NM_005249.3	flanking_5UTR	T	C	0.8224	0.02489	T	C		
	rs8012337	NM_005249.3	flanking_5UTR	G	T	0.8006	0.02766	G	T		
FSD1L fibronectin type III and SPRY domain containing 1-like	rs12335518	NM_031919.1	flanking_5UTR	C	A	1.927	0.009382	C	A		
GABRA2 gamma-aminobutyric acid (GABA) A receptor, alpha 2	rs1025852	NM_000807.1	flanking_5UTR	A	G	0.8283	0.0307	A	G	A	G
GABRB2 gamma-aminobutyric acid (GABA) A receptor, beta 2	rs6888242	NM_021911.1	flanking_5UTR	A	G	0.8441	0.04984	A	G	A	G
	rs7714930	NM_021911.1	intron	C	T	1.245	0.04743	C	T	G	A
	rs10037137	NM_021911.1	flanking_5UTR	T	G	0.8398	0.04398	T	G		
GABRB3 gamma-aminobutyric acid (GABA) A receptor, beta 3	rs10873636	NM_021912.2	intron	G	T	1.293	0.006327	G	T	C	A
	rs17738087	NM_021912.2	intron	C	T	1.334	0.04612	C	T	G	A
	rs1863459	NM_021912.2	intron	G	A	1.323	0.00925	G	A	G	A
	rs2928697	NM_021912.2	intron	T	C	1.201	0.04964	T	C	A	G
	rs4906676	NM_021912.2	flanking_3UTR	G	A	1.225	0.04231	G	A	G	A
	rs8027455	NM_021912.2	flanking_3UTR	A	C	1.289	0.002829	A	C	A	C
	rs933950	NM_021912.2	flanking_3UTR	T	C	1.288	0.003382	T	C	A	G
	rs11161324	NM_021912.2	intron	G	A	1.294	0.00613	G	A		
	rs4572353	NM_021912.2	flanking_3UTR	T	C	1.256	0.006747	T	C		
GABRG2 gamma-aminobutyric acid (GABA) A receptor, gamma 2	rs17060763	NM_198904.1	flanking_3UTR	T	G	0.6303	0.009148	T	G	A	C
	rs6881035	NM_198904.1	flanking_3UTR	G	T	0.791	0.0472	G	T	C	A
	rs7713655	NM_198904.1	flanking_5UTR	A	G	0.6052	0.01153	A	G	A	G
	rs13171710	NM_198904.1	flanking_3UTR	G	A	0.7821	0.03822	G	A		
	rs6864438	NM_198904.1	flanking_5UTR	C	T	1.255	0.03759	C	T		
GBAP1 glucosidase, beta, acid pseudogene 1	rs2049805	NR_002188.1	flanking_5UTR	G	A	0.8488	0.0468	G	A	A	G

GFAP glial fibrillary acidic protein	rs3169733	NM_002055.2	flanking_5UTR	G	A	0.8153	0.02821	G	A	G	A
	rs736866	NM_002055.2	flanking_5UTR	A	C	0.8082	0.0114	A	C	A	C
	rs3744473	NM_002055.2	flanking_3UTR	C	T	0.8239	0.03695	C	T		
	rs744281	NM_002055.2	flanking_5UTR	T	C	0.806	0.01052	T	C		
GNAI1 guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1	rs10486920	NM_002069.4	flanking_5UTR	A	G	0.7347	0.00182	A	G	A	G
	rs17802148	NM_002069.4	flanking_5UTR	C	T	0.7595	0.02124	C	T	G	A
	rs2523189	NM_002069.4	intron	T	C	1.194	0.03724	T	C	A	G
	rs2886609	NM_002069.4	flanking_3UTR	A	G	0.8164	0.01836	A	G	G	A
	rs2886611	NM_002069.4	intron	T	C	0.7558	0.006321	T	C	A	G
	rs4731111	NM_002069.4	flanking_5UTR	G	A	0.8001	0.0325	G	A	G	A
	rs6466884	NM_002069.4	flanking_5UTR	C	T	0.7516	0.01101	C	T	G	A
	rs12706724	NM_002069.4	flanking_3UTR	A	G	0.7121	0.00059	A	G		
	rs4731302	NM_002069.4	flanking_3UTR	C	T	0.7392	0.000998	C	T		
	rs7803811	NM_002069.4	flanking_5UTR	A	C	0.7485	0.0106	A	C		
	rs7805663	NM_002069.4	flanking_5UTR	G	A	0.7354	0.006723	G	A		
GNG12 guanine nucleotide binding protein (G protein), gamma 12	rs17531147	NM_018841.3	intron	A	G	1.378	0.02658	A	G		
GOT2 glutamic-oxaloacetic transaminase 2, mitochondrial	rs12709027	NM_002080.2	flanking_5UTR	A	G	0.8288	0.0269	A	G	A	G
	rs1351575	NM_002080.2	flanking_5UTR	G	T	0.8323	0.03067	G	T	C	A
	rs1445456	NM_002080.2	flanking_5UTR	A	C	1.204	0.04273	A	C	A	C
	rs13339064	NM_002080.2	flanking_5UTR	T	C	1.3	0.02053	T	C		
	rs9888768	NM_002080.2	flanking_5UTR	A	G	1.19	0.04885	A	G		
GRHPR glyoxylate reductase/hydroxypyruvate reductase	rs309453	NM_012203.1	intron	C	T	0.828	0.0319	C	T	A	G
	rs309455	NM_012203.1	intron	T	C	0.7308	0.03926	T	C	A	G
GRIA1 glutamate receptor, ionotropic, AMPA 1	rs4530817	NM_000827.2	intron	G	A	0.8415	0.04338	G	A	G	A
GRM3 glutamate receptor, metabotropic 3	rs17160519	NM_000840.2	flanking_5UTR	A	C	1.3	0.008344	A	C	A	C
	rs17315854	NM_000840.2	flanking_5UTR	C	T	0.7019	0.02115	C	T	G	A
	rs4236502	NM_000840.2	flanking_5UTR	C	A	1.2	0.03087	C	A	C	A
	rs6944937	NM_000840.2	flanking_5UTR	C	T	1.254	0.0211	C	T	G	A
	rs10499898	NM_000840.2	flanking_5UTR	T	C	0.651	0.006829	T	C		
	rs12668989	NM_000840.2	flanking_5UTR	C	T	1.262	0.0474	C	T		
	rs12673599	NM_000840.2	flanking_5UTR	T	C	1.309	0.02271	T	C		
	rs13222675	NM_000840.2	flanking_5UTR	C	T	0.6243	0.002859	C	T		
	rs13236080	NM_000840.2	flanking_5UTR	G	A	0.7174	0.02956	G	A		
	rs1527769	NM_000840.2	flanking_5UTR	T	C	0.66	0.008384	T	C		
	rs1554888	NM_000840.2	flanking_5UTR	C	T	1.295	0.02852	C	T		
	rs17161018	NM_000840.2	intron	A	G	1.271	0.04277	A	G		
	rs2373124	NM_000840.2	flanking_5UTR	G	A	0.7499	0.00497	G	A		
	rs2708553	NM_000840.2	flanking_5UTR	T	C	0.739	0.005225	T	C		
	rs41440	NM_000840.2	flanking_5UTR	G	A	0.7221	0.001126	G	A		

H2AFV H2A histone family, member V	rs3801403	NM_138635.2	intron	C	T	0.7736	0.03082	C	T		
HMGCR 3-hydroxy-3-methylglutaryl-CoA reductase	rs11742194	NM_000859.1	intron	T	C	1.313	0.03693	T	C	A	G
	rs3761740	NM_000859.1	flanking_5UTR	A	C	1.315	0.03334	A	C	A	C
	rs5909	NM_000859.1	3UTR	A	G	1.307	0.04048	A	G	A	G
HNRNPA2B1 heterogeneous nuclear ribonucleoprotein A2/B1	rs12672536	NM_031243.1	intron	C	T	0.7002	0.02132	C	T		
HSPH1 heat shock 105kDa/110kDa protein 1	rs2280059	NM_006644.2	5UTR	A	G	1.23	0.0492	A	G	A	G
	rs9591599	NM_006644.2	flanking_3UTR	C	T	0.792	0.01293	C	T	G	A
	rs1028713	NM_006644.2	flanking_3UTR	G	A	0.6879	0.01195	G	A		
	rs2224779	NM_006644.2	flanking_3UTR	A	G	0.7362	0.03786	A	G		
	rs7358865	NM_006644.2	flanking_3UTR	C	T	0.7424	0.01354	C	T		
	rs932715	NM_006644.2	flanking_3UTR	C	T	0.7269	0.01153	C	T		
HTR1B 5-hydroxytryptamine (serotonin) receptor 1B, G protein-coupled	rs1145831	NM_000863.1	flanking_3UTR	G	A	0.7971	0.02364	G	A	G	A
	rs12189558	NM_000863.1	flanking_5UTR	A	G	0.8014	0.04257	A	G	A	G
	rs12202018	NM_000863.1	flanking_3UTR	G	A	1.204	0.04768	G	A	G	A
	rs12206214	NM_000863.1	flanking_3UTR	A	G	1.181	0.04411	A	G	A	G
	rs1343485	NM_000863.1	flanking_3UTR	G	T	1.251	0.01128	G	T	C	A
	rs1548240	NM_000863.1	flanking_3UTR	A	G	1.22	0.04113	A	G	A	G
	rs17263843	NM_000863.1	flanking_3UTR	T	C	1.27	0.009987	T	C	A	G
	rs1777766	NM_000863.1	flanking_3UTR	C	T	0.7344	0.01584	C	T	G	A
	rs2504298	NM_000863.1	flanking_5UTR	C	T	1.297	0.02946	C	T	G	A
	rs4265000	NM_000863.1	flanking_3UTR	A	G	1.227	0.02681	A	G	A	G
	rs7747803	NM_000863.1	flanking_3UTR	G	A	1.185	0.04026	G	A	A	G
	rs9343725	NM_000863.1	flanking_5UTR	C	T	1.477	0.03446	C	T	G	A
	rs10484725	NM_000863.1	flanking_5UTR	G	A	1.28	0.02722	G	A		
	rs10943426	NM_000863.1	flanking_3UTR	G	A	1.19	0.03507	G	A		
	rs1213378	NM_000863.1	flanking_5UTR	A	G	0.8262	0.02151	A	G		
	rs1350406	NM_000863.1	flanking_3UTR	G	A	1.208	0.02456	G	A		
	rs17209149	NM_000863.1	flanking_5UTR	C	A	1.278	0.03586	C	A		
	rs2320160	NM_000863.1	flanking_3UTR	A	G	1.557	0.001425	A	G		
	rs4075570	NM_000863.1	flanking_3UTR	G	A	0.8462	0.04935	G	A		
	rs7753704	NM_000863.1	flanking_5UTR	C	T	1.311	0.03269	C	T		
	rs9294040	NM_000863.1	flanking_3UTR	A	G	0.8012	0.02402	A	G		
	rs9341629	NM_000863.1	flanking_3UTR	T	C	1.553	0.001517	T	C		
	rs9343544	NM_000863.1	flanking_3UTR	G	A	1.557	0.001425	G	A		
	rs9343676	NM_000863.1	flanking_5UTR	T	C	1.312	0.01347	T	C		
	rs9350721	NM_000863.1	flanking_5UTR	T	G	1.31	0.01541	T	G		
	rs9352483	NM_000863.1	flanking_5UTR	A	G	1.275	0.03796	A	G		
	rs9359291	NM_000863.1	flanking_5UTR	T	G	1.284	0.02446	T	G		

	rs9361346	NM_000863.1	flanking_5UTR	G	A	1.323	0.02385	G	A		
	rs9361361	NM_000863.1	flanking_5UTR	T	C	1.477	0.03446	T	C		
	rs9443501	NM_000863.1	flanking_5UTR	T	C	1.342	0.01713	T	C		
	rs953576	NM_000863.1	flanking_5UTR	C	T	1.283	0.02522	C	T		
HTR2A 5-hydroxytryptamine (serotonin) receptor 2A, G protein-coupled	rs1328684	NM_000621.2	intron	C	T	0.8332	0.03825	C	T	G	A
	rs1536762	NM_000621.2	flanking_5UTR	C	T	1.353	0.04417	C	T	G	A
	rs2406091	NM_000621.2	flanking_5UTR	A	G	0.7913	0.0471	A	G	A	G
	rs2802396	NM_000621.2	flanking_5UTR	C	T	0.8346	0.03658	C	T	G	A
	rs582385	NM_000621.2	intron	C	T	0.808	0.04868	C	T	G	A
	rs7333963	NM_000621.2	flanking_5UTR	T	C	0.781	0.03023	T	C	A	G
	rs9567733	NM_000621.2	flanking_3UTR	G	A	0.8126	0.03314	G	A	G	A
	rs1923885	NM_000621.2	intron	C	T	1.277	0.003812	C	T		
	rs4142745	NM_000621.2	flanking_5UTR	T	C	0.8365	0.03816	T	C		
	rs6561363	NM_000621.2	flanking_5UTR	C	A	1.84	0.03841	C	A		
	rs9526329	NM_000621.2	flanking_5UTR	T	G	1.361	0.0306	T	G		
IFITM3 interferon induced transmembrane protein 3	rs11246074	NM_021034.1	flanking_5UTR	G	A	1.263	0.03415	G	A	G	A
IGF1 insulin-like growth factor 1 (somatomedin C)	rs7306935	NM_000618.2	flanking_5UTR	C	A	1.33	0.007657	C	A	C	A
	rs1350356	NM_000618.2	flanking_5UTR	A	G	1.208	0.03705	A	G		
IGF1R insulin-like growth factor 1 receptor	rs12908437	NM_000875.2	intron	T	C	1.227	0.01555	T	C	G	A
	rs2311767	NM_000875.2	intron	T	C	0.6081	0.03992	T	C	A	G
	rs3743264	NM_000875.2	intron	G	A	1.262	0.00708	G	A	G	A
	rs4426332	NM_000875.2	intron	T	C	0.6337	0.03654	T	C	A	G
	rs4527002	NM_000875.2	flanking_5UTR	C	T	1.191	0.03549	C	T	A	G
	rs6598534	NM_000875.2	flanking_5UTR	A	G	1.214	0.0206	A	G	G	A
	rs7181975	NM_000875.2	flanking_5UTR	T	C	1.206	0.0298	T	C	A	G
	rs8038015	NM_000875.2	intron	C	T	1.18	0.04825	C	T	G	A
	rs1007212	NM_000875.2	intron	C	T	1.255	0.0086	C	T		
	rs11633294	NM_000875.2	intron	A	C	1.209	0.02995	A	C		
	rs12898502	NM_000875.2	intron	T	C	1.214	0.02518	T	C		
INSIG1 insulin induced gene 1	rs3923644	NM_198337.1	flanking_5UTR	A	C	0.8361	0.03586	A	C	A	C
IST1 increased sodium tolerance 1 homolog (yeast)	rs4788449	NM_014761.2	intron	G	A	1.253	0.01156	G	A		
JMJD8 jumonji domain containing 8	rs6597	NM_001005920.2	3UTR	G	T	1.379	0.003844	G	T	C	A
KDM4C lysine (K)-specific demethylase 4C	rs10116021	NM_015061.1	intron	A	G	0.8013	0.03837	A	G	A	G
	rs10491779	NM_015061.1	intron	A	G	0.7944	0.01191	A	G	A	G
	rs1052489	NM_015061.1	flanking_5UTR	C	T	0.7891	0.04302	C	T	G	A
	rs10815532	NM_015061.1	intron	C	T	1.225	0.01639	C	T	A	G
	rs1093707	NM_015061.1	intron	G	T	0.8012	0.02729	G	T	C	A
	rs10976080	NM_015061.1	intron	A	C	1.258	0.023	A	C	A	C

	rs10976081	NM_015061.1	intron	T	G	1.25	0.0442	T	G	A	C
	rs10976092	NM_015061.1	flanking_3UTR	G	A	1.224	0.04075	G	A	G	A
	rs10976129	NM_015061.1	flanking_3UTR	C	A	0.8385	0.04034	C	A	C	A
	rs11790877	NM_015061.1	intron	G	A	1.295	0.001616	G	A	G	A
	rs11998950	NM_015061.1	intron	T	C	1.354	0.01695	T	C	A	G
	rs12001316	NM_015061.1	intron	A	G	0.7336	0.01791	A	G	A	G
	rs12345872	NM_015061.1	intron	G	T	1.299	0.00938	G	T	C	A
	rs1570508	NM_015061.1	intron	T	G	0.7665	0.002529	T	G	A	C
	rs1570512	NM_015061.1	intron	G	A	0.7835	0.00532	G	A	G	A
	rs16925068	NM_015061.1	intron	T	C	0.6012	0.0151	T	C	A	G
	rs2381530	NM_015061.1	intron	A	C	1.32	0.002137	A	C	A	C
	rs2770759	NM_015061.1	flanking_3UTR	C	T	1.245	0.01081	C	T	G	A
	rs2792238	NM_015061.1	intron	A	G	0.8045	0.008342	A	G	A	G
	rs2820914	NM_015061.1	flanking_3UTR	C	T	1.254	0.02148	C	T	G	A
	rs2890733	NM_015061.1	intron	A	G	0.8289	0.02943	A	G	G	A
	rs2990658	NM_015061.1	intron	T	C	1.194	0.04071	T	C	A	G
	rs4742269	NM_015061.1	intron	A	G	0.5016	0.001892	A	G	A	G
	rs4742298	NM_015061.1	intron	A	G	0.7442	0.003066	A	G	A	G
	rs4742301	NM_015061.1	intron	G	A	0.7653	0.01159	G	A	G	A
	rs7019042	NM_015061.1	intron	G	A	0.804	0.01098	G	A	A	G
	rs722628	NM_015061.1	intron	A	G	0.8283	0.03019	A	G	A	G
	rs7848022	NM_015061.1	intron	T	G	1.272	0.003954	T	G	C	A
	rs818887	NM_015061.1	intron	G	A	0.7724	0.009574	G	A	G	A
	rs913588	NM_015061.1	coding	A	G	0.8409	0.03829	A	G	A	G
	rs946929	NM_015061.1	intron	T	C	0.7932	0.004648	T	C	G	A
	rs10815539	NM_015061.1	flanking_3UTR	G	A	1.222	0.04637	G	A		
	rs12353012	NM_015061.1	flanking_3UTR	G	T	0.7912	0.04072	G	T		
	rs12553351	NM_015061.1	intron	G	T	0.4783	0.000602	G	T		
	rs2381536	NM_015061.1	intron	G	A	0.7924	0.008355	G	A		
	rs2381542	NM_015061.1	intron	G	A	0.7618	0.005789	G	A		
	rs2765964	NM_015061.1	flanking_3UTR	A	G	0.7605	0.04901	A	G		
	rs2792235	NM_015061.1	intron	A	G	1.241	0.01403	A	G		
	rs7027375	NM_015061.1	intron	G	A	1.227	0.01523	G	A		
	rs7031179	NM_015061.1	intron	G	A	0.7602	0.005425	G	A		
	rs7037553	NM_015061.1	flanking_3UTR	G	A	1.238	0.04255	G	A		
	rs7045009	NM_015061.1	intron	T	C	0.7675	0.002663	T	C		
	rs876294	NM_015061.1	flanking_3UTR	C	T	1.311	0.02588	C	T		
	rs913587	NM_015061.1	intron	A	C	1.196	0.03217	A	C		
KDR kinase insert domain receptor (a type III receptor	rs2305948	NM_002253.1	coding	T	C	0.5813	0.000462	T	C	A	G
	rs17709427	NM_002253.1	flanking_3UTR	C	T	0.788	0.03393	C	T		

tyrosine kinase)	rs2305949	NM_002253.1	intron	T	C	0.7552	0.01052	T	C		
KLK6 kallikrein-related peptidase 6	rs1654531	NM_002774.3	flanking_5UTR	C	T	0.7985	0.03805	C	T	G	A
	rs4592765	NM_001012965.1	flanking_3UTR	C	T	0.8344	0.03726	C	T	G	A
	rs1654537	NM_001012965.1	intron	C	T	0.7832	0.007661	C	T		
MAN2A1 mannosidase, alpha, class 2A, member 1	rs17162540	NM_002372.2	flanking_3UTR	G	A	1.33	0.03099	G	A	G	A
	rs2416227	NM_002372.2	flanking_3UTR	C	A	1.31	0.01905	C	A	C	A
	rs6865255	NM_002372.2	flanking_3UTR	G	A	1.236	0.02691	G	A	A	G
	rs10900662	NM_002372.2	flanking_3UTR	A	G	1.235	0.02819	A	G		
MAP1B microtubule-associated protein 1B	rs1217785	NM_005909.2	flanking_5UTR	T	C	1.22	0.01705	T	C	A	G
	rs2118695	NM_032010.1	intron	A	G	0.8214	0.02542	A	G	G	A
MAPT microtubule-associated protein tau	rs11867549	NM_005910.2	intron	G	A	0.7858	0.01196	G	A	G	A
	rs4792893	NM_005910.2	intron	A	G	0.7216	0.01245	A	G	A	G
	rs4792891	NM_005910.2	intron	G	T	0.7896	0.009256	G	T		
MAT2B methionine adenosyltransferase II, beta	rs10042643	NM_013283.3	flanking_3UTR	T	C	1.239	0.02313	T	C	A	G
	rs1030674	NM_013283.3	flanking_3UTR	C	A	1.262	0.01891	C	A	C	A
	rs10515861	NM_182796.1	intron	C	T	0.805	0.0144	C	T	G	A
	rs10515902	NM_013283.3	flanking_3UTR	A	G	1.362	0.04805	A	G	A	G
	rs11740544	NM_013283.3	flanking_3UTR	T	C	1.513	0.01452	T	C	A	G
	rs1421630	NM_013283.3	flanking_3UTR	A	C	0.7929	0.01251	A	C	A	C
	rs1469064	NM_013283.3	flanking_3UTR	T	C	0.8084	0.01651	T	C	A	G
	rs1895139	NM_013283.3	flanking_3UTR	G	A	0.8257	0.02599	G	A	G	A
	rs3980160	NM_013283.3	flanking_3UTR	G	A	0.8402	0.03751	G	A	G	A
	rs6864913	NM_013283.3	flanking_3UTR	A	G	1.202	0.04408	A	G	A	G
	rs6867953	NM_013283.3	flanking_3UTR	C	A	1.195	0.03792	C	A	A	C
	rs6875996	NM_013283.3	flanking_3UTR	A	G	0.8221	0.02246	A	G	A	G
	rs7443611	NM_013283.3	flanking_3UTR	A	G	0.8214	0.02262	A	G	A	G
	rs7702763	NM_013283.3	flanking_3UTR	A	G	1.228	0.02894	A	G	A	G
	rs7719164	NM_013283.3	flanking_3UTR	G	A	1.453	0.04043	G	A	G	A
	rs1433779	NM_013283.3	flanking_3UTR	T	C	1.616	0.005647	T	C		
	rs1895210	NM_013283.3	flanking_3UTR	C	T	0.6238	0.04619	C	T		
	rs2635175	NM_013283.3	flanking_3UTR	T	C	1.181	0.04515	T	C		
	rs294587	NM_013283.3	flanking_3UTR	G	A	1.201	0.0424	G	A		
	rs297966	NM_013283.3	flanking_3UTR	C	T	1.291	0.04706	C	T		
	rs4073827	NM_013283.3	flanking_3UTR	T	C	0.826	0.02596	T	C		
	rs4868851	NM_013283.3	flanking_3UTR	G	T	0.8152	0.01707	G	T		
	rs7735401	NM_013283.3	flanking_3UTR	A	G	1.308	0.04412	A	G		
MBP myelin basic protein	rs1124941	NM_001025101.1	flanking_5UTR	T	G	1.258	0.006503	T	G	A	C
	rs1789094	NM_001025101.1	flanking_5UTR	T	C	0.776	0.02123	T	C	A	G
	rs1789103	NM_001025101.1	flanking_5UTR	T	G	0.7456	0.009002	T	G	A	C
	rs2282566	NM_001025101.1	intron	T	C	1.224	0.01571	T	C	A	G

	rs470131	NM_001025101.1	intron	C	A	1.351	0.01954	C	A	C	A
	rs4890912	NM_001025101.1	flanking_5UTR	G	A	1.202	0.03702	G	A	G	A
	rs736421	NM_001025101.1	intron	A	G	0.7983	0.006722	A	G	A	G
	rs9947485	NM_001025101.1	flanking_5UTR	T	C	1.275	0.03242	T	C	A	G
	rs9951586	NM_001025101.1	flanking_5UTR	C	T	0.7586	0.01134	C	T	G	A
	rs1015820	NM_001025101.1	flanking_5UTR	G	A	1.229	0.01371	G	A		
	rs11877526	NM_001025101.1	flanking_5UTR	G	A	1.254	0.007231	G	A		
	rs1562771	NM_001025101.1	flanking_5UTR	A	G	1.263	0.04122	A	G		
	rs1667952	NM_001025101.1	flanking_5UTR	C	T	0.7568	0.01048	C	T		
	rs1789105	NM_001025101.1	flanking_5UTR	C	T	0.7444	0.008632	C	T		
	rs1789139	NM_001025101.1	flanking_5UTR	T	C	1.407	0.009443	T	C		
	rs1812680	NM_001025101.1	flanking_5UTR	G	A	1.267	0.008193	G	A		
METAP2 methionyl aminopeptidase 2	rs10777695	NM_006838.2	flanking_5UTR	T	C	0.7518	0.000866	T	C	G	A
	rs159853	NM_006838.2	flanking_5UTR	T	C	0.7994	0.01376	T	C	G	A
	rs2727639	NM_006838.2	flanking_5UTR	G	A	0.739	0.000592	G	A	G	A
	rs282323	NM_006838.2	flanking_5UTR	G	A	0.6665	0.000198	G	A	G	A
	rs301008	NM_006838.2	flanking_5UTR	G	T	0.8035	0.03593	G	T	C	A
	rs301013	NM_006838.2	intron	A	C	0.6609	0.000546	A	C	A	C
	rs2456549	NM_006838.2	flanking_5UTR	G	T	0.7105	0.001602	G	T		
	rs2596761	NM_006838.2	flanking_5UTR	A	G	0.6723	0.004958	A	G		
	rs2769432	NM_006838.2	flanking_5UTR	A	G	0.6759	0.005181	A	G		
	rs2769436	NM_006838.2	flanking_5UTR	G	A	0.7131	0.001818	G	A		
	rs301042	NM_006838.2	intron	C	T	0.6707	0.00469	C	T		
	rs964496	NM_006838.2	flanking_5UTR	T	C	0.6723	0.004958	T	C		
MKLN1 muskelin 1, intracellular mediator containing kelch motifs	rs13237471	NM_013255.3	flanking_5UTR	A	G	0.7906	0.04599	A	G	A	G
	rs3800678	NM_013255.3	intron	G	A	0.5989	0.01405	G	A		
MOBP myelin-associated oligodendrocyte basic protein	rs2233204	NM_006501.1	intron	T	C	0.8294	0.02925	T	C	A	G
	rs562545	NM_006501.1	intron	A	G	1.241	0.01231	A	G	A	G
MOG myelin oligodendrocyte glycoprotein	rs2747442	NM_001008229.1	flanking_3UTR	G	A	1.192	0.0428	G	A	G	A
	rs3117292	NM_001008229.1	flanking_3UTR	G	A	1.226	0.01429	G	A	A	G
	rs3117294	NM_001008229.1	flanking_3UTR	G	T	1.21	0.02512	G	T	A	C
MPDZ multiple PDZ domain protein	rs10514824	NM_003829.1	flanking_3UTR	A	G	1.281	0.009825	A	G	A	G
	rs10809906	NM_003829.1	intron	A	G	0.746	0.01461	A	G	A	G
	rs1412110	NM_003829.1	intron	T	C	0.695	0.03008	T	C	A	G
	rs1889297	NM_003829.1	flanking_5UTR	T	C	1.211	0.02368	T	C	G	A
	rs3264	NM_003829.1	3UTR	C	T	1.224	0.01967	C	T	G	A
	rs10435761	NM_003829.1	flanking_3UTR	T	C	0.7788	0.01546	T	C		
	rs10514823	NM_003829.1	flanking_3UTR	C	T	1.35	0.001443	C	T		
	rs1328906	NM_003829.1	flanking_3UTR	T	C	1.263	0.01647	T	C		
MPZL2 myelin protein	rs2187557	NM_005797.2	flanking_5UTR	T	C	0.8067	0.02351	T	C	A	G

zero-like 2	rs3759001	NM_005797.2	flanking_5UTR	T	C	0.8251	0.02419	T	C	G	A
	rs7940405	NM_005797.2	flanking_5UTR	G	A	0.8302	0.0304	G	A		
MYL2 myosin, light chain 2, regulatory, cardiac, slow	rs933296	NM_000432.1	intron	T	G	1.243	0.01066	T	G	A	C
MYO1B myosin IB	rs16833762	NM_012223.2	flanking_3UTR	C	T	0.7377	0.02792	C	T	G	A
MYO1D myosin ID	rs225184	NM_015194.1	intron	G	A	1.426	0.000219	G	A	G	A
	rs225206	NM_015194.1	intron	A	C	0.7893	0.004985	A	C	A	C
	rs225209	NM_015194.1	intron	T	C	0.8016	0.008687	T	C	A	G
	rs2640828	NM_015194.1	intron	A	G	0.7093	0.000289	A	G	A	G
	rs414947	NM_015194.1	intron	G	A	1.279	0.003381	G	A	A	G
NFIB nuclear factor I/B	rs10481545	NM_005596.1	flanking_5UTR	T	C	0.8263	0.02331	T	C	G	A
	rs10810110	NM_005596.1	intron	G	A	0.6459	0.002111	G	A	G	A
	rs10810113	NM_005596.1	intron	T	C	0.6791	0.009812	T	C	A	G
	rs10810136	NM_005596.1	flanking_5UTR	G	A	1.198	0.03335	G	A	G	A
	rs11788150	NM_005596.1	flanking_3UTR	A	C	1.192	0.04248	A	C	A	C
	rs12341720	NM_005596.1	flanking_3UTR	T	C	1.224	0.02579	T	C	A	G
	rs12377502	NM_005596.1	intron	G	A	0.6485	0.004182	G	A	G	A
	rs1323339	NM_005596.1	flanking_5UTR	T	G	0.8386	0.04619	T	G	A	C
	rs1556032	NM_005596.1	flanking_5UTR	C	T	1.181	0.04329	C	T	A	G
	rs17708974	NM_005596.1	flanking_5UTR	A	C	0.8002	0.04734	A	C	A	C
	rs4740570	NM_005596.1	flanking_5UTR	T	C	1.2	0.03085	T	C	G	A
	rs548824	NM_005596.1	intron	G	A	1.289	0.02364	G	A	G	A
	rs7042750	NM_005596.1	flanking_5UTR	T	C	0.7746	0.03106	T	C	A	G
	rs7858	NM_005596.1	3UTR	A	G	1.311	0.01804	A	G	A	G
	rs7866165	NM_005596.1	flanking_3UTR	T	G	1.31	0.003853	T	G	A	C
	rs10756536	NM_005596.1	intron	G	A	0.6455	0.004939	G	A		
	rs10810098	NM_005596.1	intron	T	C	0.7369	0.002888	T	C		
	rs2382456	NM_005596.1	intron	T	C	0.6077	0.00072	T	C		
	rs2382459	NM_005596.1	intron	T	G	0.6048	0.000928	T	G		
	rs7021837	NM_005596.1	flanking_3UTR	C	T	1.243	0.0305	C	T		
NOL11 nucleolar protein 11	rs1542610	NM_015462.2	intron	T	C	1.35	0.01285	T	C	A	G
	rs4496198	NM_015462.2	flanking_5UTR	C	T	1.372	0.009119	C	T		
NRXN3 neurexin 3	rs10142370	NM_004796.3	intron	A	G	0.7015	0.01472	A	G	A	G
	rs1030127	NM_004796.3	intron	C	T	1.351	0.03401	C	T	G	A
	rs10400751	NM_004796.3	intron	C	T	1.285	0.01667	C	T	G	A
	rs10873325	NM_004796.3	flanking_3UTR	T	C	0.7509	0.001212	T	C	A	G
	rs12891137	NM_138970.2	intron	C	T	1.359	0.0213	C	T	G	A
	rs17109935	NM_004796.3	intron	C	T	1.271	0.02497	C	T	G	A
	rs17110183	NM_004796.3	flanking_3UTR	A	G	1.214	0.03996	A	G	A	G
	rs17175771	NM_004796.3	flanking_3UTR	G	T	1.237	0.04755	G	T	C	A

	rs1861957	NM_004796.3	intron	A	C	1.223	0.01556	A	C	C	A
	rs1863034	NM_004796.3	intron	C	T	0.8338	0.03839	C	T	A	G
	rs6574495	NM_004796.3	intron	T	C	1.204	0.02694	T	C	G	A
	rs766024	NM_138970.2	intron	T	C	1.451	0.001968	T	C	A	G
	rs9646166	NM_004796.3	intron	C	T	1.218	0.02751	C	T	G	A
	rs10131645	NM_004796.3	intron	G	A	1.234	0.02226	G	A		
	rs12890943	NM_004796.3	intron	C	T	1.43	0.01689	C	T		
	rs17597295	NM_138970.2	intron	T	C	0.7997	0.04926	T	C		
	rs2193135	NM_004796.3	flanking_3UTR	T	C	1.246	0.03198	T	C		
	rs4899727	NM_004796.3	intron	G	A	1.509	0.01379	G	A		
	rs725693	NM_004796.3	intron	G	T	0.8354	0.04028	G	T		
NSMAF neutral sphingomyelinase (N-SMase) activation associated factor	rs4737508	NM_003580.2	flanking_5UTR	C	T	1.271	0.04528	C	T	G	A
NTM neurotrimin	rs3099774	NM_016522.2	intron	T	C	1.342	0.01471	T	C	A	G
	rs496368	NM_016522.2	intron	C	T	1.178	0.04836	C	T	G	A
	rs7116467	NM_016522.2	intron	G	A	0.8083	0.01248	G	A	G	A
	rs3099775	NM_016522.2	intron	T	C	1.331	0.0212	T	C		
	rs3099787	NM_016522.2	intron	G	A	1.38	0.008411	G	A		
NTRK2 neurotrophic tyrosine kinase, receptor, type 2	rs3739570	NM_006180.3	3UTR	T	C	0.7144	0.04205	T	C	A	G
	rs11140659	NM_006180.3	flanking_5UTR	T	C	0.8098	0.03423	T	C		
	rs11140688	NM_006180.3	flanking_5UTR	T	C	1.371	0.009416	T	C		
	rs1866438	NM_006180.3	flanking_5UTR	T	C	0.7811	0.02894	T	C		
PDK4 pyruvate dehydrogenase kinase, isozyme 4	rs854061	NM_002612.2	flanking_5UTR	G	A	0.7013	0.03023	G	A		
PDYN prodynorphin	rs9679771	NM_024411.2	flanking_3UTR	A	G	0.6911	0.03805	A	G	A	G
PFDN6 prefoldin subunit 6	rs456261	NM_014260.2	intron	C	T	0.806	0.01101	C	T	G	A
PGM2L1 phosphoglucomutase 2-like 1	rs3193507	NM_173582.3	3UTR	A	G	1.213	0.01831	A	G		
PLD1 phospholipase D1, phosphatidylcholine-specific	rs3850186	NM_002662.2	flanking_5UTR	A	G	1.207	0.02592	A	G	A	G
	rs3850188	NM_002662.2	flanking_5UTR	C	A	1.47	0.001101	C	A	C	A
	rs6787821	NM_002662.2	flanking_5UTR	G	A	1.327	0.008373	G	A	G	A
	rs7631361	NM_002662.2	flanking_5UTR	A	G	1.375	0.004761	A	G	A	G
	rs9817267	NM_002662.2	flanking_5UTR	G	A	0.8337	0.0269	G	A	A	G
	rs9848505	NM_002662.2	flanking_5UTR	C	T	0.8298	0.0248	C	T	G	A
	rs9968111	NM_002662.2	flanking_5UTR	A	C	1.396	0.002883	A	C	A	C
	rs9849306	NM_002662.2	flanking_5UTR	A	G	1.312	0.01375	A	G		
	rs9872465	NM_002662.2	flanking_5UTR	T	C	1.312	0.01142	T	C		
PLLP plasmolipin	rs16969211	NM_015993.1	flanking_5UTR	G	A	1.214	0.03152	G	A		
PPM1B protein phosphatase, Mg2+/Mn2+ dependent, 1B	rs4952703	NM_001033557.1	intron	T	C	1.315	0.03325	T	C	A	G

PPP1R12C protein phosphatase 1, regulatory subunit 12C	rs575144	NM_017607.1	intron	T	C	1.332	0.03198	T	C		
PRKCZ protein kinase C, zeta	rs908742	NM_001033582.1	intron	A	G	1.187	0.04567	A	G	A	G
PSMA1 proteasome (prosome, macropain) subunit, alpha type, 1	rs2575850	NM_148976.1	intron	C	T	0.826	0.02751	C	T	G	A
	rs11023274	NM_148976.1	intron	A	G	0.6056	0.03217	A	G		
PSMD13 proteasome (prosome, macropain) 26S subunit, non-ATPase, 13	rs532483	NM_175932.1	flanking_3UTR	A	G	1.326	0.0359	A	G	A	G
	rs6598055	NM_175932.1	intron	A	G	0.8222	0.02309	A	G		
PTPRE protein tyrosine phosphatase, receptor type, E	rs12249657	NM_130435.2	intron	A	G	1.303	0.02074	A	G	A	G
	rs12415045	NM_006504.3	intron	T	C	1.288	0.00429	T	C	G	A
	rs12769331	NM_006504.3	intron	G	A	1.35	0.00919	G	A	G	A
	rs932837	NM_006504.3	intron	A	G	1.293	0.03254	A	G	A	G
	rs1359850	NM_006504.3	intron	C	T	1.318	0.03397	C	T		
	rs7895103	NM_006504.3	intron	C	T	0.8038	0.02354	C	T		
QDPR quinoid dihydropteridine reductase	rs2697707	NM_000320.1	flanking_3UTR	A	G	1.45	0.03686	A	G	A	G
	rs2697705	NM_000320.1	flanking_3UTR	G	A	1.254	0.02419	G	A		
RFTN1 raftlin, lipid raft linker 1	rs553681	NM_015150.1	intron	T	C	1.216	0.0204	T	C	A	G
	rs556322	NM_015150.1	intron	A	G	0.8185	0.02653	A	G	A	G
	rs7621664	NM_015150.1	intron	C	T	0.7256	0.02672	C	T	G	A
	rs7644810	NM_015150.1	intron	C	T	0.833	0.03795	C	T	G	A
	rs17272509	NM_015150.1	intron	C	A	0.7342	0.03659	C	A		
	rs9990208	NM_015150.1	intron	T	C	0.6614	0.009928	T	C		
RHOB ras homolog family member B	rs13016649	NM_004040.2	flanking_3UTR	A	C	0.8152	0.01479	A	C	A	C
	rs342056	NM_004040.2	flanking_3UTR	A	G	1.353	0.000614	A	G	A	G
	rs959058	NM_004040.2	flanking_3UTR	T	C	1.196	0.03117	T	C	G	A
	rs342092	NM_004040.2	flanking_3UTR	G	A	1.279	0.00639	G	A		
	rs6531240	NM_004040.2	flanking_3UTR	C	T	1.32	0.000836	C	T		
RXRG retinoid X receptor, gamma	rs10800098	NM_001009598.1	intron	A	G	1.583	0.005815	A	G	A	G
SCD stearoyl-CoA desaturase (delta-9-desaturase)	rs10883463	NM_005063.4	intron	C	T	1.418	0.01946	C	T		
	rs7068970	NM_005063.4	flanking_3UTR	C	A	1.452	0.01188	C	A		
SCG2 secretogranin II	rs13402129	NM_003469.3	flanking_5UTR	G	T	0.6328	0.01329	G	T	C	A
	rs6728104	NM_003469.3	flanking_3UTR	G	A	0.7689	0.04294	G	A	G	A
	rs7565690	NM_003469.3	flanking_3UTR	T	G	1.276	0.03505	T	G	A	C
	rs10172774	NM_003469.3	flanking_5UTR	T		0.8242	0.02825	T	C		
	rs17827801	NM_003469.3	flanking_3UTR	T	C	1.394	0.02867	T	C		
	rs1921634	NM_003469.3	flanking_3UTR	A	C	0.7994	0.04793	A	C		
	rs1921651	NM_003469.3	flanking_3UTR	C	T	1.208	0.02571	C	T		
SDC4 syndecan 4	rs11696248	NM_002999.2	flanking_3UTR	C	T	0.7949	0.04524	C	T	G	A
	rs2267869	NM_002999.2	intron	T	C	0.8233	0.0212	T	C	G	A

	rs2284277	NM_002999.2	intron	A	G	0.8036	0.03722	A	G	A	G
SESTD1 SEC14 and spectrin domains 1	rs12474600	NM_178123.3	flanking_3UTR	A	G	0.7581	0.04837	A	G	A	G
	rs12693178	NM_178123.3	intron	G	A	1.227	0.01364	G	A	A	G
	rs1968618	NM_178123.3	flanking_3UTR	A	G	1.265	0.007605	A	G	G	A
	rs3731739	NM_178123.3	flanking_3UTR	A	G	1.375	0.03142	A	G	A	G
	rs6433746	NM_178123.3	flanking_3UTR	G	A	1.28	0.04378	G	A	G	A
	rs17363393	NM_178123.3	intron	A	G	1.412	0.01995	A	G		
	rs2305170	NM_178123.3	intron	G	T	1.424	0.01764	G	T		
	rs2592579	NM_178123.3	flanking_3UTR	T	C	1.288	0.04496	T	C		
SETD1A SET domain containing 1A	rs2305884	NM_014712.1	intron	C	T	1.219	0.01996	C	T		
	rs897986	NM_014712.1	intron	A	G	1.221	0.01892	A	G		
SH2B3 SH2B adaptor protein 3	rs2239194	NM_005475.1	intron	A	G	1.359	0.03346	A	G	A	G
	rs10774623	NM_005475.1	flanking_5UTR	G	A	1.268	0.01612	G	A		
SLC11A1 solute carrier family 11 (proton-coupled divalent metal ion transporter), member 1	rs2290708	NM_001032220.1	intron	T	C	1.31	0.005329	T	C	A	G
	rs3816560	NM_001032220.1	intron	C	T	1.285	0.008734	C	T		
SLC12A2 solute carrier family 12 (sodium/potassium/ chloride transporter), member 2	rs36693	NM_001046.2	flanking_3UTR	T	C	0.7199	0.00193	T	C	A	G
	rs3749748	NM_001046.2	flanking_5UTR	T	C	0.7895	0.0255	T	C	A	G
	rs790155	NM_001046.2	intron	G	A	0.8097	0.04905	G	A	G	A
	rs2409037	NM_001046.2	flanking_5UTR	T	G	0.7727	0.01808	T	G		
	rs36701	NM_001046.2	flanking_3UTR	G	A	0.7625	0.0144	G	A		
	rs790154	NM_001046.2	intron	T	C	0.8075	0.04671	T	C		
SNCA synuclein, alpha (non A4 component of amyloid precursor)	rs17015888	NM_000345.2	flanking_3UTR	A	G	0.6525	0.0476	A	G	A	G
	rs17015982	NM_000345.2	flanking_3UTR	G	A	0.6919	0.02363	G	A		
	rs6532183	NM_000345.2	flanking_3UTR	G	A	0.7021	0.02535	G	A		
	rs7668883	NM_000345.2	flanking_3UTR	G	A	0.6525	0.0476	G	A		
SOX9 SRY (sex determining region Y)-box 9	rs17246175	NM_000346.2	flanking_3UTR	T	C	1.393	0.008321	T	C	A	G
	rs2430549	NM_000346.2	flanking_5UTR	C	T	0.8084	0.03832	C	T	G	A
	rs4793400	NM_000346.2	flanking_3UTR	A	C	1.382	0.006083	A	C	A	C
	rs7214255	NM_000346.2	flanking_5UTR	A	G	1.209	0.02469	A	G	A	G
	rs725545	NM_000346.2	flanking_5UTR	G	A	1.253	0.02052	G	A	G	A
	rs918077	NM_000346.2	flanking_3UTR	C	T	1.283	0.02908	C	T	G	A
	rs12449701	NM_000346.2	flanking_3UTR	G	T	1.323	0.02064	G	T		
	rs6501473	NM_000346.2	flanking_5UTR	C	T	1.267	0.02891	C	T		
	rs6501474	NM_000346.2	flanking_5UTR	C	T	1.261	0.03283	C	T		
	rs6501475	NM_000346.2	flanking_5UTR	C	T	1.262	0.03128	C	T		
	rs713158	NM_000346.2	flanking_5UTR	T	C	1.313	0.01667	T	C		
	rs8068789	NM_000346.2	flanking_3UTR	A	G	1.445	0.001764	A	G		
	rs975737	NM_000346.2	flanking_5UTR	T	C	1.273	0.02589	T	C		
SPARC secreted protein,	rs2033467	NM_003118.2	flanking_3UTR	C	T	1.18	0.04751	C	T	A	G

acidic, cysteine-rich (osteonectin)	rs7719521	NM_003118.2	intron	A	C	1.202	0.02757	A	C	A	C
	rs6861486	NM_003118.2	intron	T	C	1.202	0.02757	T	C		
STX1A syntaxin 1A (brain)	rs6951030	NM_004603.1	intron	G	T	1.269	0.02611	G	T	C	A
SYN2 synapsin II	rs17669026	NM_133625.2	flanking_5UTR	G	A	1.251	0.03978	G	A	G	A
SYT1 synaptotagmin I	rs10861755	NM_005639.1	intron	G	A	1.195	0.04617	G	A	G	A
	rs1245819	NM_005639.1	intron	G	A	1.196	0.04614	G	A	G	A
	rs1268463	NM_005639.1	intron	A	C	1.195	0.04617	A	C	A	C
	rs1569033	NM_005639.1	flanking_5UTR	A	G	0.6939	0.04794	A	G	A	G
	rs10735416	NM_005639.1	flanking_5UTR	T	C	1.193	0.04705	T	C		
	rs1245810	NM_005639.1	intron	A	C	1.2	0.04139	A	C		
	rs1245840	NM_005639.1	intron	C	T	1.193	0.04781	C	T		
SYT11 synaptotagmin XI	rs822519	NM_152280.2	intron	G	A	1.568	0.007779	G	A		
TCF12 transcription factor 12	rs10518890	NM_003205.3	intron	T	C	1.304	0.03998	T	C	A	G
TFAP2B transcription factor AP-2 beta (activating enhancer binding protein 2 beta)	rs11970011	NM_003221.2	flanking_3UTR	C	T	0.8437	0.03772	C	T	G	A
	rs2143081	NM_003221.2	flanking_5UTR	C	T	0.8263	0.02005	C	T	G	A
	rs2635727	NM_003221.2	flanking_3UTR	T	C	1.248	0.01621	T	C	A	G
	rs2817352	NM_003221.2	flanking_3UTR	A	G	1.27	0.03366	A	G	A	G
	rs4715209	NM_003221.2	flanking_3UTR	C	T	1.263	0.03789	C	T	G	A
	rs1178063	NM_003221.2	flanking_3UTR	T	C	0.8408	0.0342	T	C		
	rs1178065	NM_003221.2	flanking_3UTR	C	T	1.27	0.03366	C	T		
	rs2636900	NM_003221.2	flanking_3UTR	A	G	1.279	0.0286	A	G		
	rs2817334	NM_003221.2	flanking_3UTR	C	A	1.268	0.0346	C	A		
	rs2817351	NM_003221.2	flanking_3UTR	G	A	0.8411	0.03421	G	A		
	rs3846904	NM_003221.2	flanking_3UTR	A	G	0.8437	0.03772	A	G		
	rs4605881	NM_003221.2	flanking_3UTR	G	A	1.193	0.04987	G	A		
	rs571503	NM_003221.2	flanking_3UTR	A	G	0.8516	0.04929	A	G		
TIMP2 TIMP metallopeptidase inhibitor 2	rs7502935	NM_003255.4	intron	A	G	0.8283	0.04309	A	G		
TIMP3 TIMP metallopeptidase inhibitor 3	rs762886	NM_000362.4	flanking_5UTR	A	G	1.264	0.01733	A	G	A	G
	rs9862	NM_000362.4	coding	C	T	1.183	0.04175	C	T	A	G
	rs137487	NM_000362.4	flanking_3UTR	A	G	0.8342	0.03265	A	G		
TMED4 transmembrane emp24 protein transport domain containing 4	rs217361	NM_182547.2	intron	A	G	0.8088	0.01461	A	G	A	G
TMEM109 transmembrane protein 109	rs555835	NM_024092.1	coding	T	C	1.205	0.03234	T	C	A	G
URI1 URI1, prefoldin-like chaperone	rs11671255	NM_134447.1	flanking_3UTR	A	G	1.395	0.01216	A	G	A	G
	rs1006584	NM_134447.1	flanking_3UTR	A	G	1.189	0.03696	A	G		
USP32 ubiquitin specific peptidase 32	rs8079220	NM_032582.3	intron	T	C	1.264	0.02085	T	C	G	A
VEZF1 vascular endothelial zinc finger 1	rs9904523	NM_007146.2	flanking_3UTR	T	C	0.7591	0.03904	T	C		

VLDLR very low density lipoprotein receptor	rs10812407	NM_001018056.1	flanking_3UTR	T	C	0.8464	0.04736	T	C	G	A
	rs10967188	NM_001018056.1	flanking_5UTR	G	T	0.7074	0.002602	G	T	C	A
	rs11999258	NM_001018056.1	flanking_5UTR	A	G	1.508	0.03144	A	G	A	G
	rs1355640	NM_001018056.1	flanking_5UTR	A	G	1.364	0.04968	A	G	A	G
	rs1996055	NM_001018056.1	flanking_5UTR	G	A	1.334	0.02675	G	A	G	A
	rs4741730	NM_001018056.1	flanking_5UTR	T	G	1.262	0.01401	T	G	A	C
	rs1006575	NM_001018056.1	flanking_5UTR	A	G	1.211	0.04115	A	G		
	rs10812382	NM_001018056.1	intron	A	G	0.8163	0.01662	A	G		
	rs10967081	NM_001018056.1	flanking_5UTR	A	G	1.403	0.04761	A	G		
	rs1551411	NM_001018056.1	intron	T	C	1.211	0.03663	T	C		
	rs2242104	NM_003383.3	intron	A	G	0.8165	0.02136	A	G		
VWF von Willebrand factor	rs1063856	NM_000552.2	coding	G	A	0.8046	0.01716	G	A	G	A
	rs2239140	NM_000552.2	intron	A	G	1.197	0.03923	A	G	A	G
	rs11610629	NM_000552.2	intron	C	A	0.8052	0.03208	C	A		
	rs11611917	NM_000552.2	intron	A	G	0.8023	0.02936	A	G		
YWHAZ tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide	rs13254653	NM_003406.2	flanking_3UTR	T	C	1.253	0.008663	T	C	A	G
	rs2135121	NM_003406.2	flanking_3UTR	T	C	0.7688	0.01411	T	C	A	G
	rs3134373	NM_003406.2	flanking_3UTR	A	G	0.8141	0.0165	A	G	A	G
	rs4734487	NM_003406.2	flanking_3UTR	G	A	1.219	0.01835	G	A	A	G
	rs4734497	NM_003406.2	intron	C	T	1.194	0.04443	C	T	G	A
	rs7816400	NM_003406.2	flanking_3UTR	A	G	1.201	0.04223	A	G	G	A
	rs3134369	NM_003406.2	flanking_3UTR	G	A	0.7474	0.000535	G	A		
	rs4734494	NM_003406.2	flanking_3UTR	C	T	1.225	0.02509	C	T		

Table S5: Epistasis testing of top candidate genes for alcoholism (n=11). Best p-value SNPs from discovery GWAS1 tested in independent cohort 2. The top epistatic interactions in GWAS2 are depicted in bold and underlined and highlighted in blue. As a caveat, the p-value was not corrected for multiple comparisons. The corresponding genes merit future follow-up work to elucidate the biological and pathophysiological relevance of their interactions.

	DRD2	GFAP	GNAI1	GRM3	MBP	MOBP	MOG	RXRG	SNCA	SYT1	TIMP2
DRD2		0.4546	0.9179	0.5273	0.7124	0.7963	0.1758	0.8905	0.6744	**0.01796**	0.6186
GFAP	0.4546		0.5293	0.08965	0.773	0.4843	0.3624	0.6137	0.9092	0.2335	0.06536
GNAI1	0.9179	0.5293		0.1319	0.07375	0.1498	0.4557	0.09469	0.7188	0.1391	0.05482
GRM3	0.5273	0.08965	0.1319		0.1263	0.1912	0.3744	0.5735	0.1586	0.6271	0.937
MBP	0.7124	0.773	0.07375	0.1263		0.7845	0.4333	0.7912	0.8075	0.5515	0.3699
MOBP	0.7963	0.4843	0.1498	0.1912	0.7845		0.8991	0.3103	0.8511	0.2651	**0.01883**
MOG	0.1758	0.3624	0.4557	0.3744	0.4333	0.8991		0.4422	0.3425	0.4614	0.8411
RXRG	0.8905	0.6137	0.09469	0.5735	0.7912	0.3103	0.4422		**0.03684**	0.2641	0.4612
SNCA	0.6744	0.9092	0.7188	0.1586	0.8075	0.8511	0.3425	**0.03684**		0.9097	0.3429
SYT1	**0.01796**	0.2335	0.1391	0.6271	0.5515	0.2651	0.4614	0.2641	0.9097		0.6851
TIMP2	0.6186	0.06536	0.05482	0.937	0.3699	**0.01883**	0.8411	0.4612	0.3429	0.6851	

Table S6. Genotype for SNCA SNP RS17015888 in test cohorts 3 (alcohol dependence) and 4 (alcohol abuse). The G allele is the affected allele from the discovery GWAS.

	SNCA Genotype		%
Control	G/G	678/1261	**53.8%**
	G/A	467/1261	37.0%
	A/A	116/1261	9.2%
Alcohol Dependent	G/G	1724/2765	**62.4%**
	G/A	808/2765	29.2%
	A/A	233/2765	8.4%
Alcohol Abuse	G/G	366/600	**61.0%**
	G/A	175/600	29.2%
	A/A	59/600	9.8%

Table S7. Genetic risk prediction score (GRPS) panels for A. bipolar disorder and B. schizophrenia, tested in alcoholism (Figure 6). The affected allele was assigned based on the original bipolar and schizophrenia datasets.

A. GRPS-BP		
Gene	**SNP**	**Affected allele**
A2BP1	rs11077135	A
	rs8046170	A
ALDH1A1	rs7873724	G
APP	rs2829984	G
ARNTL	rs11022781	T
ATXN1	rs909786	C
CACNA1A	rs16016	A
CAMK2A	rs4958469	A
CD44	rs353615	G
CDH13	rs931408	C
CUGBP2	rs932918	C
	rs2378991	T
DAPK1	rs3124236	T
DCLK1	rs1171062	G
	rs7994174	C
	rs10492555	T
DIAPH1	rs11954658	C
DISC1	rs9431714	G
	rs821577	A
GNAI1	rs2523189	T
HTR2A	rs977003	C
KCND2	rs10268591	G
KLF12	rs4885151	G
MBNL2	rs6491345	G
MBP	rs12967023	A
MYT1L	rs17039396	G
NCAM1	rs4366519	T
NDUFS2	rs11421	C
	rs5085	G
NR3C1	rs17209251	A
	rs2918417	C
NRCAM	rs3763461	G
OPRM1	rs2010884	T
PRKCE	rs4557033	A

	rs2711293	T
PTN	rs6977749	G
PTPRM	rs4121619	G
PTPRT	rs1016071	G
RYR3	rs744776	A
SYN3	rs3788467	C
TIAM1	rs845945	A
TSHZ2	rs169270	A

B. GRPS-SZ

Gene	SNP	Affected allele
ADCYAP1	rs881897	C
	rs8091765	A
	rs789046	A
ANK3	rs4948256	A
	rs7087489	A
	rs10761507	C
	rs11813307	G
	rs12767186	G
	rs4948422	C
BDNF	rs6265	C
	rs11030104	A
	rs11030119	A
	rs11030182	T
	rs11030066	T
CNR1	rs9362473	T
COMT	rs1544325	A
DISC1	rs9728261	T

	rs9782927	G
	rs12087592	C
DRD2	rs12791990	C
	rs4648318	C
DTNBP1	rs1539422	T
	rs1935784	A
FABP7	rs9490546	G
GABRB3	rs8037461	C
	rs8027455	A
	rs1435831	C
	rs8025575	G
	rs11161309	C
	rs11854349	G
	rs4906680	C
	rs500951	T
GNB1L	rs17745302	T
	rs13057910	T
GRIA1	rs17115481	A
	rs4958687	A
	rs4285285	A
	rs10515671	A
	rs2199123	A
	rs12656429	C
	rs160163	T
	rs17113267	A
	rs2973138	T

	rs6877008	C
	rs1493383	T
	rs10039253	C
	rs17115298	G
	rs4958560	T
	rs4398624	T
	rs1542485	G
	rs1422337	A
	rs17568427	G
	rs11951398	T
	rs286969	A
	rs7702336	A
GRIA4	rs17104589	A
GRIN2B	rs4363703	T
	rs11055597	T
	rs11055930	T
	rs2300268	A
	rs7306014	G
	rs2300252	C
	rs10772769	A
	rs10845826	A
GRM5	rs992259	A
	rs7396702	C
	rs12362135	G
	rs10831155	C
	rs982010	C
	rs308765	T

	rs541279	G
	rs16914531	T
HINT1	rs7716702	A
	rs1363696	G
HTR2A	rs1805055	T
KALRN	rs3772753	T
	rs16835783	A
	rs3772790	G
	rs13087377	G
	rs7621976	T
KIF2A	rs6864793	C
	rs414568	T
MBP	rs12959006	T
	rs12956305	C
	rs3900176	C
MOBP	rs1768141	G
	rs538867	T
	rs2233204	C
	rs9835143	T
NCAM1	rs1945101	G
	rs10891375	G
	rs1006826	C
	rs2212450	C
	rs7105462	G
NDUFV2	rs1784773	T
	rs1573321	T
	rs16954628	G

	rs11081446	A
NR4A2	rs12465886	G
	rs2691775	A
	rs2691787	G
	rs10175502	C
NRCAM	rs4142284	C
	rs1990713	C
	rs2284284	G
NRG1	rs12541516	C
	rs954009	T
	rs16879809	T
	rs4535704	A
	rs4035323	A
	rs7818821	G
	rs11775675	C
	rs10954887	G
	rs17642273	A
	rs3847131	G
	rs1623372	A
PDE4B	rs1321172	C
	rs6700403	T
	rs17128076	G
	rs524770	A
	rs4492586	C
PRKCA	rs16959714	T
	rs4791022	T
RELN	rs2711865	G

	rs2245617	T
RGS4	rs4657235	T
	rs951438	G
SLC1A2	rs12273644	C
	rs17379710	T
	rs736374	A
SNAP25	rs6032783	T
	rs362585	G
	rs362574	A
SYN2	rs2960421	A
	rs3755724	C
TCF4	rs17594665	A
	rs10401120	T
	rs1377242	C
	rs4800933	C
	rs10164195	A
	rs2588477	A
	rs9949821	A
	rs17089826	C
TNIK	rs360414	C
	rs11927009	T
	rs902952	A
	rs9846083	C
	rs1353021	A
	rs6788198	G
	rs11916004	A
	rs9844740	C

Figure S1. Gene size vs. CFG score. No enrichment of larger –size genes by CFG. Top candidate genes for alcoholism (n=135).

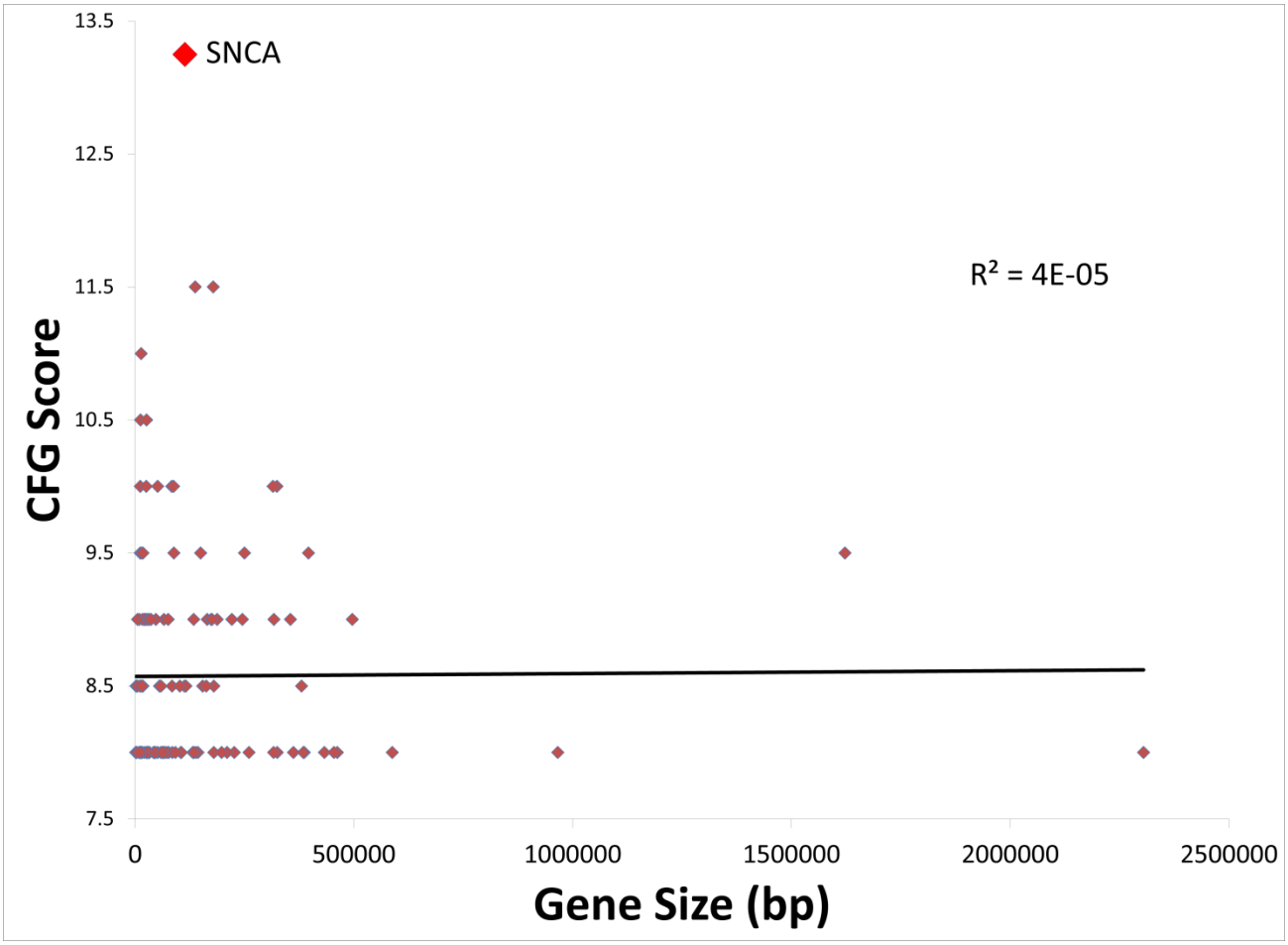
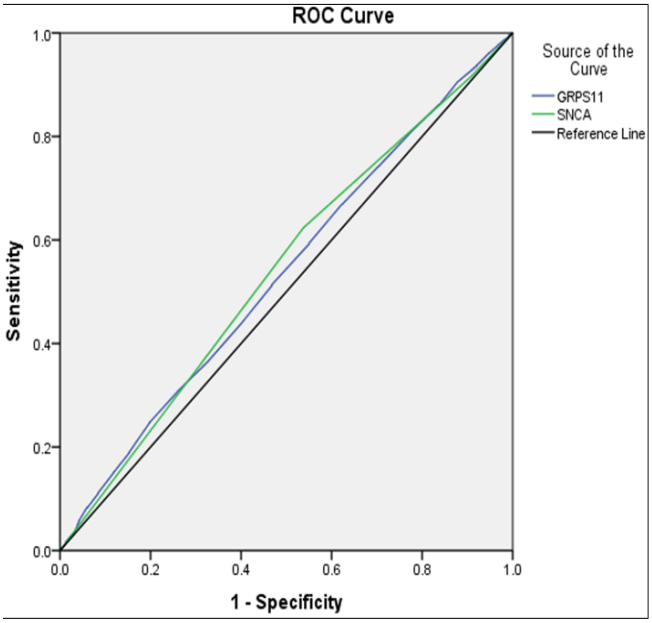
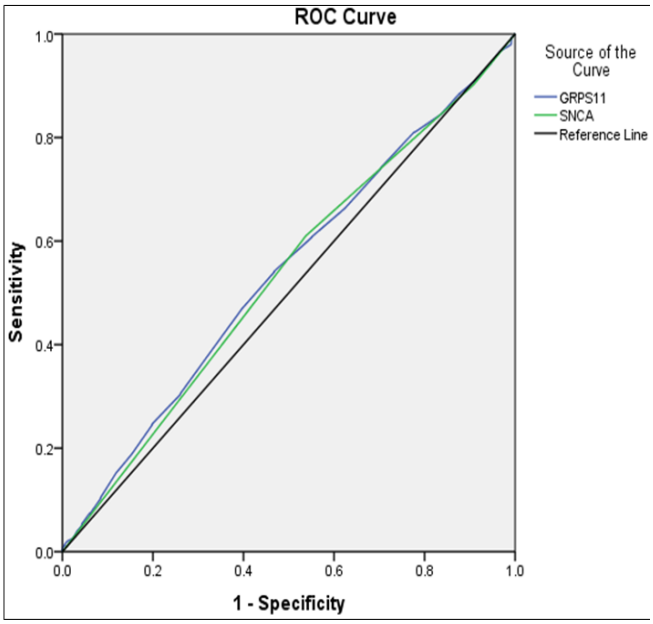


Figure S2. Genetic Risk Prediction ROC curves



Cohort 3
Alcohol Dependence

Test Result Variable(s)	Area	Std. Error	Asymptotic Significance (p-value)
SNCA	.541	.010	**.000**
GRPS11	.534	.010	**.001**



Cohort 4
Alcohol Abuse

Test Result Variable(s)	Area	Std. Error	Asymptotic Significance (p-value)
SNCA	.531	.014	**.028**
GRPS11	.537	.014	**.010**