HHS Public Access

Author manuscript

Dev Psychopathol. Author manuscript; available in PMC 2017 May 01.

Published in final edited form as:

Dev Psychopathol. 2016 November ; 28(4 Pt 1): 1089–1101. doi:10.1017/S0954579416000717.

Molecular genetic approaches to understanding the comorbidity of psychiatric disorders

Ian R. Gizer, PhD

Department of Psychological Sciences, University of Missouri, Columbia, MO 65211

Abstract

Epidemiologic studies demonstrating high rates of co-occurrence among psychiatric disorders at the population level have contributed to large literatures focused on identifying the causal mechanisms underlying the patterns of co-occurrence among these disorders. Such efforts have long represented a core focus of developmental psychopathologists and have more recently been supported by the Research Domain Criteria (RDoC) initiative developed by the National Institute of Mental Health (NIMH), which provides a further framework for how the hypothesized mechanisms can be studied at different levels of analysis. The present overview focuses on molecular genetic approaches that are being used currently to study the etiology of psychiatric disorders, and how these approaches have been applied in efforts to understand the biological mechanisms that give rise to comorbid conditions. The present report begins with a review of molecular genetic approaches used to identify individual variants that confer risk for multiple disorders and the intervening biological mechanisms that contribute to their comorbidity. This is followed by a review of molecular genetic approaches that use genetic data in aggregate to examine these questions, and concludes with a discussion of how developmental psychopathologists are uniquely positioned to apply these methods in a way that will further our understanding of the causal factors that contribute to the development of comorbid conditions.

Introduction

Given the large literature focused on using psychometric methods for understanding the underlying structure of psychopathology and the current Research Domain Criteria (RDoC) initiative developed by the National Institute of Mental Health (NIMH) (Insel et al., 2010; Sanislow et al., 2010), there is a broadening interest in identifying transdiagnostic phenotypes that can potentially explain why certain groups of psychiatric diagnoses show higher and lower rates of co-occurrence in the population (Kessler, Chiu, Demler, & Walters, 2005; Kessler et al., 1994). Though many disciplines have informed this recent interest, the field of developmental psychopathology has been particularly instrumental in demonstrating the importance of studying such phenotypes (Cicchetti & Toth, 2009). Further, as developmental psychopathologists have long argued, and the collection of papers included in this special issue demonstrate, there are many lines of research operating at different levels of analysis that can be used to address this question (Beauchaine & Gatzke-Kopp, 2012;

Correspondence regarding this article should be sent to: Ian R. Gizer, Department of Psychological Sciences, University of Missouri, 210 McAlester Hall, Columbia, MO 65211, Phone: 573-882-5427, Fax: 573-882-7710, gizeri@missouri.edu.

Cicchetti, 2008). The present report will describe current molecular genetic methods that are being used to further our understanding of the shared risk mechanisms underlying these disorders and presumably contributing to their comorbidity.

Background

The seminal work of Thomas Achenbach using factor analytic methods to delineate the internalizing and externalizing domains of childhood psychopathology (Achenbach, 1966) was integral to the formation of the field of developmental psychopathology (Achenbach, 1974; Cicchetti, 1984; Sroufe & Rutter, 1984) and has led to important inquiries and advances in the classification and development of psychiatric disorders (Cicchetti, 2013). One example that has received significant attention is the extension of the internalizing (e.g., major depressive disorder, phobias, generalized anxiety disorder) and externalizing (e.g., conduct disorder, substance use disorders, antisocial personality disorders) constructs to adult psychopathology (Kessler, et al., 2005; Krueger, Caspi, Moffitt, & Silva, 1998), and subsequent studies seeking to further refine models conceptualizing the underlying structure of psychopathology (Caspi et al., 2014). Notably, developmental psychopathologists have been particularly well-positioned to examine these patterns of co-occurrence over time through the use of longitudinal study designs, and through such efforts, have made considerable advances in identifying the risk mechanisms that underlie psychopathology (Beauchaine & Thayer, 2015; Burnette & Cicchetti, 2012).

One focus of these efforts has centered on understanding how an underlying deficit might result in certain symptoms that persist throughout development, a phenomenon termed homotypic continuity, and equally importantly, how an underlying deficit can also manifest as different though correlated symptoms across childhood, adolescence, and adulthood, a phenomenon termed heterotypic continuity (Rutter, 1989). Homotypic and heterotypic continuity have been observed for both internalizing and externalizing disorders, with some of the clearest examples stemming from research examining the course of persistent externalizing behavior over the lifespan. These studies have demonstrated that a proportion of children exhibiting symptoms of ADHD will continue to report clinically significant symptoms into adulthood (Mannuzza, Klein, Bessler, Malloy, & LaPadula, 1998). Additionally, a proportion of children exhibiting such symptoms, particularly those with an early onset, will develop symptoms of oppositional defiant disorder and conduct disorder, symptoms of one or more substance use disorders, and eventually symptoms of antisocial personality disorder in adulthood (Moffitt, 1993; Moffitt, Caspi, Harrington, & Milne, 2002).

Similar patterns of homotypic and heterotypic continuity have been observed for internalizing disorders. For example, children and adolescents diagnosed with an anxiety disorder (e.g., separation anxiety disorder, social phobia, generalized anxiety disorder) are likely to report meeting criteria for such a disorder in adulthood, though not always the same one (Gregory et al., 2007; Ormel et al., 2015), suggesting that patterns of homotypic and heterotypic continuity are characteristic of internalizing disorder as well. Notably, substantial heterotypic continuity that crosses the internalizing and externalizing spectra has also been reported (Kessler, et al., 2005), indicating that shared mechanisms are likely to

operate within and across the internalizing and externalizing domains of psychopathology to influence risk for disorder. This has led some to define continuity within the externalizing disorders or internalizing domains that emerge over time as a reflection of *broad* homotypic continuity and reserve use of the term heterotypic continuity only for observable patterns of continuity that cross the internalizing and externalizing domains (Gregory, et al., 2007).

As these studies demonstrate, developmental psychopathologists have long recognized that the mechanisms contributing to these patterns of continuity can explain much of the comorbidity observed between disorders, and importantly, further emphasize that a comprehensive understanding of these mechanisms requires their study at multiple levels of analysis (Cicchetti, 2008). This view comes from an appreciation of the transactional nature of genetic and environmental influences involved in the etiology of psychiatric disorders and has resulted in an emphasis on studying the etiology of psychiatric disorders at the genetic, physiological, neural and social and behavioral levels (Cicchetti & Dawson, 2002). This perspective has now been widely recognized and adopted in the fields of psychology and psychiatry as evidenced by the RDoC Initiative, which attempts to provide a framework for describing the levels of analysis at which mechanisms related to the development, course, and treatment of psychiatric disorders can be studied (Insel, et al., 2010; Sanislow, et al., 2010).

The following review will focus on current methods being used to further our understanding of the shared risk mechanisms underlying psychiatric disorders and presumably contribute to their comorbidity at the molecular genetic level. Investigations at this level are motivated primarily by quantitative genetic studies (e.g., family, twin, and adoption studies) that are used to estimate the heritability underlying a given trait or the genetic correlation between two or more traits (Plomin, DeFries, Knopik, & Neiderheiser, 2013). Univariate studies have consistently demonstrated that psychiatric disorders are moderately to highly heritable (h^2 = 0.30–0.80), and further that these heritability estimates tend to increase across development (Bergen, Gardner, & Kendler, 2007; Eaves, Long, & Heath, 1986). Thus, molecular geneticists have been interested in identifying the individual genetic variants that contribute to these heritability estimates and whether a subset of these variants might contribute to risk for multiple disorders.

Quantitative Genetic Studies of Comorbidity

Several seminal papers have laid out different models of how two disorders could be comorbid in an individual (e.g., Klein & Riso, 1993; Neale & Kendler, 1995), with each model making specific predictions regarding how etiological factors lead to their co-occurrence. For example, the ‘chance’ model of comorbidity stipulates that the two disorders are completely independent in terms of their etiology and is supported when the rate of co-occurrence in the population is simply the product of the prevalence rates of the two disorders. Multiformity models assert that two etiologically distinct diseases exist, but also indicate that the presence of one disorder can lead to symptoms that mimic those of the second disorder even in the absence of causal factors that define the second disorder. In contrast, correlated liability models indicate that etiologic factors are shared across the two co-occurring disorders, which gives rise to their comorbidity.

Multivariate quantitative genetic studies using the classic twin design have been elegantly applied to help disentangle which of these models provides the best fit to data observed in clinical and community populations (e.g., Agrawal et al., 2007; Rhee, Willcutt, Hartman, Pennington, & DeFries, 2008). Generally speaking, these studies have supported previous research suggesting that distinct but correlated etiological factors underlie internalizing (e.g., mood and anxiety disorders) and externalizing disorders (e.g., conduct disorder, antisocial personality disorder, alcohol and other substance use disorders), though some studies have suggested the presence of an overarching 'bifactor' that influences risk for psychopathology at a general level (Lahey, Van Hulle, Singh, Waldman, & Rathouz, 2011; Tackett et al., 2013). Despite differences in the factor structure underlying these disorders, the referenced studies suggest that a substantial proportion of genetic variation that contributes to risk for psychopathology is shared across diagnostic categories, and further suggest that this shared genetic risk likely acts through multiple transdiagnostic mechanisms that have the potential to produce the high rates of co-occurrence among disorders as observed in epidemiologic studies (Beauchaine & McNulty, 2013). It is important to note, however, that quantitative genetic studies were designed to estimate the magnitude of genetic and environmental influences that contribute to the development of a given trait or set of traits (Plomin, et al., 2013). As a result, such studies can estimate the heritability of a trait or the genetic correlations between two or more traits, but cannot identify the specific genetic variants that contribute to these estimates. Nonetheless, the described quantitative genetic studies support the argument that individual genetic variants and the genes that they influence are not related to a specific disorder, but instead influence one or more low-level biological traits that, through a complex transactional process, interact with one another and environmental mechanisms to influence risk for psychopathology (e.g., Kendler, 2005).

This view has led psychiatric geneticists to take an active interest in better understanding the phenomena of pleiotropic associations, the observation of a single variant showing a genetic association with multiple traits. Understanding why a specific instance of pleiotropy occurs can be difficult to resolve as multiple competing explanations can account for such a result. In providing a definition of pleiotropy that builds on previous work in this area (Paaby & Rockman, 2013) and is applicable to the study of psychiatric disorders, Solovieff and colleagues (2013) provide an important distinction between instances of biological pleiotropy and mediated pleiotropy, which have been previously referred to as horizontal and vertical pleiotropy, respectively (Tyler, Asselbergs, Williams, & Moore, 2009). Specifically, biological pleiotropy occurs when two seemingly independent traits are found to be directly influenced by the same genetic variant or variants within a single gene suggesting that the function of that gene influences both traits. For example, loss of function variants within the phenylalanine hydroxylase (*PAH*) gene, which encodes for an enzyme that metabolizes phenylalanine into tyrosine, are known to cause phenylketonuria (PKU). These variants result in an inability of the body to metabolize phenylalanine into tyrosine, which can lead to intellectual disability if phenylalanine levels are not monitored through diet. Notably, tyrosine also plays an important role in melanin synthesis, and thus, individuals that inherit these variants are also more likely to have reduced skin pigmentation. Thus, each trait is influenced by the same genetic locus through independent pathways (Lobo, 2008). In contrast, mediated pleiotropy occurs when a single variant or collection of variants within a

single gene directly influences one trait which then increases risk for the second trait. For example, some have suggested that variants in the *ADH1B* gene, which are robustly associated with alcohol consumption, also have an indirect influence on cardiovascular disease through increased alcohol consumption (Holmes et al., 2014).

Despite significant interest and effort, examples such as these remain fairly rare in the psychiatric genetics literature, but they serve as a useful illustration of how such studies can provide important contributions to our understanding of the mechanisms underlying comorbidity. Additionally, such findings have led to efforts to expand the tools molecular geneticists can use to explore this phenomenon. Before providing a review of these methods, however, it is important to provide a brief introduction to the field of molecular genetics.

Molecular Genetic Approaches to Studying Psychiatric Disorders

Association studies, including GWAS, use a simple case–control design to test whether a given allele at a measured genetic variant is observed at a higher frequency among affected "cases" when compared to unaffected "controls." Early association studies tended to focus on candidate genes because of the significant costs associated with genotyping a single variant at that time. As a result, candidate gene studies of psychological traits and disorders typically focused on just one or two single nucleotide variants (SNVs) or tandem repeat polymorphisms in the same limited set of 'usual suspects' that included genes involved in monoaminergic function (i.e., dopamine and serotonin transporter genes [DAT1, 5HTT], dopamine receptor genes [DRD2, DRD4], and monoamine metabolizing genes [MAOA, COMT]), given the relevance of these neurotransmitter systems to brain-behavior relations (Munafo, 2006). As a result, each of these genes has been studied in relation to multiple psychological traits as well as traits influenced by psychological factors, such as BMI, with many positive associations reported. These early candidate gene studies led some to suggest that the associations observed across phenotypes might be evidence of pleiotropy, and suggested that this might provide one explanatory mechanism for the high rates of co-occurrence among psychiatric disorders (Ebstein, Benjamin, & Belmaker, 2003). Nonetheless, the poor replication record of candidate gene studies in general also led to questions of whether the reported findings might represent false positives resulting in part from publication bias and selective reporting of results (Hirschhorn, Lohmueller, Byrne, & Hirschhorn, 2002; Munafo, 2006).

In some ways, genome-wide association (GWA) studies provided a response to the potential bias of candidate genes studies in allowing for an atheoretical approach to variant discovery through interrogation of the full genome rather than just a single gene. This is not without cost, however, as the individual examination of hundreds of thousands of variants requires a severe correction to control for Type I error with a critical p-value of 1×10^{-8} having gained acceptance as the threshold for declaring genome-wide significance (McCarthy et al., 2008). This correction has obvious implications regarding the sample sizes required to obtain adequate statistical power when conducting a GWAS, which are based on the anticipated effect sizes of the variants to be discovered.

Early GWA studies, as well as candidate gene studies, hypothesized that psychological traits would be complex in nature but have an oligogenic architecture suggesting that a relatively

small set of genes (i.e., 20–50 genes) would be involved in the etiology of the traits (Plomin & Crabbe, 2000). A “complex trait” is defined as any phenotype that does not exhibit a classic Mendelian inheritance pattern that can be attributed to a single genetic locus (Lander & Schork, 1994). As a result, multiple genes and variants contribute to a complex trait with each variant conferring only a relatively small risk for the disorder. Further, none of the risk variants by themselves are necessary nor sufficient to cause the disorder, and thus, genetic heterogeneity, where different collections of genotypes can result in the same phenotype, is characteristic of complex traits. Assuming an oligogenic architecture, early studies were conducted assuming that a portion of the variants underlying psychological traits would have modest effects on the phenotype (i.e., $R^2 = 0.02 – 0.05$). As a result, researchers were optimistic that moderately sized GWA studies with sample sizes of 6,000 cases and 6,000 controls would identify a significant portion of the genetic variation underlying these traits (Hirschhorn & Daly, 2005).

Nearly a decade of research has now shown, however, that these early oligogenic models were incorrect and that the genetic architecture underlying these traits is significantly more complex than initially anticipated. The full extent of this complexity has been made clear by large-scale GWA studies of anthropometric (e.g., height, weight), medical (e.g., autoimmune disorders, heart failure), and psychiatric traits demonstrating that individual variants typically explain less than 0.5% of variation in the trait under study (Sullivan, Daly, & O’Donovan, 2012). Further, additive models that attempt to summarize risk across a set of associated variants typically account for no more than 3% of the variation in psychiatric phenotypes, which is well short of the heritability estimates obtained from twin studies of psychiatric disorders that typically range from 30–80% (Vrieze, Iacono, & McGue, 2012). This discrepancy has been described as the ‘missing heritability’ problem (Manolio et al., 2009), and serves to highlight the difficulties in identifying individual variants that influence risk for a given psychiatric disorder, which also extend to studies examining whether a given variant is exhibiting a pleiotropic effect on more than one psychiatric disorder.

Though a complete discussion of the ‘missing heritability’ problem falls outside the scope of the current review, it is important to briefly discuss some of the potential sources of the ‘missing heritability’ relevant to the current topic as we will return to them later in the review. For example, sources of genetic variation that are not interrogated through GWAS, such as short repeat and copy number variation and rare genetic variation, are frequently cited and have been demonstrated to explain a portion of the missing heritability in several traits (Manolio, et al., 2009). In addition, gene-environment correlations (rGE) and gene × environment interactions (G×E) can lead to increased heritability estimates, and thus, have also been explored as potential contributors to the ‘missing heritability’ problem (Beauchaine & McNulty, 2013; Plomin, et al., 2013). Although each of these potential contributors are of interest to molecular geneticists attempting to identify individual genes and variants that might have a pleiotropic effect on two or more complex traits, the latter two are more directly relevant to the field of developmental psychopathology, and are therefore discussed in more detail later in the report.

In addition to insights regarding the genetic architecture of complex traits and the ‘missing heritability’ problem, a general finding to emerge from GWA studies is that pleiotropy

appears to be a widespread phenomenon. A review of published GWA studies catalogued in the National Human Genome Research Institute database suggested that 17% of genes and almost 5% of variants contained within the database have shown evidence of association with more than one trait, suggesting that pleiotropy is a relatively common phenomenon (Sivakumaran et al., 2011). Included among the results are potential pleiotropic effects between psychiatric and medical conditions such as schizophrenia and cardiovascular disease (Andreassen et al., 2013a), but the extension of these findings to pleiotropy between psychiatric disorders is limited (Andreassen et al., 2013b). Further, while we may refer to a single variant or gene influencing multiple psychiatric disorders as a pleiotropic effect, it may be more an artifact of diagnoses that are not biologically-informed separating individuals into artificial groups (Smoller, 2013). Nonetheless, such effects when observed may prove useful in refining psychiatric diagnoses. As a result, there is substantial enthusiasm in the field for studies that have sought to use molecular genetic data to better understand patterns of co-occurrence between psychiatric disorders. As noted earlier, however, the small effect sizes associated with these variants has resulted in relatively slow progress. In the following sections, we describe methods that have been used to identify variants that appear to influence multiple psychological traits, beginning with approaches that test the relations between individual variants and multiple traits and then proceeding to describe approaches that aggregate data across multiple variants to examine their collective influence on these traits.

Methods for Identifying Individual Variants Related to Multiple Traits

Cross-Disorder Approaches

One of the most commonly used approaches to studying comorbidity and cross-phenotype associations has been to conduct association studies on populations of individuals diagnosed with more than one related psychiatric disorder (e.g., multiple externalizing disorders). This approach is supported by psychometric studies investigating the latent structure of psychopathology. For example, many have argued that one explanation for the high rates of co-occurrence among psychiatric disorders is that they represent artificial distinctions of an underlying continuum, and that those at the more severely affected end of this continuum are likely to exhibit multiple related conditions (Carragher, Krueger, Eaton, & Slade, 2015). Some molecular genetic studies have built upon these conclusions by identifying individuals with two related conditions, such as ADHD and conduct disorder, with the assumption that such individuals exhibit a more severe form of externalizing psychopathology (Gizer, Otto, & Ellingson, 2015). For example, one such study conducted a GWAS of conduct disorder symptoms in a population of individuals with comorbid ADHD (Anney et al., 2008). Similar studies have been conducted looking at symptoms of oppositional defiant disorder in individuals with comorbid ADHD (Aebi et al., 2015) and aggressive behavior in individuals with comorbid ADHD (Brevik et al., 2016). Each of the described studies, however, failed to identify any genome-wide significant results, highlighting an important limitation of this approach. Specifically, it can be very difficult to collect adequate numbers of individuals exhibiting symptoms of both disorders to conduct an adequately powered study. Nonetheless, a recent GWAS meta-analysis of comorbid mood and anxiety disorders was able to amass a sample of over 18,000 individuals and successfully identified multiple

variants that are presumably related to a broad risk for internalizing disorders (Otowa et al., 2016). Thus, such studies clearly hold promise for identifying genetic variants that contribute to comorbid conditions if adequate sample sizes can be achieved.

A complementary approach that does not rely on identifying individuals with comorbid conditions has focused on the hypothesis that a broad genetic risk factor might underlie all of psychopathology as described earlier (Caspi, et al., 2014; Tackett, et al., 2013). The Cross-Disorder Group of the Psychiatric Genomics Consortium conducted the first large-scale study of this nature by conducting a GWAS meta-analysis in which they combined data from studies of schizophrenia, bipolar disorder, major depressive disorder, ADHD, and autism to contrast a broadly defined 'affected' group with a comparison group (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013b). The final sample, which included 33,332 cases and 27,888 controls, identified two SNPs in two genes, both L-type voltage-gated calcium channel subunit genes (CACNA1C and CACNB2), that were associated with increased risk for each of the five psychiatric disorders included in the meta-analysis. While this study represents one of the first successes in identifying measured variants that confer broad risk for psychopathology, it also highlights the difficulties associated with identifying such variants. Despite the large sample size, it is important to note that only two variants achieved genome-wide significance, and further, the odds ratios associated with these variants were small, ranging from 1.07 to 1.13. This result suggests that this form of cross-disorder analysis will require the same large sample sizes that have been necessary to identify variants associated with a single disorder, such as schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014).

Nonetheless, future studies using this approach have the potential to provide important insights into the genetic influences that contribute to different dimensions of psychopathology depending on their design. For example, the described study combined data across samples of individuals diagnosed with disorders characterized by symptoms of internalizing, externalizing, and thought disorders. As stated by the authors, this study design was used to identify genes relevant to all forms of psychopathology. It is easy to imagine future studies that combine data within a single domain of psychopathology in an attempt to identify variants specific to that domain. For example, data from studies focusing on various forms of externalizing, including both child and adult disorders, could be used to identify genes relevant to general disinhibition (Gizer & Ehlers, 2015).

Data Reduction Approaches

As substantial progress has been made in understanding the underlying structure of psychopathology using the described psychometric approaches, psychiatric geneticists have also begun incorporating latent variable models into their research to identify transdiagnostic phenotypes that might yield novel genetic loci associated with psychopathology and/or strengthen previously observed associations by focusing on a more biologically-informed phenotype. For example, Hicks and colleagues (2011) used factor analysis to identify a heritable behavioral disinhibition phenotype using biometric twin modeling of symptom data from the Minnesota Twin Family Study (MTFS). In a subsequent study, the authors used this derived phenotype to conduct a GWAS (McGue et al., 2013). The study identified

some suggestive associations, but unfortunately, none of them reached genome-wide significance. This result should not be surprising, however, as the study sample consisted of ~7000 participants, which is underpowered to detect variants of effect sizes typical of complex traits.

This study also highlights a potential obstacle facing the field as it considers alternatives to the DSM in terms of classification and measurement of psychopathology, and particularly those alternatives focused on the use of latent variable models to define traits underlying psychopathology. As stated, the large scale meta-analytic GWAS efforts being conducted for biometric traits, such as height and BMI and for psychiatric disorders such as schizophrenia and bipolar disorder, have repeatedly demonstrated the need for sample sizes in the 10's of thousands to identify common genetic variants associated with complex traits (Kendler, 2013). This presents unique difficulties for molecular genetic studies of psychopathology seeking to take a latent variable approach to phenotype measurement. Large-scale meta-analytic studies have been possible because researchers have focused on a single set of agreed upon criteria for a disorder in the form of the DSM diagnostic criteria as well as the same set of combinatorial rules for merging data across these criteria to reach a diagnosis (Sullivan, et al., 2012). This allows researchers working independently to integrate their datasets with some confidence in the assumption that they are measuring the same construct in a uniform manner. In contrast, researchers using latent variable approaches typically construct the models within their own dataset, and then conduct association analyses in that same dataset (e.g., Dick et al., 2008; McGue, et al., 2013). As a result, even if the same observable measures are used to assess the latent construct in two independent samples, the resulting factor models may differ in terms of item loadings and even structure based on nuances in each dataset. This makes it difficult to create datasets large enough to conduct an adequately powered GWAS.

Despite this difficulty, large-scale collaborative efforts have begun to yield some progress. For example, the Genetics of Personality Consortium consists of a large number of researchers that have collected data using commonly used self-report measures of personality, such as the NEO Personality Inventory-Revised (van den Berg et al., 2014). Using item response theory techniques to harmonize data across study sites, they were able to create a latent measure of neuroticism that was then successfully applied to GWA data for more than 60,000 individuals across 29 cohorts. This led to the identification of a single variant in the Membrane Associated Guanylate Kinase, WW And PDZ Domain Containing 1 (*MAGI1*) gene that yielded a genome-wide significant association with neuroticism (van Den Berg et al., 2015). This gene has been previously associated with schizophrenia and bipolar disorder, suggesting its possible relevance to neuroticism (Karlsson et al., 2012). Nonetheless, the fact that this was the only genome-wide significant result despite a very large combined sample size, demonstrates that we should not assume a simpler genetic architecture to emerge when using a phenotype derived from a latent variable approach relative to one derived from a set of diagnostic criteria.

Multivariate Approaches

A complementary approach has been to use multivariate linear mixed models to conduct a combined analysis of multiple traits. This approach examines the correlational structure between a set of dependent variables and then estimates how well the independent variable, in this case a genetic variant, can explain this structure (Zhou & Stephens, 2014). This approach presents some advantages over the discussed data reduction techniques. Most importantly, multivariate linear mixed models can achieve greater statistical power relative to a univariate analysis, by effectively treating each of the correlated trait values of the dependent variables as a unique observation (Stephens, 2013). This has the effect of increasing the effective sample size relative to a univariate analysis, including that of a latent factor score derived from a structural equation model. An initial study examined the potential of such an approach in the PGC and found that this method increased the sample size by at least 30% in terms of statistical power (Maier et al., 2015). Methods applying this approach have only recently been developed, and have not yet been widely applied to the field of psychiatric genetics. Thus, it will be interesting to see how these methods are utilized in the coming years.

Endophenotypes

It was an early appreciation of the genetic complexity and difficulties associated with identifying genetic variants that confer risk for complex traits that led Gottesman & Shields (1972) and later Gottesman & Gould (2003) to apply the concept of endophenotypes to psychiatric disorders. They conceptualized endophenotypes as the intervening variables that could “mark the path” between a gene and the observable disorder (Gottesman & Gould, 2003). As has been discussed, variation in a single gene can lead to multiple disorders, and endophenotypes can represent an important tool for studying the comorbidity of psychiatric disorders. For example, if an endophenotype is found to be shared across multiple disorders, identifying the genetic variants that influence that endophenotype can provide insight into how genetic influences might contribute to the comorbidity of those disorders.

The term “endophenotype” has been the subject of some controversy in the field in terms of its precise definition (see Kendler, Neale, Heath, Kessler, & Eaves, 1994; Miller & Rockstroh, 2013). For the sake of brevity, the term endophenotype will be used here to refer to heritable traits that serve to increase risk for the manifest disorder. Thus, endophenotypes can act as mediating variables through which genes and their encoded products influence biological processes to ultimately confer vulnerability to disorder. In this manner, endophenotypes represent a subset of the broader category of biomarkers that can reflect both causes and consequences of disease (Beauchaine, 2009). Because there are multiple intervening levels of analysis between genes and their encoded products and behavior, endophenotypes can be identified at each of these intervening levels. As a result, endophenotypes can be defined along a continuum with respect to their location between the gene and the manifest disorder. At the most proximal end of this continuum with respect to the studied gene, transcript levels of that gene could be used as an endophenotype. At the most distal end of this continuum, personality trait variables derived from questionnaires could be used as an endophenotype. An important implication of this continuum is the expectation that endophenotypes more proximal to the encoded gene product will have a

simpler genetic architecture, and as a consequence, will show stronger relations with individual genetic variants relative to endophenotypes that are more distal. This benefit in terms of gene discovery is not without cost, however, as the increased distance from the manifest disorder will make interpretation of how the endophenotype increases risk for the disorder more difficult to interpret. As a result, identifying endophenotypes along this continuum from gene to disorder will be critical if molecular genetic studies are going to yield gene discoveries that inform future diagnostic and classification systems. This necessity can be readily seen in the RDoC, which identifies transdiagnostic mechanisms at multiple levels of analysis that can be used to study the pathophysiology of mental illness (Insel, et al., 2010; Sanislow, et al., 2010).

To date, molecular genetic studies that have used this approach to study psychiatric disorders have focused largely on endophenotypes that are fairly distal with regard to individual genes and more proximal to the disorders under study (Gizer, et al., 2015). These include measures of personality, neuropsychological function, and even gross measures of neurological function. Despite early promises that these endophenotypes would have simpler genetic architectures and lead to stronger associations with genetic variants, these studies have largely yielded results that have been difficult to replicate and of similar magnitude to those reported for the manifest disorders themselves (Flint & Munafò, 2007). Nonetheless, large consortia have been or are in the process of being developed to study a range of endophenotypes that have been proposed for different disorders such as levels of the pro-inflammatory cytokines interleukin (IL)-1β and tumor necrosis factor alpha (TNFα) in relation to autism (Careaga et al., 2015) and schizophrenia (Miller, Buckley, Seabolt, Mellor, & Kirkpatrick, 2011) as well as structural and functional measures of brain regions relevant to these disorders such as the hippocampus, thalamus, and amygdala (Thompson et al., 2014). As the efforts of these consortia begin to yield results, it will allow for more rigorous applications of the endophenotype approach and begin to answer the question of whether these endophenotypes will lead to stronger and more robust associations. If so, these variants and the endophenotypes they undergird can be studied in more detail as potentially contributing to comorbidity between disorders.

Summary

Across the four methods described, it should be clear that regardless of the methodology used to investigate genetic influences on psychiatric traits, these traits are genetically complex. Thus, even if they could be perfectly measured, the individual variants that contribute to their development will by definition be many and of small effect. As a result, attempts to identify individual variants with pleiotropic effects on psychiatric disorders will require the same large sample sizes that have been required to identify variants associated with a single disorder, which we should expect for any complex trait regardless of how the phenotype is defined.

Methods for Studying Comorbidity Using Molecular Genetic Data in Aggregate

In addition to the described single variant approaches, psychiatric geneticists have been interested in using molecular genetic data in aggregate to better understand the relations between disorders and their comorbidity. Two methods have garnered substantial interest in this respect, and thus, are discussed in more detail. Although they are more frequently used in single variant studies, Mendelian randomization designs are also discussed in this section given recent studies suggesting polygenic risk scores can be utilized within these designs.

Genomic Similarity Methods

Twin studies use the difference in genetic relatedness between monozygotic twins, who share 100% of their genetic sequence in common, with dizygotic twins, who share 50% of their genetic sequence in common, to partition variation in a trait into genetic and environmental influences. Although unrelated individuals inherit a much smaller proportion of their genetic sequence from common ancestors compared to first degree relatives, a measurable proportion of their DNA is in fact shared, and this proportion varies across pairs of individuals drawn from the population. SNP data from a GWA microarray can be used to estimate this proportion, and thus, be used to calculate the genetic similarity (or distance) between all pairs of individuals within a sample (Kang et al., 2010; Yang, Lee, Goddard, & Visscher, 2011). These measured genetic relationships can then be included in a linear mixed model and used to partition variation in a trait into a genetic component explained by these measured genetic relationships and a residual term. This genetic component quantifies the heritable influences on a trait that are captured by the full set of variants contained on the GWA microarray, and as a result, has been referred to as a “narrow-sense” or “SNP-based” heritability estimate.

When applied to psychiatric disorders, these approaches have yielded heritability estimates that are approximately one-third to one-half the size of those estimated by twin studies. For example, a recent study estimated the “SNP-based” heritability of ADHD to be 0.25, whereas twin studies yield heritability estimates around 0.75 (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013a). As can be seen, these estimates are noticeably higher than the proportion of variance explained by a set of top results from a GWAS analysis of individual variants, suggesting that a substantial proportion of genetic risk is captured by these microarrays, but insufficient sample sizes and very small effect sizes prevent their identification. As a result, the discrepancy between the variance in a trait explained by the top GWAS SNPs and the SNP-based heritability estimate has been referred to as ‘hidden’ heritability (Wray et al., 2013). At the same time, the SNP-based heritability estimates still fall well short of the heritability estimates provided by twin studies, and thus, the discrepancy between these estimates has been referred to as “still missing” heritability.

As described earlier, one potential explanation for the ‘missing heritability,’ which would include the ‘still missing’ heritability, is the presence of gene × environment interactions (G×E), which can lead to increased heritability estimates produced by twin studies. Genomic similarity methods have been extended to estimate the proportion of variance that can be

explained by G×E at the genome-wide level due to a measured environmental variable in a manner similar to methods used in twin samples (e.g., Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003). For example, heritable variation in body mass index estimated using genomic similarity methods (BMI) has been shown to increase among college educated individuals relative to those without a college education (Boardman et al., 2014). Similarly, measured physical activity has been shown to influence the heritability of BMI with increased physical activity associated with reduced heritability (Guo, Liu, Wang, Shen, & Hu, 2015). Alternatively, gene-environment correlations (rGE) have been proposed as contributing to the increases in heritability of several traits observed over time, and have thus been investigated as potentially contributing to the 'missing heritability' problem as well (Beam & Turkheimer, 2013; Plomin, 2014). As a result, some initial studies have begun using genomic similarity methods to investigate rGE, such as a recent study suggesting that the relation between family SES and child intelligence is primarily the result of passive rGE (Trzaskowski et al., 2014). As each of these studies demonstrate, genomic similarity methods provide an innovative approach for investigating the contributions of G×E and rGE to the 'missing heritability' problem.

Of relevance to the current special issue, the use of genomic similarity methods to investigate rGE has been possible due the extension of these methods to bivariate models that estimate the shared genetic etiology between two traits of interest (Lee, Yang, Goddard, Visscher, & Wray, 2012). Thus, genomic similarity methods can be used to estimate the genetic correlation between two traits in a manner similar to that of bivariate twin analyses, and as a result, are of direct relevance to the study of comorbid conditions. This also has significant implications for the study of psychological traits and particularly for the study of rarer psychiatric conditions in which it is difficult to obtain sufficiently large twin samples of affected individuals to conduct well-powered studies. For example, prevalence estimates of schizophrenia are estimated around 1% and for bipolar I disorder the estimate is between 3–4%. As a result, twin studies investigating their shared etiology have been limited to relatively small samples (e.g., Cardno, Rijsdijk, Sham, Murray, & McGuffin, 2002), whereas population-based studies of unrelated individuals can be used to obtain much larger samples. For example, the Cross-Disorder Working Group of the PGC examined the genetic correlations between the five disorders that were the focus of their initial efforts, schizophrenia, bipolar disorder, major depressive disorder, ADHD, and autism (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013a). Bipolar disorder and schizophrenia showed the strongest genetic correlation with a co-heritability estimate of 0.68. Significant co-heritability estimates were also observed between schizophrenia and major depressive disorder (r_g = .43), bipolar disorder and major depressive disorder (r_g = .47), and ADHD and major depressive disorder (r_g = .32).

These findings provide an exciting illustration of how molecular genetic data can be used to examine the genetic contributions to the co-occurrence of psychological traits at the population level. Additionally, such methods can be applied in a longitudinal framework by examining the genetic correlation between symptoms of a disorder assessed in childhood and the symptoms of a second disorder assessed in adulthood. Nonetheless, genomic similarity methods are limited in terms of testing hypothetical models that can explain this co-occurrence. At the outset of this paper, it was noted that Klein and Riso (1993) outlined a

series of causal models offering competing explanations of how two disorders might be comorbid in an individual. Neale and Kendler (1995) extended these models by describing how family and twin study data could be used to explicitly test these models through a series of structural equation models, and their approach has been used to examine the magnitude and direction of causal factors influencing related disorders, such as ADHD and conduct disorder (Rhee, et al., 2008) and the progression of substance use (Agrawal, Neale, Prescott, & Kendler, 2004). Notably, the described genomic similarity methods cannot evaluate the fit of such models at present. Thus, extending these methods in a manner that allows for tests distinguishing between these models of comorbidity would represent an important advance that would help further our understanding of the shared genetic risk factors that contribute to psychiatric disorders.

Polygenic Risk Score Methods

Polygenic risk score methods refer to a collection of analytic approaches that attempt to collapse data across multiple measured variants. In contrast to the described genomic similarity methods which use a variance components approach to quantify the latent contribution of genetic variation to a phenotype, polygenic risk score methods combine data from measured variants in an additive fashion to create an aggregate measure of genetic risk for each participant. The measured variants can be selected either empirically from previous GWA studies or theoretically based on hypothesized relations between variants within a collection of genes and a trait of interest (e.g., dopaminergic genes and schizophrenia). For purposes of the present discussion, however, we will limit our examples to those of the former type. In such studies, data from the previously conducted GWA study or meta-analysis of GWA studies serve as a *discovery* sample and are used to identify a set of variants exhibiting *p*-values that meet a prespecified cutpoint (e.g., $p < .05$). Using GWA data from an independent sample, which serve as the *validation* sample, each participant's polygenic risk score is calculated by (1) identifying the participant's genotype at each qualifying variant, (2) counting the number of associated alleles a participant possesses at each variant (i.e., 0, 1 or 2), (3) weighting the allele count for each variant by its effect size or other measure of association as observed in the discovery sample, and (4) summing the weighted values across variants to create each participant's *polygenic risk score*. Regression analyses are then conducted to determine the amount of variation in the phenotype that is explained by this risk score in the validation sample. An important feature of this approach is that, depending on the cutpoint that is used, polygenic risk scores can allow for the inclusion of variants that may not have achieved genome-wide significance, but nonetheless, are involved in the etiology of the trait. For example, a composite score comprised from counts of risk alleles of variants with nominally significant *p*-values (i.e., $p<0.05$) will capture many such variants while the effects of spurious associations on the risk score should be randomly distributed across individuals, and thus, minimized (see Wray, et al., 2013 for a thorough review of polygenic risk score methods).

Similar to the genomic similarity methods described above, the polygenic risk score approach was initially used to assess the proportion of genetic variation that GWA study results could explain in a phenotype of interest. For example, studies of schizophrenia suggest that polygenic risk scores calculated from the results of the Schizophrenia Working

Group of the PGC can account for more than 6% of the variation in diagnostic status in independent samples (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). More recently, researchers have extended these approaches to estimate the extent to which a polygenic risk score developed for an initial trait can explain variation in a second, related trait. For example, Vink and colleagues (Vink et al., 2014) used data from the meta-analytic efforts of the Tobacco and Genetics (TAG) Consortium (Tobacco and Genetics Consortium, 2010), which conducted a GWA study of the number of cigarettes an individual smoked per day, to calculate a polygenic risk score based on these results in the Netherlands Twin Register sample. The authors used this polygenic risk score to determine whether genetic risk for increased smoking could also predict the number of alcoholic drinks consumed per week and whether a participant had ever used marijuana. Though the effect sizes were small, the authors reported significant relations between the polygenic risk score and each of the other substance use measures suggesting some overlap in the genetic influences underlying these phenotypes.

A question remaining from such studies, however, is whether the genetic effects are direct or indirect in terms of their influence on these phenotypes. In other words, do the genes that influence number of cigarettes smoked per day also directly and independently increase risk for drinking and marijuana use, or is the effect an indirect one in which the genes influence cigarettes smoked per day, which in turn, increases risk for drinking and marijuana use? To refer back to the review provided by Sovolieff et al. (2013), are the pleiotropic effects examples of direct pleiotropy or mediated pleiotropy? This is an important question when trying to understand comorbidity, and thus, genomic similarity and polygenic risk score methods need to be adapted to address this limitation.

Because polygenic risk scores provide a continuous measure of genetic liability, they have already been used to study such questions as they can be included as independent variables in a variety of psychometric models, including mediation models. To illustrate, several studies have proposed that cigarettes act as a ‘gateway drug’ for later alcohol and illicit drug use (Kandel, Yamaguchi, & Chen, 1992; Torabi, Bailey, & Majd-Jabbari, 1993). Competing models suggest that polysubstance use may result from a set of shared etiological factors rather than the effect of ‘soft’ drugs predisposing an individual for later use of ‘hard’ drugs (Degenhardt et al., 2010). Thus, a mediation analysis could be conducted to determine whether the relation between the polygenic risk score for smoking quantity and alcohol dependence was mediated by the number of cigarettes an individual smoked per day. If the genetic variants were shown to have an indirect effect on alcohol dependence, this could be seen as supporting the gateway hypothesis, which can be represented as a correlated liabilities model of comorbidity in which the symptoms of the first disorder, in turn, give rise to symptoms of the second disorder. In contrast, if mediation was not observed and the polygenic risk score provided a contribution to alcohol dependence independent of the number of cigarettes smoked per day, this would support a correlated liabilities model in which a set of shared risk factors contribute directly to risk for each disorder.

The application of polygenic risk score methods as described above are also relevant to investigations of gene-environment interplay in the form of gene-environment interaction (G×E) and gene-environment correlation (rGE) as was described for the genomic similarity

methods. In a particularly elegant example, a recent study explored how polygenic risk scores could be used to examine genetic influences on externalizing behavior over time and also whether such influences were moderated by aspects of the rearing environment (Salvatore et al., 2014). Specifically, the authors developed a polygenic risk score based on adult externalizing behavior, and then used that risk score to predict externalizing behavior in independent samples of children and young adults. A significant relation was observed between the 'adult' risk score and childhood externalizing behavior that was further moderated by parental monitoring and peer substance use such that higher levels of the former and lower levels of the latter reduced the influence of the polygenic risk score. Methods for applying polygenic risk scores to the study of rGE have also been developed, but have yet to be widely applied (Marceau et al., 2016). Nonetheless, the extension of such methods to include investigations of environmental influences represents an important area of future research that could help inform developmental models of continuity as they relate to the manifestation of psychopathology across the lifespan.

Mendelian Randomization

Another design that is being used with increased frequency to critically evaluate causal pathways between related phenotypes is Mendelian randomization (MR) (Smith & Ebrahim, 2003). The MR design is based on instrumental variable analysis, which was developed in economics and has been likened to a randomized control trial. The design rests upon the idea that if one were to hypothesize a causal pathway between two variables such as cigarette use leading to increased alcohol consumption, a simple regression analysis studying that relation would be vulnerable to all possible confounding variables. Even if one were to include the known, relevant confounding variables as covariates in the analysis, there would still be the possibility of unmeasured variables influencing the relationship. However, if a proxy for the exposure variable, in this example smoking, could be identified that was randomly distributed in the population, and thus free of these confounds, this would allow for stronger conclusions about causation to be made. Because the alleles that we inherit from our parents are randomly distributed during meiosis, genetic variants can represent such a proxy, as any possible confounders should be equally distributed across genotypes. As a result, the association between the proxy and outcome variable provides an unbiased test of whether the exposure variable is causally related to the outcome variable (see Gage, Smith, Zammit, Hickman, & Munafò, 2013 for an overview of MR and its potential applications to the study of psychiatric disorders).

As an illustration, a recent study examined the influence of cigarette smoking on body mass index (BMI) (Freathy et al., 2011). The authors used the rs1051730 variant in the *CHRNA5-A3-B4* gene cluster on chromosome 15 as the instrumental variable given its strong relation with smoking (Tobacco and Genetics Consortium, 2010). In their sample, rs1051730 genotype was associated with smoking as well as with BMI. An important assumption of MR is that the instrumental variable (i.e., genotype) only influences the outcome variable (i.e., BMI) through the exposure variable (i.e., smoking), and as a result, the presence of pleiotropy (e.g., rs1051730 influencing both phenotypes independently) would violate this assumption. To evaluate this possibility, the authors stratified their sample by smoking status and found that rs1051730 influenced BMI among smokers but not non-smokers. Thus,

rs1051730 was not associated with BMI independent of smoking, providing support for the described MR assumption as well as the conclusion that smoking is causally related to BMI. In this manner, MR studies have the potential to provide powerful investigations of the causal pathways between related disorders, and thus, provide insight into their comorbidity.

It should be noted that there are some limitations to the MR design. Most importantly, there have to be known variants robustly associated with the causal variable under study, and further, the smaller the magnitude of this relation, the larger the required sample size (Brion, Shakhbazov, & Visscher, 2013). For this reason, polygenic risk scores have been proposed as an alternative to individual variants as instrumental variables in MR studies, which is why they are being discussed at this point of the review. Nonetheless, this can make it more difficult to evaluate the assumptions underlying MR, such as the absence of pleiotropy. Studies have been conducted combining data across a small set of variants (Lawlor et al., 2013), but applications of MR using polygenic risk scores comprised of hundreds or even thousands of variants are limited. For substance use disorders, examining the relation between genetic risk and the outcome variable using unexposed individuals (e.g., nonsmokers) can serve as a test for pleiotropy, but similar examples, may not be readily available for other disorders. Nonetheless, MR studies hold particular promise and represent another important approach that molecular geneticists can use to better understand comorbidity among psychiatric disorders.

Summary

The above examples provide an important illustration of how genomic similarity methods, polygenic risk scores and MR designs can be used to the disentangle the genetic mechanisms that give rise to co-occurring disorders, and how the latter two can do so in a manner that allows for direct tests of competing models of comorbidity. Importantly, these examples can also be used to emphasize the unique opportunities for developmental psychopathologists interested in these research questions to engage in molecular genetic research. While the mediation analysis described above is intentionally simple, it is not difficult to imagine more sophisticated research questions that could be addressed using polygenic risk scores in longitudinal datasets. For example, longitudinal studies with data pertaining to the onset of different substance use disorders could be used to conduct more sophisticated tests of mediation. Additionally, causal pathways from childhood to adult disorders could be more directly tested using polygenic risk scores derived from GWAS of early onset disorders such as ADHD to determine whether a common set of genes led directly or indirectly to adult disorders such as depression or substance use disorders. In this manner, hypotheses regarding the distinction between internalizing and externalizing disorders, as well as patterns of homotypic and heterotypic continuity between them, could be tested and the results used to validate developmental models of psychopathology.

Conclusions

Despite the potential importance of each of the described methods to the study of comorbidity, there are some limitations that should be noted. First, with regards to the described phenotype refinement approaches, it should be noted that these approaches may

further our understanding of the risk factors that underlie psychiatric traits, but it is unreasonable to expect that such methods will yield stronger genetic associations with measured variants. Psychiatric traits are multifactorial in nature and highly polygenic. As has been the case for anthropometric traits such as height, which can be almost perfectly measured, the highly polygenic nature of these traits dictates that the effects of any single variant will be small in terms of explained variance at the population level. Thus, phenotype refinement approaches will likely encounter the same difficulties in terms of variant identification as those for the overarching diagnoses. As a result, methods that can facilitate such analyses across large consortia are necessary if adequate sample sizes are to be obtained.

Second, the methods used to study genetic variation in aggregate, and the genomic similarity methods in particular, are limited in that they can indicate whether a genetic correlation exists between two traits, but they are limited in their ability to describe the nature of this correlation, and thus, have not yet been used to test competing models of comorbidity. Further, while polygenic risk score methods have the potential to test such models as described above, it is important to note that these methods cannot do so in a comprehensive manner. Specifically, the polygenic risk score approach explains only a small fraction of the overall genetic risk that contributes to the development of a given disorder. As a result, the conclusions that can be drawn from such analyses only apply to the proportion of genetic risk captured by the score. For example, polygenic risk scores for number of cigarettes smoked per day explain ~5% of the variance in this trait ~0.5% of the variance in alcohol dependence diagnoses (Meyers et al., 2013; Vink, et al., 2014). Given that the heritability estimates for these traits have been reported to be much higher, the presented analyses can only be used to make conclusions regarding a fraction of the genetic risk underlying these traits. This is not true of MR studies, but, the power of the MR design lies in its ability to detect causal relationships in the absence of pleiotropy, and thus, can only be applied to a limited set of hypothesized models regarding comorbid conditions.

Despite these limitations, the molecular genetic approaches currently being applied to the study of comorbidity among psychiatric disorders have already provided important contributions to our understanding of how these disorders develop. As these methods continue to evolve, the proportion of risk captured by such models is likely to increase, and thus, they promise to provide further contributions to our understanding of the biological mechanisms that contribute to the development and co-occurrence of psychiatric disorders. Importantly, these questions are of great importance to psychopathology researchers and developmental psychopathologists, in particular, as evidenced by this special issue. The genomic similarity and polygenic risk score methods described in this report do not require the large samples needed for GWAS, and given that genotyping and sequencing technologies will continue to evolve and decline in terms of cost, developmental psychopathologists that are interested in this area of research will have access to or be able to generate the necessary data to conduct such studies. Therefore, it is important that development psychopathology as a field continues to engage and collaborate with psychiatric geneticists to address important topics in the field, including the underlying causes of comorbidity between psychiatric disorders.

Acknowledgments

This work was supported by a grant from the National Institute on Drug Abuse to IRG (DA030976).

References

Achenbach TM. The classification of children's psychiatric symptoms: a factor-analytic study. Psychological Monographs: General and Applied. 1966; 80:1–37. DOI: 10.1037/h0093906

Achenbach, TM. Developmental psychopathology. Oxford, England: Ronald Press; 1974.

Aebi M, van Donkelaar MM, Poelmans G, Buitelaar JK, Sonuga-Barke EJ, Stringaris A, et al. Gene-set and multivariate genome-wide association analysis of oppositional defiant behavior subtypes in attention-deficit/hyperactivity disorder. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2015; Advance online publication. doi: 10.1002/ajmg.b.32346

Agrawal A, Lynskey MT, Bucholz KK, Martin NG, Madden PA, Heath AC. Contrasting models of genetic co-morbidity for cannabis and other illicit drugs in adult Australian twins. Psychological Medicine. 2007; 37:49–60. DOI: 10.1017/S0033291706009287 [PubMed: 17076916]

Agrawal A, Neale MC, Prescott CA, Kendler KS. A twin study of early cannabis use and subsequent use and abuse/dependence of other illicit drugs. Psychological Medicine. 2004; 34:1227–1237. DOI: 10.1017/S0033291704002545 [PubMed: 15697049]

Andreassen OA, Djurovic S, Thompson WK, Schork AJ, Kendler KS, O'Donovan MC, et al. Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. The American Journal of Human Genetics. 2013a; 92:197–209. DOI: 10.1016/j.ajhg.2013.01.001 [PubMed: 23375658]

Andreassen OA, Thompson WK, Schork AJ, Ripke S, Mattingsdal M, Kelsoe JR, et al. Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate. PLoS Genetics. 2013b; 11:e1005544.doi: 10.1371/journal.pgen.1003455

Anney RJ, Lasky-Su J, O'Dushlaine C, Kenny E, Neale BM, Mulligan A, et al. Conduct disorder and ADHD: evaluation of conduct problems as a categorical and quantitative trait in the international multicentre ADHD genetics study. American Journal of Medical Genetics Part B Neuropsychiatric Genetics. 2008; 147B:1369–1378. DOI: 10.1002/ajmg.b.30871 [PubMed: 18951430]

Beam CR, Turkheimer E. Phenotype–environment correlations in longitudinal twin models. Development and Psychopathology. 2013; 25:7–16. DOI: 10.1017/S0954579412000867 [PubMed: 23398749]

Beauchaine TP. The role of biomarkers and endophenotypes in prevention and treatment of psychopathological disorders. Biomarkers in medicine. 2009; 3:1–3. DOI: 10.2217/17520363.3.1.1 [PubMed: 19727417]

Beauchaine TP, Gatzke-Kopp LM. Instantiating the multiple levels of analysis perspective in a program of study on externalizing behavior. Development and Psychopathology. 2012; 24:1003–1018. DOI: 10.1017/S0954579412000508 [PubMed: 22781868]

Beauchaine TP, McNulty T. Comorbidities and continuities as ontogenic processes: Toward a developmental spectrum model of externalizing psychopathology. Development and Psychopathology. 2013; 25:1505–1528. DOI: 10.1017/S0954579413000746 [PubMed: 24342853]

Beauchaine TP, Thayer JF. Heart rate variability as a transdiagnostic biomarker of psychopathology. International Journal of Psychophysiology. 2015; 98:338–350. DOI: 10.1016/j.ijpsycho.2015.08.004 [PubMed: 26272488]

Bergen SE, Gardner CO, Kendler KS. Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood: a meta-analysis. Twin Research and Human Genetics. 2007; 10:423–433. DOI: 10.1375/twin.10.3.423 [PubMed: 17564500]

Boardman JD, Domingue BW, Blalock CL, Haberstick BC, Harris KM, McQueen MB. Is the gene-environment interaction paradigm relevant to genome-wide studies? The case of education and body mass index. Demography. 2014; 51:119–139. DOI: 10.1007/s13524-013-0259-4 [PubMed: 24281739]

Brevik EJ, van Donkelaar MM, Weber H, Sánchez-Mora C, Jacob C, Rivero O, et al. Genome-wide analyses of aggressiveness in attention-deficit hyperactivity disorder. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2016; Advance online publication. doi: 10.1002/ajmg.b.32434

Brion MJA, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. International Journal of Epidemiology. 2013; 42:1497–1501. DOI: 10.1093/ije/dyt179 [PubMed: 24159078]

Burnette ML, Cicchetti D. Multilevel approaches toward understanding antisocial behavior: Current research and future directions. Development and Psychopathology. 2012; 24:703–704. DOI: 10.1017/S0954579412000314 [PubMed: 22781849]

Cardno AG, Rijsdijk FV, Sham PC, Murray RM, McGuffin P. A twin study of genetic relationships between psychotic symptoms. American Journal of Psychiatry. 2002; 159:539–545. DOI: 10.1176/appi.ajp.159.4.539 [PubMed: 11925290]

Careaga M, Rogers S, Hansen RL, Amaral DG, Van de Water J, Ashwood P. Immune Endophenotypes in Children with Autism Spectrum Disorder. Biological Psychiatry. 2015; Advance online publication. doi: 10.1016/j.biopsych.2015.08.036

Carragher N, Krueger RF, Eaton NR, Slade T. Disorders without borders: current and future directions in the meta-structure of mental disorders. Social Psychiatry and Psychiatric Epidemiology. 2015; 50:339–350. DOI: 10.1007/s00127-014-1004-z [PubMed: 25557024]

Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S, et al. The p factor one general psychopathology factor in the structure of psychiatric disorders? Clinical Psychological Science. 2014; 2:119–137. DOI: 10.1177/2167702613497473 [PubMed: 25360393]

Cicchetti D. The emergence of developmental psychopathology. Child Development. 1984; 55:1–7. DOI: 10.2307/1129830 [PubMed: 6705613]

Cicchetti, D. A multiple-levels-of-analysis perspective on research in development and psychopathology. In: Beauchaine, TP.; Hinshaw, SP., editors. Child and Adolescent Psychopathology. 1st. Hoboken, New Jersey: John Wiley & Sons; 2008. p. 27-57.

Cicchetti, D. An overview of developmental psychopathology. In: Zelazo, PD., editor. The Oxford Handbook of Developmental Psychology. Vol. 2. Oxford, England: Oxford University Press; 2013. p. 455-480.

Cicchetti D, Dawson G. Editorial: Multiple levels of analysis. Development and Psychopathology. 2002; 14:417–420. DOI: 10.1017/S0954579402003012 [PubMed: 12349866]

Cicchetti D, Toth SL. The past achievements and future promises of developmental psychopathology: The coming of age of a discipline. Journal of Child Psychology and Psychiatry. 2009; 50:16–25. DOI: 10.1111/j.1469-7610.2008.01979.x [PubMed: 19175810]

Cross Disorders Group of the Psychiatric Genomics Consortium. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nature Genetics. 2013a; 45:984–994. DOI: 10.1038/ng.2711 [PubMed: 23933821]

Cross Disorders Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet. 2013b; 381(12):1371–1379. 62129–1. DOI: 10.1016/s0140-6736 [PubMed: 23453885]

Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014; 511:421–427. DOI: 10.1038/nature13595 [PubMed: 25056061]

Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nature Genetics. 2010; 42:441–447. DOI: 10.1038/ng.571 [PubMed: 20418890]

Degenhardt L, Dierker L, Chiu WT, Medina-Mora ME, Neumark Y, Sampson N, et al. Evaluating the drug use “gateway” theory using cross-national data: consistency and associations of the order of initiation of drug use among participants in the WHO World Mental Health Surveys. Drug and Alcohol Dependence. 2010; 108:84–97. DOI: 10.1016/j.drugalcdep.2009.12.001 [PubMed: 20060657]

Dick DM, Aliev F, Wang JC, Grucza RA, Schuckit M, Kuperman S, et al. Using dimensional models of externalizing psychopathology to aid in gene identification. Archives of General Psychiatry. 2008; 65:310–318. [PubMed: 18316677]

Eaves LJ, Long J, Heath AC. A theory of developmental change in quantitative phenotypes applied to cognitive development. Behavior Genetics. 1986; 16:143–162. DOI: 10.1007/BF01065484 [PubMed: 3707482]

Ebstein, RP.; Benjamin, J.; Belmaker, RH. Behavioral genetics, genomics, and personality. In: Plomin, R.; DeFries, JC.; Craig, IW.; McGuffin, P., editors. Behavioral Genetics in the Postgenomic Era. Washington, DC: American Psychological Association; 2003. p. 365-388.

Flint J, Munafò MR. The endophenotype concept in psychiatric genetics. Psychological Medicine. 2007; 37:163–180. DOI: 10.1017/S0033291706008750 [PubMed: 16978446]

Freathy RM, Kazeem GR, Morris RW, Johnson PC, Paternoster L, Ebrahim S, et al. Genetic variation at CHRNA5-CHRNA3-CHRNB4 interacts with smoking status to influence body mass index. International Journal of Epidemiology. 2011; 40:1617–1628. DOI: 10.1093/ije/dyr077 [PubMed: 21593077]

Gage SH, Smith GD, Zammit S, Hickman M, Munafò MR. Using Mendelian randomisation to infer causality in depression and anxiety research. Depression and Anxiety. 2013; 30:1185–1193. DOI: 10.1002/da.22150 [PubMed: 23847157]

Gizer IR, Ehlers CL. Genome-wide association studies of substance use: Considerations regarding populations and phenotypes. Biological Psychiatry. 2015; 77:423–424. DOI: 10.1016/j.biopsych.2014.11.013 [PubMed: 25645273]

Gizer, IR.; Otto, JM.; Ellingson, JM. Molecular Genetic Approaches to Studying 8 the Externalizing Spectrum. In: Beauchaine, TP.; Hinshaw, SP., editors. The Oxford Handbook of Externalizing Spectrum Disorders. Oxford, England: Oxford University Press; 2015. p. 125-148.

Gottesman II, Gould TD. The Endophenotype Concept in Psychiatry: Etymology and Strategic Intentions. American Journal of Psychiatry. 2003; 160:636–645. DOI: 10.1176/appi.ajp.160.4.636 [PubMed: 12668349]

Gottesman, II.; Shields, J. Schizophrenia and Genetics: A Twin Study Vantage Point. Oxford, England: Academic Press; 1972.

Gregory AM, Caspi A, Moffitt TE, Koenen K, Eley TC, Poulton R. Juvenile mental health histories of adults with anxiety disorders. American Journal of Psychiatry. 2007; 164:301–308. [PubMed: 17267794]

Guo G, Liu H, Wang L, Shen H, Hu W. The genome-wide influence on human BMI depends on physical activity, life course, and historical period. Demography. 2015; 52:1651–1670. DOI: 10.1007/s13524-015-0421-2 [PubMed: 26319003]

Hicks BM, Schalet BD, Malone SM, Iacono WG, McGue M. Psychometric and genetic architecture of substance use disorder and behavioral disinhibition measures for gene association studies. Behavior Genetics. 2011; 41:459–475. DOI: 10.1007/s10519-010-9417-2 [PubMed: 21153693]

Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nature Reviews Genetics. 2005; 6:95–108. DOI: 10.1038/nrg1521

Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. Genetics in Medicine. 2002; 4:45–61. DOI: 10.1097/00125817-200203000-00002 [PubMed: 11882781]

Holmes MV, Dale CE, Zuccolo L, Silverwood RJ, Guo Y, Ye Z, et al. Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data. BMJ: British Medical Journal. 2014; 349:g4164.doi: 10.1136/bmj.g4164 [PubMed: 25011450]

Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. American Journal of Psychiatry. 2010; 167:748–751. DOI: 10.1176/appi.ajp.2010.09091379 [PubMed: 20595427]

Kandel DB, Yamaguchi K, Chen K. Stages of progression in drug involvement from adolescence to adulthood: further evidence for the gateway theory. Journal of Studies on Alcohol. 1992; 53:447–457. DOI: 10.15288/jsa.1992.53.447 [PubMed: 1405637]

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S-y, Freimer NB, et al. Variance component model to account for sample structure in genome-wide association studies. Nature Genetics. 2010; 42:348–354. DOI: 10.1038/ng.548 [PubMed: 20208533]

Karlsson R, Graae L, Lekman M, Wang D, Favis R, Axelsson T, et al. MAGI1 copy number variation in bipolar affective disorder and schizophrenia. Biological Psychiatry. 2012; 71:922–930. DOI: 10.1016/j.biopsych.2012.01.020 [PubMed: 22381734]

Kendler KS. "A gene for…": the nature of gene action in psychiatric disorders. American Journal of Psychiatry. 2005; 162:1243–1252. DOI: 10.1176/appi.ajp.162.7.1243 [PubMed: 15994704]

Kendler KS. What psychiatric genetics has taught us about the nature of psychiatric illness and what is left to learn. Molecular Psychiatry. 2013; 18:1058–1066. DOI: 10.1038/mp.2013.50 [PubMed: 23628988]

Kendler KS, Neale MC, Heath AC, Kessler RC, Eaves LJ. A twin-family study of alcoholism in women. American Journal of Psychiatry. 1994; 151:707–715. DOI: 10.1176/ajp.151.5.707 [PubMed: 8166312]

Kessler RC, Chiu WT, Demler O, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Archives of General Psychiatry. 2005; 62:617–627. DOI: 10.1001/archpsyc.62.6.617 [PubMed: 15939839]

Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Survey. Archives of General Psychiatry. 1994; 51:8–19. DOI: 10.1001/archpsyc.1994.03950010008002 [PubMed: 8279933]

Klein, DN.; Riso, LP. Psychiatric disorders: problems of boundaries and comorbidity. In: Costello, CG., editor. Basic Issues in Psychopathology. New York, NY: Guilford Press; 1993. p. 19-66.

Krueger RF, Caspi A, Moffitt TE, Silva PA. The structure and stability of common mental disorders (DSM-III-R): a longitudinal-epidemiological study. Journal of Abnormal Psychology. 1998; 107:216–227. DOI: 10.1037/0021-843X.107.2.216 [PubMed: 9604551]

Lahey BB, Van Hulle CA, Singh AL, Waldman ID, Rathouz PJ. Higher-order genetic and environmental structure of prevalent forms of child and adolescent psychopathology. Archives of General Psychiatry. 2011; 68:181–189. DOI: 10.1001/archgenpsychiatry.2010.192 [PubMed: 21300945]

Lander ES, Schork NJ. Genetic dissection of complex traits. Science. 1994; 265:2037–2048. DOI: 10.1126/science.8091226 [PubMed: 8091226]

Lawlor DA, Nordestgaard BG, Benn M, Zuccolo L, Tybjaerg-Hansen A, Smith GD. Exploring causal associations between alcohol and coronary heart disease risk factors: findings from a Mendelian randomization study in the Copenhagen General Population Study. European Heart Journal. 2013; 34:2519–2528. DOI: 10.1093/eurheartj/eht081 [PubMed: 23492672]

Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics (Oxford, England). 2012; 28:2540–2542. DOI: 10.1093/bioinformatics/bts474

Lobo I. Pleiotropy: one gene can affect multiple traits. Nature Education. 2008; 1:10.

Maier R, Moser G, Chen GB, Ripke S, Coryell W, Potash JB, et al. Joint analysis of psychiatric disorders increases accuracy of risk prediction for schizophrenia, bipolar disorder, and major depressive disorder. The American Journal of Human Genetics. 2015; 96:283–294. DOI: 10.1016/j.ajhg.2014.12.006 [PubMed: 25640677]

Mannuzza S, Klein RG, Bessler A, Malloy P, LaPadula M. Adult psychiatric status of hyperactive boys grown up. American Journal of Psychiatry. 1998; 155:493–498. [PubMed: 9545994]

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature. 2009; 461:747–753. DOI: 10.1038/nature08494 [PubMed: 19812666]

Marceau K, Palmer RH, Neiderhiser JM, Smith TF, McGeary JE, Knopik VS. Passive rGE or developmental gene-environment cascade? An investigation of the role of xenobiotic metabolism genes in the association between smoke exposure during pregnancy and child birth weight. Behavior Genetics. 2016; 46:365–377. DOI: 10.1007/s10519-016-9778-2 [PubMed: 26803317]

McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nature Reviews Genetics. 2008; 9:356–369. DOI: 10.1038/nrg2344

McGue M, Zhang Y, Miller MB, Basu S, Vrieze S, Hicks B, et al. A genome-wide association study of behavioral disinhibition. Behavior Genetics. 2013; 43:363–373. DOI: 10.1007/s10519-013-9606-x [PubMed: 23942779]

Meyers J, Cerdá M, Galea S, Keyes K, Aiello AE, Uddin M, et al. Interaction between polygenic risk for cigarette use and environmental exposures in the Detroit neighborhood health study. Translational psychiatry. 2013; 3:e290.doi: 10.1038/tp.2013.63 [PubMed: 23942621]

Miller BJ, Buckley P, Seabolt W, Mellor A, Kirkpatrick B. Meta-analysis of cytokine alterations in schizophrenia: clinical status and antipsychotic effects. Biological Psychiatry. 2011; 70:663–671. DOI: 10.1016/j.biopsych.2011.04.013 [PubMed: 21641581]

Miller GA, Rockstroh B. Endophenotypes in psychopathology research: where do we stand? Annual review of clinical psychology. 2013; 9:177–213. DOI: 10.1146/annurev-clinpsy-050212-185540

Moffitt TE. Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. Psychological Review. 1993; 100:674–701. DOI: 10.1037/0033-295X.100.4.674 [PubMed: 8255953]

Moffitt TE, Caspi A, Harrington H, Milne BJ. Males on the life-course-persistent and adolescence-limited antisocial pathways: Follow-up at age 26 years. Development and Psychopathology. 2002; 14:179–207. [PubMed: 11893092]

Munafo M. Candidate gene studies in the 21st century: meta-analysis, mediation, moderation. Genes, Brain and Behavior. 2006; 5:3–8. DOI: 10.1111/j.1601-183X.2006.00188.x

Neale MC, Kendler KS. Models of comorbidity for multifactorial disorders. American Journal of Human Genetics. 1995; 57:935–953. [PubMed: 7573055]

Ormel J, Raven D, van Oort F, Hartman C, Reijneveld S, Veenstra R, et al. Mental health in Dutch adolescents: a TRAILS report on prevalence, severity, age of onset, continuity and co-morbidity of DSM disorders. Psychological Medicine. 2015; 45:345–360. DOI: 10.1017/S0033291714001469 [PubMed: 25066533]

Otowa T, Hek K, Lee M, Byrne E, Mirza S, Nivard M, et al. Meta-analysis of genome-wide association studies of anxiety disorders. Molecular Psychiatry. 2016; Advance online publication. doi: 10.1038/mp.2015.197

Paaby AB, Rockman MV. The many faces of pleiotropy. Trends in Genetics. 2013; 29:66–73. DOI: 10.1016/j.tig.2012.10.010 [PubMed: 23140989]

Plomin R. Genotype-environment correlation in the era of DNA. Behavior Genetics. 2014; 44:629–638. DOI: 10.1007/s10519-014-9673-7 [PubMed: 25195166]

Plomin R, Crabbe J. DNA. Psychological Bulletin. 2000; 126:806–828. DOI: 10.1037/0033-2909.126.6.806 [PubMed: 11107877]

Plomin, R.; DeFries, JC.; Knopik, VS.; Neiderheiser, J. Behavioral genetics. 6th. New York, NY: Worth; 2013.

Rhee SH, Willcutt EG, Hartman CA, Pennington BF, DeFries JC. Test of alternative hypotheses explaining the comorbidity between attention-deficit/hyperactivity disorder and conduct disorder. Journal of Abnormal Child Psychology. 2008; 36:29–40. DOI: 10.1007/s10802-007-9157-9 [PubMed: 17636436]

Rutter M. Pathways from Childhood to Adult Life. Journal of Child Psychology and Psychiatry. 1989; 30:23–51. DOI: 10.1111/j.1469-7610.1989.tb00768.x [PubMed: 2647779]

Salvatore JE, Aliev F, Bucholz K, Agrawal A, Hesselbrock V, Hesselbrock M, et al. Polygenic risk for externalizing disorders gene-by-development and gene-by-environment effects in adolescents and young adults. Clinical Psychological Science. 2014; 3:189–201. DOI: 10.1177/2167702614534211

Sanislow CA, Pine DS, Quinn KJ, Kozak MJ, Garvey MA, Heinssen RK, et al. Developing constructs for psychopathology research: research domain criteria. Journal of Abnormal Psychology. 2010; 119:631.doi: 10.1037/a0020909 [PubMed: 20939653]

Sivakumaran S, Agakov F, Theodoratou E, Prendergast JG, Zgaga L, Manolio T, et al. Abundant pleiotropy in human complex diseases and traits. The American Journal of Human Genetics. 2011; 89:607–618. DOI: 10.1016/j.ajhg.2011.10.004 [PubMed: 22077970]

Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? International Journal of Epidemiology. 2003; 32:1–22. DOI: 10.1093/ije/dyg070 [PubMed: 12689998]

Smoller JW. Disorders and borders: psychiatric genetics and nosology. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2013; 162:559–578. DOI: 10.1002/ajmg.b.32174

Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. Nature Reviews Genetics. 2013; 14:483–495. DOI: 10.1038/nrg3461

Sroufe LA, Rutter M. The domain of developmental psychopathology. Child Development. 1984; 55:17–29. DOI: 10.2307/1129832 [PubMed: 6705619]

Stephens M. A unified framework for association analysis with multiple related phenotypes. PLoS ONE. 2013; 8:e65245.doi: 10.1371/journal.pone.0065245 [PubMed: 23861737]

Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. Nature Reviews Genetics. 2012; 13:537–551. DOI: 10.1038/nrg3240

Tackett JL, Lahey BB, Van Hulle C, Waldman I, Krueger RF, Rathouz PJ. Common genetic influences on negative emotionality and a general psychopathology factor in childhood and adolescence. Journal of Abnormal Psychology. 2013; 122:1142–1153. DOI: 10.1037/a0034151 [PubMed: 24364617]

Thompson PM, Stein JL, Medland SE, Hibar DP, Vasquez AA, Renteria ME, et al. The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. Brain Imaging and Behavior. 2014; 8:153–182. DOI: 10.1007/s11682-013-9269-5 [PubMed: 24399358]

Torabi MR, Bailey WJ, Majd-Jabbari M. Cigarette smoking as a predictor of alcohol and other drug use by children and adolescents: evidence of the "gateway drug effect". Journal of School Health. 1993; 63:302–306. DOI: 10.1111/j.1746-1561.1993.tb06150.x [PubMed: 8246462]

Trzaskowski M, Harlaar N, Arden R, Krapohl E, Rimfeld K, McMillan A, et al. Genetic influence on family socioeconomic status and children's intelligence. Intelligence. 2014; 42:83–88. DOI: 10.1016/j.intell.2013.11.002 [PubMed: 24489417]

Turkheimer E, Haley A, Waldron M, D'Onofrio B, Gottesman II. Socioeconomic status modifies heritability of IQ in young children. Psychological Science. 2003; 14:623–628. DOI: 10.1046/j.0956-7976.2003.psci_1475.x [PubMed: 14629696]

Tyler AL, Asselbergs FW, Williams SM, Moore JH. Shadows of complexity: what biological networks reveal about epistasis and pleiotropy. Bioessays. 2009; 31:220–227. DOI: 10.1002/bies.200800022 [PubMed: 19204994]

van den Berg SM, de Moor MH, McGue M, Pettersson E, Terracciano A, Verweij KJ, et al. Harmonization of Neuroticism and Extraversion phenotypes across inventories and cohorts in the Genetics of Personality Consortium: an application of Item Response Theory. Behavior Genetics. 2014; 44:295–313. DOI: 10.1007/s10519-014-9654-x [PubMed: 24828478]

van Den Berg SM, de Moor MH, Verweij KJ, Krueger RF, Luciano M, Vasquez AA, et al. Meta-analysis of Genome-Wide Association Studies for Extraversion: Findings from the Genetics of Personality Consortium. Behavior Genetics. 2015; 46:170–182. DOI: 10.1007/s10519-015-9735-5 [PubMed: 26362575]

Vink JM, Hottenga JJ, Geus EJ, Willemsen G, Neale MC, Furberg H, et al. Polygenic risk scores for smoking: predictors for alcohol and cannabis use? Addiction. 2014; 109:1141–1151. DOI: 10.1111/add.12491 [PubMed: 24450588]

Vrieze SI, Iacono WG, McGue M. Confluence of genes, environment, development, and behavior in a post Genome-Wide Association Study world. Development and psychopathology. 2012; 24:1195–1214. DOI: 10.1017/S0954579412000648 [PubMed: 23062291]

Wray NR, Yang J, Hayes BJ, Price AL, Goddard ME, Visscher PM. Pitfalls of predicting complex traits from SNPs. Nature Reviews Genetics. 2013; 14:507–515. DOI: 10.1038/nrg3457

Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. The American Journal of Human Genetics. 2011; 88:76–82. DOI: 10.1016/j.ajhg.2010.11.011 [PubMed: 21167468]

Zhou X, Stephens M. Efficient algorithms for multivariate linear mixed models in genome-wide association studies. Nature Methods. 2014; 11:407–409. DOI: 10.1038/nmeth.2848 [PubMed: 24531419]