The genomic determinants of alcohol preference in mice

Boris Tabakoff[1], **Laura Saba**[1], **Katherina Kechris**[2], **Wei Hu**[1], **Sanjiv V Bhave**[1], **Deborah A Finn**[3], **Nicholas J Grahame**[4], and **Paula L Hoffman**[1]

1*Department of Pharmacology, University of Colorado Denver School of Medicine, Aurora, CO 80045, USA.*

2*Department of Preventive Medicine and Biometrics, University of Colorado Denver School of Medicine, Aurora, CO 80045, USA.*

3*Department of Behavioral Neuroscience, Oregon Health and Science University and VAMC Research, Portland, OR 97239, USA.*

4*Department of Psychology, Indiana University School of Medicine, Indianapolis, IN 46202, USA.*

Abstract

Searches for the identity of genes which influence the levels of alcohol consumption by humans and other animals have often been driven by presupposition of the importance of particular gene products in determining positively or negatively reinforcing effects of ethanol. We have taken an unbiased approach and performed a meta-analysis across three types of mouse populations to correlate brain gene expression with levels of alcohol intake. Our studies, using filtering procedures based on QTL analysis, produced a list of eight candidate genes with highly heritable expression, which could explain a significant amount of the variance in alcohol preference in mice. Using the Allen Brain Atlas for gene expression, we noted that the candidate gene expression was localized to the olfactory and limbic areas as well as the orbitofrontal cortex. Informatics techniques and pathway analysis illustrated the role of the candidate genes in neuronal migration, differentiation and synaptic remodeling. The importance of olfactory cues, learning and memory formation (Pavlovian conditioning), and cortical executive function, for regulating alcohol intake by animals (including humans), is discussed.

Introduction

A large number of original studies and reviews (e.g., Enoch and Goldman 2001) have alluded to a "genetic" predisposition to "alcoholism" (alcohol dependence). These publications presume that the genes involved in this disorder, in combination with environmental factors, influence the susceptibility of an individual to develop dependence on alcohol, once that individual begins to drink alcohol. The fact that alcohol consumption is a prerequisite for the development of alcohol dependence may seem self-evident, but important distinctions between high alcohol intake in animal models of "alcoholism", and the signs and symptoms of alcohol dependence in humans, have many times been blurred. Alcohol dependence in humans, as defined by ICD 10 or DSM IV (American Psychiatric Association 1994; World Health Organization 2005) criteria, is a multifaceted syndrome in which diagnosis depends on the presence, in an individual, of three or more out of seven criteria, continuously over a period of twelve months. The quantitative aspects of alcohol consumption do not currently enter into the definition of alcohol dependence in humans. However, the progression from nondependent

*Correspondence: Boris Tabakoff, Department of Pharmacology, University of Colorado at Denver and Health Sciences Center, Mail Stop F-8303, P.O. Box 6511, Aurora, CO 80045-0511, USA, Fax number: 303-724-3664, Voice phone number: 303-724-3668, boris.tabakoff@ucdenver.edu.

alcohol drinking to alcohol dependence has been considered to be a dose-dependent phenomenon (i.e., higher alcohol intake triggers the neuroadaptive phenomena which then generate the physiologic state of dependence on alcohol) (Li et al. 2007). This dose-dependent, neuroadaptive phenomenon of "addiction" has been illustrated in animals in terms of alcohol tolerance (Kalant et al. 1971), alcohol withdrawal hyperexcitability (Ritzmann and Tabakoff 1976) and withdrawal-induced enhancement of alcohol consumption (relapse drinking) (Melendez et al. 2006). Therefore, one can, and should, consider the quantitative aspects of alcohol consumption as an important predisposing factor for alcohol dependence both in humans and other animals.

Studies with human twins have demonstrated a higher concordance in levels of alcohol consumption among monozygotic twins than dizygotic twins (Whitfield et al. 2004), and studies with animals have clearly shown that one can breed for differences in levels of voluntary alcohol intake in free choice situations (Grahame et al. 1999; McBride and Li 1998). Such data illustrate the fact that not only alcohol dependence, but the propensity to imbibe ethanol, has a heritable (genetic) component. The quantitative phenotype of alcohol drinking or "alcohol preference" can be measured in non-alcohol dependent animals such as mice and rats, and the genetic determinants of such behavior can be explored using currently available genetic, genomic, statistical and informatics techniques (Saba et al. 2006). Very often, the phenotype that is measured in non-human animals is alcohol consumption in a two-bottle choice paradigm, in which the animal is given a choice between various concentrations of alcohol and water, either for a limited time, or with 24-hour access, for several days or weeks (e.g., Rodriguez et al. 1994; Wahlsten et al. 2006). The measured phenotypes, which have been found to be heritable (Grahame et al. 1999; McBride and Li 1998), are either alcohol consumption (e.g., g/kg/24 hr) or alcohol preference, the ratio of alcohol to total fluid consumed. To identify genetic elements that influence the amount of non-dependent alcohol drinking by animals, we focused a genomic analysis on three types of animal populations known to display substantial variation in alcohol consumption: selectively bred, high and low alcohol-preferring mice (HAP and LAP); recombinant inbred mice (BxD RI strains); and inbred strains of mice. Our goal was to ascertain common candidate genes in the three populations which, in a quantitative way, may contribute to relatively low or high voluntary alcohol intake. We used a meta-analysis to pool the results, and utilized our previously developed methods of filtering differentially expressed genes through behavioral QTLs (bQTLs) and expression QTLs (eQTLs) (Saba et al. 2006) to narrow the focus to several likely candidates influencing alcohol drinking behavior. We also used data available in the Allen Brain Atlas (Lein et al. 2007) to define the regional localization of the differentially expressed genes, and confirmed the differences in expression by comparing data from two microarray platforms.

Materials and methods

Mice

Gene expression data were generated from brain tissue of two replicate sets of selected mouse lines, 20 inbred mouse strains, and 30 BXD recombinant inbred (RI) mouse strains. All mice were male, 70–90 days old. The two replicate lines of high alcohol-preferring (HAP) and low alcohol-preferring (LAP) mice were selectively bred in independent experiments, beginning with HS/Ibg mice (Grahame et al. 1999). Selection was based on levels of alcohol consumption in a 24 hr access two-bottle choice paradigm (10% ethanol vs. water) (Grahame et al. 1999). Six mice from each replicate line (HAP1/LAP1, generation 24; HAP2/LAP2, generation 19) were included in the gene expression analysis.

Mice from 30 BXD RI strains (BXD1, BXD2, BXD5, BXD6, BXD8, BXD9, BXD11, BXD12, BXD13, BXD14, BXD15, BXD16, BXD18, BXD19, BXD21, BXD22, BXD23, BXD24, BXD27, BXD28, BXD29, BXD31, BXD32, BXD33, BXD34, BXD36, BXD38, BXD39,

BXD40, BXD42) and 20 inbred strains (129/svlmJ, 129P3/J, A/J, AKR/J, BALB/cByJ, BALB/cJ, BTBR T+ tf/J, C3H/HeJ, C57BL/6J, C58/J, CAST/EiJ, CBA/J, DBA/2J, FVB/NJ, KK/HlJ, MOLF/EiJ, NOD/LtJ, NZW/LacJ, PWD/PhJ, SJL/J) were obtained from The Jackson Laboratory (Bar Harbor, ME). The BXD RI strains are derived from C57BL/6J (C57) and DBA/2J (DBA) inbred strains. Four to seven mice per strain were analyzed for brain gene expression levels.

The use of inbred strains allows for various measures over time on essentially the identical genetic background in each strain. Recent data indicate that the phenotypes of alcohol consumption and preference have been maintained in the same rank order across inbred strains for the past 40–50 years (Wahlsten et al. 2006). The effect of selection of the HAP and LAP mice for alcohol preference had plateaued at the time we performed our studies (Grahame et al. 1999), and measurements of brain gene expression and alcohol preference were performed on mice of the same generation of selection. It is important to note that none of the mice used for these studies had been exposed to ethanol.

Total RNA Extraction

All RNA extractions were done from whole brains of naïve (no alcohol exposure) mice. CO_2 exposure was used to sacrifice the mice. Whole brains were removed and frozen on dry ice, and kept at −70°C until RNA extraction. The RNeasy midi kit (Qiagen, Valencia, CA) was used to extract total RNA and the RNeasy mini kit (Qiagen) was used for clean-up. RNA from whole brain of each individual mouse was hybridized to a separate microarray.

Expression Analysis (Affymetrix)

Double-stranded cDNA was synthesized from 5 µg of total RNA and was processed for hybridization with the Affymetrix mouse whole-genome oligonucleotide array (MOE430 v2; Affymetrix, Santa Clara, CA) as previously described (Saba et al. 2006).

Following staining and scanning of the arrays (Saba et al. 2006), all images were evaluated in a series of tests for quality control, and only those considered to be of high quality were used in further analyses. A detailed description of the quality control process is outlined in the Supplementary Information of Saba et al. (2006) and is available at http://phenogen.uchsc.edu. As a result of this evaluation, data from a total of 20 arrays were eliminated (5 from the BXD RI strains and 15 from the inbred strains). Some concern has been raised recently that SNPs within the 25mer represented by each probe could influence the hybridization and intensity estimations for a probeset (Walter et al. 2007). Walter et al. (2007) identified 13,292 probes within 6,590 probesets that had a sequence polymorphism between C57BL/6J and DBA/2J mice. These probes were eliminated from all analyses, and entire probesets were eliminated if less than 4 probes contained no SNPs. The Robust Multichip Average (RMA) method was used to normalize the expression values of the perfect match probes and summarize these values in probesets (Saba et al. 2006). This method also includes a log base 2 transformation. Data from each of the four populations studied were normalized separately. For the correlation analysis within the BXD RI strains, only data from the 22 strains, including the two parental strains, that had alcohol consumption data available were included in the normalization. For the correlation analysis within the inbred strains, only the data from the 10 strains that had alcohol consumption data available were included in normalization. For the eQTL analysis, data from all 30 BXD RI strains and the two parental strains were normalized together.

Alcohol Consumption Phenotype

Recombinant Inbred Strains—Rodriguez et al. (1994) measured alcohol consumption using a standard 2-bottle choice paradigm with 23 strains from the BXD RI panel plus the two

parental strains. Data were collected on nine or ten male mice per strain. Alcohol consumption (g/kg) was measured as the average daily amount of ethanol solution consumed (10% ethanol vs. tap water) over a 15-day testing period, divided by the average body weight of the animal.

Inbred Strains—Wahlsten et al. (2006) and Yoneyama et al. (in press) reported alcohol consumption using a 2-bottle choice paradigm for 22 inbred mouse strains, as measured in one of our laboratories (Finn, OHSU). We used raw data in order to limit results to male mice. All mice were between the ages of 8 and 10 weeks at the time of testing. Between five and ten male mice were tested for each strain. Each animal was given the choice between two bottles, one with the ethanol solution and the other with tap water. The concentration of ethanol was increased from 3% to 6% to 10% every four days. The average daily consumption of ethanol solution on day 2 and day 4 of the 10% ethanol exposure period, adjusted for mouse body weight (g/kg), was used in our analyses.

Selected Lines of HAP and LAP Mice—Mice were selected for high or low voluntary alcohol consumption based on 24-hr access to 10% alcohol or water for 30 days. Alcohol consumption (g/kg) per day was averaged over the 30-day period.

Behavioral QTLs

The behavioral QTLs (bQTLs) for alcohol preference were obtained from two literature sources, as previously described (Saba et al. 2006).

Statistical Analyses, Filters, Gene List

Meta-Analysis—Initially, the association of brain gene expression and alcohol consumption was calculated for each of the independent experiments, and for each probeset separately. For the replicate HAP and LAP selected lines, a point biserial correlation was calculated. A point biserial correlation is a method for calculating a correlation coefficient when subjects are from two distinct groups and it is assumed that every subject within a group has the same value for the phenotype (g/kg alcohol consumed). The actual phenotypic values within the group are irrelevant in the calculation of this correlation coefficient, except that HAP mice consume more alcohol than LAP mice. For the BXD RI panel and the inbred strain panel, Pearson product-moment correlation coefficients were calculated using strain means for gene expression and alcohol consumption. Correlation coefficients were then transformed to their corresponding Z values using the Fisher transformation (Field 2001). These transformed values were combined using a weighted average where weights are based on the variance of the Z values within the study (1/(N-3)). Following the method outlined by Hedges and Vevea (1998) for a fixed effects model, the Z-score related to the average Z value and its corresponding p-value were calculated. Raw p-values were adjusted for multiple comparisons using the method for false discovery rates (FDR) outlined by Benjamini and Hochberg (1995). A list of transcripts that showed a significant meta-analysis correlation (unadjusted $p = 7.5 \times 10^{-6}$), with heritabilities passing our filter, but without eQTL/bQTL overlap (see below) is presented in Supplementary Table 2 online.

Expression QTLs—Expression QTLs (eQTLs) were calculated in a similar manner to traditional QTLs, where gene expression is the quantitative trait of interest. Data on whole brain gene expression from 30 BxD RI strains and the parental strains were used for this analysis (http://phenogen.uchsc.edu). Probes were eliminated from consideration if the probe sequence contained a SNP which distinguished C57BL/6 and DBA/2 strains. Entire probesets were eliminated if more than seven probes assigned to the probeset contained such a SNP. Although these probes and probesets were eliminated, it is still possible that probes within our final list of probesets could have SNPs that do not differentiate between C57BL/6J and DBA/2J mice, or SNPs that have yet to be discovered. The data were normalized using RMA, prior to eQTL

analysis. eQTLs were calculated for each probeset individually, using a weighted marker regression analysis, as previously described (see Saba et al. 2006). Strain means were regressed against a set of 943 markers with unique strain distributions that were derived from the set of 3,795 markers available at http://www.genenetwork.org/dbdoc/BXDGeno.html. The location of an eQTL was identified as the location of the maximum LOD score for that probeset. P-values associated with the maximum LOD score for each probeset were calculated using permutation to account for the multiple comparisons across markers (Saba et al. 2006). The number of permutations per transcript was increased until the maximum LOD score from the true data was no longer in the top ten of LOD scores from the permutation, or until 1,000,000 permutations were calculated. For eQTLs with empirical p-values less than 0.10, 95% confidence limits for location were calculated according to the bootstrap method outlined in Visscher et al. (1996), using 1,000 bootstrap samples. All eQTL calculations were carried out with the freeware QTLReaper which is downloadable at http://sourceforge.net/projects/qtlreaper/.

Heritability—Heritability of each probeset was calculated within the BXD RI panel and the inbred panel separately. A broad-sense heritability was calculated using a one-way ANOVA.

Creation of Candidate Gene List—A list of candidate genes underlying the phenotype of alcohol consumption was created using multiple filters. The first step was to limit probesets to those that had an eQTL that overlapped a bQTL for alcohol preference (Saba et al. 2006). As described in our earlier work, the rationale for using this filter is that, if a gene contributes to a complex behavior through variation in its expression levels, areas of the genome that regulate those expression levels (eQTLs), should be represented within areas of the genome that contribute to regulation of the behavior (bQTLs). Transcripts with eQTLs that overlap bQTLs are therefore considered to represent likely candidate genes for the particular behavior. A probeset had to have a significant or suggestive (unadjusted p-value < 0.10) eQTL whose 95% confidence interval for location overlapped one of the bQTL regions to pass this filter. For those probesets with overlapping eQTL and bQTL, an FDR was calculated based on the raw p-values from the meta-analysis of the correlation of gene expression with alcohol consumption. This second filter eliminated probesets with an FDR greater than 0.001. The third filter was implemented to ensure that the expression levels of the genes represented in the candidate gene list not only showed a strong genetic correlation with alcohol consumption, but also had a high biological reproducibility, as measured by heritability of expression intensities. Probesets were eliminated if their expression heritability in either the BXD RI panel or the inbred panel was lower than the median heritability of all probesets in that panel. Finally, the sequence alignment for the remaining probesets was verified through Ensembl. Any probe sequences that were questionable were verified by comparing the probe sequences against data contained in GenBank. All probes used were found to represent sequences contained in the 3'-untranslated regions of the candidate genes.

Proportion of Variance Explained and Model Building—Once a list of candidate genes was established, the proportion of the variance in alcohol consumption among strains that was explained by gene expression was examined for each transcript individually. In addition, a forward step-wise model building procedure was used to determine the best combination of transcripts to jointly describe the strain variation in alcohol consumption with a significance criterion of 0.05 for entrance into the model or elimination from the model.

Analysis of Upstream SNPs in Candidate Genes

All SNPs (and their allele information) that were contained within the 2Kb upstream region of the candidate genes were extracted using the Ensembl BioMart tool based on Mouse Version 46 (August 2007). For each probeset associated with the candidate genes, brain gene expression

data from 20 inbred strains was obtained from the PhenoGen website
(http://phenogen.uchsc.edu)

Depending on the SNP, there was allele information in BioMart available for 2 to 15 of the 20
strains. For the strains with allele information, an "eQTL analysis" was performed, using the
expression data. Since multiple strains may have the same allele, the analysis was based on a
t-test comparing the expression levels of the strains with one allele versus another. SNPs with
p-value <0.01 were considered significant (see Supplementary Table 1 online)

Localization of Significant SNPs in Upstream Regions of Candidate Genes within Transcription Factor Binding Site Motifs

Each of the upstream 2Kb sequences for the candidate genes was scanned for putative
transcription factor binding sites (TFBS) using the TRANSFAC 9.4 database of binding site
motifs and TRANSFAC "Match" software (Matys et al. 2006). Only 522 high quality vertebrate
motifs were included in the search and the cutoffs for determining a putative binding site were
selected based on the TRANSFAC pre-selected cutoffs for minimizing the number of false
positives. After identifying putative TFBS, we examined whether any of the significant SNPs
(see above) were located within the TFBS.

Confirmation of Expression Differences: Expression Analysis using CodeLink Arrays

Double-stranded cDNA was synthesized from 4 µg of total RNA, processed and hybridized
with CodeLink Mouse Whole Genome BioArrays as previously described (Bhave et al.
2006). Data were normalized using two different methods, cyclic locally weighted scatterplot
smoothing (LOESS) and quantile normalization, as outlined in the user manual for the
PhenoGen website (http://phenogen.uchsc.edu) (Bhave et al. 2007).

Results

Meta-analysis of transcripts correlated with alcohol consumption

As discussed in Methods, the approach taken to generate a list of candidate genes was to initially
identify transcripts with a significant eQTL that overlapped a behavioral (b)QTL. Without
taking into account the significance of the correlation of transcript expression with the alcohol
preference phenotype, there were 1523 transcripts from the meta-analysis that passed this filter.
On the other hand, if correlated transcripts were identified, without the bQTL/eQTL overlap
filter, there were 8,782 probesets at an FDR of 0.05. By first limiting our candidate gene list
to those transcripts with bQTL/eQTL overlap, and then requiring an FDR of ≤0.001 for the
correlation with alcohol preference in the meta-analysis, the number of candidate probesets
was reduced to 15. Although the use of the stringent FDR value is likely to result in some false
negatives, we chose to use FDR ≤0.001 in order to focus on the candidate genes most likely
to represent true positives, which can be fruitfully subjected to further investigation. The
heritabilities (Table 3) of each of the 15 candidate probesets passed the filter (i.e., heritability
higher than median heritability [$h^2 = 0.67$ for inbred strains; $h^2 = 0.52$ for BxD RI strains] in
each of the two panels [the BxD RI panel and the inbred strain panel]). From this list of 15
probesets, however, six were excluded because the probe sequence alignments could not be
verified. These filtering procedures resulted in the set of nine probesets shown in Table 1,
which represent eight unique transcripts (there were two probesets for *Gnb1* in the list). The
expression levels of all of the probesets, with the exception of the two probesets for *Gnb1*,
were positively correlated with alcohol consumption. *Gnb1* probeset expression levels were
negatively correlated with levels of alcohol consumption. The chromosomal locations of the
candidate genes, as well as the location of the β and eQTLs, are shown in Table 2. All of the
candidate transcripts were posited to be *cis*-regulated, in that the chromosomal location of the
gene is within 5 Mb of the eQTL location (Table 2). The fact that the expression levels of these

transcripts are *cis*-regulated is consistent with the high heritability of the expression of these transcripts.

The calculated heritabilities of the transcripts are shown in Table 3, as is the proportion of variance in alcohol consumption attributed to each transcript in the BxD RI and inbred strains. In the BxD RI strains, the transcript that explains the greatest proportion of variance in alcohol drinking behavior is the RIKEN cDNA 11100032A03. Other transcripts that independently explain a significant proportion of the variance include *C1qc*, *Scn4b* and *Gnb1*. In the inbred strains, the complement components (*C1qc*, *C1qb*), centrosomal protein (*Cep63*), ubiquitin conjugating enzyme (*Ube2j2*) and the RIKEN cDNA D930028F11 each independently accounted for a significant proportion of variance in the behavioral trait.

We further carried out a forward stepwise regression analysis in order to determine a multivariate model for alcohol consumption in the BxD RI and inbred strains. For the BxD RI strains, the final model contained only the main effect for the RIKEN gene 1110032A03. This final model explained 32% of the total variance between strain means for alcohol consumption. For the inbred strains, the final model contained complement component 1qc (*C1qc*) and the G protein beta 1 subunit (*Gnb1*) and their interaction. This final model explained 96% of the total variance between strain means in alcohol consumption. The fitted model was used to obtain estimates for alcohol consumption for various hypothetical scenarios, in order to illustrate the effect of the interaction between the two transcripts. It was found that mice with higher levels of *C1qc* expression and lower levels of *Gnb1* expression would be predicted to have the highest levels of alcohol consumption.

Brain regional localization of gene expression

Our analysis of gene expression intensities was carried out using whole brain tissue. In order to determine the brain regional localization of expression of the candidate genes for alcohol consumption listed in Table 1, including possible co-localization of these transcripts, we utilized the Allen Brain Atlas (www.brainatlas.org/aba/). This resource provides maps of the expression levels of about 20,000 genes in the mouse brain. Two of the transcripts of interest, *C1qb* and RIKEN cDNA D930028F11, did not display detectable levels of expression in the atlas. Therefore, we do not consider these transcripts in detail in our discussion. However, the other transcripts generally showed heterogeneous distribution of expression intensity in the mouse brain. Table 4 summarizes the levels of intensity in various brain regions for the candidate transcripts shown in Table 1. Many of the transcripts display localized, and co-localized, areas of expression. These results drew our attention in particular to the olfactory regions of the mouse brain, including the olfactory bulb, olfactory tubercle, and piriform cortex, as regions of relatively high expression levels for many of the candidate transcripts. For example, *C1qc* and RIKEN 1110032A03, which explained a significant proportion of the variance for the BXD RI and inbred strains, were both expressed at relatively high levels in the olfactory bulb, along with *Ube2j2* and *Cep63*. The RIKEN transcript, in particular, showed a very localized expression, apparently in mitral cells of the olfactory bulb. In the olfactory tubercle, there were high levels of expression of *Scn4b* and *Cep63*, while *Gnb1* was expressed at high levels in the piriform cortex, along with *Cep63*. Some of the transcripts also showed high levels of expression in regions of the limbic system, including *Scn4b* in the nucleus accumbens (and striatum); *Gnb1* in hippocampus and lateral habenula; *Ube2j2* and *Cep63* in the hippocampus and *Cep63* in the lateral habenula. Figure 1 shows the brain expression patterns of the *Scn4b*, *Cep63* and *C1qc* transcripts, with intensity levels represented by pseudocolor images. The images for the other transcripts can be accessed using the Allen Brain Atlas. The regional and overlapping localization of the candidate transcripts may provide insight into their function in predisposition to alcohol drinking (see Discussion).

Contribution of Sequence Variation to Differential Expression of Candidate Genes

SNPs present in the 2Kb upstream regions of the candidate genes were identified as described in Methods, and their association with differences in candidate gene expression among inbred mouse strains included in the correlation analysis was determined. For most of the candidate genes, there was a significant ($p<0.01$) difference in expression values associated with a large proportion of the identified SNPs (Supplementary Table 1 online). This finding is a general trend for probesets for which an eQTL (vs no eQTL) or a *cis* eQTL (vs a *trans* eQTL) can be determined (Supplementary Methods online).

As shown in Table 5, four of the candidate genes contained one or more of these "significant SNPs" within a putative transcription factor binding site. *C1qb* and *Ube2j2* each had two SNPs within a binding site motif. However, there were no SNP-containing transcription factor binding site motifs in common among the candidate genes.

Technical Verification of the Correlated Transcripts

We used a different microarray platform to verify the differential expression of the candidate genes. As shown in Table 6, comparing data from the Affymetrix arrays to data from the CodeLink arrays, for all of the transcripts except *Ube2j2*, the differences in expression levels in brains of C57BL/6 (high alcohol consumption) and DBA/2 (low alcohol consumption) mice were in the same direction, and were statistically significant using data from both array platforms. Although the difference in expression of *Scn4b* was only marginally significant using data from the CodeLink Arrays, this transcript was identified previously in a more limited meta-analysis performed by our group (Mulligan et al. 2006), and is therefore included in our current discussion. With regard to *Ube2j2*, the data from the CodeLink array did not show even a marginally significant difference in expression levels between the two strains of mice. Upon further review of the Affymetrix data, and we found that there was another probeset for *Ube2j2* on the Affymetrix array (1426138_a_at). Although the expression levels obtained with this probeset did show a positive correlation with alcohol consumption (meta-analysis correlation coefficient, 0.42, $p<0.003$), there was no significant eQTL associated with this probeset, so it did not pass our original filter. The sequences of the probe on the CodeLink array, and of the probes in probeset 1426138_a_at on the Affymetrix array, were targeted to the last exon of the *Ube2j2* transcript. On the other hand, the sequence of the probes in the probeset for *Ube2j2* for which we found a significant eQTL that overlapped with a bQTL, and for which the expression data correlated significantly with alcohol consumption (1430317_at), were targeted to an intronic region between exons 1 and 2, according to NCBI mouse build 37. It seems unlikely that we would detect expression of an intronic region of DNA in a transcript, or be able to determine a *cis* eQTL for that expression. Furthermore, the same mouse build indicates that exon two of *Ube2j2* is an alternatively spliced exon, and that both exon 1 and exon 2 are untranslated. Therefore, while *Ube2j2* may in fact be an important candidate gene for alcohol consumption, these findings inject some ambiguity into the results for this particular transcript.

The Affymetrix array included two or more probesets for some of the other candidate genes, i.e., *C1qb*, *Gnb1* and *Cep63*. The expression levels measured by the other probesets for *C1qb* correlated positively with alcohol consumption in BXD RI mice, inbred mice, and one line of HAP and LAP mice (but there was a negative correlation in the other line), and displayed eQTLs in the same region of chromosome 4 as the reported probeset. However, the FDR value for the meta-analysis did not pass the filter for either of these probesets. A second probeset for *Cep63*, and a third probeset for *Gnb1*, showed very low expression values (close to background).

Discussion

We have previously described an approach for pursuing candidate gene searches utilizing genetic and genomic methodology (Saba et al. 2006). With quantitative phenotypes, QTL analysis has been posited to identify the areas of an organism's genome which contain the genetic material responsible for a portion of the variation in the trait of interest (Lander and Botstein 1989). Therefore, our approach rests on the premise that, if differences in gene expression levels are important for influencing a quantitative phenotype, then the genetic material that contributes to the differential levels of mRNAs in organs (e.g., brain) resides within the QTLs ascertained for the quantitative trait (Chesler et al. 2005; Saba et al. 2006). This systematic approach to candidate gene identification can be generally applied to various species and complex traits, for example, by using the gene expression data and information on genomic locations of eQTLs available at http://phenogen.uchsc.edu. Our previous study of alcohol consumption/preference that used this approach (Saba et al. 2006) was limited to the selected lines of HAP and LAP mice, with confirmation of some candidate transcripts in BXD RI mice. To broaden the applicability of the approach, by including more genetic diversity, in the current study we utilized a meta-analysis of data from selectively bred animals, recombinant inbred strains and a large panel of inbred mouse strains, to assess relationships between the quantitative measures of an alcohol preference phenotype (see Methods) and the quantitative measures of brain gene expression. The meta-analysis that included the correlation of brain mRNA levels with levels of alcohol consumption in the BxD RI mice, the inbred mice, and the HAP and LAP selected lines of mice, resulted in identification 8,782 correlated transcripts in common at an FDR of 0.05. However, imposing the "filter" that required the expression (e) QTLs for candidate genes to reside in the behavioral QTLs for alcohol preference, as well as our other filters (see Results), reduced the number of candidate genes substantially. This result contrasts with the previous meta-analysis, in which we participated, of alcohol preference and brain gene expression, which utilized data from selected lines of mice and a small number of inbred strains (Mulligan et al., 2006). This prior study did not systematically apply the eQTL/ bQTL filter, and identified nearly 4000 correlated transcripts.

The eight candidate genes that we identified are all *cis*-regulated, defined by the eQTL being localized within 5 Mb of the physical location of the gene. This finding suggests that genetic variation in the region of the gene is responsible for the differential expression of the mRNA. Although sequence information is not yet available for all mouse strains used in our studies, we noted a number of SNPs that were located within the 2Kb upstream region of the candidate genes. Some of these SNPs are located within putative transcription factor binding sites in the upstream regions of the genes. These data, particularly the fact that there were no common transcription factor binding sites among the candidate genes, support the conclusion that the expression of the candidate genes may be subject to *cis*-regulation. Our findings are compatible with recent work indicating that about 25% of expressed genes exhibit *cis* sequence effects, and that these effects can account for up to 30% of variation in gene expression (Bergen et al. 2007).

One has to stress at this point that our studies are aimed at ascertaining the identity of genetic material which predisposes a mouse to select or reject ethanol solutions in a two-bottle choice situation. None of the animals in which gene expression or genetic markers were measured were ever exposed to ethanol. The information on alcohol consumption was generated from other animals of the same inbred strains or the same selected lines.

The amalgamation of a list of candidate genes for a particular trait may lead to a functional explanation of the etiology of the trait, if the function of each of the genes and their interactions are, *a priori*, well understood. However, we found that linking the products of the candidate genes that we had identified into biochemical pathways relevant to alcohol preference was

difficult without a reference framework. We decided to use information on the regional localization of gene expression in brain as our framework. Thus, we combined general knowledge of the function of a particular anatomical area of brain, our generated knowledge of differentially expressed genes related to alcohol preference, and the brain regional localization of expression of these genes. This combination of information produced insights into the function of the "candidate genes", and helped to explain the involvement of the candidate genes in predisposition to the phenotype of alcohol preference.

When we examined the brain location of the expression of the eight candidate genes utilizing the data from the Allen Brain Atlas, six of the eight genes had discernible regional expression. It became obvious that most of the candidate genes, with the notable exception of *Scn4b*, were expressed primarily in olfactory regions of mouse brain, including the olfactory bulb, the olfactory tubercle and the piriform (olfactory) cortex. *Scn4b* mRNA was, on the other hand, highly represented in limbic areas such as the nucleus accumbens, the amygdala and the ventral striatum, as well as the olfactory tubercle and hippocampus. The hippocampus was an area of concentration of the mRNA for the majority of the candidate genes as well as the mRNA of *Scn4b*. A simplistic view of the functional importance of the brain areas in which the differentially expressed candidate genes reside, would be the involvement of these brain areas in odor detection and discrimination, in determining the emotional valence and intensity of odors (and other stimuli), and in memory formation and storage (Anderson et al. 2003; Chaillan et al. 1996; Dodd 1991). When one considers further the characteristics of the olfactory areas of brain and the hippocampus, one is struck by the fact that these brain areas are highly malleable in terms of neuronal organization and reorganization, the generation of new synapses, and the incorporation of newly migrated and differentiated neural stem cells into the anatomy of these brain areas (Gould 2007). The interpretation of how the identified candidate genes can participate in generation of alcohol preference becomes more tractable in the context of the function of the olfactory system and its connections within the brain.

The olfactory system is unusual in that olfactory neurons, which are located in the nasal epithelium, are replaced throughout the life of the mature animal (Dodd 1991). These neurons extend their axons into the central nervous system and form synapses with mitral cells in the olfactory bulb. The cells in the olfactory bulb must continually accommodate to the formation of new synapses. The axons of the cells in the olfactory bulb project to the olfactory cortex, which includes the olfactory tubercle, the piriform cortex and the entorhinal cortex, which in turn projects to the hippocampus. Neurons in the olfactory tubercle project to the thalamus, and there are also olfactory pathways to the amygdala. Plasticity in the piriform cortex has been implicated in the learning of odor discrimination (Chaillan et al. 1996). Furthermore, a recent paper demonstrates integration of the activity of the olfactory bulb and the electrophysiologic activity in the ventral hippocampus during acquisition of odor discriminations, implicating the hippocampus as well as the piriform cortex in odor discrimination learning (Martin et al. 2007).

If one considers what is known about the functions of the identified candidate genes, while taking into account their regional localization and co-localization in brain, a picture emerges that draws attention to various aspects of neuronal migration, neuronal differentiation and synaptic reorganization that may be important for olfactory learning processes. For example, the complement component, *C1qc*, appears to be most intensely expressed in the olfactory bulb (Table 4). The various isoforms of integrin have been reported to be receptors for *C1q* (Zutter and Edelson 2007) and there is co-localization of integrin beta1 and *C1qc* expression in the olfactory bulb that suggests potential neural functions for *C1qc*. Integrins (and possibly, by association, C1qc) have been implicated in alteration of synaptic strength that affects olfactory learning and memory (Connolly and Tully 1998), *C1qc*, in addition to its interactions with integrin, is a ligand for proteins known as pentraxins (Mantovani et al. 2007). A novel member

of the pentraxin family, neuronal activity-regulated pentraxin (Narp), has been shown to promote the growth of dendritic processes in neurons in cortical explants (Tsui et al. 1996). In addition to Narp, neuronal pentraxin 1 (NP1) and the neuronal pentraxin receptor (NPR), contribute to synaptogenesis by affecting AMPA receptor targeting to post-synaptic membranes (Sia et al. 2007). Liu et al.(2006) reported on polymorphisms in the 3' region of the neurexin 3 (*Nrxn*3) gene which are associated with alcohol and drug abuse. Neurexin 3 functionally resembles pentraxins, particularly in its ability to promote clustering of AMPA receptors in hippocampus (Gerrow and El-Husseini 2007). Hishimoto et al. (2007) have reported that the SNPs associated with alcoholism may contribute to differential expression of the isoforms of neurexins, and particularly *Nrxn3*, in brains of human alcoholics.

Another candidate transcript that is expressed in regions associated with olfactory function, including the olfactory bulb, olfactory tubercle and piriform cortex, is the centrosome protein, *Cep63*. The positioning of the centrosome is important for the division of neuronal progenitor cells, development of polarity and neuron precursor migration (de Anda et al. 2005; Higginbotham and Gleeson 2007). Although there is little information available specifically regarding the role of CEP63, the localization of its expression in the olfactory system and hippocampus, which both display neurogenesis in the adult brain, suggest a role in the neuron replacement/reorganization process. Interestingly, mice and humans with loss-of-function mutations in the cilia-centrosomal protein CEP290 display olfactory dysfunction (McEwen et al. 2007).

The G protein beta subunit, GNB1, is expressed in many regions of brain (Liang et al. 1998), and transcript is found in the olfactory bulb and olfactory tubercle, with higher levels of expression in piriform and prefrontal cortex and hippocampus (Table 4). It is important to note that *Gnb1* expression was negatively correlated with alcohol consumption, in contrast to the other candidate genes. A lower amount of GNB1 could contribute to higher levels of free Gαo, which has been reported to promote neurite outgrowth in Neuro2A cells by a pathway that includes ubiquitination (He et al. 2006). The co-localization of expression of *Gnb1* and *Ube2j2*, a ubiquitin-conjugating enzyme, in olfactory bulb and hippocampus, is compatible with their inclusion in signaling pathways leading to neurite outgrowth, and thus contributing to neurogenesis and neuroplasticity.

The sodium channel beta 4 subunit, SCN4b, has also been linked to neurite outgrowth, both in Neuro2A cells and hippocampal neurons (Oyama et al. 2006). A major role for this subunit is to produce voltage-gated sodium channels with "resurgent" kinetics, which are specialized for high-frequency firing (Grieco et al. 2005). High-frequency firing is necessary for inducing long-term potentiation, thought to be associated with plasticity and learning (e.g., Fedulov et al. 2007). There is also significant current interest in the opposing effects of the G protein β subunits and the auxiliary sodium channel proteins, such as the product of *Scn4b*, on neuronal firing. Mantegazza et al. (2005) demonstrated that Gβ subunits (β$_2$), alone or in combination with α subunits, can maintain the voltage gated sodium channels in a persistent open state which reduces repetitive firing. GNB1 protein resembles GNB2 protein in this functional effect on the voltage sensitive sodium channels (Ma et al. 1997). An increased level of *Scn4b* transcript, in concert with reduced levels of *Gnb1* mRNA in brains of alcohol-preferring mice, could promote repeated firing, LTP formation and "learning" in limbic areas of brain, as well as olfactory tubercle (Chiang and Strowbridge 2007; Goldin 2003), where expression of *Scn4b* and *Gnb1* overlap.

The role of perception of flavor – which involves taste, olfaction and chemosensory information – as an influence on initial alcohol consumption has often been discussed. Individual and strain differences in the ability to discriminate tastes, in addition to post-ingestive consequences, have been suggested to affect the level of alcohol consumption displayed by mice and rats

(Bachmanov et al. 2003; Belknap et al. 1977; Kampov-Polevoy et al. 1990), as well as the reinstatement of ethanol-seeking behavior of rats (Maccioni et al. 2007). More recently, evidence from a number of diverse studies is accumulating for a role of the olfactory system in alcohol preference in animals and humans. For example, studies with an antagonist of the metabotropic glutamate receptor subtype 5 (mGluR5) suggested concomitant effects on alcohol consumption and on the systems involved in the processing of olfactory stimuli (Cowen et al. 2005; Schroeder et al. 2005). In human alcoholics, smelling alcohol has been found to induce self-reported alcohol craving, as well as salivary and other autonomic responses (Newlin et al. 1989; Reid et al. 2006). Thus, one has to consider that, even though alcohol may be consumed for its post-ingestive, reinforcing effects, the two-bottle choice voluntary alcohol consumption phenotype in mice, which is the basis for our studies, undoubtedly includes a salient orosensory component.

Our results demonstrate the validity of using whole brain gene expression data in mice, in combination with QTL filters and information about the regional localization of gene expression in brain, to draw novel conclusions regarding the transcriptional networks that are associated with a complex trait such as alcohol preference. Our work may apply to humans as well as mice. Although QTL regions for alcohol drinking in humans are not well-defined (e.g., Bergen et al. 2003), it may be noteworthy that several QTLs on human chromosome 1 have been suggested both for alcohol drinking and alcohol dependence. Four of the identified candidate genes are localized on the distal end of the q arm of human chromosome 1 (*Gnb1*, *C1qb*, *C1qc* and *Ube2j2*), near at least one of these human QTLs (Guerrini et al. 2005). The identification of the role of the olfactory system that can be derived from the interactions of the candidate genes is consistent with previous work, but also generates new knowledge regarding the key role of this sensory system in mediating differences in dose-dependent alcohol consumption that may influence the development of alcohol dependence.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

American Psychiatric Associaton. Diagnostic and Statistical Manual of Mental Disorders. 4 ed. Washington, DC: American Psychiatric Association; 1994.

Anderson AK, Christoff K, Stappen I, Panitz D, Ghahremani DG, Glover G, Gabrieli JD, Sobel N. Dissociated neural representations of intensity and valence in human olfaction. Nat Neurosci 2003;6:196–202. [PubMed: 12536208]

Bachmanov AA, Kiefer SW, Molina JC, Tordoff MG, Duffy VB, Bartoshuk LM, Mennella JA. Chemosensory factors influencing alcohol perception, preferences, and consumption. Alcohol Clin Exp Res 2003;27:220–231. [PubMed: 12605071]

Belknap JK, Belknap ND, Berg JH, Coleman R. Preabsorptive vs. postabsorptive control of ethanol intake in C57BL/6J and DBA/2J mice. Behav Genet 1977;7:413–425. [PubMed: 603475]

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. JR Statist Soc B 1995;57:289–300.

Bergen AW, Baccarelli A, McDaniel TK, Kuhn K, Pfeiffer R, Kakol J, Bender P, Jacobs K, Packer B, Chanock SJ, Yeager M. Cis sequence effects on gene expression. BMC Genomics 2007;8:296. [PubMed: 17727713]

Bergen AW, Yang XR, Bai Y, Beerman MB, Goldstein AM, Goldin LR. Genomic regions linked to alcohol consumption in the Framingham Heart Study. BMC Genet 2003;4:S101. [PubMed: 14975169]

Bhave SV, Hoffman PL, Lassen N, Vasiliou V, Saba L, Deitrich RA, Tabakoff B. Gene array profiles of alcohol and aldehyde metabolizing enzymes in brains of C57BL/6 and DBA/2 mice. Alcohol Clin Exp Res 2006;30:1659–1669. [PubMed: 17010133]

Bhave SV, Hornbaker C, Phang TL, Saba L, Lapadat R, Kechris K, Gaydos J, McGoldrick D, Dolbey A, Leach S, Soriano B, Ellington A, Ellington E, Jones K, Mangion J, Belknap JK, Williams RW, Hunter LE, Hoffman PL, Tabakoff B. The PhenoGen informatics website: tools for analyses of complex traits. BMC Genet 2007;8:59. [PubMed: 17760997]

Chaillan FA, Roman FS, Soumireu-Mourat B. Modulation of synaptic plasticity in the hippocampus and piriform cortex by physiologically meaningful olfactory cues in an olfactory association task. J Physiol Paris 1996;90:343–347. [PubMed: 9089510]

Chesler EJ, Lu L, Shou S, Qu Y, Gu J, Wang J, Hsu HC, Mountz JD, Baldwin NE, Langston MA, Threadgill DW, Manly KF, Williams RW. Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. Nat Genet 2005;37:233–242. [PubMed: 15711545]

Chiang E, Strowbridge BW. Diversity of neural signals mediated by multiple burst-firing mechanisms in rat olfactory tubercle neurons. J Neurophysiol 2007;98:2716–2728. [PubMed: 17855583]

Connolly JB, Tully T. Integrins: a role for adhesion molecules in olfactory memory. Curr Biol 1998;8:R386–R389. [PubMed: 9635185]

Cowen MS, Djouma E, Lawrence AJ. The metabotropic glutamate 5 receptor antagonist 3-[(2-methyl-1,3-thiazol-4-yl)ethynyl]-pyridine reduces ethanol self-administration in multiple strains of alcohol-preferring rats and regulates olfactory glutamatergic systems. J Pharmacol Exp Ther 2005;315:590–600. [PubMed: 16014750]

de Anda FC, Pollarolo G, Da Silva JS, Camoletto PG, Feiguin F, Dotti CG. Centrosome localization determines neuronal polarity. Nature 2005;436:704–708. [PubMed: 16079847]

Dodd, J.; Castellucci, VF. Smell and Taste: The Chemical Senses. In: Kandel, ER.; Schwartz, JH.; Jessell, TM., editors. Principles of neural science. New York: Elsevier; 1991. p. 512-529.

Enoch MA, Goldman D. The genetics of alcoholism and alcohol abuse. Curr Psychiatry Rep 2001;3:144–151. [PubMed: 11276410]

Fedulov V, Rex CS, Simmons DA, Palmer L, Gall CM, Lynch G. Evidence that long-term potentiation occurs within individual hippocampal synapses during learning. J Neurosci 2007;27:8031–8039. [PubMed: 17652593]

Field AP. Meta-analysis of correlation coefficients: a Monte Carlo comparison of fixed- and random-effects methods. Psychol Methods 2001;6:161–180. [PubMed: 11411440]

Gerrow K, El-Husseini A. Receptors look outward: revealing signals that bring excitation to synapses. Sci STKE 2007 2007:pe56.

Goldin AL. Mechanisms of sodium channel inactivation. Curr Opin Neurobiol 2003;13:284–290. [PubMed: 12850212]

Gould E. How widespread is adult neurogenesis in mammals? Nat Rev Neurosci 2007;8:481–488. [PubMed: 17514200]

Grahame NJ, Li TK, Lumeng L. Selective breeding for high and low alcohol preference in mice. Behav Genet 1999;29:47–57. [PubMed: 10371758]

Grieco TM, Malhotra JD, Chen C, Isom LL, Raman IM. Open-channel block by the cytoplasmic tail of sodium channel beta4 as a mechanism for resurgent sodium current. Neuron 2005;45:233–244. [PubMed: 15664175]

Guerrini I, Cook CC, Kest W, Devitgh A, McQuillin A, Curtis D, Gurling HM. Genetic linkage analysis supports the presence of two susceptibility loci for alcoholism and heavy drinking on chromosome 1p22.1–11.2 and 1q21.3–24.2. BMC Genet 2005;6:11. [PubMed: 15740611]

He JC, Neves SR, Jordan JD, Iyengar R. Role of the Go/i signaling network in the regulation of neurite outgrowth. Can J Physiol Pharmacol 2006;84:687–694. [PubMed: 16998532]

Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. Psych Methods 1998;3:486–504.

Higginbotham HR, Gleeson JG. The centrosome in neuronal development. Trends Neurosci 2007;30:276–283. [PubMed: 17420058]

Hishimoto A, Liu QR, Drgon T, Pletnikova O, Walther D, Zhu XG, Troncoso JC, Uhl GR. Neurexin 3 polymorphisms are associated with alcohol dependence and altered expression of specific isoforms. Hum Mol Genet. 2007

Kalant H, LeBlanc AE, Gibbins RJ. Tolerance to, and dependence on, some non-opiate psychotropic drugs. Pharmacol Rev 1971;23:135–191. [PubMed: 4398655]

Kampov-Polevoy AB, Kasheffskaya OP, Sinclair JD. Initial acceptance of ethanol: gustatory factors and patterns of alcohol drinking. Alcohol 1990;7:83–85. [PubMed: 2328091]

Lander ES, Botstein D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 1989;121:185–199. [PubMed: 2563713]

Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ, Chen L, Chen TM, Chin MC, Chong J, Crook BE, Czaplinska A, Dang CN, Datta S, Dee NR, Desaki AL, Desta T, Diep E, Dolbeare TA, Donelan MJ, Dong HW, Dougherty JG, Duncan BJ, Ebbert AJ, Eichele G, Estin LK, Faber C, Facer BA, Fields R, Fischer SR, Fliss TP, Frensley C, Gates SN, Glattfelder KJ, Halverson KR, Hart MR, Hohmann JG, Howell MP, Jeung DP, Johnson RA, Karr PT, Kawal R, Kidney JM, Knapik RH, Kuan CL, Lake JH, Laramee AR, Larsen KD, Lau C, Lemon TA, Liang AJ, Liu Y, Luong LT, Michaels J, Morgan JJ, Morgan RJ, Mortrud MT, Mosqueda NF, Ng LL, Ng R, Orta GJ, Overly CC, Pak TH, Parry SE, Pathak SD, Pearson OC, Puchalski RB, Riley ZL, Rockett HR, Rowland SA, Royall JJ, Ruiz MJ, Sarno NR, Schaffnit K, Shapovalova NV, Sivisay T, Slaughterbeck CR, Smith SC, Smith KA, Smith BI, Sodt AJ, Stewart NN, Stumpf KR, Sunkin SM, Sutram M, Tam A, Teemer CD, Thaller C, Thompson CL, Varnam LR, Visel A, Whitlock RM, Wohnoutka PE, Wolkey CK, Wong VY, Wood M, Yaylaoglu MB, Young RC, Youngstrom BL, Yuan XF, Zhang B, Zwingman TA, Jones AR. Genome-wide atlas of gene expression in the adult mouse brain. Nature 2007;445:168–176. [PubMed: 17151600]

Li TK, Hewitt BG, Grant BF. The Alcohol Dependence Syndrome, 30 years later: a commentary. the 2006 H. David Archibald lecture. Addiction 2007;102:1522–1530. [PubMed: 17680851]

Liang JJ, Cockett M, Khawaja XZ. Immunohistochemical localization of G protein beta1, beta2, beta3, beta4, beta5, and gamma3 subunits in the adult rat brain. J Neurochem 1998;71:345–355. [PubMed: 9648884]

Liu QR, Drgon T, Johnson C, Walther D, Hess J, Uhl GR. Addiction molecular genetics: 639,401 SNP whole genome association identifies many "cell adhesion" genes. Am J Med Genet B Neuropsychiatr Genet 2006;141:918–925. [PubMed: 17099884]

Ma JY, Catterall WA, Scheuer T. Persistent sodium currents through brain sodium channels induced by G protein betagamma subunits. Neuron 1997;19:443–452. [PubMed: 9292732]

Maccioni P, Orrú A, Korkosz A, Gessa GL, Carai MA, Colombo G, Bienkowski P. Cue-induced reinstatement of ethanol seeking in Sardinian alcohol-preferring rats. Alcohol 2007;41:31–39. [PubMed: 17452297]

Mantegazza M, Yu FH, Powell AJ, Clare JJ, Catterall WA, Scheuer T. Molecular determinants for modulation of persistent sodium current by G-protein betagamma subunits. J Neurosci 2005;25:3341–3349. [PubMed: 15800189]

Mantovani A, Garlanda C, Doni A, Bottazzi B. Pentraxins in Innate Immunity: From C-Reactive Protein to the Long Pentraxin PTX3. J Clin Immunol. 2007

Martin C, Beshel J, Kay LM. An olfacto-hippocampal network is dynamically involved in odor-discrimination learning. J Neurophysiol 2007;98:2196–2205. [PubMed: 17699692]

Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull M, Hornischer K, Voss N, Stegmaier P, Lewicki-Potapov B, Saxel H, Kel AE, Wingender E. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res 2006;34:D108–D110. [PubMed: 16381825]

McBride WJ, Li TK. Animal models of alcoholism: neurobiology of high alcohol-drinking behavior in rodents. Crit Rev Neurobiol 1998;12:339–369. [PubMed: 10348615]

McEwen DP, Koenekoop RK, Khanna H, Jenkins PM, Lopez I, Swaroop A, Martens JR. Hypomorphic CEP290/NPHP6 mutations result in anosmia caused by the selective loss of G proteins in cilia of olfactory sensory neurons. Proc Natl Acad Sci U S A 2007;104:15917–15922. [PubMed: 17898177]

Melendez RI, Middaugh LD, Kalivas PW. Development of an alcohol deprivation and escalation effect in C57BL/6J mice. Alcohol Clin Exp Res 2006;30:2017–2025. [PubMed: 17117967]

Mulligan MK, Ponomarev I, Hitzemann RJ, Belknap JK, Tabakoff B, Harris RA, Crabbe JC, Blednov YA, Grahame NJ, Phillips TJ, Finn DA, Hoffman PL, Iyer VR, Koob GF, Bergeson SE. Toward understanding the genetics of alcohol drinking through transcriptome meta-analysis. Proc Natl Acad Sci U S A (3) 2006;103:6368–6373. [PubMed: 16618939]

Newlin DB, Hotchkiss B, Cox WM, Rauscher F, Li TK. Autonomic and subjective responses to alcohol stimuli with appropriate control stimuli. Addict Behav 1989;14:625–630. [PubMed: 2618845]

Oyama F, Miyazaki H, Sakamoto N, Becquet C, Machida Y, Kaneko K, Uchikawa C, Suzuki T, Kurosawa M, Ikeda T, Tamaoka A, Sakurai T, Nukina N. Sodium channel beta4 subunit: down-regulation and possible involvement in neuritic degeneration in Huntington's disease transgenic mice. J Neurochem 2006;98:518–529. [PubMed: 16805843]

Reid MS, Flammino F, Starosta A, Palamar J, Franck J. Physiological and subjective responding to alcohol cue exposure in alcoholics and control subjects: evidence for appetitive responding. J Neural Transm 2006;113:1519–1535. [PubMed: 16604310]

Ritzmann RF, Tabakoff B. Body temperature in mice: a quantitative measure of alcohol tolerance and physical dependence. J Pharmacol Exp Ther 1976;199:158–170. [PubMed: 988176]

Rodriguez LA, Plomin R, Blizard DA, Jones BC, McClearn GE. Alcohol acceptance, preference, and sensitivity in mice. I. Quantitative genetic analysis using BXD recombinant inbred strains. Alcohol Clin Exp Res 1994;18:1416–1422. [PubMed: 7695038]

Saba L, Bhave SV, Grahame N, Bice P, Lapadat R, Belknap J, Hoffman PL, Tabakoff B. Candidate genes and their regulatory elements: alcohol preference and tolerance. Mamm Genome 2006;17:669–688. [PubMed: 16783646]

Schroeder JP, Overstreet DH, Hodge CW. The mGluR5 antagonist MPEP decreases operant ethanol self-administration during maintenance and after repeated alcohol deprivations in alcohol-preferring (P) rats. Psychopharmacology (Berl) 2005;179:262–270. [PubMed: 15717208]

Sia GM, Beique JC, Rumbaugh G, Cho R, Worley PF, Huganir RL. Interaction of the N-terminal domain of the AMPA receptor GluR4 subunit with the neuronal pentraxin NP1 mediates GluR4 synaptic recruitment. Neuron 2007;55:87–102. [PubMed: 17610819]

Tsui CC, Copeland NG, Gilbert DJ, Jenkins NA, Barnes C, Worley PF. Narp, a novel member of the pentraxin family, promotes neurite outgrowth and is dynamically regulated by neuronal activity. J Neurosci 1996;16:2463–2478. [PubMed: 8786423]

Visscher PM, Thompson R, Haley CS. Confidence intervals in QTL mapping by bootstrapping. Genetics 1996;143:1013–1020. [PubMed: 8725246]

Wahlsten D, Bachmanov A, Finn DA, Crabbe JC. Stability of inbred mouse strain differences in behavior and brain size between laboratories and across decades. Proc Natl Acad Sci U S A 2006;103:16364–16369. [PubMed: 17053075]

Walter NA, McWeeney SK, Peters ST, Belknap JK, Hitzemann R, Buck KJ. SNPs matter: impact on detection of differential expression. Nat Methods 2007;4:679–680. [PubMed: 17762873]

Whitfield JB, Zhu G, Madden PA, Neale MC, Heath AC, Martin NG. The genetics of alcohol intake and of alcohol dependence. Alcohol Clin Exp Res 2004;28:1153–1160. [PubMed: 15318113]

World Health Organization. International Statistical Classification of Diseases and Health Related Problems. 2nd edGeneva: (The) ICD-10; 2005.

Yoneyama N, Crabbe JC, Ford MM, Murillo A, Finn DA. Voluntary ethanol consumption in 22 inbred mouse strains. Alcohol. (in press)

Zutter MM, Edelson BT. The alpha2beta1 integrin: a novel collectin/C1q receptor. Immunobiology 2007;212:343–353. [PubMed: 17544819]
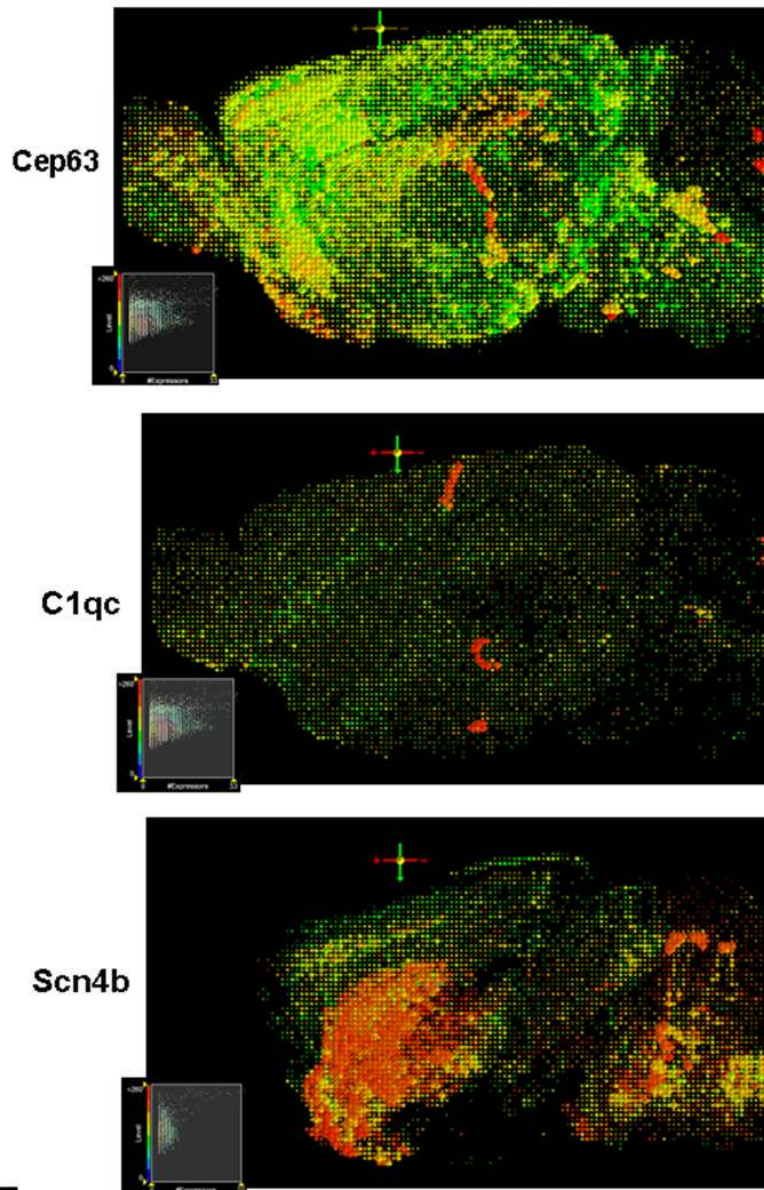
Figure 1. Brain regional localization of candidate transcripts
These figures are taken from the Allen Brain Atlas (Lein et al. 2007), and display levels of mRNA for the indicated transcripts as determined by *in situ* hybridization. The pseudocolor images correspond to the relative density of the signal.

Candidate Genes for Alcohol Preference

Table 1

All genes listed showed a significant correlation with alcohol consumption in the meta-analysis, and had an eQTL that overlapped a bQTL (see Table 2). The gene location, correlation coefficient with alcohol consumption in the BXD RIs, inbred strains and two replicate lines of HAP and LAP mice, as well as p values for these correlations, are shown. The correlation coefficient from the meta-analysis is also shown. All FDR values for the meta-analysis correlations were ≤0.001.

Gene	Gene Symbol	Gene Location: Chr (Mb)	BXD Corr Coef	Inbred Corr Coef	Corr in HAP1/LAP1	Corr in HAP2/LAP2	Meta Analysis Corr Coef
RIKEN cDNA 1110032A03 gene	1110032A03Rik	9 (50.52)	0.56	0.33	0.90	0.58	0.64
guanine nucleotide binding protein, beta 1	Gnb1	4 (154.34)	−0.45	−0.36	−0.81	−0.87	−0.65
guanine nucleotide binding protein, beta 1	Gnb1	4 (154.34)	−0.36	−0.24	−0.82	−0.89	−0.62
complement component 1, q subcomponent, beta polypeptide	C1qb	4 (136.15)	0.38	0.87	0.61	0.77	0.63
complement component 1, q subcomponent, C chain	C1qc	4 (136.16)	0.46	0.91	0.45	0.70	0.63
sodium channel, type IV, beta	Scn4b	9 (44.90)	0.46	0.49	0.84	0.59	0.60
centrosomal protein 63	Cep63	9 (102.44)	0.35	0.78	0.61	0.80	0.60
ubiquitin-conjugating enzyme E2, J2 homolog (yeast)	Ube2j2	4 (154.79)	0.38	0.72	0.50	0.84	0.59
RIKEN cDNA D930028F11 gene	D930028F11Rik	9 (47.91)	0.41	0.80	0.84	0.72	0.66

Table 2

Overlap of Expression and Behavioral QTLs for Candidate Genes for Alcohol Preference

Each candidate gene displayed a significant eQTL that overlapped a bQTL. The gene location, eQTL location and LOD score associated with the eQTL, as well as the p value for the eQTL, are shown.

Gene	Gene Symbol	Gene Location: Chr (Mb)	eQTL Location: Chr (Mb)	eQTL LOD Score	eQTL p value	bQTL Location: Chr (Mb - range)
RIKEN cDNA 1110032A03 gene	1110032A03Rik	9 (50.52)	9 (45.99)	6.17	0.001	9 (31.28 – 111.45)
guanine nucleotide binding protein, beta 1	Gnb1	4 (154.34)	4 (154.35)	17.42	<0.001	4 (128.57 – 154.00)
guanine nucleotide binding protein, beta 1	Gnb1	4 (154.34)	4 (154.35)	12.36	<0.001	4 (128.57 – 154.00)
complement component 1, q subcomponent, beta polypeptide	C1qb	4 (136.15)	4 (135.95)	4.97	0.003	4 (128.57 – 154.00)
complement component 1, q subcomponent, C chain	C1qc	4 (136.16)	4 (135.95)	6.06	<0.001	4 (128.57 – 154.00)
sodium channel, type IV, beta	Scn4b	9 (44.90)	9 (41.24)	11.37	<0.001	9 (31.28 – 111.45)
centrosomal protein 63	Cep63	9 (102.44)	9 (99.68)	3.98	0.012	9 (31.28 – 111.45)
ubiquitin-conjugating enzyme E2, J2 homolog (yeast)	Ube2j2	4 (154.79)	4 (154.70)	9.76	<0.001	4 (128.57 – 154.00)
RIKEN cDNA D930028F11 gene	D930028F11Rik	9 (47.91)	9 (45.99)	9.83	<0.001	9 (31.28 – 111.45)

Table 3

Heritability and Proportion of Variance Contributed by Candidate Genes for Alcohol Preference

The heritability of each transcript and proportion of variance (p-value) in alcohol consumption explained by each transcript were calculated separately within the BXD RI strains and the inbred strains, as described in Methods. All of the candidate genes passed our filters for heritability values (see Methods).

Gene	Gene Symbol	BXD RI Strains		Inbred Strains	
		Heritability	Proportion of Variance (p value)	Heritability	Proportion of Variance of (p value)
RIKEN cDNA 1110032A03 gene	1110032A03Rik	0.83	32% (0.0066)	0.93	11% (0.3468)
guanine nucleotide binding protein, beta 1	Gnb1	0.97	20% (0.0348)	0.97	13% (0.3066)
guanine nucleotide binding protein, beta 1	Gnb1	0.94	13% (0.1040)	0.99	6% (0.4961)
complement component 1, q subcomponent, beta polypeptide	C1qb	0.81	15% (0.0791)	0.83	76% (0.0011)
complement component 1, q subcomponent, C chain	C1qc	0.89	21% (0.0298)	0.90	83% (0.0002)
sodium channel, type IV, beta	Scn4b	0.88	21% (0.0322)	0.95	24% (0.1482)
centrosomal protein 63	Cep63	0.60	12% (0.1125)	0.81	62% (0.0072)
ubiquitin-conjugating enzyme E2, J2 homolog (yeast)	Ube2j2	0.83	15% (0.0780)	0.89	52% (0.0192)
RIKEN cDNA D930028F11 gene	D930028F11Rik	0.88	17% (0.0591)	0.95	65% (0.0050)

Table 4

Brain Regional Localization of Expression of Candidate Genes for Alcohol Preference

Brain Area	C1qc	Gnb1	Scn4b	Ube2j2	1110032A03Rik	Cep63
Olfactory bulb	++	+	0	++	+++[a]	++
Olfactory tubercle	+	+	+++	0	+	+++
Piriform cortex	±	+++	0	±	±	++
Cingulate cortex	0	0	0	0	0	0
Nucleus accumbens	+	0	+++	0	0	++
Striatum	+	0	+++	0	0	0
Hippocampus	+	+++	±	+++[b]	±±	+++[b]
Ventral tegmentum	0	0	0	0	0	0
Prefrontal cortex	0	+++	+	0	0	+
Habenula (lateral)	0	+++	0	0	0	+++
Ventral pallidum	0	++	0	0	0	0

[a] Definitive line of cells (possibly mitral) at the base of olfactory bulb and anterior olfactory nucleus (glomerular area).

[b] Subfields of the hippocampus

Table 5

Sequence Variation in Putative Transcription Factor Binding Sites in Upstream Regions of Candidate Genes

SNPs that displayed an association with expression levels (p<0.01) for the candidate genes ("significant SNPs") were identified as described in Methods. The 2 Kb upstream region of each candidate gene was scanned for putative transcription factor binding sites and the occurrence of a significant SNP within the transcription factor binding site motif is shown (in blue).

Gene Symbol	SNP	Transcription Factor	Transcription Factor Binding SiteMotif (SNP in blue)
1110032A03Rik	NT_039472.6_37357837	HNF4	tgttCTCTGaccg
C1qb	rs27625275	ZID	cGGCTCcatcatg
	rs27625276		
Ube2j2	NT_039268.4_9570463	CAAT	tcATTGGctccg
	NT_039268.4_9570468		
	NT_039268.4_9569748	MMEF	gggtcTAAAAatgcct
D930028F11Rik	NT_039472.6_34749095	CDPCR1	acaaTCTATa
	NT_039472.6_34749453	CDPCR1	gaaaTCGATc
	NT_039472.6_34750261	EVI1	tgataaaaaGAGATa
		GATA2	agaGATAAaa
		CDX2	agagATAAAaatgt
	NT_039472.6_34750326	COMP1	cactttcatagtTAATCttggtt
		HNF1	tagTTAATcttggttagc
	rs30479198	GATA4	AGAAAagaggga

HNF 4/1: hepatocyte nuclear factor 4/1; ZID: zinc finger protein with interaction domain; CAAT: CAAT box; MMEF2: related to myocyte enhancer 2; CDPCR 1: cut-like homeodomain protein CUTL 1; EVI 1: ectopic viral integration site 1 encoded factor; GATA-2/4: GATA binding protein 2/4; CDX2: caudal-related homeobox 2; COMP1: cooperates with myogenic proteins.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 6

Technical Confirmation of Differences in Candidate Gene Expression Levels

	Probe/Probeset ID	Gene Symbol	Riken cDNA Gene Name	C57BL/6J	DBA/2J	p-value
A	1417211_a_at	1110032A03Rik	RIKEN cDNA 1110032A03 gene	9.24 ± 0.11	8.53 ± 0.05	<0.0001
C	GE41487			9.90 ± 0.09	9.11 ± 0.09	<0.0001
A	1417432_a_at	Gnb1	G protein, beta 1 subunit	9.37 ± 0.18	11.06 ± 0.10	<0.0001
A	1454696_at			11.42 ± 0.31	12.76 ± 0.06	<0.0001
C	GE34498			8.87 ± 0.19	9.16 ± 0.18	0.0209
A	1417063_at	C1qb	Complement component 1, q subcomponent, beta polypeptide	9.65 ± 0.15	8.64 ± 0.09	<0.0001
C	GE35132			11.89 ± 0.13	11.17 ± 0.04	<0.0001
A	1449401_at	C1qc	Complement component 1, q subcomponent, C chain	9.69 ± 0.10	8.69 ± 0.06	<0.0001
C	GE42466			11.24 ± 0.09	10.41 ± 0.08	<0.0001
A	1434008_at	Scn4b	Sodium channel, Type IV, beta	11.04 ± 0.16	10.06 ± 0.14	<0.0001
C	GE1402939			12.36 ± 0.05	12.23 ± 0.17	0.0944
A	1438409_at	Cep63	Centrosomal protein 63	8.29 ± 0.09	8.01 ± 0.06	<0.0001
C	GE1403609			10.07 ± 0.06	7.47 ± 0.20	<0.0001
A	1430317_at	Ube2j2	Ubiquitin-conjugating enzyme E2, J2 homolog (yeast)	8.17 ± 0.06	7.86 ± 0.15	0.0007
C	GE120639			11.00 ± 0.11	11.04 ± 0.04	0.3879
A	1456111_at	D930028F11Rik	RIKEN cDNA D930028F11 gene	8.26 ± 0.27	7.40 ± 0.15	<0.0001
C	GE138361			7.51 ± 0.17	6.81 ± 0.08	<0.0001

A = Affymetrix Array; C = CodeLink Array

Data are represented as mean ± standard deviation of the log base 2 transformed expression data. Both Affymetrix and CodeLink data are from 6 mice of each strain. P-values are from two-sample t-tests assuming equal variance. The CodeLink data were normalized using the cyclic LOESS procedure. Similar results were found if quantile normalization was used. The Affymetrix data were normalized using RMA.