Human Molecular Genetics, 2006, Vol. 15, No. 9 ***1539–1549***
doi:10.1093/hmg/ddl073
Advance Access published on March 29, 2006

Association of alcohol dehydrogenase genes with alcohol dependence: a comprehensive analysis

Howard J. Edenberg[1,*], Xiaoling Xuei[1], Hui-Ju Chen[1], Huijun Tian[1], Leah Flury Wetherill[1], Danielle M. Dick[2], Laura Almasy[3], Laura Bierut[2], Kathleen K. Bucholz[2], Alison Goate[2], Victor Hesselbrock[4], Samuel Kuperman[5], John Nurnberger[1], Bernice Porjesz[6], John Rice[2], Marc Schuckit[7], Jay Tischfield[8], Henri Begleiter[6] and Tatiana Foroud[1]

[1]Indiana University School of Medicine, Indianapolis, IN, USA, [2]Washington University School of Medicine, St Louis, MO, USA, [3]Southwest Foundation for Biomedical Research, San Antonio, TX, USA, [4]University of Connecticut, Farmington, CT, USA, [5]University of Iowa Carver College of Medicine, Iowa City, IA, USA, [6]SUNY Health Sciences Center, Brooklyn, NY, USA, [7]University of California, San Diego, CA, USA and [8]Rutgers University, Piscataway, NJ, USA

Received January 18, 2006; Revised and Accepted March 22, 2006

Linkage evidence indicated that gene(s) located on chromosome 4q, in the region of the alcohol dehydrogenase (*ADH*) genes, affected risk for alcoholism. We genotyped 110 single nucleotide polymorphisms (SNPs) across the seven *ADH* genes and analyzed their association with alcoholism in a set of families with multiple alcoholic members, using the pedigree disequilibrium test. There was strong evidence that variations in *ADH4* are associated with alcoholism: 12 SNPs were significantly associated. The region of strongest association ran from intron 1 to 19.5 kb beyond the 3′ end of the gene. Haplotype tag SNPs were selected for the block in the *ADH4* gene that provided evidence of association and subsequently used in association analysis; the haplotype was significantly associated with alcoholism ($P = 0.01$) There was weaker evidence that variations in *ADH1A* and *ADH1B* might also play a role in modifying risk. Among African-Americans, there was evidence that the *ADH1B*3* allele was protective.

INTRODUCTION

Alcoholism (alcohol dependence) is a common, complex disease, with significant genetic contributions to the risk. Because drinking ethanol (beverage alcohol) is a necessary environmental condition for manifestation of this disease, the metabolism of ethanol is clearly relevant to the risk. The primary pathway of ethanol metabolism involves oxidation to acetaldehyde, catalyzed by alcohol dehydrogenases (ADHs), followed by further oxidation to acetate, catalyzed by aldehyde dehydrogenases (ALDHs) (1).

There is evidence for linkage of a region on chromosome 4q to the risk for alcoholism. The Collaborative Study on the Genetics of Alcoholism (COGA), a large study of families in which at least three individuals met diagnostic criteria for alcohol dependence, provided evidence that a broad region on chromosome 4q was linked to the risk for alcoholism (2–4). The evidence came primarily from the unaffected members of these families (2) and from reduced allele sharing among siblings discordant for the alcohol-dependence phenotype (3). Variance-component analysis of the COGA sample showed that the strongest evidence for linkage to alcoholism was in a broad region of chromosome 4q centered near the *ADH* gene cluster (4). A bivariate analysis of alcoholism and event-related potentials increased the evidence for linkage in this region (4). Analysis of a related quantitative trait, the maximum number of drinks ever consumed within a 24 h period, showed strongest linkage to a narrower region of chromosome 4q, centered near a cluster of *ADH* genes (5). In a study of a Southwestern Native American population,

*To whom correspondence should be addressed. Tel: +1 3172742353; Fax: +1 3172744686; Email: edenberg@iupui.edu

there was also evidence for linkage of alcohol dependence to markers in the ADH region of chromosome 4 (6).

Humans have seven *ADH* genes tightly clustered on chromosome 4q22 in a head-to-tail array extending over ~365 kb. The order of the genes (from 5′ to 3′) is *ADH7-ADH1C-ADH1B-ADH1A-ADH6-ADH4-ADH5*, running from qter toward the centromere. All of the ADH enzymes are broad substrate oxidoreductases that use NAD^+/NADH as cofactors (reviewed in 1). *ADH1A*, *ADH1B* and *ADH1C* encode α, β and γ subunits, respectively; these can form heterodimers, and are defined as class I ADHs. Class I ADHs have K_m for ethanol in the range of 0.05–34 mM (1). *ADH4* encodes π-ADH, a class II ADH with K_m for ethanol of 34 mM. *ADH5* encodes χ-ADH, which is also a glutathione-dependent formaldehyde dehydrogenase; χ-ADH has very low affinity for ethanol. *ADH7* encodes σ-ADH (also known as μ-ADH); it is the most efficient of these enzymes at oxidizing retinol. The protein encoded by *ADH6* has not been purified from tissue.

The primary site of ethanol oxidation is the liver, in which there are high concentrations of most of the ADHs, except σ-ADH (1,7). Consideration of the enzyme concentrations in liver and the kinetic properties of the ADHs suggest that the class I enzymes (encoded by *ADH1A*, *ADH1B* and *ADH1C*) and the class II enzyme (encoded by *ADH4*) make the most significant contribution to ethanol metabolism (1,7,8). It has been calculated that the class I enzymes contribute ~70% of the total ethanol oxidizing capacity of the liver at an ethanol concentration of 22 mM (0.1%; 0.08% is defined as legally intoxicated in virtually all states in the USA), and the class II enzyme contributes ~30% (1,8).

The pharmacokinetics of ethanol metabolism influences the risk for alcohol dependence. Many studies have shown that coding variations in the genes encoding three alcohol-metabolizing enzymes, *ADH1B*, *ADH1C* and *ALDH2*, are associated with risk for alcoholism (1,9). The *ADH1B*2* allele in which arginine 47 is replaced with histidine encodes the β2 subunit, which has a 40-fold higher V_{max} than the β1 subunit encoded by *ADH1B*1* (1,8). *ADH1B*2* is relatively common among Asians, where it has been shown to be protective against alcoholism (1,10,11); although rarer in Europeans, it has also been shown to be protective in that group (12,13). A different allele, *ADH1B*3*, encodes the β3 subunit in which arginine 269 is replaced by cysteine; the β3 subunit has a 30-fold higher V_{max} than the β1 subunit (1,8,14). *ADH1B*3* is relatively common among individuals of African ancestry, and individuals carrying this polymorphism have a higher rate of metabolizing alcohol (15). *ADH1C*, which encodes the γ subunit, has polymorphisms at amino acids 271 and 349; these are in high linkage disequilibrium (LD), with the 271Arg–349Ile form called γ1 (encoded by *ADH1C*1*) and the 271Gln–349Val called γ2 (encoded by *ADH1C*2*) (16). The V_{max} of γ1 is about twice that of γ2 (1,8). Although the *ADH1C*1* allele has a reported protective effect against alcohol dependence in Asian populations (1,10,17,18), the LD between *ADH1C*1* and *ADH1B*2* obscures the independent effect of *ADH1C*1* (19,20). In a Mexican-American population, *ADH1C*1* was protective (21), and in a Native American population, there was modest evidence of linkage between *ADH1C*1* and binge drinking and of association with alcoholism (22).

Despite their potential contribution to alcohol metabolism, there have been fewer studies of *ADH4* variants. There is a coding variation in *ADH4* in which an single nucleotide polymorphism (SNP) leads to either isoleucine or valine at position 308 of the π-ADH encoded by that gene (23). [Note that (23) and some other references use an alternate, non-standard nomenclature (24); this can be confusing. We will use the official nomenclature of the HUGO Gene Nomenclature Committee http://www.gene.ucl.ac.uk/nomenclature/genefamily/ADH.shtml and NCBI.] The kinetic properties with ethanol are very similar for the two forms, but the Val308 form is less stable than the Ile308 form (23). A functional mutation in the *ADH4* promoter has also been reported. The −136A allele (numbered from the translational start site in the currently recommended manner; this corresponds to the −75A allele numbered from the transcription start site in the original publication) has a 2-fold higher promoter activity in transfected cells (25).

The role of ADH enzymes in the pharmacokinetics of alcohol metabolism, along with the linkage results for alcoholism and alcoholism-related phenotypes, make the *ADH* genes prime candidates for the genes in the chromosome 4q region that influence risk for alcoholism. Rather than limiting our analysis to the previously reported coding and promoter variants in a few of the genes, we systematically analyzed the association of all seven *ADH* genes with alcoholism, selecting multiple SNPs to query each of the genes.

RESULTS

One hundred and ten SNPs were genotyped in the 417 kb region containing the 364 kb *ADH* gene cluster plus 21.5 kb downstream and 31.2 kb upstream; the genes in this cluster are all transcribed in the direction from qter to pter. Genotyping was concentrated within and flanking the genes and at lower density between them (Table 1, Fig. 1). Over 80% of the SNPs had minor allele frequencies (MAFs) > 0.2 in the COGA European American sample, and over 88% had MAF > 0.15. There were differences between the European-American and African-American samples in the MAFs; in 88 of the 110 SNPs, the difference was greater than 0.05 (Table 1). The extent of LD among the SNPs genotyped across the *ADH* gene cluster was examined in the European-American samples (Fig. 1). Overall, there was high LD within each gene and lower LD between genes. Within each gene, D' estimates for adjacent SNPs were greater than 0.80 for 87.6% of the adjacent comparisons, excluding those comparisons which included an SNP with MAF < 0.05. The extent of coverage was compared with the data in HapMap using Tagger (26). Tagger only analyzed 69 of the SNPs we used (others were not genotyped in HapMap or had MAF < 0.05). This subset of SNPs showed an average r^2 of 0.79 with the 388 SNPs (MAF > 0.05) in a 378 kb region; 68% of the SNPs in the region had $r^2 > 0.8$ with one of our SNPs. Within genes (and the 10 kb flanking each end), the coverage is even better. In *ADH4*, the 11 SNPs Tagger analyzed (out of 18 genotyped) had mean r^2 of 0.91 with the HapMap SNPs and had $r^2 > 0.8$ with 87% of the 76 known SNPs. Therefore, the association analyses of the tested SNPs carry information on a very large fraction

Table 1. SNPs, positions and MAFs

	SNP	Chromosome location[a]	Gene	Gene location[b]	CA MAF[c]	AA MAF[d]
1	rs894363	100, 734, 024		Intergenic	0.379	0.165
2	rs1154476	100, 717, 905	*ADH7*	Upstream	0.384	0.138
3	rs1154473	100, 717, 090	*ADH7*	Upstream	0.379	0.451
4	SB_P1[e]	100, 713, 762	*ADH7*	Upstream	0.057	0.010
5	SB_P2[e]	100, 713, 644	*ADH7*	5′-UTR	0.130	0.051
6	rs1154470	100, 713, 515	*ADH7*	IVS1	0.327	0.165
7	rs1154468	100, 711, 435	*ADH7*	IVS1	0.328	0.161
8	rs971074	100, 699, 039	*ADH7*	Exon 6	0.104	0.177
9	rs1154459	100, 698, 281	*ADH7*	IVS6	0.342	0.126
10	rs2584464	100, 696, 227	*ADH7*	IVS7	0.484	0.322
11	rs1154454	100, 695, 520	*ADH7*	IVS7	0.182	0.443
12	**rs284779**[f]	**100**, **695**, **439**	*ADH7*	**IVS7**	0.458	0.170
13	*rs284786*[f]	*100*, *691*, *155*	*ADH7*	*3′-UTR*	0.272	0.459
14	rs2584463	100, 689, 875	*ADH7*	Downstream	0.207	0.154
15	rs1553434	100, 681, 048		Intergenic	0.004	0.210
16	rs284794	100, 680, 207		Intergenic	0.109	0.157
17	rs1442484	100, 663, 354		Intergenic	0.216	0.353
18	rs283406	100, 655, 649		Intergenic	0.079	0.082
19	rs2165671	100, 642, 326		Intergenic	0.385	0.172
20	rs1229849	100, 641, 863		Intergenic	0.280	0.147
21	rs980972	100, 636, 425	*ADH1C*	Upstream	0.386	0.171
22	rs1662037	100, 635, 997	*ADH1C*	Upstream	0.275	0.153
23	rs1789924	100, 631, 464	*ADH1C*	Upstream	0.389	0.179
24	rs4147541	100, 631, 335	*ADH1C*	Upstream	0.285	0.405
25	rs3133158	100, 627, 790	*ADH1C*	IVS1	0.285	0.160
26	rs1789915	100, 623, 549	*ADH1C*	Exon 4	0.261	0.144
27	rs1693426	100, 623, 508	*ADH1C*	IVS4	0.387	0.171
28	rs2241894	100, 623, 311	*ADH1C*	Exon 5	0.238	0.412
29	rs1631460	100, 621, 798	*ADH1C*	IVS5	0.389	0.169
30	rs1693482	100, 621, 143	*ADH1C*	Exon 6 Q272R	0.383	0.178
31	rs904096	100, 620, 762	*ADH1C*	IVS6	0.388	0.173
32	rs1789903	100, 619, 219	*ADH1C*	IVS6	0.389	0.169
33	rs2298755	100, 618, 216	*ADH1C*	IVS7	0.388	0.170
34	rs698	100, 617, 967	*ADH1C*	Exon 8 *ADH1C*2* I349V	0.389	0.185
35	P351T	100, 617, 961	*ADH1C*	Exon 8 P351T	0.005	0
36	rs1662060	100, 617, 019	*ADH1C*	IVS8	0.383	0.165
37	rs1614972	100, 615, 333	*ADH1C*	IVS8	0.330	0.493
38	rs1789896	100, 614, 162	*ADH1C*	Downstream	0.450	0.275
39	rs1229863	100, 609, 564		Intergenic	0.179	0.045
40	rs2866152	100, 608, 245		Intergenic	0.216	0.123
41	rs1789891	100, 607, 597		Intergenic	0.175	0.044
42	*rs1229982*	*100*, *601*, *110*	*ADH1B*	*Upstream*	0.220	0.443
43	*rs1159918*	*100*, *600*, *187*	*ADH1B*	*Upstream*	0.348	0.355
44	*rs1353621*	*100*, *598*, *753*	*ADH1B*	*IVS1*	0.370	0.123
45	rs1229983	100, 597, 180	*ADH1B*	Exon 2	0.024	0.044
46	rs1229984	100, 596, 497	*ADH1B*	Exon 3 *ADH1B*2* R47H	0.034	0.019
47	rs2066701	100, 595, 591	*ADH1B*	IVS3	0.304	0.122
48	rs1789883	100, 593, 553	*ADH1B*	IVS5	0.027	0.015
49	rs10033960	100, 588, 520	*ADH1B*	IVS8	0.305	0.124
50	rs2066702	100, 586, 195	*ADH1B*	Exon 9 *ADH1B*3* R369C	0.004	0.165
51	rs17033	100, 586, 123	*ADH1B*	3′-UTR	0.082	0.050
52	rs1042026	100, 585, 644	*ADH1B*	3′-UTR	0.306	0.132
53	*rs1826909*	*100*, *574*, *921*		*Intergenic*	0.373	0.171
54	rs904092	100, 571, 342	*ADH1A*	Upstream	0.180	0.194
55	rs1229966	100, 570, 611	*ADH1A*	Upstream	0.346	0.395
56	*rs4147531*	*100*, *569*, *375*	*ADH1A*	*Upstream*	0.457	0.233
57	rs3805325	100, 568, 577	*ADH1A*	IVS1	0.078	0.128
58	rs931635	100, 568, 025	*ADH1A*	IVS1	0.222	0.065
59	rs1229967	100, 564, 756	*ADH1A*	IVS3	0.223	0.055
60	rs6828526	100, 563, 065	*ADH1A*	Exon 4	0.001	0.009
61	rs1229970	100, 561, 558	*ADH1A*	IVS5	0.221	0.056
62	rs1229976	100, 559, 256	*ADH1A*	IVS6	0.224	0.073
63	rs3819197	100, 557, 687	*ADH1A*	IVS8	0.251	0.275
64	*rs2866151*	*100*, *555*, *690*	*ADH1A*	*IVS8*	0.456	0.233
65	rs1618572	100, 552, 299	*ADH1A*	Downstream	0.224	0.056
66	rs1039151	100, 549, 960	*ADH1A*	Downstream	0.257	0.290

Continued

Table 1. *Continued.*

	SNP	Chromosome location[a]	Gene	Gene location[b]	CA MAF[c]	AA MAF[d]
67	rs1230024	100, 546, 168		Intergenic	0.058	0.071
68	rs1230026	100, 542, 797		Intergenic	0.223	0.074
69	rs4699733	100, 494, 712	*ADH6*	IVS1	0.290	0.342
70	rs9307238	100, 493, 360	*ADH6*	IVS2	0.488	0.359
71	rs6833176	100, 488, 341	*ADH6*	IVS5	0.461	0.186
72	rs3857224	100, 486, 863	*ADH6*	IVS6	0.333	0.453
73	rs4147545	100, 485, 931	*ADH6*	IVS6	0.342	0.453
74	rs1893883	100, 481, 894	*ADH6*	Exon 9	0.460	0.189
75	rs2097122	100, 472, 408	*ADH6*	Downstream	0.239	0.370
76	rs4699726	100, 456, 691		Intergenic	0.489	0.179
77	rs1540053	100, 439, 332		Intergenic	0.238	0.395
78	**rs1984362**	**100, 428, 151**		**Intergenic**	0.294	0.094
79	**rs4699718**	**100, 424, 969**	***ADH4***	**Upstream**	0.298	0.122
80	rs2226896	100, 424, 515	*ADH4*	Upstream	0.077	0.035
81	rs4148884[g]	100, 423, 465	*ADH4*	Upstream	0.084	0.220
82	**rs3762894**[g]	**100, 423, 262**	**ADH4**	**Upstream**	0.166	0.225
83	rs1800760	100, 422, 804	*ADH4*	Promoter	0.427	0.086
84	rs1800759	100, 422, 687	*ADH4*	Promoter-136	0.419	0.233
85	**rs4148886**	**100, 421, 827**	**ADH4**	**IVS1**	0.242	0.490
86	**rs4699714**[g]	**100, 417, 716**	**ADH4**	**IVS3**	0.304	0.119
87	**rs7694646**	**100, 416, 910**	**ADH4**	**IVS4**	0.294	0.135
88	rs1126670	100, 409, 911	*ADH4*	Exon 6	0.315	0.217
89	rs1126671	100, 405, 592	*ADH4*	Exon 7 I308V	0.319	0.223
90	**rs1126672**	**100, 404, 990**	**ADH4**	**Exon 8**	0.305	0.165
91	**DWSHpy188I**	**100, 404, 001**	**ADH4**	**IVS8**	0.295	0.135
92	rs1126673	100, 402, 794	*ADH4*	Exon 9 I373V	0.331	0.239
93	**rs1042364**	**100, 402, 752**	**ADH4**	**3′-UTR (G388R)**	0.294	0.138
94	**rs1042365**	**100, 402, 678**	**ADH4**	**3′-UTR**	0.294	0.138
95	**rs2602866**	**100, 392, 175**		**Intergenic**	0.295	0.135
96	**rs2602846**	**100, 382, 329**		**Intergenic**	0.299	0.123
97	rs7667261	100, 368, 478	*ADH5*	Upstream	0.044	0.212
98	rs1154400	100, 367, 188	*ADH5*	5′-UTR	0.328	0.341
99	rs1154401	100, 366, 916	*ADH5*	IVS1	0.355	0.364
100	rs7683704	100, 361, 404	*ADH5*	IVS2	0.108	0.268
101	rs1154412	100, 358, 425	*ADH5*	IVS4	0.190	0.031
102	rs4699700	100, 355, 513	*ADH5*	IVS4	0.148	0.261
103	rs4699699	100, 354, 357	*ADH5*	IVS6	0.108	0.071
104	rs7683802	100, 352, 316	*ADH5*	IVS7	0.102	0.107
105	rs12697	100, 350, 554	*ADH5*	3′-UTR	0.149	0.282
106	rs13832	100, 350, 253	*ADH5*	3′-UTR	0.351	0.420
107	rs1061187	100, 350, 031	*ADH5*	3′-UTR	0.107	0.074
108	rs1230155	100, 346, 437	*ADH5*	Downstream	0.345	0.318
109	rs1238741	100, 340, 490	*METAP1*	Untranslated	0.115	0.352
110	rs1230205	100, 328, 021	*METAP1*	Intron	0.158	0.444

Bold rs numbers, significantly associated with alcoholism (DSM-IV); italic, significantly associated with alcoholism by the alternative diagnostic criteria (COGA).
[a]Position in nucleotides from chromosome 4pter, as estimated in dbSNP (build 124) or by blasting against the NCBI Human Genome assembly (build 35.1).
[b]Position within or near gene; SNPs are listed in direction of transcription, which is from qter toward pter.
[c]MAF in European Americans, calculated from the COGA dataset.
[d]MAF in African-Americans, calculated from the COGA data set.
[e]Reference (47).
[f]Reference (48).
[g]htSNPs.

of untested variations within each gene. Similar patterns of LD were observed in the smaller number of African-American samples, although the blocks of LD tended to be smaller.

The pedigree disequilibrium test (PDT) (27) was employed to test for association between each *ADH* SNP and the phenotype of alcohol dependence as defined by the DSM-IV criteria (28); DSM-IV criteria gave the strongest linkage signal in prior experiments (4). The most consistent evidence of association ($P < 0.05$) was observed using the PDT_{sum} option for the SNPs in the *ADH4* gene, extending into the intergenic region between *ADH4* and *ADH5* (Fig. 2). SNPs extending over 39 kb, from rs4148886 (in intron 1) through rs2602846 (19.5 kb downstream of exon 9), were individually associated with alcohol dependence, as were three SNPs in the upstream region of *ADH4* (Figs 2 and 3, Table 1). The SNP showing the greatest evidence of association, rs4148886, yielded a *P*-value of 0.0042; permutation testing resulted in a global significance of 0.036. Among these SNPs are two

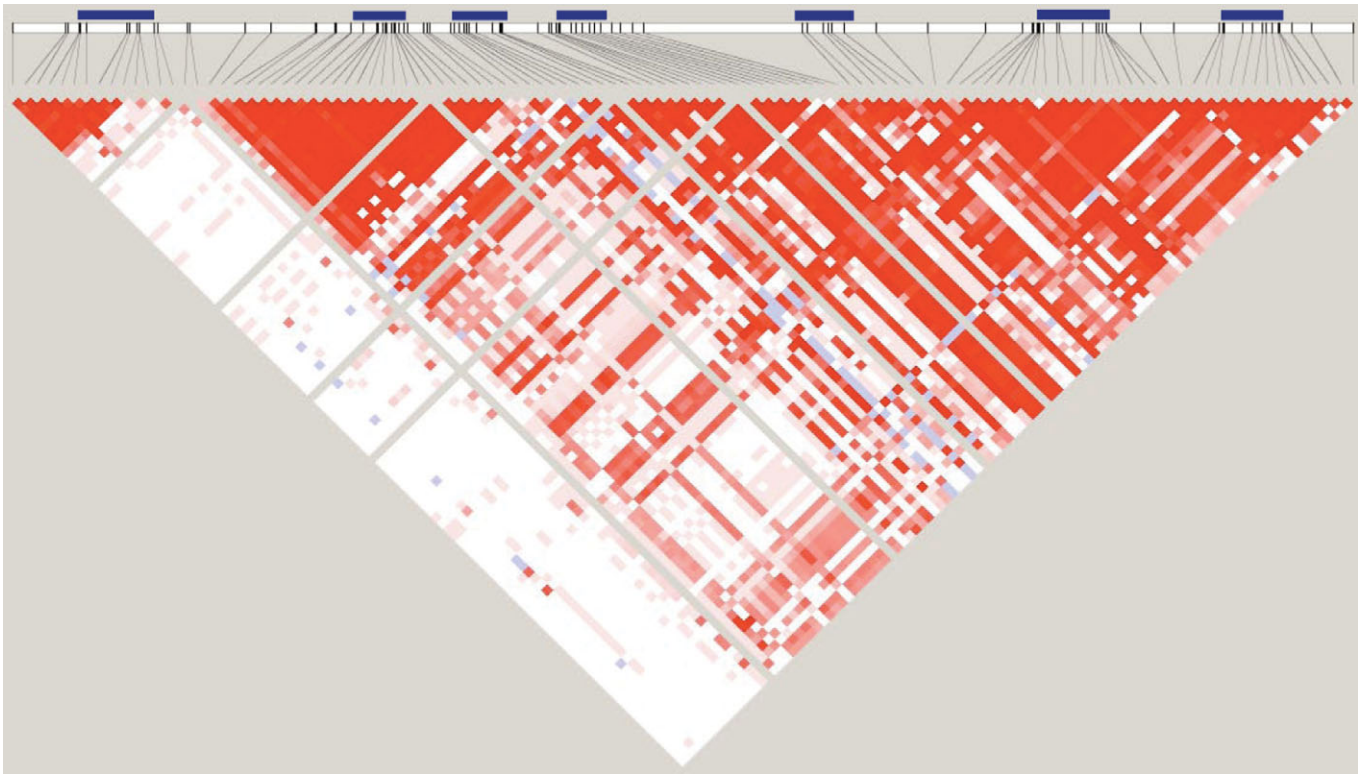

Figure 1. LD (D') between the SNPs genotyped in the *ADH* cluster: Haploview. Blue lines at top show positions of the *ADH* genes (*ADH7*, *ADH1C*, *ADH1B*, *ADH1A*, *ADH6*, *ADH4*, *ADH5* in order from qter toward cen); transcription runs from left to right.

non-synonymous SNPs in the coding region, which were associated with alcohol dependence: rs1126671 (I308V, $P = 0.06$) and rs1126673 (I373V, $P = 0.06$); rs1042364 is listed as a non-synonymous SNP in dbSNP, but lies in the 3′-non-translated region of mRNA encoding the normal peptide; it was also significant ($P = 0.02$). These were not significant when analyzed by $PDT_{average}$. The functional promoter SNP (25), rs1800759, was not significantly associated with alcohol dependence. In the remainder of the *ADH* region, only one other SNP (in *ADH7*) was significant.

Because the evidence of association with alcohol dependence was greatest in and around the *ADH4* gene, we evaluated the evidence of association of haplotypes in this region. The block structure of the SNPs that spanned the 90 kb region extending from rs2097122 to rs2602846 (SNPs 75–96, Table 1), which includes the entire *ADH4* gene plus 19.5 kb downstream and 49.8 kb upstream, was examined. One block containing 18 SNPs (rs4699718 to rs2602846) included 11 of the 12 significant ($P \leq 0.05$) individual SNPs and three marginally significant ones ($P < 0.07$). Three haplotype tag SNPs (htSNPs) were identified (rs4699714, rs3762894 and rs4148884), which successfully tagged the four haplotypes with a frequency of $\geq 5\%$ in this block. Only one of the observed htSNP haplotypes, with a frequency of 0.35%, could not be uniquely tagged with these three htSNPs. Three haplotypes were not observed in our sample. There was significant evidence for association with the htSNP haplotype ($P = 0.01$).

Exploratory genetic analyses were performed using two other definitions of alcohol dependence using a similar analytic approach to that described above. Analyses using the narrower ICD-10 criteria did not produce consistent, significant results. Analyses using the broader COGA definition of alcohol dependence [DSM-IIIR (29) plus Feighner definite alcoholism (30)] resulted in evidence of association ($P < 0.05$) with scattered SNPs within and upstream of *ADH1A* (rs2866151, rs4147531 and rs1826909) and with a series of three consecutive SNPs within the 5′ and upstream region of *ADH1B* (rs1353621, rs1159918 and rs1229982). One of these, rs1159918, was also significant with the ICD-10 definition ($P = 0.04$). A single SNP in *ADH7* (rs284786) was significant with the COGA definition.

We genotyped several coding SNPs that have been analyzed in different populations (10–13,31).The SNP that distinguishes *ADH1B*1* from *ADH1B*2* (rs1229984; Arg47 or His47 in the mature protein) has MAFs in both European-American and African-American populations <0.04 (Table 1); we found no association with alcohol dependence. Rs2066702, the SNP that distinguishes *ADH1B*1* from *ADH1B*3* (Arg369 or His369 in the mature protein) has an extremely low MAF in European-Americans and could not effectively be tested in that population. In African-Americans, rs2066702 is associated with alcohol dependence defined by ICD-10 ($P = 0.029$ and 0.046 with $PDT_{average}$ and PDT_{sum}, respectively) and with DSM-IV ($P = 0.039$ and 0.105 with $PDT_{average}$ and PDT_{sum}, respectively); the allele encoding *ADH1B*3* was the low-risk

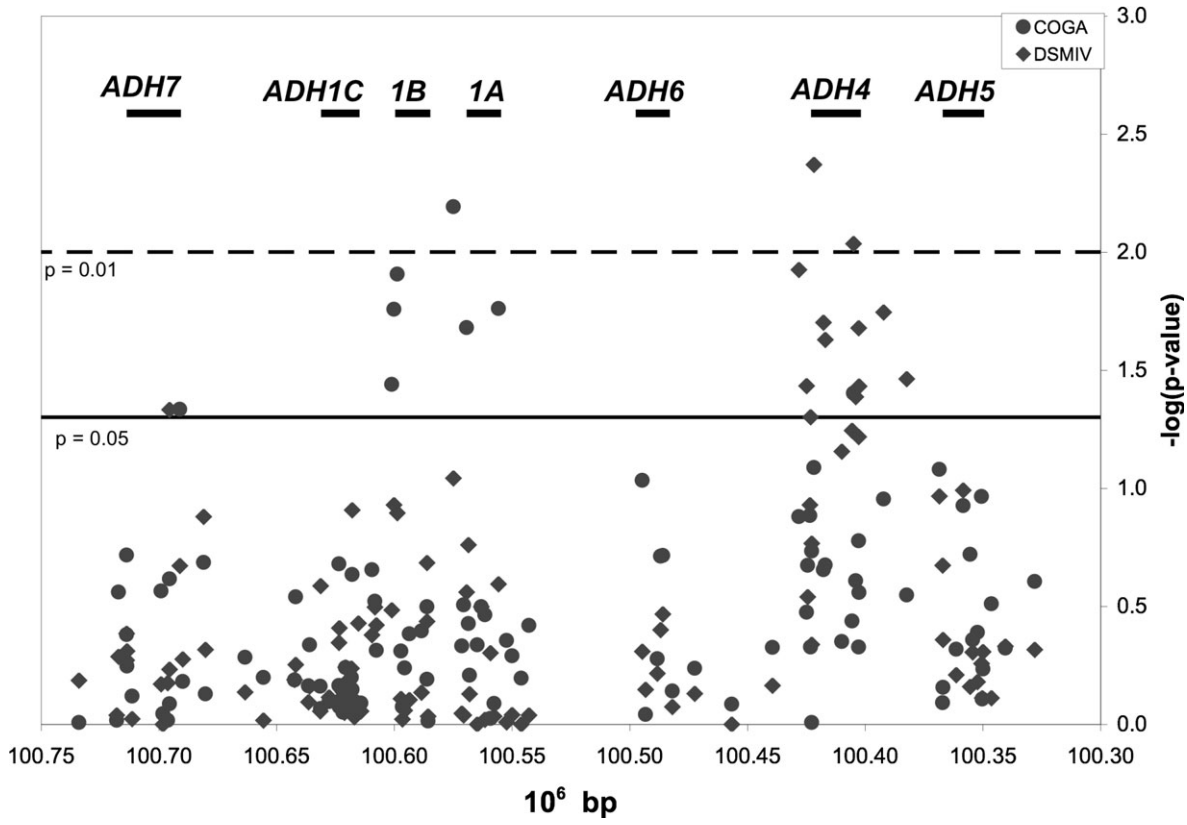

Figure 2. Association of SNPs across the *ADH* gene cluster with alcohol dependence. Results of the PDT_{sum} are plotted as $-\log$ (*P*-value). Diamonds represent results with DSM-IV definition of alcohol dependence and circles represent results using COGA criteria. Locations of the *ADH* genes are shown as lines across the top; genes are transcribed from left to right; abscissa runs from qter toward cen.

allele. In *ADH1C*, there are two coding SNPs in very high LD; most previous studies have only measured one (rs698) as a proxy for the pair. In our data on unrelated European-Americans (data not shown) or the European-American subset of COGA families, the LD between rs698 and rs1693482 is nearly complete ($D' = 1.0$, $r^2 = 0.97–1.0$, based on Haploview). Neither of these SNPs was significantly associated with alcohol dependence under any of the three definitions we examined, in either European-American or African-American families. *ADH1C-P351T* was reported in Native American populations (32); we found MAF of 0.005 and 0 in European-Americans and African-Americans, respectively.

DISCUSSION

This paper presents the most comprehensive testing to date of the association between alcohol dependence and SNPs across the entire *ADH* gene cluster. We examined 110 SNPs covering a 417 kb region, with concentrations in and near all seven of the *ADH* genes (Table 1, Fig. 1). Previous studies have focussed on a few functional SNPs in *ADH1B* and *ADH1C*. Osier *et al*. (33). reported moderate LD in a small part of this region extending from the class I genes to *ADH7*. Our data show that there is much LD between the SNPs in this region, which lie in blocks of restricted haplotypes (Fig. 1); each of the seven *ADH* genes lie in one to three blocks, as defined by Gabriel *et al*. (34). This pattern of LD affects interpretations of earlier work on coding SNPs; the coding SNPs not only affect the function of the ADH proteins, but are also in LD with potential regulatory SNPs that might affect gene expression.

The strongest finding in this analysis is the association between many SNPs and the htSNP haplotype in *ADH4* with alcohol dependence. The strongest region of associated SNPs extends 39.5 kb from intron 1 through 19.5 kb downstream of exon 9. There are also associated SNPs in the 5′ region of *ADH4*, extending 5.6 kb upstream from the initiation codon. This association was detected using the PDT_{sum} option, in which large families contribute more to the statistic. $PDT_{average}$ did not show this association.

We reported preliminary results on the first 58 SNPs, showing that eight SNPs in *ADH4* were associated with alcohol dependence (35). Prior to that time, associations with *ADH1B* and *ADH1C* had been reported, but *ADH4* had not been studied. Edenberg *et al*. (25). reported that SNP rs1800759 (then described as at −75 bp numbered from the transcription start site; now numbered −136 bp based on the translation start site) was functional, with A in that position having twice the promoter activity of C. Two other nearby SNPs, rs1800761 (then −159, now −220) and rs1800760 (then −192, now −253) did not detectably affect promoter activity (25). Edenberg *et al*. (25) predicted that the lower activity allele at rs1800759, C, increased the risk for alcohol dependence; this prediction was based on the fact that lower

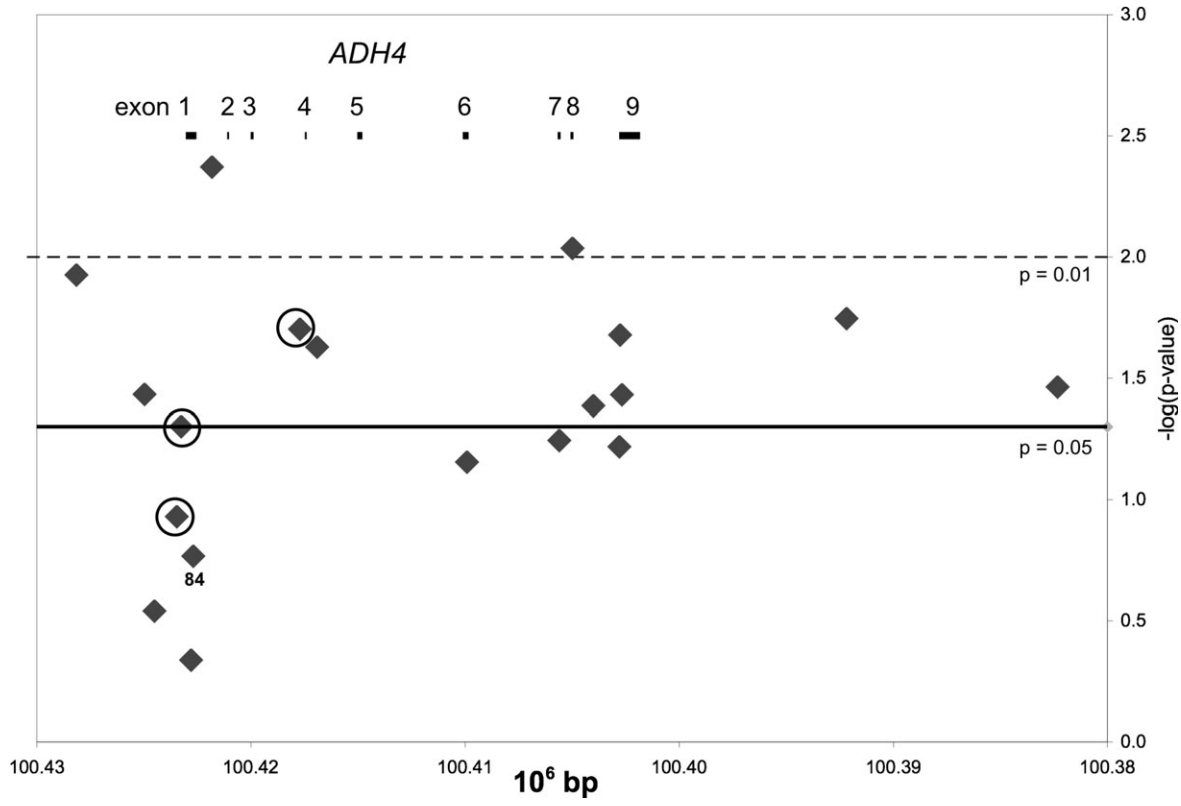

Figure 3. Association of SNPs across *ADH4* with alcohol dependence defined by DSM-IV. Results of the PDT_{sum} are plotted as −log (*P*-value). The exons are indicated across the top. Nineteen SNPs in and immediately flanking *ADH4* are shown (78–96 in Table 1; three SNPs further upstream were not significant). Circles mark the three SNPs that tag the associated haplotype, which runs from the second SNP through the last. SNP 84 is noted; this is the promoter SNP at −136 bp.

activity coding variations in *ADH1B* and *ADH1C* are the high-risk alleles. Guindalini *et al*. (36) reported that two of the three *ADH4* promoter SNPs that had been reported by Edenberg *et al*. (25). (rs1800759 and rs1800761) and several three-SNP haplotypes were associated with alcohol dependence in both European-Brazilians and African-Brazilians. The effect of rs1800759 was greater in their sample; however, rs1800759 was in strong Hardy–Weinberg disequilibrium in the European-Brazilian controls, with a large excess of heterozygotes. SNPs rs1800759, rs1800761 and rs1800760 are within 117 bp, so we did not genotype rs1800761 (the middle SNP); the LD (*D′*) between rs1800759 and rs1800760 (33 bp beyond rs1800761 was 1.0 in both European-American and African-American COGA samples and both were in Hardy–Weinberg equilibrium in our population. In our study, neither of these alleles was significantly associated with alcohol dependence.

While this paper was being prepared, there have been two reports of association of SNPs in *ADH4* with alcohol dependence and drug dependence (37,38). These studies genotyped and analyzed seven SNPs across *ADH4* in a case–control population that included many people with both alcohol and drug (primarily cocaine and opioids) dependence and analyzed the data both as case–control and cases only (the latter by Hardy–Weinberg disequilibrium). Their data did not show significant differences in allele frequencies or haplotype frequencies between cases and controls, but did show differences in genotype frequencies and highly significant Hardy–Weinberg disequilibrium in the cases-only analysis that suggested a recessive effect on both alcohol dependence and cocaine dependence in this population (37). Six of the seven SNPs they analyzed were in our preliminary analysis (35) and are reported here, along with many other SNPs that more thoroughly cover the gene (Table 1). Of these, we found four (rs1042364, rs1126671, rs7694646, rs1984362) were nominally significant and one (rs1126670) marginal using the PDT; surprisingly, the SNP they found closest to the 'functional risk locus', rs1800759, was not significant. The greater density of SNPs we used to examine *ADH4*, and the larger number of individuals, showed that what Luo *et al*. (37) considered one haplotype block appears to be split into three. Following up on their initial analysis, Luo *et al*. (38) carried out a structured association analysis of the case–control sample and a TDT analysis of a small set of nuclear families; both supported the association between alcohol dependence and SNPs in *ADH4*. The fact that their very different sample and study design showed significant association between *ADH4* and alcohol dependence strongly supports our results.

We did not find significant association with the coding SNPs in *ADH1B* in the European-American families; these SNPs had very low allele frequencies in those families (Table 1). We did find association with three adjacent SNPs in *ADH1B*, extending from intron 1 through the promoter

region to 1.5 kb upstream of the initiation codon; association in this region was with the COGA definition of alcoholism rather than DSM-IV. Thus, there is evidence that variations in *ADH1B* affect the risk for alcoholism in a population in which the coding polymorphisms are uncommon.

We found significant association with rs2066702, the SNP that defines *ADH1B*3*, in African-American families. Association was strongest with the ICD-10 definitions of alcoholism, with $P = 0.046$ for PDT_{sum} and 0.029 for $PDT_{average}$; it was also significant for DSM-IV for $PDT_{average}$ ($P = 0.039$). The high-risk allele was the C that encodes *ADH1B*1*, and the low-risk allele encodes *ADH1B*3*. The fact that the allele encoding the higher V_{max} enzyme was protective is consistent with the pattern seen for *ADH1B*2*. This supports the role of *ADH1B*3* as a functional polymorphism in affecting risk for alcoholism.

In addition, we found association with three SNPs distributed across *ADH1A* from the upstream region through exon 8, also with the COGA definition of alcoholism. These SNPs span two haplotype blocks and the intermediate SNPs do not show association. There have been no previous reports of association of *ADH1A* with alcoholism and no reports of coding variation in the protein it encodes (although a non-synonymous SNP is listed in dbSNP, rs1041977, this non-validated SNP lies in a region identical in *ADH1A* and *ADH1B* and nearly identical in *ADH1C*, suggesting that it might be a sequencing error).

Surprisingly, we did not find association with any SNPs in or near *ADH1C*, despite evidence that coding variations in this gene are associated with risk for alcoholism in several populations (21,22). We analyzed the coding SNP most frequently studied (rs698, Ile349Val) and the coding SNP in very high LD with it (rs1693482, Gln271Arg), both of which have high MAF. The effect of these coding SNPs has always been relatively small in other populations; this might explain our failure to find association in our families.

Osier *et al*. (39) confirmed the protective effect associated with the *ADH1B*2* allele in a Taiwanese sample, but found that in their sample the effect was restricted to one of two common haplotypes identical at *ADH1B* but differing at a StyI intronic SNP (rs1154458) in *ADH7*. They did not find evidence of LD across the segment between the *ADH1B* site and the *ADH7* site in that Asian population. Our data on European-Americans (and on African-Americans) also show little LD between SNPs in *ADH1B* and *ADH7* (Fig. 1). Osier *et al*. (39) suggested that their result might be due to epistasis between a variation in strong LD with the *ADH7* SNP and the *ADH1B*2* allele or to the possibility that the combination of SNPs demarks a chromosome containing protective alleles.

Birley *et al*. (40) analyzed the time course of blood and breath alcohol levels in a set of monozygotic and dizygotic twins and found that the *ADH* region of chromosome 4 contained a quantitative trait locus that accounted for 64% of the additive genetic covariation common to blood and breath alcohol at the first time point measured. They did not detect more than a very minor contribution of haplotypes for *ADH1B* and *ADH1C* to the genetic variation in metabolism at this initial time point; their population, of European descent, has a very low frequency of *ADH1B*2* alleles. They suggested that other genetic variation in the *ADH* region must explain the genetic effects on metabolism and suggested both *ADH4* and regulatory variation in general as possibilities.

Our comprehensive assessment of SNPs across the *ADH* region shows that the strongest association with alcoholism is due to variations in the *ADH4* gene and weaker association with SNPs in *ADH1A* and *ADH1B*. The associations are with non-coding SNPs. In the case of *ADH4*, the association is strongest with the SNPs extending from intron 1 to 19.5 kb past the 3′ end of the gene. Within the blocks of restricted haplotypes, it is difficult to assign which SNP(s) are 'functional' and which are merely traveling on the same chromosomes. These findings, along with the potential for epistatic interactions among SNPs, must be considered in analyzing and interpreting association data. The breadth of the original linkage peak (2–4,6) suggests that additional genes might also affect risk for alcoholism.

MATERIALS AND METHODS

Sample

The COGA is a multi-site study recruiting families at six centers across the USA: Indiana University, State University of New York Health Science Center, University of Connecticut, University of Iowa, University of California/San Diego and Washington University, St Louis (41–43). The institutional review boards of all participating institutions approved the study. Probands were identified through inpatient or outpatient alcohol treatment programs. Probands and their families were administered a poly-diagnostic instrument, the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) interview (44,45). The families that participated in the genetic phase of this study included a proband and at least two first-degree relatives who met both DSM-IIIR criteria (29) for alcohol dependence and Feighner *et al*. (30). criteria for definite alcoholism; this combination is called COGA criteria. The SSAGA also allows derivation of diagnoses based on DSM-IV (28) and ICD-10 (46) criteria. Details of the ascertainment and assessment have previously been published (41–43).

SNP genotyping

SNPs throughout the *ADH* gene cluster were mainly selected from public databases, primarily dbSNP (www.ncbi.nlm.nih.gov/SNP/); several were from Alfred (47), Buervenich *et al*. (48) and three were identified in our laboratory (25,49). Key SNPs encoding *ADH1B* and *ADH1C* coding variants and an *ADH4* promoter variant (25,49) were genotyped, despite their low MAFs. Additional SNPs were chosen throughout each of the *ADH* genes and at lower density in regions between the genes. At the time SNPs were selected, allele frequencies for most were not available, so SNPs were genotyped on two sets of 40 unrelated individuals from the Coriell European- and African-American samples to determine approximate allele frequencies. SNPs with high heterozygosities were preferentially genotyped in the full sample if they were in Hardy–Weinberg equilibrium in both test populations. Locations of the SNPs were in most cases determined

from the annotations in the NCBI human genome assembly (build 35.1); in some cases, position was determined by BLASTing the sequence against the human genome assembly. Annotations were based on BLAST alignments; in some cases, they do not match the annotations in NCBI databases, which contain some errors. Most SNPs were not in coding regions, but rather were located in intronic, 5′ and 3′ regions of the genes. SNPs are presented along the coding strand (opposite to the direction on chromosome 4q).

Genotyping was done using a modified single nucleotide extension reaction, with allele detection by mass spectrometry (Sequenom MassArray system; Sequenom, San Diego, CA, USA). All SNP genotypes were checked for Mendelian inheritance using the program PEDCHECK (50). Marker allele frequencies and heterozygosities were computed separately in the European and African-American families using the program USERM13 (51).

Statistical analyses

Because of the known ethnic differences in *ADH* allele frequencies for the functional SNPs (1,33) and our determination of frequency differences in other SNPs, genetic analyses were separately performed in European-American families (1860 individuals, 218 families) and in African-American families (279 individuals, 35 families). Families were classified based on the racial assignment of the genetically informative portions of the pedigree. Because of the small number of African-American families, we focussed on the results in the European-Americans.

To ensure that the SNP density was sufficient to evaluate the evidence of association between the *ADH* gene cluster and alcohol dependence, LD between SNPs in the same gene and across genes was evaluated using the program Haploview (52). A gene was considered sufficiently genotyped when D' between adjacent SNPs was greater than 0.80 for at least 75% of the adjacent pairwise comparisons. The haplotype block structure in this region was examined using Haploview, with blocks defined as a set of contiguous SNPs whose average D' exceeds a predetermined threshold (34,53).

The PDT (27) as implemented in the program UNPHASED (54) was used to test for association in the extended, multiplex COGA pedigrees. The PDT utilizes data from all available trios in a family, as well as discordant sibships. It produces two statistics: the '$PDT_{average}$', which averages the association statistic across all families, and the 'PDT_{sum}', which gives greater weight to families with a larger number of informative trios and discordant sibships. Our primary phenotype was alcoholism as defined by DSM-IV criteria (28), because that gave the strongest signal in previous analyses (4). The permutation test implemented in UNPHASED was employed to obtain a global level of significance for the individual SNP analyses.

The block structure of 22 SNPs in and flanking ADH4 was determined using Haploview (52). htSNPs were selected for that block of the *ADH4* gene providing evidence of association with alcoholism (SNPs 77–96, Table 1 and Fig. 3). htSNPs were selected such that haplotypes with a frequency of $\geq 5\%$ could be uniquely identified. The htSNPs were then employed to perform family-based association (PDT) analysis using haplotypes rather than single SNPs.

ACKNOWLEDGEMENTS

We thank Jinghua Zhao, Gayathri Rajan, Christopher Rush and Robert George for technical assistance with genotyping and data organization. Genotyping facilities were provided by the Center for Medical Genomics at Indiana University School of Medicine, supported in part by the Indiana 21st Century Research and Technology Fund and the Indiana Genomics Initiative (INGEN, supported in part by the Lilly Endowment, Inc.). The Collaborative Study on the Genetics of Alcoholism (COGA) (Principal Investigator: H. Begleiter; Co-Principal Investigators: L. Bierut, H. Edenberg, V. Hesselbrock, B. Porjesz) includes nine different centers where data collection, analysis and storage take place. The nine sites and Principal Investigators and Co-investigators are: University of Connecticut (V. Hesselbrock); Indiana University (H.J. Edenberg, J. Nurnberger Jr, P.M. Conneally, T. Foroud); University of Iowa (S. Kuperman, R. Crowe); SUNY Downstate Medical Center (B. Porjesz, H. Begleiter); Washington University in St Louis (L. Bierut, A. Goate, J. Rice); University of California at San Diego (M. Schuckit); Howard University (R. Taylor); Rutgers University (J. Tischfield); Southwest Foundation (L. Almasy). Zhaoxia Ren serves as the NIAAA Staff Collaborator. This national collaborative study is supported by the NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA). In memory of Theodore Reich, M.D., Co-Principal Investigator of COGA since its inception and one of the founders of modern psychiatric genetics, we acknowledge his immeasurable and fundamental scientific contributions to COGA and the field.

Conflict of Interest statement. None declared.

REFERENCES

1. Hurley, T.D., Edenberg, H.J. and Li, T.-K. (2002) The Pharmacogenomics of alcoholism. *Pharmacogenomics*: *the Search for Individualized Therapies*, Wiley-VCH, Weinheim, Germany, pp. 417–441.
2. Reich, T., Edenberg, H.J., Goate, A., Williams, J.T., Rice, J.P., Van Eerdewegh, P., Foroud, T., Hesselbrock, V., Schuckit, M.A., Bucholz, K. *et al*. (1998) Genome-wide search for genes affecting the risk for alcohol dependence. *Am. J. Med. Genet.*, **81**, 207–215.
3. Reich, T. (1996) A genomic survey of alcohol dependence and related phenotypes: results from the Collaborative Study on the Genetics of Alcoholism (COGA). *Alcohol. Clin. Exp. Res.*, **20** (suppl), 133A–137A.
4. Williams, J.T., Begleiter, H., Porjesz, B., Edenberg, H.J., Foroud, T., Reich, T., Goate, A., Van Eerdewegh, P., Almasy, L. and Blangero, J. (1999) Joint multipoint linkage analysis of multivariate qualitative and quantitative traits. II. Alcoholism and event-related potentials. *Am. J. Hum. Genet.*, **65**, 1148–1160.
5. Saccone, N.L., Kwon, J.M., Corbett, J., Goate, A., Rochberg, N., Edenberg, H.J., Foroud, T., Li, T.K., Begleiter, H., Reich, T. *et al*. (2000) A genome screen of maximum number of drinks as an alcoholism phenotype. *Am. J. Med. Genet.*, **96**, 632–637.
6. Long, J.C., Knowler, W.C., Hanson, R.L., Robin, R.W., Urbanek, M., Moore, E., Bennett, P.H. and Goldman, D. (1998) Evidence for genetic linkage to alcohol dependence on chromosomes 4 and 11 from and autosome-wide scan in an American Indian population. *Am. J. Med. Genet. (Neuropsychiatry Genet.)*, **81**, 216–221.
7. Maly, I.P., Toranelli, M. and Sasse, D. (1999) Distribution of alcohol dehydrogenase isoenzymes in the human liver acinus. *Histochem. Cell Biol.*, **111**, 391–397.

8. Lee, S.L., Hoog, J.O. and Yin, S.J. (2004) Functionality of allelic variations in human alcohol dehydrogenase gene family: assessment of a functional window for protection against alcoholism. *Pharmacogenetics*, **14**, 725–732.
9. Li, T.-K. (2000) Pharmacogenetics of responses to alcohol and genes that influence alcohol drinking. *J. Stud. Alcohol*, **61**, 5–12.
10. Thomasson, H.R., Edenberg, H.J., Crabb, D.W., Mai, X.L., Jerome, R.E., Li, T.K., Wang, S.P., Lin, Y.T., Lu, R.B. and Yin, S.J. (1991) Alcohol and aldehyde dehydrogenase genotypes and alcoholism in Chinese men. *Am. J. Hum. Genet.*, **48**, 677–681.
11. Thomasson, H.R., Crabb, D.W., Edenberg, H.J., Li, T.-K., Hwu, H.-G., Chen, C.C., Yeh, E.K. and Yin, S.-J. (1994) Low frequency of the ADH2*2 allele among Atayal natives of Taiwan with alcohol use disorders. *Alcohol. Clin. Exp. Res.*, **18**, 640–643.
12. Whitfield, J.B. (2002) Alcohol dehydrogenase and alcohol dependence: variation in genotype-associated risk between populations. *Am. J. Hum. Genet.*, **71**, 1247–1250; author reply 1250–1251.
13. Borras, E., Coutelle, C., Rosell, A., Fernandez-Muixi, F., Broch, M., Crosas, B., Hjelmqvist, L., Lorenzo, A., Gutierrez, C., Santos, M. *et al*. (2000) Genetic polymorphism of alcohol dehydrogenase in Europeans: the ADH2*2 allele decreases the risk for alcoholism and is associated with ADH3*1. *Hepatology*, **31**, 984–989.
14. Burnell, J.C., Carr, L.G., Dwulet, F.E., Edenberg, H.J., Li, T.-K. and Bosron, W.F. (1987) The human beta 3 alcohol dehydrogenase subunit differs from beta 1 by a Cys for Arg-369 substitution which decreases NAD(H) binding. *Biochem. Biophys. Res. Commun.*, **146**, 1127–1133.
15. Thomasson, H.R., Beard, J.D. and Li, T.-K. (1995) *ADH2* gene polymorphisms are determinants of alcohol pharmacokinetics. *Alcohol. Clin. Exp. Res.*, **19**, 1494–1499.
16. Hoog, J.O., Heden, L.O., Larsson, K., Jornvall, H. and von Bahr-Lindstrom, H. (1986) The gamma 1 and gamma 2 subunits of human liver alcohol dehydrogenase. cDNA structures, two amino acid replacements, and compatibility with changes in the enzymatic properties. *Eur. J. Biochem.*, **159**, 215–218.
17. Thomasson, H.R., Crabb, D.W., Edenberg, H.J. and Li, T.-K. (1993) Alcohol and aldehyde dehydrogenase polymorphisms and alcoholism. *Behav. Genet.*, **23**, 131–136.
18. Chen, W.J., Loh, E.W., Hsu, Y.P., Chen, C.C., Yu, J.M. and Cheng, A.T. (1996) Alcohol-metabolising genes and alcoholism among Taiwanese Han men: independent effect of ADH2, ADH3 and ALDH2. *Br. J. Psychiatry*, **168**, 762–767.
19. Osier, M., Pakstis, A.J., Kidd, J.R., Lee, J.-F., Yin, S.-J., Ko, H.-C., Edenberg, H.J., Lu, R.-B. and Kidd, K.K. (1999) Linkage disequilibrium at the ADH2 and ADH3 loci and risk for alcoholism. *Am. J. Hum. Genet.*, **64**, 1147–1157.
20. Chen, C.-C., Lu, R.-B., Chen, Y.-C., Wang, M.-F., Chang, Y.-C., Li, T.-K. and Yin, S.-J. (1999) Interaction between the functional polymorphisms of the alcohol-metabolism genes in protection against alcoholism. *Am. J. Hum. Genet.*, **65**, 795–807.
21. Konishi, T., Calvillo, M., Leng, A.S., Feng, J., Lee, T., Lee, H., Smith, J.L., Sial, S.H., Berman, N., French, S. *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.
22. Mulligan, C.J., Robin, R.W., Osier, M.V., Sambuughin, N., Goldfarb, L.G., Kittles, R.A., Hesselbrock, D., Goldman, D. and Long, J.C. (2003) Allelic variation at alcohol metabolism genes (*ADH1B*, *ADH1C*, *ALDH2*) and alcohol dependence in an American Indian population. *Hum. Genet.*, **113**, 325–336.
23. Stromberg, P., Svensson, S., Hedberg, J.J., Nordling, E. and Hoog, J.O. (2002) Identification and characterisation of two allelic forms of human alcohol dehydrogenase 2. *Cell. Mol. Life Sci.*, **59**, 552–559.
24. Duester, G., Farres, J., Felder, M.R., Holmes, R.S., Hoog, J.O., Pares, X., Plapp, B.V., Yin, S.J. and Jornvall, H. (1999) Recommended nomenclature for the vertebrate alcohol dehydrogenase gene family. *Biochem. Pharmacol.*, **58**, 389–395.
25. Edenberg, H.J., Jerome, R.E. and Li, M. (1999) Polymorphism of the human alcohol dehydrogenase 4 (ADH4) promoter affects gene expression. *Pharmacogenetics*, **9**, 25–30.
26. de Bakker, P.I., Yelensky, R., Pe'er, I., Gabriel, S.B., Daly, M.J. and Altshuler, D. (2005) Efficiency and power in genetic association studies. *Nat. Genet.*, **37**, 1217–1223.
27. Martin, E.R., Bass, M.P. and Kaplan, N.L. (2001) Correcting for a potential bias in the pedigree disequilibrium test. *Am. J. Hum. Genet.*, **68**, 1065–1067.
28. American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. American Psychiatric Association Press, Washington, DC.
29. American Psychiatric Association (1987) *Diagnostic and Statistical Manual of Mental Disorders*, 3rd edn, revised. American Psychiatric Association Press, Washington, DC.
30. Feighner, J.P., Robins, E., Guze, S.B., Woodruff, R.A., Winokur, G. and Munoz, R. (1972) Diagnostic criteria for use in psychiatric research. *Arch. Gen. Psychiatry*, **26**, 57–63.
31. Shen, Y.-C., Fan, J.-F., Edenberg, H.J., Li, T.-K., Cui, Y.-H., Wang, Y.-F., Tian, C.-H., Zhou, C.-F., Zhou, R.-L., Wang, J. *et al*. (1997) Polymorphism of *ADH* and *ALDH* genes among four ethnic groups in China and effects upon the risk for alcoholism. *Alcohol. Clin. Exp. Res.*, **21**, 1272–1277.
32. Osier, M.V., Pakstis, A.J., Goldman, D., Edenberg, H.J., Kidd, J.R. and Kidd, K.K. (2002) A proline–threonine substitution in codon 351 of ADH1C is common in Native Americans. *Alcohol. Clin. Exp. Res.*, **26**, 1759–1763.
33. Osier, M.V., Pakstis, A.J., Soodyall, H., Comas, D., Goldman, D., Odunsi, A., Okonofua, F., Parnas, J., Schulz, L.O., Bertranpetit, J. *et al*. (2002) A global perspective on genetic variation at the *ADH* genes reveals unusual patterns of linkage disequilibrium and diversity. *Am. J. Hum. Genet.*, **71**, 84–99.
34. Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M. *et al*. (2002) The structure of haplotype blocks in the human genome. *Science*, **296**, 2225–2229.
35. Foroud, T., Dick, D., Xuei, X., Goate, A., Porjesz, B., Begleiter, H., Reich, T. and Edenberg, H.J. (2003) Association of SNPs in the *ADH* gene cluster with alcoholism. *Am. J. Hum. Genet.*, **73**, 521.
36. Guindalini, C., Scivoletto, S., Ferreira, R.G., Breen, G., Zilberman, M., Peluso, M.A. and Zatz, M. (2005) Association of genetic variants in alcohol dehydrogenase 4 with alcohol dependence in Brazilian patients. *Am. J. Psychiatry*, **162**, 1005–1007.
37. Luo, X., Kranzler, H.R., Zuo, L., Lappalainen, J., Yang, B.Z. and Gelernter, J. (2005) *ADH4* gene variation is associated with alcohol dependence and drug dependence in European Americans: results from HWD tests and case–control association studies. *Neuropsychopharmacology*, Epub Oct 12.
38. Luo, X., Kranzler, H.R., Zuo, L., Yang, B.Z., Lappalainen, J. and Gelernter, J. (2005) *ADH4* gene variation is associated with alcohol and drug dependence: results from family controlled and population-structured association studies. *Pharmacogenet. Genomics*, **15**, 755–768.
39. Osier, M.V., Lu, R.B., Pakstis, A.J., Kidd, J.R., Huang, S.Y. and Kidd, K.K. (2004) Possible epistatic role of ADH7 in the protection against alcoholism. *Am. J. Med. Genet. B Neuropsychiatry Genet.*, **126**, 19–22.
40. Birley, A.J., Whitfield, J.B., Neale, M.C., Duffy, D.L., Heath, A.C., Boomsma, D.I. and Martin, N.G. (2005) Genetic time-series analysis identifies a major QTL for *in vivo* alcohol metabolism not predicted by *in vitro* studies of structural protein polymorphism at the ADH1B or ADH1C loci. *Behav. Genet.*, **35**, 509–524.
41. Begleiter, H., Reich, T., Hesselbrock, V., Porjesz, B., Li, T.-K., Schuckit, M.A., Edenberg, H.J. and Rice, A.P. (1995) The Collaborative Study on the Genetics of Alcoholism. *Alcohol Health Res. World*, **19**, 228–236.
42. Reich, T., Edenberg, H.J., Goate, A., Williams, J.T., Rice, J.P., Van Eerdewegh, P., Foroud, T., Hesselbrock, V., Schuckit, M.A., Bucholz, K. *et al*. (1998) A genome-wide search for genes affecting the risk for alcohol dependence. *Am. J. Med. Genet. (Neuropsychiatry Genet.)*, **81**, 207–215.
43. Foroud, T., Edenberg, H.J., Goate, A., Rice, J., Flury, L., Koller, D.L., Bierut, L.J., Conneally, P.M., Nurnberger, J.I., Bucholz, K.K. *et al*. (2000) Alcoholism susceptibility loci: confirmation studies in a replicate sample and further mapping. *Alcohol. Clin. Exp. Res.*, **24**, 933–945.
44. Bucholz, K.K., Cadoret, R., Cloninger, C.R., Dinwiddie, S.H., Hesselbrock, V.M., Nurnberger, J.I.J., Reich, T., Schmidt, I. and Schuckit, M.A. (1994) A new semi-structured psychiatric interview for use in genetic linkage studies: A report of the reliability of the SSAGA. *J. Stud. Alcohol*, **55**, 149–158.

45. Hesselbrock, M., Easton, C., Bucholz, K.K., Schuckit, M. and Hesselbrock, V. (1999) A validity study of the SSAGA—a comparison with the SCAN. *Addiction*, **94**, 1361–1370.
46. World Health Organization (1993) *International Classification of Disease*, 10th edn. World Health Organization, Geneva.
47. Osier, M.V., Cheung, K.H., Kidd, J.R., Pakstis, A.J., Miller, P.L. and Kidd, K.K. (2002) ALFRED: an allele frequency database for anthropology. *Am. J. Phys. Anthropol.*, **119**, 77–83.
48. Buervenich, S., Sydow, O., Carmine, A., Zhang, Z., Anvret, M. and Olson, L. (2000) Alcohol dehydrogenase alleles in Parkinson's disease. *Mov. Disord.*, **15**, 813–818.
49. Li, M. and Edenberg, H.J. (1998) Function of *cis*-acting elements in human alcohol dehydrogenase 4 (ADH4) promoter, and role of C/EBP proteins in gene expression. *DNA Cell Biol.*, **17**, 387–397.
50. O'Connell, J.R. and Weeks, D.E. (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am. J. Hum. Genet.*, **63**, 259–266.
51. Boehnke, M. (1991) Allele frequency estimation from data on relatives. *Am. J. Hum. Genet.*, **48**, 22–25.
52. Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.
53. Reich, D.E., Cargill, M., Bolk, S., Ireland, J., Sabeti, P.C., Richter, D.J., Lavery, T., Kouyoumjian, R., Farhadian, S.F., Ward, R. *et al.* (2001) Linkage disequilibrium in the human genome. *Nature*, **411**, 199–204.
54. Dudbridge, F. (2003) Pedigree disequilibrium tests for multilocus haplotypes. *Genet. Epidemiol.*, **25**, 115–121.