Sex chromosome-wide association analysis suggested male-specific risk genes for alcohol dependence

Lingjun Zuo[a], Kesheng Wang[e], Xiangyang Zhang[f], Xinghua Pan[b], Guilin Wang[c], John H. Krystal[a], Heping Zhang[d] and Xingguang Luo[a]

Background **Alcohol dependence is more common among men than among women. Potential explanations for this include the role of genes in sex chromosomes (X and Y). In the present study, we scanned the entire Y chromosome and its homologs on the X chromosome in men to identify male-specific risk genes for alcohol dependence.**

Methods **Two thousand nine hundred and twenty-seven individuals in two independent cohorts were analyzed. The European-American male cohort (883 cases with alcohol dependence and 445 controls) served as the discovery cohort and the European-American female cohort (526 cases and 1073 controls) served as a contrast group. All individuals were genotyped on the Illumina Human 1M beadchip. Two thousand two hundred and twenty-four single nucleotide polymorphisms (SNPs) on the Y chromosome or in the homologs on the X chromosome were analyzed. The allele frequencies were compared between cases and controls within each cohort using logistic regression analysis.**

Results **We found that, after experiment-wide correction, two SNPs on the X chromosome were associated significantly with alcohol dependence in European-American men ($P=1.0\times10^{-4}$ for rs5916144 and $P=5.5\times10^{-5}$ for rs5961794 at 3′ UTR of *NLGN4X*), but not in the women. A total of 26 SNPs at 3′UTR of or within *NLGN4X* were nominally associated with alcohol dependence in men ($5.5\times10^{-5}\leq P\leq0.05$), all of which were not statistically significant in women.**

Conclusion **We conclude that *NLGN4X* was a significant male-specific risk gene for alcohol dependence in European-Americans. *NLGN4X* might harbor a causal variant(s) for alcohol dependence. A defect of synaptogenesis in neuronal circuitry caused by *NLGN4X* mutations is believed to play a role in alcohol dependence.** ***Psychiatr Genet*** **23:233–238** © **2013 Wolters Kluwer Health | Lippincott Williams & Wilkins.**

Psychiatric Genetics 2013, **23**:233–238

Keywords: alcohol dependence, homolog, male specificity, *NLGN4X*, synaptogenesis, Y chromosome

Departments of [a]Psychiatry and [b]Genetics, Yale University School of Medicine, [c]Department of Genetics, Yale Center for Genome Analysis, Yale University School of Medicine, [d]Department of Biostatistics, Yale University School of Public Health, New Haven, Connecticut, [e]Department of Biostatistics and Epidemiology, College of Public Health, East Tennessee State University, Johnson City, Tennessee and [f]Menninger Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, Texas, USA

Correspondence to Xingguang Luo, MD, Department of Psychiatry, Yale University School of Medicine, West Haven, CT 06516, USA
Tel: +1 203 932 5711 x5745; fax: +1 203 937 4741;
e-mails: xingguang.luo@yale.edu, lingjun.zuo@yale.edu

Received 3 October 2012 **Revised** 7 January 2013 **Accepted** 7 April 2013

Introduction

Studies in many cultures have found that the prevalence of alcohol dependence and heavy drinking generally is higher among men than among women (Heath *et al.*, 1998). Potential explanations for this include the role of genes in sex chromosomes (X and Y) or genes on autosomes that are related to the function of sex-related hormones or intrauterine or postnatal hormonal influences. However, the role of sex chromosomes in alcohol dependence has not been investigated comprehensively so far. In the present study, we scanned the entire Y chromosome and its homologs on the X chromosome in men to identify male-specific risk genes for alcohol dependence.

The Y chromosome comprises two pseudoautosomal regions (PARs) at both ends (5%) and a male-specific region (MSY) in between (95%). Sequences in PARs are the only portions on the Y chromosome that are identical (~100%) to the corresponding PARs on the X chromosome, reflecting the exchange of DNA between X and Y (crossover) that occurs during sperm production. The PAR1 at Yp11.32/Xp22.33 (2.7 Mb) is the result of an obligatory recombination in male meiosis; gene loci in the PAR1 are present in two copies in both men and women and are not subject to dosage compensation by X chromosome inactivation (Helena Mangs and Morris, 2007). These genes in the PAR1 include *ASMT*, *ASMTL*, *CD99*, *CRLF2*, *CSF2RA*, *SFRS17A*, *DHRS*, *GTPBP6*, *IL3RA*, *P2RY8*, *PLCXD1*, *PPP2R3B*, *SHOX*, *SLC25A6*, *XG*, and *ZBED1*. The PAR2 at Yq12/Xq28 (330 kb) was created by duplication of material from X to Y since the divergence of human and chimpanzee lineages about 6 million years ago. The genes in the PAR2 include *SPRY3*, *VAMP7*, *SYBL1*, and *IL9R*, some of which are subject to X chromosome inactivation (Helena Mangs and Morris, 2007). MSY consists of heterochromatin and euchromatin. Heterochromatin hugs the centromere and the Yq12

0955-8829 © 2013 Wolters Kluwer Health | Lippincott Williams & Wilkins DOI: 10.1097/YPG.0b013e328364b8c7

region (without PAR2) at the telomere, which are tightly coiled and enriched with repetitive DNA sequences. Heterochromatin maintains the chromosome's structural integrity. It is interspersed with euchromatin along the rest of MSY. All active genes are located in the euchromatin. Euchromatin is classified to an X-transposed region (XTR), an X-degenerate region (XDR), an ampliconic region, and others. (a) The XTR (3.38 Mb) is located at Yp11.2, next to the PAR1. Its sequences are 99% identical to DNA sequences in Xq21.3 (3.91 Mb), a band in the midst of the Xq chromosome. This XTR was transposed from X to Y about 6 million years ago, after the divergence of the human and chimpanzee lineages. It escapes X chromosome inactivation and does not require dosage compensation. This XTR contains two known pairs of functional genes, that is *PCDH11X* and *PCDH11Y*, a pair of protocadherin genes expressed significantly in the adult human brain, and *TGIF2LX* and *TGIF2LY*, a pair of transforming growth factor-beta-induced factor 2-like genes expressed exclusively in testes. This region is hominid specific and absent from the Y chromosome of nonhuman primates (Blanco *et al.*, 2000; Williams *et al.*, 2006; Wilson *et al.*, 2006). It might thus play a role in hominid-specific characteristics including the cerebral asymmetry (brain lateralization), a basis of human-specific cognitive functions such as language (adaptive traits). (b) The X-degenerate segments on Y are not concentrated in one region. They are separated into 12 blocks by some single-copy genes or pseudogenes. However, the homologs of these 12 blocks on the X chromosome are concentrated in a 6 Mb continuous region around Xp22.32, next to the PAR1 (Xp22.33). XDRs show 60–96% of sequence identity between X and Y. They are surviving relics of ancient autosomes from which the X and Y chromosome coevolved. XDRs hug *SOX3/SRY*, *RPS4X/Y1*, *ZFX/Y*, *AMELX/Y*, *TBL1X/Y*, *PRKX/Y*, *USP9X/Y*, *DBX/Y*, *UTX/Y*, *TMSB4X/Y*, *NLGN4X/Y*, *CXorf15/CYorf15A*, *CXorf15/CYorf15B*, *SMCX/Y*, *EIF1AX/Y*, and *RPS4X/Y2*. Most of these genes are ubiquitous expressed, but only *NLGN4X/Y* are significantly expressed in the adult human brain (Lin *et al.*, 2012). (c) The ampliconic regions are nonrecombinant and Y-specific, that is, they are single-copy and do not have homologous regions on X. The genes in these regions are expressed exclusively in testes. Deletions in these regions are a recognized cause of male infertility. That is, the PAR1, XTR, and XDR are more likely to be relevant to brain disorders.

In summary, a total of 54 functional homologous genes have been identified so far, including 24 in PAR1, five in PAR2, three in XTR, 15 in XDR, and seven in other regions. It is estimated that ~15 protein-coding genes on the Y chromosome are Y-specific and have no detectable X chromosome homolog. In the present study, all homologous genes on both X and Y chromosomes, in addition to other Y-specific genes on Y, were tested. These homologs are believed to balance their dosage between men and women. Therefore, as a contrast, these homologs were also tested in women to determine whether the risk genes among those homologs are male specific or shared by women. Because the phasing information of these homologs was not available in the microarray platforms, we focused on individual variants but not multilocus haplotypes in the present study.

Materials and methods

Participants

Two thousand nine hundred and twenty-seven individuals in two independent cohorts were analyzed. The European-American male cohort [883 cases with alcohol dependence *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed. (DSM-IV) and 445 controls] served as the discovery cohort and the European-American female cohort (526 cases and 1073 controls) served as a contrast group. In addition, two smaller African-American cohorts including a male cohort (428 cases and 169 controls) and a female cohort (253 cases and 339 controls) were also analyzed as contrast groups on an exploratory basis. Affected individuals fulfilled the lifetime DSM-IV criteria (American Psychiatric Association, 1994) for alcohol dependence. Controls were defined as individuals who had been exposed to alcohol (and possibly to other drugs), but had never become addicted to alcohol or other illicit substances (lifetime diagnoses). Controls were also screened to exclude individuals with major axis I disorders, including schizophrenia, mood disorders, and anxiety disorders. More detailed demographic data have been published previously (Bierut *et al.*, 2010; Edenberg *et al.*, 2010; Heath *et al.*, 2011; Zuo *et al.*, 2011a, 2011b, 2012, 2013a, 2013b). All participants were deidentified in this study. All participants were genotyped on the Illumina Human 1M beadchip (Illumina Inc., San Diego, California, USA).

Data cleaning

Individuals with poor genotypic data and questionable diagnostic information, those with allele discordance, duplicated IDs, potential sample misidentification, sample relatedness, sample misspecification, sex anomalies, chromosome anomalies (such as aneuploidy and mosaic cell populations), missing race, non-European and non-African ethnicity, population group outliers, individuals with a mismatch between self-identified and genetically inferred ethnicity, and those with a missing genotype call rate of at least 2% across all SNPs examined were excluded step by step (Zuo *et al.*, 2011a). Furthermore, only the SNPs on the Y chromosome and the SNPs in the homologous genes on the X chromosome were extracted from the 1M beadchip for association analysis. Because some homologous genes, especially in XDR, were not completely identical in sequence between X and Y chromosomes, some of the SNPs in these genes appeared only on the X chromosome. Among these extracted SNPs, those with allele discordance or chromosomal anomalies were excluded. We then filtered out the SNPs with an

overall missing genotype call rate of at least 2% and the SNPs with minor allele frequencies less than 0.01 in either European-Americans or African-Americans. The SNPs on the X chromosome in women and the SNPs in the homologous regions in men that deviated from the Hardy–Weinberg equilibrium ($P < 10^{-4}$) within European-American or African-American controls were also excluded. This selection process yielded a total of 2224 SNPs for analysis, including 1481 in 17 homologous genes on the X chromosome only, 480 SNPs in 34 genes on the Y chromosome only, and 263 SNPs in the recombinant regions (i.e. XY regions) of both X and Y chromosomes [including 16 genes in the PAR1 (*ASMT*, *ASMTL*, *CD99*, *CRLF2*, *CSF2RA*, *CXYorf2*, *DHRSXY*, *DXYS155E*, *GTPBP6*, *IL3RA*, *P2RY8*, *PLCXD1*, *PPP2R3B*, *SHOX*, *SLC25A6*, *XG*, and *ZBED1*), three genes in the PAR2 (*IL9R*, *SPRY3*, and *SYBL1*), two genes in the XTR (*PCDH11X/Y* and *TGIF2LX/Y*), 12 genes in XDR (*AMELX/Y*, *DDX3X/Y*, *EIF1AX/Y*, *NLGN4X/Y*, *PRKX/Y*, *RPS4X/Y*, *SMCX/Y*, *TBL1X/Y*, *TMSB4X/Y*, *USP9X/Y*, *UTX/Y*, and *ZFX/Y*), and one gene in another region (*NXF2*)].

Association analysis

The allele frequencies were compared between cases and controls within each cohort using logistic regression analysis implemented in the program PLINK (Purcell *et al.*, 2007). Diagnosis served as the dependent variable, alleles served as the independent variables, and age and the first 10 principal components served as the covariates. The principal component scores of our samples were derived from all autosomal SNPs across the genome using principal component analysis implemented in the software package EIGENSTRAT (Price *et al.*, 2006). Each individual received scores on each principal component. The first 10 principal component scores accounted for greater than 95% of variance. These principal components reflected the population structure of our samples. These principal components serving as covariates in the regression model controlled for the population stratification and admixture effects on association analysis.

The experiment-wide significance levels (α) were corrected for the numbers of effective markers that were calculated using the program SNPSpD (Nyholt, 2004), which is an adjusted Bonferroni correction taking the linkage disequilibrium structure into account. Approximately 379 and 494 effective SNPs that had an r^2 of up to 0.1 between any two of them captured most of the information of all variants examined in cohorts of European and African descent, respectively. Thus, the corrected significance levels (α) for association tests were set at 1.3×10^{-4} in cohorts of European descent and 1.0×10^{-4} in cohorts of African descent, respectively.

Power analysis

We computed the power of test using the R package 'pwr', which uses the effect sizes and the notations from Cohen (1988). Given the effect size (w), the total number of observations (n), the degree of freedom (*d.f.*), and the significance level (α), the statistical power of test can be determined from these parameters. The effect size is $w = \sum (O-E)^2/E$, where O is the observed allele frequency of the targeted SNP and E is its expected allele frequency (Table 1). The α was corrected for multiple testing as described above.

Results

We found that 79 (62 on X, four on Y, 13 on XY) and 132 (50 on X, 73 on Y, nine on XY) SNPs were nominally associated with alcohol dependence in European-American men and African-American men, respectively ($P < 0.05$), and 51 (37 on X and 14 on XY) and 69 (62 on X and seven on XY) SNPs were nominally associated with alcohol dependence in European-American women and African-American women, respectively ($P < 0.05$). After experiment-wide correction, two SNPs on X were associated significantly with alcohol dependence in European-American men ($P = 1.0 \times 10^{-4}$ for rs5916144 and $P = 5.5 \times 10^{-5}$ for rs5961794 at 3′UTR of *NLGN4X*), but not in other cohorts.

Given the allele frequencies discovered in the European-American male cohort listed in Table 1, $w = 0.167$, *d.f.* $= 1$, and $\alpha = 10^{-4}$, the European-American female cohort ($n = 1599$) had 99.8% of power to detect the two most significant risk markers. However, the African-American male and female cohorts had only 68.1 and 51.4% of power to significantly detect them, respectively; thus, the findings in African-American cohorts were considered exploratory.

We found that 26 of 444 SNPs at 3′UTR or within *NLGN4X* were nominally associated with alcohol dependence in European-American men ($5.5 \times 10^{-5} \leq P \leq 0.05$), including those two significant ones. All of these associations were not statistically significant in European-American women, and most of them were not significant in African-Americans men and women either (Table 1).

Discussion

We found that *NLGN4X* on the X chromosome was a significant male-specific risk gene for alcohol dependence in European-Americans. *NLGN4X* might harbor a causal variant(s) for alcohol dependence.

NLGN4X, spanning 338 kb, maps to Xp22.32 on XDR, proximal to the PAR1 of the X chromosome. It contains six exons and codes for neuroligin 4 (with 816 amino acids) (Ichtchenko *et al.*, 1996). Its homolog on the Y chromosome, that is, *NLGN4Y* (at Yq11.221), is an X-degenerate gene. Forty million years ago, during the process of divergence of Old World from New World Monkeys, *NLGN4X* and *NLGN4Y* were introduced into X and Y chromosomes, respectively, originating from the common ancient autosomes (De Oliveira *et al.*, 2012).

Table 1 Associations between the *NLGN4X* gene and alcohol dependence

				Risk allele frequency		Association			Contrast groups		
SNP	Position	Location	Risk allele	Cases	Controls	OR	95% CI	*P* (EAM)	*P* (EAF)	*P* (AAM)	*P* (AAF)
rs5915821	4746402	3′UTR	T	0.532	0.442	1.45	1.14–1.80	0.001	0.574	0.166	0.012
rs5915826	4752475	3′UTR	A	0.532	0.442	1.45	1.14–1.80	0.001	0.575	0.177	0.006
rs5915829	4758997	3′UTR	G	0.646	0.562	1.43	1.13–1.80	0.002	0.901	0.854	0.199
rs4826671	4771654	3′UTR	C	0.306	0.232	1.51	1.13–1.90	0.002	0.205	0.450	0.550
rs17305063	5396913	3′UTR	G	0.822	0.769	1.40	1.05–1.83	0.017	0.684	0.699	0.313
rs5916106	5440260	3′UTR	A	0.827	0.771	1.43	1.07–1.87	0.012	0.588	0.803	0.446
rs5916133	5464521	3′UTR	T	0.826	0.776	1.40	1.04–1.82	0.021	0.548	0.580	0.288
rs1564673	5486848	3′UTR	G	0.686	0.586	1.58	1.22–1.95	1.5×10^{-4}	0.762	0.527	0.366
rs5916144	5489611	3′UTR	A	0.731	0.633	1.62	1.24–2.01	1.0×10^{-4}	0.870	0.614	0.182
rs5915572	5500890	3′UTR	G	0.740	0.662	1.48	1.13–1.85	0.002	0.770	0.568	0.195
rs5916157	5503597	3′UTR	G	0.689	0.625	1.35	1.05–1.69	0.013	0.814	0.424	0.794
rs5961740	5509538	3′UTR	G	0.689	0.626	1.35	1.04–1.68	0.015	0.811	0.418	0.922
rs4826792	5518048	3′UTR	A	0.729	0.668	1.35	1.05–1.71	0.016	0.508	0.243	0.913
rs5916176	5525522	3′UTR	A	0.706	0.648	1.32	1.03–1.66	0.024	0.595	0.474	0.546
rs5916183	5568175	3′UTR	T	0.481	0.423	1.31	1.01–1.59	0.022	0.808	0.939	0.856
rs9698745	5573674	3′UTR	C	0.725	0.640	1.50	1.16–1.89	0.001	0.407	0.656	0.275
rs7051992	5576009	3′UTR	T	0.720	0.636	1.49	1.16–1.88	0.001	0.412	0.548	0.262
rs5961794	5584328	3′UTR	A	0.721	0.614	1.64	1.28–2.06	5.5×10^{-5}	0.245	0.511	0.792
rs882145	5597394	3′UTR	A	0.631	0.563	1.31	1.05–1.67	0.021	0.412	0.628	0.279
rs2369265	5600519	3′UTR	G	0.632	0.564	1.31	1.05–1.67	0.022	0.408	0.738	0.328
rs11094800	5621715	3′UTR	G	0.490	0.429	1.30	1.02–1.61	0.027	0.659	0.701	0.171
rs5961840	5712122	3′UTR	C	0.798	0.741	1.40	1.06–1.81	0.014	0.437	0.031	0.255
rs5961851	5722793	3′UTR	A	0.802	0.741	1.43	1.08–1.85	0.009	0.617	0.059	0.904
rs5961861	5738176	3′UTR	T	0.674	0.609	1.36	1.05–1.68	0.010	0.961	0.288	0.162
rs16997675	5952368	Intron4	G	0.201	0.145	1.44	1.09–2.02	0.022	0.919	0.689	0.242
rs7052989	5954178	Intron4	A	0.202	0.150	1.39	1.05–1.94	0.037	0.836	0.237	0.523

Other protein-coding gene closest to these markers, that is *PRKX*, is 2176 kb far away from the 3′ of *NLGN4X*.
AAF, African-American females; AAM, African-American males; CI, confidence interval; EAF, European-American females; EAM, European-American males; OR, odds ratio; SNP, single nucleotide polymorphism; UTR, untranslated region.

Females have two Xs and males have one X and one Y. To make the gene dosage equivalent between both sexes, females inactivate almost an entire X chromosome. However, about 15% of the genes on the inactive X chromosome, mostly situated on the short arm (Xp) including PAR1 and XDR, permanently escape inactivation (Carrel and Willard, 2005). Consistent with the dosage compensation rule, these 15% of genes have homologs on the Y chromosome, so that their dosages are still equivalent between males and females. *NLGN4X* is among these 15% of genes; therefore, *NLGN4X* and *NLGN4Y* are supposed to be paired, diploid-acting genes on chromosomes X and Y, respectively. However, unlike the genes in the PARs and XTRs that have sequence identity of at least 99%, *NLGN4X* and *NLGN4Y* are XDR genes and are ~97.5% similar at the amino acid level and less similar at the DNA sequence level ($\leq$ 96%) (Jamain *et al.*, 2003), so that some variants at *NLGN4X* are located on the X chromosome only (although most variants at *NLGN4Y* are located on both X and Y chromosomes according to dbSNPs). Interestingly, all SNPs at *NLGN4X* significantly or nominally associated with alcohol dependence in men identified in the present study are located on a single copy of X (Table 1). Those *NLGN4* variants located on both X and Y chromosomes were not associated with alcohol dependence in men. In women, all of the *NLGN4X* markers located on two copies of X chromosomes were not associated with alcohol dependence either (after correction). That is, only haploid *NLGN4X* variants on X in men showed associations with alcohol dependence. Those diploid *NLGN4* SNPs paired between two X chromosomes (in women) or between X and Y chromosomes (in men) were not associated with alcohol dependence.

Two copies of *NLGN4X* in women escape the inactivation and thus both express protein product. When one of the pair has a defective effect, another copy might compensate the defective effect of the risk copy, and thus might protect against diseases. Similarly, when the variants are paired between *NLGN4X* and *NLGN4Y* in men, the defective effects of risk alleles could also be compensated. Only when the defective effects come from the single-copy of X (i.e. in men) could these effects be invariably expressed, increasing the risk for diseases, which was exactly what we observed in this study. This is why many male-dominant diseases are caused by genes on the X chromosome, and men are more likely to have X chromosome-related diseases (Zinn *et al.*, 2007).

The male-specific associations between *NLGN4X* and alcohol dependence are paralleled by male-dominant inheritance for other disorders where *NLGN4X* is implicated, particularly autism (Jamain *et al.*, 2003). *NLGN4X* is also named as autism susceptibility X-linked type 2 gene (*AUTSX2*). Given the significant phenotypic overlap with autism and the putative role for synaptic plasticity in the etiology of alcohol dependence, it is understandable that *NLGN4X* is also a risk gene for

alcohol dependence. Interestingly, other genes have been linked to both, the heritable risk for alcoholism and that for autism. For example, the autism susceptibility candidate gene 2 (*AUTS2*) was reported to be a risk gene for alcohol consumption by a genome-wide association study (Schumann *et al.*, 2011). In addition to autism and alcohol dependence, *NLGN4X* has also been associated with several other cognitive disorders including Asperger syndrome, Tourette syndrome, mental retardation, selective nonverbal deficits (e.g. impaired visual-spatial abilities), learning disability, and/or schizophrenia (Jamain *et al.*, 2003; Laumonnier *et al.*, 2004; Yan *et al.*, 2005; Zinn *et al.*, 2007; Lawson-Yuen *et al.*, 2008; Walsh *et al.*, 2008).

The brain processes information by transmitting signals at synapses, which connect neurons into vast networks of communicating cells. In these networks, synapses not only transmit signals but also transform and refine them. All processing of information in the brain involves synapses, and almost all abnormalities in brain function have a direct or an indirect effect on synaptic function. Neurexins and neuroligins are enriched specifically in the membrane of glutamatergic synapses, especially in the dendritic spines in the hippocampal neurons (Chih *et al.*, 2005). They are arguably the best-characterized synaptic cell-adhesion molecules, and they are the only ones for which a specifically synaptic function has been established (Dean and Dresbach, 2006; Craig and Kang, 2007). Neuroligin–neurexin complexes connect presynaptic and postsynaptic neurons at synapses, mediate signaling across the synapse, and shape the properties of neural networks by specifying synaptic functions, playing a central role in the brain's ability to process information during nervous system development. Neuroligin deficiency in the brain caused by *NLGN4X* mutations may lead to abnormal development of synaptic structures and may have adverse effects on communication processes and cognitive development; thus, neuroligin is a key target in the pathogenesis of cognitive diseases (Sudhof, 2008).

Imbalance between neuronal excitation and inhibition could lead to epilepsy; the prevalence of epilepsy among alcoholics is at least triple that in the general population (Chan, 1985). It is hypothesized that at least some forms of alcohol dependence might be caused by an imbalance of neuronal excitation/inhibition. Neuroligin 4 is localized to excitatory glutamatergic axons and plays a role in adjusting the balance between excitatory and inhibitory synapses (Graf *et al.*, 2004; Prange *et al.*, 2004; Chih *et al.*, 2005). Neuroligin defects caused by *NLGN4X* mutations may abolish the formation, stabilization, or recognition of specific synapses essential for the communication processes, and subsequently lead to selective loss of excitatory function and abnormal excitatory/inhibitory balance in neurons (Jamain *et al.*, 2003). Such a defect of synaptogenesis in neuronal circuitry is believed to play a role in alcohol dependence.

NLGN4X conferring a risk for alcohol dependence might be population specific (i.e. European-Americans only). However, this specificity should be confirmed in the future by expanding the sample sizes of African-American cohorts to achieve sufficient study power.

Acknowledgements

The authors thank NIH GWAS Data Repository, the Contributing Investigator(s) (Drs. Bierut, Edenberg, etc.) who contributed the phenotype and genotype data (SAGE and COGA) from his/her original study, and the primary funding organization that supported the contributing study. Assistance with data cleaning was provided by the National Center for Biotechnology Information. Genotyping was performed at the Johns Hopkins University Center for Inherited Disease Research or at deCODE. Grant sponsor: National Alliance for Research on Schizophrenia and Depression (NARSAD) Award (L.Z.); Grant number: 17616; Grant sponsor: ABMRF/The Foundation for Alcohol Research (L.Z.); Grant sponsor: National Institute on Drug Abuse (NIDA); Grant numbers: K01 DA029643, K24 DA017899, R01 DA016750, K02 DA026990, R01 DA013423; Grant sponsor: National Institute on Alcohol Abuse and Alcoholism (NIAAA); Grant numbers: R01 AA016015, R21 AA020319, K24 AA013736, P50 AA012870, U10 AA008401; Grant sponsor: Genes, Environment and Health Initiative (GEI); Grant numbers: U01HG004422, U01HG004438; Grant sponsor: GENEVA Coordinating Center; Grant number: U01 HG004446; Grant sponsor: National Cancer Institute; Grant number: P01 CA089392; Grant sponsor: NIH contract 'High throughput genotyping for studying the genetic contributions to human disease'; Grant number: HHSN268200782096C; Grant sponsor: Department of Veterans Affairs (the VA Alcohol Research Center, the VA National Center for PTSD, and the Depression REAP).

Conflicts of interest

There are no conflicts of interest.

References

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Press.

Bierut LJ, Agrawal A, Bucholz KK, Doheny KF, Laurie C, Pugh E, *et al.* (2010). A genome-wide association study of alcohol dependence. *Proc Natl Acad Sci USA* **107**:5082–5087.

Blanco P, Sargent CA, Boucher CA, Mitchell M, Affara NA (2000). Conservation of PCDHX in mammals; expression of human X/Y genes predominantly in brain. *Mamm Genome* **11**:906–914.

Carrel L, Willard HF (2005). X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* **434**:400–404.

Chan AW (1985). Alcoholism and epilepsy. *Epilepsia* **26**:323–333.

Chih B, Engelman H, Scheiffele P (2005). Control of excitatory and inhibitory synapse formation by neuroligins. *Science* **307**:1324–1328.

Cohen J (1988). *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum.

Craig AM, Kang Y (2007). Neurexin-neuroligin signaling in synapse development. *Curr Opin Neurobiol* **17**:43–52.

De Oliveira EH, Neusser M, Muller S (2012). Chromosome evolution in new world monkeys (platyrrhini). *Cytogenet Genome Res* **137**:259–272.

Dean C, Dresbach T (2006). Neuroligins and neurexins: linking cell adhesion, synapse formation and cognitive function. *Trends Neurosci* **29**:21–29.

Edenberg HJ, Koller DL, Xuei X, Wetherill L, McClintick JN, Almasy L, *et al.* (2010). Genome-wide association study of alcohol dependence implicates a region on chromosome 11. *Alcohol Clin Exp Res* **34**:840–852.

Graf ER, Zhang X, Jin SX, Linhoff MW, Craig AM (2004). Neurexins induce differentiation of GABA and glutamate postsynaptic specializations via neuroligins. *Cell* **119**:1013–1026.

Heath AC, Slutske W, PAF Madden (1998). Gender differences in the genetic contribution to alcoholism risk and alcohol consumption patterns. In: Wilsnack RW, Wilsnack SC, editors. *Gender and alcohol: individual and social perspectives*. Rutgers, NJ: Rutgers University Press. pp. 114–149.

Heath AC, Whitfield JB, Martin NG, Pergadia ML, Goate AM, Lind PA, *et al.* (2011). A quantitative-trait genome-wide association study of alcoholism risk in the community: findings and implications. *Biol Psychiatry* **70**: 513–518.

Helena Mangs A, Morris BJ (2007). The Human Pseudoautosomal Region (PAR): origin, function and future. *Curr Genomics* **8**:129–136.

Ichtchenko K, Nguyen T, Sudhof TC (1996). Structures, alternative splicing, and neurexin binding of multiple neuroligins. *J Biol Chem* **271**: 2676–2682.

Jamain S, Quach H, Betancur C, Rastam M, Colineaux C, Gillberg IC, *et al.* (2003). Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet* **34**:27–29.

Laumonnier F, Bonnet-Brilhault F, Gomot M, Blanc R, David A, Moizard MP, *et al.* (2004). X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. *Am J Hum Genet* **74**:552–557.

Lawson-Yuen A, Saldivar JS, Sommer S, Picker J (2008). Familial deletion within NLGN4 associated with autism and Tourette syndrome. *Eur J Hum Genet* **16**:614–618.

Lin M, Hrabovsky A, Pedrosa E, Wang T, Zheng D, Lachman HM (2012). Allele-biased expression in differentiating human neurons: implications for neuropsychiatric disorders. *PLoS One* **7**:e44017.

Nyholt DR (2004). A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* **74**:765–769.

Prange O, Wong TP, Gerrow K, Wang YT, El-Husseini A (2004). A balance between excitatory and inhibitory synapses is controlled by PSD-95 and neuroligin. *Proc Natl Acad Sci USA* **101**:13915–13920.

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**:904–909.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al.* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**:559–575.

Schumann G, Coin LJ, Lourdusamy A, Charoen P, Berger KH, Stacey D, *et al.* (2011). Genome-wide association and genetic functional studies identify autism susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. *Proc Natl Acad Sci USA* **108**:7119–7124.

Sudhof TC (2008). Neuroligins and neurexins link synaptic function to cognitive disease. *Nature* **455**:903–911.

Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, *et al.* (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* **320**:539–543.

Williams NA, Close JP, Giouzeli M, Crow TJ (2006). Accelerated evolution of protocadherin11X/Y: a candidate gene-pair for cerebral asymmetry and language. *Am J Med Genet B Neuropsychiatr Genet* **141B**:623–633.

Wilson ND, Ross LJ, Crow TJ, Volpi EV (2006). PCDH11 is X/Y homologous in *Homo sapiens* but not in *Gorilla gorilla* and *Pan troglodytes*. *Cytogenet Genome Res* **114**:137–139.

Yan J, Oliveira G, Coutinho A, Yang C, Feng J, Katz C, *et al.* (2005). Analysis of the neuroligin 3 and 4 genes in autism and other neuropsychiatric patients. *Mol Psychiatry* **10**:329–332.

Zinn AR, Roeltgen D, Stefanatos G, Ramos P, Elder FF, Kushner H, *et al.* (2007). A Turner syndrome neurocognitive phenotype maps to Xp22.3. *Behav Brain Funct* **3**:24.

Zuo L, Gelernter J, Zhang CK, Zhao H, Lu L, Kranzler HR, *et al.* (2011a). Genome-wide association study of alcohol dependence implicates KIAA0040 on chromosome 1q. *Neuropsychopharmacology* **37**:557–566.

Zuo L, Zhang CK, Wang F, Li CS, Zhao H, Lu L, *et al.* (2011b). A novel, functional and replicable risk gene region for alcohol dependence identified by genome-wide association study. *PLoS One* **6**:e26726.

Zuo L, Zhang F, Zhang H, Zhang XY, Wang F, Li CS, *et al.* (2012). Genome-wide search for replicable risk gene regions in alcohol and nicotine co-dependence. *Am J Med Genet B Neuropsychiatr Genet* **159B**:437–444.

Zuo L, Wang F, Zhang XY, Li CSR, Lu L, Ye L, *et al.* (2013a). Genome-wide significant association signals in IPO11-HTR1A region specific for alcohol and nicotine co-dependence. *Alcohol Clin Exp Res* **37**:730–739, .

Zuo L, Zhang H, Malison RT, Li CSR, Zhang XY, Wang F, *et al.* (2013b). Rare ADH variant constellations are specific for alcohol dependence. *Alcohol Alcohol* **48**:9–14.