The Dissertation Committee for Kerrie Hall Lodowski Certifies that this is the approved version of the following dissertation:

Alcohol-induced temporal transcriptome remodeling in the prefrontal cortex in a mouse model of alcohol dependence

Committee:

Susan E. Bergeson, Supervisor

Dean R. Appling

Karen J. Artzt

R. Adron Harris

R. Dayne Mayfield

Alcohol-induced temporal transcriptome remodeling in the prefrontal cortex in a mouse model of alcohol dependence

by

Kerrie Hall Lodowski, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2006

Dedication

To my parents, Randy and Mary Frances Hall, who always encouraged me in all of my endeavors.

and

To my husband, David, whose love and support was instrumental in my success as a graduate student.

Acknowledgements

I would like to thank…

My mentor, Dr. Susan Bergeson, for her insatiable enthusiasm for science and constant encouragement. I would not have the confidence I do today as an independent scientist if it were not for her.

The current and former members of the Bergeson lab. Dr. Ari Berman is a wonderful colleague and has become one of my closest friends. Pat Levin, Megan Mulligan, Julie Owen, Dr. Gwen Schafer and Amy Wolfrey created an intellectually stimulating and fun environment in which to work. Farah Hameed, Chris Hornsby and Mahesh Thenappan were three undergraduate students who contributed to the success of my project.

The members of the Waggoner Center for their constructive comments at lab meetings. A special thanks to Dr. Igor Ponomarev and Elizabeth Osterndorff-Kahanek.

Dr. Vishwanath Iyer and the members of the Iyer lab, especially Patrick Killion, for their assistance with the production of the microarrays and maintenance of the LAD.

The members of my committee for all their constructive comments over the years. A special thanks to Dr. R. Dayne Mayfield for sage scientific advice and for agreeing to join my committee at such a late stage in my graduate career.

M. June and J. Virgil Waggoner for their generous donation to UT to create the Waggoner Center for Alcohol and Addiction Research.

My and my husband's families for their unending encouragement and support throughout this whole process.

My husband David, for his patience, his willingness to listen, his love and his commitment. I cannot imagine having done this without him by my side.

Alcohol-induced temporal transcriptome remodeling in the prefrontal cortex in a mouse model of alcohol dependence

Publication No._____

Kerrie Hall Lodowski, Ph.D.

The University of Texas at Austin, 2006

Supervisor: Susan E. Bergeson

Alcohol dependence (alcoholism) is a complex disease influenced by both environmental factors and genetic predisposition. Mouse models have been used to study many alcohol dependence-related traits and the genetics that underlie them. Two of the most commonly used mice in alcohol research are the C57BL/6J (B6) and DBA/2J (D2) inbred strains, which diverge on several alcohol-related traits including the development of acute physical dependence. Here we utilized the B6 and D2 mice as a genetic model of acute physical dependence, coupled with mRNA Differential Display (DD) and cDNA microarray analysis, to uncover the transcriptional response of the brain to an acute dose of alcohol as a function of time. About 150 genetically divergent and alcohol-responsive genes were identified between the whole brains of B6 and D2 mice using DD and were added as additional targets to the mouse microarrays. Microarray analysis of the prefrontal cortex of B6 and D2 mice revealed strain-specific, acute alcohol-responsive transcriptome remodeling manifested as temporal patterns of gene expression. Distinct expression patterns were identified for physiologically relevant alcohol-related

consequences including intoxication, withdrawal and neuroadaptation. *In silico* characterization of the differentially expressed genes showed genotype dependent and independent transcriptional regulation and functional classification. In addition, categorization of differentially expressed genes by their cellular profiles revealed that some of the genes were known to be more highly expressed in either excitatory or inhibitory neuronal cell types. Our results indicate that the B6 and D2 prefrontal cortices have very different cellular and molecular responses to acute alcohol exposure. The specific roles that the genes identified in this study may play in mediating the divergent alcohol-related behavior between the strains warrant further study.

TABLE OF CONTENTS

List of Tables

List of Figures

Rationale

Alcohol dependence (alcoholism) is a complex disease influenced by both environmental factors and genetic predisposition. The genes contributing to the risk for developing alcohol dependence are numerous and of small individual effect. Several behavioral, pharmacological and molecular methodologies have been used over the years to identify genes contributing to every aspect of this disease. The current hypothesis is that alcohol disrupts the normal homeostasis of neurons in the brain, affecting neuronal communication and ultimately neuronal circuits between brain regions. Over time, the brain adapts to the presence of alcohol such that it becomes physically dependent on the alcohol and functions poorly without it. Gene expression is thought to play an important role in this brain neuroadaptation. Although progress has been made in identifying individual genes associated with alcohol dependence, to truly understand the molecular basis of a disease as complex as alcohol dependence, studies must go beyond that of individual genes. This study combines a genetically divergent mouse model of physical dependence, mRNA Differential Display, cDNA microarrays and a time course following alcohol administration to simultaneously identify thousands of genes that are differentially expressed as a result of genetic predisposition, alcohol and time. Genes that are differentially expressed can be classified according to their function, common transcription factor binding sites or cellular identities and specific, testable hypotheses can be developed about how genes or gene families contribute to the development of alcohol dependence.

Hypotheses

Alcohol-related and genetically divergent, alcohol-related gene expression changes occur in C57BL/6J and DBA/2J mouse brain as a result of acute alcohol exposure.

Gene expression differences exist between the brains of C57BL/6J and DBA/2J mice.

CHAPTER 1: INTRODUCTION

Alcohol dependence

Alcohol dependence is a faceless disease. It does not care about nationality, race, gender or age. The moment someone takes his or her first drink of alcohol, the possibility exists that he or she may become dependent on alcohol. Risk for developing the disease is influenced by genetics and the environment (Aston and Hill, 1990; Crabbe and Goldman, 1992). There is no cure for alcohol dependence and although treatment options exist, they are often population- or person-specific in their effectiveness.

In the 1980s, the American Psychiatric Association (APA) divided the term "alcoholism" into two categories—"alcohol abuse" and "alcohol dependence" (APA, 1980). The Diagnostic and Statistical Manual of Mental Disorders-IV-TR (APA, 2000) defines these terms as follows:

(1) Alcohol abuse: "A maladaptive pattern of drinking, leading to clinically significant impairment or distress as manifested by at least one of the following in a 12-month period: failure to fulfill major role obligations at work, school or home, recurrent alcohol-related legal, social or interpersonal problems or recurrent participation in physically hazardous situations."

(2) Alcohol dependence: "A maladaptive pattern of drinking, leading to clinically significant impairment or distress as manifested by 3 or more of the following in a 12-month period: tolerance, withdrawal, impaired control, a great deal of time spent in alcohol-related activities or continued use despite negative consequences."

As the names imply, dependence is the key difference between the two categories. Chronic alcohol abuse causes a "re-wiring" of the brain, creating a state of dependence where the body functions poorly without the presence of alcohol.

PREVALENCE

The prevalence of alcohol dependence and abuse places a tremendous social and economic strain on the United States each year. Roughly 4-14% of the U.S. population meet the criteria for alcohol abuse or dependence (Regier et al., 1990; Grant et al., 1994; Grant et al., 2004). According to the most recent National Survey on Drug Use and Health performed by the Substance Abuse and Mental Services Administration (SAMHSA), the percentage of current alcohol drinkers in the U.S. (ages 12 and older) is 50% (SAMHSA, 2004).

Alcohol consumption has been linked to the development chronic diseases such as cancer, diabetes, cardiovascular disease, cirrhosis of the liver, depression and epilepsy. The more alcohol consumed on a daily basis, the greater the risk for disease development. In addition, alcohol consumption contributes to acute injuries, such as automobile accidents. Even moderate doses of alcohol have effects on reaction time, coordination and cognitive processing (Rehm et al., 2003). Alcohol use also increases the risk of injury from automobile and bicycle accidents, falls, fires, sports activities and fights by increasing their prevalence and preventing the body from recovering as quickly after incurring damage.

It is not surprising that the economic impact of alcohol abuse and dependence in the U.S. is substantial. The estimated cost of alcohol abuse in 1998 was $185 billion (Harwood, 1998). By applying the same formula used to generate this estimate, the cost in 2006 would be $249 billion. That estimate equates to $833 per person, based on the U.S. population in July 2006 (http://www.cia.gov). Therefore, each U.S. citizen is

contributing over $800 per year to the cost of alcohol abuse treatment, prevention, research and training, medical consequences of alcohol consumption, lost productivity at work, crime and property damage (Harwood, 1998). Obviously this is an overwhelming burden to society that calls for a better understanding of the biology of the disease so that more effective prevention and treatment options are available.

TREATMENT

Treatment options for alcohol abuse and dependence fall into two broad categories: non-pharmacotherapy and pharmacotherapy. Non-pharmacotherapy includes a wide range of behavior modification techniques and programs, including Cognitive-Behavioral Therapy, Motivational Enhancement Therapy, Couples Therapy, Brief Interventions and 12-step self-help programs such as Alcoholics Anonymous (AA) (Fuller and Hiller-Sturmhofel, 1999; Steinglass, 1999).

Only three medications are currently approved by the U.S. Food and Drug Administration (FDA) for treating alcohol dependence. They are disulfiram, naltrexone and acamprosate. Disulfiram is an aldehyde dehydrogenase blocker, which causes an increase in acetaldehyde concentrations in the body, resulting in unpleasant side effects (Heilig and Egli, 2006). Naltrexone and acamprosate both act on receptor systems in the brain. Naltrexone is an opiate antagonist that blocks the rewarding effects of alcohol and acamprosate is a glutamate antagonist thought to decrease the effects of the excitatory neurotransmitter glutamate, thereby increasing abstinence (Volpicelli et al., 1992; Swift, 1995; Dahchour and De Witte, 2000; Littleton and Zieglgansberger, 2003; Heilig and Egli, 2006). A number of other drugs are currently being tested to treat dependence. These include baclofen, which activates $GABA_B$ receptors and reduces alcohol craving and consumption (Addolorato et al., 2002), topiramate, which reduces craving by acting on AMPA-kainate receptors to decrease glutamate activity while increasing GABA

activity (Johnson et al., 2003) and ondansetron, which reduces the activity of the serotonin (5-HT_3) receptor and reduces the desire to drink (Johnson et al., 2000). In addition, a recent study has tested the Chinese herbal root, kudzu, for its effectiveness in reducing alcohol intake in humans (Lukas et al., 2005). The study found that participants drank less beer in a single sitting after being treated with kudzu extract for seven days.

Despite the wide mechanistic range of current medications to treat dependence, none are universally effective. This disparity is likely the result of genetic makeup. The effectiveness of the drug ondansetron is a perfect example of how a drug can selectively work on a subset of people. Ondansetron is more effective in early-onset alcoholics (alcohol-dependent before age 25) than late-onset alcoholics (alcohol-dependent after age 25) (Johnson et al., 2000). The hypothesis currently being tested is that this observation may be the result of a polymorphism in the promoter region of the 5-HT_3 receptor gene between these two types of alcoholics. Therefore, treatments may need to be tailored to individuals or subsets of alcoholics to be most effective. It is crucial to identify both "predisposition" genes as well as genes affected by alcohol, so that clinicians may be able to identify individuals at risk for developing the disease and treat them with the most appropriate drug.

PHYSIOLOGY

Alcohol Metabolism

Alcohol is metabolized in a two-step process in the liver. Alcohol dehydrogenase (ADH) converts alcohol to acetaldehyde, which is then converted to acetate by aldehyde dehydrogenase (ALDH). Acetate subsequently enters the Krebs cycle to become carbon dioxide and water. Depending on the quantity consumed and the rate of consumption, a

small portion of the alcohol that is imbibed does not get metabolized in the liver and is instead excreted through the breath and urine (Wallgren, 1970).

The metabolism of alcohol in the body can be described as zero order metabolism, meaning the rate of metabolism does not change with quantity consumed. It takes 2-3 hours for a male's blood alcohol concentration (BAC) to return to zero after one standard drink (12 ounces of beer, 5 ounces of wine, or 1.5 ounces of 80 proof distilled spirits) on an empty stomach. It takes seven hours after four standard drinks. As a consequence of this slow rate of metabolism, alcohol can accumulate in the body, affecting organ systems and producing a state of intoxication (Wilkinson et al., 1977).

Alcohol metabolism is affected by genetic makeup. A polymorphism in the ALDH gene (called ALDH2*2) causes the enzyme to degrade acetaldehyde at a slower rate than normal. Since acetaldehyde is mildly toxic, many unpleasant side effects, such as an increased pulse, dizziness, headaches and nausea can occur when alcohol is consumed by someone with this mutant allele. Consequently, this allele appears to have a protective effect against the development of alcohol dependence (Kwon and Goate, 2000).

Chronic exposure

Chronic alcohol abuse affects almost every organ in the body. In addition to the myriad effects alcohol has on the brain, it can also have serious effects on the liver, the immune, cardiovascular and skeletal systems. More people die in the U.S. from liver disease caused by long term, heavy alcohol use than any other form of liver disease (DeBakey et al., 1996). Alcohol can decrease immune system efficiency, making abusers more susceptible to infection from bacterial pneumonia, pulmonary tuberculosis and hepatitis C (Cook, 1998). Cardiovascular illnesses such as cardiomyopathy, coronary heart disease, high blood pressure, arrhythmia and stroke can all be the result of chronic

heavy drinking (Lucas et al., 2005). Heavy drinking can also decrease bone mass and lead to osteoporosis (Chakkalakal, 2005).

Even though long term alcohol abuse can have severe consequences on the body as a whole, alcohol dependence is a disease of the brain. At the cellular level, alcohol is thought to alter brain gene expression and protein production as well as alter the structure of pre- and post-synaptic cells. When these changes become permanent, they are referred to as a "neuroadaptation" of the brain to the presence of alcohol. Neuropathological changes are also evident after chronic abuse. Magnetic Resonance Imaging (MRI) studies of human subjects who drank heavily for 1-2 decades show loss of cortical gray and white matter (Pfefferbaum et al., 1992) and a reduction in brain volume (Agartz et al., 1999). Post-mortem studies of human brain tissue from chronic alcoholics confirm this finding, revealing neuronal loss in the frontal cortex (Kril et al., 1997).

Chronic alcoholics often suffer from vitamin deficiency, especially of thiamine (vitamin B1). Up to 80% of alcoholics have this deficiency (Morgan, 1982) and many develop Wernicke-Korsakoff syndrome (WKS) as a result (Martin et al., 2003). The symptoms of this syndrome include mental confusion, muscle incoordination and problems with learning and memory (Victor et al., 1989). Brain damage can also occur indirectly as a result of liver damage. For example, cirrhosis of the liver can cause the liver to function improperly, which leads to a condition known as hepatic encephalopathy. Some of the symptoms of this disorder include anxiety, depression, cognitive deficits, personality changes and incoordination (Butterworth, 2003).

Acute exposure

Binge drinking is defined as drinking five or more drinks on the same occasion, within a couple hours of each other (SAMHSA, 2004). Almost 23% of Americans aged 12 or older reported having had at least one binge drinking episode in the 30 days prior to

the 2004 SAMHSA survey. College age students (18-25 years old) report a higher frequency of binge drinking than any other age group (SAMHSA, 2004).

Although some of the immediate effects of high alcohol consumption may be desirable, such as its anxiolytic and sedative effects, drinking excessive amounts of alcohol in a short period of time can cause severe consequences. For example, acute alcohol intoxication can result in loss of balance and motor coordination, cognitive dysfunction, memory impairments, vomiting (with possible suffocation if person is unconscious) and blackouts (White, 2003).

A number of studies have been done in rats that support both the behavioral and cellular impairments seen in humans after acute intoxication. High doses of alcohol administered intragastrically to adult rats for 4-10 successive days caused neurodegeneration in the olfactory bulbs, piriform cortex, perirhinal cortex, entorhinal cortex and dentate gyrus (Collins et al., 1998; Obernier et al., 2002b). The neurodegeneration occurring in these areas was attributed to necrotic cell death, as opposed to apoptosis (Obernier et al., 2002a). Furthermore, as exposure increased over time, an increase in the number of regions affected was observed. Additionally, Obernier et al. discovered that after the alcohol-treated animals experienced withdrawal, they performed more poorly on the reversal learning task than the control animals (Obernier et al., 2002b). The authors propose that this behavior could be a result of cognitive dysfunction caused by the binge alcohol paradigm. Therefore, long term alcohol abuse is not necessary to produce physical damage to the brain, and consequently, cognitive impairment.

Neurobiology of alcohol dependence

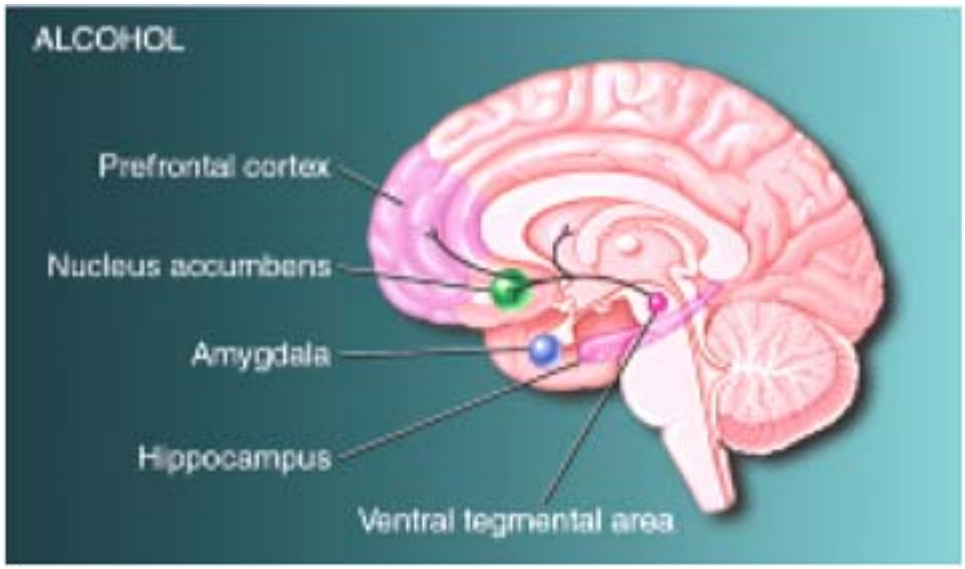
BEHAVIORAL AND PHYSICAL MANIFESTATIONS

The key to becoming dependent on alcohol is believed to involve neuroadaptation of the brain to the presence of alcohol. Neuroadaptation most commonly occurs after repeated exposure to alcohol over many years (Nestler and Aghajanian, 1997), but it is possible to become dependent on alcohol after just a single exposure (Friedman, 1980; Kosobud and Crabbe, 1986; McCaul et al., 1990; Newlin and Pretorius, 1990; Crabbe et al., 1991) Stable changes in brain chemistry (neuroadaptation) can be manifested as behaviors associated with alcohol abuse and dependence. These complex behaviors include reward, craving, sensitization, tolerance and physical dependence (withdrawal) (Nestler and Aghajanian, 1997).

The high that is often felt after consuming alcohol contributes to the compulsion to drink again. Alcohol has positive reinforcing properties that act on the brain's natural "reward" pathway (Figure 1.1). This pathway is normally stimulated by natural rewards such as food and sex, but most drugs of abuse, including alcohol, stimulate it as well (Kelley and Berridge, 2002). This pathway, called the mesocorticolimbic dopamine (DA) system, is comprised of DA neurons in the ventral tegmental area (VTA) projecting mainly to the nucleus accumbens (Nacc), amygdala and prefrontal cortex (PFC) (Kelley and Berridge, 2002). A mechanism known to be involved in motivation and learning has been hypothesized to also contribute to the rewarding and reinforcing effects of alcohol. Co-activation of DA D1 receptors and glutamate N-methyl-D-aspartate (NMDA) receptors in the Nacc and PFC causes transcriptional consequences proven necessary for reward learning (Kelley and Berridge, 2002). Since alcohol targets NMDA receptors and stimulates DA release, it is likely that these same pathways are at least partly responsible

for alcohol's rewarding effects. In addition, the constant stimulation of the reward pathway seen with chronic alcohol abuse can cause DA hypofunction. Therefore one must consume more alcohol to make up for the decrease in DA release, thereby promoting reward and reinforcement (Weiss and Porrino, 2002). Reward in animal models of alcohol dependence is often measured using two-bottle choice or lever press or by performing conditioned place preference. A correlation has been shown between the level to which the animal self-administers or seeks out the environmental cues associated with receiving alcohol and the perceived reward (Crabbe et al., 1994). The greater the reward, the more likely the animal (or human) will drink again. Repeated alcohol exposure promotes brain neuroadaptation and eventually dependence.

Figure 1.1: Reward pathway for alcohol in human brain. The mesocorticolimbic dopamine (DA) system is comprised of DA neurons in the ventral tegmental area (VTA) projecting mainly to the nucleus accumbens (Nacc), amygdala and prefrontal cortex (PFC).



(Neuroscience, 2005)

A number of drugs of abuse produce a stimulant response upon exposure (Wise and Bozarth, 1987). Cocaine, amphetamine and morphine are most commonly associated with this response, however, alcohol can cause this effect as well (Phillips et al., 1997). The stimulant response is augmented upon repeated exposure to the drug, a process termed behavioral sensitization (Phillips et al., 1997). One way to assess behavioral sensitization in mice is by measuring locomotor response. It has been hypothesized that behavioral sensitization accompanies euphoria and reward of the drug, possibly leading to uncontrolled drinking and the development of alcohol abuse and dependence. Brain neuroadaptation may also be occurring during sensitization and may be a contributing factor to the feelings of euphoria and reward (Hunt and Lands, 1992; Schmidt et al., 2000). Evidence for neuroadaptation comes from studies where the sensitization phenotype has persisted for 3 months to 1 year, depending on the drug of abuse (Lessov and Phillips, 1998), suggesting that stable changes in brain chemistry may be occurring to maintain the phenotype.

While sensitization refers to an increase in response to alcohol after repeated exposure, tolerance refers to a decrease in response. Tolerance is defined as the reduction in response to alcohol after repeated exposure or the need for increased amounts of alcohol to achieve the same effects as when it was first introduced (Littleton et al., 1980; Le et al., 1992). Basically, a higher dose of the drug is required to produce the same behavioral effects. A rapid form of tolerance, known as acute (functional) tolerance, develops quickly (within minutes) (Palmer et al., 1985) after a single exposure to alcohol (Mellanby, 1919; LeBlanc et al., 1975; Erwin and McClearn, 1981; Erwin and Deitrich, 1996). Tolerance can be assessed using any behavioral, physiological or biochemical process that is affected by alcohol. Examples in rodents include loss of righting reflex, motor incoordination, balance and hypothermia (Le et al., 1992). The development of a

functional tolerance to the behavioral effects of alcohol implies that at least some neuroadaptation is occurring, suggesting that tolerance may be a prerequisite to a physically dependent state (Grieve and Littleton, 1979).

Alcohol withdrawal is a clinical syndrome that can affect people who regularly consume alcohol and who decrease or stop their alcohol consumption suddenly (Saitz, 1998). Both acute and chronic alcohol use causes neural cells, circuits and systems to adapt to the presence of alcohol, such that when it is removed, withdrawal symptoms occur (Hyman and Malenka, 2001). For example, long term alcohol abuse leads to an up-regulation of glutamate receptors and a down-regulation of GABA receptors. Upon cessation of alcohol's presence, more glutamate can bind to the abundance of glutamate receptors and less GABA can bind to the decreased number of GABA receptors. The consequence of this is over-excitation in the brain, which is what causes withdrawal seizures (Oscar-Berman and Marinkovic, 2003). The severity of the withdrawal syndrome in rodents can be assessed using the Handling Induced Convulsion (HIC) scoring system (Goldstein and Pal, 1971; Goldstein, 1972; Crabbe et al., 1991; Terdal and Crabbe, 1994). Two of the more common methods for eliciting withdrawal symptoms in mice are exposing them to alcohol vapor for three days (Goldstein and Pal, 1971; Goldstein, 1972; Terdal and Crabbe, 1994) and injecting them with a single, high dose of alcohol (Crabbe et al., 1991; Metten and Crabbe, 1994).

In conclusion, the factors that contribute to becoming dependent on alcohol are part of a vicious cycle. As soon as alcohol is imbibed and its rewarding effects are felt, a concomitant change in brain chemistry and desire to continue drinking occur. If this pattern continues over a long period of time, tolerance can develop and more alcohol will be necessary to maintain the same euphoric effect. The body eventually becomes physically dependent on the alcohol and therefore if consumption stops, severe emotional

and physical consequences are felt, including withdrawal symptoms. It is not the case that anyone who imbibes alcohol is destined to follow this pattern and become dependent on alcohol. Genetic makeup plays a major role in the risk for developing dependence, which will be discussed in detail later.

CELLULAR AND MOLECULAR CONSEQUENCES

Cell-cell communication between neurons in the brain allows the brain to store information, regulate basic body functions and direct behavior. Neurons communicate via neurotransmitters and their receptors. Neurotransmitters are small molecules or peptides that are released by a pre-synaptic neuron, travel across the synapse (the small gap between neurons) and bind specific receptors on the post-synaptic neuron. A series of molecular events follow inside the post-synaptic cell, depending on what type of receptor is activated.

Alcohol disrupts the normal homeostasis of neuronal communication by altering the location and production of neurotransmitters, receptors, transporters, signaling molecules, second messengers and transcription factors (For a review, see (Diamond and Gordon, 1997)). Alcohol affects several key neurotransmitters and neuromodulators in the brain, including gamma-aminobutyric acid (GABA) (Mihic and Harris, 1997), glutamate (Gonzales and Jaworski, 1997), dopamine (Di Chiara, 1997), serotonin (Lovinger, 1997) and opioid peptides (Froehlich, 1997). Two of the more well-studied systems are glutamate and the NMDA receptor and GABA and the $GABA_A$ receptor. Glutamate is the major excitatory neurotransmitter in the brain and GABA is the major inhibitory one. Alcohol inhibits neuronal communication by decreasing the excitatory actions of glutamate and increasing the inhibitory actions of GABA (For a review see (Diamond and Gordon, 1997)). NMDA receptor activity is altered by chronic alcohol consumption and studies have linked altered receptor activity to fetal alcohol syndrome,

physical dependence and cognitive impairment (Gonzales and Jaworski, 1997). Both acute and chronic alcohol exposure can alter $GABA_A$ receptor function and several studies have shown that $GABA_A$ receptors may play a role in the development of alcohol tolerance and dependence (Mihic and Harris, 1997). Dopamine is a neuromodulator involved in motivation and reinforcement. Alcohol increases the firing rate of dopamine neurons in the VTA of the midbrain thereby stimulating the release of dopamine in the Nacc. The activation of this reward pathway by alcohol contributes to the compulsion to drink again (Di Chiara, 1997). Serotonin and its receptors have been implicated in alcohol consumption and reinforcement (Valenzuela, 1997). Serotonin receptor antagonists reduce alcohol drinking in certain animal models (Diamond and Gordon, 1997) and specific receptor subtypes can stimulate DA release in the Nacc when activated by alcohol (Valenzuela, 1997). Opioid peptides have also been implicated in alcohol consumption and reinforcement (Froehlich, 1997). Opioid system inhibitors decrease alcohol self-administration in animals and alcohol consumption in humans (Froehlich, 1997). Acute alcohol consumption stimulates the release of endogenous opioids (Gianoulakis, 2001) that bind to opioid receptors, which in turn activates the mesolimbic DA reward pathway (Herz, 1997).

In addition to altering communication between cells and brain regions, alcohol also alters communication within cells. When neurotransmitters bind to metabotropic receptors on the surface of neurons, a signal transduction cascade is begun inside the neuron. Molecules called second messengers, kinases and phosphatases (the latter two alter protein phosphorylation) are a part of this signaling cascade. Some of these signals are transmitted to the nucleus to regulate gene transcription. For example, many transcription factors are regulated by phosphorylation. Acute and chronic exposure to alcohol causes cyclic AMP (cAMP) (a second messenger) levels to increase and

decrease, respectively. cAMP regulates the activity of protein kinase A (PKA), and it too has been implicated in mediating cellular responses to alcohol. Numerous studies have demonstrated that the activity of another kinase, protein kinase C (PKC), is also affected by alcohol (Diamond and Gordon, 1997). Acute alcohol exposure has been shown to cause an increase in CREB (a transcription factor) phosphorylation and CRE-mediated gene transcription, whereas chronic exposure decreased CREB activation (Pandey et al., 2005). Partial deletion of the CREB gene has been shown to increase alcohol drinking in mice (Pandey et al., 2004). Some members of the *fos* and *jun* families of immediate early genes are activated in response to acute alcohol administration in rodents (Rhodes and Crabbe, 2005). By altering transcription factors such as *fos* and *jun*, alcohol can indirectly affect the expression of hundreds of other genes as well.

Finally, a number of studies have shown that several proteins have alcohol-sensitive binding sites. Mihic et al. demonstrated that an alcohol binding pocket exists between transmembrane segments 2 and 3 (TM2 and TM3) of the $GABA_A$ and glycine receptors' subunits (Mihic et al., 1997). Furthermore, it was discovered that the size of the amino acid (and consequently, the binding pocket itself) at a specific site on TM2 was directly proportional to the magnitude and direction of alcohol's effects (Ye et al., 1998). Two amino acid sites on the $\alpha 2\beta 4$ nicotinic acetylcholine receptor (nAChR) were found to be sensitive to short- and long-chain alcohols, producing enhancement or inhibition of the receptor, respectively (Borghese et al., 2003). In addition, alcohol potentiation of DA transporter (DAT) function can be abolished by introducing specific amino acid mutations in the second intracellular loop between TM2 and TM3 of the protein (Maiya et al., 2002).

Based on all of these studies, it is clear that alcohol affects a wide range of receptors and neurochemical systems in the brain. The brain starts to adapt to these

molecular and cellular alterations almost immediately. The current hypothesis is that over time, some of these changes become permanent. Stable changes in receptor amount, intracellular signaling, projections between brain regions and ultimately, basic brain function are thought to be caused, in part, by changes in gene expression. Stable alterations in gene expression are thought to contribute to the manifestations of alcohol dependence, including tolerance, withdrawal symptoms and the ultimate loss of control felt with regard to alcohol consumption (Hyman and Malenka, 2001; Nestler, 2001; Nestler and Landsman, 2001).

THE PREFRONTAL CORTEX (PFC)

The prefrontal cortex (PFC) in humans is located in the frontal lobes, anterior to the motor cortices. The PFC is responsible for higher cognitive thinking, such as problem solving, critical thinking, planning, judgment and impulse control (Funahashi, 2001). The ability to make appropriate decisions regarding everyday behavior comes from the PFC's ability to maintain a "working memory" (Baddeley, 1986). When faced with a new environment, the PFC recalls a number of stored memories simultaneously, guiding the decision-making process and preventing an inappropriate behavioral response (Krawczyk, 2002). The PFC has been defined in the rat, both structurally and functionally (Heidbreder and Groenewegen, 2003; Uylings et al., 2003). Similar to humans, the rat PFC has been implicated in working memory, initiation of response and management of emotions (Heidbreder and Groenewegen, 2003). Topographically, the mouse PFC closely resembles that of the rat (Guldin et al., 1981), although less is known about its function. Insight into the molecular and cellular characteristics of the mouse PFC, however, has been gained through studies of gene expression (Kerns et al., 2005) and neurotransmitter release (Ihalainena et al., 1999; Douglas et al., 2001, 2002).

Both short and long-term alcohol abuse have been shown to damage the PFC and surrounding areas in humans and rats. The frontal lobes are particularly susceptible to the effects of chronic alcohol abuse in humans, where brain shrinkage as well as a loss of cortical neurons have been observed (Kril et al., 1997; Kril and Halliday, 1999). Acute alcohol exposure was found to impair planning and spatial recognition in social drinkers (Weissenborn and Duka, 2003). In rodent models of binge drinking, neurodegeneration was observed in the olfactory bulbs and selected regions of the cerebral cortex (piriform cortex, perirhinal cortex, entorhinal cortex and dentate gyrus) (Obernier et al., 2002b).

The PFC is an important brain region with which to study the development of dependence due to its connection to the mesocorticolimbic DA pathway. DA is released into the PFC upon activation of the VTA by alcohol (Bowirrat and Oscar-Berman, 2005). The increase in DA is associated with feelings of euphoria, thereby reinforcing the positive effects of alcohol exposure.

The effects of alcohol exposure on gene expression in the mouse PFC has only recently been explored (Kerns et al., 2005). It is obvious that this brain region is affected in many ways by alcohol consumption in humans. Deficits in decision-making and judgment along with feelings of euphoria and reward all contribute to the motivation to drink again, even after negative consequences are felt. Evidence suggests that even a few binge drinking episodes can alter brain function both physically and mentally, priming the brain for an eventual state of alcohol dependence. By identifying gene expression changes in the mouse PFC after acute alcohol administration over time, insight may be gained into the neuroadaptation occurring in this part of the reward pathway to enhance reinforcement and contribute to the development of dependence.

Genetics of alcohol dependence

Alcohol dependence is a complex polygenic disease. The risk for becoming dependent on alcohol is influenced by both environmental factors and genetic makeup. Several other common diseases, such as cancer, heart disease, diabetes and obesity are also complex in nature. The prevalence, severity and complexity of all these diseases make the science behind them difficult to solve. One of the most challenging obstacles faced when trying to identify genes that contribute to a disease phenotype is the fact that hundreds of genes may only individually exert a small effect on the disease. Alcohol dependence is no exception.

HUMAN

Several twin and adoption studies have explored the genetic contribution to alcohol dependence. Studies of adoptees raised apart from affected and unaffected parents found the disease to be about 40% heritable (Goodwin et al., 1973; Cloninger et al., 1981). Twin studies using both identical and fraternal same sex twins (male and female), found a range of genetic contribution, from 35% to 64% (Pickens et al., 1991; Kendler et al., 1992; Heath et al., 1997; Kendler et al., 1997).

Based on the results of these studies, the Collaborative Studies on Genetics of Alcoholism (COGA) was founded in 1989 to identify and map susceptibility genes for alcohol dependence in humans (Begleiter et al., 1995). With approximately 300 families consisting of greater than 3000 individuals (Edenberg, 2002), the COGA has provided alcohol researchers with a large genetic human "database" with which to perform candidate gene and linkage studies to identify genes and markers that may be associated with increased risk for developing dependence. For example, electroencephalographic (EEG) patterns and specific personality traits are two heritable endophenotypes (a trait

that has been shown to be associated with a specific condition) that have been shown to be correlated to alcohol use (Reich et al., 1999). The amplitude of the P3 event-related brain potential (ERP) and four stable quantitative dimensions of personality (Cloninger et al., 1993) can be measured and potentially linked to chromosomal regions, thereby indicating where susceptibility genes for dependence reside (Reich et al., 1999). EEG patterns and personality traits have been used to identify genetic linkage on chromosomes 2, 6, 5 and 13 (Begleiter et al., 1998) and chromosome 8 (Cloninger et al., 1998), respectively.

Alcohol dependence is a human disease with characteristics that are difficult to fully recapitulate in other organisms. However, linkage studies in humans often lack the power to find the definitive loci associated with a particular trait. Therefore, genetic studies using animal models of alcohol dependence serve as complementary methods to human studies in detecting the genes that may contribute to alcohol dependence.

MOUSE

The mouse genome was recently sequenced (Waterston et al., 2002), allowing researchers to directly compare the genetic makeup of the mouse to human (Lander et al., 2001; Venter et al., 2001). As of 2002, the mouse genome contained about 30,000 genes, of which 99% have direct counterparts in humans. Therefore, genes found in the mouse to be affected by alcohol or predisposing to dependence may directly correlate with genes in human. The mouse has become a valuable tool with which to study the genetics of alcohol dependence due to its ability to model specific phenotypes of the disease and the wide variety of experimental techniques that exist to study those phenotypes.

Experimental techniques

Mice have been used to study the inheritance of traits for at least 3000 years (Beck et al., 2000). In the early 1900s, mice started to become a valuable tool in the laboratory to study inheritance. Clarence C. Little established the first inbred strain of mouse. An inbred strain is created by performing brother x sister matings for 20 or more consecutive generations (Beck et al., 2000). Now, over 450 inbred strains exist (Beck et al., 2000). They are an invaluable tool to genetic researchers because all of the same sex mice in an inbred strain are genetically identical to each other and 98.6% of the loci in each mouse are homozygous (Beck et al., 2000). This means that experiments done with a particular inbred strain from 1980 in Portland, Oregon can be compared to data from experiments performed in 2005 in Austin, Texas.

Another breeding technique that started being utilized in addiction studies in the 1940s is selective breeding (Crabbe, 2002a). Heterogeneous mice can be screened for specific behaviors and then bred according to the intensity of those behaviors. So, all of the mice with the most or least severe form of the trait are bred (Grahame, 2000). Furthermore, selected lines of mice can be inbred as well to make them genetically identical.

Alcohol dependence is a quantitative trait, meaning multiple genes contribute to the disease and the severity in people varies continuously across a population, often following a normal distribution (Crabbe et al., 1994). Regions of chromosome that contain a gene or genes influencing a quantitative trait are called Quantitative Trait Loci (QTL) (Grisel, 2000). A technique called QTL analysis is used to identify chromosomal regions that are correlated with a specific trait. The development of Recombinant Inbred (RI) strains has made QTL analysis much easier. The BxD RI strains were created by crossing C57BL/6J (B6) mice with DBA/2J (D2) inbred strains of mice (Taylor, 1978).

In a B6 x D2 cross, all of the F1s are identical to each other, possessing one allele from the B6 mouse and one from the D2 mouse. In order to create genetic diversity, the F1s are crossed. As a result of recombination, individual F2s contain a unique pattern of parental alleles. Brother-sister pairs of F2s are mated to create inbred strains of each of the F2s. Therefore, each inbred strain has a unique combination of B6 and D2 alleles fixed in a certain place on the chromosome. 26 inbred strains were created using this technique, with over 1300 polymorphic alleles genotyped to date (Grisel, 2000). A correlation between genotyped alleles and specific biochemical or behavioral responses are used to identify a QTL (Tabakoff and Hoffman, 2000). Once a QTL has been identified, the next step is to identify candidate genes within the QTL that are influencing the trait of interest directly or through epistasis.

The mouse genome can be manipulated to study specific genes of interest by creating transgenic, knockout and knock-in mice. To create a transgenic mouse, a foreign gene is introduced into the mouse genome, which often leads to an over-expression of the mRNA and protein product of the inserted gene. Although it is possible to control the timing and/or location of the gene's expression by using a specific promoter, it is impossible to predict where the gene will be introduced into the genome. The transgene can randomly insert anywhere and may disrupt other genes, promoters or intron/exon boundaries (Bowers, 2000). To create a knockout or knock-in mouse, the gene of interest is mutated to render it non-functional (knockout) or alter its function from wild type (knock-in). It is then inserted into the genome at its normal location on the chromosome using targeted mutagenesis. If the wild-type version of the gene is vital to development, the mouse may not live to adulthood (Bowers, 2000).

Models of alcohol-related behavior

Alcohol dependence is a disease that only truly affects humans. Therefore, humans are theoretically the best subjects with which to study the disease and its genetic component. However, it is not possible to perform certain types of experiments using humans as the subjects. The mouse offers a viable alternative, due to its short gestation period, size, availability and known genome that can be manipulated with relative ease. Although no mouse model exists that fully recapitulates all aspects of alcohol dependence, many alcohol-related traits have been modeled and studied successfully in mice using inbred strains, selected lines or genetically altered mice.

A difference in alcohol preference was observed between the B6 and D2 inbred strains in 1959 by McClearn (McClearn, 1959). The B6 mice preferred drinking a 10% alcohol solution over water while the D2 mouse avoided the alcohol solution. A correlation between alcohol consumption and withdrawal severity has been shown in both inbred strains and selected lines. The mice that drink more tend to have less severe withdrawal symptoms and vice-versa (Metten et al., 1998). Selected lines of mice have been bred to study initial sensitivity to alcohol via its hypnotic effects (Long-Sleep (LS) and Short-Sleep (SS)), severity of withdrawal from chronic alcohol exposure (Withdrawal Seizure-Prone (WSP) and Withdrawal-Seizure Resistant (WSR) mice), degree of acute functional tolerance to ataxia caused by alcohol exposure (High/Low Acute Functional Tolerance (HAFT1, -2 and LAFT1, -2) and preference for a 10% alcohol solution over water (High/Low Alcohol-Preferring mice (HAP1, -2 and LAP1, -2)) (Crabbe et al., 1994).

QTL analysis has been used in combination with the BxD RI strains to identify QTL for alcohol preference drinking on chromosomes 1, 2, 4 and 9 and acute alcohol withdrawal on chromosomes 1, 4 and 11 (Crabbe et al., 1999). Other QTL for chronic

alcohol withdrawal, conditioned taste aversion, conditioned place preference, tolerance to the rotarod and hypothermia have also been identified (Crabbe, 2002b).

Several transgenic and knockout mice have been created to study the effects of specific genes on alcohol-related phenotypes. Some of the genes chosen for study were 5-HT_3 receptor, neuropeptide Y, dopamine D_1 and D_2 receptors, fyn kinase, protein kinase A and C and a selection of $GABA_A$ receptor subunits (Bowers, 2000; Boehm et al., 2004).

What all of these studies demonstrate is that there is a clear genetic component to many if not all alcohol-related phenotypes. All of the models used to study these phenotypes can give insight into the genetics behind them, and hopefully by extension, the genes contributing to the development of alcohol dependence. Ideally, given the polygenic nature of the disease, these mouse models would be combined with other techniques, such as gene expression microarrays, to increase the number of genes studied at once.

C57BL/6J (B6) and DBA/2J (D2) strains

Two of the most commonly used inbred strains of mice are B6 and D2. The D2 strain was the first inbred strain created, produced by Clarence Little in 1909 (Silver, 1995). The B6 strain was created soon after D2. As with all inbred strains, due to their fixed genetic makeup, B6 and D2 mice have some abnormal characteristics, such as a high susceptibility to diet-induced obesity (B6), late-onset hearing loss (B6), high-frequency hearing loss (D2) and progressive eye abnormalities (D2) (For more information, see http://www.jax.org). Fortunately for alcohol researchers, they have divergent phenotypes in many alcohol-related traits as well. It was McClearn who first noticed in 1959 that the B6 strain preferred to take nearly all of its daily fluid from a 10% alcohol solution, as opposed to tap water, and that the D2 mouse was just the opposite

(McClearn, 1959). One of the most striking differences between these two strains is their susceptibility to physical dependence, as evidenced by withdrawal symptoms following alcohol removal. B6 mice experience extremely mild withdrawal symptoms, while D2 mice experience severe withdrawal symptoms (Crabbe et al., 1983; Metten and Crabbe, 1994). Alcohol withdrawal severity also varies widely in human alcoholics, with manifestations ranging from mild insomnia to delirium tremens or death (Saitz, 1998).

The B6 and D2 strains are easily obtainable, commonly used mice that differ greatly in alcohol preference and withdrawal severity, making them ideal models with which to study the genetics of alcohol dependence.

Gene expression and alcohol dependence

The sequencing of the human genome was one of the most important scientific achievements of the last 100 years (Lander et al., 2001; Venter et al., 2001). Knowing the sequence of nucleotides along each chromosome, however, is only the first step in understanding how those nucleotides combine to form genes (and ultimately, how those genes function in the cell). The central dogma of molecular biology is that DNA is transcribed to mRNA (genes), which is translated to protein. It is estimated that the human genome contains 25,000 (protein coding) genes (IHGSC, 2004), a much smaller number than the once thought 100,000 genes. On average, each gene is translated into 2-3 different proteins, a consequence of alternative splicing. Differential gene expression in cells at various stages of development and throughout life is the mechanism responsible for specific cell types, tissues and organs in the body.

GENE EXPRESSION DETECTION AND ANALYSIS

The combination of gene expression changes occurring in a cell at any given time can be referred to as its gene expression profile. Alterations in gene expression profiles

during the normal stages of development, as a result of a diseased state or due to an environmental stimulus can give insight into the types of genes working in concert to produce a particular phenotype. Gene expression analysis can be used to study all of the following phenomena: 1) What are the gene expression differences between the normal or unaltered state of one type of cell, tissue, organ or organism versus another? Examples include: a muscle cell versus a neuron, the liver versus the brain and a chimp versus a human. 2) What are the gene expression differences between the normal or unaltered state of one type of cell, tissue, organ or organism versus the altered or diseased state? Examples include: a healthy cell versus a cancerous one, the brain of a non-alcohol drinker versus the brain of a chronic alcoholic and a wild type mouse versus a transgenic or knock-out. 3) What are the gene expression differences between the normal or unaltered state of one type of cell, tissue, organ or organism versus one that has been exposed to some type of insult or treatment? Examples include: healthy tissue versus tissue that has been exposed to a deleterious or potentially beneficial drug. 4) What are the gene expression differences between the normal or unaltered state of one type of cell, tissue, organ or organism at time point A versus time point B? Examples include: the M phase versus the S phase of the cell cycle and a fetal cell versus an aged cell (Stears et al., 2003). Identifying what genes are regulated by all these processes is the first step in understanding the mechanisms by which they are regulated and generating hypotheses about how and why those mechanisms create the phenotype of interest.

Non-microarray methods

A variety of methods are available to study gene expression. Some methods such as Northern blot analysis, RNase protection assay, Quantitative Reverse Transcription-PCR (RT-PCR), real time RT-PCR and *in situ* hybridization detect the expression of one (or a few) gene(s) at a time; while others can detect the expression of hundreds or

thousands of genes simultaneously, including mRNA Differential Display (DD), Serial Analysis of Gene Expression (SAGE), subtractive hybridization and microarray analysis.

In Northern blot analysis, RNase protection assay and *in situ* hybridization a labeled (with radioactivity or chemiluminescence) RNA or cDNA probe is generated from the gene(s) of interest. The probe is then hybridized to the sample RNA which is immobilized on a blot transferred from an electrophoretic gel (Northern blot analysis) (Alwine et al., 1977; Principalle et al., 2005), in a solution containing RNases (RNase protection assay) (Horii, 1996) or within a slice of tissue (*in situ* hybridization) (Gall and Pardue, 1969; Hofler et al., 1998; Principalle et al., 2005). In quantitative RT-PCR (Freeman et al., 1999; Joyce, 2002) or real-time RT-PCR (Bustin et al., 2005; Wong and Medrano, 2005), primers are generated for the gene of interest, the RNA samples are reverse transcribed and the cDNA is amplified. The amount of PCR product reflects the expression level of the gene of interest in the original RNA samples. Although the same amount of starting material is amplified, the amplification efficiency of different samples can vary, causing problems with quantification. There is no way to account for this inequality of efficiency in quantitative RT-PCR. However, in real-time PCR, a third probe is generated that is tagged with a fluorescent label, thereby allowing the amplification step to be visualized and the PCR product to be more accurately measured. All of the discussed methods are semi-quantitative (Cunningham, 2001).

Subtractive hybridization, SAGE and mRNA DD are three methods that have the benefit of not only being able to detect the expression of hundreds of genes at once, but also the ability to detect novel transcripts. A sequence specific probe is not used; therefore no prior knowledge of a gene or gene sequence is necessary for these techniques. In subtractive hybridization, two RNA samples are combined and common transcripts hybridize together. These common transcripts are subtracted out of the mix,

leaving only the differentially expressed genes. A cDNA library is constructed from the unique genes. In SAGE, the differentially expressed transcripts are reverse transcribed, cleaved and ligated together to form a single long piece of DNA that can easily be sequenced to determine gene identity (Cunningham, 2001). mRNA DD is the most commonly used method of the three. It was developed by Liang and Pardee in 1992 (Liang and Pardee, 1992). In DD, RNA samples are reverse transcribed into cDNA, which is then PCR amplified using a unique set of primers. The 3' primer is an anchored T-mer that contains 11 Ts and a single base anchor (A, G or C) at the primer's 3' end. The 5' primer is one of 80 13-mers of arbitrary sequence (Liang and Pardee, 1998). Each RNA sample is therefore amplified 240 times. This unique design should statistically represent the majority of expressed eukaryotic mRNAs in a cell. The distance separating the 3' primer from the annealing location of the 5' primer varies, usually producing PCR products of 100-500 base pairs in length (Liang and Pardee, 1992). The PCR products are resolved on a denaturing polyacrylamide gel and appear as multiple bands of different sizes in each lane. By comparing the location (size), presence (or absence) and intensity of bands in different samples, genes that are differentially expressed between the samples can easily be detected. Radioactively labeled dATP is added to the mix of dNTPs so that the bands can be visualized on the gel. Differentially expressed genes are excised from the gel and sequenced to determine gene identity. DD has several advantages over other gene expression detection methods, including: 1) DD can detect previously unidentified transcripts, 2) DD can detect both rare and abundant transcripts, 3) DD can detect insertions and deletions in transcripts and 4) DD can theoretically detect every differentially expressed gene between RNA samples, based on the use of the arbitrary and anchored T-mer primers (For more detail, see http://www.genhunter.com). Although

these advantages do make DD an attractive method for gene expression detection, DD can be very time consuming and difficult to perform well.

Microarray analysis is a fourth method of gene expression detection that has continued to gain popularity since its inception 11 years ago. It is discussed in the next section.

Microarray analysis

In 1995, a group of scientists at Stanford published a paper using a new method of gene expression detection they had developed called cDNA microarray analysis that would revolutionize the way high-throughput gene expression profiling was performed (Schena et al., 1995). They printed 48 cDNAs from *Arabidopsis thaliana* onto a glass microscope slide, isolated RNA from a sample of *Arabidopsis* and reverse transcribed it to make cDNA, labeled the cDNA with a fluorescent tag and hybridized it to the microscope slide to determine which genes were expressed in the original *Arabidopsis* RNA sample. Since that time, cDNA microarrays have expanded to include as many as 50,000 cDNAs printed on a single glass microscope slide. In addition, several variants of cDNA microarray analysis have been developed and include oligonucleotide, carbohydrate and protein arrays to measure expression, Single Nucleotide Polymorphism (SNP) arrays to screen for DNA mutations and polymorphisms, Comparative Genomic Hybridization (CGH) arrays identify difference in DNA copy number and Chromatin ImmunoPrecipitation (ChIP) arrays to measure DNA methylation and histone acetylation (Goldsmith and Dhanasekaran, 2004).

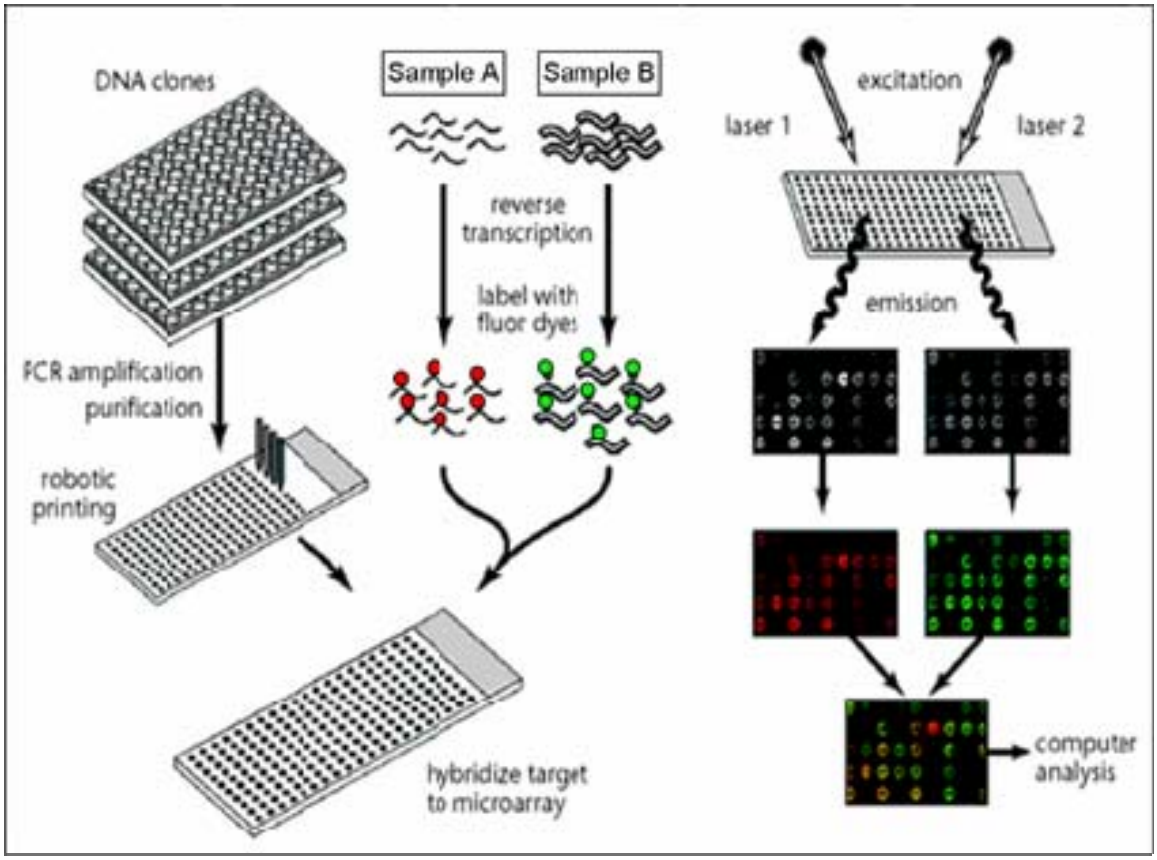
cDNA microarray analysis was originally developed to identify gene expression differences between two (or more) samples of RNA. The basic steps of cDNA microarray analysis are the following: 1) A collection of cDNAs (can be known genes or expressed sequence tags (ESTs)) are placed on a solid support (usually a glass

microscope slide or nylon membrane) at defined locations, 2) RNA is isolated from cells, tissue or an organism and reverse transcribed to create cDNA, 3) The cDNA is labeled (usually with a fluorescent tag or radioactivity) and allowed to hybridize to the solid support, 4) Based on the signal intensity of each cDNA spot on the array, information is gained about the amount and type of genes expressed in the original RNA sample. Usually, the cDNA hybridized to the slide is from two different sources of RNA. The two samples that are co-hybridized represent one of two different experimental designs: 1) A direct or "within slide" comparison, where both samples are experimental – one is a control (sample A) and the other is treated (sample B) or 2) An indirect or "between slide" comparison, where one sample is experimental (sample A or sample B) and the other is a reference (Yang and Speed, 2002). The second experimental design that employs the use of a reference is ideal for large scale studies that involve many treatment groups for example. Theoretically, an unlimited number of experimental samples could be included in the design and all of them would be comparable if they had all been hybridized to the same reference sample. An ideal reference sample would be one that is: 1) homogenous, 2) of unlimited supply and 3) contains every gene or EST printed on the microarray (so that the reference channel always has an intensity associated with it) (Churchill, 2002).

For cDNA microarrays that use fluorescently labeled cDNA and glass microscope slides as the solid support, the co-hybridized RNA (cDNA) samples must be fluorescently labeled with two different dyes. The most commonly used dyes are cyanine-3 (Cy-3) and cyanine-5 (Cy-5). After the fluorescently labeled cDNA samples are co-hybridized to the array, the array must be scanned using a machine with a pair of lasers that excite and detect the emissions of the Cy-3 and Cy-5 dyes. The cDNA spots on the scanned array image show up as green (Cy-3), red (Cy-5) or yellow (Cy-3 and Cy-

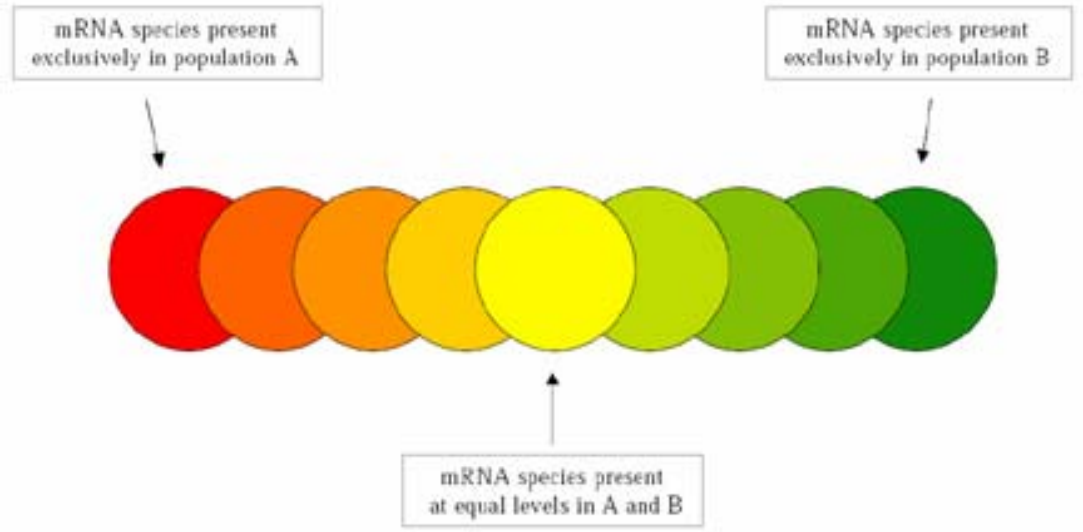
5 together), which is indicative of the gene expression profile of each of the co-hybridized samples (Figure 1.2). The color of the spot indicates in which of the two samples the gene was more highly expressed. For example, if sample A was labeled with Cy-5 and sample B with Cy-3 (Figures 1.2 and 1.3), then a red spot would indicate that the cDNA in that spot hybridized only to cDNA in sample A and that that particular gene or EST was not present in sample B. A green spot would indicate just the opposite—that the cDNA in that spot hybridized only to cDNA in sample B and that that particular gene or EST was not present in sample A. A yellow spot would indicate that that gene or EST was present in equal amounts in both sample A and B and a spot with no color would indicate that that gene or EST was not present in either sample A or B (Figures 1.3 and 1.4).

Figure 1.2: cDNA microarray overview. Two RNA samples (A and B) are reverse transcribed and labeled with Cy-5 and Cy-3 (respectively). They are co-hybridized to a microscope slide containing thousands of cDNA clones that have been printed with a robot. Lasers excite the dyes and detect their emission to produce an image with colors that indicate which genes were differentially expressed and to what level (between the two RNA samples).



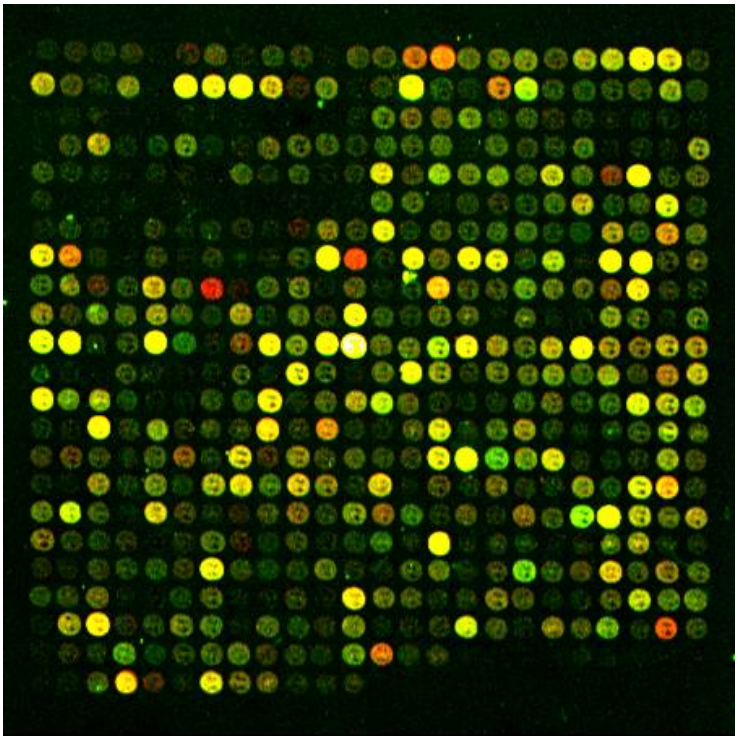
(Duggan et al., 1999)

Figure 1.3: Interpretation of colors on a scanned microarray image. The color of the spot indicates in which of the two samples the gene was more highly expressed. If sample A was labeled with Cy-5 and sample B with Cy-3, then a red spot would indicate that the cDNA in that spot hybridized only to cDNA in sample A and that that particular gene or EST was not present in sample B. A green spot would indicate just the opposite—that the cDNA in that spot hybridized only to cDNA in sample B and that that particular gene or EST was not present in sample A. A yellow spot would indicate that that gene or EST was present in equal amounts in both sample A and B.



(Nuwaysir, 2000)

Figure 1.4: Example of a cDNA microarray. A single block (over 500 genes) of a scanned microarray slide is shown. To continue with the examples using samples A and B from Figures 1.2 and 1.3, sample A is B6 reference and is labeled with Cy-5 (red) while sample B is D2 male cerebellum at the 0 hour time point after chronic alcohol exposure (72 hours in alcohol vapor chamber) and is labeled with Cy-3 (green). A spectrum of colors and intensity is present, indicating many genes are differentially expressed between these two samples.



(A.E. Berman; Genisphere 3DNA Array 350 kit; August 20, 2003)

The gene identification, color and intensity information is gathered for each spot on the array using software designed specifically for that purpose. The intensity value of each spot can range from 0 to 65,535 intensity units (Stears et al., 2003) per channel (red and green). The background-subtracted red and green intensity values from each spot are saved in a tab delimited file for further processing. The expression level for each gene or EST on the array is represented as a ratio of red to green or green to red, depending on the experimental design. When a reference sample is used, the intensity value of the reference channel is the denominator in the ratio. The expression ratios are often log (base 2) transformed because this allows up- and down-regulated genes to be treated equally—the magnitude of regulation is symmetric on either side of zero. For example, a ratio of 2:1 represents a two-fold increase in expression whereas a ratio of 1:2 represents a two-fold decrease. The quotient of 2:1 is 2, but the quotient of 1:2 is 0.5. They both represent the same absolute value of fold change but are not symmetrical with respect to zero. However, if the ratios are log transformed, 2:1 becomes 1 and 1:2 becomes -1. These values are symmetrical with respect to zero (Quackenbush, 2002; Goldsmith and Dhanasekaran, 2004).

In order to identify the genes that are differentially expressed on the arrays, it is necessary to separate technical variability from true biological differences. Technical variability can be reduced by excluding low quality or un-resolvable spots from further analysis, normalizing the data (see the Materials and Methods section for a discussion about array normalization) and filtering out low intensity or outlier values. Once a set of genes exists that truly represents the biology of the original RNA samples, a series of statistical tests, visualization tools and dimension reducers can be applied to discover the gene expression profiles of those samples (Quackenbush, 2001; Stears et al., 2003).

The ultimate goals of microarray analysis extend beyond lists of differentially expressed genes. Identifying the gene expression profiles of a particular phenotype or condition of an organism is only the first step in understanding the biology behind that phenotype or condition. Genes can be grouped by their cellular function or location, by common transcription factor binding sites or by the biological pathways to which they belong—all to gain a better understanding of how hundreds of concomitant transcriptional changes translate into phenotypic changes. For the research scientist, gene expression arrays create innumerable hypotheses to test; for the clinician, they can identify potential disease markers, differentiate between a healthy cell and a cancerous one or elucidate the biochemical basis of a drug response (Stears et al., 2003). It is conceivable that drugs could be tailored to specific subsets of the population, depending on their gene expression profiles…or that patients could purchase home testing kits to determine if they are at risk for breast cancer or cystic fibrosis (Stears et al., 2003). Several studies have already demonstrated the successful use of expression arrays to: 1) Identify prognostic biomarkers for breast cancer and predict response to therapy (Jeffrey et al., 2005), 2) Differentiate between patients with cardiovascular disease versus healthy patients (Archacki and Wang, 2004), 3) Screen patients for possible Human Papilloma Virus (HPV) infection (An et al., 2003; Kim et al., 2003) and 4) Identify the virus responsible for Severe Acute Respiratory Syndrome (SARS) (Wang et al., 2002; Ksiazek et al., 2003; Rota et al., 2003).

Of all the gene expression detection methods available, microarrays are the most attractive due to the fact that: 1) Tens of thousands of genes can be screened simultaneously, 2) Many RNA samples can be screened in a relatively short amount of time, 3) Less than 500ng of starting material is required, 4) Sophisticated instrumentation for manufacture, detection and data analysis is available, 5) The technology for

manufacturing, hybridizing and scanning the arrays, processing the data and analyzing the data is constantly being improved and updated and 6) Hundreds of types of microarrays are commercially available (Stears et al., 2003).

APPLICATION OF GENE EXPRESSION METHODS TO STUDY ALCOHOL DEPENDENCE

Studies using non-microarray methods

It is well established that alcohol dependence is about 50% heritable (Goodwin et al., 1973; Cloninger et al., 1981; Pickens et al., 1991; Kendler et al., 1992; Heath et al., 1997; Kendler et al., 1997). One approach used to identify the genetic factors that may contribute to the development of the alcohol dependence is to study gene expression.

Individual genes have been identified as differentially expressed between (alcohol-naïve) inbred strains and selected lines of mice and rats that are divergent in alcohol-related phenotypes. Both neuropeptide Y (*NPY*) and alpha-synuclein genes were identified as differentially expressed between the naïve inbred alcohol-preferring (iP) and inbred alcohol-nonpreferring (iNP) rat lines. *NPY* expression was decreased in several brain regions of the iP versus iNP rat (Spence et al., 2005), whereas alpha-synuclein expression was increased in the hippocampus of the iP versus the iNP rat (Liang et al., 2003). Three of the $GABA_A$-receptor subunit genes (alpha-1, alpha-3 and alpha-6) were less expressed in naïve WSP mice compared to WSR mice (Buck et al., 1991; Keir and Morrow, 1994).

Alcohol consumption in mice and rats alters gene expression as well. After being fed an alcohol-containing liquid diet, the brains of WSP and WSR mice were compared for differences in gene expression of $GABA_A$-receptor subunits. A decrease in the alpha-1 subunit gene was observed in the WSP line, but not the WSR line, whereas a decrease in the alpha-6 subunit gene was observed in the WSR line, but not the WSP line (Buck et

al., 1991). The expression of microtubule-associated protein 2 (*MAP2*) gene decreased in several brain regions of the alcohol-preferring AA (Alko Alkohol) rats after 16 months of voluntary alcohol consumption compared to control (Putzke et al., 1998).

Both acute and chronic paradigms of alcohol exposure have been used to detect changes in gene expression. Neural cell lines were used to identify three alcohol-responsive genes that were up-regulated compared to control: heat shock cognate 70 (*Hsc70*) (Miles et al., 1991), tyrosine hydroxylase (*TH*) (Gayer et al., 1991) and dopamine beta-hydroxylase (*DBH*) (Hassan et al., 2003). Metallothionein II (*MT-II*) gene was up-regulated in the brains of D2, A/J, BALB/cJ and B6 strains of mice after a single high injection of alcohol (acute paradigm) compared to control (Loney et al., 2003). Another study by the same group identified retinaldehyde binding protein 1 (*Rlbp1*) gene as less abundant in the brains of alcohol-treated B6 and D2 strains compared to control and syntaxin 12 (*Stx12*) gene as less abundant in the brains of alcohol-treated D2 mice compared to control, again using a single high injection of alcohol (Treadwell et al., 2004). Schafer et al. identified neuroendocrine-specific protein (*NSP*)/reticulon 1 (*RTN1*) gene as differentially expressed in the brain between the WSP and WSR lines of mice after chronic alcohol exposure (Schafer et al., 1998). Furthermore, this gene was found to be more highly expressed in the hippocampus and cortex of all alcohol-treated mice (WSP/WSR and B6/D2) over controls and less expressed in the cerebellum of WSP and D2 mice, but not WSR or B6, compared to controls (Schafer et al., 2001). Prodynorphin (*Pdyn*) gene expression was increased during chronic alcohol withdrawal in several brain regions of WSP-1 mice (Beadles-Bohling and Wiren, 2005).

Finally, studies of chronic alcoholics also reveal differential brain gene expression. Mitochondrial 12S rRNA gene was identified as more highly expressed in

the Nacc and hippocampus of post-mortem human alcoholic brains (compared to control), but not in several other regions tested (Fan et al., 1999). *hNP22* gene expression was increased in the superior frontal cortex of human alcoholics compared to controls (Fan et al., 2001).

In all of the above studies, techniques like mRNA Differential Display, real time PCR, Northern blot analysis and *in situ* hybridization were used to identify differentially expressed genes. The common theme among all of these techniques is that the amount of labor required to identify more than just a few differentially expressed genes is not time- or cost-effective. Granted, many of these studies were "unbiased," meaning candidate genes were not chosen to study based on previous knowledge about their involvement in alcohol dependence. Given the complexity of the disease, it is novel to use models of dependence to search for new genes that could never have been intuited. However, to search effectively for genes involved in a complex, polygenic disease such as alcohol dependence, fast high-throughput methods are ideal. Microarray analysis has become an increasingly popular method to detect differential gene expression in models of alcohol dependence.

Studies using microarrays

A number of microarray studies have explored the genetics of predisposition to alcohol dependence using (alcohol-naïve) inbred strains and selected lines of mice and rats that are divergent in alcohol-related phenotypes. Candidate genes for alcohol sensitivity (Xu et al., 2001) and acute functional tolerance (Tabakoff et al., 2003) were identified in the whole brains of mouse models for each (inbred long-sleep (ILS) and inbred short-sleep (ISS), high acute functional tolerance (HAFT) and low acute functional tolerance (LAFT), respectively) using microarrays. The genetics of alcohol drinking was explored in a meta-analysis of microarray results from three selected lines

and six isogenic strains of mice that differ in voluntary alcohol consumption (Mulligan et al., 2006). Genes were identified that may contribute to the development of alcohol tolerance by comparing the hippocampal transcription profile of inbred alcohol-preferring (iP) and -nonpreferring (iNP) rats (Edenberg et al., 2005). Finally, two studies compared the transcription profile of brain regions involved in the reward pathway (cingulate cortex, nucleus accumbens, amygdala and hippocampus (Arlinde et al., 2004) and frontal cortex (Worst et al., 2005)) to search for alcohol preference genes using the alcohol-preferring AA (Alko, alcohol) and alcohol-avoiding ANA (Alko, non-alcohol) rat strains.

Gene expression changes elicited by alcohol can be explored using a wide range of organisms/models, from the simple to the complex. Alcohol-related genes have been identified using simple systems of mammalian cells and *Saccharomyces cerevisiae* (yeast) coupled with microarray analysis. Human neuroblastoma cells were treated with alcohol for three days to identify alcohol-sensitive genes. 42/6000 genes on the oligonucleotide array were up- or down-regulated by alcohol (Thibault et al., 2000). Cortical neurons isolated from B6 fetuses were treated with alcohol for five days and the isolated RNA was applied to a pathway-specific cDNA microarray containing 638 clones. Most of the differentially expressed genes were down-regulated (56) compared to only a few that were up-regulated (10) (Gutala et al., 2004). Yeast cells were used to identify genes differentially expressed as a result of 30 minutes of alcohol stress. About 6% of the genes on the 6138 clone-containing cDNA array were up- or down-regulated (3% each) (Alexandre et al., 2001).

As the models used move from single-celled to multi-cellular, phenotypes such as alcohol-related behavior can be considered in addition to alcohol-sensitive changes in gene expression. *Caenorhabditis elegans* (nematode) was used as a model system to study the effects of acute alcohol administration on gene expression. The nematode

displays similar behavioral responses to acute alcohol as does a mouse or human, mainly hypersensitivity, uncoordination and sedation. The nematodes were exposed to alcohol for 15 minutes, 30 minutes and 6 hours. In total, 230 genes on the cDNA array were affected by alcohol. Most of the genes were up- or down-regulated at the later time point, as opposed to the earlier time points (Kwon et al., 2004).

Several studies have used mouse and rat strains that diverge in alcohol-related behaviors to study the effects of alcohol administration on gene expression. Daniels and Buck used cDNA microarrays to compare gene expression changes in the hippocampus of B6 and D2 mice after chronic (72 hour vapor chamber with 0 hour sacrifice) or acute (4 g/kg alcohol injection with 7 hour sacrifice) withdrawal (Daniels and Buck, 2002). Of the 7634 clones on the cDNA array, 2% were differentially expressed in D2 and less than 1% in B6 during chronic withdrawal (compared to control). Less than 1% were differentially expressed in D2 during acute withdrawal (compared to control). Upon functional classification of the differentially expressed genes, many were found to be associated with the Ras/mitogen-activated protein kinase (*MAPK*), Janus kinase/signal transducers and activators of transcription (*JAK/STAT*) and Akt/phophatidylinositol 3-kinase (*PI 3-kinase*) signal transduction pathways. Treadwell and Singh also explored the brain expression profiles of B6 and D2 mice after an acute injection of alcohol (6 g/kg) (Treadwell and Singh, 2004). They identified 61 genes (out of 24,000) that were differentially expressed six hours after alcohol exposure in one or both strains (compared to control). The alcohol-related genes have roles in cell signaling, gene regulation and homeostasis/stress response. Unlike the two previous studies, where the administration of alcohol resulted in a physically dependent mouse, the following study identified gene expression changes in the B6 and D2 mouse brain after a stimulating dose of alcohol. Kerns et al. examined the regions composing the brain's reward pathway (nucleus

accumbens (Nacc), prefrontal cortex (PFC) and ventral tegmental area (VTA)) for differences in gene expression four hours after an acute injection of alcohol (2 g/kg) (Kerns et al., 2005). They identified 307 alcohol-sensitive genes (out of 12,000) in the three brain regions of the B6 or D2 mice (compared to control). The Nacc had the most gene changes (~78 in B6; ~90 in D2), followed by the PFC (~23 in B6; ~87 in D2) and VTA (~37 in B6; ~10 in D2)). The regulated genes in each brain region belonged to discrete functional groups: neuropeptide signaling and development in the Nacc, glucocorticoid signaling, neurogenesis and myelination in the PFC and retinoic acid signaling in the VTA. Rimondini et al. subjected Wistar rats to repeated cycles of withdrawal and intoxication for a seven week period, followed by a three week recovery period (Rimondini et al., 2002). After the three week recovery period, the rats consumed higher amounts of alcohol than before, as measured by a two-bottle choice paradigm. Microarray analysis was done on the cingulate frontal cortex and amygdala of the alcohol-exposed rats (versus control) following the three week recovery period (designed to be in parallel with the induction of alcohol preference in the rats given the two-bottle choice paradigm). Gene expression changes were identified (19 in the cortex and 13 in the amygdala) that belonged to the following functional categories: neurotransmission, signal transduction and synaptic plasticity. Finally, Saito et al. examined the gene expression profile of the hippocampus of female Lewis rats fed a 12% alcohol solution as their only source of liquid for 15 months. They identified 164 (out of 5000) differentially expressed genes between the alcohol fed and control rats. Well represented functional categories included metabolism, signal transduction, cell surface antigens, protein synthesis and synaptic/neuronal developmental proteins (Saito et al., 2002).

Microarray analysis has been used to survey alterations in brain gene expression of chronic alcoholics. The regions of the frontal lobe (including the prefrontal and motor

cortices) are susceptible to physical damage following long-term alcohol consumption (Kril et al., 1997; Kril and Halliday, 1999) and are therefore often the focus of microarray studies in human alcoholics (Lewohl et al., 2000; Mayfield et al., 2002; Iwamoto et al., 2004; Liu et al., 2004; Flatscher-Bader et al., 2005; Liu et al., 2006). Several commonalities existed among the results of the microarray analyses comparing post-mortem tissue of chronic alcoholics versus control subjects; including: 1) More genes were down-regulated than up-regulated in alcoholic versus control tissue (specifically, the frontal cortex) (Lewohl et al., 2000; Mayfield et al., 2002; Sokolov et al., 2003; Liu et al., 2004; Flatscher-Bader et al., 2005; Liu et al., 2006), 2) A large number of genes expressed in the alcoholic tissue were myelin-related (Lewohl et al., 2000; Mayfield et al., 2002; Iwamoto et al., 2004; Liu et al., 2004; Flatscher-Bader et al., 2005; Liu et al., 2006) and 3) There were many overlapping functional groups represented in the differentially expressed genes: signal transduction (Lewohl et al., 2000; Mayfield et al., 2002; Sokolov et al., 2003; Iwamoto et al., 2004; Liu et al., 2004; Flatscher-Bader et al., 2005), transcription/translation (Lewohl et al., 2000; Mayfield et al., 2002; Iwamoto et al., 2004; Flatscher-Bader et al., 2005; Liu et al., 2006), metabolism (Lewohl et al., 2000; Mayfield et al., 2002; Liu et al., 2004; Flatscher-Bader et al., 2005; Liu et al., 2006), cell cycle regulation and growth (Lewohl et al., 2000; Mayfield et al., 2002; Liu et al., 2004; Flatscher-Bader et al., 2005), cell adhesion (Lewohl et al., 2000; Mayfield et al., 2002; Sokolov et al., 2003; Iwamoto et al., 2004; Liu et al., 2004; Flatscher-Bader et al., 2005; Liu et al., 2006) and mitochondrial (Sokolov et al., 2003; Flatscher-Bader et al., 2005; Liu et al., 2006). Two recently published studies have noted brain region- and treatment-specific functional categories. Flatscher-Bader et al. compared the gene expression profiles of the Nacc and the PFC of chronic alcoholics and found that only 6% of the alcohol-sensitive differentially expressed genes were common to both regions. In

addition, the predominant functional groups assigned to the altered genes in each brain region were different from each other (Flatscher-Bader et al., 2005). Liu et al. used Principal Components Analysis to separate controls and alcoholics in four different functional categories (Liu et al., 2006).

In conclusion, alcohol affects the expression of a large number of genes simultaneously, often to only a small degree. It is likely that these changes are not all direct effects of alcohol exposure, but also include secondary effects such as the downstream consequences of an altered transcription factor, membrane receptor, signal transduction cascade member or biochemical pathway member. The types of functional categories represented by the alcohol-related genes are similar across animal models and alcohol paradigms indicating that alcohol affects a wide variety of cellular processes but in a consistent manner. Finally, microarray analyses have contributed a great deal of knowledge about how alcohol affects central nervous system gene expression. The next step is to examine the complex network of gene interactions and decipher how and why they may be contributing to the development of alcohol dependence.

Summary

Alcohol dependence is a mental disease caused by a combination of environmental factors and genetic makeup. The disease costs the United States billions of dollars annually and places a tremendous strain on society. Several treatment options exist however none are universally effective—a phenomenon that is likely due to genetic makeup. By understanding more about the genetic component of alcohol dependence, medication and treatment could be tailored to individuals or populations based on their genotype.

Alcohol affects a wide range of receptors and neurochemical systems in the brain. The brain starts to adapt to these molecular and cellular alterations almost immediately. The current hypothesis is that over time, some of these changes become permanent. Stable changes in receptor amount, intracellular signaling, projections between brain regions and ultimately, basic brain function are thought to be caused in part by changes in gene expression. Stable alterations in gene expression are thought to contribute to the manifestations of alcohol dependence, including tolerance, physical dependence (withdrawal) and the ultimate loss of control felt with regard to alcohol consumption.

Both chronic (lifelong alcohol consumption) and acute (binge drinking) alcohol exposure alter brain gene expression and elicit signs of dependence. Mice are a useful model with which to study the effects of chronic and acute alcohol exposure on brain gene expression given that 99% of the genes in the mouse genome have direct counterparts in humans. In addition, almost all aspects of the disease have been modeled and studied successfully in mice using inbred strains, selected lines or genetically altered mice. The B6 and D2 strains are easily obtainable, commonly used mice that differ

greatly in alcohol preference and withdrawal severity, making them ideal models with which to study the genetics of alcohol dependence.

Several studies have used mouse and rat strains that diverge in alcohol-related behaviors to study the effects of alcohol administration on brain gene expression (Daniels and Buck, 2002; Saito et al., 2002; Treadwell and Singh, 2004; Kerns et al., 2005). While all of the aforementioned studies (and many others) have contributed to our knowledge of both predisposition and alcohol-related gene expression, the story is still incomplete. Two key components are missing in all of the above studies—microarrays containing all of the clones representing the entire mouse transcriptome (the studies above used 7634, 24,000, 12,000 and 5000 clones on their microarrays, respectively) and a time course of gene expression changes following alcohol administration. To truly discover the complete transcriptional response of the brain to alcohol, it is better to monitor expression changes over a time course following alcohol administration. Gene expression is a dynamic process, with different genes having different half-lives and potentially regulating the transcription of other genes, therefore it is important to survey the alcohol-induced expression changes at multiple time points following alcohol exposure.

In order to gain a better understanding of the complete transcriptional response of the mouse prefrontal cortex to acute alcohol exposure, the following four Specific Aims were proposed:

Specific Aims

1) To identify genetic and alcohol-related gene expression in C57BL/6J (B6) and DBA/2J (D2) whole brain using mRNA Differential Display (DD)

2) To generate additional targets for cDNA microarrays from genes identified in the DD screen

3) To identify genetic and alcohol-related temporally regulated gene expression in B6 and D2 prefrontal cortex (PFC) using cDNA microarray analysis

Hypotheses to be tested:

- Acute alcohol exposure changes PFC gene expression temporally.
- Acute alcohol exposure causes genotype-independent changes in PFC gene expression.
- Genetic differences in gene expression exist between B6 and D2 PFC.

4) To use computational tools to characterize differentially expressed genes identified in the cDNA microarray analysis

Hypotheses to be tested:

- Co-expressed genes share common promoter region sequences.
- Alcohol-related differentially expressed belong to different functional categories based on genotype and time post alcohol exposure
- Genetic and alcohol-related differentially expressed genes can be categorized by specific neuronal cell type

CHAPTER 2: MATERIALS AND METHODS

Animals and animal husbandry

mRNA DIFFERENTIAL DISPLAY

Male C57BL/6J (B6) and DBA/2J (D2) mice, 60-90 days old, were group-housed under a 12 hour light/dark cycle with food and water available *ad libitum*. All animal procedures were IACUC (Institutional Animal Care and Use Committee) approved according to NIH/AAALAC (National Institutes of Health/Association for Assessment and Accreditation of Laboratory Animal Care) guidelines.

cDNA MICROARRAYS

Male C57BL/6J (B6) and DBA/2J (D2) mice, 70-100 days old, were group-housed under a 12 hour light/dark cycle with food and water available *ad libitum*. All animal procedures were IACUC approved according to NIH/AAALAC guidelines.

Experimental design and tissue collection

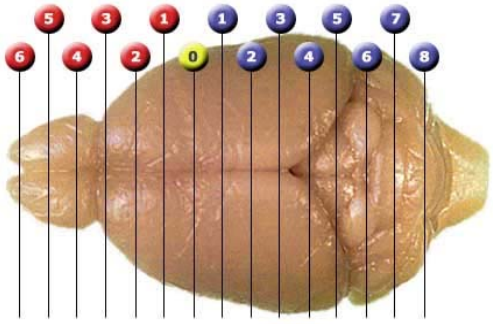
mRNA DIFFERENTIAL DISPLAY

B6 and D2 mice (n = 2 per strain) were injected intraperitoneally (i.p.) with an acute dose of 4 g/kg ethanol (200 proof, diluted to 20% v/v in 0.9% sodium chloride) or an equivalent volume of saline (0.9% sodium chloride). Animals were sacrificed by cervical dislocation and whole brains removed at 2 hours and 6 hours post-injection, immediately frozen in liquid nitrogen and stored at -70°C.

cDNA MICROARRAYS

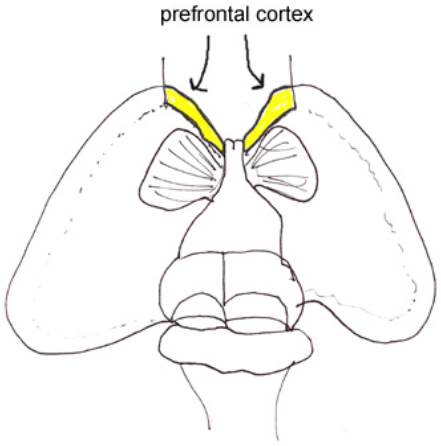
B6 and D2 mice were habituated to injection for at least three days with a 0.5 ml intraperitoneal (i.p.) injection of saline (0.9% sodium chloride). On the day of the experiment, the mice were injected in random order with 4 g/kg, i.p. ethanol (200 proof, diluted to 20% v/v in 0.9% sodium chloride) or an equivalent volume of saline. Alcohol-injected mice were sacrificed by cervical dislocation at 2, 7, 12 and 24 hours post-injection (n = 7 per group) while saline control mice were sacrificed at 2 hours post-injection (n = 7 per group). Seven brain regions, including the prefrontal cortex, were dissected from whole brain and immediately frozen in liquid nitrogen within two minutes of sacrifice and three minutes of removal from home cage. Prefrontal cortex was isolated by making an incision along the *fissura longitudinalis cerebri*, peeling the cortex flat and dissecting the anterior portion of the brain back 1.7 mm posterior, approximately 1.5 mm rostral to bregma, including portions of the following anatomical regions: agranular, cingular, frontal association, motor, orbital, pre- and infralimbic cortex (Franklin and Paxinos, 1997). See Figures 2.1 and 2.2 for more detail. Samples were stored at -70°C.

Figure 2.1: Dorsal view of mouse brain. The cortex was dissected along the *fissura longitudinalis cerebri* from #3 (red) to #3 (blue) such that the left and right cortical hemispheres could be splayed flatly to either side of the brain. An anterior section of the brain, from #3.2 (red) to #1.5 (red), was dissected out as the prefrontal cortex. The zero coordinate is the bregma point, with the red numbers running anterior/rostral to bregma and the blue numbers running posterior/caudal to bregma. The lines/numbers are spaced 1 mm apart.



(http://www.mbl.org)

Figure 2.2: Schematic showing prefrontal cortex dissection. A dorsal view of the mouse brain, with the left and right cortical hemispheres splayed flatly to either side of the brain. The yellow color indicates the regions that were dissected out as the prefrontal cortex.



RNA isolation and quantification

mRNA DIFFERENTIAL DISPLAY

Total RNA was isolated from whole brain using RNA STAT-60 (Tel-Test, Inc., Friendswood, TX) according to the manufacturer's protocol. Individual brains were homogenized in 6 ml of RNA STAT-60 using a Polytron homogenizer (Kinematica, Lucerne, Switzerland). RNA was resuspended in DEPC-treated water and quantified with the DU640 UV/Vis spectrophotometer (Beckman Coulter, Inc., Fullerton, CA). Samples were DNase treated with 1 U DNase I/μg RNA (Gibco-BRL) for 15 minutes at 37°C, re-extracted with acid-phenol/chloroform, resuspended in DEPC-treated water and quantified using the spectrophotometer. Samples were stored at -70°C.

cDNA MICROARRAYS

Total RNA was isolated from prefrontal cortex using RNA STAT-60 (Tel-Test, Inc., Friendswood, TX) according to the manufacturer's protocol. Individual prefrontal cortices were homogenized in 1 ml of RNA STAT-60 using a Polytron homogenizer (Kinematica, Lucerne, Switzerland). RNA was resuspended in 15 μl of nuclease-free water (Ambion, Austin, TX) and quantified with a NanoDrop ND-1000 UV/Vis spectrophotometer (NanoDrop Technologies, Wilmington, DE). The average 260/280 ratio was 1.86 and the average 260/230 ratio was 2.06; indicating that the small tissue mass and low resuspension volume did not negatively affect the RNA quality. In addition, the RNA was checked for degradation using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). See Figures 2.3 and 2.4 for more detail. Samples were stored at -70°C.

Figure 2.3: Quality determination of total RNA isolated from prefrontal cortices. A pseudo-gel created from the Agilent 2100 Bioanalyzer of ten prefrontal cortex total RNA samples. The two bands in each sample represent 18S and 28S rRNA. The ladder is the RNA 6000 ladder from Ambion (Austin, TX). The bands are sharp and distinct with no signs of RNA degradation.

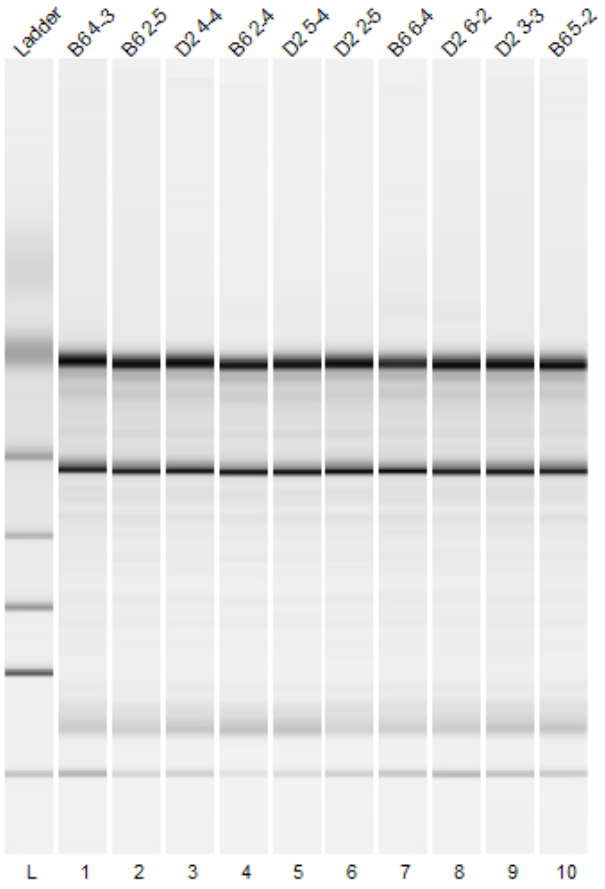
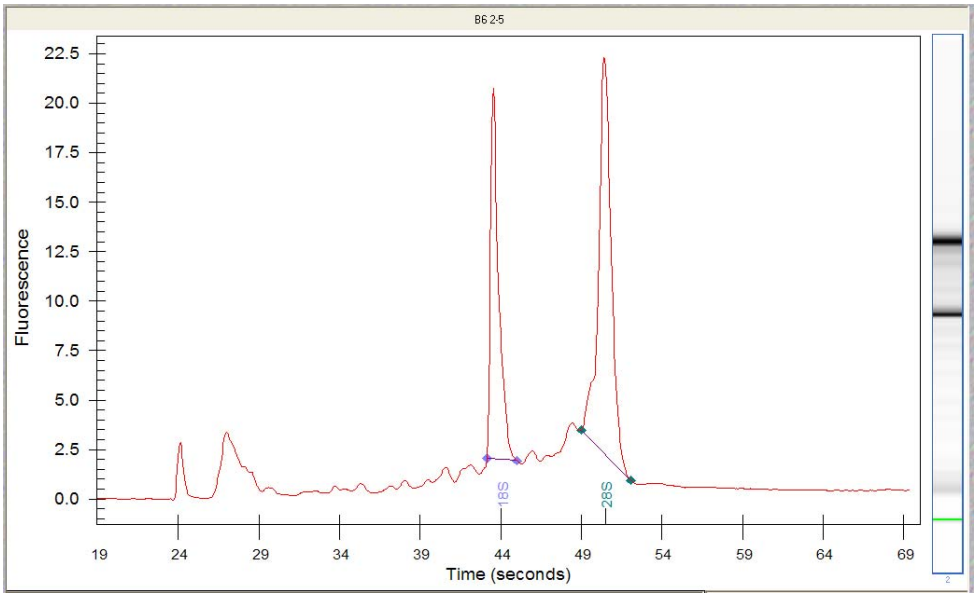
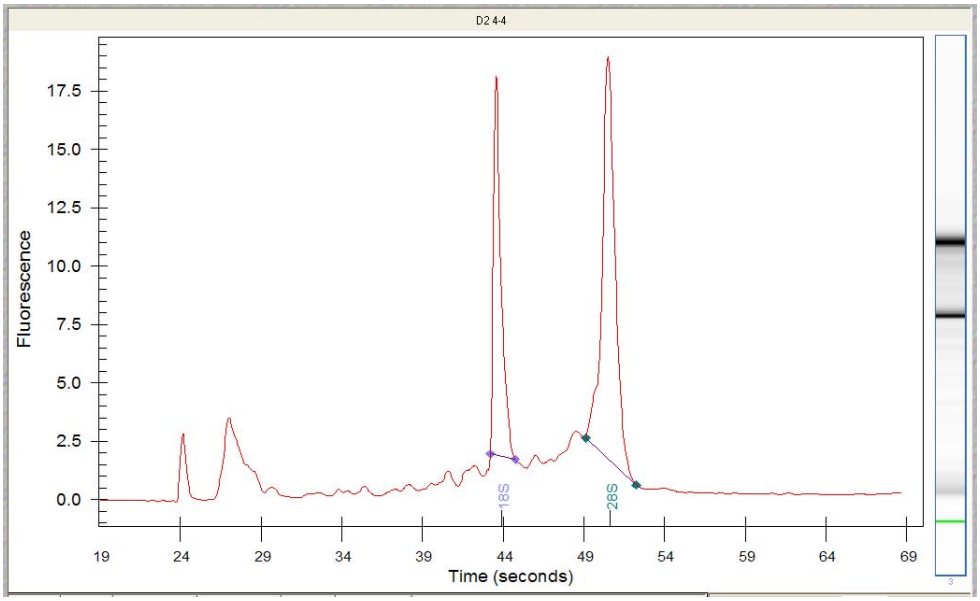


Figure 2.4: Electropherograms of two prefrontal cortex total RNA samples. The top (sample B6 2-5) and bottom (sample D2 4-4) electropherograms represent two single wells from the pseudo-gel in Figure 2.3. The two tall peaks in the electropherograms represent the 18S and 28S rRNA bands, respectively. A relatively flat, consistent baseline reading indicates no degradation is present.





mRNA Differential Display (DD)

REVERSE TRANSCRIPTION

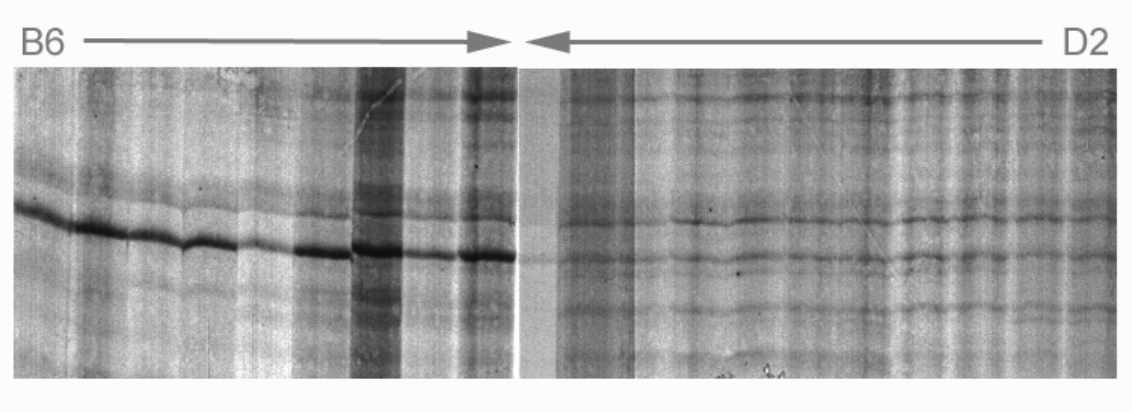
5 μg of total RNA was reverse transcribed using Superscript™ II RNase H- reverse transcriptase (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol, with the following modifications: 1) A one-base anchored oligo dT primer, H-T_{11}M (GenHunter, Brookline, MA; H = *Hin*dIII restriction sequence, M = A, G or C) replaced the oligo $(dT)_{12-18}$ primer, thereby creating three different cDNAs from every sample of RNA, (2) An initial denaturation incubation at 75°C for 10 minutes followed by icing was performed, (3) The final dNTP concentration was increased to 0.5 mM from 25 μM as this gave substantially better DD banding results, (4) The 5X 1st strand buffer, DTT (dithiothreitol) and dNTPs were added in a single pipetting from a master mix in order to minimize pipetting errors and (5) A 2 minute incubation at 42°C allowed good annealing of primer, as did pre-warming of tubes prior to the addition of enzyme. Incubation at 42°C proceeds for an additional 50 minutes and is followed by thermal termination of enzyme activity at 75°C for 15 minutes. Synthesized cDNA was resuspended in 260 μl DEPC-treated water and stored at -20°C.

POLYMERASE CHAIN REACTION (PCR) AMPLIFICATION

Each of the three types of cDNAs created in the reverse transcription (using A, G, or C anchored primers) was PCR amplified using each of the 80 arbitrary forward primers (H-AP; AP = 13mer of arbitrary sequence) and one of the three reverse primers, H-T_{11}M (M = A, G or C, depending on the cDNA being amplified) in the RNAimage™ kit (GenHunter) for a total of 240 PCR reactions per original sample of RNA. Reactions were performed as described in Liang and Pardee (Liang and Pardee, 1998) with the

following exceptions: 1) α-[^{33}P-dATP] (>2000Ci/mmol) replaced α-[^{35}S-dATP] (Callard et al., 1994), 2) Denaturation temperature was increased, 3) Annealing temperatures were adjusted based on the sequence of the forward primer to enhance specificity of the reaction and 4) Extension time was increased. Together these changes seem to have lead to a reduction in false positives and larger sized fragments displayed. Samples are run using Robbins Scientific (Sunnyvale, CA) cycle plates having 192 split wells which were verified to work well (Chen et al., 1994). The PCR conditions for thermocycler runs (GeneAmp PCR System 9600, Perkin Elmer Cetus) were optimized to include a 2 minute "hot start" at 96°C to ensure maximal denaturation. PCR cycles were: 40 cycles of denaturation at 96°C for 30 seconds, annealing at 35-48°C for 1 minute (based on optimal binding for each primer, determined using PC-based program Oligo, Oxford Molecular), and extension at 72°C for 1 minute. A final 7 minute extension at 72°C followed by an indefinite 4°C incubation until removal of the samples completes the PCR. PCR products were heat denatured for 2 minutes at 96°C, placed on ice and loaded on a 6% polyacrylamide/urea sequencing gel (SequaGel-6, National Diagnostics, Atlanta, GA). Samples were not loaded on the gel until the gel reached 50°C. The gels were run at 55-60°C with the voltage prevented from rising above 1700. Wattage was usually 85, considerably higher than recommended. The gel was transferred to 3M filter paper, dried and exposed to X-ray film (Kodak, Rochester, NY) overnight. Film was developed using a Kodak X-OMAT 2000A processor machine. Autoradiograms of the gels were examined for genes that were differentially expressed between control mice, as a result of alcohol exposure or both (based on intensity and/or presence of bands). See Figure 2.5 for an example of what a differential display gel looks like.

Figure 2.5: Example of a differentially expressed gene on a differential display gel. The banding pattern difference between the B6 and D2 samples indicates this gene is expressed at a higher level in the B6 mouse than the D2 mouse, and is therefore genetically divergent.



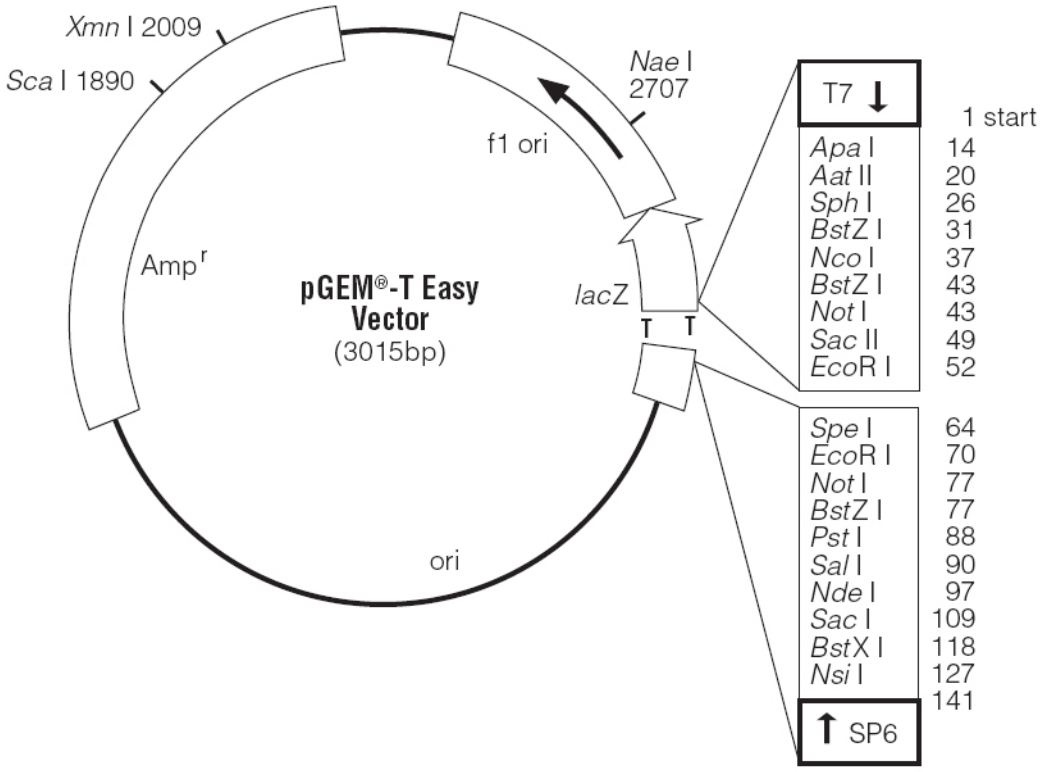
CLONING

Bands of interest were outlined on the backside of the autoradiogram with a fine pointed woman's eyeliner (Max Factor). The autoradiogram was then easily aligned to the blotted gel using several unique spots of sample containing radioactivity and dye (that were spotted prior to film exposure). The autoradiogram was pressed down on the dried gel to transfer the eyeliner outline to the gel to demarcate where the band was on the dried gel. The band was excised from the dried gel using a clean razor blade. The band was soaked in 30-50 μl of filtered water for 10 minutes, boiled for 15 minutes and spun down. The DNA in the band's "supernatant" was ethanol precipitated and resuspended in 10 μl of filtered water. The DNA was PCR amplified using the same forward and reverse primers, temperatures and cycle lengths as in the differential display (DD) PCR reaction. The PCR product was run on a 1% agarose gel, the band excised with a clean razor blade and the DNA extracted using the freeze-squeeze method (Islam et al., 2002). In this method, the agarose is subjected to multiple freeze-thaw cycles, placed inside a piece of parafilm and squeezed. The supernatant was ethanol precipitated and resuspended in 5 μl of filtered water.

The PCR product was subcloned into a T/A cloning vector (pGEM®-T Easy Vector, Promega, Madison, WI) (see Figure 2.6) according to the manufacturer's protocol, except that the volumes of all of the reagents were cut in half. 1 μl of ligation reaction was added to 50 μl of electrocompetent cells (XL1-Blue, Stratagene, La Jolla, CA), which were transformed using a GenePulser II System (Biorad, Hercules, CA) according to the manufacturer's protocol. Electroporated cells were gently shaken at 37°C for 30 minutes to 1 hour in 1 ml of SOC medium, spun down, resuspended in a small volume of SOC and spread on agar plates (containing X-gal/IPTG/ampicillin)

which were incubated overnight (inverted) at 37°C. Positively transformed colonies were chosen based on blue/white color screening and the Wizard® Plus SV Minipreps DNA Purification System (Promega) was used to isolate the DNA. A sample of the DNA was cut with 50 U/μl of *Eco*RI to ensure presence of the insert.

Figure 2.6: pGEM-T Easy Vector. The PCR products generated by the differential display screen were subcloned into the multiple cloning site of the vector.



CLONE IDENTIFICATION

Clones were sequenced with an automated DNA sequencer (ABI 373A, using Taq dye terminator chemistry; VAMC Molecular Biology Core Facility, Portland, OR), using the T7 primer. Sequences were entered into Omiga 2.0 (Genetics Computer Group, Madison, WI) and aligned with the T7 and SP6 primer sequences so that the primer sequences could be deleted from the clone sequence. Truncated clone sequences were entered into the Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) and based on their E score, identified as known genes, expressed sequence tags (ESTs) or unknown genes.

cDNA microarrays

MICROARRAY FABRICATION

Clone preparation

Clones printed on the University of Texas at Austin custom cDNA microarrays included the NIH BMAP UniGene library (~11,000 clones), Sequence Verified IMAGE clones (~5000) (ResGen/Invitrogen, (Carlsbad, CA), The National Institute of Aging clone set (~22,000 clones) and the clones (~300) generated from the mRNA differential display screen. Based on assignment to Unigene clusters, approximately 25,000 distinct genes are represented in the first three clone sets. The differential display clones represent an additional 150 unique genes or ESTs.

The three purchased clone sets were shipped as frozen glycerol stocks of *E. coli*. A small portion of each frozen clone was distributed into individual wells of 96 well plates so each could be amplified by PCR. PCR amplification was performed in a

Perkin-Elmer 9600 thermocycler using commercially available primers (primer pairs were chosen based on the vectors used in each clone set) in 50 µl reactions containing PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 10 pmol of each primer, 5 U Taq polymerase, and 0.5 µl of the bacterial culture. PCR cycles included annealing at 55°C for 20 seconds and extension at 72°C for 90 seconds for a total of 30 cycles. 5 µl of each reaction was then analyzed on a 1% agarose gel to verify the success of the PCR.

Primers were designed (TAEasy5' and TAEasy3') to PCR amplify the clones generated by the mRNA differential display. DD clones were amplified in 100 µl reactions using standard concentrations of $MgCl_2$, dNTPs, primers, MasterAmp™ Tfl DNA polymerase and MasterAmp™ Tfl 20X PCR buffer (EPICENTRE Biotechnologies, Madison, WI). The PCR conditions were as follows: (1) A 2 minute "hot start" at 94°C, (2) 40 cycles of 94°C for 30 seconds, 62°C for 30 seconds, 72°C for 60 seconds, (3) A 7 minute extension at 72°C and (4) An indefinite hold at 4°C. The amplification was performed on a PTC-100 PCR machine (MJ Research, Inc. (Biorad), Waltham, MA). A small portion of the PCR product was analyzed on a 1% agarose gel to ensure the presence of a single band (Figure 3.3). For each unique gene or EST found to be differentially expressed in the DD screen, two clones representing the same DNA sequence were chosen to add as targets to the microarray. The PCR products of these clones were transferred to U bottom plates and precipitated. All clones (purchased and DD) were precipitated in the same manner, using isopropanol and sodium acetate. The precipitated clones were resuspended in water, transferred to 384 well plates (Genetix USA Inc., Boston, MA), spun down in an Eppendorf 5810R centrifuge (Westbury, NY), dried down in a Savant Speedvac® Plus SC210A (ThermoQuest (Thermo Electron Corporation), Waltham, MA) and stored at room temperature. Several copies of each 384

well plate were created for longevity of the clones. Prior to a print, the DNA in one set of 384 well plates was resuspended in 6 µl of 3X saline-sodium citrate (SSC) buffer.

Printing the microarrays

Glass slides (Gold Seal, Erie Scientific Co., Portsmouth, NH) were cleaned for two hours in a solution of 2.5 N NaOH in 60% ethanol, followed by a water rinse. Slides were soaked for 30 minutes in a solution of 0.017% w/v poly-L-lysine (Sigma-Aldrich Corp., St. Louis, MO) in 0.1X phosphate buffered saline (PBS), rinsed with water, and spun dry at 600 rpm in a GH3.8 rotor in an Allegra 6R centrifuge (Beckman Coulter Inc., Fullerton, CA) for 10 minutes. Slides were stored in a cool, dry place for a minimum of 10 days before printing.

All clones were stored in 384 well plates as dried pellets. Prior to each print, clones were resuspended in 6 µl of water or 3X SSC. Resuspension was performed using a Biomek FX Laboratory Automation Workstation (Beckman Coulter, Inc., Fullerton, CA) one day prior to printing. The 384 well plates were covered with tape to form an air-tight seal and stored at 4^0C overnight. The next day the plates were loaded consecutively on the custom built arraying robot. The arraying robot used to print the arrays was built by and is housed in the Iyer lab on the UT campus. It is similar to that of (DeRisi et al., 1997). The proper alignment, calibration and movement of the robot were controlled by the ArrayMaker software (http://derisilab.ucsf.edu/arraymaker.shtml). The 48 stainless steel printing tips (MicroQuill® 3000 Array Pin, Majer Precision Engineering, Tempe, AZ) were individually checked for sharpness under a magnifying glass prior to printing. 260 poly-L-lysine coated slides were taped down on the arrayer platform, the first plate of cDNA clones was placed on the arrayer and the print started. The print head containing the steel printing tips descended eight times into each 384 well plate, captured picoliter amounts of cDNA and printed the cDNA on each slide. Plates were changed

every 20 minutes for approximately 36 hours until all 39,000 cDNA clones were printed on each of the 260 glass slides. Slides were removed from the arrayer platform and stored in slide boxes in a cool dry area.

MICROARRAY HYBRIDIZATION

Post processing the microarrays

Slides were hydrated by holding them one at a time, array side up, above a 90°C water bath for 10 seconds. The steam from the water bath hydrates each spot of dried cDNA to allow it to expand and form a uniform circle on the slide. Slides were dried by gently waving them in the air for 10 seconds. Printed cDNA was crosslinked to the slide in a UV Stratalinker 1800 (Stratagene, La Jolla, CA) for 1.5 minutes. The UV light creates a covalent bond between the positively charged NH_2 group on the poly-L-lysine and the negatively charged phosphate backbone of the cDNA. The slides were transferred to a slide rack, immersed in isopropanol and gently shaken for 15 minutes. Afterwards, the slide rack was transferred to a boiling water bath for 10 minutes to denature the cDNA. Excess water was removed from the slides and slide rack before the rack was placed in a GH3.8 rotor in the Allegra 6R centrifuge and spun dry at 800 rpm for 10 minutes. Slides were hybridized within two days of post-processing.

cDNA labeling and hybridization

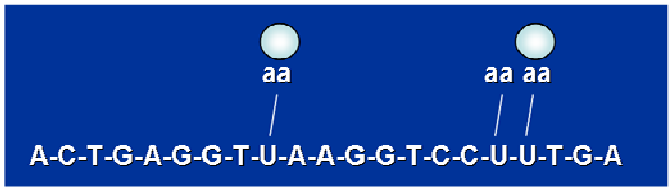
cDNA was synthesized and dye-labeled from total RNA using the 3DNA Array 900 kit (Genisphere, Hatfield, PA). The 3DNA technology uses an end-labeling technique to attach fluorescent dyes to the cDNA. The benefits of this technology over more traditional labeling methods are outlined in Figure 2.7.

The 3DNA Array 900 kit protocol is described here briefly. 500 ng of total RNA from the alcohol-treated or saline control mouse or 1 μg of total RNA from the B6

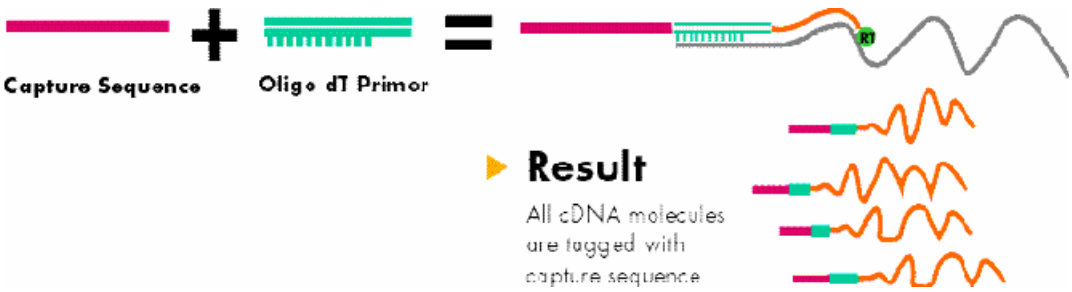
reference pool (whole brain RNA isolated from 100 male B6 mice) was added to 1 pmole of Cyanine-3- (Cy-3-) or Cyanine-5- (Cy-5-) -specific RT primer (respectively) and nuclease free water to a final volume of 6 μl. The Cyanine dyes are bright, intense dyes that have sharp, distinct absorption and emission spectra, allowing good spectral separation when used in co-labeling experiments (Amersham Pharmacia Biotech bulletin). The RNA-primer mix was heated to 80°C for 5 minutes, then immediately transferred to ice. RNA was reverse transcribed for 2 hours at 42°C in a reaction containing 5X SuperScript™ II First Strand Buffer and SuperScript™ II RNase H- reverse transcriptase (Invitrogen, Carlsbad, CA), 0.1 M DTT, Superase-in RNAse Inhibitor, and dNTP mix. The reverse transcription was stopped with 1 μl of 1 M NaOH/100 mM EDTA, incubated at 65°C for 10 minutes, and neutralized with 1.2 μl of 2 M Tris-HCl, pH 7.5.

Figure 2.7: A comparison of methods for labeling cDNA with fluorescent dyes: the amino-allyl indirect labeling method versus the Genisphere 3DNA end-labeling technology. The Genisphere technology is advantageous over the more traditional indirect labeling method for several reasons (outlined below).

1) *Amino-allyl indirect labeling method*: In this method, amino-allyl dNTPs (AA-dUTP) are incorporated during cDNA synthesis, followed by the chemical coupling of amine reactive dyes (Cy-3 or Cy-5) (see below). As a result, the fluorescent signal is dependent on the base composition and length of the cDNA. In addition, each cDNA molecule may contain only 2-50 fluors, requiring that a large amount (15 µg) of RNA be labeled in order to detect a signal. Finally, dye spacing may negatively affect hybridization efficiency.

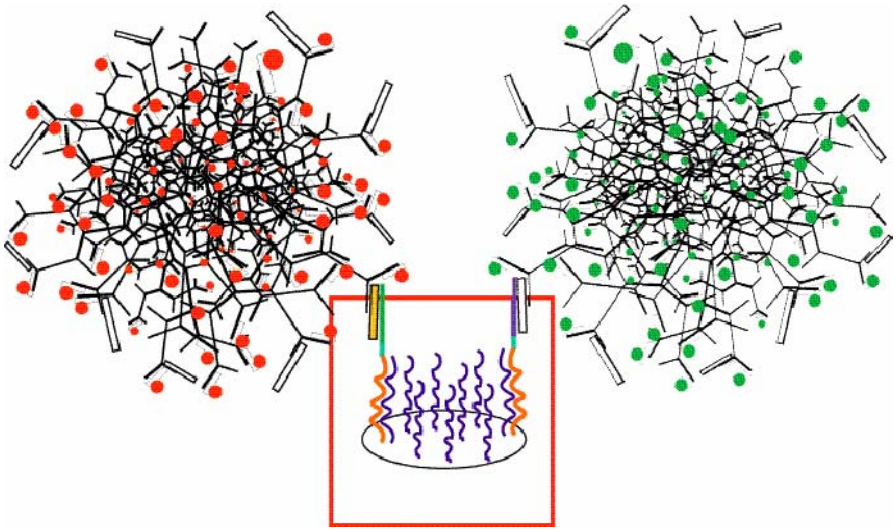


2) *Genisphere 3DNA end-labeling technology*: In this method, each RNA sample is reverse transcribed using an oligo dT primer with a unique capture sequence attached to its 5' end.



(http://www.genisphere.com)

The capture sequence is complementary to one of two fluorescently labeled dendrimers. The dendrimers are 3-dimensional structures composed of multiple layers of DNA monomers. Each monomer is composed of two DNA strands that are complementary to each other in the center portion of the strands. Each dendrimer has 850 (Array 900 kit) fluorescent dyes (Cy-3 or Cy-5) attached to it. Such a large amount of fluorescent dye associated with each molecule of cDNA allows a small amount of RNA to be labeled (500 ng).



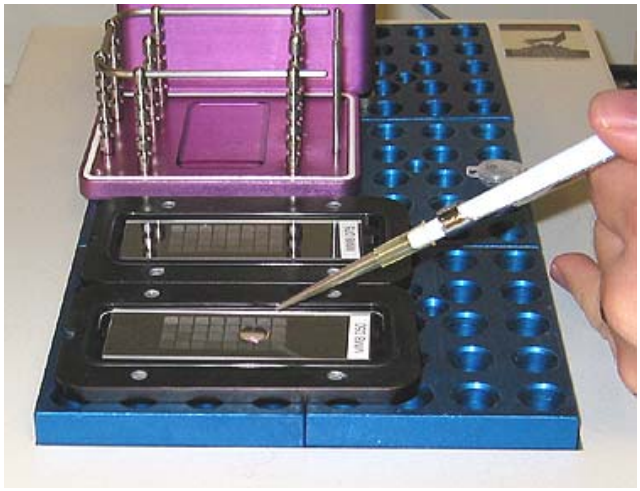
(http://www.genisphere.com)

The cDNA containing the capture sequence is hybridized first to the microarray slide, followed by the dendrimers containing the fluorescently labeled dyes. Therefore, each cDNA is represented by only one dendrimer, producing a base-composition- and length-independent signal.

cDNA from an alcohol-treated or saline control animal was always hybridized against cDNA from a reference. In the Genisphere protocol, the cDNA is hybridized to the array slide first, followed by the fluorescent dyes. This two-step hybridization process is described here. Each experimental cDNA (alcohol treated or saline control) was combined with B6 reference cDNA and added to 2X Hybridization Buffer, LNA dT Blocker, 1 µl of 100 ng/µl COT-1 DNA and nuclease free water, to a total volume of 65 µl. This mix was incubated for 10 minutes at 80°C, then at 65°C until it was applied to the microarray slide. Each microarray slide was pre-warmed to 65°C before the cDNA mix was added to it. The cDNA mix was pipetted onto the microarray and a microscope coverslip was gently place on top of the mix. The slide was then inserted into a humidified hybridization chamber (Genetix Limited, United Kingdom), which remained on the 65°C heat block while the other slides were prepared (Figure 2.8). As many as 10 slides can fit into each Genetix chamber at one time and no more than two chambers were used simultaneously in an experiment. The Genetix chamber(s) were submerged in a 63°C water bath overnight. The next day, each slide was removed one at a time and placed in a slide rack submerged in a 65°C, 2X SSC, 0.2% SDS wash solution. Once all the slides were immersed in this solution, they were shaken for 15 minutes at 65°C. Two more 15 minute room temperature washes followed, in 2X SSC and 0.2X SSC, respectively. The slide rack was removed from the final wash solution and spun in an Allegra 6R centrifuge for 10 minutes at 800 rpm to dry. All of the following steps were performed in the dark. A 3DNA hybridization mix was made that contained a Cy-3 and a Cy-5 3DNA Array 900 Capture Reagent, 2X Hybridization Buffer, 1 µl of 100 ng/µl COT-1 DNA and nuclease free water, to a total volume of 65 µl. This mix was incubated for 10 minutes at 80°C, then at 65°C until it was applied to the microarray slide. The

same technique was used to apply this mix to the slides as was used in the cDNA hybridization. The slides were incubated for 4-5 hours in a 63°C water bath. The slides were removed from the Genetix chamber(s), washed, and spun dry using the same techniques listed above. Once all the slides were dry, they were stored under argon gas during scanning to prevent oxidation of the Cy-5 dye by ozone in the air.

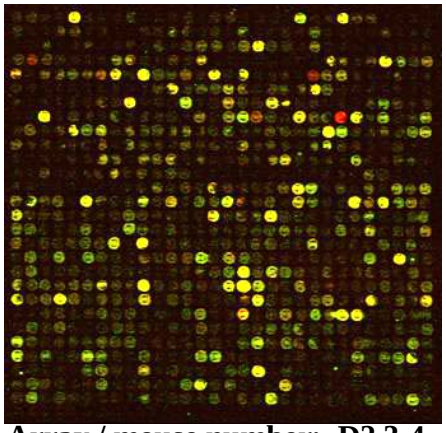
Figure 2.8: cDNA hybridization step of microarray process. Each microarray slide was pre-warmed to 65°C before the cDNA mix was added to it. The cDNA mix was pipetted onto the microarray and a microscope coverslip was gently place on top of the mix. The slide was then inserted into a humidified hybridization chamber (Genetix Limited, United Kingdom), which remained on the 65°C heat block while the other slides were prepared. As many as 10 slides can fit into each Genetix chamber at one time and no more than two chambers were used simultaneously in an experiment. Performing 20 array hybridizations simultaneously decreases technical variability among the arrays.



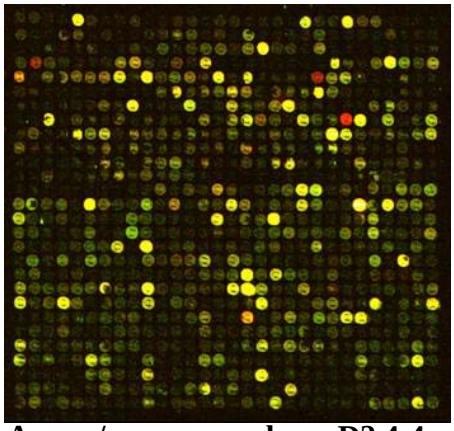
The combination of the Genisphere 3DNA technology, high-throughput hybridization techniques and minor changes to many other aspects of the microarray protocol resulted in high quality microarrays—with little background and strong signal intensity (Figure 2.9).

Figure 2.9: Reproducibility of microarrays using Genisphere 3DNA Array 900 kit.

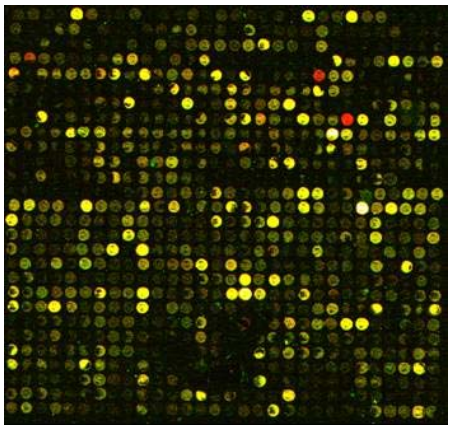
A single block (#31) is shown from four different microarray hybridizations representing 4 n mice. cDNA synthesized from 500 ng of total RNA (labeled with Cy-3) from the prefrontal cortex of a D2 mouse sacrificed 12 hour post alcohol injection was hybridized against 1 μg of B6 reference RNA (labeled with Cy-5). The same spot color and intensity is seen in all four blocks.



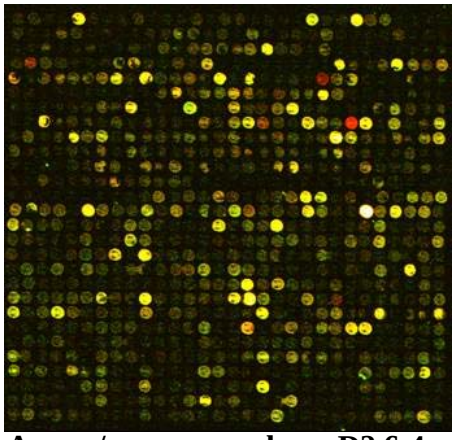
Array / mouse number: D2 3-4



Array / mouse number: D2 4-4



Array / mouse number: D2 5-4



Array / mouse number: D2 6-4

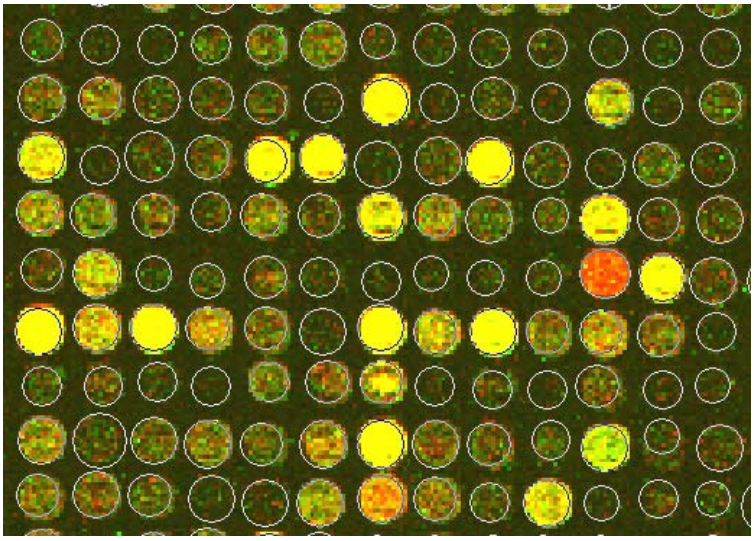
Array scanning and gridding

Slides were scanned using a GenePix 4000A or B scanner (Axon Instruments, Inc., Union City, CA). The scanner has two lasers, 532 nm and 635 nm, that excite Cy-3 (green) and Cy-5 (red) respectively. The emission filters are 575DF35 (green; ~557-592 nm) and 670DF40 (red; ~650-690 nm). Two photomultiplier tubes (PMTs) inside the scanner detect the fluorescent light emitted from the dyes. The sensitivity of the PMTs can be adjusted by increasing or decreasing the PMT gain (voltage). The PMT gain for each laser was adjusted during scanning to achieve an intensity ratio of 1.00. The intensity ratio is the total number of intensity quanta (counts) on the array from the red channel over the total number of intensity quanta from the green channel. An intensity ratio of 1.00 ensures that the intensities from both lasers are balanced, that the red and green traces in the intensity histogram overlap and that the full range of intensities (1-65,535) are utilized.

The array spots were labeled and quantified using SpotReader (Niles Scientific, Portola Valley, CA) and/or GenePix Pro 5.1 (Axon Instruments, Inc., Union City, CA). To label and quantify the spots, a grid of open circles was placed on top of the scanned array image (Figure 2.10). The grid contained clone identification information for each spot. The size and shape of the open circles was manually adjusted to encompass each spot on the array. A spot was flagged (excluded from further analysis) if it ran together with another spot, was incomplete, was non-existent due to poly-L-lysine peeling or had false color (where the signal was not from a true hybridization). As few spots as possible were flagged.

After gridding is complete, a GenePix Results (GPR) file is created that contains clone identification and red/green intensity information for every spot on the array.

Figure 2.10: Example of a grid placed over a scanned microarray slide to label and quantify the spots. A snapshot of spots from slide #85, sample B6 8-3 (seven hours post alcohol exposure). The size of the white circles surrounding the spots is manually adjusted to encompass each spot. All of the spots in this picture have similar color distribution within the spot and are not compromised in any way (necessitating a "flag").



ARRAY ANALYSIS

Normalization, standardization and filtering

One assumption common to all microarray experiments is that the expression levels of most of the genes in the two samples hybridized against each other are equal, thereby producing an array with mostly yellow spots (it is important to note that this assumption only holds true for microarrays that contain a large number of random cDNAs—not arrays that contain a custom set of genes known to be differentially expressed in the hybridized samples). Based on this assumption, several steps are taken to balance the signal contribution and intensity of the red and green channels on every array so that all arrays in an experiment can be directly compared to each other. The first step is during scanning when the gain of the PMTs is adjusted to produce an array with equal signal contribution from the Cy-3 and Cy-5 dyes. In reality, it is difficult to produce an intensity ratio of 1.00 (representative of equal signal contribution), with ratios usually ranging from 0.85 to 1.15. However this minor imbalance can be corrected by the second step, normalization.

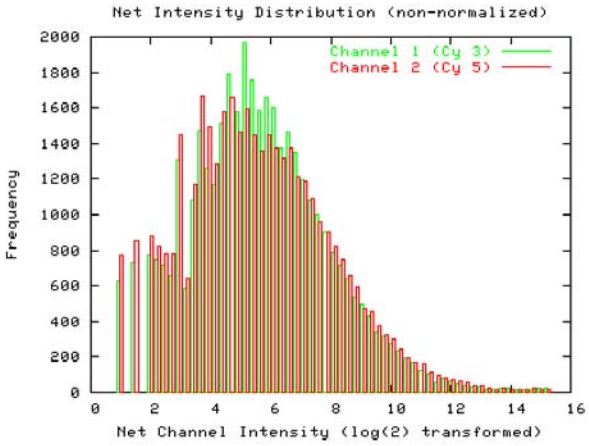
Normalization is a mathematical process used to remove systematic variation from array experiments; meaning, variation that arises from the technology rather than from biological differences between the RNA samples (Smyth and Speed, 2003). Examples of such systematic variation are differences in labeling efficiency between the two fluorescent dyes, physical properties of the dyes (heat, light and ozone sensitivity, relative half-life) and scanner settings (Yang et al., 2002; Fare et al., 2003). A common method of normalization is to apply a normalization factor to the red or green (or both) intensity of every gene on the array, such that the average ratio of every gene on the array

becomes 1 (Quackenbush, 2002). Again, this method relies on the assumption that most genes do not respond to experimental conditions. A different normalization factor is applied to each array but the end result is the same, allowing all arrays in an experiment to be directly comparable. However this global, linear method of normalization does not take into account intensity- or spatially-dependent dye biases, two commonly occurring sources of systematic variation on arrays (Yang et al., 2002). For example, print tips can produce spots of varying sizes between blocks on the array (a single print tip produces all the spots in a single block), the amount of cDNA printed on the slides can vary depending on the slide's position on the arrayer and the hybridization of the dyes can be non-uniform across the array. All of these factors can negatively affect the calculation and application of the global normalization factor to all the genes on the array. A within-print tip, intensity-dependent method of normalization was developed to account for these issues. First, the intensities and intensity ratios of every gene in a single block on the array are plotted against each other to reveal intensity-specific artifacts in the ratios. Second, these artifacts are corrected by the application of a locally weighted linear regression (lowess), resulting in an intensity-specific normalization of all the genes in a single block on the array (Quackenbush, 2002; Smyth and Speed, 2003).

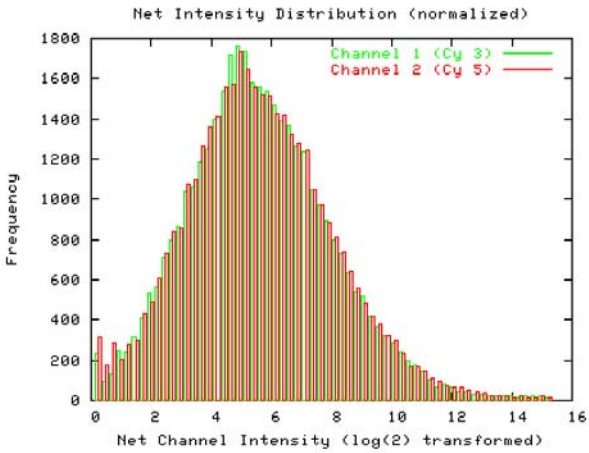
Microarrays were normalized within the R environment (Ihaka and Gentleman, 1996) using the Statistics for Microarray Analysis (SMA) add-on package (Yang et al., 2002). Specifically, the within-print tip non-linear lowess normalization function in SMA was used for normalization (Figure 2.11). A program developed in our lab that automates the normalization procedure was used to apply the normalized values to the GPR files.

Figure 2.11: Histogram of two microarrays before and after normalization. The net (background-subtracted) intensity distribution of the red and green channels on each array was plotted on a histogram prior to and after array normalization to ensure the normalization process produced normally distributed net intensities. On the x-axis is the range of intensity values possible for any given spot on the array (1-65,535 ($\log_2$ transformed)). On the y-axis is the frequency of each intensity value.

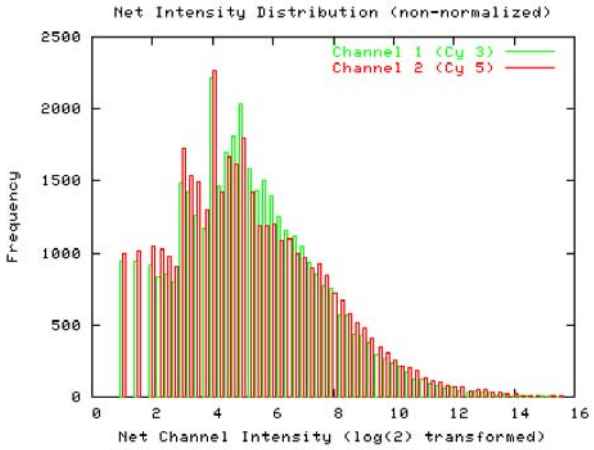
1) Mouse/array number: B6 6-5 (pre-normalization)



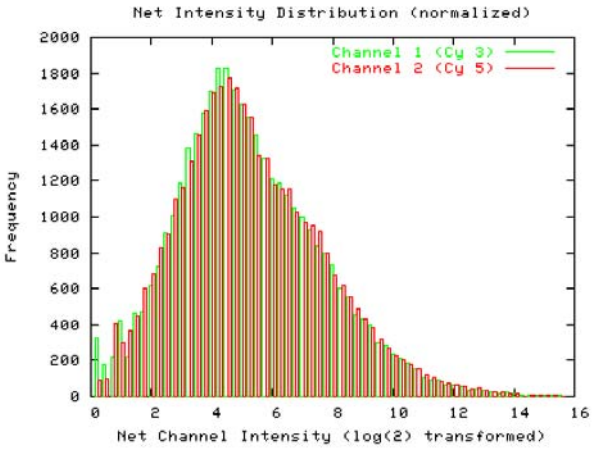
2) Mouse/array number: B6 6-5 (post-normalization)



3) Mouse/array number: D2 3-1 (pre-normalization)



4) Mouse/array number: D2 3-1 (post-normalization)



The normalized GPR files were uploaded and stored in the Longhorn Array Database (LAD) (Killion et al., 2003). Information such as gene identifiers (name, clone ID, accession number, chromosome number), intensity values and $\log_2$ ratios can be retrieved from the LAD for a single array or a pre-defined group of arrays. Microarray data was retrieved from the LAD as net (background-subtracted) median intensity values. The Channel 1 (Ch1 (alcohol-treated or control Cy-3 labeled sample)) and Channel 2 (Ch2 (reference Cy-5 labeled sample)) intensities of every gene (39,335) on all 70 arrays were downloaded as two separate text files. Obtaining Ch1 and Ch2 intensity values for each gene (as opposed to the more traditional $\log_2$ ratio values) has several benefits, including: 1) Genes that are turned on or off as a result of alcohol exposure are maintained and 2) Intensity information is maintained. Genes that go from "on" to "off" or vice versa can have very low intensity values, which when converted to $\log_2$ ratios can produce artificially high, low or undefined values depending on whether the intensity value was in the numerator or denominator. Artificially high or low values may skew the data and undefined values would be removed from further analysis. This problem can be avoided by keeping gene values in the intensity format.

The two text files were converted to $\log_2$ scale so that the following equation could be applied (on a per gene basis) to standardize all arrays to each other:

$$Ch1_{adj} = Ch1 + (Ch2_{med} - Ch2)$$

$Ch1_{adj}$ is the adjusted Ch1 $\log_2$ intensity for each spot, Ch1 is the original Ch1 $\log_2$ intensity for that spot, $Ch2_{med}$ is the median $\log_2$ Ch2 intensity of all of the spots for that gene, and Ch2 is the individual $\log_2$ Ch2 intensity of each of the spots for that gene. Since all of the microarrays were hybridized with the same sample in Ch2, the Ch1 intensities can be directly compared when adjusted proportionally with the median (or

average) Ch2 intensity across all arrays being compared. This technique allows the microarrays to be analyzed without the use of ratios, and allows for the comparison of any condition to any other condition and across experiments as long as the same reference sample was used on all of the arrays. The output of the standardization procedure is a single file with adjusted Ch1 intensities for all of the microarrays originally downloaded.

A conservative selection of filters was applied to the genes downloaded from the LAD in order to maintain as many genes for statistical analysis as possible, while not compromising gene reliability. A text file with the adjusted Ch1 intensity values for 39,335 genes on 70 arrays was opened in Excel (Microsoft, Redmond, WA) for filtering. An outlier test was run in Excel on each group of 7 n arrays (10 groups total) to remove values that were greater than or less than +/- two times the standard deviation of the mean. Blank cells replaced outlier values in Excel. Genes were selected for further analysis if they had at least 6 out of 7 n values present. After this filter, 36,694 genes remained. Finally, an intensity filter was applied to maintain only those genes with at least one group of 6-7 n having an average net normalized and standardized intensity of 25 or greater, leaving 26,418 genes for statistical analysis.

Statistics

One of the major advantages of performing microarray experiments is that thousands of gene expression differences can be detected simultaneously. The goal is to identify which of those thousands of gene expression differences are truly "significant." Until a few years ago, it was common practice to identify significant genes by their fold change values. For example, any gene that had a two-fold increase or decrease in expression from control was deemed a significantly differentially expressed gene. Although this method of detection did produce gene lists that are often validated by more

stringent criteria today, it did not take into account the variance in the system. In addition, it failed to account for those genes with a less than two-fold change in expression that may have also contributed to the phenotype of interest. To solve this problem, statistical methods started to be incorporated into microarray analysis. A common statistical test used to identify differentially expressed genes between conditions A and B is the Student's *t*-test (developed by W.S. Gossett). The output of a *t*-test is a probability (p) value describing the difference between conditions A and B, while taking into account the variance in the samples representing each condition. As the number of samples representing an experimental condition increases, so does the accuracy of the estimate of experimental variance. For any given gene on a microarray, the higher the ratio of the expression difference between samples to the variance within samples, the lower the p-value describing the two experimental conditions being compared (and the more likely the detected gene expression difference is real). The samples themselves can be either technical replicates (running the same sample from a single organism on multiple arrays), biological replicates (running multiple samples from multiple organisms on multiple arrays) or a combination of both (Pavlidis, 2003). Technical replicates are often used for quality control purposes only, as they cannot accurately represent an experimental condition as biological replicates can. Ideally, a large number of biological replicates would be used to represent each experimental condition being compared in a microarray study. However, due to issues such as the cost of the arrays and/or limited sample availability, such a large sample size is not feasible. The general consensus is that a minimum of five biological replicates is necessary to identify differential gene expression using microarrays (Reimers, 2005).

With any method of statistics applied to a set of scientific data, the goal is always to find the appropriate balance of Type I and Type II errors. In a Type I error, the null

hypothesis is rejected when it should not be, generating false positives. In a Type II error, the null hypothesis is not rejected when it should be, generating false negatives (Zar, 1999). In trying to balance the generation of false positives and false negatives, scientists often settle on a false positive rate of 5% (p-value of 0.05), meaning that 5% of the time, the null hypothesis will be rejected erroneously. Although controlling both types of errors is important in microarray data, one could argue that the generation of false positives may be more detrimental considering microarrays are often used as hypothesis-generating techniques and therefore, it is often the case that genes identified as differentially regulated by microarrays are further characterized in the system of interest. Therefore, pursuing a gene that is not actually differentially regulated (a false positive) may lead to wasted time and money. With regard to experiments that test only one hypothesis at a time, a false positive rate of 5% is usually acceptable. However with microarray experiments, thousands of hypotheses are being tested simultaneously. Therefore, a p-value of 0.05 generates a large number of false positives (with 40,000 genes, 2000 false positives are generated). Several options exist for decreasing the number of false positives generated in a microarray experiment. Option 1 is to decrease the p-value cutoff, making the criteria for calling a gene "differentially expressed" more stringent. Option 2 is to apply a correction for multiple testing to the list of differentially expressed genes. A Bonferroni correction is the traditional method used to decrease the false positive rate in scientific data; however, it has been deemed too conservative for microarray data because it often generates false negatives at the expense of controlling false positives (Manly et al., 2004). Therefore, the focus has shifted in microarray experiments from controlling the false positive rate to controlling the false discovery rate (FDR). The FDR is the proportion of significant features (in the case of microarrays, genes identified as significantly differentially expressed) that are truly null (Benjamini

and Hochberg, 1995; Storey, 2003). In other words, an FDR of 0.3 means that out of 100 features (genes) deemed "significant" (using a specific p-value cutoff), 30 are potentially false positives. Storey proposed a modified version of the FDR called the positive FDR (pFDR) that generates a q-value for each feature as a measure of the feature's probability of being a false positive when called significant (Storey, 2003). The q-value is calculated based on the distribution of p-values, simultaneously considering that there are thousands of hypotheses being tested at once (Storey and Tibshirani, 2003). The pFDR correction can be applied to existing microarray data sets through the program QVALUE, which is freely available at http://faculty.washington.edu/~jstorey/qvalue/.

An extension of the *t*-test is the ANalysis Of VAriance (ANOVA) or *F*-test. Regarding microarray data, an *F*-test can be used to identify differentially expressed genes between three or more experimental conditions. Although it is possible to do a series of pair wise *t*-tests among the three or more experimental conditions (each represented by n samples) to identify differentially expressed genes, the chance of making a Type I error (generating false positives) increases as the number of *t*-tests increases. An *F*-test controls this error by comparing the means (and variances) of all n samples at once. As with the *t*-test, if the variance among samples is greater than the variance within them, it indicates that at least one of the experimental conditions differs (significantly) from all the rest (but it does not indicate which condition is different) (Churchill, 2004).

Not all microarray experimental designs are as simple as comparing two or three experimental conditions to one another. Experimental designs that incorporate a time course (such as studying the cell cycle, circadian rhythms or the effects of a drug over time) present a unique challenge when it comes to applying the appropriate statistical tests to them to identify differentially expressed genes. Genes are no longer just up- or

down-regulated compared to control—they are now regulated at one or several time points, that when taken together, represent expression *patterns*. Some genes will be up- or down-regulated at a single time point while others will be regulated at several time points. While an *F*-test could be applied to temporal data to identify gene expression differences among all the time points, it would not identify at which time points the expression differences existed. What is needed is a statistical test, that when applied to the microarray dataset, could classify differentially expressed genes as a function of time, according to their specific pattern of expression. Johansson et al. used Partial Least Squares (PLS) regression analysis (developed by Herman Wold) to accomplish this (Johansson et al., 2003). They were interested in identifying genes whose transcription was coupled to the yeast cell cycle. They used PLS to derive a response model that described an expected expression profile over time for periodically transcribed genes. Their model included two possible expression patterns—one matching a sine curve and the other matching a cosine curve. They applied the microarray results of three different yeast studies to their cell cycle models and found the same group of genes to be cyclically expressed in all three data sets. These same principles could be applied to time course data following the administration of a drug that causes known physiological and behavioral changes in the organism at specific time points. A defined set of biologically relevant expression patterns could be modeled using PLS, then the gene expression data could be fit to these patterns, producing a list of genes whose expression matched one of the pre-defined patterns.

The *t*-test, *F*-test and PLS analysis are three statistical tests that can perform complementary functions when applied to microarray data, especially when a time course is involved. It is important to apply the appropriate statistics with enough biological replicates to truly detect differential gene expression in a microarray experiment. This is

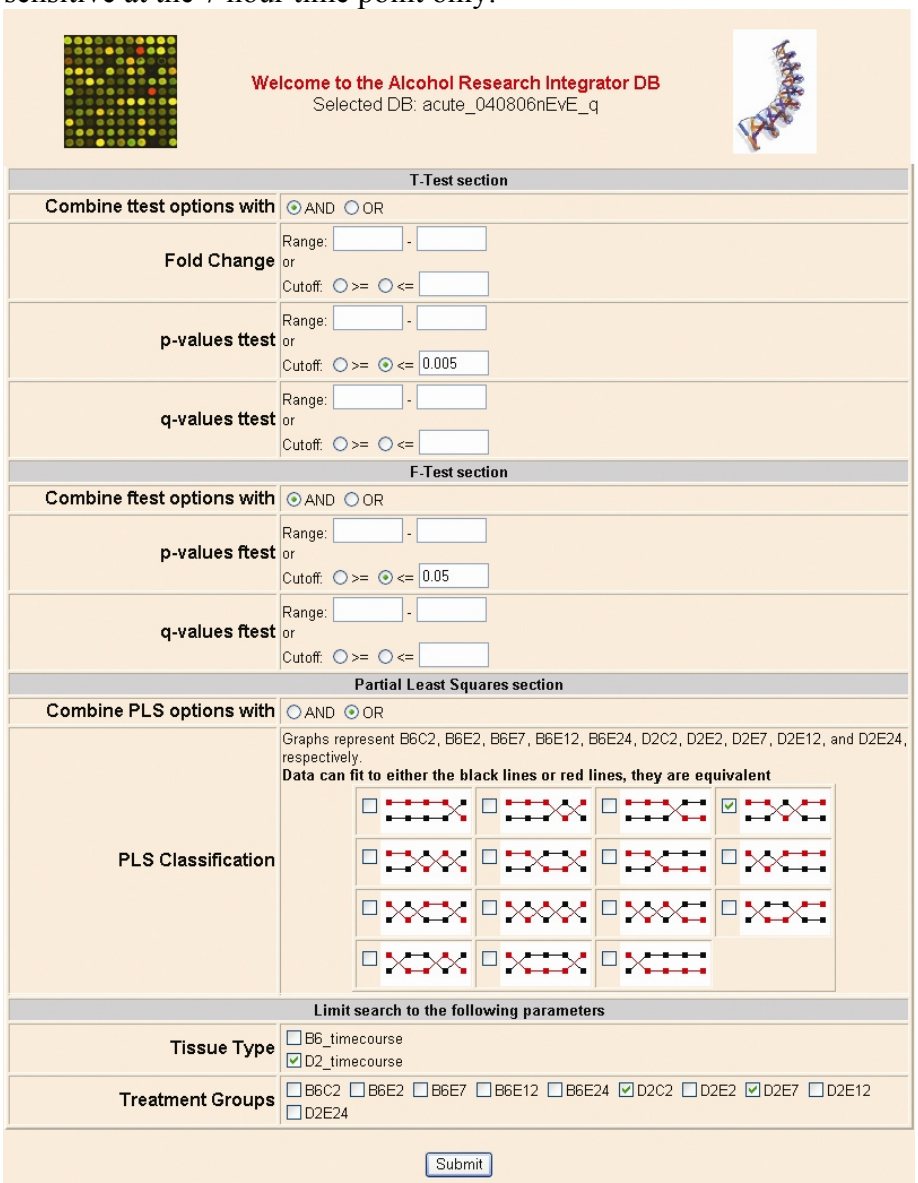
especially true when it comes to studying diseases of the brain. Important gene expression changes in the brain can be subtle, due to the changes occurring in a less well-represented cell type or as a result of an insult such as alcohol (it is well documented that alcohol causes many small changes in brain gene expression, as opposed to few large changes) (Nisenbaum, 2002; Pavlidis, 2003; Mirnics and Pevsner, 2004). Therefore, the ability to identify small but significant changes in gene expression is key in order to prevent false negatives and generate new hypotheses about the biology of the nervous system.

Statistical analysis of the microarray data included the application of *t*-tests, *F*-tests and Partial Least Square (PLS) analyses to identify significant changes in gene expression. All statistical analyses were performed using a combination of internally developed algorithms, R and R packages (Gentleman et al., 2004) from CRAN (Comprehensive R Archive Network (http://cran.r-project.org/)) and BioConductor (Gentleman et al., 2004; Carey et al., 2005). More specifically: Thirteen *t*-test comparisons (7 n arrays per group) were run including B6 control versus D2 control and B6 or D2 control versus a single B6 or D2 alcohol time point (respectively). Non-paired, two-tailed, equal variance *t*-tests were run using the mt.teststat vignette within the multtest package (v1.5.2) in R. Two *F*-tests were run, one for B6 and one for D2. Each *F*-test consisted of a control group and four alcohol treatment groups, with 7 n arrays per group. Two-tailed *F*-tests were run using the mt.teststat vignette within the multtest package (v1.5.2) in R. To correct for multiple testing in the *t*- and *F*-tests, the FDR was calculated and represented as a q-value for each gene (Storey and Tibshirani, 2003). Q-values were calculated based on the distribution of p-values for each *t*-test comparison and for each *F*-test group comparison (QVALUE software; http://faculty.washington.edu/~jstorey/qvalue/). Fold change was calculated as follows:

fold change = 2^((average of $\log_2$ (intensity) of 7 n of second group in pair wise comparison) – (average of $\log_2$ (intensity) of 7 n of first group in pair wise comparison)). If fold change < 1, then fold change = (-1)/fold change. In the B6 control versus D2 control *t*-test, the first group was B6 and the second was D2. In all control versus alcohol *t*-test comparisons, the first group was control and the second was alcohol. Prior to running the PLS analysis, missing values were imputed using a k-nearest neighbor algorithm within R using the Bioconductor package "impute" and its function "impute.knn." The PLS analysis will fail if missing data points are present. The PLS analysis was run using the pls2 function within the pls package in R. PLS was run on the B6 and D2 time courses separately, with each PLS analysis consisting of a control time point and four alcohol time points. With five time points a total of 30 possible gene expression patterns exist. A binary counting algorithm was used to create 30 dummy codes describing each possible gene expression pattern (e.g. 00100 and 11011). Each gene's pattern of expression was fit to one of the 30 possible patterns of expression and a loading score was calculated which is indicative of how well the two match. The loading scores allowed assignment of a single best fit pattern of expression to each gene. The loading scores are converted to Z scores, then to p-values. The smaller the p-value, the better the data fits the expression profile in question relative to all other genes in the system. The results of the *t*-test, *F*-test, fold change, False Discovery Rate (FDR) and PLS analyses were loaded into the Alcohol Research Integrator (ARI) gene database (Bergeson et al., 2005). The ARI database was developed to provide a means with which to search for specific genes, p-values, q-values, fold changes, *t*-test comparisons, *F*-test comparisons and PLS patterns (Figure 2.12). Gene lists matching the search criteria can be downloaded from the database as a text file for further analysis.

To generate lists of statistically significant differentially expressed genes, gene lists were downloaded from the ARI database that matched the following search criteria: (1) For the B6 control versus D2 control query, a *t*-test p-value of 0.001 was used, (2) For the B6 and D2 control versus the two hour alcohol time point query, a *t*-test p-value of < 0.05 was used and (3) For the control versus alcohol time point(s) within a single genotype query, a *t*-test p-value of 0.005 (connected by "OR" for multiple time points), *F*-test p-value of 0.05 and single PLS pattern were used. For query #1, the genes were sorted by fold change (descending). The gene list of the intoxication-related genes (#2 above) was further filtered to maintain only those genes with directionality agreement in both strains. For query #3, the 20 separate gene lists incorporating 10 different PLS patterns (10 patterns x 2 genotypes) represented biologically relevant patterns of gene expression across time, according to our model of alcohol dependence.

Figure 2.12: Alcohol Research Integrator (ARI) gene database search page. An example of specific search criteria used to generate gene lists is shown below. The gene list generated from this criteria would be the result of: 1) A *t*-test comparison between D2 control and D2 alcohol treatment at the 7 hour time point, with a p-value cutoff of 0.005, 2) An *F*-test comparison of D2 control versus all D2 alcohol time points with a p-value cutoff of 0.05 and 3) A PLS pattern within D2 where the genes are alcohol-sensitive at the 7 hour time point only.



Clustering

Microarray experiments can produce thousands to tens of thousands of data points, making data analysis particularly challenging. One method used to reduce the large amount of data to a more easily interpretable form is to organize it into meaningful structures. Cluster analysis, as it applies to microarray data, accomplishes this task by grouping genes with similar expression profiles together (Slonim, 2002). Several different types of cluster analysis exist. The two more commonly used clustering algorithms applied to microarray data are hierarchical clustering and K-means clustering. In both cases, a distance or similarity metric is used to determine how similar (or dissimilar) the expression profiles are of the genes being clustered (McShane et al., 2003). In hierarchical clustering, the expression profile of every gene is compared to the expression profile of every other gene to identify the two genes with the most similar expression profile, which join to form a cluster. This process proceeds iteratively, each time comparing all existing clusters for similarity, then joining the most similar ones to create a new cluster. The process stops when all the genes have been assigned to a cluster (Brazma and Vilo, 2000; Quackenbush, 2001). Hierarchical clustering results are often depicted by a tree structure called a dendrogram (McShane et al., 2003). The distance from one branch of the tree to the next represents the degree of similarity between the two elements in the branches (Quackenbush, 2001). In hierarchical clustering, the number of clusters formed depends on the data and is not pre-specified, unlike in K-means clustering, where the algorithm is told how many clusters to generate. In K-means clustering, genes are randomly assigned to one of the pre-specified number of clusters. Genes are moved from one cluster to another while the algorithm measures the similarity between their expression profiles to determine if they should remain in the

new cluster or be moved to a different one. All the genes in each cluster must have similar expression profiles to each other and different profiles than the genes in the other clusters. The process stops when the intra-cluster similarity and inter-cluster dissimilarity are greatest (Quackenbush, 2001). Due to the nature of K-means clustering in that the data must be fit to a predefined number of clusters, K-means cluster analysis is often combined with techniques such as Principal Component Analysis (PCA) or PLS analysis that can estimate the number of clusters to which the data should be fit (Quackenbush, 2001). The results of cluster analysis are often displayed as a heat map, which is a rectangular array of colored boxes where each color represents a gene expression value in the sample (McShane et al., 2003; Shannon et al., 2003). Each column of boxes represents an array while each row of boxes represents a gene. The heat map contains patches of color that correspond to the clusters formed and are indicative of the expression levels of the genes within each cluster.

One of the major reasons why cluster analysis is applied to microarray data is to determine the functional and/or regulatory relationship(s) between genes that cluster together. Many of the cDNAs surveyed on microarrays are actually ESTs whose function is unknown. By grouping ESTs with genes of known function, the function of the ESTs can be inferred. The hypothesis is that genes in a cluster must share some common function, be part of the same pathway or have similar upstream regulatory elements (Quackenbush, 2001). Two keys studies from 1998 were the first to demonstrate a functional and/or regulatory relationship between genes in the same cluster (Eisen et al., 1998; Spellman et al., 1998).

Although cluster analysis has been applied to hundreds if not thousands of microarray data sets to identify groups of co-regulated genes, cluster analysis does not have a probabilistic foundation and is therefore only an exploratory tool meant to identify

candidate genes for further study (Heyer et al., 1999; Shannon et al., 2003). It cannot identify statistically significant differences in gene expression between experimental conditions (Reimers, 2005). In addition, especially with regards to hierarchical clustering, clustering unrelated data will still produce clusters, which would represent noise as opposed to biological meaning (Quackenbush, 2001; McShane et al., 2003).

To visualize the patterns of expression within the 22 gene lists downloaded from the ARI database, a K-means clustering algorithm was applied to each of the lists. K-means clustering was chosen over hierarchical clustering because each gene list contained two distinct patterns of expression: (1) For the control versus alcohol lists, the pattern contained genes that were up- or down-regulated compared to control and (2) For the control versus control list, the pattern contained genes that were up-regulated in either B6 or D2 control mice. Prior to the application of the clustering algorithm, duplicate genes within each of the 20 gene lists were removed. In addition, 31 genes were removed from the B6 control versus D2 control gene list because they were also alcohol-sensitive (present in one of the 20 control versus alcohol gene lists).

All 22 gene lists were uploaded into the LAD to obtain $\log_2$ of R/G normalized ratio (median) values for each gene, which are required for the clustering program to run properly. The lists were queried separately in the LAD, with the following criteria: (1) For the B6 control versus D2 control gene list and for the B6 and D2 control versus the two hour alcohol time point gene list, the $\log_2$ ratios were obtained using B6 and D2 arrays, in chronological order (with B6 arrays listed first). (2) For the control versus alcohol time points within a single genotype gene list, the $\log_2$ ratios for all 10 B6 gene lists were obtained using only B6 arrays, in chronological order. The $\log_2$ ratios for all 10 D2 gene lists were obtained using only D2 arrays, in chronological order. The $\log_2$ ratios for each gene were centered by median values (which mimics the gene-wise

standardization procedure we use on intensity values). The $\log_2$ ratio data was downloaded as .pcl files (22 files, one for each gene list) from the LAD.

The .pcl files were opened one at a time in Cluster 3.0 (http://bonsai.ims.u-tokyo.ac.jp/~mdehoon/software/cluster/software.htm#ctv) (de Hoon et al., 2004) and K-means clustering was applied to each. Clustering was performed on the x-axis only (genes, not arrays) with a specification of 2 clusters (because each PLS pattern is bi-directional with respect to control or bidirectional between genotypes) and a similarity metric of (Pearson) correlation (uncentered). The number of runs specified was 100 and the algorithm was run several times to obtain the highest number of solutions found (the number of times the same cluster result was found). The output file from Cluster 3.0 is a .cdt file. Each .cdt file was opened in Java TreeView (Saldanha, 2004) to visualize and process each cluster. For aesthetic purposes, the red/green orientation of some of the clusters had to be reversed. The reversal of the red and green colors did not affect the data. Once the clusters were in the correct orientation, x and y coordinates were specified to size the cluster correctly, at which point the clusters were exported as .png thumbnail images.

To generate the images of non-clustered genes that correlate with the clustered images, 20 gene lists were re-loaded into the LAD. The 20 gene lists were either the original gene list or the newly sorted list, based on whether the cluster image had to be reversed. The $\log_2$ of R/G normalized ratio (median) values were downloaded from the LAD for each gene in all lists in a similar fashion as stated above, with one exception: The $\log_2$ ratios for all 10 B6 gene lists were obtained using only D2 arrays, whereas the $\log_2$ ratios for all 10 D2 gene lists were obtained using only B6 arrays. 20 “fake” .cdt (non-clustered) files were visualized in TreeView and saved in a similar fashion as above. Thus, the order of genes in the clustered images matches that of the non-clustered images.

BIOINFORMATICS

Functional analysis

The differential gene expression profile of an organism in response to an experimental condition can only begin to give insight into the biology of that condition. It is often difficult to determine what type(s) of biological processes are occurring at the cellular level based on gene names alone, especially if hundreds of genes are differentially expressed. Functional classification of co-expressed genes is a method toward understanding more about the global cellular response to a particular experimental condition. Functional classification can accomplish the following: 1) Determine whether the co-expressed genes are functionally related, 2) Predict the function of uncharacterized genes and ESTs, 3) Correlate the identified functions with physiological or behavioral phenotypes occurring in the organism at the same time and 4) Give insight into which genes should be prioritized for further analysis.

Several years ago the Gene Ontology (GO) Consortium was created to produce a "structured, precisely defined, common, controlled vocabulary for describing the roles of genes and gene products in any organism" (Ashburner et al., 2000; Harris et al., 2004). Within the GO database, genes are classified in three different ways: according to their molecular function, biological process or cellular location. The GO database (http://www.geneontology.org) is a useful tool to functionally classify large numbers of genes simultaneously. Whereas the GO database classifies genes according to their function, the Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.jp/keg) (Kanehisa, 1997; Ogata et al., 1999) and BioCarta (http://www.biocarta.com) databases classify genes according to which biochemical pathway they belong. All three of these databases, and many others like them, provide useful insight into the functional nature of a co-expressed list of genes.

In recent years, statistics began to be incorporated into the functional classification of genes as a measure of how significant the identified functional groups were. The functional groups associated with a reference group of genes (all the genes on the array or all the genes in an organism, for example) are compared to the functional groups associated with list of differentially expressed genes. For each functional group associated with the list of expressed genes, a p-value is calculated that represents the probability that that functional group could have been identified by chance, considering its prevalence in the reference list of genes (Wu et al., 2002; Beissbarth and Speed, 2004).

WebGestalt (WEB-based GEne SeT AnaLysis Toolkit; http://genereg.ornl.gov/webgestalt) (Zhang et al., 2005) is an online data mining system for large gene sets that incorporates statistical analysis into the GO, KEGG and BioCarta database results. Two options for statistical tests are available: the hypergeometric test and Fisher's exact test. The tests are used to determine whether the functional categories present in the gene list are significant or occur by chance due to their prevalence in the reference gene list. The user can select at what p-value the functional annotations are deemed significant and also the minimum number of genes that must compose a functional category before it can be deemed significant. Significant functional groups are colored red in the results output of the GO, KEGG and BioCarta database queries.

Transcripts that passed the statistical thresholds were functionally annotated using WebGestalt followed by overrepresentation analysis. The 20 B6 and D2 gene lists were combined (within genotype only) in different ways to create new gene lists, each representing a different biological process (intoxication, withdrawal or neuroadaptation) occurring in the mouse as a result of the experimental paradigm. See Table 4.2 for the

PLS patterns composing each biological process. The new gene lists are genotype-specific and contain both up- and down-regulated genes compared to saline control.

Three gene set organization sub-modules within WebGestalt were used to functionally classify gene lists: 1) GOTree (Zhang et al., 2004), 2) KEGG Table and Maps and 3) BioCarta Table and Maps. The gene identifier used to query functional information in WebGestalt was gene symbol. WebGestalt compared the lists of differentially expressed genes with a reference list of genes to identify overrepresented functional categories. The reference list was all the genes entered into the ARI database. This list contained 10,977 unique gene symbols. Significance was set at a hypergeometric test p-value cutoff of 0.05 and a minimum of three genes within each functional group annotation. Within the GOTree module, the classification type chosen to represent function was Biological Process. Significant functional categories that were non-specific (eg "cellular physiological process") were excluded from the results.

Transcription Factor Binding Site (TFBS) analysis

Transcription factors regulate gene expression by binding to specific DNA sequences upstream of the transcription start site and initiating transcription. These sequences are referred to as Transcription Factor Binding Sites (TFBSs). It has been hypothesized that genes that are co-expressed (share the same expression profile) are also (potentially) co-regulated (share the same TFBSs) (Brazma and Vilo, 2000; Allocco et al., 2004). Several studies have explored this relationship and found evidence to support this hypothesis. Hannenhalli and Levy showed that functionally linked interacting proteins had a higher proportion of shared TFBSs than proteins that do not interact experimentally (Hannenhalli and Levy, 2003). Allocco et al. found that genes with highly similar expression profiles or functional annotations were more likely to share a common transcription factor than those that were unrelated (Allocco et al., 2004).

Several investigators have provided indirect evidence for this hypothesis by demonstrating that clustered genes often share common upstream sequence motifs (Allocco et al., 2004).

By searching for common TFBSs among genes with similar expression profiles, insight can be gained into the composition and function of the underlying regulatory networks responsible for altering gene expression as a result of a particular experimental condition. The presence of a TFBS upstream of a gene does not necessarily mean that its concomitant transcription factor is important for the regulation of that gene, but it does guide further analysis.

As with functional analysis, TFBS analysis has also been recently subjected to the application of statistics to identify significant TFBSs in a differentially expressed set of genes relative to a reference set (all the genes on the array or all the genes in an organism, for example). oPOSSUM v1.3 is an online program that identifies statistically overrepresented TFBSs in gene promoters (Ho Sui et al., 2005). The program identifies TFBSs by using Position Specific Scoring Matrices (PSSMs) and phylogenetic footprinting. PSSMs are a measure of the binding specificity of a transcription factor to its DNA binding motif. Because TFBS motifs are often short and degenerate, the use of PSSMs alone to identify them often results in a high false positive rate. To address this issue, oPOSSUM searches for TFBSs only in conserved and non-coding regions. This tactic, called phylogenetic footprinting, is based on the assumption that functional non-coding sequences are more likely to be conserved across closely related species than non-functional sequences. This method eliminates about 80% of the sequence that would otherwise be used to search for TFBS; however, it has been estimated that when comparing mouse and human sequences, the proportion of TFBSs occurring within conserved regions is about 70% (Ho Sui et al., 2005).

oPOSSUM uses two different complementary statistical tests to identify overrepresented TFBSs: a Z-score calculation and a one-tailed Fisher exact probability. The results output includes a list of the transcription factors that bind to the overrepresented TFBSs along with the differentially expressed genes from the input list that contain that TFBS sequence. The transcription factors are ranked in two ways—according to their Z-score and their p-value (from the Fisher exact probability). Several user-defined parameters, including the level of conservation, matrix match threshold and number of base pairs upstream and downstream from the transcription start site, guide oPOSSUM in its search for TFBSs.

TFBS analysis was performed using oPOSSUM to determine whether genotype-specific differentially expressed genes contained unique TFBSs and therefore may be under the control of unique transcription factors. Intoxication-related, strain-specific gene lists were divided into four groups: (1) B6-up, (2) B6-down, (3) D2-up and (4) D2-down and entered into oPOSSUM. oPOSSUM compared the lists of differentially expressed genes with a reference gene list (all the genes entered into the ARI database) to identify overrepresented TFBSs. A custom analysis was run in oPOSSUM with the following parameters: 1) Taxonomic supergroup: vertebrate, 2) conservation level: 30%, 3) matrix match score: 85%, 4) upstream sequence length: 2000 base pairs, 5) downstream sequence length: 500 base pairs. Significance was set at a Fisher p-value of < 0.05 and/or a Z-score of 7 or greater. It should be noted that some TFBSs (e.g. HNF-3 beta and HFH-3) have similar consensus sequences.

Cellular classification of differentially expressed genes

Differentially expressed genes were examined for their enrichment in different neuronal populations. The method used was similar to that of (Ponomarev et al., 2006) and is briefly described here. Significant genes were queried in the Mouse Neuronal

Expression Database (MNED, http://mouse.bio.brandeis.edu) and their values in seven neuronal cell types (three excitatory (glutamate) and four inhibitory (GABA)) from cerebral cortex were retrieved (Sugino et al., 2006). Genes that had a difference in expression of 2-fold or higher across the seven neuronal populations in the MNED were clustered using K-means cluster analysis. A chi square analysis was applied to the genes enriched in the excitatory and inhibitory cell populations to determine whether a significant difference existed in the PLS pattern distribution between the two populations.

CHAPTER 3: GENERATION OF MICROARRAY TARGETS BY DIFFERENTIAL DISPLAY

Introduction

The human and mouse genomes both contain 25,000 to 30,000 genes, of which 99% are similar in sequence. Therefore it is likely the expression of those genes that makes the human and mouse appear so different (Waterston et al., 2002).

The complete sequence of the mouse (C57BL/6J) genome was not published until late 2002 (Waterston et al., 2002). Therefore, the initial clone sets purchased for the cDNA microarray analysis were incomplete and did not represent the entire mouse transcriptome. An mRNA differential display (DD) screen was performed to search for novel differentially expressed genes in the B6 and D2 mice as a result of acute alcohol exposure. The goal was to generate previously uncharacterized genetically divergent and alcohol-sensitive transcripts that could be added to the cDNA microarrays as additional targets. Several attributes of DD made it an attractive choice for this purpose: 1) it is an unbiased technique that allows novel transcripts to be detected that may not have been intuited by a hypothesis-driven approach, 2) it is highly sensitive and can detect expression differences as small as 20%, 3) it allows a reasonably complete representation of all the genes in a sample to be seen, including some "rare" transcripts, 4) it can detect sequence polymorphisms, insertions and deletions, as well as absolute differences in expression.

Two biologically relevant time points (two hours and six hours) were chosen to search for genetically divergent, alcohol-sensitive changes in gene expression as a result of acute alcohol administration. The two hour time point represents the state of

intoxication and the six hour time point represents the state of physical dependence (6-8 hours post acute alcohol administration is the peak of withdrawal).

Results

The DD screen identified differentially regulated known genes, ESTs and previously uncharacterized transcripts between the B6 and D2 mice and as a result of alcohol exposure. Approximately 1.3% percent of the total genes screened were differentially regulated. Of these, about 24% represented genetic differences between the strains, 45% represented alcohol-sensitive differences and 36% were a combination of both. Figures 3.1 and 3.2 show two of the differentially expressed genes identified in the DD screen.

Figure 3.1: Example of an alcohol-related gene identified by DD. 2-oxoglutarate dehydrogenase is expressed 2 hours and 6 hours post acute alcohol administration. This gene was not included in any of the purchased cDNA clone sets and therefore represents a novel additional target printed on the microarrays as a direct result of the DD screen.

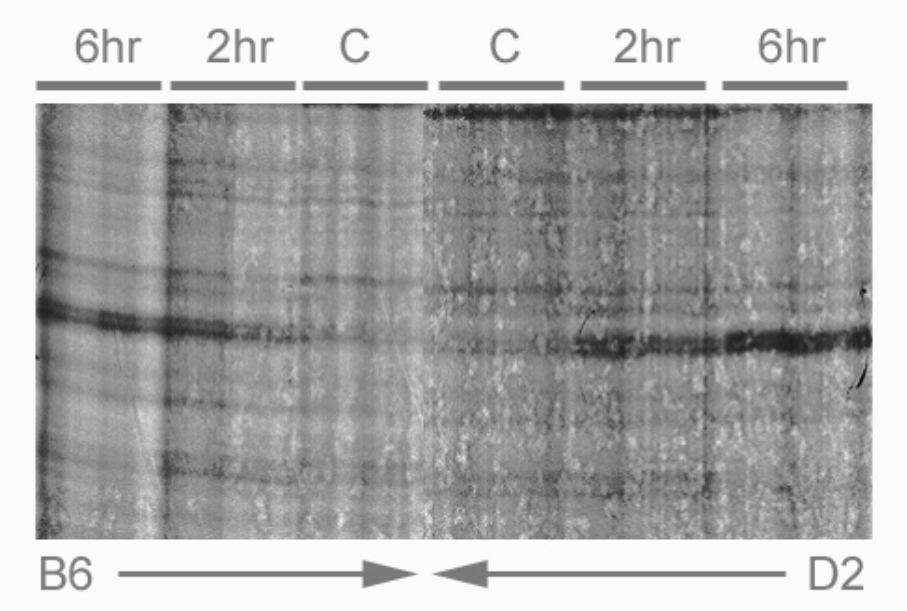
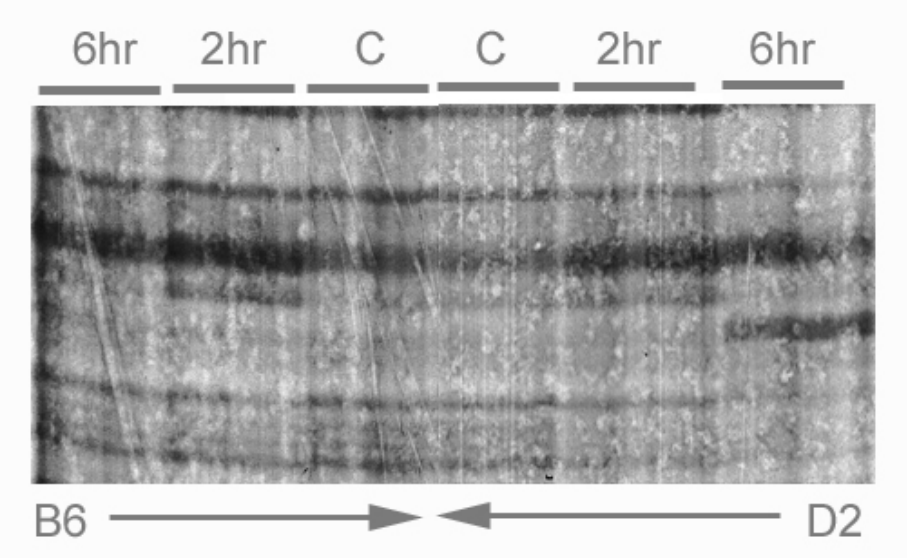
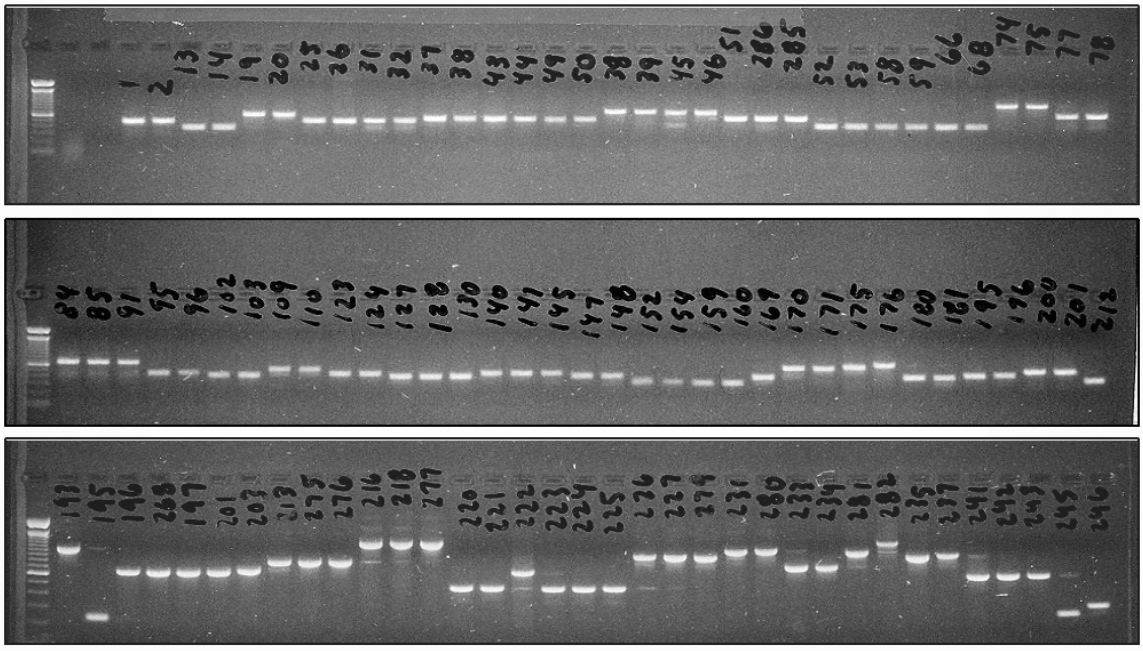


Figure 3.2: Example of a genetically divergent, alcohol-related gene identified by DD. N-deacetylase/N-sulfotransferase is expressed at 2 hours in both strains, with slightly higher expression in the B6 mouse. It is also expressed at 6 hours post acute alcohol administration, but only in the D2 strain.



The clones isolated from the DD screen were PCR amplified prior to their addition as targets on the microarrays. A small portion of the PCR product was run on an agarose gel to ensure the presence of a single band (and therefore no contamination). Any clone with greater than one band was either re-run or discarded as a potential target. Figure 3.3 shows a sampling of DD clones on an agarose gel, most of which were added to the microarrays.

Figure 3.3: Agarose gel with PCR products of amplified DD clones. Most of the clones contained a single band and were therefore cleared as new targets to be printed on the microarrays.



The DD clones added to the microarrays represent about 150 unique known genes, ESTs and unknown sequences (Table A.1). Many of them are not included in any of the purchased clone sets and therefore are unique targets that are only present on the arrays as a result of the DD screen. These novel genes include: calcyphosine, 2-oxoglutarate dehydrogenase, aczonin, catenin alpha 2, myelin proteolipid protein, minor histocompatibility antigen precursor, transmembrane 4 superfamily member 8, myotonic dystrophy kinase-related Cdc42-binding kinase and TU3A, which is a human orthologue of a novel mouse gene that has not yet been characterized.

Discussion

The DD screen identified several novel genetically divergent and alcohol-sensitive transcripts that would not have been included in the cDNA microarray analysis if they had not been discovered through DD. Therefore, the DD clones represent valuable reagents that enhanced the search for gene expression differences between the B6 and D2 mice over a time course following an acute dose of alcohol. DD served as a complementary method to the cDNA microarray analysis, allowing more genes to be identified and serve as candidates to hypothesis-test for a potential role in the development of alcohol dependence.

CHAPTER 4: ALCOHOL-INDUCED TEMPORAL TRANSCRIPTOME REMODELING IN MOUSE PREFRONTAL CORTEX

Abstract

The role specific anatomical brain regions play in the manifestation of alcohol-related traits and endophenotypes is poorly understood. Autopsies have shown that the prefrontal cortex is selectively damaged in long-term alcoholics (Kril and Halliday, 1999). Therefore, we chose to study the prefrontal cortex and its role in the initiation of alcohol dependence using the inbred mouse strains, C57BL/6J (B6) and DBA/2J (D2). B6 and D2 mice differ for numerous measures of alcohol-induced outcomes, including alcohol dependence. Nearly 3 million cDNA microarray data points were analyzed across genotype and a 24 hour time course for changes in gene expression following intoxication-related anesthesia and acute physical dependence induced by a single large dose of alcohol (4 g/kg, i.p.). Unique genotype-specific temporal patterns of gene expression were identified for physiologically relevant alcohol-related consequences including: intoxication, withdrawal and neuroadaptation. *In silico* analysis of alcohol-sensitive genes revealed genotype-specific patterns of transcriptional regulation and functional activity, which may contribute to strain differences in alcohol-related behaviors. Categorization of alcohol-sensitive genes by their cellular profiles demonstrated a divergence in gene expression between excitatory and inhibitory neurons. The results of the study identified genes of both known and unknown function as important candidates underlying the complex molecular responses to alcohol.

Introduction

Alcohol is one of the most widely used drugs. Its abuse can lead to the development of alcoholism (alcohol dependence), a disease with a nearly 20% lifetime risk in the United States (Hasin et al., 2005). Alcoholism is classified clinically as including physical dependence, tolerance and a maladaptive pattern of alcohol use (APA, 2000). It is a complex disease influenced by both environmental factors and genetic predisposition (Goodwin et al., 1973; Cloninger et al., 1981; Enoch and Goldman, 2001; Crabbe, 2002a). Rodent models have long been used to study specific alcohol-related phenotypes and are now considered valuable molecular tools. C57BL/6J (B6) and DBA/2J (D2) inbred strains of mice are the most widely used mouse model to study alcoholism as they differ for nearly every alcohol response (Crabbe et al., 1999), including acute physical dependence (Buck et al., 1997).

Gene expression analyses of several rodent models have led to a better understanding of alcohol-induced transcriptional response and genetic influences on alcohol-related traits (Xu et al., 2001; Rimondini et al., 2002; Hoffman et al., 2003; Saito et al., 2004; Treadwell and Singh, 2004; Bergeson et al., 2005; Edenberg et al., 2005; Kerns et al., 2005; MacLaren and Sikela, 2005; McBride et al., 2005; MacLaren et al., 2006; Mulligan et al., 2006). Each of the studies identified candidate genes, which individually are likely to contribute to the overall complexity of a specific alcohol-related behavior or physiological effect. The results of a recent meta-analysis for alcohol preference underscore the value of transcriptome analysis in understanding the complex nature of the genetics of alcohol-related traits (Mulligan et al., 2006). Ultimately, important allelic differences that contribute to unique, innate responses to alcohol need to

be identified. For example, a role for *Mpdz* in alcohol withdrawal symptoms has been successfully established (Fehr et al., 2004).

In humans, the prefrontal cortex (PFC) is associated with higher cognitive functions such as problem solving, critical thinking, planning, judgment and impulse control (Funahashi, 2001), all of which can be negatively affected by alcohol use. The PFC is part of the mesocorticolimbic dopamine (DA) pathway, which is activated by alcohol (Bowirrat and Oscar-Berman, 2005). Positive reinforcing properties act in the "reward" pathway by causing an increase in DA leading to the compulsion to drink (Koob, 1992). Over time, alcohol-related selective damage to the PFC occurs, as is well documented in post-mortem analyses of chronic alcoholics (Kril et al., 1997; Kril and Halliday, 1999). Several post-mortem expression studies in humans have shown significant alcoholism-related gene changes in cortex (Lewohl et al., 2000; Mayfield et al., 2002; Sokolov et al., 2003; Iwamoto et al., 2004; Flatscher-Bader et al., 2005; Liu et al., 2006).

The primary goal of the current study was to use a standard B6,D2 mouse model of acute physical dependence (Buck et al., 1997) coupled with cDNA microarray analysis to monitor changes in PFC gene transcription over a 24 hour time course following alcohol administration. Incorporation of a time course allows the gene expression changes concomitant with intoxication, withdrawal and neuroadaptation—three biologically relevant stages of acute alcohol response—to be compared. A genotype × time paradigm allowed the molecular fingerprinting of genetic divergence within the PFC transcriptome, which may underlie phenotypic differences in the initiation and development of alcohol dependence between the B6 and D2 strains.

Materials and Methods

ANIMALS AND ANIMAL HUSBANDRY

Male C57BL/6J (B6) and DBA/2J (D2) mice, 70-100 days old, were group-housed under a 12 hour light/dark cycle with food and water available *ad libitum*. All animal procedures were IACUC approved according to NIH and AAALAC guidelines.

EXPERIMENTAL DESIGN AND TISSUE COLLECTION

B6 and D2 mice were habituated to injection for at least three days with a 0.5 ml i.p. injection of saline (0.9% sodium chloride). On the day of the experiment, the mice were injected in random order with 4 g/kg, i.p. ethanol (20% v/v in 0.9% sodium chloride) or an equivalent volume of saline. Alcohol-injected mice were sacrificed by cervical dislocation at 2, 7, 12 and 24 hours post-injection (n = 7 per group) while saline control mice were sacrificed at 2 hours post-injection (n = 7 per group). The prefrontal cortex was dissected from whole brain and immediately frozen in liquid nitrogen within two minutes of sacrifice and three minutes of removal of the mouse from its home cage. Prefrontal cortex was isolated by making an incision along the *fissura longitudinalis cerebri*, peeling the cortex flat and dissecting the anterior portion of the brain back 1.7 mm posterior, approximately 1.5 mm rostral to bregma, including portions of the following anatomical regions: agranular, cingular, frontal association, motor, orbital, pre- and infralimbic cortex (Franklin and Paxinos, 1997).

RNA ISOLATION AND QUANTIFICATION

Individual brain samples were homogenized in 1 ml of RNA STAT-60 (Tel-Test, Inc., Friendswood, TX) using a Polytron homogenizer (Kinematica, Lucerne, Switzerland). Total RNA was isolated according to the manufacturer's protocol. RNA

was resuspended in 15 µl of nuclease-free water (Ambion, Austin, TX) and quantified with a NanoDrop ND-1000 UV/Vis Spectrophotometer (NanoDrop Technologies, Wilmington, DE). The quality of the RNA was determined using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Samples were stored at –70°C.

cDNA MICROARRAY FABRICATION AND HYBRIDIZATION

Clones printed on the University of Texas at Austin custom cDNA microarrays included the NIH BMAP UniGene library (~11,000 clones), Sequence Verified IMAGE clones (~5000) (ResGen/Invitrogen, (Carlsbad, CA) and The National Institute of Aging (~22,000) cDNA clone set. Based on assignment to Unigene clusters, approximately 25,000 distinct genes were represented in the clone sets. cDNA clones were amplified by polymerase chain reaction (PCR) and printed on poly-L-lysine coated glass microscope slides (Gold Seal, Erie Scientific Co., Portsmouth, NH) using an arraying robot as previously reported (DeRisi et al., 1997).

Cy-dye-labeled cDNA was generated from 500 ng of total RNA using the 3DNA Array 900 kit (Genisphere, Hatfield, PA) according to the manufacturer's protocol. Sample cDNA from alcohol-treated or saline control animals (Cy3-labeled) was always hybridized against a Cy5-labeled standard reference (whole brain RNA isolated from 100 male B6 mice). The dye-labeling procedure and all subsequent steps were performed under minimal lighting conditions.

Hybridization took place in humidified chambers that fit up to 10 microarray slides (Genetix Limited, United Kingdom). Chambers were balanced for groups and no more than two chambers were used simultaneously. Ozone degradation of dye was prevented by keeping slides in an argon atmosphere following hybridization and prior to scanning. Slides were scanned using a GenePix 4000A or B scanner and spots were

labeled and quantified using SpotReader (Niles Scientific, Portola Valley, CA) and/or GenePix Pro 5.1 (Axon Instruments, Inc., Union City, CA).

The overall rate of post-hoc confirmation of single gene expression for our cDNA platform is greater than 80% (Bergeson lab, unpublished results and see also (Ponomarev et al., 2006)).

NORMALIZATION AND STORAGE OF MICROARRAY DATA

Microarray normalization was performed using the R (Ihaka and Gentleman, 1996) package: Statistics for Microarray Analysis (Yang et al., 2002). Specifically, the lowess within-print-tip-group function was used for normalization. A program developed in our lab that automates the normalization procedure was used to apply the normalized values to GenePix Results (GPR) files. The normalized GPR files were uploaded and stored in the Longhorn Array Database (LAD) (Killion et al., 2003).

RETRIEVAL AND STANDARDIZATION OF MICROARRAY DATA

Microarray data was retrieved from the LAD as net (background-subtracted) median intensity values. The Channel 1 (Ch1 (alcohol-treated or control Cy-3 labeled sample)) and Channel 2 (Ch2 (reference Cy-5 labeled sample)) intensities of every gene (39,335) on all 70 arrays were downloaded as two separate text files. The two files were converted to $\log_2$ scale so that the following equation could be applied (on a per gene basis) to standardize all arrays to each other:

$$Ch1_{adj} = Ch1 + (Ch2_{med} - Ch2)$$

$Ch1_{adj}$ is the adjusted Ch1 $\log_2$ intensity for each spot, Ch1 is the original Ch1 $\log_2$ intensity for that spot, $Ch2_{med}$ is the median $\log_2$ Ch2 intensity of all of the spots for that gene, and Ch2 is the individual $\log_2$ Ch2 intensity of each of the spots for that gene. Since all of the microarrays were hybridized with the same sample in Ch2, the Ch1

intensities can be directly compared when adjusted proportionally with the median (or average) Ch2 intensity across all arrays being compared. This technique allows the microarrays to be analyzed without the use of ratios, and allows for the comparison of any condition to any other condition and across experiments as long as the same reference sample was used on all of the arrays. The output of the standardization procedure is a single file with adjusted Ch1 intensities for all of the microarrays originally downloaded. An outlier test was run for each group to remove spots that were +/- two standard deviations from the group mean. Outliers were replaced with blanks and genes with fewer than six values were removed leaving 36,694 clones. Finally, an intensity filter was applied to maintain only those genes with at least one group mean intensity of 25 or greater, leaving 26,418 genes for further analysis.

DATA ANALYSIS AND STATISTICAL METHODS

Statistical analysis of the microarray data included the application of *t*-tests, *F*-tests and Partial Least Square (PLS) analyses to identify significant changes in gene expression. The *t*-tests (non-paired, two-tailed, homoscedastic) were used to compare control versus control or control versus alcohol samples. The *F*-tests (two-tailed) were used to compare all time points (samples of control through 24 hours) within a single genotype. PLS analysis was performed using all possible fit models (patterns) for the defined data set, which was calculated using a binary counting algorithm that yielded dummy codes to describe each pattern for each test (i.e. – 00011 for five comparisons, with each number corresponding to a control or alcohol-related time point). All statistical analyses were performed using a combination of internally developed algorithms, R and R packages (Gentleman et al., 2004) from CRAN (http://cran.r-project.org) and BioConductor (Gentleman et al., 2004; Carey et al., 2005). Fold change was calculated by comparing group averages of control arrays to alcohol arrays (or control to control).

The results of these analyses were loaded into the Alcohol Research Integrator database (ARIdb) (Bergeson et al., 2005). K-means cluster analysis was performed using Cluster 3.0 (de Hoon et al., 2004) (http://bonsai.ims.u-tokyo.ac.jp/~mdehoon/software/cluster/software.htm) and cluster images were generated using Java TreeView (Saldanha, 2004).

Gene lists were downloaded from the ARIdb that matched the following search criteria: (1) For the B6 control versus D2 control query, a *t*-test p-value of < 0.001 was used, (2) For the B6 and D2 control versus the two hour alcohol time point query, a *t*-test p-value of < 0.05 was used and (3) For the control versus alcohol time point(s) within a single genotype query, a *t*-test p-value of < 0.005, *F*-test p-value of < 0.05 and single PLS patterns were used. The gene list of the intoxication-related genes (#2 above) was further filtered to maintain only those genes with directionality agreement in both strains. For query #3, the 20 separate gene lists incorporating 10 different PLS patterns (10 patterns x 2 genotypes) represented biologically relevant patterns of gene expression across time, according to our model of alcohol dependence.

TRANSCRIPTION FACTOR BINDING SITE (TFBS) OVERREPRESENTATION ANALYSIS

The oPOSSUM database was used to analyze transcription factor binding sites (TFBS) that were overrepresented in up- or down-regulated intoxication-related genes (Vadigepalli et al., 2003; Ho Sui et al., 2005). The following parameters were chosen for each gene list entered into oPOSSUM: 1) Taxonomic supergroup: vertebrate, 2) conservation level: 30%, 3) matrix match score: 85%, 4) upstream sequence length: 2000 base pairs, 5) downstream sequence length: 500 base pairs. Overrepresented binding sites were determined for each group of genes based on one-tailed Fisher exact probabilities (Fisher p-values) and Z-scores. Significance was set at Fisher $p < 0.05$ and/or Z score > 7. Determination of a non-random association of the TFBS used Fisher

p-values for a comparison of the proportion of target genes (co-expressed genes) containing a specific TFBS to the proportion of background genes (all genes entered into the ARIdb) for that TFBS. The number of times a TFBS appeared in a promoter region was irrelevant. The Z-score, however, compared the rate of every occurrence of a TFBS in target genes with the rate of each occurrence in background genes. It should be noted that some TFBS (*e.g.* HNF-3 beta and HFH-3) have similar consensus sequences.

Note that the gene lists used to query the oPOSSUM program were last annotated in July 2005.

FUNCTIONAL OVERREPRESENTATION ANALYSIS

Transcripts that passed the statistical thresholds were functionally annotated using WebGestalt (WEB-based Gene SeT AnaLysis Toolkit, http://genereg.ornl.gov/webgestalt) followed by overrepresentation analysis (Zhang et al., 2005). Functional annotations were based on Gene Ontology (GO) Consortium (http://www.geneontology.org) (Ashburner et al., 2000; Harris et al., 2004), BioCarta pathways (http://www.biocarta.com/genes/index.asp) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.jp/kegg) (Kanehisa, 1997; Ogata et al., 1999). WebGestalt used a hypergeometric test to determine whether the functional categories represented in the gene lists were significant or occurred by chance based on their prevalence in the background gene list (all genes entered into the ARIdb). Significance was set at $p < 0.05$ and a minimum of three genes within each functional group annotation. Significant functional categories that were non-specific (eg "cellular physiological process") were excluded from the results.

Note that the gene lists used to query WebGestalt were last annotated in July 2005.

CELLULAR CLASSIFICATION OF DIFFERENTIALLY EXPRESSED GENES

Differentially expressed genes were examined for their enrichment in different neuronal populations. The method used was similar to our previous report (Ponomarev et al., 2006) and is briefly described here. Significant genes were queried in the Mouse Neuronal Expression Database (MNED, http://mouse.bio.brandeis.edu) and their values in seven neuronal cell types (three excitatory (glutamate) and four inhibitory (GABA)) from cerebral cortex were retrieved (Sugino et al., 2006). Genes that had a difference in expression of 2-fold or higher across the seven neuronal populations in the MNED were clustered using K-means cluster analysis. A chi square analysis was applied to the genes enriched in the excitatory and inhibitory cell populations to determine whether a significant difference existed in the PLS pattern distribution between the two populations.

Note that the gene lists used to query the MNED were last annotated in July 2005.

Results and Discussion

STRAIN SPECIFIC TRANSCRIPTOME FINGERPRINTS RESULT FROM ALCOHOL EXPOSURE

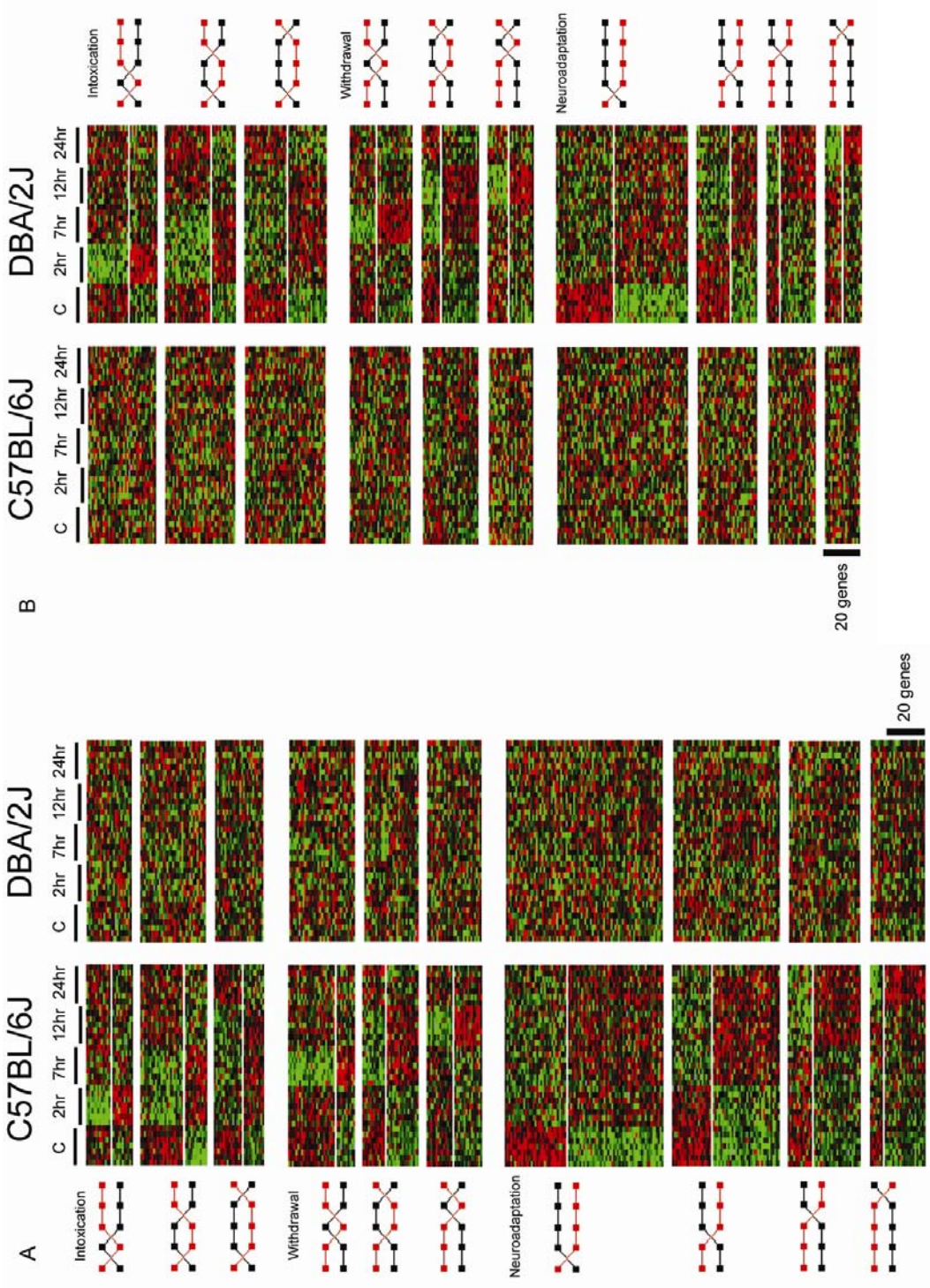
Transcriptional responses to alcohol, regardless of dose or the mode of exposure, encompass a range of time points, from initial exposure and intoxication to the time the alcohol clears the system to the short and long-term consequences that follow. Previous studies examining the effects of alcohol exposure on the rodent brain transcriptome have focused on a single time point following a specific alcohol treatment (Daniels and Buck, 2002; Rimondini et al., 2002; Saito et al., 2004; Treadwell and Singh, 2004; Kerns et al., 2005; Rulten et al., 2006). The goal of the present study was to identify PFC-specific temporal patterns of alcohol-related gene expression in a standard mouse model of acute dependence (Buck et al., 1997). Microarray analyses of 2, 7, 12 and 24 hour post-alcohol injection time points were completed, which reflect the biological processes related to intoxication, peak of withdrawal, recovery and neuroadaptation, respectively. At two hours post injection of 4 g/kg i.p. ethanol, B6 and D2 mice were visibly intoxicated, showing little movement but when active, a severely impaired gait. Seven hours is within the published peak of withdrawal for the acute dependence model with handling-induced convulsion (HIC) scores returning to baseline by 12 hours (Buck et al., 1997). None of the mice used were tested for withdrawal signs so that HIC-related confounds would not alter gene expression. In order to assay for longer-term or neuroadaptive expression changes, a 24 hour time point was included.

PLS analysis coupled with statistical tests for significance (*t*-test and *F*-test) were used to fit ten biologically relevant patterns of expression, with the nodes in each PLS pattern corresponding to the control, 2, 7, 12 and 24 hour time points, respectively

(Figure 4.1). Cluster analysis within each PLS class revealed genetically distinct alcohol expression responses, which are shown as a pseudo-color raster display in Figure 4.1. The genes significantly changed in B6 mice (386) are shown in the left panel of Figure 4.1A while those changed in D2 mice (354) are displayed in the right panel of Figure 4.1B. Gene identities are listed in Table A.2. The striking patterns of change indicate that genetically divergent transcriptome remodeling occurred continuously across the assayed time course. Categorization of the genes according to their unique transcriptional profile allowed genes that change only at the two hour time point (intoxication) to be distinguished from those that change at the two hour time point and remain changed throughout the time course (neuroadaptation). The PLS patterns allowed unique insight into candidate genes likely to play a role in intoxication, withdrawal, recovery and neuroadaptation. The significance of the findings was that considerable transcriptome remodeling occurs in the PFC following a single alcohol exposure.

Figure 4.1: Alcohol-related strain differences in gene expression as a function of time. A total of 386 (A; left panel) and 354 (B; right panel) genes were differentially expressed as a result of alcohol exposure in B6 and D2 mice respectively. K-means cluster analysis revealed the up- (green) and down- (red) regulation of differentially expressed genes within each temporal pattern of expression. Values of clustered genes are displayed in the opposite strand (A; right panel and B; left panel) to show that expression patterns are strain-specific. Rows represent genes and columns represent arrays, with each time point corresponding to seven arrays. Gene names are listed in Table A.2 and are current as of July 31, 2006.

Figure 4.1: Alcohol-related strain differences in gene expression as a function of time.

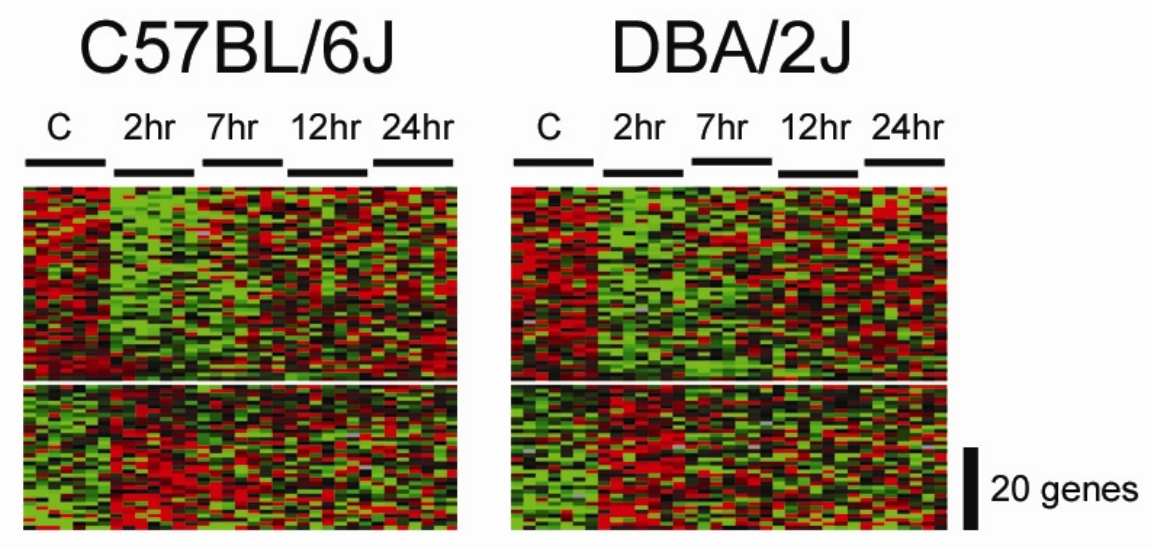


To determine whether the differentially expressed genes in one genotype followed the same temporal expression pattern in the other genotype, the values of the significant genes in the B6 and D2 strains were queried in the D2 and B6 strains (respectively), maintaining gene order. Interestingly, very few genes (~30) overlapped between the B6 and D2 mice at the stringent significance level chosen ($p < 0.005$). The right panel of Figure 4.1A and the left panel of Figure 4.1B show the gene values in the opposite or non-clustered strain. No expression patterns were apparent, indicating that the two mouse strains have unique alcohol-associated transcriptional fingerprints. The genes that contribute to the distinct patterns between B6 and D2 represent candidate genes that may play a role in the dependence-related phenotypic differences between the strains.

A genetic divergence clearly existed between the transcriptional consequences of an acute dose of alcohol in B6 and D2 mice. However, previous studies have shown that the strains contain genetic factors that contribute to both high and low responses for several alcohol-related phenotypes, suggesting that at least some genes should be expressed similarly between the strains in response to alcohol (Phillips et al., 1994; Buck et al., 1997; Metten et al., 1998; Crabbe et al., 2005). We searched for genes with similar expression in B6 and D2 PFC in response to alcohol intoxication (*t*-test p-value < 0.05 and fold change directionality agreement in both strains). A total of 84 alcohol-dependent, genotype-independent changes in gene expression were observed. Cluster analysis indicated that the genes were similarly expressed in both strains across all time points following alcohol exposure (Figure 4.2; see Table A.3 for the gene names). The patterns of B6 and D2 gene expression were consistent with previous behavioral data that suggests that neither strain can be simply or categorically labeled as sensitive or resistant to the effects of alcohol (Crabbe et al., 2005). The similarly expressed genes may

explain, in part, non-maximal phenotypic dissimilarity between the two inbred strains for acute alcohol withdrawal and other alcohol-related responses.

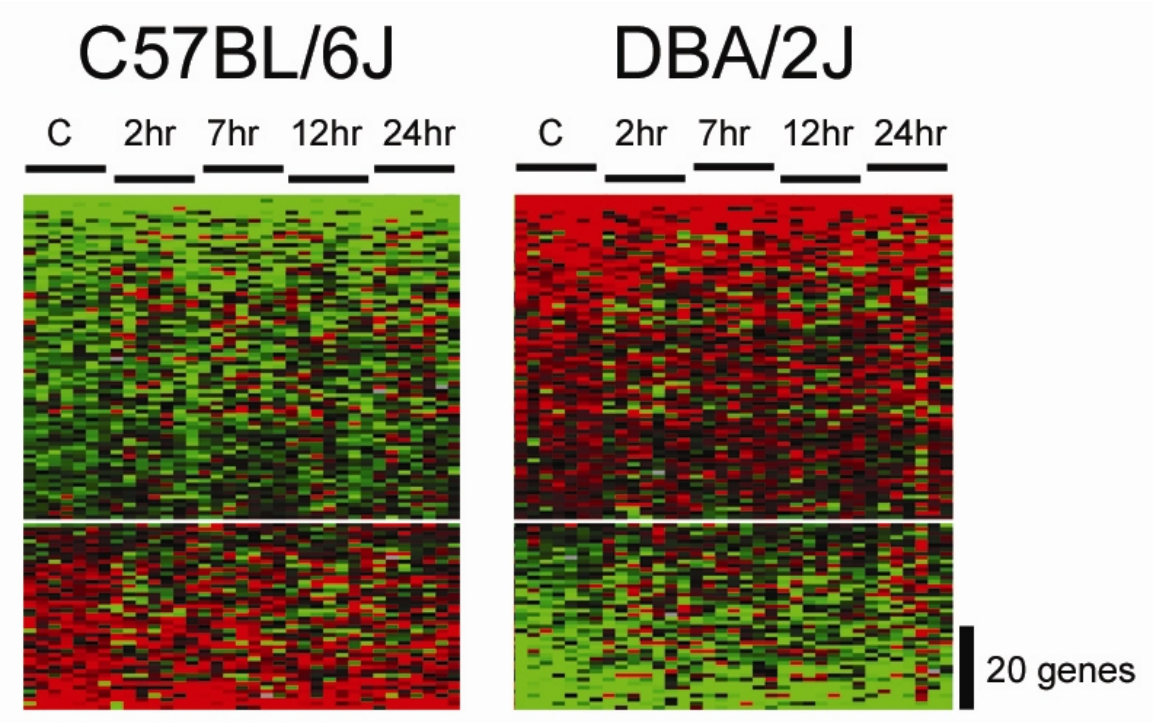
Figure 4.2: General expression responses following intoxication. K-means cluster analysis was applied to 84 genes differentially expressed in both B6 and D2 two hours post-alcohol exposure (*t*-test p-value < 0.05) with fold change directionality agreement. A pattern indicative of genotype-independent alcohol-related gene expression was revealed. Green and red boxes represent genes up- or down-regulated by alcohol compared to control, respectively. Rows represent genes and columns represent arrays, with each time point corresponding to seven arrays. Gene names are listed in Table A.3 and are current as of July 31, 2006.



Alcohol insensitive, genetically divergent expression can contribute to alcohol-related traits, e.g. as a genetic predisposition. Naïve strains and lines of mice have provided insight into the genetics of alcohol-related traits such as sensitivity and acute functional tolerance (Xu et al., 2001; Tabakoff et al., 2003). We compared B6 and D2 control mice to identify genetic differences in gene expression between the PFC of the two strains. One hundred twenty-six genes showed significant genetic divergence between B6 and D2 mice with no alcohol-associated expression response (Figure 4.3; see Table A.4 for the gene names). Approximately 60% of the expression differences were higher in B6 than D2 mice which might indicate the presence of allelic differences in general control factors.

A recent meta-analysis of nine mouse models of alcohol preference identified specific candidate genes for alcohol consumption (Mulligan et al., 2006). Thirty-one of the B6/D2 alcohol non-responsive genes identified in our study are consistent with these findings. The genetically divergent expression differences identified in our study may be predisposition genes for alcohol-related traits such as susceptibility to physical dependence; however, at least some of the genes identified likely play a role in other B6, D2 divergent phenotypes.

Figure 4.3: Genetic differences in gene expression between B6 and D2 mice. One hundred twenty-six genes showed significant genetic divergence between B6 and D2 mice with no alcohol-associated expression response (*t*-test p-value < 0.001). Green and red boxes represent genes more highly expressed in B6 or D2, respectively. Rows represent genes and columns represent arrays, with each time point corresponding to seven arrays. Gene names are listed in Table A.4 and are current as of July 31, 2006.



IN SILICO PROMOTER ANALYSIS IDENTIFIES A REGULATORY TFBS PATTERN FOR ALCOHOL INTOXICATION-ASSOCIATED GENE EXPRESSION.

Microarray hybridization assays the relative amounts of thousands of individual transcripts in a sample but gives no direct information on the mechanisms responsible for the level of transcript detected. Regulation of transcript concentration can take place at several levels; through binding of a complex of transcription factors (with either positive or negative activity), by chromatin remodeling, which can be mediated by covalent modification of DNA or histones, or through controlled stability or degradation of the message. Bioinformatic tools capable of analyzing DNA sequences for transcription factor binding sites (TFBS) now exist and are a valuable tool for understanding the molecular mechanisms underlying alcohol-associated phenotypes, including acute alcohol withdrawal.

Genes that show similar alcohol-related temporal expression patterns may be co-regulated by common transcription factors. The presence of specific transcriptional regulatory sequences in the 5' upstream region of a gene indicates the potential complement of transcription factors that may be regulating a particular gene. oPOSSUM was used to identify overrepresented transcription factor binding sites (TFBS) in genes expressed during intoxication. Genes expressed at two hours post alcohol exposure were divided into 4 groups: 1) B6 up-regulated genes, 2) B6 down-regulated genes, 3) D2 up-regulated genes, 4) D2 down-regulated genes, to separate out the effect of genotype and direction of regulation.

oPOSSUM analysis uncovered genotype-dependent and -independent TFBS usage for intoxication-related increases and decreases in gene expression in the PFC (Table 4.1). Alcohol-related gene expression appeared to be driven by unique sets of transcription factors in B6 mice; whereas in the D2 mouse, few specific TFBS were

identified. In addition, in B6 mice, distinct differences were apparent in the regulation of genes with increased or decreased expression in response to alcohol intoxication. The TFBS for Sp1 was present in over 50% of the genes up-regulated at two hours while c-Myb, Gfi and Yin-Yang were abundant in the down-regulated genes. Previous studies have implicated Sp1 in transcriptional control of alcohol mediated expression changes (Wilke et al., 2000; Rulten et al., 2006) while c-Myb, Gfi and Yin-Yang represent novel candidates for further study.

General, genotype-independent, alcohol-related up-regulation shared several control factors in common between the mouse strains. Increased gene expression in B6 and D2 following exposure to alcohol was mediated, in part, by CREB, c-REL, NF-kappaB and p65. CREB and NF-kappaB have been confirmed in other studies (Spitzer et al., 2002; Hassan et al., 2003; Gukovskaya et al., 2004; Nestler, 2005) supporting our *in silico* findings. Overall, our results indicated that genetic differences in transcriptional regulation following alcohol treatment exist. While alcohol-related increases in expression are less genotype-dependent, decreased expression in B6 mice may be influenced more by the activity of unique transcription factors.

It should be noted that each subcategory of genes (from B6 up, B6 down, D2 up or D2 down) highlights TFBSs that are overrepresented within that subcategory of genes only (compared to reference), as denoted by a colored "X" (Table 4.1). In other words, genes within each subcategory may contain TFBSs other than those listed with an "X," but they are just not overrepresented.

Table 4.1: Transcription factor binding site (TFBS) analysis of intoxication-associated gene expression. oPOSSUM was used to detect overrepresentation of TFBSs in the 2000 base pair region upstream of the start site of all genes differentially expressed two hours following alcohol treatment. Significance for overrepresented TFBSs in up- and down-regulation in B6 and D2 mice was set by Fisher's $p < 0.05$ and/or Z score > 7. The y-axis (blue cells) is transcription factors that bind to (overrepresented) TFBSs. The x-axis (pink cells) is the genes within each subcategory (B6 up, B6 down, D2 up or D2 down) that contain at least one of the overrepresented TFBSs identified within that subcategory. The vertical double line denotes general alcohol effects to the left and genetically divergent TFBS usage to the right.

Table 4.1: Transcription factor binding site (TFBS) analysis of intoxication-associated gene expression.

Table 1. Transcription factor binding site (TFBS) analysis of intoxication-associated gene expression.

			TFs													
		Genes	CREB	c-REL	NF-kappaB	p65	SP1	SRF	cEBP	c-MYB_1	Gfi	Spz1	Yin-Yang	HNF-1	HLF	TEF-1
C57BL/6J	up	*1810037C20Rik*		X	X	X										
		3110050N22Rik					X									
		Ccrn4l		X	X	X										
		Cyp51	X				X									
		Dscr1l2					X									
		Fos	X	X		X	X	X								
		Gadd45b	X	X	X	X										
		Heyl	X		X											
		Junb					X	X								
		Kcnip1					X									
		Npr1					X									
		Pbx3	X													
		Rps6ka4	X				X									
		Smarcad1	X	X	X	X										
		Sos1		X		X	X									
		Zap70		X		X										
C57BL/6J	down	*2500003M10Rik*											X			
		Abcf2									X		X			
		Bat8								X	X		X			
		BC018601								X			X			
		Calr											X			
		Capza2											X			
		Cct5									X		X			
		Ches1								X	X	X	X			
		Copeb								X	X	X				
		Dnaja1								X						
		Donson											X			
		E430002G05Rik											X			
		Eif4g2									X		X			
		Hrmt1l1											X			
		Hsp105									X		X			
		Idb3							X		X		X			
		Lgtn									X		X			
		Lmx1a								X	X					
		Opa1							X				X			
		Pcnxl3							X	X			X			
		Phf12							X	X			X			
		Pogz								X	X					
		Postn								X			X			
		Ptpla							X							
		Sfrs7								X		X	X			
		Xrn2							X	X			X			
DBA/2J	up	*1300002F13Rik*												X		
		2610528M18Rik	X													
		4632411B12Rik		X												
		4930438O05Rik	X													
		9430029K10Rik				X										
		Bteb1												X		
		Ccrn4l		X	X	X										
		Cyp51	X													
		D430028G21Rik			X	X										
		Idb4	X													
		Nfkbia		X	X	X										
		Rdx		X		X										
		Sncg	X													
		Sp3		X												
		Taf11	X													
		Tgfb1i4		X	X	X										
DBA/2J	down	*1110001J03Rik*	X													
		Amn													X	
		Bat8	X													
		BC030335	X												X	
		BC038311													X	
		D2Ertd750e														X
		Facl4	X													
		Pogz	X													X
		Tax1bp1	X													

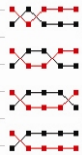
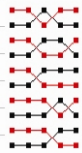
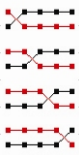
FUNCTIONAL ANALYSIS UNCOVERS GENETICALLY DIVERSE, ALCOHOL-RELATED OVERREPRESENTATION OF BIOLOGICAL PROCESSES.

The discovery of a strong genotype-dependent transcriptional response to a single high dose of alcohol led us to ask whether the response was genotype-dependent on a functional level as well. Overrepresentation analysis for function was performed within WebGestalt (Zhang et al., 2005) using the GO Tree, KEGG and Biocarta databases. Genes showing significant changes in expression (Table A.2) were grouped into three general classes: 1) intoxication, 2) withdrawal and 3) neuroadaptation according to best fit PLS patterns (Table 4.2). A dichotomy in functional themes between the B6 and D2 mice was seen. Several overrepresented categories in the B6 strain contained genes involved in signaling, while a different trend was seen in the D2 strain, where metabolism-related categories were more abundant. The genetic difference may represent a consequence of strain divergent alcohol-related compensatory mechanisms in the PFC.

Differences in functional categories were greater between genotypes than across classifications (intoxication, withdrawal, neuroadaptation) within genotype (Table 4.2). With the exception of RNA splicing and nucleic acid metabolism, no functional categories were consistent between the two strains. However several categories overlapped between classifications within the strains. While this may be due, in part, to common PLS patterns between classifications, it suggests that the cellular functional response to alcohol was more dependent on genotype than on the basic physiological effects elicited by alcohol's presence.

Table 4.2: Overrepresented functional categories in B6 and D2 mice. Genes from Table A.2 were divided into three groups (per genotype) to represent the physiologically relevant alcohol consequences of intoxication, withdrawal and neuroadaptation. Groups contained both up- and down-regulated genes and the PLS patterns comprising each group are shown on the left. The Gene Ontology, KEGG and BioCarta databases were queried within WebGestalt to identify overrepresented functional categories within each group of genes. Significance was set at $p < 0.05$ with a minimum requirement of three genes per functional group. The p-values for each overrepresented functional group are shown on the right.

Table 2. Overrepresented functional categories in B6 and D2 mice.

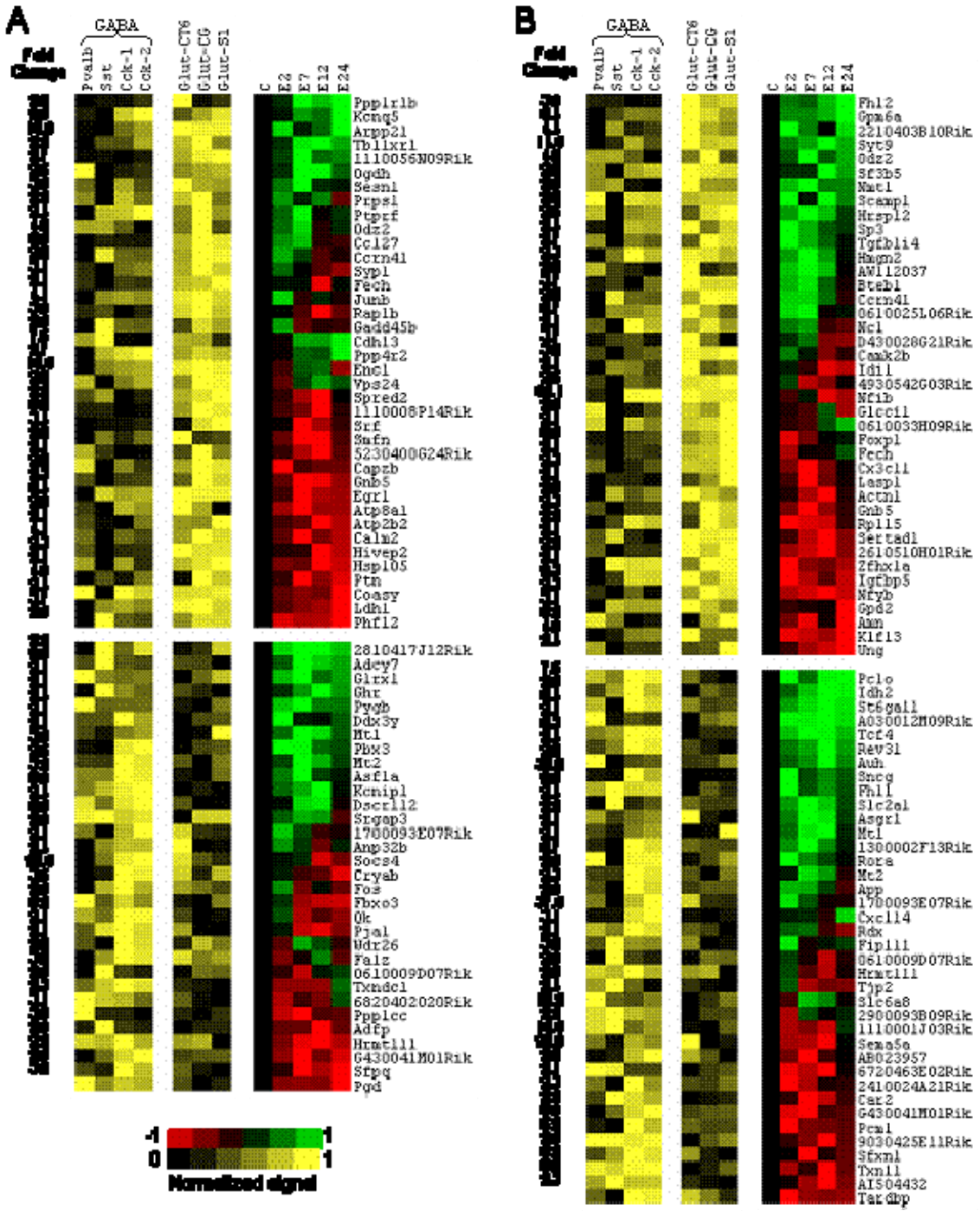
Classification	Strain	Overrepresented functional categories	p value	Strain	Overrepresented functional categories	p value
Intoxication	C57BL/6J	ERK1/ERK2 MAPK signaling pathway	0.003	DBA/2J	Myeloid cell differentiation	0.007
		T cell receptor signaling pathway	0.005		Exocytosis	0.015
		Insulin signaling pathway	0.008		Tube morphogenesis	0.015
		Nucleic acid metabolism	0.017		RNA metabolism	0.042
		Second-messenger-mediated signaling	0.021			
		RNA splicing	0.025			
		Actin cytoskeleton organization and biogenesis	0.030			
		Gap junction	0.030			
		Translation	0.033			
		MAPK signaling pathway	0.038			
		Nucleotide biosynthesis	0.042			
		Steroid metabolism	0.048			
Withdrawal	C57BL/6J	Growth hormone signaling pathway	0.002	DBA/2J	TCA cycle	0.002
		Jak-STAT signaling pathway	0.003		Coenzyme metabolism	0.007
		Sensory perception of sound	0.006		Cell proliferation	0.018
		Intracellular signaling cascade	0.011		Sodium ion transport	0.019
		Lipid transport	0.019		RNA splicing	0.023
		MAPK signaling pathway	0.029		Angiogenesis	0.027
		Tissue development	0.049		Nucleic acid metabolism	0.039
Neuroadaptation	C57BL/6J	RNA splicing	0.016	DBA/2J	TCA cycle	0.002
		Intracellular signaling cascade	0.046		Coenzyme metabolism	0.005
		Nucleic acid metabolism	0.047		Nucleic acid metabolism	0.014
		Tissue development	0.048		DNA repair	0.024
					Main pathways of carbohydrate metabolism	0.028
					Cytokine-cytokine receptor interaction	0.038
					Vesicle-mediated transport	0.042
					Regulation of transcription	0.044
					Tight junction	0.049

ALCOHOL-RELATED GENES HAVE UNIQUE CELLULAR PROFILES IN EXCITATORY AND INHIBITORY NEURONS.

To determine whether the alcohol-related genes in this study could be categorized by specific neuronal subtypes, we examined their expression across seven neuronal populations in the cerebral cortex: three excitatory (glutamate) and four inhibitory (GABA). We searched the Mouse Neuronal Expression Database (MNED) (Sugino et al., 2006) for the presence of differential neuronal expression of our alcohol-related genes (from Table A.2). A total of 148 alcohol-related genes (70 in B6 and 78 in D2) were found to have an average expression difference of 2-fold or higher across the seven neuronal subtypes (Figure 4.4). A K-means cluster analysis of the expression values from the MNED clustered the alcohol-related genes according to their enrichment in excitatory or inhibitory neuronal populations, suggesting that some of the genes identified as differentially expressed in the PFC have unique cellular profiles within that brain region (Figure 4.4). It should be noted however, that the expression of many of the alcohol-related genes was cell-type specific and was not uniform across all three excitatory or all four inhibitory cell populations—a result that was not unexpected based on the findings of (Sugino et al., 2006). A pseudo-color rastergram of the differentially expressed alcohol-related genes shows an approximate 1:1 ratio of up- to down-regulation in both B6 and D2 within each neuronal population, suggesting that no bias existed towards directionality of gene expression within excitatory or inhibitory neuronal populations (Figure 4.4).

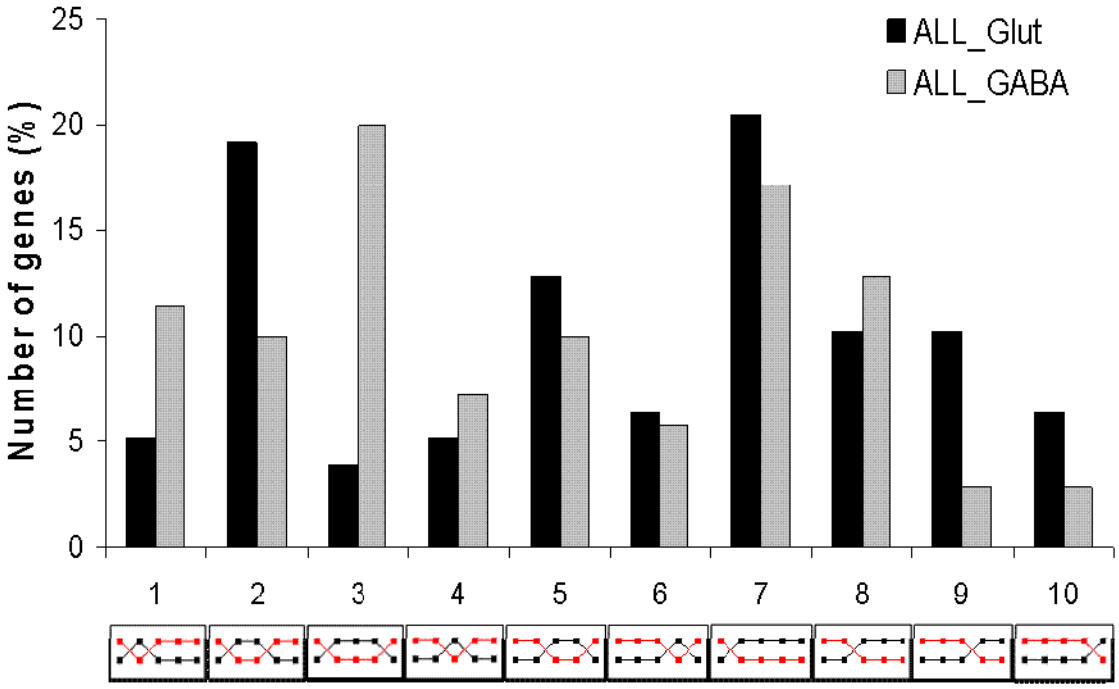
Figure 4.4: Cellular classification of differentially expressed genes. Alcohol-related genes (green and red panels) clustered according to their enrichment in different neuronal populations (black and yellow panels). Expression of alcohol-sensitive genes was examined across seven neuronal populations in the cortex: three excitatory (glutamate (Glut)) and four inhibitory (GABA). Gene symbols of alcohol-sensitive genes from Table A.2 were matched to data from the Mouse Neuronal Expression Database (MNED) (Sugino et al., 2006), and genes present in both datasets and that also showed substantial heterogeneity in expression among the seven neuronal populations (difference of 2-fold and higher) were selected. Our search revealed 70 genes regulated in B6 (A) and 78 genes regulated in D2 (B). K-means clustering clearly distinguished expression patterns characteristic of glutamate (upper yellow panels in both A and B) and GABA (lower yellow panels in both A and B) neurons. Green and red panels show time course data from the present study (C=control, E=ethanol at different time points). The expression values of alcohol-sensitive genes were normalized to show an up- (green) or down- (red) regulation compared to control. Fold change refers to the difference between lowest (black) and highest (bright yellow) expressing neurons.

Figure 4.4: Cellular classification of differentially expressed genes.



The expression patterns of the genes enriched in the glutamate and GABA neuronal populations represented all 10 PLS patterns, to varying degrees (Figure 4.5). A difference in the distribution of PLS patterns existed between the glutamate and GABA neuronal populations ($\chi^2(9) = 17.2$; $p < 0.05$). This difference was manifested in at least two different ways, as seen in Figure 4.5: (1) A large difference in the number of genes expressed in excitatory versus inhibitory neurons was evident for genes with expression patterns matching PLS patterns 1, 2, 3 and 9 and (2) PLS patterns of intoxication and early withdrawal (1-4) have a greater percentage of genes expressed in inhibitory neurons while patterns of late withdrawal and neuroadaptation (5-10) have a greater percentage expressed in excitatory neurons. Our findings suggest that excitatory and inhibitory neurons in the PFC require different mechanisms for molecular and cellular adaptation to alcohol and that the cellular temporal adaptation to the presence of alcohol is cell-type specific.

Figure 4.5: Frequency distribution of PLS patterns (x axis) of genes shown in Figure 4.4. Genes enriched in either glutamate (ALL_Glut) or GABA (ALL_GABA) neuronal populations are shown according to their PLS patterns from Table A.2 (data collapsed across strains). A comparison of the PLS pattern distribution between glutamate and GABA populations revealed differences ($\chi^2(9) = 17.2$; $p < 0.05$).



Conclusions

Physiological and behavioral responses to alcohol in both humans and mice are complex. A central hypothesis of alcohol dependence focuses on the idea that as the brain adapts to the presence of alcohol (during a binge episode or over the course of chronic exposure), changes in signal transduction, synaptic activity and gene transcription elicit downstream effects that rewire the molecular and anatomical neurocircuitry of the brain. Eventually, neuroadaptation to the alcohol alters brain physiology to the point that normal brain function becomes physically impossible in the absence of alcohol. Using microarray analysis of a common mouse model of acute alcohol dependence, we show that a large dose of alcohol results in a complex molecular response detected as a temporal pattern of transcriptome remodeling in the PFC. The analysis detected over 700 differentially expressed genes that showed genetically divergent alcohol responsiveness in PFC across the stages of intoxication, withdrawal, recovery and neuroadaptation. In addition, 84 genes showed strain-independent expression changes throughout the 24 hour time course. The candidate genes fit 10 unique PLS temporal patterns giving, for the first time, specific information about alcohol-mediated transcriptional neuroadaptation on a genomic scale.

As expected, genetic, non-alcohol responsive differences in brain (PFC) gene expression between the two strains were detected, consistent with previous comprehensive studies using the B6, D2 mouse model (Chesler et al., 2003; Kerns et al., 2005). Gene expression studies using alcohol-naïve animals are an important contribution to the growing list of candidate genes for alcohol-related traits. A list of 3800 candidate genes for predisposition to alcohol consumption was achieved using nine

naïve mouse models that diverge in alcohol preference, demonstrating the use of naïve strains to identify potential predisposition genes for alcohol-related traits (Mulligan et al., 2006).

In silico promoter and functional analyses demonstrated that the genetic divergence in the molecular consequences of an acute dose of alcohol extend beyond the level of gene transcription. Genotype- and direction-dependent transcriptional control was evident during intoxication, as shown by the presence of unique TFBSs upstream of the genes with increased and decreased expression in the two strains. As well, very little overlap existed between overrepresented functional categories in the B6 and D2 mice, with a trend towards signaling in B6 and metabolism in D2, suggesting that the strains diverge in their alcohol-related compensatory mechanisms in the PFC. Several functional categories identified in the current study were similar to those identified in previous studies using rodent models of alcohol dependence; including metabolism, signal transduction, development and transcription/translation (Daniels and Buck, 2002; Saito et al., 2002; Treadwell and Singh, 2004; Kerns et al., 2005). Of note is the MAPK pathway; identified in this study as overrepresented during intoxication and withdrawal in the B6 mouse and previously implicated in hippocampal gene expression changes following chronic alcohol exposure in B6 mice (Daniels and Buck, 2002) as well as overrepresented in several lines and strains of naïve alcohol-preferring mice (Mulligan et al., 2006). Further *in silico* analysis provided likely cellular identities within the PFC for genes that are expressed as a result of acute alcohol exposure, giving insight into the next level of specificity—beyond a single brain region, and providing another step towards understanding the underlying mechanisms of alcohol dependence.

In summary, microarray analysis of the PFC transcriptome in B6 and D2 mice revealed a striking genotype- and time-dependent remodeling by alcohol intoxication and

its subsequent physiological consequences. Nearly 1000 differentially expressed genes were narrowed by functional analysis to highlight potential candidate genes for further study. It is important to recognize; however, that these are candidate genes only and that their expression at specific time points does not necessarily indicate that they play a direct role in mediating the corresponding temporal physiological response(s) to alcohol. Further studies are warranted to test whether the transcriptional responses are a direct effect of intoxication or withdrawal, for example, or whether they are downstream effects of signal transduction or biochemical pathways affected at a previous time point. In addition, the current bioinformatics tools for functional characterization of genes are incomplete; therefore, results obtained from them can only provide partial insight into the potentially important functional groups and TFBS usage related to acute alcohol dependence.

The PFC is an important brain region in mediating the rewarding effects of alcohol and is a region susceptible to damage in chronic alcoholics. We have shown that the PFC undergoes a complex transcriptome remodeling in response to alcohol. New hypotheses about the role the genes identified in our study may play in mediating alcohol consequences, such as tolerance and dependence, can now be tested.

CHAPTER 5: DISCUSSION

Summary

Alcohol dependence is a complex polygenic disease influenced by both environmental factors and genetic makeup. Mouse models of alcohol-related traits are commonly used to identify the genetic factors that contribute to specific physiological and behavioral consequences of alcohol exposure. The B6 and D2 inbred strains of mice, which diverge in several alcohol-related traits, were used in two complementary gene expression studies to identify potential candidate genes for alcohol dependence. The mice were given a single, high dose of alcohol and brain gene expression changes were monitored over a time course using mRNA Differential Display (DD) and cDNA microarray analysis. The DD study was performed first in order to generate additional clones for the cDNA microarrays. 150 new unique brain targets were added to the cDNA clone sets as a result of the DD screen. Over 700 genetically divergent, alcohol-sensitive gene expression differences between the prefrontal cortices (PFC) of the B6 and D2 mice were identified in the cDNA microarray analysis. A combination of statistical tests (*t*-test, *F*-test and PLS analysis) and K-means cluster analysis revealed striking strain-specific temporal patterns of gene expression induced by alcohol. In addition, genotype-independent changes in gene expression were identified as a result of intoxication in both strains, as well as baseline genetic differences between the PFCs of B6 and D2. *In silico* analysis of the genotype-dependent differentially expressed genes demonstrated genetically divergent trends in function, transcription factor binding site (TFBS) usage and cell-type classification. The results indicate that the brains of B6 and D2 mice have very different cellular and molecular responses to acute alcohol exposure.

mRNA Differential Display as a complement to cDNA microarray analysis

One of the major advantages to cDNA microarray analysis over more traditional gene expression detection techniques is that tens of thousands of genes can be screened at once for changes in expression. However one of the drawbacks to using microarrays is that the results are dependent on the clones printed on the array. A less biased method of gene expression detection is mRNA Differential Display (DD). In DD, primer pairs are used to amplify every (theoretically) expressed gene at any given time in an organism. Therefore, DD was used as a complementary technique to the microarrays in order to increase the number of brain-specific, genetic and/or alcohol-sensitive genes on the arrays. About 150 DD clones were combined with the existing microarray clones to become additional targets on the arrays.

Our experimental design of combining DD and cDNA microarrays to increase the number of brain-specific, genetic and/or alcohol-sensitive genes did not prove to be as beneficial as we had anticipated for several reasons. First, DD is labor and time intensive. Each sample (mouse brain for example) must be PCR amplified 240 times and each of those reactions must be run on a sequencing gel to identify differentially expressed genes. The gels must be examined, the bands excised, subcloned and sequenced, and the sequencing results must be queried using a database of known genes and ESTs. Depending on the size of the experiment, this process can take several months to complete. Second, although 150 new clones were added to the arrays as a direct result of the DD screen, only about 10% of the named genes were novel to the array clone sets. Third, the purpose of the DD screen was to generate novel targets that would contribute to and enhance the microarray results by increasing the number of brain-specific, genetic

and/or alcohol-sensitive genes on the arrays, yet very few of the DD clones were differentially expressed in the microarray experiment. For example, of the ~30 genes that were differentially expressed between the whole brains of the saline-treated (control) B6 and D2 mice in the DD experiment, only three were found to be differentially expressed in the microarray experiment (*t*-test p-value of 0.001 between B6 and D2 saline treated animals). Furthermore, none of the 700 genetically divergent, alcohol-sensitive genes identified in the microarray experiment were DD clones. The following paragraph provides several possible explanations for the disparity between the two experiments.

One of the major differences between the DD and microarray studies was the tissue type used in each. Whole brain was examined in the DD study while PFC was examined in the array study. It is likely that only a small fraction of the differentially expressed genes from the DD study were expressed in the PFC, given that it is such a small brain region. Therefore, it may be invalid to assume that the majority of the DD clones would be detected in the array experiment. Brain gene expression was examined at two time points (2 hours and 6 hours) following alcohol administration in the DD experiment while in the array experiment, four time points of 2, 7, 12 and 24 hours were studied. The 700 genes identified as differentially expressed in the array experiment included genes that increased or decreased in expression at 12 hours and/or 24 hours only. Since these two time points were not included in the DD experiment, the total number of differentially expressed genes across time is not directly comparable between the two studies. Finally, methodologically speaking, DD is a more sensitive technique than arrays in terms of detecting small quantities of RNA while it is a less sensitive technique in terms of the criteria required to deem a gene differentially expressed. To elaborate, as PCR amplification is the basis of DD, a very small amount of RNA can be present in the cell and it will still be detected on the sequencing gel using radioactivity.

On the other hand, for microarrays, an equivalent small amount of RNA (cDNA) must find its complement on a microscope slide containing tens of thousands of other cDNA species, bind well to that complementary strand, emit enough fluorescence to be detected by the scanners and be given a high enough intensity value (during the scanning process) such that it is interpretable by the array software. In the DD experiment, an n of 2 was sufficient to determine whether a gene was differentially expressed or not. If, in those two samples, the amount of radioactivity was different between saline treated animals and alcohol treated animals, for example, then the gene passed the criteria for differential expression. However, in the array experiment, that same gene must have been present (good spot morphology and high enough intensity value) on 6 out of 7 n arrays for both the saline and alcohol treated samples, had low enough variance to pass strict p-value cutoffs when compared to each other and had to have been appointed a pattern of expression through PLS analysis that matched one of ten pre-determined biologically relevant temporal expression patterns.

A comparison of studies

One of the goals of the study was to identify PFC-specific temporal patterns of alcohol-related gene expression in a standard mouse model of acute dependence (Buck et al., 1997) using cDNA microarrays. PLS analysis coupled with statistical tests for significance (*t*-test and *F*-test) were used to fit 10 biologically relevant patterns of expression encompassing the first 24 hours after alcohol exposure. About 825 genes were differentially expressed in one or both of the strains (B6 and D2) over the 24 hour time course, which included a 2, 7, 12 and 24 hour time point following alcohol exposure.

Several previous studies have examined the effects of alcohol exposure on the rodent brain transcriptome (Daniels and Buck, 2002; Rimondini et al., 2002; Saito et al., 2004; Treadwell and Singh, 2004; Kerns et al., 2005; Rulten et al., 2006). Consistent with our findings, about 1-3% of the genes surveyed (i.e., genes present on the microarray) in these studies were differentially expressed as a result of alcohol exposure. However, a comparison of the differentially expressed genes from the present study with those found in three previous studies (Daniels and Buck, 2002; Rimondini et al., 2002; Kerns et al., 2005) revealed very few overlapping genes. The three studies selected for comparison are similar to the present study in genotype (Daniels and Buck, 2002; Kerns et al., 2005), dose of alcohol (Daniels and Buck, 2002), time point after alcohol exposure (Daniels and Buck, 2002) or brain region (Rimondini et al., 2002; Kerns et al., 2005). A total of 36 genes were common to at least two of the four studies.

There are a number of possible reasons why so few overlapping genes were found. The first is that the arrays were done in different laboratories, using different array platforms (cDNA and oligonucleotide) with different criteria for determination of gene expression "significance." Several studies have examined the agreement of microarray results across lab and across platform and found the following: 1) The agreement among platforms (short oligo, long oligo, cDNA) was greater within the same lab than it was between labs (Bammler et al., 2005; Wang et al., 2005), 2) cDNA arrays showed poor concordance with Affymetrix and spotted oligonucleotide arrays (Woo et al., 2004) and 3) Short oligo, long oligo and cDNA arrays produced similar results only if the criterion for differential gene expression was the direction, not the magnitude, of change (Petersen et al., 2005). In addition, in the Petersen et al. study, 6430 genes were common among the three platforms…a condition that was not true for the four alcohol-related studies compared here. Therefore, the second reason why such disagreement

exists among the four studies being compared is the difference in gene composition of the arrays. Thirdly, the entity used to compare array results was the gene symbol. Importantly, not all genes have gene symbols and some genes can be represented by more than one symbol. In addition, with time, more unknown genes were assigned symbols and the studies compared here were done several years apart.

Biological reasons exist as well to explain the discrepancy in array results. Although commonalities exist among the compared studies, no single study mimics the genotype, dose of alcohol, brain region and time points surveyed post-alcohol exposure of the present study. One of the biggest differences that sets the current study apart from all previous ones is the utilization of a time course to examine gene expression changes after alcohol exposure. All of the genes deemed significantly differentially expressed in the present study had to match one of 10 predetermined temporal expression profiles. In all previous studies, genes labeled as differentially expressed by alcohol had to be altered at a single time point only. Such a difference in criteria for differential expression could create very different results.

Computational tools as a means to further characterize differentially expressed genes

The present study demonstrated that a single, high injection of alcohol in B6 and D2 mice results in a striking strain-specific pattern of transcriptome remodeling in the prefrontal cortex. It has been hypothesized that genes with similar (alcohol-related) temporal expression patterns may be co-regulated by common transcription factors (Altman and Raychaudhuri, 2001; Schulze and Downward, 2001). This hypothesis was tested by searching for overrepresented transcription factor binding sites (TFBS) in the upstream region of genes expressed in B6 or D2 during intoxication. oPOSSUM analysis

uncovered genotype-dependent and -independent TFBS usage for intoxication-related increases and decreases in gene expression in the PFC, suggesting that some transcription factors were specifically regulating gene expression in only one of the genotypes in response to the alcohol exposure while other transcription factors were regulating gene expression based on alcohol exposure alone. Considering the minimal overlap in genes expressed during intoxication between B6 and D2 (and therefore in the gene lists uploaded to the oPOSSUM program), the existence of strain-specific transcriptional regulation is not surprising. However, common overrepresented TFBSs were found in genes up-regulated by alcohol in both B6 and D2, suggesting that even though very few (if any) actual genes were common in the lists of up-regulated genes in B6 and D2 during intoxication, the transcription factors that regulate the expression of those up-regulated genes may be the same. The disparity between the strain-specificity of the up-regulated genes and the non-strain-specificity of the overrepresented TFBSs suggests that some transcription factors are activated in response to alcohol, but then activate different individual genes. Of note is that several of the genotype-independent TFBSs are sequence motifs for common transcription factors. NF-kappaB activates gene expression in response to inflammation while CREB activates neuronal gene expression in response to growth factors, hormones and neurotransmitters. With such general requirements for transcriptional activation, it is plausible if not highly likely that alcohol could activate these two transcription factors as a general response mechanism to alcohol, in a genotype-independent manner.

Although TFBS analysis of the intoxication-related genes from the present study has given insight into the potential complement of transcription factors that may be regulating those genes, the methodology is still in its infancy and therefore the results should be interpreted accordingly. For example, in the present study, only about 30% of

the significant genes uploaded into the oPOSSUM program were recognized and used in the overrepresentation analysis. In addition, when a less stringent set of criteria were used in a different promoter analysis program (MATCH) to search for TFBSs, most of the genes contained hundreds if not thousands of potential TFBSs, indicating that setting limitations on programs to ensure the output is of manageable size and interpretation, some information might be lost.

The use of *in silico* analyses to extend or support microarray data is becoming more common (Mulligan et al., 2006; Ponomarev et al., 2006). However, the databases currently available for promoter analysis (Vadigepalli et al., 2003; Ho Sui et al., 2005) and functional annotation of genes are incomplete. For example, when querying the GO, KEGG and Biocarta databases within WebGestalt to functionally annotate the significant genes from this study, only 85% of the genes had gene symbols assigned to them. Of the genes with symbols, only 85% of them were recognized by the WebGestalt program. Taken together, only 72% of the significant genes were used to assign overrepresented functional categories to the entire list. In addition, the fact that a gene has a gene symbol assigned to it does not guarantee that its function is known, adding another layer of uncertainty to the final functionally annotated gene list. Therefore, caution should be exercised when drawing conclusions about the nature of the functional annotations assigned to lists of significant genes. Although they can provide insight into the functional trends among significant genes, they cannot provide solid conclusions.

Narrowing the focus: from whole brain to brain region

The brain is an incredibly heterogeneous and complex tissue. Billions of neurons and glial cells, encompassing a wide range of cell types makes gene expression studies of

the brain particularly challenging (Nisenbaum, 2002). It is not uncommon for genes of functional significance to be expressed at low levels and in very specific subpopulations of cells (Nisenbaum, 2002; Mirnics and Pevsner, 2004). As a result, significant changes in gene expression may be difficult to detect in whole brain. By focusing on a single brain region, small but significant changes in gene expression may be more easily detectable due to the fact that they are not being masked by an abundance of cell types and/or more highly expressed transcripts (Mirnics and Pevsner, 2004).

A second advantage to focusing brain gene expression studies on a particular region is that individual brain regions often differ in their anatomical, physical, cellular, electrophysiological and functional characteristics. Assessing the transcriptional profiles of single brain regions offers the potential to link the gene expression changes to one or more of these characteristics (Letwin et al., 2006). For example, it has been established that the prefrontal cortex is selectively damaged by long term alcohol exposure (Kril et al., 1997; Kril and Halliday, 1999; Obernier et al., 2002a; Obernier et al., 2002b). Monitoring the transcriptional changes over time as the physical damage is occurring may give insight into the mechanisms underlying the damage.

The results of the present study demonstrate that a strong transcriptional response to alcohol exists in the prefrontal cortices of B6 and D2 mice. Furthermore, the genes identified as differentially expressed in the whole brains of these mice using DD differ from the genes identified as differentially expressed in the PFCs using microarrays. In addition, a separate experiment performed in the Bergeson lab that was designed to test the effects of chronic alcohol exposure on D2 mice in a brain-region dependent manner found very little overlap among the genes expressed in the six brain regions tested (olfactory bulb, striatum, hippocampus, cortex, cerebellum and midbrain) (Bergeson, lab, unpublished data), indicating that each region had its own unique expression profile in

response to chronic alcohol exposure. The results of the acute and chronic studies suggest that more insight may be gained into the molecular mechanisms underlying a disease such as alcohol dependence by studying the expression profiles of individual brain regions as opposed to the whole brain.

Limitations of study

As discussed previously, no single animal model can completely recapitulate all the symptoms of alcohol dependence. Inbred strains and selected lines of mice have been successfully used to model specific clinically relevant traits. For example, a single injection of a large dose of alcohol (4 g/kg) has been shown to cause withdrawal (a sign of physical dependence) in D2, and to a lesser extent, B6 mice (Buck et al., 1997). The present study utilized this model to identify gene expression changes occurring concomitantly with intoxication, withdrawal and neuroadaptation over a 24 hour period after a 4 g/kg injection of alcohol. While the acute physical dependence model can be used to identify the transcriptional response in brain to such a large dose of alcohol, it cannot give insight into the mechanisms underlying the choice to drink the equivalent of such a large dose. Voluntary oral consumption of alcohol in mice rarely produces substantial blood ethanol concentrations (BECs), nor do the mice exhibit the loss of control often seen in human alcoholics (Finn et al., 2005; Rhodes et al., 2005). Therefore, the limitations of this model include the degree of voluntary consumption that is present in human alcoholics. Although progress has been made in creating drinking paradigms (Finn et al., 2005; Rhodes et al., 2005) and new mouse models (Blednov et al., 2005) to increase voluntary alcohol consumption, no drinking model yet exists that can mimic the physiological effects of a single 4 g/kg dose.

The gene expression changes occurring across time in this study parallel important molecular and physiological changes that result from acute alcohol exposure, including intoxication and withdrawal. It is important to realize, however, that the genes identified in this study are only candidate genes, and that further experiments are necessary to determine whether the genes directly contribute to intoxication, withdrawal or neuroadaptation after a single, high dose of alcohol. To date, no connection has been made between prefrontal cortex function and acute alcohol withdrawal. Therefore, the gene expression changes occurring at each time point in the time course (2, 7, 12 and 24 hours) following alcohol exposure may or may not be directly related to the concurrent alcohol response. It is possible that the results of this study revealed a previously unknown connection between the prefrontal cortex and acute alcohol dependence; however, that hypothesis would have to be specifically tested. What can be concluded is that acute alcohol exposure induces a strong transcriptional response in the PFC that is strain-specific and time-dependent.

Future directions

Microarray analysis has the potential to provide hundreds if not thousands of genes for further study. In our study, differentially expressed genes were narrowed by classification into functional categories, by TFBS usage and by cellular profiles. Several other types of bioinformatics tools are also available, including programs that assign genes to pathways (PathwayArchitect™) and online tools such as mouse brain atlases that allow the user to search for specific genes among hundreds of digital *in situ* hybridization images (http://www.brainatlas.org/ and (Carson et al., 2005)). These tools provide a mechanism by which to build gene networks, highlight biochemical pathways

and map cellular locations of genes—all of which help to give a more three-dimensional view of gene expression changes and possibly better insight into the molecular consequences of alcohol exposure.

Previous studies have demonstrated a clear difference in gene expression profiles among several brain regions at the basal level (Letwin et al., 2006) and as a result of chronic alcohol exposure (Bergeson lab, unpublished data). It is likely that acute alcohol exposure also causes a unique, brain region-specific transcriptional response as well. Six other brain regions—olfactory bulb, posterior cortex, striatum, hippocampus, midbrain and cerebellum—were dissected along with the prefrontal cortex after a single, high injection of alcohol in the B6 and D2 mice. Microarray analyses are currently underway to identify the alcohol-induced temporal expression patterns in these regions. Once completed, a comparative analysis could be done to identify brain-region specific gene expression patterns following acute alcohol exposure.

Conclusions

The results of this study demonstrated that a single, high dose of alcohol was capable of causing considerable transcriptome remodeling in the prefrontal cortices of B6 and D2 mice. Furthermore, over 700 transcriptional changes identified were genotype- and time-dependent, demonstrating the dichotomy in acute alcohol-induced brain gene expression between the B6 and D2 strains.

This study was the first to monitor gene expression over a time course following acute alcohol exposure in mice. Transcriptional responses to alcohol, regardless of dose or the mode of exposure, encompass a range of time points, from initial exposure and intoxication to the time the alcohol clears the system to the short and long-term

consequences that follow. By focusing solely on gene expression changes during intoxication or withdrawal, as previous studies have, other important events that may explain the transition from intoxication to withdrawal or from withdrawal to neuroadaptation may be missed. By identifying gene expression changes as a function of time, a better understanding of the progression of events that may ultimately protect B6 from severe physical dependence while allowing D2 to be susceptible to it may be achieved.

Appendix

Table A.1: Differential display clones printed on the cDNA microarrays. The clones isolated from the DD screen that were added as targets to the microarrays are listed. About 150 unique known genes, ESTs and unknown sequences were isolated. The column labeled "regulation" lists the expression profile of each clone (G = genetically divergent expression between B6 and D2 control mice, E = ethanol-sensitive, non-genetically divergent expression and G/E = genetically divergent, ethanol-sensitive expression). The column labeled "oligo" lists the number (1-80) of the forward primer and the letter (A,G,C) of the anchor on the reverse primer. EST = expressed sequence tag. E/S = E score from BLAST results.

Table A.1: Differential display clones printed on the cDNA microarrays.

#	SEQ ID	REGULATION	OLIGO	GENE NAME	ACCESSION
1	109603	G/E	38A	N-deacetylase/ N-Sulfotransferase 4	NM_022565
2	109604	G/E	38A	N-deacetylase/ N-Sulfotransferase 4	NM_022565
3	109611		42A	EST	AK009203
4	109612		42A	EST	AK009203
5	109613		39A	bad E/S	
6	109614		39A	NADH dehydrogenase subunit 6	NP_008119
7	109615	G/E	48A	EST	AW493321
8	109618	G/E	48A	EST	AW493321
9	109619	G/E	39A	myotonic dystrophy kinase-related Cdc42-binding kinase	AF021935
10	109621	G/E	39A	myotonic dystrophy kinase-related Cdc42-binding kinase	AF021935
11	109622	G/E	39A	myotonic dystrophy kinase-related Cdc42-binding kinase	AF021935
12	109623	G/E	39A	EST	BG078468
13	109624	G/E	39A		
14	110524	G/E	17G	cyclic GMP-stimulated phosphodiesterase mRNA	BC006845
15	110525	G/E	17G		
16	110529	E	54G	EST	AA049037
17	110530	E	54G	EST	AK011541
				calcyphosine gene	Y09919
18	113212	E	54G	EST	AK011541
				calcyphosine gene	Y09919
19	110531	E	54G	EST	AK008255
20	110533	E	54G	EST	AK008255
21	110535	G	52G	EST	AK008255
22	110536	G	52G	EST	AK008255
23	110537	G	52G	26S proteasome subunit p97 mRNA	D78151
				tumor necrosis factor type 1 receptor assoc'd protein mRNA	U12596
24	110538	G	52G	26S proteasome subunit p97 mRNA	D78151
				tumor necrosis factor type 1 receptor assoc'd protein mRNA	U12596
25	110539	G	21G	EST	AK011102
				beta-globin complex DNA for ..... genes and pseudogenes	X14061
				G protein Gi2 alpha	S71213
26	110540	G	21G	EST	AK011102
				beta-globin complex DNA for ..... genes and pseudogenes	X14061
				G protein Gi2 alpha	S71213
27	110542	E	38A	2-oxoglutarate dehydrogenase gene	D32064
28	110544	E	38A	2-oxoglutarate dehydrogenase gene	D32064
29	110545	E	38A	EST	BB283632
30	110546	E	38A	EST	
31	110547	G	34A	PFTAIRE protein kinase 1 mRNA	NM_012395
				KIAA0834 protein mRNA	AB020641
32	110548	G	34A	PFTAIRE protein kinase 1 mRNA	NM_012395
				KIAA0834 protein mRNA	AB020641
33	110549	G/E	48A	EST	BE285624
34	110550	G/E	48A	EST	BE285624
35	110554	E	41A	endogenous retroviral-like element MuERVC-C105	AF049340
36	110555	E	41A	endogenous retroviral-like element MuERVC-C105	AF049340
37	110556	G	41A	EST	AW493451
38	110557	G	41A	EST	AW493451
39	110558	G	41A	junction-mediating and regulatory protein mRNA	NM_021310

Table A.1: Differential display clones printed on the cDNA microarrays.

#	SEQ ID	REGULATION	OLIGO	GENE NAME	ACCESSION
				p300 transcriptional cofactor JMY mRNA	AF201390
40	110562	E	42A	EST	AA833409
41	110624	E	42A	EST	AA833409
42	110625	E	42A	EST	AK009068
				EST	AK018233
				proteasome 26S subunit mRNA	NM_016883
				gankyrin mRNA	AB022022
43	110626	E	42A	hect domain and RCC1-like domain 1 mRNA	NM_003922
				guanine nucleotide exchange factor p532 mRNA	U50078
44	110627	E	42A	hect domain and RCC1-like domain 1 mRNA	NM_003922
				guanine nucleotide exchange factor p532 mRNA	U50078
45	110628	G	42A	EST	AK008776
				EST	BC004572
46	110629	G	42A	EST	AK008776
				EST	BC004572
47	110630	G	42A	EST	AK008776
				EST	BC004572
48	110631	G	42A	EST	AK008776
				EST	BC004572
49	111822	E	39A	olfactomedin related ER localized protein mRNA	NM_019498
				pancortin-1 mRNA	D78262
				pancortin-3 mRNA	D78264
50	111823	E	39A	alpha-subunit of Gi protein mRNA (bovine)	X03642
51	111824	E	39A		
52	111825	G	43A	EST	BF532435
				TBP-associated factor 170 partial mRNA	AJ001017
53	111826	G	43A	EST	BG069663
54	111827	G	45A	ribosomal protein L3 mRNA	NM_013762
				J1 protein, yeast ribosomal protein L3 homologue mRNA	Y00225
55	111828	G	45A	ribosomal protein L3 mRNA	NM_013762
				J1 protein, yeast ribosomal protein L3 homologue mRNA	Y00225
				chromosome X clone	AC091473
56	112801	G/E	21G	calmodulin synthesis cDNA	M27844
57	112802	G/E	21G	synaptojanin mRNA	U45479
58	112803	E	2G	rS-Rex-b mRNA	U17604
				C1-13 gene product mRNA	X52817
				neuroendocrine-specific protein gene	L49137
				reticulon 1 mRNA	XM_007381
59	112804	E	2G	rS-Rex-b mRNA	U17604
				C1-13 gene product mRNA	X52817
				neuroendocrine-specific protein gene	L49137
				reticulon 1 mRNA	XM_007381
60	112807	E	49A	human clone	AC007429
61	112808	E	49A	human clone	AC007429
62	112811	E	58G	msy-3 gene	L14609
63	112812	E	58G	EST	AK011718
64	112815	E	55G	26S proteasome, subunit 112, mRNA	AJ006340
65	112816	E	55G	26S proteasome, subunit 112, mRNA	AJ006340
66	112817	E	55G	Atrn gene for membrane attractin	AB049248

Table A.1: Differential display clones printed on the cDNA microarrays.

#	SEQ ID	REGULATION	OLIGO	GENE NAME	ACCESSION
				attractin mRNA	NM_031351
67	112818	E	55G	attractin mRNA	NM_031351
68	112821	E	55G	EST	AK011123
				CD36 antigen-like 2 mRNA	NM_007644
				mLGP85/LIMP II mRNA	AB008553
				85kDa sialoglycoprotein mRNA	D10587
				lysosomal membrane protein mRNA	M68965
69	112822	E	55G	EST	AK011123
				CD36 antigen-like 2 mRNA	NM_007644
				mLGP85/LIMP II mRNA	AB008553
				85kDa sialoglycoprotein mRNA	D10587
				lysosomal membrane protein mRNA	M68965
70	112823	E	55G	mito	J01420
71	112824	E	55G		
72	112825	G/E	74A	EST	BE691495
73	112826	G/E	74A	EST	BE691495
74	112829	E	74G	microtubule-associated protein 2 mRNA	U30938
75	113211	E	74G	microtubule-associated protein 2 mRNA	U30938
76	113217	G	32A	EST	AK012950
77	113038	E	57C	EST	AV112875
78	113228	G/E	13A	EST	BF455791
				human mRNA	AL136558
79	113229	G/E	13A	EST	BB593088
80	113235	G	17A	kinesin family member C2 mRNA	NM_010630
				kinesin motor protein KIFC2 mRNA	U92949
81	113242	G	16A	EST	BG069663
82	113245	G	26A	EST	AI427639
				EST	BB272018
83	113246	G	26A	EST	AI427639
				EST	BB272018
84	111817	G	31A	EST	BB523722
85	111818	G	31A	EST	BB523722
86	111820	G	31A	mito	J01420
87	111821	G	31A	mito	J01420
88	113059			mito	J01420
89	113218			EST	AV272356
90	113223			EST	AI740442
91	113225			EST	BB306521
92	113845			mouse chromosome 16 clone	AC005816
93	113849			mito	AB049357
				NADH dehydrogenase subunit 4 (BLASTX)	BAA95649
94	113851			myelin basic protein mRNA	M15060
95	113856			EST	AI642058
96	113652		36G	DNA ligase I mRNA	U04674
97	113653		36G	mito (cyto c)	AB042432
98	113655	E	36G	EST	AK017863
				TDRM1 protein mRNA	AF251787
				hepatoma-derived growth factor mRNA	NM_008231
99	113656	E	36G	EST	AK017863

Table A.1: Differential display clones printed on the cDNA microarrays.

#	SEQ ID	REGULATION	OLIGO	GENE NAME	ACCESSION
				TDRM1 protein mRNA	AF251787
				hepatoma-derived growth factor mRNA	NM_008231
100	113659	E	36G	mitochondrial gene for subunit I of cytochrome c oxidase	X57780
101	113660	E	36G	mitochondrial gene for subunit I of cytochrome c oxidase	X57780
102	113755		36G	EST	AK004489
				major histocompatibility complex class I region	AF111102
103	113756		36G	mouse clone mRNA	BC004056
104	113757	E	37G	mito	L07096
				NADH dehydrogenase subunit 4 (BLASTX)	BAA95649
105	113759	E	37G	mito	L07096
				NADH dehydrogenase subunit 4 (BLASTX)	BAA95649
106	113859		37G	EST	AK005256
				U2 small nuclear ribonucleoprotein polypeptide A'	NM_021336
107	113861	G/E	26G	EST	AK010605
				RPS3a gene	Z83368
108	113862	G/E	26G	bad E/S	
109	114123	E	35G	AU RNA-binding protein/enoyl-coenzyme A hydratase mRNA	NM_016709
110	114124	E	35G	glycerol phosphate dehydrogenase 1 mRNA	BC005756
				glycerol-3-phosphate dehydrogenase gene	M25558
				glycerophosphate dehydrogenase gene	M13366
111	114125	E	35G	EST	AK013744
				hypothetical protein mRNA	AJ250345
112	114126	E	35G	EST	AK013744
				hypothetical protein mRNA	AJ250345
113	114132		6G	mito	AB042809
				NADH dehydrogenase subunit 2 (BLASTX)	BAA95617
114	114134		6G	mito	AB042809
				NADH dehydrogenase subunit 2 (BLASTX)	BAA95617
115	114135	E	1G	EST	BB472735
				EST	AV331125
116	114137	E	1G	bad E/S	
117	114139	E	1G	glutamate transporter mRNA	AF330257
118	114140	E	1G	glutamate transporter mRNA	AF330257
119	114141	E	1G	adenine nucleotide translocase-1 mRNA	U27315
				adenine nucleotide translocase 1 gene	AF240002
				ANC1 mRNA for adenine nucleotide carrier	X74510
120	114142	E	1G	adenine nucleotide translocase-1 mRNA	U27315
				adenine nucleotide translocase 1 gene	AF240002
				ANC1 mRNA for adenine nucleotide carrier	X74510
121	115322	G	8G	EST	AV349480
122	115324	G	8G	EST	AV349480
123	115325	E	8G	mito	AB049357
124	115326	E	8G	mito	AB042432
125	115328	E	11G	mito	AB042432
126	115329	E	11G	mito	AB049357
127	115331		11G	EST	AK004399
128	115334	G/E	11G	bad E/S	
129	115336	G/E	11G	EST	AK004399
130	115338	E	13G	aczonin mRNA	NM_011995

Table A.1: Differential display clones printed on the cDNA microarrays.

#	SEQ ID	REGULATION	OLIGO	GENE NAME	ACCESSION
				multidomain presynaptic cytomatrix protein Piccolo mRNA	NM_020098
				aczonin (BLASTX)	NP_036125
				multidomain presynaptic cytomatrix protein Piccolo (BLASTX)	AAF07822
131	115339	E	13G	EST	NM_026252
				EST	AK015401
				EST	AK021394
132	115340	G/E	13G	EST	AK009092
133	115341	G/E	13G	EST	AK009092
134	115348	E	56G	EST	AW492077
				EST	AA222116
135	115349	E	57G	EST	AK017338
136	115351	E	57G	EST	AK017338
137	115353	E	57G	EST	AK017338
138	115354	E	57G	EST	AK017338
139	115355		57G	EST	AI661025
140	115356		57G	twisted gastrulation protein mRNA	AJ297390
141	115231		57G	bad E/S	
142	115232	G/E	61G	similar to euk. translation initiation factor 3, subnt. 1 mRNA	BC006055
143	115235	E	61G	PHR1 isoform 4 mRNA	AF071000
				PHR1 isoform 2 mRNA	AF101053
				PHR1 isoform 3 mRNA	AF100613
				PHR1 isoform 1 mRNA	AF000272
				PH domain containing protein in retina 1 mRNA	NM_013746
144	115236	E	61G	PHR1 isoform 4 mRNA	AF071000
				PHR1 isoform 2 mRNA	AF101053
				PHR1 isoform 3 mRNA	AF100613
				PHR1 isoform 1 mRNA	AF000272
				PH domain containing protein in retina 1 mRNA	NM_013746
145	115238	G/E	61G	EST	BB195730
146	115109	G/E	61G	EST	BB195730
147	115112		61G	bad E/S	
148	115113		61G	EST	AW556334
149	115121	G/E	64G	EST	AK010388
150	115122	G/E	64G	EST	AK010388
151	115124	G/E	3G	EST	AK004133
				EST	AK013803
152	115125	G/E	3G	EST	AK004133
				EST	AK013803
153	115126	G/E	3G	minor histocompatibility antigen precursor mRNA	AF335543
154	115128	G/E	3G	minor histocompatibility antigen precursor mRNA	AF335543
155	115130	G/E	3G	minor histocompatibility antigen precursor mRNA	AF335543
				EST	AK005204
156	115131	G/E	3G	minor histocompatibility antigen precursor mRNA	AF335543
				EST	AK005204
157	115132	E	4G	EST	AK013661
158	115134	E	4G	EST	AK013661
159	115135		4G	EST	BB084706
160	115137		4G	EST	BB084706
161	115139	G/E	4G	bad E/S	

Table A.1: Differential display clones printed on the cDNA microarrays.

#	SEQ ID	REGULATION	OLIGO	GENE NAME	ACCESSION
162	115140	G/E	4G	EST	AK011778
				EST	AW050204
				EST	BF456090
163	115142	G/E	4G	EST	AK003519
				EST	AK002263
164	115143	G/E	4G	EST	AK003519
				EST	AK002263
165	115144	G/E	4G	bad E/S	
166	115146	G/E	4G	EST	AK019054
				regulator of G-protein signaling 11 mRNA (RGS 11)	AF061934
167	115150	E	53G	EST	AW490056
168	115151	E	53G	EST	AW490056
169	115153	G/E	53G	bad E/S	
170	115154	G/E	53G	bad E/S	
171	115155	G/E	53G	RAS-like protein expressed in neuron mRNA	NM_009065
				rin mRNA	U71202
172	115156	G	G3-1	EST	BG695121
				EST	BG075146
				EST	AW555499
173	115157	G	G3-1	EST	BE631083
174	115162	G/E	G46-1	EST	AK018188
				solute carrier family 1, member 1, mRNA	NM_009199
				glutamate transporter mRNA	U73521
175	115163	G/E	G46-1	EST	AK018188
				solute carrier family 1, member 1, mRNA	NM_009199
				glutamate transporter mRNA	U73521
176	115166	G/E	G46-1	EST	AK003863
				EST	W09074
177	115167	G/E	G46-1	EST	AK003863
				human kiaa-iso protein mRNA	XM_003228
178	115170	G/E	G47-1	EST	BE947793
179	115171	G/E	G48-1	no matches	
180	115173	G/E	G48-1	EST	AK014603
				EST	AK018329
181	115175	G/E	G48-1	ring 3 mRNA	AF318183
				ring 3 mRNA	AB010246
				bromodomain-containing 2 mRNA	NM_010238
				female sterile homeotic-related protein Frg-1 mRNA	AF045462
182	115176	G/E	G48-1	bad E/S	
183	115178	G/E	G48-1	EST	AK004507
				protein kinase BRPK mRNA	AF316872
184	115181	G/E	G48-1	bad E/S	
185	115182	G/E	G48-1	bad E/S	
186	115241	G	G11-1	myelin proteolipid protein mRNA	M15442
				PLP gene	X07221
187	115242	G	G11-1	myelin proteolipid protein mRNA	M15442
				PLP gene	X07221
188	115246	G/E	G17-1	bad E/S	
189	115248	G/E	G17-1	EST	BE648468

Table A.1: Differential display clones printed on the cDNA microarrays.

#	SEQ ID	REGULATION	OLIGO	GENE NAME	ACCESSION
				EST	AW546677
				EST	AW554687
				EST	AW909028
190	120002	G/E	G17-1	EST	AW546677
191	120003	G/E	G17-1	EST	AW546677
192	120007		G22-1	mito	AB049357
193	120008		G22-1	mito	AB049357
194	120010	G/E	G27-1	EST	AW047303
195	120011	G/E	G27-1	EST	AW047303
196	119241	G/E	G27-2	TU3A protein	XM_002850
197	119242	G/E	G27-2	TU3A protein	XM_002850
198	120014	G/E	G35-1	EST	AK020026
199	120015	G/E	G35-1	EST	AK020026
200	119253	E	G36-1	transmembrane 4 superfamily member 8	NM_019793
201	120016	E	G36-1	EST	AK006798
				transmembrane 4 superfamily member 8	NM_019793
202	119254	G	G38-1	EST	BE995089
203	119256	G	G38-1	EST	BE995089
204	120017	G	G38-2	mouse clone from chromosome 2 (lots of poss. genes)	AL450341
205	120018	G	G38-2	mouse clone from chromosome 2 (lots of poss. genes)	AL450341
206	120021	E/G	G66-1	EST	AK018317
207	120022	E/G	G66-1	EST	AK018317
208	119260	G	G67-1	EST	BE944676
209	120023	G	G67-1	low E/S	
210	119262	G/E	G67-3	low E/S	
211	119229	G/E	G67-3	mito	AB049357
212	120025		G72-1	EST	AK005205
213	119313		G72-1	EST	AK005205
214	119317		G71-1	heat shock cognate protein 70 mRNA	BC006722
215	119318		G71-1	heat shock cognate protein 70 mRNA	BC006722
216	119319	G	G24-2	lysosomal trafficking regulator mRNA	NM_010748
217	119320	G	G24-2	myelin and lymphocyte protein; T-cell differentiation protein	BC006826
218	119325	E	G50-1	mito	AB042432
219	119330			low E/S	
220	119331			low E/S	
221	109605	G/E	42A	bad E/S	
222	109606	G/E	42A	bad E/S	
223	109607	G/E	42A	bad E/S	
224	110527	E	17G	bad E/S	
225	110551	G	48A	bad E/S	
226	110552	G	48A	bad E/S	
227	110553	G	32A	bad E/S	
228	110561	G	42A	bad E/S	
229	111829		45A		
230	111830		45A		
231	112805	E	49A		
232	112806	E	49A		
233	112809	G/E	58G		
234	112810	G/E	58G		

Table A.1: Differential display clones printed on the cDNA microarrays.

#	SEQ ID	REGULATION	OLIGO	GENE NAME	ACCESSION
235	112813	G/E	58G		
236	112814	G/E	58G		
237	112827	E	74A		
238	113227	G/E	13A	bad E/S	
239	113236	G	17A	bad E/S	
240	113237	G	25A	bad E/S	
241	113238	G	25A	bad E/S	
242	113239	G	25A	bad E/S	
243	113220			bad E/S	
244	113221			bad E/S	
245	113841			bad E/S	
246	113842			bad E/S	
247	113846			bad E/S	
248	113848			bad E/S--all vector?	
249	113838	G	6A	bad E/S	
250	113839	G	6A	bad E/S	
251	114129		42G	bad E/S	
252	114130		42G	bad E/S	
253	115347	E	56G	bad E/S--all vector?	
254	115114		61G	not in database	
255	115115		61G	not in database	
256	115117		61G	bad E/S	
257	115119		61G	bad E/S	
258	115148		4G	bad E/S	
259	115160	G	G44-1	bad E/S	
260	115183	G	G9-1	bad E/S	
261	115185	G	G9-1	bad E/S	
262	115186	G	G9-2	bad E/S	
263	115191	G	G9-3	bad E/S	
264	115193	G	G10-2	bad E/S	
265	119248		G34-2	low E/S	
266	119323	E	G50-1	HepA-related protein HARP gene	AF209773
267	119324	E	G50-1	low E/S	
268	119326	E	G37-1	low E/S	
269	123116	G/E	58G	catenin alpha 2 (Catna2) mRNA	NM 009819
				alpha N-catenin II mRNA	D25282
				alpha N-catenin I mRNA	D25281
270	122935	G/E	58G	catenin alpha 2 (Catna2) mRNA	NM 009819
				alpha N-catenin II mRNA	D25282
				alpha N-catenin I mRNA	D25281

Table A.2: Alcohol-related strain differences in gene expression. The alcohol-related differentially expressed genes in B6 and D2 are listed. The gene order corresponds to the gene order in Figures 4.1A (B6) and 4.1B (D2). The B6 genes are listed first and the D2 genes are listed second. Note that the genes names are current as of July 31, 2006.

Table A.2: Alcohol-related strain differences in gene expression.

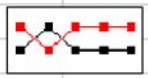
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
1	C57Bl/6J		up	0.00979	*Ankrd27*	Ankyrin repeat domain 27 (VPS9 domain)	7	602239	AI449697
2			up	0.01393				H3041E06	
3			up	0.00011	*3110050N22Rik*	RIKEN cDNA 3110050N22 gene	3	AI849782	AI849782
4			up	0.00433	*Rab31*	RAB31, member RAS oncogene family	17	614369	AI448835
5			up	0.00001	*Cyp51*	Cytochrome P450, family 51	5	H4065G08	BQ560678.1
6			up	0.00006	*Junb*	Jun-B oncogene	8	AI850555	AI850555
7			up	0.00078	*Trio*	Triple functional domain (PTPRF interacting)	15	H3005C05	BG063282.2
8			up	0.03751	*Gadd45b*	Growth arrest and DNA-damage-inducible 45 beta	10	AI848850	AI848850
9			up	0.00251				H4016E05	
10			up	0.00277	*Ddx3y*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked	Y	H4032D03	BQ555138.1
11			up	0.00612	*Ints3*	Integrator complex subunit 3	3	H3029E02	BG065296.2
12			up	0.00762	*Fos*	FBJ osteosarcoma oncogene	12	H3084G03	BG070196.2
13			up	0.00179	*Ralgds*	Ral guanine nucleotide dissociation stimulator	2	AI843629	AI843629
14			down	0.00020	*Capzb*	Capping protein (actin filament) muscle Z-line, beta	4	H3148E03	BG075541.2
15			down	0.00951	*LOC668212*	Similar to RIKEN cDNA C920006C10	12	H4053E10	BQ558658.1
16			down	0.03759	*A530088H08Rik*	RIKEN cDNA A530088H08 gene	11	H4025D04	BQ553979.1
17			down	0.00656	*Cc2d1b*	Coiled-coil and C2 domain containing 1B	4	AI835541	AI835541
18			down	0.00058	*Id3*	Inhibitor of DNA binding 3	4	AI839283	AI839283
19			down	0.00428	*Pcnxl3*	Pecanex-like 3 (Drosophila)	19	H4044C06	BQ557134.1
20			down	0.00260	*Ptrf*	Polymerase I and transcript release factor	11	H3030F05	BG065396.2
21			down	0.01371	*Sfrs2ip*	Splicing factor, arginine/serine-rich 2, interacting protein	15	H4009H05	BQ551529.1
22			down	0.03889	*Znrf3*	Zinc and ring finger 3	11	H4073G09	BQ562042.1
23			down	0.01818	*Falz*	Fetal Alzheimer antigen	11	H3087A06	BG070407.2
24			down	0.00011	*Capg*	Capping protein (actin filament), gelsolin-like	6	H3010G06	CK334228.1
25	C57Bl/6J		up	0.02718	*2310035C23Rik*	RIKEN cDNA 2310035C23 gene	1	AI843794	AI843794
26			up	0.01102	*Sos1*	Son of sevenless homolog 1 (Drosophila)	17	H3029E04	BG065298.2
27			up	0.01061	*D5Wsu178e*	DNA segment, Chr 5, Wayne State University 178, expressed	5	636058	AI451319
28			up	0.04390	*Ccl27*	Chemokine (C-C motif) ligand 27	4	H3159A01	BG076373.2
29			up	0.02632	*Asf1a*	ASF1 anti-silencing function 1 homolog A (S. cerevisiae)	10	AI854496	AI854496
30			up	0.00133	*Odz2*	Odd Oz/ten-m homolog 2 (Drosophila)	11	AI837170	AI837170

Table A.2: Alcohol-related strain differences in gene expression.

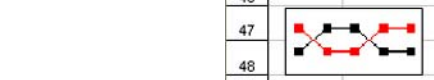
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
31			up	0.02377	*D130020G16Rik*	RIKEN cDNA D130020G16 gene	16	H4013B07	BQ552053.1
32			up	0.00204	*0610031J06Rik*	RIKEN cDNA 0610031J06 gene	3	AI848304	AI848304
33			up	0.00589	*Ddx28*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 28	8	600869	AI449652
34			up	0.01026	*Pbx3*	Pre B-cell leukemia transcription factor 3	2	H3045B05	BG066541.2
35			up	0.00008	*Insig1*	Insulin induced gene 1	5	AI848986	AI848986
36			up	0.00071	*Sesn1*	Sestrin 1	10	AI843965	AI843965
37			up	0.00067	*9130204K15Rik*	RIKEN cDNA 9130204K15 gene	11	H3141C01	BG074939.2
38			up	0.00369	*Fnbp4*	Formin binding protein 4	2	AI835645	AI835645
39			up	0.04769	*A930015D03Rik*	RIKEN cDNA A930015D03 gene	Un	AI848402	AI848402
40			up	0.00403	*Akp2*	Alkaline phosphatase 2, liver	4	AI851320	AI851320
41			up	0.00144	*Npr1*	Natriuretic peptide receptor 1	3	H4043F01	BQ557019.1
42			up	0.01196	*Mknk2*	MAP kinase-interacting serine/threonine kinase 2	10	AI845732	AI845732
43			up	0.00004	*Pbx1*	Pre B-cell leukemia transcription factor 1	1	H4007G01	BQ551166.1
44			up	0.00408	*Rhob*	Ras homolog gene family, member B	12	H3008A09	BG063507.2
45			up	0.03120	*Ccrn4l*	CCR4 carbon catabolite repression 4-like (S. cerevisiae)	3	AI848832	AI848832
46			up	0.02634				H3084C10	
47			up	0.00001	*Tsc22d3*	TSC22 domain family 3	X	476319	AI326808
48			down	0.00326	*Cmpk*	Cytidylate kinase	4	H4001C10	BQ550088.1
49			down	0.00302	*2500003M10Rik*	RIKEN cDNA 2500003M10 gene	3	AI843731	AI843731
50			down	0.00829	*Ppp1cc*	Protein phosphatase 1, catalytic subunit, gamma isoform	5	H3103D01	BG071790.2
51			down	0.04929	*Btbd12*	BTB (POZ) domain containing 12	16	H4057F07	BQ559337.1
52			down	0.02755	*Ttc15*	Tetratricopeptide repeat domain 15	12	H4022D08	BQ553496.1
53			down	0.03796	*9430020K01Rik*	RIKEN cDNA 9430020K01 gene	18	H4041G09	BQ556727.1
54			down	0.00006	*Hspd1*	Heat shock protein 1 (chaperonin)	11	H3030A03	BG078626.2
55			down	0.02474	*Postn*	Periostin, osteoblast specific factor	3	H4027G04	BQ554373.1
56			down	0.00059	*Sfrs7*	Splicing factor, arginine/serine-rich 7	17	AI836707	AI836707
57			down	0.00468	*Fgfr1*	Fibroblast growth factor receptor 1	8	H4013H01	BQ552166.1
58			down	0.00535	*Pogz*	Pogo transposable element with ZNF domain	3	H3087B02	BG070415.2
59			down	0.00011	*Sfrs6*	Splicing factor, arginine/serine-rich 6	2	AI846595	AI846595
60	C57Bl/6J		up	0.01277		EST		AI837335	AI837335

Table A.2: Alcohol-related strain differences in gene expression.

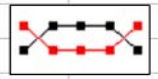
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
61			up	0.00895	*Dscr1l2*	Down syndrome critical region gene 1-like 2	4	H3152G06	BG075865.2
62			up	0.02111	*Prps1*	Phosphoribosyl pyrophosphate synthetase 1	X	H3024G02	BG064866.2
63			up	0.00179	*Tfb2m*	Transcription factor B2, mitochondrial	1	575114	AI323620
64			up	0.00493	*Rps6ka4*	Ribosomal protein S6 kinase, polypeptide 4	19	AI848992	AI848992
65			up	0.02296	*Zfp609*	Zinc finger protein 609	9	AI848174	AI848174
66			up	0.00273	*Kcnip1*	Kv channel-interacting protein 1	11	H4032D06	BQ555143.1
67			up	0.01082	*6430511F03*	Hypothetical protein 6430511F03	X	AI851885	AI851885
68			up	0.04585	*Pygb*	Brain glycogen phosphorylase	2	AI846739	AI846739
69			up	0.02132	*Dnajc7*	DnaJ (Hsp40) homolog, subfamily C, member 7	11	H4004G08	BQ550675.1
70			up	0.04003	*Btbd11*	BTB (POZ) domain containing 11	10	AI847060	AI847060
71			up	0.03846	*St6gal2*	Beta galactoside alpha 2,6 sialyltransferase 2	17	H4034B06	BQ555448.1
72			up	0.01394	*Srgap3*	SLIT-ROBO Rho GTPase activating protein 3	6	AI849784	AI849784
73			up	0.00673	*Hrmt1l3*	Heterogeneous nuclear ribonucleoprotein methyltransferase-like 3 (S. cerevisiae)	7	AI854416	AI854416
74			up	0.02823	*Rsbn1*	Rosbin, round spermatid basic protein 1	3	636872	AI450166
75			down	0.03034	*Klf6*	Kruppel-like factor 6	13	608451	AI448727
76			down	0.00067		EST		AI836068	AI836068
77			down	0.04696	*Phf20*	PHD finger protein 20	2	H4037H02	BQ556064.1
78			down	0.01315	*Txndc1*	Thioredoxin domain containing 1	12	H3015G06	BG064124.2
79			down	0.01790	*Gnaq*	Guanine nucleotide binding protein, alpha q polypeptide	19	H4067F01	BQ560978.1
80			down	0.00925	*Ezh1*	Enhancer of zeste homolog 1 (Drosophila)	11	H3154F11	BG076024.2
81			down	0.01304	*Xrn2*	5'-3' exoribonuclease 2	2	H3031H09	BG065507.2
82			down	0.02529	*LOC670176*	Similar to thymine DNA glycosylase isoform 2	7	AI843711	AI843711
83			down	0.01053	*Mtus1*	Mitochondrial tumor suppressor 1	8	H3140C10	CK335090.1
84			down	0.01853	*E430002G05Rik*	RIKEN cDNA E430002G05 gene	2	318668	AI414397
85			down	0.00573	*Ehmt2*	Euchromatic histone lysine N-methyltransferase 2	17	H3150B12	BG088216.2
86	C57Bl/6J		up	0.01658	*Ucp2*	Uncoupling protein 2 (mitochondrial, proton carrier)	7	608716	AI325229
87			up	0.00101	*Adora1*	Adenosine A1 receptor	1	AI848715	AI848715
88			up	0.00034	*Cpt2*	Carnitine palmitoyltransferase 2	4	AI853250	AI853250
89			up	0.03948	*A730041O15Rik*	RIKEN cDNA A730041O15 gene	5	AI848145	AI848145
90			up	0.00015	*Zfp423*	Zinc finger protein 423	8	555773	AI428511

Table A.2: Alcohol-related strain differences in gene expression.

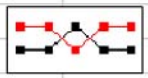
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
91			up	0.00000	*Pnpla2*	Patatin-like phospholipase domain containing 2	7	H3089B03	BG070602.2
92			up	0.02960	*2600014M03Rik*	RIKEN cDNA 2600014M03 gene	11	AI848298	AI848298
93			up	0.00172				H3101F10	BG084503.2
94			up	0.00235	*AI467657*	Expressed sequence AI467657	9	AI853366	AI853366
95			up	0.03707	*Camkk1*	Calcium/calmodulin-dependent protein kinase kinase 1, alpha	11	AI846603	AI846603
96			up	0.01471	*Pcbp4*	Poly(rC) binding protein 4	9	373664	AI414246
97			up	0.00378	*Cri1*	CREBBP/EP300 inhibitory protein 1	2	AI844939	AI844939
98			up	0.00184	*Mal*	Myelin and lymphocyte protein, T-cell differentiation protein	2	AI839913	AI839913
99			up	0.01781	*Btbd10*	BTB (POZ) domain containing 10	7	AI836768	AI836768
100			up	0.00165	*Tef*	Thyrotroph embryonic factor	15	615611	AI448996
101			up	0.04430	*Nfrkb*	Nuclear factor related to kappa B binding protein	9	AI838631	AI838631
102			up	0.00001	*Dock6*	Dedicator of cytokinesis 6	9	H4056A09	BQ559082.1
103			up	0.00082	*Gpr22*	G protein-coupled receptor 22	12	AI848708	AI848708
104			up	0.01490	*Chchd2*	Coiled-coil-helix-coiled-coil-helix domain containing 2	X	AI835937	AI835937
105			up	0.00953	*Ptprf*	Protein tyrosine phosphatase, receptor type, F	4	AI848471	AI848471
106			up	0.00207				H3069A05	BG068758.2
107			up	0.02661	*Eif4g3*	Eukaryotic translation initiation factor 4 gamma, 3	4	AI838775	AI838775
108			up	0.00328	*Epb4.1l1*	Erythrocyte protein band 4.1-like 1	2	AI854440	AI854440
109			up	0.04404				H3034B04	
110			up	0.00176	*Gtpbp3*	GTP binding protein 3	8	AI852334	AI852334
111			down	0.00551				H3121H08	
112			down	0.00717	*Mat2a*	Methionine adenosyltransferase II, alpha	6	AI845231	AI845231
113			down	0.02252	*Nfkbil2*	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 2	15	AI851335	AI851335
114			down	0.00453	*Tmem55a*	Transmembrane protein 55A	4	AI837052	AI837052
115			down	0.00266	*0610009D07Rik*	RIKEN cDNA 0610009D07 gene	12	H3159G04	BG076503.2
116			down	0.01771	*Myo9b*	Myosin IXb	8	H4029H01	BQ554706.1
117			down	0.00222	*Rpl30*	Ribosomal protein L30	15	H3030F06	BG065397.2
118			down	0.02398	*AW495713*	Expressed sequence AW495713	8	H4071E03	BQ561631.1
119			down	0.02822	*Zfp277*	Zinc finger protein 277	12	H3099C11	BG071467.2
120			down	0.00129	*2900083I11Rik*	RIKEN cDNA 2900083I11 gene	5	AI853104	AI853104

Table A.2: Alcohol-related strain differences in gene expression.

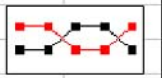
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
121	C57Bl/6J		up	0.01803	*Wdr26*	WD repeat domain 26	1	H3101E04	BG071643.1
122			up	0.02404				H3063E01	
123			up	0.02030	*Enc1*	Ectodermal-neural cortex 1	13	AI853828	AI853828
124			up	0.03130	*Fen1*	Flap structure specific endonuclease 1	19	H3009B04	BG063590.2
125			up	0.01280	*Rpl14*	Ribosomal protein L14	9	AI839125	AI839125
126			up	0.00007	*Mbd2*	Methyl-CpG binding domain protein 2	18	H4002F09	BQ550320.1
127			up	0.04570	*Lrrc47*	Leucine rich repeat containing 47	4	AI849627	AI849627
128			up	0.00030	*Rasa2*	RAS p21 protein activator 2	9	AI844397	AI844397
129			up	0.00191	*Ogdh*	Oxoglutarate dehydrogenase (lipoamide)	11	H3083F12	CK334672.1
130			up	0.00027	*Slc44a5*	Solute carrier family 44, member 5	3	539387	AI427060
131			up	0.00077	*Mt2*	Metallothionein 2	8	H3010E09	CK334225.1
132			up	0.00762	*Mt1*	Metallothionein 1	8	H3020C02	BG064480.2
133			down	0.00914	*Stam2*	Signal transducing adaptor molecule (SH3 domain and ITAM motif) 2	2	H3007H10	BG063497.2
134			down	0.02201	*Slc16a10*	Solute carrier family 16 (monocarboxylic acid transporters), member 10	10	H3044E10	BG066654.2
135			down	0.00218	*Usp46*	Ubiquitin specific peptidase 46	5	AI852059	AI852059
136			down	0.02858	*Srf*	Serum response factor	17	AI846181	AI846181
137			down	0.00383	*Bet1l*	Blocked early in transport 1 homolog (S. cerevisiae)-like	7	AI836964	AI836964
138			down	0.00235	*Mvd*	Mevalonate (diphospho) decarboxylase	8	H3041B11	BG066289.2
139			down	0.04975	*Armet*	Arginine-rich, mutated in early stage tumors	9	H3009A06	BG063580.2
140			down	0.00596	*1110014N23Rik*	RIKEN cDNA 1110014N23 gene	19	AI845926	AI845926
141			down	0.04350				H3053H08	BG067390.2
142			down	0.00182	*Atf4*	Activating transcription factor 4	15	AI850284	AI850284
143			down	0.00500	*Snrpa*	Small nuclear ribonucleoprotein polypeptide A	7	616400	AI324697
144			down	0.00039	*Pja1*	Praja1, RING-H2 motif containing	X	H3070G09	BG068926.2
145			down	0.00064	*Spred2*	Sprouty-related, EVH1 domain containing 2	11	AI851250	AI851250
146			down	0.04145	*Gnb5*	Guanine nucleotide binding protein, beta 5	9	AI847689	AI847689
147			down	0.00347	*R3hdm2*	R3H domain containing 2	10	AI838049	AI838049
148			down	0.00073	*Rexo2*	REX2, RNA exonuclease 2 homolog (S. cerevisiae)	9	AI839882	AI839882
149			down	0.01671	*Hyou1*	Hypoxia up-regulated 1	9	AI847189	AI847189
150	C57Bl/6J		up	0.00332	*Ftsj1*	Ftsj homolog 1 (E. coli)	X	H4057C08	BQ559284.1

Table A.2: Alcohol-related strain differences in gene expression.

	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
151			up	0.02497	*Ywhab*	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide	2	H4007D04	BQ551106.1
152			up	0.01293	*Sema6a*	Sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A	18	H4040D06	BQ556481.1
153			up	0.03846	*Grb2*	Growth factor receptor bound protein 2	11	H3153D02	BG075917.2
154			up	0.00732	*Cd276*	CD276 antigen	9	352990	AI415625
155			up	0.00881				H4024G10	
156			up	0.02746	*Dnajb14*	DnaJ (Hsp40) homolog, subfamily B, member 14	3	AI843057	AI843057
157			up	0.00404	*C330002I19Rik*	RIKEN cDNA C330002I19 gene	12	H3152D03	BG075834.2
158			up	0.00038				H3083C08	
159			up	0.00348	*Hnrpl*	Heterogeneous nuclear ribonucleoprotein L	7	H3137D02	BG074625.2
160			up	0.02377	*Rab12*	RAB12, member RAS oncogene family	17	AI848325	AI848325
161			up	0.01224		EST		AI839478	AI839478
162			up	0.04102	*Agtrap*	Angiotensin II, type I receptor-associated protein	4	H3027C07	BG065093.2
163			up	0.00261	*Atp5a1*	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1	18	H3116G12	CK334843.1
164			down	0.00951	*Fkbp4*	FK506 binding protein 4	6	H3015G10	BG064128.2
165			down	0.04293	*Adfp*	Adipose differentiation related protein	4	AI848880	AI848880
166			down	0.00472	*Anp32b*	Acidic nuclear phosphoprotein 32 family, member B	4	AI842771	AI842771
167			down	0.00037	*Fech*	Ferrochelatase	18	H3124D02	BG086309.2
168			down	0.00625	*Rap1b*	RAS related protein 1b	10	AI838463	AI838463
169			down	0.00835	*5730403M16Rik*	RIKEN cDNA 5730403M16 gene	7	H3096D12	BG071243.1
170			down	0.00800	*Cdc2l1*	Cell division cycle 2-like 1	4	596211	AI528628
171			down	0.01405	*Selpl*	Selectin, platelet (p-selectin) ligand	5	AI838678	AI838678
172			down	0.00081	*Foxk2*	Forkhead box K2	11	AI837643	AI837643
173			down	0.04930				H3030D02	BG065374.2
174			down	0.01536	*1110008P14Rik*	RIKEN cDNA 1110008P14 gene	2	H3009G09	BG063647.1
175			down	0.00000				H3102A01	
176			down	0.00209	*A930017K11Rik*	RIKEN cDNA A930017K11 gene	17	AI846462	AI846462
177			down	0.00311	*Aff3*	AF4/FMR2 family, member 3	1	H4074C11	BQ562137.1
178			down	0.00780	*Atp2b2*	ATPase, Ca++ transporting, plasma membrane 2	6	H4037F12	BQ556039.1
179	C57Bl/6J		up	0.01401	*LOC619719*	Hypothetical protein LOC619719	3	AI853688	AI853688
180			up	0.00082		Transcribed sequences		AI839852	AI839852

Table A.2: Alcohol-related strain differences in gene expression.

	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
181			up	0.03384	*Clic1*	Chloride intracellular channel 1	17	AI838299	AI838299
182			up	0.00829	*1810037C20Rik*	RIKEN cDNA 1810037C20 gene	X	AI854852	AI854852
183			up	0.03601	*Tmem38a*	Transmembrane protein 38a	8	367610	AI413399
184			up	0.00011	*5330417C22Rik*	RIKEN cDNA 5330417C22 gene	3	AI844689	AI844689
185			up	0.03480	*Sco1*	SCO cytochrome oxidase deficient homolog 1 (yeast)	11	372377	AI415359
186			up	0.00608	*Tmem57*	Transmembrane protein 57	4	H3126D02	BG073638.2
187			up	0.01615	*Mrps26*	Mitochondrial ribosomal protein S26	2	AI836090	AI836090
188			up	0.00807	*4833417A11Rik*	RIKEN cDNA 4833417A11 gene	1	608273	AI450947
189			up	0.03165				H3070A01	
190			up	0.00020				AI837136	AI837136
191			up	0.00222				H3079E10	
192			up	0.03018		Adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B230215D24 product:unknown EST, full insert sequence		AI849106	AI849106
193			up	0.01834				551313	AI428261
194			up	0.02517				H3034D10	
195			up	0.01063	*Heyl*	Hairy/enhancer-of-split related with YRPW motif-like	4	H3010G03	BG063796.2
196			up	0.00749	*1700023B02Rik*	RIKEN cDNA 1700023B02 gene	2	329609	AI414443
197			up	0.03947	*Gja6*	Gap junction membrane channel protein alpha 6	X	AI848688	AI848688
198			up	0.02936				H3093F07	
199			up	0.01768	*Slc29a1*	Solute carrier family 29 (nucleoside transporters), member 1	17	AI848257	AI848257
200			up	0.02612	*Kcnq5*	Potassium voltage-gated channel, subfamily Q, member 5	1	H4020G04	BQ553225.1
201			up	0.03814	*Mtdh*	Metadherin	15	H3015H08	BG077482.2
202			up	0.02162	*Smarcad1*	SWI/SNF-related, matrix-associated actin-dependent regulator of chromatin, subfamily a, containing DEAD/H box 1`	6	H4054C06	BQ558778.1
203			up	0.01882	*2810417J12Rik*	RIKEN cDNA 2810417J12 gene	17	AI848613	AI848613
204			up	0.03917	*Adcy5*	Adenylate cyclase 5	16	AI844317	AI844317
205			up	0.04191	*Tpbpb*	Trophoblast specific protein beta	13	H3003A10	BG063011.2
206			up	0.00758	*Zap70*	Zeta-chain (TCR) associated protein kinase	1	576335	AI327364
207			up	0.03103	*1500041N16Rik*	RIKEN cDNA 1500041N16 gene	8	AI842657	AI842657
208			up	0.04655	*Mll3*	Myeloid/lymphoid or mixed-lineage leukemia 3	5	H3057C11	BG067709.2
209			up	0.00744	*A730013G04*	Hypothetical protein A730013G04	11	H3057G03	BG067747.2
210			up	0.03366	*Glrx*	Glutaredoxin	13	H3030A12	BG065350.2

Table A.2: Alcohol-related strain differences in gene expression.

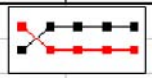
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
211			up	0.00204				H3050C04	BG067071.2
212			down	0.00932	*AI841971*	Expressed sequence AI841971	13	AI841971	AI841971
213			down	0.02461	*Pgd*	Phosphogluconate dehydrogenase	4	H4070F03	BQ561486.1
214			down	0.01456				AI850898	AI850898
215			down	0.01736	*2010012C16Rik*	RIKEN cDNA 2010012C16 gene	6	H4005G06	BQ550841.1
216			down	0.00888	*A030003K02Rik*	RIKEN cDNA A030003K02 gene	11	H3147D02	BG075449.2
217			down	0.03146	*Gm440*	Gene model 440, (NCBI)	5	H3153G09	BG075955.2
218			down	0.02585	*Capza2*	Capping protein (actin filament) muscle Z-line, alpha 2	6	H3085F12	BG070286.2
219			down	0.02158	*Lgtn*	Ligatin	1	H3138F09	BG074737.2
220			down	0.00956	*Tufm*	Tu translation elongation factor, mitochondrial	7	H3033F01	BG065643.2
221			down	0.02923		EST		AI842418	AI842418
222			down	0.01990				H3145F01	
223			down	0.04799				H3066G01	
224			down	0.00549	*5230400G24Rik*	RIKEN cDNA 5230400G24 gene	1	H3092H05	BG070850.2
225			down	0.03026	*Pycard*	PYD and CARD domain containing	7	AI843240	AI843240
226			down	0.04968	*Jakmip2*	Janus kinase and microtubule interacting protein 2	18	H4032G10	BQ555215.1
227			down	0.00192	*Ches1*	Checkpoint suppressor 1	12	AI849096	AI849096
228			down	0.02058	*Ppap2b*	Phosphatidic acid phosphatase type 2B	4	AI847054	AI847054
229			down	0.04760	*Ppp2r1a*	Protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), alpha isoform	17	H3028E04	BG065208.2
230			down	0.01390	*AI875142*	Expressed sequence AI875142	11	AI846413	AI846413
231			down	0.01676				H3025A07	
232			down	0.03333	*Son*	Son cell proliferation protein	16	AI845729	AI845729
233			down	0.01989	*Opa1*	Optic atrophy 1 homolog (human)	16	H3148C03	BG075519.2
234			down	0.00438	*BC018601*	CDNA sequence BC018601	11	H3053D09	BG067350.2
235			down	0.00138				H4077F10	BQ562737.1
236			down	0.00573	*Calr*	Calreticulin	8	AI848391	AI848391
237			down	0.00267	*Praf2*	PRA1 domain family 2	X	H3038G12	BG079358.2
238			down	0.00509	*Sfrs3*	Splicing factor, arginine/serine-rich 3 (SRp20)	3	H3025G01	BG064951.2
239			down	0.02064	*Atp8a1*	ATPase, aminophospholipid transporter (APLT), class I, type 8A, member 1	5	AI853962	AI853962
240			down	0.00423	*Snta1*	Syntrophin, acidic 1	2	H3028E02	BG065206.2

Table A.2: Alcohol-related strain differences in gene expression.

	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
241			down	0.03666				H3032D10	
242			down	0.01159	*4933433P14Rik*	RIKEN cDNA 4933433P14 gene	12	AI848430	AI848430
243			down	0.02811	*Cct5*	Chaperonin subunit 5 (epsilon)	15	H3023F07	BG064769.2
244			down	0.01762	*Lmx1a*	LIM homeobox transcription factor 1 alpha	1	317647	W11485.1
245			down	0.00405	*Lcat*	Lecithin cholesterol acyltransferase	8	AI848758	AI848758
246			down	0.00386	*Eif4g2*	Eukaryotic translation initiation factor 4, gamma 2	7	H4078G01	BQ562918.1
247			down	0.00032	*Ptpla*	Protein tyrosine phosphatase-like (proline instead of catalytic arginine), member a	2	H4026F10	BQ554197.1
248			down	0.00398				H3057D07	
249			down	0.00007	*Tmem106c*	Transmembrane protein 106C	15	AI851309	AI851309
250			down	0.03793	*Abcf2*	ATP-binding cassette, sub-family F (GCN20), member 2	5	AI837318	AI837318
251			down	0.02085	*Dnaja1*	DnaJ (Hsp40) homolog, subfamily A, member 1	4	AI841066	AI841066
252			down	0.00112	*D12Ertd553e*	DNA segment, Chr 12, ERATO Doi 553, expressed	12	H3064E09	BG068352.2
253			down	0.00267	*Prmt2*	Protein arginine N-methyltransferase 2	10	H3059E02	BG067900.2
254			down	0.01932	*Hsp110*	Heat shock protein 110	5	AI845946	AI845946
255			down	0.02293	*Calm2*	Calmodulin 2	17	H3006H06	CK334135.1
256			down	0.00625	*Unc13a*	Unc-13 homolog A (C. elegans)	8	AI843770	AI843770
257			down	0.01858	*Hivep2*	Human immunodeficiency virus type I enhancer binding protein 2	10	H4045A11	BQ557273.1
258			down	0.00381	*Lrrc28*	Leucine rich repeat containing 28	7	AI852668	AI852668
259			down	0.00035				H3104E08	
260			down	0.04676	*4930534B04Rik*	RIKEN cDNA 4930534B04 gene	12	372704	AI414212
261			down	0.00335	*Chordc1*	Cysteine and histidine-rich domain (CHORD)-containing, zinc-binding protein 1	9	H3044D10	BG066642.2
262			down	0.00596	*1500010G04Rik*	RIKEN cDNA 1500010G04 gene	6	H3072D10	BG069073.2
263	C57Bl/6J		up	0.00075	*Eml2*	Echinoderm microtubule associated protein like 2	7	H3073G09	BG069197.2
264			up	0.00005	*Sin3b*	Transcriptional regulator, SIN3B (yeast)	8	AI834948	AI834948
265			up	0.00606	*4930471M23Rik*	RIKEN cDNA 4930471M23 gene	5	AI847696	AI847696
266			up	0.00250	*Gprasp1*	G protein-coupled receptor associated sorting protein 1	X	H3149C09	BG075607.2
267			up	0.04665	*Cct4*	Chaperonin subunit 4 (delta)	11	H3044B01	BG066612.1
268			up	0.01255	*Tbl1xr1*	Transducin (beta)-like 1X-linked receptor 1	3	H4077A10	BQ562637.1
269			up	0.00412	*Zcchc6*	Zinc finger, CCHC domain containing 6	13	638204	AI450318
270			up	0.00524	*Adcy7*	Adenylate cyclase 7	8	596455	AI528639

Table A.2: Alcohol-related strain differences in gene expression.

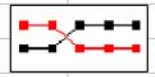
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
271			up	0.00144	*Rufy3*	RUN and FYVE domain containing 3	5	AI843460	AI843460
272			up	0.02334	*Pak6*	P21(CDKN1A)-activated kinase 6	2	AI850458	AI850458
273			up	0.01991				H3037E05	BG065976.2
274			up	0.00256	*Bckdha*	Branched chain ketoacid dehydrogenase E1, alpha polypeptide	7	314098	AI323918
275			up	0.01296	*Ppp1r1b*	Protein phosphatase 1, regulatory (inhibitor) subunit 1B	11	AI839758	AI839758
276			up	0.00483	*Lmbr1l*	Limb region 1 like	15	H3075D12	BG069429.2
277			up	0.00909	*Ghr*	Growth hormone receptor	15	H4008B07	BQ551242.1
278			up	0.02934	*Nek11*	NIMA (never in mitosis gene a)-related expressed kinase 11	9	H3073A08	BG069128.2
279			up	0.02461	*Slc7a8*	solute carrier family 7 (cationic amino acid transporter, y+ system), member 8	14	AI849037	AI849037
280			up	0.00018	*Cdh13*	Cadherin 13	8	AI842303	AI842303
281			up	0.00588	*1810009B06Rik*	RIKEN cDNA 1810009B06 gene	15	AI847108	AI847108
282			up	0.00151				H3051D10	BG067175.2
283			up	0.01129	*Slc35c2*	Solute carrier family 35, member C2	2	AI838675	AI838675
284			down	0.00870	*3110021N24Rik*	RIKEN cDNA 3110021N24 gene	4	H3146A09	BG087883.2
285			down	0.01969	*Hs3st2*	Heparan sulfate (glucosamine) 3-O-sulfotransferase 2	7	H4038G04	BQ556211.1
286			down	0.00886	*Lass6*	Longevity assurance homolog 6 (S. cerevisiae)	2	H3148G11	BG075568.2
287			down	0.00142		ESTs		AI847070	AI847070
288			down	0.00468		Transcribed sequences		AI849045	AI849045
289			down	0.01535	*Zfp68*	Zinc finger protein 68	5	H3050G01	CK334552.1
290			down	0.00000	*Hspa5*	Heat shock 70kD protein 5 (glucose-regulated protein)	2	AI843553	AI843553
291			down	0.00004				H3096C04	
292			down	0.00003	*Mpdu1*	Mannose-P-dolichol utilization defect 1	11	H3090D12	BG070723.2
293			down	0.00039	*Egr1*	Early growth response 1	18	H3092E09	BG070825.2
294			down	0.00274				H3017B03	
295			down	0.00896	*Oas1c*	2'-5' oligoadenylate synthetase 1C	5	H3058H01	BG067845.2
296			down	0.00380	*Xbp1*	X-box binding protein 1	11	H3097C04	BG071303.2
297			down	0.04280	*2310002J21Rik*	RIKEN cDNA 2310002J21 gene	19	AI848627	AI848627
298			down	0.00111	*Sfpq*	Splicing factor proline/glutamine rich (polypyrimidine tract binding protein associated)	4	H3141G07	BG074984.2
299			down	0.04745	*Hnrpc*	Heterogeneous nuclear ribonucleoprotein C	14	AI836825	AI836825
300			down	0.00881	*Ptn*	Pleiotrophin	6	AI846004	AI846004

Table A.2: Alcohol-related strain differences in gene expression.

	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
301			down	0.00019	*Arfip2*	ADP-ribosylation factor interacting protein 2	7	AI852332	AI852332
302			down	0.03950	*Fem1b*	Feminization 1 homolog b (C. elegans)	9	H3073H05	BG069205.2
303			down	0.01280	*Qk*	Quaking	17	AI837287	AI837287
304			down	0.02407	*Mrpl23*	Mitochondrial ribosomal protein L23	7	H3101A08	BG071604.2
305			down	0.00680	*Ldha*	Lactate dehydrogenase A	7	H3023H12	BG064797.2
306			down	0.00353	*Mterfd1*	MTERF domain containing 1	13	652668	AI465491
307			down	0.00389	*Bst2*	Bone marrow stromal cell antigen 2	8	H3056G04	BG067655.2
308			down	0.03527	*Apoa4*	Apolipoprotein A-IV	9	H3025G09	BG064959.2
309			down	0.00409	*Pde4dip*	Phosphodiesterase 4D interacting protein (myomegalin)	3	H4028D02	BQ554471.1
310			down	0.02078	*4930427A07Rik*	RIKEN cDNA 4930427A07 gene	12	H3089C06	BG070617.2
311			down	0.03757	*Klhl22*	Kelch-like 22 (Drosophila)	16	H4003C04	BQ550415.1
312			down	0.01420		EST		AI837058	AI837058
313			down	0.00797	*Efna5*	Ephrin A5	17	H3141D02	BG074951.2
314			down	0.00383	*Fbxo3*	F-box only protein 3	2	H3102A06	BG071684.2
315			down	0.01342	*Ppp2ca*	Protein phosphatase 2 (formerly 2A), catalytic subunit, alpha isoform	11	H3007C02	BG063435.1
316			down	0.00687	*Ctcf*	CCCTC-binding factor	8	H3040A12	BG066197.2
317			down	0.03154	*Ppp1r3d*	Protein phosphatase 1, regulatory subunit 3D	2	AI851283	AI851283
318			down	0.02127	*Fcho2*	FCH domain only 2	13	H4021E10	BQ553369.1
319			down	0.00877	*Atg5*	Autophagy-related 5 (yeast)	10	H3052C04	BG067250.2
320	C57Bl/6J		up	0.02845	*Ninj1*	Ninjurin 1	13	H3072B10	BG069050.2
321			up	0.00869	*Cirbp*	Cold inducible RNA binding protein	10	H3124B02	BG073558.2
322			up	0.00535	*Slc6a6*	Solute carrier family 6 (neurotransmitter transporter, taurine), member 6	6	H4016E11	BQ552559.1
323			up	0.02187	*2310047O13Rik*	RIKEN cDNA 2310047O13 gene	2	H3058F06	BG067830.2
324			up	0.01210	*Tkt*	Transketolase	14	H3011G11	CK334246.1
325			up	0.04652	*Tbc1d24*	TBC1 domain family, member 24	17	AI843603	AI843603
326			up	0.00268	*Ifi203*	Interferon activated gene 203	1	H4066H11	BQ560864.1
327			up	0.03184	*Vps24*	Vacuolar protein sorting 24 (yeast)	6	H3124B11	BG073566.2
328			up	0.00132				H3155F06	CK335249.1
329			up	0.00012	*1500034J01Rik*	RIKEN cDNA 1500034J01 gene	8	H3083H11	BG070124.2
330			up	0.00532	*Hdlbp*	High density lipoprotein (HDL) binding protein	1	AI844871	AI844871

Table A.2: Alcohol-related strain differences in gene expression.

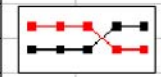
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
331			up	0.01224	*Sfmbt2*	Scm-like with four mbt domains 2	2	H4007B05	BQ551067.1
332			up	0.00403	*Ube2h*	Ubiquitin-conjugating enzyme E2H	6	H3057B09	BG067695.2
333			down	0.02677	*Mrpl21*	Mitochondrial ribosomal protein L21	19	H3005C03	BG076737.2
334			down	0.00015	*Hivep2*	Human immunodeficiency virus type I enhancer binding protein 2	10	H4028B01	BQ554429.1
335			down	0.00952	*Rwdd1*	RWD domain containing 1	10	AI853842	AI853842
336			down	0.04300	*Blvrb*	Biliverdin reductase B (flavin reductase (NADPH))	7	AI843182	AI843182
337			down	0.00130	*C030002B11Rik*	RIKEN cDNA C030002B11 gene		H3097G02	BG071345.1
338			down	0.00003	*Mllt1*	Myeloid/lymphoid or mixed lineage-leukemia translocation to 1 homolog (Drosophila)	17	387728	AI604771
339			down	0.00103	*Daam1*	Dishevelled associated activator of morphogenesis 1	12	659710	AI464307
340			down	0.00574	*Ubl4*	Ubiquitin-like 4	X	H4021D12	BQ553350.1
341			down	0.00010	*Centg2*	Centaurin, gamma 2	1	642365	AI451534
342			down	0.00037	*Dpp9*	Dipeptidylpeptidase 9	17	H4038F07	BQ556195.1
343			down	0.00277	*Hcfc1*	Host cell factor C1	7	AI836062	AI836062
344			down	0.00661	*Zfp1*	Zinc finger protein 1	8	AI854760	AI854760
345			down	0.00013	*Chd5*	Chromodomain helicase DNA binding protein 5	4	AI844094	AI844094
346			down	0.00206	*Sypl*	Synaptophysin-like protein	12	H3021B04	BG064554.2
347			down	0.03845	*Anapc11*	Anaphase promoting complex subunit 11 homolog (yeast)	11	AI850597	AI850597
348			down	0.00811	*Ywhag*	3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide	5	H3014H04	BG064043.2
349			down	0.00969	*Mpp5*	Membrane protein, palmitoylated 5 (MAGUK p55 subfamily member 5)	12	H4037E10	BQ556012.1
350			down	0.01150	*Cbxn1*	Cortexin 1	8	H4034F01	BQ555529.1
351			down	0.01454	*Ebag9*	Estrogen receptor-binding fragment-associated gene 9	15	AI835379	AI835379
352			down	0.03565	*Sh3glb1*	SH3-domain GRB2-like B1 (endophilin)	3	H3086A04	BG070314.2
353			down	0.01569	*Dscr1*	Down syndrome critical region homolog 1 (human)	16	AI846152	AI846152
354			down	0.04649	*Mical2*	Microtubule associated monoxygenase, calponin and LIM domain containing 2	7	H4038F08	BQ556196.1
355			down	0.00602	*4930556P03Rik*	RIKEN cDNA 4930556P03 gene	15	658268	AI465223
356			down	0.00104	*Socs4*	Suppressor of cytokine signaling 4	14	H3106B03	BG072049.2
357			down	0.03469	*Mypn*	Myopalladin	10	AI853556	AI853556
358	C57Bl/6J		up	0.00202	*Arl6*	ADP-ribosylation factor-like 6	16	476610	AI326814
359			up	0.04258	*Gse1*	Genetic suppressor element 1	8	AI854638	AI854638
360			up	0.00034	*Ppp4r2*	protein phosphatase 4, regulatory subunit 2	6	574655	AI448654

Table A.2: Alcohol-related strain differences in gene expression.

	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
361			up	0.00306	*Plac9*	Placenta specific 9	14	AI852603	AI852603
362			up	0.02560	*Arpp21*	Cyclic AMP-regulated phosphoprotein, 21	9	718431	AI666401
363			up	0.03335	*2810417M05Rik*	RIKEN cDNA 2810417M05 gene	17	AI851143	AI851143
364			up	0.02126	*Jarid2*	Jumonji, AT rich interactive domain 2	13	H3045G12	BG066599.2
365			down	0.01054	*Scamp2*	Secretory carrier membrane protein 2	9	H3142G07	BG075058.2
366			down	0.00997	*Pnkd*	Paroxysmal nonkinesiogenic dyskinesia	1	H3152B02	BG075810.2
367			down	0.01210		EST		AI841830	AI841830
368			down	0.00556	*Sri*	Sorcin	5	H3010E05	BG063700.2
369			down	0.00039	*Gmps*	Guanine monphosphate synthetase	3	AI840562	AI840562
370			down	0.03454	*Ngly1*	N-glycanase 1	14	H4041C01	BQ556623.1
371			down	0.01675	*Ppt1*	Palmitoyl-protein thioesterase 1	4	H4076G09	BQ562591.1
372			down	0.00834	*Gnpda1*	Glucosamine-6-phosphate deaminase 1	18	H3023F09	BG064771.2
373			down	0.04378	*Nfkb2*	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 2, p49/p100	19	AI843566	AI843566
374			down	0.02231	*Stmn3*	Stathmin-like 3	2	H3033G01	BG065654.2
375			down	0.00487	*Ptprg*	Protein tyrosine phosphatase, receptor type, G	14	AI836685	AI836685
376			down	0.02866	*6-Mar*	Membrane-associated ring finger (C3HC4) 6	15	H4014C12	BQ552232.1
377			down	0.00602	*Ubadc1*	Ubiquitin associated domain containing 1	2	AI840235	AI840235
378			down	0.03814	*Hpcal4*	Hippocalcin-like 4	4	AI846570	AI846570
379			down	0.01725	*Tex261*	Testis expressed gene 261	6	H3007E03	BG063459.2
380			down	0.01142	*Ttc3*	Tetratricopeptide repeat domain 3	16	H3071D09	BG082005.2
381			down	0.03150	*2300008B03Rik*	RIKEN cDNA 2300008B03 gene	1	H4068H02	BQ561194.1
382			down	0.00383	*Orc5l*	Origin recognition complex, subunit 5-like (S. cerevisiae)	5	H3098G07	BG071423.2
383			down	0.00567	*Coasy*	Coenzyme A synthase	11	AI851009	AI851009
384			down	0.00499	*Rasip1*	Ras interacting protein 1	7	AI853551	AI853551
385			down	0.00388	*Stmn2*	Stathmin-like 2	3	H4039C08	BQ556296.1
386			down	0.00022	*Cryab*	Crystallin, alpha B	9	H3143B04	BG075084.2
387	DBA/2J		up	0.03652	*Islr2*	Immunoglobulin superfamily containing leucine-rich repeat 2	9	385503	AI413097
388			up	0.01382	*Zfp36*	Zinc finger protein 36	7	338091	AI893411
389			up	0.00566	*Ccdc86*	Coiled-coil domain containing 86	19	H3047D04	BG066826.2
390			up	0.00887				AI838368	AI838368

Table A.2: Alcohol-related strain differences in gene expression.

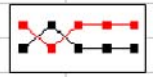
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
391			up	0.00749	*5830446M03Rik*	RIKEN cDNA 5830446M03 gene	6	H3027C11	BG065097.2
392			up	0.02945	*Psmd1*	Proteasome (prosome, macropain) 26S subunit, non-ATPase, 1	1	H3031D07	BG065461.2
393			up	0.00120				H3100G04	
394			up	0.00253	*C79267*	Expressed sequence C79267	4	H3019D12	CK334360.1
395			up	0.00062	*Kif18a*	Kinesin family member 18A	2	H4076F05	BQ562552.1
396			up	0.04913	*Fip1l1*	FIP1 like 1 (S. cerevisiae)	5	AI841692	AI841692
397			up	0.00933	*4632411B12Rik*	RIKEN cDNA 4632411B12 gene	1	AI841107	AI841107
398			up	0.00242	*Hba-a1*	Hemoglobin alpha, adult chain 1	11	H3140G04	BG074904.2
399			up	0.00031	*Vangl1*	Vang, van gogh-like 1 (Drosophila)	3	303868	AI893759
400			up	0.02075	*Tbp*	TATA box binding protein	17	H3026B01	BG064985.2
401			up	0.00147	*Insig1*	Insulin induced gene 1	5	AI848986	AI848986
402			up	0.02645	*Hsp90aa1*	Heat shock protein 90kDa alpha (cytosolic), class A member 1	11	H3009C07	BG063605.2
403			up	0.00215	*Narg2*	NMDA receptor-regulated gene 2	9	H3063E05	BG068262.2
404			up	0.00504	*Cyp51*	Cytochrome P450, family 51	5	H4065G08	BQ560678.1
405			up	0.04420	*2210403B10Rik*	RIKEN cDNA 2210403B10 gene	11	573949	AI448488
406			up	0.00067	*Klf2*	Kruppel-like factor 2 (lung)	8	642661	AI451567
407			up	0.00339	*Cenpq*	Centromere protein Q	17	H3138D04	BG074710.2
408			up	0.00250	*Sncg*	Synuclein, gamma	14	AI852799	AI852799
409			down	0.00004	*2900093B09Rik*	RIKEN cDNA 2900093B09 gene	7	H4035C03	BQ555635.1
410			down	0.00812	*Fech*	Ferrochelatase	18	H3124D02	BG086309.2
411			down	0.02149	*4930504E06Rik*	RIKEN cDNA 4930504E06 gene	3	H3109D02	BG072333.2
412			down	0.01872	*Bbs4*	Bardet-Biedl syndrome 4 homolog (human)	9	AI850809	AI850809
413			down	0.00073	*Tax1bp1*	Tax1 (human T-cell leukemia virus type I) binding protein 1		717773	AI661905
414			down	0.01582	*Clk1*	CDC-like kinase 1	1	H3027D01	BG065099.2
415			down	0.00966	*6530406A20Rik*	RIKEN cDNA 6530406A20 gene	12	H4042A08	BQ556764.1
416			down	0.00189	*Flnb*	Filamin, beta	14	H3021B06	BG064556.2
417			down	0.00001	*Id3*	Inhibitor of DNA binding 3	4	AI839283	AI839283
418			down	0.04721				H4023F04	
419			down	0.00017	*Ube4a*	Ubiquitination factor E4A, UFD2 homolog (S. cerevisiae)	9	H4052E12	BQ558502.1
420			down	0.00000				H4029A06	

Table A.2: Alcohol-related strain differences in gene expression.

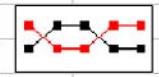
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
421			down	0.00578	*Pldn*	Pallidin	2	H4043A05	BQ556926.1
422			down	0.00043	*Sfrs7*	Splicing factor, arginine/serine-rich 7	17	AI836707	AI836707
423			down	0.01419				H3086E08	
424	DBA/2J		up	0.02101	*Ppil2*	Peptidylprolyl isomerase (cyclophilin)-like 2	16	H3024A01	BG064798.2
425			up	0.03399	*Limd2*	LIM domain containing 2	11	H3078A06	BG069592.2
426			up	0.01281	*Cul3*	Cullin 3	1	AI840051	AI840051
427			up	0.00345	*Rdx*	Radixin	9	AI851579	AI851579
428			up	0.00811	*Sp3*	Trans-acting transcription factor 3	2	H3083C05	BG069978.2
429			up	0.00009	*Mgea5*	Meningioma expressed antigen 5 (hyaluronidase)	19	H3004C04	BG063211.2
430			up	0.00173	*Kcnq1*	Potassium voltage-gated channel, subfamily Q, member 1	7	H3093G10	BG071018.2
431			up	0.00403	*Lars*	Leucyl-tRNA synthetase	18	H3026C04	BG064998.2
432			up	0.00296	*Tsc22d1*	TSC22 domain family, member 1	14	H3066H08	BG068563.2
433			up	0.00002	*Bteb1*	Basic transcription element binding protein 1	19	AI848050	AI848050
434			up	0.00000				H3154H07	
435			up	0.04002	*Rora*	RAR-related orphan receptor alpha	9	H4066C07	BQ560752.1
436			up	0.04130	*Ncl*	Nucleolin	1	H3025E03	BG064932.1
437			up	0.00047	*Rab31*	RAB31, member RAS oncogene family	17	614369	AI448835
438			up	0.00049	*Ssr3*	Signal sequence receptor, gamma	3	AI835359	AI835359
439			up	0.00208	*Camk2b*	Calcium/calmodulin-dependent protein kinase II, beta	11	AI842756	AI842756
440			up	0.03026	*D430028G21Rik*	RIKEN cDNA D430028G21 gene	2	H4042C08	BQ556806.1
441			up	0.00051	*Nfkbia*	Nuclear factor of kappa light chain gene enhancer in B-cells inhibitor, alpha	12	H3026A08	BG064981.2
442			up	0.00289	*MGC40675*	hypothetical protein MGC40675	7	AI854879	AI854879
443			up	0.00463	*Cd9*	CD9 antigen	6	AI854515	AI854515
444			up	0.03501	*Hel308*	Helicase, mus308-like (Drosophila)	5	H3060H11	BG068026.2
445			up	0.00496	*Hmgn2*	High mobility group nucleosomal binding domain 2	4	AI836129	AI836129
446			up	0.00170				H3053C11	
447			up	0.00001	*Pnpla2*	Patatin-like phospholipase domain containing 2	7	H3017F03	BG064267.2
448			up	0.00000	*Tsc22d3*	TSC22 domain family 3	X	476319	AI326808
449			down	0.01039	*Csk*	C-src tyrosine kinase	9	H3012F10	BG063779.2
450			down	0.02622	*Xab1*	XPA binding protein 1	5	636618	AI450139

Table A.2: Alcohol-related strain differences in gene expression.

	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
451			down	0.00070	*Dnm2*	Dynamin 2	9	424653	AI327127
452			down	0.00250	*H2afy2*	H2A histone family, member Y2	10	478045	AI430911
453			down	0.00413	*Gna-rs1*	Guanine nucleotide binding protein, related sequence 1	17	H4042C02	BQ556796.1
454			down	0.03706	*Pcm1*	Pericentriolar material 1	8	H4075G03	BQ562397.1
455			down	0.00494	*Foxp1*	Forkhead box P1	6	AI851964	AI851964
456			down	0.00050	*Fdps*	Farnesyl diphosphate synthetase	3	440407	AI324075
457			down	0.00032	*Cml1*	Camello-like 1	6	722321	AI661237
458			down	0.04311	*Dcps*	Decapping enzyme, scavenger	9	H3004F06	BG063230.1
459			down	0.01701	*2410024A21Rik*	RIKEN cDNA 2410024A21 gene	12	AI835528	AI835528
460			down	0.00007	*1500010G04Rik*	RIKEN cDNA 1500010G04 gene	6	H3072D10	BG069073.2
461			down	0.00014	*Sfrs6*	Splicing factor, arginine/serine-rich 6	2	AI846595	AI846595
462	DBA/2J		up	0.00048				H3056H08	
463			up	0.00229				H3056D05	
464			up	0.00347	*Idb4*	Inhibitor of DNA binding 4	13	AI838339	AI838339
465			up	0.00487				H3037B07	
466			up	0.00026	*Slc2a1*	Solute carrier family 2 (facilitated glucose transporter), member 1	4	H3015G08	BG064126.2
467			up	0.00085	*Ccrn4l*	CCR4 carbon catabolite repression 4-like (S. cerevisiae)	3	AI848832	AI848832
468			up	0.00422	*Mtmr7*	Myotubularin related protein 7	8	AI844209	AI844209
469			up	0.00280	*Errfi1*	ERBB receptor feedback inhibitor 1	4	H3017B12	BG064235.2
470			up	0.00505				H3067H02	
471			up	0.03397				H3013B05	
472			up	0.00032	*App*	Amyloid beta (A4) precursor protein	16	AI843768	AI843768
473			up	0.00707	*Fkbp5*	FK506 binding protein 5	17	H3138G12	BG087373.2
474			up	0.03640	*Pclo*	Piccolo (presynaptic cytomatrix protein)	5	AI852175	AI852175
475			up	0.00667	*Cnot6*	CCR4-NOT transcription complex, subunit 6	11	AI853946	AI853946
476			up	0.01881				H3024H11	
477			up	0.00961	*Gm996*	Gene model 996, (NCBI)	2	AI846620	AI846620
478			up	0.01468	*Sgcb*	Sarcoglycan, beta (dystrophin-associated glycoprotein)	5	AI844132	AI844132
479			up	0.04499	*Gramd1b*	GRAM domain containing 1B	9	AI844516	AI844516
480			up	0.00411	*Asgr1*	Asialoglycoprotein receptor 1	11	AI845829	AI845829

Table A.2: Alcohol-related strain differences in gene expression.

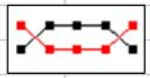
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
481			up	0.04036	*Spg3a*	Spastic paraplegia 3A homolog (human)	12	H3068F02	BG068719.2
482			up	0.00669	*Shank3*	SH3/ankyrin domain gene 3	15	H3069C04	BG068781.2
483			up	0.04094	*Setdb2*	SET domain, bifurcated 2	14	H4073H05	BQ562058.1
484			up	0.00760	*AW112037*	Expressed sequence AW112037	1	H3104A09	BG071867.2
485			down	0.00003	*Rpl15*	Ribosomal protein L15	14	AI843598	AI843598
486			down	0.02135	*Cox18*	COX18 cytochrome c oxidase assembly homolog (S. cerevisiae)	5	H4030B04	BQ554756.1
487			down	0.00099	*R3hdm2*	R3H domain containing 2	10	AI838049	AI838049
488			down	0.00686	*Pogz*	Pogo transposable element with ZNF domain	3	H3087B02	BG070415.2
489			down	0.00372	*AW536275*	Expressed sequence AW536275	2	AI846248	AI846248
490			down	0.01668	*Map2k2*	Mitogen activated protein kinase kinase 2	10	AI836622	AI836622
491			down	0.00163	*Mybbp1a*	MYB binding protein (P160) 1a	11	H3008A10	BG063508.2
492			down	0.02586	*Grm5*	Glutamate receptor, metabotropic 5	7	AI850523	AI850523
493			down	0.00004	*Acsl4*	Acyl-CoA synthetase long-chain family member 4	X	524762	AI426379
494			down	0.02292	*Atp6v0a2*	ATPase, H+ transporting, lysosomal V0 subunit A2	5	AI848884	AI848884
495			down	0.03699	*6720463E02Rik*	RIKEN cDNA 6720463E02 gene	11	AI836322	AI836322
496			down	0.03135	*Thyn1*	Thymocyte nuclear protein 1	9	AI848392	AI848392
497			down	0.00988	*Sema5a*	Sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5A	15	H3070G08	BG068925.2
498			down	0.03415	*9430020K01Rik*	RIKEN cDNA 9430020K01 gene	18	AI838016	AI838016
499			down	0.01135	*Cacybp*	Calcyclin binding protein	1	H3079F05	BG069742.2
500			down	0.03930	*Sart3*	Squamous cell carcinoma antigen recognized by T-cells 3	5	AI846150	AI846150
501			down	0.01204	*D2Ertd750e*	DNA segment, Chr 2, ERATO Doi 750, expressed	2	575429	AI447547
502			down	0.03196	*1110001J03Rik*	RIKEN cDNA 1110001J03 gene	6	AI848967	AI848967
503			down	0.00978	*Ccdc72*	Coiled-coil domain containing 72	9	315756	AI414814
504			down	0.00209	*Med19*	Mediator of RNA polymerase II transcription, subunit 19 homolog (yeast)	2	H4061A03	BQ559897.1
505			down	0.00830	*Sars*	Seryl-aminoacyl-tRNA synthetase	3	AI841802	AI841802
506	DBA/2J		up	0.00013	*Tes*	Testis derived transcript	6	582953	AI325946
507			up	0.01414	*Slc6a8*	Solute carrier family 6 (neurotransmitter transporter, creatine), member 8	X	H3102C06	BG071707.2
508			up	0.02370	*2010305C02Rik*	RIKEN cDNA 2010305C02 gene	11	AI840174	AI840174
509			up	0.01380	*1110063F24Rik*	RIKEN cDNA 1110063F24 gene	4	AI843727	AI843727
510			up	0.00018	*Atp6ap1*	ATPase, H+ transporting, lysosomal accessory protein 1	X	AI853870	AI853870

Table A.2: Alcohol-related strain differences in gene expression.

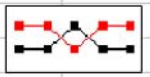
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
511			up	0.00715	*Odz2*	odd Oz/ten-m homolog 2 (Drosophila)	11	AI837170	AI837170
512			up	0.01453				517580	AI644471
513			up	0.00093	*Kctd10*	Potassium channel tetramerisation domain containing 10	5	H3056H01	BG067663.2
514			up	0.02094	*Nmt1*	N-myristoyltransferase 1	11	H3026D05	BG065011.2
515			up	0.02875	*Cpeb4*	Cytoplasmic polyadenylation element binding protein 4	11	AI845839	AI845839
516			up	0.01975	*Arhgap17*	Rho GTPase activating protein 17	7	H3100H08	BG084440.2
517			up	0.01352	*Dgka*	Diacylglycerol kinase, alpha	10	H3029G05	BG065321.2
518			up	0.00059	*Foxk2*	Forkhead box K2	11	AI846986	AI846986
519			up	0.01010	*Slc44a5*	Solute carrier family 44, member 5	3	539387	AI427060
520			down	0.03943	*Apcdd1*	Adenomatosis polyposis coli down-regulated 1	18	H3099A04	BG084307.2
521			down	0.04005				H3004F03	
522			down	0.02764				AI836069	AI836069
523			down	0.00044	*Gtf2h2*	General transcription factor II H, polypeptide 2	13	AI848115	AI848115
524			down	0.01997	*Zfp282*	Zinc finger protein 282	6	H3049G11	BG067034.2
525			down	0.00095	*Myst1*	MYST histone acetyltransferase 1	7	H3084C09	BG070156.2
526			down	0.04231	*Lphn1*	Latrophilin 1	8	AI854553	AI854553
527			down	0.00926	*Abi1*	Abelson helper integration site	10	H4075B03	BQ562282.1
528			down	0.02160	*AA407452*	EST AA407452	6	H4014F03	BQ552271.1
529			down	0.01897	*Sfxn1*	Sideroflexin 1	13	617904	AI450790
530			down	0.02975	*Mapk14*	Mitogen activated protein kinase 14	17	H3140A08	BG087465.2
531			down	0.01096				AI838428	AI838428
532			down	0.00849				H3064G01	
533			down	0.00074	*Kif21b*	Kinesin family member 21B	1	AI846241	AI846241
534			down	0.01357	*Pdcd11*	Programmed cell death protein 11	19	AI843772	AI843772
535			down	0.00437	*Lasp1*	LIM and SH3 protein 1	11	H3017H10	BG077548.2
536			down	0.02911	*LOC629986*	Similar to similar to 4933409K07Rik protein		H4034B09	BQ555453.1
537			down	0.02407	*Hivep2*	Human immunodeficiency virus type I enhancer binding protein 2	10	H4012F11	BQ551981.1
538			down	0.00696	*Gnb5*	Guanine nucleotide binding protein, beta 5	9	AI847689	AI847689
539	DBA/2J		up	0.04237	*4732435N03Rik*	RIKEN cDNA 4732435N03 gene	8	H3086F08	BG070374.2
540			up	0.04048	*Slc15a2*	Solute carrier family 15 (H+/peptide transporter), member 2	16	H3073B03	BG069134.2

Table A.2: Alcohol-related strain differences in gene expression.

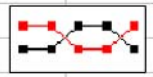
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
541			up	0.00119	*Sepp1*	Selenoprotein P, plasma, 1	15	AI838693	AI838693
542			up	0.03483	*Syt9*	Synaptotagmin IX	7	AI842687	AI842687
543			up	0.00843		EST		AI842735	AI842735
544			up	0.00001	*1810011O16Rik*	RIKEN cDNA 1810011O16 gene	12	H3020D03	BG064492.2
545			up	0.00295	*1810012N18Rik*	RIKEN cDNA 1810012N18 gene	16	AI839212	AI839212
546			up	0.00163	*Cmtm3*	CKLF-like MARVEL transmembrane domain containing 3	8	AI841689	AI841689
547			up	0.00000	*Mt2*	Metallothionein 2	8	H3013D11	BG063925.2
548			up	0.00001	*Mt1*	Metallothionein 1	8	H3020C02	BG064480.2
549			down	0.04748	*Morf4l1*	Mortality factor 4 like 1	16	H3014G01	BG064029.2
550			down	0.04849	*Hspc171*	HSPC171 protein	8	AI848260	AI848260
551			down	0.02222	*Actn1*	Actinin, alpha 1	12	H3018E09	BG064350.2
552			down	0.00625	*Mirg*	MiRNA containing gene	12	AI852992	AI852992
553			down	0.01905	*0610042C05Rik*	RIKEN cDNA 0610042C05 gene	10	514366	AI427913
554			down	0.03959	*Slc25a18*	Solute carrier family 25 (mitochondrial carrier), member 18	6	AI839063	AI839063
555			down	0.00694	*Cdc2l5*	Cell division cycle 2-like 5 (cholinesterase-related cell division controller)	13	617527	AI450703
556			down	0.00904	*Slc34a1*	Solute carrier family 34 (sodium phosphate), member 1	13	661943	AI464529
557			down	0.00076	*Cttnbp2*	Cortactin binding protein 2	6	AI840625	AI840625
558			down	0.00671	*AI836003*	Expressed sequence AI836003	15	AI836003	AI836003
559			down	0.01360	*Hnrpc*	Heterogeneous nuclear ribonucleoprotein C	14	AI836825	AI836825
560			down	0.01246	*Ppap2b*	Phosphatidic acid phosphatase type 2B	4	AI847054	AI847054
561			down	0.00015	*Nanos1*	Nanos homolog 1 (Drosophila)	19	H3145F03	BG075296.1
562			down	0.00097	*Fkbp4*	FK506 binding protein 4	6	AI848455	AI848455
563			down	0.01996	*Jph4*	Junctophilin 4	14	H4030C10	BQ554788.1
564			down	0.00117	*Ptges3*	Prostaglandin E synthase 3 (cytosolic)		H3009H08	BG077044.2
565			down	0.00448	*Edg1*	Endothelial differentiation sphingolipid G-protein-coupled receptor 1	3	H4072B03	BQ561727.1
566			down	0.00000	*Tubb2c*	Tubulin, beta 2c	2	H3024D02	BG064832.2
567			down	0.00000	*Serpinh1*	Serine (or cysteine) peptidase inhibitor, clade H, member 1	7	475256	AI326777
568			down	0.00173	*Car2*	Carbonic anhydrase 2	3	AI841511	AI841511
569	DBA/2J		up	0.00048	*Cutl1*	Cut-like 1 (Drosophila)	5	H3003E01	BG063135.2
570			up	0.01885	*Nat12*	N-acetyltransferase 12	14	H3085E05	BG070269.2

Table A.2: Alcohol-related strain differences in gene expression.

	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
571			up	0.01310	*Slc12a7*	Solute carrier family 12, member 7	13	H3045B12	BG066547.2
572			up	0.00099	*Bag4*	BCL2-associated athanogene 4	8	H3042A05	BG066354.2
573			up	0.00209	*Dsp*	Desmoplakin	13	H4018H07	BQ552906.1
574			up	0.01720				H3089B07	BG070606.2
575			up	0.00912	*Lyar*	Ly1 antibody reactive clone	5	H3155B08	BG076061.2
576			up	0.03922	*Wdr73*	WD repeat domain 73	7	H3051C11	BG067164.2
577			up	0.04158	*LOC434130*	Similar to Probable coiled-coil domain containing protein 8 homolog	7	AI837167	AI837167
578			up	0.03202	*Fbxo21*	F-box only protein 21	5	H3158E10	BG076338.2
579			up	0.00389	*BC020077*	CDNA sequence BC020077		H4020D10	BQ553167.1
580			down	0.00144	*Prmt2*	Protein arginine N-methyltransferase 2	10	AI845257	AI845257
581			down	0.04264	*Col6a2*	Procollagen, type VI, alpha 2	10	H3152G04	BG075864.2
582			down	0.00161	*Ttc3*	Tetratricopeptide repeat domain 3	16	H3071D09	BG082005.2
583			down	0.02732	*Btg4*	B-cell translocation gene 4	9	H3073E04	BG069170.2
584			down	0.00698	*2310044H10Rik*	RIKEN cDNA 2310044H10 gene	7	AI849984	AI849984
585			down	0.00080				H3098D07	
586			down	0.02949	*9030425E11Rik*	RIKEN cDNA 9030425E11 gene	9	H3154B11	BG075988.2
587			down	0.01416	*0610009D07Rik*	RIKEN cDNA 0610009D07 gene	12	H3159G04	BG076503.2
588			down	0.01847	*Lrrfip1*	Leucine rich repeat (in FLII) interacting protein 1	1	H3072C08	BG069059.2
589			down	0.00658	*Tead1*	TEA domain family member 1	7	H3159B06	BG076381.2
590			down	0.00436	*Sdf4*	Stromal cell derived factor 4	4	H3022F01	BG064675.1
591			down	0.03908	*Prkcq*	Protein kinase C, theta	2	582973	AI528522
592			down	0.04161	*Txnl1*	Thioredoxin-like 1	18	AI837780	AI837780
593	DBA/2J		up	0.00000	*6230416A05Rik*	RIKEN cDNA 6230416A05 gene	13	H3060F03	BG067998.2
594			up	0.00008	*Taf11*	TAF11 RNA polymerase II, TATA box binding protein (TBP)-associated factor	17	AI845542	AI845542
595			up	0.00062	*Fhl1*	Four and a half LIM domains 1	X	H3141F12	BG074978.2
596			up	0.00256	*Gnb1*	Guanine nucleotide binding protein, beta 1	4	H3094D02	BG071068.2
597			up	0.02315	*Rsad2*	Radical S-adenosyl methionine domain containing 2	12	H4008A06	BQ551218.1
598			up	0.02988	*Gpm6a*	Glycoprotein m6a	8	H4045G08	BQ557384.1
599			up	0.04694	*Rae1*	RAE1 RNA export 1 homolog (S. pombe)	2	576512	AI327368
600			up	0.04473	*St6gal1*	Beta galactoside alpha 2,6 sialyltransferase 1	16	H3152A02	BG075800.2

Table A.2: Alcohol-related strain differences in gene expression.

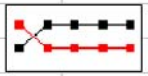
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
601			up	0.01645	*Lrrc8e*	Leucine rich repeat containing 8 family, member E	8	H3057G04	BG067748.2
602			up	0.02501	*Hrsp12*	Heat-responsive protein 12	15	AI850467	AI850467
603			up	0.00401	*Dusp9*	Dual specificity phosphatase 9	X	H3002F03	BG062971.2
604			up	0.01098	*Yif1b*	Yip1 interacting factor homolog B (S. cerevisiae)	7	AI854242	AI854242
605			up	0.01998	*1190007F08Rik*	RIKEN cDNA 1190007F08 gene	4	H3158D01	BG076322.2
606			up	0.00466	*Gm129*	Gene model 129, (NCBI)	3	AI843058	AI843058
607			up	0.02525	*MGC107702*	Similar to proteasome (prosome, macropain) subunit, beta type 7	X	651384	AI465359
608			up	0.00362		Transcribed sequences		599032	AI447691
609			up	0.00843	*D930020B18Rik*	RIKEN cDNA D930020B18 gene	10	H3147H09	BG075494.2
610			up	0.04883				H3064B12	
611			up	0.00389				H3090F05	BG070739.2
612			up	0.02745	*Tcf4*	Transcription factor 4	18	H4037F08	BQ556031.1
613			up	0.02420	*Pcmtd1*	Protein-L-isoaspartate (D-aspartate) O-methyltransferase domain containing 1	1	H4005B12	BQ550744.1
614			up	0.00625	*Scamp1*	Secretory carrier membrane protein 1	13	AI851575	AI851575
615			up	0.02831	*Thrb*	Thyroid hormone receptor beta	14	AI852182	AI852182
616			up	0.03571	*Mier2*	Mesoderm induction early response 1, family member 2	10	AI850156	AI850156
617			up	0.02871				H3146H08	
618			up	0.03468		OCIA domain containing 2	5	337767	AI415308
619			up	0.00557	*Adamtsl5*	ADAMTS-like 5	10	AI852777	AI852777
620			up	0.03318	*Ube2a*	Ubiquitin-conjugating enzyme E2A, RAD6 homolog (S. cerevisiae)	X	H3046G09	BG066776.2
621			up	0.01650	*1700025G04Rik*	RIKEN cDNA 1700025G04 gene	1	H3153B01	BG075895.2
622			up	0.02752	*6330505F04Rik*	RIKEN cDNA 6330505F04 gene	X	642750	AI451578
623			up	0.02778	*Epsti1*	Epithelial stromal interaction 1 (breast)	14	H4078G08	BQ562930.1
624			down	0.04030	*Ehmt2*	Euchromatic histone lysine N-methyltransferase 2	17	H3150B12	BG088216.2
625			down	0.02694	*Ubr2*	Ubiquitin protein ligase E3 component n-recognin 2	17	H3079H03	BG069763.2
626			down	0.01127	*Mapre2*	Microtubule-associated protein, RP/EB family, member 2	18	AI836597	AI836597
627			down	0.03509	*Coq6*	Coenzyme Q6 homolog (yeast)	12	H4021B09	BQ553299.1
628			down	0.00187	*Tardbp*	TAR DNA binding protein	4	H4052F09	BQ558518.1
629			down	0.00005				H3060F05	
630			down	0.00194	*Klf13*	Kruppel-like factor 13	7	AI851693	AI851693

Table A.2: Alcohol-related strain differences in gene expression.

	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
631			down	0.00059	*Las1l*	LAS1-like (S. cerevisiae)	X	H3153C01	BG075905.2
632			down	0.01345	*Igfbp5*	Insulin-like growth factor binding protein 5	1	AI843940	AI843940
633			down	0.00160	*Amn*	Amnionless	12	AI853281	AI853281
634			down	0.00257	*Vdp*	Vesicle docking protein	5	AI850786	AI850786
635			down	0.00181	*Txndc10*	Thioredoxin domain containing 10	18	H3137E01	BG087261.2
636			down	0.01330	*Rai17*	Retinoic acid induced 17	14	AI845939	AI845939
637			down	0.02383	*Peg3*	Paternally expressed 3	7	H3156C02	BG076150.2
638			down	0.02056	*0610011D08Rik*	RIKEN cDNA 0610011D08 gene	4	AI844568	AI844568
639			down	0.00373	*Bnip3l*	BCL2/adenovirus E1B interacting protein 3-like	14	H3103B07	BG071794.2
640			down	0.02767	*2610020H08Rik*	RIKEN cDNA 2610020H08 gene	7	H4053F06	BQ558674.1
641			down	0.01057	*Cx3cl1*	Chemokine (C-X3-C motif) ligand 1	8	AI848747	AI848747
642			down	0.04944	*Cacna1i*	Calcium channel, voltage-dependent, alpha 1I subunit	15	H4030F01	BQ554836.1
643			down	0.00402	*Ascc3*	Activating signal cointegrator 1 complex subunit 3	10	H4078H08	BQ562950.1
644			down	0.00315		EST		AI854792	AI854792
645			down	0.01880	*Phyhipl*	Phytanoyl-CoA hydroxylase interacting protein-like	10	H4064B03	BQ560403.1
646			down	0.00520	*Mllt1*	Myeloid/lymphoid or mixed lineage-leukemia translocation to 1 homolog (Drosophila)	17	H3001E03	BG063049.2
647			down	0.00107	*Nfyb*	Nuclear transcription factor-Y beta	10	H3146B03	BG075341.2
648			down	0.00051	*Rab9*	RAB9, member RAS oncogene family	X	H3102H06	BG071759.2
649			down	0.03825				H4067D07	BQ560942.1
650			down	0.02557	*Eif4enif1*	Eukaryotic translation initiation factor 4E nuclear import factor 1	11	H3035H06	BG065845.2
651			down	0.01387				H3080B12	
652			down	0.00213	*D130029J02Rik*	RIKEN cDNA D130029J02 gene	8	AI850250	AI850250
653			down	0.00708	*Pum2*	Pumilio 2 (Drosophila)	12	H4075F08	BQ562384.1
654			down	0.01929	*LOC666825*	Hypothetical protein LOC666825	12	H3079G05	BG069753.2
655			down	0.04927	*C330019G07Rik*	RIKEN cDNA C330019G07 gene	5	H4069B11	BQ561249.1
656			down	0.00260	*Gm740*	Gene model 740, (NCBI)	11	H3090B06	BG070698.2
657			down	0.01855	*Rcc2*	Regulator of chromosome condensation 2	4	AI849056	AI849056
658			down	0.01066	*Rgl1*	Ral guanine nucleotide dissociation stimulator,-like 1	1	H4040G07	BQ556549.1
659			down	0.03817	*Slc12a5*	Solute carrier family 12, member 5	2	H4032E08	BQ555168.1
660			down	0.02718	*Zfhx1a*	Zinc finger homeobox 1a	18	H4067F04	BQ560984.1

Table A.2: Alcohol-related strain differences in gene expression.

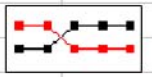
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
661			down	0.02274	*Ube2j2*	Ubiquitin-conjugating enzyme E2, J2 homolog (yeast)	4	576930	AI449455
662			down	0.00450	*BC013481*	CDNA sequence BC013481	5	H3036A10	BG065860.1
663			down	0.04919	*Sppl3*	Signal peptide peptidase 3	5	AI849165	AI849165
664	DBA/2J		up	0.02473	*D14Ertd231e*	DNA segment, Chr 14, ERATO Doi 231, expressed	14	H3009E01	BG076982.2
665			up	0.04072	*Upp1*	Uridine phosphorylase 1	11	H3027E05	BG065114.2
666			up	0.00285				AI842916	AI842916
667			up	0.00361	*D030007L05Rik*	RIKEN cDNA D030007L05 gene	13	H3065C12	BG068423.2
668			up	0.00220	*Rev3l*	REV3-like, catalytic subunit of DNA polymerase zeta RAD54 like (S. cerevisiae)	10	AI846010	AI846010
669			up	0.04238	*Idh2*	Isocitrate dehydrogenase 2 (NADP+), mitochondrial	7	H3062E02	BG081213.2
670			up	0.02888	*Pdk2*	Pyruvate dehydrogenase kinase, isoenzyme 2	11	AI848783	AI848783
671			up	0.00693	*C630002B14Rik*	RIKEN cDNA C630002B14 gene	10	H4040D05	BQ556479.1
672			up	0.03760	*Auh*	AU RNA binding protein/enoyl-coenzyme A hydratase	13	H4061G12	BQ560021.1
673			up	0.02491	*Plxna3*	Plexin A3	X	AI852941	AI852941
674			up	0.00005	*Lrba*	LPS-responsive beige-like anchor	3	621451	AI451097
675			up	0.02390	*Adora1*	Adenosine A1 receptor	1	AI848715	AI848715
676			up	0.00349	*BC055107*	CDNA sequence BC055107	14	AI836637	AI836637
677			up	0.01421				H3083E08	
678			up	0.00607	*Eml2*	Echinoderm microtubule associated protein like 2	7	H3073G09	BG069197.2
679			up	0.00477	*Sepx1*	Selenoprotein X 1	17	AI840037	AI840037
680			up	0.04123	*Ppil3*	Peptidylprolyl isomerase (cyclophilin)-like 3	1	H3016C01	BG077504.2
681			up	0.00000	*BQ952480*	Expressed sequence BQ952480	6	AI844065	AI844065
682			down	0.02287	*Pnkp*	Polynucleotide kinase 3'- phosphatase	7	H3053G12	BG067384.2
683			down	0.01340	*Ppp4r1*	Protein phosphatase 4, regulatory subunit 1	17	H3022D11	BG064662.2
684			down	0.00931	*Il6st*	Interleukin 6 signal transducer	13	H3086G09	BG070387.2
685			down	0.04153	*Igf2bp3*	Insulin-like growth factor 2 mRNA binding protein 3	6	AI853878	AI853878
686			down	0.00001	*Hras1*	Harvey rat sarcoma virus oncogene 1	7	H3091H05	BG070933.2
687			down	0.00528	*Tjp2*	Tight junction protein 2	19	AI843238	AI843238
688			down	0.02862	*Angpt1*	Angiopoietin 1	15	H4074C09	BQ562133.1
689			down	0.02591	*Ung*	Uracil-DNA glycosylase	5	H3090B05	BG070697.2
690			down	0.00417	*Setdb1*	SET domain, bifurcated 1	3	H3062C11	BG068152.2

Table A.2: Alcohol-related strain differences in gene expression.

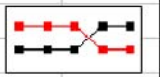
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
691			down	0.03551	*Osgep*	O-sialoglycoprotein endopeptidase	14	AI840215	AI840215
692			down	0.00652	*Ccdc64*	Coiled-coil domain containing 64	5	H3129B02	BG073949.2
693			down	0.01452	*Sertad1*	SERTA domain containing 1	7	AI847051	AI847051
694			down	0.00560	*Sf3a3*	Splicing factor 3a, subunit 3	4	H3032D12	BG065547.2
695			down	0.00232	*Itgb8*	Integrin beta 8	12	539225	AI427072
696	DBA/2J		up	0.00120	*D030074E01Rik*	RIKEN cDNA D030074E01 gene	18	H3071C07	BG068969.2
697			up	0.00075	*Aurkaip1*	Aurora kinase A interacting protein 1	4	H3045H03	BG066690.2
698			up	0.00629	*Fhl2*	Four and a half LIM domains 2	1	459229	AI323335
699			up	0.00817	*E130309D14Rik*	RIKEN cDNA E130309D14 gene	11	H3079A12	BG069691.2
700			up	0.03390	*Sf3b5*	Splicing factor 3b, subunit 5	10	H3134G08	BG074410.2
701			up	0.00488	*Mrps31*	Mitochondrial ribosomal protein S31	8	H4054C03	BQ558773.1
702			up	0.00452	*Sgcg*	Sarcoglycan, gamma (dystrophin-associated glycoprotein)	14	H4023F12	BQ553691.1
703			down	0.00558	*Olig1*	Oligodendrocyte transcription factor 1	16	AI836478	AI836478
704			down	0.01127	*2310014H01Rik*	RIKEN cDNA 2310014H01 gene	17	H3003H06	BG063171.2
705			down	0.03514				H3068F10	
706			down	0.00706	*Acly*	ATP citrate lyase	11	AI841630	AI841630
707			down	0.00967	*Ddx55*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 55	5	H3060B06	BG067958.2
708			down	0.04512	*Mphosph6*	M phase phosphoprotein 6	8	AI853569	AI853569
709			down	0.01053	*Prrt1*	Proline-rich transmembrane protein 1	17	AI849130	AI849130
710			down	0.00508	*Rnf185*	Ring finger protein 185	11	AI848225	AI848225
711			down	0.00266	*Suclg1*	Succinate-CoA ligase, GDP-forming, alpha subunit	6	AI852717	AI852717
712			down	0.04498	*Samd14*	Sterile alpha motif domain containing 14	11	AI839049	AI839049
713			down	0.01713	*8430427H17Rik*	RIKEN cDNA 8430427H17 gene	2	H3137C04	BG074615.2
714			down	0.01194	*Fbxl13*	F-box and leucine-rich repeat protein 13	5	H3094A12	BG071043.2
715			down	0.00231	*Daam1*	Dishevelled associated activator of morphogenesis 1	12	659710	AI464307
716			down	0.04588	*Bmp2k*	BMP2 inducible kinase	5	H4023B09	BQ553611.1
717			down	0.00074	*AI504432*	Expressed sequence AI504432	3	575326	AI323624
718			down	0.00084	*Idi1*	Isopentenyl-diphosphate delta isomerase	13	H3046C10	BG066732.2
719			down	0.00099	*Nfib*	Nuclear factor I/B	4	H3069A09	BG068762.2
720			down	0.00524	*AI849156*	Expressed sequence AI849156	8	AI849156	AI849156

Table A.2: Alcohol-related strain differences in gene expression.

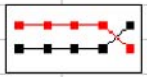
	Strain	PLS pattern	Ethanol / Control	F-test p value	Gene Symbol	Gene Name	Chromosome	Clone ID	Accession number
721			down	0.00176	*Tmem132b*	Transmembrane protein 132B	5	AI854797	AI854797
722	DBA/2J		up	0.03346	*Gga1*	Golgi associated, gamma adaptin ear containing, ARF binding protein 1	15	AI841123	AI841123
723			up	0.00042	*Cxcl14*	Chemokine (C-X-C motif) ligand 14	13	334419	AI414372
724			up	0.00987	*Dst*	Dystonin	1	AI841793	AI841793
725			up	0.00143	*Dcun1d1*	DCUN1D1 DCN1, defective in cullin neddylation 1, domain containing 1 (S. cerevisiae)	13	H4076D06	BQ562510.1
726			up	0.00280	*Bbs2*	Bardet-Biedl syndrome 2 homolog (human)	8	349934	AI413220
727			up	0.00241	*Ylpm1*	YLP motif containing 1	12	AI851834	AI851834
728			up	0.00645	*D4Ertd429e*	DNA segment, Chr 4, ERATO Doi 429, expressed	4	AI854558	AI854558
729			up	0.00795	*Aff3*	AF4/FMR2 family, member 3	1	H4074C12	BQ562139.1
730			up	0.04009	*H19*	H19 fetal liver mRNA	7	H3133G06	BG074329.2
731			down	0.00005				H3126H02	
732			down	0.01364	*Sox1*	SRY-box containing gene 1	8	AI837754	AI837754
733			down	0.00885	*Glcci1*	Glucocorticoid induced transcript 1	6	H3082A03	BG070033.2
734			down	0.00405	*Cttn*	Cortactin	7	660123	AI464324
735			down	0.02009	*Abl1*	V-abl Abelson murine leukemia oncogene 1	2	H3019F05	BG064435.2
736			down	0.00861	*Gpd2*	Glycerol phosphate dehydrogenase 2, mitochondrial	2	AI849040	AI849040
737			down	0.00949	*Trak1*	Trafficking protein, kinesin binding 1	9	H4056H02	BQ559215.1
738			down	0.00685	*Ufd1l*	Ubiquitin fusion degradation 1 like	16	AI850551	AI850551
739			down	0.00658	*Ankhd1*	Ankyrin repeat and KH domain containing 1	18	H4018G12	BQ552892.1
740			down	0.00294	*Clint1*	Clathrin interactor 1	11	H3098A04	BG071368.2

Table A.3: Intoxication-related differentially expressed genes in B6 and D2 mice. The genes in this list correspond to Figure 4.2. Note that the genes names are current as of July 31, 2006.

Table A.3: Intoxication-related differentially expressed genes in B6 and D2 mice.

	Gene Symbol	Gene Name	t-test p value (B6 C vs B6 E-2hr)	Fold change (B6 E-2hr / B6 C)	t-test p value (D2 C vs D2 E-2hr)	Fold change (D2 E-2hr / D2 C)	Chromosome	Clone ID	Accession number
1			1.24E-02	5.83	3.52E-02	2.89		H3062B04	BG068133.2
2	*Pbx1*	Pre B-cell leukemia transcription factor 1	2.58E-03	5.12	3.92E-02	2.66	1	H4007G01	BQ551166.1
3	*Ppp2r1b*	Protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), beta isoform	2.51E-05	3.78	2.26E-02	1.49	9	H3051B08	BG067151.2
4	*Fos*	FBJ osteosarcoma oncogene	9.27E-04	2.72	2.58E-02	2.17	12	H3084G03	BG070196.2
5	*Junb*	Jun-B oncogene	6.11E-04	2.52	2.69E-02	1.93	8	AI850555	AI850555
6	*Dbndd1*	Dysbindin (dystrobrevin binding protein 1) domain containing 1	6.24E-03	2.51	1.57E-02	2.03	8	H3065H11	BG068474.2
7	*Nedd4l*	Neural precursor cell expressed, developmentally down-regulated gene 4-like	4.04E-02	2.47	1.94E-02	2.58	18	H4035F01	BQ555694.1
8	*Nfkbia*	Nuclear factor of kappa light chain gene enhancer in B-cells inhibitor, alpha	1.09E-02	2.24	9.35E-03	2.12	12	H3026A08	BG064981.2
9	*Cd28*	CD28 antigen	8.26E-03	2.2	4.22E-02	2.4	1	576501	AI327367
10	*Btg2*	B-cell translocation gene 2, anti-proliferative	2.53E-02	2.17	4.19E-02	1.48	1	H3059F09	BG067918.2
11	*Ptprn2*	Protein tyrosine phosphatase, receptor type, N polypeptide 2	7.62E-03	2.12	4.80E-02	1.66	12	AI854394	AI854394
12	*Slc35e4*	Solute carrier family 35, member E4	1.18E-02	2.11	6.71E-03	2.51	11	332704	AI414849
13	*Pfkp*	Phosphofructokinase, platelet	1.37E-02	2.08	1.46E-02	2.69	13	H4002A03	BQ550203.1
14	*Dut*	Deoxyuridine triphosphatase	2.55E-02	2.04	3.86E-02	3.05	2	H3074C12	BG069247.2
15	*Lpin1*	Lipin 1	1.22E-02	2.04	3.83E-02	1.83	12	H4039H03	BQ556391.1
16	*Errfi1*	ERBB receptor feedback inhibitor 1	2.83E-02	2.03	7.88E-03	1.85	4	H3017B12	BG064235.2
17	*Klf2*	Kruppel-like factor 2 (lung)	2.76E-02	1.95	8.84E-05	2.36	8	642661	AI451567
18	*5133400G04Rik*	RIKEN cDNA 5133400G04 gene	3.12E-02	1.95	2.35E-02	2.24	18	H3088B06	BG070511.2
19	*Dstn*	Destrin	2.26E-02	1.9	4.31E-02	1.26	2	579759	AI325494
20	*Myd116*	Myeloid differentiation primary response gene 116	8.37E-03	1.88	1.88E-02	2.15	7	598433	AI528708
21	*Mvd*	Mevalonate (diphospho) decarboxylase	1.49E-02	1.85	2.29E-02	1.93	8	479201	AI427478
22	*Mknk2*	MAP kinase-interacting serine/threonine kinase 2	3.95E-03	1.83	4.30E-02	1.56	10	AI845732	AI845732
23	*Gas2l3*	Growth arrest-specific 2 like 3	3.56E-02	1.8	1.68E-03	3.92	10	H3070C02	BG068872.2
24	*Nrxn2*	Neurexin II	1.68E-02	1.8	6.68E-03	1.78	19	AI849944	AI849944
25			7.64E-03	1.72	2.78E-03	2.75		H3100G04	
26	*Tsc22d3*	TSC22 domain family 3	2.91E-03	1.72	2.17E-04	2.21	X	476319	AI326808
27			2.55E-02	1.71	3.04E-04	2.01		H3056H08	
28	*Rab31*	RAB31, member RAS oncogene family	3.72E-03	1.7	4.00E-03	2.09	17	614369	AI448835
29	*2410015N17Rik*	RIKEN cDNA 2410015N17 gene	1.08E-02	1.67	2.60E-02	1.59	7	H3012D09	BG063758.2

Table A.3: Intoxication-related differentially expressed genes in B6 and D2 mice.

	Gene Symbol	Gene Name	t -test p value (B6 C vs B6 E-2hr)	Fold change (B6 E-2hr / B6 C)	t -test p value (D2 C vs D2 E-2hr)	Fold change (D2 E-2hr / D2 C)	Chromosome	Clone ID	Accession number
30	*Foxo3a*	Forkhead box O3a	1.18E-02	1.67	1.31E-02	1.54	10	AI849135	AI849135
31	*Insig1*	Insulin induced gene 1	4.11E-05	1.6	3.69E-03	1.67	5	AI848986	AI848986
32	*1500041N16Rik*	RIKEN cDNA 1500041N16 gene	3.61E-04	1.59	1.78E-02	1.57	8	AI842657	AI842657
33	*Hmgcl*	3-hydroxy-3-methylglutaryl-Coenzyme A lyase	6.24E-03	1.59	4.67E-02	1.26	4	575476	AI323628
34	*4921525H12Rik*	RIKEN cDNA 4921525H12 gene	4.59E-02	1.54	1.05E-02	1.65	3	H4065E03	BQ560628.1
35	*Cyp51*	Cytochrome P450, family 51	4.43E-03	1.5	3.66E-03	1.5	5	H4065G08	BQ560678.1
36	*Cmtm3*	CKLF-like MARVEL transmembrane domain containing 3	4.30E-02	1.49	3.95E-02	1.59	8	AI841689	AI841689
37			2.50E-02	1.47	3.01E-02	3.13		H3038D04	BG066054.2
38	*Mrps10*	Mitochondrial ribosomal protein S10	1.15E-02	1.47	3.27E-02	1.41	17	H3014A01	BG063968.1
39	*X83328*	EST X83328	2.77E-02	1.45	1.16E-03	1.82	11	H3009H03	BG063653.2
40	*Rhob*	Ras homolog gene family, member B	2.48E-03	1.45	4.51E-02	1.32	12	H3008A09	BG063507.2
41	*C79267*	Expressed sequence C79267	2.10E-02	1.44	2.05E-03	1.8	4	H3019D12	CK334360.1
42	*Lrrc40*	Leucine rich repeat containing 40	2.02E-02	1.4	5.83E-03	1.71	3	H3027F04	BG065125.2
43	*Glrx*	Glutaredoxin	1.31E-02	1.4	1.17E-02	1.47	13	H3030A12	BG065350.2
44	*Ccrn4l*	CCR4 carbon catabolite repression 4-like (S. cerevisiae)	4.76E-03	1.3	4.76E-03	1.47	3	AI848832	AI848832
45	*Slc44a5*	Solute carrier family 44, member 5	4.26E-02	1.27	2.12E-02	1.34	3	539387	AI427060
46	*2810055G22Rik*	RIKEN cDNA 2810055G22 gene	4.70E-02	1.25	2.08E-02	1.24	5	H4013D02	BQ552085.1
47	*Mt1*	Metallothionein 1	3.48E-03	1.23	8.63E-03	1.41	8	H3020C02	BG064480.2
48	*BC026585*	CDNA sequence BC026585	3.50E-02	1.19	1.71E-02	1.23	1	H3153G04	CK335223.1
49	*Ptp4a2*	Protein tyrosine phosphatase 4a2	4.66E-02	-1.25	3.89E-02	-1.28	4	AI838635	AI838635
50	*Pls3*	Plastin 3 (T-isoform)	2.21E-02	-1.29	4.95E-02	-1.27	X	AI849034	AI849034
51	*Naca*	Nascent polypeptide-associated complex alpha polypeptide	1.12E-02	-1.32	1.71E-02	-1.24	10	H3157F04	BG076267.2
52	*Tpk1*	Thiamin pyrophosphokinase	3.75E-02	-1.33	2.78E-02	-1.38	6	H3044C10	BG066630.2
53	*1500010G04Rik*	RIKEN cDNA 1500010G04 gene	2.14E-02	-1.35	2.24E-04	-1.56	6	H3072D10	BG069073.2
54	*Lemd2*	LEM domain containing 2	8.59E-03	-1.39	3.21E-02	-1.47	17	AI851024	AI851024
55			3.52E-02	-1.4	2.24E-04	-1.54		H4029A06	
56	*Hmgb4*	High-mobility group box 4	6.29E-03	-1.46	1.97E-03	-1.88	4	AI850936	AI850936
57			1.91E-02	-1.49	3.86E-02	-1.2		H3080A10	
58	*Akr1a4*	Aldo-keto reductase family 1, member A4 (aldehyde reductase)	2.77E-02	-1.49	3.85E-02	-1.38	4	H3145A08	BG075245.2

Table A.3: Intoxication-related differentially expressed genes in B6 and D2 mice.

	Gene Symbol	Gene Name	t-test p value (B6 C vs B6 E-2hr)	Fold change (B6 E-2hr / B6 C)	t-test p value (D2 C vs D2 E-2hr)	Fold change (D2 E-2hr / D2 C)	Chromosome	Clone ID	Accession number
59	*Myo1b*	Myosin IB	4.71E-02	-1.5	2.44E-03	-2.67	1	H3003B02	BG063103.2
60	*9430020K01Rik*	RIKEN cDNA 9430020K01 gene	2.81E-02	-1.52	1.19E-03	-1.3	18	AI838016	AI838016
61	*1700112E06Rik*	RIKEN cDNA 1700112E06 gene	1.82E-02	-1.52	4.77E-02	-2.74	14	H3148E04	BG075542.2
62			1.94E-02	-1.53	4.23E-02	-1.6		H3041C11	
63	*Ryk*	Receptor-like tyrosine kinase	1.72E-02	-1.57	1.37E-02	-1.35	9	H3118F02	BG073091.2
64	*Eif2s3x*	Eukaryotic translation initiation factor 2, subunit 3, structural gene X-linked	2.77E-02	-1.6	1.42E-02	-1.99	X	H3132A12	BG074201.1
65	*Ppp3ca*	Protein phosphatase 3, catalytic subunit, alpha isoform	3.25E-02	-1.64	4.63E-02	-1.58	3	AI843781	AI843781
66	*1500001M20Rik*	RIKEN cDNA 1500001M20 gene	2.99E-02	-1.68	7.77E-03	-1.68	6	H3013C05	BG063913.2
67	*C130051G18Rik*	RIKEN cDNA C130051G18 gene	4.40E-02	-1.69	1.52E-02	-2.16	3	H3045F11	BG066586.2
68	*Id3*	Inhibitor of DNA binding 3	3.38E-04	-1.73	1.03E-04	-1.84	4	AI839283	AI839283
69	*Tcf4*	Transcription factor 4	3.40E-02	-1.74	2.36E-02	-1.5	18	H4055G08	BQ559044.1
70	*Auts2*	Autism susceptibility candidate 2	3.56E-02	-1.74	4.93E-03	-3.8	5	H3149C06	BG075604.2
71	*Pgs1*	Phosphatidylglycerophosphate synthase 1	2.20E-02	-1.76	3.29E-02	-1.68	11	H3022A04	BG064623.2
72	*Mak10*	MAK10 homolog, amino-acid N-acetyltransferase subunit, (S. cerevisiae)	2.19E-02	-1.77	2.19E-02	-3.71	13	H3091D12	BG070892.2
73	*Ehmt2*	Euchromatic histone lysine N-methyltransferase 2	1.82E-03	-1.78	7.42E-03	-2.13	17	H3150B12	BG088216.2
74	*Sfrs7*	Splicing factor, arginine/serine-rich 7	9.96E-04	-1.9	2.06E-03	-1.87	17	AI836707	AI836707
75	*Sfrs6*	Splicing factor, arginine/serine-rich 6	5.18E-05	-1.97	3.61E-05	-1.84	2	AI846595	AI846595
76	*Pogz*	Pogo transposable element with ZNF domain	4.83E-03	-2.15	2.36E-02	-1.48	3	H3087B02	BG070415.2
77	*Gsk3b*	Glycogen synthase kinase 3 beta	3.38E-02	-2.19	2.40E-02	-2.21	16	AI848867	AI848867
78	*Coq6*	Coenzyme Q6 homolog (yeast)	2.39E-03	-2.21	1.03E-02	-2.07	12	H4021B09	BQ553299.1
79	*Klf13*	Kruppel-like factor 13	1.51E-02	-2.25	9.14E-03	-4.73	7	AI851693	AI851693
80	*Mapkapk5*	MAP kinase-activated protein kinase 5	2.61E-02	-2.53	1.56E-02	-3.5	5	H4017H04	BQ552742.1
81	*Mtmr3*	Myotubularin related protein 3	4.47E-02	-3.19	3.71E-02	-2.64	11	H3150D05	BG075687.2
82			4.38E-03	-3.2	2.88E-02	-2.31		H4077F10	BQ562737.1
83	*4930534B04Rik*	RIKEN cDNA 4930534B04 gene	4.17E-03	-3.81	1.85E-03	-2.73	12	372704	AI414212
84	*Pbxip1*	Pre-B-cell leukemia transcription factor interacting protein 1	3.23E-02	-4.75	6.01E-03	-2.51	3	H3148F09	CK335156.1

Table A.4: Differentially expressed genes between B6 and D2 control mice. The genes in this list correspond to Figure 4.3. Note that the genes names are current as of July 31, 2006.

Table A.4: Differentially expressed genes between B6 and D2 control mice.

	Gene Symbol	Gene Name	t-test p value	Fold change D2 C / B6 C	Chromosome	Clone ID	Accession number
1	*Camk2n1*	Calcium/calmodulin-dependent protein kinase II inhibitor 1	1.45E-09	-14.44	4	AI848599	AI848599
2	*4933409K07Rik*	RIKEN cDNA 4933409K07 gene	5.73E-08	-8.36	4	H4012H09	BQ552019.1
3	*Pttg1*	Pituitary tumor-transforming 1	1.68E-09	-7.56	11	H3073E02	BG069168.2
4			5.84E-09	-5.62		H3072F07	BG069093.2
5	*4933409K07Rik*	RIKEN cDNA 4933409K07 gene	9.28E-05	-5.11	4	AI849820	AI849820
6	*Ccl21b*	Chemokine (C-C motif) ligand 21b	6.03E-05	-3.9	4	AI838322	AI838322
7	*D7Ertd715e*	DNA segment, Chr 7, ERATO Doi 715, expressed	1.06E-04	-3.61	7	H4048E08	BQ557833.1
8	*2210010A19Rik*	RIKEN cDNA 2210010A19 gene	4.19E-04	-3.44	16	AI849181	AI849181
9			3.23E-07	-3.25		H3144D01	
10	*Msr2*	Macrophage scavenger receptor 2	2.10E-04	-3.1	3	H4045B02	BQ557278.1
11	*Ammecr1*	Alport syndrome, mental retardation, midface hypoplasia and elliptocytosis chromosomal region gene 1 homolog (human)	3.42E-04	-2.88	X	H4070D11	BQ561457.1
12	*Slc4a4*	Solute carrier family 4 (anion exchanger), member 4	1.72E-04	-2.87	5	661621	AI465149
13		ESTs, Weakly similar to IGEB_MOUSE IGE-BINDING PROTEIN [M.musculus]	3.78E-05	-2.62		597699	AI326000
14	*Chd9*	Chromodomain helicase DNA binding protein 9	1.30E-05	-2.62		AI837598	AI837598
15	*2210012G02Rik*	RIKEN cDNA 2210012G02 gene	7.02E-04	-2.49	4	H3149C12	BG075610.2
16	*5830457H20Rik*	RIKEN cDNA 5830457H20 gene	1.13E-06	-2.47	12	AI841954	AI841954
17	*Entpd4*	Ectonucleoside triphosphate diphosphohydrolase 4	6.34E-05	-2.32	14	H3072D08	BG069071.2
18	*Ogfrl1*	Opioid growth factor receptor-like 1	2.08E-04	-2.31	1	AI846434	AI846434
19	*Spock1*	Sparc/osteonectin, cwcv and kazal-like domains proteoglycan 1	4.85E-04	-2.27	13	456265	AI323313
20	*Comt*	Catechol-O-methyltransferase	8.95E-04	-2.19	16	H4003F09	BQ550492.1
21		Transcribed sequence with strong similarity to protein sp:P00722 (E. coli) BGAL_ECOLI Beta-galactosidase	1.98E-05	-2.18		AI853048	AI853048
22	*Fus*	Fusion, derived from t(12;16) malignant liposarcoma (human)	2.02E-05	-2.16	3	H3028E01	BG065205.2
23	*BC056474*	CDNA sequence BC056474	1.05E-04	-2.11	8	H3030A06	CK334442.1
24	*Pkmyt1*	Protein kinase, membrane associated tyrosine/threonine 1	3.08E-05	-2.07	17	H4058B05	BQ559423.1
25	*Timp4*	Tissue inhibitor of metalloproteinase 4	7.94E-04	-2.07	6	AI844247	AI844247
26	*H47*	Mm histocompatibility 47 (H47) mRNA; Rn Ratsg2 mRNA	8.87E-04	-2.06		UTDDEST115130	
27	*Dnajc11*	DnaJ (Hsp40) homolog, subfamily C, member 11	5.06E-04	-1.99	4	H3037G09	BG066002.2
28	*Thada*	Thyroid adenoma associated	7.49E-05	-1.97	17	H4077F04	BQ562726.1
29	*Arpc5*	Actin related protein 2/3 complex, subunit 5	1.21E-04	-1.93	1	H3013G08	BG077307.2
30	*Pgbd5*	PiggyBac transposable element derived 5	3.05E-04	-1.91	8	AI854313	AI854313

	Gene Symbol	Gene Name	t-test p value	Fold change D2 C / B6 C	Chromosome	Clone ID	Accession number
31	Son	Son cell proliferation protein	2.46E-04	-1.89	16	AI854469	AI854469
32	Bean	Brain expressed, associated with Nedd4	9.06E-05	-1.87	8	AI851802	AI851802
33	Oxr1	Oxidation resistance 1	3.83E-04	-1.83	15	H3133G01	BG074326.2
34	Api5	Apoptosis inhibitor 5	3.11E-04	-1.78	2	H3147D12	BG075457.2
35	TU3A	Hs TU3A protein (TU3A) mRNA; Hs DRR1 (DRR1) mRNA	5.21E-04	-1.78		UTDDEST120010	
36	Snap25	Synaptosomal-associated protein 25	8.56E-04	-1.75	2	AI841525	AI841525
37	Schip1	Schwannomin interacting protein 1	2.04E-04	-1.74	3	H4033C10	BQ555299.1
38	Kcnc1	Potassium voltage gated channel, Shaw-related subfamily, member 1	9.72E-04	-1.74	7	AI844638	AI844638
39	Recql	RecQ protein-like	8.66E-04	-1.74	6	AI853438	AI853438
40	D430039N05Rik	RIKEN cDNA D430039N05 gene	1.93E-04	-1.71	2	H4037E05	BQ556002.1
41	Pdzrn3	RIKEN cDNA 9530086O07 gene	4.98E-05	-1.71	6	H4063H06	BQ560367.1
42	Dctn3	Dynactin 3	2.53E-05	-1.7	4	AI850004	AI850004
43	Camk4	Calcium/calmodulin-dependent protein kinase IV	1.70E-04	-1.69	18	AI840829	AI840829
44	Atp6v1g1	ATPase, H+ transporting, lysosomal V1 subunit G1	1.10E-04	-1.68	4	596308	AI326121
45	Akr1b3	Aldo-keto reductase family 1, member B3 (aldose reductase)	1.88E-04	-1.68	1	AI844017	AI844017
46	Kifap3	Kinesin-associated protein 3	5.57E-04	-1.65	1	H3018F02	BG064354.2
47	Pgm2l1	Phosphoglucomutase 2-like 1	6.50E-04	-1.64	7	AI845739	AI845739
48	Tmem68	Transmembrane protein 68	9.49E-04	-1.58	4	H3004E01	BG063217.1
49	Cep152	Centrosomal protein 152	7.03E-04	-1.58	2	AI851464	AI851464
50	Pcmt1	Protein-L-isoaspartate (D-aspartate) O-methyltransferase 1	6.62E-04	-1.58	10	AI841469	AI841469
51	Klk6	Kallikrein 6	5.15E-04	-1.58	7	AI849898	AI849898
52	Bex2	Brain expressed X-linked 2	8.33E-05	-1.58	X	AI836711	AI836711
53	2310016C08Rik	RIKEN cDNA 2310016C08 gene	9.47E-04	-1.56	6	H4006G02	BQ551002.1
54	Khdrbs3	KH domain containing, RNA binding, signal transduction associated 3	3.74E-04	-1.55	15	H4018F12	BQ552869.1
55		EST	2.83E-04	-1.54		AI839642	AI839642
56	Rps2	Ribosomal protein S2	1.09E-04	-1.52	17	H3004D12	BG063216.2
57	Psma1	Proteasome (prosome, macropain) subunit, alpha type 1	9.56E-04	-1.5	7	AI836804	AI836804
58	Gas5	Growth arrest specific 5	6.72E-04	-1.5	1	AI848205	AI848205
59	Cnbp1	Cellular nucleic acid binding protein 1	4.83E-05	-1.49	6	AI849259	AI849259
60	Stxbp1	Syntaxin binding protein 1	4.65E-04	-1.48	2	H4031G03	BQ555034.1

Table A.4: Differentially expressed genes between B6 and D2 control mice.

	Gene Symbol	Gene Name	t-test p value	Fold change D2 C / B6 C	Chromosome	Clone ID	Accession number
61		EST	3.78E-04	-1.47		AI845979	AI845979
62	*9630041G16Rik*	RIKEN cDNA 9630041G16 gene	2.10E-04	-1.45	4	AI848669	AI848669
63	*Gtpbp3*	GTP binding protein 3	3.76E-04	-1.44	8	H4024G12	BQ553885.1
64	*Rpsa*	Ribosomal protein SA	3.38E-04	-1.43	9	AI836492	AI836492
65	*Fbxo2*	F-box only protein 2	3.72E-04	-1.42	4	AI849925	AI849925
66	*HSP8*	Mm heat shock protein 8, mRNA; Mm bone marrow macrophage cDNA (product: HSP8)	6.01E-04	-1.42		UTDDEST119317	
67	*Slc25a3*	Solute carrier family 25 (mitochondrial carrier, phosphate carrier), member 3	5.34E-04	-1.41	10	H4058G05	BQ559528.1
68	*Skp1a*	S-phase kinase-associated protein 1A	8.98E-04	-1.4	11	H3023E01	BG064752.2
69	*Rps3*	Ribosomal protein S3	2.93E-04	-1.39	7	H3021F02	CK334393.1
70	*Golga3*	Golgi autoantigen, golgin subfamily a, 3	2.83E-05	-1.39	5	AI837809	AI837809
71	*Glul*	Glutamate-ammonia ligase (glutamine synthetase)	3.40E-05	-1.39	1	AI850098	AI850098
72	*Pls3*	Plastin 3 (T-isoform)	5.78E-05	-1.39	X	559843	AI448193
73	*Ndg1*	Nur77 downstream gene 1	6.80E-04	-1.37	1	H3091B03	BG070783.2
74		EST	5.61E-04	-1.36		AI849910	AI849910
75	*C1qb*	Complement component 1, q subcomponent, beta polypeptide	6.87E-04	-1.35	4	AI836786	AI836786
76	*Pabpc1*	Poly A binding protein, cytoplasmic 1	4.39E-04	-1.34	15	H3012E10	BG063770.2
77	*Brd2*	Bromodomain containing 2	5.20E-04	-1.34	17	616492	AI450617
78	*Hdac3*	Histone deacetylase 3	2.91E-04	-1.25	18	H3142E03	BG075034.2
79			6.82E-04	-1.22		H3092H11	
80	*Peli1*	Pellino 1	8.79E-04	-1.19	11	718768	AI666723
81	*Hax1*	HCLS1 associated X-1	1.49E-05	-1.35	3	H3024C08	BG064826.2
82	*Klhl15*	Kelch-like 15 (Drosophila)	7.80E-04	1.21	X	H3073G12	BG069200.2
83	*Ube2d2*	Ubiquitin-conjugating enzyme E2D 2	3.40E-04	1.27	18	H3018G09	BG077711.2
84	*2600010E01Rik*	RIKEN cDNA 2600010E01 gene	6.99E-04	1.27	2	H3055F02	BG067547.2
85			8.61E-04	1.33		H3104A06	BG071864.2
86	*Ccnc*	Cyclin C	4.65E-05	1.33	4	H3080H04	BG069856.2
87			5.85E-04	1.36		H4015G04	BQ552422.1
88		Transcribed sequences	4.24E-04	1.36		621008	AI451016
89	*2610203E10Rik*	RIKEN cDNA 2610203E10 gene	2.22E-04	1.46	2	H4019G02	BQ553048.1
90	*Tmepai*	Transmembrane, prostate androgen induced RNA	4.49E-04	1.46	2	AI841538	AI841538

Table A.4: Differentially expressed genes between B6 and D2 control mice.

	Gene Symbol	Gene Name	t-test p value	Fold change D2 C / B6 C	Chromosome	Clone ID	Accession number
91	*Rcbtb1*	Regulator of chromosome condensation (RCC1) and BTB (POZ) domain containing protein 1	8.71E-04	1.51	14	H3027H03	BG065148.2
92	*Rnase4*	Ribonuclease, RNase A family 4	5.67E-05	1.52	14	H4013G01	BQ552148.1
93	*Foxn2*	Forkhead box N2	3.69E-04	1.55	17	H3068A05	BG068663.2
94	*Cpne8*	Copine VIII	4.60E-05	1.57	15	H4022A05	BQ553434.1
95			3.16E-04	1.57		AI847856	AI847856
96			3.96E-04	1.67		H3088C08	BG070525.2
97			6.89E-04	1.78		H3002B12	
98	*1110008L16Rik*	RIKEN cDNA 1110008L16 gene	8.21E-04	1.79	12	H4076E01	BQ562522.1
99	*Rpa3*	Replication protein A3	1.16E-04	1.79	6	AI848299	AI848299
100			7.76E-05	1.85		H3060E10	
101	*Lypla1*	Lysophospholipase 1	1.04E-04	1.86	1	H3101E10	BG071649.1
102	*Itk*	IL2-inducible T-cell kinase	8.07E-04	1.92	11	H4078F03	BQ562899.1
103	*Spp1*	Secreted phosphoprotein 1	5.13E-04	1.92	5	AI847805	AI847805
104	*Tmem18*	Transmembrane protein 18	7.66E-04	1.93	12	AI852422	AI852422
105			3.40E-06	2.01		AI838238	AI838238
106	*Hsp90b1*	Heat shock protein 90kDa beta (Grp94), member 1	9.80E-04	2.04	10	AI836536	AI836536
107	*Ncoa2*	Nuclear receptor coactivator 2	2.47E-05	2.06	1	H3056A03	BG080723.2
108	*Eef2*	Eukaryotic translation elongation factor 2	8.93E-04	2.15	11	557887	AI448101
109		Mus musculus transcribed sequence	3.93E-04	2.17		585286	AI447764
110	*Klf12*	Kruppel-like factor 12	1.84E-04	2.44	14	AI853380	AI853380
111	*Glo1*	Glyoxalase 1	2.13E-04	2.44	17	AI852001	AI852001
112	*1190005F20Rik*	RIKEN cDNA 1190005F20 gene	3.38E-04	2.45	1	H3068H06	BG068747.2
113	*Glo1*	Glyoxalase 1	7.56E-07	2.46	17	H3012F03	BG063773.1
114	*4932416N17Rik*	RIKEN cDNA 4932416N17 gene	4.31E-05	2.48	16	AI852034	AI852034
115	*Alad*	Aminolevulinate, delta-, dehydratase	3.16E-06	2.62	4	H3013F01	BG063937.2
116			4.73E-05	2.78		H3064G12	BG068379.1
117	*Ccdc73*	Coiled-coil domain containing 73	4.59E-05	2.86	2	H3007D11	BG063455.2
118	*Unc13c*	Unc-13 homolog C (C. elegans)	9.08E-05	2.89	9	H3076H09	BG069400.2
119	*Lpl*	Lipoprotein lipase	2.34E-05	3.11	8	475661	AI326787
120	*Pisd*	Phosphatidylserine decarboxylase	7.36E-04	3.34	5	H3028E03	BG065207.2

Table A.4: Differentially expressed genes between B6 and D2 control mice.

	Gene Symbol	Gene Name	t-test p value	Fold change D2 C / B6 C	Chromosome	Clone ID	Accession number
121	*Slc7a2*	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 2	2.77E-04	3.67	8	H3050G05	BG067114.2
122	*Pisd*	Phosphatidylserine decarboxylase	8.87E-07	4.19	5	602514	AI449740
123	*Kirrel3*	Kin of IRRE like 3 (Drosophila)	1.40E-04	4.32	9	H4051B02	BQ558250.1
124	*Rbpsuh*	Recombining binding protein suppressor of hairless (Drosophila)	4.96E-06	5.71	5	H3049D12	BG067001.2
125	*Hddc3*	HD domain containing 3	8.76E-06	5.93	7	AI840260	AI840260
126	*Csnk1a1*	Casein kinase 1, alpha 1	6.06E-05	12.64	18	H3083D12	BG069996.2

References

Addolorato G, Caputo F, Capristo E, Domenicali M, Bernardi M, Janiri L, Agabio R, Colombo G, Gessa GL, Gasbarrini G (2002) Baclofen efficacy in reducing alcohol craving and intake: A preliminary double-blind randomized controlled study. Alcohol and Alcoholism 37:504-508.

Agartz I, Momenan R, Rawlings RR, Kerich MJ, Hommer DW (1999) Hippocampal volume in patients with alcohol dependence. Arch Gen Psychiatry 56:356-363.

Alexandre H, Ansanay-Galeote V, Dequin S, Blondin B (2001) Global gene expression during short-term ethanol stress in Saccharomyces cerevisiae. FEBS Lett 498:98-103.

Allocco DJ, Kohane IS, Butte AJ (2004) Quantifying the relationship between co-expression, co-regulation and gene function. BMC Bioinformatics 5:18.

Altman RB, Raychaudhuri S (2001) Whole-genome expression analysis: challenges beyond clustering. Curr Opin Struct Biol 11:340-347.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403-410.

Alwine JC, Kemp DJ, Stark GR (1977) Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proc Natl Acad Sci U S A 74:5350-5354.

An HJ, Cho NH, Lee SY, Kim IH, Lee C, Kim SJ, Mun MS, Kim SH, Jeong JK (2003) Correlation of cervical carcinoma and precancerous lesions with human papillomavirus (HPV) genotypes detected with the HPV DNA chip microarray method. Cancer 97:1672-1680.

APA (1980) Diagnostic and Statistical Manual of Mental Disorders, Third Edition: Washington, D.C.: The Association.

APA (2000) Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR, Fourth Edition: Washington, D.C.: The Association.

Archacki S, Wang Q (2004) Expression profiling of cardiovascular disease. Hum Genomics 1:355-370.

Arlinde C, Sommer W, Bjork K, Reimers M, Hyytia P, Kiianmaa K, Heilig M (2004) A cluster of differentially expressed signal transduction genes identified by microarray analysis in a rat genetic model of alcoholism. Pharmacogenomics J 4:208-218.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene

ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25-29.

Aston CE, Hill SY (1990) Segregation analysis of alcoholism in families ascertained through a pair of male alcoholics. Am J Hum Genet 46:879-887.

Baddeley A (1986) Working memory.: Oxford: Oxford University Press.

Bammler T, Beyer RP, Bhattacharya S, Boorman GA, Boyles A, Bradford BU, Bumgarner RE, Bushel PR, Chaturvedi K, Choi D, Cunningham ML, Deng S, Dressman HK, Fannin RD, Farin FM, Freedman JH, Fry RC, Harper A, Humble MC, Hurban P, Kavanagh TJ, Kaufmann WK, Kerr KF, Jing L, Lapidus JA, Lasarev MR, Li J, Li YJ, Lobenhofer EK, Lu X, Malek RL, Milton S, Nagalla SR, O'Malley J P, Palmer VS, Pattee P, Paules RS, Perou CM, Phillips K, Qin LX, Qiu Y, Quigley SD, Rodland M, Rusyn I, Samson LD, Schwartz DA, Shi Y, Shin JL, Sieber SO, Slifer S, Speer MC, Spencer PS, Sproles DI, Swenberg JA, Suk WA, Sullivan RC, Tian R, Tennant RW, Todd SA, Tucker CJ, Van Houten B, Weis BK, Xuan S, Zarbl H (2005) Standardizing global gene expression analysis between laboratories and across platforms. Nat Methods 2:351-356.

Beadles-Bohling AS, Wiren KM (2005) Alteration of kappa-opioid receptor system expression in distinct brain regions of a genetic model of enhanced ethanol withdrawal severity. Brain Res 1046:77-89.

Beck JA, Lloyd S, Hafezparast M, Lennon-Pierce M, Eppig JT, Festing MF, Fisher EM (2000) Genealogies of mouse inbred strains. Nat Genet 24:23-25.

Begleiter H, Reich T, Hesselbrock V, Porjesz B, Li T-K, Schuckit MA, Edenberg HJ, Rice JP (1995) The Collaborative Study on the Genetics of Alcoholism. Alcohol Res Health 19.

Begleiter H, Porjesz B, Reich T, Edenberg HJ, Goate A, Blangero J, Almasy L, Foroud T, Van Eerdewegh P, Polich J, Rohrbaugh J, Kuperman S, Bauer LO, O'Connor SJ, Chorlian DB, Li TK, Conneally PM, Hesselbrock V, Rice JP, Schuckit MA, Cloninger R, Nurnberger J, Jr., Crowe R, Bloom FE (1998) Quantitative trait loci analysis of human event-related brain potentials: P3 voltage. Electroencephalogr Clin Neurophysiol 108:244-250.

Beissbarth T, Speed TP (2004) GOstat: find statistically overrepresented Gene Ontologies within a group of genes. Bioinformatics 20:1464-1465.

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological) 57:289-300.

Bergeson SE, Berman AE, Dodd PR, Edenberg HJ, Hitzemann RJ, Lewohl JM, Lodowski KH, Sommer W (2005) Expression profiling in alcoholism research. Alcoholism: Clinical and Experimental Research 29:1066-1073.

Blednov YA, Metten P, Finn DA, Rhodes JS, Bergeson SE, Harris RA, Crabbe JC (2005) Hybrid C57BL/6J x FVB/NJ mice drink more alcohol than do C57BL/6J mice. Alcohol Clin Exp Res 29:1949-1958.

Boehm SL, 2nd, Ponomarev I, Jennings AW, Whiting PJ, Rosahl TW, Garrett EM, Blednov YA, Harris RA (2004) gamma-Aminobutyric acid A receptor subunit mutant mice: new perspectives on alcohol actions. Biochem Pharmacol 68:1581-1602.

Borghese CM, Henderson LA, Bleck V, Trudell JR, Harris RA (2003) Sites of excitatory and inhibitory actions of alcohols on neuronal alpha2beta4 nicotinic acetylcholine receptors. J Pharmacol Exp Ther 307:42-52.

Bowers BJ (2000) Applications of transgenic and knockout mice in alcohol research. Alcohol Res Health 24:175-184.

Bowirrat A, Oscar-Berman M (2005) Relationship between dopaminergic neurotransmission, alcoholism, and Reward Deficiency syndrome. Am J Med Genet B Neuropsychiatr Genet 132:29-37.

Brazma A, Vilo J (2000) Gene expression data analysis. FEBS Lett 480:17-24.

Buck KJ, Hahner L, Sikela J, Harris RA (1991) Chronic ethanol treatment alters brain levels of gamma-aminobutyric acidA receptor subunit mRNAs: relationship to genetic differences in ethanol withdrawal seizure severity. J Neurochem 57:1452-1455.

Buck KJ, Metten P, Belknap JK, Crabbe JC (1997) Quantitative trait loci involved in genetic predisposition to acute alcohol withdrawal in mice. J Neurosci 17:3946-3955.

Bustin SA, Benes V, Nolan T, Pfaffl MW (2005) Quantitative real-time RT-PCR--a perspective. J Mol Endocrinol 34:597-601.

Butterworth RF (2003) Hepatic encephalopathy--a serious complication of alcoholic liver disease. Alcohol Res Health 27:143-145.

Callard D, Lescure B, Mazzolini L (1994) A method for the elimination of false positives generated by the mRNA differential display technique. Biotechniques 16:1096-1097, 1100-1093.

Carey VJ, Gentry J, Whalen E, Gentleman R (2005) Network structures and algorithms in Bioconductor. Bioinformatics 21:135-136.

Carson JP, Ju T, Lu HC, Thaller C, Xu M, Pallas SL, Crair MC, Warren J, Chiu W, Eichele G (2005) A Digital Atlas to Characterize the Mouse Brain Transcriptome. PLoS Comput Biol 1:e41.

Chakkalakal DA (2005) Alcohol-induced bone loss and deficient bone repair. Alcohol Clin Exp Res 29:2077-2090.

Chen Z, Swisshelm K, Sager R (1994) A cautionary note on reaction tubes for differential display and cDNA amplification in thermal cycling. Biotechniques 16:1002-1004, 1006.

Chesler EJ, Wang J, Lu L, Qu Y, Manly KF, Williams RW (2003) Genetic correlates of gene expression in recombinant inbred strains: a relational model system to explore neurobehavioral phenotypes. Neuroinformatics 1:343-357.

Churchill GA (2002) Fundamentals of experimental design for cDNA microarrays. Nat Genet 32 Suppl:490-495.

Churchill GA (2004) Using ANOVA to analyze microarray data. Biotechniques 37:173-175, 177.

Cloninger CR, Bohman M, Sigvardsson S (1981) Inheritance of alcohol abuse. Cross-fostering analysis of adopted men. Arch Gen Psychiatry 38:861-868.

Cloninger CR, Svrakic DM, Przybeck TR (1993) A psychobiological model of temperament and character. Arch Gen Psychiatry 50:975-990.

Cloninger CR, Van Eerdewegh P, Goate A, Edenberg HJ, Blangero J, Hesselbrock V, Reich T, Nurnberger J, Jr., Schuckit M, Porjesz B, Crowe R, Rice JP, Foroud T, Przybeck TR, Almasy L, Bucholz K, Wu W, Shears S, Carr K, Crose C, Willig C, Zhao J, Tischfield JA, Li TK, Conneally PM, et al. (1998) Anxiety proneness linked to epistatic loci in genome scan of human personality traits. Am J Med Genet 81:313-317.

Collins MA, Zou JY, Neafsey EJ (1998) Brain damage due to episodic alcohol exposure in vivo and in vitro: furosemide neuroprotection implicates edema-based mechanism. Faseb J 12:221-230.

Cook RT (1998) Alcohol abuse, alcoholism, and damage to the immune system--a review. Alcohol Clin Exp Res 22:1927-1942.

Crabbe J, Goldman D (1992) Alcoholism: A complex genetic disease. Alcohol Health Res World 16.

Crabbe JC (2002a) Genetic contributions to addiction. Annu Rev Psychol 53:435-462.

Crabbe JC (2002b) Alcohol and genetics: new models. Am J Med Genet 114:969-974.

Crabbe JC, Merrill C, Belknap JK (1991) Acute dependence on depressant drugs is determined by common genes in mice. J Pharmacol Exp Ther 257:663-667.

Crabbe JC, Belknap JK, Buck KJ (1994) Genetic animal models of alcohol and drug abuse. Science 264:1715-1723.

Crabbe JC, Metten P, Cameron AJ, Wahlsten D (2005) An analysis of the genetics of alcohol intoxication in inbred mice. Neurosci Biobehav Rev 28:785-802.

Crabbe JC, Phillips TJ, Buck KJ, Cunningham CL, Belknap JK (1999) Identifying genes for alcohol and drug sensitivity: recent progress and future directions. Trends Neurosci 22:173-179.

Crabbe JC, Jr., Young ER, Kosobud A (1983) Genetic correlations with ethanol withdrawal severity. Pharmacol Biochem Behav 18 Suppl 1:541-547.

Cunningham BA (2001) Assessing differential gene expression. The Scientist 15:27.

Dahchour A, De Witte P (2000) Ethanol and amino acids in the central nervous system: assessment of the pharmacological actions of acamprosate. Prog Neurobiol 60:343-362.

Daniels GM, Buck KJ (2002) Expression profiling identifies strain-specific changes associated with ethanol withdrawal in mice. Genes Brain Behav 1:35-45.

de Hoon MJ, Imoto S, Nolan J, Miyano S (2004) Open source clustering software. Bioinformatics 20:1453-1454.

DeBakey SF, Stinson FS, Grant BF, Dufour MC (1996) Liver Cirrhosis Mortality in the United States, 1970-93. In. Bethesda, MD: NIAAA.

DeRisi JL, Iyer VR, Brown PO (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. Science 278:680-686.

Di Chiara G (1997) Alcohol and dopamine. Alcohol Health Res World 21:108-114.

Diamond I, Gordon AS (1997) Cellular and molecular neuroscience of alcoholism. Physiol Rev 77:1-20.

Douglas CL, Baghdoyan HA, Lydic R (2001) M2 muscarinic autoreceptors modulate acetylcholine release in prefrontal cortex of C57BL/6J mouse. J Pharmacol Exp Ther 299:960-966.

Douglas CL, Baghdoyan HA, Lydic R (2002) Prefrontal cortex acetylcholine release, EEG slow waves, and spindles are modulated by M2 autoreceptors in C57BL/6J mouse. J Neurophysiol 87:2817-2822.

Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM (1999) Expression profiling using cDNA microarrays. Nat Genet 21:10-14.

Edenberg HJ (2002) The collaborative study on the genetics of alcoholism: an update. Alcohol Res Health 26:214-218.

Edenberg HJ, Strother WN, McClintick JN, Tian H, Stephens M, Jerome RE, Lumeng L, Li TK, McBride WJ (2005) Gene expression in the hippocampus of inbred alcohol-preferring and -nonpreferring rats. Genes Brain Behav 4:20-30.

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95:14863-14868.

Enoch MA, Goldman D (2001) The genetics of alcoholism and alcohol abuse. Curr Psychiatry Rep 3:144-151.

Erwin VG, McClearn GE (1981) Genetic influences on alcohol consumption and actions of alcohol. Curr Alcohol 8:405-420.

Erwin VG, Deitrich RA (1996) Genetic selection and characterization of mouse lines for acute functional tolerance to ethanol. J Pharmacol Exp Ther 279:1310-1317.

Fan L, Jaquet V, Dodd PR, Chen W, Wilce PA (2001) Molecular cloning and characterization of hNP22: a gene up-regulated in human alcoholic brain. J Neurochem 76:1275-1281.

Fan L, van der Brug M, Chen W, Dodd PR, Matsumoto I, Niwa S, Wilce PA (1999) Increased expression of mitochondrial genes in human alcoholic brain revealed by differential display. Alcohol Clin Exp Res 23:408-413.

Fare TL, Coffey EM, Dai H, He YD, Kessler DA, Kilian KA, Koch JE, LeProust E, Marton MJ, Meyer MR, Stoughton RB, Tokiwa GY, Wang Y (2003) Effects of atmospheric ozone on microarray data quality. Anal Chem 75:4672-4675.

Fehr C, Shirley RL, Metten P, Kosobud AE, Belknap JK, Crabbe JC, Buck KJ (2004) Potential pleiotropic effects of Mpdz on vulnerability to seizures. Genes Brain Behav 3:8-19.

Finn DA, Belknap JK, Cronise K, Yoneyama N, Murillo A, Crabbe JC (2005) A procedure to produce high alcohol intake in mice. Psychopharmacology (Berl) 178:471-480.

Flatscher-Bader T, van der Brug M, Hwang JW, Gochee PA, Matsumoto I, Niwa S, Wilce PA (2005) Alcohol-responsive genes in the frontal cortex and nucleus accumbens of human alcoholics. J Neurochem 93:359-370.

Franklin KBJ, Paxinos G (1997) The mouse brain in stereotaxic coordinates. New York: Academic Press.

Freeman WM, Walker SJ, Vrana KE (1999) Quantitative RT-PCR: pitfalls and potential. Biotechniques 26:112-122, 124-115.

Friedman HJ (1980) Assessment of physical dependence on and withdrawal from ethanol in animals. In: Alcohol Tolerance and Dependence (Rigter H, Crabbe JC, eds), pp 93-121. Amsterdam: Elsevier/North-Holland Biomedical Press.

Froehlich JC (1997) Opioid peptides. Alcohol Health Res World 21:132-136.

Fuller RK, Hiller-Sturmhofel S (1999) Alcoholism Treatment in the United States. Alcohol Research and Health 23.

Funahashi S (2001) Neuronal mechanisms of executive control by the prefrontal cortex. Neurosci Res 39:147-165.

Gall JG, Pardue ML (1969) Formation and detection of RNA-DNA hybrid molecules in cytological preparations. Proc Natl Acad Sci U S A 63:378-383.

Gayer GG, Gordon A, Miles MF (1991) Ethanol increases tyrosine hydroxylase gene expression in N1E-115 neuroblastoma cells. J Biol Chem 266:22279-22284.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F,

Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5:R80.

Gianoulakis C (2001) Influence of the endogenous opioid system on high alcohol consumption and genetic predisposition to alcoholism. J Psychiatry Neurosci 26:304-318.

Goldsmith ZG, Dhanasekaran N (2004) The microrevolution: applications and impacts of microarray technology on molecular biology and medicine (review). Int J Mol Med 13:483-495.

Goldstein DB (1972) Relationship of alcohol dose to intensity of withdrawal signs in mice. J Pharmacol Exp Ther 180:203-215.

Goldstein DB, Pal N (1971) Alcohol dependence produced in mice by inhalation of ethanol: grading the withdrawal reaction. Science 172:288-290.

Gonzales RA, Jaworski JN (1997) Alcohol and glutamate. Alcohol Health Res World 21:120-127.

Goodwin DW, Schulsinger F, Hermansen L, Guze SB, Winokur G (1973) Alcohol problems in adoptees raised apart from alcoholic biological parents. Arch Gen Psychiatry 28:238-243.

Grahame NJ (2000) Selected lines and inbred strains. Tools in the hunt for the genes involved in alcoholism. Alcohol Res Health 24:159-163.

Grant BF, Harford TC, Dawson DA, Chou P, Dufour MC, Pickering RP (1994) Prevalence of DSM-IV alcohol abuse and dependence: United States, 1992. Alcohol Health Res World 18:243-248.

Grant BF, Dawson DA, Stinson FS, Chou SP, Dufour MC, Pickering RP (2004) The 12-month prevalence and trends in DSM-IV alcohol abuse and dependence: United States, 1991-1992 and 2001-2002. Drug Alcohol Depend 74:223-234.

Grieve SJ, Littleton JM (1979) The rapid development of functional tolerance to ethanol by mice. J Pharm Pharmacol 31:605-610.

Grisel JE (2000) Quantitative trait locus analysis. Alcohol Res Health 24:169-174.

Gukovskaya AS, Hosseini S, Satoh A, Cheng JH, Nam KJ, Gukovsky I, Pandol SJ (2004) Ethanol differentially regulates NF-kappaB activation in pancreatic acinar cells through calcium and protein kinase C pathways. Am J Physiol Gastrointest Liver Physiol 286:G204-213.

Guldin WO, Pritzel M, Markowitsch HJ (1981) Prefrontal cortex of the mouse defined as cortical projection area of the thalamic mediodorsal nucleus. Brain Behav Evol 19:93-107.

Gutala R, Wang J, Kadapakkam S, Hwang Y, Ticku M, Li MD (2004) Microarray analysis of ethanol-treated cortical neurons reveals disruption of genes related to

the ubiquitin-proteasome pathway and protein synthesis. Alcohol Clin Exp Res 28:1779-1788.

Hannenhalli S, Levy S (2003) Transcriptional regulation of protein complexes and biological pathways. Mamm Genome 14:611-619.

Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, Richter J, Rubin GM, Blake JA, Bult C, Dolan M, Drabkin H, Eppig JT, Hill DP, Ni L, Ringwald M, Balakrishnan R, Cherry JM, Christie KR, Costanzo MC, Dwight SS, Engel S, Fisk DG, Hirschman JE, Hong EL, Nash RS, Sethuraman A, Theesfeld CL, Botstein D, Dolinski K, Feierbach B, Berardini T, Mundodi S, Rhee SY, Apweiler R, Barrell D, Camon E, Dimmer E, Lee V, Chisholm R, Gaudet P, Kibbe W, Kishore R, Schwarz EM, Sternberg P, Gwinn M, Hannick L, Wortman J, Berriman M, Wood V, de la Cruz N, Tonellato P, Jaiswal P, Seigfried T, White R (2004) The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res 32:D258-261.

Harwood H (1998) Updating Estimates of the Economic Costs of Alcohol Abuse in the United States: Estimates, Update Methods, and Data. In. Bethesda, MD: The Lewin Group for the NIAAA.

Hasin DS, Goodwin RD, Stinson FS, Grant BF (2005) Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. Arch Gen Psychiatry 62:1097-1106.

Hassan S, Duong B, Kim KS, Miles MF (2003) Pharmacogenomic analysis of mechanisms mediating ethanol regulation of dopamine beta-hydroxylase. J Biol Chem 278:38860-38869.

Heath AC, Bucholz KK, Madden PA, Dinwiddie SH, Slutske WS, Bierut LJ, Statham DJ, Dunne MP, Whitfield JB, Martin NG (1997) Genetic and environmental contributions to alcohol dependence risk in a national twin sample: consistency of findings in women and men. Psychol Med 27:1381-1396.

Heidbreder CA, Groenewegen HJ (2003) The medial prefrontal cortex in the rat: evidence for a dorso-ventral distinction based upon functional and anatomical characteristics. Neurosci Biobehav Rev 27:555-579.

Heilig M, Egli M (2006) Pharmacological treatment of alcohol dependence: Target symptoms and target mechanisms. Pharmacol Ther 111:855-876.

Herz A (1997) Endogenous opioid systems and alcohol addiction. Psychopharmacology (Berl) 129:99-111.

Heyer LJ, Kruglyak S, Yooseph S (1999) Exploring expression data: identification and analysis of coexpressed genes. Genome Res 9:1106-1115.

Ho Sui SJ, Mortimer JR, Arenillas DJ, Brumm J, Walsh CJ, Kennedy BP, Wasserman WW (2005) oPOSSUM: identification of over-represented transcription factor binding sites in co-expressed genes. Nucleic Acids Res 33:3154-3164.

Hoffman PL, Miles M, Edenberg HJ, Sommer W, Tabakoff B, Wehner JM, Lewohl J (2003) Gene expression in brain: a window on ethanol dependence, neuroadaptation, and preference. Alcohol Clin Exp Res 27:155-168.

Hofler H, Mueller J, Werner M (1998) Principles of in situ hybridization. In: In Situ Hybridization. Principles and Practice., Second Edition (Polak JM, McGee JOD, eds), pp 1-19. Oxford: Oxford Medical Publications.

Horii A (1996) [RNase protection analysis]. Tanpakushitsu Kakusan Koso 41:561-563.

Hunt WA, Lands WE (1992) A role for behavioral sensitization in uncontrolled ethanol intake. Alcohol 9:327-328.

Hyman SE, Malenka RC (2001) Addiction and the brain: the neurobiology of compulsion and its persistence. Nat Rev Neurosci 2:695-703.

Ihaka R, Gentleman R (1996) R: A language for data analysis and graphics. Journal of Computational and Graphical Studies 5:299-314.

Ihalainena JA, Riekkinen Jr. P, Feenstra MGP (1999) Comparison of dopamine and noradrenaline release in mouse prefrontal cortex, striatum and hippocampus using microdialysis. Neuroscience Letters 277:71-74.

IHGSC (2004) Finishing the euchromatic sequence of the human genome. Nature 431:931-945.

Islam MR, Rodova M, Calvet JP (2002) A fast and efficient method of DNA fragment isolation from agarose gels without using commercial kits. American Biotechnology Laboratory 20:18.

Iwamoto K, Bundo M, Yamamoto M, Ozawa H, Saito T, Kato T (2004) Decreased expression of NEFH and PCP4/PEP19 in the prefrontal cortex of alcoholics. Neurosci Res 49:379-385.

Jeffrey SS, Lonning PE, Hillner BE (2005) Genomics-based prognosis and therapeutic prediction in breast cancer. J Natl Compr Canc Netw 3:291-300.

Johansson D, Lindgren P, Berglund A (2003) A multivariate approach applied to microarray data for identification of genes with cell cycle-coupled transcription. Bioinformatics 19:467-473.

Johnson BA, Ait-Daoud N, Bowden CL, DiClemente CC, Roache JD, Lawson K, Javors MA, Ma JZ (2003) Oral topiramate for treatment of alcohol dependence: a randomised controlled trial. Lancet 361:1677-1685.

Johnson BA, Roache JD, Javors MA, DiClemente CC, Cloninger CR, Prihoda TJ, Bordnick PS, Ait-Daoud N, Hensler J (2000) Ondansetron for reduction of drinking among biologically predisposed alcoholic patients: A randomized controlled trial. Jama 284:963-971.

Joyce C (2002) Quantitative RT-PCR. A review of current methodologies. Methods Mol Biol 193:83-92.

Kanehisa M (1997) A database for post-genome analysis. Trends Genet 13:375-376.

Keir WJ, Morrow AL (1994) Differential expression of GABAA receptor subunit mRNAs in ethanol-naive withdrawal seizure resistant (WSR) vs. withdrawal seizure prone (WSP) mouse brain. Brain Res Mol Brain Res 25:200-208.

Kelley AE, Berridge KC (2002) The neuroscience of natural rewards: relevance to addictive drugs. J Neurosci 22:3306-3311.

Kendler KS, Prescott CA, Neale MC, Pedersen NL (1997) Temperance board registration for alcohol abuse in a national sample of Swedish male twins, born 1902 to 1949. Arch Gen Psychiatry 54:178-184.

Kendler KS, Heath AC, Neale MC, Kessler RC, Eaves LJ (1992) A population-based twin study of alcoholism in women. Jama 268:1877-1882.

Kerns RT, Ravindranathan A, Hassan S, Cage MP, York T, Sikela JM, Williams RW, Miles MF (2005) Ethanol-responsive brain region expression networks: implications for behavioral responses to acute ethanol in DBA/2J versus C57BL/6J mice. J Neurosci 25:2255-2266.

Killion PJ, Sherlock G, Iyer VR (2003) The Longhorn Array Database (LAD): an open-source, MIAME compliant implementation of the Stanford Microarray Database (SMD). BMC Bioinformatics 4:32.

Kim CJ, Jeong JK, Park M, Park TS, Park TC, Namkoong SE, Park JS (2003) HPV oligonucleotide microarray-based detection of HPV genotypes in cervical neoplastic lesions. Gynecol Oncol 89:210-217.

Koob GF (1992) Drugs of abuse: anatomy, pharmacology and function of reward pathways. Trends Pharmacol Sci 13:177-184.

Kosobud A, Crabbe JC (1986) Ethanol withdrawal in mice bred to be genetically prone or resistant to ethanol withdrawal seizures. J Pharmacol Exp Ther 238:170-177.

Krawczyk DC (2002) Contributions of the prefrontal cortex to the neural basis of human decision making. Neurosci Biobehav Rev 26:631-664.

Kril JJ, Halliday GM (1999) Brain shrinkage in alcoholics: a decade on and what have we learned? Prog Neurobiol 58:381-387.

Kril JJ, Halliday GM, Svoboda MD, Cartwright H (1997) The cerebral cortex is damaged in chronic alcoholics. Neuroscience 79:983-998.

Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, Tong S, Urbani C, Comer JA, Lim W, Rollin PE, Dowell SF, Ling AE, Humphrey CD, Shieh WJ, Guarner J, Paddock CD, Rota P, Fields B, DeRisi J, Yang JY, Cox N, Hughes JM, LeDuc JW, Bellini WJ, Anderson LJ (2003) A novel coronavirus associated with severe acute respiratory syndrome. N Engl J Med 348:1953-1966.

Kwon JM, Goate AM (2000) The candidate gene approach. Alcohol Res Health 24:164-168.

Kwon JY, Hong M, Choi MS, Kang S, Duke K, Kim S, Lee S, Lee J (2004) Ethanol-response genes and their regulation analyzed by a microarray and comparative genomic approach in the nematode Caenorhabditis elegans. Genomics 83:600-614.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409:860-921.

Le AD, Mihic SJ, Wu PH (1992) Alcohol Tolerance: Methodological and Experimental Issues. In: Neuromethods, Vol. 24: Animal Models of Drug Addiction (Boulton A, Baker G, Wu PH, eds): The Humana Press Inc.

LeBlanc AE, Kalant H, Gibbins RJ (1975) Acute tolerance to ethanol in the rat. Psychopharmacologia 41:43-46.

Lessov CN, Phillips TJ (1998) Duration of sensitization to the locomotor stimulant effects of ethanol in mice. Psychopharmacology (Berl) 135:374-382.

Letwin NE, Kafkafi N, Benjamini Y, Mayo C, Frank BC, Luu T, Lee NH, Elmer GI (2006) Combined application of behavior genetics and microarray analysis to identify regional expression themes and gene-behavior associations. J Neurosci 26:5277-5287.

Lewohl JM, Wang L, Miles MF, Zhang L, Dodd PR, Harris RA (2000) Gene expression in human alcoholism: microarray analysis of frontal cortex. Alcohol Clin Exp Res 24:1873-1882.

Liang P, Pardee AB (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257:967-971.

Liang P, Pardee AB (1998) Differential display. A general protocol. Mol Biotechnol 10:261-267.

Liang T, Spence J, Liu L, Strother WN, Chang HW, Ellison JA, Lumeng L, Li TK, Foroud T, Carr LG (2003) alpha-Synuclein maps to a quantitative trait locus for

alcohol preference and is differentially expressed in alcohol-preferring and -nonpreferring rats. Proc Natl Acad Sci U S A 100:4690-4695.

Littleton J, Zieglgansberger W (2003) Pharmacological mechanisms of naltrexone and acamprosate in the prevention of relapse in alcohol dependence. Am J Addict 12 Suppl 1:S3-11.

Littleton JM, John GR, Jones PA, Grieve SJ (1980) The rapid onset of functional tolerance to ethanol--role of different neurotransmitters and synaptosomal membrane lipids. Acta Psychiatr Scand Suppl 286:137-151.

Liu J, Lewohl JM, Dodd PR, Randall PK, Harris RA, Mayfield RD (2004) Gene expression profiling of individual cases reveals consistent transcriptional changes in alcoholic human brain. J Neurochem 90:1050-1058.

Liu J, Lewohl JM, Harris RA, Iyer VR, Dodd PR, Randall PK, Mayfield RD (2006) Patterns of gene expression in the frontal cortex discriminate alcoholic from nonalcoholic individuals. Neuropsychopharmacology 31:1574-1582.

Loney KD, Uddin KR, Singh SM (2003) Strain-specific brain metallothionein II (MT-II) gene expression, its ethanol responsiveness, and association with ethanol preference in mice. Alcohol Clin Exp Res 27:388-395.

Lovinger DM (1997) Serotonin's role in alcohol's effects on the brain. Alcohol Health Res World 21:114-120.

Lucas DL, Brown RA, Wassef M, Giles TD (2005) Alcohol and the cardiovascular system research challenges and opportunities. J Am Coll Cardiol 45:1916-1924.

Lukas SE, Penetar D, Berko J, Vicens L, Palmer C, Mallya G, Macklin EA, Lee DY (2005) An extract of the Chinese herbal root kudzu reduces alcohol drinking by heavy drinkers in a naturalistic setting. Alcohol Clin Exp Res 29:756-762.

MacLaren EJ, Sikela JM (2005) Cerebellar gene expression profiling and eQTL analysis in inbred mouse strains selected for ethanol sensitivity. Alcohol Clin Exp Res 29:1568-1579.

MacLaren EJ, Bennett B, Johnson TE, Sikela JM (2006) Expression profiling identifies novel candidate genes for ethanol sensitivity QTLs. Mamm Genome 17:147-156.

Maiya R, Buck KJ, Harris RA, Mayfield RD (2002) Ethanol-sensitive sites on the human dopamine transporter. J Biol Chem 277:30724-30729.

Manly KF, Nettleton D, Hwang JT (2004) Genomics, prior probability, and statistical tests of multiple hypotheses. Genome Res 14:997-1001.

Martin PR, Singleton CK, Hiller-Sturmhofel S (2003) The role of thiamine deficiency in alcoholic brain disease. Alcohol Res Health 27:134-142.

Mayfield RD, Lewohl JM, Dodd PR, Herlihy A, Liu J, Harris RA (2002) Patterns of gene expression are altered in the frontal and motor cortices of human alcoholics. J Neurochem 81:802-813.

McBride WJ, Kerns RT, Rodd ZA, Strother WN, Edenberg HJ, Hashimoto JG, Wiren KM, Miles MF (2005) Alcohol effects on central nervous system gene expression in genetic animal models. Alcohol Clin Exp Res 29:167-175.

McCaul ME, Turkkan JS, Svikis DS, Bigelow GE (1990) Alcohol and secobarbital effects as a function of familial alcoholism: acute psychophysiological effects. Alcohol Clin Exp Res 14:704-712.

McClearn G (1959) The genetics of mouse behavior in novel situations. J Comp Physiol Psychol 52:62-67.

McShane LM, Shih JH, Michalowska AM (2003) Statistical issues in the design and analysis of gene expression microarray studies of animal models. J Mammary Gland Biol Neoplasia 8:359-374.

Mellanby E (1919) Alcohol: Its absorption into and disappearance from the blood under different conditions. In: Special Report Series No. 31. London, UK: Medical Research Committee.

Metten P, Crabbe JC (1994) Common genetic determinants of severity of acute withdrawal from ethanol, pentobarbital and diazepam in inbred mice. Behav Pharmacol 5:533-547.

Metten P, Phillips TJ, Crabbe JC, Tarantino LM, McClearn GE, Plomin R, Erwin VG, Belknap JK (1998) High genetic susceptibility to ethanol withdrawal predicts low ethanol consumption. Mamm Genome 9:983-990.

Mihic SJ, Harris RA (1997) GABA and the GABAA receptor. Alcohol Health Res World 21:127-131.

Mihic SJ, Ye Q, Wick MJ, Koltchine VV, Krasowski MD, Finn SE, Mascia MP, Valenzuela CF, Hanson KK, Greenblatt EP, Harris RA, Harrison NL (1997) Sites of alcohol and volatile anaesthetic action on GABA(A) and glycine receptors. Nature 389:385-389.

Miles MF, Diaz JE, DeGuzman VS (1991) Mechanisms of neuronal adaptation to ethanol. Ethanol induces Hsc70 gene transcription in NG108-15 neuroblastoma x glioma cells. J Biol Chem 266:2409-2414.

Mirnics K, Pevsner J (2004) Progress in the use of microarray technology to study the neurobiology of disease. Nat Neurosci 7:434-439.

Morgan MY (1982) Alcohol and nutrition. Br Med Bull 38:21-29.

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 (2006) Toward understanding the genetics of alcohol drinking through transcriptome meta-analysis. Proc Natl Acad Sci U S A 103:6368-6373.

Nestler EJ (2001) Molecular basis of long-term plasticity underlying addiction. Nat Rev Neurosci 2:119-128.

Nestler EJ (2005) Is there a common molecular pathway for addiction? Nat Neurosci 8:1445-1449.

Nestler EJ, Aghajanian GK (1997) Molecular and cellular basis of addiction. Science 278:58-63.

Nestler EJ, Landsman D (2001) Learning about addiction from the genome. Nature 409:834-835.

Neuroscience Sf (2005) Brain Facts: A primer on the brain and nervous system.

Newlin DB, Pretorius MB (1990) Sons of alcoholics report greater hangover symptoms than sons of nonalcoholics: a pilot study. Alcohol Clin Exp Res 14:713-716.

Nisenbaum LK (2002) The ultimate chip shot: can microarray technology deliver for neuroscience? Genes Brain Behav 1:27-34.

Nuwaysir EF (2000) Design, generation and use of cDNA microarrays on glass. In: Gene Expression Center at NIEHS.

Obernier JA, Bouldin TW, Crews FT (2002a) Binge ethanol exposure in adult rats causes necrotic cell death. Alcohol Clin Exp Res 26:547-557.

Obernier JA, White AM, Swartzwelder HS, Crews FT (2002b) Cognitive deficits and CNS damage after a 4-day binge ethanol exposure in rats. Pharmacol Biochem Behav 72:521-532.

Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M (1999) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 27:29-34.

Oscar-Berman M, Marinkovic K (2003) Alcoholism and the brain: an overview. Alcohol Res Health 27:125-133.

Palmer MR, Basile AS, Proctor WR, Baker RC, Dunwiddie TV (1985) Ethanol tolerance of cerebellar purkinje neurons from selectively outbred mouse lines: in vivo and in vitro electrophysiological investigations. Alcohol Clin Exp Res 9:291-296.

Pandey SC, Roy A, Zhang H, Xu T (2004) Partial deletion of the cAMP response element-binding protein gene promotes alcohol-drinking behaviors. J Neurosci 24:5022-5030.

Pandey SC, Chartoff EH, Carlezon WA, Jr., Zou J, Zhang H, Kreibich AS, Blendy JA, Crews FT (2005) CREB gene transcription factors: role in molecular mechanisms of alcohol and drug addiction. Alcohol Clin Exp Res 29:176-184.

Pavlidis P (2003) Using ANOVA for gene selection from microarray studies of the nervous system. Methods 31:282-289.

Petersen D, Chandramouli GV, Geoghegan J, Hilburn J, Paarlberg J, Kim CH, Munroe D, Gangi L, Han J, Puri R, Staudt L, Weinstein J, Barrett JC, Green J, Kawasaki ES (2005) Three microarray platforms: an analysis of their concordance in profiling gene expression. BMC Genomics 6:63.

Pfefferbaum A, Lim KO, Zipursky RB, Mathalon DH, Rosenbloom MJ, Lane B, Ha CN, Sullivan EV (1992) Brain gray and white matter volume loss accelerates with aging in chronic alcoholics: a quantitative MRI study. Alcohol Clin Exp Res 16:1078-1089.

Phillips TJ, Roberts AJ, Lessov CN (1997) Behavioral sensitization to ethanol: genetics and the effects of stress. Pharmacol Biochem Behav 57:487-493.

Phillips TJ, Crabbe JC, Metten P, Belknap JK (1994) Localization of genes affecting alcohol drinking in mice. Alcohol Clin Exp Res 18:931-941.

Pickens RW, Svikis DS, McGue M, Lykken DT, Heston LL, Clayton PJ (1991) Heterogeneity in the inheritance of alcoholism. A study of male and female twins. Arch Gen Psychiatry 48:19-28.

Ponomarev I, Maiya R, Harnett MT, Schafer GL, Ryabinin AE, Blednov YA, Morikawa H, Boehm SL, 2nd, Homanics GE, Berman AE, Lodowski KH, Bergeson SE, Harris RA (2006) Transcriptional signatures of cellular plasticity in mice lacking the alpha1 subunit of GABAA receptors. J Neurosci 26:5673-5683.

Princivalle AP, Parker RM, Dover TJ, Barnes NM (2005) mRNA: detection by in situ and northern hybridization. Methods Mol Biol 306:51-91.

Putzke J, De Beun R, Schreiber R, De Vry J, Tolle TR, Zieglgansberger W, Spanagel R (1998) Long-term alcohol self-administration and alcohol withdrawal differentially modulate microtubule-associated protein 2 (MAP2) gene expression in the rat brain. Brain Res Mol Brain Res 62:196-205.

Quackenbush J (2001) Computational analysis of microarray data. Nat Rev Genet 2:418-427.

Quackenbush J (2002) Microarray data normalization and transformation. Nat Genet 32 Suppl:496-501.

Regier DA, Farmer ME, Rae DS, Locke BZ, Keith SJ, Judd LL, Goodwin FK (1990) Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study. Jama 264:2511-2518.

Rehm J, Gmel G, Sempos CT, Trevisan M (2003) Alcohol-Related Morbidity and Mortality. Alcohol Research and Health 27.

Reich T, Hinrichs A, Culverhouse R, Bierut L (1999) Genetic studies of alcoholism and substance dependence. Am J Hum Genet 65:599-605.

Reimers M (2005) Statistical analysis of microarray data. Addict Biol 10:23-35.

Rhodes JS, Crabbe JC (2005) Gene expression induced by drugs of abuse. Curr Opin Pharmacol 5:26-33.

Rhodes JS, Best K, Belknap JK, Finn DA, Crabbe JC (2005) Evaluation of a simple model of ethanol drinking to intoxication in C57BL/6J mice. Physiol Behav 84:53-63.

Rimondini R, Arlinde C, Sommer W, Heilig M (2002) Long-lasting increase in voluntary ethanol consumption and transcriptional regulation in the rat brain after intermittent exposure to alcohol. Faseb J 16:27-35.

Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Penaranda S, Bankamp B, Maher K, Chen MH, Tong S, Tamin A, Lowe L, Frace M, DeRisi JL, Chen Q, Wang D, Erdman DD, Peret TC, Burns C, Ksiazek TG, Rollin PE, Sanchez A, Liffick S, Holloway B, Limor J, McCaustland K, Olsen-Rasmussen M, Fouchier R, Gunther S, Osterhaus AD, Drosten C, Pallansch MA, Anderson LJ, Bellini WJ (2003) Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300:1394-1399.

Rulten SL, Ripley TL, Hunt CL, Stephens DN, Mayne LV (2006) Sp1 and NFkappaB pathways are regulated in brain in response to acute and chronic ethanol. Genes Brain Behav 5:257-273.

Saito M, Smiley J, Toth R, Vadasz C (2002) Microarray analysis of gene expression in rat hippocampus after chronic ethanol treatment. Neurochem Res 27:1221-1229.

Saito M, Szakall I, Toth R, Kovacs KM, Oros M, Prasad VV, Blumenberg M, Vadasz C (2004) Mouse striatal transcriptome analysis: effects of oral self-administration of alcohol. Alcohol 32:223-241.

Saitz R (1998) Introduction to alcohol withdrawal. Alcohol Health Res World 22:5-12.

Saldanha AJ (2004) Java Treeview--extensible visualization of microarray data. Bioinformatics 20:3246-3248.

SAMHSA (2004) Substance Abuse and Mental Health Services Administration: 2004 National Survey on Drug Use and Health. In: U.S. Department of Health and Human Services.

Schafer GL, Crabbe JC, Wiren KM (1998) Identification of neuroendocrine-specific protein as an ethanol-regulated gene with mRNA differential display. Mamm Genome 9:979-982.

Schafer GL, Crabbe JC, Wiren KM (2001) Ethanol-regulated gene expression of neuroendocrine specific protein in mice: brain region and genotype specificity. Brain Res 897:139-149.

Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467-470.

Schmidt LG, Dufeu P, Kuhn S, Smolka M, Rommelspacher H (2000) Transition to alcohol dependence: clinical and neurobiological considerations. Compr Psychiatry 41:90-94.

Schulze A, Downward J (2001) Navigating gene expression using microarrays--a technology review. Nat Cell Biol 3:E190-195.

Shannon W, Culverhouse R, Duncan J (2003) Analyzing microarray data using cluster analysis. Pharmacogenomics 4:41-52.

Silver LM (1995) Mouse Genetics: Concepts and Applications: Oxford University Press.

Slonim DK (2002) From patterns to pathways: gene expression data analysis comes of age. Nat Genet 32 Suppl:502-508.

Smyth GK, Speed T (2003) Normalization of cDNA microarray data. Methods 31:265-273.

Sokolov BP, Jiang L, Trivedi NS, Aston C (2003) Transcription profiling reveals mitochondrial, ubiquitin and signaling systems abnormalities in postmortem brains from subjects with a history of alcohol abuse or dependence. J Neurosci Res 72:756-767.

Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D, Futcher B (1998) Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell 9:3273-3297.

Spence JP, Liang T, Habegger K, Carr LG (2005) Effect of polymorphism on expression of the neuropeptide Y gene in inbred alcohol-preferring and -nonpreferring rats. Neuroscience 131:871-876.

Spitzer JA, Zheng M, Kolls JK, Vande Stouwe C, Spitzer JJ (2002) Ethanol and LPS modulate NF-kappaB activation, inducible NO synthase and COX-2 gene expression in rat liver cells in vivo. Front Biosci 7:a99-108.

Stears RL, Martinsky T, Schena M (2003) Trends in microarray analysis. Nat Med 9:140-145.

Steinglass P (1999) Family therapy: Alcohol. In: The American Psychiatric Press Textbook of Substance Abuse Treatment, 2nd Edition (Galanter M, Kleber HD, eds). Washington, DC: American Psychiatric Association.

Storey JD (2003) The positive false discovery rate: A bayesian interpretation and the q-value. The annals of statistics 31:2013-2035.

Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 100:9440-9445.

Sugino K, Hempel CM, Miller MN, Hattox AM, Shapiro P, Wu C, Huang ZJ, Nelson SB (2006) Molecular taxonomy of major neuronal classes in the adult mouse forebrain. Nat Neurosci 9:99-107.

Swift RM (1995) Effect of naltrexone on human alcohol consumption. J Clin Psychiatry 56 Suppl 7:24-29.

Tabakoff B, Hoffman PL (2000) Animal models in alcohol research. Alcohol Res Health 24:77-84.

Tabakoff B, Bhave SV, Hoffman PL (2003) Selective breeding, quantitative trait locus analysis, and gene arrays identify candidate genes for complex drug-related behaviors. J Neurosci 23:4491-4498.

Taylor BA (1978) Origins of Inbred Mice. In: (Morse HC, ed), pp 423-438. New York: Academic Press.

Terdal ES, Crabbe JC (1994) Indexing withdrawal in mice: matching genotypes for exposure in studies using ethanol vapor inhalation. Alcohol Clin Exp Res 18:542-547.

Thibault C, Lai C, Wilke N, Duong B, Olive MF, Rahman S, Dong H, Hodge CW, Lockhart DJ, Miles MF (2000) Expression profiling of neural cells reveals specific patterns of ethanol-responsive gene expression. Mol Pharmacol 58:1593-1600.

Treadwell JA, Singh SM (2004) Microarray analysis of mouse brain gene expression following acute ethanol treatment. Neurochem Res 29:357-369.

Treadwell JA, Pagniello KB, Singh SM (2004) Genetic segregation of brain gene expression identifies retinaldehyde binding protein 1 and syntaxin 12 as potential contributors to ethanol preference in mice. Behav Genet 34:425-439.

Uylings HB, Groenewegen HJ, Kolb B (2003) Do rats have a prefrontal cortex? Behav Brain Res 146:3-17.

Vadigepalli R, Chakravarthula P, Zak DE, Schwaber JS, Gonye GE (2003) PAINT: a promoter analysis and interaction network generation tool for gene regulatory network identification. Omics 7:235-252.

Valenzuela CF (1997) Alcohol and neurotransmitter interactions. Alcohol Health Res World 21:144-148.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, et al. (2001) The sequence of the human genome. Science 291:1304-1351.

Victor M, Davis RD, Collins GH (1989) The Wernicke-Korsakoff Syndrome and Related Neurologic Disorders Due to Alcoholism and Malnutrition. Philadelphia: F.A. Davis.

Volpicelli JR, Alterman AI, Hayashida M, O'Brien CP (1992) Naltrexone in the treatment of alcohol dependence. Arch Gen Psychiatry 49:876-880.

Wallgren H (1970) Absorption, diffusion, distribution and elimination of ethanol: Effect on biological membranes. In: International Encyclopedia of Pharmacology and Therapeutics, pp 161-188: Oxford: Pergamon.

Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, Ganem D, DeRisi JL (2002) Microarray-based detection and genotyping of viral pathogens. Proc Natl Acad Sci U S A 99:15687-15692.

Wang H, He X, Band M, Wilson C, Liu L (2005) A study of inter-lab and inter-platform agreement of DNA microarray data. BMC Genomics 6:71.

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyras E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, et al. (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420:520-562.

Weiss F, Porrino LJ (2002) Behavioral neurobiology of alcohol addiction: recent advances and challenges. J Neurosci 22:3332-3337.

Weissenborn R, Duka T (2003) Acute alcohol effects on cognitive function in social drinkers: their relationship to drinking habits. Psychopharmacology (Berl) 165:306-312.

White AM (2003) What happened? Alcohol, memory blackouts, and the brain. Alcohol Res Health 27:186-196.

Wilke N, Sganga MW, Gayer GG, Hsieh KP, Miles MF (2000) Characterization of promoter elements mediating ethanol regulation of hsc70 gene transcription. J Pharmacol Exp Ther 292:173-180.

Wilkinson PK, Sedman AJ, Sakmar E, Kay DR, Wagner JG (1977) Pharmacokinetics of ethanol after oral administration in the fasting state. J Pharmacokinet Biopharm 5:207-224.

Wise RA, Bozarth MA (1987) A psychomotor stimulant theory of addiction. Psychol Rev 94:469-492.

Wong ML, Medrano JF (2005) Real-time PCR for mRNA quantitation. Biotechniques 39:75-85.

Woo Y, Affourtit J, Daigle S, Viale A, Johnson K, Naggert J, Churchill G (2004) A comparison of cDNA, oligonucleotide, and Affymetrix GeneChip gene expression microarray platforms. J Biomol Tech 15:276-284.

Worst TJ, Tan JC, Robertson DJ, Freeman WM, Hyytia P, Kiianmaa K, Vrana KE (2005) Transcriptome analysis of frontal cortex in alcohol-preferring and nonpreferring rats. J Neurosci Res 80:529-538.

Wu LF, Hughes TR, Davierwala AP, Robinson MD, Stoughton R, Altschuler SJ (2002) Large-scale prediction of Saccharomyces cerevisiae gene function using overlapping transcriptional clusters. Nat Genet 31:255-265.

Xu Y, Ehringer M, Yang F, Sikela JM (2001) Comparison of global brain gene expression profiles between inbred long-sleep and inbred short-sleep mice by high-density gene array hybridization. Alcohol Clin Exp Res 25:810-818.

Yang YH, Speed T (2002) Design issues for cDNA microarray experiments. Nat Rev Genet 3:579-588.

Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res 30:e15.

Ye Q, Koltchine VV, Mihic SJ, Mascia MP, Wick MJ, Finn SE, Harrison NL, Harris RA (1998) Enhancement of glycine receptor function by ethanol is inversely correlated with molecular volume at position alpha267. J Biol Chem 273:3314-3319.

Zar JH (1999) Biostatistical analysis, 4th Edition. Upper Saddle River, NJ: Prentice-Hall, Inc.

Zhang B, Kirov S, Snoddy J (2005) WebGestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic Acids Res 33:W741-748.

Zhang B, Schmoyer D, Kirov S, Snoddy J (2004) GOTree Machine (GOTM): a web-based platform for interpreting sets of interesting genes using Gene Ontology hierarchies. BMC Bioinformatics 5:16.

Vita

Kerrie Hall Lodowski was born in Lake Charles, Louisiana on March 13, 1975, the daughter of Randall Mac Hall and Mary Frances Hall. After graduating from Barbe High School in Lake Charles in 1993, she entered The University of Texas at Austin in Austin, Texas. She received a Bachelor of Science degree in Molecular Biology from The University of Texas at Austin in May 1997. During the following two years she was employed at The University of Texas Health Science Center in San Antonio as a Research Scientist in the laboratory of Dr. Chris Walter. In August 1999, she entered the Graduate School of The University of Texas at Austin as a student in the Cellular and Molecular Biology program. She is the primary author on a manuscript that was recently submitted to the *Journal of Neurochemistry*. She is also co-author on two other manuscripts published during her tenure as a graduate student. In addition, she has two co-author publications under Dr. Chris Walter. She has been a Teaching Assistant for three undergraduate classes during graduate school—Genetics (BIO325), Microbiology lab (MIC129K) and an Intensive Chemistry Seminar associated with Chemistry 301. She married David Thomas Lodowski on April 26, 2003.

Permanent address: 2703 Nottoway Lane, Lake Charles, Louisiana, 70605

This dissertation was typed by the author.