

ADH4 Gene Variation is Associated with Alcohol Dependence and Drug Dependence in European Americans: Results from HWD Tests and Case–Control Association Studies

Xingguang Luo^{1,2}, Henry R Kranzler³, Lingjun Zuo^{1,2}, Jaakko Lappalainen^{1,2}, Bao-zhu Yang^{1,2} and Joel Gelernter^{*,1,2}

¹Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA; ²VA Connecticut Healthcare System, West Haven Campus, CT, USA; ³Department of Psychiatry, Alcohol Research Center, University of Connecticut School of Medicine, Farmington, CT, USA

The alcohol dehydrogenase (ADH) family constitutes one of the key sets of enzymes responsible for the oxidation of alcohol. The *ADH4* gene, an important member of this family, is a functional and positional candidate for alcohol dependence. The present study aimed to investigate the relationship between *ADH4* gene variation and alcohol dependence and drug dependence in European-Americans (EAs) and African-Americans (AAs). Seven single nucleotide polymorphisms (SNPs) spanning the *ADH4* gene were genotyped in 365 healthy controls (317 EAs and 48 AAs) and 561 subjects (400 EAs and 161 AAs) affected with alcohol dependence and/or drug dependence (436 with alcohol dependence; 356 with drug dependence). Hardy–Weinberg equilibrium (HWE) for the genotype frequency distributions of these markers was tested in all phenotype groups to evaluate association between *ADH4* gene variation and phenotypes and to fine-map the disease risk locus. The allele, genotype, and haplotype frequency distributions of these markers were compared between cases and controls to confirm the associations. The genotype frequency distributions of *ADH4* markers were in HWE in EA controls, but were in Hardy–Weinberg disequilibrium (HWD) (ie, deviation from HWE) in EA cases. Among all markers, SNP2 (rs1042363) at exon 9 or SNP6 (rs1800759) at the promoter showed the greatest degree of HWD, among patients with either alcohol dependence or drug dependence. Significant differences between EA cases and controls were seen for genotype ($10^{-6} < \text{global } p < 0.044$), but not any allele or haplotype, frequency distributions for all seven *ADH4* markers. These findings suggest that *ADH4* genotypes predispose to alcohol dependence and drug dependence in a recessive manner, a predisposition that is population specific. SNP2 or SNP6 was the marker genetically closest to the functional risk loci for both alcohol dependence and drug dependence.

Neuropsychopharmacology (2006) 31, 1085–1095. doi:10.1038/sj.npp.1300925; published online 12 October 2005

Keywords: *ADH4* gene; alcohol dependence; drug dependence; Hardy–Weinberg disequilibrium (HWD)

INTRODUCTION

Ethanol, absorbed through the gastrointestinal tract, enters the liver via the portal vein. After being oxidized in the liver, ethanol metabolites and any unmetabolized ethanol enter the general circulation. Two kinds of enzymes are primarily responsible for the oxidation of alcohol. The first enzyme family is the alcohol dehydrogenases (ADH), which convert alcohol to acetaldehyde (which is toxic). The second enzyme family is the aldehyde dehydrogenases (ALDH), which

quickly convert acetaldehyde into acetate. Acetate is then oxidized into CO₂ and H₂O via the tricarboxylic acid cycle.

The development of alcohol dependence is related to an individual's ethanol consumption, which is influenced by ADH activity. Increased ADH activity can increase the rate of alcohol metabolism, thereby increasing the production of acetaldehyde. If the acetaldehyde cannot be metabolized quickly, the accumulation of acetaldehyde will result in an aversive 'flushing reaction,' characterized by facial flushing, headache, increased skin temperature, dizziness, anxiety, nausea, vomiting, decreased blood pressure with resultant tachycardia and, rarely, severe cardiovascular complications. The aversive effects of acetaldehyde can limit alcohol consumption, and, it is hypothesized, thereby reduce the risk of alcohol dependence (Thomasson *et al*, 1994; Muramatsu *et al*, 1995; Hasin *et al*, 2002).

Active ADH enzyme is composed of two subunits, which can be, but are not necessarily, identical. Humans express

*Correspondence: Dr J. Gelernter, Yale University School of Medicine, VA Psychiatry 116A2, 950 Campbell Avenue, West Haven, CT 06516, USA, Tel: +1 203 932 5711 ext. 3590, Fax: +1 203 937 3897, E-mail: joel.gelernter@yale.edu
Received 4 April 2005; revised 19 July 2005; accepted 30 August 2005
Online publication: 2 September 2005 at <http://www.acnp.org/citations/Npp090205050219/default.pdf>

at least nine possible different ADH subunits, each with slightly different properties. These isoforms include α , β , γ , π , ζ , ADH6 isoenzyme, μ , and σ . A unique gene encodes each isoform. A human *ADH* gene cluster maps to chromosome 4 (4q21–25), extending a total of 364 128 base pairs (bp) and including genes encoding five classes, in the following order: IV (*ADH7*), I- γ (*ADH1C*), I- β (*ADH1B*), I- α (*ADH1A*), V (*ADH6*), II (*ADH4*), and III (*ADH5*).

The human π subunit is encoded by the *ADH4* gene. In adults, the π subunit mainly contributes to liver ADH activity, catalyzing the oxidation or reduction of a wide range of substrates at moderate and high concentrations, including ethanol, primary long-chain aliphatic alcohols, and aromatic aldehydes (Li *et al*, 1977; Li and Bosron, 1987; Edenberg *et al*, 1999). Variation in *ADH4* could change the function of the π subunit and affect the risk for alcohol dependence by modulating alcohol metabolism. For example, the –75A allele has promoter activity that is more than twice that of the –75C allele; this –75A/C substitution (SNP6: rs1800759) (see Figure 1 for single nucleotide polymorphism (SNP) definitions and locations) at the promoter could significantly affect π subunit expression (Edenberg *et al*, 1999). Considering the important role of the *ADH4* gene in alcohol metabolism, it is a functional candidate for alcohol dependence.

Several linkage studies, including the Collaborative Study on the Genetics of Alcoholism (COGA) (Reich *et al*, 1998) and a study by investigators at the National Institute on Alcohol Abuse and Alcoholism (NIAAA) (Long *et al*, 1998), have demonstrated that the *ADH* gene cluster is coincident with a possible linkage signal for alcohol dependence. Reich *et al* (1998) performed a genome-wide screen by examining 291 markers in 987 individuals from 105 affected families and found suggestive evidence of a protective locus for alcohol dependence on chromosome 4 near the *ADH* gene cluster. Zinn-Justin and Abel (1999) performed an autosome-wide scan in the COGA data using the method of weighted pairwise correlation incorporating identity-by-descent (WPC-IBD), and located a risk locus for alcohol dependence at a 50 cM region spanning from *GABRB1* to D4S1651, a region harboring the *ADH* genes. Williams *et al* (1999) applied a variance-component method to analyze COGA data and located a risk locus for alcohol dependence at a chromosome 4 region near the *ADH* gene cluster. Saccone *et al* (2000) carried out Haseman–Elston-based regression analyses on the COGA data set and mapped susceptibility genes for a quantitative alcoholism-related phenotype, that is, ‘maximum number of drinks consumed in a 24-h period,’ which is correlated with the alcohol dependence diagnosis, to the region of the *ADH* gene cluster. In the study from NIAAA, Long *et al* (1998) performed a whole-autosomal genome scan with 517 microsatellite markers in several extended pedigrees and found that three loci in the *ADH* gene cluster

showed evidence for linkage with alcohol dependence. In summary, the *ADH* gene cluster is in a region that has been positionally implicated in alcohol dependence. Since the *ADH4* gene maps within this *ADH* gene cluster, *ADH4* gene is also a positional candidate for alcohol dependence.

There are also case-control association data implicating the same region of chromosome 4 (but different specific loci) by showing associations of *ADH* genes to alcohol dependence (eg, Thomasson *et al*, 1994; Hasin *et al*, 2002; Konishi *et al*, 2003). Foroud *et al* (2003) analyzed 58 markers at the *ADH* gene cluster in COGA alcoholic families using the pedigree disequilibrium test and primarily identified eight variants (published abstract only; specific marker names were unavailable to us) in the *ADH4* gene associated with the risk for alcohol dependence in European-Americans (EAs).

The present study aimed to extend these findings by examining the relationship between *ADH4* gene variants and alcohol dependence in several populations, including EAs and African-Americans (AAs). We focused on the *ADH4* gene in this study because of the functional importance of its protein product, remarkable for its activity at high ethanol concentrations that might be of particular relevance in the context of alcohol dependence risk (Li *et al*, 1977); because of the preliminary COGA results (Foroud *et al*, 2003); and because our preliminary case-control survey of variants at many *ADH* loci immediately identified this as a gene of interest (Luo *et al*, unpublished data).

Drug dependence, which mainly includes cocaine dependence and opioid dependence in our sample, is one of the most common phenotypes comorbid with alcohol dependence (Gossop, 2004). Many studies have demonstrated that drug dependence has a number of features in common with alcohol dependence, including symptomatology (Glatt, 1967; Heil *et al*, 2001), neuropsychological impairments (Robinson *et al*, 1999), pathogenetic mechanisms (Anokhina *et al*, 1990; O'Neill *et al*, 2001), and response to specific treatments, especially the effects of disulfiram, an ALDH blocker, which though approved for the treatment of alcohol dependence, has also been shown to be efficacious in the treatment of cocaine dependence (Carroll *et al*, 1993, 1998, 2000, 2004; George *et al*, 2000; Petrakis *et al*, 2000). Drug dependence may share susceptibility genes with alcohol dependence (reviewed by Kreek *et al*, 2004), for example, *OPRM1* gene variation moderated susceptibility to alcohol dependence and/or drug dependence (Hoehe *et al*, 2000; Szeto *et al*, 2001; Schinka *et al*, 2002; Luo *et al*, 2003). To answer the question of whether *ADH4* variants influence risk for alcohol dependence, drug dependence, or both of these traits, we examined the association of *ADH4* gene with drug dependence diagnoses in this study.



Figure 1 Map of markers at the *ADH4* gene.

MATERIALS AND METHODS

Subjects

A total of 926 unrelated subjects (557 male and 369 female subjects), including 400 EA cases, 317 EA healthy controls, 161 AA cases, and 48 AA controls, were included in this study. The cases met lifetime DSM-IV criteria (American Psychiatric Association, 1994) for alcohol dependence, drug (cocaine or opioid) dependence, or a combination of these disorders. Among the cases, 436 subjects had a diagnosis of alcohol dependence (334 EAs, 102 AAs) and 356 had a diagnosis of drug dependence (207 EAs, 149 AAs), including 318 with cocaine dependence (178 EAs, 140 AAs) and 152 with opioid dependence (95 EAs, 57 AAs). The control subjects were screened to exclude major Axis I mental disorders, including alcohol or drug abuse or dependence, psychotic disorders (including schizophrenia or schizophrenia-like disorders), mood disorders, and anxiety disorders. Male subjects constituted 73.2% of the cases and 40.0% of the controls. The average ages were 39.4 ± 9.2 years for cases and 28.1 ± 9.1 years for controls.

The subjects were recruited at the University of Connecticut Health Center or the VA Connecticut Healthcare System-West Haven Campus. All subjects gave informed consent before participating in the study, which was approved by the Institutional Review Board at the respective institutions.

Genotyping

Marker inclusion (see Figure 1). Seven known SNPs (all of those available from the public Applied Biosystems (Foster City, CA) website (<http://www.appliedbiosystems.com/>) at the time of genotyping; designated by hcv# or rs#) were included in this study (see Figure 1). These SNPs were selected based on the following findings or hypotheses: (1) the promoter SNP (SNP6: rs1800759) is known to be a functional polymorphism, altering promoter activity (Edenberg *et al*, 1999); (2) SNP2 (rs1042363) at exon 9 and SNP3 (rs1126671) at exon 7, which result in the amino-acid substitutions Gly→Arg and Val→Ile, respectively, might also affect the function of the π subunit; and (3)

some 'silent' SNPs such as SNP1 (hcv2033010) at the 3' end, SNP4 at exon 6, SNP5 (rs7694646) at intron 4, and SNP7 (rs1984362) at the 5' end might influence π subunit expression as well, either directly (eg, through effects on regulation or mRNA stability) or indirectly via linkage disequilibrium (LD) with other functional variants such as SNP2, SNP3, or SNP6.

The allele frequencies and PCR conditions for these markers have been validated by ABI. The heterozygosity of each of these markers in the ABI database is >0.375 in an EA population, which corresponds to a minor allele frequency >0.25 . These markers, spanning the full length of *ADH4* gene, cover a range of 28 752 bp, with an average spacing of 4792 bp.

Genotyping. Genomic DNA was extracted from peripheral blood by standard methods. The seven *ADH4* SNPs were genotyped with a fluorogenic 5' nuclease assay method, that is, the TaqMan technique (Shi *et al*, 1999), using the ABI PRISM[®] 7900 Sequence Detection System (ABI, Foster City, CA, USA). PCR was performed in a final volume of 3 μ l, including 1.5 μ l $2 \times$ TaqMan Universal PCR Master Mix (ABI), 0.075 μ l $20 \times$ Assays-on-Demand or 0.0375 μ l $40 \times$ Assays-by-Design (ABI), 0.03 μ l $100 \times$ BSA (New England Biolabs Inc., Beverly, MA), and 1 ng DNA. All genotyping was performed in duplicate and compared to ensure validity of the data. Mismatched genotypes, which constituted $<0.5\%$ of the total number of duplicate genotypes performed, were discarded.

Statistical Analysis

Linkage disequilibrium analysis. Different populations have different generations of ancestries and different degrees of admixture, which may result in different LD decays from the initial generations to the current generations. Different LD decay is one factor that results in different LD block sizes between populations. Thus, in the present study, the pairwise LD between any two *ADH4* markers was analyzed separately for EAs and AAs. The D' value for each LD pair was calculated and visualized through the program Haploview (Barrett *et al*, 2005) (see Figure 2a and b).

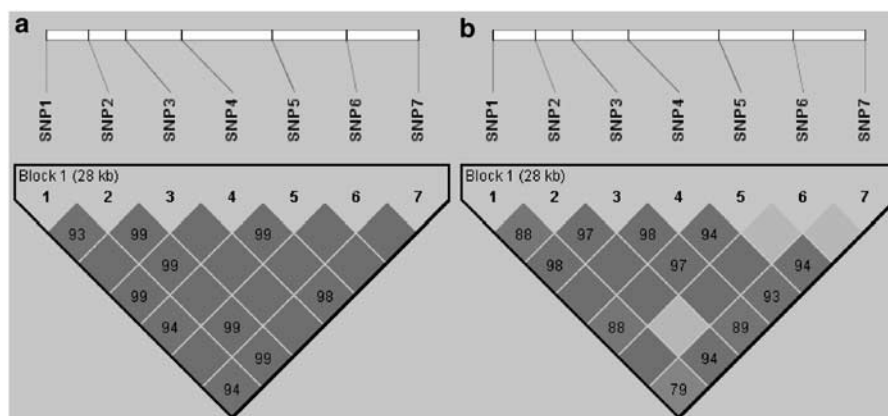


Figure 2 LDs among seven *ADH4* markers in (a) unrelated EA subjects and (b) unrelated AA subjects. $D' = 1.00$ in the blank red squares; the numbers inside the darker squares are $D' \times 100\%$; the lighter squares represent low r^2 values.

Hardy–Weinberg equilibrium test. Deviation from Hardy–Weinberg equilibrium (HWE) expectations, that is, Hardy–Weinberg disequilibrium (HWD), can indicate a valid gene–phenotype association (Feder *et al*, 1996; Nielsen *et al*, 1998; Jiang *et al*, 2001; Hoh *et al*, 2001; Lee, 2003; Hao *et al*, 2004). The HWD test as an association method in cases only has been reported to be more powerful than a case–control study, especially for a recessive model (Nielsen *et al*, 1998). HWE was tested within populations, and separately for cases and controls. The test in a mixed population was conducted as an exploratory evaluation of population stratification effects. HWE tests in the groups with mixed affection status (ie, those resulting from combining cases and controls) were used to determine whether the markers were in HWE within each population. HWE for each marker was tested using goodness-of-fit implemented in the program PowerMarker (Liu and Muse, 2004). Genotype frequency distributions of the seven markers within each phenotype group were also considered together and tested by an omnibus HWD test via the program ‘min_p_test’ (see below).

In addition to the use of the HWD test to evaluate associations between genes and diseases, the HWD test must be applied to determine what analysis methods (software programs) are valid for haplotype reconstruction (see below), because some methods are valid only in the presence of HWE.

Two HWD measures (ie, F and J) can be used to fine-map the disease risk locus. These measures are calculated using the following equations: $F = \Delta/H_e$ (Feder *et al*, 1996) and $J = \Delta/2H'_e$ (Jiang *et al*, 2001), where $\Delta = H_e - H_o$, H_e and H_o are the expected and the observed frequencies, respectively, of the heterozygotes in cases (assuming HWE), and H'_e is the expected frequency of the non-disease-associated homozygotes in cases. F and J are decreasing functions of θ , where θ is the recombination rate between the marker and the disease susceptibility locus. Higher values of F and J reflect greater LD between marker and disease locus. The marker locus with either the highest F or J value was taken as the marker closest to the disease susceptibility locus. The value of F depends not only on the genetic distance between the marker locus of interest and the disease susceptibility locus, but also on the allele frequencies at the marker locus. J does not hinge on the allele frequencies at the marker locus.

Case–control comparisons for allele and genotype frequencies. The allele and genotype frequencies of the *ADH4* markers among EAs and AAs are shown in Tables 1 and 2. Associations between the alleles, genotypes, and phenotypes were tested by comparing the allele and genotype frequency distributions between cases and controls with the exact tests in the program PowerMarker.

To avoid type I error due to multiple comparisons without information loss due to the overly conservative Bonferroni correction, the genotypes of these seven markers were considered together in an omnibus test using the program ‘min_p_test’ (Hao *et al*, 2004; personal communication), to obtain adjusted global p -values and potentially to identify real susceptibility SNPs among these markers. ‘min_p_test’ provides three p -values based on three

different statistics including the Pearson χ^2 test statistic (PE) for genotypewise analysis, the Cochran–Armitage χ^2 trend test statistic (CA) for allelewise analysis, and the HWD statistic for HWD test.

Haplotype Frequency Analysis

Multilocus haplotypes usually are more informative than any single marker. However, the *ADH4* markers are in HWD in cases, violating the assumption of the expectation–maximization (EM) algorithm upon which many haplotype-reconstruction programs are based. The Bayesian approach and the Partition Ligation algorithm may still be applied to reconstruct haplotypes even when HWE is violated (Niu *et al*, 2002). The program PHASE 2.0.2, based on a combination of the Bayesian statistical method and the so-called ‘divide-and-conquer’ (ie, Partition Ligation) strategy, which is claimed to be more accurate in reconstructing haplotypes than the EM algorithm, was used for this purpose (Stephens *et al*, 2001; Stephens and Donnelly, 2003). PHASE not only provides the overall haplotype frequency distribution in the whole sample, but also directly compares the haplotype frequency distributions between cases and controls.

The haplotype frequency distribution depends on the LD among markers, which varies by population. In order to infer the haplotypes more accurately, haplotypes were reconstructed separately for EAs and AAs.

RESULTS

(1) Pairwise analysis in different populations showed that the *ADH4* markers were in strong LD. In EAs, the seven *ADH4* markers were in strong LD ($D' > 0.9$) (see Figure 2a). In AAs, six *ADH4* markers were in strong LD ($D' > 0.9$), with SNP1 in weaker LD with the others ($D' > 0.79$) (Figure 2b). These seven markers can be taken as representing a single haplotype block (Gabriel *et al*, 2002). There were no significant differences in LD among the *ADH4* markers between cases and controls (within each population).

(2) The genotype frequency distributions of the seven *ADH4* markers were in HWD in the EA patients with alcohol dependence or cocaine dependence, but were in HWE in the EA healthy controls (see Table 3).

In EAs, the genotype frequency distributions of all seven *ADH4* markers were in HWD in patients with alcohol dependence or cocaine dependence, but were in HWE in patients with opioid dependence and in controls. After Bonferroni correction, where α was set at 0.007 ($= 0.05/7$; 7 is the number of markers), five markers (SNP2–6) and seven markers remained in HWD in patients with alcohol dependence and cocaine dependence, respectively. SNP2 was in HWD in subjects with drug dependence without comorbid alcohol dependence, even after Bonferroni correction. As mentioned above, the Bonferroni correction is very conservative, as the markers were in LD and their genotypes were correlated. The omnibus HWD test, which obviates the Bonferroni correction, showed that the genotype frequency distributions of *ADH4* markers were in HWD in all affected groups of EAs (see global p -values for HWD tests in Table 3).

Table 1 Genotype and Allele Frequencies of Seven ADH4 Markers in EAs

	AD (n = 334)		Coc (n = 178)		Op (n = 95)		DD (n = 207)		DD-only (n = 66)		Con (n = 317)	
	N	f _o	N	f _o	N	f _o	N	f _o	N	f _o	N	f _o
<i>SNP1</i>												
T/T	38	0.119	25	0.144	10	0.108	26	0.129	10	0.152	24	0.078
T/C	116	0.363	55	0.316	30	0.323	65	0.323	24	0.364	140	0.458
C/C	166	0.519	94	0.540	53	0.570	110	0.547	32	0.485	142	0.464
T	192	0.300	105	0.302	50	0.269	117	0.291	44	0.333	188	0.307
C	448	0.700	243	0.698	136	0.731	285	0.709	88	0.667	424	0.693
<i>SNP2</i>												
T/T	34	0.107	26	0.149	9	0.098	27	0.135	12	0.185	18	0.060
T/C	107	0.338	46	0.264 ^a	27	0.293 ^a	54	0.270 ^a	20	0.308 ^a	131	0.440
C/C	176	0.555	102	0.586	56	0.609	119	0.595	33	0.508	149	0.500
T	175	0.276	98	0.282	45	0.245	108	0.270	44	0.338	167	0.280
C	459	0.724	250	0.718	139	0.755	292	0.730	86	0.662	429	0.720
<i>SNP3</i>												
A/A	44	0.135	26	0.148	10	0.105	27	0.132	10	0.154	24	0.078
A/G	119	0.365	56	0.318	32	0.337	67	0.328	24	0.369	147	0.476
G/G	163	0.500	94	0.534	53	0.558	110	0.539	31	0.477	138	0.447
A	207	0.317	108	0.307	52	0.274	121	0.297	44	0.338	195	0.316
G	445	0.683	244	0.693	138	0.726	287	0.703	86	0.662	423	0.684
<i>SNP4</i>												
C/C	43	0.134	25	0.146	10	0.108	26	0.131	10	0.159	24	0.077
A/C	114	0.356	53	0.310	30	0.323	63	0.318	23	0.365	145	0.468
A/A	163	0.509	93	0.544	53	0.570	109	0.551	30	0.476	141	0.455
C	200	0.313	103	0.301	50	0.269	115	0.290	43	0.341	193	0.311
A	440	0.688	239	0.699	136	0.731	281	0.710	83	0.659	427	0.689
<i>SNP5</i>												
A/A	37	0.114	24	0.138	9	0.096	25	0.124	10	0.154	18	0.058
A/T	110	0.340	48	0.276	29	0.309	58	0.287	22	0.338	139	0.450
T/T	177	0.546	102	0.586	56	0.596	119	0.589	33	0.508	152	0.492
A	184	0.284	96	0.276	47	0.250	108	0.267	42	0.323	175	0.283
T	464	0.716	252	0.724	141	0.750	296	0.733	88	0.677	443	0.717
<i>SNP6</i>												
A/A	67	0.209	39	0.223	15	0.163	40	0.199	16	0.242	44	0.145
A/C	122	0.381 ^{a,b}	60	0.343 ^b	35	0.380 ^b	74	0.368 ^b	25	0.379 ^b	154	0.507
C/C	131	0.409	76	0.434	42	0.457	87	0.433	25	0.379	106	0.349
A	256	0.400	138	0.394	65	0.353	154	0.383	57	0.432	242	0.398
C	384	0.600	212	0.606	119	0.647	248	0.617	75	0.568	366	0.602
<i>SNP7</i>												
A/A	35	0.108	22	0.127	9	0.096	23	0.114	10	0.156	20	0.065
A/G	110	0.340	50	0.289	29	0.309	60	0.299	22	0.344	134	0.434
G/G	179	0.552	101	0.584	56	0.596	118	0.587	32	0.500	155	0.502
A	180	0.278	94	0.272	47	0.250	106	0.264	42	0.328	174	0.282
G	468	0.722	252	0.728	141	0.750	296	0.736	86	0.672	444	0.718

AD, alcohol dependence; Coc, cocaine dependence; Op, opioid dependence; DD, drug dependence (ie, cocaine dependence and/or opioid dependence); DD-only, drug dependence without comorbid alcohol dependence; Con, controls; n, sample size; N, individual numbers (for genotype) or chromosome numbers (for alleles); f_o, observed genotype or allele frequency.

^aThe marker with highest *F* value among all markers. $F = (H_e - H_o)/H_e$, where H_e and H_o are the expected and the observed frequencies of heterozygote of a marker.

^bThe marker with highest *J* value among all markers. $J = (H_e - H_o)/2H_e$, where H_e is the expected frequency of the non-disease-associated homozygote of a marker.

Table 2 Genotype and Allele Frequencies of Seven *ADH4* Markers in AAs and HWD Tests in AAs, EAs and EAs+AAs

AA														EA	EA+AA
AD (n = 102)		Coc (n = 140)		Op (n = 57)		DD (n = 149)		DD-only (n = 59)		Con (n = 48)		Case+Con (n = 209)	Case+Con (n = 717)	Case+Con (n = 926)	
N	f _o	N	f _o	N	f _o	N	f _o	N	f _o	N	f _o	p	p	p	
SNP1															
T/T	1	0.010	2	0.015	1	0.018	2	0.014	1	0.017	1	0.021	0.188	0.177	0.172
T/C	33	0.330	40	0.296	12	0.218	42	0.286	14	0.237	14	0.292			
C/C	66	0.660	93	0.689	42	0.764	103	0.701	44	0.746	33	0.688			
T	35	0.175	44	0.163	14	0.127	46	0.156	16	0.136	16	0.167			
C	165	0.825	226	0.837	96	0.873	248	0.844	102	0.864	80	0.833			
SNP2															
T/T	0	0.000	1	0.007	1	0.018	1	0.007	1	0.017	1	0.021	1.000	0.074	0.014
T/C	19	0.188	25	0.184	9	0.161	26	0.176	12	0.203	10	0.208			
C/C	82	0.812	110	0.809	46	0.821	121	0.818	46	0.780	37	0.771			
T	19	0.094	27	0.099	11	0.098	28	0.095	14	0.119	12	0.125			
C	183	0.906	245	0.901	101	0.902	268	0.905	104	0.881	84	0.875			
SNP3															
A/A	2	0.020	3	0.022	2	0.036	3	0.021	1	0.017	1	0.021	0.323	0.242	0.141
A/G	32	0.320	41	0.306	15	0.268	43	0.295	15	0.259	15	0.313			
G/G	66	0.660	90	0.672	39	0.696	100	0.685	42	0.724	32	0.667			
A	36	0.180	47	0.175	19	0.170	49	0.168	17	0.147	17	0.177			
G	164	0.820	221	0.825	93	0.830	243	0.832	99	0.853	79	0.823			
SNP4															
C/C	1	0.010	2	0.015	1	0.018	2	0.014	1	0.017	1	0.021	0.213	0.140	0.115
A/C	34	0.337	43	0.316	15	0.263	45	0.304	16	0.271	14	0.298			
A/A	66	0.653	91	0.669	41	0.719	101	0.682	42	0.712	32	0.681			
C	36	0.178	47	0.173	17	0.149	49	0.166	18	0.153	16	0.170			
A	166	0.822	225	0.827	97	0.851	247	0.834	100	0.847	78	0.830			
SNP5															
A/A	1	0.010	2	0.015	2	0.036	2	0.014	1	0.017	1	0.021	0.473	0.163	0.024
A/T	19	0.188	24	0.178	9	0.161	25	0.170	11	0.190	9	0.191			
T/T	81	0.802	109	0.807	45	0.804	120	0.816	46	0.793	37	0.787			
A	21	0.104	28	0.104	13	0.116	29	0.099	13	0.112	11	0.117			
T	181	0.896	242	0.896	99	0.884	265	0.901	103	0.888	83	0.883			
SNP6															
A/A	59	0.590	82	0.607	34	0.618	87	0.596	34	0.586	33	0.688	0.553	0.015	0.000
A/C	36	0.360	47	0.348	17	0.309	51	0.349	21	0.362	15	0.313			
C/C	5	0.050	6	0.044	4	0.073	8	0.055	3	0.052	0	0.000			
A	154	0.770	211	0.781	85	0.773	225	0.771	89	0.767	81	0.844			
C	46	0.230	59	0.219	25	0.227	67	0.229	27	0.233	15	0.156			
SNP7															
A/A	1	0.010	1	0.007	1	0.018	2	0.014	1	0.017	0	0.000	0.394	0.097	0.006
A/G	16	0.160	21	0.157	7	0.127	21	0.144	9	0.155	10	0.208			
G/G	83	0.830	112	0.836	47	0.855	123	0.842	48	0.828	38	0.792			
A	18	0.090	23	0.086	9	0.082	25	0.086	11	0.095	10	0.104			
G	182	0.910	245	0.914	101	0.918	267	0.914	105	0.905	86	0.896			

AA, African-Americans; EA, European-Americans; AD, Coc, Op, DD; DD-only, Con, n, N, f_o, see Table 1 footnote; Case, all patients; p, p-value for HWE test (goodness-of-fit). Italic numbers denote p < 0.05.

Table 3 *p*-Values for Allelewise and Genotypewise Case-Control Comparisons and for HWD Tests in EAs

	AD			Coc			Op			DD			DD-only			Con	
	Allele ^a	Genotype ^a	HWD	Allele	Genotype	HWD	Allele	Genotype	HWD	Allele	Genotype	HWD	Allele	Genotype	HWD	HWD	HWD
SNP1	0.755	0.031	0.011	0.813	0.003	0.001	0.361	0.055	0.107	0.554	0.005	0.003	0.597	0.109	0.062	0.184	
SNP2	0.905	0.011	0.005	1.000	1.5E-04	5.0E-06	0.400	0.029	0.051	0.733	7.5E-05	1.7E-05	0.170	0.004	0.004	0.108	
SNP3	0.950	0.005	0.005	0.772	8.8E-04	0.001	0.270	0.052	0.062	0.498	0.002	0.003	0.563	0.099	0.162	0.067	
SNP4	0.976	0.005	0.002	0.738	0.001	0.001	0.251	0.040	0.110	0.460	0.002	0.002	0.464	0.078	0.094	0.122	
SNP5	0.952	0.003	0.002	0.730	3.1E-05	9.5E-05	0.354	0.035	0.102	0.565	2.2E-04	2.1E-04	0.383	0.023	0.082	0.073	
SNP6	0.900	0.004	2.0E-04	0.875	0.002	1.9E-04	0.310	0.085	0.115	0.645	0.007	0.002	0.423	0.077	0.077	0.313	
SNP7	0.891	0.023	0.008	0.721	0.002	0.001	0.466	0.085	0.097	0.554	0.005	0.001	0.277	0.046	0.105	0.238	
Global ^b	0.993	0.007(7)	0.001(7)	0.969	<10 ⁻⁶ (7)	<10 ⁻⁶ (7)	0.326	0.044(7)	0.022(7)	0.672	0.0004(7)	<10 ⁻⁶ (7)	0.400	0.008(1)	0.014(1)		

AD, Coc, Op, DD, DD-only, Con, see Table 1 footnote. 'E', scientific format of number, that is, 'E-n' = $\times 10^{-n}$. Italic numbers denote $p < 0.05$.

^aAllelewise and genotypewise analyses compared the allele and genotype frequency distributions between cases and controls, respectively. HWD tests were performed within each phenotype subgroup.

^bGlobal *p*-values were derived by the program 'min_p_test', which is a simulation program with three χ^2 statistics (CA, Cochran-Armitage χ^2 , corresponding to the allelewise analysis; PE, Pearson χ^2 , corresponding to the genotypewise analysis; HWD, corresponding to HWD test); the numbers in parentheses are the susceptibility SNP numbers.

Among these markers, SNP6 and SNP2 were in most significant HWD in cases with alcohol dependence and cocaine dependence, respectively (see Table 3). A goodness-of-fit test showed that these deviations from HWE were attributable to the excess of the observed homozygote frequency (f_o) (see Table 1) over the expected homozygote frequency (f_e) (data not shown).

According to the values of *F*, SNP6 and SNP2 were the markers closest to the risk loci for alcohol dependence and drug dependence (including cocaine and opioid dependence), respectively. But after correcting for the effects of allele frequencies of markers on HWD, that is, as reflected in the values of *J*, SNP6 is the marker closest to the risk loci for both alcohol dependence and drug dependence (*J* values not shown).

In AAs, all of the markers were in HWE in patients and controls. Although SNP5 and SNP7 were nominally in HWD in patients with opioid dependence ($p = 0.020$ and 0.029 , respectively), the effects were not significant after Bonferroni correction (see Table 2, *p*-values not listed).

Within EAs and AAs (mixing cases and controls; see second-to-last and third-to-last columns in Table 2), most *ADH4* markers were in HWE, except SNP6 in EAs ($p = 0.015$). When mixing EAs and AAs (see the last column in Table 2), over half of these markers were in HWD, which was expected given that the population was not stratified.

(3) The genotypes of *ADH4* markers were associated with alcohol dependence and drug dependence in EAs (*p*-values listed in Table 3).

The allele and genotype frequencies of the seven *ADH4* markers are listed in Tables 1 and 2. In EAs, there were nominally significant differences in genotype, but not in any allele, frequency distributions between controls and subjects with alcohol dependence, cocaine dependence, opioid dependence, or any drug dependence without comorbid alcohol dependence (see Table 3). All of these associations resulted from the higher frequency of homozygotes of each marker in cases than in controls (see Table 1). Before Bonferroni correction, the genotypes of all seven SNPs were associated with alcohol dependence and cocaine dependence ($p < 0.05$); three SNPs, that is, SNP2, SNP4, and SNP5, were associated with opioid dependence; and three SNPs, that is, SNP2, SNP5, and SNP7, were associated with drug dependence without comorbid alcohol dependence. After Bonferroni correction ($\alpha = 0.05/7 = 0.007$), four SNPs (ie, SNP3–6) were associated with alcohol dependence, all seven SNPs were associated with cocaine dependence, and one SNP (ie, SNP2) was associated with drug dependence without comorbid alcohol dependence.

The omnibus tests using the 'min_p_test' program showed compelling evidence that (1) all seven SNPs were susceptibility markers for both alcohol dependence ($P_{PE} = 0.007$) and drug dependence ($P_{PE} = 0.0004$) (highly significant in cocaine dependence ($P_{PE} < 10^{-6}$) and modestly significant in opioid dependence ($P_{PE} = 0.044$)), and (2) one of the SNPs (SNP2) was a susceptibility marker for drug dependence without comorbid alcohol dependence ($P_{PE} = 0.008$), an association identified despite the small sample size available for this subgroup.

In AAs, no significant differences were found in either allele or genotype frequencies between cases and controls.

(4) The program PHASE was used to compute estimated haplotype frequency distributions and directly compare these distributions between cases and controls. There were no significant differences in the seven-variant haplotype frequency distributions between cases and controls, in either EAs or AAs (data not shown, all $p > 0.05$).

DISCUSSION

The present study provides strong evidence supporting the hypothesis that *ADH4* is a susceptibility gene for alcohol dependence, and presents the first evidence that it is also a susceptibility gene for drug dependence. These findings were obtained using the HWD test in a case-only design and were confirmed using a case-control design.

In this study, the average age for controls was 28.1 years, more than 10 years younger than that (39.4 years) for cases. Many healthy controls have not completely passed the risk age for manifesting alcohol dependence or drug dependence. The healthy controls have a probability (\approx prevalence of disease) to develop disease in their lifetime, and this probability increases when the prevalence is moderate (eg, alcohol dependence or major depression), so that the case-control association design becomes less powerful than a case-only study. That is, some associations that can be discovered by the use of a case-only study might not be detected using a case-control design. This is due to the fact that there are likely to be some control subjects who, although presently unaffected, have the genetic liability for affection and may be affected in the future. Therefore, under these circumstances, the case-only study design is more sensitive for the detection of allelic association between susceptibility alleles and diseases. The HWD test is an association method for case-only study, which has been demonstrated to be more powerful than case-control study in some circumstances. This seems to be particularly true for recessive diseases (Nielsen *et al*, 1998). Because, to our knowledge, this method has not been previously employed in psychiatric genetics (except in one previous study, Luo *et al*, 2005), it was necessary to confirm the results obtained with this method using a conventional case-control association study as a positive control.

A disease has the potential to exert strong selection on the genome, which can result in HWD for some specific loci. This deviation from HWE could be very significant if it occurs right at the disease risk locus (Nielsen *et al*, 1998; Wittke-Thompson *et al*, 2005). The markers in LD with the risk locus may be in HWD as well; alternatively, HWD may indicate that the markers in HWD are in close proximity to the disease risk locus. The detection of HWD has recently been shown to be useful as a method for detecting gene-phenotype association (Feder *et al*, 1996; Nielsen *et al*, 1998; Jiang *et al*, 2001; Hoh *et al*, 2001; Lee, 2003; Hao *et al*, 2004; Wittke-Thompson *et al*, 2005; Luo *et al*, 2005). The present study provides an additional illustration of this approach. We found that the *ADH4* markers were in HWE in EA healthy controls, but were in HWD in EA cases with alcohol dependence or drug (ie, cocaine or opioid) dependence, which strongly suggests that the HWD reflects a gene-phenotype association. When considering all the markers together in an omnibus test (using the program

'min_p_test') to avoid multiple comparisons, we obtained global p -values that ranged from 10^{-6} to 0.022 for HWD tests in all phenotype groups (see P_{HWD} in Table 3), which is a strong indication that *ADH4* is in LD with the putative disease risk locus (or may be a risk locus itself). In other words, there are strong associations between diseases (including alcohol dependence and drug dependence) and *ADH4* in EAs.

To confirm the findings we obtained using HWD tests, we also compared cases with controls on the allele, genotype, and haplotype frequency distributions within EA and AA populations. We found that (1) the allelwise analysis (using the program PowerMarker) did not detect any significant difference in the allele frequency distribution for any marker between cases and controls within EAs or AAs. Considering all of the markers together in an omnibus test to avoid multiple tests, we obtained global p -values > 0.05 for all allelwise analyses (see P_{CA} in Table 3). These negative results may be attributable to HWD, because under HWD, the marker alleles are not independent of one another, which leads to the case that the effects of disease-predisposing alleles may be 'masked' by other non-disease-predisposing alleles (Cordell, 2002). This reduces the power for the allelwise analysis (Sasieni, 1997; Nielsen *et al*, 1998; Luo *et al*, 2005) and may even invalidate it. (2) A multilocus haplotype is actually the subset of every single-locus allele; both allele and haplotype reflect the features of chromosomes in the population. Similar to the allelwise analysis, the haplotypewise analysis (using the program PHASE) in this study showed significant differences in the haplotype frequency distributions between cases and controls neither in EAs nor in AAs. Although these negative results (directly derived from the program PHASE) might be partly attributable to the low power of PHASE for case-control comparison, we speculate that they more likely result from HWD, because under HWD, as with the allelwise analysis, the haplotypes are not independent of one another. Consequently, disease-predisposing haplotypes may be 'masked' by other non-disease-predisposing haplotypes, which thus reduces the power for, or even invalidates, the haplotypewise analysis. We have also demonstrated that the haplotypewise analysis is potentially less powerful than the diplotypewise analysis (Luo *et al*, 2005). A diplotype (ie, a haplotype pair) is the subset of every single-locus genotype; both genotype and diplotype represent the types of chromosome pairs in each individual. Diplotype is predicted by PHASE too, but diplotypewise analysis is exempt from the HWE requirement and thus preserves power. (3) Genotypewise analysis is valid regardless of the presence of HWE and thus should preserve statistical power even in the presence of strong HWD. Very interestingly, as predicted, the genotypewise analysis showed that there were significant differences in genotype frequency distributions for all seven *ADH4* markers between cases and controls in EAs (see Table 3). After Bonferroni correction, the genotypes of four *ADH4* markers were significantly associated with alcohol dependence; seven markers were associated with drug dependence (mainly cocaine dependence); and one marker was associated with drug dependence without comorbid alcohol dependence. Bonferroni correction treats every association test as an independent test, but since the seven *ADH4* markers are in strong LD, this

approach is clearly too conservative for setting the significance level, which could result in information loss. To avoid multiple comparisons and the concomitant inflation of type I error, we considered all of the markers together in an omnibus test using the program 'min_p_test'; this approach identified seven susceptibility markers for alcohol dependence (global $P_{PE} = 0.007$) and drug dependence (cocaine dependence, global $P_{PE} < 10^{-6}$; opioid dependence, global $P_{PE} = 0.044$) and one susceptibility marker for drug dependence without alcohol dependence (global $P_{PE} = 0.008$). These associations are population specific, that is, they were seen only in EAs, not in AAs. (We did not exclude the possibility that other unknown *ADH4* markers might be associated with alcohol dependence or drug dependence in AAs, and did not exclude the possibility that negative results in AAs are attributable to small sample size for that population.) Therefore, in the present study, the allele and haplotype analyses were considered exploratory, and we emphasized the genotype-wise analysis, whence we drew conclusions.

Comparing the genotypewise analyses and the HWD tests, we find that (1) the findings from the genotypewise analyses basically coincide with those obtained from the HWD tests. That is, the findings in the case-only study are confirmed by the case-control association study, which is *post hoc* evidence that the HWD test is a valid association method, and which is also *post hoc* evidence that makes other explanations (except for this indication of association) for HWD highly unlikely. (2) Most associations from the HWD tests are more significant than those from the case-control comparisons (with lower *p*-values—compare the *p*-values between two tests or compare P_{HWD} with P_{PE} in Table 3). In other words, the HWD method can be more powerful than the case-control association method, supporting the conclusion by Nielsen *et al* (1998).

The fact that the *ADH4* markers were in HWE in healthy controls but in HWD in cases not only indicates a gene-phenotype association, but is also consistent with a recessive mode of inheritance of these markers (Feder *et al*, 1996; Nielsen *et al*, 1998). The fact that the associations were seen between the genotypes and the phenotypes, but not between the alleles and the phenotypes, also suggests a recessive mode of inheritance of these susceptibility markers, because when a trait-associated marker acts via a recessive mode of inheritance and is under HWD, the effect of the recessive allele (ie, the disease-risk allele) can be 'masked' by the dominant allele (ie, the non-disease-risk allele).

Both the HWD test and the case-control design as association methods are potentially vulnerable to population stratification, including admixture effects, which could result in spurious associations. For example, when EAs and AAs are mixed, over half of the markers are observed to be in HWD (see the last column in Table 2), which results from population stratification. Thus, we guarded against this confound by testing for HWD and making comparisons only within each population. Additionally, in a second study (Luo *et al*, in press), we used the structured association method to stratify our sample, and obtained results consistent with those reported here. Similarly, in the follow-up study (Luo *et al*, in press), the associations from the present case-only and case-control studies were

confirmed using the transmission disequilibrium test (TDT) in an independent sample of small nuclear families.

An outlier marker with extreme HWD usually results from genotyping error, so HWD has also been used to detect such errors (Hoh *et al*, 2001). However, the genotyping technique applied in this study has a low measured error rate (<0.5%) and all the genotyping was performed in duplicate. Furthermore, none of the seven markers is an apparent outlier according to the HWD statistics, so HWD resulting from genotyping error is not a likely explanation for the observed results.

HWD in cases can also be useful for fine-mapping a disease locus. According to the *F* values (Feder *et al*, 1996), which reflect the extent of HWD, SNP6 and SNP2 were the markers closest to the risk loci for alcohol dependence and drug dependence (including cocaine and opioid dependence), respectively. After correcting for the influence of marker allele frequency on HWD, that is, according to the *J* values (Jiang *et al*, 2001), which also reflect the extent of HWD and are equivalent to P_{excess} in a case-control design (Devlin and Risch, 1995), SNP6 is the marker closest to the risk loci for both alcohol dependence and drug dependence.

Association between the *ADH4* gene and alcohol dependence accords with the functional role of the π subunit in the development of alcohol dependence. However, we noted that drug dependence was also associated with *ADH4* gene polymorphisms. Interestingly, both in the HWD test and in the case-control comparison, the association of *ADH4* gene with drug dependence reached levels of statistical significance that were at least as great as those for alcohol dependence. This suggests two major possibilities: (1) the susceptibility to drug dependence that is attributable to *ADH4* is shared with alcohol dependence susceptibility (ie, it is independently a risk factor for each of the two phenotypes); or (2) the associations between the *ADH4* gene and drug dependence are driven by the actual associations between *ADH4* and alcohol dependence, because many patients with drug dependence have comorbid alcohol dependence. In addressing this question, we noted that some of the less robust associations, which we presumed to be due to smaller sample size, were found between the *ADH4* gene and drug dependence only (ie, without comorbid alcohol dependence). This supports the first explanation, although further study of this question with a larger sample of such individuals is warranted. This explanation may help to shed light on the pharmacological mechanism by which disulfiram, an ALDH blocker that is thought to exert its treatment effect in alcohol dependence by interfering with acetaldehyde metabolism, decreases the relapse rate among cocaine-dependent individuals, even in the absence of comorbid alcohol dependence (Carroll *et al*, 1993, 1998, 2000, 2004; George *et al*, 2000; Petrakis *et al*, 2000).

In summary, findings from this study show that the seven markers at *ADH4* are associated with alcohol dependence and drug dependence in EAs. Our results suggest that, of these SNPs, the functional variant SNP2 (rs1042363) at exon 9 and SNP6 (rs1800759) in the promoter region are the most important and may affect the risk for both alcohol dependence and drug dependence. Other susceptibility SNPs are in strong LD with SNP2 or SNP6. If the disease risk locus is not SNP2 or SNP6 itself, the putative disease

risk locus may reside within an LD block containing SNP2 or SNP6 at *ADH4* gene.

ACKNOWLEDGEMENTS

This work was supported in part by funds from the US Department of Veterans Affairs (the VA Medical Research Program, the VA Connecticut–Massachusetts Mental Illness Research, Education and Clinical Center (MIRECC), and the VA Research Enhancement Award Program (REAP) research center), NIH Grants K02-MH01387, R01-DA12849, R01-DA12690, K24-DA15105, R01-AA11330, P50-AA12870, K08-13732, K24-AA13736, and M01-RR06192 (University of Connecticut General Clinical Research Center), a National Alliance for Research on Schizophrenia and Depression (NARSAD) Young Investigator Award (X Luo), and the NIH Biological Sciences Training Program (BSTP) MH14276 (BZ Yang). Ann Marie Lacobelle and Greg Kay provided excellent technical assistance. Dr Ke Hao (Harvard EPH) kindly provided access to his software. Dr Hongyu Zhao (Yale EPH) provided helpful comments on the manuscript.

REFERENCES

- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. American Psychiatric Press: Washington, DC.
- Anokhina IP, Kogan BM, Man'kovskaia IV, Reshchikova EV, Stanishvskaya AV (1990). The similarity of the pathogenetic mechanisms of alcoholism and drug dependence and the means of searching for agents for treating these diseases. *Farmakol Toksikol* 53: 4–9 (Russian).
- Barrett J, Fry B, Maller J, Daly J (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263–265.
- Carroll KM, Fenton LR, Ball SA, Nich C, Frankforter TL, Shi J et al (2004). Efficacy of disulfiram and cognitive behavior therapy in cocaine-dependent outpatients: a randomized placebo-controlled trial. *Arch Gen Psychiatry* 61: 264–272.
- Carroll KM, Nich C, Ball SA, McCance-Katz EF, Frankforter T, Rounsaville BJ (2000). One year follow-up of disulfiram and psychotherapy for cocaine-alcohol abusers: sustained effects of treatment. *Addiction* 95: 1335–1349.
- Carroll KM, Nich C, Ball SA, McCance-Katz EF, Rounsaville BJ (1998). Treatment of cocaine and alcohol dependence with psychotherapy and disulfiram. *Addiction* 93: 713–727.
- Carroll KM, Ziedonis D, O'Malley SS, McCance-Katz E, Gordon L, Rounsaville BJ (1993). Pharmacologic interventions for abusers of alcohol and cocaine: a pilot study of disulfiram versus naltrexone. *Am J Addict* 2: 77–79.
- Cordell HJ (2002). Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum Mol Genet* 11: 2463–2468.
- Devlin B, Risch N (1995). A comparison of linkage disequilibrium measures for fine scale mapping. *Genomics* 29: 311–322.
- Edenberg HJ, Jerome RE, Li M (1999). Polymorphism of the human alcohol dehydrogenase 4 (*ADH4*) promoter affects gene expression. *Pharmacogenetics* 9: 25–30.
- Feder JN, Gnirke A, Thomas W, Tsuchihashi Z, Ruddy DA, Basava A et al (1996). A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet* 13: 399–408.
- Foroud T, Dick D, Xuei X, Goate A, Porjesz B, Begleiter H et al (2003). Association of SNPs in the ADH gene cluster with alcoholism. *Am J Hum Genet* 73(5 Suppl): 521 (meeting abstract).
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B et al (2002). The structure of haplotype blocks in the human genome. *Science* 296: 2225–2229.
- George TP, Chawarski MC, Pakes J, Carroll KM, Kosten TR, Schottenfeld RS (2000). Disulfiram versus placebo for cocaine dependence in buprenorphine-maintained subjects: a preliminary trial. *Biol Psychiatry* 47: 1080–1086.
- Glatt MM (1967). Problems common to alcoholism and drug dependence. *WHO Chron* 21: 293–303.
- Gossop M (2004). Multiple substance use and multiple dependencies. In: Kranzler HR, Tinsley JA (eds). *Dual Diagnosis and Psychiatric Treatment*. Marcel Dekker: New York. pp 129–156.
- Hao K, Xu X, Laird N, Wang X, Xu X (2004). Power estimation of multiple SNP association test of case-control study and application. *Genet Epidemiol* 26: 22–30.
- Hasin D, Aharonovich E, Liu X, Mammen Z, Matseoane K, Carr L et al (2002). Alcohol and ADH2 in Israel: Ashkenazis, Sephardics, and recent Russian immigrants. *Am J Psychiatry* 159: 1432–1434.
- Heil SH, Badger GJ, Higgins ST (2001). Alcohol dependence among cocaine-dependent outpatients: demographics, drug use, treatment outcome and other characteristics. *J Stud Alcohol* 62: 14–22.
- Hoehe MR, Kopke K, Wendel B, Rohde K, Flachmeier C, Kidd KK et al (2000). Sequence variability and candidate gene analysis in complex disease: association of mu opioid receptor gene variation with substance dependence. *Hum Mol Genet* 19: 2895–2908.
- Hoh J, Wille A, Ott J (2001). Trimming, weighting, and grouping SNPs in human case-control association studies. *Genome Res* 11: 2115–2119.
- Jiang R, Dong J, Wang D, Sun FZ (2001). Fine-scale mapping using Hardy-Weinberg disequilibrium. *Ann Hum Genet* 65: 207–219.
- Konishi T, Calvillo M, Leng AS, Feng J, Lee T, Lee H et al (2003). The ADH3*2 and CYP2E1 c2 alleles increase the risk of alcoholism in Mexican American men. *Exp Mol Pathol* 74: 183–189.
- Kreek MJ, Nielsen DA, LaForge KS (2004). Genes associated with addiction: alcoholism, opiate, and cocaine addiction. *Neuromol Med* 5: 85–108.
- Lee WC (2003). Searching for disease-susceptibility loci by testing for Hardy-Weinberg disequilibrium in a gene bank of affected individuals. *Am J Epidemiol* 158: 397–400.
- Li TK, Bosron WF (1987). Distribution and properties of human alcohol dehydrogenase isoenzymes. *Ann NY Acad Sci* 492: 1–10.
- Li TK, Bosron WF, Dafeldecker WP, Lange LG, Vallee BL (1977). Isolation of PI-alcohol dehydrogenase of human liver: is it a determinant of alcoholism? *Proc Nat Acad Sci USA* 74: 4378–4381.
- Liu K, Muse S (2004). PowerMarker: new genetic data analysis software. Version 3.0. Free program distributed by the author over the internet from <http://www.powermarker.net>. Accessed October 1, 2004.
- Long JC, Knowler WC, Hanson RL, Robin RW, Urbanek M, Moore E et al (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.
- Luo X, Kranzler HR, Zuo L, Yang BZ, Lappalainen J, Gelernter J (2004). ADH4 gene variation is associated with alcohol dependence and drug dependence in European Americans: results from family-controlled and population-structured association studies. *Pharmacogenet Genomics* (in press).
- Luo X, Kranzler HR, Zhao H, Gelernter J (2003). Haplotypes at the OPRM1 locus are associated with susceptibility to substance

- dependence in European-Americans. *Am J Med Genet (Neuropsychiatric Genet)* **120B**: 97–108.
- Luo X, Kranzler HR, Zuo L, Wang S, Blumberg HP, Gelernter J (2005). CHRM2 gene predisposes to alcohol, drug dependence and affective disorder: results from an extended population structured association study. *Hum Mol Genet* **14**: 2421–2434.
- Muramatsu T, Wang ZC, Fang YR, Hu KB, Yan H, Yamada K *et al* (1995). Alcohol and aldehyde dehydrogenase genotypes and drinking behavior of Chinese living in Shanghai. *Hum Genet* **96**: 151–154.
- Nielsen DM, Ehm MG, Weir BS (1998). Detecting marker-disease association by testing for Hardy–Weinberg disequilibrium at a marker locus. *Am J Hum Genet* **63**: 1531–1540.
- Niu T, Qin ZD, Xu X, Liu JS (2002). Bayesian haplotype inference for multiple linked single-nucleotide polymorphisms. *Am J Hum Genet* **70**: 157–169.
- O'Neill J, Cardenas VA, Meyerhoff DJ (2001). Separate and interactive effects of cocaine and alcohol dependence on brain structures and metabolites: quantitative MRI and proton MR spectroscopic imaging. *Addict Biol* **6**: 347–361.
- Petrakis IL, Carroll KM, Nich C, Gordon LT, McCance-Katz EF, Frankforter T *et al* (2000). Disulfiram treatment for cocaine dependence in methadone maintained opioid addicts. *Addiction* **95**: 219–228.
- Reich T, Edenberg HJ, Goate A, Williams JT, Rice JP, Van Eerdewegh P *et al* (1998). Genome-wide search for genes affecting the risk for alcohol dependence. *Am J Med Genet* **81**: 207–221.
- Robinson JE, Heaton RK, O'Malley SS (1999). Neuropsychological functioning in cocaine abusers with and without alcohol dependence. *J Int Neuropsychol Soc* **5**: 10–19.
- Saccone NL, Kwon JM, Corbett J, Goate A, Rochberg N, Edenberg HJ *et al* (2000). A genome screen of maximum number of drinks as an alcoholism phenotype. *Am J Med Genet* **96**: 632–637.
- Sasieni PD (1997). From genotypes to genes: doubling the sample size. *Biometrics* **53**: 1253–1261.
- Schinka JA, Town T, Abdullah L, Crawford FC, Ordorica PI, Francis E *et al* (2002). A functional polymorphism within the mu-opioid receptor gene and risk for abuse of alcohol and other substances. *Mol Psychiatry* **7**: 224–228.
- Shi MM, Myrand SP, Bleavins MR, de la Iglesia FA (1999). High throughput genotyping for the detection of a single nucleotide polymorphism in NAD(P)H quinone oxidoreductase (DT diaphorase) using TaqMan probes. *Mol Pathol* **52**: 295–299.
- Stephens M, Donnelly P (2003). A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* **73**: 1162–1169.
- Stephens M, Smith NJ, Donnelly P (2001). A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**: 978–989.
- Szeto CY, Tang NL, Lee DT, Stadlin A (2001). Association between mu opioid receptor gene polymorphisms and Chinese heroin addicts. *Neuroreport* **6**: 1103–1106.
- Thomasson HR, Crabb DW, Edenberg HJ, Li TK, Hwu HG, Chen CC *et al* (1994). Low frequency of the ADH2*2 allele among Atayal natives of Taiwan with alcohol use disorders. *Alcohol Clin Exp Res* **18**: 640–643.
- Williams JT, Begleiter H, Porjesz B, Edenberg HJ, Foroud T, Reich T *et al* (1999). Joint multipoint linkage analysis of multivariate qualitative and quantitative traits. II. Alcoholism and event-related potentials. *Am J Hum Genet* **65**: 1148–1160.
- Wittke-Thompson JK, Pluzhnikov A, Cox NJ (2005). Rational inferences about departures from Hardy–Weinberg equilibrium. *Am J Hum Genet* **76**: 967–986.
- Zinn-Justin A, Abel L (1999). Genome search for alcohol dependence using the weighted pairwise correlation linkage method: interesting findings on chromosome 4. *Genet Epidemiol* **17**(Suppl 1): S421–S426.