Discovery of Distinct Mechanisms Underlying the Relationship Between Drug Taking and Predisposing Behaviors

Tyler A. Roy
The Jackson Laboratory, tyler.roy@jax.org

Follow this and additional works at: https://digitalcommons.library.umaine.edu/etd

Part of the Animals Commons, Behavior and Behavior Mechanisms Commons, Genetic Phenomena Commons, and the Mental Disorders Commons

Recommended Citation
https://digitalcommons.library.umaine.edu/etd/3168

This Open-Access Thesis is brought to you for free and open access by DigitalCommons@UMaine. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of DigitalCommons@UMaine. For more information, please contact um.library.technical.services@maine.edu.
DISCOVERY OF DISTINCT MECHANISMS UNDERLYING THE RELATIONSHIP
BETWEEN DRUG TAKING AND PREDISPOSING BEHAVIORS

By

Tyler A. Roy

B.A. & B.S. University of Maine, 2015

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Biochemistry)

The Graduate School

The University of Maine

May 2020

Advisory Committee:

Elissa J. Chesler, Professor at The Jackson Laboratory, Advisor

Alan M. Rosenwasser, Professor of Psychology

Greg A. Cox, Associate Professor The Jackson Laboratory
Drug addiction is a heritable disease characterized by compulsive drug use. The biological mechanisms driving addiction remain largely unknown. Previous studies show shared genetic mechanisms underlying addiction risk phenotypes such as anxiety, depression, and novelty/sensation seeking. Therefore, high-throughput behavioral screening of these traits in single gene knockout mice can allow for the rapid detection of addiction risk candidate genes and mechanisms. Many of these traits are represented in the Knock-Out Mouse Program (KOMP) phenotyping pipeline. Of the initial two hundred twenty-one strains screened in this program, we tested nineteen phenodeviant knock-out mouse strains with C57BL/6NJ controls (N = 951) for effects on drug consumption and preference using a two-bottle choice paradigm with either ethanol (EtOH), methamphetamine (MA), or nicotine. Initial screening confirmed that fifteen of the nineteen gene deletions significantly affected EtOH consumption or preference (EtOH-related traits), MA consumption or preference (MA-related traits), or both. Thirteen strains exhibited drug specific effects and two exhibited significantly altered patterns of consumption, preference, or both for MA and EtOH. To investigate the shared relationships underlying drug consumption and predisposing drug-naïve phenotypes, we ran a principle component analysis. This revealed a
complex relationship among predisposing behaviors and their effects on drug-related phenotypes across different strains, instead of consistent predictive relationships between predisposing behaviors and drug consumption phenotypes. Using a multivariate strategy, a second screening approach was performed based on multidimensional phenodeviance across predisposing traits. In this analysis, based on four hundred two inbred strains, fifteen KOMP lines chosen based on phenodeviance, in addition to six expert nominated strains, and were assessed with controls (N=608) using both methamphetamine two-bottle choice and ethanol drinking-in-the-dark paradigms. Results from second screening revealed ten more genes affecting drug phenotypes, three of which altered both methamphetamine and ethanol related traits. This shows that using a multidimensional assessment of predisposing traits enriched our candidate genes with a higher rate of effects across multiple drugs. Collectively this project provided twenty-five new confirmed gene deletion mutants with addiction risk effects, representing multiple distinct, novel biological mechanisms of addiction.
# TABLE OF CONTENTS

LIST OF FIGURES .............................................................................................................................. v

INTRODUCTION .................................................................................................................................. 1

1.1. Addiction ...................................................................................................................................... 1

1.2. Neurocircuitry of Addiction ........................................................................................................... 2

1.3. SUDs Comorbidity with Mental Health Disorders .......................................................................... 3

1.4. Role of Genetics in Addiction .......................................................................................................... 5

1.5. Current Treatments and Potential for Pharmacotherapeutic Treatments ..................................... 7

1.6. Leveraging Mouse Models to Study Addiction Risk Behaviors ................................................... 8

1.7. Predisposing Drug-Naïve Behaviors .............................................................................................. 10

EXPERIMENTAL STRATEGY, MATERIALS, AND METHODS .......................................................... 12

2.1. Overarching Experimental Strategy .............................................................................................. 12

2.2. Animal Husbandry: Behavioral Phenotyping ................................................................................ 16

2.3. General Testing Procedures ........................................................................................................... 16

2.4. Open Field .................................................................................................................................. 17

2.5. Light/Dark ................................................................................................................................... 17

2.6. Holeboard .................................................................................................................................. 18

2.7. Acoustic Startle/Pre-pulse Inhibition ............................................................................................. 18

2.8. Tail Suspension ............................................................................................................................. 19
2.9. Identification of Predictive Behavioral Phenotypes: Initial Screening ........................................ 19

2.10. Subjects: Initial Screening ........................................................................................................ 19

2.11. Two Bottle Choice: Animal Husbandry .................................................................................. 21

2.12. Two Bottle Choice: Protocol .................................................................................................. 21

2.13. Data Analysis: Two Bottle Choice ......................................................................................... 22


2.15. Calculating Mahalanobis Distance of Predisposing Behaviors ............................................ 23

2.16. Subjects: Second Screen ......................................................................................................... 24


2.18. Drinking in the Dark: Protocol .................................................................................................. 26


2.20. Data Quality Control and Analysis of the Second Screen ...................................................... 27

RESULTS ........................................................................................................................................ 29

3.1. Results: Initial Screening of Drug-Naïve Behavioral Phenotyping ............................................ 29

3.2. Results: Initial Two Bottle Choice Screening............................................................................. 30

3.3. Results: Principle Component Analysis.................................................................................... 32

3.4. Results: Phenodeviant Strain Identification in Second Screening ............................................. 35

3.5. Results: Second Methamphetamine Two Bottle Choice Screening ......................................... 36

3.6. Results: Drinking in the Dark ..................................................................................................... 40

3.7. Results: Genes with Drug Specific or Multi-Drug Effects .......................................................... 43
4.1. Identification of Novel Addiction Risk Genes

4.2. Selection of Phenodeviant Strains: Initial Screening

4.3. Validation of Strain Identification Using Two Bottle Choice Paradigm

4.4. Functional Analysis of Novel Candidate Genes

4.5. Heterogeneous Behavioral Profiles Underlying Drug-Related Phenotypes

4.6. Multidimensional Assessment of Predisposing Behaviors

4.7. Refining Assessment of Drug-Related Phenotypes

4.8. Improved Identification of Genes Altering Methamphetamine-Related Phenotypes

4.9. Strains Assessed for Ethanol-Related Phenotypes Using DID Protocol

4.10. Identification of Phenodeviant Strains Using Multidimensional Analysis Improved our Ability to Identify Genes with Multi-Drug Effects

4.11. Conclusions and Future Directions

REFERENCES

APPENDIX

BIOGRAPHY OF THE AUTHOR
LIST OF FIGURES

Figure 1. Phenotypic allele variation in drug-naïve behaviors for initial screening……30
Figure 2. Shared associations of gene deletion mutations with alcohol, methamphetamine and nicotine intake phenotypes ……………………………32
Figure 3. Shared relationships among drug- naïve behaviors with drug-related phenotypes…………………………………………………………34
Figure 4. Multidimensional assessment of phenodeviance in drug-naïve behaviors …36
Figure 5. Single gene knockouts resulting in significant methamphetamine effects……39
Figure 6. Single gene knockouts and their effects on binge drinking .........................42
Figure 7. Bipartite graph depicting significantly altered drug-related phenotypes…….44
CHAPTER 1

INTRODUCTION

1.1. Addiction

Addiction is a complex brain disease that manifests both behaviorally as well as physiologically.\textsuperscript{4,5} Addiction is characterized as the compulsive seeking of rewarding stimulation despite reduction in euphoric effects and other negative consequences that may accompany this compulsive seeking for the individual.\textsuperscript{4–7} While addiction is not a disease that is part of the Diagnostic and Statistical Manual, version 5 (DSM-5), many of the defining traits associated with addiction are covered by the diagnostic criteria for Substance Use Disorders (SUD). Addiction is often used in correspondence with severe SUDs, but making a precise diagnosis of the severity of an SUD can be difficult even with the clearly defined diagnostic criteria set forth by the (DSM-5). The difficulty establishing a precise diagnosis is that diagnostic criteria are based on symptom counts, and it has been shown that while two individuals may check off the same number of diagnostic criteria for an SUD they can drastically differ on the severity of their addiction depending on which criteria they endorse.\textsuperscript{8} With the numerous biological and societal factors that go into each individual case of addiction or SUD, studying their basic underlying features of these diseases is critical to improve our understanding of them.

With increased prevalence of mental health issues and drug availability SUDs are highly prevalent diseases that cause devastating effects to both the individual and society. In 2018, the Substance Abuse and Mental Health Services Administration (SAMHSA) conducted a national survey that estimated 164.8 million (60.2\%) Americans had used alcohol, nicotine, or illicit drugs within the past month, and 20.3 million people ages twelve and up are currently suffering from an
SUD related to their use of alcohol or illicit drugs. The National Institute on Drug Abuse (NIDA) estimated that this use of drugs and alcohol costs Americans more than $700 billion and contributes to 570,000 deaths per year. Collectively, these data illustrate that as a result of increased access to drugs and alcohol SUDs are becoming drastically more prevalent diseases with devastating effects on the individual and society.

1.2. Neurocircuitry of Addiction

Repeated exposure to drugs of abuse, used by people with SUDs, lead to compulsive behavior through their pharmacological effects on dopaminergic neurons in the reward pathway common to all mammals. The regions of the brain most commonly associated with addiction and SUDs are the prefrontal cortex (PFC), extended amygdala, and the ventral and dorsal striatum, which include components of the mesolimbic pathway (reward pathway). Due to evolutionary importance these reward pathways have been highly conserved across many different species. These pathways likely evolved to increase Darwinian fitness through incentive salience (motivation) and hedonic pleasure (liking), but because drugs of abuse activate habit formation pathways, the motivation for drugs can persist even after the liking subsides.

The NAc is an ancient forebrain structure which may have originally developed to increase Darwinian fitness through incentive salience for rewarding stimuli such as food and sex. Today highly refined drugs of abuse hijack this system with over stimulation of dopaminergic neurons falsely signifying large benefits to our survival. The excessive release of dopamine from the terminals of VTA neurons within the NAC leads to motivational arousal in the individual that creates a drive to procure the drug again. Along with innervating the NAc, the VTA also innervates regions of the PFC that in conjunction with the amygdala make environmental and emotional cues
respectively to guide further decision about the motivation to procure the drug.\textsuperscript{5,16} As drug use continues, these cues and associations get stronger and incentive salience for the drugs continues to escalate as more cues become associated with their use. In addition to the increasing motivation for drug use, dopamine neurons in the substantia nigra (SN) translate the recurring reward signals into habitual actions that become increasingly insensitive to the reduction of hedonic response to continued drug use and instead use the reinforcement associated with prior experiences\textsuperscript{17,18}. Often at this point the euphoric feelings that originally were associated with drug use have subsided and further administration of drugs merely reduces the depressive feelings caused by changes to normal dopamine levels in the reward pathway so the individual can return to an emotional and physiological baseline \textsuperscript{7,15}. This dysregulation of the reduced hedonic response in the NAc and increased incentive salience, driven by positively reinforcing cues stored in the PFC and amygdala, can lead to emergence of the compulsive drug taking habits which underlie addiction and SUDS.\textsuperscript{5,7,11,12,17} With the multitude of neural processes and brain regions working in unison there are many mechanisms by which genetic and biological variations can alter their function. The slight alterations to different brain regions and neural processes can manifest themselves in many different behavioral and physiological manners leading to many mechanisms of addiction vulnerability in different people.

1.3. SUDs Comorbidity with Mental Health Disorders

Results from the 2018 Substance Abuse and Mental Health Service Administration (SAMHSA) survey show a high rate of comorbidity between SUDs and people that suffer from mental illness.\textsuperscript{9} This study estimated that currently 9.2 million people are suffering from both mental illness and an SUD. This estimate represents about twenty percent of all people suffering
from mental illnesses. The high rates of comorbidity between mental illness and SUDs are indicative of potential genetic, neurobiological, and behavioral traits underlying both of these complex diseases. Some believe these diseases are connected through the self-medication hypothesis, which posits that people take drugs as negative reinforcers that relieve symptoms of mental illness. However, self-administration of drugs also alters natural levels of neurotransmitters such as dopamine or GABA in regions of the brain, leading to or exacerbating mental health problems. Although the interplay of the disorders and trajectory between are not fully elucidated, it is clear that SUDs and mental health disorders are two strongly intertwined diseases. Using the connections between SUDs and many mental illnesses, studies have shown that many predisposing drug-naïve behaviors, personality traits, and co-occurring psychological conditions connected to mood disorders such as anxiety, depression, impulsivity, and sensation/novelty seeking are associated with the risk for and are also consequences of many aspects of SUDs. These connections between SUDs and mental illnesses can make it a vicious cycle of one disease predisposing individuals to the other disease leading to a vicious cycle that is hard to break. Although these traits and behaviors may be putting individuals at greater risk, many other factors must be taken into account when determining the overall risk of an individual developing an SUD. These factors may include socioeconomic status, family/peer structure, age, sex, genetic vulnerability, and other extrinsic influences. Finding the genetic and biological factors that underlie substance use and many of the traits associated with mental illness can provide biomarkers for screenings and potential therapeutic targets for treating these diseases.
1.4. Role of Genetics in Addiction

Many factors play a role in determining whether an individual will develop an SUD and become addicted to that substance. An individual who lives in an environment where there is no access to drugs is highly unlikely to develop an addiction to a drug, but few such places exist. Beyond environmental access to drugs and predisposing behaviors associated with drug use, not all individuals who initiate drug use become addicted to the substances they use implying that there must be some underlying biological or genetic vulnerability to addiction. Many studies in humans have shown that on average the heritability of substance addiction is around 50% \(^{12,27,29,30}\) with hallucinogen use being the least heritable 39% and cocaine use being the most heritable 72% according to data from twin studies.\(^1\)

Although genetic vulnerability to addiction and substance use is relatively high across many substances it can be challenging to identify and examine the effects of any particular candidate gene. Due to their highly heterogeneous nature and numerous confounding gene × environment interactions, it is therefore a daunting task to identify genes that underlie the vulnerability to SUDs and addictions. Historically, researchers have undertaken complex hypothesis driven studies to examine the role of a particular gene in often a specific substance use paradigm, leading to long, expensive studies that slowly advance our knowledge of the genetic factors playing a role in addiction based on already well-studied mechanisms. Much focus has gone into studying genetic components underlying substance-specific effects because these genetic components can play crucial roles in development of treatments for drug-specific SUDs. In addition to studying these substance-specific gene additional focus is being put on identifying and characterizing substance-non-specific genes which may underlie generalized substance use. These genes that affect use of multiple substances are important as well because they could play a role in
developing more generalized substance-use treatments.\textsuperscript{31} Additionally, it has been shown that genes predisposing for addiction risk traits like high intake or preference drinking differ from the genes involved in transition from substance use to addiction.\textsuperscript{32} Distinguishing the genetic factors that underlie addiction risk factors like high initial intake from those that underlie susceptibility to transition into addiction is a challenging task further slowing the pace for discovery.

In contrast, discovery approaches including Genome Wide Association Studies (GWAS) allow for the discovery of previously unknown genes in addiction. While GWA studies do not have the precise control over factors such as environment and experimental conditions, these studies can work directly with vast quantities of human data and therefore, genetic candidates are known to be involved in human disease, in contrast to findings from animal models, which need to be evaluated from a comparative perspective. Now, using the mass amount of data compiled in the human, GWAS researchers can use results from addiction phenotypes and biomarkers to identify genetic loci, individual genes, and sometimes even single nucleotide polymorphisms (SNPs) that can explain some of the heritability for addiction.\textsuperscript{27,28} While GWAS are a promising start, data from hundreds of thousands to millions of individuals is required to explore the genetics underlying addiction in humans, making the studies of many drugs and many populations cost-prohibitive.

Many of the factors that limit genetic discovery in human studies, including environmental factors, genetic background, medical history, and sample size are easily controlled when using animal models. The use of animal models, such as mice, can serve as a powerful research tool for studying many of the genetic and biological mechanisms underlying complex disease phenotypes.\textsuperscript{33–35} Animal models, in combination with advancing genetic engineering technologies, have been used to evaluate and characterize hypothesized gene disease associations from human
studies and provide therapeutic targets through the use of behavioral assays proven to have translational applications.\textsuperscript{35} Additionally, advancing high-diversity mouse populations such as the Diversity Outbred (DO) and Collaborative Cross (CC) stocks greatly increase the power to identify genetic factors and biological mechanisms underlying complex diseases when used in a systems genetics approach.\textsuperscript{33} Using these ever advancing animal models and technologies also allow for the discovery of high order interactions such as gene × sex interactions\textsuperscript{34} and gene × environment interactions\textsuperscript{36} which most human studies are underpowered to detect. Undoubtedly, complementary findings from both human and animal models will be required to identify and characterize genetic risk factors underlying our susceptibility to addictions and SUDs.

1.5. Current Treatments and Potential for Pharmacotherapeutic Treatments

Sadly, of the nearly twenty million people with SUDs, SAMSHA estimated that in 2018 90-97\% of people who needed treatment received no treatment of any type including medical, psychiatric, or support groups.\textsuperscript{10,37} Among the small percentage of people who do seek treatment, there is a 40-60\% relapse rate across all SUDs which is comparable to the relapse rates of many other well established chronic diseases.\textsuperscript{4,38} Due to the highly heterogeneous causes of SUDs and addictions, lack of precise diagnostic criteria for assessing true severity of the disease\textsuperscript{39}, and inadequate training of medical professionals in the field of substance use\textsuperscript{40} it is not surprising that there are extremely high rates of relapse.

Current treatments for addictions and SUDs are limited and often consist of a combination of behavior therapy, counseling, and medications depending on what substance an individual is using. Research for new pharmacotherapeutic treatments is ongoing, but currently the FDA has only approved medications for the treatment of alcohol, nicotine, and opioid use.\textsuperscript{5} New techniques
with promising results for the treatment of addictions and SUDs are being developed at increasing rates as the cost of these diseases continues to rise. One of these promising techniques include: transdiagnostic treatments, which look to treat the underlying conditions shared across behavioral and substance related addictions\textsuperscript{41}. One promising pharmacological treatment is the development of the drug Topiramate that shows efficacy across SUDs\textsuperscript{5,42}

One area of interest is the identification of genes and biological mechanisms that could lead to the discovery of novel therapeutic agents. These novel biomarkers could allow medical professionals to assess a person’s vulnerability to SUDs, identify which pharmacotherapeutic techniques could be the most beneficial to an afflicted individual\textsuperscript{43}, or even help determine the vulnerability to relapse\textsuperscript{44}. Through the discovery of novel addiction genes and biological mechanisms, there is the potential to accelerate the rate at which we can produce improved therapeutic treatments, both general and individualized.

1.6. Leveraging Mouse Models to Study Addiction Risk Behaviors

The identification of the biological mechanisms and genetic factors underlying all stages of addiction and SUDs is critical for the advancement of current treatments and prevention methods. The vulnerability to these diseases, and an individual’s trajectory of progression from abuse to addiction, is strongly influenced by both genetic and environmental factors.\textsuperscript{45,46} A combination of family and association studies have led to limited identification of genetically heritable factors underlying drug-specific\textsuperscript{47} and generalized\textsuperscript{48} addiction at various stages.\textsuperscript{45–47} Previous studies have shown that \textit{ADH1B}, \textit{ADH1C}, and \textit{ALDH} alter a person’s initial vulnerability to alcoholism\textsuperscript{47}; meanwhile, other studies have shown that \textit{DRD2} and \textit{DRD4} are associated with later stages of alcohol use such as alcohol dependence.\textsuperscript{45} Additionally, \textit{DRD4} is associated with dependence of
multiple other drugs including MA, suggesting that it potentially may be a genetic factor underlying general mechanisms of addiction. In addition to their associations with various stages of drug use and addiction, some of the identified gene variants such as (SERT, THP1, and TPH2) have also been independently linked to personality traits of comorbid diseases which increase their vulnerability to addiction.

Identifying these mechanisms and candidate genes in humans is a daunting task due the heterogeneous etiology for each individual afflicted with these diseases. With human studies it is almost impossible to separate a person from the innumerable environmental factors that go into the formation of their addictive disorder and study predisposing biological and behavioral factors separate from effects of the drug use. While, as stated earlier, GWAS have started to identify some promising candidate genes for biomarkers of addiction and SUDs they really has only started to scratch the surface of the genetic factors effecting addiction.

Because there is conservation of the addiction risk reward circuitry across species animal models can be used to help fill many of the gaps in knowledge left by the limited information available from human studies. Using an animal models, including mice, affords many complementary benefits to human studies. While the total scope of addiction or SUDs could never be perfectly modeled using a single mouse model due to the complexity of the diseases. Mouse models allow control over genetic background and environmental factors, and they allow access to neurobiological processes to study precisely defined addiction risk phenotypes. Leveraging the valuable tool of mouse genetics allows for a more rapid and precise discovery of new biomarkers and biological mechanisms of addiction risk behaviors.
1.7. Predisposing Drug-Naïve Behaviors

Genetic and genomic screens have been previously employed in mutant mouse lines in the hopes of identifying novel genetic biomarkers related to addiction and SUDs. One major historical challenge with these programs has been that they required a separate drug-naïve cohort of mice to avoid the effects of drug exposure on subsequent physiology and behaviors. These procedures, designed to screen candidates for their role in addiction risk phenotypes, required the application of expensive and complex experimental paradigms for the study of drug consumption and drug effects. In more recent high-throughput discovery-based approaches, by the International Mouse Phenotyping Consortium, large-scale screens which employ a single unified test battery are being used to characterize behavioral and physiological phenotypes of single gene C57BL/6NJ knockout strains.

Risk for and consequences of many aspects of addiction and SUDs are associated with other predisposing drug-naïve behaviors, personality traits, and co-occurring psychological conditions in humans. Mouse behavioral tests can be used to precisely model many aspects of these predisposing or co-occurring traits related to anxiety, depression, impulsivity, and novelty seeking. Previous studies have shown that these predisposing drug-naïve behaviors, which include approach-avoidance tasks, ‘behavioral despair’ assays, tolerance for delayed reward, and novelty seeking tasks can be used to predict drug-related phenotypes in the laboratory mouse. In a similar vein, studies using inbred mouse populations have revealed that there are shared genetic mechanisms driving predisposing behaviors and drug-related phenotypes across distinct drug classes. However, many of the genes underlying the shared genetic variation among drugs, alcohol, and predisposing phenotypes such as novelty seeking remain unknown. This study attempts to exploit these proven relationships traits to identify novel addiction risk genes and
potential addiction-related genes through the screening of single-gene knockout strains to identify
strains with extreme responses on phenotypes that are have been shown to be genetically correlated
with drug-related phenotypes.\textsuperscript{51}
CHAPTER 2

EXPERIMENTAL STRATEGY, MATERIALS, AND METHODS

2.1. Overarching Experimental Strategy

First, two hundred twenty-one KO strains (8F, 8M per strain) matched with C57BL/6NJ controls were tested in each cohort on a broad behavioral phenotyping pipeline as part of the JAX KOMP2 project (Sup Fig. 1). This phenotyping pipeline included open field, light dark, hole board, acoustic startle, and tail suspension; phenotypes from these five behavioral assays have been shown to predict substance use in mice.52,53,58–60 Next, from this initial group of two hundred twenty-one KO mouse strains, we identified the subset of KO strains (n = 19) that exhibited a significant difference from B6NJ controls on at least one of these five behavioral assays that predict substance use in mice. The measures we used to select lines and rationale for each is as follows:

1) Nose-pokes (Holeboard) - The total number of nose-pokes into the sixteen holes in the hole board testing arena during the single twenty-minute testing session. Nose-pokes in a hole board is one of several genetically distinct indexes of novelty seeking53 which are positive predictors of substance use in mice.56 High nose-pokes reflect high novelty seeking.

2) Time in light (LightDark) - The total amount of time, represented as a percentage of total testing time, during which the mouse spent on the light side of the two-chambered light dark apparatus during the single twenty minute testing session. Time in light is an index of anxiety52 and is positively genetically correlated with substance use in mice.57 Low time in light reflects high anxiety and more time in light reflects excessive risk-taking.
3) Time immobile (Tail Suspension) - The total amount of time a mouse is immobile while suspended by its tail during the five minute testing session. Time immobile is recognized as an animal model for efficacy of antidepressants\textsuperscript{58} which has been shown to increase chances of drug acquisition and maintenance.\textsuperscript{25} While there may be controversy over the validity of this model as a depression-like behavior, many researchers still use time immobile due to its predictive validity.\textsuperscript{61,62} A greater time immobile is considered a measure of higher levels of depression-like behavior.

4) Latency to immobility (Tail Suspension) - The total amount of time a mouse is actively moving while suspended by its tail before it becomes immobile during the five minute testing session. Latency to immobility is recognized as an animal model for the study of depression\textsuperscript{58} which has been shown to predict increased drug acquisition and maintenance.\textsuperscript{25} Shorter latency to immobility reflects greater depression-like behavior.

5) Percent PPI (Acoustic Startle) - Percentage of baseline startle response when lower-intensity ‘prepulse’ sounds precede a louder ‘pulse’ sound. Reduced PPI is a Research Diagnostic Criteria (RDOC) and endophenotype for multiple neuropsychiatric disorders\textsuperscript{59} including panic disorder (anxiety) which is positively genetically correlated with substance use in mice\textsuperscript{57}. Reduced %PPI is an index for increased anxiety.

6) Total rearing (Open Field) - The total amount of times a mouse rears or jumps during the twenty minute testing session in the open field arena. This endpoint serves as a measure for anxiety\textsuperscript{60,63} which is positively genetically correlated with substance use in mice.\textsuperscript{57} Decreased total rearing is an index for increased anxiety.

7) Total distance travelled (Open Field) - Total distance travelled during the twenty minute testing session in the open field arena. Distance traveled serves as a measure of
anxiety\textsuperscript{60,63} which is positively genetically correlated with substance use in mice.\textsuperscript{57} Decreased total distance travelled is an index of increased anxiety.

8) Slope of distance travelled (Open Field) - Slope of the best fit line measuring the change in distance traveled over the twenty minute test which is broken into four five-minute time bins. This measure is calculated to show habituation to the novelty which can be used to measure anxiety or exploratory/risk taking phenotypes.\textsuperscript{64} These behaviors are genetically correlated with substance use in mice.\textsuperscript{57}

9) Increased slope (increased habituation) is an index for decreased anxiety. Distance first five minutes (Open Field) - Total distance travelled traveled during the first 5 minutes of the twenty minute test in the open arena. Distance traveled in the first five minutes of open field serves as a measure of novelty reactivity and anxiety\textsuperscript{65} which has been shown to be predictive of initiation of drug use and progression to compulsive drug use\textsuperscript{66} High distance travelled reflects high novelty reactivity and low anxiety.

10) Center time (Open field) - The total amount of time spent in the center 40\% of total surface area in the open arena during twenty minute testing. Time spent in the center servers as a measure for anxiety\textsuperscript{60,66} and is positively genetically correlated with substance use in mice.\textsuperscript{57} High time in the center reflects low anxiety.

Although some strains had more extreme phenotypic values than those chosen, the validation of our primary screening experiment was limited to those lines that were maintained as live stocks and were available at the time of testing. Using a two-bottle choice assay, we quantified substance use (MA, EtOH, or nicotine) in mice from these nineteen KO strains (N=951) (Sup Table 1). The 2BC paradigm was chosen because it allows for the measurement of drug consumption (mg of drug consumed corrected for kg of mouse), drug preference (the percentage of milliliters of drug
consumed out of total consumption), along with total-drinking (total amount of milliliters consumed) phenotypes. The total-drinking phenotype is of interest because natural variation in fluid consumption may serve as a biomarker of future drug use. Perturbations of the caudate have been associated with both changes in fluid consumption such as hyperdipsia\textsuperscript{67} and also changes in drug consumption\textsuperscript{68,69} From this group of nineteen KO strains we identified strains that exhibited a significant difference from B6NJ controls on two-bottle choice of MA, EtOH, or nicotine. We then performed a functional analysis using GeneWeaver\textsuperscript{70} of gene deletions which significantly affected a drug-related phenotype. Next, we ran a principle component analysis in order to investigate the underlying shared correlation structure across addiction risk phenotypes and predisposing baseline behaviors.

Following completion and validation of the initial nineteen strains, the selection process was modified to identify the strains that displayed the most phenodeviance overall. The use of a multidimensional assessment of predisposing behaviors was based on our finding that most extreme neurobehavioral effects of gene deletion are also associated with drug-related phenotypes, and that rather than make specific univariate predictions, a single multivariate based prediction was more straightforward and allowed us to identify the most extreme deletion effects. We hypothesized that looking at overall deviance of strains could improve our selection criteria and produce more strains with drug-related phenotypes and a greater percentage of genes that have significant effects on multiple drugs instead of drug specific effects. Using a Mahalanobis distance calculation to calculate overall phenodeviant strains, we applied this multi-dimensional approach to the four hundred two KO strains tested in KOMP. With these calculations we identified and rederived fifteen of the most deviant strains and two expert nominated strains for assessment of
drug-related phenotypes. Using MA 2BC and EtOH DID protocols, we screened knockouts and controls (N=608) (Sup Table 2, 3) for addiction risk phenotypes.

2.2. Animal Husbandry: Behavioral Phenotyping

Mice of both sexes were obtained from The Jackson Laboratory Repository at 4 to 6 weeks of age and transferred to the Research Animal Facility housing area by a wheeled cart. All mice were tested on the adult JAX KOMP2 pipeline starting at 8 weeks of age (www.mousephenotype.org/impress/procedures/12). Upon completion of each test, mice were returned to their home cage and brought back to the mouse housing room unless otherwise stated. Mice were housed in duplex polycarbonate cages with bedding consisting of pine shavings in pressurized individual ventilated (PIV) racks. The housing room was climate-controlled 71± 3° F humidity range 50 ± 20% set to standard 12:12 light-dark cycle and mice were provided access to food (NIH315K52 chow, Lab Diet 6%/PM Nutrition, St. Louis, MO, USA) and acidified water *ad-libitum*. All procedures and protocols were approved by The Jackson Laboratory Animal Care and Use Committee, and were conducted in compliance with the National Institutes of Health Guideline for Care and Use of Laboratory Animals. All testing was conducted between hours 7am and 5pm of the light phase.

2.3. General Testing Procedures

Mice were subjected to a battery of tests of biological and behavioral endpoints.51 Tests were arranged by perceived stressfulness in an effort to minimize potential carry-over effects. For exact details on all fifteen weeks of testing see Supplemental Fig. 1. All runs within the phenotyping pipeline were run in a sex specific manner, e.g. each run would either consist of all males or all females. Upon completion of each test all chambers were cleaned with 70% ethanol
between mice, and then cleaned with a scent free 1% Virkon solution at the end of the day, followed by a 70% ethanol rinse to remove the residue left by the Virkon. Protocols match SOPs used at the time of testing (2013-2015) and current updated versions can be found at https://www.mousephenotype.org/impress/PipelineInfo?id=7.

2.4. Open Field

The open field arena is a clear Plexiglas or acrylic box (42x42x42cm) with a dark gray floor, illuminated at 150-200 LUX in a 10 × 9 ft. room. Zones of the arena were set as follows – center 40% of total surface area; four corner zones measuring 58.06 cm² each; peripheral zone measures 7.62 cm from the edge of the arena walls. The test was performed using eight Versamax Animal Activity Monitors (AccuScan Instruments, Inc; Columbus OH). The Friday prior to the testing week, mice were moved in their home cage from main housing room and placed on identical PIV racks within the testing room and allowed to habituate to the room for 3 days. On the day of testing, up to eight mice were placed randomly in one of eight empty arenas and allowed to explore for twenty minutes undisturbed. Activity during the test was recorded using AccuScan Versamax Animal Activity Monitors. Upon completion of the twenty minutes mice were removed from testing arena and returned to their home cage and placed back on PIV rack in testing room.

2.5. Light/Dark

The light-dark box consists of the AccuScan open field arena with an insert evenly dividing the arena into two equally sized compartments with the light side illuminated at 150-200 lux and the dark chamber at 4-7 lux. Mice were placed randomly in one of eight empty arenas on the light side and allowed to explore for twenty minutes undisturbed. Upon completion of the twenty
minutes mice were removed from testing arena and returned to their home cage and placed back on PIV rack in testing room.

2.6. Holeboard

The holeboard test serves to monitor the exploratory and inhibitory phenotypes over a defined period of time. The holeboard apparatus (AccuScan Instruments, Inc; Columbus OH) consists of an insert that fits into the AccuScan open field chamber consisting of 16 holes in a 4 × 4 grid. Ambient lighting (150-200 lux) was used to encourage exploration. Subjects were randomly placed into 1 of 4 empty chambers and allowed to explore for 10 minutes undisturbed. Upon completion of the 10 minutes mice were removed from testing arena and returned to their home cage and placed back in the housing room.

2.7. Acoustic Startle/Pre-pulse Inhibition

The experimental apparatus consists of a load cell platform below a restraint container linked to a piezoelectric transducer and amplifier and housed in sound attenuated chamber (San Diego Instruments, Inc; San Diego, CA). A sound generator interfaced to software generate audio stimuli. The testing procedure includes a 5 minutes acclimation period of background noise followed by the presentation of trial types of the startle response (120 dB, 40 msec duration) presented alone or in combination with prepulses at either xx, yy, or zz dB (20 msec duration) and presented in a pseudorandom order (Supplemental Table 4). % prepulse inhibition is calculated as the % change of the paired prepulse + pulse stimulus combination relative to the pulse alone.
2.8. Tail Suspension

The tail suspension test consists of a metal bar suspended on a clamp stand 36.6 cm above the lab bench in front of a neutral color background. Noldus Ethovision XT (Noldus, Inc; Leesburg, VA) behavioral tracking software interfaced with video camera was used to record individual mice during the test. A plastic cone cut from a 10 mL Falcon tube was placed over the tail to minimize the ability of mice to climb their tails and masking tape was used to gently attach the tail to the bar. The mouse was exposed to the suspension and recorded for a single 5 minute trial. Latency to immobility and time spent immobile were recorded.

2.9. Identification of Predictive Behavioral Phenotypes: Initial Screening

We performed analysis on traits that had at least a certain number of observations as stipulated by R/PhenStat bioconductor package (v 1.0.0).\textsuperscript{71} PhenStat is built on linear mixed-effects model where date of test is considered the random effect with sex, genotype and the interaction of sex and genotype information are fixed effects terms. Missing values were ignored. Individual animals were tested on each assay once. Mutants were compared statistically to C57BL6/NJ controls found in their own testing batch. Selection of knockout strains for further screening on addiction risk phenotypes was based on the criteria that they at least one significant hit (genotype p-value < 0.05) across the ten phenotypes in addition to availability of homozygous null mutant lines.

2.10. Subjects: Initial Screening

Nineteen strains that represented the range of observed phenotypes were selected for the two bottle choice protocol. Strains selected from KOMP Pipeline were as follows: C57BL/6NJ-
Btg2^tm1b(KOMP)Mbp^/2J, B6N(Cg)-C1qa^tm1b(EUCOMM)Wtsi^/3J, B6N(Cg)-C9^tm1.1(KOMP)Vlcg^/J, B6N(Cg)-Cfb^tm1.1(KOMP)Wtsi^/J, B6N(Cg)-Cp^tm1b(KOMP)Wtsi^/J, B6N(Cg)-Dnajb3^tm1.1(KOMP)Vlcg^/J, C57BL/6NJ-Dnase1I^em1(IMPC)^/J, B6N(Cg)-Epb41l4a^tm1b(KOMP)Mbp^/2J, B6N(Cg)-Far2^tm2b(KOMP)Wtsi^/2J, B6N(Cg)-Gipc3^tm1b(KOMP)Wtsi^/J, B6N(Cg)-Hdac10^tm1.1(KOMP)Mbp^/J, B6N(Cg)-Hspb2^tm1.1(KOMP)Vlcg^/J, B6N(Cg)-Htr1a^tm1.1(KOMP)Vlcg^/J, B6N(Cg)-Il12rb2^tm1.1(KOMP)Vlcg^/J, B6N(Cg)-Lpar6^tm1.1(KOMP)Vlcg^/J, B6N(Cg)-Parp8^tm1.1(KOMP)Wtsi^/J, B6N(Cg)-Pitx3^tm1.1(KOMP)Vlcg^/J, B6N(Cg)-Pnmt^tm1.1(KOMP)Vlcg^/J, B6N(Cg)-Rilpl2^tm1b(KOMP)Wtsi^/J. All strains were homozygous for their gene deletions. All mutants were tested relative to sex and age-matched control C57BL/6NJ mice.

All C57BL/6NJ controls were purchased from JAX production. C57BL/6NJ were maintained as their own line for the duration of our project to use as controls. While non-littermate controls are considered approximate controls, many of the historical problems with using non-littermate controls in testing were addressed in this project. JAX has rigorous genetic quality control and mutant gene genotyping programs to ensure the genetic background of JAX® Mice strains. In addition to the rigorous quality control JAX also employs to maintain integrity of the background strains these quality control measures are also employed to maintain integrity of the genotypes of strains with identified molecular mutations. All KO strains used in this project were created using C57BL/6NJ embryonic stem cells backcrossed to at least N5 leaving no flanking DNA that differs from controls and mutants. Similarly, all endonuclease modified lines used have no flanking DNA which differs from control strain. All strains in our initial screen were received directly from production directly for testing. This ensured that all strains tested met all of JAX requirements for rigorous quality control of background and mutations.
2.11. Two Bottle Choice: Animal Husbandry

Experimental mice were obtained from the JAX Repository and transferred to the JAX housing and phenotyping facility. Mice were group housed in duplex polycarbonate cages prior to testing. One day prior to testing, mice were individually rehoused in duplex polycarbonate cages with a single Shepherd Shack® and Nestlet® for the duration of testing. Duplex cages were all placed on conventional mouse racks. Cages were set up with two 50 mL sipper tubes on each half of the cage top. The sipper tubes were constructed from sterile 50 mL Polypropylene centrifuge tubes (#430291 Corning, Corning NY), fitted with a #6 one-holed rubber stopper (#14-135J Fischer-Scientific) and a 2.5 inch stainless steel sipper (Cat# SPS-SM tube 2.5 w/ball, Sta-Pure Systems, Carnegie PA). Mice were maintained in a climate-controlled room under a standard 12:12 light-dark cycle (lights on at 0600 hours and off at 1800 hours) and mice were provided free access to food (NIH315K52 chow, Lab Diet 6%/PM Nutrition, St. Louis, MO, USA) All procedures and protocols were approved by JAX Animal Care and Use Committee, and were conducted in compliance with the National Institutes of Health Guideline for Care and Use of Laboratory Animals.

2.12. Two Bottle Choice: Protocol

The 2BC paradigm used three different drugs at varying concentrations: EtOH (3%, 6%, 12%, 15%), MA (10 mg/L, 20 mg/L, 40 mg/L, 80 mg/L), and nicotine (10 mg/L, 20 mg/L, 40 mg/L, 80 mg/L). Individual mice were only exposed to one drug for testing. Drugs were presented in ascending order by concentration. Subjects (N=951) normal water was replaced with sipper tubes filled with water for a minimum of one day to acclimate. On day one of the testing protocol, one sipper of water was replaced with a sipper filled with one of the three drugs diluted in sterilized
acidified (pH 2.5-3) water (nicotine solution also contained 20 g/L saccharine sodium salt hydrate), while the second sipper remained filled with water and also 20 g/L saccharine sodium salt hydrate (if paired with nicotine as drug). Both sippers were then weighed using an Ohaus AV212 field weigh scale with a tolerance of .01g (OHAUS, Corp; Parsippany, NJ). Mice were given access to the two bottles for 48 hours. On day three, drug and water sippers were weighed for every mouse and switched to the opposite side of the cage to avoid side bias. Mice were given access to both bottles again for 48 hours. On day five, both drug and water sippers were weighed and replaced with new sippers. The water sipper was replaced with a freshly filled sipper of water and the drug sipper was replaced with the next dose in the curve. This process was repeated for all four doses, giving the mice access to the drug on each side of the cage for 48 hours.

Interpretation of 2BC results are as follows:

Significant consumption phenotype in the absence of preference phenotype suggests that there is an effect for general fluid intake usually accompanied by total-drinking phenotype suggestive of dysregulation of fluid intake. Significant consumption and preference phenotype suggests that there is an alteration to avidity for the drug and selective effect on drug intake as opposed to general fluid intake. Finally significant preference phenotype in the absence of a consumption phenotype indicates a selective effect on water intake in the presence of the drugs.

2.13. Data Analysis: Two Bottle Choice

Phenodeviant KO strains were assessed for significantly altered drug-related phenotypes in the EtOH, MA, and Nicotine 2BC protocols using a repeated measures ANOVA. KO strains were assessed for significant effects of strain, sex, dose, and the interaction between them relative to the C57BL/6NJ controls. A threshold of $\text{FDR} < 0.05$ was used to test for significance of terms in the model. Strain and sex were between-subjects factors and dose (3%, 6%, 12%, 15%) or (10 mg/L,
20 mg/L, 40 mg/L, 80 mg/L) were within-subjects factors. All values were computed using the average of two 48 hour sessions per dose. Data lost from one session, due to leaks, was not corrected for and the remaining value was used.

### 2.14. Principle Component Analysis

Investigation of the underlying shared correlation structure across drug consumption phenotypes and baseline behaviors was performed using a principle component analysis. All genotype difference estimates were subjected to a Van der Weerden (RankZ) transformation and subjected to principle component analysis using R (V 3.4.4) / factoextra _1.0.5, fviz_pca_biplot(). We extracted the first two PCs to assess the relationships among the predisposing behaviors, drug related behaviors and to assess the location of each knockout strain in this two-dimensional space constructed from the measured variables.

### 2.15. Calculating Mahalanobis Distance of Predisposing Behaviors

To assess overall phenodeviance measures from Open Field, Light/Dark, Holeboard, and Acoustic startle were used to calculate Mahalanobis distance for all tested strains compared to C57BL/6NJ controls. Mahalanobis identifies strains that are multivariate outliers through the calculation of the distance from the centroid, representative of control strain values, in a multidimensional space using the values from the nine predisposing behavioral traits. We used R/Phenstat bioconductor package (v 1.0.0) for modeling the association between trait and genotype. Genotype effect estimates were then used for Mahalanobis distance calculations following a rankZ transformation. Multivariate outliers were identified using the Mahalanobis distance. Outliers were defined by the distance of the data point from the calculated centroid of the other observations in multivariate space. The centroid is defined as the intersection of the mean
of the variables being assessed. The Mahalanobis distance follows a $\chi^2$ distribution, which is used to assess statistical significance.\textsuperscript{72} Tail suspension was not used in Mahalanobis calculations because it was dropped from KOMP2 testing and would have greatly reduced the number of strains with complete data for the second screening analysis. Using these results, we identified one hundred thirty-three as significantly phenodeviant from the C57BL/6NJ controls.

### 2.16. Subjects: Second Screen

Fifteen strains were selected for their significant phenodeviance: B6N(Cg)-

- $\textit{Elol}^{tm1.1(KOMP)Wtsi/J^{+/+}}$, C57BL/6NJ-$\textit{Stk36}^{em1l(IMPC)J/J^{+/+}}$, C57BL/6NJ-$\textit{Myh10}^{em1l(IMPC)J/J^{+/+}}$,
- B6N(Cg)-$\textit{Dnmt3a}^{tm1b(KOMP)Wtsi/J^{+/+}}$, B6N(Cg)-$\textit{Cp}^{tm1b(KOMP)Wtsi/J^{+/+}}$, B6N(Cg)-
- $\textit{Zbtb4}^{tm1.1(KOMP)Wtsi/J^{+/+}}$, B6N(Cg)-$\textit{Dnaja4}^{tm1b(KOMP)Wtsi/J^{+/+}}$, B6N(Cg)-$\textit{Irf8}^{tm1b(KOMP)Wtsi/J^{+/+}}$,
- B6N(Cg)-$\textit{Htr7}^{tm1b(KOMP)Wtsi/J^{+/+}}$, B6N(Cg)-$\textit{Gpr142}^{tm1.1(KOMP)Wtsi/J^{+/+}}$, B6N(Cg)-$\textit{C3}^{tm1.1(KOMP)Wtsi/J^{+/+}}$,
- B6N(Cg)-$\textit{Rap2b}^{tm1.1(KOMP)Wtsi/J^{+/+}}$, B6N(Cg)-$\textit{Tmod2}^{tm1b(KOMP)Wtsi/J^{+/+}}$, B6N(Cg)-
- $\textit{Stx19}^{tm1.1(KOMP)Wtsi/J^{+/+}}$, B6N(Cg)-$\textit{Lrrc15}^{tm1b(KOMP)Wtsi/J^{+/+}}$. Six expert nominated strains were also tested.

Expert nominated strains derived from the C57BL/6NJ line were B6N(Cg)-

- $\textit{Nr2f1}^{tm1.1(KOMP)Mbp/J^{+/+}}$, C57BL/6NTac-$\textit{Trib3}^{tm1a(EUCOMM)Hmg4/JcsOrl/Ejc^{+/+}}$. In addition to C57BL/6NJ strains expert nominated strains derived from the C57BL/6J line were C57BL/6J-$\textit{Bmerb1}^{tm1Ejc/J^{+/+}}$, C57BL/6J-$\textit{Taar1}^{tm1Ejc/J^{+/+}}$ (B6J-$\textit{Taar1}$) and C57BL/6J-$\textit{Fam53b}^{tm5Ejc/J^{+/+}}$. All C57BL/6J derived strains were tested against contemporaneously characterized C57BL/6J controls. Expert nominated strain derived from the DBA/2J line was DBA/2J-$\textit{Taar1}^{tm2Ejc/J^{+/+}}$ (B6KI) (DBA-$\textit{Taar1}$). All DBA/2J derived strains where tested against DBA/2J controls. C57BL/6J-$\textit{Taar1}^{(D2KI)}$ and DBA/2J-$\textit{Taar1}^{(B6KI)}$ are both CRISPR knock-in (KI) strains with functional versions of $\textit{Taar1}$ alleles swapped.
All mutants were tested using sex and age matched C57BL/6NJ controls that were housed in the same breeding room as the mutant colonies. All mutants in the second screen were procured in the same manner as strains in our initial screen with the exception that mutants were received as heterozygous mutants and bred in lab to homozygous when possible. Strain maintained in our colony were bred to no greater than F5 to avoid genetic drift from production colonies.

2.17. Drinking in the Dark/ Second 2BC: Animal Husbandry

Chosen strains were obtained from the JAX Breeding and Rederivation Services and transferred to the JAX housing and phenotyping facility where they were bred to testable cohort size through pair and trio mating. Homozygous viable strains were bred using hom × hom breeding pairs or trios (1M, 2F). Homozygous lethal strains were bred using het × het breeding pairs or trios. Mice were housed in duplex polycarbonate cages with bedding consisting of pine shavings in pressurized individual ventilated (PIV) racks. The housing room was climate-controlled 71± 3° F humidity range 50 ± 20% set to standard 12:12 light-dark cycle and mice were provided access to food (NIH315K52 chow, Lab Diet 6%/PM Nutrition, St. Louis, MO, USA) and acidified water ad-libitum. All procedures and protocols were approved by JAX Animal Care and Use Committee, and were conducted in compliance with the National Institutes of Health Guideline for Care and Use of Laboratory Animals. Mice were group housed in duplex polycarbonate cages prior to testing. Twenty-four days prior to testing all mice were housed individually in duplex polycarbonate cages with a single Shepherd Shacks® and Nestlets® and remained so throughout all testing. Duplex cages were all placed on conventional mouse racks. Cages were set up with standard Lab Diet 6%/PM Nutrition, and acidified water ad libitum. Mice were maintained in a climate-controlled room under a reversed 12:12 light-dark cycle (lights on at 2000 hours and off
at 800 hours). All procedures and protocols were approved by JAX Animal Care and Use Committee, and were conducted in compliance with the National Institutes of Health Guideline for Care and Use of Laboratory Animals. At time of testing acidified water was replaced with 20% EtOH sipper for two to four hours. Sippers were created from 2.5 inch stainless steel sipper (Cat# SPS-SM tube 2.5 w/ball, Sta-Pure Systems, Carnegie PA), 10 mL Disposable Polystyrene Serological Pipette (Fischer Scientific, Agawam MA), and shrink wrap.

2.18. Drinking in the Dark: Protocol

All strains chosen for this second screening were tested for MA and EtOH phenotypes. The MA 2BC protocol was left unchanged, but for EtOH the Drinking in the Dark (DID) paradigm was used instead. DID is a protocol that was developed as a binge drinking paradigm in mice. Binge drinking is classified as dangerous amounts of alcohol consumption leading to a blood ethanol content (BEC) 80 mg/dL which leads to an increased risk of alcohol use disorder. Normally, mice do not drink sufficient EtOH to become intoxicated. Using limited access during a time in the circadian cycle when mice are behaviorally active results in a level of consumption leading to intoxication. From this protocol we can get measures of four day consumption and measures of BEC.

Testing was done using a modified DID protocol. All testing is done in a dark room, with red lights. The DID paradigm used 20% EtOH made from mix of 100% EtOH and acidified drinking water. Individually housed transgenic mice and controls (N=608) chosen for testing were habituated for twenty-four days to reverse light cycle and weighed before lights off on first of four day of testing. Weights were recorded for all mice on the first day of the protocol during the lights on phase. New EtOH solution was prepared every day of testing. At 1100 hours the initial starting volume for each sipper were measured to the closest .05 mL along with initial starting weights
which were measured to the nearest .01 g and recorded in excel. Following initial recordings of volume and weight, sippers were put into cages in place of water with a two minute stagger between mice. Once all mice had received sippers, two control sippers were set up in an empty cage to be used as standards for evaporation and leakage. On days one through three, the mice had access to EtOH for two hours. At 1300 hours sippers were collected, weighed, and acidified water was returned to cage. Sippers were collected and weighed at the same pace they were distributed. On day four mice had access to EtOH for four hours. Sippers were administered to mice one at a time with a three minute stagger between mice. At 1500 hours sippers were collected for assessment of final volumes and weights. Finally, as sippers were removed, blood samples were taken from mice using a cheek bleed technique for assessment of BEC.


A minimum of 50 µL of blood was collected into a microtainer (purchased from VWR, cat.# VT365956) immediately following ethanol removal. Blood samples were then spun at 13,300 RPM for 11 minutes, and serum was pipetted into separate Eppendorf tubes on dry ice within one hour of collection. Once all samples were collected they were transferred to a -80°C freezer within three hours of blood being drawn. Samples were brought to JAX Clinical Chemistry for analysis.

2.20. Data Quality Control and Analysis of the Second Screen

Removal of outliers was done on a strain × sex × dose basis. Means were calculated on a strain × sex × dose basis and data points which fell more than two standard deviations outside the mean were considered outliers and removed.

We applied a repeated measures ANOVA to each of the MA phenotypes to assess the strain × sex × dose effects, in addition to assessing strain and strain × sex effects. After fitting each
model, we obtained the least squares mean difference between each knock-out relative to the C57BL/6NJ controls. A threshold of FDR $< 0.05$ was used to test for significance of terms in the model. Of particular interest were strain, strain $\times$ dose, and strain $\times$ dose $\times$ sex effects.

A repeated-measures ANOVA was also applied to EtOH consumption in DID protocol to assess the strain $\times$ sex effects, in addition to assessing strain effects. Following the model fit, we obtained least squares mean difference between each knock-out relative to the C57BL/6NJ control. A threshold of FDR $< 0.05$ was used to determine significance of terms in the model. Of particular interest were strain, and strain $\times$ sex effects.

Analysis of BEC was done using a two-way ANOVA. Tukey’s Honestly Significant Difference test was used to test for significant difference between each knock-out relative to the C57BL/6NJ control.
CHAPTER 3

RESULTS

3.1. Results: Initial Screening of Drug-Naïve Behavioral Phenotyping

To assess performances on the ten predisposing behaviors measured in the KOMP2 pipeline, we performed analysis using R/PhenStat bioconductor package (v 1.0.0) on strains which met the threshold for minimum number of observations for each test. Strains with insufficient replicates were omitted from the analysis. Results from the initial screen of predisposing behaviors revealed that about two thirds of the strains (64.7%; n= 143) were phenodeviant (deviant from C57BL/6NJ controls on at least one behavioral trait) on at least one of the ten predictive behaviors p< .05. Strains with extreme phenotypes and or multiple phenodeviant behaviors were prioritized in our selection process, but selection was limited to strains that were alive and available at the time of testing. Initial screening of available KOMP2 data led to the selection of nineteen strains to assess for drug-related phenotypes. While selected strains do not represent the most phenodeviant extremes in all cases, they do show a representative spectrum of observed phenodeviant behaviors (Fig.1).
Results: Initial Two Bottle Choice Screening

The selection process successfully identified single gene KO strains with differential drug-related phenotypes relative to controls. Fifteen of the nineteen strains resulted in a significant phenotypic difference of strain, or strain × dose across either MA or EtOH. (Fig 1a,b). Of the
eighteen strains tested on the EtOH 2BC protocol, slightly over three quarters of them (n=14; 77.8%) displayed significantly altered EtOH-related phenotypes either through altered preference or consumption (mg/kg). Six of these fourteen strains showed significant changes in both consumption and preference. Single gene deletions led to varying profiles of EtOH consumption, *Il12rb2* and *Far2* exemplify the range of diverse effects measured in the 2BC protocol (Fig. 1c-f). Deletion of *Il12rb2* resulted in increased EtOH preference (*F* strain (1, 90) = 35.7, FDR= 8.63E-07) and also increased EtOH consumption (*F* strain (1, 90) = 88.48, FDR=9.28E-14) compared to C57BL/6NJ controls (Fig. 1c,d). Deletion of *Far2* resulted in the opposite effects and significantly decreased both EtOH preference (*F* strain (1,87) = 12.31, FDR=3.40E-3) and EtOH consumption (Sup Panel 1). (*F* strain (1,87) = 9.3, FDR=6.40E-3). For most strains, females displayed higher levels of preference and consumption for EtOH than males (Sup Panel 2).

The initial screening process was also successful at selecting strains which displayed altered MA use. Seventeen strains were tested for altered MA-related phenotypes and just under one fifth of the strains (n=3; 17.8%) showed significantly altered preference (Sup Panel 3) or consumption (Sup Panel 4). While *Il12rb2* and *Far2* deletions led to opposing EtOH phenotypes, both deletions led to significantly increased consumption of MA (*F* strain (1, 71) = 9.88, FDR=2.19E-02) and (*F* strain (1, 69)= 10.59, FDR=2.19E-02) respectively. (Fig. 2 g, h) The only deletion to result in an altered MA preference was *Dnase1l2* (*F* strain × dose (3, 207) = 5.07, FDR=3.75E-02) which drastically reduced preference for MA at the initial dose.

Results from initial 2BC screening indicate that none of the gene knockouts altered any nicotine related phenotypes in the 2BC protocol. Knockouts showed non-significantly altered nicotine consumption, preference, or change in total amount consumed between nicotine and water.
3.3. Results: Principle Component Analysis

A principle component analysis was run to assess the relationships within and between our ten measured predisposing traits and our three drug-related traits across three drugs. A bi-plot clustering was then performed using the effect sizes across predisposing and drug-related traits of the sixteen knockout strains which were tested on all nineteen measures. Figure 3 depicts the
results of our PCA where each gene knockout falls in the multivariate space of our two strongest principal components score vectors. Results from our PCA reveal that using our first two principal components we can account for 21.1\% of the variance with principal component 1 (PC1) and 17.8\% with principal component two (PC2). Using both these PCs we can account for ~ 39\% of the variation seen in our sixteen strains across the twenty-two measures. The correlations within and between any of the predisposing behavioral and drug-related traits can be assessed using the angle which separates any two vectors (Fig. 3). Vectors which fall close to one another are strongly positively correlated, vectors which fall 180° apart are strongly negatively correlated, and vectors which fall 90° apart are independent of one another. Using these relationships we can use the PC1 axis to separate our tested strains into ones which increased EtOH consuming/preferring phenotypes from those who decreased EtOH consuming/preferring phenotypes. Along the PC2 axis we see clustering of predisposing traits which can be used to divide our strains into different baseline behavioral profiles i.e. “low anxiety” and “exploratory” profiles. Also along PC2 we see a close relationship which can separate our MA consuming strains from our non-consuming strains. Scores for each of the strains are obtained by multiplying the PC loadings by the strain means, allowing strains to be plotted in the two dimensional space. Strains with high absolute scores on PC1 such as Hdac10 and Dnajb3 have strong EtOH preferring/consumption phenotypes. Strains such as Il12rb2 and Far2 whose variation is strongly explained by PC2 would be strains with stronger MA phenotypes.
Figure 3. Shared relationships among drug-naïve behaviors with drug related phenotypes. Principal components analysis was used to assess shared variance among predictive behavioral traits and drug related phenotypes. Each point represents a KO strain, while arrows represent drug-naïve and drug intake related traits. We observe clusters of gene KOs along with clusters of drug-naïve and drug intake traits indicative of shared relationships. PC1 explains 21.1% of the shared variance and helps differentiate KO strains with ethanol drinking from those that display a non-ethanol drinking phenotype. PC2, which explains 17.8% of the variance, allows us to categorize and relate ethanol drinking KO strains to drug-naïve behavioral profiles of risk taking versus exploration. These relationships reveal the multiple mechanisms by which basal behavioral variations are associated with drug intake behaviors. This shared variation allows us to exploit baseline behaviors towards identification of novel drug intake genes. Analysis was run using only the sixteen strains that were tested on all behavioral phenotyping measures along with being tested on all three drugs. (Colors are added to drug phenotypes to match color scheme from figure 2)

3.4. Results: Phenodeviant Strain Identification in Second Screening

Knockout strains were selected for second screening based on overall phenodeviance across traits from the open field, light dark, holeboard, and acoustic startle. As of April 2016, four hundred two strains had completed the phenotyping pipeline and were included in the analysis. Calculation of overall phenodeviance was done using Mahalanobis Distances\textsuperscript{72} of effect sizes relative to C57BL/6NJ controls. Results from analysis indicate that of the four hundred two single gene knockout strains tested, one hundred twenty-three strains were significantly phenodeviant with scores ranging from 24.1- 2038.9. Strains with the highest scores were prioritized for rederivation and screening for drug-related phenotypes. Thirteen of the most extremely phenodeviant strains were successfully rederived and bred for testing (Fig. 4). Two additional KOMP strains on the lower end of the significant scores (288 and 115) were selected for a combination of phenodeviance and prior biological evidence for their role in drug-related phenotypes. Finally, six other expert nominated strains were chosen for testing.
3.5. Results: Second Methamphetamine Two Bottle Choice Screening

In our initial screening of the seventeen strains tested on MA 2BC, three strains had altered consumption or preference attributed to main effects of strain or strain × dose respectively (FDR < 0.05). Sex had no significant effect on any of the measures of MA use; sex only had a main effect on the total-drinking phenotypes (mLs/kg). Results from second screening for predicted MA
2BC effects indicates that eight of the fifteen tested strains had significantly altered patterns of MA use, either manifested through a main effect of strain and or strain × dose across preference and consumption (mg/kg) phenotypes (FDR <.05). Six of the fifteen gene deletions resulted in a main effect of strain and increased preference for MA in various patterns compared to the C57BL/6NJ strain. Five gene deletions resulted in strain × dose effects, three of which also had main effects of strain, altered preference to MA in a dose dependent manner compared to the control strain.

Three gene deletions, Rap2b, Tmod2, and Irf8 all showed an increase in MA consumption compared to the wild type control. Deletion of Irf8 gene resulted in increased preference for MA across doses (F_{strain} (1, 51) = 12.34, FDR= 5.33E-03) compared to control strains (Fig 5A). Deletion of Irf8 also significantly increased preference to initial lower doses of 10 mg/L and 20 mg/L 55.1% ±6.5 and 40.0% ±6.4 respectively, and did not display differences from control strains at higher doses (F_{strain × dose} (3, 140) = 3.69, FDR= 4.59E-02). While all three strains had increased overall patterns of consumption and were also strains which had significantly increased overall preference, the gene deletion which most significantly increased MA consumption was Rap2b that displayed increased consumption on every dose (F_{strain} (1, 51) = 15.14, FDR= 4.95E-03) (Fig 5b). In total, eight of the fifteen strains altered MA-related phenotypes, either through changes in preference or consumption, and an additional five knockouts altered total-drinking phenotype (FDR <.05). One of the two expert nominated C57BL/6NJ lines display significant alterations to MA consumption or preference. Trib3 displayed increased consumption at the final dose compared to controls (F_{strain × dose} (3, 152) = 3.09, p= 2.88E-03). Nr2f1 strain displayed no significant MA phenotypes.

Of the five non C57BL/6NJ expert nominated strains, three strains had significantly altered MA phenotypes. DBA/2J-Taar1^{B6KI} and C57BL/6J-Taar1^{D2KI} strains each displayed significant
alterations of both MA consumption (mg/kg) and preference (p<.05) (Fig 5C-F). DBA/2J-Taar1\textsuperscript{(B6K)} female KIs rapidly decreased their preference to MA as doses increased, while male preference only gradually decreased as doses increased (F\textsubscript{sex} (1, 28) = 4.82, p= 3.65E-02). KI of B6 alleles into the DBA/2J strain also resulted in a significantly decreased preference to MA (F\textsubscript{strain} (1, 28) = 17.66, p= 2.43E-04). (Fig. 5C) No main effect of sex was detected for MA consumption. DBA/2J-Taar1\textsuperscript{(B6K)} had significantly decreased MA consumption compared to control DBA/2J strains (F\textsubscript{strain} (1, 28) = 31.79, p= 4.86E-06). As drug concentration increased Taar1 mutants had a decrease in consumption compared to the DBA/2J control (F\textsubscript{strain × dose} (3, 78) = 7.39, p= 2.01E-04) (Fig.5D).

KI of D2 alleles into the C57BL/6J strain resulted in a significantly increased preference for MA, (F\textsubscript{strain} (1, 19) = 20.63, p= 2.23E-04) (Fig. 5E). C57BL/6J-Taar1\textsuperscript{(D2K)} displayed increased preference relative to controls as the dose increased. (F\textsubscript{strain × dose} (3, 55) = 4.7, p= 5.40E-03). No main effect of sex was detected for MA consumption. C57BL/6J-Taar1\textsuperscript{(D2K)} also resulted in significantly increased MA consumption compared to C57BL/6J control strain (F\textsubscript{strain} (1, 19)= 31.5, p= 2.06E-05). The C57BL/6J-Taar1\textsuperscript{(D2K)} strain showed an increase of MA consumption compared to the C57BL/6J controls in a strain by dose fashion (F\textsubscript{strain × dose} (3, 53) = 14.63, p= 4.63E-07) (Fig.5F).

Expert nominated strain Bmerb1 showed increased consumption at the 40 mg/L dose and decreased at 80 mg/L, 2.65 ± 0.46 mg/kg and 3.29 ± 0.39 mg/kg respectively, compared to the control strain 1.80 ± 0.16 mg/kg and 4.60 ± 0.45 mg/kg respectively (F\textsubscript{strain × dose} (3, 54) = 3.32, p= 2.64E-02) (Sup Panel 5). Additionally, Bmerb1 displayed increased preference at 40 mg/L and decreased preference at 80 mg/L, 32.7 ± 2.3% and 16.2 ± 2.2% respectively compared to controls 20.7 ± 7.1% and 28.8 ± 4.2% respectively (F\textsubscript{strain × dose} (3, 53) = 4.5, p= 6.85E-03) (Sup Panel 6).
Finally, *Bmerb1* was the only strain to display a sex difference on a MA-related phenotype. Male *Bmerb1* had an overall increased preference for MA compared to females ($F_{sex} (1, 18) = 8.68$, $p=8.64E-03$)

![Figure 5: Single gene knockouts resulting in significant methamphetamine effects.](image)

A,B, Line graphs of preference for MA by dose. Single gene deletions of *Tmod2* and *Rap2b* both resulted in increased preference for MA with main effects of strain. FDR<.05 C,D, Swapping of the *Taar1* allele from non-functional to functional version in DBA/2J background alleviated increased MA preference and consumption across all doses. E,F, Swapping of the *Taar1* allele from functional to non-functional version in C57BL/6J induced the increased preference and consumption across all doses normally observed in the high MA consuming DBA/2J lines.
3.6. Results: Drinking in the Dark

Analysis of DID results indicate that six of the fifteen phenodeviant strains displayed significantly altered consumption (mg/kg). EtOH consumption (mg/kg) was strongly influenced by sex in all strains with a minimum main effect of sex ($F_{\text{sex}} (1, 47) = 40.33$, FDR= 7.92E-08). Two phenodeviant strains, *Irf8* (Fig 6A) and *Elof1* (Sup Panel 7), had significantly altered EtOH consumption across the four days of access ($F_{\text{strain}} (1, 48) = 13.38$, FDR= 5.37E-03) and ($F_{\text{strain}} (1, 48) = 7.6$, FDR= 4.66E-02) respectively. Additionally for two strains, *Zbtb4* (Fig 6B) and *Tmod2*, EtOH consumption was influenced by strain and sex ($F_{\text{strain} \times \text{sex}} (1, 48) = 11.72$, FDR= 2.17E-02) and ($F_{\text{strain} \times \text{sex}} (1, 47) = 8.39$, FDR= 4.84E-02) respectively. Similar to what was observed in EtOH consumption phenotype, BEC was also influenced by sex in all strains as indicated by a minimum main effect of sex ($F_{\text{sex}} (1, 46) = 6.31$, p= 1.56E-02). Of the fifteen phenodeviant strains chosen for testing in the DID paradigm no strain resulted in a significantly altered BEC.

Of the five expert nominated strains, only *Nr2f1* affected EtOH phenotypes in the DID paradigm. Deletion of *Nr2f1* resulted in significantly decreased consumption compared to control strain ($F_{\text{strain}} (1, 47) = 38.21$, FDR= 2.43E-06), and also significantly decreased BEC following four hour access on the final day of testing ($F_{\text{strain}} (1, 42) = 5.35$, p= 2.51E-02).

All C57BL/6J derived strains reach BEC levels above 100 mg/dL, but no strain was significantly different from control strains as indicated by significant effect of strain. The only observed difference was increased BEC levels in Male *Fam53b* males ($F_{\text{strain} \times \text{sex}} (1, 29) = 4.46$, p= 4.34E-02). This resulted in an increase in average BEC from 54.3±17.6 mg/dL in controls to 105.2±13.5 mg/dL in KOs.

The DID protocol was developed to induce binge-like drinking in non-dependent C57BL/6J mice. In our experiment C57BL/6J consumed on average 7.21±0.54 mg/kg on day four.
resulting in an average BEC of 108.7 ± 16.8. On the final day of testing, female C57BL/6J consumed more and resulted in higher BEC 9.58 ± 0.37 mg/kg and 163.2 ± 15.0 mg/dL respectively, than males 4.84 ± 0.39 mg/kg and 54.3 ± 17.6 mg/dL respectively. Results from our study show the protocol did not induce C57BL/6NJ to reach translationally relevant levels of BEC. C57BL/6NJ consumed on average 6.38 ± 0.54 mg/kg on day four resulting in an average BEC of 51. ± 8.5 mg/dL. Similar to C57BL/6J, female C57BL/6J consumed more and resulted in higher BEC 9.03 ± 0.44 mg/kg and 76.8 ± 18.1 mg/dL respectively, than males 3.73 ± 0.37 mg/kg and 25.2 ± 5.95 mg/dL respectively. For overall consumption, final day consumption, and BEC, measures from C57BL/6J were significantly higher than C57BL/6NJ (F_{strain \ (1, \ 52)} = 24.05, p=9.64E-06), (F_{strain \ (1, \ 52)} = 11.69, p=7.49E-04), and (F_{strain \ (1, \ 52)} = 20.32, p=3.76E-05) respectively.
Figure 6. **Single gene knockouts and their effects on binge drinking** A,B. Line graph of average strain consumption (mg/kg) by day in the DID protocol separated by sex. Both gene deletions also significantly altered total ethanol consumption across the four day DID protocol. Deletion of *Irfl8* resulted in a decrease consumption across both strains (main effect of strain FDR < .05) which was more extreme in the females. Deletion of *Zbtb4* had a significant strain x sex effect and decrease female consumptions while mildly increasing male consumption. FDR < .05. C. Two dimensional circle graph summarizing the effects of the 17 single gene deletions on average BEC’s following DID protocol. X dimension represents mean male BEC, and Y dimension represents mean Female BEC average. Grid lines intersect at the average female and male BECs for C57BL/6NJ controls. The radius of the circle equals the standard error for each strain black outlines for females and colored outlines for males. Star represent significant strain effect for BEC measurements.
3.7. Results: Genes with Drug Specific or Multi-Drug Effects

Using a multidimensional assessment of phenodeviance across predisposing drug-naïve behaviors we identified fifteen single gene KO strains as phenodeviant, and tested them for altered patterns of drug-related phenotypes. Results from MA 2BC and EtOH DID revealed that two thirds (66.6%, 10/15) of the identified single gene deletions altered drug-related phenotypes. Further assessment of the ten gene deletions that altered drug-related phenotypes reveals that thirty percent (30%, 3/10) had multi-drug effects across both drugs and testing protocols. Results from the second screening indicate that ten strains showed altered patterns of drug-related phenotypes, seven of which were drug specific, and three of which were significant effects across both drugs and testing paradigms (Fig 7).

Expert nominated strains Bmerbl and Fam53b were only tested on one of the two protocols and could not be assessed for multi-drug effects. Of the six expert nominated strains, four had drug specific effects altering either MA or EtOH consumption. None of the expert nominated strains tested across both drugs exhibited multi-drug effects (Fig 7).
Figure 7. Bipartite graph depicting significantly altered drug use phenotypes. Bipartite graph displays significant hits across the five measured phenotypes. Graph depicts significant effect of strain. Significant associations are represented by the thickness of the edge connecting the two nodes. Edge weights are inversely proportional to the $-\log_{10} p$-value of the association.
CHAPTER 4

DISCUSSION

4.1. Identification of Novel Addiction Risk Genes

Through this project, we were able to successfully circumvent many of the challenges that genomic assays have historically faced and identified novel addiction risk genes in a truly discovery based manner. Leveraging the high-throughput KOMP phenotyping pipeline in combination with the known relationship underlying drug-naïve behaviors and drug-related phenotypes, we successfully identified twenty-five novel genes influencing both affective behavior and initial, non-dependent drug-related phenotypes. Fifteen of our novel addiction risk genes came from our initial screening where we made uniformed predictions that specific predisposing traits would lead to specific drug-related phenotypes. Interestingly, while we did successfully identify novel addiction risk genes at a high percentage, our predictions of their specific drug-related phenotypes were not as accurate. These results informed us that there are not uniform connections between drug-related phenotypes and their predisposing traits, but a diverse multidimensional nature of the relations. Using these complex relations we were able to modify our selection process leading to the identification ten more novel addiction risk genes.

While our prediction rate was not as high as our initial screening we were able to identify genes which significantly altered drug-related phenotypes across multiple drugs. This is particularly important because the multi-drug effects mean that these genes may play roles underlying generalized drug-related phenotypes.
4.2. Selection of Phenodeviant Strains: Initial Screening

In our initial screening for addiction risk candidate genes we analyzed drug-naïve phenotypes gathered from all two hundred twenty-one mice tested in the KOMP2 behavioral phenotyping pipeline and identified nineteen strains which exhibited a significantly deviant (p < .05) novelty-response, anxiety-related, and or depression-related phenotypes. Although many strains had more extreme phenotypic values than those chosen for further testing, our validation experiments were limited to those lines that were available at the time of testing. While strains chosen for testing in initial screening were not the most phenodeviant strains, they display a representative range of phenotypes observed on all the measures used (Fig.1).

Using the predictive traits like anxiety, depression, exploration/novelty seeking we hoped to identify genes that had drug specific effects based on their baseline behavioral profile. While strains were selected for phenodeviance on a single trait strains of the greatest interest include those that were phenodeviant across multiple phenotypes i.e. Far2, which showed phenodeviance in five separate traits across three separate behavioral tests. We hypothesized that animals that showed many anxiety behaviors would manifest EtOH phenotypes, depression to nicotine, and novelty seeking to MA. Genes like Far2 were predicted to result in multiple phenotypes.

4.3. Validation of Strain Identification Using Two Bottle Choice Paradigm

The 2BC paradigm, also known as free choice drinking, was chosen because of its high face and construct validity for voluntary moderate alcohol use\(^{85,86}\), and it is a high throughput protocol which allows for the measurement multiple drug intake phenotypes. Using the 2BC protocol, we can measure drug consumption (mg of drug consumed corrected for kg of mouse), drug preference
(the percentage of milliliters of drug consumed out of total consumption), along with total-drinking (total amount of milliliters consumed).

Results from the initial 2BC screening indicated that eighteen of the nineteen tested knockouts exhibited a significant effect of either strain or strain × dose (FDR < 0.05) on at least one of the three measured phenotypes (Fig 2; Supplementary table 2). All significant effects either altered EtOH phenotypes, MA phenotypes, or both. Using the 2BC protocol we were unable to detect any significant genotype associated effects for oral nicotine related phenotypes. These findings could be due to the aversive taste of oral nicotine, a confounding of saccharine with the nicotine, or could reflect low genetic effect size for the measured phenotypes most likely due to complexity of the nicotine’s pharmacology both in terms of dose-response but also temporal patterns. Only one gene, Hdac10, exhibited significant genotype × sex effect on the ethanol total-drinking phenotype.

In our initial screening, fifteen of the nineteen tested strains had a significant effect on the MA, EtOH, or both phenotypes either through alteration of their consumption or preference. Thirteen of the strains exhibited drug specific effects, and two had multi-drug effects altering both alcohol and MA phenotypes. In the field of addiction, much of the focus has gone into studying genetic components underlying drug-specific effects through alteration of drug-specific metabolism or drug receptors in the reward pathway. These genetic components can play crucial roles in development of treatments for drug-specific SUDs. Along with these genes that play a role in drug specific responses, we are also interested in our genes which had multi-drug effects, implying they may play larger roles underlying generalized drug-related phenotypes.

As a good example of multi-drug effects, the Il12rb2 knockout strain was associated with significant genotype and genotype × dose effect across EtOH preference, ethanol consumption,
and MA consumption. This suggests that while Il12rb2 has mainly been annotated for its role in immune response and T-helper cell differentiation\textsuperscript{93}, it may also have a significant role underlying the reaction to drugs through unknown mechanisms and pathways.

Similar to what is found in the literature\textsuperscript{94,95} and results from SAMHSA’s 2018 findings that one third of all people suffering from illicit drug use disorders (IDUD) also suffer from AUDs, we found that two thirds of the knockout strains which exhibited significant effects on MA drinking also had a significant effect on alcohol related traits. This is consistent with epidemiological studies which reveal a high rate of polysubstance use such that alcohol use accompanies most other drug use.\textsuperscript{96} All eighteen of the knockout strains that had significant drug or alcohol phenotypes warrant further characterization with advanced neuroscientific and addiction assays. This is especially the case for genes such as Il12rb2 and Far2, which were found to have significant effects on both drug-seeking and addiction risk phenotypes across both ethanol and MA, suggesting that these may play an important underlying role in drug-related behaviors.

4.4. Functional Analysis of Novel Candidate Genes

To establish the validity of our significant gene deletions, we searched for genomic assays that corroborated our findings and identified plausible biological mechanisms in which the genes could have affected drug-related phenotypes. To accomplish this, we performed a functional analysis of all our significant genes using GeneWeaver.\textsuperscript{70} There, we performed a systematic search of the genes which altered 2BC phenotypes to determine whether they were represented in previous curated genomic data sets from studies of humans, mice, and rats. Among the data resources used in the analysis were the following: Medical Subject Headings (MeSH)\textsuperscript{97} related to drugs or addiction, Gene Ontoloy (GO)\textsuperscript{98,99} terms related to drugs or addiction, Quantitative Trait Loci (QTL)\textsuperscript{100} genesets related to drugs or addiction, Kyoto Encyclopedia of Genes and Genomes
pathways related to addiction and alcoholism, Neuroinformatics Framework Drug Related Genes (DRG), and Genomewide Association Studies (GWAS) of alcohol and substance use related traits (Sup Table 5).

All fifteen of the genes were supported by additional evidence from at least one of the searched databases establishing prior connections to drug related studies, either through expression data, QTL mapping, or connections to drug related biological mechanisms. For example, Hdac10 is part of the curated KEGG pathway for alcoholism in humans (hsa05034), mice (mmu05034), and rats (rno05034); it is also shown to be upregulated in the nucleus accumbens of cocaine-treated mice relative to their saline controls (GS87011). In addition to the functional analysis of drug related gene sets in GeneWeaver, our bioinformatics approach also assessed our significant genes for overlap in biological pathways. Using Gene Set Enrichment Analysis (GSEA) a systematic search of Canonical, KEGG, GO biological or cellular pathways revealed none of our genes were annotated to the same pathways. (Accession date 02/04/2020). These results suggest that these single gene knockouts may all alter drug-related phenotypes through multiple independent biological pathways and mechanisms.

4.5. Heterogeneous Behavioral Profiles Underlying Drug-Related Phenotypes

Results from our initial screening corroborate those from previous studies which found shared genetic components underlying predisposing and subsequently drug-related phenotypes. These shared genetic components led others to find predictive relationships between the drug-naïve behaviors of inbred mice and their drug-related phenotypes. To build upon these studies, we looked to further examine these relationships in a multidimensional manner using the drug-naïve and drug-related phenotypes. This analysis was done by preforming a principal component
analysis (PCA)\textsuperscript{109} of effect sizes on all drug-naïve and drug-related traits using the data from the sixteen strains tested on phenotypes and all three drugs (Fig 3). Results from the analysis indicates that principal component 1 (PC1) explains 21.1% of the variance in these traits and separates the mice that have ethanol preference and consumption phenotypes from those that do not. The second principal component (PC2) explains 17.8% of the variance and can be used to separate the different baseline behavioral profiles that underlie the ethanol-preferring strains, namely, low anxiety/high risk taking strains from the high exploration/ high activity strains (e.g., center permanence time and total number of rears in open field respectively).

Examining the results from the PCA, strains with significant drug-seeking phenotypes can be found in all four quadrants of the graph, each representing a different baseline behavioral profile predictive of different drug related phenotypes. Two ethanol preferring strains that exemplify different baseline behavioral profiles are \textit{Il12rb2} (found among strains with risk taking/low avoidance behaviors) and \textit{Hspb2} (found among strains with high exploratory behaviors). Although both of these knockout strains showed an ethanol-preferring phenotype, the different behavioral profiles segregated along PC2, which also correlated with MA consumption phenotype. The 2BC choice data reveals that \textit{Il12rb2} knockout mice have a significant MA consumption phenotype whereas \textit{Hspb2} knockouts do not. Thus, results from the principal component analysis show the diverse multidimensional nature of the relations underlying the many predisposing behaviors and their predicted drug-related phenotypes. Rather than reflecting a uniform predictive relationship between each behavioral phenotype and its predisposing effect on drug intake, \textsuperscript{50,82–84} these findings indicate a complex interaction of all the predisposing behaviors and their effects on drug-related phenotypes across different drugs, and that many biological mechanisms support the distinct relations among baseline behaviors and drug-related phenotypes.
Through deeper exploration of these relationships, one can gain a better understanding of the specific relationships among biological pathways and behavioral processes that lead to heterogeneous behavioral and genetic mechanisms of addiction and substance use. Relationships resulting in apparently similar on drug self-administration in particular individuals would be difficult to discern from GWAS where a uniform population-wide relationship between genetic variants and addiction risk traits is assumed.

4.6. Multidimensional Assessment of Predisposing Behaviors

In order to reflect the complex multidimensional relations underlying drug-related phenotypes and their predisposing drug-naïve phenotypes we used a multidimensional measure of phenodeviance to identify our next round of candidate genes for screening. To assess the 402 strains tested in the KOMP2 pipeline for overall phenodeviance we performed a Mahalanobis calculation for all strains compared to the C57BL/6NJ control using drug-naïve behavioral traits from the open field, light dark, holeboard, and acoustic startle tests.

Using traits from these four behavioral tests, we identified one hundred twenty-three strains which displayed overall significantly phenodeviant behavior from the control strain. Of the one hundred twenty-three phenodeviant strains, we prioritized the rederivation of the most extreme strains. Using findings from the initial screen we hypothesized that the mice that showed the greatest overall phenodeviance would manifest multi-drug phenotypes at a higher rate since where were combining all predisposing behaviors into a single predictive value.

Due to availability of sperm, success of in vitro fertilization, and viability of the lines for producing viable cohorts, we were able to test thirteen of the most extreme phenodeviant strains. These strains were rederived and tested solely due to the Mahalanobis score ranging from 517-1983(Fig. 4) and had no prior evidence suggesting they may alter drug-related phenotypes. In
addition to the thirteen strains chosen solely for their Mahalanobis score, two other expert nominated KOMP lines were chosen for testing. These two strains, Rap2b and Tmod2, were among the one hundred twenty-three strains found to be significantly deviant, but were found to be on the moderate to lower end of phenodeviance spectrum (288 and 115 respectively). Finally, along with the KOMP single gene knockout lines, six other expert nominated strains were chosen for testing.

While in the initial screening we were limited to strains which were alive and available at the time of testing, our second screening allowed for the rederivation of the most overall phenodeviant strains based on their Mahalanobis score. We hypothesized that rederivation of the strains which were the most overall deviant across all predisposing traits would enrich our ability to detect more genes that underlie drug-related phenotypes and identify more candidate genes which have multi-drug effects.

4.7. Refining Assessment of Drug-Related Phenotypes

To assess the chosen strains for significantly altered patterns of drug-related phenotypes, we used the same MA 2BC protocol and updated our EtOH protocol to a more translationally relevant binge drinking DID paradigm. This protocol is a limited access protocol which has been refined to induce mice to drink to levels of intoxication (~100 mg/dL), where they would not normally in a 2BC protocol since EtOH is only moderately rewarding. This protocol allows for similar overall consumption measures (mg/kg) of EtOH across the four days of limited exposure, measurement of binge drinking phenotypes, and potential difference in EtOH sensitivity which could affect pleasurable and rewarding responses to drug use. As previously stated, there is a significant comorbidity of alcohol and MA use but newer evidence suggests that binge drinking has a significantly higher comorbidity rate and is a better predictor for MA use than moderate drinking. Using these two protocols were able to screen our extreme single gene knockout strains
for drug-specific effects and also strong multi-drug effects that alter drug-related phenotypes across drugs, testing procedures, and have strong translational relevance to human patterns of MA and alcohol use.

4.8. Improved Identification of Genes Altering Methamphetamine Use

Using a multidimensional analysis to identify extreme overall phenodeviant strains greatly increased our ability to identify strains which had significantly altered patterns of MA use. Of the fifteen strains chosen for their Mahalanobis score, 12 showed significant differences on one of our three measured drug-related phenotypes, and eight directly altered MA use. The use of Mahalanobis distance to identify strains with predisposing behaviors increased our ability to identify genes which affect methamphetamine use by ~ 36% from a 17.6% (3/17) success rate to 53.3% (8/15). A success rate of 53% for identifying genes that altered MA-related phenotypes from only predisposing drug-naïve behaviors, while not as high a rate as expert nominated strain 80% (4/5), is a notable accomplishment.

In our MA 2BC we observed no main effect of sex across any measure of MA-related phenotypes which corroborates previous MA studies that also detected no significant interactions of sex\textsuperscript{91,92}. Six of our fifteen phenodeviant knockouts resulted in a main effect of strain which increased general preference for MA compared to controls. (Fig 5A) Three of these six strains which showed increased preference for MA also had a significant main effect of strain, which significantly increased their consumption of MA (mg/kg). (Fig 5b)

Five of the screened knockout strains resulted in significant strain × dose interactions. For the control strains and the knockouts, preference for MA trended down as doses increased, with the initial dose of 10 mg/L being their most preferred dose. Control strain had an average initial preference for MA 31.5±3.3, and most significant knockouts had initial preferences ranging from
41.5%-57.8%. Myh10, Dnaja4, Tmod2, and Rap2b knockouts resulted in a shift in their dose response curves and had the highest average preference for the 20 mg/L dose 48.0% and 47.4% respectively. (Sup Panel 6) Other significant strain × dose interactions revealed increased sensitivity to increasing dosages on MA. While Dnmt3a, Cp, and Lrrc15 displayed initially increased preference for MA at the initial dose compared to control, they all also showed decreased preference to MA at higher doses, specifically at the 40 mg/L dose. These results show that the multidimensional approach to identifying phenodeviant strains greatly enriched our ability to detect gene which had an effect on MA compared to the initial screen.

All of our expert nominated strains were previously identified based on human or model organism stimulant studies. Of the two expert nominated C57BL/6NJ strains, only Trib3 resulted in any significant alterations of MA use. C57BL/6J derived strain, Bmerb1 displayed significantly increased MA consumption at the highest dose compared to controls. This suggests that deletion of this gene may potentially lead to reduced efficacy of MA resulting in increased consumption at higher doses that are aversive to their control strain. In comparison to these results, deletion of the Bmerb1 gene in the C57BL/6J strain resulted in increased preference and consumption and the 40 mg/L dose and decreased both at the 80/ mg/L dose compared to controls. Deletion of Bmerb1 resulted in the only significant effect of sex for all MA phenotypes. Differences in preference and consumption at 40 mg/L were driven by the males who displayed increased preference and consumption compared to females and controls. Finally, while Bmerb1 preference and consumption of MA equal to or greater than controls at the first three doses, the KOs displayed increased aversion to the final 80 mg/L dose. This finding suggests that Bmerb1 may a role regulating the rewarding and aversive effects of MA-related phenotypes in a dose dependent manner.
The expert nominated strains of particular interest in our MA 2BC paradigm are the C57BL/6J-Taar1\textsuperscript{(D2Ki)} and DBA/2J-Taar1\textsuperscript{(B6Ki)} strains with the allele swap. The trace amine-associated receptor 1 (Taar1) gene was identified as a potentially important gene involved in MA intake in a quantitative trait loci (QTL) mapping study using MA high drinking (MAHDR) and low drinking (MALDR) strains from the BXD RI panel.\textsuperscript{112} In this study Phillips and colleagues identified a QTL on chromosome 10 that accounted for 50-60% of genetic variance of MA intake.\textsuperscript{91,112} In the initial study using MAHDR and MALDR, it was found that the MAHDR phenotype segregated with the non-functional version of the Taar1 gene, making it a risk factor for MA use.\textsuperscript{112} Further examining the BXD RI lines and their history, it was found that BXD lines which were older and derived earlier in BXD projects only had B6 functional versions of the Taar1 gene, but some of the newer derived strains made using DBA/2J as their progenitors had the non-functional of the gene.\textsuperscript{92,113} This finding led to the sequencing of the non-functional Taar1 gene to find a single nucleotide polymorphism (SNP) at position two hundred twenty-nine. At position two hundred twenty-nine a cytosine is altered to an adenosine, which results in a missense mutation changing a non-polar Proline to a polar Threonine.\textsuperscript{113,114} This missense mutation occurs in Taar1’s only exon and inhibits proper folding in Taar1’s transmembrane domain, leading to a lack of function.\textsuperscript{113,114}

To validate that this exact SNP is responsible for the increased MA intake observed in the DBA/2J and the BXD lines which inherited its Taar1 allele, we performed a gene swap with C57BL/6J. In our gene-swap KI strains, we recreated the SNP at position two hundred twenty-nine in codon seventy-seven. This changes a Proline to a Threonine in our C57BL/6J line (C57BL/6J-Taar1\textsuperscript{(D2Ki)}), and restored function to Taar1 gene in our DBA/2J lines DBA/2J-Taar1\textsuperscript{(B6Ki)}.
Results from our screening using a 24-hour access 2BC protocol (Fig 5C-F) validate that this SNP is responsible for the observed increased MA intake, and restoration of its function can alleviate risk factors for MA use. This mutation was originally discovered using a limited access 2BC paradigm which was necessary for high levels of MA intake. While our DBA/2J and C57BL/6J-Taar1(D2KI) strains did not consume to the excess that the MAHDR strains did in the limited access paradigms, it speaks to the strength of the phenotype observed with the Taar1 mutation. Our results show that the mutation still has significant effects and increases consumption when MA is available ad libitum, and it can be parsed out from the limited access paradigm which drives increased MA consumption. Validation of this SNP as the causal variant leading to increased MA consumption provides a strong potential target out of the two hundred variants observed in TAAR1. Further research of this variant could lead to potential treatments for MA use in humans through TAAR1 agonists in addition to serving as potential screening marker for assessing vulnerability to MA use in humans.

4.9. Strains Assessed for Ethanol Related Phenotypes Using DID Protocol

In our second screening we augmented our protocol for EtOH drinking from an ad libitum 2BC protocol to a limited access DID protocol, with the hopes of identifying genes which significantly affected the more translationally relevant binge drinking phenotype. It has been shown that EtOH consumption is only moderately rewarding from a physiological standpoint in mice. Throughout the course of the ad libitum 2BC protocol BEC changes rapidly, and mice do not sustain pharmacologically relevant levels of BEC. Although preference drinking is a widely used and valid partial model for alcohol use, we felt using the DID protocol to measure consumption and binge drinking phenotypes would improve the translational relevance of our project.
Findings from our project corroborate those of previous studies on EtOH binge-like experiments, which found that female mice consume higher levels of EtOH than male mice.\textsuperscript{119-121} Another closely related study hypothesized that this may be due to an inherently heightened risk for excessive EtOH consumption in females related to the drugs ability to reduce a naturally more active stress response\textsuperscript{122}. In our study, we found that for every strain tested there was a significant main effect of sex, and females drank more than males when corrected for body weight (mg/kg) in every case (Sup Panel 7).

Using the DID protocol, we found that four of fifteen phenodeviant strains had significant EtOH phenotypes. All four gene deletions altered EtOH consumption over the four days of limited access. \textit{Irf8} and \textit{Elof1} strains both displayed reduced EtOH consumption for both male and females compared to control strain. \textit{Tmod2} and \textit{Zbtb4} both had significant strain $\times$ sex interactions as indicated by decreased consumption for the female lines but no difference in male consumption compared to controls. This suggests that unlike deletion of the \textit{Irf8} gene, which has the same effect on both sexes across both drugs, deletion of these three genes results in sex differences in EtOH consumption.

We found that our measure of EtOH intake is a moderately strong predictor of BEC with an $r^2 = 0.49$ across all twenty strains and both sexes. This correlation for consumption on day four and the resulting BEC is comparable to one found in the creation of the DID protocol which was $r^2 = 0.53$ using only female C57BL/6J from different sites or $r^2 = 0.50$ when using both males and females.\textsuperscript{77} This indicates our measures of BEC are reliable even with our altered process of blood collection based on refinement of the protocol in Thiele \textit{et al.} 2016.

The findings of sex differences in responses of \textit{Tmod2} and \textit{Zbtb4} to EtOH are particularly interesting because they both also significantly altered MA intake but no effect of sex or strain $\times$
sex was detected. The sex differences in Tmod2 phenotypes corroborate findings from previous studies done using BXD RI lines which found sex differences in gene expression in various locations throughout the reward pathway following drug exposure. Through a functional analysis of all our significant genes, it was found that the gene sets from this study included Tmod2 as a gene which was differentially expressed following drug exposure in a sex dependent manner. These data indicate deletion of Zbtb4 and Tmod2 may alter the response of the reward pathway in a drug by sex specific manner.

Of the fifteen phenodeviant strains chosen for screening in the DID paradigm, none of the strains resulted in a significant BEC phenotype following multiple testing corrections. This may in part be due their C57BL/6NJ background which was found to have low BEC levels in our experiment. An important finding in this study was the differences observed between the control C57BL/6J and C57BL/6NJ control strains. As previously stated, this protocol was developed to induce pharmacologically relevant BEC levels in C57BL/6J strains. Our results indicate that this was true for all our C57BL/6J derived strains but was unsuccessful for all of our C57BL/6NJ strains. In our experiment, C57BL/6J mice consumed on average 13% more ethanol than C57BL/6NJ mice, 6% and 30% for females and male respectively, on the final day of the DID protocol. While C57BL/6J mice only drank on average 13% more across both sexes than the C57BL/6NJ mice, this resulted in a 113% increase in BEC levels. This is suggestive of potential differences in EtOH metabolism between the two control strains. While only drinking slightly less than their C57BL/6J counterparts, C57BL/6NJ mice did not reach pharmacologically relevant levels of BEC, possibly due to increased rate of EtOH metabolism. These decreased consumption and BEC levels were consistent across all C57BL6/NJ-derived lines. This increased metabolism could have contributed to the decreased number of significant strain effects we were able to detect due to a floor effect.
Since our control strain started with a lower average BEC and most deletions either did not alter BEC or tended to decrease BEC levels further, this protocol was less suited to detecting strain effects in the C57BL/6NJ lines.

Nr2f1 was the only expert nominated strain to result in a significant effect of strain for both EtOH consumption and Average BEC measurements. Deletion of the Nr2f1 gene significantly lowered EtOH consumption across both sexes and consequently was the only gene which significantly altered average BEC for both sexes. Deletion of Nr2f1 resulted in a decreased consumption on day four from control levels of 9.03 ± 0.44 mg/kg and 3.73 ± .45 mg/kg for females and males respectively to 6.21 ± 0.42 mg/kg and 1.05 ± .38 mg/kg respectively. This resulted in a decrease of average BEC from control levels of 76.8 ± 18.1 mg/dL and 25.2 ± 6.0 mg/dL for females and males respectively to 42.45 ± 15.0 mg/dL and 1.69 ± .64 mg/dL respectively. Nr2f1 has previously been connected to optic atrophy and intellectual disability\textsuperscript{123}, it has also been found to be one of 475 genes which were differentially expressed in the striatum of C57BL/6J mice following drinking to intoxication in a DID paradigm.\textsuperscript{124} This suggests that the decreased consumption and corresponding BEC may not be due to vision or intellectual problems, but deletion of the Nr2f1 gene may alter response to EtOH in the reward pathway.

While females from only two C57BL/6NJ strains (Lrrc15 and Trib3) reached binge drinking levels with average BECs above 100 mg/dL, the results from our DID screening are still important for examining the role our candidate genes play in EtOH drinking and susceptibility.

4.10. Identification of Phenodeviant Strains Using Multidimensional Analysis Improved our Ability to Identify Genes with Multi-Drug Effects

In our second screening, of the fifteen strains which were selected for being extremely phenodeviant, ten strains resulted in significantly altered patterns of drug-related phenotypes.
Three of our KO strains resulted in significant effects on both MA and EtOH phenotypes. While the total percentage of genes we tested that resulted in at least one significant effect was not as high as our original screening (79%, 15/19), we greatly increased our ability to identify genes which effected MA to 53.3% (8/15) from ~18% (3/17). Importantly, we also increased our ability to identify genes which had multi-drug effects on both MA and EtOH from 12.5% (2/16) to 20% (3/15). Reduced number of strains which had significant EtOH phenotypes is presumably due to the refined EtOH protocol used in the second screening with a more translationally relevant binge-like phenotype. Overall, the use of a multidimensional assessment of predisposing behaviors improved our identification of gene KOs which altered MA-related phenotypes in addition to improving our ability to identify genes such as *Irf8*, *Tmod2*, and *Zbtb4* which significantly altered phenotypes on multiple drugs.

### 4.11. Conclusions and Future Directions

Overall, our data indicates the utility of leveraging the known complex relationships among predisposing drug-naïve behaviors and addiction risk phenotypes. In this project, we used and refined our understanding of these relationships in combination with the high-throughput JAX-KOMP2 program to identify thirty-three plausible single gene KO strains predictive of addiction risk phenotypes. Of those thirty-three plausible candidates, twenty-five gene KOs significantly altered drug-related phenotypes. Following screening for drug-related phenotypes, all significant genes were validated through functional analysis for having plausible connections and mechanisms to affect drug-related phenotypes. Further analysis through GSEA indicated no overlapping pathways among our candidate genes that could have possibly affected drug-related phenotypes. This indicates that these genes could represent multiple diverse pathways for roles in addiction risk. This strategy was therefore successful and circumvented the effects of drug exposure on
subsequent physiological testing in the screening program; that is, it allowed us to discriminate risk from consequences of drug exposure. An approach that uses a drug-naïve screen is efficient, but it will necessarily miss those genes with addiction risk effects that are not manifested in predisposing behaviors. Nevertheless, through this study combining publicly available data, multiple novel candidate genes, high through-put testing using multiple drugs, and functional analysis of multiple genomic databases, we have now identified twenty-five new addiction risk genes amenable for detailed characterization in viable mutant mice.

While addiction risk genes are often independent from the genes predisposing for transition from initial use to addiction, further characterization of these genes and associated pathways could elucidate their or related pathway gene’s distinct roles in the process of transition addiction. Our findings suggest that the continued phenotyping of inbred single gene knockout mice using the broad neurobehavioral screen from KOMP will allow for the continued identification on novel addiction risk genes. Through our particular project we were able to detect multiple genes affecting drug using through a diverse set of biological pathways. Each of these genes would only accounts for small proportions of the genetic variation and would often be missed using GWA studies. Hopefully moving forward, building off the findings from this study, the continual screening of KO mice for predisposing drug-naïve behaviors can lead to further discovery of more novel addiction risk genes that would not be discovered easily using hypothesis-driven or other discovery based techniques. This continual identification and characterization of novel addiction risk genes will hopefully lead to better screening for vulnerability to SUDs and addictions as well as potential therapeutic targets for improved treatments.
REFERENCES


65


71

APPENDIX

All supplemental materials can be found online in the Digital Commons at the University of Maine along with this manuscript.
BIOGRAPHY OF THE AUTHOR

Tyler Roy was born in Lewiston, Maine on January 30, 1992 to Denise M. and Robert R. Roy. Tyler was raised in Lewiston and graduated from Lewiston High School in 2010. Following high school Tyler attended the University of Maine for his undergraduate career where he earned a B.A. in Psychology and B.S. in Biochemistry with a minor in Chemistry and graduated with High Honors from The University of Maine Honors College in 2015. Following graduate Tyler joined the lab of Elissa Chesler at The Jackson Laboratory in Bar Harbor, Maine studying the genetic of addiction. Tyler is a candidate for the Masters of Science degree in Biochemistry from the University of Maine in May 2020.