Published in final edited form as:

Alcohol Clin Exp Res. 2013 August ; 37(8): 1295–1303. doi:10.1111/acer.12100.

Selection for Drinking in the Dark Alters Brain Gene Coexpression Networks

Ovidiu Dan Iancu¹, Denesa Oberbeck¹, Priscila Darakjian¹, Pamela Metten^{1,4}, Shannon McWeeney^{2,3}, John C. Crabbe^{1,4}, and Robert Hitzemann^{1,4}

¹Department of Behavioral Neuroscience, Oregon Health & Science University, Portland, Oregon 97239-3098

²Oregon Clinical and Translational Research Institute, Oregon Health & Science University, Portland, Oregon 97239-3098

³Division of Biostatistics, Public Health & Preventative Medicine, Oregon Health & Science University, Portland, Oregon 97239-3098

⁴Portland Alcohol Research Center, Research Service, Veterans Affairs Medical Center, Portland, Oregon 97239

Abstract

Background—Heterogeneous stock (HS/NPT) mice have been used to create lines selectively bred in replicate for High Drinking in the Dark (HDID) (Crabbe et al., 2009; Crabbe et al., 2011b). Both selected lines routinely reach a blood ethanol concentration of 1.00 mg/ml or greater at the end of the day 2 four hour period of access. The mechanisms through which genetic differences influence DID are currently unclear. Therefore, the current study examines the transcriptome, the first stage at which genetic variability affects neurobiology. Rather than focusing solely on differential expression, we also examine changes in the ways that gene transcripts collectively interact with each other, as revealed by changes in coexpression patterns.

Methods—Naïve mice (N=48/group) were genotyped using the Mouse Universal Genotyping Array, which provided 3683 informative markers. QTL analysis used a marker-by-marker strategy with the threshold for a significant LOD set at 10.6. Gene expression in the ventral striatum was measured using the Illumina Mouse 8.2 array. Differential gene expression and the Weighted Gene Coexpression Network Analysis (WGCNA) were implemented largely as described elsewhere (Iancu et al., 2010; Iancu et al., 2012b).

Results—Significant quantitative trait loci (QTLs) for elevated blood ethanol concentrations after DID were detected on chromosomes 4, 14 and 16; the latter two were associated with gene poor regions. None of the QTLs overlapped with known QTLs for ethanol preference drinking. Ninety- four transcripts were detected as being differentially expressed in both selected lines vs HS controls; there was no overlap with known preference genes. The WGCNA revealed 2 modules as showing significant effects of both selections on intramodular connectivity. A number of genes known to be associated with ethanol phenotypes (e.g. *Gabarg1*, *Gla2*, *Grik1*, *Npy2r* and *Nts*) showed significant changes in connectivity.

Discussion—We found marked and consistent effects of selection on coexpression patterns; differential expression changes were more modest and less concordant. The QTLs and differentially expressed genes detected here are distinct from the preference phenotype (Mulligan

et al., 2006). This is consistent with behavioral data and suggests that the DID and preference phenotypes are markedly different genetically.

Keywords

alcohol; weighted gene coexpression network analysis; drinking in the dark; binge; mouse; microarray; selection

INTRODUCTION

Beginning with (Lewohl et al., 2000) there has been a continuous effort to align data on excessive ethanol consumption and differential brain gene expression. In rodents, this was focused on two-bottle choice, where alcohol is continuously available for days. The phenotype is easily implemented, there is variability among mouse strains and selection studies have generated preferring and non-preferring lines of animals e.g. (Belknap et al., 1997; Grahame et al., 1999). Apparently identical QTLs for ethanol preference have been detected in independent experiments (Belknap and Atkins, 2001). Using a genetical genomics approach, QTL and brain gene expression have been integrated to detect putative quantitative trait genes (Mulligan et al., 2006; Saba et al., 2011).

We address the relationship between genes and ethanol consumption from a different perspective. First, the animals used were selectively bred for HDID (Crabbe et al., 2009). Briefly, the phenotype involves replacing the water bottle with an ethanol (20% v/v) solution during the circadian dark phase. Animals drink to intoxication during the second day of limited access exposure (Rhodes et al., 2005; 2007). For selection, heterogeneous stock (HS/NPT) mice were used as the founders. This HS was formed by crossing eight laboratory mouse inbred strains (Hitzemann et al., 1994); the cross has a moderate genetic diversity (Roberts et al., 2007). Selection has been unidirectional since HS/NPT mice exhibit very low DID; the selection phenotype is the blood ethanol concentration (BEC). Details of the first selection (HDID-1) through the first eleven generations (S_{11}) are found in (Crabbe et al., 2009); at S_{22} , more than 80% of the animals had a BEC of 100 mg/ml or greater at the end of the four hour access period on Day 2. A second selection (HDID-2) was initiated two years later; at S_{15} , 60% of the animals exceeded the 1.00 mg/ml threshold. The National Institute on Alcohol Abuse and Alcoholism (NIAAA) criterion for binge drinking is a pattern that results in $BEC \geq 0.80$ mg/ml. The HDID-1 mice do not show markedly greater preference than HS and thus, there is an expectation that HDID and two-bottle preference QTLs will differ (Crabbe et al., 2009).

For analysis of gene expression data, the emphasis is on changes in coexpression patterns. Gene coexpression analysis is based on evaluation of pairwise correlations between pairs of genes. Classical differential expression focuses on the expression levels of N genes. In contrast, coexpression analysis focuses on analysis of an N by N coexpression matrix, also denoted as the network. The nodes of the network are the gene transcripts, while the edges are coexpression values. Important nodes (hubs) have high connectivity due to many strong connections. One seminal finding in biological networks is that the distribution of connectivity is not normal but exponential, with few hubs and many nodes with small connectivity.

Here we use WGCNA (Zhang and Horvath, 2005), which we have successfully applied for coexpression in complex murine crosses (Iancu et al., 2010) and selection for haloperidol response (Iancu et al., 2012a). Expression data (HDID-1, HDID-2 and HS/NPT) were collected from the ventral striatum from alcohol-naïve animals.

METHODS AND MATERIALS

Animals

For gene expression analysis, S₂₂ HDID-1 and S₁₅ HDID-2 animals were used in the current study; HS/NPT animals from the 72nd filial generation (G₇₂) were used as the controls. For QTL analysis, S₁₁ HDID-1 and -2 animals were compared with the independent HS/NPT founders (G₅₅ for the HDID-1 and G₆₄ for the HDID-2). Details concerning the selection of the HDID animals are found in (Crabbe et al., 2009). The HDID-1 selected line began with a “within family” selection design but was switched to mass selection at S₅. The HDID-2 selection used only mass selection. All animals used in the study were sacrificed between 10 AM and 2 PM; the light cycle for all animals was between 6 AM and 6 PM. All animal care, breeding, and testing procedures were approved by the Institutional Animal Care and Use Committees at the Veterans Affairs Medical Center, Portland, OR, and the Oregon Health & Science University, Portland, OR.

Genotyping the HDID-1 and HDID-2 selectively bred lines and the HS/NPT founders

We genotyped forty-eight animals (half male and half female) from each category (two selected lines and founders) as described in (Malmanger et al., 2006). These animals were also used for gene expression analysis. We used the Mouse Universal Genotyping Array (MUGA) (Geneseek, Lincoln, NE) which has 7,851 SNP markers, with an average spacing of 325 + 191 kb. After elimination of non-polymorphic or low frequency (below 2.5%) SNPs, the data contained 3683 markers further analyzed using a marker by marker approach (Hitzemann et al., 2009; Malmanger et al., 2006).

Quantification of genetic variability

Genome wide genetic distances between the animals were calculated using the genotype data. This measure quantifies genome-wide differences between pairs of animals, regardless of line. At each marker, genetic differences were computed based on the number of identical alleles; total genetic distances sums over all markers. Next, the between-group and within-group distances were calculated, following the Analysis of Molecular Variance (AMOVA) approach (Excoffier et al., 1992).

Dissection of tissue and extraction of RNA

Mice were euthanized, brains removed and immediately frozen on dry ice. Frozen brains were slightly thawed and dissected by hand under RNase-free conditions. Using the optic chiasm as the rostral marker, a 2 mm coronal slice of brain tissue was isolated. Following the partial cone shape of the striatum, the dissection moved dorsal 1 mm, followed by a cut to the lateral boundary of the striatum, with a final cut following the lateral-ventral boundary. The isolated tissue was immediately placed into 100 ml of Trizol (Invitrogen, Carlsbad, CA). On average, the isolated striatal tissue samples weighed 3 mg. Further details are found in (Malmanger et al., 2006).

Gene expression data processing

Gene expression data were obtained using the Illumina WG 8.2 array exactly as described by the manufacturer. Data were imported into the R application environment (<http://www.r-project.org>) using the lumi package (Du et al., 2008) and processed using the “lumiExpresso” function, with “quantile” normalization. The lumi package computes the probability of detection for each probe/samples; this value represents the probability that a particular probe/sample is not expressed above noise level. Using a cutoff threshold of 0.01, all probes not expressed above noise level in at least a quarter of the samples were removed. After these filtering steps, 9393 transcripts (7,380 unique genes) were retained.

Coexpression network construction and validation

This procedure quantifies connection strengths between transcripts, which become network edges. We use Pearson correlation between transcripts, computed over samples. A second step is detection of subnetworks/modules. Because there are many more genes than samples, procedures for validating the network structure are detailed below.

We constructed the coexpression network following the WGCNA approach (Langfelder and Horvath, 2008; Zhang and Horvath, 2005). Expression values for transcripts i, j , which iterate over all transcripts, were correlated across samples. The resulting correlation values a_{ij} were collected into the adjacency matrix A . The final connection strength between transcripts i, j , becomes $a_{ij} = |\text{corr}(x_i, x_j)|^\beta$; $\beta=10$ was selected in accordance to the scale-free topology criterion (Zhang and Horvath, 2005). The adjacency matrix was further processed into a topological overlap matrix (TOM); topological overlap between two transcripts i, j was computed as

$$\omega_{ij} = \frac{l_{ij} + a_{ij}}{\min\{k_i, k_j\} + 1 - a_{ij}}, l_{ij} = \sum_{u \neq i, j} a_{iu} a_{uj}$$

where $l_{ij} = \sum_u a_{iu} a_{uj}$ represents the number of genes connected to both gene i and gene j , while u indexes all the genes in the network. Using the topological overlap measure minimizes the effects of spurious connection strengths.

Detection of modules or groups of co-expressed transcripts was accomplished by clustering the TOM matrix using the “dynamic tree cut” algorithm (Langfelder et al., 2008). Transcript connectivity is defined as the sum of all connection strengths for a given transcript. For intramodular connectivity (kIM) the summation is restricted to a transcript’s own module. Transcript correlation with module eigengene (first principal component) results in module eigengene connectivity (kME). Module composition was validated by comparing module statistics against random groups of transcripts. Network statistics are defined in (Langfelder et al., 2011), equations 1–20. We computed intramodular connectivity (kIM), total network eigengene connectivity (kMEAll), module eigengene connectivity (kME), clustering coefficient (clusterCoeff) and maximum adjacency ratio (MAR). These quantities characterize individual nodes; for a module the nodes were arranged as a vector. Vectors originating from two different networks (HS/NPT and HDID) were correlated, resulting in cor.kIM, cor.kMEAll, cor.kME, cor.clusterCoeff and cor.MAR; and cor.ADJ. Statistical significance of preservation values was quantified using a Z score (Langfelder et al., 2011):

$$Z_a = \frac{obs_a - \mu_a}{\sigma_a}$$

where obs_a corresponds to preservation for a network statistic and module, μ_a and σ_a correspond to mean and standard deviation of preservation values for random groups of transcripts.

Module disruption between HS/NPT and HDID animals

Module preservation procedures were adapted to detect module disruption. Separate networks were formed using the HS/NPT and each HDID line’s expression data; differences between these networks were evaluated against random changes. An empirical distribution of random changes was generated by constructing networks (N=1000) using a mixture of

samples from the HS/NPT and HDID animals. Bootstrapping and statistical significance assessment was performed over samples. Disruption Z scores are negative, reflecting that preservation between different biological conditions is lower than preservation between networks created using a mixture of samples: obs_d is smaller than μ_a . Following (Langfelder et al., 2011), Z scores lower than -2 were considered disrupted and Z scores lower than -10 , strongly disrupted.

RESULTS

QTL analysis of HDID-1 and HDID-2 selected lines

QTL analysis was performed as described elsewhere (Hitzemann et al., 2009). The LOD threshold corrected for the number of markers entered into the analysis (3683) and corrected for replication was 5.3. However, because of the genetic drift, an additional correction was made (Hitzemann et al., 2009) which raised the LOD threshold to 10.6. A QTL exceeding this threshold in one of the selected lines and confirmed in the other line with a LOD >5.3 was considered significant. The corrected combined LOD score threshold is 14.3. QTLs meeting these criteria were detected on Chr 4 (87.1 to 90.5 Mbp), Chr 14 (83.0 – 87.1 Mbp) and Chr 16 (69.0 – 73.9 Mbp) (Supplemental Table 1). The haplotype structure of the QTLs was estimated on the basis of the peak SNP haplotype. For example, the Chr 14 QTL corresponds to the LP/J strain being different from the other 7 founder strains. The Chr 14 and 16 QTLs are associated with gene poor regions; e.g. there is only one known gene in the Chr 14 interval (protocadherin17) and only four known genes in the Chr 16 interval. In contrast, the Chr 16 interval and to a lesser extent the Chr 14 interval, are enriched in non-coding RNAs. The Chr 4 interval is gene rich, although poorly annotated in terms of function.

Genetic distances among HDID-1, HDID-2 and HS/NPT animals

The genetic distances among samples were calculated and plotted (Figure 1). The greater dispersion among the unselected HS/NPT controls compared to the selected lines was expected. Genetic distances were contracted in both selected lines, reflecting the effects of both selection and genetic drift. The 8 inbred strains used to form the HS/NPT are also plotted in Figure 1. Note that 7 of the strains cluster together and differ from the C57BL/6J. The data also illustrate that there were some substantial genetic differences between HDID-1 and HDID-2 animals.

Construction and validation of the consensus gene co-expression network

A consensus gene network application of the WGCNA was used to detect effects of selection on network structure (Iancu et al., 2012a; Langfelder and Horvath, 2008). The consensus network was constructed by combining all data. Clustering the consensus data identified 21 modules (sub-networks) designated with arbitrary colors; unassigned genes were denoted as “grey”. Module sizes varied between 46 (“dark-magenta”) and 1023 (“dark-red”). The quality of the detected modules (the strength of module connectedness), was evaluated using the procedure outlined in (Langfelder et al., 2011). The majority of the modules had high (above 10) quality Z scores (Supplemental Table 2), in contrast with low quality values (< 0) for the “grey” and “gold” (random selection of 1000 transcripts).

GO annotation of modules revealed association with distinct biological processes (Supplemental Table 3). Two modules of interest (see below), “black” and “magenta,” displayed rich neurobiological functional annotations. Module membership was also cross-referenced with markers of brain cell types (Cahoy et al., 2008). Six modules including “black” and “magenta” were enriched with neuronal cell markers, two modules with

oligodendrocyte markers and one module with astrocyte markers (Bonferroni corrected $p < 0.05$) (Supplemental Table 4).

Differential expression between HDID and HS/NPT animals

Differential striatal gene expression (DE) between HS/NPT controls and the selected lines was calculated using the eBayes modified t-statistic and further adjusted using the False Discovery Rate (FDR) (Benjamini and Hochberg, 1995). At a FDR 0.1, we detected 1430 (HDID-1 vs HS/NPT) and 301 (HDID-2 vs HS/NPT) differentially expressed transcripts. 104 transcripts were differentially expressed in both comparisons; 94 of these had the same directionality and are listed in Supplemental Table 5. A majority of the DE transcripts (85 out of 94) were found among the “grey” transcripts (Fisher exact test, $p < 10^{-20}$). The identity and GO annotation of the DE transcripts are found in Supplemental Tables 5 and 6. No differentially expressed transcripts were found in the “black” and “magenta” modules.

Differential network connectivity between HDID and HS/NPT animals

The network data originating from the HS/NPT and the selected lines is visualized in Figure 2. Gene transcripts with similar expression profile are placed closer together and are grouped in “modules” distinguished by color. Second, the figure illustrates the network changes in response to selection, with some modules appearing either more dispersed (“dark-red”) or more compacted (“magenta”). In this section, we focus on changes in module structure, we quantify and we evaluate their statistical significance.

Two distinct types of changes in network structure were evaluated. The first evaluates modules as whole units and quantifies changes using Z scores. Large negative Z scores signify statistically significant changes. A second type of change evaluated is transcript-specific and quantifies changes in connectivity. Here, the Z scores reported can be either positive or negative, signifying increased or decreased connectivity. Estimation of both types of Z scores is based on a bootstrapping (random sampling) procedure as outlined below.

Using the consensus network as the template, separate networks were constructed, using the HDID-1,2 and HS/NPT data. To determine the threshold for statistical difference in connectivity, a set of network comparisons ($N=1000$) were constructed using random mixtures of HDID and HS/NPT data. The first step was to evaluate the changes in connectivity at the level of whole modules (see Methods). A number of module preservation values were relatively low between the HDID and HS/NPT animals; these low-preservation values (or alternatively, high module disruption) were reflected in low negative Z scores (Supplemental Table 7). In particular, the “black” and “magenta” modules had markedly low cor.kME Z values in both selections. To a smaller extent (more moderate Z scores), the “dark-red”, “green” and “cyan” modules showed marked and consistent disruptions as well.

To focus on individual transcripts, the changes in transcript kME between HDID and HS/NPT networks were computed; effect size and statistical significance are summarized using Z scores (Supplemental Table 8). Eleven of the 173 transcripts in the “black” module (6.4%) showed significant changes in kME across both selections. Eight of the eleven transcripts showed an increase in connectivity while three showed a decrease. The positive connectivity changes for the *Dgkz* gene transcript are illustrated in Figure 3. For all transcripts in the “black” module, the changes in individual transcript kME were compared for both selections (Figure 4A); the correlation was significant ($r^2 = 0.41$; $p < 3.0 \times 10^{-14}$).

Within the “magenta” module 57 of the 265 transcripts (22%) showed a significant change in connectivity (Supplemental Table 8); 48 transcripts significantly increased connectivity kME and only 9 showed decreased connectivity. This result is concordant with the

visualization of the “magenta” module in Figure 2, showing a more compact structure in the selection networks. In contrast, the “dark-red” module appears more dispersed in the selected lines, concordant with the fact that most transcripts decreased connectivity (Supplemental Table 8). The connectivity values for a sample of significantly affected transcripts from the “magenta” module are found in Table 1. Note the particularly marked changes in connectivity for *Gabrg1*, *Glnr2*, *Npyr2* and *Nts*. For all transcripts in the “magenta” module, the change in individual transcript kME were compared for both selections (Figure 4B); the correlation was significant ($r^2 = 0.64$; $p < 6.5 \times 10^{-31}$), signifying that the two selections affected module structure in a very consistent manner.

DISCUSSION

There are several methods for inducing binge-like ethanol consumption in mice (Crabbe et al., 2011a); these include DID, intragastric delivery of ethanol (Fidler et al., 2012) and scheduled high alcohol consumption (SHAC) (Finn et al., 2005). Of these, the DID model has been most widely used and generated a review on the underlying neurobiology (Sprow and Thiele, 2012). As noted by (Sprow and Thiele, 2012) studies of the DID model have largely depended on using the C57BL/6J (B6) mouse which has a proclivity for ethanol consumption. This feature of the B6 mouse is shared across strains from the C57/C58 mouse lineage but is not widely understood and thus may limit the extent to which DID data can be generalized. One solution to this issue is to selectively breed for a high DID phenotype and use genetically diverse heterogeneous stock (HS) founders. The general argument here is that one is not constrained by the unique features of the B6 background and selection will primarily force the segregation of those alleles associated with the DID phenotype. One problem with selective breeding is that with limited population sizes, genetic drift will lead to the fixation of some alleles unrelated to the selection phenotype. Assuming the drift effects are stochastic, one solution to the drift problem is an independent replicate selection. The value of the replicate selection was repeatedly observed in the current study. When comparing the HDID-1 and HDID-2 lines, which show relatively similar high BEC and amount of ethanol consumption, marked differences in allele segregation, differential gene expression and effects on network structure were detected. For example, when compared to the HS/NPT, there were 1430 and 301 differentially expressed transcripts in the HDID-1 and HDID-2 animals, respectively. However, there were only 94 common transcripts differentially expressed in the same direction.

The value of the replicate design was also evidenced in the QTL analysis. In addition to the common QTLs (Chr 4, 14 and 16), significant QTLs exceeding the LOD 10.6 threshold were found for HDID-1 on Chr 2, 6, 15, 18 and 19 and for HDID-2 on Chr 7, 10 and 11 (data not shown). The common HDID QTLs detected in the study are very different from those that have been detected for 2-bottle preference drinking (Belknap and Atkins, 2001; Hitzemann et al., 2009) and thus further support the idea that binge and preference phenotypes are mostly genetically different (Crabbe et al., 2011b), even though some genetic similarity is indicated by significant genetic correlations among inbred strain values for preference drinking and DID (Crabbe et al., 2012). Perhaps the most notable preference QTL not detected was the preference QTL found on Chr 9 and centered at approximately 50 Mbp. This QTL has not only been detected in genotypes derived from crosses of the high-preferring B6 and low-preferring DBA/2 inbred strains (Belknap and Atkins, 2001) but also in the analysis of the High Alcohol Preferring/Low Alcohol Preferring mice that are derived by selective breeding from an HS with six strains in common with the HS/Ibg founders (Bice et al., 2006).

Related to the QTL analysis, selection had marked effects on genetic structure (as measured by genetic distance) in both the HDID-1 and HDID-2 lines (Figure 1). Given the limited

number of families used for breeding and the associated inbreeding, the decrease in genetic variance was not surprising. Similar results have been obtained elsewhere (Iancu et al., 2012a). The variance can be expected to continue to decrease as selection proceeds for additional generations. The data also illustrate some substantial genetic differences between the two selections, likely due to the random nature of the drift-related inbreeding. Given small population sizes, the specific constellations of genes whose alleles are changed by directional selection are bound to be different as there are no doubt many genes with the potential to influence the DID trait.

Lists of genes associated with excessively high ethanol consumption have been largely generated from two sources: microarray and gene knockout data - see e.g. (Blednov et al., 2012; Crabbe et al., 2006; Mulligan et al., 2006; Saba et al., 2011). One such list has been generated by the Integrative Neuroscience Initiative on Alcoholism (G. Koob and R. Harris, personal communication - Supplemental Table 9). This list has no overlap with the 94 genes differentially expressed in both selections. The lack of overlap is not entirely surprising since the INIA list was largely generated from 2-bottle preference data. While the HDID selected lines show roughly 4-fold greater BECs after DID than HS controls, their g/kg intake is only about 50% greater. As other brain regions are surveyed, some overlap may emerge. In the current study, the differences in expression were generally very small (10 to 20%). A transcript for *Slc25a18* (mitochondrial glutamate transporter) showed the greatest difference (+67% HDID-1/HS/NPT and +57% HDID-2/HS/NPT).

The WGCNA is one of several gene clustering based approaches that have been used to examine the functional organization of the transcriptome. Genes that cluster together are assumed to share some similar functions. Given that many genes are poorly annotated (see above for the genes within the Chr 4 QTL interval) or are poorly annotated for the organ of interest, the network analysis allows one to infer new functionalities. The biological plausibility of modules can be ascertained by GO annotation, overlap with cell type markers, enrichment in protein interactions and spatial colocalization (Iancu et al., 2010). The network emphasizes the importance of relatively few highly connected hubs (Zhang and Horvath, 2005). These hubs, as opposed to isolated nodes, are key targets for experimental manipulation.

In the current study, a consensus network approach (Iancu et al., 2012a; Langfelder and Horvath, 2007) was used to evaluate the effects of selection on transcriptome organization. By comparing the results from the HDID-1 and HDID-2 selections, it was possible to conclude that selection markedly disrupted the structure of 2 modules ("black" and "magenta"); "cyan", dark red" and "green" were affected to a smaller extent. The "black", "green" and "magenta" modules were 3 of the 6 modules enriched in neuronal cell type genes, suggesting some specificity of the selection effects. Our analysis focused on the "black" and "magenta" modules and the effects of selection on intramodular connectivity. The direction of connectivity change was remarkably consistent (Figure 4). The "black" module contained 173 transcripts, 11 of which showed a significant change in their correlation to the module eigengene. A gene with no known annotation, *EG574403* (Chr 11:34214822) showed the most significant change. Of the remaining genes, *Dgkz*, has potential behavioral links. For example, (Kim et al., 2009) have shown that the synaptic removal of diacylglycerol by DGKzeta and PSD-95 regulates synaptic spine maintenance. Chronic ethanol exposure has been shown to affect spine density and/or morphology (Carpenter-Hyland and Chandler, 2006; Zhou et al., 2007). Further, in comparison to ethanol-avoiding NP rats, ethanol-preferring P rats have a lower dendritic spine density in the amygdala (Moonat et al., 2011).

Compared to the “black” module, the “magenta” module was more significantly affected by selection (compare Figures 4A and 4B). Fifty-eight of 265 transcripts showed a significant change; eight of these are highlighted in Table 1 and will be briefly discussed. To our knowledge, there is no reference linking *Camk2n1* (calcium/calmodulin kinase II inhibitor alpha) or its gene product to alcohol related phenotypes. There is however, a rich literature describing the relationship between alcohol and Ca^{+2} /calmodulin kinase II (CaMKII). CaMKII has a well established role in synaptic plasticity (Lisman et al., 2012). Chronic ethanol exposure can increase CaMKII activity which in turn modulates adaptive responses (Christian et al., 2012). Ca^{+2} /calmodulin kinase kinase 2 beta (encoded by *Camkk2*) has a role in hippocampal dependent memory (Peters et al., 2003). This kinase also has a key role in ghrelin signaling; however, the systemic administration of ghrelin has not been found to have an effect on DID (Lyons et al., 2008). DGKkappa has no known function in the brain but is the second member of the DGK family to show a significant change in connectivity. In human genetic studies, the gamma-1 subunit of the GABA-A receptor has been shown to be associated with alcohol response (Ray and Hutchison, 2009), alcohol dependence (Enoch et al., 2009) and alcohol withdrawal (Wang et al., 2012). Further, chronic ethanol exposure increases *Gabrg1* expression by > 70% in the rat cerebral cortex (Devaud et al., 1996). It is well established that ethanol interacts with a number of pentameric ligand gated ion channels, including the glycine receptor and many details of the structural interactions are known (see (Howard et al., 2011) and ref therein). The alpha2 subunit is the most common in the rodent forebrain (Jonsson et al., 2009); however, alpha2 homomer receptors are less sensitive than alpha1 homomer receptors to ethanol effects (potentiating) (Davies et al., 2004). Jonsson et al. (2009) have found in both the AA and ANA rat selected lines that ethanol consumption down-regulated *Gla2* expression in the amygdala and anterior hypothalamus. Polymorphisms in *Grik1*, which encodes for the ionotropic GluR5 kainate receptor, have been linked to alcohol dependence (Kranzler and Edenberg, 2010) and topiramate, a modulator of the GluR5 receptor, has been used to treat alcohol dependence (Johnson et al., 2007). A polymorphism in *Grik1* predicts adverse events from topiramate treatment (Kranzler and Edenberg, 2010). Sparrow et al. (2012) have found that the neuropeptide Y2R antagonist, BIIE-0246, reduces DID ethanol consumption. Further, chronic binge alcohol consumption increases Y2R immunoreactivity in the central nucleus of the amygdala. The neurotensin (NTS) system has been repeatedly linked to voluntary ethanol consumption - e.g. (Ehlers et al., 1999; Lee et al., 2011); however, there appear to be no studies that specifically focus on the DID phenotype. However, it was the *Nts* transcript that showed the greatest selection-induced change in connectivity.

(Mulligan et al., 2011) have examined gene expression and DID using a similar network approach. These authors were interested in understanding at the gene expression level the marked differences in ethanol consumption among B6 mice (N=21); differences in ethanol consumption were approximately 100% and the differences in BECs were even greater (see figure 1 – (Mulligan et al., 2011)). Of relevance to the current study, (Mulligan et al., 2011) observed that one striatal module was significantly correlated with BECs and was enriched in neuronal transcripts, similar to the affected modules in the present study. Our study shares with (Mulligan et al., 2011) the general network approach to transcriptome analysis, but it also differs in several essential aspects. First, our study focuses much more closely on the striatal transcriptome and (Mulligan et al., 2011) analyzed several brain regions. Second, the genetic background of the HS/NPT and HDID selected lines is highly diverse, as opposed to the C57BL/6J inbred strain used by (Mulligan et al., 2011). Third, animals analyzed here were never exposed to alcohol while the (Mulligan et al., 2011) study focused on transcriptional changes induced after ethanol exposure. In summary, our study aims to elucidate genetic and transcriptional underpinnings of *propensity* for the post-DID high BEC phenotype, while (Mulligan et al., 2011) analyzes the after-effects of ethanol exposure. Despite these differences, both studies generally find that alcohol-related phenotypes

correspond to system or network-level effects at the level of the transcriptome, which are both more subtle and more extensive than simple changes in expression levels of a few genes.

The data presented here and in (Mulligan et al., 2011) illustrate the value of a network-based approach to investigating gene expression and binge-like ethanol consumption. We emphasized changes in intramodular connectivity associated with HDID selection and we found that differential expression changes were relatively modest and less concordant. One interesting feature of differential expression is the concentration of differential expression in the “grey” category, which is outside the modules. We hypothesize that large, but uncorrelated changes in expression levels render these genes less likely to be assigned to specific modules. It is also possible that selective pressure could generate two separate effects on the transcriptome, with a small number of isolated (“grey”) genes having large changes in expression levels while larger groups of genes (modules) changing transcription profiles in a more moderate but synergistic/correlated fashion. Higher reproducibility in connectivity changes suggests that the latter mechanism is more important. Further, the genes showing the most marked connectivity changes have been previously aligned with ethanol phenotypes, including binge consumption.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported in part by AA 10760, AA 13484, AA 13519, AA 20245 and DA005228 and grants from the Department of Veterans Affairs.

REFERENCES

- Belknap JK, Atkins AL. The replicability of QTLs for murine alcohol preference drinking behavior across eight independent studies. *Mamm Genome*. 2001; 12:893–899. [PubMed: 11707775]
- Belknap JK, Richards SP, O'Toole LA, Helms ML, Phillips TJ. Short-term selective breeding as a tool for QTL mapping: ethanol preference drinking in mice. *Behav Genet*. 1997; 27:55–66. [PubMed: 9145544]
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol*. 1995; 57:289–300.
- Bice PJ, Foroud T, Carr LG, Zhang L, Liu L, Grahame NJ, Lumeng L, Li TK, Belknap JK. Identification of QTLs influencing alcohol preference in the High Alcohol Preferring (HAP) and Low Alcohol Preferring (LAP) mouse lines. *Behav Genet*. 2006; 36:248–260. [PubMed: 16482403]
- Blednov YA, Mayfield RD, Belknap J, Harris RA. Behavioral actions of alcohol: phenotypic relations from multivariate analysis of mutant mouse data. *Genes Brain Behav*. 2012; 11:424–435. [PubMed: 22405477]
- Cahoy JD, Emery B, Kaushal A, Foo LC, Zamanian JL, Christopherson KS, Xing Y, Lubischer JL, Krieg PA, Krupenko SA, Thompson WJ, Barres BA. A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J Neurosci*. 2008; 28:264–278. [PubMed: 18171944]
- Carpenter-Hyland EP, Chandler LJ. Homeostatic plasticity during alcohol exposure promotes enlargement of dendritic spines. *Eur J Neurosci*. 2006; 24:3496–3506. [PubMed: 17229098]
- Christian DT, Alexander NJ, Diaz MR, Robinson S, McCool BA. Chronic intermittent ethanol and withdrawal differentially modulate basolateral amygdala AMPA-type glutamate receptor function and trafficking. *Neuropharmacology*. 2012; 62:2430–2439. [PubMed: 22387532]
- Crabbe JC, Harris RA, Koob GF. Preclinical studies of alcohol binge drinking. *Ann N Y Acad Sci*. 2011a; 1216:24–40. [PubMed: 21272009]

- Crabbe JC, Metten P, Rhodes JS, Yu CH, Brown LL, Phillips TJ, Finn DA. A line of mice selected for high blood ethanol concentrations shows drinking in the dark to intoxication. *Biol Psychiatry*. 2009; 65:662–670. [PubMed: 19095222]
- Crabbe JC, Phillips TJ, Harris RA, Arends MA, Koob GF. Alcohol-related genes: contributions from studies with genetically engineered mice. *Addict Biol*. 2006; 11:195–269. [PubMed: 16961758]
- Crabbe JC, Spence SE, Brown LL, Metten P. Alcohol preference drinking in a mouse line selectively bred for high drinking in the dark. *Alcohol*. 2011b; 45:427–440. [PubMed: 21194877]
- Davies DL, Crawford DK, Trudell JR, Mihic SJ, Alkana RL. Multiple sites of ethanol action in $\alpha 1$ and $\alpha 2$ glycine receptors suggested by sensitivity to pressure antagonism. *J Neurochem*. 2004; 89:1175–1185. [PubMed: 15147510]
- Devaud LL, Purdy RH, Finn DA, Morrow AL. Sensitization of gamma-aminobutyric acidA receptors to neuroactive steroids in rats during ethanol withdrawal. *J Pharmacol Exp Ther*. 1996; 278:510–517. [PubMed: 8768698]
- Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics*. 2008; 24:1547–1548. [PubMed: 18467348]
- Ehlers CL, Somes C, Li TK, Lumeng L, Kinkead B, Owens MJ, Nemeroff CB. Neuronten studies in alcohol naive, preferring and non-preferring rats. *Neuroscience*. 1999; 93:227–236. [PubMed: 10430486]
- Enoch MA, Hodgkinson CA, Yuan Q, Albaugh B, Virkkunen M, Goldman D. GABRG1 and GABRA2 as independent predictors for alcoholism in two populations. *Neuropsychopharmacology*. 2009; 34:1245–1254. [PubMed: 18818659]
- Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. 1992; 131:479–491. [PubMed: 1644282]
- Fidler TL, Powers MS, Ramirez JJ, Crane A, Mulgrew J, Smitasin P, Cunningham CL. Dependence induced increases in intragastric alcohol consumption in mice. *Addict Biol*. 2012; 17:13–32. [PubMed: 21955048]
- Finn DA, Belknap JK, Cronise K, Yoneyama N, Murillo A, Crabbe JC. A procedure to produce high alcohol intake in mice. *Psychopharmacology (Berl)*. 2005; 178:471–480. [PubMed: 15765261]
- Grahame NJ, Li TK, Lumeng L. Selective breeding for high and low alcohol preference in mice. *Behav Genet*. 1999; 29:47–57. [PubMed: 10371758]
- Hitzemann B, Dains K, Kanes S, Hitzemann R. Further studies on the relationship between dopamine cell density and haloperidol-induced catalepsy. *J Pharmacol Exp Ther*. 1994; 271:969–976. [PubMed: 7965819]
- Hitzemann R, Edmunds S, Wu W, Malmanger B, Walter N, Belknap J, Darakjian P, McWeeney S. Detection of reciprocal quantitative trait loci for acute ethanol withdrawal and ethanol consumption in heterogeneous stock mice. *Psychopharmacology (Berl)*. 2009; 203:713–722. [PubMed: 19052728]
- Howard RJ, Murail S, Ondricek KE, Corringer PJ, Lindahl E, Trudell JR, Harris RA. Structural basis for alcohol modulation of a pentameric ligand-gated ion channel. *Proc Natl Acad Sci U S A*. 2011; 108:12149–12154. [PubMed: 21730162]
- Iancu OD, Darakjian P, Malmanger B, Walter NA, McWeeney S, Hitzemann R. Gene networks and haloperidol-induced catalepsy. *Genes Brain Behav*. 2012a; 11:29–37. [PubMed: 21967164]
- Iancu OD, Darakjian P, Walter NA, Malmanger B, Oberbeck D, Belknap J, McWeeney S, Hitzemann R. Genetic diversity and striatal gene networks: focus on the heterogeneous stock-collaborative cross (HS-CC) mouse. *BMC Genomics*. 2010; 11:585. [PubMed: 20959017]
- Iancu OD, Kawane S, Bottomly D, Searles R, Hitzemann R, McWeeney S. Utilizing RNA-Seq data for de novo coexpression network inference. *Bioinformatics*. 2012b; 28:1592–1597. [PubMed: 22556371]
- Johnson BA, Rosenthal N, Capece JA, Wiegand F, Mao L, Beyers K, McKay A, Ait-Daoud N, Anton RF, Ciraulo DA, Kranzler HR, Mann K, O'Malley SS, Swift RM. Topiramate for treating alcohol dependence: a randomized controlled trial. *JAMA*. 2007; 298:1641–1651. [PubMed: 17925516]
- Jonsson S, Kerekes N, Hyytia P, Ericson M, Soderpalm B. Glycine receptor expression in the forebrain of male AA/ANA rats. *Brain Res*. 2009; 1305:S27–S36. [PubMed: 19781529]

- Kim K, Yang J, Zhong XP, Kim MH, Kim YS, Lee HW, Han S, Choi J, Han K, Seo J, Prescott SM, Topham MK, Bae YC, Koretzky G, Choi SY, Kim E. Synaptic removal of diacylglycerol by DGKzeta and PSD-95 regulates dendritic spine maintenance. *Embo J*. 2009; 28:1170–1179. [PubMed: 19229292]
- Kranzler HR, Edenberg HJ. Pharmacogenetics of alcohol and alcohol dependence treatment. *Curr Pharm Des*. 2010; 16:2141–2148. [PubMed: 20482509]
- Langfelder P, Horvath S. Eigengene networks for studying the relationships between co-expression modules. *BMC Syst Biol*. 2007; 1:54. [PubMed: 18031580]
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008; 9:559. [PubMed: 19114008]
- Langfelder P, Luo R, Oldham MC, Horvath S. Is my network module preserved and reproducible? *PLoS Comput Biol*. 2011; 7:e1001057. [PubMed: 21283776]
- Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics*. 2008; 24:719–720. [PubMed: 18024473]
- Lee MR, Hinton DJ, Unal SS, Richelson E, Choi DS. Increased ethanol consumption and preference in mice lacking neurotensin receptor type 2. *Alcohol Clin Exp Res*. 2011; 35:99–107. [PubMed: 21039631]
- Lewohl JM, Wang L, Miles MF, Zhang L, Dodd PR, Harris RA. Gene expression in human alcoholism: microarray analysis of frontal cortex. *Alcohol Clin Exp Res*. 2000; 24:1873–1882. [PubMed: 11141048]
- Lisman J, Yasuda R, Raghavachari S. Mechanisms of CaMKII action in long-term potentiation. *Nat Rev Neurosci*. 2012; 13:169–182. [PubMed: 22334212]
- Lyons AM, Lowery EG, Sparta DR, Thiele TE. Effects of food availability and administration of orexigenic and anorectic agents on elevated ethanol drinking associated with drinking in the dark procedures. *Alcohol Clin Exp Res*. 2008; 32:1962–1968. [PubMed: 18782340]
- Malmanger B, Lawler M, Coulombe S, Murray R, Cooper S, Polyakov Y, Belknap J, Hitzemann R. Further studies on using multiple-cross mapping (MCM) to map quantitative trait loci. *Mamm Genome*. 2006; 17:1193–1204. [PubMed: 17143586]
- Moonat S, Sakharikar AJ, Zhang H, Pandey SC. The role of amygdaloid brain-derived neurotrophic factor, activity-regulated cytoskeleton-associated protein and dendritic spines in anxiety and alcoholism. *Addict Biol*. 2011; 16:238–250. [PubMed: 21182574]
- 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*. 2006; 103:6368–6373. [PubMed: 16618939]
- Mulligan MK, Rhodes JS, Crabbe JC, Mayfield RD, Adron Harris R, Ponomarev I. Molecular profiles of drinking alcohol to intoxication in C57BL/6J mice. *Alcohol Clin Exp Res*. 2011; 35:659–670. [PubMed: 21223303]
- Peters M, Mizuno K, Ris L, Angelo M, Godaux E, Giese KP. Loss of Ca²⁺/calmodulin kinase kinase beta affects the formation of some, but not all, types of hippocampus-dependent long-term memory. *J Neurosci*. 2003; 23:9752–9760. [PubMed: 14586002]
- Ray LA, Hutchison KE. Associations among GABRG1, level of response to alcohol, and drinking behaviors. *Alcohol Clin Exp Res*. 2009; 33:1382–1390. [PubMed: 19426171]
- Rhodes JS, Best K, Belknap JK, Finn DA, Crabbe JC. Evaluation of a simple model of ethanol drinking to intoxication in C57BL/6J mice. *Physiol Behav*. 2005; 84:53–63. [PubMed: 15642607]
- Rhodes JS, Ford MM, Yu CH, Brown LL, Finn DA, Garland T Jr, Crabbe JC. Mouse inbred strain differences in ethanol drinking to intoxication. *Genes Brain Behav*. 2007; 6:1–18. [PubMed: 17233637]
- Roberts A, Pardo-Manuel de Villena F, Wang W, McMillan L, Threadgill DW. The polymorphism architecture of mouse genetic resources elucidated using genome-wide resequencing data: implications for QTL discovery and systems genetics. *Mamm Genome*. 2007; 18:473–481. [PubMed: 17674098]

- Saba LM, Bennett B, Hoffman PL, Barcomb K, Ishii T, Kechris K, Tabakoff B. A systems genetic analysis of alcohol drinking by mice, rats and men: influence of brain GABAergic transmission. *Neuropharmacology*. 2011; 60:1269–1280. [PubMed: 21185315]
- Sparrow AM, Lowery-Gionta EG, Pleil KE, Li C, Sprow GM, Cox BR, Rinker JA, Jijon AM, Pena J, Navarro M, Kash TL, Thiele TE. Central neuropeptide Y modulates binge-like ethanol drinking in C57BL/6J mice via Y1 and Y2 receptors. *Neuropsychopharmacology*. 2012; 37:1409–1421. [PubMed: 22218088]
- Sprow GM, Thiele TE. The neurobiology of binge-like ethanol drinking: evidence from rodent models. *Physiol Behav*. 2012; 106:325–331. [PubMed: 22245775]
- Wang KS, Liu X, Zhang Q, Wu LY, Zeng M. Genome-wide association study identifies 5q21 and 9p24.1 (KDM4C) loci associated with alcohol withdrawal symptoms. *J Neural Transm*. 2012; 119:425–433. [PubMed: 22072270]
- Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol*. 2005; 4 Article17.
- Zhou FC, Anthony B, Dunn KW, Lindquist WB, Xu ZC, Deng P. Chronic alcohol drinking alters neuronal dendritic spines in the brain reward center nucleus accumbens. *Brain Res*. 2007; 1134:148–161. [PubMed: 17198693]

Figure 1.

Genome-wide genetic distances between the HDID-1 and -2 selected lines, HS/NPT animals and the inbred strains used to form the HS/NPT. Details of the animals used are found in the Methods. Data are presented as a multidimensional scaling (MDS) plot. Note the greater dispersion in the HS/NPT animals when compared to the HDID-1 and HDID-2 selected lines and the differences between the two selected lines. Also note that among the inbred strains the C57BL/6J is distinct from the other 7 founders.

Figure 2.

Multidimensional scaling (MDS) plots of the coexpression networks in HS/NPT (A), HDID-2 (B) and HDID-1 (C) datasets. For visual clarity, only the 4 modules most consistently affected by selection (“black”, “magenta”, “dark-red” and “green”) are depicted. Each dot represents a transcript, with colors corresponding to module assignments. The distances between points correspond to network adjacency. The figure illustrates 1) the modularity of the networks, with similar colors clustered together and 2) the effect of selection on the network structure, with HDID-1,2 diverging from the original HS/NPT network structure. In particular, the “dark-red” module appears more dispersed, while the “magenta” module appears more compacted in the selection networks.

Figure 3.

The effects of selection on intra-modular connectivity for *Dgkz*. *Dgkz* is found in the “black” module. Edge thickness and opacity are proportional with network adjacency between *Dgkz* and other module transcripts. The intra-modular connectivity of the other module genes is reflected in the node size. A: HS/NPT network connectivity. B: HDID-2 network connectivity. C: HDID-1 network connectivity. Note the more pronounced increase in connectivity in the HDID-1 as compared to the HDID-2 animals.

Figure 4. Reproducibility of effects on intramodular connectivity in the “black” and “magenta” modules. A: Correlation of changes in intramodular connectivity for “black” module transcripts (HDID-1/HS/NPT vs HDID-2/HS/NPT). B: Correlation of changes in intramodular connectivity for “magenta” module transcripts (HDID-1/HS/NPT vs HDID-2/HS/NPT).

Table 1

Selected Genes From the Magenta Module that Show Significant Changes in Connectivity in both Selections

Symbol	kMENPT	kMEH1	zH1	kMEH2	zH2
<i>Camk2n1</i>	-0.32	-0.77	-5.62	-0.68	-2.06
<i>Camkk2</i>	-0.62	-0.89	-5.80	-0.87	-3.43
<i>Dgkk</i>	0.30	0.62	2.72	0.64	2.33
<i>Gabrg1</i>	0.21	0.78	5.51	0.68	2.79
<i>Gtra2</i>	0.27	0.84	7.20	0.84	4.85
<i>Grik1</i>	0.55	0.77	3.07	0.80	2.58
<i>Npy2r</i>	-0.05	0.59	4.65	0.74	4.89
<i>Nts</i>	-0.27	0.54	5.56	0.56	3.33

kME is the correlation between the gene transcript and the module eigengene. zH is the significance of the difference in the correlation between the HS/NPT and the HDID-1 and -2 selected lines. Z scores between 2 and 10 are considered to be moderately significant (see Methods). Symbols: calcium/calmodulin-dependent protein kinase 2 inhibitor 1 (*Camk2n1*), calcium/calmodulin-dependent protein kinase kinase 2, beta (*Camkk2*), diacylglycerol kinase kappa (*Dgkk*), GABA receptor subunit gamma 1 (*Gabrg1*), glycine receptor, alpha2 subunit (*Gtra2*), glutamate receptor, ionotropic, kainate (*Grik1*), neuropeptide Y receptor Y2 (*Npy2r*), neurotensin (*Nts*).