

Analysis of Rare Variants in the Alcohol Dependence Candidate Gene *GATA4*

Franziska Degenhardt, Laurenz Krämer, Josef Frank, Jens Treutlein, Stefanie Heilmann-Heimbach, Julian Hecker, Heide Löhlein Fier, Maren Lang, Stephanie H. Witt, Anna C. Koller, Karl Mann, Sabine Hoffmann, Falk Kiefer, Rainer Spanagel, Marcella Rietschel, and Markus M. Nöthen

Background: Common variants in the gene GATA binding protein 4 (*GATA4*) show association with alcohol dependence (AD). The aim of this study was to identify rare variants in *GATA4* in order to elucidate the role of this gene in AD susceptibility. Identification of rare variants may provide a more complete picture of the allelic architecture at this risk locus.

Methods: Sanger sequencing of all 6 coding exons of *GATA4* was performed in 528 patients and 517 controls. Four in silico prediction tools were used to determine the effect of a DNA variant on the amino acid sequence and protein function. Five variants were included in the replication step. Of these, 4 were successfully genotyped in our replication cohort of 655 patients and 1,501 controls. All patients fulfilled DSM-IV criteria for AD, and all individuals were of German descent.

Results: In the discovery step, 19 different heterozygous variants were identified. Four patient-specific and potentially functionally relevant variants were followed up. Only the variant S379S (c.1137C>T) remained patient specific (1/1,166 patients vs. 0/1,997 controls). None of the variants showed a statistically significant association with AD.

Conclusions: The present study elucidated the role of *GATA4* in AD susceptibility by identifying rare variants via Sanger sequencing and subsequent replication. Although novel patient-specific rare variants of *GATA4* were identified, none received support in the independent replication step. However, given previous robust findings of association with common variants, *GATA4* remains a promising candidate gene for AD.

Key Words: *GATA4*, Genetic Risk Factor, Rare Variants, Common Variants, Alcohol Dependence.

From the Institute of Human Genetics (FD, LK, SH-H, JH, HLF, ACK, MMN), University of Bonn, Bonn, Germany; Department of Genomics (FD, LK, SH-H, JH, HLF, ACK, MMN), Life and Brain Center, University of Bonn, Bonn, Germany; Department of Genetic Epidemiology in Psychiatry (JF, JT, ML, SHW, MR), Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany; Institute of Genomic Mathematics (JH, HLF), University of Bonn, Bonn, Germany; Harvard T.H. Chan School of Public Health (HLF), Harvard University, Boston, Massachusetts; Department of Addictive Behavior and Addiction Medicine (KM, SH, FK), Central Institute of Mental Health, Heidelberg University, Mannheim, Germany; and Institute of Psychopharmacology (RS), Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany.

Received for publication December 7, 2015; accepted May 8, 2016.

Reprint requests: Franziska Degenhardt, MD, Institute of Human Genetics, Department of Genomics, Life and Brain Center, Sigmund-Freud-Straße 25, 53127 Bonn, Germany; Tel.: +49-(0)228-6885-433; Fax: +49-(0)228-6885-401; E-mail: franziska.degenhardt@uni-bonn.de

Copyright © 2016 The Authors. *Alcoholism: Clinical and Experimental Research* published by Wiley Periodicals, Inc. on behalf of Research Society on Alcoholism.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

DOI: 10.1111/acer.13125

Alcohol Clin Exp Res, Vol 40, No 8, 2016; pp 1627–1632

ALCOHOL DEPENDENCE (AD) has a multifactorial etiology and an estimated heritability of 40 to 60% (Agrawal and Lynskey, 2008; Enoch and Goldman, 2001; Rietschel and Treutlein, 2013). Candidate-, linkage-, and genome-wide association studies (GWAS) of AD have identified multiple candidate genes (Samochowiec et al., 2014). At the time of writing, the best replicated risk factors in the European population are genetic variants in the alcohol dehydrogenase (*ADH*) gene cluster. Identification of these variants was initiated more than 15 years ago, when large-scale linkage studies implicated a risk locus on chromosome 4q (Long et al., 1998; Prescott et al., 2006; Reich et al., 1998; Samochowiec et al., 2014). Subsequent GWAS then pinpointed the association signal to the *ADH* gene cluster on chromosome 4q23. The importance of this gene cluster has since been confirmed in several independent GWAS (Frank et al., 2012; Gelernter et al., 2014; Park et al., 2013; Treutlein et al., 2009). Besides providing further genetic evidence for genes already implicated in AD pathogenesis, GWAS facilitate the unraveling of novel genetic risk factors.

One gene of interest is GATA binding protein 4 (*GATA4*) on chromosome 8p23.1, which was first identified in the GWAS of AD by Treutlein and colleagues (2009). The

discovery step of this GWAS included 487 AD patients and 1,358 controls, and the subsequent follow-up study investigated the 139 most promising single nucleotide polymorphisms (SNPs) in 1,024 AD patients and 996 controls. The intronically located *GATA4* variant rs13273672 was among the 15 variants with at least nominal significance in the replication cohort.

Subsequent studies have provided further evidence that *GATA4* is a promising candidate gene for AD. First, the association reported by Treutlein and colleagues (2009) was replicated in an independent GWAS performed by Edenberg and colleagues (2010). In a subcohort comprising patients with early onset AD (≤ 22 years), the SNP rs13273672 achieved a *p*-value of 1.6×10^{-03} (Edenberg et al., 2010). Second, Karpyak and colleagues (2014) genotyped 816 AD patients and 1,248 controls for a total of 11 SNPs located within *GATA4*. Five of the 11 tested SNPs in *GATA4* showed a nominally significant association with AD, although no result withstood correction for multiple testing. Furthermore, a global test performed using a principle component analysis revealed a significant association at the gene level ($p = 0.009$; Karpyak et al., 2014). Third, in the PRE-DICT study, Kiefer and colleagues (2011) found that the *GATA4* variant rs13273672 showed a nominally significant association with relapse to heavy drinking within 12 weeks of treatment. This randomized, double-blind, placebo-controlled multicenter trial included 374 AD patients (Kiefer et al., 2011). Fourth, Jorde and colleagues (2014) genotyped rs13273672 in 81 AD patients, and identified *GATA4* genotype-dependent differences in alcohol cue-induced amygdala activity.

The search for rare variants in *GATA4* may provide a more complete picture of the allelic architecture at this risk locus and identify variants with higher penetrance. The latter might be better suited for functional follow-up studies than common variants with lower penetrance. The aim of this study was to elucidate the role of *GATA4* in AD susceptibility by identifying rare variants. All protein-coding exons of *GATA4* were sequenced in 528 AD patients and 517 controls of German descent. Variants that were both exclusive to patients and predicted by in silico tools to be functionally relevant were then genotyped in an independent cohort of 655 patients and 1,501 controls.

MATERIALS AND METHODS

The study was approved by the respective ethics committees, and all participants provided written informed consent prior to inclusion. All study procedures were performed in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). All participants were of German descent according to self-reported ancestry.

Sample Description-Sanger Sequencing and Replication Cohort

The majority of study participants were genome-wide genotyped as part of previously published studies (see Frank et al., 2012; Treutlein et al., 2009). For these individuals, principal component

analysis or multidimensional scaling was performed, respectively. No population substructure was identified.

Patient Sample

The Sanger sequencing cohort comprised 528 AD patients. The replication study cohort comprised 655 independent AD patients. Patients were recruited through consecutive admissions to psychiatry and addiction medicine departments of several German psychiatric hospitals as described elsewhere (see Frank et al., 2012; Treutlein et al., 2009). All patients fulfilled the DSM-IV criteria (American Psychiatric Association, 1994) for AD and had a history of hospitalization for the treatment or prevention of severe withdrawal symptoms. A more detailed phenotypic description of the sample is provided in Table 1a and 1b.

Control Sample

The Sanger sequencing cohort comprised 517 controls. The replication step cohort comprised 1,501 independent controls. The population-based control samples were recruited from volunteer blood donors in collaboration with the Baden-Württemberg State Association of the German Red Cross. A more detailed phenotypic description of the sample is provided in Table 1a and 1b.

Sanger Sequencing

Primer design was based on the NCBI37/hg19 reference sequence of the longest *GATA4* protein-coding isoform (ENST00000335135). The *GATA4* gene comprises 7 exons. Six of these are coding exons. The 6 coding exons and their flanking sequences (± 10 bp of each exon) were amplified. Sanger sequencing was performed using the service offered by Beckman Coulter Genomics (Takeley, UK). The identified variants were confirmed at the Institute of Human Genetics in Bonn by sequencing the complementary strand of a second, independent amplicon. The obtained nucleotide sequences were analyzed using Seqman™ II (DNASTAR, Madison, WI). Primer sequences are obtainable upon request.

In Silico Analysis of Sequence Variants

To determine the effect of a DNA variant on the amino acid sequence and protein function, the following in silico prediction

Table 1. (a) Discovery Sample-Sample Characteristics. (b) Replication Sample-Sample Characteristics

(a) Variable	Patients (<i>n</i> = 528)	Controls (<i>n</i> = 517)
Mean age (SD)	45.1 (9.0)	45.3 (11.7)
Gender distribution in % female/male	21/79	22/78
Mean severity of AD as measured by ADS (SD)	15.1 (6.7)	N/A
Proportion of individuals with positive family history of AD in %	48	n.a.
(b) Variable	Patients (<i>n</i> = 651)	Controls (<i>n</i> = 1497)
Mean age (SD)	41.0 (10.6)	46.1 (15.1)
Gender distribution in % female/male	0/100	63/37
Mean severity of AD as measured by ADS (SD)	n.a.	N/A
Proportion of individuals with positive family history of AD in %	n.a.	n.a.

n.a., not available; N/A, not applicable; SD, standard deviation; AD, alcohol dependence; ADS, alcohol dependence scale.

tools were used: (i) MutationTaster (Schwarz et al., 2010; <http://mutationtaster.org/>; NCBI37/Ensembl 69); (ii) PolyPhen-2 version 2.2.2 (Adzhubei et al., 2013; [http://genet\(19\)ics.bwh.harvard.edu/pph2/](http://genet(19)ics.bwh.harvard.edu/pph2/)); (iii) SIFT (Ng and Henikoff, 2001; <http://sift.jcvi.org/>; Ensemble 66); (iv) and PROVEAN (Choi and Chan, 2015; <http://provean.jcvi.org/index.php>; Ensemble 66).

Replication Analysis-Selection of Variants and Genotyping

Five variants were selected for genotyping using the following criteria: (i) the variant was not detected in the Sanger sequencing control cohort; and (ii) the variant was predicted to be functionally relevant by at least 1 prediction tool. The primer molecules for the multiplex reaction were designed using the Assay Design Suite tool (www.agenacx.com; Agena Bioscience, Inc., San Diego, CA). Assay designs were successful for all 5 variants. Primer sequences and assay conditions are available upon request.

Replication Analysis-Quality Control

Five variants (c.939G>T, c.627C>T, c.699G>A, c.1137C>T, and g.73282C>A) were genotyped using Agena Bioscience’s Mass Array System and iPlex Gold reagents in accordance with the manufacturer’s instructions. Genotyping was performed at the Department of Genomics, Life & Brain Center at the University of Bonn.

Only variants with a call rate of >90% were included in the subsequent analyses. Individuals with a DNA call rate of <90% were removed from the data set. The positive control for the variant g.73282C>A was not detected in all 384-well plates, and this variant was therefore excluded from the study. After quality control, data for 4 SNPs, as obtained from 651 patients and 1,497 controls, were available for analysis.

Statistical Analysis

Fisher’s exact test (2 sided) was used to test for an association between the single variants and AD.

RESULTS

Sanger Sequencing

All 6 coding exons of *GATA4* (i.e., exons 2 to 7) were sequenced. High-quality sequence data were available from: (i) 484 patients and 513 controls for exon 2; (ii) 513 patients and 511 controls for exon 3; (iii) 518 patients and 501 controls for exon 4; (iv) 515 patients and 511 controls for exon 5; (v) 515 patients and 500 controls for exon 6; and (vi) 507 patients and 513 controls for exon 7.

Overall, 19 different heterozygous variants were identified. One of these is a known common SNP (rs3729856). No association was detected between rs3729856 and AD. The remaining 18 variants are listed in Table 2. These comprised: (i) 8 nonsynonymous variants; (ii) 7 synonymous variants; and (iii) 3 intronic variants (Table 2). Of these, the following 7 were detected exclusively among our discovery patient cohort: (i) 3 nonsynonymous variants (E313D, A343T, and A346V); (ii) 3 synonymous variants (D209D, T233T, and S379S); and (iii) 1 intronic variant (g.73282C>A).

Table 2. GATA4 Variants Identified in the Sanger Sequencing Cohort

Exon	Position	Nucleotide change	Amino acid change	Patients	Controls	PolyPhen-2 HumVar	PolyPhen-2 HumDiv	MutationTaster	SIFT	PROVEAN	Functionally relevant	
2	chr8:11566585G>C	Nonsynonymous c.34G>C	G12R	0/484	1/513	Prob. dam.	Dis. caus.	Dis. caus.	Damaging	Deleterious	5/5	
			A66P	0/484	1/513	Benign	Polymorphism	Tolerated	Neutral	0/5		
			G88R	0/484	1/513	Pos. dam.	Polymorphism	Damaging	Neutral	3/5		
			T277P	0/518	1/501	Prob. dam	Dis. caus.	Damaging	Deleterious	5/5		
			E313D	1/515	0/511	Benign	Dis. caus.	Damaging	Deleterious	3/5		
			A343T	1/515	0/500	Benign	Polymorphism	Tolerated	Neutral	0/5		
			A346V	1/515	0/500	Benign	Polymorphism	Tolerated	Neutral	0/5		
2	chr8:11566169C>T	Synonymous c.348C>T	S116S	0/484	1/513	N/A	Dis. caus.	Dis. caus.	N/A	N/A	1/1	
			A139A	0/484	1/513	N/A	Polymorphism	Polymorphism	N/A	N/A	0/1	
			D209D	1/513	0/511	N/A	Dis. caus.	Damaging	N/A	N/A	1/1	
			T233T	8/513	0/511	N/A	Dis. caus.	Damaging	N/A	N/A	1/1	
			C224C	1/518	4/501	N/A	Dis. caus.	Damaging	N/A	N/A	1/1	
			S379S	1/515	0/500	N/A	Dis. caus.	Damaging	N/A	N/A	1/1	
			A388A	0/507	1/513	N/A	Dis. caus.	Damaging	N/A	N/A	1/1	
3	chr8:116066602	Intronic g.72135C>T		0/513	1/511	N/A	Polymorphism	Polymorphism	N/A	N/A	0/1	
				chr8:11607749	g.73282C>A		1/518	0/501	N/A	N/A	N/A	1/1
				chr8:11614440	g.79973C>G		0/515	1/500	N/A	N/A	N/A	1/1

A total of 18 different heterozygous rare variants were identified in the discovery step. All chromosomal positions are given according to NCBI37/hg19. N/A, not applicable; prob. dam., probably damaging; dis. caus., disease causing; functionally relevant, number of in silico tools that predict the variant to be at least probably disease causing.

Of the 7 patient-specific variants, the nonsynonymous variant E313D in exon 5 was predicted to be functionally relevant by SIFT, PROVEAN, and MutationTaster. The remaining 2 patient-specific, nonsynonymous variants were predicted to be benign by all prediction tools applied. All 3 patient-specific synonymous variants were predicted to be functionally relevant by MutationTaster. The PolyPhen-2, SIFT, and PROVEAN tools are unable to predict the potential functional relevance of synonymous or intronic variants. MutationTaster predicted the intronic variant g.73282C>A to be disease causing.

With the exception of the synonymous variant T233T (detected in 8/513 patients; 1.56%), all patient-specific variants were single observations. No patient or control carried more than 1 variant in *GATA4*.

The following 5 patient-specific and potentially functionally relevant variants were selected for follow-up in the independent replication cohort: (i) D209D and T233T in exon 3; (ii) E313D in exon 5; (iii) S379S in exon 6; and (iv) the intronic variant g.73282C>A.

Replication Analysis

Four of the 5 selected variants were genotyped successfully. The assay for g.73282C>A did not detect the positive control reliably, and this variant was therefore excluded from the downstream analysis. The results of the replication analysis are shown in Table 3. Only the synonymous variant S379S (c.1137C>T) remained patient specific (1/1,166 patients vs. 0/1,997 controls in the combined sample; $p > 0.05$). No other patient from the replication cohort carried the S379S variant. In the replication cohort, 12 T233T (c.699G>A) carriers were detected among the patients (1.84%). In total, 17 T233T variant carriers were detected among the controls (1.14%). For D209D and E313D, no additional variant carriers were detected in the replication patient cohort. However, D209D and E313D variant carriers were identified in the replication control cohort. None of the 4 investigated variants showed a nominally significant association with AD in the replication cohort.

DISCUSSION

The present study detected no significant association between AD and rare variants in *GATA4*. The only variant found exclusively in patients was S379S (c.1137C>T) (1/1,166 patients vs. 0/1,997 controls in the combined sample). This mutation is not listed in ClinVar (<http://www.ncbi.nlm.nih.gov/clinvar/>). Within the cohorts reported in the Exome Aggregation Consortium (ExAC), the variant c.1137C>T was identified in 1/33,075 non-Finnish European samples (Exome Aggregation Consortium (ExAC), Cambridge, MA (URL: <http://exac.broadinstitute.org>) [(September, 2015) accessed]). Given the rarity of c.1137C>T, no definite conclusions can yet be drawn concerning its relevance to AD susceptibility. Replication analyses in independent study cohorts are warranted to elucidate its role in AD pathogenesis.

Notably, a marked difference in the frequency of the variant c.669G>A was observed between the Sanger sequencing control cohort (0/511) and the replication control cohort (17/1,497). Although this variant was associated with AD in the Sanger sequencing cohort, this association was not replicated. To test whether this frequency difference was attributable to the different genotyping methods, we first re-genotyped the Sanger sequencing control cohort for c.669G>A using the Sequenom assay applied in the replication study and then Sanger-sequenced exon 3 in all c.669G>A variant carriers identified in the replication cohort. The re-genotyping of the Sanger sequencing control cohort confirmed the absence of c.669G>A variant carriers, and all variant carriers identified in the replication study were unambiguously detected using Sanger sequencing. The possibility that the frequency difference was due to the different genotyping methods can therefore be excluded. Within the cohorts reported in ExAC, the variant c.669G>A was identified in 263/33,364 non-Finnish European samples (including 2 homozygous variants). The possibility that (as yet unknown) confounding factors explain the absence of the variant c.669G>A in the Sanger sequencing cohort cannot be excluded. However, no pronounced phenotypic differences

Table 3. Replication Analysis of Selected *GATA4* Variants

Exon	Position	Nucleotide change	Amino acid change	Patients-Sanger sequencing	Patients-replication	Controls-Sanger sequencing	Controls-replication
5	chr8:11612584G>T	<i>Nonsynonymous</i> c.939G>T	E313D	1/515	0/651	0/511	1/1,497
3	chr8:11606438C>T	<i>Synonymous</i> c.627C>T	D209D	1/513	0/651	0/511	2/1,497
3	chr8:11606510G>A	c.699G>A	T233T	8/513	12/651	0/511	17/1,497
6	chr8:11614583C>T	c.1137C>T	S379S	1/515	0/651	0/500	0/1,497
4	chr8:11607749C>A	<i>Intronic</i> g.73282C>A		1/518	Technical failure	0/501	Technical failure

Five variants were included in the replication analysis. The intronic variant was excluded from downstream analysis due to technical failure. All chromosomal positions are given according to NCBI37/hg19.

were present in the investigated inhouse control cohorts. The frequency differences observed between the present Sanger sequencing control cohort, the present replication-genotyping control cohort, and the cohorts of the ExAC demonstrate the importance of replication analyses.

The present study had several limitations. First, due to the limited sample size of our initial Sanger sequencing cohort and the genotyping of only selected variants in the replication cohort, it is likely that only a small fraction of the rare variants that contribute to the allelic spectrum of *GATA4* were detected. Sequencing studies in larger AD cohorts are warranted to provide a more complete overview of rare sequencing variants in this gene. Second, none of the controls were screened for AD, which may have led to false-negative findings for some genuinely associated regions. Third, our analysis focused on exonic variants. Further studies are required to elucidate the contribution of noncoding variants to the allelic spectrum of *GATA4* in patients with AD. Fourth, as the aim of the present study was to identify rare risk-associated variants, the study design selected against variants identified in the Sanger sequencing control cohort. As a consequence of this, the study design precluded the identification of potentially protective variants in *GATA4*. Fifth, we did not perform burden analysis in the Sanger sequencing cohort to find the excess of risk or protective variants in the AD patients as our study design did not allow for a replication of the results.

In conclusion, although patient-specific rare variants of *GATA4* were identified in similarly sized patient versus control samples, none received support in an independent replication step. However, given previous robust association findings for common variants (12 to 16), *GATA4* remains a promising candidate gene for AD, and thus warrants further investigation.

ACKNOWLEDGMENTS

We thank all patients and controls for their participation. The authors thank the Exome Aggregation Consortium and groups that provided exome variant data for comparison. A full list of contributing groups can be found at <http://exac.broadinstitute.org/about>. We thank Christine Schmäler for her careful proofreading of the manuscript.

AUTHOR CONTRIBUTIONS

FD, MR, RS, and MMN were responsible for the study concept and design. JF, JT, SHW, ML, KM, SH, FK, and MR contributed to the recruitment of the patient and control cohorts. FD, LK, SH-H, and ACK performed the sequencing and genotyping analyses. HLF and JH assisted with the data analyses. FD drafted the manuscript. MR and MMN provided a critical revision of the manuscript for intellectual content. All authors have critically reviewed the content of the final manuscript and approved its submission for publication.

FINANCIAL SUPPORT AND DECLARATION OF COMPETING INTERESTS

This study was funded by the German Federal Ministry of Education and Research (BMBF) through the Integrated Network sysMedAlcoholism, under the auspices of the e:Med Programme; grant 01ZX1311A to Markus M. Nöthen and Marcella Rietschel and grant 01ZX1311A to Rainer Spanagel. Markus M. Nöthen received support from the Alfried Krupp von Bohlen und Halbach-Stiftung, and is a member of the DFG-funded Excellence Cluster ImmunoSensation. Franziska Degenhardt, Heide Löhlein Fier, and Stefanie Heilmann-Heimbach received support from the BONFOR Programme of the University of Bonn, Germany. These funding sources had no involvement in the study design; the collection, analysis, and interpretation of data; the writing of the report; or the decision to submit the paper for publication.

REFERENCES

- Adzhubei I, Jordan DM, Sunyaev SR (2013) Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet* Chapter 7:Unit7.20.
- Agrawal A, Lynskey MT (2008) Are there genetic influences on addiction: evidence from family, adoption and twin studies. *Addiction* 103:1069–1081.
- American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. American Psychiatric Association, Washington, DC.
- Choi Y, Chan AP (2015) PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* 31:2745–2747.
- Edenberg HJ, Koller DL, Xuei X, Wetherill L, McClintick JN, Almasy L, Bierut LJ, Bucholz KK, Goate A, Aliev F, Dick D, Hesselbrock V, Hinrichs A, Kramer J, Kuperman S, Nurnberger JI Jr, Rice JP, Schuckit MA, Taylor R, Todd Webb B, Tischfield JA, Porjesz B, Foroud T (2010) Genome-wide association study of alcohol dependence implicates a region on chromosome 11. *Alcohol Clin Exp Res* 34:840–852.
- Enoch MA, Goldman D (2001) The genetics of alcoholism and alcohol abuse. *Curr Psychiatry Rep* 3:144–151.
- Frank J, Cichon S, Treutlein J, Ridinger M, Mattheisen M, Hoffmann P, Herms S, Wodarz N, Soyka M, Zill P, Maier W, Mössner R, Gaebel W, Dahmen N, Scherbaum N, Schmäler C, Steffens M, Lucae S, Ising M, Müller-Myhsok B, Nöthen MM, Mann K, Kiefer F, Rietschel M (2012) Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster. *Addict Biol* 17:171–180.
- Gelernter J, Kranzler HR, Sherva R, Almasy L, Koesterer R, Smith AH, Anton R, Preuss UW, Ridinger M, Rujescu D, Wodarz N, Zill P, Zhao H, Farrer LA (2014) Genome-wide association study of alcohol dependence: significant findings in African- and European-Americans including novel risk loci. *Mol Psychiatry* 19:41–49.
- Jorde A, Bach P, Witt SH, Becker K, Reinhard I, Vollstädt-Klein S, Kirsch M, Hermann D, Charlet K, Beck A, Wimmer L, Frank J, Treutlein J, Spanagel R, Mann K, Walter H, Heinz A, Rietschel M, Kiefer F (2014) Genetic variation in the atrial natriuretic peptide transcription factor *GATA4* modulates amygdala responsiveness in alcohol dependence. *Biol Psychiatry* 75:790–797.
- Karpyak VM, Winham SJ, Biernacka JM, Cunningham JM, Lewis KA, Geske JR, Colby CL, Abulseoud OA, Hall-Flavin DK, Loukianova LL, Schneekloth TD, Frye MA, Heit JA, Mrazek DA (2014) Association of *GATA4* sequence variation with alcohol dependence. *Addict Biol* 19:312–315.

- Kiefer F, Witt SH, Frank J, Richter A, Treutlein J, Lemenager T, Nöthen MM, Cichon S, Batra A, Berner M, Wodarz N, Zimmermann US, Spanagel R, Wiedemann K, Smolka MN, Heinz A, Rietschel M, Mann K (2011) Involvement of the atrial natriuretic peptide transcription factor GATA4 in alcohol dependence, relapse risk and treatment response to acamprosate. *Pharmacogenomics J* 11:368–374.
- Long JC, Knowler WC, Hanson RL, Robin RW, Urbanek M, Moore E, Bennett PH, Goldman D (1998) Evidence for genetic linkage to alcohol dependence on chromosomes 4 and 11 from an autosome-wide scan in an American Indian population. *Am J Med Genet* 81:216–221.
- Ng PC, Henikoff S (2001) Predicting deleterious amino acid substitutions. *Genome Res* 11:863–874.
- Park BL, Kim JW, Cheong HS, Kim LH, Lee BC, Seo CH, Kang TC, Nam YW, Kim GB, Shin HD, Choi IG (2013) Extended genetic effects of ADH cluster genes on the risk of alcohol dependence: from GWAS to replication. *Hum Genet* 132:657–668.
- Prescott CA, Sullivan PF, Kuo PH, Webb BT, Vittum J, Patterson DG, Thiselton DL, Myers JM, Devitt M, Halberstadt LJ, Robinson VP, Neale MC, van den Oord EJ, Walsh D, Riley BP, Kendler KS (2006) Genome-wide linkage study in the Irish affected sib pair study of alcohol dependence: evidence for a susceptibility region for symptoms of alcohol dependence on chromosome 4. *Mol Psychiatry* 11:603–611.
- Reich T, Edenberg HJ, Goate A, Williams JT, Rice JP, Van Eerdewegh P, Foroud T, Hesselbrock V, Schuckit MA, Bucholz K, Porjesz B, Li TK, Conneally PM, Nurnberger JI Jr, Tischfield JA, Crowe RR, Cloninger CR, Wu W, Shears S, Carr K, Crose C, Willig C, Begleiter H (1998) Genome-wide search for genes affecting the risk for alcohol dependence. *Am J Med Genet* 81:207–215.
- Rietschel M, Treutlein J (2013) The genetics of alcohol dependence. *Ann N Y Acad Sci* 1282:39–70.
- Samochowiec J, Samochowiec A, Puls I, Bienkowski P, Schott BH (2014) Genetics of alcohol dependence: a review of clinical studies. *Neuropsychobiology* 70:77–94.
- Schwarz JM, Rödelsperger C, Schuelke M, Seelow D (2010) MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods* 7:575–576.
- Treutlein J, Cichon S, Ridinger M, Wodarz N, Soyka M, Zill P, Maier W, Moessner R, Gaebel W, Dahmen N, Fehr C, Scherbaum N, Steffens M, Ludwig KU, Frank J, Wichmann HE, Schreiber S, Dragano N, Sommer WH, Leonardi-Essmann F, Lourdasamy A, Gebicke-Haerter P, Wienker TF, Sullivan PF, Nöthen MM, Kiefer F, Spanagel R, Mann K, Rietschel M (2009) Genome-wide association study of alcohol dependence. *Arch Gen Psychiatry* 66:773–784.