Am. J. Hum. Genet. 64:1147–1157, 1999

Linkage Disequilibrium at the *ADH2* and *ADH3* Loci and Risk of Alcoholism

Michael Osier,[1] Andrew J. Pakstis,[1] Judith R. Kidd,[1] Jia-Fu Lee,[3,5] Shih-Jiun Yin,[4] Huei-Chen Ko,[6] Howard J. Edenberg,[7] Ru-Band Lu,[2] and Kenneth K. Kidd[1]

[1]Department of Human Genetics, Yale University, New Haven, CT; [2]Department of Psychiatry, Tri-Service General Hospital, [3]Institute of Medical Sciences, and [4]Department of Biochemistry, National Defense Medical Center, and [5]Military Psychiatric Center, Taipei, Taiwan, R.O.C.; [6]Graduate Institute of Behavioral Medicine, National Cheng Kung University Medical College, Tainan, Taiwan, R.O.C.; and [7]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis

Summary

Two of the three class I alcohol dehydrogenase (ADH) genes (*ADH2* and *ADH3*) encode known functional variants that act on alcohol with different efficiencies. Variants at both these genes have been implicated in alcoholism in some populations because allele frequencies differ between alcoholics and controls. Specifically, controls have higher frequencies of the variants with higher V_{max} (*ADH2*2* and *ADH3*1*). In samples both of alcoholics and of controls from three Taiwanese populations (Chinese, Ami, and Atayal) we found significant pairwise disequilibrium for all comparisons of the two functional polymorphisms and a third, presumably neutral, intronic polymorphism in *ADH2*. The class I ADH genes all lie within 80 kb on chromosome 4; thus, variants are not inherited independently, and haplotypes must be analyzed when evaluating the risk of alcoholism. In the Taiwanese Chinese we found that, only among those chromosomes containing the *ADH3*1* variant (high V_{max}), the proportions of chromosomes with *ADH2*1* (low V_{max}) and those with *ADH2*2* (high V_{max}) are significantly different between alcoholics and controls ($P < 10^{-5}$). The proportions of chromosomes with *ADH3*1* and those with *ADH3*2* are not significantly different between alcoholics and controls, on a constant *ADH2* background (with *ADH2*1*, $P = .83$; with *ADH2*2*, $P = .53$). Thus, the observed differences in the frequency of the functional polymorphism at *ADH3*, between alcoholics and controls, can be accounted for by the disequilibrium with *ADH2* in this population.

Received August 7, 1998; accepted for publication January 22, 1999; electronically published March 8, 1999.

Address for correspondence and reprints: Dr. Kenneth K. Kidd, Department of Human Genetics, SHM I-353, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520-8005. E-mail: kidd@biomed.med.yale.edu

© 1999 by The American Society of Human Genetics. All rights reserved.

0002-9297/99/6404-0027$02.00

Introduction

Of the seven known human genes coding for the alcohol dehydrogenases (ADHs), four code for enzymes catalyzing the initial step in the pathway for the metabolism of ethanol: *ADH1, ADH2, ADH3,* and *ADH4* (Edenberg and Bosron 1997). The class I ADHs (*ADH1, ADH2,* and *ADH3*) exist in a cluster on the long arm of chromosome 4. *ADH4* is in the same region but is at an unknown distance (Riess et al. 1994). The enzymes are very similar in sequence and structure but differ in preferred substrates (Edenberg and Bosron 1997). Two of the three class I genes are known to have alleles producing enzymes that catalyze the oxidation of ethanol at different rates. The two most common variants of *ADH2* are *ADH2*2,* which produces an enzyme with a high V_{max} (340 μM/min), and *ADH2*1,* which produces an enzyme with a lower V_{max} (9 μM/min) (Edenberg and Bosron 1997). In Asian populations, the *ADH2*2* allele is found at higher frequencies in controls than in alcoholics (Thomasson et al. 1991; Chen et al. 1996; Shen et al. 1997; Tanaka et al. 1997). These functionally different alleles are the result of a single amino acid substitution at residue 47 in exon 3 (arginine in *ADH2*1* and histidine in *ADH2*2*), which is the result of a single-nucleotide difference. A third variant, *ADH2*3,* is caused by a single-nucleotide difference in codon 369 (exon 9), which results in a substitution of cysteine for arginine (Burnell et al. 1987). *ADH2*3* has a V_{max} similar to that of *ADH2*2* (320 μM/min) but a higher Michaelis-Menten constant (K_m) for ethanol (36 mM for *ADH2*3;* 0.94 mM for *ADH2*2*). Likewise, the enzyme produced by *ADH3*1* has a higher V_{max} (88 μM/min) than that for the enzyme produced by *ADH3*2* (35 μM/min), which also is the result of a single-nucleotide difference that causes an amino acid substitution in exon 8 (isoleucine in *ADH3*1;* valine in *ADH3*2*) (Edenberg and Bosron 1997). The *ADH3*1* allele is generally found at higher frequencies in controls than in alcoholics (Chen et al. 1996).

The common ADH variants show distinct population

distributions. *ADH2*2* is present at frequencies >.59 in East Asian populations, with the exception of a frequency of .33 among Thais. It is found at frequencies <.25 in populations from other regions (Goedde et al. 1992; Neumark et al. 1998). The very high frequency of the *ADH2*2* allele in East Asian populations, relative to frequencies in other regions of the world, has long been a puzzle. Did it reach such high frequencies because of some past or present selection, or did it just drift to high frequency in an ancestral population? The question has potential health implications if there is some selective agent that has not been identified. *ADH2*3* is present primarily in African populations and has not been observed in Asian populations (Edenberg and Bosron 1997). *ADH3*2* is present in East Asia at frequencies of .12–.24 (Thomasson et al. 1991; Chen et al. 1996; Shen et al. 1997), whereas in a sample of mixed Europeans it was reported at a frequency of .39 (Whitfield et al., in press). The hypothesized medical implication of these functional variants is that individuals with the *ADH2*2* and/or *ADH3*1* allele(s) (the isozymes with a higher V_{max}) are more prone to adverse reactions from the consumption of ethanol, which reduces the chance of such individuals becoming alcoholic (Thomasson et al. 1994; Chen et al. 1996; Tanaka et al. 1997). An ADH with a higher V_{max} should increase the rate at which ethanol is converted into acetaldehyde, a toxin.

Because the *ADH2* and *ADH3* loci are only ∼15 kb apart (Yasunami et al. 1990) and because our previous studies of Asian populations had shown that linkage disequilibrium often extends over such a physical distance in these populations (Lu et al. 1996; Tishkoff et al. 1996, 1998; Kidd et al. 1998), we examined linkage disequilibrium at these loci in these populations. If strong linkage disequilibrium exists, additional factors could be involved in the phenotypic consequences beyond the differences in enzyme activity resulting from the individual amino acid substitutions. For example, a regulatory difference in strong disequilibrium with an amino acid difference could have an important role but may not have been identifiable in the studies performed to date. Alzheimer disease recently has provided an example of this phenomenon at the apolipoprotein E locus, where combinations of alleles in the regulatory region and in the coding region appear to increase disease risk (Bullido et al. 1998). Nickerson et al. (1998) discuss the difficulties of assigning causation to a single nucleotide when many nucleotide differences exist between alleles, as appears to be common over short molecular distances. Linkage disequilibrium has been reported between the two ADH functional variants in a sample of normal Chinese individuals (Chen et al. 1997), but haplotype frequencies have not been published. Moreover, functionally neutral polymorphisms, such as previously defined RFLPs (Smith 1986), could provide additional information; neutral polymorphisms in the ADH genes have not been previously studied as haplotypes in Asian populations, but some show disequilibrium in European populations (Edman and Maret 1992). Valdes and Thomson (1997) recently proposed a method for the analysis of different variants in haplotypes to determine which variants are likely to be causative of a disease and which are merely associated with a disease because of linkage disequilibrium.

In an initial attempt to study the population genetics and evolution of the ADH cluster of genes, we investigated as haplotypes the functional variants at *ADH2* and *ADH3,* as well as an intronic *Rsa*I RFLP (Smith 1986), in four populations: three Taiwanese populations and one Amerindian population. Samples of alcoholics and nonalcoholics were studied for the Taiwanese populations. The sample of Amerindians was a random sample of Maya undiagnosed for alcoholism and was used primarily for the analysis of disequilibrium. As background, we precisely mapped the *Rsa*I restriction site polymorphism at *ADH2,* converted it to PCR-based typing, and collected data for this single-nucleotide polymorphism (SNP) in several other populations. Through analysis of the haplotype data, we find that the observed differences in frequency of the functional polymorphism at *ADH3* in alcoholics and controls can be accounted for by the polymorphism's disequilibrium with *ADH2,* in the Taiwanese Chinese population.

Subjects, Material, and Methods

Patients and Controls

All patients were diagnosed, by clinical interview, as having severe alcohol dependence, according to the DSM-III-R (*Diagnostic and Statistical Manual of Mental Disorders,* 3d ed., revised) criteria (APA 1987). See the report by Lu et al. (1996) for more details regarding sample ascertainment, diagnostic procedures, and the measures of severity used for alcohol dependence. Normal controls were obtained from the same ethnic groups, including some medical students among the Taiwanese Chinese controls, and live in the same geographic areas as the patients. Men were predominantly included in this study. All samples were collected with informed consent. Some of these samples had been studied previously for *DRD2* haplotypes (Lu et al. 1996) and *DRD4* polymorphisms (Chang et al. 1997), and no associations with alcoholism were found.

Population Samples

The two Taiwanese aboriginal populations, the Ami and Atayal, were studied previously for the *DRD2* (Lu et al. 1996) and *DRD4* (Chang et al. 1997) loci and for a possible joint effect of *ADH2, ADH3,* and *DRD2* (Lee

et al., in press). The Atayal in this study were sampled from the eastern region of Taiwan. The samples of Taiwanese Chinese included those individuals studied along with the Ami and Atayal in the three studies mentioned above and an additional 107 alcoholics and 110 nonalcoholics included in the study by Lee et al. (in press). The sample of Taiwanese Chinese controls included some students. Since alcoholism generally is diagnosed later in life in Taiwan, some of these controls may be diagnosed in the future as alcoholics. Therefore, the only effect of an age-matching discrepancy should be a reduction of the statistical power of this study. All the Chinese are descendants of migrants from the South Fujien province ~400 years ago, and the typing at multiple loci of a subset of these individuals demonstrated that they tend to have a lower average heterozygosity than do the mainland Chinese, supporting their isolation and the absence of significant admixture (K. K. Kidd, unpublished data). At loci other than the ADH loci, we have not detected differences in allele frequencies between alcoholics and controls (e.g., see Lu et al. 1996; Chang et al. 1997).

The sample of Maya has been studied for many different genetic markers (see Castiglione et al. 1995; Chang et al. 1996; Tishkoff et al. 1996, 1998; Calafell et al. 1998) and is described in detail in the report by Castiglione et al. (1995). The individuals were normal, apparently healthy volunteers (76% female) living in the state of Campeche, Mexico; no diagnoses for alcoholism or related disorders were performed.

The other populations studied only for the *Rsa*I SNP are similar to the Maya in that they represent random population samples and no diagnostic information is available. The Biaka, Mbuti, Druze, Danish, Finnish, mixed European, San Francisco (SF) Chinese, Adygei, Yemenite Jewish, Rondonian Surui, and Karitiana samples are described in the report by Castiglione et al. (1995). The Japanese and Cambodian samples are described in the report by Mountain and Cavalli-Sforza (1997). The Nasioi sample is described in the report by Bowcock et al. (1991). The Cheyenne and Jemez Pueblo samples are described in the report by Goldman et al. (1993). The Ticuna sample is described in the report by Schurr et al. (1990). The Arizona Pima sample is described in the report by Price et al. (1992); a subset of this sample was studied here. The Ethiopian Jewish sample is described in the report by Zoossmann-Diskin et al. (1991). The specific samples of Russians, Taiwanese Hakka, Yakut, and Mexican Pima will be described in more detail at a later date. Additional information on all samples studied can be found at the Kidd Lab website (see Electronic-Database Information).

DNA from the samples of Taiwanese aborigines, the original Taiwanese Chinese (Lu et al. 1996), the Maya, and the other samples typed only for the *Rsa*I site polymorphism was extracted, by standard phenol/chloroform methods, from lymphoblastoid cell lines (Sambrook et al. 1986). DNA from the remaining Taiwanese Chinese was extracted from lymphocytes by use of the guanidine method (Anderson and Gusella 1984; Sambrook et al. 1986; Bowcock et al. 1987).

Mapping and Sequencing

The *Rsa*I site polymorphism originally was identified as an RFLP by Southern blotting of *Rsa*I-digested genomic DNA with clone pADH36 (Smith 1986). This clone was reported to include exon 3 of *ADH2* and some flanking intronic segments, for a total of 1.3 kb of insert (Smith 1986). The two alleles of the original RFLP are seen as bands of 1 kb and 500 bp (*ADH2*A1* and *ADH2*A2,* respectively; Willard et al. 1985), with constant bands of 1.2 kb, 0.62 kb, and 0.46 kb. Since the smaller allelic band, seen for the *ADH2*A2* allele, is only 500 bp, the polymorphic site must be either within the segment of DNA corresponding to the clone or ⩽500 bp from this segment. Therefore, we sequenced the clone, designed PCR primers for different sections of the region, and amplified the segments from individuals of known genotype for the *Rsa*I RFLP. Once the segment containing the site was identified by digestion of the PCR products and then by running them on an agarose gel, the PCR products from the two alternate homozygotes (identified in earlier RFLP studies) were sequenced to identify the nucleotide that differed. The polymorphic site was within the intron 3 portion of the cloned sequence. Note that *ADH2* is very similar in sequence to *ADH1* and to *ADH3,* with considerable similarity in the introns as well as near identity of the exons (Ikuta et al. 1986; Edenberg and Bosron 1997; M. Osier and K. K. Kidd, unpublished data). As such, care must be taken to ensure the specificity of the PCR reaction. The initial sequence from a gene-specific clone (pADH36) was confirmed by the sequencing of PCR products generated from genomic DNA by use of gene-specific primers.

Sequencing of the clone and the PCR product of genomic DNA was performed by cycle sequencing with an ABI PRISM Dye Terminator Cycle Sequencing Core Kit on an ABI 373S automatic DNA sequencer. For sequencing of the clone, vector primers were used to obtain initial sequences at each end of the insert. Primers then were designed to sequence further in each direction. After overlapping sequences were obtained, additional primers were designed to sequence the strand complementary to each end of the insert.

Marker Typing

Typing of the *Rsa*I SNP was performed with primers A2IN3DW1 (5′-TAA CTA GGA CAT TGC CAT ACC-

3′) and A2IN3UP2 (5′-GAG CTA AAA CAT ACT TTG GAT AG-3′). The PCR consisted of 4–8 ng genomic DNA template/μl, 100 ng each primer/μl, 200 μM dNTP, 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris HCl (pH 8.4), and 0.5 U AmpliTaq DNA polymerase (Perkin Elmer), in a total volume of 25 μl. The PCR was performed on a Perkin Elmer 9600 at 95°C for 5 min, followed by 30 cycles of 95°C for 40 s, 60°C for 30 s, and 72°C for 45 s and then by a single incubation at 72°C for 10 min. Primers used for the typing of the *Rsa*I site were sufficiently gene specific to result in a robust test under the given PCR conditions. Following PCR amplification, the unpurified product was digested with 5 U *Rsa*I restriction enzyme (Boehringer) at 37°C for 3 h. Digested product was electrophoresed on a 1.5% agarose gel and was stained with ethidium bromide, for visualization. The uncut fragment was 236 bp in length, and the cut fragments were 176 bp and 60 bp in length.

For the Maya, typing of the *ADH2* functional polymorphism in exon 3 and the *ADH3* functional polymorphism in exon 8 was performed as described in the report by Thomasson et al. (1994). For the samples of Taiwanese, the PCR methods of typing the *ADH2* functional polymorphism in exon 3 and the *ADH3* functional polymorphism in exon 8 were performed as described in the reports by Couzigou et al. (1990) and Walzer et al. (1993).

Statistical Analyses

Genotype and allele frequencies for each polymorphism were calculated by direct counting. The HAPLO computer program (Hawley and Kidd 1995) was used to estimate haplotype frequencies from the typing data for the three polymorphic sites. The program also used the jackknife procedure to estimate standard errors for the frequency estimates. Disequilibrium was calculated by use of the LINKD program (A. J. Pakstis, unpublished data), which calculates the standard disequilibrium coefficient D', as described elsewhere (Lewontin 1964). The haplotype method developed by Valdes and Thomson (1997) was used to evaluate the relative importance of the polymorphic sites in the determination of susceptibility to alcoholism.

Results

Mapping and Sequencing

The sequence for clone pADH36 is given in figure 1 (GenBank accession number AF040967). The insert is 1,318 bp in length, excluding the base pairs composing the *Eco*RI cloning sites. Interestingly, the *Eco*RI cloning site at the upstream end of this insert, although clearly present in the clone, was not seen in the sequence of genomic PCR product spanning the cloning site, whereas the downstream *Eco*RI cloning site was present in the

```
   1 gaattcatgc ttctctttat tctgtagatg gtggctgtag gaatctgtcg cacagatgac
  61 cacgtggtta gtggcaacct ggtgaccccc cttcctgtga ttttaggcca tgaggcagcc
 121 ggcatcgtgg agagtgttgg agaaggggtg actacagtca aaccaggtac aggattcaca
 181 ctcagggaac acgtgtggtt caccatccag gatttcccag cctggataag gaaaccaagg
 241 caatgagaga cgaaagggct tgcacaaggt caccgcgcag ccaggacttc aggagtttcc
 301 tctttccctc tctgcctgcc tgacccaagc atgtatgcat tcaggcactc atgtattgta
 361 cccttcaacc attatgtacc gagtatctac tagaagccag gcaccaggtt agagcctgga
 421 gatatgtgag gatcaaaaga gacacagtgc ctaccagtgt ggcattcata gtccagtaat
 481 tctgtgttag tgttggctga acactgtctt ctgtcagggc ccagatagaa gtagaaacgg
 541 gaggagggaa aaaaagtaaa tgtggaaact gttctgtgct gagttcacaa aatggaggct
 601 acatggctac ggatgaataa acatgacctt ttcttaggtg acctatgcag ctccttcacc
 661 aatttcctct tccaacatgg tgtctcttcc tgagcatgga atcaatgaaa caggctaggt
 721 gaaaggaaat gaaagaaacg tagaaggaac catcccacca tcaacaatgt ctgccttgtt
 781 accttcagcc atgcttaaat gcaactggtt acagattact aattatttca tcttcaatat
 841 ttattttacc ctaaacttat gtattaatgt aaatactttt agaagtaaag ggttaataac
 901 taggacattg ccataccatc attttcacat aaactttacc tataaaatgt taatgCacca
 961 acaattttc ttgatcatgt ggaatgtata gcataggtga ttagaaataa aaattatgtt
1021 taattttaaa aagttgcttt agcctgtaga ttcaccctgc cccttacca aataccaatt
1081 taaactgtga acagaaaata ataaaataaa ctatccaaag tatgttttag ctctattttc
1141 tttagttaga aataaattaa caacaaaatg acagccttac tttttacccc tactccctcc
1201 accataacag gaaaagtaga taattcgcaa attatcagct aaaaatattt gttaagtcca
1261 ttcttatcct aaaagaaagt gcctaacttg ttgtgaagag ctgaagtttt cagagaaatt
1321 tctggaattc
```

Figure 1 Sequence of pADH36. The location of exon 3 is underlined, and the polymorphic nucleotide for the *Rsa*I SNP is capitalized and double underlined (nucleotide 956; C is site absent, and T is site present). The position of the functional exon 3 SNP is nucleotide 50.

sequence of genomic PCR product spanning the cloning site (data not shown). The insert starts 21 bp upstream of exon 3 and continues 1,157 bp from the downstream end of exon 3, again excluding the cloning site. The sequence of exon 3 does not show any deviation from published sequences (Ikuta et al. 1986; multiple other references are summarized in Edenberg 1991). The nucleotide responsible for the *Rsa*I SNP is 906 bp downstream from the varying nucleotide in the functional exon 3 SNP.

Global *Rsa*I Frequencies

Frequencies of the *Rsa*I SNP for various populations are given in table 1. In general, global frequencies of the *ADH2*A1* allele are >.70, except in Southeast Asian populations. In most Southeast Asian populations (Taiwanese Chinese, SF Chinese, Hakka, Japanese, Ami, and Atayal) the nearly opposite appears to be true, with *ADH2*A1* frequencies within the range .16–.35 (table 1), which is roughly the same as the frequencies of the *ADH2*A2* allele in African, European, Siberian (Yakut), and Melanesian (Nasioi) populations. The Cambodians were the only population in our study with an "intermediate" allele frequency for the *ADH2*A1* allele (.40; table 1). In all three samples of South American Amerindian populations, *ADH2*A2* alleles were not observed. Although the sample sizes were too small to determine whether *ADH2*A2* alleles were not present, the true frequency is certainly quite low (upper 95% confidence limit: Surui, .049; Karitiana, .029; and Ticuna, .022).

Haplotype Frequencies and Disequilibrium

The observed numbers of individuals with different three-site phenotypes are given in table 2. The data on

Table 1

Allele Frequencies for the *ADH2 Rsa*I SNP Polymorphism

Population	Frequency of *ADH2*A1* Allele	Standard Error	No. of *ADH2*A1* Alleles	Total No. of Chromosomes
Biaka	.93	.02	129	138
Mbuti	.97	.02	76	78
Ethiopians	.93	.05	28	30
Yemenite Jews	.91	.03	71	78
Adygei	.72	.04	75	104
Druze	.74	.04	118	160
Danes	.71	.05	61	86
Finns	.81	.05	52	64
Russians	.74	.05	70	94
Mixed Europeans	.70	.04	105	150
SF Chinese	.19	.04	19	102
Taiwanese Chinese[a]	.19	.03	14	270
Hakka	.20	.04	16	82
Ami[a]	.33	.07	28	40
Atayal[a]	.21	.06	15	42
Japanese	.16	.04	15	96
Cambodians	.40	.07	19	48
Nasioi	.70	.07	32	46
Yakut	.76	.04	74	98
Cheyenne	.98	.01	110	112
Jemez Pueblo	.92	.03	83	90
Pima, Arizona	1.00	0	100	100
Pima, Mexico	.99	.01	178	180
Maya	.99	.01	99	100
Ticuna	1.00	0	132	132
Rondonian Surui	1.00	0	60	60
Karitiana	1.00	0	106	106

[a] Frequencies in the samples of nonalcoholic controls. See tables 2 and 3.

the two functional polymorphisms are complete. There are several instances of missing *Rsa*I data, but this site was typed for 87.5% of the sample of alcoholics and for 94.1% of the sample of controls. Genotypes were counted directly (homozygotes and single-site heterozygotes) for 87 (68.0%) of 128 alcoholics and for 84 (62.2%) of 135 controls. Only 6 (5.7%) of the 128 alcoholics were triple heterozygotes or possible triple heterozygotes because of missing *Rsa*I data; only 11 (8.2%) of the 135 controls had such genotypes. Estimated haplotype frequencies and the nomenclature used for haplotypes are given in table 3.

Undue consideration should not be given to the large numbers of haplotypes with a frequency of 0. Owing to the somewhat small sample sizes of the Ami and Atayal, haplotypes at low frequencies may not have been readily detectable. The 1-2-A2 haplotype (*ADH3*1-ADH2*2-ADH2*A2*) was the most common in all three Asian populations. The allele frequencies among the Atayal controls in this study were significantly different from those reported by Thomasson et al. (1994) for an independent sample of Atayal. This may be the result of sampling from different geographic regions for the Atayal, since the smaller sample size in this study does not explain the differences. Our sample was selected from the Taroko Gorge–Hualien area in the eastern region of Taiwan, which belongs to one of the subtribes of the Atayal population. The sample studied by Thomasson et al. (1994) was from the Taipei area in the northern region of Taiwan.

The pairwise disequilibrium values (D') and significance values for all polymorphic sites in this study are listed in table 4. When calculation was possible, pairwise disequilibrium was significant ($P < .005$) between all pairs of sites for all populations, except the Maya. For the Maya there may not have been sufficient discriminating power to observe significant disequilibrium, because frequencies for two of the loci (*ADH3* and the *Rsa*I site) were nearly fixed. As was expected, disequilibrium between the functional *ADH2* polymorphism and the *Rsa*I SNP (the two closest loci) was greater than that for the other two pairwise comparisons. There was no significant difference in linkage disequilibrium between alcoholics and controls for any pair of loci in any of the three populations diagnosed.

Predisposition to and Protection against Alcoholism

The Taiwanese Chinese alcoholics have allele frequencies significantly different from those of the Taiwanese Chinese controls, for all three polymorphic sites (table 5). This finding agrees with those of previous stud-

Table 2

Joint Distribution at Polymorphic Sites, for *ADH2, Rsa*I, and *ADH3*

			No. of Taiwanese Chinese		No. of Ami		No. of Atayal		No. of Maya
ADH2	*Rsa*I	*ADH3*	Alcoholics	Controls	Alcoholics	Controls	Alcoholics	Controls	
22	A2A2	11	35	62	6	11	14	13	0
22	A2A2	12	2	2	0	0	0	0	0
22	A1A2	11	3	0	4	1	1	0	0
22	A1A1	11	1	0	1	2	0	0	0
22	?	11	7	8	0	0	0	0	0
22	?	12	1	1	0	0	0	0	0
12	A2A2	11	13	12	2	0	1	1	0
12	A2A2	12	2	3	0	0	0	0	0
12	A1A2	11	9	20	4	4	5	5	0
12	A1A2	12	5	9	0	0	0	0	0
12	A1A2	22	0	1	0	0	0	0	0
12	A1A1	11	1	3	0	1	0	1	5
12	A1A1	12	0	2	0	0	0	0	1
12	?	11	7	4	0	0	0	0	0
12	?	12	1	2	0	0	0	0	0
12	?	22	1	0	0	0	0	0	0
11	A2A2	11	4	1	0	0	0	0	0
11	A2A2	12	1	0	0	0	0	0	0
11	A1A2	11	8	2	1	0	0	0	1
11	A1A2	12	2	0	0	0	0	0	0
11	A1A2	22	1	0	0	0	0	0	0
11	A1A1	11	5	0	2	1	0	0	28
11	A1A1	12	9	1	0	0	0	0	10
11	A1A1	22	4	1	0	0	0	0	1
11	?	11	1	1	0	0	0	0	0
11	?	12	5	0	0	0	0	0	0
13	A1A1	11	0	0	0	0	0	0	1
Total (*N*)			128	135	20	20	21	20	47

ies of *ADH3* and *ADH2* alleles that have demonstrated an association between alcoholism and the functional polymorphisms and is expected for the *Rsa*I SNP, since tight disequilibrium exists among all three polymorphisms. The existence of disequilibrium means that the frequency differences of individual sites are not independent and that causation cannot be inferred from statistical significance when the sites are considered singly.

Valdes and Thomson (1997) showed that neutral sites in disequilibrium with a disease-predisposing site should have the same relative frequencies in samples of affecteds and controls when the state of the predisposing site is held constant. Table 6 gives the haplotype-frequency ratios for the Taiwanese Chinese, for the *ADH3* and *ADH2* sites. Statistical significance is difficult to address directly, since we only have estimates of haplotype frequencies, obtained by use of the expectation-maximization (EM) algorithm. Instead, we used the contingency-table approach—described by Valdes and Thomson (1997) for the two-site model of one susceptibility site and one neutral site—as an approximation, under the assumption that the counts are directly calculable from the estimates of haplotype frequency.

In our analyses, we considered only the expressed sites at *ADH2* and *ADH3*. The intronic *Rsa*I site was omitted as almost certainly neutral on the basis of prior considerations and comparable contingency-table analyses (data not shown) indicating no effect on alcoholism. Any error in the allocation of chromosomes by the EM algorithm was minimal and did not affect the contingency analysis, since >93% of all haplotypes in alcoholics and >88% of all haplotypes in controls were counted directly (which is discussed below in more detail).

Only one of the four ratios is significant, namely, the ratio comparing *ADH2* alleles on the *ADH3*1* background. Thus, we can reject the hypothesis that the *ADH2* alleles are neutral when we consider the *ADH3*1* background. However, we cannot reject the hypothesis that *ADH2* alleles are neutral when we consider the *ADH3*2* background, because that contingency table did not yield a significant χ^2 value. This finding likely was the result of reduced power, owing to small numbers (single individuals) in the contingency table, since the difference between alcoholics and controls was nearly twofold, which is similar to that seen on the *ADH3*1* background. In contrast, both ratios for *ADH3* alleles, when either *ADH2* background was considered, are nonsignificant; thus, there is no evidence that the *ADH3*

Table 3

Estimates of Haplotype Frequency

Population and Group	$2N$	Frequency (Jackknife Standard Error) of Haplotype[a]							
		1-1-A1	2-1-A1	1-1-A2	2-1-A2	1-2-A1	2-2-A1	1-2-A2	2-2-A2
Taiwanese Chinese:									
Alcoholics	256	.174 (.026)	.119 (.024)	.153 (.025)	.019 (.010)	.031 (.014)	.0 (.0)	.485 (.037)	.019 (.009)
Controls	270	.110 (.019)	.060 (.017)	.071 (.017)	.011 (.007)	.023 (.012)	.0 (.0)	.707 (.030)	.018 (.010)
Ami:									
Alcoholics	40	.214 (.081)	.0 (.0)	.086 (.047)	.0 (.0)	.161 (.069)	.0 (.0)	.539 (.083)	.0 (.0)
Controls	40	.175 (.066)	.0 (.0)	.0 (.0)	.0 (.0)	.150 (.073)	.0 (.0)	.675 (.091)	.0 (.0)
Atayal:									
Alcoholics	42	.118 (.048)	.0 (.0)	.025 (.025)	.0 (.0)	.025 (.025)	.0 (.0)	.833 (.053)	.0 (.0)
Controls	40	.149 (.052)	.0 (.0)	.026 (.026)	.0 (.0)	.026 (.026)	.0 (.0)	.799 (.068)	.0 (.0)
Maya:									
Random	94	.788 (.045)	.063 (.027)	.011 (.011)	.0 (.0)	.138 (.038)	.0 (.0)	.0 (.0)	.0 (.0)

[a] Eight haplotypes can be produced by combinations of the three biallelic markers; of these, seven are estimated to be present in at least one population. The haplotypes are labeled as the sites occur, in order from 5′ to 3′, in the ADH cluster: *ADH3,* functional alleles 1 or 2 in exon 8; *ADH2,* functional alleles 1 or 2 in exon 3; and *ADH2,* alleles A1 (site absent) or A2 (site present) at the *Rsa*I SNP in intron 3. For ease of reading, the sites are separated by hyphens. Thus, 1-2-A1 indicates allele 1 at *ADH3,* allele 2 at *ADH2,* and the site-absent allele at the *Rsa*I site in *ADH2.* The jackknife standard errors usually are somewhat larger (i.e., ⩽30% larger) than the binomial standard errors calculated from the sample size and the haplotype-frequency estimate.

polymorphism has any effect on susceptibility to alcoholism. From this analysis we can conclude that *ADH2*1* does confer a different level of susceptibility than *ADH2*2* and that *ADH2*1* may be the only significant variant.

Discussion

In the unusual condition in which genes coding for very similar activity are molecularly close to each other, examination of the consequences of the linkage between the genes becomes important. One consequence may be linkage disequilibrium, the nonrandom occurrence of sets of alleles on the chromosomes in the population. Linkage disequilibrium of a disease gene with a strong effect to a disease gene with a slightly weaker effect may distort statistical calculations of individual gene effects by violation of the assumption of independence. The only way to determine the individual effects is by examination of the context in which each gene appears. In other words, haplotypes must be analyzed in order to gain an appropriate understanding of the effect of an individual gene.

In East Asia, the *ADH2*A1* allele at the *Rsa*I site occurs at frequencies <.40; in the rest of the world, frequencies are >.70. This highly significant difference is similar to that seen for the *ADH2*1* allele (Goedde et al. 1992). Instead of merely reflecting the *ADH2* genotypings, however, the *Rsa*I site provides more information by subdividing some of the *ADH2-ADH3* haplotypes. This is most obvious for the frequencies of the 1-1-A1 and 1-1-A2 haplotypes. The *Rsa*I site also provides a theoretically neutral polymorphism to contrast with the functional variants in the haplotype analysis.

As demonstrated by the disequilibrium values, variants at the two ADH loci do not occur independently in these populations. One particular haplotype (1-2-A2) of the eight possible (including the presumably neutral *Rsa*I SNP) showed the greatest difference between alcoholics and nonalcoholics. This haplotype occurred at a frequency of nearly 71% in the control sample of Taiwanese Chinese but at only ~49% in the alcoholics, a highly significant difference ($P < .0001$, two-proportion z-test). There was no comparable reduction for the 1-2-A1 haplotype, which occurred at a frequency of 2.3% in controls and 3.1% in alcoholics, a nonsignificant difference in the opposite direction. This comparison of the 1-2-A1 haplotype has almost no power to detect a dif-

Table 4

Pairwise Disequilibrium for Loci

Population and Group	D' for Paired Loci[a]		
	*Rsa*I, *ADH2*	*ADH3, ADH2*	*ADH3, Rsa*I
Taiwanese Chinese:			
Alcoholics	.821	−.774	−.642
Controls	.841	−.730	−.596
Ami:			
Alcoholics	.541	NP	NP
Controls	1.000	NP	NP
Atayal:			
Alcoholics	.796	NP	NP
Controls	.820	NP	NP
Maya:			
Random	−1.000	−1.000	−1.000

[a] *ADH2* and *ADH3* refer to the functional polymorphisms in these genes, and *Rsa*I refers to the SNP in intron 3 of *ADH2*. The D' values given are for significant pairwise disequilibrium ($P[D' = 0] < .005$, except for the Maya [see Results]). NP = calculation not possible (because one locus was fixed).

Table 5

Ratio of Allele Frequencies among Alcoholics to Frequencies among Controls

Population	Ratio (*P* Value) of Allele Frequencies[a]		
	*ADH3*1*	*ADH2*2*	*ADH2*A2*
Taiwanese Chinese	.925 (.0087)	.715 (<.0001)	.838 (.0003)
Ami	1.000 (NA)	.848 (.0951)	.926 (.3192)
Atayal	1.000 (NA)	1.040 (.3409)	1.040 (.3409)

[a] Significance measured by a two-proportion *z*-test of the null hypothesis that the frequency in both samples is the same. NA = not applicable.

ference of the magnitude seen for the 1-2-A2 haplotype, because of the rarity of 1-2-A1. However, three other haplotypes demonstrated a significant difference between alcoholics and controls, in the opposite direction of haplotype 1-2-A2 (1-1-A1, $P = .018$; 2-1-A1, $P = .008$; and 1-1-A2, $P = .001$ [all *P* values determined by a two-proportion *z*-test]). These estimates are likely to be conservative because students were included among the controls. Alcoholism usually is diagnosed later in life in Taiwan; therefore, the power to determine differences should be somewhat reduced, since some younger controls may reveal a susceptibility later in life. A sample of controls matched more closely in age to the sample of alcoholics may show an even greater difference in haplotype frequencies. Haplotype comparisons between alcoholics and controls had low power within the Ami and Atayal populations, owing to the small sample sizes, and are not considered further in this article.

Although not challenging the clear existence of the functional variation observed at *ADH2* and *ADH3* (Thomasson et al. 1991; Chao et al. 1994; Chen et al. 1996), the nonrandom association in the Taiwanese Chinese leaves open the question of exactly what is responsible for the different susceptibilities to alcoholism. One possibility is that there may be other functionally relevant variables on those chromosomes in linkage disequilibrium. For example, in addition to the higher metabolic rate for ethanol for *ADH2*2,* there may be regulatory differences for *ADH2* as well, and the allele with the more active transcription may exist only in combination with one haplotype or group of haplotypes with a nonrandom representation of ability to metabolize ethanol. This possibility cannot now be distinguished from the effects of the amino acid substitution on the risk of alcoholism.

The demonstration of strong disequilibrium between the functional variants in *ADH3* and *ADH2* raises the question of whether the two *ADH3* variants, which have a smaller difference in V_{max}, have an independent effect on the risk of alcoholism in the Taiwanese Chinese. The overall frequency of *ADH3*1* was .84 in alcoholics and .91 in controls, but the reduced frequency in alcoholics was entirely attributable to the 1-2-A2 haplotype: the other six observed haplotypes were more common in alcoholics than in controls. The frequencies of three of the haplotypes with *ADH2*1* were significantly different between alcoholics and controls, in the opposite direction of haplotype 1–2–A2, as indicated in the discussion above (higher frequency in alcoholics). The other three observed haplotypes have nonsignificant differences between alcoholics and controls.

The haplotype analysis of Valdes and Thomson (1997) does not reject the null hypothesis that *ADH3* alleles are neutral; *ADH2* allelic variation is sufficient to explain the different levels of susceptibility to alcoholism. Although comparisons of *ADH2* frequencies on an *ADH3*2* background and of *ADH3* frequencies on an *ADH2*2* background have reduced power due to small numbers of individuals in some cells, there is no such problem with the comparison of *ADH2* frequencies on an *ADH3*1* background and *ADH3* frequencies on an *ADH2*1* background. All comparisons were statistically consistent with no effect of *ADH3* alleles. There was, however, a hint of a very weak effect of *ADH3*. When only those haplotypes with the *ADH2*1* allele were considered, the *ADH3*1* allele occurred in alcoholics 0.979 times as frequently as in controls, and the *ADH3*2* allele occurred in alcoholics 1.05 times as frequently as in controls. Thus, numerically, *ADH3*1* showed a small decrease in alcoholics, in accordance with *ADH3*1* having a small protective effect independent of *ADH2*2,* but the effect, if real, is very small and far from being statistically significant.

The question arises of whether any of these analyses of alcoholics owe their significance to use of the EM algorithm to estimate haplotype frequencies, since Hardy-Weinberg (H-W) expectations clearly were violated in the sample of alcoholics. Although an additional level of uncertainty is added, it actually is quite small for the analyses in table 5, as can be seen from the "raw" data in table 2 and from the following discussion. Using extreme ranges of the possible numbers does not markedly alter the significance levels and leaves the conclusions unaltered.

Except for the multiple heterozygotes, the EM algorithm uses gene counting, which is not affected by departures from H-W ratios. When only the functional sites are considered, almost 12% of controls are double heterozygotes, but there is no deviation from H-W ratios; therefore, the use of the EM algorithm to allocate these counts is appropriate. Only 6.25% of the alcoholics are double heterozygotes, and there is a clear departure from H-W ratios. There are two components to H-W ratios, the relative proportions of homozygotes and heterozygotes and the relative proportions of individual genotypes within each category. Use of H-W ratios within the EM algorithm involves estimation of only the

Table 6

Relative Haplotype Frequencies among Alcoholics and Controls, for the Taiwanese Chinese

Haplotype Ratio	Alcoholics	Controls	*P* Value[a]
$\frac{f(ADH3^*1, ADH2^*1)}{f(ADH3^*1, ADH2^*2)}$	$\frac{83.7/256}{132.1/256} = \frac{.327}{.516} = .64$	$\frac{48.9/270}{197.1/270} = \frac{.181}{.730} = .25$	7×10^{-6}
$\frac{f(ADH3^*2, ADH2^*1)}{f(ADH3^*2, ADH2^*2)}$	$\frac{35.3/256}{4.9/256} = \frac{.138}{.019} = 7.00$	$\frac{19.2/270}{4.9/270} = \frac{.071}{.018} = 3.80$	.37
$\frac{f(ADH3^*1, ADH2^*1)}{f(ADH3^*2, ADH2^*1)}$	$\frac{83.7/256}{35.3/256} = \frac{.327}{.138} = 2.40$	$\frac{48.9/270}{19.2/270} = \frac{.181}{.071} = 2.55$	.83
$\frac{f(ADH3^*1, ADH2^*2)}{f(ADH3^*2, ADH2^*2)}$	$\frac{132.1/256}{4.9/256} = \frac{.516}{.019} = 27.16$	$\frac{197.1/270}{4.9/270} = \frac{.730}{.018} = 40.56$	.53

NOTE.—Frequencies listed as estimated number of chromosomes/total number of chromosomes (e.g., 84 estimated chromosomes with the haplotype from alcoholics/256 total chromosomes from alcoholics).

[a] Significance measured by use of the 2 × 2 contingency table of Valdes and Thomson (1997). The method tests the null hypothesis that, in the subsample with a specific allele at one site, the ratio of the number of chromosomes with the two alleles at the second site is the same for alcoholics and controls.

relative proportions of different heterozygous genotypes, except when data are missing. When only the two-site data for the two expressed polymorphisms at *ADH3* and *ADH2* are considered, the EM allocates counts for the double heterozygotes in the proportion $p_ip_j{:}p_kp_l$, where *i* and *j* index the two *cis* haplotypes and *k* and *l* index the two *trans* haplotypes; the frequencies of these four haplotypes are largely based on the direct counts from homozygotes and single-site heterozygotes. Unless there is a *cis-trans* effect or a functional variant affecting the two genotypes differently (such as disequilibrium of one haplotype with an allele at a third site), this is a simple sampling model and is not affected by population levels of inbreeding. This model is affected if one of the homozygous classes is differentially included in the sample, and, therefore, the underlying frequency of that haplotype in the sample is different from what it was in the population of zygotes. Certain ADH genotypes clearly are less likely to be found among alcoholics, as indicated by the data in table 2. However, an improved estimate will be possible only if the correct model of disequilibrium and/or selection is incorporated into the EM algorithm. Since there is no suitable model to use at this time, the EM algorithm incorporating H-W ratios provides a reasonable method for estimation of haplotype frequencies. For alcoholics, the expected ratio $p_ip_j{:}p_kp_l$ based on all genotypes except those of the double heterozygotes was 1:10.26 for the *cis:trans* genotypes among the double heterozygotes ([*ADH3*1, ADH2*1/ADH3*2, ADH2*2*]:[*ADH3*1, ADH2*2/ADH3*2, ADH2*1*]). For the controls, the same ratio was 1:3.79. We considered the numeric extremes for the allocation of the double heterozygotes: the results for *ADH2* on an *ADH3*1* background remained highly significant, and the conclusions were not altered (analysis not shown).

On the basis of the contingency-table analyses, the *Rsa*I SNP appears to be functionally neutral (analyses not shown). Although the frequencies of the *Rsa*I SNP on an *ADH3*1* background yielded nominally significant results ($P = .016$), without a Bonferroni correction, the analyses of the frequencies at the *Rsa*I SNP on either *ADH2* background were statistically insignificant. Therefore, the observed significance of the *Rsa*I SNP on an *ADH3*1* background is likely due to disequilibrium with the *ADH2* exon 3 SNP.

The three Taiwanese populations were not tested for the *ADH2*3* allele, which is determined by a nucleotide difference in exon 9 that changes an arginine to a cysteine in position 369 of the protein (Burnell et al. 1987). This variant has not been found previously in East Asian populations. The samples of Mayan DNA were tested for this variant, and one individual was heterozygous. In fact, since the *ADH2*3* allele is determined by a separate site, it can be ignored in an analysis of the site determining the *ADH2*2* allele. In this case, the rest of the haplotype was 1-1-A1, and it was counted as such.

The existence of the *ADH2*3* allele in the sample of Maya possibly reflects a small amount of African admixture in the post-Colombian period. Such admixture at <5% in this specific sample of Maya has been suggested by findings at other loci (K. K. Kidd, unpublished data). However, the existence of the *ADH2*3* allele at a low frequency in a heterogeneous sample of Amerindians has been documented by Wall et al. (1997). Although none of the individuals was of purely Amerindian descent, none was of documented African descent. Thus, a low frequency of this allele possibly is an inherent characteristic of Amerindians and is not the result of recent African admixture. In either case, the finding of one occurrence of the *ADH2*3* allele in the sample of Maya is not especially unusual.

When haplotypes rather than expressed variants are considered, nomenclature for alleles becomes more complex. As noted above, the *ADH2*2* and *ADH2*3* alleles are not really variants at the same position. As such, it is not logical to use the haplotype nomenclature that we used (table 3) and simply substitute allele 3 for allele 2 at *ADH2:* the sites are on opposite sides of the *Rsa*I site in intron 3, and the double-mutant haplotype is possible,

although it may not actually exist. One logical nomenclature would be to use the amino acid differences to label the functional alleles and to consider what are now called *"ADH2*1," "ADH2*2,"* and *"ADH2*3"* as the haplotypes representing three of the four possible combinations of the particular amino acids in the protein. This system would be more consistent with the recommendations of Antonarakis and the Nomenclature Working Group (1998).

Despite the complications of nomenclature and the unknown metabolic effects of multiple functional polymorphisms in each gene, haplotype analysis is more informative than analysis of the polymorphisms individually, especially since the three class I ADH genes (*ADH1, ADH2,* and *ADH3*) are within a region spanning only 80 kb (Yasunami et al. 1990). Because linkage disequilibrium can extend over dozens of kilobases, as shown by our haplotype analyses, differences in allele frequencies at one site, between alcoholics and controls, can be explained fully by linkage disequilibrium with another nearby site. Haplotype information requires more attention than it has been given previously.

Acknowledgments

We would like to thank William C. Speed, for sequencing data. This study was supported, in part, by the Department of Health, Republic of China (grant DOH 86-HR-612 to R.-B.L. and H.-C.K.); by the National Science Council, Republic of China (grant NSC 87-2314-B016-024 to R.-B.L. and grant NSC 87-2316-B016-006Y to S.-J.Y.); by U.S. Public Health Service grants AA09379-05, MH30929, and MH39239 (to K.K.K.); and by U.S. National Science Foundation grant SBR9632509 (to J.R.K.).

Electronic-Database Information

Accession number and URLs for data in this article are as follows:

GenBank, http://www.ncbi.nlm.nih.gov/Web/Genbank (for pADH36 sequence [AF040967])
Kidd Lab, Department of Human Genetics, Yale University, http://info.med.yale.edu/genetics/kkidd (for supplemental population information)

References

American Psychiatric Association (APA) (1987) Diagnostic and statistical manual of mental disorders, 3d ed, rev. APA, Washington DC

Anderson MA, Gusella JF (1984) Use of cyclosporin A in establishing Epstein-Barr virus–transformed human lymphoblastoid cell lines. In Vitro 20:856–858

Antonarakis SE, Nomenclature Working Group (1998) Recommendations for a nomenclature system for human gene mutations. Hum Mutat 11:1–3

Bowcock AM, Bucci C, Hebert JM, Kidd JR, Kidd KK, Friedlaender JS, Cavalli-Sforza LL (1987) Study of 47 DNA markers in five populations from four continents. Gene Geogr 1: 47–64

Bowcock AM, Kidd JR, Mountain JL, Hebert JM, Carotenuto L, Kidd KK, Cavalli-Sforza LL (1991) Drift, admixture, and selection in human evolution: a study with DNA polymorphisms. Proc Natl Acad Sci USA 88:839–843

Bullido MJ, Artiga MJ, Recuero M, Sastre I, Garcia MA, Aldudo J, Lendon C, et al (1998) A polymorphism in the regulatory region of *APOE* associated with risk for Alzheimer's dementia. Nat Genet 18:69–71

Burnell JC, Carr LG, Dwulet FE, Edenberg HJ, Li TK, Bosron WF (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

Calafell F, Shuster A, Speed WC, Kidd JK, Kidd KK (1998) Short tandem repeat polymorphism evolution in humans. Eur J Hum Genet 6:38–49

Castiglione CM, Deinard AS, Speed WC, Sirugo G, Rosenbaum HC, Zhang Y, Grandy DK, et al (1995) Evolution of haplotypes at the DRD2 locus. Am J Hum Genet 57: 1445–1456

Chang FM, Kidd JR, Livak KJ, Pakstis AJ, Kidd KK (1996) The world-wide distribution of allele frequencies at the human dopamine D4 receptor locus. Hum Genet 98:91–101

Chang FM, Ko H-C, Lu R-B, Pakstis AJ, Kidd KK (1997) The dopamine D4 receptor gene (*DRD4*) is not associated with alcoholism in three Taiwanese populations: six polymorphisms tested separately as haplotypes. Biol Psychiatry 41: 394–405

Chao Y-C, Liou S-R, Chung Y-Y, Tang H-S, Hsu C-T, Li TK, Yin S-J (1994) Polymorphism of alcohol and aldehyde dehydrogenase genes and alcoholic cirrhosis in Chinese patients. Hepatology 19:360–366

Chen WJ, Loh EW, Hsu YP, Chen CC, Yu JM, Cheng AT (1996) Alcohol-metabolizing genes and alcoholism among Taiwanese Han men: independent effect of *ADH2, ADH3* and *ALDH2*. Br J Psychiatry 168:762–767

Chen WJ, Loh EW, Hsu YP, Cheng AT (1997) Alcohol dehydrogenase and aldehyde dehydrogenase genotypes and alcoholism among Taiwanese aborigines. Biol Psychiatry 41: 703–709

Couzigou P, Fleury B, Groppi A, Cassaigne A, Begueret J, Iron A (1990) Genotyping study of alcohol dehydrogenase class I polymorphism in French patients with alcoholic cirrhosis: The French Group for Research on Alcohol and Liver. Alcohol Alcohol 25:623–626

Edenberg HJ (1991) Molecular biological approaches to studies of alcohol-metabolizing enzymes. In: Crabbe JC Jr, Harris RA (eds) The genetic basis of alcohol and drug actions. Plenum Press, New York, pp 165–223

Edenberg HJ, Bosron WF (1997) Alcohol dehydrogenases. In: Guengerich FP (ed) Biotransformation. Vol 3 in: Sipes IG, McQueen CA, Gandolfi AJ (eds) Comprehensive toxicology. Pergamon, New York, pp 119–131

Edman K, Maret W (1992) Alcohol dehydrogenase genes: restriction fragment length polymorphisms for *ADH4* (π-

ADH) and *ADH5* (χ-ADH) and construction of haplotypes among different ADH classes. Hum Genet 90:395–401

Goedde HW, Agarwal DP, Fritze G, Meier-Tackmann D, Singh S, Beckmann G, Bhatia K, et al (1992) Distribution of *ADH2* and *ALDH2* genotypes in different populations. Hum Genet 88:344–346

Goldman D, Brown GL, Albaugh B, Robin R, Goodson S, Trunzo M, Akhtar L, et al (1993) *DRD2* dopamine receptor genotype, linkage disequilibrium, and alcoholism in American Indians and other populations. Alcohol Clin Exp Res 17:199–204

Hawley ME, Kidd KK (1995) HAPLO: a program using the EM algorithm to estimate the frequencies of multi-site haplotypes. J Hered 86:409–411

Ikuta T, Szeto S, Yoshida A (1986) Three human alcohol dehydrogenase subunits: cDNA structure and molecular and evolutionary divergence. Proc Natl Acad Sci USA 83: 634–638

Kidd KK, Morar B, Castiglione CM, Zhao H, Pakstis AJ, Speed WC, Bonne-Tamir B, et al (1998) A global survey of haplotype frequencies and linkage disequilibrium at the DRD2 locus. Hum Genet 103:211–227

Lee J-F, Lu R-B, Ko H-C, Chang F-M, Yin S-J, Pakstis AJ, Kidd KK. No association between *DRD2* locus and alcoholism after controlling the *ADH* and *ALDH* genotypes in a Chinese Han population. Alcohol Clin Exp Res (in press)

Lewontin RC (1964) The interaction of selection and linkage. I. General considerations: heterotic models. Genetics 49: 49–67

Lu R-B, Ko H-C, Chang F-M, Castiglione CM, Schoolfield G, Pakstis AJ, Kidd JR, et al (1996) No association between alcoholism and multiple polymorphisms at the dopamine D2 receptor gene (*DRD2*) in three distinct Taiwanese populations. Biol Psychiatry 39:419–429

Mountain JL, Cavalli-Sforza LL (1997) Multilocus genotypes, a tree of individuals, and human evolutionary history. Am J Hum Genet 61:705–718

Neumark YD, Friedlander Y, Thomasson HR, Li T-K (1998) Association of the *ADH2*2* allele with reduced ethanol consumption in Jewish men in Israel: a pilot study. J Stud Alcohol 59:133–139

Nickerson DA, Taylor SL, Weiss KM, Clark AG, Hutchinson RG, Stengard J, Salomaa V, et al (1998) DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene. Nat Genet 19:233–240

Price RA, Lunetta K, Ness R, Charles MA, Saad MF, Ravussin E, Bennett PH, et al (1992) Obesity in Pima Indians: distribution characteristics and possible thresholds for genetic studies. Int J Obes Relat Metab Disord 16:851–857

Riess O, Winkelmann B, Epplen JT (1994) Toward the complete genomic map and molecular pathology of human chromosome 4. Hum Genet 94:1–18

Sambrook J, Hanahan D, Rodgers L, Gething MJ (1986) Expression of human tissue–type plasminogen activator from lytic viral vectors and in established cell lines. Mol Biol Med 3:459–481

Schurr TG, Ballinger SW, Gan Y-Y, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, et al (1990) Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. Am J Hum Genet 46:613–623

Shen Y-C, Fan J-H, Edenberg HJ, Li T-K, Cui Y-H, Wang Y-F, Tian C-H, 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

Smith M (1986) Genetics of human alcohol and aldehyde dehydrogenases. Adv Hum Genet 15:249–290

Tanaka FY, Shiratori Y, Yokosuka O, Imazeki F, Tsukada Y, Omata M (1997) Polymorphism of alcohol-metabolizing genes affects drinking behavior and alcoholic liver disease in Japanese men. Alcohol Clin Exp Res 21:596–601

Thomasson HR, Crabb DW, Edenberg HJ, Li TK, Hwu HG, Chen CC, Yeh EK, 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

Thomasson HR, Edenberg HJ, Crabb DW, Mai X-L, Jerome RE, Li T-K, Wang S-P, et al (1991) Alcohol and aldehyde dehydrogenase genotypes and alcoholism in Chinese men. Am J Hum Genet 48:677–681

Tishkoff SA, Goldman A, Calafell F, Speed WC, Deinard AS, Bonne-Tamir B, Kidd JR, et al (1998) A global haplotype analysis of the myotonic dystrophy locus: implications for the evolution of modern humans and for the origin of myotonic dystrophy mutations. Am J Hum Genet 62: 1389–1402

Tishkoff SA, Ruano G, Kidd JR, Kidd KK (1996) Distribution and frequency of a polymorphic Alu insertion at the plasminogen activator locus in humans. Hum Genet 97:759–764

Valdes AM, Thomson G (1997) Detecting disease-predisposing variants: the haplotype method. Am J Hum Genet 60: 703–716

Wall T, Garcia-Andrade C, Thomasson HR, Carr LG, Ehlers CL (1997) Alcohol dehydrogenase polymorphisms in Native Americans: identification of the *ADH2*3* allele. Alcohol Alcohol 32:129–132

Walzer C, Turler H, Balant L, Golaz O, Hochstrasser DF, Monteiro M, von Wartburg JP (1993) Determination of human alcohol dehydrogenase and acetaldehyde dehydrogenase genotypes by single strand conformation polymorphism in discontinuous buffer electrophoresis. Electrophoresis 14: 566–569

Whitfield JB, Nightingale BN, Bucholz KK, Madden PAF, Heath AC, Martin NG. *ADH* genotypes and alcohol use and dependence in Europeans. Alcohol Clin Exp Res (in press)

Willard HF, Skolnick MH, Pearson PL, Mandel J-L (1985) Report of the Committee on Human Gene Mapping by Recombinant DNA Techniques. Cytogenet Cell Genet 40: 360–489

Yasunami M, Kikuchi I, Sarapata D, Yoshida A (1990) The human class I alcohol dehydrogenase gene cluster: three genes are tandemly organized in an 80-kb–long segment of the genome. Genomics 7:152–158

Zoossmann-Diskin A, Ticher A, Hakim I, Goldwitch Z, Rubinstein A, Bonne-Tamir B (1991) Genetic affinities of Ethiopian Jews. Isr J Med Sci 27:245–251