

## ARTICLE

## Open Access

# Genome-wide association study identifies glutamate ionotropic receptor *GRIA4* as a risk gene for comorbid nicotine dependence and major depression

Hang Zhou<sup>1</sup>, Zhongshan Cheng<sup>1</sup>, Nicholas Bass<sup>2</sup>, John H. Krystal<sup>1,3,4</sup>, Lindsay A. Farrer<sup>5,6,7,8,9</sup>, Henry R. Kranzler<sup>10,11</sup> and Joel Gelernter<sup>1,3,12,13</sup>

## Abstract

Smoking and major depression frequently co-occur, at least in part due to shared genetic risk. However, the nature of the shared genetic basis is poorly understood. To detect genetic risk variants for comorbid nicotine dependence (ND) and major depression (MD), we conducted genome-wide association study (GWAS) in two samples of African-American participants (Yale-Penn 1 and 2) using linear mixed model, followed by meta-analysis. 3724 nicotine-exposed subjects were analyzed: 2596 from Yale-Penn-1 and 1128 from Yale-Penn-2. Continuous measures (*Fagerström Test for Nicotine Dependence* (FTND) scores and *DSM-IV* MD criteria) rather than disorder status were used to maximize the power of the GWAS. Genotypes were ascertained using the Illumina HumanOmni1-Quad array (Yale-Penn-1 sample) or the Illumina HumanCore Exome array (Yale-Penn-2 sample), followed by imputation based on the 1000 Genomes reference panel. An intronic variant at the *GRIA4* locus, rs68081839, was significantly associated with ND–MD comorbidity ( $\beta = 0.69$  [95% CI, 0.43–0.89],  $P = 1.53 \times 10^{-8}$ ). *GRIA4* encodes an AMPA-sensitive glutamate receptor that mediates fast excitatory synaptic transmission and neuroplasticity. Conditional analyses revealed that the association was explained jointly by both traits. Enrichment analysis showed that the top risk genes and genes co-expressed with *GRIA4* are enriched in cell adhesion, calcium ion binding, and synapses. They also have enriched expression in the brain and they have been implicated in the risk for other neuropsychiatric disorders. Further research is needed to determine the replicability of these findings and to identify the biological mechanisms through which genetic risk for each condition is conveyed.

## Introduction

Substance use is highly associated with other psychiatric illnesses<sup>1–5</sup>. For instance, substance use disorders (SUDs) and major depression (MD) are highly comorbid in the general population<sup>2,3</sup>, and strong associations between alcohol misuse and other psychiatric disorders were

observed in a U.S. Army cohort<sup>1</sup>. Clinical outcome is worse in the patients with comorbid psychiatric disorders and SUDs than in each disorder separately<sup>6</sup>. The causes of this comorbidity are poorly understood, and a better understanding of the causal relationship and etiology may provide opportunities for risk mitigation. In recent years, genetic associations (pleiotropy) between specific substance use and psychiatric disorders have been investigated by genome-wide approaches<sup>7–9</sup> and some specific genome-wide significant (GWS) loci that affect SUD/psychiatric comorbidity have been identified in our previous study<sup>9</sup>.


Correspondence: Joel Gelernter ([joel.gelernter@yale.edu](mailto:joel.gelernter@yale.edu))

<sup>1</sup>Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA

<sup>2</sup>Molecular Psychiatry Laboratory, Division of Psychiatry, University College London, London, UK

Full list of author information is available at the end of the article.

© The Author(s) 2018

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

The association between cigarette smoking and MD is a particularly well-studied comorbidity, with several epidemiological studies showing co-occurrence<sup>3,10–12</sup>. Smoking initiation, daily smoking, persistent daily smoking, and heavy smoking were significantly associated with increased risk of MD, and the association also applies to nicotine dependence (ND)<sup>13–15</sup>. Different hypotheses have been proposed to explain the association. It has been suggested that depression may result from the neuropharmacological effects of nicotine or nicotine withdrawal<sup>12,16,17</sup>, or alternatively, that depression may cause smoking as an attempt at self-medication of negative feelings<sup>18,19</sup> or that there are bidirectional causal effects linking smoking and depression<sup>15</sup>. Genetic risk variants for ND (as well as smoking-related behaviors) and MD (as well as depressive symptoms) have been separately identified in large cohorts by genome-wide association study (GWAS)<sup>20–25</sup>. Common risk factors or shared etiology for smoking and depression have also been suggested<sup>26,27</sup>, and genetic factors that predispose to both smoking and MD were also suggested in a study of female twins<sup>28</sup>.

To detect shared genetic variants that predispose to comorbid ND and MD, we conducted GWAS and meta-analysis on criterion counts comprised of *Fagerström Test for Nicotine Dependence* (FTND) scores and *DSM-IV* MD criteria in two African-American samples. A variant in *GRIA4*, the gene that codes for the subunit 4 of the  $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) glutamate receptor, showed genome-wide significance for association with ND–MD comorbidity.

## Materials and methods

### Participants and diagnostic procedures

A total of 4944 African American (AA) subjects were recruited for the Yale-Penn genetics of substance dependence study from 2000 to 2013, as previously described<sup>22,29</sup>. The subjects were grouped into two sets, Yale-Penn-1 (3227) and Yale-Penn-2 (1717), based on their epoch of recruitment and the genotyping platforms used. All subjects provided written informed consent, and certificates of confidentiality were obtained from National Institute on Drug Abuse (NIDA) and National Institute on Alcohol Abuse and Alcoholism (NIAAA). All subjects were interviewed using the Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA)<sup>30</sup>. Lifetime FTND scores<sup>31</sup> and criterion counts for MD from the *DSM-IV* (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition)<sup>32</sup> were derived. Six items were assessed for the FTND, generating scores from 0 to 10 (the higher the score, the more severe the nicotine use), and nine criteria were assessed for MD, generating scores from 0 to 9. We scaled the FTND scores uniformly using the same range as for MD criteria so as to weight them comparably for the GWAS<sup>9</sup>. Then, the comorbid

(summed) criterion counts (ranging from 0 to 18) were treated as the outcomes, representing the overall severity of comorbidity. Subjects who were not exposed to tobacco (i.e., who answered “no” to the question, *have you ever tried any form of tobacco?*) were excluded from the 4994 participants, leaving 3724 eligible subjects, 2596 from Yale-Penn-1 and 1128 from Yale-Penn-2.

### Genotyping, quality control, and imputation

The Yale-Penn-1 sample was genotyped using the Illumina HumanOmni1-Quad array containing ~988,000 SNPs. The Yale-Penn-2 sample was genotyped using the Illumina HumanCore Exome array containing ~266,000 exonic SNPs and ~240,000 tagging SNPs for genome-wide imputation. Individuals and SNPs with genotype call rates <98%, and SNPs with minor allele frequency (MAF) <1% were removed from downstream analyses. Yale-Penn-1 and Yale-Penn-2 data were analyzed separately.

To correct any misclassification from self-reported race, we conducted principal component (PC) analysis<sup>33</sup> on SNPs common to both the two Yale-Penn genotype datasets and the 1000 Genomes phase 3 reference panel which contains African, American, Asian, and European populations<sup>34</sup>. SNPs were pruned based on LD ( $r^2 < 0.2$ ) using PLINK<sup>35</sup>. Yale-Penn subjects were clustered into different groups by the Euclidean distances to the reference populations (based on the first 3 PCs). For this study, subjects that clustered with non-African populations were removed from the downstream analyses. We then conducted a second PC analysis within the remaining Yale-Penn subjects and removed any outliers beyond three standard deviations from the mean. The first 10 PCs were used in all subsequent analyses to correct for residual population stratification.

We imputed additional single nucleotide variants (SNVs) using Minimac3 implemented in Michigan Imputation Server (<https://imputationserver.sph.umich.edu/index.html>)<sup>36</sup> based on the 1000 Genomes phase 3 reference panel<sup>34</sup>. SNVs with Hardy–Weinberg equilibrium  $P$  values <10<sup>-5</sup>, imputation accuracy <0.8, or MAF <1% were excluded from downstream analyses. In the Yale-Penn-1 sample, 14,778,319 SNVs were included in the association analyses; in the Yale-Penn-2 sample, 9,658,251 SNVs were analyzed. 9,520,174 SNPs common in two samples were meta-analyzed.

### Phenotype imputation

ND scores or the set of MD criteria were incomplete in a small proportion of the sample: 4.7% (121) of the Yale-Penn-1 and 4.6% (52) of the Yale-Penn-2 subjects. To address this without the power reduction that would result from simply excluding these subjects, we used PHENIX<sup>37</sup>, a variational Bayesian method fitting in a Bayesian multiple-phenotype mixed model, to impute the

**Table 1 Demographic characteristics of the samples**

	Yale-Penn-1	Yale-Penn-2	Total
Sample size (female %)	3227 (47.1)	1717 (41.7)	4944 (45.2)
GWAS			
Tobacco exposed (female %)	2596 (44.3)	1128 (33.2)	3724 (40.9)
Age, mean (SD), years	41.5 (8.2)	42.0 (10.4)	41.7 (8.9)
Subjects with partial missing (%)	121 (4.7)	52 (4.6)	173 (4.6)
–with partial missing of ND	20	11	31
–with partial missing of MD	101	41	142
Imputation correlation of ND	0.74	0.71	
Imputation correlation of MD	0.86	0.86	
Median (IQR) of ND+MD	7.9 (4.5–12.5)	8.8 (3.8–12.4)	8.1 (4.5–12.5)
Median (IQR) of ND	5 (3–6)	4 (3–6)	5 (3–6)
Median (IQR) of MD	5 (0–8)	6 (0–8)	5 (0–8)
Correlation between ND and MD	0.15	0.20	0.16

ND nicotine dependence, MD major depression, IQR interquartile range

missing criteria. ND and MD were imputed separately in the two datasets, using the correlation matrix of the subjects derived from genome-wide efficient mixed model association (GEMMA)<sup>38</sup>.

### Statistical analysis

We performed association tests for the ND+MD criterion counts (ranging from 0 to 18). All SNVs, both genotyped and imputed, were tested using a linear mixed model (GEMMA), adjusted by age, sex, and the first 10 PCs. Analyses were performed separately within each dataset. The association results were meta-analyzed across the two datasets, using the inverse variance method implemented in the program METAL<sup>39</sup>. Regional associations were plotted using LocusZoom<sup>40</sup>.

### Functional annotation and enrichment analysis

Functional annotations for the top variants and genes were explored from the literature and from expression databases, including Gene-Tissue Expression (GTEx, <https://www.gtexportal.org/home/>) for gene-tissue expression<sup>41</sup> and BrainSpan (<http://www.brainspan.org/>) for information regarding the transcriptome across human brain development<sup>42</sup>. Genes co-expressed with the target gene were identified using COXPRESdb v6.0 ([\[coxpresdb.jp/\]\(http://coxpresdb.jp/\)\)<sup>43</sup>, a coexpression database of DNA-microarray and RNAseq-based expression data. Disease enrichment among the co-expressed genes was assessed using WebGestalt \(<http://www.webgestalt.org/option.php>\)<sup>44</sup>, a functional enrichment analysis web tool. Gene ontology \(GO\) enrichment of the genes mapped to the top SNVs \( \$P\$ -value  \$< 1 \times 10^{-4}\$  in either dataset or meta-analysis\) was analyzed using the web-based gene set analysis tool Gorilla \(<http://cbl-gorilla.cs.technion.ac.il/>\)<sup>45</sup>. Terms with false discovery rate \(FDR\)  \$< 0.05\$  were considered to be significantly enriched.](http://</a></p>
</div>
<div data-bbox=)

### Results

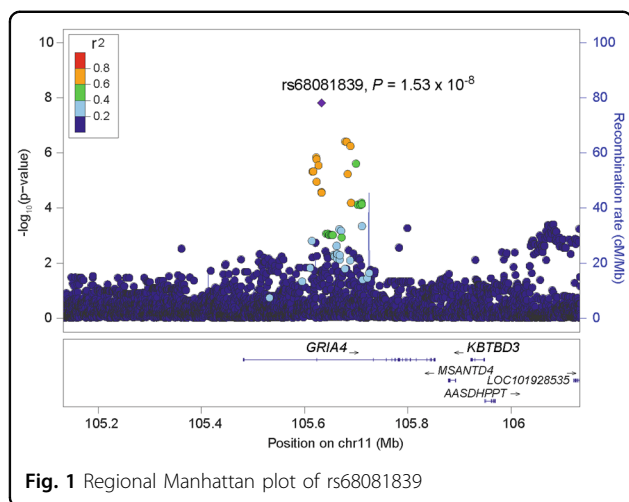
In total, 3724 AA subjects (mean age, 42 years [SD, 8.9]; 1523 women [40.9%]) were included in the analysis (2596 from Yale-Penn-1 and 1128 from Yale-Penn-2). There were 173 subjects (4.6%) with partially missing ND or MD criteria, which were imputed. Among subjects with imputed phenotype data, the average number of items that needed to be imputed was 1.2 for ND and 2.1 for MD. For ND, the imputation correlation was 0.74 for Yale-Penn-1 and 0.71 for Yale-Penn-2; for MD, the imputation correlation (between imputed phenotypes and their true hidden values) was 0.86 for both datasets. The median comorbid criterion count was 8.1 (interquartile range [IQR], 4.5–12.5) (Table 1). The distributions of comorbid criterion counts are shown in Figure S1.

### Genome-wide significant association

GWAS was performed in each dataset, followed by meta-analysis (Figure S2; SNVs with  $P$ -values  $< 1 \times 10^{-4}$  (in either individual sample or the meta-analysis) are listed in Table S1). No GWS signals were detected in either sample analyzed individually. In the meta-analysis, a significant association was detected in *GRIA4* (rs68081839, a single nucleotide deletion, –/T; the frequency of the risk allele (–) is 0.68, beta coefficient  $[\beta] = 0.69$  [95% CI, 0.43–0.89],  $P = 1.53 \times 10^{-8}$ , Fig. 1). This variant was well imputed in both Yale-Penn-1 (INFO = 0.91) and Yale-Penn-2 (INFO = 0.87) samples. rs68081839 was nominally associated in both the Yale-Penn-1 ( $P = 1.17 \times 10^{-5}$ ) and Yale-Penn-2 ( $P = 2.95 \times 10^{-4}$ ) samples.

### Conditional analyses of rs68081839

We tested the association of rs68081839 with FTND scores (controlling for MD criterion counts) and MD criterion counts (controlling for FTND scores) to determine whether the association was being driven by a single disorder. Both traits were nominally associated with rs68081839 ( $P = 7.11 \times 10^{-3}$  for ND and  $P = 7.34 \times 10^{-6}$  for MD), indicating an additive or synergistic association for ND–MD comorbidity: i.e., the risk allele contributes to the risk of each trait taken individually. To test whether the association effect was age- or sex-related, we split the

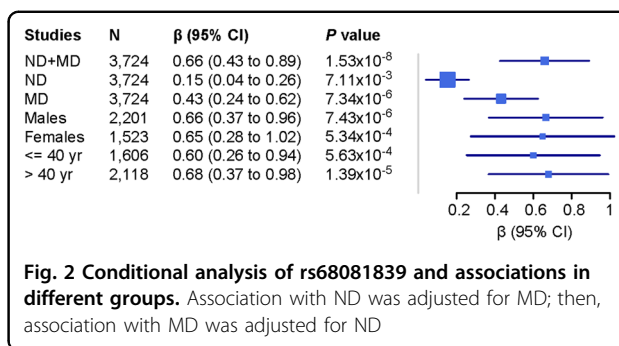


sample into older (>40 years old) and younger groups ( $\leq 40$ ), adjusting for sex and 10 PCs, and into male and female groups, adjusting for age and 10 PCs. Similar associations between rs68081839 and ND+MD were observed with each of these approaches, indicating a consistent effect in all of the subgroups (Fig. 2). We also tested whether rs68081839 has pleiotropic effects with other substance dependence traits including alcohol, cocaine, marijuana, and opioids, and found no evidence for association (all  $P$ -values  $> 0.5$ ).

#### Functional assessment of GRIA4

*GRIA4* codes for subunit 4 of the AMPA glutamate receptor and it is implicated in glutamate signaling and neuroplasticity<sup>46,47</sup>. It is involved in several KEGG pathways (e.g., amphetamine addiction, nicotine addiction, the cAMP signaling pathway, neuroactive ligand–receptor interaction, glutamatergic synapses, dopaminergic synapses<sup>48</sup>). We explored the gene expression profiles of *GRIA4* in different tissues from GTEx<sup>41</sup>, where it is shown to be widespread and primarily expressed in human brain (Figure S3). We then evaluated the spatio-temporal transcriptome of *GRIA4* in human brain<sup>42</sup>. High expression of *GRIA4* across several brain regions was observed in adulthood, increasing from the early fetal periods (Figure S4). The consistent high level of expression in brain supports the functional relevance of *GRIA4* in psychiatric traits.

To investigate the functional relevance of *GRIA4* further, the top 100 genes co-expressed with *GRIA4* were derived from COXPRESSdb<sup>43</sup> (Table S2). These include *NLGN1*, *KCND2*, *ELAVL4*, *NXP1*, *GRM5*, and *GABRB1*. We assessed the disease enrichment of the co-expressed genes using web-based tool WebGestalt<sup>44</sup> and found that mental disorders, depression, bipolar disorder, and anxiety disorder were significantly enriched (FDR  $< 0.05$ , Table S3).



#### Gene ontology enrichment analysis

The top SNVs ( $P < 1 \times 10^{-4}$ ) were mapped to 223 genes (Table S1). GO enrichment analysis using the GOrilla web tool<sup>45</sup> showed these genes to be enriched for cell adhesion, calcium ion binding, synapse, and plasma membrane (Table 2). We also tested the tissue expression enrichment using DAVID<sup>49,50</sup>; it showed significant enrichment in the brain ( $P = 5.44 \times 10^{-6}$ , FDR =  $6.29 \times 10^{-3}$ ). Disease enrichment analysis using WebGestalt showed that the top genes are enriched in various psychiatric disorders including bipolar disorder, anxiety disorder, depression, and substance-related disorders (Table S4). Taken together, the enrichments in signal transduction, synapse, and mental disorders support the interpretation that the polygenic risk of ND+MD is related to neural functions.

#### Discussion

ND and MD are among the most common psychiatric disorders worldwide and are associated with substantial morbidity and mortality<sup>51</sup>. The association between ND (as well as smoking) and MD has been well established, and both shared and distinct etiologies have been postulated. GWAS have identified risk or protective variants for ND and MD individually. To our knowledge, this is the first study of the shared genetic risks for ND and MD comorbidity. To accomplish this, we employed a dimensional approach using our phenotype data collected using the SSADDA. We found one SNP to be significantly associated with ND+MD comorbidity ( $\beta = 0.69$  [95% CI, 0.43–0.89],  $P = 1.53 \times 10^{-8}$ , Fig. 1). rs68081839 is a single nucleotide deletion in the *GRIA4* gene. Conditional analyses showed that the association was not driven by ND or MD alone; instead, there is an additive or synergistic effect of ND and MD. In our dataset, the contribution from MD is greater than that from ND (Fig. 2). There was no evidence of pleiotropy with other substance dependence traits (based on direct testing for association).

*GRIA4* (glutamate ionotropic receptor AMPA type subunit 4)—also referred to as GluR-D or GluR4—is a member of the AMPA-selective glutamate receptor family (AMPA receptors). AMPARs are expressed ubiquitously in the central nervous system and are the predominant

**Table 2 Gene ontology enrichment of the genes mapped to top SNVs ( $P < 1 \times 10^{-4}$ )**

Category	Term	P	FDR
Biological process	GO:0098742—cell–cell adhesion via plasma-membrane adhesion molecules	$1.81 \times 10^{-12}$	$2.67 \times 10^{-8}$
	GO:0007156—homophilic cell adhesion via plasma membrane adhesion molecules	$6.14 \times 10^{-11}$	$4.52 \times 10^{-7}$
	GO:0098609—cell–cell adhesion	$4.14 \times 10^{-8}$	$2.03 \times 10^{-4}$
	GO:0007155—cell adhesion	$5.98 \times 10^{-8}$	$2.20 \times 10^{-4}$
	GO:0022610—biological adhesion	$6.52 \times 10^{-8}$	$1.92 \times 10^{-4}$
Molecular function	GO:0005509—calcium ion binding	$1.43 \times 10^{-6}$	$6.29 \times 10^{-3}$
Cellular component	GO:0045202—synapse	$1.60 \times 10^{-6}$	$2.89 \times 10^{-3}$
	GO:0044459—plasma membrane part	$2.67 \times 10^{-6}$	$2.42 \times 10^{-3}$
	GO:0005887—integral component of plasma membrane	$2.05 \times 10^{-5}$	$1.24 \times 10^{-2}$
	GO:0031226—intrinsic component of plasma membrane	$2.23 \times 10^{-5}$	$1.01 \times 10^{-2}$
	GO:0005886—plasma membrane	$8.50 \times 10^{-5}$	$3.08 \times 10^{-2}$
	GO:0031224—intrinsic component of membrane	$8.92 \times 10^{-5}$	$2.69 \times 10^{-2}$

excitatory neurotransmitter receptors in the mammalian brain. They are localized at the postsynaptic membrane and are essential for synaptic plasticity<sup>52</sup>. The most thoroughly characterized examples of synaptic plasticity are long-term potentiation (LTP) and long-term depression (LTD), widely believed to be the cellular basis of learning and memory<sup>53</sup>. Further studies have shown that LTP and LTD participate in pathological processes such as Alzheimer's disease, schizophrenia, and addiction<sup>52,54</sup>. *GRIA4* functions as a ligand-gated ion channel in the central nervous system and plays an important role in fast excitatory synaptic transmission<sup>46</sup>. Expression of *GRIA4* is sufficient to alter the signaling requirements for LTP during a critical period of synapse development<sup>47</sup>, and the membrane proximal region of *GRIA4*, needed for receptor trafficking and synaptic plasticity, is essential for long-term fear memory formation<sup>55</sup>.

Changes in *GRIA4* expression have been associated with both depression and stress. Postmortem studies showed *GRIA4* upregulation in depressed patients. Expression of *GRIA4* in Brodmann area 10 and amygdala was increased in subjects who died by suicide during an episode of MD compared to subjects who died by suicide without depression, or controls who died suddenly from other causes and had no history of suicidal behavior<sup>56</sup>. Higher expression of *GRIA4* in the dorsolateral prefrontal cortex in female patients with MD than that in female controls has been reported<sup>57</sup>. In relation to stress, *Gria4* was upregulated in the hippocampus in stressed rats, and this could be reversed by the antidepressant drug venlafaxine. *Gria4* expression was also increased by chronic treatment with corticosterone, the major stress hormone<sup>58</sup>. An opposite effect was observed in the ventral (but not

dorsal) hippocampus in rats which were treated by neonatal handling<sup>59</sup>.

Synaptic plasticity is known to play a key role in drug addiction. Indeed, addiction has been conceptualized as a pathological form of learning and memory, as they share synaptic plasticity mechanisms. Synaptic plasticity may contribute to different aspects of addiction, including craving, withdrawal, and relapse<sup>60,61</sup>. Altered expression of *GRIA4* and other glutamatergic genes in postmortem hippocampus was observed after chronic exposure to alcohol or cocaine<sup>62</sup>, perhaps contributing to the development of craving<sup>63</sup>. Studies in mouse models showed that AMPARs and N-methyl-D-aspartate receptors (NMDAR) in the ventral tegmental area (VTA) are involved in behavioral sensitization, thus playing key roles in the development of addiction<sup>64,65</sup>. For example, a single exposure to cocaine in vivo can increase the AMPAR/NMDAR ratio in the VTA, which may be involved in an early stage of drug addiction<sup>66</sup>. Along with addictive substances such as cocaine<sup>64</sup> and morphine<sup>65</sup>, nicotine activates nicotinic acetylcholine receptors (nAChRs) in the VTA to reinforce smoking behavior<sup>67,68</sup>. We therefore speculate that the synaptic plasticity effects of *GRIA4* may explain its contribution to the risk of ND–MD comorbidity.

Besides the significant finding at *GRIA4*, we performed enrichment analyses taking the genes identified by top SNPs ( $P < 1 \times 10^{-4}$ ) as a whole. The enriched GO terms include cell adhesion in biological processes, calcium ion binding in molecular function, and synapse and plasma membrane in cellular component (Table 2). A significant enrichment of tissue expression was reported in DAVID using the same list of genes. Disease-level enrichment is more informative than GO level enrichment, in this

instance: the disease enrichment analyses for the top genes or top coexpressed genes with *GRIA4* shows that the significant enriched disease traits are mainly related to mental disorders (Table S3 and S4). All the reported terms are significant after multiple testing correction ( $FDR < 0.05$ ). In GWAS with limited sample size, it is very common that no significant enrichment can be detected or the enriched terms cannot be linked to the study trait in an obvious way. Here, despite the sample size limitation, we observed consistent GO or disease enrichments using different web-based tools or different gene lists, indicating that the nominally significant findings ( $P < 1 \times 10^{-4}$ ) and the coexpressed genes with *GRIA4* are robustly related to the ND+MD trait.

This study has important limitations including modest sample size and the lack of a replication sample. Further studies to understand the biological mechanisms of the genetic risk loci we identified are also warranted.

In conclusion, we identified variation at *GRIA4*, a gene that codes for an AMPA glutamate receptor subunit, as a genetic risk factor for ND and MD comorbidity. This provides initial evidence that variation in the glutamatergic system may underlie the common etiology of these highly comorbid disorders. Thus, the glutamatergic system may thus be a target for treatment of ND<sup>69</sup>, as it is already for MD<sup>70</sup>; especially so in clinical contexts where the two traits are comorbid.

#### Acknowledgements

We appreciate the work of recruitment and assessment by James Poling, PhD, at Yale University School of Medicine and the APT Foundation; by Roger Weiss, MD, at McLean Hospital; by Kathleen Brady, MD/PhD and Raymond Anton, MD, at the Medical University of South Carolina; and David Oslin, MD at the University of Pennsylvania. Genotyping services for a part of our GWAS were provided by the Center for Inherited Disease Research and Yale University (Center for Genome Analysis), which is fully funded by Federal contract N01-HG-65403 from the NIH to The Johns Hopkins University. We thank Ann Marie Lacobelle, MS, at the VA CT Healthcare Center, and Christa Robinson, BS, at the VA CT Healthcare Center, who provided technical assistance. We thank John Farrell, PhD, Section of Biomedical Genetics, Boston University School of Medicine, who provided database management assistance. This study was supported by National Institutes of Health grants RC2 DA028909, R01 DA12690, R01 DA12849, R01 DA18432, R01 AA11330, R01 AA017535, 2P50-AA012870, VA Connecticut Healthcare Center, Philadelphia VA MIRECCS, and National Center for Post Traumatic Stress Disorder. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

#### Author details

<sup>1</sup>Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA. <sup>2</sup>Molecular Psychiatry Laboratory, Division of Psychiatry, University College London, London, UK. <sup>3</sup>Department of Neuroscience, Yale University School of Medicine, New Haven, CT, USA. <sup>4</sup>Clinical Neurosciences Division, VA National Center for PTSD, VA CT Healthcare System, West Haven, CT, USA. <sup>5</sup>Department of Medicine (Biomedical Genetics), Boston University School of Medicine, Boston, MA, USA. <sup>6</sup>Department of Neurology, Boston University School of Medicine, Boston, MA, USA. <sup>7</sup>Department of Ophthalmology, Boston University School of Medicine, Boston, MA, USA. <sup>8</sup>Department of Genetics and Genomics, Boston University School of Medicine, Boston, MA, USA. <sup>9</sup>Department of Epidemiology and Biostatistics, Boston University School of Public Health, Boston, MA, USA. <sup>10</sup>Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA. <sup>11</sup>VISN 4

MIRECC, Crescenz VA Medical Center, Philadelphia, PA, USA. <sup>12</sup>Department of Genetics, Yale University School of Medicine, New Haven, CT, USA.

<sup>13</sup>Department of Psychiatry, VA CT Healthcare Center, West Haven, CT, USA

#### Conflict of interest

Dr. Kranzler has been a consultant or advisory board member for Indivior and Lundbeck. He is also a member of the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative, which was supported for the last 3 years by AbbVie, Alkermes, Ethypharm, Indivior, Lilly, Lundbeck, Otsuka, Pfizer, Arbor, and Amygdala Neurosciences. Dr. Krystal has stock in ArRETT Neuroscience and Biohaven Pharmaceuticals Medical Sciences, and has stock options in Biohaven Pharmaceuticals Medical Sciences, Blackthorn Therapeutics, and Luc Therapeutics. Dr. Krystal also has patents and inventions 5,447,948, 8,778,979 B2, 14/197,767, 14/306,382, 61/973,961, Y0087.70116US00, and 62/444,552. The other authors have no biomedical financial interests or potential conflicts of interest.

#### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Supplementary Information** accompanies this paper at (<https://doi.org/10.1038/s41398-018-0258-8>).

Received: 11 December 2017 Revised: 21 February 2018 Accepted: 11 May 2018

Published online: 04 October 2018

#### References

- Stein, M. B. et al. Alcohol misuse and co-occurring mental disorders among new soldiers in the U.S. Army. *Alcohol. Clin. Exp. Res.* **41**, 139–148 (2017).
- Edwards, A. C. et al. Adolescent alcohol use is positively associated with later depression in a population-based U.K. cohort. *J. Stud. Alcohol Drugs* **75**, 758–765 (2014).
- Grant, B. F. Comorbidity between DSM-IV drug use disorders and major depression: results of a national survey of adults. *J. Subst. Abuse* **7**, 481–497 (1995).
- Regier, D. A. et al. Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study. *JAMA* **264**, 2511–2518 (1990).
- Kessler, R. C. et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch. Gen. Psychiatry* **51**, 8–19 (1994).
- Volkow, N. D. Substance use disorders in schizophrenia—clinical implications of comorbidity. *Schizophr. Bull.* **35**, 469–472 (2009).
- Edwards, A. C. et al. Genome-wide association study of comorbid depressive syndrome and alcohol dependence. *Psychiatr. Genet.* **22**, 31–41 (2012).
- Andersen, A. M. et al. Polygenic scores for major depressive disorder and risk of alcohol dependence. *JAMA Psychiatry* <https://doi.org/10.1001/jamapsychiatry.2017.2269>. (2017).
- Zhou, H. et al. Genetic risk variants associated with comorbid alcohol dependence and major depression. *JAMA Psychiatry* **74**, 1234–1241 (2017).
- Bakhshai, J., Zvolensky, M. J. & Goodwin, R. D. Cigarette smoking and the onset and persistence of depression among adults in the United States: 1994–2005. *Compr. Psychiatry* **60**, 142–148 (2015).
- Steuber, T. L. & Danner, F. Adolescent smoking and depression: which comes first? *Addict. Behav.* **31**, 133–136 (2006).
- Klungsoyr, O., Nygard, J. F., Sorensen, T. & Sandanger, I. Cigarette smoking and incidence of first depressive episode: an 11-year, population-based follow-up study. *Am. J. Epidemiol.* **163**, 421–432 (2006).
- Hu, M. C., Davies, M. & Kandel, D. B. Epidemiology and correlates of daily smoking and nicotine dependence among young adults in the United States. *Am. J. Public Health* **96**, 299–308 (2006).
- Grant, B. F., Hasin, D. S., Chou, S. P., Stinson, F. S. & Dawson, D. A. Nicotine dependence and psychiatric disorders in the United States: results from the national epidemiologic survey on alcohol and related conditions. *Arch. Gen. Psychiatry* **61**, 1107–1115 (2004).

15. Breslau, N., Kilbey, M. M. & Andreski, P. Nicotine dependence and major depression. New evidence from a prospective investigation. *Arch. Gen. Psychiatry* **50**, 31–35 (1993).
16. Pasco, J. A. et al. Tobacco smoking as a risk factor for major depressive disorder: population-based study. *Br. J. Psychiatry* **193**, 322–326 (2008).
17. Covey, L. S., Glassman, A. H. & Stetner, F. Major depression following smoking cessation. *Am. J. Psychiatry* **154**, 263–265 (1997).
18. Patton, G. C. et al. Is smoking associated with depression and anxiety in teenagers? *Am. J. Public Health* **86**, 225–230 (1996).
19. Crone, M. R. & Reijneveld, S. A. The association of behavioural and emotional problems with tobacco use in adolescence. *Addict. Behav.* **32**, 1692–1698 (2007).
20. Thorgeirsson, T. E. et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* **452**, 638–642 (2008).
21. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat. Genet.* **42**, 441–447 (2010).
22. Gelernter, J. et al. Genome-wide association study of nicotine dependence in American populations: identification of novel risk loci in both African-Americans and European-Americans. *Biol. Psychiatry* **77**, 493–503 (2015).
23. Converge Consortium Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* **523**, 588–591 (2015).
24. Hyde, C. L. et al. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat. Genet.* **48**, 1031–1036 (2016).
25. Okbay, A. et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* **48**, 624–633 (2016).
26. Fergusson, D. M., Lynskey, M. T. & Horwood, L. J. Comorbidity between depressive disorders and nicotine dependence in a cohort of 16-year-olds. *Arch. Gen. Psychiatry* **53**, 1043–1047 (1996).
27. Breslau, N., Peterson, E. L., Schultz, L. R., Chilcoat, H. D. & Andreski, P. Major depression and stages of smoking. A longitudinal investigation. *Arch. Gen. Psychiatry* **55**, 161–166 (1998).
28. Kendler, K. S. et al. Smoking and major depression. A causal analysis. *Arch. Gen. Psychiatry* **50**, 36–43 (1993).
29. Sherva, R. et al. Genome-wide association study of Cannabis dependence severity, novel risk variants, and shared genetic risks. *JAMA Psychiatry* **73**, 472–480 (2016).
30. Pierucci-Lagha, A. et al. Diagnostic reliability of the Semi-structured Assessment for Drug Dependence and Alcoholism (SSADDA). *Drug Alcohol Depend.* **80**, 303–312 (2005).
31. Heatherton, T. F., Kozlowski, L. T., Frecker, R. C. & Fagerstrom, K. O. The Fagerstrom test for nicotine dependence: a revision of the Fagerstrom Tolerance Questionnaire. *Br. J. Addict.* **86**, 1119–1127 (1991).
32. American Psychiatric Association *Diagnostic and Statistical Manual of Mental Disorders*. 4th edn, (American Psychiatric Press, Washington, DC, USA, 1994).
33. Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
34. Auton, A. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
35. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
36. Das, S. et al. Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
37. Dahl, A. et al. A multiple-phenotype imputation method for genetic studies. *Nat. Genet.* **48**, 466–472 (2016).
38. Zhou, X. & Stephens, M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* **11**, 407–409 (2014).
39. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
40. Pruim, R. J. et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).
41. GTEx Consortium Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
42. Kang, H. J. et al. Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).
43. Okamura, Y. et al. COXPRESdb in 2015: coexpression database for animal species by DNA-microarray and RNAseq-based expression data with multiple quality assessment systems. *Nucleic Acids Res.* **43**, D82–D86 (2015).
44. Wang, J., Duncan, D., Shi, Z. & Zhang, B. WEB-based Gene Set Analysis Toolkit (WebGestalt): update 2013. *Nucleic Acids Res.* **41**, W77–W83 (2013).
45. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48 (2009).
46. Zhu, J. J., Esteban, J. A., Hayashi, Y. & Malinow, R. Postnatal synaptic potentiation: delivery of GluR4-containing AMPA receptors by spontaneous activity. *Nat. Neurosci.* **3**, 1098–1106 (2000).
47. Luchkina, N. V. et al. Developmental switch in the kinase dependency of long-term potentiation depends on expression of GluA4 subunit-containing AMPA receptors. *Proc. Natl. Acad. Sci. USA* **111**, 4321–4326 (2014).
48. Ogata, H. et al. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **27**, 29–34 (1999).
49. da Huang, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
50. da Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
51. Murray, C. J. & Lopez, A. D. Measuring the global burden of disease. *N. Engl. J. Med.* **369**, 448–457 (2013).
52. Henley, J. M. & Wilkinson, K. A. Synaptic AMPA receptor composition in development, plasticity and disease. *Nat. Rev. Neurosci.* **17**, 337–350 (2016).
53. Bliss, T. V. & Collingridge, G. L. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* **361**, 31–39 (1993).
54. Huganir, R. L. & Nicoll, R. A. AMPARs and synaptic plasticity: the last 25 years. *Neuron* **80**, 704–717 (2013).
55. Ganea, D. A., Dines, M., Basu, S. & Lamprecht, R. The membrane proximal region of AMPA receptors in lateral amygdala is essential for fear memory formation. *Neuropsychopharmacology* **40**, 2727–2735 (2015).
56. Sequeira, A. et al. Global brain gene expression analysis links glutamatergic and GABAergic alterations to suicide and major depression. *PLoS ONE* **4**, e6585 (2009).
57. Gray, A. L., Hyde, T. M., Deep-Soboslay, A., Kleinman, J. E. & Sodhi, M. S. Sex differences in glutamate receptor gene expression in major depression and suicide. *Mol. Psychiatry* **20**, 1057–1068 (2015).
58. Martisova, E. et al. Long lasting effects of early-life stress on glutamatergic/GABAergic circuitry in the rat hippocampus. *Neuropharmacology* **62**, 1944–1953 (2012).
59. Katsouli, S. et al. Sexually dimorphic long-term effects of an early life experience on AMPA receptor subunit expression in rat brain. *Neuroscience* **257**, 49–64 (2014).
60. Dani, J. A., Ji, D. & Zhou, F. M. Synaptic plasticity and nicotine addiction. *Neuron* **31**, 349–352 (2001).
61. Kauer, J. A. & Malenka, R. C. Synaptic plasticity and addiction. *Nat. Rev. Neurosci.* **8**, 844–858 (2007).
62. Enoch, M. A. et al. Expression of glutamatergic genes in healthy humans across 16 brain regions; altered expression in the hippocampus after chronic exposure to alcohol or cocaine. *Genes Brain Behav.* **13**, 758–768 (2014).
63. Robinson, T. E. & Berridge, K. C. The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res. Brain Res. Rev.* **18**, 247–291 (1993).
64. Harris, G. C. & Aston-Jones, G. Critical role for ventral tegmental glutamate in preference for a cocaine-conditioned environment. *Neuropsychopharmacology* **28**, 73–76 (2003).
65. Harris, G. C., Wimmer, M., Byrne, R. & Aston-Jones, G. Glutamate-associated plasticity in the ventral tegmental area is necessary for conditioning environmental stimuli with morphine. *Neuroscience* **129**, 841–847 (2004).
66. Ungless, M. A., Whistler, J. L., Malenka, R. C. & Bonci, A. Single cocaine exposure in vivo induces long-term potentiation in dopamine neurons. *Nature* **411**, 583–587 (2001).
67. Mansvelder, H. D. & McGehee, D. S. Long-term potentiation of excitatory inputs to brain reward areas by nicotine. *Neuron* **27**, 349–357 (2000).
68. Pidoplichko, V. I., DeBiasi, M., Williams, J. T. & Dani, J. A. Nicotine activates and desensitizes midbrain dopamine neurons. *Nature* **390**, 401–404 (1997).
69. Liechti, M. E. & Markou, A. Role of the glutamatergic system in nicotine dependence: implications for the discovery and development of new pharmacological smoking cessation therapies. *CNS Drugs* **22**, 705–724 (2008).
70. Krystal, J. H. et al. Glutamate and GABA systems as targets for novel antidepressant and mood-stabilizing treatments. *Mol. Psychiatry* **7**(Suppl. 1), S71–S80 (2002).