Association of the gastric alcohol dehydrogenase gene ADH7 with variation in alcohol metabolism

Andrew J. Birley1,†, Michael R. James1,‡, Peter A. Dickson1,†, Grant W. Montgomery1, Andrew C. Heath2, John B. Whitfield1 and Nicholas G. Martin1,*

1Genetic Epidemiology Unit, Queensland Institute of Medical Research, Brisbane, Australia and 2Midwest Alcoholism Research Center, Washington University, St Louis MO, USA

Received October 1, 2007; Revised and Accepted October 3, 2007

Seven alcohol-metabolizing enzymes are encoded by the human alcohol dehydrogenase (ADH) gene cluster on chromosome 4q22–23. One of these genes, ADH7, is uniquely expressed in the stomach mucosa and can influence metabolism of alcohol before its absorption into the blood. However, the contribution of ADH7 to the overall genetic variation in alcohol oxidation in vivo is unknown. Data on in vivo alcohol metabolism were obtained for 206 Australian twin pairs of Caucasian ancestry, following ingestion of a standard dose (0.75 g kg⁻¹ body weight) of alcohol. Twenty-five single nucleotide polymorphisms that cover the ADH7 encoding region were genotyped. The patterns of linkage disequilibrium among these SNPs identified a recombinational hotspot within intron 7 of the ADH7 gene. A model for the absorption and elimination of alcohol from the body led to the identification of haplotypes associated with inter-individual variation in the early stages of alcohol metabolism. These are within a 35 kb DNA tract contained in the region 5’ of intron 7 in the ADH7 gene. The region accounts for 18% of the linkage for alcohol concentration associated with the ADH region, or ~11% of the genetic variance.

INTRODUCTION

Variation in human alcohol (ethanol) metabolism is of interest because of known associations between alcohol dependence and polymorphisms in the alcohol-metabolizing enzymes alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH). The disease associations of ADH gene variation have been the subject of extensive and continuing research (1–8), and the ADH allelic associations found are presumed to be mediated by effects on alcohol metabolism. Linkage analysis using blood and breath alcohol concentration (BAC, BrAC) data from the Alcohol Challenge Twin Study (ACTS) (9) and subsequent microsatellite genotyping have shown strong evidence for a major quantitative trait locus (QTL) influencing alcohol metabolism, linked to the ADH gene region of chromosome 4. This QTL has been estimated to account for 64% of the genetic variance in alcohol metabolism (10). The coding polymorphisms ADH1B Arg47His and ADH1C Val349Ile, which are known to affect enzyme activity in vitro, accounted for only ~1% of the genetic variance in BAC in vivo. Therefore, this QTL must be attributable to other polymorphism(s) within or near the ADH gene cluster. Association studies with closely spaced single-nucleotide polymorphism (SNP) markers offer a way to locate and identify them.

The rate-limiting conversion of alcohol to acetaldehyde is catalysed by ADHs in the human stomach and liver. Seven ADH genes are in a cluster (5’-ADH7-ADH1C-ADH1B-ADH1A-ADH6-ADH4-ADH5-3’) on chromosome 4q22–23 (11–13). ADH7 is only expressed in the oesophagus and gastric mucosa (14–18), whereas ADH4, ADH6, ADH1A, ADH1B and ADH1C are mainly expressed in the liver and account for ~80% of post-absorptive alcohol metabolism (19).

The ADH-related QTL affecting breath and blood alcohol levels (10) appears to act early in the time course of alcohol metabolism, soon after ingestion when ethanol concentrations are highest. This suggests that variation in ADH enzymes with a comparatively high $K_m$ (low affinity) for alcohol, present in

---

†To whom correspondence should be addressed: Genetic Epidemiology Unit, Queensland Institute of Medical Research, PO Royal Brisbane Hospital, Queensland 4029, Australia. Tel: +61 7 33620278; Fax: +61 7 33620101; Email nick.martin@qimr.edu.au
‡Peter Dickson died on 15th October 2005.
#The Author 2007. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org
the stomach or liver, contributes to the QTL. Gastric ADH activity is primarily due to ADH7. The Class I genes ADH1B and ADH1C are also expressed in the gastric mucosa, and ADH1C is present in stomach muscle. However, the in vitro kinetic properties of ADH1B and ADH1C enzymes imply a smaller contribution to gastric metabolism than from ADH7, due to substrate inhibition of ADH class I enzymes by ethanol (20,21). Several authors have noted that at high alcohol concentration ADH7 has the highest activity of all stomach-expressed ADHs for ethanol (15,16,20–22), and therefore represents a good candidate for study. We have typed a total of 104 SNPs across the ADH gene cluster; this paper reports results from 25 SNPs within and flanking the ADH7 gene. We have tested for association between these ADH7 SNPs and variation in alcohol metabolism, and estimated the contribution of ADH7 gene variation to the QTL previously described.

RESULTS
Linkage disequilibrium in the ADH7 region
Two remarkably polarized haplotype blocks span the ADH7 gene region, and separately include sequences flanking the 5’ and 3’ ends of the ADH7 transcription unit, as shown in Figure 1. The short region between the haplotype blocks is marked by rs1154454 which is not in linkage disequilibrium (LD) with any other SNP. The 3’ haplotype block contains ten typed SNPs (rs1348276 to rs284784; see Table 1). SNPs rs1154458 to rs1583971 form a 5’ block. This striking discontinuity of LD within the ADH7 gene was also noted by Han et al. (22) in a study of 38 distinct populations. The 5’ and 3’ blocks of LD are consistent with a recombination hotspot within the transcribed ADH7 gene. Since SNP rs1154454 (Intron 7) is not in LD with either the 5’ or 3’ blocks, it is likely that the hotspot is within Intron 7 where there would be no disruption to the coding sequences. Nevertheless, we cannot exclude Exon 7 and sequences in Intron 6 or the region 5’ of rs284784 in Intron 8 as boundaries of the hotspot. The preliminary scan of LD across the ADH7 region showed a tendency for D’ values between SNPs located 5’ of Exon 1 (Fig. 1) to be less than unity, and the most 3’-located of the 25 genotyped SNPs, rs283406 through rs2032350, were likely that the hotspot is within Intron 7 where there would be no disruption to the coding sequences. Nevertheless, we cannot exclude Exon 7 and sequences in Intron 6 or the region 5’ of rs284784 in Intron 8 as boundaries of the hotspot. The preliminary scan of LD across the ADH7 region showed a tendency for D’ values between SNPs located 5’ of Exon 1 (Fig. 1) to be less than unity, and the most 3’-located of the 25 genotyped SNPs, rs283406 through rs2032350, were part of another distinct haplotype block not involving ADH7. Cluster analysis [MERLIN 1.0.1 (23)] for all SNPs in the ADH7 region also identified the two main 5’ and 3’ blocks, originally defined with HAPLOVIEW.

The discontinuities of LD in the ADH7 region were confirmed in data within the Hapmap project (http://www.hapmap.org/cgi-perl/gbrowse/hapmap_B35;/NCBI B25 assembly, dbSNP125). The region in intron 7 marked by rs1154454 coincided with a hotspot (estimated recombination rate, 3.5 cM/Mb), whereas the ~11.6 kb region between ADH7 and ADH1C and respectively bounded by rs2032350 and rs1348276, corresponded to a further recombination hotspot (31.7 cM/Mb). In summary, the ADH7 gene is characterized by two distinct regions of high LD which are genetically uncorrelated with each other, and the discontinuity of LD marking two haplotype blocks is within intron 7. In addition, the Hapmap data show very little LD between ADH7 and ADH1C. This is consistent with the activity of recombination hotspots that have eroded the LD between ADH7 and other members of the ADH gene family during recent primate evolution.

The presence of three segregating units within the ADH7 gene—haplotypes for the 5’ and 3’ blocks and the intervening region marked by SNP rs1154454—allowed us to minimize the number of tests for association of haplotypes with alcohol metabolism. Haplotype frequencies within the 5’ haplotype block were estimated for the cluster of adjacent SNPs identified when $r^2 \geq 0.6$. Within this block, the most frequent haplotype (haplotype 5’a in Table 2) includes all of the common alleles and there are six rarer haplotypes (haplotypes 5’b to 5’g). Three haplotypes (5’e, 5’f and 5’g) had frequencies of <5% and their effects on alcohol metabolism were too small to be estimated.

Nine haplotypes (3’a through 3’i in Table 2) were identified for the nine SNPs in the 3’ haplotype block using a threshold value of $r^2=0.6$ for adjacent SNPs. Within this block, we observed a common haplotype including all common alleles (haplotype 3’a) and eight further haplotypes. Haplotypes 3’b, 3’d and 3’f had frequencies of <5% and their effects were not estimated in the analysis of phenotypic associations.

The detection of phenotypic effects associated with the ADH7 gene/region was made with reference to the most frequent and clearly defined haplotypes within the 5’ and 3’ blocks, as well as individual SNPs.

Preliminary tests for effects of ADH7 variation on alcohol metabolism
The constrained or time-averaged effect reached the 5% level of statistical significance (Fig. 2) for BAC levels with SNPs rs1154461, rs1154468 and rs894363 and for BrAC levels with these three and also rs1154470. All four of these SNPs are 5’ proximal to the hotspot for recombination within ADH7. The effects were generally largest at the first BAC or BrAC reading and were statistically consistent over time for BrAC level, but were heterogeneous for BAC levels and fell with time for rs1154468 and rs1154470. The non-synonymous coding SNP rs1573496, flanked by rs1154461 and rs1154468, is at a lower frequency than the remaining SNPs in the 5’ haplotype block (Table 1), and was not significantly associated with alcohol levels. Similarly, SNP rs971074 formed part of the 5’ haplotype block, yet was not associated with alcohol levels. Evidence of phenotypic association was only seen for the four most frequent SNPs in the 5’ block. Association in the 3’ haplotype block was confined to rs994772 which is proximal to, but 3’ of, the ADH7 transcription unit. Allelic association was also present for rs283406, located 3’ of the ADH7 region in a further small haplotype block.

Goodness of fit to the pharmacokinetic model
The solutions for mean sex-adjusted BAC and BrAC values, using Eq. (1) (see Materials and Methods section), were:

$$BAC(t) = 11.70(1 - e^{-2.51t}) - 1.51t$$
and
$$BrAC(t) = 9.80(1 - e^{-2.99t}) - 1.30t$$

These blood and breath parameter estimates are in reasonable agreement for both the absorption rate constant, $k_1$ ($-2.51$ and $-2.99$), and the elimination rate $k_2$ (1.51 and 1.30). The difference in $A_0$ (11.7 against 9.8) is consistent with the BAC
and BrAC means; the calibration of the breath analyser is known to produce under-estimation of the blood alcohol concentrations. Comparison of the goodness of fit of this model with an analysis in which only the means are estimated by maximum likelihood resulted in $\chi^2 = 4.73; P = 0.19$ for blood and $\chi^2 = 45.46; P = 1.1 \times 10^{-7}$ for breath alcohol readings. While the kinetic model did not yield a poorer goodness of fit than the empirical analysis of means at each time point for BAC readings, a much poorer fit to the observed readings was found for BrAC readings. This discrepancy may in part be due to the greater power to detect a departure from the model-free estimates of means provided by the longer time span of readings covered by the ten BrAC time points, and remained when the model was extended to include the large differences due to sex (Fig. 3). The deviation between observed and expected means in BrAC readings is seen for both sexes during the first three readings, and for males during the linear stage of alcohol elimination. Nonetheless, the magnitude of the discrepancy between observed and expected readings is small (see Fig. 3) and the kinetic model is a good approximation to the BAC and BrAC curves.

Parameter estimates for effects of sex and age on BAC and BrAC curves following the kinetic model (Table 3) show the effects of age are small and confined to $A_0$. Although the subjects were within a narrow age range (18–34 years, mean 23 years), there was a small but significant positive regression of age on $A_0$ for both BAC and BrAC (Table 3). An effect of sex upon BAC and BrAC was detected for $A_0$, while the rate of absorption also depends upon sex for BrAC but not for BAC profiles. The kinetic model was therefore parameterized in the model for the means as:

\[
BAC(t) = (A_0 + A_{0A} + A_{0S})(1 - e^{k_1t}) + k_2t
\]

\[
BrAC(t) = (A_0 + A_{0A} + A_{0S})(1 - e^{(k_1 + k_2)t}) + k_2t
\]

where $A_{0A}$ and $A_{0S}$ are the effects of age and sex (male deviation) and $k_1$ is the male deviation for the rate of absorption. These adjustments for age and sex were included in the estimation of SNP and haplotype effects in the ADH7 region.
Associations between haplotypes or SNPs and BAC or BrAC levels

The effects of the rare haplotypes 5'e, 5'f and 5'g in the 5' haplotype block (Table 2) were not estimated. Haplotypes 5'c and 5'd differed from the consensus haplotype 5'a by two and one nucleotides respectively and were lower in frequency than haplotype 5'b (Table 2). Haplotypes 5'c and 5'd were not associated with BAC or BrAC levels (Table 4). Haplotype 5'b was more strongly associated with BAC than with BrAC.

Table 1. Typed SNPs within and flanking the ADH7 gene

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Chromosome Position (bp)</th>
<th>Distance next SNP (bp)</th>
<th>Ancestral allele</th>
<th>Poly morphism</th>
<th>MAF</th>
<th>Functionality/placement</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1354368</td>
<td>100 799 065</td>
<td>34 864</td>
<td>G</td>
<td>G/A</td>
<td>0.350</td>
<td>DKFZP434G072</td>
</tr>
<tr>
<td>rs1583971</td>
<td>100 764 201</td>
<td>15 918</td>
<td>T</td>
<td>A/T</td>
<td>0.107</td>
<td>Upstream ADH7</td>
</tr>
<tr>
<td>rs2851024</td>
<td>100 748 283</td>
<td>14 258</td>
<td>T</td>
<td>G/T</td>
<td>0.480</td>
<td>Upstream ADH7</td>
</tr>
<tr>
<td>rs894363</td>
<td>100 734 025</td>
<td>20 509</td>
<td>C</td>
<td>C/T</td>
<td>0.403</td>
<td>Upstream ADH7</td>
</tr>
<tr>
<td>rs1154470</td>
<td>100 713 516</td>
<td>2080</td>
<td>G</td>
<td>G/A</td>
<td>0.335</td>
<td>Intron 1</td>
</tr>
<tr>
<td>rs1154468</td>
<td>100 711 436</td>
<td>4589</td>
<td>A</td>
<td>A/T</td>
<td>0.335</td>
<td>Intron 1</td>
</tr>
<tr>
<td>rs1154466</td>
<td>100 707 847</td>
<td>6766</td>
<td>C</td>
<td>G/C</td>
<td>0.100</td>
<td>Intron 5</td>
</tr>
<tr>
<td>rs1348726</td>
<td>100 682 809</td>
<td>11 568</td>
<td>G</td>
<td>G/A</td>
<td>0.110</td>
<td>Exon 6 (Arg:Arg)</td>
</tr>
<tr>
<td>rs894369</td>
<td>100 691 025</td>
<td>1339</td>
<td>G</td>
<td>G/A</td>
<td>0.100</td>
<td>Intron 6</td>
</tr>
<tr>
<td>rs1154548</td>
<td>100 697 091</td>
<td>2180</td>
<td>C</td>
<td>C/G</td>
<td>0.413</td>
<td>Intron 6</td>
</tr>
<tr>
<td>rs1154545</td>
<td>100 695 217</td>
<td>2468</td>
<td>T</td>
<td>T/C</td>
<td>0.182</td>
<td>Intron 7</td>
</tr>
<tr>
<td>rs284784</td>
<td>100 693 053</td>
<td>1897</td>
<td>G</td>
<td>G/T</td>
<td>0.236</td>
<td>Intron 8</td>
</tr>
<tr>
<td>rs284786</td>
<td>100 691 156</td>
<td>45</td>
<td>T</td>
<td>A/T</td>
<td>0.309</td>
<td>3'-UTR</td>
</tr>
<tr>
<td>rs3805331</td>
<td>100 691 111</td>
<td>86</td>
<td>A</td>
<td>A/G</td>
<td>0.073</td>
<td>3'-UTR</td>
</tr>
<tr>
<td>rs894369</td>
<td>100 691 025</td>
<td>234</td>
<td>C</td>
<td>C/G</td>
<td>0.229</td>
<td>3'-UTR</td>
</tr>
<tr>
<td>rs3805329</td>
<td>100 690 791</td>
<td>56</td>
<td>T</td>
<td>T/C</td>
<td>0.074</td>
<td>3'-UTR</td>
</tr>
<tr>
<td>rs284787</td>
<td>100 690 735</td>
<td>289</td>
<td>C</td>
<td>C/T</td>
<td>0.236</td>
<td>3'-UTR</td>
</tr>
<tr>
<td>rs729147</td>
<td>100 690 446</td>
<td>3574</td>
<td>A</td>
<td>G/A</td>
<td>0.228</td>
<td>Downstream ADH7</td>
</tr>
<tr>
<td>rs9690804</td>
<td>100 686 872</td>
<td>2029</td>
<td>—</td>
<td>T/A</td>
<td>0.392</td>
<td>Downstream ADH7</td>
</tr>
<tr>
<td>rs994772</td>
<td>100 684 843</td>
<td>2034</td>
<td>G</td>
<td>G/A</td>
<td>0.118</td>
<td>Downstream ADH7</td>
</tr>
<tr>
<td>rs1348276</td>
<td>100 682 809</td>
<td>11 568</td>
<td>T</td>
<td>T/G</td>
<td>0.394</td>
<td>Downstream ADH7</td>
</tr>
<tr>
<td>rs2032350</td>
<td>100 671 241</td>
<td>7886</td>
<td>C</td>
<td>C/T</td>
<td>0.197</td>
<td>Downstream ADH7</td>
</tr>
<tr>
<td>rs1442484</td>
<td>100 663 355</td>
<td>5128</td>
<td>T</td>
<td>T/C</td>
<td>0.200</td>
<td>Downstream ADH7</td>
</tr>
<tr>
<td>rs1826906</td>
<td>100 658 227</td>
<td>2577</td>
<td>C</td>
<td>T/C</td>
<td>0.314</td>
<td>Downstream ADH7</td>
</tr>
<tr>
<td>rs283406</td>
<td>100 655 650</td>
<td>0</td>
<td>C</td>
<td>C/T</td>
<td>0.052</td>
<td>Downstream ADH7</td>
</tr>
</tbody>
</table>

bNCBI Build 36.1.
cBased on chimpanzee (NCBI)—indicates unavailable.
dMajor allele—Minor allele.

Table 2. Haplotype frequencies for the ADH7 5' and 3'-haplotype blocks

<table>
<thead>
<tr>
<th>Haplotype Name</th>
<th>Frequency</th>
<th>rs971 074</th>
<th>rs1154 461</th>
<th>rs1154 468</th>
<th>rs1154 470</th>
<th>rs894 363</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'a</td>
<td>0.478</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5'b</td>
<td>0.330</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5'c</td>
<td>0.112</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5'd</td>
<td>0.067</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5'e</td>
<td>0.010</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5'f</td>
<td>0.001</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5'g</td>
<td>0.001</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5'-Sub-block</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
</tbody>
</table>

Haplotypes depicted in terms of major (1) and minor (2) alleles as shown in Table 1.

4Haplotype sub-blocks revealed with $r^2 > 0.6$.

Associations between haplotypes or SNPs and BAC or BrAC levels

The effects of the rare haplotypes 5'e, 5'f and 5'g in the 5' haplotype block (Table 2) were not estimated. Haplotypes 5'c and 5'd differed from the consensus haplotype 5'a by two and one nucleotides respectively and were lower in frequency than haplotype 5'b (Table 2). Haplotypes 5'c and 5'd were not associated with BAC or BrAC levels (Table 4). Haplotype 5'b was more strongly associated with BAC than with BrAC.
levels (Table 4, Fig. 4). The signs of the effect sizes of haplotype 5\textsuperscript{0}b upon \(A_0\) and \(k_1\) (level and rate of absorption, respectively) were the same, in accordance with expectations of a first-order kinetic process. The joint effect of the two parameters \(A_0\) and \(k_1\) for 2 df, measuring the effects of the early stages of absorption, and the rate of elimination \(k_2\) (1 df) were used to compare haplotype effects. No effects for alcohol elimination were detected for \(k_2\) but \(A_0\) and \(k_1\), were jointly significant for BAC level and, less strongly, for BrAC (Table 4). Compared to haplotype 5\textsuperscript{0}a, haplotype 5\textsuperscript{0}b is associated with the rare allele at the four most common SNPs rs1154461, rs1154468, rs1154470 and rs894363. Haplotype 5\textsuperscript{0}c is complementary to haplotype 5\textsuperscript{0}b (Table 2) and was not associated with alcohol levels; its effect size is not significantly different from that of haplotype 5\textsuperscript{0}a. The effect on BAC and BrAC levels is most likely associated with rs1154461, rs1154468, rs1154470 or rs894363.

Figure 2. Probabilities for tests of SNP association with BAC and BrAC levels across the ADH7 region. See Table 1 for SNP details.

Figure 3. Time course for mean BAC and BrAC levels by sex. The means for the ‘Empirical model’ of mean BAC and BrAC levels are compared to the estimates from the ‘Kinetic model’ [Eq. (1)].

Extensive tests were performed for effects of individual haplotypes and combinations of haplotypes within the 3’ block on the various parameters of alcohol metabolism, but no significant effects were found (data not shown).

Table 3. Maximum likelihood estimates (\(\theta\)) of the effects of sex and age upon the parameters of the kinetic equation \(A_0\), \(k_1\) and \(k_2\)

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Parameter</th>
<th>BAC (\hat{\theta})</th>
<th>(\hat{\theta})</th>
<th>BrAC (\hat{\theta})</th>
<th>(\hat{\theta})</th>
<th>(\chi^2)</th>
<th>(P)</th>
<th>(\chi^2)</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>(A_0)</td>
<td>0.079</td>
<td>4.48</td>
<td>0.034</td>
<td>0.045</td>
<td>4.77</td>
<td>0.029</td>
<td>0.003</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>(k_1)</td>
<td>0.009</td>
<td>0.12</td>
<td>0.735</td>
<td>0.035</td>
<td>1.77</td>
<td>0.184</td>
<td>0.003</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>(k_2)</td>
<td>-0.023</td>
<td>1.53</td>
<td>0.217</td>
<td>-0.003</td>
<td>0.21</td>
<td>0.649</td>
<td>0.003</td>
<td>0.21</td>
</tr>
<tr>
<td>Sex</td>
<td>(A_0)</td>
<td>-1.187</td>
<td>13.83</td>
<td>2.00e-04</td>
<td>-1.281</td>
<td>47.61</td>
<td>5.21e-12</td>
<td>0.345</td>
<td>2.35</td>
</tr>
<tr>
<td></td>
<td>(k_1)</td>
<td>0.345</td>
<td>2.35</td>
<td>0.125</td>
<td>0.996</td>
<td>17.61</td>
<td>2.71e-05</td>
<td>0.143</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>(k_2)</td>
<td>0.143</td>
<td>0.89</td>
<td>0.347</td>
<td>0.094</td>
<td>3.19</td>
<td>0.074</td>
<td>0.003</td>
<td>0.21</td>
</tr>
</tbody>
</table>
Contribution of haplotypes in the 5' block of ADH7 to linkage in the ADH gene region

The models for the effects of the 5' block upon mean BrAC or BAC levels were used to estimate the contribution of haplotypes in the 5'0 ADH7 region to linkage of in vivo metabolism to the ADH region on chromosome 4. The effect of haplotype 5'b (Table 2) was estimated as a deviation from the baseline effects of all remaining haplotypes. Since the effects of haplotypes 5'c through 5'g (Table 2) were homogenous with those of the most common haplotype 5'a, our statistical comparison of haplotype effects is as previously described, a deviation of the effect of haplotype 5'b from the baseline haplotype 5'a. In the bivariate simplex model, the covariance due to the QTL and the effect of haplotypes in the means part of the equation are jointly estimated, in the model for the multivariate normal distribution.

The contribution of the major haplotype 5'b to the QTL for alcohol concentration was estimated by dropping its effects from the bivariate simplex model. The change in the log-likelihood was \( \chi^2 = 22.7 (P = 0.0009) \) for the kinetic parameters, \( A_0 \), \( k_1 \) and \( k_2 \), corresponding to the effects of haplotypes upon breath alcohol (almost entirely due to \( A_0 \)) and blood alcohol level (almost entirely due to \( k_1 \)). Direct calculation of the joint effect of allelic deviations (in the means part of the multivariate likelihood equation) for both blood and breath alcohol levels gave an estimate of 18% for the effect of the 5' haplotypes, and the difference in means between haplotypes 5'a and 5'b at the first blood alcohol measurement time was 5.8 mg/100 ml.

DISCUSSION

There are a number of practical reasons to study interindividual variation in alcohol pharmacokinetics. The peak concentration achieved after any specified dose and the rate of metabolism will affect the duration and intensity of intoxication, and the rate of production and peak concentration of alcohol metabolites such as acetaldehyde. There is some evidence to suggest that variation in alcohol pharmacokinetics affects alcohol dependence risk (24), and a large amount of evidence implicating alcohol-metabolizing enzymes in alcohol dependence risk and in the development of other alcohol-related disease (25–28). Many of the published twin studies have been under-powered to detect a significant genetic component to variation in alcohol metabolism (29–32), but Martin et al. (9) showed heritability of 0.62 for peak blood alcohol and 0.49 for rate of elimination. So far, there is little information on the exact genes and polymorphisms involved. Studies on the most obvious candidates, the long-recognized polymorphisms in \( ADH1B \) (Arg47His and Arg369Cys) and \( ADH1C \) (Ile 349Val/ Arg271Gln) have revealed only minor \( \chi^2 \) effects (10,33,34), although in vitro kinetic properties are substantially different (35). Despite the negative results with known \( ADH \) variants, we were able to show (10) that a substantial QTL for alcohol metabolism exists at the \( ADH \) gene cluster on chromosome 4. This could of course be due to a single polymorphism or to additive or interaction effects of several; the latter would be harder to detect but should be approachable through haplotype association studies.

Table 4. Effects of haplotypes in the 5' Haplotyp Block of ADH7 upon BAC and BrAC levels

<table>
<thead>
<tr>
<th>Model</th>
<th>versus</th>
<th>BAC (-2LL)</th>
<th>df</th>
<th>( \chi^2 )</th>
<th>df</th>
<th>( P )</th>
<th>BrAC (-2LL)</th>
<th>df</th>
<th>( \chi^2 )</th>
<th>df</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Haplotypes b, c and d</td>
<td>5489.64</td>
<td>2038</td>
<td></td>
<td>5206.80</td>
<td>3367</td>
<td></td>
<td>184 Human Molecular Genetics, 2008, Vol. 17, No. 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Drop haplotypes c and d</td>
<td>5500.33</td>
<td>2044</td>
<td>10.696</td>
<td>6</td>
<td>0.0982</td>
<td>7211.20</td>
<td>3373</td>
<td>4.400</td>
<td>6</td>
<td>0.6227</td>
</tr>
<tr>
<td>3</td>
<td>Drop haplotype b</td>
<td>5516.96</td>
<td>2047</td>
<td>16.628</td>
<td>3</td>
<td>0.0008</td>
<td>7220.38</td>
<td>3376</td>
<td>9.185</td>
<td>3</td>
<td>0.0269</td>
</tr>
<tr>
<td>4</td>
<td>( A_0 )</td>
<td>5500.34</td>
<td>2045</td>
<td>0.010</td>
<td>1</td>
<td>0.9203</td>
<td>7215.63</td>
<td>3374</td>
<td>4.437</td>
<td>1</td>
<td>0.0352</td>
</tr>
<tr>
<td>5</td>
<td>( k_1 )</td>
<td>5509.65</td>
<td>2045</td>
<td>9.321</td>
<td>1</td>
<td>0.0023</td>
<td>7212.83</td>
<td>3374</td>
<td>1.633</td>
<td>1</td>
<td>0.2013</td>
</tr>
<tr>
<td>6</td>
<td>( k_2 )</td>
<td>5500.55</td>
<td>2045</td>
<td>0.213</td>
<td>1</td>
<td>0.6444</td>
<td>7212.73</td>
<td>3374</td>
<td>1.536</td>
<td>1</td>
<td>0.2152</td>
</tr>
<tr>
<td>7</td>
<td>Drop ( A_0, k_1 )</td>
<td>5513.69</td>
<td>2046</td>
<td>13.358</td>
<td>2</td>
<td>0.0013</td>
<td>7220.03</td>
<td>3375</td>
<td>8.833</td>
<td>2</td>
<td>0.0121</td>
</tr>
<tr>
<td>8</td>
<td>Drop ( k_2 )</td>
<td>5515.96</td>
<td>2047</td>
<td>2.270</td>
<td>1</td>
<td>0.1319</td>
<td>7220.38</td>
<td>3376</td>
<td>0.352</td>
<td>1</td>
<td>0.5530</td>
</tr>
</tbody>
</table>

All effects are deviations from the baseline value of haplotype 5'a.
We have now shown that SNP haplotypes across approximately 35 kb of DNA in ADH7, spanning a region just 5' of exon 1 and extending into intron 6 of the ADH7 coding region, are related to blood and breath alcohol concentrations in vivo. Results for blood and breath data differ in some details, but Table 4 shows that although the 5' haplotype effect is associated with $k_i$ for blood alcohol and $A_0$ for breath alcohol, the joint effects are significant for both. This is evident in Figure 4, which illustrates that early alcohol concentrations differ by 5' haplotype for both blood and breath measurements. Although blood and breath alcohol measurements are both subject to measurement error, they point to a similar conclusion.

This result defines a new region of significance for the in vivo metabolism of alcohol, although our best estimate is that it accounts for only 18% of the previously reported linkage to the ADH gene region, or 11% of the total genetic variance. This is consistent with the expectation of multiple additive or interaction effects at several loci within or close to the ADH gene cluster.

Definition of the boundaries of the region responsible for this effect is aided by the pattern of LD between adjacent SNPs across the ADH7 gene. We identified two major haplotype blocks, which show a marked discontinuity of LD in a region centered on SNP rs1154454, within intron 7, and flanked by rs1154458 (5') and rs248784 (3'). The most likely explanation for this is that there is a recombination hotspot within intron 7 of ADH7. This was also observed by Han et al. (22) in a world-wide study of patterns of allele frequency and LD. They used different markers, but included rs971074 and rs1154458 (3') and rs284784 and rs284786 (5'). The 5' haplotype block was observed by Luo et al. (12), in D' values for rs971074, rs11573496 and rs1154470 in Americans of European descent. While only one SNP (rs284786) in the 3' block was genotyped by Luo et al., it was not in LD with any of the 5' located ADH7 SNPs. D' values for SNPs in the ADH7 region chosen by Edenberg et al. (8) also identified the 5' LD block, but their coverage of the 3' block adjacent to rs1154454 was less complete.

Two major haplotypes in the 5' haplotype block were found to be associated with in vivo metabolism. The nucleotide composition of these haplotypes was, in 5’–3’ order, GCCAGC (frequency 0.48) and GCCTAT (0.33), which are distinguished by four nucleotide substitutions. Of these, we can tentatively exclude the effect of GCCAGC as compared to GCCTAT since the effect of the rare haplotype AGGTAT was not different from that of GCCAGC in the analysis of 5' block effects. Because this haplotype has a frequency of only 0.11, we should regard this conclusion with some caution. However, the second nucleotide in the two most frequent haplotypes GCCAGC and GCCTAT differs. The relevant SNP is rs1154461 within intron 5, which has a minor allele frequency (MAF) of 0.3. This is immediately 5' of rs1573496 which is responsible for the non-synonymous substitution ADH7Ala80Gly (with a MAF ~0.1) which is exclusively part of haplotype 5'c. However, SNP rs1573496 is the same (C) in both haplotypes 5'a and 5'b so cannot account for the effect of the 5' haplotype block on in vivo metabolism.

Another SNP in this region, which we did not type, may affect gene expression. A T/C substitution at +25 bp in the 5'-untranslated region changes a TATA sequence to TACA (36,37). This variant is within the 5' haplotype block and probably in strong LD with our typed SNPs. It may account for the variation in alcohol metabolism, but the individual effects of SNPs in strong LD are difficult to differentiate by allelic association studies.

In contrast to the 5'-haplotype block, the 3'-haplotype block shows very little evidence of association with in vivo metabolism. Although tests of haplotypes did reveal indications of effects, they were not consistent for BAC and BrAC and we conclude that any effects are smaller than those seen for the 5'-haplotype block.

Turning to the issue of pre-absorptive alcohol metabolism, the contribution of gastric ADH has been recognized in the literature, but with variable results that may be confounded by study method (38,39). As noted above, gastric ADH is primarily coded by the ADH7, ADH1B and ADH1C genes which are expressed in the gastric mucosa, with ADH1C expressed in stomach muscle. In vitro kinetic properties imply a smaller contribution of ADH1B and ADH1C enzymes to gastric metabolism than of ADH7, due to substrate inhibition of ADH class I enzymes (20,21).

As well as in vitro study of gastric isoenzymes and gene expression, there has been extensive study in vivo of ‘first pass metabolism’ (FPM) of ethanol. This after absorption, or occurring before absorption, and assessed from the difference in blood alcohol concentrations between oral and intravenous alcohol administration. Its contribution to total alcohol metabolism in Caucasians ranges from 1 to 10% (40). Ethnic differences between Caucasians and Asian peoples support a major role of ADH7 in FPM in Caucasians; FPM is very low in Japanese because of an absence of expressed ADH7 enzyme (39,41). At very high alcohol concentrations, Japanese and Caucasian subjects both show an additional contribution to FPM from ADH5, which has a poor substrate affinity for ethanol (39). The contribution of gastric alcohol metabolizing enzymes (ADH7 and ADH5) to FPM in both Japanese and Caucasian subjects has been estimated to reach 40% at high alcohol concentration (39), although this estimate has been criticized as too large (19,40,42). Since FPM almost disappears on gastrectomy (43), gastric rather than hepatic ADHs must account for the bulk of FPM in Caucasians (44).

Our study shows that ADH7 variation has a significant impact on in vivo alcohol metabolism, commencing at the early times of blood alcohol measurement and carrying forward to the later times (Fig. 4). Since the SNPs on ADH7 5' haplotype block are not associated through LD with those in the 3' block, it is unlikely that we are measuring correlated effects due to LD elsewhere in the ADH7 region. The result also has implications for ethnic differences in the role of alcohol oxidation to acetaldehyde. Since ADH7 apparently is not expressed in a substantial proportion of people from East Asian populations (20,21,41) the effects of genetic variability in ADH7 upon initial acetaldehyde levels and consequences for alcohol dependence may be more important in Caucasians.

The magnitude of this effect has been estimated at 18% of the QTL or 11% of the genetic variance. Comparing the effect size for haplotype 5'b with the standard deviation for the peak blood alcohol concentration, we can express it as a z-score of 0.36. We conclude that although the contribution of genetic variation at
Further investigation of the effects of genetic variation in other ADHs on blood and breath alcohol concentrations in these Alcohol Challenge subjects must be complemented by SNP typing in our larger cohorts with data on alcohol use and dependence, and by existing published reports on related topics. As mentioned previously, there are a few genetic linkage or association studies on alcohol metabolism or pharmacokinetics. There are, however, a substantial number which test for effects of variation in or near ADH genes on related phenotypes, including alcohol dependence or ADH gene expression. These may be useful as a guide to genes, haplotypes or polymorphisms which affect alcohol metabolism. Until recently, most studies examined ADH1B Arg47His and ADH1C Val349Ile, but other suggestive or significant reports for alcohol dependence include effects within ADH1A, ADH1B, ADH5 and ADH7 (12); ADH1A, ADH1B and ADH4 (8); ADH7 (45) and the broad region of chromosome 4 spanning the ADH cluster (28). The report on ADH7’s suggested interaction between a polymorphism in this gene and the ADH1B Arg47His polymorphism, producing differences in risk of alcohol dependence among Chinese subjects. It will be more difficult to test for such interaction in European subjects, because of the lower frequency of the ADH1B Arg47His allele, but interactions across the ADH gene region need to be considered more generally when adequate data become available.

Expression of ADH genes has been shown to be affected by sites within the ADH gene cluster (46–48), or elsewhere on chromosome 4 (49). These diverse reports suggest that additive or epistatic effects of sequence variation in the ADH gene cluster combine to produce the effect found in our previous linkage analysis, and evaluating these will require multiple studies on ADH variation and its effects on transcription, enzyme activity, alcohol use and dependence and alcohol-related disease.

MATERIALS AND METHODS

Alcohol challenge study: subjects and phenotypes

Alcohol metabolism data were available for 206 pairs of twins, aged 18–34 years (mean 23), who took part in the ACTS in 1979–81 (9). Ten timed measurements of breath alcohol concentration and six of blood alcohol concentration were obtained, starting 40 min after ingestion of alcohol (0.75 g.kg⁻¹ body weight). Breath alcohol readings were converted to blood alcohol by a conversion factor of 2100:1; this is conservative in that it under-estimates blood alcohol concentrations. Full details are given in Martin et al. (9).

Blood samples and SNP assays

Some 10–20 years after completion of the ACTS, twins were recontacted to obtain blood samples for DNA extraction; at this time, we also collected blood samples from parents and/or non-twin siblings of the twins. The zygosity of ACTS twin pairs was originally determined by self-report and a limited range of serological and isoenzyme markers; extensive genotyping since the initial study has shown that some re-assignment is necessary. The revised zygosity status of participants is 91 monozygotic (MZ) twin pairs (46 female and 45 male) and 115 dizygotic (DZ) twin pairs (41 female, 35 male and 39 opposite sex).

A set of 25 polymorphic SNPs related to ADH7 (Table 1) was typed in 817 individuals from 196 of the original 206 twin families. SNP typing was performed with the MassARRAY genotyping platform (Sequenom Inc., San Diego, CA, USA). DNA was available for 220 non-twin siblings and 215 parents, in addition to the twins, in 112 DZ and 84 MZ families. At least one parent and non-twin sibling were present in 106 families, 24 more families had at least one non-twin sibling and 37 had one or both parents; 28 families only contained twins. The extended families provided additional information for the determination of haplotype phase. The ancestry of ACTS participants, based upon reported country of origin of their four grandparents, was estimated at 87% ‘Northern European’ (including Australia and New Zealand), and 13% ‘other’, nearly all Southern European/Mediterranean (50).

Genotype quality control

Family relationships were confirmed with PREST ver. 3.02 (51) and the SNPs conformed to the expectations of Mendelian Inheritance [PEDSTATS ver. 0.4.6 (52)]. A genotyping error rate of 0.12% was estimated from 65 MZ families for which both twins were genotyped. Eleven dispersed genotyping errors or spurious point double recombinants were identified by coinheritance checks [MERLIN ver. 1.0.1 (53)]. They were confined to six SNP loci that were sporadically distributed across the ADH7 region and not due to incorrect map locations (map order was from NCBI Build 35.1 (http://www.ncbi.nlm.nih.gov/projects/SNP/)). The error rate was in the range 0.05–0.49% for each of these six SNPs. The distribution of errors by family (eight families) showed they were sporadic in this respect and they were recoded as missing. Minor allele frequencies for the 25 SNPs were obtained from three databases; Caucasians/Northern Europeans in HapMap [Public Release No. 19, The International HapMap Consortium, 2003 (54)], the CEU population (CEPH Utah Residents with Northern and Western European Ancestry), NCBI Build 35.1 (http://www.ncbi.nlm.nih.gov/projects/SNP/) and from ALFRED (http://alfred.med.yale.edu/alfred/index.aspx). The mean difference in frequency of minor alleles as genotyped in the present study and published MAFs was small (0.012 ± 0.062).

Linkage disequilibrium

Pairwise estimates of LD across the ADH7 region were obtained from the SNP-typed data using HAPLOVIEW (ver. 3.2; http://www.broad.mit.edu/mpg/haploview/). The most likely haplotypes were identified using MERLIN 1.0.1 (53) which takes into account within-family segregation patterns and the extent of LD within the region. The definition of the haplotype block was resolved using MERLIN 1.0.1 by varying the threshold value for the genetic correlation.
needed to define a haplotype block separately for the 5’ and 3’ blocks using the ‘- -’ cluster option.

SNP associations with the alcohol phenotype

Two approaches were taken to obtain the maximum information from the data on blood and breath alcohol concentrations. These involved firstly, an empirical approach based on the multiple measures of blood and breath alcohol obtained during the study and, secondly, estimation of kinetic parameters of a physiological model of alcohol absorption, distribution and metabolism.

Effect sizes for SNP alleles were estimated for each time point in the means part of the model for the multivariate normal equation using Mx (55) in a preliminary screen of the ADH7 region. The expected sib pair covariance matrices were modeled jointly with means and were parameterized as additive and specific environmental covariances by Cholesky decomposition, taking into account the zygosity of twin pairs. Initial analysis included the effects of sex, age and the interaction of SNP effects with sex in the model for mean breath or blood alcohol levels. The interaction effects proved to be trivial and the final model only retained the effects of sex and age. Initial analysis of the association of SNP genotypes and alcohol levels showed no evidence for dominance effects.

An allelic association model was used in the preliminary analysis to test for the effect of the minor (less common) allele at each SNP as a deviation from the overall mean (effect of the common SNP allele). A profile of allelic effects against the location of each SNP in the ADH7 region was used to identify possible sites/regions of interest. The overall variance in blood or breath alcohol levels associated with SNPs was partitioned into the average effect or consistency of allelic effects for all time-points, and the heterogeneity (time-change) of allelic effects over time.

Pharmacokinetic model

Three main characteristics of the blood or breath alcohol/time curves describe the salient features of alcohol absorption and metabolism. There is an early absorptive phase when alcohol concentrations rise, a peak alcohol concentration reached when rates of absorption and elimination balance, and an elimination phase when alcohol levels fall. We have used a general approach based upon the physiological and biochemical processes of absorption and elimination that underlie BAC and BrAC curves. The model was described by Martin et al. (56) and used previously with the present data (9). It predicts concentration of alcohol, C(t), at time t following ingestion, using Eq. (1).

\[ C(t) = A_0(1 - e^{-k_1t}) + k_2t \]  

\( A_0 \) represents the concentration of alcohol that would be achieved if all ingested alcohol were absorbed and evenly distributed throughout the whole body or tissue volume, \( k_1 \) the rate of absorption, and \( k_2 \) the rate of elimination. The rate of elimination is a linear term that reflects substrate saturation or independence between the rate of reaction and ethanol concentration. This has been shown to be valid with a dose of 0.67 g kg\(^{-1}\) (close to the 0.75 g kg\(^{-1}\) used in our study) and persists until alcohol levels fall to 20 mg/100 ml (57). The early stages of alcohol absorption are concentration dependent and estimated by \( A_0 \) and \( k_1 \). Maximum likelihood estimates of the three parameters \( A_0, k_1 \) and \( k_2 \) were obtained in Mx (55) by specifying Eq. (1) for the concentration of alcohol at time \( t \) after ingestion, in the time-related means vector of the multivariate normal equation.

Haplotype association

ADH7 activity variants can be expected to affect both \( A_0 \) and \( k_1 \) by way of enzyme activity in the gastric mucosa during absorption. Both \( A_0 \) and \( k_1 \) will be biologically correlated, as too are the conventional measures of alcohol pharmacokinetics: time to peak alcohol concentration and peak alcohol concentration following drinking. In the absence of genetic association (LD) or epistasis between gastric ADH7 variants and other sites affecting alcohol metabolism, any effects detected by the parameters \( A_0 \) and \( k_1 \) will not affect the rate of elimination (\( k_2 \)) at high ethanol concentration.

The effects of haplotypes were estimated as deviations for \( A_0 \), \( k_1 \) and \( k_2 \) from the overall mean (defined by the most frequent haplotype pattern in the 5′ and 3′ blocks) (Table 2). Preliminary analysis did not show evidence for haplotype by sex interaction. Haplotype analysis for association was carried out separately for the 5′-haplotype block and 3′-haplotype block, defined above. We also tested for effects at SNP loci that were not in strong LD with the 5′- and 3′-haplotype blocks, where the overall reduction in log-likelihood indicated the possibility of an effect in at least one of the three estimated parameters.

Linkage and association

A joint test of linkage and association (10) allowed the estimation of the contribution of haplotypes in the ADH7 region to the covariation in the timed series of ten BrAC and six BAC readings. In this method, the effects of the covariance between breath and blood alcohol levels are jointly estimated with the effects of the measured genotypes or haplotypes in the means part of the multivariate normal equation, by maximum likelihood in the structural equation modeling package, Mx. The chosen model followed a bivariate simplex design (10) and showed that the joint genetic covariance of BrAC and BAC levels was largely initiated before the time of the first measurements in the time series (about 40 min after completion of drinking). Similarly, the effect of linkage was only identified with this time point, and only seen for the joint covariance in BrAC and BAC levels, not the covariance specific to either BrAC or BAC levels. The contribution of the 5′ haplotypes to the covariance between blood and breath readings, due to the QTL associated with the ADH gene family, was estimated following the method described by Fulker et al. (58). The effects of the 5′ haplotypes were estimated in the means model of the bivariate simplex model jointly with the covariance between BAC and BrAC levels due to the linked QTL. The model for the 5′ haplotypes is described in the Results section. The effect of the 5′ haplotypes upon linked covariance was estimated from the increase
REFERENCES


