Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Sonam Nitin Patel Date

Interaction of Cannabis Use Genes and Implicated Genes in Schizophrenia and Bipolar Disorder in a Molecular Pathway Analysis

By

Sonam Nitin Patel

Master of Public Health

Epidemiology

Brad Pearce

Faculty Thesis Advisor

Interaction of Cannabis Use Genes and Implicated Genes in Schizophrenia and Bipolar Disorder in a Molecular Pathway Analysis

By

Sonam Nitin Patel

B.S., B.A., University of California, Irvine, 2013

Thesis Committee Chair: Brad Pearce, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2015

Abstract

Interaction of Cannabis Use Genes and Implicated Genes in Schizophrenia and Bipolar Disorder in a Molecular Pathway Analysis

By Sonam Patel

Cannabis is the most commonly used illicit drug worldwide. With debate surrounding cannabis legalization, it has become an integral area of research in mental health. Recreational use is commonly implicated as a risk factor for schizophrenia and bipolar disorder, yet the vast majority of users will not develop either disorder in their lifetime. The potential causal role between cannabis and these disorders remains uncertain. Broadly, the purpose of this thesis is to examine the relationship between genetic sequence variation in these mental disorders and in cannabis abuse, and more specifically, to discern molecular pathways involving relevant susceptibility genes and potential causal pathways. Genetic variants implicated in schizophrenia, bipolar disorder and cannabis use were compiled from the Psychiatric Genetics Consortium and the National Institutes of Health Database of Genetic Phenotypes. Single nucleotide polymorphisms (SNPs) were mapped to specific genes for each dataset. Several shared genes were discovered in comparisons between cannabis use and schizophrenia susceptibility variants and between cannabis use and bipolar disorder variants. To better understand the physiological relevance of these genes, I conducted a gene set enrichment and molecular pathway analysis for genes that showed overlap between cannabis and schizophrenia (44 genes), and between cannabis and bipolar disorder (42 genes). Predictive molecular network analyses attested to the biological role of susceptibility genes within networks underlying cannabis use, schizophrenia and bipolar disorder. In particular, the cannabis schizophrenia connection involved networks that included *DLG2* and ubiquitin. Acknowledgement of this gene-environment interaction is a significant factor when determining impact of environmental risk factors on the disorders, including cannabis use.

Interaction of Cannabis Use Genes and Implicated Genes in Schizophrenia and Bipolar Disorder in a Molecular Pathway Analysis

By

Sonam Nitin Patel

B.S., B.A.
University of California, Irvine
2013

Faculty Thesis Advisor: Brad Pearce, Ph.D.

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2015

Table of Contents

Acknowledgements

I would like to acknowledge my advisor, Brad Pearce, for his contribution, advice and assistance. I would like to thank my friends and family, who provided encouragement throughout my time at the Rollins School of Public Health.

To my parents, Bhavini and Nitin Patel, everything is because of you. Thank you for always believing in my passion and dedication to a career in public health and for providing me the opportunity to pursue a graduate degree and my dreams. To my brother, Sanum, thank you for all your support these last two years. I owe so much to my family for their ongoing encouragement.

To my dear friends back home in California, thank you for being my cheerleaders.

Thank you to my Atlanta family, who made my two years at RSPH an incredible, fulfilling experience that I will always cherish. Lastly, thank you to Dr. David Timberlake, who taught me that the most rewarding thing about public health is the service itself.

Introduction:

Bipolar disorders (BD) are characterized by cycles of mania (grandiosity, decreased need for sleep, pressure to keep talking, flight of ideas, distractibility) and depression. Schizoaffective disorder is a condition where a person experiences a combination of mood symptoms (depression or mania) and psychosis (hallucinations, delusions). Those with a bipolar disorder have a high degree of psychiatric comorbidity including substance use disorders and all-cause mortality (Lewinsohn, Klein, & Seeley, 1995). The demarcation between bipolar disorder and schizoaffective disorder is that psychotic symptoms in schizoaffective disorder occur in the absence of significant mood symptoms. Several studies have shown that schizoaffective disorder bipolar type can be classified as a phenotypic intermediate point between bipolar I disorder and schizophrenia (Craddock, O'Donovan, & Owen, 2009).

Cannabis is the most widely used illicit drug in the world. There is increasing evidence to suggest cannabis is an important risk factor for psychotic disorders (Caspi et al., 2005). Evidence suggests cannabis use is a modest statistical risk factor for emergence of psychosis, including emergence of hallucinations and delusions, bipolar disorders, and schizophrenia-associated disorders (Wittchen et al., 2007). Studies also suggest cannabis use during adolescence and young adulthood further compounds risk associated with psychotic disorders and mental illnesses. Literature on substance abuse epidemiology proposes that adolescents are most vulnerable to deleterious effects of cannabis due to timing of use during a critical brain development period (Schneider & Koch, 2003). Cannabis use significantly declines after the age of 21 in the U.S. (Kandel, Chen, Warner, Kessler, & Grant, 1997).

Tone theory suggests that cannabis use during the critical period of adolescence results in long-term epigenetic alterations that influence neurobiological process involved in bipolar disorder and schizoaffective disorder pathologies (Rutten & Mill, 2009). Patton et al. (2002) found that cannabis use in adolescence predisposed individuals to higher rates of depression and anxiety, symptoms commonly present in bipolar disorder, in adulthood.

Cannabis dependence, as defined by criteria of the fourth edition of The Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), may also manifest in a dose-response relationships in symptom progression for those with schizoaffective disorder. A dose-response relationship with cannabis and risk for schizophrenia has been substantiated in a few independent studies. For example, a 2002 study found that heavy cannabis users (defined as smoking at least once a day for at least a month) were six times more likely than non-users to receive a diagnosis of schizophrenia (Zammit et al., 2007). Although the association between cannabis use and increased risk of bipolar disorder and schizophrenia has been observed, the majority of the majority of adolescents who use cannabis do not develop psychosis. This suggests the relationship is modified by underlying risk factors, genetic and social, that increase vulnerability to psychosis and mental illness (Casadio, Fernandes, Murray, & Di Forti, 2011).

The presence of such a gene by environment interaction is indicated by evidence that the association between cannabis and psychosis outcomes is most significant in subjects with an established vulnerability to psychosis (Henquet et al., 2004). As populations with psychotic disorders and populations with bipolar share genetic characteristics, cannabis may increase risk for bipolar disorders through the same

pathways as it does with schizophrenia. Accordingly, it is commonly theorized that bipolar disorder and schizophrenia patients use cannabis to self-medicate their symptoms, but there are conflicting findings to support this theory.

Multiple environmental factors affect risk for psychotic disorders, many of which are common among schizophrenia, schizoaffective and bipolar I disorders. Socioeconomic status, experience of childhood abuse, past experience of sexual abuse, urbanicity, gender, race, education, income, alcohol use, tobacco use and other illicit drug use have all been proposed to affect individual vulnerability to psychotic disorders (Aleman, Kahn, & Selten, 2003; Alvarado-Esquivel et al., 2011; March et al., 2008); Creemers et al. (2011); Ellett, Freeman, and Garety (2008); Peters et al. (2008); Hall and Dengenhardt (2008); Ringen et al. (2008); Tedla et al. (2011)). Traumatic brain injury during a critical development period of adolescence has also been speculated as a risk factor for psychosis (Molloy, Conroy, Cotter, & Cannon, 2011). Evidence suggests that identification as a first-generation immigrant increases risk of schizophrenia, and this risk persists into the second generation (Bourque, Van der Ven, & Malla, 2011). These factors may also contribute to or result in earlier onset of cannabis use or dependence and may also directly increase risk for bipolar I, schizophrenia, and schizoaffective disorders, regardless of cannabis use.

Schizophrenia, schizoaffective, and bipolar 1 disorders are multifactorial and include a heritable component (Thaker, 2008). Causal chains for these mental illnesses include environmental factors, which could act during gestation, at the time of birth, during childhood, or over the course of adolescence. Cannabis abuse or dependence has

been explored as a significant environmental factor that contributes to schizophrenia, but also has a heritable component (Jutras-Aswad et al., 2012).

Genome-wide association studies (GWAs) are conducted to examine many common genetic variants in different individuals derive conclusions of associations between variants and traits. Findings of some GWAS have documented associations of single nucleotide polymorphisms (SNPs) with psychosis and other mental illnesses. The differential alleles in the Val66Met polymorphism in the brain-derived neurotropic gene (BDNF) may adversely affect synaptic plasticity and neuron survival, dysregulation of which is linked to bipolar disorder (Nakata et al., 2003). The short allele of the 5-HTTL serotonin transporter gene polymorphism, which codes for a functional 44 base pair insertion or deletion in the gene's promoter region, is also linked to bipolar disorder (Chen et al., 2013). Furthermore, Patients presenting with the single copy of the MTHFR 677TT allele presented with worse cognitive function than patients homozygous for the allele in a 2012 study(Peerbooms et al., 2012). Methylentetrahydrofolate reductase (MTHFR) catalyzes the conversion of 5,10-methlenetetrahydrofolate (5,10-MTHF) to 5-methyltetrahydrofolate (5-MTFHF). MTHFR is a crucial enzyme in methylation of DNA (Peerbooms et al., 2012). Single copy of the MTHFR 677T allele causes a 35% reduction of enzyme activity. Reduction of MTHFR enzyme activity has been shown to be associated with development of psychotic symptoms and bipolar disorder I (Kempisty et al., 2006). Combined effects of individual SNP variants which confer in biological conditions or mechanisms, such as decreased synaptic plasticity, may interact together to synergistically increase risk for schizophrenia and associated disorders (Allen et al., 2012; Epstein & Kumra, 2014).

A well-known source of bias in genomic data association analyses is confounding from population stratification. If not properly accounted for, spurious associations may occur in GWAS due to confounding factors rather than the true association between tested genomic factors and the trait of interest. Various methods can be used to account for population stratification including using genomic controls, implementing a principal components approach (Alkes L Price et al., 2006), matching, and Bayesian smoothing.

Genome-wide analysis metadata from the Psychiatric Genetics Consortium and the National Institutes of Health Database of Genes and Phenotypes Study of Addiction and Genetics (SAGE) was analyzed to discern these interactions. Population stratification can present a problem with combination of these datasets, as data sources included several combined samples of subjects of various ethnic ancestries. Data cleaning processes prior to manipulation of datasets for this thesis primarily included a principal components approach and implementing genomic controls to correct for population stratification.

In consideration of such factors, the purpose of this thesis is to examine the relationship between genetic sequence variation in these mental disorders and in cannabis abuse. More specifically, this thesis will aim to discern common liability molecular pathways involving relevant susceptibility genes and potential causal pathways in schizophrenia, schizoaffective, and bipolar I disorders. By examining the differential causal pathways, there may be evidence to suggest effect measure modification by cannabis use occurring among those genetically susceptible to the aforementioned mental illnesses. Other pathways may contain cannabis use as an intermediate, or as not part of any direct significant causal pathways. The hypothesis guiding this proposal is that there

is significant effect modification by cannabis occurring among those genetically susceptible to schizoaffective disorder and bipolar disorder, which will be reflected in various potential molecular causal pathways. Results from this thesis may support suggested theories that such effect modification by cannabis use is more extensive among individuals with greater genetic susceptibility to schizoaffective, bipolar I, and schizophrenia disorders.

A Review of the Literature

Several mechanisms should be explored further when considering potential causal pathways. Pharmacological properties of cannabis could result in altered brain circuitry and lead to psychotic illness. A 2014 case-control study examining effects of cannabis use on brain morphology revealed that regular cannabis use is associated with alterations in medial temporal, frontal and cerebellar brain regions (Løberg et al., 2014). A systematic review of functional neuroimaging studies suggest that prefrontal blood flow is lower in chronic cannabis users than in controls (Wrege et al., 2014). Reduction of white matter volume is also associated with onset of depressive symptoms. Medina, Nagel, Park, McQueeny, and Tapert (2007) found that subjects who were cannabis dependent showed more depressed symptoms than controls, and had smaller white matter volume in their cohort of adolescent cannabis users. Results of this study acknowledged that marijuana use and resulting reduced white matter volume were additive and interactive in predicting depressive symptoms among adolescents(Medina et al., 2007).

Genetic variation does not completely predict onset of schizophrenia, suggesting there may be other risk factors that affect the relationship between cannabis and psychotic illness. Some aforementioned risk factors may significantly increase risk for

schizophrenia, which could interact with or negate any genetic variations. For example, experience of all forms of childhood abuse may significantly increase risk for psychosocial problems. It may lead to cannabis use during adolescence or adulthood as a means of self-medicating and relieving effects of sustained psychological trauma. Studies have shown that there are significant differences in comorbid schizophrenia, cannabis dependent patients reporting experience of childhood abuse compared to patients without a diagnosis of cannabis dependence (Compton, Furman, & Kaslow, 2004) (Houston, Murphy, Shevlin, & Adamson, 2011) (Holowka, King, Saheb, Pukall, & Brunet, 2003). With this potential causal pathway, cannabis use represents an effect modifier between the relationship of childhood abuse and schizophrenia, schizoaffective and bipolar I disorders. Some causal pathways may not even significantly involve key genetic variants tested in this thesis.

Gene mapping showing SNPs and associated coded proteins will reveal the presence or absence of association of candidate genes with the mental illnesses in question. If genetic variation is not relevant to the hypothesis in question, there will be no genetic variants connected to susceptibility proteins that alter risk for the mental illnesses in question.

Since the vast majority of cannabis users do not develop psychosis, certain individuals may be genetically vulnerable to the deleterious effects of cannabis as well as schizophrenia, bipolar I, and schizoaffective disorders (Shaun M Purcell et al., 2009). Despite the strong association between cannabis use and psychotic disorders acknowledged in the literature, it is conceivable that cannabis has no connection to psychosis. Thus, there may be external underlying factors that predisopose individuals to

both cannabis use and either schizophrenia or bipolar disorder. The set of genes and associated proteins involved in cannabis dependence and those involved in the aforementioned mental illnesses could meaningfully overlap. For instance, a polymorphism in the catechol-O-methyltransferase (COMT) gene has been shown to moderate the influence of adolescent cannabis use on developing adult psychosis (Caspi et al. 2005) (Costas et al., 2011). Carriers of the COMT valine158 allele were more likely to exhibit psychotic symptoms and to develop schizophreniform disorder if they used cannabis than non-carriers. Cannabis use had no such adverse influence on individuals with two copies of the methionine allele for the COMT gene. Verdejo-Garcia et al.'s study further explored the impact of genetic variants in COMT on cognition and memory (Verdejo-Garcia et al., 2013). Results acknowledged that carriers of the Val allele were most sensitive to the Δ-9-tetrahydrocannabinol-induced psychotic experiences, which was conditional on prior evidence of risk for psychosis (Bilder et al., 2002). Carriers of the Val allele were also more sensitive to Delta-9-THC-induced memory and attention impairments compared to carriers of the Met allele (Bilder et al., 2002). Additionally, there is evidence to suggest that age of illness onset for individuals being homozygous for the methionine allele of the COMT gene is greater than homozygous Val carriers (Mannisto and Kaakkola (1999); Pelayo-Teran et al. (2010); Estrada et al. (2011)).

Genetic variants that increase risk for schizophrenia, bipolar I, and schizoaffective disorders can interact to additively increase risk for such disorders. To assess this cumulative effect of numerous polymorphisms, a polygenic risk score, a quantitative measure that sums the weights of effect of genetic variants, can be used (Purcell, Cherny, & Sham, 2003). The risk score relative to a certain threshold of significance, as

determined by the researcher, will determine an individual's risk status for schizophrenia, schizoaffective and bipolar I disorders. The effect of cannabis may alter the risk score to transcend beyond the threshold. This effect could potentially be tested by comparing the statistical significance in the difference of the polygenic risk scores before and after consideration of cannabis use genetic variants.

Although evidence suggests a strong association between cannabis use and schizophrenia and bipolar disorders, it has been widely theorized that affected patients may be using cannabis as a form of self-medication, suggesting reverse causation (Griffith-Lendering et al., 2013). There are potentially two causal pathways that may link cannabis use and psychosis. First, cannabis use may lead to an increased susceptibility to psychotic symptoms via changes in brain chemistry. Alternatively, those developing psychosis may have an increased susceptibility to using cannabis as a consequence of their psychological state. Furthermore, cannabis use and psychosis may be associated with one other reciprocally by a feedback loop in which the use of cannabis increases risks of psychosis and the onset of psychosis simultaneously leads to an increased use of cannabis. Structural equation models can be used to address possibility of reverse causality by devising statistical models that permit reciprocal relationships between cannabis use and psychosis and using these models to provide a guide to probable patterns of causation (Fergusson, Horwood, & Ridder, 2005).

Predisposition to cannabis use and/or dependence may be due both to genetic variants that are common to all addictions and to those specific to a particular addiction. For example, there are different influences of environment versus genetic factors on the transitions from initiation of drug use, to consistent drug use, to drug

addiction/dependence and then potentially to relapse (Palmer et al., 2014). The genetics of addiction encompasses heritable factors that influence the different stages in the trajectory of initiation and progression to drug addiction, including severity of dependence or withdrawal and risk of relapse (Iwasaki, Ishiguro, Higuchi, Onaivi, & Arinami, 2007). Variation in personality dimensions, such as impulsivity, risk taking and novelty seeking, may affect the initiation of drug use as well as the transitions from initial use to regular use to addiction (DeRosse, Kaplan, Burdick, Lencz, & Malhotra, 2010).

Longitudinal studies have reported an increased likelihood for developing schizophrenia and other psychotic illnesses after cannabis use, especially when cannabis use was initiated in early adolescence (Løberg et al., 2014). Several studies confirm that cannabis use is approximately two times more frequent among schizophrenia patients than in the general population (J. van Os et al., 2002). Varying pharmacological properties of cannabis, such as its effects on dopamine release, can result in manifestation of greater biological vulnerability to schizophrenia. For example, there is evidence to suggest that the primary psychoactive ingredient in cannabis, delta-(9)-tetrahydrocannabinol (THC), influences the endogenous cannabinoid and dopamine systems, via cannabinoid receptors (D'Souza et al., 2005). Such receptors which are highly distributed in cerebral cortex, including brain regions implicated in schizophrenia, and influence dopamine systems and uptake (D'Souza et al., 2005). THC is a cannabinoid receptor 1 (CB1) agonist, and Casadio et al. (2011) suggests that cannabis produces its effects via influence on CB1 receptors on GABA and glutamate, which regulate excitability of midbrain dopamine neurons and prefrontal cortical pyramidal cells. THC may worsen dopaminergic imbalances by increasing dopaminergic tone in striatal regions

of the brain (Kuepper et al., 2010). Repeated increases in dopaminergic tone in such regions can decrease dopaminergic levels in prefrontal regions of the brain via sensitization processes, resulting in expressions of a psychotic disorder (van Winkel, 2011); (Miyamoto, Miyake, Jarskog, Fleischhacker, & Lieberman, 2012). Moreover, studies have implicated cannabis in resulting in poorer white matter fiber, which is associated with greater risk for development of schizophrenia (Jacobus et al., 2009). Deficits in executive and motor functioning in patients with first-episode psychosis are associated with reductions in white matter integrity in major fasciculi connecting frontal and temporal cortices, and between cortical and subcortical regions (Rocío Pérez-Iglesias, 2010). Localized reduction of fractional anisotropy, a measure of white matter changes, in white matter underlying left parietal lobe was found among schizophrenia patients with 22q11.2 deletion syndrome in a 2012 study (Kikinis et al., 2012).

Impaired cognitive function has been recognized as a core feature of schizophrenia. Changes in neurocognitive functioning due to cannabis use differ significantly across the schizophrenia spectrum disorders (Makkos et al., 2011). Ringen et al. (2010) found that cannabis use in bipolar disorder subjects was associated with better neurocognitive function but the opposite was the case for schizophrenia subjects in the sample (Ringen et al., 2010). A 2009 case-control study found no significant differences in cognitive performance between healthy non-users of cannabis and healthy users. Researchers found a schizophrenia-like increase in preservation on the Wisconsin Card Sort Test (WCST), however, which tests for cognitive flexibility (Scholes & Martin-Iverson, 2010). Results also indicated no significant differences between schizophrenic cannabis users and schizophrenic non-users (Scholes & Martin-Iverson,

2010). Findings indicate that there is significant variation in cognitive function among schizophrenia patients with a comorbid cannabis use disorder. Thus, other factors that affect cognitive function among schizophrenia patients should be considered in addition to cannabis use.

With findings that negate complete causation between cannabis use and psychosis, other environmental factors must be interacting with one another or with underlying genetic susceptibility to increase risk of psychosis. Many risk factors cited are related to family history of psychosis, urbanicity, upbringing and exposure to *Toxoplasma gondii* (*T. gondii*) parasites and cytomegalovirus (CMV) (Torrey, Bartko, & Yolken, 2012). *T. gondii* is a coccidian protozoa of the apicomplexa family. When it infects pregnant women, it leads to a variety of severe birth defects such as deafness, retinal damage, seizures, and mental retardation. Numerous case-control studies from different countries found significant association between *T. gondii* infection and onset of psychosis, schizophrenia, or bipolar disorder. A 2011 study conducted on a Mexican population of Mestizo ethnicity found that seroprevalence of *T. gondii* IgG antibodies were higher in schizophrenic patients than in control subjects (OR=4.44, 95% CI: 1.49, 13.37) (Alvarado-Esquivel et al., 2011). An Ethiopian study which ventured to determine the magnitude of *T. gondii* infection in individuals with schizophrenia and bipolar disorder found that seroprevalence was higher in schizophrenia patients (OR= 4.7; 95% CI: 1.5, 15.1) and bipolar disorder (OR= 3.0; 95% CI: 1.1, 8.6) than in controls (Tedla et al., 2011). Other women have provided significant evidence that fetal exposure to cytomegalovirus leads to abnormal temporal lobe development (Biegon, Grossman,

Bokov, Lipitz, & Hoffmann, 2011). A large U.S. cohort study found an association between bipolar disorder and *T. gondii* (Pearce, Kruszon-Moran, & Jones, 2012)

Neuroimaging studies have implicated dysfunction of the medial temporal lobe (MTL) and dopamine system in psychosis, especially considering successful completion of certain cognitive verbal tasks (Allen et al., 2012). Childhood abuse (both sexual and physical types) is significantly associated with onset of psychotic symptoms, according to multiple studies (Kelleher et al. (2008); Lataster et al. (2006)). Childhood abuse can also interact with cannabis use to increase risk for schizophrenia. Findings from Harley et al. (2010) suggested that the presence of both childhood trauma and early cannabis use significantly increased the risk for psychotic symptoms beyond the risk posed by either risk factor alone (Harley et al., 2009). Thus, interaction between childhood trauma and cannabis use. Meta-analyses acknowledge a dose-response association with urban environment across wide definitions of urbanicity (McGrath et al. (2004); March et al. (2008); Krabbendam and Van Os (2005); Kelly et al. (2010)). Findings from the studies reveal that higher risk for schizophrenia was conferred with greater number of years lived in urbanicity and higher degree of urbanicity. The conferred stress associated with social experience of urbanicity may interact with preexisting genetic vulnerability (as expressed by several genetic polymorphisms that interact to increase risk) to overall increase risk for psychosis.

Longitudinal studies show that change in environment, such as moving from an urban to a rural environment in childhood, brings about a corresponding decrease in risk for psychotic outcome, implicating urbanicity as an significant factor in epigenetic interaction (Pedersen and Mortensen (2001); Jim van Os, Kenis, and Rutten (2010)).

Findings from a 2011 meta-analysis indicated that there is a significant association between traumatic brain injury and schizophrenia (OR=1.65; 95% 1.17, 2.23) but did not find a statically significant dose-response relationship between severity of head injury and subsequent risk of schizophrenia (Molloy et al., 2011). Evidence has also been presented to implicate pregnancy and delivery complications, some of which result in traumatic brain injury, in an increased risk for schizophrenia(Murray & Lewis, 1987) (Dalman, Allebeck, Cullberg, Grunewald, & Köster, 1999). Authors suggested that the effect appears larger in those who have family history of schizophrenia, or a genetic predisposition to psychosis, acknowledging possible interaction between genetic vulnerability and brain injury.

Other familial risk factors, such as greater paternal age, may substantially increase risk for psychosis. Miller et al. (2011) reported a minor effect size for fathers aged 35 or older, and a minor to fair effect size for fathers aged 50 years or older, when compared to fathers aged 25–29 years old. Evidence summarizing the effect of migrant status seems to be inefficient to derive conclusions about second generation immigrants and immigrants from developing countries, compared to non-immigrants, with regards to risk for psychosis (Bourque et al., 2011). Differences in psychosis risk also differ by gender, with a recent meta-analysis reporting a small effect size of increased risk for males compared to females, even after controlling for sampling bias (Aleman et al., 2003); (Matheson, Shepherd, Laurens, & Carr, 2011).

Considering the strong evidence for confounding factors, it is uncertain to what extent cannabis use predicts onset of psychosis. Cannabis use may only lead to psychosis in individuals who are genetically predisposed to mental illnesses, and the set of genes

that constitute this genetic susceptibility. Polymorphisms of psychosis vulnerability genes can potentially moderate the interaction between cannabis use and confounding factors on psychotic symptoms in the general population. The effect of such gene-environment interactions may vary significantly among populations deemed more vulnerable to psychosis and bipolar disorder due to other environmental factors.

Childhood abuse and gender are two such factors that can meaningfully moderate the association between implicated SNPs in psychosis and bipolar disorder and cannabis use (Vinkers et al., 2013). A 2011 study examined possible interaction between childhood maltreatment, cannabis use and the BDNF-Val66Met polymorphism (Alemany et al., 2011). The study concluded that individuals exposed to childhood abuse are more likely to report positive psychotic-like experiences. Results showed that Met carriers reported more positive psychotic-like experiences when exposed to childhood abuse than did individuals carrying the homozygous Val genotype. Experience of psychotic symptoms were enhanced with greater use and frequency of cannabis use, as well (Alemany et al., 2011). Similar conclusions were drawn from another 2011, which retrospectively assessed expression and frequency of BDNF Val66Met and cannabis use among 585 schizophrenia patients (Decoster et al., 2011). No evidence was found for a significant BDNF-cannabis interaction, but significant interaction between gender, BDNF and cannabis use was found. In female patients, cannabis use was associated with earlier age of onset of psychotic disorder in BDNF Met-carriers but not among homozygous Val carriers (Decoster et al., 2011). Cannabis use was also significantly associated with earlier age of onset of psychotic disorder, indicating a dose-response effect with higher frequency and earlier age at first use.

Other studies have explored density and frequency of cannabis use among those with varying 5-HTTL genotypes. 5-HTT is considered a candidate gene for bipolar disorder, and many studies have examined its functional polymorphism (5-HTTLPR) with regards to its effect on emotional development(De Pradier, Gorwood, Beaufils, Adès, & Dubertret, 2010). 5-HTTLPR represents a serotonin transporter gene polymorphism which may moderate psychopathological reactions to stressful experiences(Hariri et al., 2002). De Pradier et al. (2010) found that the short allele form of 5-HTT and cannabis abuse were significantly more frequent among patients with psychotic symptoms than in those without ($p=0.01$ and p= 0.004, respectively). This study also attributed a significant gene-environment interaction between the presence of the *s* allele and childhood sexual abuse as strong risk factors for cannabis abuse or dependence in its cohort of bipolar patients.

There may be significant cognitive differences to cannabis users across the different polymorphisms implicated in cannabis-schizophrenia and cannabis-bipolar disorder interactions. For example, (Verdejo-Garcia et al., 2013) concluded that cannabis users carrying the COMT val/val genotype exhibited lower accuracy of sustained attention than val/val non-users. Cannabis users carrying the COMT Val allele also committed more monitoring/shifting errors than cannabis users carrying the met/met genotype (Verdejo-Garcia et al., 2013). Furthermore, the gene encoding the CB2 cannabinoid receptor, CNR2, has been shown to be associated with drug addiction and mood fluctuations characteristic of bipolar disorder. A 2011 study found a statistically significant association between bipolar disorder and the rs41311993 (524C>A; Leu133Ile) polymorphism (Minocci et al., 2011). This missense polymorphism encodes a

base pair change that codes for either leucine or isoleucine. Patients with the 524C>A; Leu133Ile polymorphism presented with significant lesions in white matter, occurrence of which has been associated with cognitive dysfunction (Voss et al., 2013).

There are several genes and polymorphisms implicated in the cannabis-psychosis association. A polygenic risk score, which reflects the cumulative burden of risk alleles carried by an individual as identified in a GWAS, is commonly utilized to estimate additive effect (Power et al., 2014); (Shaun M Purcell et al., 2009). It is theorized that the risk for either cannabis use or psychosis is greater with increasing polygenic risk score, if significant. The threshold for significance is at discretion of the researcher, and it could be noteworthy to examine whether effect of cannabis could potentially drive an individual's risk (as determined by preexisting genetic vulnerability) beyond the threshold (Dudbridge & Gusnanto, 2008). As cannabis use is often associated with tobacco use, illicit drug use and general polydrug use it should be necessary to evaluate genetic susceptibility of cannabis use with that of several other substances commonly abused (Malmberg et al., 2010). However, a 2014 study found that polygenic scores for age at onset of smoking were significantly correlated with age at regular drinking (P=0.001, R2=1.1-1.5%) while scores for smoking cessation did not significantly predict cannabis use (Vink et al., 2014). Graphical comparison of ranges of mean standardized schizophrenia polygenic risk scores can assist in evaluating the risk imposed by genetic susceptibility by cannabis, in addition to the preexisting genetic susceptibility to psychosis (Derks et al., 2012). Polygenic risk scores were used to detect significance between individual's burden of schizophrenia risk alleles and use of cannabis in Power et al. (2014). Significant association was found for comparing individuals who have ever

versus never used cannabis ($P= 2.6 \times 10^{-4}$) and for quantity of use within users ($P=3.0 \times 10^{-3}$). Significance of the association varied across different significance thresholds, however.

Existing literature is lacking information on bipolar disorder and cannabis use in relation to genetic susceptibility. The connection between schizophrenia and bipolar disorder has been widely examined, but some schizophrenia symptoms may meaningfully overlap with bipolar disorder. Common symptoms among the two disorders may be due to shared liability genetic pathways. Cannabis use can exacerbate or alleviate such symptoms or may also have similar liability molecular pathways. Hence, it may be of interest to further explore the connection between genes implicated in bipolar disorder and those underlying cannabis use.

In consideration of existing literature, schizophrenia and bipolar disorder, this thesis will aim to discern molecular pathways involving relevant susceptibility genes and potential causal pathways in schizophrenia, schizoaffective, and bipolar I disorders to determine presence of interaction of biological pathways implicated in these variables of interest.

Methods

Study Design and Population

To test for interaction between the molecular pathways of schizophrenia, bipolar disorder, and cannabis use, a list of single nucleotide polymorphisms (SNPs) associated with each outcome was derived from two genome wide association studies from the Psychiatric Genetics Consortium (PGC) and one GWAS from NIH dbGaP (Pamela Sklar et al., 2011); (Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011);

(S. P. G.-W. A. S. Consortium, 2011); (dbGAP, 2009a). A list of SNPs for cannabis use was derived from the National Institute of Health dbGaP Study of Addiction: Genetics and Environment (SAGE) Illumina SNP batch (dbGAP, 2009a) . SAGE is a large case-control study which aims to detect susceptible genetic variants for addiction. ("PGC Downloads," 2013); (Pamela Sklar et al., 2011). The Psychiatric GWAS Consortium (PGC) has performed multistage mega-analyses for several psychiatric disorders, including schizophrenia and bipolar disorder.

For schizophrenia and bipolar disorder, we uploaded the most recent dataset referenced from the Cross-Disorder Group of the Psychiatric Genomics Consortium (Cross-Disorder Group of the Psychiatric Genomics, 2013). In this thesis, I did not endeavor to compare bipolar and SZ directly, but rather to examine the relationship between susceptibility variants for each of these disorders in the context of cannabis abuse and its underlying genetic labiality.

For schizophrenia and bipolar disorder, we downloaded full datasets from http://www.med.unc.edu/pgc/downloads using the “Cross-disorder Full” data set (versus the “clump” dataset). This had been processed through the PGC pipeline (Cross-Disorder Group of the Psychiatric Genomics, 2013). The primary references for the schizophrenia dataset is Schizophrenia Psychiatric Genome-Wide Association Study (2011). For the BP dataset the primary reference is (Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011). These papers have extensive discussion of issues related to population stratification and statistical consideration genomic inflation. A prominent point of this discussion concerns the polygenic nature of SZ and BP, and the difficulty in discerning genomic inflation (confounding) due to population stratification versus the

expected statistics based on a polygenic model of inheritance. Thus, the p-values and odds ratios for association of SNPs with these disorders were not adjusted for these factors beyond those performed by the respective PGC studies.

Schizophrenia

The characteristics of the variants and subjects of this dataset are shown in Tables 1 and 4. Approximately 43% of the combined sample are cases (n=9,394), and case inclusion criteria was broadly operationalized as schizophreniform and schizoaffective disorder were also included with schizophrenia as possible outcomes. Schizophreniform is characterized has having presence of symptoms of schizophrenia but is distinguished from that condition by its shorter duration, usually 1 to 6 months. DSM-IV criteria, the standard classification of mental disorders used by mental health professionals in the United States, ICD-10, the 10th revision of the International Statistical Classification of Diseases and Related Health Problems, and RDC, a standardized collection of influential psychiatric diagnostic criteria published in the 1970s were used as diagnostic criteria instruments across all samples. Age at onset is operationalized as the age at illness onset, or the age of first psychiatric contact or impairment. Average age of onset across all studies is 24.5 years (range= 14.6-34.0).

The data downloaded for this thesis consisted of columns for SNPs, chromosome number, number of base pairs for the polymorphism, allelic variations, the odds ratio and associated standard error, and p-value. An example of the primary data layout is provided in Figure 1.

Figure 1. Original Data Layout for PCG Metadata on Schizophrenia

SNP_ID	Chromosome	Base pairs	a1	a2	OR	SE	p-value
rs3131972	1	742584	A	G	1	0.0966	0.9991
rs3131969	1	744045	A	G	1	0.0925	0.9974

Figure 2 and Table 4 show there were 1,237,958 variants from the originally downloaded dataset for schizophrenia. SNPs were presented alphanumerically, with their *rs* identification number in front of the physical gene location.

Bipolar Disorder

Characteristics of subjects and variants from this PGC dataset are shown in Tables 2 and 4. The original GWAS that gave rise to this list consisted of 11 different study samples, many of which contain populations of largely European and American ancestry (see Table 2). Approximately 45% of the combined sample are cases (n=7,481), and case inclusion criteria was broadly operationalized as bipolar disorder, types 1 and 2, schizoaffective-bipolar disorder, and bipolar disorder not otherwise specified as possible outcomes. DSM-IV criteria, DSM-IIR, a revision of the APA's second edition of DSM criteria, and RDC, a standardized collection of influential psychiatric diagnostic criteria published in the 1970s were used as diagnostic criteria instruments across all samples. The majority of the overall sample was diagnosed with bipolar disorder, Type 1 (84%), followed by bipolar disorder, Type 2 (11%), schizoaffective-bipolar disorder (3.5%), and then bipolar disorder, not otherwise specified (1.5%).

Figure 3 and Table 4 show there were 2,427,220 variants from the originally downloaded dataset for bipolar disorder. SNPs were presented alphanumerically, with their *rs* identification number in front of the physical gene location.

Cannabis use

Characteristics of subjects and variants from the dbGaP SAGE dataset are shown in Tables 3 and 4. SAGE consists of three large, complementary datasets: the Collaborative Study on the Genetics of Alcoholism (COGA), the Family Study of Cocaine Dependence (FSCD), and the Collaborative Genetic Study of Nicotine Dependence (COGEND).

COGA was initiated in 1989 and is a large-scale family study that has had as its primary aim the identification of genes that contribute to alcoholism susceptibility and related characteristics. Subjects were recruited from 7 sites across the U.S. Alcohol dependent probands were recruited from treatment facilities and assessed by personal interview. The FCSD was initiated in 2000 as a case-control study of cocaine dependence funded through NIDA. Cocaine dependent individuals were systematically recruited from public, private or outpatient chemical dependency treatment units in the greater St. Louis Metropolitan area. Controls were matched through a Missouri Driver's License Registry, matched by age, race, gender and residential zip code. COGEND was initiated in 2001 as a project grant funded through the National Cancer Institute. Nicotine dependent cases and non-dependent smoking controls were identified and recruited from Detroit and St. Louis. The overarching goal of SAGE is to identify novel genetic factors that contribute to addiction through a large-scale SWAS of DSM-IV alcohol or illicit drug dependent cases and non-dependent control subjects of European and African American descent.

Considering cannabis use commonly co-occurs with tobacco and other illicit drug use, it is difficult to sample population-based cannabis users, who also isolate themselves from other forms of substance abuse. Many substance use disorders are enriched and

selected with other disorders, such as alcoholism. The fact that substance users have a tendency to become involved with multiple substances allows for broader sampling methods to include both single and polydrug users in substance abuse epidemiological studies. Utilization of a dataset that operationalizes cannabis use as occurring in isolation or enriched with other drug use was ideal for this thesis.

In lieu of this, the NIH Study of Addiction and Genetics (dbGAP, 2009a) was used to derive candidate genes for cannabis use as it is the largest cohort of complementary data sets that is publicly available. The study includes extensive information on users that abuse alcohol, cocaine, cannabis and nicotine among other drugs. By selecting for and assessing a sample not limited to isolated cannabis users, statistical power is largely enhanced, thereby increasing our ability to detect a true association between susceptible genes in cannabis use with those implicated in schizophrenia and bipolar disorder. Recent, relevant studies also reference use of dbGaP SAGE, using comparable justification. Palmer et al. (2014) estimated the aggregate effect of common SNPs on multiple indicators of comorbid drug problems using a subset of 2,596 subjects from the original SAGE sample of 4,121 subjects. Justification for use of SAGE included explanation of the study representing the most comprehensive dataset accessible to the public. Assessment of both single and polydrug users were also significant in attaining higher statistical power in Palmer et al. (2014), as it is in this analysis.

The data were downloaded from dbGaP SAGE (study accession phs000092.v1.p1) (dbGAP, 2009a). The subjects were recruited from eight study sites in seven states and the District of Columbia in the United States. All subjects' life time dependencies on these six dependencies are diagnosed by using the Diagnostic and Statistical Manual of Mental

Disorders, Fourth Edition (DSM-IV). All samples were genotyped on ILLUMINA Human 1 M platform at the Center for Inherited Disease Research in Johns Hopkins University.

Before statistical analysis and release to the public, data from dbGaP SAGE had been extensively cleaned. Before being released to the public, the COGEND project had 4,324 subjects, COGA project had 1,989 subjects, and FSCD had 1,267 subjects. Non-duplicate, unrelated cannabis-exposed individuals self-identified as EA or AA were selected for analysis. SNP QC filters for analysis inclusion were a call rate of at least 98%, a minor allele frequency > 1%, and a HWE p-value > 1x10-4. The analysis of genotype effect on cannabis dependence was modeled with an adjusted logistic regression model performed with PLINK. SNP genotypes were coded log-additively as 0, 1, or 2 copies of the minor allele (based on the allele frequencies in the self-identified EAs). Covariates included in each model were participant's gender, age at interview (collapsed into quartiles), source study (COGA, COGEND, or FSCD), and self-identified race (EA or AA). The age quartile variables were defined as 34 years and younger (reference), 35-39 years, 40-44 years, and 45 years and older. The reference level for source study indicator variables was COGEND. The mean non-Y SNP call rate and mean sample call rate was 99.7% for the released dataset, and the study duplicate reproducibility was 99.98%. The genotype concordance rate in the overlapping subjects (n=1,477) between COGEND, FSCD and COGA was 99.98%.

All substance dependence symptom counts were residualized over sex, age quartiles (Laura J Bierut et al., 2010), primary study source (COGA, COGEND or FSCD (reference)), and ancestry. Quality control processes included various steps. Samples from COGEND, FSCD, and COGA with heterozygosity below 0.32 were removed from the filtered release set but are present in the unfiltered release set. Trios were tested for Mendelian errors. No trio in

the filtered or unfiltered datasets has Mendelian errors for >2% of markers, confirming reported relationships between individuals. Samples with overall genotype call rates below 95% were not included in the filtered release set.

Investigators for SAGE adjusted for population stratification through a series of steps. In the event that a particular subject was included in both the COGEND and COGA studies, COGA subjects were retained; and finally, only 480 COGA subjects having genotype data were merged into the SAGE sample. Then, subjects with missing or inconclusive diagnostic information that prevented them from being reliably classified as cases or controls were excluded. The subjects with allele discordance, duplicated IDs, potential sample misidentification, sample relatedness, other sample misspecification, gender anomalies, chromosome anomalies (such as aneuploidy and mosaic cell populations), missing race, non-EA and non-AA ethnicity, and population group outliers were also screened out sequentially.

A procedure described by Alkes L Price et al. (2006), a principal components analysis, was utilized to classify subjects on the basis of genetic ethnicity and to filter out subjects (n=12) where there was a mismatch between self-identified and genetically-inferred ethnicity. Two main principal components, corresponding to European versus African ancestry (PC1) and Hispanic versus non-Hispanic ancestry (PC2), resulted from this procedure (Bierut et al. 2010). Both principal components were included as covariates over which symptom counts were residualized. Population structure was first evaluated using PCA implemented in the software package EIGENSTRAT (A. L. Price et al., 2006) (with all autosomal SNPs having a call rate >95%). Each individual received scores on each principal component. Because related subjects, non-EA and non-AA subjects, and any population group outliers were excluded first, their effect on PCA was removed. The first principal component (PC1)

separated the self-identified EA and AA subjects very well, which was highly consistent with results from L. J. Bierut et al. (2010) study. PC1 was used to measure the continuous ethnicity variance for EAs and AAs. The cut-off value (=0) of PC1 separated the "genetic" EAs and AAs. A total of 12 subjects were mismatched between "genetic" and self-identified ethnicity. The second principal component separated the self-identified Hispanic subjects from the non-Hispanic subjects.

European Americans (EA) and African American (AA) subjects constituted groups with high levels of ancestral homogeneity (see Supplementary Figure 1). After filtering out the subjects with a missing genotype call rate ≥2% across all markers, the final sample included 4,121 individuals, including 1,413 EA cases, 1,518 EA controls, 681 AA cases and 508 AA controls. Investigators filtered out the markers on all chromosomes with an overall missing genotype call rate ≥2%, the monomorphic markers, and the SNPs with minor allele frequencies (MAFs) <0.01 in either EAs or AAs. The SNPs that deviated from HWE ($P < 10^{-4}$) within EA or AA controls were also excluded. This selection process yielded 640,020 markers in EAs and 273,010 markers in AAs (DbGaP, 2009b). Investigators computed from p-values of association between cannabis use and genetic variants a low genomic inflation factor (GIF) of 1.07 in EAs and 1.03 in AAs (DbGaP, 2009b). This suggests a relatively high level of homogeneity among the samples.

The majority of the original SAGE sample is White (65%), followed by African Americans (32%), Hispanic/Latinos (3.3%) and then Asians, and those of unknown and mixed race (0.02%). Average age of the sample was 39 years (s.e. = 9.1) and largely consisted of males (males=93.4%, females= 6.6%). Approximately 40% of subjects had some high school education or a high school diploma ('Some high school' = 16.6%;

'High school (diploma)' = 23.4%). Half of the sample had some college education or a Bachelor's degree ('Some college' = 29.9%; 'Bachelor's degree' = 20.0%). Only 10.2% (n=388) of the sample had some post-graduate education.

Information on cannabis use was ascertained via yes/no questions on a survey (e.g.: 'Have you ever used marijuana or hashish?'). An overwhelming majority of the sample had ever used marijuana (n=3,180 (77.2%)) and non-users represented 23% of the sample (n=938). There were 2 people (0.05%), defined as 'Others', who declined to answer questions regarding their marijuana use. Mean age at initiation of cannabis use was 16.7 years (s.e. =4.4) and mean age of marijuana use recency was 28.8 years (s.e. = 9.3). DSM-IV criteria was used as diagnostic criteria across all studies for assessment of cannabis dependence. Only 18.3% (n=753) of the sample were diagnosed as cannabis dependent, with 81.6% (n=3,361) of the sample yielding negative diagnosis and 17% (n=7) having inconclusive diagnosis results. Mean age of onset of cannabis dependence was 19 years (s.e. = 5.8) and mean age of recency of dependence was 28.3 years (s.e. =8.6). Dependence, on average, occurs roughly three years after initiation of cannabis use, and recency of dependence almost occurs simultaneously with age of marijuana use recency, on average.

There were 917,030 SNPs from the original downloaded dbGaP SAGE dataset (see Tables 3 and 4 and Figure 4). SNPs were presented alphanumerically, with their *rs* identification number in front of the physical gene location. Information on SNPs included associated p-value, allelic variants and odds ratios.

Variant Isolation and Gene Annotation

Schizophrenia Variant Isolation

Columns of variants for schizophrenia were isolated by varying alpha (α) levels of statistical significance into different Microsoft Excel spreadsheets. Variants for each variable were separated by significance cutoffs 0.1, 0.01, 0.001, 0.0001, and 0.00001, using p-value cutoffs as Power et al. (2014) utilized in calculating polygenic risk scores for the association between cannabis use and schizophrenia.

These lists were created by importing the original dataset from PGC as a text file into SAS 9.4. The dataset was initially sorted by p-value, and then subsetted into five different datasets according to the p-value cutoff. For example, the first dataset created only had variants with p-values less than 0.1. As expected, with increasing α significance threshold, the number of eligible variants for pathway analysis reduced significantly (see Table 5). These five datasets were then exported from SAS 9.4 as Excel spreadsheets, saved into one Excel file. There were 1,237,958 primary eligible variants to be assessed in molecular pathway analysis (see Table 4 and Figure 2).

To further prepare for analysis, each list was truncated to only include variables "SNP_ID" and "p-value." This would better facilitate analysis through Ingenuity Pathway Analysis, the software which would derive molecular pathways for variants.

Bipolar Disorder Variant Isolation

The same process was conducted for isolating and annotating variants for bipolar disorder. There were 2,427,220 primary eligible variants to be assessed in molecular pathway analysis (see Table 4 and Figure 3). Counts of eligible variants at each significance level are shown in Table 5.

Cannabis Use Variant Isolation

The NIH dataset detailing cannabis use was downloaded as a .gz zip file and extracted using 7-Zip, an open source Windows utility for manipulating archives. A tab-delimited text file was extracted and then imported into Excel for data cleaning. Annotation of the dataset above variables was copied and pasted into a separate spreadsheet, apart from the variants, providing information of the descriptive information of the dataset. Variables denoting variant ID and associated p-values were renamed to match the variable names of the PGC datasets for schizophrenia and bipolar disorder. There were 917,030 eligible variants to be assessed in molecular pathway analysis (see Table 4 and Figure 4).

The process of isolating variants for analysis was the same for cannabis use as it was for bipolar disorder and schizophrenia. Counts of variants at each significance level for cannabis use are shown in Table 5.

Assessment of Molecular Pathways

Lists of variants with p-values less than $\alpha = 0.001$ were chosen for analysis to minimize number of false positives in multiple comparisons analyses, and to generate lists of potential analysis candidates that had at least 1,000 variants with associated genes. This thesis utilizes the same criteria for the significance cutoffs for variant selection as was used in Agrawal and Lynskey (2009). Selection of lists of variants with p-values less than 0.0001 or 0.00001 would lead to too few variants (at $\alpha = 0.00001$, there are only 15 variants for cannabis use) to be included in analysis (see Table 5). The final lists of annotated genes were uploaded into Ingenuity Pathway Analysis, which identifies

molecular relationships, biological mechanisms, functions and potential causal pathways between variants, genes and neurotransmitters.

Gene Annotation

SNP array data from Illumina or Affymetrix platforms, or directly uploaded NCBI dbSNP IDs for SNPs of interest can be uploaded into Ingenuity Pathway Analysis (IPA) (Guda, 2013). IPA automatically maps SNPs falling within or near gene-coding regions to the relevant gene ortholog for subsequent pathways analysis and exploration. A SNP was mapped to a gene if the SNP fell within the gene-coding region or within the 2 kilobases upstream or 0.5 kilobases downstream range of the gene-coding region (Guda, 2013).

Additional Mapping of Genes

Mapping of the dbSNP IDs to Ingenuity genes was done through Entrez Gene IDs. The mappings between dbSNP IDs and Entrez Gene IDs were generated based on the Entrez Gene database. This was performed using a configuration and tested for any SNPs not annotated in IPA.

Analytical Plan

Only variants that have p-values less than $a = 0.001$ were included in analysis, and counts of those variants stratified by each outcome are provided in Table 6. There were 4,680, 6,884 and 1,243 variants isolated for schizophrenia, bipolar disorder and cannabis use respectively for purposes of analysis. Ingenuity Pathway Analysis was utilized to identify and compare common genes among the variants submitted across cannabis use and schizophrenia and cannabis use and bipolar disorder.

Only those genes that overlap in each comparison were mapped in a molecular pathway network diagram, and unmatched genes (and associated variants) would be further analyzed in dbSNP variant submission to verify gene annotation.

Regarding mapping of genes, the software assigned statistical scores to pathways, taking into account the genes used in analysis, network size, and the total number of molecules in Ingenuity Knowledge Base. The network score is the negative logarithm of p-value, which reflects the probability of finding the focus molecules in a given network by random chance. The identified networks were then presented as graphs, which demonstrate the molecular relationships between gene products (Ingenuity, 2015).

Table 1: Description of 2011 PGC GWA Meta-Analysis of Schizophrenia (N= 21,856)						
Sample [1]	***Ancestry***	***Case (n)***	***Control (n)***	***Diagnostic Criteria*** [2]	***Diagnosis*** [3]	***Age at Onset*** [4]
Cardiff UK	UK	472	2,394	DSM-IV	SZ	23.8 (7.9)
CATIE	United States	402	207	DSM-IV	SZ	14.6 (11.0)
ISC - Aberdeen	UK	720	698	DSM-IV, ICD-10	SZ	23.9 (10.0)
ISC - Cardiff	Bulgaria	527	609	DSM-IV	SZ	24.5 (7.2)
ISC - Dublin	Ireland	270	860	DSM-IV, ICD-10	SZ	24.7 (7.2)
ISC - Edinburgh	UK	368	284	DSM-IV	SZ	24.3 (8.3)
ISC - London	UK	518	491	RDC	SZ	23.6 (7.4)
ISC - Portugal	Portugal	346	215	DSM-IV	SZ or SA depressed	23.2 (7.3)
ISC - SW1	Sweden	168	167	DSM-IV	SZ	33.4 (10.1)
ISC - SW2	Sweden	390	229	DSM-IV	SZ	34.0 (11.3)
MGS	United States, Australia	2,679	2,484	DSM-IV	SZ or SA (10% SA)	21.4 (7.0)
SGENE - Bonn	Germany	474	1,304	DSM-IV	SZ	21.9 (6.5)
SGENE - Copenhagen	Denmark	482	457	ICD-10	SZ or SA	26.7 (9.3)
SGENE - Munich	Germany	434	351	DSM-IV	SZ	23.6 (8.6)
SGENE - TOP3	Norway	248	351	DSM-IV	SZ, schizophreniform (6%), or SA (17%)	24.9 (8.9)
SGENE - UCLA	The Netherlands	704	631	DSM-IV	SZ	26.3 (9.3)
Zucker Hillside	United States	192	190	DSM-IV	SZ, schizophreniform, or SA	21.7 (5.9)
TOTAL		**9,394**	**12,462**			

[1] *Primary publication reporting individual sample level genotypes for bipolar disorder are listed.*

[2] *DSM-IV is the Diagnostic and Statistical Manual of Mental Disorders (DSM) is the standard classification of mental disorders used by mental health professionals in the United States. ICD-10 is the 10th revision of the International Statistical Classification of Diseases and Related Health Problems. RDC are a colelction of influential*

[3] *SZ= Schizophrenia, SA= schizoaffective disorder*

[4] *Age at onset is usually age at illness onset, though occasionally is age at first psychiatric contact or age at impairment*

Table 2: Description of 2011 PGC GWA Meta-Analysis of Bipolar Disorder (N=16,731)

Sample[1]	*Ancestry*	*Case (n)*[2]	*Control (n)*	*BD1*	*BD2*	*SAB*	*BD-NOS*[3]	*Diagnostic Criteria*	λ^4
BOMA-Bipolar Study, University of Bonn and CIMH Mannheim[a,c,d,e]	German	675	1,297	673 (99.7%)	2 (0.3%)	0	0	DSM-IV	1.04
Genetic Association Information Network/ Bipolar Genome Study[a,c,d,e]	European-American	542	649	516 (95%)	0	26 (5%)	0	DSMIIR & IV	1.03
GlaxoSmithKline (GSK)[c]	British, Canadian or Scottish	890	902	632 (71%)	80 (9%)	134 (15%)	44 (5%)	DSM-IV	1.03
Pritzker Neuropsychiatric Disorders Research Consortium[c]	European-American	1,130	718	1130 (100%)	0	0	0	DSMIIR & IV	1.02
Systematic Treatment Enhancement Program for Bipolar Disorder (STEP1)[a,d]	European-American	922	645	922 (100%)	0	0	0	DSM-IV	1.03
Systematic Treatment Enhancement Program for Bipolar Disorder (STEP2)[a]	European-American	659	192	108 (16%)	551 (84%)	0	0	DSM-IV	1.02
Thematically Organized Psychosis (TOP) Study[a]	Norwegian	203	349	119 (59%)	58 (29%)	6[5] (3%)	20 (9%)	DSM-IV	1.03
Trinity College Dublin[a]	Irish	150	797	150 (100%)	0	0	0	DSM-IV	1.02
University College London[a,d]	British	457	495	457 (100%)	0	0	0	DSM-IV	1.01
University of Edinburgh[a]	Scottish	282	275	282 (100%)	0	0	0	DSM-IV	1.03
Wellcome Trust Case-Control Consortium[a,c,f]	British	1,571	2,931	1300 (83%)	133 (8%)	97 (6%)	41 (3%)	RDC	1.08
TOTAL		**7,481**	**9,250**	**6,289 (84%)**	**824 (11%)**	**263 (3.5%)**	**105 (1.5%)**		

[1] *Primary publication reporting individual sample level genotypes for bipolar disorder are listed.*

[2] *Cases include BD1, BD2, SAB, BD-NOS.*

[3] *BD-NOS includes manic disorder*

[4] *Genomic control lambda*

[5] *Includes psychotic depression (n=3)*

Table 2. References:

[a] *Djurovic et al. (2010)*

[b] *Ferreira et al. (2008)*

[c] *Scott et al. (2009)*

[d] *P. Sklar et al. (2008)*

[e] *The Wellcome Trust Case Control Consortium (2007)*

Table 3. Descriptive Statistics of sample for cannabis use traits (N=4,121)[a]			
Demographics			
		Mean (s.e.) / N(%)	
Mean age[1]		39.0 (9.1)	
Education[2,3]			
Some high school		633 (16.6)	
High school (diploma)		892 (23.4)	
Some college[4]		1,114 (29.9)	
Bachelor's degree		765 (20.0)	
Post graduate education (M.S., M.A., J.D., M.D., Ph.D)		388 (10.2)	
Gender			
Female		274 (6.6)	
Male		3,847 (93.4)	
Race[2]			
Asian		1 (0.02)	
Black		1,340 (31.5)	
Hispanic/Latino		140 (3.3)	
Mixed		1 (0.02)	
White		2,772 (65.1)	
Unknown		1 (0.02)	
Cannabis Use			
	Users	***Non-users***	***Others***
Ever used marijuana	3180 (77.0%)	938 (22.7%)	13 (0.3%)
Mean age at initiation (s.e.)	16.7 (4.4)	--	--
Mean age when last used marijuana (s.e.)	28.8 (9.3)	--	--
Cannabis Dependence			
	Positive	***Negative***	***Unknown***
DSM-IV cannabis dependence diagnosis	753 (18.3%)	3361 (81.6%)	7 (17.0%)
Mean age of onset (s.e.)		19.0 (5.8)	
Mean age of recency of dependence (s.e.)		28.3 (8.6)	

[1] *Recorded at time of interview*

[2] *Information stratified by marijuana use status not available*

[3] *Highest level of education completed*

[4] *defined as at least one year of college or technical school*

Table 3. References:

[a] *dbGAP (2009a)*

Table 4. Summary Statistics of Annotated Variants[1]		
Variable of Interest	*Data Source*	*Primary Variants Eligible (n)*
Schizophrenia	Schizophrenia Psychiatric Genetics Consortium[a]	1,237,958
Bipolar Disorder	Psychiatric GWAS Consortium Bipolar Disorder Working Group[b]	2,427,220
Cannabis Use	NIH dbGAP Study of Addiction: Genes and Environment[c]	917,030

[1] Numbers of variants presented are derived directly from their respective datasets (downloaded from their respective references. This data is not consistent with the data presented in original publications for each dataset.

Table 4. References:

[a] *Consortium (2011)*
[b] *Psychiatric GWAS Consortium Bipolar Disorder Working Group (2011)*
[c] *dbGAP (2009a)*

Table 5. Number of Variants and Candidate Genes for Variables at Varying Significance Levels					
Variable of Interest	*α=0.1*	*α=0.01*	*α=0.001*	*α=0.0001*	*α=0.00001*
Schizophrenia Variants[a]	75,869	20,903	4,680	1,252	518
Bipolar Disorder Variants[b]	306,823	43,729	6,884	1,195	358
Cannabis Use Variants[c]	100,713	10,993	1,243	141	15

Table 5. References:

[a] *Consortium (2011)*
[b] *Psychiatric GWAS Consortium Bipolar Disorder Working Group (2011)*
[c] *dbGAP (2009a)*

Table 6. Number of Variants to be Input into Ingenuity Pathway Analysis	
Variable of Interest	*α=0.001*
Schizophrenia Variants[a]	4,680
Bipolar Disorder Variants[b]	6,884
Cannabis Use Variants[c]	1,243

Table 6. References:

[a] *Consortium (2011)*
[b] *Psychiatric GWAS Consortium Bipolar Disorder Working Group (2011)*
[c] *dbGAP (2009a)*

Figure 2

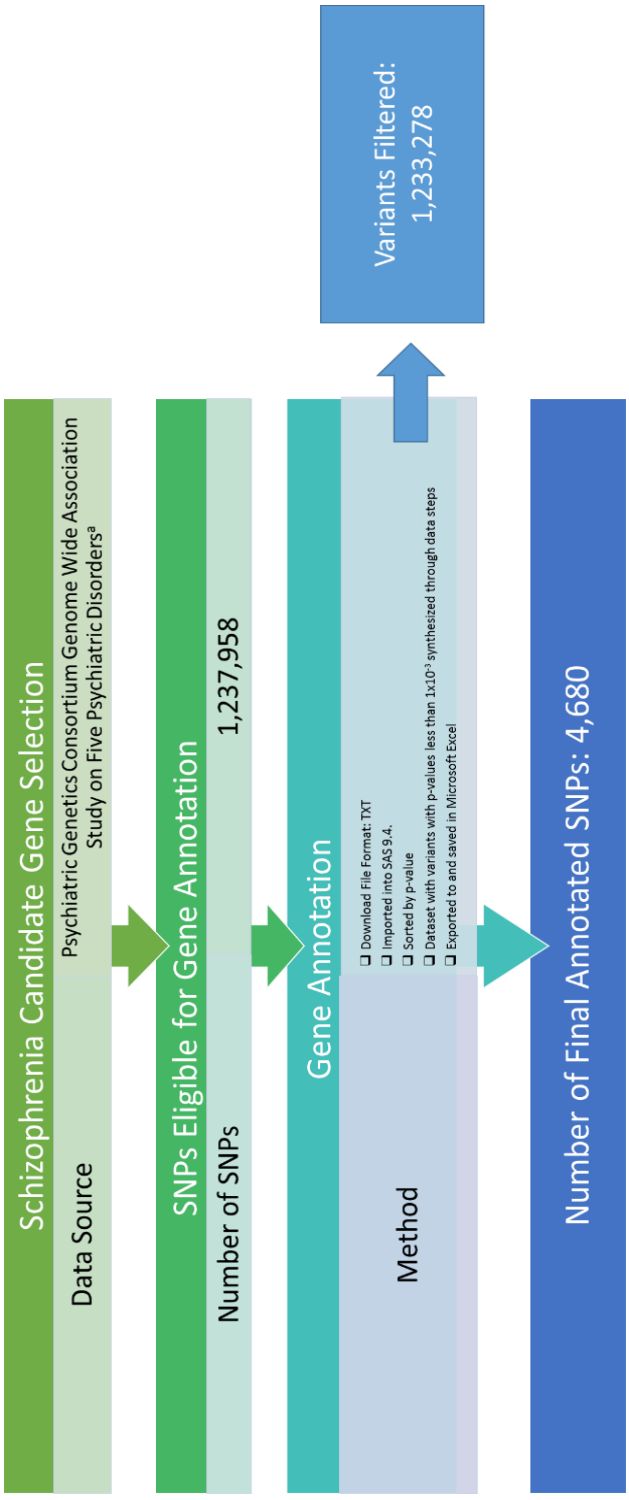


Figure 2. References:

[a] *(Consortium, 2011)*

Figure 3

Figure 3 References:

[a] *Psychiatric GWAS Consortium Bipolar Disorder Working Group (2011)*

Figure 4

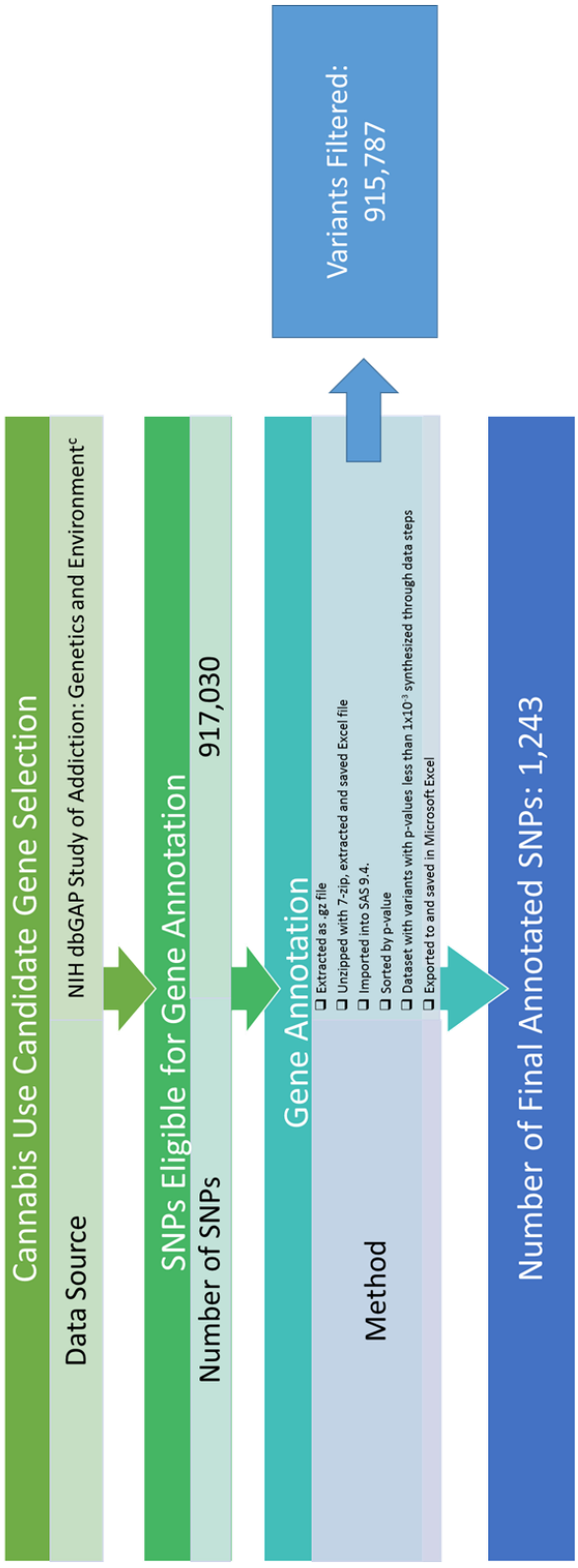


Figure 4 References:

[a] *dbGAP (2009a)*

Results Figures

Table 7. Summary of Mapped Variants from Ingenuity Pathway Analysis					
Variable	*Uploaded SNPs*[1]	*Unmapped SNPs (n/%)*	*Mapped SNPs (n/ %)*	*Duplicate SNPs (n/ %)*	*Unique SNPs mapped (n/ %)*
Schizophrenia	4,680	2,300 (49.1%)	2,380 (50.9%)	0 (0%)	2,380 (50.9%)
Bipolar Disorder	6,884	3,302 (48.0%)	3,582 (52.0%)	34 (0.5%)	3,547 (51.5%)
Cannabis Use	1,243	570 (54.1%)	673 (54.1%)	6 (0.5%)	667 (53.7%)

[1] *SNPs had p-value less than* $1x10^{-3}$

Table 8. Ingenuity Pathway Analysis Gene Summary		
Variable	*Unique Genes Mapped and Analysis Ready (n)*	*Overlapping Genes with Cannabis Use (n)*
Schizophrenia	689	44
Bipolar Disorder	639	42
Cannabis Use	342	--

Figure 5. Cross-Comparison of Genes Between Schizophrenia and Canabis Use

Overlapping Mapped Genes from Ingenuity Pathway Analysis

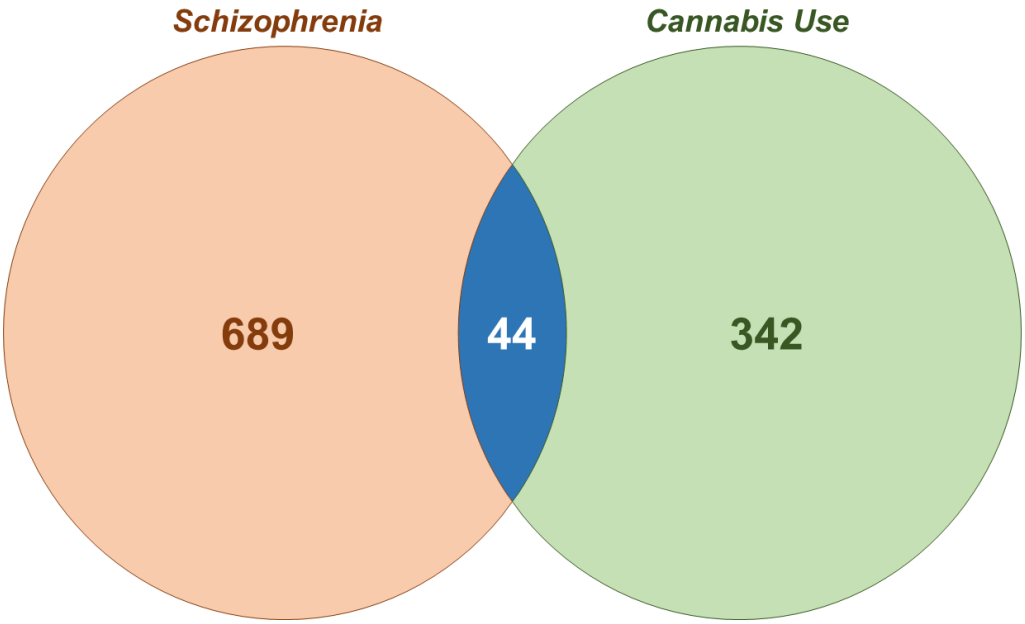


Figure 6. Cross-Comparison of Genes Between Bipolar Disorder and Cannabis Use

Overlapping Mapped Genes from Ingenuity Pathway Analysis

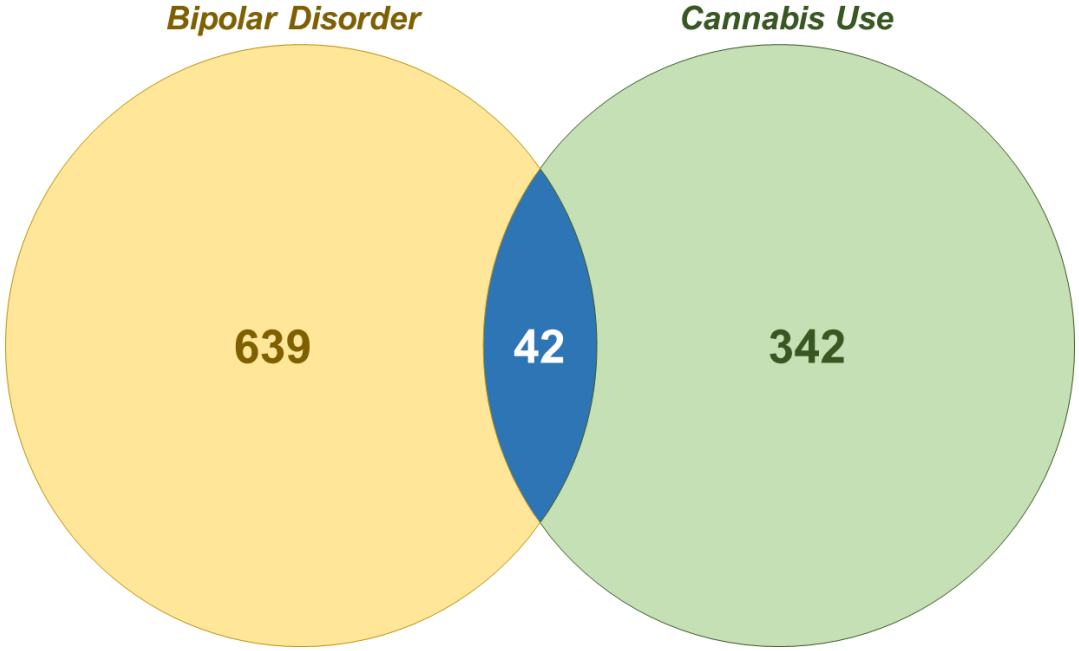


Table 9. Shared Genes and Associated Variants Between Schizophrenia and Cannabis Use from Ingenuity Pathway Analysis

Intersection Cannabis-Schizophrenia	SNPs in Cannabis use	SNPs in SZ
ADAMTS17	rs8029650	
AGBL1	rs2034633	rs16977933
	rs11073678	
	rs959095	
B3GALT1	rs1897339	rs17643505
BICD1	rs792853	rs4931619
	rs326633	rs7961369
C9orf72	rs3739526	rs10967991
		rs10757668
		rs12347201
		rs12349820
CACNA1D	rs3774533	rs219847
	rs1380605	
CCDC85A	rs13404821	rs6760801
		rs17047819
CDH4	rs6129093	rs6760801
		rs17047819
		rs2427104
		rs6062121
		rs17734201
		rs12479835
		rs1892320
		rs6061402
		rs6061412
		rs6062000
		rs2427101
		rs6061762
		rs4925349
		rs2427106
		rs16985459
		rs2427094
CDH10	rs10069640	rs13174538
		rs11747412
		rs10942057
		rs12516067
CDH13	rs8057717	rs8045995
		rs3845200
CNTNAP2	rs1860681	rs1548743
	rs700305	rs1014686

	rs2906310	rs10242598
	rs2906308	rs740809
	rs2972126	rs6464781
	rs2906300	rs1466971
	rs2373346	rs1002975
CSMD1	rs4875772	rs11136729
	rs6996668	rs10503253
		rs6558863
		rs11993860
		rs10503256
		rs10108725
		rs6558872
		rs1594352
		rs10098869
		rs1875897
		rs10107472
		rs3990909
		rs10086105
		rs1594353
		rs7834964
		rs1583130
		rs10102768
		rs10866968
		rs7839613
		rs10105113
		rs10104209
		rs10091134
		rs1430447
		rs4568637
		rs1595468
		rs17070107
DLG2	rs7951686	rs3815986
		rs11233649
		rs2009715
		rs1469609
FMN1	rs16960151	rs3817591
	rs7165427	rs1258749
	rs12594394	
	rs12595789	
	rs1562930	
	rs4780055	
GABBR2	rs10818743	rs914665
	rs995213	rs914662

		rs2808539
GREM2	rs12024485	rs2185283
		rs3748538
KAZN	rs7551069	rs2294886
		rs11579756
		rs7542378
		rs6668635
		rs10927473
		rs16850635
		rs2294888
		rs12758763
		rs761191
		rs761192
		rs7553790
		rs705582
LINGO2	rs2183826	rs2150861
	rs10126008	
	rs4407982	
	rs10121454	
LOC100996630	rs431750	rs2068336
	rs2236390	
	rs17113351	
LRP1B	rs6745610	rs7582294
	rs13398962	rs1474406
		rs7603711
		rs1486963
MAD1L1	rs4721441	rs3800917
		rs10224497
		rs3778969
		rs3800913
		rs3800882
		rs1107592
		rs10239050
		rs3779003
		rs4721441
		rs4721441
		rs10226475
		rs4721295
		rs12666575
		rs3778991
		rs3778994
		rs4719457
		rs3800924

		rs10275045 rs4721190 rs10257990 rs11772205 rs6952727 rs2280550 rs3757440 rs6461233 rs2056480 rs12699477 rs4721184 rs12537914 rs1637759
MUC22	rs3869095 rs12198448 rs12198448	rs1634721
NAALADL2	rs12631086 rs10936794 rs10936797 rs4318562 rs938441 rs1515595 rs9823267 rs6780717	rs1381122 rs1461250
NGEF	rs6719766	rs938569 rs778370 rs1083522 rs2675954 rs748002 rs709937 rs2592114 rs4973569 rs778364 rs1996342 rs778347
NRG2	rs11746363	rs197197
NTM	rs1106362	rs1448363 rs1550976
OPCML	rs10894628	rs12417381 rs505350 rs11223408 rs3016388 rs3016389

		rs476840
		rs3016384
		rs12418625
		rs3018393
		rs9971380
		rs4937758
PCDH15	rs11004025	rs1930147
	rs1911382	rs11004153
	rs11004384	rs1930146
PHACTR1	rs9395495	rs9296494
PKNOX2	rs4537771	rs10790734
	rs4550233	rs671789
	rs1426153	rs689051
	rs11220015	rs11220057
	rs11602925	rs11220064
	rs750338	rs10893378
	rs12273605	rs1044314
	rs10893365	rs10790735
	rs10893366	rs11220058
	rs12284594	rs2321158
PTPRD	rs10977456	rs7848469
PTPRE	rs7081735	rs11018424
	rs7088062	rs12412502
		rs4751554
R3HDM1	rs1446585	rs6745540
RBFOX1	rs4786804	rs4627375
	rs7196961	rs10852677
	rs2178721	
SEPSECS	rs13109061	rs16876882
		rs3796795
SGCZ	rs13281297	rs7010173
SLC25A21	rs10143714	rs712349
	rs10130470	
SOX2-OT	rs13086529	rs12485391
	rs9838604	rs13061117
	rs9882028	rs9833280
	rs7617521	rs6782970
		rs7373337
		rs4855028
		rs13071279
		rs1351235
		rs7611361
		rs13100468

		rs35589824
		rs4855026
		rs13065466
		rs2216427
		rs4854914
		rs13077245
		rs9858313
		rs3860517
		rs13072212
		rs6801189
		rs2718788
		rs3109469
		rs13086738
		rs1479176
		rs4855015
		rs4854912
		rs7619173
		rs13100379
		rs6784620
		rs11716918
		rs1806190
		rs2567665
		rs4456860
		rs4855017
SPATA13	rs12584822	rs9511117
	rs9805786	rs4770620
SPOCK1	rs2348183	rs2974499
	rs6880363	
TCF4	rs4468713	rs4801156
	rs3760600	rs2958172
	rs2924328	rs17512836
	rs1377243	rs1377242
		rs17597926
		rs17594665
		rs17509991
		rs17594721
		rs17089826
		rs11152369
		rs17594526
		rs17511376
		rs587136
		rs9646596
		rs17596267

		rs658977
		rs17514172
		rs12457949
		rs12455205
		rs8098032
		rs17596974
TMEM110-MUSTN1	rs11130329	rs3733047
		rs6445538
		rs6445539
		rs6803519
		rs2276825
WWOX	rs17572451	rs6564580
		rs6564576
ZNF385D	rs6785629	rs1490157
		rs1032314
		rs12491351
		rs7619318

Table 10. Shared Genes and Associated Variants Between Bipolar Disorder and Cannabis Use from Ingenuity Pathway Analysis

Intersection Cannabis-Bipolar Disorder	SNPs in Cannabis use	SNPs in Bipolar Disorder
AGBL1	rs2034633 rs11073678 rs959095	rs981076
CACNA1D	rs3774533 rs1380605	rs3774609 rs3774608 rs3774601 rs3774604 rs719260 rs3774574 rs3774573 rs3774605 rs11720848 rs2289212 rs2253795 rs2680648 rs3774583 rs2077460 rs17053472 rs3796347 rs2612012 rs2612012 rs3774614 rs3774581 rs893363 rs4687586 rs6776947 rs3774570 rs877484 rs1401495 rs2359133 rs14165 rs9311514 rs1401497 rs11705918
CDH13	rs8057717	rs4238724 rs2156 rs1155970
CNTN4	rs2727927	rs1153512

		rs6786554
		rs2018016
		rs9870617
CNTN5	rs10450631	rs12786005
	rs7937128	
CNTNAP2	rs1860681	rs4130001
	rs700305	rs12703863
	rs2906310	rs11773694
	rs2906308	rs12703869
	rs2972126	rs940455
	rs2906300	rs10081247
	rs2373346	rs12703848
		rs11765622
		rs12703852
		rs10500168
		rs13230428
		rs12703865
		rs12703850
		rs10952664
		rs10245776
		rs13233790
		rs12703871
		rs1524340
		rs12703872
		rs1524337
		rs10241470
		rs12703847
		rs10267864
		rs10238991
		rs1916946
		rs1524339
		rs10251563
CSMD1	rs4875772	rs10503283
	rs6996668	rs12682116
		rs7010443
		rs12675806
		rs7011467
		rs12679612
		rs12675866
		rs10111469
		rs11775421
		rs11784939
		rs11782229

		rs10088637
		rs2656298
		rs4875384
		rs1457184
		rs11781735
		rs11778154
		rs17432254
		rs11775382
		rs10081548
		rs11775007
		rs11778566
DLEU1	rs9316497	rs1262774
		rs183950
		rs495838
		rs1753633
		rs3116597
		rs1638703
		rs157164
		rs1262778
		rs1262776
		rs1262775
		rs192492
		rs201762
DLG2	rs7951686	rs17807611
		rs4145049
		rs17807712
		rs1430952
		rs17734964
		rs10501568
DPP6	rs38981	rs7788310
		rs10251606
EPB41L4B	rs4978783	rs12349447
		rs10979751
ERBB4	rs961593	rs12623444
FAM107B	rs1869228	rs7077412
		rs7096784
		rs6602750
		rs7919661
FTO	rs8062891	rs8063241
		rs4389136
		rs12597422
		rs7202620
		rs12445591

HDAC9	rs1469364	rs17140431 rs4141042
KCND3	rs12408551 rs11102355	rs10399721
KIF26B	rs7523252	rs1173829 rs1173838 rs1173837 rs1173835 rs654873 rs1093941
LINGO2	rs2183826 rs10126008 rs4407982 rs10121454	rs1412226 rs1029062 rs1412227 rs6476063
LRP1B	rs6745610 rs13398962	rs10928093
MAD1L1	rs4721441	rs4332037 rs6461009 rs6947019 rs10267593 rs1403175 rs11773627 rs3757440 rs11764590 rs868754 rs11762636 rs6461233 rs3996329 rs3996329 rs10243920 rs11770148 rs7788921 rs4721185 rs12699449 rs733611 rs10230383 rs10227517 rs11772205 rs11762834 rs11764337 rs6944877 rs11764124 rs6952727

		rs3823624
		rs10278591
		rs4721295
		rs10275045
		rs4719336
		rs11772627
		rs6967442
		rs10244946
		rs4721121
		rs12699404
		rs2280548
		rs6951956
		rs2280550
		rs6461049
		rs12537914
		rs4721184
		rs6950627
MUC22	rs3869095	rs9262549
	rs12198448	
	rs12198448	
NGEF	rs6719766	rs778347
OPCML	rs10894628	rs4937708
		rs10894573
		rs10894575
		rs10894574
		rs2212487
		rs4936169
PHACTR1	rs9395495	rs13191496
		rs6914233
		rs6900427
		rs6914467
		rs13194950
		rs13198167
		rs16873893
		rs416852
		rs16873462
		rs426518
		rs7759256
		rs8180628
		rs393032
		rs429029
		rs6930661
		rs13211722

		rs396360
		rs1706941
		rs595864
		rs13215290
		rs521960
		rs546673
		rs4715168
		rs4715154
		rs4715155
		rs1569418
		rs1014820
		rs4715157
PTPRD	rs10977456	rs7865023
		rs7040193
PTPRE	rs7081735	rs4462251
	rs7088062	rs10829323
		rs10764743
PTPRT	rs6103012	rs6102917
		rs6102942
		rs6102944
		rs6102948
		rs6102941
		rs6102946
		rs6102940
		rs6102939
		rs877440
		rs6102945
		rs8123391
		rs2057072
		rs6102943
		rs2057071
		rs6102947
		rs6030378
		rs933240
		rs761010
		rs1157668
		rs761009
		rs6030384
		rs6030341
		rs17841999
		rs6030385
		rs2223541
		rs6016834

		rs6030390
		rs3746537
RBFOX1	rs4786804	rs17138946
	rs7196961	rs17144513
	rs2178721	
RPS6KA2	rs16900973	rs763193
	rs845641	rs960145
	rs845671	rs3823198
	rs845674	rs3778386
	rs1099646	rs3799600
	rs2242573	rs3799598
		rs3799603
		rs3799597
		rs9457187
		rs3778385
		rs6928849
RSU1	rs7092024	rs3740170
	rs7893556	rs17156952
RYR2	rs268786	rs1833419
	rs12063070	rs10495392
		rs1421207
SELP	rs3917739	rs3917843
SGCZ	rs13281297	rs7841764
SLC25A21	rs10143714	rs1048200
	rs10130470	rs1884213
		rs4605059
SOX2-OT	rs13086529	rs9823623
	rs9838604	rs13081234
	rs9882028	rs9842957
	rs7617521	rs7426901
		rs9818320
		rs9290723
		rs4282105
		rs12487748
		rs13086529
		rs12497248
SOX5	rs725124	rs17399946
		rs10842241
		rs7136659
		rs7136898
		rs10842323
		rs16915574
		rs12811046

		rs11047161
		rs4963720
		rs12296192
STX8	rs12452147	rs9903924
		rs11078799
		rs9900532
		rs11078800
		rs17206891
		rs9902096
SYNE1	rs1000864	rs9371601
		rs6557230
		rs6557229
		rs4318888
		rs9383995
		rs214976
		rs214952
		rs551900
		rs551900
		rs4523096
		rs1203233
		rs20585
		rs214944
		rs214945
		rs70018
		rs70015
		rs70017
		rs2623968
		rs551681
		rs169974
		rs214941
		rs214942
		rs521514
		rs492233
		rs2623970
		rs525210
		rs214969
		rs549981
		rs214962
		rs214961
		rs214963
		rs7759578
		rs506181
		rs553642

rs7763880
rs127196
rs7756410
rs579464
rs488673
rs177330
rs2449116
rs490448
rs215001
rs1949962
rs214943
rs522437
rs214954
rs9397512
rs548400
rs544098
rs2746418
rs7774755
rs6557226
rs544125
rs2348798
rs550685
rs9371603
rs1830219
rs1527369
rs2623958
rs214997
rs214993
rs1738438
rs527021
rs502268
rs214994
rs214987
rs4509131
rs2623932
rs214981
rs214980
rs214978
rs544863
rs2623981
rs548978
rs7751588
rs9397510

		rs9478332
		rs7747960
		rs9479313
		rs7740022
		rs9397105
		rs2295041
		rs12055686
		rs214992
		rs7744498
		rs215006
		rs17082664
		rs6557219
		rs169976
		rs9479314
		rs6909684
		rs6913500
		rs9478329
		rs6935362
		rs169977
		rs9479307
		rs554608
		rs214968
		rs7738528
		rs742784
		rs9397514
		rs12661143
		rs9479316
		rs1949963
		rs761408
		rs17545391
		rs1489586
		rs214972
		rs214972
		rs2623966
		rs2695263
		rs12526147
		rs1534298
		rs2623972
		rs2623971
		rs11755737
		rs9371604
TMEM110-MUSTN1	rs11130329	rs6445539
		rs3733047

		rs9682464
		rs6445538
		rs6795646
		rs6803519
		rs2276825
		rs9860296
		rs9810807
		rs9880978
		rs2336668
		rs3821873
		rs4687682
		rs9850563
		rs4687680
		rs12492391
		rs2336664
		rs9836499
		rs6769789
		rs13089851
		rs9844736
		rs6445547
		rs4302374
		rs9876403
TMEM178B	rs6975836	rs1860752
		rs6967482
		rs722219
		rs11762357
		rs4726268
		rs4726259
		rs216996
TTN	rs16866488	rs13398235
		rs10176708
		rs13417645
		rs2366753
		rs10179811
		rs10183237
		rs10183361
		rs4893852
		rs11888217
		rs17355446
		rs16866473
		rs1905520
		rs1905520
		rs4894045

		rs4893853
		rs7572955
		rs4894048
		rs16866519
		rs11897386
		rs1484120
		rs4894041
		rs4894040
		rs10497522
WWOX	rs17572451	rs11640201
		rs2161636
		rs9319534
		rs9319535

Figure 7. Enrichment Analysis of Schizophrenia-Cannabis Comparison (with Benjamini-Hochberg P-value Adjustment)

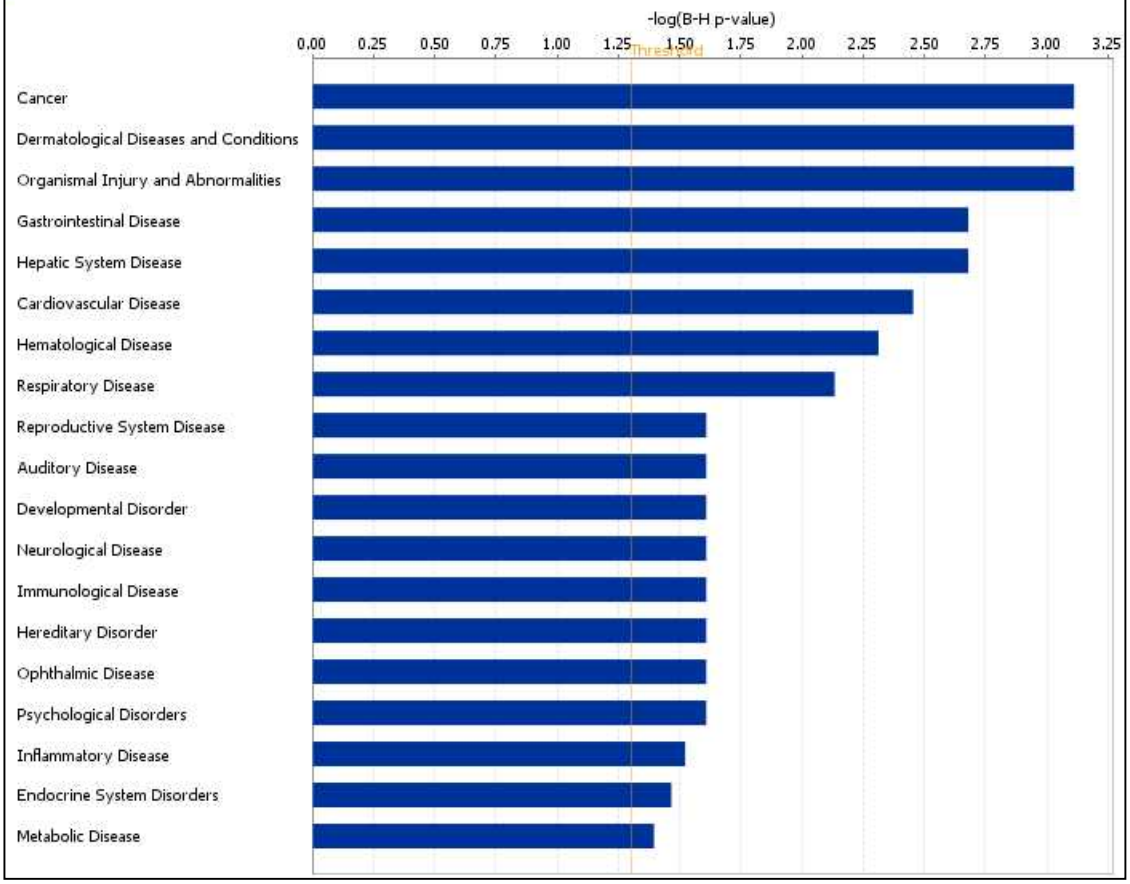


Figure 8. Cellular and Physiological Pathway Enrichment Analysis of Schizophrenia-Cannabis Comparison (with Benjamini Hochberg P-value Adjustment)

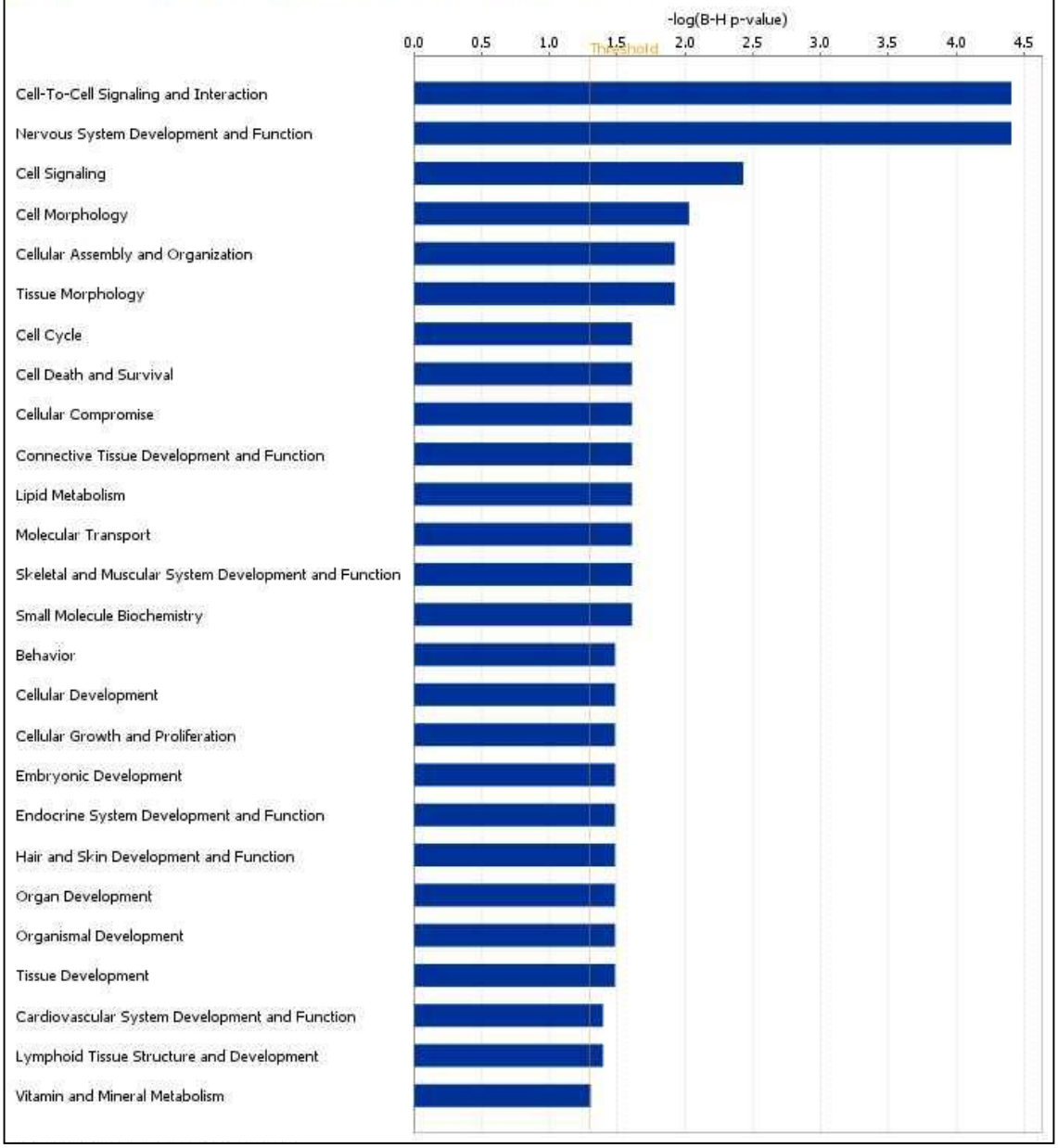


Figure 9. Disease Enrichment Analysis of Bipolar Disorder-Cannabis Comparison (with Benjamini Hochberg P-value Adjustment)

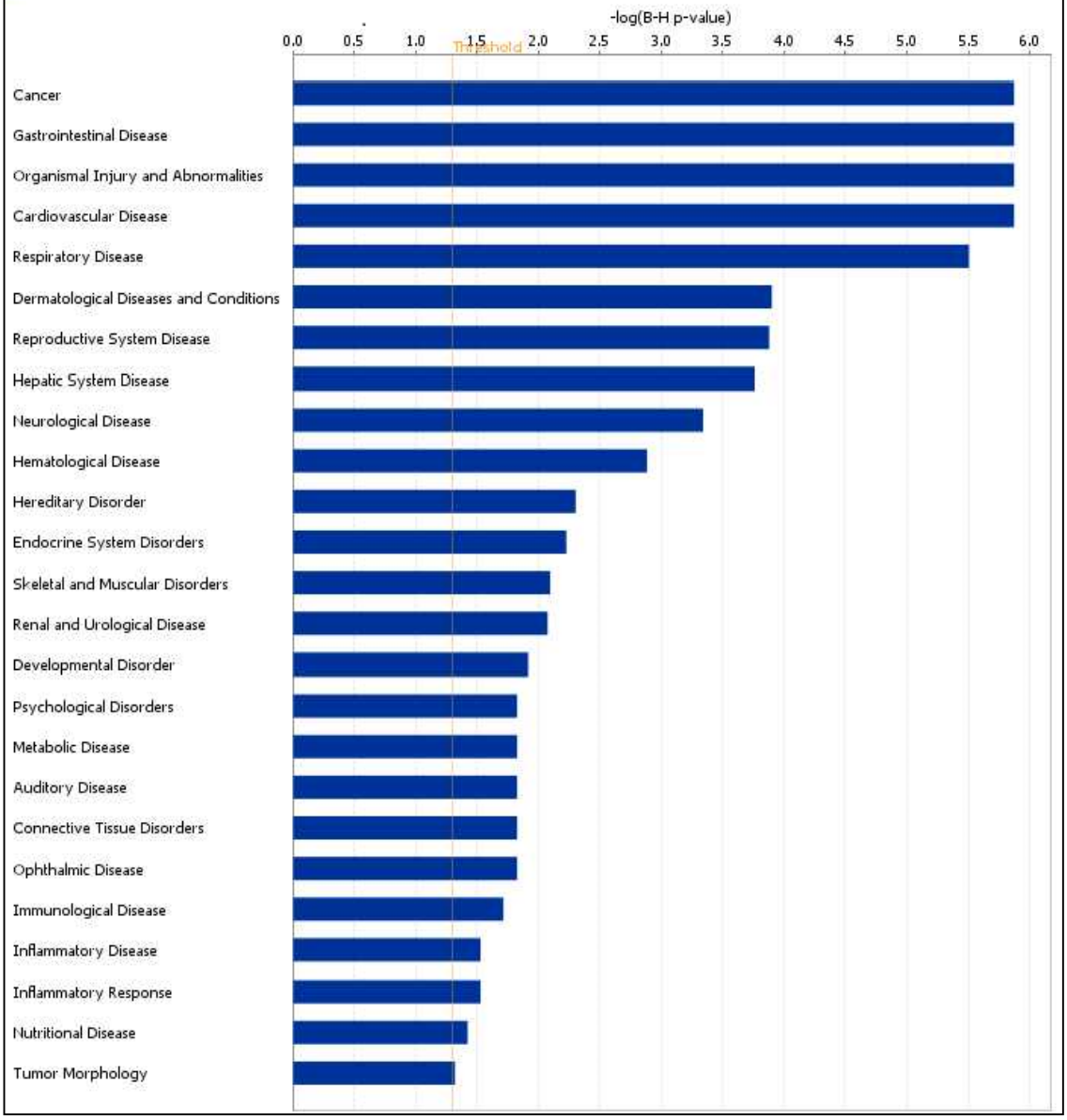


Figure 10. Cellular and Physiological Pathway Enrichment Analysis of Bipolar Disorder-Cannabis Comparison (with Benjamini Hochberg P-value Adjustment)

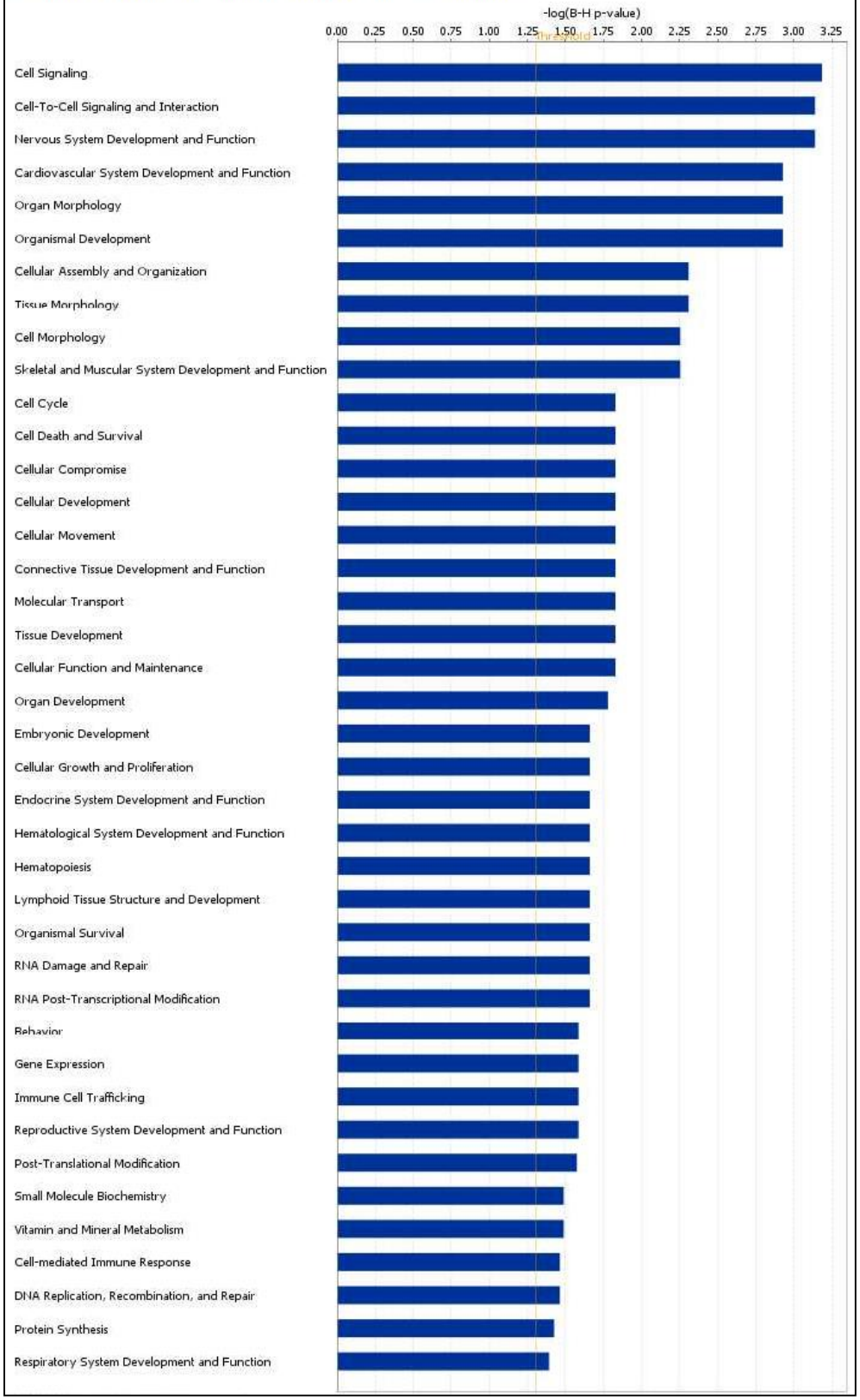
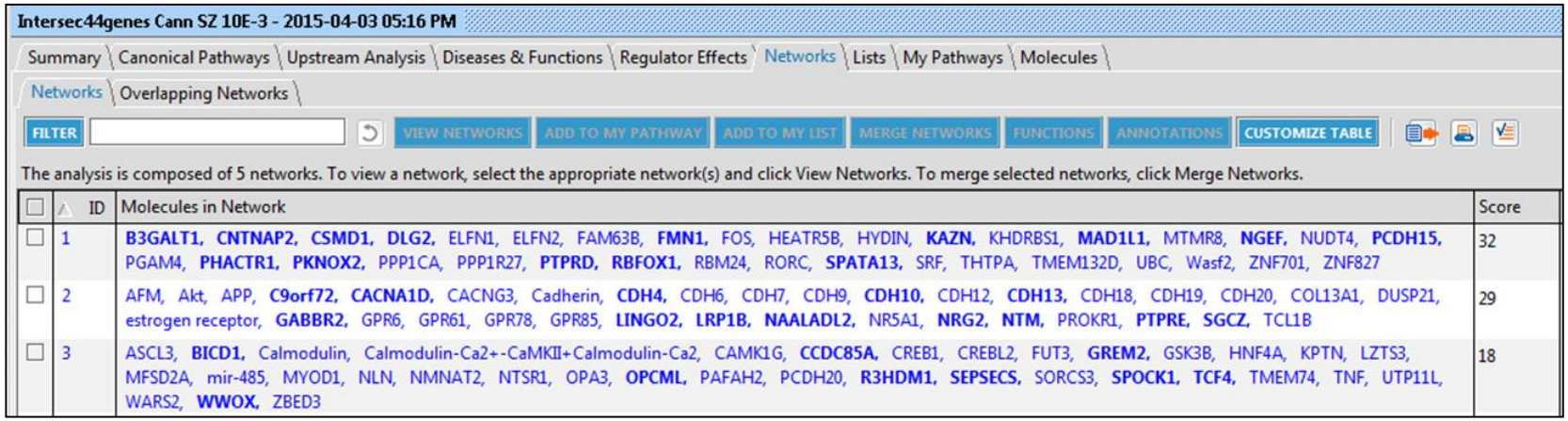


Figure 11. Most Significant Predicted Molecular Pathway Networks for Schizophrenia-Cannabis Comparison



Intersec44genes Cann SZ 10E-3 - 2015-04-03 05:16 PM

Summary | Canonical Pathways | Upstream Analysis | Diseases & Functions | Regulator Effects | Networks | Lists | My Pathways | Molecules

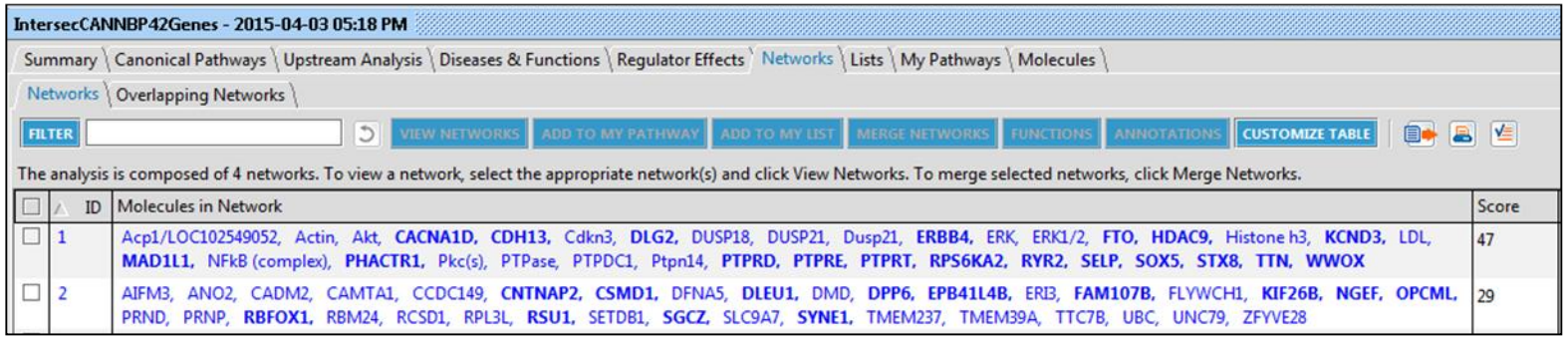
Networks | Overlapping Networks

FILTER | VIEW NETWORKS | ADD TO MY PATHWAY | ADD TO MY LIST | MERGE NETWORKS | FUNCTIONS | ANNOTATIONS | CUSTOMIZE TABLE

The analysis is composed of 5 networks. To view a network, select the appropriate network(s) and click View Networks. To merge selected networks, click Merge Networks.

ID	Molecules in Network	Score
1	**B3GALT1, CNTNAP2, CSMD1, DLG2,** ELFN1, ELFN2, FAM63B, **FMN1,** FOS, HEATR5B, HYDIN, **KAZN,** KHDRBS1, **MAD1L1,** MTMR8, **NGEF,** NUDT4, **PCDH15,** PGAM4, **PHACTR1, PKNOX2,** PPP1CA, PPP1R27, **PTPRD, RBFOX1,** RBM24, RORC, **SPATA13,** SRF, THTPA, TMEM132D, UBC, Wasf2, ZNF701, ZNF827	32
2	AFM, Akt, APP, **C9orf72, CACNA1D,** CACNG3, Cadherin, **CDH4,** CDH6, CDH7, CDH9, **CDH10,** CDH12, **CDH13,** CDH18, CDH19, CDH20, COL13A1, DUSP21, estrogen receptor, **GABBR2,** GPR6, GPR61, GPR78, GPR85, **LINGO2, LRP1B, NAALADL2,** NR5A1, **NRG2, NTM,** PROKR1, **PTPRE, SGCZ,** TCL1B	29
3	ASCL3, **BICD1,** Calmodulin, Calmodulin-Ca2+-CaMKII+Calmodulin-Ca2, CAMK1G, **CCDC85A,** CREB1, CREBL2, FUT3, **GREM2,** GSK3B, HNF4A, KPTN, LZTS3, MFSD2A, mir-485, MYOD1, NLN, NMNAT2, NTSR1, OPA3, **OPCML,** PAFAH2, PCDH20, **R3HDM1, SEPSECS,** SORCS3, **SPOCK1, TCF4,** TMEM74, TNF, UTP11L, WARS2, **WWOX,** ZBED3	18

Figure 12. Most Significant Predicted Molecular Pathway Networks for Bipolar Disorder-Cannabis Comparison



IntersecCANNBP42Genes - 2015-04-03 05:18 PM

Summary | Canonical Pathways | Upstream Analysis | Diseases & Functions | Regulator Effects | Networks | Lists | My Pathways | Molecules

Networks | Overlapping Networks

FILTER | VIEW NETWORKS | ADD TO MY PATHWAY | ADD TO MY LIST | MERGE NETWORKS | FUNCTIONS | ANNOTATIONS | CUSTOMIZE TABLE

The analysis is composed of 4 networks. To view a network, select the appropriate network(s) and click View Networks. To merge selected networks, click Merge Networks.

ID	Molecules in Network	Score
1	Acp1/LOC102549052, Actin, Akt, **CACNA1D, CDH13,** Cdkn3, **DLG2,** DUSP18, DUSP21, Dusp21, **ERBB4,** ERK, ERK1/2, **FTO, HDAC9,** Histone h3, **KCND3,** LDL, **MAD1L1,** NFkB (complex), **PHACTR1,** Pkc(s), PTPase, PTPDC1, Ptpn14, **PTPRD, PTPRE, PTPRT, RPS6KA2, RYR2, SELP, SOX5, STX8, TTN, WWOX**	47
2	AIFM3, ANO2, CADM2, CAMTA1, CCDC149, **CNTNAP2, CSMD1,** DFNA5, **DLEU1,** DMD, **DPP6, EPB41L4B,** ERI3, **FAM107B,** FLYWCH1, **KIF26B, NGEF, OPCML,** PRND, PRNP, **RBFOX1,** RBM24, RCSD1, RPL3L, **RSU1,** SETDB1, **SGCZ,** SLC9A7, **SYNE1,** TMEM237, TMEM39A, TTC7B, UBC, UNC79, ZFYVE28	29

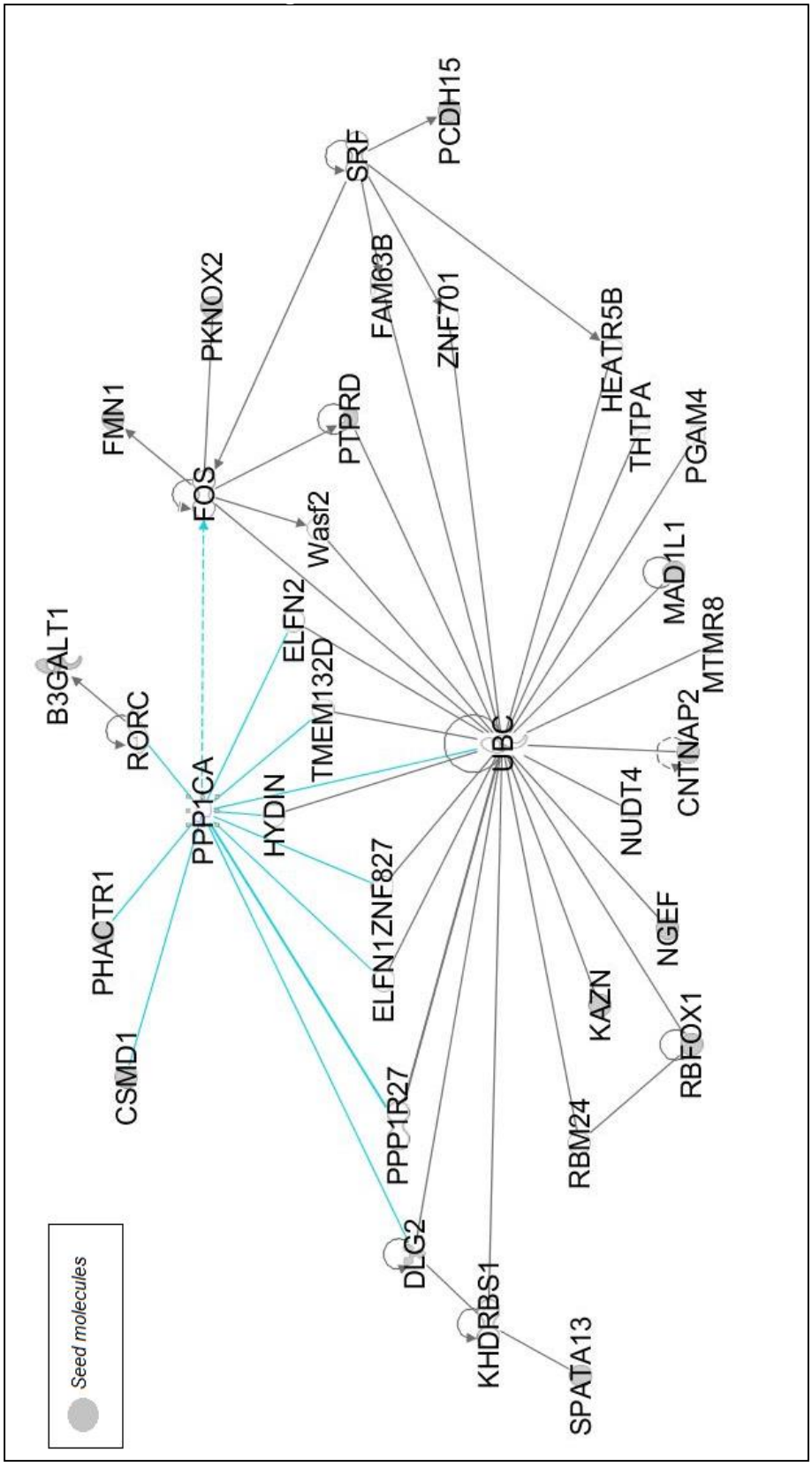


Figure 13. Predicted Molecular Network for Schizophrenia-Cannabis Comparison (ID=1, Score= 32)

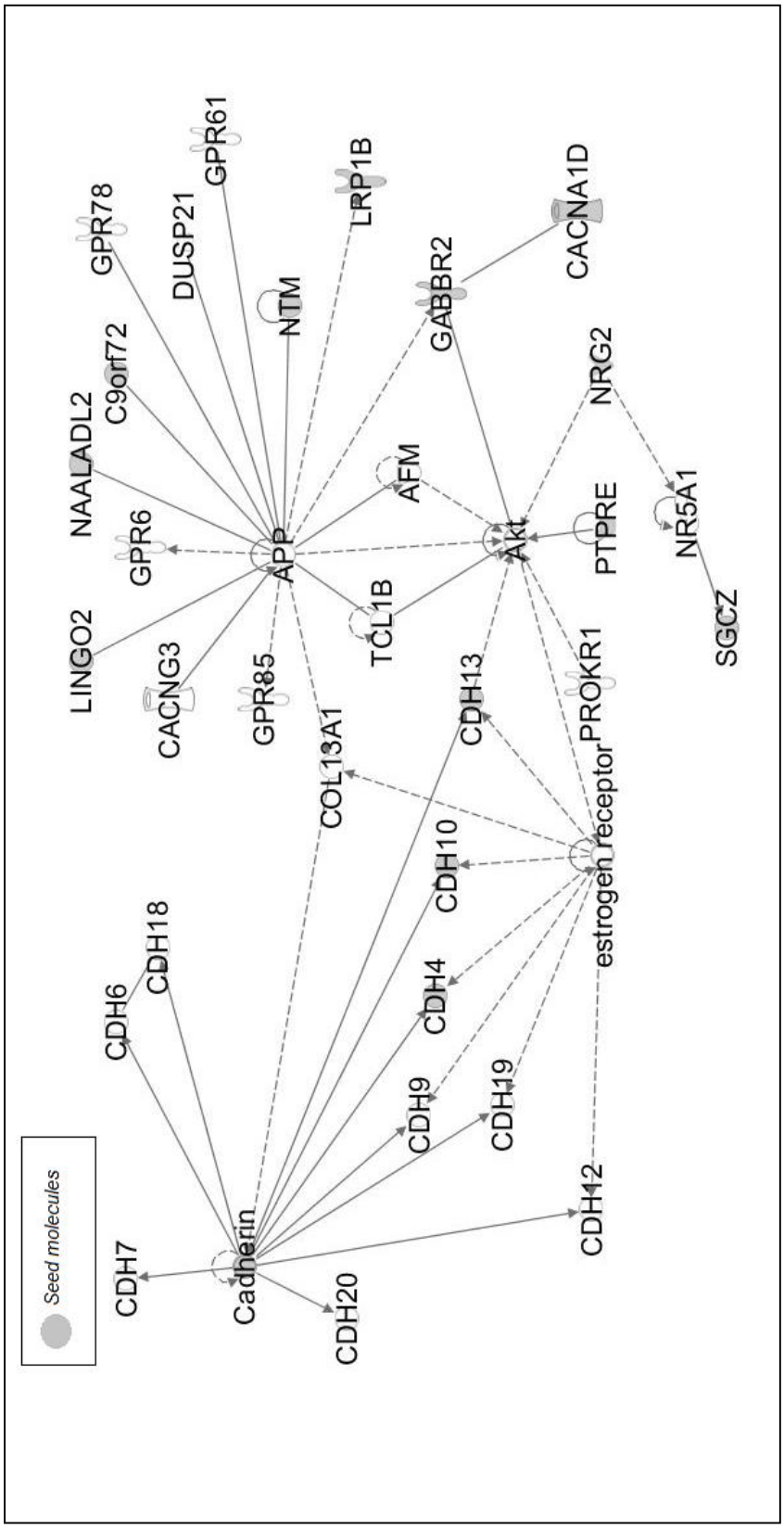


Figure 14. Predicted Molecular Network for Schizophrenia-Cannabis Comparison (ID=2, Score= 29)

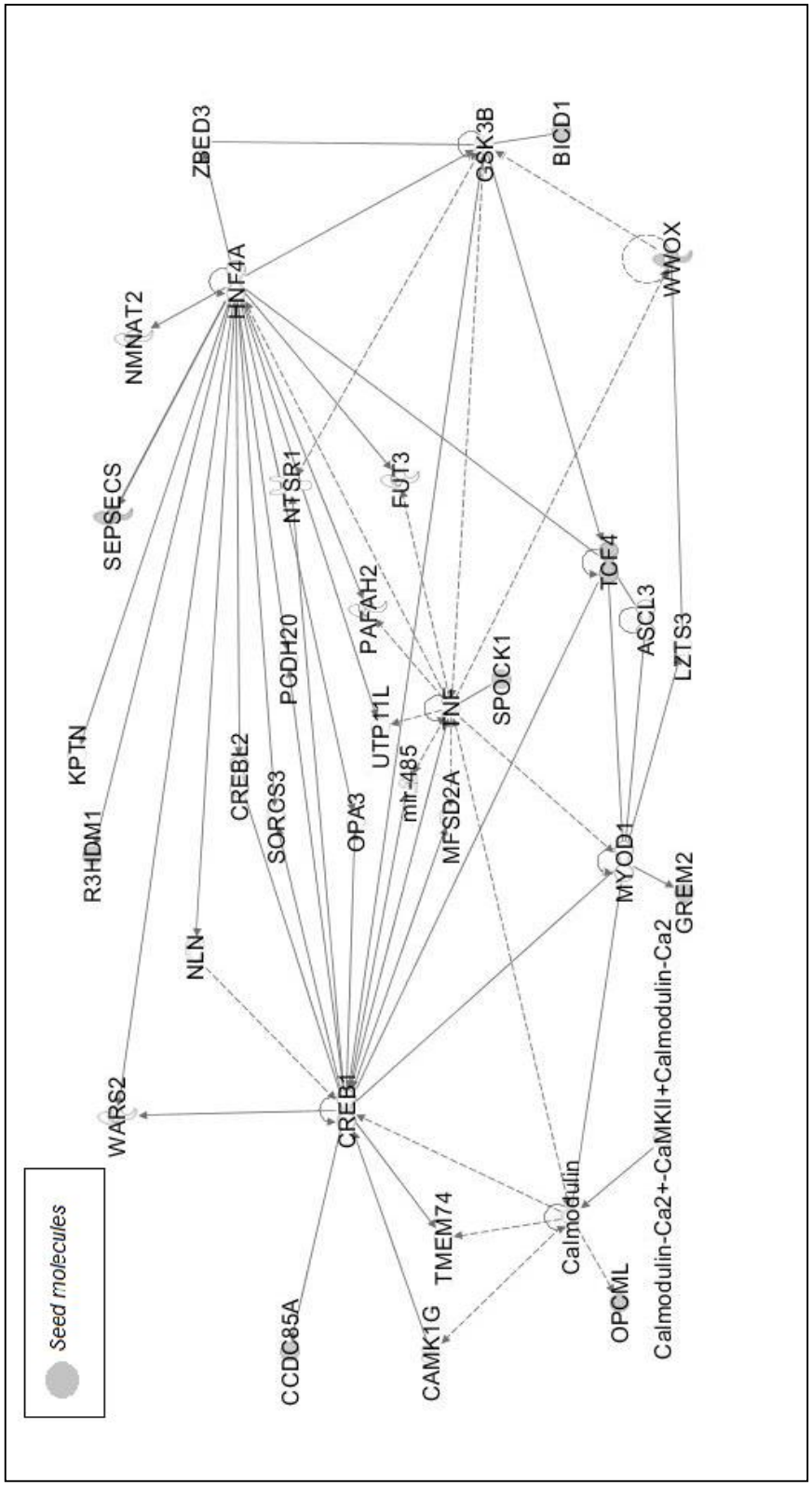


Figure 15. Predicted Molecular Network for Schizophrenia-Cannabis Comparison (ID=3, Score= 15)

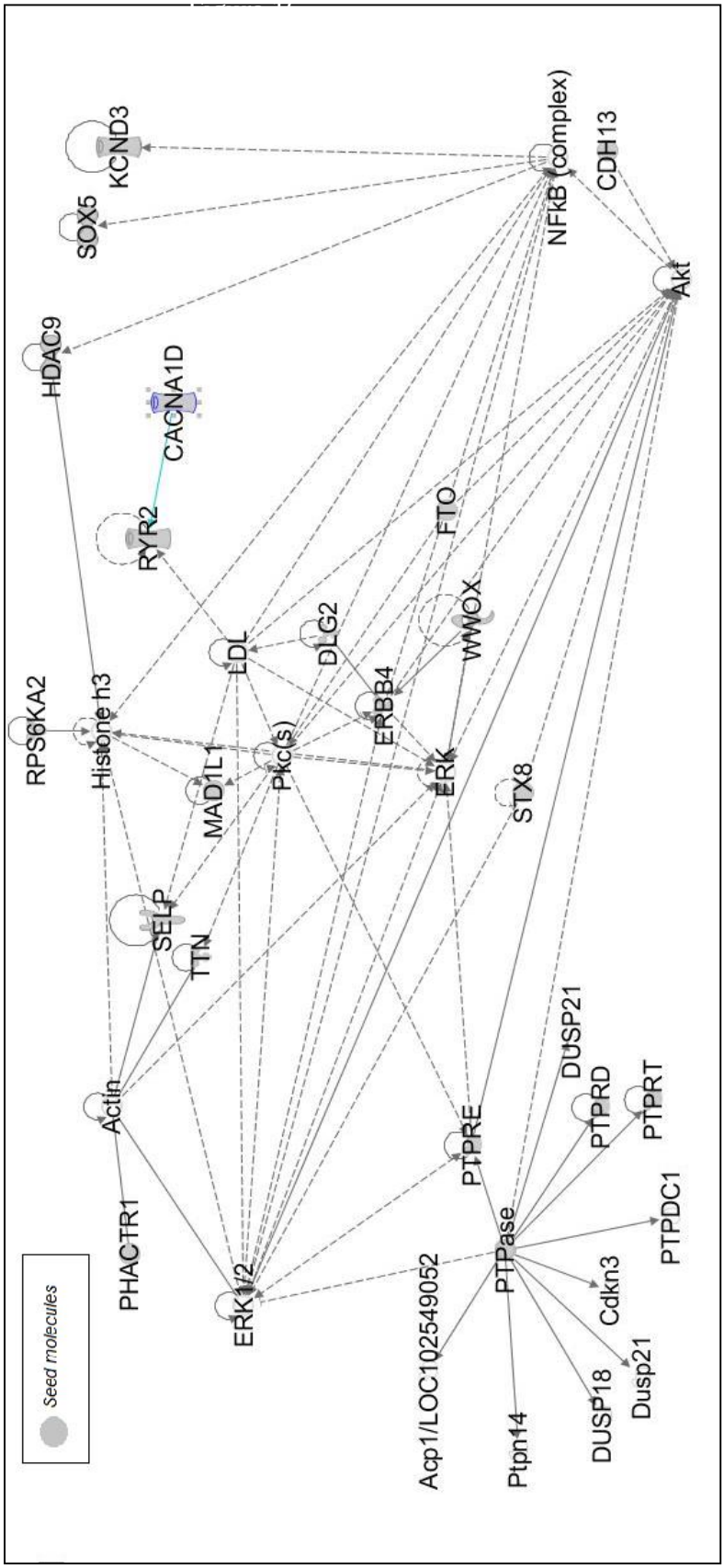


Figure 16. Predicted Molecular Network for Bipolar Disorder-Cannabis Comparison (ID=1, Score= 47)

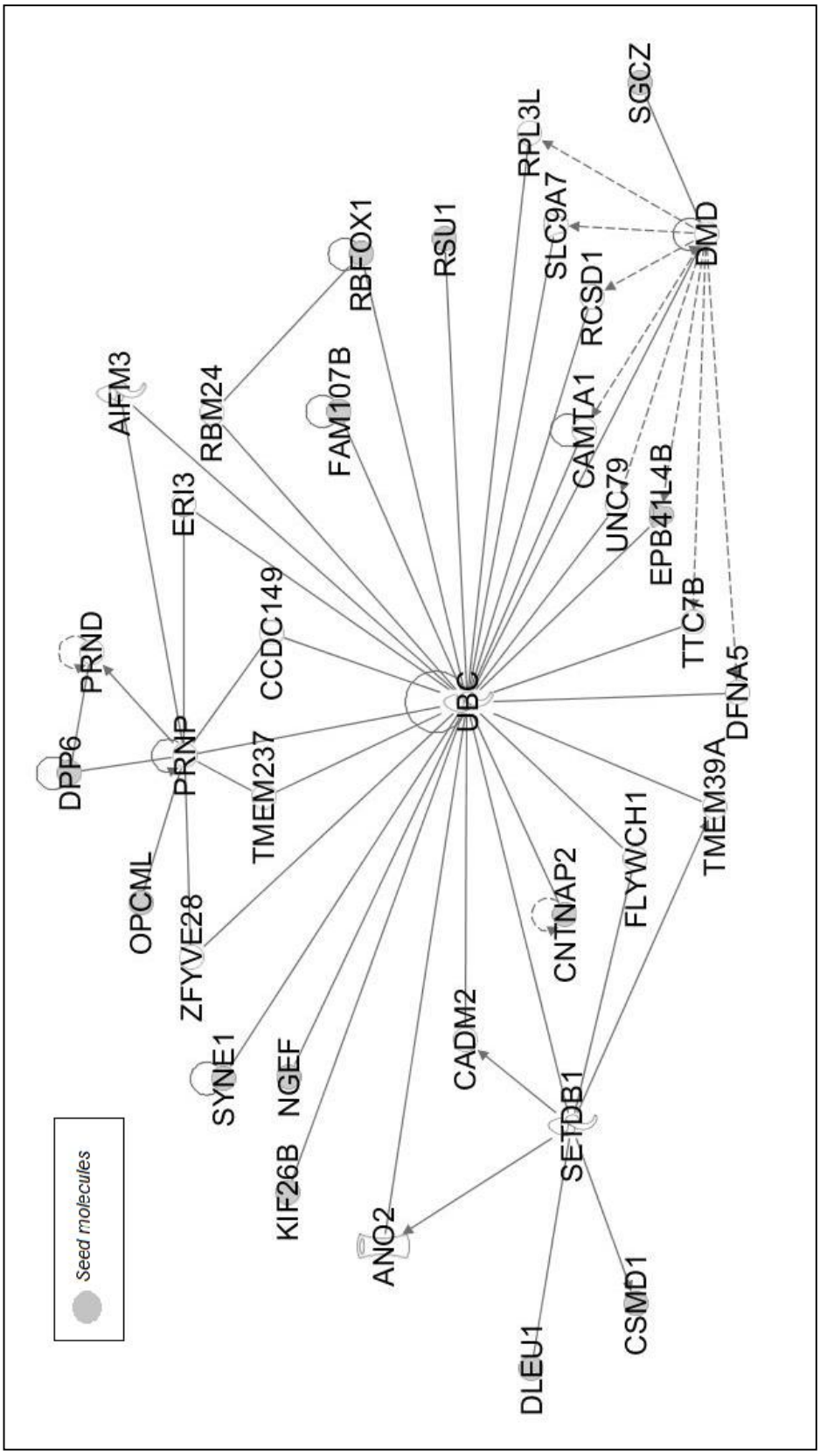


Figure 17. Predicted Molecular Network for Bipolar Disorder-Cannabis Comparison (ID=2, Score= 29)

Results

Overall, there were 2,380 (50.9%), 3,547 (51.5%) and 667 (53.7%) unique variants for schizophrenia, bipolar disorder and cannabis use for which IPA derived genes in mapping networks (see Table 7). Roughly half of the variants for each variable of interest were unmapped, as IPA could not match certain variants to appropriate genes (schizophrenia: n= 2,300 (50%); bipolar disorder: n= 3,302 (48%); cannabis use: n= 570 (45.9%)). Of the 3,582 and 673 variants that were mapped by IPA initially for bipolar disorder, and cannabis use, 35 and 6 genes had duplicate *rs* identifiers. After eliminating duplicates, there were 3,547 (51.5%) unique variants for bipolar disorder and 667 (53.7%) unique variants for and cannabis use.

All three variables had between 384 to 689 genes that were considered "ready for analysis" and could be mapped in molecular pathways (see Table 8); 689, 639 and 342 unique genes were matched to variants for schizophrenia, bipolar disorder and cannabis use, respectively (see Table 8).

To identify potential sources of association between each outcome variable and cannabis use, IPA was further utilized to identify shared genes between each outcome and cannabis use. IPA identified 44 shared genes in the schizophrenia-cannabis use cross-comparison and 42 shared genes in the schizophrenia-cannabis use cross-comparison (see Tables 8-10 and Figures 5-6).

An enrichment analysis, an IPA illustration of most common diseases and cellular pathways related to the genes mapped after Benjamini Hochberg (FDR) p-value adjustment, was also conducted. IPA synthesizes results from relevant peer-review journals to establish association between certain genes and health outcomes and/or

physiological processes. Association is established based on a statistical p-value IPA assigns to each comparison, which is adjusted using the FDR method. This method, usually less conservative than the Bonferroni adjustment, is utilized to correct for inflation of false positive rate that occurs with multiple statistical comparisons. Adjustment occurs with comparison of the corrected p-value relative to the expected p-value (Gietzen, 2010). In this method, p-values are sorted and ranked. The smallest value gets ranked first, the next smallest is ranked second, and the largest receives rank N. Subsequently, each p-value is multiplied by N and divided by its assigned rank to yield adjusted p-values (DNAStar). Outcomes and processes with logarithmic (log) adjusted p-values greater than 1.30 (equivalent to log of α=0.05) are displayed in a bar graph (Gietzen, 2010).

Figure 7. shows results of the disease enrichment analysis for the schizophrenia-cannabis use comparison. Cancer is the most prevalent outcome associated with the analyzed 44 common genes, followed by dermatological diseases, organismal injuries and abnormalities, gastrointestinal disease, and hepatic system disease for the top five processes. Psychological disorders rank 16th in the list of 19 diseases, with an adjusted p-value of roughly Log-1.67 (see Figure 7).

Figure 8 displays the significant cellular and physiological processes associated with the analyzed shared genes. Cell-to-cell signaling, nervous system development and function, cell signaling, cell morphology, and cellular assembly and organization represent the five most prevalent physiological processes. Of specific interest to the hypothesis guiding this thesis are nervous system development and function and behavior which rank 2nd and 15th among the 27 processes shown (see Figure 8).

Enrichment analyses were also conducted for the bipolar disorder-cannabis comparison. Cancer, gastrointestinal disease, organismal injury and abnormalities, cardiovascular disease and respiratory disease represent the most significant illnesses derived from analysis of 42 common genes (see Figure 9). Psychological disorders are among the bottom half of derived significant outcomes (rank = 16^{th}).

Cell signaling, cell-to-cell signaling, nervous system development and function, cardiovascular system development and function, and organ morphology characterize the five most significant physiological processes for the bipolar disorder-cannabis comparison (see Figure 10). Nervous system development and function and behavior rank 3^{rd} and 30^{th} of the list of 40 cellular and physiological functions.

Molecular Pathway Modeling

To further investigate the nature of the molecular pathways of genes shared by schizophrenia and cannabis use and by bipolar disorder and cannabis use, IPA elicits a predictive molecular network feature. This tool identifies *N* genes from the list of common genes for each comparison, called "seeds." Networks of those seeds with external genes show to have molecular and cellular intersections identified in the literature of genome-wide association studies, are developed. Such pathways are important features to consider in discussion of effect of external upstream and downstream molecular pathways on common genes. Identified pathways are given a score, which denotes comparative potential impact, and scores less than 30 are usually considered trivial and scores higher than this warrant consideration. The score is a weighted Kolmogorov-Smirnov-like running sum statistic which describes the

overrepresentation of the N_H genes at the top of the entire ranked list of genes (Ngwa et al., 2011).

Three significant predictive pathways, with scores of 32, 29 and 18 were discovered in analysis of the common genes between schizophrenia and cannabis use (see Figure 11). The pathway with the highest score is shown in Figure 13. Catalytic Protein Phosphatase 1 (PPP1CA) and Ubiquitin C (UBC) denote two predicted molecules which high frequency of connections to seed molecules (see Figure 13). The pathway with intermediate significance (score = 29), shown in Figure 14, identified Cadherin 1 (Cadherin), Amyloid Beta (A4) Precursor Protein (APP), and V-Akt Murine Thymona Viral Oncogene (Akt) as molecules with greatest frequency of connections to seed molecules. CAMP Responsive Element Binding Protein 1 (CREB1) and Hepatocyte Nuclear Factor 4, Alpha (HNF4A) (shown in Figure 15) could be of particular interest in discussing Pathway 3 (from Figure 11).

Two significant predictive pathways, with scores of 47 and 29 were derived from the common genes in the bipolar disorder and cannabis connection (see Figure 12). Of particular interest if the network of Calcium Channel L Type, Alpha-1 (CACNA1D) which has been implicated in bipolar disorder previously in the literature (Network 1 in Figure 12; see Figure 16) (Pinggera et al., 2015). Two relevant molecules, UBC and SET Domain Bifurcated (SETDB1) are of particular interest in the bipolar disorder-cannabis association (see Pathway no. 2 from Figure 12 and Figure 17). SETDB1 histone methyltransferase has been suggested to regular mood-related behaviors and expression of NMDA receptor subunit NR2B (Jiang et al., 2010).

Discussion

Shared Genes

There is considerable overlap between genes implicated in schizophrenia and cannabis use and those implicated in bipolar disorder and cannabis use (Table 8). Figures 5-6 show more than 40 genes shared by both schizophrenia and bipolar disorder with cannabis use. 44 genes are shared by schizophrenia and cannabis use (see Table 8 and Figure 5) and 42 genes are shared by bipolar disorder and cannabis use (see Table 8 and Figure 6). Shared genes and associated variants are outlined in Tables 9-10. Such extensive overlap between schizophrenia and cannabis use is consistent with findings of Power et al. (2014). Similar to the conclusions drawn from this thesis, findings regarding direction of association were inconclusive (Power et al. (2014)). Power et al. (2014) speculates that a bidirectional association could be plausible, as the genes that are associated with cannabis use could predispose an individual to use cannabis, which may increase risk for developing psychosis or bipolar disorder. Consequently, it is plausible that individuals may use cannabis following diagnosis of either schizophrenia or bipolar disorder to relieve psychotic symptoms or prevent mood fluctuations. If these shared genes are integral components of reciprocal feedback loops between cannabis use and schizophrenia and cannabis use and bipolar disorder, it would be interesting to investigate whether these genes confer greater risk for addiction to multiple substances.

A considerable body of research points to the use of cannabis among schizophrenia patients to self-medicate and alleviate psychotic symptoms (Stone et al., 2014). Fergusson et al. (2005) predict that individuals with an increased genetic predisposition to schizophrenia are both more likely to use cannabis. Strakowski and

DelBello (2000) explains four potential hypotheses for common co-occurrence of bipolar disorder and cannabis use: (a) cannabis dependence occurs as a symptom of bipolar disorder; (b) substance abuse is an attempt by bipolar patients to self-medicate symptoms; (c) substance abuse causes bipolar disorder and (d) cannabis dependence and bipolar disorder share a common risk factor. In context of this thesis, however, it should be noted that variants for cannabis use were much fewer than variants for schizophrenia and bipolar disorder (see Table 7), and the population sample derived for cannabis use was also considerably smaller than the populations for the outcome samples (see Tables 1-3). This suggests that the genes shared by schizophrenia and cannabis use and those shared among bipolar disorder and cannabis use may have greater impact on characteristics related to cannabis use than on the disorders themselves. Table 1 describes the average age of onset of schizophrenia as ranging from 14 to 34, which largely overlaps with the range from mean age of imitation of cannabis use to mean age of recency of cannabis use presented in Table 3. Concurring time periods of onset of illness and exposure confound temporality of the association, and allude to the possibility that cannabis use and schizophrenia share common genetic risk factor(s).

Enrichment Analyses

Enrichment analyses for both outcomes with respect to shared genes displayed diseases and cellular processes seemingly unrelated to the outcomes of interest (Figures 7-10). Cancer is the most significant health outcome in both comparisons, suggesting the genes implicated in cannabis use, schizophrenia, and bipolar disorder may be oncogenic. A possible explanation for this phenomenon is that cannabis use often co-occurs with tobacco use (Agrawal, Budney, & Lynskey, 2012). The high commonality of dual

tobacco and marijuana use among adolescents and young adults (Ramo, Liu, & Prochaska, 2012) suggests there are mutual "addiction" genes associated with both forms of substance abuse. Greater significance of cancer and cardiovascular disease over psychological outcomes may be attributed to smoking as a potent risk factor for numerous cancers and cardiovascular disease. These enrichment analysis findings foreshadow the findings in the cannabis-bipolar disorder comparison.

Similar findings are replicated in enrichment of cellular processes for the bipolar disorder-cannabis cross-comparison with respect to cardiovascular function (Figure 10). Several of the top processes found for the schizophrenia-cannabis cross-comparison are involved in normal cell function, but nervous system development and function may be of specific interest to this thesis. Traumatic brain injury during critical periods of development can also delay or stunt brain development. As aforementioned, occurrence of traumatic brain injury during childhood or adolescence may be associated with schizophrenia in multivariate pedigree analyses (Malaspina et al., 2014). Gender differences for prevalence of traumatic brain injury (more common among men than women) is also consistent with the pattern of gender differences among schizophrenia patients (Malaspina et al. (2014); Ochoa, Usall, Cobo, Labad, and Kulkarni (2012)). Accordingly, traumatic brain injury may represent a risk factor for schizophrenia, but may also predispose individuals to cannabis use as means of self-medication. Moreover, traumatic brain injury may represent a driving environmental factor for cannabis use and may explain why the vast majority of cannabis users do not develop psychosis.

Predicted Molecular Networks

Results from predictive molecular networks for the schizophrenia-cannabis connection acknowledge two prominent predictive molecules and associated networks, one involving PP1CA and another involving UBC as hubs (see Figure 13). There is evidence to suggest PPP1CA is significantly correlated with DISC1-biomarker isoforms in CD34+ cells and immature dendritic cells (Young et al., 2009). DISC1 is a well-validated schizophrenia risk gene (Xie et al., 2014), as DISC1 splice variants are upregulated in schizophrenia and present in higher proportions in the hippocampus of schizophrenia patients (Nakata et al., 2009). The network also shows association between Discs Large Homolog 2 (DLG2), and PPP1CA. DLG2 represents a seed molecule and is known to induce hypofunction of NMDA receptor signaling (Carlisle, Fink, Grant, & O'Dell, 2008), a process implicated in schizophrenia (Javitt, 2007). Expression of c-Fos, a nuclear phosphoprotein oncogene and another molecule associated with PPP1CA (see Figure 13), has been found to be altered in rats after acute treatment with cannabis (Wegener & Koch, 2009). Relevant studies conclude that chronic stimulation of the cannabinoid receptor CB_1 during the rats' puberty not only leads to persistent behavioral changes but also long-term adaptations in c-Fos immunoreactivity within brain regions critical for neuropsychiatric disease and drugs of abuse (Wegener & Koch, 2009). It is possible that PPP1CA acts as an addiction gene intermediate between DLG2 and FOS, or that dysfunction of PPP1CA and DISC1 are common genetic abnormalities among both cannabis use and schizophrenia.

Expression of UBC, another significant molecule in the pathway, in peripheral blood is correlated to positive symptoms of psychosis (Bousman et al., 2010) is present in

the pathway. UBC is involved in protein modification along with cellular proliferation and regulation of DNA repair (Quinn et al., 2007). Previous studies have noted altered regulation of ubiquitin-conjugating enzyme 24 hours after Δ^9-THC administration to rats (Kittler et al. (2000);Grigorenko et al. (2002)). UBC expression may induce disruption of cell proliferation and repair resulting from drug exposure during the adolescent period. Genes for ubiquitin-conjugating enzymes are also decreased in human schizophrenic hippocampal samples (Altar et al., 2005). Cannabis use has been associated with a deteriorated course of schizophrenic illness (Arseneault et al., 2002). Consequently, it is uncertain whether cannabis use predicts such UBC expression or whether patients choose to self-medicate with cannabis to relieve symptoms arising from altered UBC expression.

Another significant predictive pathway (Figure 14) displays cadherin as a significant hub. Common cannibnoid receptor CB1 agonists mimic the N-cadherin response in axonal growth in vitro at a step downstream of fibroblast growth factor receptor 2 (FGFR2) activation (Williams, Walsh, & Doherty, 2003). Endocannabinoids released from depolarized post-synaptic neuron bind to CB_1 receptors in the pre-synaptic neuron and cause a reduction in GABA release. Dopaminergic neurons can normally be inhibited by GABA expression, but cannabis removes this inhibition by the GABA neurons and activates the dopamine reward circuit. In chronic consumers of cannabis, the loss of CB1 receptors in the brain's arteries reduces the flow of blood to the brain, resulting in memory loss and attention deficits. CACNA1D, a voltage-gated calcium channel, is a common susceptibility gene for bipolar disorder, schizophrenia, and major depressive disorder (Cross-Disorder Group of the Psychiatric Genomics Consortium). NRG2, a common gene implicated in schizophrenia, a gene related to expression of

NRG1, a common gene implicated in cannabis use-schizophrenia interaction, is shown as a seed molecular in the pathway. It has shown to influence Akt, which downstream regulates GABBR2, eventually regulating CACNA1D. GABBR2 is associated with nicotine dependence in European and African-American populations (Li et al., 2009), which are represented heavily in the population samples from the three datasets.

An important molecule to note in the third pathway (Figure 15) is cyclic AMP-responsive element-binding protein (CREB), involved in intracellular signal transduction pathways used by most dopamine receptor subtypes (Kawanishi, Harada, Tachikawa, Okubo, & Shiraishi, 1999). Mice studies indicate phosphorylation of CREB may underlie impaired long-term synaptic plasticity induced by repeated *in vivo* exposure to Δ9-THC (Fan, Yang, Zhang, & Chen, 2010). This pathway suggests cannabis use as a biologically predictive factor preceding onset of schizophrenia. ATK1 commonly leads to phosphorylation of CREB (Chrivia et al., 1993), indicating that future directions in research may target efforts towards characterizing nature of molecules that phosphorylate CREB.

CACNA1D is prominent in characterizing risk for bipolar disorder and therefore is an important molecule to consider in pathway 1 (see Figures 12 and 16) in the bipolar disorder-cannabis comparison. DLG2, is shown regulating expression of LDL, a cholesterol receptor, which affects expression of RYR2, a cardiac cell receptor, eventually affecting CACNA1D. CACNA1D is a subunit gene of the CACNA1C gene. The *A* allele of the CACNA1C rs1006737 polymorphism serves as an effect modifier of the interaction between hypoactivation of prefrontal cortex cognitive systems and disinhibited automatic emotional regulation networks (Radua et al., 2013).

The second significant pathway for the bipolar disorder-cannabis connection denotes UBC as the major hub. Ubiquitin signaling system abnormalities have been frequently present in postmortem brains from subjects with a history of cannabis dependence. Of particular interest is the molecular SETDB1, a histone methyltransferase, which partially regulates expression of UBC (see Figure 17). Results from a NIH study revealed reduced mRNA expression of SETDB1 in children exposed in utero to cannabis (Morris, Dinieri, Szutorisz, & Hurd, 2013).

Limitations

Several variants remain unmatched to genes for schizophrenia, bipolar disorder and cannabis use. This may occur for a variety of reasons. The extensive nature of linkage disequilibrium can confound the interpretation of a gene association signal as the variant or variants can lie at substantial distance from the initial association signal (Johnson et al., 2008). If multiple variants are in strong linkage disequilibrium, pathway analysis may tag multiple variants to the same gene, resulting in redundant information for gene mapping. The nature and sequences of these unmapped regions may also be dependent on the array platform used. Both PGC datasets utilized Affymetrix Genome-Wide Human SNP Array 6.0, Affymetrix Genome-Wide Human SNP Array 5.0, and Illumina HumanHap 550. Illumina 100K was the array platform used for analysis of molecular data from the dbGaP dataset. Affymetrix and Illumina use different types of probes and different measures to assess signal quality, so the quality filter setting differs for these two platforms. Affymetrix uses multiple probes for each gene along with one-base mismatch probes intended as controls for non-specific hybridization. In contrast, the randomly generated Illumina arrays yield on the order of 30 copies of the same

oligonucleotide on the array, which provide an internal technical replication that Affymetrix lacks (Barnes, Freudenberg, Thompson, Aronow, & Pavlidis, 2005). Due to these differences in genome sequencing across datasets, even if variants have strong linkage disequilibrium, they may tag different sequences. Accordingly, mapping of non-functional intergenic variants can be confounded. Consequently, characteristics of these unmatched genes can be very informative in determining possible functional effect of intergenic variants.

The varying methods for sequencing genomes utilized in all of the samples from the PGC datasets may result in genes for which there are multiple identical *rs* identifiers. Results indicate presence of 35 bipolar disorder genes for which there are duplicate *rs* identifiers and 6 cannabis use genes for which there are duplicate *rs* identifiers for which there are duplicate *rs* identifiers. Most of these duplicates represent exact pairs, but it is unknown whether duplicates of a set tag the same sequence. Interpretation of potential causal molecular pathways conferred by analyzed genes may heavily depend on precise definition and annotation of the variants.

This thesis was primarily limited by lack of individual-level data relating schizophrenia and bipolar disorder to cannabis use. Accordingly, data incorporated several different population samples, and environmental factors such as other drug use, were not consistently collected across samples. As a result, predictive models could not be utilized to control for environmental risk factors related to schizophrenia and bipolar disorder that may be possible confounders. Moreover, due to the lack of individual-level data, causality of cannabis use to schizophrenia or bipolar disorder cannot be attributed. It remains unclear whether cannabis use precedes schizophrenia or bipolar disorder or

whether cannabis use wholly predicts onset of either disorder. Furthermore, varying control populations were used for each data sample, with controls for the dB GAP NIH study primarily being American, and controls for the PGC data sources being primarily of European-American descent. Several controls for the dbGaP NIH study were also African Americans recruited from St. Louis, Missouri. This may be problematic considering the health disparities unique to African Americans, and perhaps varying demographic characteristics between African Americans and European Americans.

A major limitation of this analysis was that cases and controls were not selected using population-based samples for the cannabis use dataset. Subjects were recruited from various places, including substance abuse treatment catchment centers located across the United States, chemical dependency units in the greater St. Louis metropolitan area and through a community based case-control family study based in Detroit, Michigan and St. Louis. Such treatment centers and placement sites can attract a variety of patients from varying demographic backgrounds and unique substance use patterns. Nevertheless, substance abuse disorders have been traditionally enriched and selected for in other, broader disorders, such as alcoholism in large, genome-wide association studies. Due to the high frequency of polydrug use with cannabis and other substances, recruiting for samples of isolated cannabis users may garner inadequate statistical power for our projected analyses (Palmer et al., 2014).

The variation in sampling sites could create bias and issues with generalizability undermine the credibility of associations between genotype and risk of disease. Associations between genotype and outcome may be confounded by unrecognized population stratification. Family-based designs, as implemented in COGA, have been an

alternative method of controlling for such confounding bias in studies (Allison, 1997). Processes previously described to assure homogeneity of samples, however, may have greatly reduced the risk of bias, although there is great heterogeneity between samples combined in dbGaP SAGE (Allison, 1997). In many genome-wide association studies, population stratification cannot be avoided, but the use of a principal component analysis to correct for stratification has been suggested (Patterson, Price, & Reich, 2006). The principal components method used in cleaning dbGaP data, one component comparing EA versus AA ancestry and another comparing Hispanic versus non-Hispanic ancestry, were used as regression covariates in the association model (Alkes L Price et al., 2006). The result of including such components could greatly minimize the risk of population stratification prior to analysis.

Public Health Implications and Future Directions in Research

Overall, these results highlight the blurring between behavioral phenotypes and environment, and have wider implications for perceptions of environmental risks for schizophrenia and bipolar disorder. Environmental factors can be both individually-based and hereditary. Several studies have shown that implicated risk factors for both schizophrenia and bipolar disorder such as urbanicity, religiosity and other illicit drug use have heritable components to them (Murray, Mehta, & Di Forti, 2014). Further research is needed to explain the overlap between genes implicated in cannabis use and other forms of substance abuse, as well as potential epigenetic interactions between cannabis use and common risk factors that confer increased risk. Exploration of common addiction phenotypes among polydrug users in relation to schizophrenia and bipolar disorder should also be considered.

References

Agrawal, A., Budney, A. J., & Lynskey, M. T. (2012). The co-occurring use and misuse of cannabis and tobacco: a review. *Addiction, 107*(7), 1221-1233.

Agrawal, A., & Lynskey, M. T. (2009). Candidate genes for cannabis use disorders: findings, challenges and directions. *Addiction, 104*(4), 518-532.

Aleman, A., Kahn, R. S., & Selten, J.-P. (2003). Sex differences in the risk of schizophrenia: evidence from meta-analysis. *Archives of general psychiatry, 60*(6), 565-571.

Alemany, S., Arias, B., Aguilera, M., Villa, H., Moya, J., Ibáñez, M. I., . . . Fañanás, L. (2011). *Childhood abuse, the BDNF-Val66Met polymorphism and adult psychotic-like experiences* (Vol. 199).

Allen, P., Chaddock, C. A., Howes, O. D., Egerton, A., Seal, M. L., Fusar-Poli, P., . . . McGuire, P. K. (2012). Abnormal Relationship Between Medial Temporal Lobe and Subcortical Dopamine Function in People With an Ultra High Risk for Psychosis. *Schizophrenia bulletin, 38*(5), 1040-1049. doi: DOI 10.1093/schbul/sbr017

Allison, D. B. (1997). Transmission-disequilibrium tests for quantitative traits. *American journal of human genetics, 60*(3), 676.

Altar, C. A., Jurata, L. W., Charles, V., Lemire, A., Liu, P., Bukhman, Y., . . . Brockman, J. A. (2005). Deficient hippocampal neuron expression of proteasome, ubiquitin, and mitochondrial genes in multiple schizophrenia cohorts. *Biol Psychiatry, 58*(2), 85-96. doi: 10.1016/j.biopsych.2005.03.031

Alvarado-Esquivel, C., Urbina-Álvarez, J. D., Estrada-Martínez, S., Torres-Castorena, A., Molotla-de-León, G., Liesenfeld, O., & Dubey, J. P. (2011). Toxoplasma gondii infection and schizophrenia: A case control study in a low Toxoplasma seroprevalence Mexican population. *Parasitology international, 60*(2), 151-155.

Arseneault, L., Cannon, M., Poulton, R., Murray, R., Caspi, A., & Moffitt, T. E. (2002). *Cannabis use in adolescence and risk for adult psychosis: longitudinal prospective study* (Vol. 325).

Barnes, M., Freudenberg, J., Thompson, S., Aronow, B., & Pavlidis, P. (2005). Experimental comparison and cross-validation of the Affymetrix and Illumina gene expression analysis platforms. *Nucleic acids research, 33*(18), 5914-5923.

Biegon, A., Grossman, R., Bokov, I., Lipitz, S., & Hoffmann, C. (2011). Abnormal temporal lobe development in cytomegalovirus (CMV) infected foetuses: quantitative MRI study in utero. *International Clinical Psychopharmacology*, *26*, e53-e54.

Bierut, L. J., Agrawal, A., Bucholz, K. K., Doheny, K. F., Laurie, C., Pugh, E., . . . Bertelsen, S. (2010). A genome-wide association study of alcohol dependence. *Proceedings of the National Academy of Sciences*, *107*(11), 5082-5087.

Bierut, L. J., Agrawal, A., Bucholz, K. K., Doheny, K. F., Laurie, C., Pugh, E., . . . Rice, J. P. (2010). A genome-wide association study of alcohol dependence. *Proc Natl Acad Sci U S A, 107*(11), 5082-5087.

Bilder, R. M., Volavka, J., Czobor, P. á., Malhotra, A. K., Kennedy, J. L., Ni, X., . . . Lieberman, J. A. (2002). Neurocognitive correlates of the COMT Val158Met

polymorphism in chronic schizophrenia. *Biological Psychiatry*, *52*(7), 701-707. doi: http://dx.doi.org/10.1016/S0006-3223(02)01416-6

Bourque, F., Van der Ven, E., & Malla, A. (2011). A meta-analysis of the risk for psychotic disorders among first-and second-generation immigrants. *Psychological medicine*, *41*(05), 897-910.

Bousman, C. A., Chana, G., Glatt, S. J., Chandler, S. D., May, T., Lohr, J., . . . Everall, I. P. (2010). Positive symptoms of psychosis correlate with expression of ubiquitin proteasome genes in peripheral blood. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *153B*(7), 1336-1341. doi: 10.1002/ajmg.b.31106

Carlisle, H. J., Fink, A. E., Grant, S. G., & O'Dell, T. J. (2008). Opposing effects of PSD-93 and PSD-95 on long-term potentiation and spike timing-dependent plasticity. *J Physiol*, *586*(Pt 24), 5885-5900. doi: 10.1113/jphysiol.2008.163469

Casadio, P., Fernandes, C., Murray, R. M., & Di Forti, M. (2011). Cannabis use in young people: The risk for schizophrenia. *Neuroscience & Biobehavioral Reviews*, *35*(8), 1779-1787. doi: http://dx.doi.org/10.1016/j.neubiorev.2011.04.007

Caspi, A., Moffitt, T. E., Cannon, M., McClay, J., Murray, R., Harrington, H., . . . Braithwaite, A. (2005). Moderation of the effect of adolescent-onset cannabis use on adult psychosis by a functional polymorphism in the catechol-O-methyltransferase gene: longitudinal evidence of a gene X environment interaction. *Biological Psychiatry*, *57*(10), 1117-1127.

Chen, D., Jiang, X., Akula, N., Shugart, Y., Wendland, J., Steele, C., . . . Jamain, S. (2013). Genome-wide association study meta-analysis of European and Asian-

ancestry samples identifies three novel loci associated with bipolar disorder. *Molecular psychiatry, 18*(2), 195-205.

Chrivia, J. C., Kwok, R. P., Lamb, N., Hagiwara, M., Montminy, M. R., & Goodman, R. H. (1993). Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature, 365*(6449), 855-859.

Compton, M. T., Furman, A. C., & Kaslow, N. J. (2004). Lower negative symptom scores among cannabis-dependent patients with schizophrenia-spectrum disorders: preliminary evidence from an African American first-episode sample. *Schizophrenia Research, 71*(1), 61-64.

Consortium, S. P. G.-W. A. S. (2011). Genome-wide association study identifies five new schizophrenia loci. *Nature genetics, 43*(10), 969-976.

Costas, J., Sanjuán, J., Ramos-Ríos, R., Paz, E., Agra, S., Tolosa, A., . . . Arrojo, M. (2011). Interaction between COMT haplotypes and cannabis in schizophrenia: a case-only study in two samples from Spain. *Schizophrenia Research, 127*(1), 22-27.

Craddock, N., O'Donovan, M. C., & Owen, M. J. (2009). Psychosis genetics: modeling the relationship between schizophrenia, bipolar disorder, and mixed (or "schizoaffective") psychoses. *Schizophrenia bulletin*, sbp020.

Creemers, H., Harakeh, Z., Dick, D., Meyers, J., Vollebergh, W., Ormel, J., . . . Huizink, A. (2011). DRD2 and DRD4 in relation to regular alcohol and cannabis use among adolescents: does parenting modify the impact of genetic vulnerability? The TRAILS study. *Drug and alcohol dependence, 115*(1), 35-42.

Cross-Disorder Group of the Psychiatric Genomics, C. (2013). Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *The Lancet, 381*(9875), 1371-1379.

Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *The Lancet, 381*(9875), 1371-1379. doi: http://dx.doi.org/10.1016/S0140-6736(12)62129-1

D'Souza, D. C., Abi-Saab, W. M., Madonick, S., Forselius-Bielen, K., Doersch, A., Braley, G., . . . Krystal, J. H. (2005). Delta-9-tetrahydrocannabinol effects in schizophrenia: Implications for cognition, psychosis, and addiction. *Biological Psychiatry, 57*(6), 594-608. doi: http://dx.doi.org/10.1016/j.biopsych.2004.12.006

Dalman, C., Allebeck, P., Cullberg, J., Grunewald, C., & Köster, M. (1999). Obstetric complications and the risk of schizophrenia: a longitudinal study of a national birth cohort. *Archives of general psychiatry, 56*(3), 234-240.

dbGAP, N. I. o. H. (2009a, 2007). DbSNP Short Genetic Variations. from http://www.ncbi.nlm.nih.gov/SNP/snp_viewBatch.cgi?sbid=33668

DbGaP, N. I. o. H. (2009b). [Readme file for dbGaP Study of Addiction and Genetics].

De Pradier, M., Gorwood, P., Beaufils, B., Adès, J., & Dubertret, C. (2010). Influence of the serotonin transporter gene polymorphism, cannabis and childhood sexual abuse on phenotype of bipolar disorder: A preliminary study. *European Psychiatry, 25*(6), 323-327. doi: http://dx.doi.org/10.1016/j.eurpsy.2009.10.002

Decoster, J., van Os, J., Kenis, G., Henquet, C., Peuskens, J., De Hert, M., & van Winkel, R. (2011). Age at onset of psychotic disorder: Cannabis, BDNF Val66Met, and

sex-specific models of gene–environment interaction. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 156*(3), 363-369.

Derks, E. M., Vorstman, J. A., Ripke, S., Kahn, R. S., Ophoff, R. A., & Consortium, S. P. G. (2012). Investigation of the genetic association between quantitative measures of psychosis and schizophrenia: a polygenic risk score analysis. *PloS one, 7*(6), e37852.

DeRosse, P., Kaplan, A., Burdick, K. E., Lencz, T., & Malhotra, A. K. (2010). Cannabis use disorders in schizophrenia: Effects on cognition and symptoms. *Schizophrenia Research, 120*(1–3), 95-100. doi: http://dx.doi.org/10.1016/j.schres.2010.04.007

Djurovic, S., Gustafsson, O., Mattingsdal, M., Athanasiu, L., Bjella, T., Tesli, M., . . . Morken, G. (2010). A genome-wide association study of bipolar disorder in Norwegian individuals, followed by replication in Icelandic sample. *Journal of Affective Disorders, 126*(1), 312-316.

DNAStar. ArrayStar Help. from http://www.dnastar.com/arraystar_help/index.html#!Documents/fdrbenjaminihochberg1.htm

Dudbridge, F., & Gusnanto, A. (2008). Estimation of significance thresholds for genomewide association scans. *Genet. Epidemiol., 32*, 227-234.

Ellett, L., Freeman, D., & Garety, P. A. (2008). The psychological effect of an urban environment on individuals with persecutory delusions: the Camberwell walk study. *Schizophrenia Research, 99*(1), 77-84.

Epstein, K. A., & Kumra, S. (2014). White matter fractional anisotropy over two time points in early-onset schizophrenia and adolescent cannabis use disorder: A

naturalistic diffusion tensor imaging study. *Psychiatry Research: Neuroimaging*(0). doi: http://dx.doi.org/10.1016/j.pscychresns.2014.10.010

Estrada, G., Fatjó-Vilas, M., Munoz, M., Pulido, G., Minano, M., Toledo, E., . . . Miret, S. (2011). Cannabis use and age at onset of psychosis: further evidence of interaction with COMT Val158Met polymorphism. *Acta Psychiatrica Scandinavica, 123*(6), 485-492.

Fan, N., Yang, H., Zhang, J., & Chen, C. (2010). Reduced expression of glutamate receptors and phosphorylation of CREB are responsible for in vivoΔ9-THC exposure-impaired hippocampal synaptic plasticity. *Journal of Neurochemistry, 112*(3), 691-702. doi: 10.1111/j.1471-4159.2009.06489.x

Fergusson, D. M., Horwood, L. J., & Ridder, E. M. (2005). Tests of causal linkages between cannabis use and psychotic symptoms. *Addiction, 100*(3), 354-366.

Ferreira, M. A. R., O'Donovan, M. C., Meng, Y. A., Jones, I. R., Ruderfer, D. M., Jones, L., . . . Green, E. K. (2008). Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nature genetics, 40*(9), 1056-1058.

Gietzen, D. (2010). Ingenuity Pathways Analysis (IPA) Of Large Datasets. from http://www.usc.edu/hsc/nml/assets/bioinfo/IPA/Data%20Analysis%20training%20Handouts.pdf

Griffith-Lendering, M. F., Wigman, J. T., Prince van Leeuwen, A., Huijbregts, S. C., Huizink, A. C., Ormel, J., . . . Vollebergh, W. A. (2013). Cannabis use and vulnerability for psychosis in early adolescence—a TRAILS study. *Addiction, 108*(4), 733-740.

Grigorenko, E., Kittler, J., Clayton, C., Wallace, D., Zhuang, S., Bridges, D., . . . Deadwyler, S. (2002). Assessment of cannabinoid induced gene changes: tolerance and neuroprotection. *Chem Phys Lipids, 121*(1-2), 257-266.

Guda, B. (2013). Tools and Algorithms in Bioinformatics. Retrieved March 15th, 2015, from http://unmc.edu/bsbc/docs/Week10.pdf

Hall, W., & Dengenhardt, L. (2008). Cannabis use and the risk of developing a psychotic disorder. *World Psychiatry, 7*(2), 68-71.

Hariri, A. R., Mattay, V. S., Tessitore, A., Kolachana, B., Fera, F., Goldman, D., . . . Weinberger, D. R. (2002). Serotonin transporter genetic variation and the response of the human amygdala. *Science, 297*(5580), 400-403.

Harley, M., Kelleher, I., Clarke, M., Lynch, F., Arseneault, L., Connor, D., . . . Cannon, M. (2010). Cannabis use and childhood trauma interact additively to increase the risk of psychotic symptoms in adolescence. *Psychological medicine, 40*(10), 1627-1634.

Henquet, C., Krabbendam, L., Spauwen, J., Kaplan, C., Lieb, R., Wittchen, H.-U., & Van Os, J. (2004). Prospective cohort study of cannabis use, predisposition for psychosis, and psychotic symptoms in young people. *Bmj, 330*(7481), 11.

Holowka, D. W., King, S., Saheb, D., Pukall, M., & Brunet, A. (2003). Childhood abuse and dissociative symptoms in adult schizophrenia. *Schizophrenia Research, 60*(1), 87-90.

Houston, J., Murphy, J., Shevlin, M., & Adamson, G. (2011). Cannabis use and psychosis: re-visiting the role of childhood trauma. *Psychological medicine, 41*(11), 2339-2348.

Ingenuity. (2015). Advanced Analytics for Scientists.

Iwasaki, S., Ishiguro, H., Higuchi, S., Onaivi, E. S., & Arinami, T. (2007). Association study between alcoholism and endocannabinoid metabolic enzyme genes encoding fatty acid amide hydrolase and monoglyceride lipase in a Japanese population. *Psychiatric genetics*, *17*(4), 215-220.

Jacobus, J., McQueeny, T., Bava, S., Schweinsburg, B. C., Frank, L. R., Yang, T. T., & Tapert, S. F. (2009). White matter integrity in adolescents with histories of marijuana use and binge drinking. *Neurotoxicology and teratology*, *31*(6), 349-355.

Javitt, D. C. (2007). Glutamate and schizophrenia: phencyclidine, N-methyl-D-aspartate receptors, and dopamine-glutamate interactions. *Int Rev Neurobiol, 78*, 69-108. doi: 10.1016/s0074-7742(06)78003-5

Jiang, Y., Jakovcevski, M., Bharadwaj, R., Connor, C., Schroeder, F. A., Lin, C. L., . . . Akbarian, S. (2010). Setdb1 histone methyltransferase regulates mood-related behaviors and expression of the NMDA receptor subunit NR2B. *The Journal of Neuroscience*, *30*(21), 7152-7167.

Johnson, A. D., Handsaker, R. E., Pulit, S. L., Nizzari, M. M., O'Donnell, C. J., & de Bakker, P. I. (2008). SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*, *24*(24), 2938-2939. doi: 10.1093/bioinformatics/btn564

Jutras-Aswad, D., Jacobs, M. M., Yiannoulos, G., Roussos, P., Bitsios, P., Nomura, Y., . . . Hurd, Y. L. (2012). Cannabis-dependence risk relates to synergism between

neuroticism and proenkephalin SNPs associated with amygdala gene expression: case-control study. *PloS one, 7*(6), e39243.

Kandel, D., Chen, K., Warner, L. A., Kessler, R. C., & Grant, B. (1997). Prevalence and demographic correlates of symptoms of last year dependence on alcohol, nicotine, marijuana and cocaine in the US population. *Drug and alcohol dependence, 44*(1), 11-29.

Kawanishi, Y., Harada, S., Tachikawa, H., Okubo, T., & Shiraishi, H. (1999). Novel variants in the promoter region of the CREB gene in schizophrenic patients. *Journal of human genetics, 44*(6), 428-430.

Kelleher, I., Harley, M., Lynch, F., Arseneault, L., Fitzpatrick, C., & Cannon, M. (2008). Associations between childhood trauma, bullying and psychotic symptoms among a school-based adolescent sample. *The British Journal of Psychiatry, 193*(5), 378-382.

Kelly, B. D., O'Callaghan, E., Waddington, J. L., Feeney, L., Browne, S., Scully, P. J., . . . Morgan, M. G. (2010). Schizophrenia and the city: A review of literature and prospective study of psychosis and urbanicity in Ireland. *Schizophrenia Research, 116*(1), 75-89.

Kempisty, B., Mostowska, A., Górska, I., Łuczak, M., Czerski, P., Szczepankiewicz, A., . . . Jagodziński, P. P. (2006). Association of 677C>T polymorphism of methylenetetrahydrofolate reductase (MTHFR) gene with bipolar disorder and schizophrenia. *Neuroscience Letters, 400*(3), 267-271. doi: http://dx.doi.org/10.1016/j.neulet.2006.02.055

Kikinis, Z., Asami, T., Bouix, S., Finn, C., Ballinger, T., Tworog-Dube, E., . . . Kubicki, M. (2012). Reduced fractional anisotropy and axial diffusivity in white matter in 22q11. 2 deletion syndrome: a pilot study. *Schizophrenia Research, 141*(1), 35-39.

Kittler, J. T., Grigorenko, E. V., Clayton, C., Zhuang, S. Y., Bundey, S. C., Trower, M. M., . . . Deadwyler, S. (2000). Large-scale analysis of gene expression changes during acute and chronic exposure to [Delta]9-THC in rats. *Physiol Genomics, 3*(3), 175-185.

Krabbendam, L., & Van Os, J. (2005). Schizophrenia and urbanicity: a major environmental influence—conditional on genetic risk. *Schizophrenia bulletin, 31*(4), 795-799.

Kuepper, R., Morrison, P. D., van Os, J., Murray, R. M., Kenis, G., & Henquet, C. (2010). Does dopamine mediate the psychosis-inducing effects of cannabis? A review and integration of findings across disciplines. *Schizophrenia Research, 121*(1–3), 107-117. doi: http://dx.doi.org/10.1016/j.schres.2010.05.031

Lataster, T., van Os, J., Drukker, M., Henquet, C., Feron, F., Gunther, N., & Myin-Germeys, I. (2006). Childhood victimisation and developmental expression of non-clinical delusional ideation and hallucinatory experiences. *Social Psychiatry and Psychiatric Epidemiology, 41*(6), 423-428. doi: 10.1007/s00127-006-0060-4

Lewinsohn, P. M., Klein, D. N., & Seeley, J. R. (1995). Bipolar disorders in a community sample of older adolescents: prevalence, phenomenology, comorbidity, and course. *Journal of the American Academy of Child & Adolescent Psychiatry, 34*(4), 454-463.

Li, M. D., Mangold, J. E., Seneviratne, C., Chen, G.-B., Ma, J. Z., Lou, X.-Y., & Payne, T. J. (2009). Association and interaction analyses of GABBR1 and GABBR2 with nicotine dependence in European-and African-American populations. *PloS one, 4*(9), e7055.

Løberg, E.-M., Helle, S., Nygård, M., Berle, J. Ø., Kroken, R. A., & Johnsen, E. (2014). The cannabis pathway to non-affective psychosis may reflect less neurobiological vulnerability. *Frontiers in Psychiatry, 5*. doi: 10.3389/fpsyt.2014.00159

Makkos, Z., Fejes, L., Inczédy-Farkas, G., Kassai-Farkas, A., Faludi, G., & Lazary, J. (2011). Psychopharmacological comparison of schizophrenia spectrum disorder with and without cannabis dependency. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 35*(1), 212-217.

Malaspina, D., Goetz, R. R., Friedman, J. H., Kaufmann, C. A., Faraone, S. V., Tsuang, M., . . . Blehar, M. C. (2014). Traumatic brain injury and schizophrenia in members of schizophrenia and bipolar disorder pedigrees.

Malmberg, M., Overbeek, G., Monshouwer, K., Lammers, J., Vollebergh, W. A., & Engels, R. C. (2010). Substance use risk profiles and associations with early substance use in adolescence. *Journal of behavioral medicine, 33*(6), 474-485.

Mannisto, P. T., & Kaakkola, S. (1999). Catechol-O-methyltransferase (COMT): biochemistry, molecular biology, pharmacology, and clinical efficacy of the new selective COMT inhibitors. *Pharmacol Rev, 51*(4), 593-628.

March, D., Hatch, S. L., Morgan, C., Kirkbride, J. B., Bresnahan, M., Fearon, P., & Susser, E. (2008). Psychosis and place. *Epidemiologic reviews, 30*(1), 84-100.

Matheson, S. L., Shepherd, A. M., Laurens, K. R., & Carr, V. J. (2011). A systematic meta-review grading the evidence for non-genetic risk factors and putative antecedents of schizophrenia. *Schizophrenia Research, 133*(1), 133-142.

McGrath, J., Saha, S., Welham, J., El Saadi, O., MacCauley, C., & Chant, D. (2004). A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. *BMC medicine, 2*(1), 13.

Medina, K. L., Nagel, B. J., Park, A., McQueeny, T., & Tapert, S. F. (2007). Depressive symptoms in adolescents: associations with white matter volume and marijuana use. *Journal of Child Psychology and Psychiatry, 48*(6), 592-600.

Miller, B., Messias, E., Miettunen, J., Alaräisänen, A., Järvelin, M.-R., Koponen, H., . . . Kirkpatrick, B. (2011). Meta-analysis of paternal age and schizophrenia risk in male versus female offspring. *Schizophrenia bulletin, 37*(5), 1039-1047.

Minocci, D., Massei, J., Martino, A., Milianti, M., Piz, L., Di Bello, D., . . . Nieri, P. (2011). Genetic association between bipolar disorder and 524A>C (Leu133Ile) polymorphism of CNR2 gene, encoding for CB2 cannabinoid receptor. *Journal of Affective Disorders, 134*(1–3), 427-430. doi: http://dx.doi.org/10.1016/j.jad.2011.05.023

Miyamoto, S., Miyake, N., Jarskog, L. F., Fleischhacker, W. W., & Lieberman, J. A. (2012). Pharmacological treatment of schizophrenia: a critical review of the pharmacology and clinical effects of current and future therapeutic agents. *Mol Psychiatry, 17*(12), 1206-1227.

Molloy, C., Conroy, R. M., Cotter, D. R., & Cannon, M. (2011). Is traumatic brain injury a risk factor for schizophrenia? A meta-analysis of case-controlled population-based studies. *Schizophrenia bulletin, 37*(6), 1104-1110.

Morris, C., Dinieri, J., Szutorisz, H., & Hurd, Y. L. (2013). Prenatal cannabis exposure leads to long-term disturbance of striatal

chromatin modifying enzymes from http://icahn.mssm.edu/static_files/MSSM/Files/Research/Institutes/Brain%20Institute/Retreat%202013%20pamphlet.pdf

Murray, R. M., & Lewis, S. W. (1987). Is schizophrenia a neurodevelopmental disorder? *Bmj, 295*(6600), 681-682.

Murray, R. M., Mehta, M., & Di Forti, M. (2014). Different dopaminergic abnormalities underlie cannabis dependence and cannabis-induced psychosis. *Biological Psychiatry, 75*(6), 430-431.

Nakata, K., Lipska, B. K., Hyde, T. M., Ye, T., Newburn, E. N., Morita, Y., . . . Kleinman, J. E. (2009). DISC1 splice variants are upregulated in schizophrenia and associated with risk polymorphisms. *Proceedings of the National Academy of Sciences, 106*(37), 15873-15878. doi: 10.1073/pnas.0903413106

Nakata, K., Ujike, H., Sakai, A., Uchida, N., Nomura, A., Imamura, T., . . . Kuroda, S. (2003). Association study of the brain-derived neurotrophic factor (BDNF) gene with bipolar disorder. *Neuroscience letters, 337*(1), 17-20.

Ngwa, J. S., Manning, A. K., Grimsby, J. L., Lu, C., Zhuang, W. V., & DeStefano, A. L. (2011). *Pathway analysis following association study.* Paper presented at the BMC proceedings.

Ochoa, S., Usall, J., Cobo, J., Labad, X., & Kulkarni, J. (2012). Gender differences in schizophrenia and first-episode psychosis: a comprehensive literature review. *Schizophrenia research and treatment, 2012*.

Palmer, R. H., Brick, L., Nugent, N. R., Bidwell, L. C., McGeary, J. E., Knopik, V. S., & Keller, M. C. (2014). Examining the role of common genetic variants on alcohol, tobacco, cannabis and illicit drug dependence: genetics of vulnerability to drug dependence. *Addiction, 110*, 530-537.

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS genetics, 2*(12), e190.

Patton, G. C., Coffey, C., Carlin, J. B., Degenhardt, L., Lynskey, M., & Hall, W. (2002). Cannabis use and mental health in young people: cohort study. *Bmj, 325*(7374), 1195-1198.

Pearce, B. D., Kruszon-Moran, D., & Jones, J. L. (2012). The relationship between Toxoplasma gondii infection and mood disorders in the third National Health and Nutrition Survey. *Biological Psychiatry, 72*(4), 290-295.

Pedersen, C. B., & Mortensen, P. B. (2001). Evidence of a dose-response relationship between urbanicity during upbringing and schizophrenia risk. *Archives of general psychiatry, 58*(11), 1039-1046.

Peerbooms, O., Rutten, B., Collip, D., Lardinois, M., Lataster, T., Thewissen, V., . . . Van Os, J. (2012). Evidence that interactive effects of COMT and MTHFR moderate psychotic response to environmental stress. *Acta Psychiatrica Scandinavica, 125*(3), 247-256.

Pelayo-Teran, J. M., Perez-Iglesias, R., Mata, I., Carrasco-Marin, E., Vazquez-Barquero, J. L., & Crespo-Facorro, B. (2010). Catechol-O-Methyltransferase (COMT) Val158Met variations and cannabis use in first-episode non-affective psychosis: clinical-onset implications. *Psychiatry Res, 179*(3), 291-296. doi: 10.1016/j.psychres.2009.08.022

Peters, B. D., de Haan, L., Dekker, N., Blaas, J., Becker, H. E., Dingemans, P. M., . . . den Heeten, G. J. (2008). White matter fibertracking in first-episode schizophrenia, schizoaffective patients and subjects at ultra-high risk of psychosis. *Neuropsychobiology, 58*(1), 19.

PGC Downloads. (2013). Retrieved February 4, 2015, 2014, from http://www.med.unc.edu/pgc/downloads

Pinggera, A., Lieb, A., Benedetti, B., Lampert, M., Monteleone, S., Liedl, K. R., . . . Striessnig, J. (2015). CACNA1D De Novo Mutations in Autism Spectrum Disorders Activate Cav1.3 L-Type Calcium Channels. *Biological Psychiatry, 77*(9), 816-822. doi: http://dx.doi.org/10.1016/j.biopsych.2014.11.020

Power, R., Verweij, K., Zuhair, M., Montgomery, G., Henders, A., Heath, A., . . . Martin, N. (2014). Genetic predisposition to schizophrenia associated with increased use of cannabis. *Molecular psychiatry, 19*(11), 1201-1204.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet, 38*(8), 904-909.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics, 38*(8), 904-909.

Psychiatric GWAS Consortium Bipolar Disorder Working Group. (2011). Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet, 43*(10), 977-983. doi: 10.1038/ng.943

Purcell, S., Cherny, S. S., & Sham, P. C. (2003). Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics, 19*, 149-150.

Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., . . . Morris, D. W. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature, 460*(7256), 748-752.

Quinn, H. R., Matsumoto, I., Callaghan, P. D., Long, L. E., Arnold, J. C., Gunasekaran, N., . . . McGregor, I. S. (2007). Adolescent Rats Find Repeated [Delta]9-THC Less Aversive Than Adult Rats but Display Greater Residual Cognitive Deficits and Changes in Hippocampal Protein Expression Following Exposure. *Neuropsychopharmacology, 33*(5), 1113-1126.

Radua, J., Surguladze, S. A., Marshall, N., Walshe, M., Bramon, E., Collier, D. A., . . . McDonald, C. (2013). The impact of CACNA1C allelic variation on effective connectivity during emotional processing in bipolar disorder. *Mol Psychiatry, 18*(5), 526-527. doi: 10.1038/mp.2012.61

Ramo, D. E., Liu, H., & Prochaska, J. J. (2012). Tobacco and marijuana use among adolescents and young adults: a systematic review of their co-use. *Clinical Psychology Review, 32*(2), 105-121.

Ringen, P., Melle, I., Birkenaes, A., Engh, J., Faerden, A., Vaskinn, A., . . . Andreassen, O. A. (2008). The level of illicit drug use is related to symptoms and premorbid functioning in severe mental illness. *Acta Psychiatrica Scandinavica, 118*(4), 297-304.

Ringen, P., Vaskinn, A., Sundet, K., Engh, J., Jonsdottir, H., Simonsen, C., . . . Andreassen, O. (2010). Opposite relationships between cannabis use and neurocognitive functioning in bipolar disorder and schizophrenia. *Psychological medicine, 40*(08), 1337-1347.

Rocío Pérez-Iglesias, D. T.-G., Philip K. McGuire, Gareth J. Barker, Roberto Roiz-Santiañez, Ignacio Mata, Enrique Marco de Lucas, Jose Manuel Rodríguez-Sánchez, Rosa Ayesa-Arriola, M.Psych.; Jose L. Vazquez-Barquero, Benedicto Crespo-Facorro. (2010). White Matter Integrity and Cognitive Impairment in First-Episode Psychosis. *American Journal of Psychiatry, 167*(4), 451-458. doi: doi:10.1176/appi.ajp.2009.09050716

Rutten, B. P., & Mill, J. (2009). Epigenetic mediation of environmental influences in major psychotic disorders. *Schizophrenia bulletin, 35*(6), 1045-1056.

Schizophrenia Psychiatric Genome-Wide Association Study, C. (2011). Genome-wide association study identifies five new schizophrenia loci. *Nature genetics, 43*(10), 969-976.

Schneider, M., & Koch, M. (2003). Chronic pubertal, but not adult chronic cannabinoid treatment impairs sensorimotor gating, recognition memory, and the performance in a progressive ratio task in adult rats. *Neuropsychopharmacology*.

Scholes, K., & Martin-Iverson, M. (2010). Cannabis use and neuropsychological performance in healthy individuals and patients with schizophrenia. *Psychological medicine, 40*(10), 1635-1646.

Scott, L. J., Muglia, P., Kong, X. Q., Guan, W., Flickinger, M., Upmanyu, R., . . . Absher, D. (2009). Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry. *Proceedings of the National Academy of Sciences, 106*(18), 7501-7506.

Sklar, P., Ripke, S., Scott, L. J., Andreassen, O. A., Cichon, S., Craddock, N., . . . Blackwood, D. (2011). Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nature genetics, 43*(10), 977.

Sklar, P., Smoller, J. W., Fan, J., Ferreira, M. A. R., Perlis, R. H., Chambert, K., . . . Purcell, S. M. (2008). Whole-genome association study of bipolar disorder. *Mol Psychiatry, 13*(6), 558-569. doi: http://www.nature.com/mp/journal/v13/n6/suppinfo/4002151s1.html

Stone, J. M., Fisher, H. L., Major, B., Chisholm, B., Woolley, J., Lawrence, J., . . . Johnson, S. (2014). Cannabis use and first-episode psychosis: relationship with manic and psychotic symptoms, and with age at presentation. *Psychological medicine, 44*(03), 499-506.

Strakowski, S. M., & DelBello, M. P. (2000). The co-occurrence of bipolar and substance use disorders. *Clinical Psychology Review, 20*(2), 191-206. doi: http://dx.doi.org/10.1016/S0272-7358(99)00025-2

Tedla, Y., Shibre, T., Ali, O., Tadele, G., Woldeamanuel, Y., Asrat, D., . . . Alem, A. (2011). Serum antibodies to Toxoplasma gondii and Herpesvidae family viruses in individuals with schizophrenia and bipolar disorder: a case-control study. *Ethiopian medical journal, 49*(3), 211-220.

Thaker, G. K. (2008). Neurophysiological endophenotypes across bipolar and schizophrenia psychosis. *Schizophrenia bulletin, 34*(4), 760-773.

The Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature, 447*(7145), 661-678. doi: http://www.nature.com/nature/journal/v447/n7145/suppinfo/nature05911_S1.html

Torrey, E. F., Bartko, J. J., & Yolken, R. H. (2012). Toxoplasma gondii and other risk factors for schizophrenia: an update. *Schizophrenia bulletin, 38*(3), 642-647.

van Os, J., Bak, M., Hanssen, M., Bijl, R. V., de Graaf, R., & Verdoux, H. (2002). Cannabis Use and Psychosis: A Longitudinal Population-based Study. *American Journal of Epidemiology, 156*(4), 319-327. doi: 10.1093/aje/kwf043

van Os, J., Kenis, G., & Rutten, B. P. (2010). The environment and schizophrenia. *Nature, 468*(7321), 203-212.

van Winkel, R. (2011). Family-based analysis of genetic variation underlying psychosis-inducing effects of cannabis: sibling analysis and proband follow-up. *Archives of general psychiatry, 68*(2), 148-157.

Verdejo-Garcia, A., Fagundo, A. B., Cuenca, A., Rodriguez, J., Cuyas, E., Langohr, K., . . . de la Torre, R. (2013). COMT val158met and 5-HTTLPR genetic polymorphisms moderate executive control in cannabis users. *Neuropsychopharmacology, 38*(8), 1598-1606. doi: 10.1038/npp.2013.59

Vink, J. M., Hottenga, J. J., de Geus, E. J., Willemsen, G., Neale, M. C., Furberg, H., & Boomsma, D. I. (2014). Polygenic risk scores for smoking: predictors for alcohol and cannabis use? *Addiction, 109*(7), 1141-1151. doi: 10.1111/add.12491

Vinkers, C. H., Van Gastel, W. A., Schubart, C. D., Van Eijk, K. R., Luykx, J. J., Van Winkel, R., . . . Wiersma, D. (2013). The effect of childhood maltreatment and cannabis use on adult psychotic symptoms is modified by the COMT Val(1)(5)(8)Met polymorphism. *Schizophr Res, 150*(1), 303-311. doi: 10.1016/j.schres.2013.07.020

Voss, M. W., Heo, S., Prakash, R. S., Erickson, K. I., Alves, H., Chaddock, L., . . . White, S. M. (2013). The influence of aerobic fitness on cerebral white matter integrity and cognitive function in older adults: Results of a one-year exercise intervention. *Human brain mapping, 34*(11), 2972-2985.

Wegener, N., & Koch, M. (2009). Behavioural disturbances and altered Fos protein expression in adult rats after chronic pubertal cannabinoid treatment. *Brain Research, 1253*(0), 81-91. doi: http://dx.doi.org/10.1016/j.brainres.2008.11.081

Williams, E.-J., Walsh, F. S., & Doherty, P. (2003). The FGF receptor uses the endocannabinoid signaling system to couple to an axonal growth response. *The Journal of Cell Biology, 160*(4), 481-486. doi: 10.1083/jcb.200210164

Wittchen, H.-U., Fröhlich, C., Behrendt, S., Günther, A., Rehm, J., Zimmermann, P., . . . Perkonigg, A. (2007). Cannabis use and cannabis use disorders and their relationship to mental disorders: a 10-year prospective-longitudinal community study in adolescents. *Drug and alcohol dependence, 88*, S60-S70.

Wrege, J., Schmidt, A., Walter, A., Smieskova, R., Bendfeldt, K., Radue, E.-W., . . . Borgwardt, S. (2014). Effects of cannabis on impulsivity: a systematic review of neuroimaging findings. *Current pharmaceutical design, 20*(13), 2126.

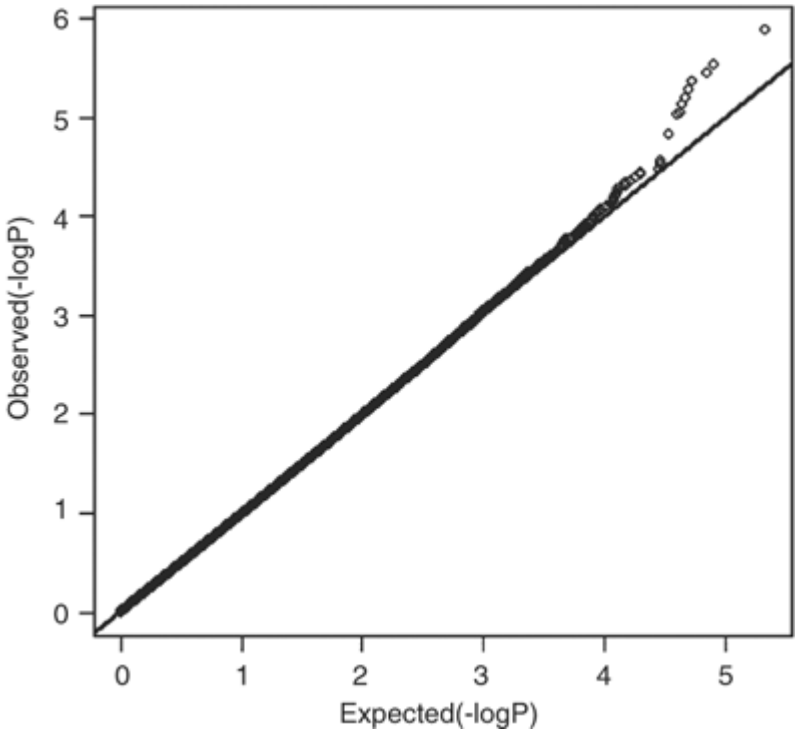
Xie, P., Kranzler, H. R., Krystal, J. H., Farrer, L. A., Zhao, H., & Gelernter, J. (2014). Deep resequencing of 17 glutamate system genes identifies rare variants in DISC1 and GRIN2B affecting risk of opioid dependence. *Addiction biology, 19*(5), 955-964. doi: 10.1111/adb.12072

Young, R., Lawford, B., Morris, C. P., Voisey, J., Hughes, I., & Swagell, C. D. (2009). Diagnostic methods and agents: Google Patents.

Zammit, S., Spurlock, G., Williams, H., Norton, N., Williams, N., O'DONOVAN, M. C., & Owen, M. J. (2007). Genotype effects of CHRNA7, CNR1 and COMT in schizophrenia: interactions with tobacco and cannabis use. *The British Journal of Psychiatry, 191*(5), 402-407.

Supplementary Figures

Figure S1. HWE Q-Q Plot of EA and AA sample comparison for dbGAP SAGE[1]



[1] *Reference: DbGaP (2009b)*