Published in final edited form as:
Am J Addict. 2014 July ; 23(4): 411–414. doi:10.1111/j.1521-0391.2013.12115.x.

Common *PTP4A1-PHF3-EYS* variants are specific for alcohol dependence

Lingjun Zuo, MD, PhD[1,*], **Kesheng Wang, PhD**[2], **Guilin Wang, PhD**[3], **Xinghua Pan, PhD**[4], **Xiangyang Zhang, MD, PhD**[5], **Heping Zhang, PhD**[6], and **Xingguang Luo, MD, PhD**[1,*]

[1]Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA

[2]Department of Biostatistics and Epidemiology, College of Public Health, East Tennessee State University, Johnson City, TN, USA

[3]Yale Center for Genome Analysis, Departments of Genetics, Yale University School of Medicine, Orange, CT, USA

[4]Departments of Genetics, Yale University School of Medicine, New Haven, CT, USA

[5]Menninger Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, Texas, USA

[6]Department of Biostatistics, Yale University School of Public Health, New Haven, CT, USA

Abstract

Background and Objectives—We previously reported a risk genomic region (i.e., *PTP4A1-PHF3-EYS*) for alcohol dependence in a genome-wide association study (GWAS). We also reported a rare variant constellation across this region that was significantly associated with alcohol dependence. In the present study, we significantly increased the marker density within this region and examined the specificity of the associations of common variants for alcohol dependence.

Methods—One African-American discovery sample (681 cases with alcohol dependence and 508 controls), one European-American replication sample (1,409 alcohol dependent cases and 1,518 controls), and one European-Australian replication sample (a total of 6,438 family subjects with 1,645 alcohol dependent probands) underwent association analysis. A total of 38,714 subjects from 18 other cohorts with 10 different neuropsychiatric disorders served as contrast groups.

Results—We found 289 SNPs that were nominally associated with alcohol dependence in the discovery sample ($p<0.05$). Fifty-six associations of them were significant after correction ($1.9\times10^{-6}\leq p\leq1.6\times10^{-5}$). No markers were significantly associated with other neuropsychiatric disorders after experiment-wide correction.

[*]Corresponding authors and reprints: Xingguang Luo, MD and Lingjun Zuo, MD; Yale University School of Medicine; 950 Campbell Avenue, West Haven, CT 06516. Xingguang.Luo@yale.edu and Lingjun.Zuo@yale.edu; Tel: 203-932-5711 ext 3590; Fax: 203-937-4741.

Conflict of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this paper.

Conclusions and Scientific Significance—We confirmed with our previous findings that *PTP4A1-PHF3-EYS* variants were significantly associated with alcohol dependence, which were replicable across multiple independent populations and were specific for alcohol dependence. These findings suggested that this region might harbor a causal variant(s) for alcohol dependence.

Keywords

Common variants; Alcohol dependence; PTP4A1; PHF3; EYS

We previously reported a novel, functional and replicable risk gene region [i.e., protein tyrosine phosphatase type IVA gene, member 1 - Plant HomeoDomain (PHD) finger protein 3 gene - eyes shut homolog gene (*PTP4A1-PHF3-EYS*)] for alcohol dependence in a genome-wide association study (1). This 765kb region (Chr6: 64,066,604-64,831,120) included *PTP4A1, PHF3,* and their flanking regions, and part of *EYS* (close to 3' of *PHF3*). It was enriched with common risk variants [minor allele frequency (MAF) > 0.05] for alcohol dependence across African-Americans (see Supplementary Figure S1B), European-Americans (Figure S1D) and European-Australians (Figure S1F). Within 90Mb range surrounding this region in the discovery sample, all variants with $p<10^{-4}$ were concentrated in this region. Most of these risk variants had significant *cis*-acting regulatory effects on mRNA expression. The distributions of –log(p) values for association and functional signals in this region were highly consistent across six independent populations. Additionally, we tested 1,896 rare SNPs (MAF < 0.05) within this 765kb region in another association study (2) and found 22 ($9.5\times10^{-4}\leq p\leq0.05$), 17 ($0.015\leq p\leq0.05$) and 9 ($0.006\leq p\leq0.05$) individual rare SNPs that were nominally associated with alcohol dependence in above three populations, respectively. Furthermore, a rare variant constellation across the entire 765kb region was significantly associated with alcohol dependence in European-Australians ($p=4.2\times10^{-3}$). We speculated that this region might harbor a causal variant(s) for alcohol dependence.

These findings were novel. The possibility that these genes are associated with other neuropsychiatric diseases, especially those comorbid with alcohol dependence, cannot be excluded until it has been tested. In the present study, we imputed this *PTP4A1-PHF3-EYS* region across 21 independent cohorts with 11 different neuropsychiatric disorders (Table 1). We examined the associations between common *PTP4A1-PHF3-EYS* variants [minor allele frequency (MAF) > 0.05 in both cases and controls] and these disorders, in order to test whether this *PTP4A1-PHF3-EYS* region is specific for alcohol dependence. The data of these disorders were all of those with neuropsychiatric and neurological disorders available for us from the dbGaP database (http://www.ncbi.nlm.nih.gov/gap/).

A total of 49,268 subjects in these 21 cohorts were analyzed (Table 1), including one African-American discovery cohort [681 cases with alcohol dependence (DSM-IV) and 508 controls], one European-American replication cohort [1,409 alcohol dependent cases and 1,518 controls], and one European-Australian replication cohort [a total of 6,438 family subjects with 1,645 alcohol dependent probands]. 38,714 subjects in other 18 non-alcoholism cohorts served as contrast. Detailed demographic information for these samples has been published (1-9).

We imputed the missing SNPs across the entire *PTP4A1-PHF3-EYS* region using the same reference panels (1,000 Genome Project and HapMap 3). We used the same strategies as previously to maximize the success rate and accuracy of imputation, to stringently clean the phenotype and genotype data, and then to test the variant-disease associations (7). Finally, a total of 477-1,095 SNPs with MAF>0.05 in both cases and controls were extracted for association analysis. The MAFs and minimal p values of the most significant risk SNPs are shown in Table 1. The experiment-wide significance level (α) was set at 1.6×10^{-5} via correction for the number of the cohorts (i.e., n=21) and the number of effective markers (i.e., n=144) that were calculated from the entire marker set by the adjusted Bonferroni-type program SNPSpD (10).

We found that among a total of 1,095 common SNPs in the African-American cohort, 289 SNPs were nominally associated with alcohol dependence ($p<0.05$). Fifty-six associations of them were significant after correction ($1.9\times10^{-6}\leq p\leq1.6\times10^{-5}$) (Table 1). Most top-ranked SNPs ($p<10^{-5}$) were located in the 5' regulatory region of *PTP4A1* (Figure S1A). Among a total of 477-1,042 common SNPs in other cohorts, 1-272 SNPs were nominally associated with diseases ($p<0.05$); however, none of them were significant after experiment-wide correction (Table 1; Figure S1). Furthermore, 200 SNPs were nominally associated with alcohol dependence both in African-Americans ($1.9\times10^{-6}\leq p<0.05$) and European-Americans ($3.9\times10^{-4}\leq p<0.05$); 15 SNPs were nominally associated with alcohol dependence across African-Americans ($2.3\times10^{-4}\leq p<0.05$), European-Americans ($0.006\leq p<0.05$) and European-Australians ($0.009\leq p<0.05$). Most of the replicable SNPs were located in *EYS* (data not shown).

By expanding the marker set, we confirmed with our previous findings that common *PTP4A1-PHF3-EYS* variants were significantly associated with alcohol dependence, which were replicable across African-Americans, European-Americans and European-Australians. Furthermore, by testing 10 other non-alcoholism neuropsychiatric disorders, we found that common *PTP4A1-PHF3-EYS* variants were specific to alcohol dependence; that is, they were not significantly associated with any other disorder examined. This study supports our previous conclusion that this region might harbor a causal variant(s) for alcohol dependence. PTP4A1 protein may interact with an activating transcription factor 7 (ATF7), which is a cAMP responsive element (CRE) binding protein and may interact with FOSB. CRE and FOSB have been implicated in addiction, including alcohol dependence. Functional connections of other genes within this region with alcohol dependence warrant further investigation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by National Institute on Drug Abuse (NIDA) grants K01 DA029643 and R01DA016750, National Institute on Alcohol Abuse and Alcoholism (NIAAA) grants R01 AA016015, R21 AA021380 and R21 AA020319, ABMRF/The Foundation for Alcohol Research (L.Z.) and the National Alliance for Research on Schizophrenia and Depression (NARSAD) Award 17616 (L.Z.). We thank NIH GWAS Data Repository, the Contributing Investigator(s) who contributed the phenotype and genotype data from his/her original

study (e.g., Drs. Bierut, Edenberg, Heath, Singleton, Hardy, Foroud, Myers, Gejman, Faraone, Sonuga-Barke, Sullivan, Nurnberger, Devlin, Monaco, etc.), and the primary funding organization that supported the contributing study. Funding and other supports for phenotype and genotype data were provided through the National Institutes of Health (NIH) Genes, Environment and Health Initiative (GEI) (U01HG004422, U01HG004436 and U01HG004438); the GENEVA Coordinating Center (U01HG004446); the NIAAA (U10AA008401, R01AA013320, P60AA011998); the NIDA (R01DA013423); the National Cancer Institute (P01 CA089392); the Division of Neuroscience, the NIA National Institute of Neurological Disorders and Stroke (NINDS); the NINDS Human Genetics Resource Center DNA and Cell Line Repository; the NIH contract "High throughput genotyping for studying the genetic contributions to human disease" (HHSN268200782096C); the Center for Inherited Disease Research (CIDR); a Cooperative Agreement with the Division of Adult and Community Health, Centers for Disease Control and Prevention; the NIH Office of Research on Women's Health (ORWH) (R01NS45012); the Department of Veterans Affairs; the University of Maryland General Clinical Research Center (M01RR165001), the National Center for Research Resources, NIH; the National Institute of Mental Health (K01MH086621, R01MH059160, R01MH59565, R01MH59566, R01MH59571, R01MH59586, R01MH59587, R01MH59588, R01MH60870, R01MH60879, R01MH61675, R01MH62873, R01MH081803, R01MH67257, R01MH81800, U01MH46276, U01MH46282, U01MH46289, U01MH46318, U01MH79469, U01MH79470 and R01MH67257); the NIMH Genetics Initiative for Bipolar Disorder; the Genetic Association Information Network (GAIN); the Genetic Consortium for Late Onset Alzheimer's Disease; the Autism Genome Project, the MARC: Risk Mechanisms in Alcoholism and Comorbidity; the Molecular Genetics of Schizophrenia Collaboration; the Medical Research Council (G0601030) and the Wellcome Trust (075491/Z/04), University of Oxford; the Netherlands Scientific Organization (904-61-090, 904-61-193, 480-04-004, 400-05-717, NWO Genomics, SPI 56-464-1419) the Centre for Neurogenomics and Cognitive Research (CNCR-VU); Netherlands Study of Depression and Anxiety (NESDA) and the Netherlands Twin Register (NTR); and the European Union (EU/WLRT-2001-01254), ZonMW (geestkracht program, 10-000-1002). Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the Genetic Consortium for Late Onset Alzheimer's Disease, the GENEVA Coordinating Center (U01 HG004446), and the National Center for Biotechnology Information. Genotyping was performed at the Johns Hopkins University Center for Inherited Disease Research, and GlaxoSmithKline, R&D Limited. The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gap. The dbGaP accession numbers include phs000125.v1.p1, phs000021.v3.p2, phs000021.v3.p2, phs000167.v1.p1, phs000167.v1.p1, phs000267.v1.p1, phs000016.v2.p2, phs000092.v1.p1, phs000092.v1.p1, phs000181.v1.p1, phs000020.v2.p1, phs000017.v3.p1, phs000017.v3.p1, phs000017.v3.p1, phs000168.v1.p1, phs000219.v1.p1, phs000101.v3.p1, phs000292.v1.p1, phs000292.v1.p1, phs000102.v1.p1, phs000196.v2.p1, phs000126.v1.p1, phs000089.v3.p2, phs000089.v3.p2, phs000089.v3.p2 and phs000089.v3.p2.

References

1. Zuo L, Zhang CK, Wang F, Li CS, Zhao H, Lu L, et al. A novel, functional and replicable risk gene region for alcohol dependence identified by genome-wide association study. PLoS One. 2011; 6:e26726. [PubMed: 22096494]
2. Zuo L, Zhang X, Deng HW, Luo X. Association of rare PTP4A1-PHF3-EYS variants with alcohol dependence. J Hum Genet. 2013; 58:178–9. [PubMed: 23324950]
3. Bierut LJ, Agrawal A, Bucholz KK, Doheny KF, Laurie C, Pugh E, et al. A genome-wide association study of alcohol dependence. Proc Natl Acad Sci U S A. 2010; 107:5082–7. [PubMed: 20202923]
4. Edenberg HJ, Koller DL, Xuei X, Wetherill L, McClintick JN, Almasy L, et al. Genome-wide association study of alcohol dependence implicates a region on chromosome 11. Alcohol Clin Exp Res. 2010; 34:840–52. [PubMed: 20201924]
5. Heath AC, Whitfield JB, Martin NG, Pergadia ML, Goate AM, Lind PA, et al. A quantitative-trait genome-wide association study of alcoholism risk in the community: findings and implications. Biol Psychiatry. 2011; 70:513–8. [PubMed: 21529783]
6. Zuo L, Gelernter J, Zhang CK, Zhao H, Lu L, Kranzler HR, et al. Genome-wide association study of alcohol dependence implicates KIAA0040 on chromosome 1q. Neuropsychopharmacology. 2012; 37:557–66. [PubMed: 21956439]
7. Zuo L, Wang F, Zhang XY, Li CSR, Lu L, Ye L, et al. Genome-wide significant association signals in IPO11-HTR1A region specific for alcohol and nicotine co-dependence. Alcohol Clin Exp Res. 2013; 37:730–9. [PubMed: 23216389]
8. Zuo L, Zhang F, Zhang H, Zhang XY, Wang F, Li CS, et al. Genome-wide search for replicable risk gene regions in alcohol and nicotine co-dependence. Am J Med Genet B Neuropsychiatr Genet. 2012; 159B:437–44. [PubMed: 22488850]

9. Zuo L, Zhang H, Malison RT, Li CSR, Zhang XY, Wang F, et al. Rare ADH variant constellations are specific for alcohol dependence. Alcohol Alcohol. 2013; 48:9–14. [PubMed: 23019235]

10. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity (Edinb). 2005; 95:221–7. [PubMed: 16077740]

Table 1

Associations between common *PTP4A1-PHF3-EYS* variants and different neuropsychiatric disorders

Human Diseases	Ethnicity	Design	Dataset name	SNP # (total)	SNP # (p<0.05)	SNP # (p<α)	Minimal p value	Most sig. SNP	Gene	Affected N	Affected MAF	Unaffected N	Unaffected MAF
Alcoholism	AA	CC	SAGE+COGA	1095	289	56	1.9×10^{-6}	rs7742595	*5' to PTP4A1*	681	0.249	508	0.167
Alcoholism	EA	CC	SAGE+COGA	762	272	0	3.9×10^{-4}	rs10755416	*5' to PTP4A1*	1409	0.456	1518	0.408
Alcoholism	EAu	Fam	OZ-ALC	734	42	0	4.5×10^{-3}	rs2347978	*5' to PTP4A1*	1645	0.074	1645	0.135
Bipolar Disorder	EA	CC	BDO+GRU	719	248	0	2.4×10^{-5}	rs504776	*EYS*	368	0.316	1034	0.417
Bipolar Disorder	EA	CC	BARD+GRU	718	183	0	3.5×10^{-4}	rs1057530	*PHF3-EYS*	653	0.419	1034	0.491
Bipolar Disorder	AA	CC	BARD+GRU	825	4	0	0.026	rs1482444	*EYS*	141	0.177	671	0.391
ADHD	CA	Fam	IMAGE	768	63	0	9.2×10^{-5}	rs10943832	*5' to PTP4A1*	924	0.137	924	0.421
Schizophrenia	EA	CC	MGS_nonGAIN	702	6	0	0.022	rs114419825	*EYS*	1437	0.092	1347	0.056
Schizophrenia	EA	CC	GAIN	717	56	0	5.4×10^{-4}	rs1723533	*5' to PTP4A1*	1351	0.064	1378	0.101
Schizophrenia	AA	CC	GAIN	805	58	0	3.4×10^{-4}	rs76384923	*EYS*	1195	0.476	954	0.330
Autism	EA	Fam	AGP	720	22	0	1.7×10^{-4}	rs9351126	*EYS*	1330	0.169	1330	0.252
Major Depression	CA	CC	PRSC	730	57	0	2.5×10^{-4}	rs7753631	*EYS*	1805	0.282	1820	0.324
Alzheimer's Disease	CA	Fam	LOAD × 4	802	6	0	0.015	rs2624662	*EYS*	2298	0.288	2298	0.313
Alzheimer's disease	EA	CC	GenADA	477	1	0	0.045	rs1779776	*5' to PTP4A1*	806	0.057	782	0.075
ALS	CA	CC	GRU	540	79	0	0.008	rs1711920	*5' to PTP4A1*	261	0.250	246	0.330
Early Onset Stroke	EA	CC	GEOS × 3	749	29	0	0.007	rs6915363	*EYS*	372	0.114	430	0.064
Early Onset Stroke	AA	CC	GEOS × 3	1042	10	0	0.014	rs9362331	*EYS*	309	0.353	290	0.435
Ischemic Stroke	CA	CC	ISGS	722	3	0	0.020	rs3003669	*EYS*	219	0.370	266	0.298
Parkinson's Disease	CA	CC	NGRC	753	4	0	0.004	rs6900114	*EYS*	2000	0.099	1986	0.125
Parkinson's Disease	CA	CC	PDRD+GRU	711	1	0	0.046	rs1681939	*5' to PHF3*	900	0.093	867	0.114
Parkinson's Disease	CA	CC	lng_coriell_pd	765	254	0	0.004	rs13213141	*3' to PTP4A1*	940	0.200	801	0.249

Only the most significant risk markers are listed. The significance level (α) is set at 1.6×10^{-5} based on correction for the numbers of effective genetic markers (calculated by SNPSpD) and the number of cohorts (i.e., 21). N, sample size; MAF, minor allele frequency; AA, African-American; EA, European-American; EAu, European-Australian; CA, Caucasian; CC, case-control design; Fam, family-based design. ADHD, Attention deficit hyperactivity disorder; ALS, Amyotrophic Lateral Sclerosis. Dataset names correspond to dbGaP. In family-based cohorts, N= sample size of affected offspring; "affected MAF"="transmitted MAF", "unaffected MAF"="untransmitted MAF" in offspring.