

Genome-wide Association Study of Alcohol Dependence

Jens Treutlein, PhD*; Sven Cichon, PhD*; Monika Ridinger, MD*; Norbert Wodarz, MD; Michael Soyka, MD; Peter Zill, PhD; Wolfgang Maier, MD; Rainald Moessner, MD; Wolfgang Gaebel, MD; Norbert Dahmen, MD; Christoph Fehr, MD; Norbert Scherbaum, MD; Michael Steffens, MD; Kerstin U. Ludwig, MSc; Josef Frank, MA; H. Erich Wichmann, MD, PhD; Stefan Schreiber, MD; Nico Dragano, PhD; Wolfgang H. Sommer, MD, PhD; Fernando Leonardi-Essmann, MA; Anbarasu Lourdasamy, PhD; Peter Gebicke-Haerter, PhD; Thomas F. Wienker, MD; Patrick F. Sullivan, MD; Markus M. Nöthen, MD; Falk Kiefer, MD; Rainer Spanagel, PhD*; Karl Mann, MD*; Marcella Rietschel, MD*

Context: Alcohol dependence is a serious and common public health problem. It is well established that genetic factors play a major role in the development of this disorder. Identification of genes that contribute to alcohol dependence will improve our understanding of the mechanisms that underlie this disorder.

Objective: To identify susceptibility genes for alcohol dependence through a genome-wide association study (GWAS) and a follow-up study in a population of German male inpatients with an early age at onset.

Design: The GWAS tested 524 396 single-nucleotide polymorphisms (SNPs). All SNPs with $P < 10^{-4}$ were subjected to the follow-up study. In addition, nominally significant SNPs from genes that had also shown expression changes in rat brains after long-term alcohol consumption were selected for the follow-up step.

Setting: Five university hospitals in southern and central Germany.

Participants: The GWAS included 487 male inpatients with alcohol dependence as defined by the DSM-IV and an age at onset younger than 28 years and 1358 population-based control individuals. The follow-up study included 1024 male inpatients and

996 age-matched male controls. All the participants were of German descent.

Main Outcome Measures: Significant association findings in the GWAS and follow-up study with the same alleles.

Results: The GWAS produced 121 SNPs with nominal $P < 10^{-4}$. These, together with 19 additional SNPs from homologues of rat genes showing differential expression, were genotyped in the follow-up sample. Fifteen SNPs showed significant association with the same allele as in the GWAS. In the combined analysis, 2 closely linked intergenic SNPs met genome-wide significance (rs7590720, $P = 9.72 \times 10^{-9}$; rs1344694, $P = 1.69 \times 10^{-8}$). They are located on chromosome region 2q35, which has been implicated in linkage studies for alcohol phenotypes. Nine SNPs were located in genes, including the *CDH13* and *ADH1C* genes, that have been reported to be associated with alcohol dependence.

Conclusions: This is the first GWAS and follow-up study to identify a genome-wide significant association in alcohol dependence. Further independent studies are required to confirm these findings.

Arch Gen Psychiatry. 2009;66(7):773-784

ALCOHOL DEPENDENCE IS characterized by a cluster of cognitive, behavioral, and physiologic symptoms, with an affected individual continuing to drink despite significant alcohol-induced impairment or distress. Because alcohol affects most human organs, its abuse is associated with a wide range of types of physical, mental, and social harm. According to the World Health Organization,¹ alcohol abuse constitutes a serious public health problem worldwide, accounting for 4% of the global burden, a burden comparable with the

death and disability attributable to tobacco use and hypertension.^{1,2}

Alcohol dependence is a phenotypically heterogeneous disorder that runs in families and has high genetic loading. Twin and adoption studies³⁻⁵ have shown that 40% to 60% of the interindividual phenotypic variance is accounted for by genetic factors. In congruence with the observed phenotypic heterogeneity of alcohol dependence, vulnerability to alcohol dependence on the molecular level is thought to be mediated by many genetic loci of small to modest effects in Europeans.⁶⁻¹⁴ Despite strenuous efforts and numerous link-

Author Affiliations are listed at the end of this article.

* Indicates co-primary authorship.

age and candidate gene studies, identification of the underlying susceptibility genes has proved to be difficult. Genome-wide association studies (GWASs) conducted in other complex disorders have been shown to be a successful tool in identifying underlying susceptibility genes (for all published GWASs, see <http://www.genome.gov/26525384>) and have already resulted in the identification of genetic susceptibility variants for psychiatric disorders such as nicotine addiction, schizophrenia, and bipolar disorder,¹⁵⁻²⁴ which were mostly not detected using other diagnosis-based discovery approaches. For alcohol dependence, only one GWAS using pooled DNA samples has been published to date, and it has proposed several new susceptibility loci for alcohol dependence. Their products are implicated in cellular signaling, gene regulation, development, and cell adhesion.²⁵ Convergent translational approaches integrate genetic findings from animal models with a candidate gene or GWAS approach in humans and have proved to be very successful.^{26,27} The goal of the present study is to conduct a GWAS and a follow-up study for alcohol dependence using individual genotyping in samples stratified for homogeneity with respect to sex, ethnicity, age at onset, and recruitment procedures. To increase the explanatory power of the findings, we applied a convergent functional genomics²⁸ approach to integrate findings from gene expression data in alcohol-dependent rats with the GWAS findings.

METHODS

HUMAN STUDIES

Patients

Patients were recruited from consecutive admissions to the psychiatric and addiction medicine departments of 5 different study centers at university hospitals across southern and central Germany: Regensburg, Mannheim, Munich, Bonn/Essen/Dusseldorf/Homburg, and Mainz. These centers are members of the German Addiction Research Network (GARN; <http://www.bw-suchtweb.de>), for which one of us (K.M.) is the spokesperson. All the patients included in the study had alcohol dependence of such severity that hospitalization for the treatment or prevention of severe withdrawal symptoms was warranted. All the patients received a diagnosis of alcohol dependence (per *DSM-IV* criteria) by the consensus of 2 clinical psychiatrists. In the German Addiction Research Network study, *DSM-IV* criteria were ascertained systematically through the use of semistructured and independently rated interviews conducted by trained staff members.

In Munich and Bonn, patients were assessed using the Semi-Structured Assessment for the Genetics of Alcoholism²⁹; in Mainz, the Composite International Diagnostic Interview³⁰ was used; and in Regensburg, Essen/Dusseldorf/Homburg, and Mannheim, the *Structured Clinical Interview for DSM-IV*³¹ was used. The latter tool was also applied in Munich. All the patients and controls were of self-reported German ancestry. The study was approved by all the relevant local ethics committees, and all the participants provided written informed consent.

To increase the power of the study, we increased the homogeneity of the patient sample by including male patients only. Furthermore, only patients with an early age at onset of alcohol dependence were included, a feature that

has higher heritability in males.^{32,33} Patients included in the GWAS (Mannheim, *n* = 98; Bonn/Essen/Dusseldorf/Homburg, *n* = 30; Mainz, *n* = 30; Munich, *n* = 53; and Regensburg, *n* = 276) had an age at onset younger than 28 years (median, 20 years; mean [SD], 21 [4.1] years), defined as the age at which *DSM-IV* criteria for alcohol dependence were fulfilled for the first time.

The follow-up sample consisted of 1024 German males recruited at the same sites and under the same protocol as the patients in the GWAS (Mannheim, *n* = 256; Bonn/Essen/Dusseldorf/Homburg, *n* = 149; Mainz, *n* = 63; Munich, *n* = 136; and Regensburg, *n* = 420). Because most of the patients with an earlier age at onset had already been included in the GWAS, age at onset for the follow-up study was necessarily less stringent and was defined as younger than 45 years (median, 30 years; mean [SD], 29.3 [7] years). Phenotypes and genotypes resulting from the GWAS are stored in a comprehensive database at the Institute for Medical Biometry, Informatics, and Epidemiology in Bonn.

Control Subjects

A total of 1358 control subjects were included in the GWAS. Controls were taken from 3 population-based epidemiologic studies: 487 from PopGen,³⁴ 488 from KORA-gen,³⁵ and 383 from the Heinz Nixdorf Risk Factors, Evaluation of Coronary Calcium, and Lifestyle (RECALL) Study.³⁶ These 3 recruitment areas are located in Schleswig-Holstein (northern Germany), Augsburg (southern Germany), and Essen/Bochum/Mulheim (Ruhr Area, western Germany). The Population-Based Recruitment of Patients and Controls for the Analysis of Complex Genotype-Phenotype Relationships (PopGen) project (http://www.science.ngfn.de/10_233.htm; <http://www.popgen.de>) was initiated to provide all locally prevalent cases with the diseases in question for the disease-oriented projects from the National Genome Research Project (NGFN) and population-based control samples that will ultimately be comprised of 7200 persons. The Cooperative Health Research in the Region of Augsburg (KORA-gen) project (http://www.science.ngfn.de/10_234.htm; <http://www.gsf.de/KORA>), which has evolved from the World Health Organization MONICA (Monitoring of Trends and Determinants of Cardiovascular Disease) study, contains the biosamples, phenotypic characteristics, and environmental parameters of 18 000 adults from Augsburg and the surrounding regions. The biological specimen bank was established to enable researchers to perform epidemiologic research regarding molecular and genetic factors. The Heinz Nixdorf RECALL Study (<http://www.recall-studie.uni-essen.de>) was started in the year 2000 as a prospective cohort study to determine predictors of coronary heart disease. It includes the biosamples, phenotypic characteristics, and environmental parameters of 4500 adults. PopGen and KORA-gen are the two major German biobanking resources available to participants in the NGFN (<http://www.ngfn.de/>) for use as universal controls,³⁷ and they have already been used as samples in other GWASs.^{37,38}

Male control individuals for the follow-up study were drawn from the KORA-gen study (*n* = 382) and from a population-based sample (*n* = 614) that had been collected in the area of Bonn by one of us (M.R.) for use as a control sample for association studies in the framework of the NGFN. Controls were not stratified for alcohol abuse or dependence, which may introduce a conservative bias. Tests for genetic differentiation between German populations have shown low levels of population substructure, thus demonstrating that the German population is an appropriate source for use in association studies of complex diseases.³⁹

DNA Preparation, Genotyping, and Statistical Analysis

Genomic DNA was prepared from whole blood according to standard procedures. For all samples, DNA concentrations were quantified in triplicate measurements using the PicoGreen dsDNA Quantitation Kit (Invitrogen Corporation, Carlsbad, California) and normalized to 50 ng/ μ L at the Central Institute of Mental Health Molecular Genetics Laboratory. For the GWAS, all samples were genotyped individually using HumanHap 550 BeadChips (Illumina Inc, San Diego, California). The patient and the PopGen and KORA-gen samples were genotyped at Illumina Inc, and the Heinz Nixdorf RECALL Study sample was genotyped at the Department of Genomics at the Life and Brain Center, University of Bonn. Genotyping of the follow-up sample was performed by means of primer extension reaction chemistry with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry using the iPLEX Assay (Sequenom Inc, San Diego) at the Life and Brain Center, University of Bonn. Genotyping results were imported into a central computer system for statistical analysis.

Data Analysis and Quality Control

Data analysis and quality control (QC) were performed using the software packages R version 2.5.1 (<http://www.R-project.org>) and PLINK version 1.0.3.⁴⁰ In the GWAS, genotype data were cleaned before analysis by removing single-nucleotide polymorphisms (SNPs) or individuals not fulfilling the QC criteria, which included a SNP call proportion of at least 95%, a subject completeness proportion of at least 95%, a SNP minor allele frequency of at least 0.01, and SNP conformity with Hardy-Weinberg equilibrium expectations ($P \geq .01$ in controls). To correct for cryptic relatedness, all pairs of individuals displaying an identity-by-state value larger than 1.6 were marked, and for each pair, the individual with the lower typing rate was removed from the analysis. Cochran-Armitage trend statistics were used to calculate significant association for autosomal SNPs (eTable 1; for all supplementary materials, see http://www.zi-mannheim.de/pub_gwas.html). To visualize the outcome of the QC steps, Cochran-Armitage P values were depicted in a quantile-quantile plot (Figure 1). We observed good adherence of P values to the line of expectance, which implies that potential spurious associations characterized by an inflation of highly significant P values were successfully removed by the QC measures. The remaining slight deviations from the line of expectance are interpreted to include true genetic effects. Further correction for λ^{41} improved the quantile-quantile plot of Cochran-Armitage P values (eFigure 1).

ANIMAL STUDIES

Subjects

Three groups of 2- to 3-month-old alcohol-preferring rats were used for long-term alcohol consumption and gene expression profiling: male P rats ($n = 15$) (Indiana University, Indianapolis), male HAD rats ($n = 13$) (Indiana University), and male AA rats ($n = 14$) (National Public Health Institute, Helsinki, Finland). Each rat strain shows alcohol preference due to various neurochemical alterations, as indicated by the abbreviations P (preferring), HAD (high alcohol drinking), and AA (Alko alcohol addicted).^{42,43} All experimental procedures were approved by the Committee on Animal Care and Use (Regierungspräsidium Karlsruhe) and were performed in accordance with the local Animal Welfare Act and the European Communities Council Directive of November 24, 1986 (86/609/EEC).

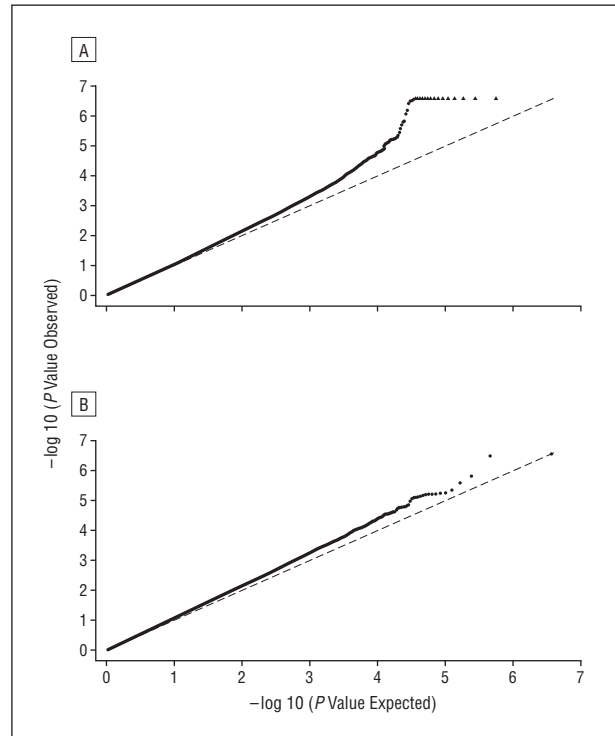


Figure 1. Before (A) and after (B) quality control quantile-quantile plots of Cochran-Armitage trend test P values of autosomal single-nucleotide polymorphisms (SNPs). The SNPs at which the test statistic exceeds 30 are represented by triangles at the top of the plot.

According to the protocol of Vengeliene et al,^{44,45} 8 P rats, 7 HAD rats, and 7 AA rats were given ad libitum access to tap water and to 5% and 20% ethanol solution (vol/vol). All the rats underwent a 2-week deprivation cycle after 8 weeks of continuous alcohol availability. After the deprivation period, rats were given access to alcohol again, and 3 more 2-week deprivation periods were introduced in a random manner (the duration between deprivation periods varied between 4 and 16 weeks). The long-term voluntary alcohol-drinking procedure, including all deprivation phases, lasted 52 weeks. Total ethanol intake (grams per kilogram of body weight per day) was calculated as the daily average across 7 measuring days. For comparison, 3 age- and weight-matched control groups, consisting of 7 P rats, 6 HAD rats, and 7 AA rats, experienced identical handling procedures for the entire duration of the experiment but did not receive alcohol.

Gene Expression Profiling in Alcohol-Preferring Rat Strains

Preparation of brain samples and RNA isolation are described in the eText. Target preparation was performed for individual samples from the caudate putamen and the amygdala using 5 μ g of total RNA. Hybridization to RG_U34A arrays, staining, washing, and scanning of the chips were performed according to the manufacturer's technical manual (Affymetrix, Santa Clara, California).

Data Mining

Microarray Suite 5.0 (Affymetrix)-derived cell intensity files were processed using the R 2.1.1 language and environment (<http://www.R-project.org>) and Bioconductor 1.6⁴⁶ packages. Each array was inspected for regional hybridization bias and

Table 1. SNPs Confirmed in the Follow-up Study: Location According to Chromosomal Bands and Gene Annotation

SNP	Chromosomal Band	Genes ^a
rs1344694	2q35	NA
rs7590720	2q35	NA
rs705648	2q35	Peroxisomal trans-2-enoyl-CoA reductase (<i>PECR</i>)
rs1614972 ^b	4q23	Alcohol dehydrogenase 1C (class I), gamma polypeptide (<i>ADH1C</i>)
rs13362120	5q15	Calpastatin (<i>CAST</i>)
rs13160562	5q15	Endoplasmic reticulum aminopeptidase 1 (<i>ERAP1</i>); calpastatin (<i>CAST</i>)
rs1864982	5q32	Protein phosphatase 2 (formerly 2A), regulatory subunit B, beta isoform (<i>PPP2R2B</i>)
rs6902771	6q25.1	Estrogen receptor 1 (<i>ESR1</i>)
rs729302	7q32.1	NA
rs13273672 ^b	8p23.1	GATA binding protein 4 (<i>GATA4</i>)
rs1487814	11p14.3	NA
rs7138291	12q22	Coiled-coil domain containing 41 (<i>CCDC41</i>)
rs36563	14q24.2	NA
rs11640875 ^b	16q23.3	Cadherin 13, H-cadherin (heart) (<i>CDH13</i>)
rs12388359	Xp22.2	NA

Abbreviations: CoA, coenzyme A; NA, not applicable; SNP, single-nucleotide polymorphism.

^aAnnotation according to SNP database build 129.

^bSelected following the strategy of "rodent candidate gene."

QC parameters as recently described.⁴⁷ Fifty-three arrays (27 from the caudate putamen and 26 from the amygdala) passed through the quality filter and were included in the statistical analysis. Of the 8799 probe sets on the RG_U34A arrays, only those with intensity values greater than 100 in at least 25% of the samples were retained (6344 probe sets). A 3-way analysis of variance (ANOVA) was used to identify differentially expressed genes across strain, brain region, and treatment. The list of genes affected by long-term ethanol consumption included those with a $P < .05$ for treatment from the 3-way analysis of variance. Post hoc analysis was performed via template matching across strains. Correlation coefficients (r) were calculated for consistent upregulation or downregulation by ethanol in the caudate putamen, the amygdala, or both.

Integration of Animal Data Into the Human Study

The SNPs found to be associated with the disorder in the GWAS with $P < 1 \times 10^{-3}$ or found to lie in a SNP cluster (defined as ≥ 2 SNPs with $P \leq 1 \times 10^{-2}$ and at least 1 SNP with $P < 1 \times 10^{-3}$ in a distance of < 30 kilobase [kb]) (examples of which can be seen in eFigures 2, 3, and 4) were taken forward to the follow-up study if they were located in a homologue of a rat gene showing differential brain expression with long-term high ethanol consumption compared with their respective controls without ethanol access. Only 1 SNP was selected for each gene. If more than 1 SNP in the gene fulfilled the previously mentioned criteria, the SNP with the highest number of assigned transcripts (Sullivan et al annotation file: <https://slep.unc.edu/evidence/>) was taken forward. In cases of similar transcript numbers, the SNP with the lowest P value was chosen.

In the animal study, we used a convergent translational approach with 2 strategic lines to integrate the rat data:

1. a qualitative approach looking for orthologous or paralogous genes in the GWAS and expression profiling in rats (see the list of genes in eTable 2 and eTable 3).
2. a (semiquantitative) ranking strategy, as theoretically described by Bertsch et al,²⁸ in which those orthologous or paralogous genes were included. In this strategy, a weighting ratio of 1 (both data sets with equal weights) was used, multiplying the P values of respective rat and human genes.

Using strategy 1, 19 additional SNPs were identified and were ranked according to strategy 2. Because the total number of SNPs

derived through the animal approach was only 19, we eventually decided to include all of them in the follow-up study instead of selecting a small subset based on the ranking order of genes. The 3 genes that were confirmed by the follow-up study (**Table 1**) had been ranked No. 7 (*ADH*) (alcohol dehydrogenase 1C, gamma polypeptide) (OMIM *103730), No. 12 (*CDH*) (cadherin 13) (OMIM *601364), and No. 19 (*GATA*) (gata-binding protein 4) (OMIM *600576) using approach 2. We want to conclude from these data that it is hard to introduce quantitative evaluations in the convergent approaches. Nevertheless, convergent approaches as pursued in this study can provide valuable additional information and apparently become increasingly more popular.

We searched the Gene Ontology database (<http://amigo.geneontology.org/cgi-bin/amigo/go.cgi>) for the annotation of gene products for genes in which confirmed SNPs of the follow-up study are located. The Gene Ontology database provides a standardized vocabulary describing gene products to ensure uniformity across databases.⁴⁸ We investigated whether these genes share distribution patterns in Gene Ontology categories by mapping them to Gene Ontology annotations.

RESULTS

HUMAN STUDIES

We excluded 11 of the 487 GWAS cases with alcohol dependence from further analyses. Of these, 6 DNA samples failed to genotype, and, in 5 individuals, the genome-wide identity-by-state score was greater than 1.60, indicating possible cryptic relatedness. Of the 561 466 SNPs genotyped per individual, 99.6% passed QC (**Figure 2**). The final data set comprised 524 396 SNPs (511 701 autosomal and 12 695 sex-linked markers) in 476 male cases and 1358 male controls. The mean (SD) call rate of the final set of SNPs (cases and controls) was 99.76% (0.45%). A genome-wide overview of the GWAS association results is shown in Figure 2.

We used 2 complementary SNP selection strategies, a "lowest P value" strategy and a "rodent candidate gene" strategy, to prioritize SNPs for follow-up. In total, 139

SNPs were carried forward for genotyping in the follow-up sample, including all 121 SNPs with a $P < 10^{-4}$ (Cochran-Armitage trend test for autosomes and allelic test for X chromosome), of which an assay design was possible for 120 (eTable 1), and an additional 19 SNPs with at least nominal significance, which would not have been accounted for by the lowest P value approach. These 19 SNPs are among 22 SNPs (3 of which were already represented in the aforementioned lowest P value selection) located in human homologues of rat genes showing differential expression in the rat brain after long-term alcohol consumption (for more details see the eText) and, therefore, have a higher a priori probability of being involved in the etiology of alcohol dependence.

In the follow-up study, 16 SNPs showed association with at least nominal significance ($P < .05$, 2-sided), 15 of which (9 intragenic and 6 intergenic) were associated with the same allele as in the GWAS. The number of significantly replicated SNPs is higher than expected by chance ($P = 2.45 \times 10^{-6}$). Three of the intragenic SNPs are derived from the 22 SNPs selected on the basis of the animal model (Table 1). This number is also significantly higher than that expected by chance ($P = 1.69 \times 10^{-2}$).

Combining the data across the 2 samples showed genome-wide significance (Bonferroni-adjusted α level .05/524 000 [$P < 9.5 \times 10^{-8}$])^{49,50} for 2 SNPs, rs7590720 ($P = 9.72 \times 10^{-9}$) and rs1344694 ($P = 1.69 \times 10^{-8}$). These SNPs map to the 3'-flanking region of the gene encoding peroxisomal trans-2-enoyl-coA [coenzyme A] reductase (*PECR*), located on chromosome region 2q35. They are in linkage disequilibrium with each other and with another SNP, rs705648 ($P = 1.78 \times 10^{-6}$), located in the *PECR* gene (rs1344694-rs7590720: $D' = 1.0$, $r^2 = 0.739$; rs1344694-rs705648: $D' = 0.943$, $r^2 = 0.568$; and rs7590720-rs705648: $D' = 0.948$, $r^2 = 0.776$; International HapMap Project).

A further 8 SNPs are located in genes: rs13362120 (combined $P = 1.85 \times 10^{-5}$) in the calpastatin (*CAST*) gene; rs13160562 (combined $P = 7.09 \times 10^{-6}$) in the endoplasmic reticulum aminopeptidase 1 (*ERAPI1*) and calpastatin (*CAST*) genes; rs1864982 (combined $P = 3.46 \times 10^{-6}$) in the protein phosphatase 2 (formerly 2A), regulatory subunit B, beta isoform (*PPP2R2B*) gene; rs6902771 (combined $P = 8.30 \times 10^{-6}$) in the estrogen receptor 1 (*ESR1*) gene; rs7138291 (combined $P = 3.68 \times 10^{-5}$) in the coiled-coil domain containing the 41 (*CCDC41*) gene; rs1614972 (combined $P = 1.41 \times 10^{-4}$) in the alcohol dehydrogenase 1C (*ADH1C*) gene; rs13273672 (combined $P = 4.75 \times 10^{-4}$) located in the GATA binding protein 4 (*GATA4*) gene, and rs11640875 (combined $P = 1.84 \times 10^{-3}$) in the cadherin 13 (*CDH13*) gene. The first 5 SNPs were selected for the follow-up study using the lowest P value strategy and the latter 3 via the rodent candidate gene approach. These findings and those of the intergenic SNPs are provided in Table 1, **Table 2**, eTable 4, and eTable 5.⁵¹ **Figures 3, 4, and 5** show P values of the SNPs located in ± 1 -Mb intervals around the most significant findings at 2q35, 5q32, and Xp22.2.

Specific Gene Ontology database biological process, cellular component, and molecular function terms remained largely unique for each of the genes and were not

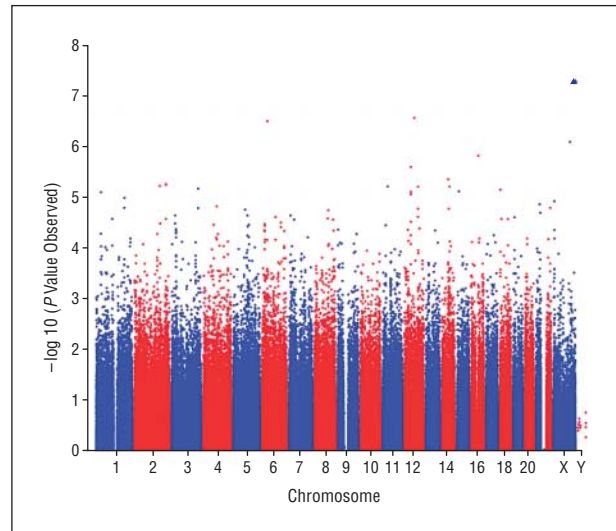


Figure 2. Genome-wide overview of the genome-wide association study findings. The single-nucleotide polymorphisms at which the test statistic exceeds 30 are represented by the triangle at the top of the plot.

found to be enriched in the genes analyzed. No specific coherent underlying process could be deduced from Gene Ontology database term associations. Consequently, we considered the literature-derived molecular function of each of the gene products of intragenic SNPs separately.

ANIMAL STUDIES

The GeneChip Rat Genome U34 Set (Affymetrix) provides gene expression data for more than 24 000 known genes and expressed sequence tag (EST) clusters for comprehensive coverage of the rat genome. The 2 brain regions were evaluated separately. Three-way analysis of variance revealed 542 genes differentially expressed at $P < .05$ in the caudate putamen, the amygdala, or both brain regions of alcohol-preferring P, HAD, and AA rats after 1 year of ethanol consumption. A detailed description of the rodent candidate gene strategy for selection of markers for the follow-up study is given in the "Integration of Animal Data Into the Human Study" subsection of the "Methods" section.

COMMENT

The present GWAS and follow-up study of alcohol dependence detected and replicated evidence for association with 15 markers. Two of these markers, rs7590720 and rs1344694, remained significant after genome-wide correction for multiple testing in the combined sample of 1460 patients and 2332 controls.

These 2 markers are located approximately 5 kb apart in chromosomal region 2q35, which has been implicated in alcohol dependence in previous linkage studies. Linkage to this region was found in a genome-wide search in 2282 individuals from 262 families with a high density of alcohol dependence in the Collaborative Study on the Genetics of Alcoholism (COGA).⁵² Near microsatellite marker D2S1371, the highest logarithm of odds (LOD) score for the comorbid alcoholism and depression phenotype was

Table 2. SNPs Confirmed in the Follow-up Study: Comparison of Results of the GWAS, Replication Study, and Pooled Analysis

SNP	Allele A ^d	Allele B ^d	GWAS ^a				Replication Sample ^b				Pooled Sample ^c			
			Frq A Ca	Frq A Co	P Value ^e	Odds Ratio (95% CI)	Frq A Ca	Frq A Co	P Value ^e	Odds Ratio (95% CI)	Frq A Ca	Frq A Co	P Value ^e	Odds Ratio (95% CI)
rs1344694	G	T	0.603	0.685	5.57E-6	0.70 (0.60-0.81)	0.613	0.659	3.36E-3	0.82 (0.72-0.93)	0.610	0.674	1.69E-8	0.76 (0.69-0.83)
rs7590720	A	G	0.645	0.724	5.70E-6	0.69 (0.59-0.81)	0.661	0.713	6.68E-4	0.79 (0.68-0.90)	0.656	0.719	9.72E-9	0.74 (0.67-0.82)
rs705648	C	T	0.313	0.244	2.68E-5	1.41 (1.20-1.67)	0.292	0.256	1.24E-2	1.19 (1.04-1.37)	0.299	0.249	1.78E-6	1.29 (1.16-1.43)
rs1614972	C	T	0.754	0.690	2.84E-4	1.37 (1.16-1.61)	0.721	0.690	3.62E-2	1.16 (1.01-1.33)	0.732	0.690	1.41E-4	1.23 (1.10-1.36)
rs13362120	C	T	0.258	0.329	5.72E-5	0.71 (0.60-0.84)	0.296	0.333	1.25E-2	0.84 (0.73-0.96)	0.283	0.331	1.85E-5	0.80 (0.72-0.88)
rs13160562	A	G	0.252	0.326	2.31E-5	0.70 (0.59-0.83)	0.283	0.316	2.47E-2	0.85 (0.75-0.98)	0.273	0.322	7.09E-6	0.79 (0.71-0.88)
rs1864982	A	C	0.181	0.129	9.71E-5	1.49 (1.22-1.82)	0.163	0.130	4.47E-3	1.30 (1.09-1.56)	0.169	0.130	3.46E-6	1.36 (1.20-1.55)
rs6902771	C	T	0.591	0.517	9.31E-5	1.35 (1.16-1.56)	0.552	0.505	3.15E-3	1.20 (1.06-1.37)	0.565	0.512	8.30E-6	1.24 (1.13-1.36)
rs729302	A	C	0.599	0.671	6.19E-5	0.73 (0.63-0.85)	0.641	0.677	1.84E-2	0.85 (0.75-0.97)	0.627	0.673	4.02E-5	0.82 (0.74-0.90)
rs13273672	C	T	0.364	0.312	2.18E-3	1.27 (1.09-1.47)	0.342	0.311	3.61E-2	1.15 (1.01-1.32)	0.350	0.312	4.75E-4	1.19 (1.08-1.31)
rs1487814	A	G	0.589	0.511	3.58E-5	1.37 (1.18-1.59)	0.544	0.508	2.48E-2	1.15 (1.02-1.30)	0.559	0.510	3.49E-5	1.22 (1.11-1.34)
rs7138291	C	T	0.142	0.093	2.46E-5	1.61 (1.28-2.00)	0.115	0.095	4.10E-2	1.23 (1.01-1.54)	0.124	0.094	3.68E-5	1.36 (1.18-1.58)
rs36563	A	C	0.210	0.147	6.21E-6	1.54 (1.28-1.85)	0.184	0.159	3.83E-2	1.19 (1.01-1.41)	0.192	0.152	4.57E-6	1.33 (1.18-1.50)
rs11640875	A	G	0.389	0.326	5.25E-4	1.32 (1.12-1.54)	0.372	0.332	9.85E-3	1.19 (1.04-1.35)	0.377	0.329	1.84E-5	1.24 (1.12-1.36)
rs12388359	G	T	0.811	0.900	1.21E-5	0.47 (0.34-0.67)	0.840	0.879	1.25E-2	0.72 (0.56-0.93)	0.831	0.888	3.57E-6	0.62 (0.50-0.76)

Abbreviations: CI, confidence interval; Frq, frequency; GWAS, genome-wide association study; SNP, single-nucleotide polymorphism.

^aAnalysis was based on 476 cases and 1358 controls.

^bAnalysis was based on 984 cases and 974 controls.

^cAnalysis was based on 1460 cases and 2332 controls.

^dAllele names are assigned in alphabetical order.

^eCochran-Armitage test for trend for autosomal markers; allele-based χ^2 test for X chromosomal marker.

4.12 (this LOD score, however, was obtained in only one data set, ie, in the replication data set, whereas in the initial and combined data sets, the LOD scores were 0.00 and 2.16, respectively).⁵² Marker D2S1371 is located approximately 1.4 Mb from rs7590720 and rs1344694. Another COGA project⁵³ found linkage with a maximum LOD score of 2.4 between a low level of response to alcohol and the microsatellite marker D2S434, which is only approximately 1.7 Mb from markers rs7590720 and rs1344694. The low level of response to alcohol is an endophenotype related to heavy drinking and alcohol problems.⁵⁴ It is genetically affected in animals⁵⁵⁻⁵⁷ and humans,^{58,59} with an estimated heritability in humans of 40% to 50%.⁶⁰ Two further linkage studies of the COGA project^{61,62} reported linkage with LOD scores of 3.28 and 3.0, respectively, between the same marker D2S434 and the P3(00) component of the event-related potential. Heritability of the P3 amplitude ranges from 50% to 80% (for a review, see the study by Porjesz et al⁶²), and P3 deficits have been repeatedly reported to be a strong vulnerability factor for alcohol dependence, especially in males with high familial loading of alcohol dependence.^{63,64} In a linkage study conducted by Hill et al⁶⁵ in 330 members of 65 families with a high density of alcohol dependence, linkage to 2q35 with marker D2S2382 was identified. This marker is located only approximately 0.2 Mb from the 2 main SNPs. A LOD score of 3.68 was observed in a 4-parameter model that included measurement of age, gender, and alcohol dependence along with measurement of constraint, a personality construct⁶⁶ that seems to overlap with risk-taking behavior.

These linkage findings, which all satisfied the Lander and Kruglyak⁶⁷ criteria for suggestive linkage, underline the potential importance of this chromosomal region. However, because no meta-analysis of linkage studies in alcohol addiction exists, the possibility that these findings are merely due to random noise cannot be ruled out.

The gene that is closest to the peak association signal and that also harbors another associated SNP, rs705648 ($P=1.78 \times 10^{-6}$), is *PECR*, which seems to be a key enzyme of fatty acid metabolism as it catalyzes the reduction of medium-chain (unsaturated) enoyl-CoAs to saturated acyl-CoAs, which may then undergo oxidation in mitochondria for energy production. This pathway, originating from triglycerides, is particularly important in conditions of starvation, when the energy supply is switched from the use of glucose to fatty acids. In this condition, *PECR* needs to be upregulated. Triglyceride levels are typically increased in the blood of alcohol-dependent individuals.^{68,69} *PECR* reduces C-C double bonds in not only even-chain but also in odd-chain fatty acids. On degradation to acetyl-CoA, odd-chain fatty acids end up with propionyl-CoA, which requires carboxylation to an even-chain fatty acid. This step is mediated through methylmalonyl-CoA, an intermediate formed by the essential cofactors methyl-tetrahydrofolate as the methyl donor and vitamin B₁₂. Folate deficiency, often observed in alcohol-dependent patients,⁷⁰ may, therefore, interfere with this pathway and prevent the availability of a sufficient supply of acetyl-CoA. Expression of *PECR* is highest in the liver, followed by the kidney, muscle tissue, the lungs, and the heart but barely, if at all, in the brain.⁷¹ If this gene is involved in alcohol dependence, it must act in the periphery rather than in the central nervous system. It could, therefore, be expected that we detected *PECR* only in the GWAS, which identifies susceptibility genes independent of organ-specific expression, but not in the animal approach, which is brain specific.

A further 4 SNPs are intergenic, and another 8 are located in genes. No ontologic relatedness was detected between these genes. That is not surprising given the assumed genetic heterogeneity of alcohol addiction. In the eText, we discuss data from linkage findings, biochemistry, and animal studies that have provided evidence that

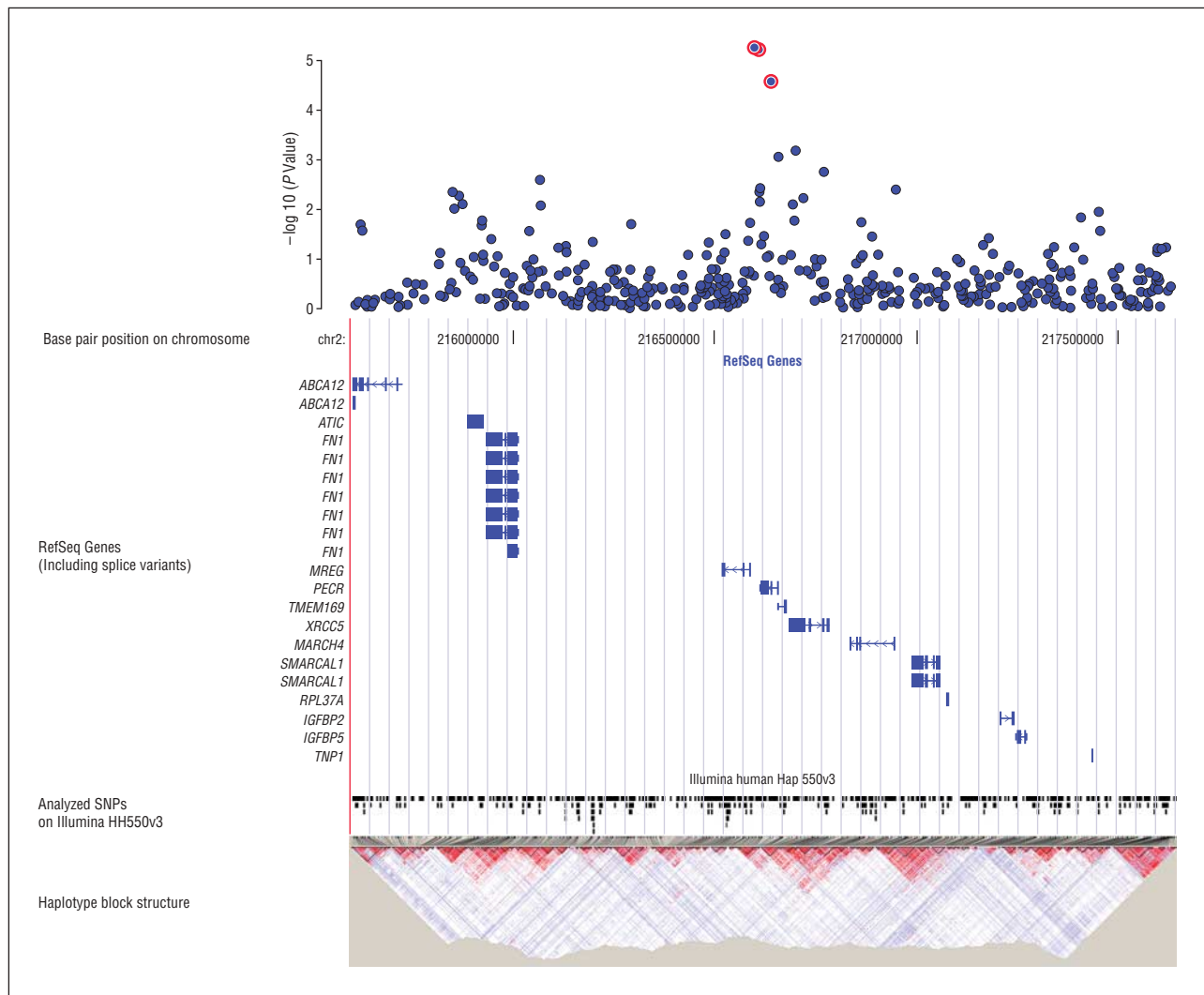


Figure 3. Illustration of P values in a ± 1 -megabase interval around the most significant findings at 2q35: rs1344694, rs7590720, and rs705648. Gene annotations are from University of California Santa Cruz RefSeq genes; linkage disequilibrium maps are from the International HapMap Project. SNP indicates single-nucleotide polymorphism.

these loci may be involved in alcohol dependence. However, none of these markers achieved genome-wide significance in this study. This does not necessarily disqualify them. Alcohol dependence is a complex and heterogeneous disorder that is affected by multiple genetic, biological, psychological, and sociocultural factors. Formal genetic and linkage studies do not indicate the existence of genes with major effects in the European population. Thus, genome-wide significant findings in moderately sized samples will be rare, but the number of replicable associations may actually be much larger. One replicated association that did not withstand correction for multiple testing was obtained with the SNP rs1614972 in the *ADH1C* gene. The *ADH* gene cluster is located on chromosomal region 4q, which is among the most consistently replicated loci contributing to alcohol phenotypes.⁷² The *ADH1C* gene belongs to the biologically relevant alcohol-metabolizing alcohol dehydrogenase (*ADH*) genes. The P value of this SNP, which was at position 282 of the top-down ranking, achieved a $P = 2.84 \times 10^{-4}$ in the GWAS. It would not have been carried forward into the follow-up study if it had not

been supported by evidence from the animal study. To achieve a power of 80% to detect significant association withstanding correction for multiple testing, we would have needed to include 5573 patients and 8900 controls (regarding actual odds ratio and minor allele frequency of this SNP). The same is true for the replicated association to marker rs11640875, which was ranked 510 in the GWAS findings. It is located in the *CDH13* gene, a gene that has been described in previous studies as a susceptibility locus for alcohol dependence²⁵ and methamphetamine dependence.⁷³ Increasing sample sizes is often not feasible and bears the risk of increasing clinical and genetic heterogeneity. Experience from GWASs of diabetes, for example, has shown that differences in assessment can jeopardize even the most robust findings.⁷⁴ Obtaining very large sample sizes may necessitate the inclusion of samples from different cultural and ethnic backgrounds, which can also increase heterogeneity.

The *ADH1C* gene belongs to the *ADH* class I genes (*ADH1A*, *ADH1B*, and *ADH1C*), which have been extensively investigated for the risk of alcohol dependence, ini-

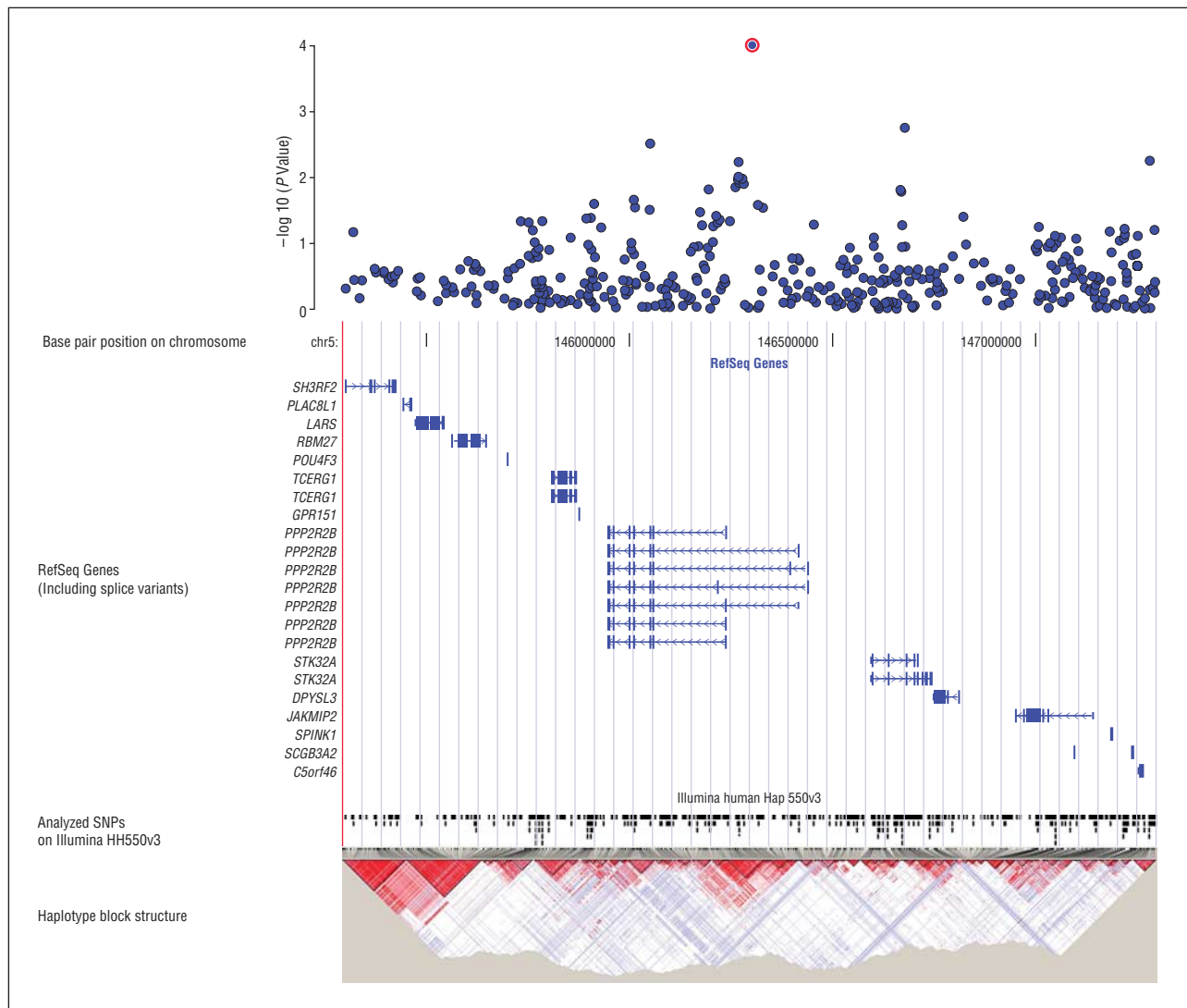


Figure 4. Illustration of P values in a ± 1 -megabase interval around the most significant finding, rs1864982, at 5q32. Gene annotations are from University of California Santa Cruz RefSeq genes; linkage disequilibrium maps are from the International HapMap Project. SNP indicates single-nucleotide polymorphism.

tially in Asian⁷⁵⁻⁷⁷ and subsequently in patients with African ancestry^{78,79} and ethnic European⁸⁰⁻⁸⁴ populations. A meta-analysis⁸⁵ reported a significantly higher risk of alcohol dependence in carriers of the *ADH1C*2* allele in East Asian populations (odds ratio, 1.91), but its role in the European population is less clear. The associated marker rs1614972 is in complete linkage disequilibrium ($D' = 1.0$, $r^2 = 0.311$, International HapMap Project) with 1 of the 2 polymorphisms of the *ADH1C*2* protein isoform that has recently been investigated⁸⁴ in 575 patients and 530 controls from the Irish Affected Sib Pair Study of Alcohol Dependence sample. The 2-marker haplotype composing markers rs1614972 and rs1693482 yielded evidence for association with alcohol dependence. Single-marker analysis with rs1614972, however, produced no evidence. Although the allele frequencies in ethnic Irish and ethnic German patients are the same, they differ between Irish and German control samples, with the allele frequencies between the German GWAS and the follow-up study being the same. This finding underlines the importance of sample homogeneity.

To increase the genetic homogeneity of the sample, and thus increase the power of the study, we attempted to select patients using as rigorous a set of criteria as possible. Patients of German descent were recruited in a joint research project and were stratified for male sex and an early age at onset, a combination of traits associated with increased heritability.^{32,33} Previous studies^{39,86,87} have shown that population stratification in Germans is small. Thus, at the time when SNPs for the follow-up study were chosen, we had not applied a formal test to control for population stratification. A post hoc principal components analysis^{88,89} (EigenSoft v1.01; <http://genepath.med.harvard.edu/~reich/Software.htm>) and genomic control analysis⁴¹ (showing a λ of 1.099) had no major impact on the significance level of the 15 SNPs (eTable 4), thus confirming these previous observations.

Alcohol dependence is a complex disorder, and association studies with more biologically refined phenotypes will help to disentangle this heterogeneity. The present approach of selecting more genetically homogeneous groups by stratifying for male sex and an earlier

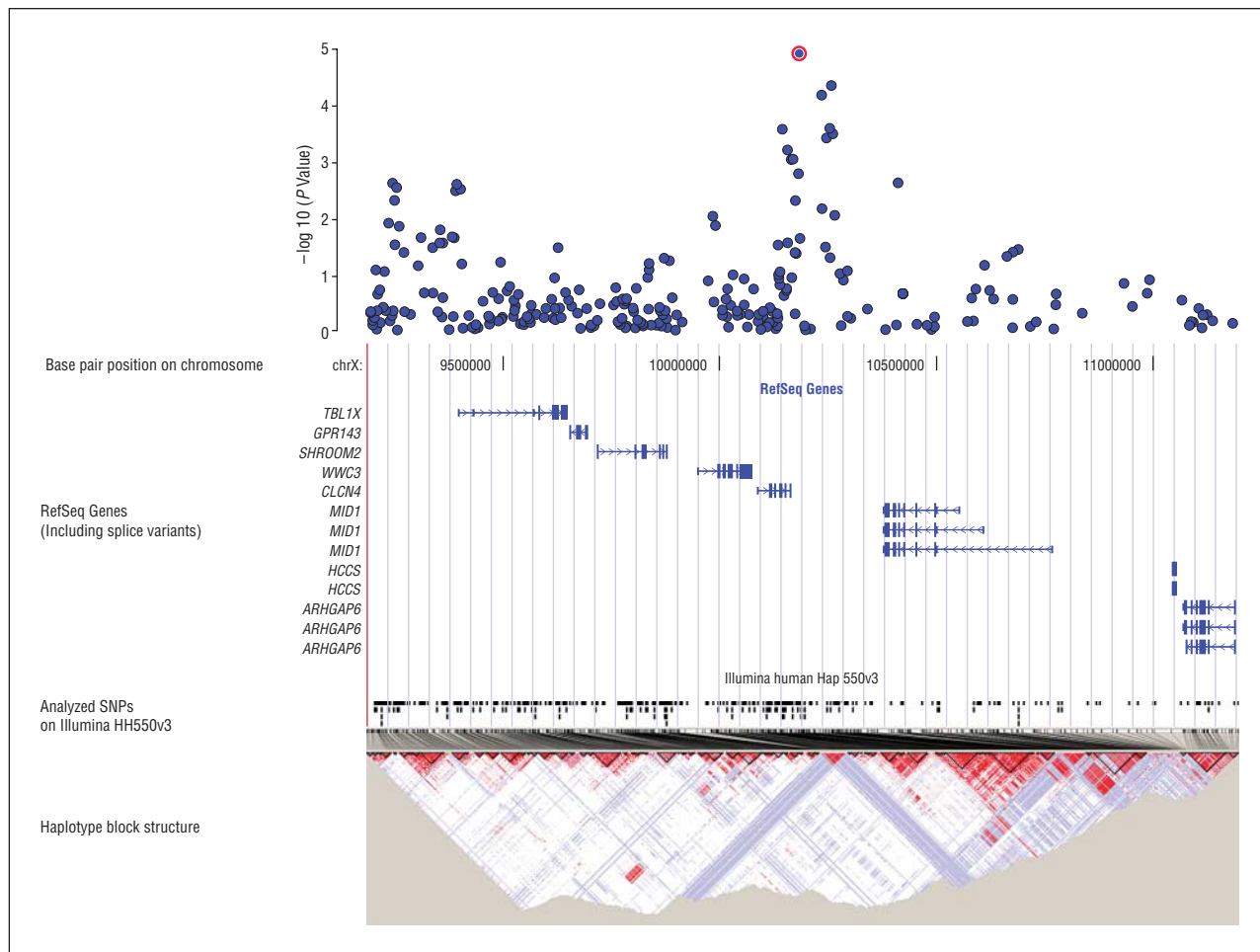


Figure 5. Illustration of P values in a ± 1 -megabase interval around the most significant finding, rs12388359, at Xp22.2. Gene annotations are from University of California Santa Cruz RefSeq genes; linkage disequilibrium maps are from the International HapMap Project. SNP indicates single-nucleotide polymorphism.

age at onset is a first step in this direction. We must point out, however, that to achieve an overall sample of more than 1500 patients, it was necessary to relax the early-age-at-onset criteria. Whereas in GWAS patients the upper inclusion limit was 28 years (median, 20 years), the upper inclusion limit in the replication sample had to be extended to 45 years (median, 30 years). Although, this cannot be considered as a specifically early age at onset, the sample as a whole and especially the GWAS sample had clearly been stratified for early age at onset.

In conclusion, we identified genome-wide significant association with 2 SNPs located in a chromosomal region for which linkage and animal studies provide compelling evidence for the presence of susceptibility variants predisposing an individual to alcohol addiction. Further independent studies are required to confirm these findings. The *PECR* gene, which is located close to these SNPs, is a plausible candidate and merits further investigation. We showed that the GWAS is a powerful tool for producing valuable findings in even a moderately sized sample when it has been carefully selected for genetically homogeneous cases. In addition, this study underlines the value of a convergent approach between animal studies and systematic genetic screens in humans to deal with the multiple testing problem in GWAS.

Submitted for Publication: August 29, 2008; final revision received December 9, 2008; accepted January 21, 2009.
Author Affiliations: Departments of Genetic Epidemiology in Psychiatry (Drs Treutlein and Rietschel and Mr Frank), Psychopharmacology (Drs Sommer, Gebicke-Haerter, and Spanagel and Mr Leonardi-Essmann), and Addictive Behavior and Addiction Medicine (Drs Treutlein, Kiefer, and Mann), Central Institute of Mental Health, Mannheim, Germany; Department of Genomics, Life and Brain Center (Drs Cichon and Nöthen and Ms Ludwig), Institute of Human Genetics (Drs Cichon and Nöthen), Department of Psychiatry (Drs Maier and Moessner), and Institute for Medical Biometry, Informatics and Epidemiology (Drs Steffens and Wienker), University of Bonn, Bonn, Germany; Department of Psychiatry, Mental State Hospital, University of Regensburg, Regensburg, Germany (Drs Ridinger and Wodarz); Private Hospital Meiringen, Meiringen, Switzerland (Dr Soyka); Department of Psychiatry, University of Munich, Munich, Germany (Drs Soyka and Zill); Department of Psychiatry and Psychotherapy, University of Dusseldorf, Dusseldorf, Germany (Dr Gaebel); Department of Psychiatry, University of Mainz, Mainz, Germany (Drs Dahmen and Fehr); Addiction Research Group at the Department of Psychiatry and Psychotherapy, Rhine State

Hospital Essen, University of Duisburg-Essen, Essen, Germany (Dr Scherbaum); Institute of Epidemiology, Helmholtz Center, Munich (Dr Wichmann); Institute of Medical Data Management, Biometrics, and Epidemiology, Chair of Epidemiology, Ludwig Maximilians University, Munich (Dr Wichmann); Institute for Clinical Molecular Biology and Department of General Internal Medicine, University of Kiel, Kiel, Germany (Dr Schreiber); Department of Medical Sociology, University of Dusseldorf for the Heinz Nixdorf Risk Factors, Evaluation of Coronary Calcium, and Lifestyle (RECALL) Study Group, Dusseldorf (Dr Dragan); National Institute on Alcohol Abuse and Alcoholism, Bethesda, Maryland (Dr Sommer); Dipartimento di Medicina Sperimentale e Sanità Pubblica, Università di Camerino, Camerino, Italy (Dr Lourdasamy); and Department of Genetics, University of North Carolina at Chapel Hill (Dr Sullivan).

Correspondence: Marcella Rietschel, MD, Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health Mannheim, University of Heidelberg, J5, 68159 Mannheim, Germany (marcella.rietschel@zi-mannheim.de).

Author Contributions: Drs Treutlein, Cichon, Ridinger, Spanagel, Mann, and Rietschel are co-primary authors of this article. They are responsible for distinct, but integral, parts of the study: the clinical study with homogeneous assessment of the patients, the animal study, and the genetic study. They had full access to all of the data in the study and take joint responsibility for the integrity of the data and the accuracy of the data analysis.

Financial Disclosure: None reported.

Funding/Support: This work was supported by grants from the German Federal Ministry of Education and Research: NGFN2 and NGFN-Plus FKZ 01GS0117 and 01GS08152 (Drs Mann, Nöthen, Rietschel, Schreiber, Spanagel, and Wichmann) and grants FKZ EB01011300 and 01EB0410 (Drs Mann and Spanagel), and by a grant from the Sixth Framework Program of the European Commission: Integrated Project IMAGEN IP-I3250 (Drs Mann, Rietschel, and Spanagel). Drs Cichon and Nöthen received support from the Alfred Krupp von Bohlen und Halbach-Stiftung. The Heinz Nixdorf RECALL Study was supported by a grant from the Heinz Nixdorf Foundation.

Previous Presentations: This study was presented at Deutscher Suchtkongress; June 14, 2008; Mannheim, Germany; at Drei-Länder-Symposium für Biologische Psychiatrie; October 9, 2008; Gottingen, Germany; at the XVI World Congress on Psychiatric Genetics; October 14, 2008; Osaka, Japan; and at the 1st Annual Meeting of NGFN-Plus and NGFN-Transfer in the Program of Medical Genome Research, Helmholtz Zentrum München/Neuherberg; December 12, 2008; Munich-Neuherberg, Germany.

Additional Information: The eText, eFigures, and eTables are available at http://www.zi-mannheim.de/pub_gwas.html.

Additional Contributions: Marina Füg, Christine Hohmeyer, and Rosa Ferrando provided expert technical assistance; Thomas G. Schulze, MD, Central Institute of Mental Health/National Institute of Mental Health Mood and Anxiety Disorders Program, provided helpful discussions; and Christine Schmael, MD, Central Institute

of Mental Health, critically read the manuscript. The rats were provided by TK Li, PhD (Department of Psychiatry, Institute of Psychiatric Research, Indiana University School of Medicine) and D Sinclair, PhD (Department of Mental Health and Alcohol Research, National Public Health Institute).

REFERENCES

1. World Health Organization. *The World Health Report 2002: Reducing Risks, Promoting Healthy Life*. Geneva, Switzerland: World Health Organization; 2002.
2. World Health Organization. *WHO Global Status Report on Alcohol 2004*. Geneva, Switzerland: Department of Mental Health and Substance Abuse, World Health Organization; 2004.
3. Radel M, Goldman D. Pharmacogenetics of alcohol response and alcoholism: the interplay of genes and environmental factors in thresholds for alcoholism. *Drug Metab Dispos*. 2001;29(4, pt 2):489-494.
4. Tyndale RF. Genetics of alcohol and tobacco use in humans. *Ann Med*. 2003;35(2):94-121.
5. Goldman D, Oroszi G, Ducci F. The genetics of addictions: uncovering the genes. *Nat Rev Genet*. 2005;6(7):521-532.
6. Cloninger CR. Neurogenetic adaptive mechanisms in alcoholism. *Science*. 1987;236(4800):410-416.
7. Enoch MA. Pharmacogenomics of alcohol response and addiction. *Am J Pharmacogenomics*. 2003;3(4):217-232.
8. Dick DM, Foroud T. Candidate genes for alcohol dependence: a review of genetic evidence from human studies. *Alcohol Clin Exp Res*. 2003;27(5):868-879.
9. Dick DM, Bierut LJ. The genetics of alcohol dependence. *Curr Psychiatry Rep*. 2006;8(2):151-157.
10. Higuchi S, Matsushita S, Kashima H. New findings on the genetic influences on alcohol use and dependence. *Curr Opin Psychiatry*. 2006;19(3):253-265.
11. Edenberg HJ, Foroud T. The genetics of alcoholism: identifying specific genes through family studies. *Addict Biol*. 2006;11(3-4):386-396.
12. Porjesz B, Ranganaswamy M. Neurophysiological endophenotypes, CNS disinhibition, and risk for alcohol dependence and related disorders. *Scientific World J*. 2007;7:131-141.
13. Gorwood P, Schumann G, Treutlein J, Ades J. Pharmacogenetics of alcohol dependence. In: Gorwood P, Hamon M, eds. *Psychopharmacogenetics*. Heidelberg, Germany: Springer; 2006:177-202.
14. Mayfield RD, Harris RA, Schuckit MA. Genetic factors influencing alcohol dependence. *Br J Pharmacol*. 2008;154(2):275-287.
15. Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, Manolescu A, Thorleifsson G, Stefansson H, Ingason A, Stacey SN, Bergthorsson JT, Thorlacius S, Gudmundsson J, Jonsson T, Jakobsdottir M, Saemundsdottir J, Olafsdottir O, Gudmundsson LJ, Bjornsdottir G, Kristjansson K, Skuladottir H, Isaksson HJ, Gudbjartsson T, Jones GT, Mueller T, Gottsäter A, Flex A, Aben KK, de Vegt F, Mulders PF, Isla D, Vidal MJ, Asin L, Saez B, Murillo L, Blondal T, Kolbeinnsson H, Stefansson JG, Hansdottir I, Runarsdottir V, Pola R, Lindblad B, van Rij AM, Dieplinger B, Haltmayer M, Mayordomo JI, Kiemeny LA, Matthiasson SE, Oskarsson H, Tyrifingsson T, Gudbjartsson DF, Gulcher JR, Jonsson S, Thorsteinsdottir U, Kong A, Stefansson K. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature*. 2008;452(7187):638-642.
16. Uhl GR, Liu QR, Drgon T, Johnson C, Walther D, Rose JE, David SP, Niaura R, Lerman C. Molecular genetics of successful smoking cessation: convergent genome-wide association study results. *Arch Gen Psychiatry*. 2008;65(6):683-693.
17. Mah S, Nelson MR, Delisi LE, Reneland RH, Markward N, James MR, Nyholt DR, Hayward N, Handoko H, Mowry B, Kammerer S, Braun A. Identification of the semaphorin receptor *PLXNA2* as a candidate for susceptibility to schizophrenia. *Mol Psychiatry*. 2006;11(5):471-478.
18. Lencz T, Morgan TV, Athanasiou M, Dain B, Reed CR, Kane JM, Kucherlapati R, Malhotra AK. Converging evidence for a pseudoautosomal cytokine receptor gene locus in schizophrenia. *Mol Psychiatry*. 2007;12(6):572-580.
19. Sullivan PF, Lin D, Tzeng JY, van den Oord E, Perkins D, Stroup TS, Wagner M, Lee S, Wright FA, Zou F, Liu W, Downing AM, Lieberman J, Close SL. Genome-wide association for schizophrenia in the CATIE study: results of stage 1. *Mol Psychiatry*. 2008;13(6):570-584.
20. O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskvina V, Nikolov I, Hamshere M, Carroll L, Georgieva L, Dwyer S, Holmans P, Marchini JL, Spencer C, Howie B, Leung H-T, Hartmann AM, Möller H-J, Morris DW, Shi Y, Feng G, Hoffmann P, Propping P, Vasilescu C, Maier W, Rietschel M, Zammit S,

- Schumacher J, Quinn EM, Schulze TG, Williams NM, Giegling I, Iwata N, Ikeda M, Darvasi A, Shifman S, He L, Duan J, Sanders AR, Levinson DF, Gejman PV; Molecular Genetics of Schizophrenia Collaboration. Identification of novel schizophrenia loci by genome-wide association and follow-up. *Nat Genet.* 2008;40(9):1053-1055.
21. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007;447(7145):661-678.
 22. Baum AE, Akula N, Cabanero M, Cardona I, Corona W, Klemens B, Schulze TG, Cichon S, Rietschel M, Nöthen MM, Georgi A, Schumacher J, Schwarz M, Abou Jamra R, Höfels S, Propping P, Satagopan J, Detera-Wadleigh SD, Hardy J, McMahon FJ. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. *Mol Psychiatry.* 2008;13(2):197-207.
 23. Sklar P, Smoller JW, Fan J, Ferreira MA, Perlis RH, Chambert K, Nimgaonkar VL, McQueen MB, Faraone SV, Kirby A, de Bakker PI, Ogdie MN, Thase ME, Sachs GS, Todd-Brown K, Gabriel SB, Sougnez C, Gates C, Blumenstiel B, Defelice M, Ardlie KG, Franklin J, Muir WJ, McGhee KA, MacIntyre DJ, McLean A, VanBeck M, McQuillin A, Bass NJ, Robinson M, Lawrence J, Anjorin A, Curtis D, Scolnick EM, Daly MJ, Blackwood DH, Gurling HM, Purcell SM. Whole-genome association study of bipolar disorder. *Mol Psychiatry.* 2008;13(6):558-569.
 24. Ferreira MA, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, Fan J, Kirov G, Perlis RH, Green EK, Smoller JW, Grozeva D, Stone J, Nikolov I, Chamberlain K, Hamshere ML, Nimgaonkar VL, Moskvina V, Thase ME, Caesar S, Sachs GS, Franklin J, Gordon-Smith K, Ardlie KG, Gabriel SB, Fraser C, Blumenstiel B, Defelice M, Breen G, Gill M, Morris DW, Elkin A, Muir WJ, McGhee KA, Williamson R, Macintyre DJ, Maclean AW, St Clair D, Robinson M, Van Beck M, Pereira AC, Kandaswamy R, McQuillin A, Collier DA, Bass NJ, Young AH, Lawrence J, Nicol Ferrier I, Anjorin A, Farmer A, Curtis D, Scolnick EM, McGuffin P, Daly MJ, Corvin AP, Holmans PA, Blackwood DH, Gurling HM, Owen MJ, Purcell SM, Sklar P, Craddock N; Wellcome Trust Case Control Consortium. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet.* 2008;40(9):1056-1058.
 25. Johnson C, Drgon T, Liu Q-R, Walther D, Edenberg H, Rice J, Foroud T, Uhl GR. Pooled association genome scanning for alcohol dependence using 104,268 SNPs: validation and use to identify alcoholism vulnerability loci in unrelated individuals from the collaborative study on the genetics of alcoholism. *Am J Med Genet B Neuropsychiatr Genet.* 2006;141B(8):844-853.
 26. Spanagel R, Pendyala G, Abarca C, Zghoul T, Sanchis-Segura C, Magnone MC, Lascorz J, Depner M, Holzberg D, Soyka M, Schreiber S, Matsuda F, Lathrop M, Schumann G, Albrecht U. The clock gene *Per2* influences the glutamatergic system and modulates alcohol consumption. *Nat Med.* 2005;11(1):35-42.
 27. Rodd ZA, Bertsch BA, Strother WN, Le-Niculescu H, Balaraman Y, Hayden E, Jerome RE, Lumeng L, Nurnberger JI Jr, Edenberg HJ, McBride WJ, Niculescu AB. Candidate genes, pathways and mechanisms for alcoholism: an expanded convergent functional genomics approach. *Pharmacogenomics J.* 2007;7(4):222-256.
 28. Bertsch B, Ogden CA, Sidhu K, Le-Niculescu H, Kuczenski R, Niculescu AB. Convergent functional genomics: a Bayesian candidate gene identification approach for complex disorders. *Methods.* 2005;37(3):274-279.
 29. Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI Jr, Reich T, Schmidt I, Schuckit MA. A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. *J Stud Alcohol.* 1994;55(2):149-158.
 30. Wittchen HU. Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): a critical review. *J Psychiatr Res.* 1994;28(1):57-84.
 31. First MB, Spitzer RL, Gibbon M, Williams J. *Structured Clinical Interview for Axis I DSM-IV Disorders.* New York, NY: Biometrics Research; 1994.
 32. Pickens RW, Svikis DS, McGue M, Lykken DT, Heston LL, Clayton PJ. Heterogeneity in the inheritance of alcoholism: a study of male and female twins. *Arch Gen Psychiatry.* 1991;48(1):19-28.
 33. Liu IC, Blacker DL, Xu R, Fitzmaurice G, Tsuang MT, Lyons MJ. Genetic and environmental contributions to age of onset of alcohol dependence symptoms in male twins. *Addiction.* 2004;99(11):1403-1409.
 34. Krawczak M, Nikolaus S, von Eberstein H, Croucher PJP, El Mokhtari NE, Schreiber S. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet.* 2006;9(1):55-61.
 35. Wichmann H-E, Gieger C, Illig T; MONICA/KORA Study Group. KORA-gen: resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen.* 2005;67(suppl 1):S26-S30.
 36. Schmermund A, Möhlenkamp S, Stang A, Grönemeyer D, Seibel R, Hirche H, Mann K, Siffert W, Lauterbach K, Siegrist J, Jöckel KH, Erbel R. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. *Am Heart J.* 2002;144(2):212-218.
 37. Luca D, Ringquist S, Klei L, Lee AB, Gieger C, Wichmann HE, Schreiber S, Krawczak M, Lu Y, Styche A, Devlin B, Roeder K, Trucco M. On the use of general control samples for genome-wide association studies: genetic matching highlights causal variants. *Am J Hum Genet.* 2008;82(2):453-463.
 38. Winkelmann J, Schormair B, Lichtner P, Ripke S, Xiong L, Jalilzadeh S, Fulda S, Pütz B, Eckstein G, Hauk S, Trenkwalder C, Zimprich A, Stiasny-Kolster K, Oertel W, Bachmann CG, Paulus W, Peglau I, Eisenhauer I, Montplaisir J, Turecki G, Rouleau G, Gieger C, Illig T, Wichmann HE, Holsboer F, Müller-Myhsok B, Meitinger T. Genome-wide association study of restless legs syndrome identifies common variants in three genomic regions. *Nat Genet.* 2007;39(8):1000-1006.
 39. Steffens M, Lamina C, Illig T, Bettecken T, Vogler R, Entz P, Suk EK, Toliat MR, Klopp N, Caliebe A, König IR, Köhler K, Ludemann J, Diaz Lacava A, Fimmers R, Lichtner P, Ziegler A, Wolf A, Krawczak M, Nürnberg P, Müller J, Schreiber S, Meitinger T, Wichmann HE, Roeder K, Wienker TF, Baur MP. SNP-based analysis of genetic substructure in the German population. *Hum Hered.* 2006;62(1):20-29.
 40. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81(3):559-575.
 41. Devlin B, Roeder K. Genomic control for association studies. *Biometrics.* 1999;55(4):997-1004.
 42. Bell RL, Rodd ZA, Lumeng L, Murphy JM, McBride WJ. The alcohol-preferring P rat and animal models of excessive alcohol drinking. *Addict Biol.* 2006;11(3-4):270-288.
 43. Sommer W, Hyytiä P, Kiianmaa K. The alcohol-preferring AA and alcohol-avoiding ANA rats: neurobiology of the regulation of alcohol drinking. *Addict Biol.* 2006;11(3-4):289-309.
 44. Vengeliene V, Siegmund S, Singer MV, Sinclair JD, Li T-K, Spanagel R. A comparative study on alcohol-preferring rat lines: effects of deprivation and stress phases on voluntary alcohol intake. *Alcohol Clin Exp Res.* 2003;27(7):1048-1054.
 45. Vengeliene V, Leonardi-Essmann F, Perreau-Lenz S, et al. The dopamine D3 receptor plays an essential role in alcohol-seeking and relapse. *FASEB J.* 2006;20(13):2223-2233.
 46. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 2004;5(10):R80.
 47. Reimers M, Heilig M, Sommer WH. Gene discovery in neuropharmacological and behavioral studies using Affymetrix microarray data. *Methods.* 2005;37(3):219-228.
 48. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G; Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat Genet.* 2000;25(1):25-29.
 49. Dudbridge F, Gusnanto A. Estimation of significance thresholds for genome-wide association scans. *Genet Epidemiol.* 2008;32(3):227-234.
 50. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genome-wide association studies of nearly all common variants. *Genet Epidemiol.* 2008;32(4):381-385.
 51. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet.* 2005;76:887-893.
 52. Nurnberger JI Jr, Foroud T, Flury L, Su J, Meyer ET, Hu K, Crowe R, Edenberg H, Goate A, Bierut L, Reich T, Schuckit M, Reich W. Evidence for a locus on chromosome 1 that influences vulnerability to alcoholism and affective disorder. *Am J Psychiatry.* 2001;158(5):718-724.
 53. Schuckit MA, Edenberg HJ, Kalmijn J, Flury L, Smith TL, Reich T, Bierut L, Goate A, Foroud T. A genome-wide search for genes that relate to a low level of response to alcohol. *Alcohol Clin Exp Res.* 2001;25(3):323-329.
 54. Schuckit MA, Smith TL, Trim R, Kreikebaum S, Hinga B, Allen R. Testing the level of response to alcohol-based model of heavy drinking and alcohol problems in offspring from the San Diego prospective study. *J Stud Alcohol Drugs.* 2008;69(4):571-579.
 55. Baldwin HA, Wall TL, Schuckit MA, Koob GF. Differential effects of ethanol on punished responding in the P and NP rats. *Alcohol Clin Exp Res.* 1991;15(4):700-704.
 56. Moore MS, DeZazzo J, Luk AY, Tully T, Singh CM, Heberlein U. Ethanol intoxication

- cation in *Drosophila*: genetic and pharmacological evidence for regulation by the cAMP signaling pathway. *Cell*. 1998;93(6):997-1007.
57. Li TK. Pharmacogenetics of responses to alcohol and genes that influence alcohol drinking. *J Stud Alcohol*. 2000;61(1):5-12.
 58. Heath AC, Martin NG. Genetic differences in psychomotor performance decrement after alcohol: a multivariate analysis. *J Stud Alcohol*. 1992;53(3):262-271.
 59. Madden PA, Heath AC, Starmer GA, Whitfield JB, Martin NG. Alcohol sensitivity and smoking history in men and women. *Alcohol Clin Exp Res*. 1995;19(5):1111-1120.
 60. Martin NG. Twin studies of alcohol consumption, metabolism and sensitivity. *Aust Drug Alcohol Rev*. 1988;7:9-12.
 61. Begleiter H, Porjesz B, Reich T, Edenberg HJ, Goate A, Blangero J, Almasy L, Foroud T, Van Eerdewegh P, Polich J, Rohrbaugh J, Kuperman S, Bauer LO, O'Connor SJ, Chorlian DB, Li TK, Conneally PM, Hesselbrock V, Rice JP, Schuckit MA, Cloninger R, Nurnberger J Jr, Crowe R, Bloom FE. Quantitative trait loci analysis of human event-related brain potentials: P3 voltage. *Electroencephalogr Clin Neurophysiol*. 1998;108(3):244-250.
 62. Porjesz B, Begleiter H, Wang K, et al. Linkage and linkage disequilibrium mapping of ERP and EEG phenotypes. *Biol Psychol*. 2002;61(1-2):229-248.
 63. Polich J, Pollock VE, Bloom FE. Meta-analysis of P300 amplitude from males at risk for alcoholism. *Psychol Bull*. 1994;115(1):55-73.
 64. Benegal V, Jain S, Subbukrishna DK, Channabasavanna SM. P300 amplitudes vary inversely with continuum of risk in first-degree male relatives of alcoholics. *Psychiatr Genet*. 1995;5(4):149-156.
 65. Hill SY, Shen S, Zezza N, Hoffman EK, Perlin M, Allan W. A genome wide search for alcoholism susceptibility genes. *Am J Med Genet B Neuropsychiatr Genet*. 2004;128B(1):102-113.
 66. Tellegen A, Waller NG. *Exploring Personality Through Test Construction: Development of the Multidimensional Personality Questionnaire*. Minneapolis: University of Minnesota Press; 1982.
 67. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet*. 1995;11(3):241-247.
 68. Nestel PJ, Hirsch EZ. Mechanism of alcohol-induced hypertriglyceridemia. *J Lab Clin Med*. 1965;66(3):357-365.
 69. Sun GY, Rush A, Chin PC, Gorka C, Lahiri S, Wood WG. Serum lipids and acyl group composition of alcoholic patients. *Alcohol*. 1988;5(2):153-157.
 70. Hamid A, Kaur J. Decreased expression of transporters reduces folate uptake across renal absorptive surfaces in experimental alcoholism. *J Membr Biol*. 2007;220(1-3):69-77.
 71. Das AK, Uhler MD, Hajra AK. Molecular cloning and expression of mammalian peroxisomal *trans-2-enoyl-coenzyme A reductase* cDNAs. *J Biol Chem*. 2000;275(32):24333-24340.
 72. Prescott CA, Sullivan PF, Kuo P-H, 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. Genomewide 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*. 2006;11(6):603-611.
 73. Uhl GR, Drgon T, Liu Q-R, Johnson C, Walther D, Komiyama T, Harano M, Sekine Y, Inada T, Ozaki N, Iyo M, Iwata N, Yamada M, Sora I, Chen CK, Liu HC, Ujike H, Lin SK. Genome-wide association for methamphetamine dependence: convergent results from 2 samples. *Arch Gen Psychiatry*. 2008;65(3):345-355.
 74. Ioannidis JPP, Patsopoulos NA, Evangelou E. Heterogeneity in meta-analyses of genome-wide association investigations. *PLoS ONE*. 2007;2(9):e841.
 75. Chen WJ, Loh EW, Hsu YP, Chen CC, Yu JM, Cheng AT. Alcohol-metabolising genes and alcoholism among Taiwanese Han men: independent effect of ADH2, ADH3 and ALDH2. *Br J Psychiatry*. 1996;168(6):762-767.
 76. Chen CC, Lu RB, Chen YC, Wang MF, Chang YC, Li TK, Yin SJ. Interaction between the functional polymorphisms of the alcohol-metabolism genes in protection against alcoholism. *Am J Hum Genet*. 1999;65(3):795-807.
 77. Osier M, Pakstis AJ, Kidd JR, Lee JF, Yin SJ, Ko HC, Edenberg HJ, Lu RB, Kidd KK. Linkage disequilibrium at the ADH2 and ADH3 loci and risk of alcoholism. *Am J Hum Genet*. 1999;64(4):1147-1157.
 78. Ehlers CL, Gilder DA, Harris L, Carr L. Association of the ADH2*3 allele with a negative family history of alcoholism in African American young adults. *Alcohol Clin Exp Res*. 2001;25(12):1773-1777.
 79. Ehlers CL, Montane-Jaime K, Moore S, Shafe S, Joseph R, Carr LG. Association of the ADH1B*3 allele with alcohol-related phenotypes in Trinidad. *Alcohol Clin Exp Res*. 2007;31(2):216-220.
 80. Whitfield JB, Nightingale BN, Bucholz KK, Madden PAF, Heath AC, Martin NG. ADH genotypes and alcohol use and dependence in Europeans. *Alcohol Clin Exp Res*. 1998;22(7):1463-1469.
 81. Borràs E, Coutelle C, Rosell A, Fernández-Muixi F, Broch M, Crosas B, Hjelmqvist L, Lorenzo A, Gutiérrez C, Santos M, Szczepanek M, Heilig M, Quattrocchi P, Farrés J, Vidal F, Richart C, Mach T, Bogdal J, Jörnvall H, Seitz HK, Couzigou P, Parés X. Genetic polymorphism of alcohol dehydrogenase in Europeans: the ADH2*2 allele decreases the risk for alcoholism and is associated with ADH3*1. *Hepatology*. 2000;31(4):984-989.
 82. Edenberg HJ, Xuei X, Chen H-J, Tian H, Wetherill LF, Dick DM, Almasy L, Bierut L, Bucholz KK, Goate A, Hesselbrock V, Kuperman S, Nurnberger J, Porjesz B, Rice J, Schuckit M, Tischfield J, Begleiter H, Foroud T. Association of alcohol dehydrogenase genes with alcohol dependence: a comprehensive analysis. *Hum Mol Genet*. 2006;15(9):1539-1549.
 83. Luo X, Kranzler HR, Zuo L, Wang S, Schork NJ, Gelernter J. Diplotype trend regression analysis of the ADH gene cluster and the ALDH2 gene: multiple significant associations with alcohol dependence. *Am J Hum Genet*. 2006;78(6):973-987.
 84. Kuo P-H, Kalsi G, Prescott CA, Hodgkinson CA, Goldman D, van den Oord EJ, Alexander J, Jiang C, Sullivan PF, Patterson DG, Walsh D, Kendler KS, Riley BP. Association of ADH and ALDH genes with alcohol dependence in the Irish Affected Sib Pair Study of Alcohol Dependence (IASPSAD) sample. *Alcohol Clin Exp Res*. 2008;32(5):785-795.
 85. Zintzaras E, Stefanidis I, Santos M, Vidal F. Do alcohol-metabolizing enzyme gene polymorphisms increase the risk of alcoholism and alcoholic liver disease? *Hepatology*. 2006;43(2):352-361.
 86. Neale BM, Sham PC, Purcell S, Banaschewski T, Buitelaar J, Franke B, Sonuga-Barke E, Ebstein R, Eisenberg J, Mulligan A, Gill M, Manor I, Miranda A, Mulas F, Oades RD, Roeyers H, Rothenberger A, Sergeant J, Steinhausen HC, Taylor E, Thompson M, Chen W, Zhou K, Asherson P, Faraone SV. Population differences in the International Multi-Centre ADHD Gene Project. *Genet Epidemiol*. 2008;32(2):98-107.
 87. Novembre J, Johnson T, Bryc K, et al. Genes mirror geography within Europe. *Nature*. 2008;456(7218):98-101.
 88. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38(8):904-909.
 89. Buch S, Schafmayer C, Völzke H, Becker C, Franke A, von Eller-Eberstein H, Kluck C, Bässmann I, Brosch M, Lammert F, Miquel JF, Nervi F, Wittig M, Rosskopf D, Timm B, Höll C, Seeger M, ElSharawy A, Lu T, Egberts J, Fändrich F, Fölsch UR, Krawczak M, Schreiber S, Nürnberg P, Tepel J, Hampe J. A genome-wide association scan identifies the hepatic cholesterol transporter ABCG8 as a susceptibility factor for human gallstone disease. *Nat Genet*. 2007;39(8):995-999.