CHAPTER EIGHT

Opportunities for Bioinformatics in the Classification of Behavior and Psychiatric Disorders

Elissa J. Chesler[*,1], **Ryan W. Logan**[*,†]
[*]The Jackson Laboratory, Bar Harbor, Maine, USA
[†]University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA
[1]Corresponding author: e-mail address: elissa.chesler@jax.org

Contents

International Review of Neurobiology, Volume 104
ISSN 0074-7742
http://dx.doi.org/10.1016/B978-0-12-398323-7.00008-2

Abstract

A bioinformatics approach to behavioral neuroscience provides both unique opportunities and challenges for research on behavior. A major challenge has been to describe, define, and discriminate among abstract behavioral processes, in large part by distinguishing among the biological mechanisms of unique but not entirely discrete, entities of behavior. Understanding the complexity of neurobiology and behavior requires integration of data across diverse biological systems, types of data, and levels of scale. With the perspective and application of bioinformatics, we can uncover the relationships among these systems and take steps forward in realizing the common and distinct bases of psychiatric disease.

1. INTRODUCTION

As the final chapter in this volume on the informatics of behavior, we here, expand on the historical challenges of behavioral neuroscience to define, characterize, and classify psychiatric disorders, and elaborate on ways in which the tools and analyses of bioinformatics are able to advance behavioral investigation (Fig. 8.1). In psychiatry, classification schemes have been largely based on clusters of symptoms of seemingly related overt phenotypes. Unfortunately, for many disorders, the resulting diagnostic criteria provide poor classification with limited implications for research and therapeutics. An alarming reduction in investments in behavioral science by industry is a telling indicator of the challenges that have been faced in psychopharmacology (Brunner, Balci, & Ludvig, 2012) and calls for a pharmacologically relevant nosology have been made (Ban, 2006). Each of the preceding chapters highlights the diverse technologies and methods for multilevel data integration and large-scale data analysis that can be brought to bear in the application of bioinformatics to the intersection of behavioral neuroscience and psychiatry.

2. CURRENT CLASSIFICATION SYSTEMS IN PSYCHIATRY

2.1. Brief historical overview of psychiatric classification systems

Early psychiatric classifications, similar to other "medicalized" disorders, were based on broad aggregates of undesirable and maladaptive characteristics, or "habits." These systems essentially attributed abnormal behavior to individual responsibility. Classifications and diagnosis of psychiatric and

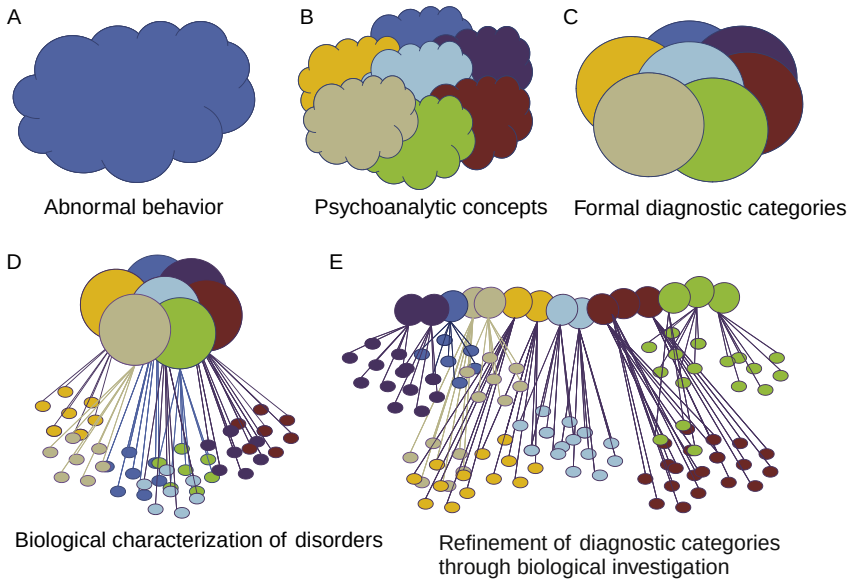

Figure 8.1 A schematic illustration of progress in classification of psychiatric disorders and the developing role of integrative bioinformatics. (A) Early conceptions of abnormal behavior later gave rise to (B) conceptual descriptions of disordered behavior. It is not until medicalization that these fuzzy concepts (represented by clouds) are set into (C) a formal classification scheme. (D) Modern biology has developed many means of identifying biological substrates (small circles) to these diagnostic categories (larger circles), but the categories themselves remain deeply heterogeneous and overlapping. (E) Emerging research efforts are using this data to test and modify disease categories.

behavioral disorders originally stem from qualitative interpretation of the maladies of individuals, tanamount to "folk psychology" (Slavney, 1992) reliant on the psychiatrists own "theory of mind." Folk psychology entails construction of the roles of thoughts and emotions in directing behavioral outcomes. These constructs are typically devoid of biological foundations and lack mechanistic description of how complex mechanisms interact across multiple biological and environmental levels to influence overt behaviors (Coley, 1995; Cosmides & Tooby, 1994; Stich & Ravenscroft, 1994). The eventual alignment of "mind" disorders with other health-related diseases seemed to be purely circumstantial and political (Hirshbein, 2011; Mack, Forman, Brown, & Frances, 1994). In part to aid the Census Bureau analysis, the first psychiatric classification system was formulated in 1918, consisting of 22 separate groups of mental disorders (American Medico-Psychological Association, 1918; Hyman, 2007; Sanders, 2011).

During and following World War II, the importance of psychiatric classification, diagnosis, and treatment reemerged. The first and second editions of the DSM (DSM-I,II) aimed to have clinical diagnostic utility (American Psychiatric Association, 1952). However, these manuals were formulated under the popular psychodynamic and psychoanalytic theoretical orientations of the era, and consequently diverged from biological explanations of psychiatric disorders (Hyman, 2007; Rogler, 1997). Dominated by subjective interpretations of disorders, unreliable diagnoses, and poor treatment outcomes, psychiatry required a complete theoretical transformation toward empirical bases (Feighner et al., 1972; Kendler, Munoz, & Murphy, 2010). The third and fourth editions of DSM were more standardized and comprehensive. The DSM-III represented a dramatic theoretical shift toward comprehensive dimensionality (American Psychiatric Association, 1987; Decker, 2007), and entailed descriptive criteria of discrete psychiatric categories accompanied by a multiaxial system (American Psychiatric Association, 2000). The multiaxial approach added comprehensive diagnostic value, although it has also further complicated diagnostic decisions. Importantly, it is the first attempt to systematically address issues of symptom overlap, disease heterogeneity, and comorbidity in clinical practice (Kawa & Giordano, 2012). As psychiatry strives to incorporate increased dimensionality, spectrums, and gradients of disorders in order to redefine, restructure, and reformulate categorization and diagnostic criteria (Regier, Narrow, Kuhl, & Kupfer, 2009; Robbins, Gillan, Smith, de Wit, & Ersche, 2012), a basis in neuroscience and integration across disciplines becomes essential to this process (Craddock & Owen, 2007; Hyman, 2007; Insel & Wang, 2010; Morris & Cuthbert, 2012).

2.2. Challenges to the classification of psychiatric disorders

With each revision of the DSM, there has been gradual refinement to both conceptual and diagnostic frameworks (e.g., Hilsenroth et al., 2000; Kotov et al., 2011; Kraemer, Kupfer, Clarke, Narrow, & Regier, 2012). However, this ever-evolving set of criteria and the perceived instability of psychiatric diagnoses have led to much criticism of current classification systems, even from within the field of psychiatry (Katschnig, 2010; Mayou, Kirmayer, Simon, Kroenke, & Sharpe, 2005; Miller, 2012). Discrete categories and constructs of psychiatric disorders are considered to be imprecise and arbitrary (Hyman, 2007), and extensive research indicates emotion, cognition, and behavior may better be conceptualized as dynamic

continuums (Varga, 2011). Classification of psychiatric disorders is also complicated by high comorbidity among various conditions. A particular disorder may be a risk factor for other disorders, such as anxiety and addiction (Goldman, Oroszi, & Ducci, 2005), or clusters of symptoms may commonly overlap between two or more distinct disorders, resembling similar underlying biological processes (Cerda, Sagdeo, Johnson, & Galea, 2010; Enoch, White, Waheed, & Goldman, 2008; Lawford, Young, Noble, Kann, & Ritchie, 2006; Molina et al., 2011). The interactions among a multitude of biological and environmental factors across development also produces varying degrees of "symptomology." Although technologies are advancing along with our understanding of complex psychiatric disorders, the exact method by which to incorporate findings on the roles of genetic, pathophysiological, and environmental factors into a clinically useful framework remains unclear (Enoch et al., 2008; Hyman, 2007; Lawford et al., 2006; Molina et al., 2011). The vague nature of psychiatric constructs renders alignment of biology and therapeutics with diagnostics quite challenging.

3. A BIOINFORMATICS APPROACH TO CLASSIFYING BEHAVIORAL AND PSYCHIATRIC DISORDERS

3.1. Challenges of understanding psychiatric disorders

Classification of psychiatric and behavioral disorders has been challenging because of the complexity and heterogeneity of the disorders. Ultimately, this results in difficulty naming and identifying discrete entities of behavioral function and presents challenges for research, diagnostics, and therapeutics. These challenges are eloquently described in a recent review of translational studies of alcohol use disorders (Crabbe, 2012).

Transforming the approach to psychiatric classification from primarily sociocultural and subjective externalities to the biological basis of disorders is expected to facilitate research, diagnosis, and treatment of mental illness (Craddock & Owen, 2010). Quite often these two approaches are seen as oppositional, although emerging evidence is beginning to bridge the gap between biomedical and psychological therapy, as well as behavioral interventions, even in animal models (Karpova et al., 2011). A recent commentary describes psychotherapy as an "epigenetic drug," modifying brain circuits and neurochemistry on the road to behavioral change (Stahl, 2012) and researchers are uncovering the biological effects of widely used

and newly developed cognitive behavioral therapies (Bryant et al., 2008; Davidson & McEwen, 2012; Huyser et al., 2012; Kobayashi et al., 2005).

Rapidly advancing tools in neuroscience and genetics enable a paradigm shift in psychiatric medicine toward a classification based on the biological foundations of behavioral disorders (Insel & Wang, 2010). Such an effort requires comprehensive association of behavioral variation to biological mechanism. The data-rich methods of modern high-throughput biological techniques coupled with bioinformatics methods can be brought to bear for developing comprehensive strategies for analyzing and integrating large volumes of biological data to define and classify psychiatric disorders.

3.2. Defining, characterizing, and redefining behavioral disorders

One can think of the research enterprise as creating an ontological framework, annotating information to the ontology, and testing the validity, reliability, coherence, and utility of the ontology itself. Well-defined, discrete entities are the anchors of conventional data integration strategies (see Volume 103, Chapter 2) and can serve to organize efforts to identify biological substrates of behavioral disorders. However, behavioral science has historically been challenged by its lack of discrete, reliably defined phenomena. Ontology development addresses this challenge with the creation of structured disease classification systems. Other research efforts create meaningful categories of disorders that can be used as the basis for classifying individuals, experiments, methods, and other aspects of research.

An ontology enables computational aggregation of data and knowledge around particular disease identities (see Volume 103, Chapter 5), behavioral processes, and model organism behavioral phenotypes (see Volume 103, Chapter 4), and thus enables assessment of the consistency, validity, and robustness of behavioral categories. These efforts also greatly facilitate endeavors to catalogue what is known about each of the underlying processes of behavior in model organism databases including the Mouse Genome Database (see Chapter 4), Rat Genome Database (see Chapter 2), and human disease resources including Online Mendelian Inheritance in Man; (Amberger, Bocchini, & Hamosh, 2011; McKusick, 2007) and NCBI's dbGAP database of genes and phenotypes (Mailman et al., 2007).

Bioinformatics techniques can transform our current understanding of psychiatric diseases by comprehensively integrating biological data with psychiatric diseases and behavioral characteristics. There are emerging basic and clinical research efforts put forth by National Institutes of Mental Health to

help define such characteristics, including the Research Domain Criteria (Morris & Cuthbert, 2012). The new developments in bioinformatics are poised to redefine disease classifications based on underlying biological entities and disease processes. This can be achieved through integration of large heterogeneous biological data sets. Convergent biological findings may be used to define essential biological processes underlying brain and behavior. The result may ultimately be a more holistic and precise approach to understand the relations among brain processes involved in psychiatric diseases.

3.3. The role of informatics

The scale and scope of the biological literature vastly exceeds that which can be mastered within even a single disease area (Fraser & Dunstan, 2010), much less to find patterns and organize or reorganize knowledge and disease frameworks by reading alone. High-throughput biological assays in basic and clinical research along with availability, diversity, and standardization of large-scale data resources will provide inputs to our understanding of the biology of complex diseases. For the research community to make full use of these data, they must be disseminated, harmonized (see Chapter 1) and integrated in a meaningful form for use by diverse investigators. Bringing together neuroscience resources in a computable form is a heroic task (Marenco, Nadkarni, Martone, & Gupta, 2007; Martone, Gupta, & Ellisman, 2004) but is only a step in the long process. A computational integration strategy must be applied, and for this, one needs a driving, falsifiable biological question or a methodological approach that results in falsifiable assertions. There have been promising developments in approaches that integrate data either semantically or through quantitative analysis of biological data as described extensively throughout 2012.

Due to the complex nature of psychiatric disease, those engaged in basic research and clinical medicine use an incredibly broad set of biological research tools. Modern bioinformatics has evolved well beyond its early roots in sequence analysis to embrace the challenge of deep data integration from functional genomics to a wealth of other areas, such as network modeling, functional and predictive biology, enabling representation and integration of biological knowledge. Comprehensive approaches combining content-rich biological quantitation and importantly, well-founded and informative behavioral phenotypes may some day provide comprehensive and deep systems biological analyses into mechanisms involved in thought, emotion, and behavior (Akil, Martone, & Van Essen, 2011; Markram, 2007).

4. THE BIOLOGICAL APPROACH TO UNDERSTANDING BEHAVIOR AND BEHAVIORAL DISORDERS

4.1. Finding the biological correlates of behavior

The fundamental objective and challenge of biological psychology has been to reliably map behavioral states and traits onto biological mechanisms and processes. Early philosophers could merely ponder the connections between biology, cognition, and behavior from the scant evidence presumably provided by gross injury. These unfortunate explorations eventually led to the recognition that the seat of thought and emotion belonged in the brain, not the heart. Human consciousness was considered far too abstract, and therefore, the mind was not subject to mechanistic laws of the universe despite awareness that the link between mind and body was the brain. Descartes considered this dualism an essential property separating humans from other animals, an idea that has implicit and explicit ramifications to this day.

However, comparative psychology took root, enabling the study of behavioral mechanisms through experimental biology. Charles Darwin provided evidence for phenotypic similarities between humans and other animals and the important role of inheritance. Support for the concept of inherited conservation of behavior ebbed and flowed throughout the nineteenth and twentieth centuries. It eventually found momentum in the parallel emergence of cognitive psychology and behavioral neuroscience (Fisch, 2007). Extensive research on human behavior began to reveal biological and genetic underpinnings of neuropsychiatric disorders. Under the assumptions that behavior could be reduced to interactions among brain mechanisms and that behavior in animals resembled aspects of human behaviors, a neuroscientific approach emerged. Early studies using animal models focused on consequences of lesions to specific brain areas on behavioral tasks. Animal models evolved from these early studies to comparatively sophisticated resources targeting genes, neural mechanisms, and interactions with environmental perturbations.

4.2. Animal models in psychiatric research and the importance of endophenotypes

Much modern research into the biological basis of behavioral disorders involves the use of model organisms from diverse species, including flies, mice, rats, and nonhuman primates. The advent of relatively low-cost massively parallel sequencing technologies has expanded this list. A host of less

well-developed non-conventional model organisms are now amenable for behavioral genomics study, including a growing number of insect species (Robinson et al., 2011). The power of these organisms for enabling insight into the biological mechanisms of behavior is virtually irrefutable, but great challenges lie in understanding whether and how the behaviors and biological correlates and mediators are indeed comparable to human characteristics. The major types of validity used to assess the translational mapping from animal to human phenotypes are face validity, which is the extent to which a behavior resembles the human condition, and pharmacological validity, which is the extent to which a behavior responds to existing pharmaceuticals used to treat a given human condition. A major challenge and goal is to steer away from symptom-based models that may have high face validity, but poor predictive and construct validity (Crabbe, 2012; Edwards & Koob, 2012). Animal models that are able to capture a single human behavioral disorder in its entirety are practically unattainable because the vast majority of disorders result from a tremendously complex, dynamic, interacting network of genetic, physiological, developmental, and environmental factors (Seong, Seasholtz, & Burmeister, 2002).

There is a growing appreciation and formalization of the concept that complex psychiatric and behavioral disorders can best be conceived as interacting and overlapping "building blocks," or endophenotypes, amenable to biological investigation through experimentally useful animal models for psychiatric and behavioral disorders (Crabbe, 2012; Gould & Gottesman, 2006; Kaffman & Krystal, 2012; Kalueff, Ren-Patterson, LaPorte, & Murphy, 2008; Robbins, 2012). Over 40 years ago, endophenotypes were described as "internal" phenotypes that can bridge gaps between available disease classifications with genetic and biological mechanisms of the disease process (Gottesman & Shields, 1973). This idea is a behavioral analogy to the concept of a "phenotypic profile," as a composite of "atomic" anatomical phenotypes for other kinds of disease and disorders (Washington et al., 2009). Endophenotypes are intended to represent putatively elementary phenomena, as opposed to behavioral macros, which provide a means for identification of associations between behavioral traits or states and genetic and other biological factors (Insel & Cuthbert, 2009), and may be distinguished from biological markers of the disorder (Gould & Gottesman, 2006). While identifying and defining robust endophenotypes may be in its infancy in psychiatry, there has been progress (e.g., Geyer, Olivier, Joels, & Kahn, 2012; Gotlib & Hamilton, 2012; Light et al., 2012; Matsuo et al., 2012; Nenadic, Gaser, & Sauer, 2012; Powell, Weber, &

Geyer, 2012). Several endophenotypes have been identified for affective and mood disorders (see Gould & Gottesman, 2006; Hasler et al., 2006). For example, circadian disruption is a common endophenotype among individuals with bipolar disorder (Murray & Harvey, 2010), which has been associated with several polymorphisms in circadian genes (McCarthy, Nievergelt, Kelsoe, & Welsh, 2012; McClung, 2011; Partonen, 2012) and preclinical models provided evidence these genes regulate specific neurochemical pathways to modulate behaviors with high face validity to the human condition (McClung, 2007; Mukherjee et al., 2010; Roybal et al., 2007). This translational approach has led to the development of intervention strategies and neuropharmacology targeting circadian mechanisms for the treatment of mood disorders (Arey & McClung, 2012; Coogan & Thome, 2011; Kozikowski et al., 2011).

4.3. The logic of double dissociation

Biological approaches to behavior, as many other biomedically relevant investigations, have the fundamental aim of identifying the biological substrate of disease. Learning which pathway and process is altered in normal versus abnormal states is the conventional challenge of those seeking new diagnostics and therapeutics in most areas of modern medicine. Behavioral science has a significant additional challenge that is to define which behaviors and diagnostic categories are discrete entities, subserved by distinct systems. Such categories are critical to the aggregation of information about the disorder, differential diagnosis, and alignment with therapeutic interventions. Yet, in behavioral science we recognize full well that many behavioral disorders may never be able to be defined as discrete entities.

Experimental psychology and behavioral neuroscience have formalized a process and logic for the identification of distinct psychological processes. The basic idea is to determine which behaviors are capable of being perturbed by distinct manipulations. The primary means by which this is performed is through double dissociation of behaviors by manipulation of distinct biological substrates. A disruption or lesion of one brain structure, cell type, or network should affect one process but not the other if the two processes are indeed distinct. Examples of this strategy include early work on different aphasias caused by lesions to specific speech and language processing centers, and the classic study by White and McDonald in which three distinct brain regions were mapped onto three distinct types of memory tasks (McDonald & White, 1993). Naturally, the earliest use of these

methods has literally relied on a highly modularized view of brain function, making the somewhat inappropriate analogy of brain structures to encapsulated organs of the rest of the body. However, more modern thinking about cognitive and other processes recognized that the substrate, whether it be discrete processing modules or a connected network, is the entity that must be dissociated to define distinct behavioral processes (Plaut, 1995).

4.4. Bioinformatics for behavioral classification

Due to the complexity and heterogeneity of behavioral disorders, investigation with both targeted mechanistic studies, and holistic, integrative strategies are required to relate basic biological foundations to behavior and clinical disorders. The double dissociation research strategy suggests a global approach to the classification of behavior that could account for the involvement and interaction among the multitude of biological and environmental factors influencing the development and trajectories of psychiatric disorders. In such an approach, large-scale comparison of behavioral processes and their associated biological substrates may be employed to test the legitimacy of psychiatric categories. Legitimate concern will most likely persist that a fundamentally biopsychosocial problem cannot be completely understood solely through a biological perspective. Most critics emphasize failures of the field manifest in poor treatment efficacy despite the identification of numerous neuropharmacological targets (McMahon & Insel, 2012). Understanding and classifying behavioral disorders through integrative biological research provides a targeted approach to the pathophysiology of neuropsychiatric illnesses (Binder & Ressler, 2012; Taber, Hurley, & Yudofsky, 2010).

The integration of diverse sets of data is critical to this endeavor, creating a unique challenge for bioinformatics and systems biology to provide the requisite tools to systematically harmonize data across studies and perform analyses that match sets of biological entities to sets of behavioral characters. This approach is experimentally intensive, but in principle can be performed *en masse* through the large-scale correlation of behaviors to global variation or manipulation of biological processes. The preceding chapters describe many of the requisite strategies, tools, and resources to perform the global association of biological substrate to behavior. Efforts such as the screening of behavioral phenotypes in model organisms with single gene perturbations, and to assess differential expression of thousands of genes following behavioral manipulations, enable the discovery of the biological

underpinnings. Systematic storage and comparison of these results in integrative systems such as GeneWeaver (Baker, Jay, Bubier, Langston, & Chesler, 2012) and by leveraging ontological description and similarity of phenotypes (Chen et al., 2012; Washington et al., 2009) describe ways in which these substrates can be compared.

5. DATA INTENSIVE METHODS FOR MAPPING BIOLOGICAL SUBSTRATE TO BEHAVIORAL FUNCTION

Genomics and bioinformatics present new technologies and experimental methods for the global mapping of biological substrates onto psychological functions and characteristics. Experimental technologies have rendered it feasible to measure the abundance of tens of thousands of biological molecules, image *in situ* the expression of transcripts in three-dimensional space, map large numbers of human functional images onto common coordinates, and enable integration of diverse experimental data types in semantic frameworks and data-mining enabled structures.

5.1. Gene annotation

Gene annotation properly refers to the identification of the boundaries of a coding region of a gene, its isoforms and structural variants, exons and introns, and regulatory sequences. These sequence analysis approaches are among the earliest function of bioinformatics and give rise to the important result of uniquely identified sequence features, typically aligned and displayed on a "genome browser," such as the UCSC Genome Browser (Karolchik, Hinrichs, & Kent, 2011) or Ensembl (Fernandez-Suarez & Schuster, 2010). Individual species databases make use of the GBrowse generic system (Stein et al., 2002) and the new updated version that is now available, JBrowse (Skinner, Uzilov, Stein, Mungall, & Holmes, 2009). Harmonizing the various identifiers for genes, gene products, and homologs within and across species for integrative analysis is an ongoing task of all model organism databases (see Chapter 1 and Volume 103, Chapter 7). Recent advances in sequencing have resulted in genome-wide sequence analysis across individuals (The 1000 Genomes Project Consortium, 2010), and strains of laboratory mouse (Danecek et al., 2012; Keane et al., 2011; Nellaker et al., 2012; Yalcin et al., 2011, 2012). These efforts will ultimately enable researchers to connect not just genes and gene products, but known sequence variants, isoforms, and other genomic and transcriptomic attributes with behavior. With these features in hand, one

has only a "biological parts list." Identifying, naming, characterizing, and cataloging the functions of these parts are the roles of many of the key resources in the bioinformatics of behavior.

5.2. Integrating model organism research and model organism databases

Few, if any, data-rich genomic methodologies are possible without extensive development and characterization of model organisms. Behavioral scientists make widespread use of various rodents, zebrafish, and drosophila, each of which have their own model organism databases built on the Generic Model Organism Databases system (O'Connor et al., 2008). Bult (See Chapter 4) and Shimoyama (see Chapter 2) describe data resources, applications, and curation issues in model organism databases for the mouse (Mouse Genome Database) and the rat (Rat Genome Database). These databases each present a single portal to genome centered data in an important laboratory species and include results of genetic mapping studies, gene annotations, homologies, anatomical, and functional characteristics that can be indexed to gene function and gene expression. Using behavioral phenotype ontologies including Mammalian Phenotype Ontology (Smith & Eppig, 2009) or the entity–quality based ontology (Gkoutos et al., 2004; Mungall et al., 2010), genes and gene products are associated to behavioral characters. Some of these characters bear great construct and face validity for neurobehavioral disorders, whereas others are species-specific behaviors and their importance to clinical investigation is unknown. Species-specific databases are a phenomenal resource for finding the role of genes annotated to behavioral processes, identifying mutant models of behavioral disorders, and finding overlapping genetic mapping results for related traits. Bridging the gap from model organisms to human in behavioral neuroscience will require functional mapping of gene and gene products across species.

5.3. Functional annotation in mutation screens

Another approach to associate genes to behavior is through systematic perturbation and functional characterization. Mutation, knockout, knockdown, and other techniques of gene manipulation have been in widespread use in individual labs. Results of these individual studies are meticulously curated to functionally annotate the genome (Knowlton et al.,

2008; Smith, Goldsmith, & Eppig, 2005) and can be found in the model organism databases, though primary data from these screens are often found in project-specific databases (Morgan et al., 2010). Several major efforts in mice have perturbed large catalogues of genes through mutagenesis (Bult et al., 2004; Goldowitz et al., 2004) and targeted deletion (Austin et al., 2004). This type of effort was pioneered in yeast but has been extended to mutagenesis of drosophila, and a variety of manipulations of the zebrafish (Bedell, Westcot, & Ekker, 2011; Clark, Urban, Skuster, & Ekker, 2011; Ekker et al., 2007; Gerlai, 2003; Klee, Ebbert, Schneider, Hurt, & Ekker, 2011; Petzold et al., 2009; Sivasubbu, Balciunas, Amsterdam, & Ekker, 2007). Individuals with these perturbations are then systematically screened for behavioral phenotypes to provide rapid gene-behavior annotation. Using phenotypic screens that are aligned to specific terms in behavioral ontologies, rapid, large-scale annotation is performed and "phenotypic alleles" are rapidly entered into model organism databases. However, the typical objective of these projects to study as many genes/constructs/individuals as possible has unfortunately led to a somewhat limited depth of behavioral analysis. Although a great many genes and/or strains are often characterized, very little is learned about the behaviors from the initial results. However, collections of models are identified for further functional study, and result in a host of new pathways to pursue, and recent studies demonstrate the utility of integrating focused collections of well-characterized mutants to find relations among phenotypes (Blednov, Mayfield, Belknap, & Harris, 2012).

5.4. Gene expression analysis

Gene expression analysis has been widely used to associate genes throughout the genome to various behavioral processes. Tissues collected following exquisitely defined behavioral manipulations or developmental stages are profiled in various brain regions. Laser capture microdissection and cell sorting provide techniques for gaining tremendous anatomical precision to these assays. The earliest studies made use of gene expression microarrays. Now, next-generation sequencing is an emerging technology of choice for whole transcriptome profiling, though many technical hurdles remain in the processing and analysis of sequencing data before the technique can be widely deployed. Initial results and stringent statistical filtering left some wondering whether transcriptomics provides more questions than answers, but later developments in profiling technology, experimental design,

and how biologists conceive of and interpret biological networks, has been improving the outlook for these techniques as part of a new powerful tool to complement biological investigation (see Chapter 5). Another use of sequence and gene expression array analysis is to quantify abundance of specific sequence in genomic DNA, enabling analysis of the number of copies of particular transcripts found across different individuals. This strategy has been employed to detect copy number variants in anxiety (Williams et al., 2009), alcoholism (Boutte et al., 2012), aggression in mice and humans (Velez, Sokoloff, Miczek, Palmer, & Dulawa, 2010; Vu, Coccaro, Eichler, & Girirajan, 2011), and autism (Cook, 2010; Pinto et al., 2010; van Daalen et al., 2011). The large number of experimental data sets generated by RNA quantitation has been aggregated in databases such as the Gene Expression Omnibus (Barrett & Edgar, 2006). Analytic approaches to integrate data across studies use statistical meta-analysis (Mulligan et al., 2006), combinatorics (Baker et al., 2009), Bayesian approaches (Guan, Ackert-Bicknell, Kell, Troyanskaya, & Hibbs, 2010). These approaches are aimed at finding the cohesive set of genes underlying shared processes in a purely empirical fashion, and using the methods of "guilt-by-association" and cognate approaches to identify genes which may share a function with previously well-characterized genes known to be involved in disease. Combined Bayesian-ontological similarity approaches, such as that recently described in Bauer, Kohler, Schulz, and Robinson (2012), hold great promise for leveraging semantic inference together with probabilistic modeling in the identification of the genetic basis of behavioral disease.

5.5. Finding the source of genetic variation in behavior

Finding genes that share a role in multiple related but seemingly disparate processes may help identify the mechanisms and nature of particular comorbidities. Using population genetics strategies including quantitative trait locus mapping in model organisms, genetic loci that modulate behavioral disorders have been mapped for a number of functions. Often related traits can be mapped to overlapping loci, but the loci are quite large (Flint, 2003) and it is thus challenging to refine to precise causal polymorphisms (Milner & Buck, 2010). This is clearly an area where integration of bioinformatic strategies and resources has been a critical complement to genetic strategies.

Human genetic analysis to find variants associated with behavioral disorders relies on the assessment of known genetic markers. The earliest studies attempted to identify regions of the genome that were linked to the

occurrence of the disorder in families. Association studies identify large cohorts of unrelated cases and controls to test the association of predicted causal variants with the disorder. With sufficiently large sample sizes, this method can be applied genome-wide in a systematic analysis of the correlation of allelic variants at specific loci to disease related phenotypes. Major collaborative efforts have been formed to study a host of behavioral traits including alcoholism and substance use disorders. Genome Wide Association Studies from the 1000 Genomes project have attempted to examine the "big five" personality traits (Liu et al., 2010; Terracciano et al., 2010). While there have been compelling successes, the cost of these studies has been of major concern. In particular, this is because the variants found do not account for much of the variation in behavior. This so-called missing heritability problem has many explanations, largely related to the simplicity of analysis methods relative to the complexity of genetic variation and the diversity of the human population. Deep sequencing of affected individuals for genes known to be involved in behavioral disorders enables the discovery of novel variants and segregating alleles that cause behavioral and other pathology in affected individuals. But, many have noted that a fundamental problem for psychiatric genetics may primarily be the challenge of classifying cases into discrete diagnostic categories, as described above.

Most of the techniques that provide an association of behavior to a gene or region of genome identify a limited number of putative regulators. Identifying and prioritizing them is facilitated by aggregation of functional information (Saccone, 2012). While these genetic strategies enable the discovery of biological mechanism of natural behavioral variation, they are generally less useful for learning about the relations among various disorders (Lee, Woon, Teo, & Sim, 2012). Compounding the problem is that the discovery of a genetic predictor of behavioral variation often leads to additional low-powered tests of the role of the same gene or locus in multiple other behaviors, creating a confusing array of associations rather than a holistic understanding of the shared role of endophenotypic processes in multiple behaviors.

5.6. Trait correlation, gene expression correlation and systems genetics

Complex behavioral traits are often inadequately assessed through a single dimension of behavior or a single experimental paradigm. Reference populations are panels of individuals that can be broadly profiled across many traits because the entire panel consists of a set of isogenic stocks that can be

reproduced and characterized indefinitely. Although this strategy had been in use long before widespread use of bioinformatics Web services, a wealth of mouse inbred strain data in the Mouse Phenotype Database (see Chapter 9) and GeneNetwork.org (see Chapter 12) system have enabled behavioral neuroscientists to correlate large numbers of measures across genetically diverse populations studied in multiple laboratories or environments, to identify traits which may share regulation by common genetic variants.

Genetic correlation of transcript to transcript, transcript to behavior, and among behavioral processes has given rise to an emerging marriage of systems biology and population genetics of behavior—systems genetics. Systems biologists construct large networks from high-throughput molecular and functional data to develop a causal model of a biological system. These networks consist of nodes, typically representing biological entities or measured traits, and edges, typically representing the causal or associative relations among the nodes. Systematic perturbation of the nodes enables one to test the relations among them and to define the direction of causality among them. Naturally occurring genetic polymorphisms provide genetic perturbations of biological systems. By constructing networks entirely out of trait correlations, one can identify a set of putative network nodes, while simultaneously identifying causal nodes, represented at quantitative trait loci.

The earliest of these systems-level studies were small and involved a limited number of mouse strains (Carter et al., 2001); larger studies using recombinant inbred populations (Chesler et al., 2003) have coupled QTL and microarray gene expression methods to identify the cause of variation in transcript abundance and complex traits. Studies by Hovatta et al. (2005) identified an anxiety related locus, *Glo1*, later found to be driven by a CNV (Williams et al., 2009).This approach has been extended to many species and populations (Kahsai & Zars, 2011; Morozova et al., 2009; Park et al., 2011) and the sophistication of the network modeling approach continues to increase.

Although use of genetic correlation to find common regulators of related brain and behavioral processes is promising, the results have thus far been mixed. A major challenge has been that among the inbred strains, few correlations among candidate related behaviors have been reported (e.g., Kliethermes & Crabbe, 2006). One explanation for this may be the inherent risk of face valid assays—a fundamental misalignment of model organism behavioral measures with the behavioral functions they assess. Another compelling explanation may be that the existing model organism populations in which much of this research has been conducted has been subject to

bottlenecking events, selection, and inbreeding depression, thereby limiting the range of behavioral variation and covariation detectable. Efforts to collect, create, and refine model organism populations for the improved execution of genetic mapping studies and systems genetics analysis are in progress. Some of these populations have already been shown to increase precision for genetic mapping (Philip et al., 2011) and gene co-expression analysis (Iancu et al., 2010, Iancu et al., 2012).

5.7. Integrative functional genomics and comparative psychology

Each of the techniques described above provides new ways to attach biological substrate to behavioral process in a global manner (Fig. 8.1D). Comparing these substrates within and among known behavioral processes can enable a test of the validity of classifications—to assess whether two named behaviors share a substrate or whether they are indeed functionally dissimilar (Fig. 8.1E). Convergence may also occur, resulting in the collapse of currently distinct categories or disorders. GeneWeaver.org (see Chapter 1) and its underlying Ontological Discovery Environment toolkit (Baker et al., 2009) provide an approach to perform this type of analysis from gene-set centered data. This system makes use of integrated functional genomics data across species and diverse experiments to find common substrates assumed to be related, whether they be a model organism phenotype and psychiatric disorder, facets of a single disorder, or two hypothetically distinct disorders. By matching data from empirical studies, the similarity of constructs can be tested, free of semantic knowledge and perspective biases. The concept and many of the analysis tools can be extended to other types of biological entities.

5.8. Challenges: Time, space, modularity

Most techniques in functional genomics, particularly when applied to brain and behavior are faced with the challenges of temporal and spatial resolution. The practicalities of most high-throughput technologies are such that only a discrete snapshot in time or space can be obtained. Decisions must be made regarding when in time samples are to be obtained relative to development, environmental exposure, and behavioral manipulation or other life history event. Likewise, equally challenging decisions are made about which organ, tissue, or compartment is to be obtained for characterization. The consequence is an unfortunate return to modularity in which gross brain regions

are studied with respect to particular functions, reminiscent of early "grind and bind" neurochemistry. Thus, despite the high granularity with which individual biological experiments are carried out, high-throughput experimentation and subsequent bioinformatic analysis relies necessarily on the extrapolation from snapshots in time and space. Increasingly, technologies and analysis methods are enabling an extension of high-throughput analytics to precise temporal and spatial events.

Resources like the Allen Brain Atlas (see Chapter 7) have provided a means for identifying transcripts highly expressed in particular regions, and tissues in which particular transcripts are highly expressed. This resource is greatly expanding in temporal resolution, focusing on the developing brain. Many new methods of time series analysis of gene expression will enable the specific association of gene expression in time and space. Measurement technologies in functional imaging, multiple unit electrophysiological recording, deep video analysis of behavior and other areas enable the temporal assessment and functional correlation of multiple behaviors, gene products, cells, or brain regions to enable interpretation of the relations among brain structures to one another and to processes of behavior. In the field of ontology development, there is a movement toward development of "application ontologies" which essentially obtain cross-products of "reference ontologies" such as those representing anatomy, cell type, time, and process to provide a computable description of the conditions under which annotated biological entities are associated to a particular set of terms. Recent efforts (Maynard, Mungall, Lewis, Imam, & Martone, 2012) hold promise for similarly describing and relating behavioral phenotypes to varying anatomical levels of granularity. In short, our technologies for measuring ensemble biological activity are rapidly improving along with the analysis methodologies required to track multiple measures in time and space. Bioinformatics techniques are being developed to represent, store, and share the information coming from these integrative studies, enabling a more comprehensive integration of the function of the central nervous system and how the aspects of central nervous system function relate to psychiatric conditions.

6. CONCLUSION: THE PROMISE OF RECONSTRUCTING BEHAVIOR THROUGH BIOLOGY

Bioinformatics and complementary advances in high-throughput assessment of brain and behavior have delivered technologies for rapidly identifying and characterizing the role of biological systems in behavioral

processes. This has enabled the discovery of new molecular targets for investigation, diagnostics, and therapeutics. While much of this work is in early stages, compelling advances are being made and translation to practice is already occurring. A major opportunity enabled by the application of bioinformatics to behavior is the potential reconstruction of behavior from its biological underpinnings. Ontology development creates a structure and platform for data integration to build from the ground up a data driven classification of behavioral processes. High-throughput biology generates data that can be rapidly annotated to these ontologies to provide a way to uncover relations such as the pleiotropy of gene action. Strategies to aggregate single findings enable researchers to make connections across disease related phenomena. These phenomena may not be the currently named diseases themselves, but rather the endophenotypes which form the robustly measurable aspects of disorders. The challenge is to transition from finding the substrate for the disorder to defining the disorder by its substrate.

Real experimental data, often collected using high-throughput measurement systems, provide the inputs and the validating data. The notion is that "real" classifications are those which can be found in the common biology of related disorders and compared to the distinct biology of distinct disorders. Bioinformatics presents an opportunity to define, categorize, and structure knowledge of psychiatric disorders and their component processes and features. Furthermore, and critical to any scientific endeavor, it presents a technology and framework that can be applied to test these structures, allowing falsifiability of the classification scheme.

ACKNOWLEDGMENTS

The authors gratefully acknowledge helpful suggestions by M. H. Haendel and funding from NIH RO1 AA18776 and The Jackson Laboratory.

REFERENCES

Akil, H., Martone, M. E., & Van Essen, D. C. (2011). Challenges and opportunities in mining neuroscience data. *Science*, *331*(6018), 708–712. http://dx.doi.org/10.1126/science.1199305.

Amberger, J., Bocchini, C., & Hamosh, A. (2011). A new face and new challenges for Online Mendelian Inheritance in Man (OMIM(R)). *Human Mutation*, *32*(5), 564–567. http://dx.doi.org/10.1002/humu.21466.

American Medico-Psychological Association. (1918). Statistical manual for the use of institutions for the insane. From National Committee for Mental Hygiene.

American Psychiatric Association, (1952). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association.

American Psychiatric Association, (1987). *Diagnostic and statistical manual of mental disorders—Revised* (3rd ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association, (2000). *Diagnostic and statistical manual of mental disorders—Text revision* (4th ed.). Washington, DC: American Psychiatric Association.

Arey, R., & McClung, C. A. (2012). An inhibitor of casein kinase 1 epsilon/delta partially normalizes the manic-like behaviors of the ClockDelta19 mouse. *Behavioural Pharmacology*, *23*(4), 392–396. http://dx.doi.org/10.1097/FBP.0b013e32835651fd.

Austin, C. P., Battey, J. F., Bradley, A., Bucan, M., Capecchi, M., Collins, F. S., et al. (2004). The knockout mouse project. *Nature Genetics*, *36*(9), 921–924. http://dx.doi.org/10.1038/ng0904-921.

Baker, E. J., Jay, J. J., Bubier, J. A., Langston, M. A., & Chesler, E. J. (2012). GeneWeaver: A web-based system for integrative functional genomics. *Nucleic Acids Research*, *40*(Database issue), D1067–D1076. http://dx.doi.org/10.1093/nar/gkr968.

Baker, E. J., Jay, J. J., Philip, V. M., Zhang, Y., Li, Z., Kirova, R., et al. (2009). Ontological Discovery Environment: A system for integrating gene-phenotype associations. *Genomics*, *94*(6), 377–387. http://dx.doi.org/10.1016/j.ygeno.2009.08.016.

Ban, T. A. (2006). Academic psychiatry and the pharmaceutical industry. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *30*(3), 429–441. http://dx.doi.org/10.1016/j.pnpbp. 2005.11.014.

Barrett, T., & Edgar, R. (2006). Gene expression omnibus: Microarray data storage, submission, retrieval, and analysis. *Methods in Enzymology*, *411*, 352–369. http://dx.doi.org/10.1016/S0076-6879(06)11019-8.

Bauer, S., Köhler, S., Schulz, M. H., & Robinson, P. N. (2012). Bayesian ontology querying for accurate and noise-tolerant semantic searches. *Bioinformatics*, *28*(19), 2502–2508. Epub 2012 Jul 26. PubMed PMID: 22843981; PubMed Central PMCID: PMC3463114.

Bedell, V. M., Westcot, S. E., & Ekker, S. C. (2011). Lessons from morpholino-based screening in zebrafish. *Briefings in Functional Genomics*, *10*(4), 181–188. http://dx.doi.org/10.1093/bfgp/elr021.

Binder, E. B., & Ressler, K. J. (2012). Emerging methods in the molecular biology of neuropsychiatric disorders. *Handbook of Clinical Neurology*, *106*, 191–209. http://dx.doi.org/10.1016/B978-0-444-52002-9.00011-5.

Blednov, Y. A., Mayfield, R. D., Belknap, J., & Harris, R. A. (2012). Behavioral actions of alcohol: Phenotypic relations from multivariate analysis of mutant mouse data. *Genes, Brain, and Behavior*, *11*(4), 424–435. http://dx.doi.org/10.1111/j.1601-183X.2012.00780.x.

Boutte, D., Calhoun, V. D., Chen, J., Sabbineni, A., Hutchison, K., & Liu, J. (2012). Association of genetic copy number variations at 11 q14.2 with brain regional volume differences in an alcohol use disorder population. *Alcohol*, *46*(6), 519–527. http://dx.doi.org/10.1016/j.alcohol.2012.05.002.

Brunner, D., Balci, F., & Ludvig, E. A. (2012). Comparative psychology and the grand challenge of drug discovery in psychiatry and neurodegeneration. *Behavioural Processes*, *89*(2), 187–195. http://dx.doi.org/10.1016/j.beproc.2011.10.011.

Bryant, R. A., Felmingham, K., Kemp, A., Das, P., Hughes, G., Peduto, A., et al. (2008). Amygdala and ventral anterior cingulate activation predicts treatment response to cognitive behaviour therapy for post-traumatic stress disorder. *Psychological Medicine*, *38*(4), 555–561. http://dx.doi.org/10.1017/S0033291707002231.

Bult, C., Kibbe, W. A., Snoddy, J., Vitaterna, M., Swanson, D., Pretel, S., et al. (2004). A genome end-game: Understanding gene function in the nervous system. *Nature Neuroscience*, 7(5), 484–485. http://dx.doi.org/10.1038/nn0504-484.

Carter, T. A., Del Rio, J. A., Greenhall, J. A., Latronica, M. L., Lockhart, D. J., & Barlow, C. (2001). Chipping away at complex behavior: transcriptome/phenotype correlations in the mouse brain. *Physiology & Behavior*, *73*(5), 849–857. PubMed PMID: 11566218.

Cerda, M., Sagdeo, A., Johnson, J., & Galea, S. (2010). Genetic and environmental influences on psychiatric comorbidity: A systematic review. *Journal of Affective Disorders*, *126*(1–2), 14–38. http://dx.doi.org/10.1016/j.jad.2009.11.006.

Chen, C. K., Mungall, C. J., Gkoutos, G. V., Doelken, S. C., Kohler, S., Ruef, B. J., et al. (2012). MouseFinder: Candidate disease genes from mouse phenotype data. *Human Mutation*, *33*(5), 858–866. http://dx.doi.org/10.1002/humu.22051.

Chesler, E. J., Wang, J., Lu, L., Qu, Y., Manly, K. F., & Williams, R. W. (2003). Genetic correlates of gene expression in recombinant inbred strains: a relational model system to explore neurobehavioral phenotypes. *Neuroinformatics*, *1*(4), 343–357. PubMed PMID: 15043220.

Clark, K. J., Urban, M. D., Skuster, K. J., & Ekker, S. C. (2011). Transgenic zebrafish using transposable elements. *Methods in Cell Biology*, *104*, 137–149. http://dx.doi.org/10.1016/B978-0-12-374814-0.00008-2.

Coley, J. D. (1995). Emerging differentiation of folkbiology and folkpsychology: Attributions of biological and psychological properties to living things. *Child Development*, *66* (6), 1856–1874.

Coogan, A. N., & Thome, J. (2011). Chronotherapeutics and psychiatry: Setting the clock to relieve the symptoms. *The World Journal of Biological Psychiatry*, *12*(Suppl. 1), 40–43. http://dx.doi.org/10.3109/15622975.2011.598389.

Cook, E. H., Jr. (2010). Autism spectrum disorder: Unbroken mirror neurons; rare copy number variants. *Autism Research*, *3*(4), 196–197. http://dx.doi.org/10.1002/aur.149.

Cosmides, L., & Tooby, J. (1994). Beyond intuition and instinct blindness: Toward an evolutionarily rigorous cognitive science. *Cognition*, *50*(1–3), 41–77.

Crabbe, J. C. (2012). Translational behaviour-genetic studies of alcohol: Are we there yet? *Genes, Brain, and Behavior*, *11*(4), 375–386. http://dx.doi.org/10.1111/j.1601-183X.2012.00798.x.

Craddock, N., & Owen, M. J. (2007). Rethinking psychosis: The disadvantages of a dichotomous classification now outweigh the advantages. *World Psychiatry*, *6*(2), 84–91.

Craddock, N., & Owen, M. J. (2010). The Kraepelinian dichotomy—Going, going . . . but still not gone. *The British Journal of Psychiatry*, *196*(2), 92–95. http://dx.doi.org/10.1192/bjp.bp.109.073429.

Danecek, P., Nellaker, C., McIntyre, R. E., Buendia-Buendia, J. E., Bumpstead, S., Ponting, C. P., et al. (2012). High levels of RNA-editing site conservation amongst 15 laboratory mouse strains. *Genome Biology*, *13*(4), r26. http://dx.doi.org/10.1186/gb-2012-13-4-r26.

Davidson, R. J., & McEwen, B. S. (2012). Social influences on neuroplasticity: Stress and interventions to promote well-being. *Nature Neuroscience*, *15*(5), 689–695. http://dx.doi.org/10.1038/nn.3093.

Decker, H. S. (2007). How Kraepelinian was Kraepelin? How Kraepelinian are the neo-Kraepelinians?–from Emil Kraepelin to DSM-III. *History of Psychiatry*, *18*(71 Pt 3), 337–360.

Edwards, S., & Koob, G. F. (2012). Experimental psychiatric illness and drug abuse models: From human to animal, an overview. *Methods in Molecular Biology*, *829*, 31–48. http://dx.doi.org/10.1007/978-1-61779-458-2_2.

Ekker, S. C., Stemple, D. L., Clark, M., Chien, C. B., Rasooly, R. S., & Javois, L. C. (2007). Zebrafish genome project: Bringing new biology to the vertebrate genome field. *Zebrafish*, *4*(4), 239–251. http://dx.doi.org/10.1089/zeb.2007.9979.

Enoch, M. A., White, K. V., Waheed, J., & Goldman, D. (2008). Neurophysiological and genetic distinctions between pure and comorbid anxiety disorders. *Depression and Anxiety*, *25*(5), 383–392. http://dx.doi.org/10.1002/da.20378.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Jr., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, *26*(1), 57–63.

Fernandez-Suarez, X. M., & Schuster, M. K. (2010). Using the ensembl genome server to browse genomic sequence data. *Current Protocols in Bioinformatics*, *chap. 1*, unit 1 15. http://dx.doi.org/10.1002/0471250953.bi0115s30.

Fisch, G. S. (2007). Animal models and human neuropsychiatric disorders. *Behavior Genetics*, *37*(1), 1–10. http://dx.doi.org/10.1007/s10519-006-9117-0.

Flint, J. (2003). Analysis of quantitative trait loci that influence animal behavior. *Journal of Neurobiology*, *54*(1), 46–77. PubMed PMID: 12486698.

Fraser, A. G., & Dunstan, F. D. (2010). On the impossibility of being expert. *British Medical Journal*, *341*, c6815.

Gerlai, R. (2003). Zebra fish: An uncharted behavior genetic model. *Behavior Genetics*, *33*(5), 461–468.

Geyer, M. A., Olivier, B., Joëls, M., & Kahn, R. S. (2012). From antipsychotic to anti-schizophrenia drugs: role of animal models. *Trends in Pharmacological Sciences*, *33*(10), 515–521. http://dx.doi.org/10.1016/j.tips.2012.06.006. Epub 2012 Jul 16. PubMed PMID: 22810174; PubMed Central PMCID: PMC3461097.

Gkoutos, G. V., Green, E. C., Mallon, A. M., Blake, A., Greenaway, S., Hancock, J. M., et al. (2004). Ontologies for the description of mouse phenotypes. *Comparative and Functional Genomics*, *5*(6–7), 545–551. http://dx.doi.org/10.1002/cfg.430.

Goldman, D., Oroszi, G., & Ducci, F. (2005). The genetics of addictions: Uncovering the genes. *Nature Reviews. Genetics*, *6*(7), 521–532. http://dx.doi.org/10.1038/nrg1635.

Goldowitz, D., Frankel, W. N., Takahashi, J. S., Holtz-Vitaterna, M., Bult, C., Kibbe, W. A., et al. (2004). Large-scale mutagenesis of the mouse to understand the genetic bases of nervous system structure and function. *Brain Research. Molecular Brain Research*, *132*(2), 105–115. http://dx.doi.org/10.1016/j.molbrainres.2004.09.016.

Gotlib, I. H., & Hamilton, J. P. (2012). Bringing genetics back to psychiatric endophenotypes. *Biological Psychiatry*, *71*(1), 2–3. http://dx.doi.org/10.1016/j.biopsych.2011.10.033.

Gottesman, I. I., & Shields, J. (1973). Genetic theorizing and schizophrenia. *The British Journal of Psychiatry*, *122*(566), 15–30.

Gould, T. D., & Gottesman, I. I. (2006). Psychiatric endophenotypes and the development of valid animal models. *Genes, Brain, and Behavior*, *5*(2), 113–119. http://dx.doi.org/10.1111/j.1601-183X.2005.00186.x.

Guan, Y., Ackert-Bicknell, C. L., Kell, B., Troyanskaya, O. G., & Hibbs, M. A. (2010). Functional genomics complements quantitative genetics in identifying disease-gene associations. *PLoS Computational Biology*, *6*(11), e1000991. http://dx.doi.org/10.1371/journal.pcbi.1000991.

Hasler, G., Drevets, W. C., Gould, T. D., Gottesman, I. I., & Manji, H. K. (2006). Toward constructing an endophenotype strategy for bipolar disorders. *Biological Psychiatry*, *60*(2), 93–105. Epub 2006 Jan 9. Review. PubMed PMID: 16406007.

Hilsenroth, M. J., Ackerman, S. J., Blagys, M. D., Baumann, B. D., Baity, M. R., Smith, S. R., et al. (2000). Reliability and validity of DSM-IV axis V. *The American Journal of Psychiatry*, *157*(11), 1858–1863.

Hirshbein, L. (2011). The American Psychiatric Association and the history of psychiatry. *History of Psychiatry*, *22*(87 Pt 3), 302–314.

Hovatta, I., Tennant, R. S., Helton, R., Marr, R. A., Singer, O., Redwine, J. M., et al. (2005). Glyoxalase 1 and glutathione reductase 1 regulate anxiety in mice. *Nature*, *438*(7068), 662–666. http://dx.doi.org/10.1038/nature04250.

Huyser, C., van den Heuvel, O. A., Wolters, L. H., de Haan, E., Boer, F., & Veltman, D. J. (2012). Increased orbital frontal gray matter volume after cognitive behavioural therapy in paediatric obsessive compulsive disorder. *The World Journal of Biological Psychiatry*. http://dx.doi.org/10.3109/15622975.2012.674215.

Hyman, S. E. (2007). Can neuroscience be integrated into the DSM-V? *Nature Reviews. Neuroscience*, *8*(9), 725–732. http://dx.doi.org/10.1038/nrn2218.

Iancu, O. D., Darakjian, P., Malmanger, B., Walter, N. A., McWeeney, S., & Hitzemann, R. (2012). Gene networks and haloperidol-induced catalepsy. *Genes, Brain, and Behavior*, *11*(1), 29–37. http://dx.doi.org/10.1111/j.1601-183X.2011.00736.x.

Iancu, O. D., Darakjian, P., Walter, N. A., Malmanger, B., Oberbeck, D., Belknap, J., et al. (2010). Genetic diversity and striatal gene networks: Focus on the heterogeneous stock-collaborative cross (HS-CC) mouse. *BMC Genomics*, *11*, 585. http://dx.doi.org/10.1186/1471-2164-11-585.

Insel, T. R., & Cuthbert, B. N. (2009). Endophenotypes: Bridging genomic complexity and disorder heterogeneity. *Biological Psychiatry*, *66*(11), 988–989. http://dx.doi.org/10.1016/j.biopsych.2009.10.008.

Insel, T. R., & Wang, P. S. (2010). Rethinking mental illness. *JAMA: The Journal of the American Medical Association*, *303*(19), 1970–1971. http://dx.doi.org/10.1001/jama.2010.555.

Kaffman, A., & Krystal, J. H. (2012). New frontiers in animal research of psychiatric illness. *Methods in Molecular Biology*, *829*, 3–30. http://dx.doi.org/10.1007/978-1-61779-458-2_1.

Kahsai, L., & Zars, T. (2011). Learning and memory in Drosophila: Behavior, genetics, and neural systems. *International Review of Neurobiology*, *99*, 139–167. http://dx.doi.org/10.1016/B978-0-12-387003-2.00006-9.

Kalueff, A. V., Ren-Patterson, R. F., LaPorte, J. L., & Murphy, D. L. (2008). Domain interplay concept in animal models of neuropsychiatric disorders: A new strategy for high-throughput neurophenotyping research. *Behavioural Brain Research*, *188*(2), 243–249. http://dx.doi.org/10.1016/j.bbr.2007.11.011.

Karolchik, D., Hinrichs, A. S., & Kent, W. J. (2011). The UCSC Genome Browser. *Current Protocols in Human Genetics*, *chap. 18*, unit 18.6. http://dx.doi.org/10.1002/0471142905.hg1806s71.

Karpova, N. N., Pickenhagen, A., Lindholm, J., Tiraboschi, E., Kulesskaya, N., Agustsdottir, A., et al. (2011). Fear erasure in mice requires synergy between antidepressant drugs and extinction training. *Science*, *334*(6063), 1731–1734. http://dx.doi.org/10.1126/science.1214592.

Katschnig, H. (2010). Are psychiatrists an endangered species? Observations on internal and external challenges to the profession. *World Psychiatry*, *9*(1), 21–28.

Kawa, S., & Giordano, J. (2012). A brief historicity of the Diagnostic and Statistical Manual of Mental Disorders: Issues and implications for the future of psychiatric canon and practice. *Philosophy, Ethics, and Humanities in Medicine*, *7*, 2. http://dx.doi.org/10.1186/1747-5341-7-2.

Keane, T. M., Goodstadt, L., Danecek, P., White, M. A., Wong, K., Yalcin, B., et al. (2011). Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature*, *477*(7364), 289–294. http://dx.doi.org/10.1038/nature10413.

Kendler, K. S., Munoz, R. A., & Murphy, G. (2010). The development of the Feighner criteria: A historical perspective. *The American Journal of Psychiatry*, *167*(2), 134–142. http://dx.doi.org/10.1176/appi.ajp.2009.09081155.

Klee, E. W., Ebbert, J. O., Schneider, H., Hurt, R. D., & Ekker, S. C. (2011). Zebrafish for the study of the biological effects of nicotine. *Nicotine & Tobacco Research*, *13*(5), 301–312. http://dx.doi.org/10.1093/ntr/ntr010.

Kliethermes, C. L., & Crabbe, J. C. (2006). Genetic independence of mouse measures of some aspects of novelty seeking. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(13), 5018–5023. http://dx.doi.org/10.1073/pnas.0509724103.

Knowlton, M. N., Li, T., Ren, Y., Bill, B. R., Ellis, L. B., & Ekker, S. C. (2008). A PATO-compliant zebrafish screening database (MODB): Management of morpholino knockdown screen information. *BMC Bioinformatics*, *9*, 7. http://dx.doi.org/10.1186/1471-2105-9-7.

Kobayashi, K., Shimizu, E., Hashimoto, K., Mitsumori, M., Koike, K., Okamura, N., et al. (2005). Serum brain-derived neurotrophic factor (BDNF) levels in patients with panic disorder: As a biological predictor of response to group cognitive behavioral therapy. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *29*(5), 658–663. http://dx.doi.org/10.1016/j.pnpbp. 2005.04.010.

Kotov, R., Ruggero, C. J., Krueger, R. F., Watson, D., Yuan, Q., & Zimmerman, M. (2011). New dimensions in the quantitative classification of mental illness. *Archives of General Psychiatry*, *68*(10), 1003–1011. http://dx.doi.org/10.1001/archgenpsychiatry.2011.107.

Kozikowski, A. P., Gunosewoyo, H., Guo, S., Gaisina, I. N., Walter, R. L., Ketcherside, A., et al. (2011). Identification of a glycogen synthase kinase-3beta inhibitor that attenuates hyperactivity in CLOCK mutant mice. *ChemMedChem*, *6*(9), 1593–1602. http://dx.doi.org/10.1002/cmdc.201100188.

Kraemer, H. C., Kupfer, D. J., Clarke, D. E., Narrow, W. E., & Regier, D. A. (2012). DSM-5: How reliable is reliable enough? *The American Journal of Psychiatry*, *169*(1), 13–15. http://dx.doi.org/10.1176/appi.ajp. 2011.11010050.

Lawford, B. R., Young, R., Noble, E. P., Kann, B., & Ritchie, T. (2006). The D2 dopamine receptor (DRD2) gene is associated with co-morbid depression, anxiety and social dysfunction in untreated veterans with post-traumatic stress disorder. *European Psychiatry*, *21*(3), 180–185. http://dx.doi.org/10.1016/j.eurpsy.2005.01.006.

Lee, K. W., Woon, P. S., Teo, Y. Y., & Sim, K. (2012). Genome wide association studies (GWAS) and copy number variation (CNV) studies of the major psychoses: What have we learnt? *Neuroscience and Biobehavioral Reviews*, *36*(1), 556–571. http://dx.doi.org/10.1016/j.neubiorev.2011.09.001.

Light, G. A., Swerdlow, N. R., Rissling, A. J., Radant, A., Sugar, C. A., Sprock, J., et al. (2012). Characterization of neurophysiologic and neurocognitive biomarkers for use in genomic and clinical outcome studies of schizophrenia. *PloS One*, *7*(7), e39434. http://dx.doi.org/10.1371/journal.pone.0039434.

Liu, J. Z., Tozzi, F., Waterworth, D. M., Pillai, S. G., Muglia, P., Middleton, L., et al. (2010). Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nature Genetics*, *42*(5), 436–440. http://dx.doi.org/10.1038/ng.572.

Mack, A. H., Forman, L., Brown, R., & Frances, A. (1994). A brief history of psychiatric classification. From the ancients to DSM-IV. *The Psychiatric Clinics of North America*, *17*(3), 515–523.

Mailman, M. D., Feolo, M., Jin, Y., Kimura, M., Tryka, K., Bagoutdinov, R., et al. (2007). The NCBI dbGaP database of genotypes and phenotypes. *Nature Genetics*, *39*(10), 1181–1186. http://dx.doi.org/10.1038/ng1007-1181.

Marenco, L., Nadkarni, P., Martone, M., & Gupta, A. (2007). Interoperability across neuroscience databases. *Methods in Molecular Biology*, *401*, 23–36. http://dx.doi.org/10.1007/978-1-59745-520-6_2.

Markram, H. (2007). Bioinformatics: Industrializing neuroscience. *Nature*, *445*(7124), 160–161. http://dx.doi.org/10.1038/445160a.

Martone, M. E., Gupta, A., & Ellisman, M. H. (2004). E-neuroscience: Challenges and triumphs in integrating distributed data from molecules to brains. *Nature Neuroscience*, *7*(5), 467–472. http://dx.doi.org/10.1038/nn1229.

Matsuo, K., Kopecek, M., Nicoletti, M. A., Hatch, J. P., Watanabe, Y., Nery, F. G., et al. (2012). New structural brain imaging endophenotype in bipolar disorder. *Molecular Psychiatry*, *17*(4), 412–420. http://dx.doi.org/10.1038/mp.2011.3.

Maynard, S. M., Mungall, C. J., Lewis, S. E., Imam, F., & Martone, M. E. (2012). A knowledge based approach to matching human neurodegenerative disease and animal models. *BMC Bioinformatics*. (in press).

Mayou, R., Kirmayer, L. J., Simon, G., Kroenke, K., & Sharpe, M. (2005). Somatoform disorders: Time for a new approach in DSM-V. *The American Journal of Psychiatry*, *162*(5), 847–855. http://dx.doi.org/10.1176/appi.ajp.162.5.847.

McCarthy, M. J., Nievergelt, C. M., Kelsoe, J. R., & Welsh, D. K. (2012). A survey of genomic studies supports association of circadian clock genes with bipolar disorder spectrum illnesses and lithium response. *PloS One*, *7*(2), e32091. http://dx.doi.org/10.1371/journal.pone.0032091.

McClung, C. A. (2007). Role for the Clock gene in bipolar disorder. *Cold Spring Harbor Symposia on Quantitative Biology*, *72*, 637–644. http://dx.doi.org/10.1101/sqb.2007.72.031.

McClung, C. A. (2011). Circadian rhythms and mood regulation: Insights from pre-clinical models. *European Neuropsychopharmacology*, *21*(Suppl. 4), S683–S693. http://dx.doi.org/10.1016/j.euroneuro.2011.07.008.

McDonald, R. J., & White, N. M. (1993). A triple dissociation of memory systems: Hippocampus, amygdala, and dorsal striatum. *Behavioral Neuroscience*, *107*(1), 3–22.

McKusick, V. A. (2007). Mendelian Inheritance in Man and its online version, OMIM. *American Journal of Human Genetics*, *80*(4), 588–604. http://dx.doi.org/10.1086/514346.

McMahon, F. J., & Insel, T. R. (2012). Pharmacogenomics and personalized medicine in neuropsychiatry. *Neuron*, *74*(5), 773–776. http://dx.doi.org/10.1016/j.neuron.2012.05.004.

Miller, G. (2012). Psychiatry. Criticism continues to dog psychiatric manual as deadline approaches. *Science*, *336*(6085), 1088–1089. http://dx.doi.org/10.1126/science.336.6085.1088.

Milner, L. C., & Buck, K. J. (2010). Identifying quantitative trait loci (QTLs) and genes (QTGs) for alcohol-related phenotypes in mice. *International Review of Neurobiology*, *91*, 173–204. http://dx.doi.org/10.1016/S0074-7742(10)91006-4.

Molina, E., Cervilla, J., Rivera, M., Torres, F., Bellon, J. A., Moreno, B., et al. (2011). Polymorphic variation at the serotonin 1-A receptor gene is associated with comorbid depression and generalized anxiety. *Psychiatric Genetics*, *21*(4), 195–201. http://dx.doi.org/10.1097/YPG.0b013e3283457a48.

Morgan, H., Beck, T., Blake, A., Gates, H., Adams, N., Debouzy, G., et al. (2010). EuroPhenome: A repository for high-throughput mouse phenotyping data. *Nucleic Acids Research*, *38*(Database issue), D577–D585. http://dx.doi.org/10.1093/nar/gkp1007.

Morozova, T. V., Ayroles, J. F., Jordan, K. W., Duncan, L. H., Carbone, M. A., Lyman, R. F., et al. (2009). Alcohol sensitivity in Drosophila: Translational potential of systems genetics. *Genetics*, *183*(2), 733–745. http://dx.doi.org/10.1534/genetics.109.107490.

Morris, S. E., & Cuthbert, B. N. (2012). Research Domain Criteria: Cognitive systems, neural circuits, and dimensions of behavior. *Dialogues in Clinical Neuroscience*, *14*(1), 29–37.

Mukherjee, S., Coque, L., Cao, J. L., Kumar, J., Chakravarty, S., Asaithamby, A., et al. (2010). Knockdown of Clock in the ventral tegmental area through RNA interference results in a mixed state of mania and depression-like behavior. *Biological Psychiatry*, *68*(6), 503–511. http://dx.doi.org/10.1016/j.biopsych.2010.04.031.

Mulligan, M. K., Ponomarev, I., Hitzemann, R. J., Belknap, J. K., Tabakoff, B., Harris, R. A., et al. (2006). Toward understanding the genetics of alcohol drinking through transcriptome meta-analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(16), 6368–6373. http://dx.doi.org/10.1073/pnas.0510188103.

Mungall, C. J., Gkoutos, G. V., Smith, C. L., Haendel, M. A., Lewis, S. E., & Ashburner, M. (2010). Integrating phenotype ontologies across multiple species. *Genome Biology*, *11*(1), R2. http://dx.doi.org/10.1186/gb-2010-11-1-r2.

Murray, G., & Harvey, A. (2010). Circadian rhythms and sleep in bipolar disorder. *Bipolar Disorders*, *12*(5), 459–472. http://dx.doi.org/10.1111/j.1399-5618.2010.00843.x.

Nellaker, C., Keane, T. M., Yalcin, B., Wong, K., Agam, A., Belgard, T. G., et al. (2012). The genomic landscape shaped by selection on transposable elements across 18 mouse strains. *Genome Biology*, *13*(6), R45. http://dx.doi.org/10.1186/gb-2012-13-6-r45.

Nenadic, I., Gaser, C., & Sauer, H. (2012). Heterogeneity of brain structural variation and the structural imaging endophenotypes in schizophrenia. *Neuropsychobiology*, *66*(1), 44–49. http://dx.doi.org/10.1159/000338547.

O'Connor, B. D., Day, A., Cain, S., Arnaiz, O., Sperling, L., & Stein, L. D. (2008). GMODWeb: A web framework for the Generic Model Organism Database. *Genome Biology*, *9*(6), R102. http://dx.doi.org/10.1186/gb-2008-9-6-r102.

Park, C. C., Gale, G. D., de Jong, S., Ghazalpour, A., Bennett, B. J., Farber, C. R., et al. (2011). Gene networks associated with conditional fear in mice identified using a systems genetics approach. *BMC Systems Biology*, *5*, 43.

Partonen, T. (2012). Clock gene variants in mood and anxiety disorders. *Journal of Neural Transmission*, *119*(10), 1133–1145. Epub 2012 Apr 27. PubMed PMID: 22538398.

Petzold, A. M., Balciunas, D., Sivasubbu, S., Clark, K. J., Bedell, V. M., Westcot, S. E., et al. (2009). Nicotine response genetics in the zebrafish. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(44), 18662–18667. http://dx.doi.org/10.1073/pnas.0908247106.

Philip, V. M., Sokoloff, G., Ackert-Bicknell, C. L., Striz, M., Branstetter, L., Beckmann, M. A., et al. (2011). Genetic analysis in the collaborative cross breeding population. *Genome Research*, *21*(8), 1223–1238. Epub 2011 Jul 6. PubMed PMID: 21734011; PubMed Central PMCID: PMC3149490.

Pinto, D., Pagnamenta, A. T., Klei, L., Anney, R., Merico, D., Regan, R., et al. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*, *466*(7304), 368–372. http://dx.doi.org/10.1038/nature09146.

Plaut, D. C. (1995). Double dissociation without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, *17*(2), 291–321. http://dx.doi.org/10.1080/01688639508405124.

Powell, S. B., Weber, M., & Geyer, M. A. (2012). Genetic models of sensorimotor gating: Relevance to neuropsychiatric disorders. *Current Topics in Behavioral Neurosciences*, *12*, 251–318. http://dx.doi.org/10.1007/7854_2011_195.

Regier, D. A., Narrow, W. E., Kuhl, E. A., & Kupfer, D. J. (2009). The conceptual development of DSM-V. *The American Journal of Psychiatry*, *166*(6), 645–650. http://dx.doi.org/10.1176/appi.ajp.2009.09020279.

Robbins, T. W. (2012). Animal models of neuropsychiatry revisited: A personal tribute to Teitelbaum. *Behavioural Brain Research*, *231*(2), 337–342. http://dx.doi.org/10.1016/j.bbr.2012.03.005.

Robbins, T. W., Gillan, C. M., Smith, D. G., de Wit, S., & Ersche, K. D. (2012). Neurocognitive endophenotypes of impulsivity and compulsivity: Towards dimensional psychiatry. *Trends in Cognitive Sciences*, *16*(1), 81–91. http://dx.doi.org/10.1016/j.tics.2011.11.009.

Robinson, G. E., Hackett, K. J., Purcell-Miramontes, M., Brown, S. J., Evans, J. D., Goldsmith, M. R., et al. (2011). Creating a buzz about insect genomes. *Science*, *331* (6023), 1386. http://dx.doi.org/10.1126/science.331.6023.1386.

Rogler, L. H. (1997). Making sense of historical changes in the Diagnostic and statistical manual of mental disorders: Five propositions. *Journal of Health and Social Behavior*, *38*(1), 9–20.

Roybal, K., Theobold, D., Graham, A., DiNieri, J. A., Russo, S. J., Krishnan, V., et al. (2007). Mania-like behavior induced by disruption of CLOCK. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(15), 6406–6411. http://dx.doi.org/10.1073/pnas.0609625104.

Sanders, J. L. (2011). A distinct language and a historic pendulum: The evolution of the Diagnostic and Statistical Manual of Mental Disorders. *Archives of Psychiatric Nursing*, *25*(6), 394–403. http://dx.doi.org/10.1016/j.apnu.2010.10.002.

Seong, E., Seasholtz, A. F., & Burmeister, M. (2002). Mouse models for psychiatric disorders. *Trends in Genetics*, *18*(12), 643–650.

Sivasubbu, S., Balciunas, D., Amsterdam, A., & Ekker, S. C. (2007). Insertional mutagenesis strategies in zebrafish. *Genome Biology*, *8*(Suppl. 1), S9. http://dx.doi.org/10.1186/gb-2007-8-s1-s9.

Skinner, M. E., Uzilov, A. V., Stein, L. D., Mungall, C. J., & Holmes, I. H. (2009). JBrowse: A next-generation genome browser. *Genome Research*, *19*(9), 1630–1638. http://dx.doi.org/10.1101/gr.094607.109.

Slavney, P. R. (1992). Belief and behavior: The role of "folk psychology" in psychiatry. *Comprehensive Psychiatry*, *33*(3), 166–172.

Smith, C. L., & Eppig, J. T. (2009). The mammalian phenotype ontology: Enabling robust annotation and comparative analysis. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine*, *1*(3), 390–399. http://dx.doi.org/10.1002/wsbm.44.

Smith, C. L., Goldsmith, C. A., & Eppig, J. T. (2005). The Mammalian Phenotype Ontology as a tool for annotating, analyzing and comparing phenotypic information. *Genome Biology*, *6*(1), R7. http://dx.doi.org/10.1186/gb-2004-6-1-r7.

Stahl, S. M. (2012). Psychotherapy as an epigenetic 'drug': Psychiatric therapeutics target symptoms linked to malfunctioning brain circuits with psychotherapy as well as with drugs. *Journal of Clinical Pharmacy and Therapeutics*, *37*(3), 249–253. http://dx.doi.org/10.1111/j.1365-2710.2011.01301.x.

Stein, L. D., Mungall, C., Shu, S., Caudy, M., Mangone, M., Day, A., et al. (2002). The generic genome browser: A building block for a model organism system database. *Genome Research*, *12*(10), 1599–1610. http://dx.doi.org/10.1101/gr.403602.

Stich, S., & Ravenscroft, I. (1994). What is folk psychology? *Cognition*, *50*(1–3), 447–468.

Taber, K. H., Hurley, R. A., & Yudofsky, S. C. (2010). Diagnosis and treatment of neuropsychiatric disorders. *Annual Review of Medicine*, *61*, 121–133.

Terracciano, A., Sanna, S., Uda, M., Deiana, B., Usala, G., Busonero, F., et al. (2010). Genome-wide association scan for five major dimensions of personality. *Molecular Psychiatry*, *15*(6), 647–656. http://dx.doi.org/10.1038/mp.2008.113.

Terracciano, A., Tanaka, T., Sutin, A. R., Sanna, S., Deiana, B., Lai, S., et al. (2010). Genome-wide association scan of trait depression. *Biological Psychiatry*, *68*(9), 811–817. http://dx.doi.org/10.1016/j.biopsych.2010.06.030.

The 1000 Genomes Project Consortium, (1000). A map of human genome variation from population-scale sequencing. *Nature*, *467*(7319), 1061–1073. http://dx.doi.org/10.1038/nature09534.

van Daalen, E., Kemner, C., Verbeek, N. E., van der Zwaag, B., Dijkhuizen, T., Rump, P., et al. (2011). Social Responsiveness Scale-aided analysis of the clinical impact of copy number variations in autism. *Neurogenetics*, *12*(4), 315–323. http://dx.doi.org/10.1007/s10048-011-0297-2.

Varga, S. (2011). Defining mental disorder. Exploring the 'natural function' approach. *Philosophy, Ethics, and Humanities in Medicine*, *6*, 1. http://dx.doi.org/10.1186/1747-5341-6-1.

Velez, L., Sokoloff, G., Miczek, K. A., Palmer, A. A., & Dulawa, S. C. (2010). Differences in aggressive behavior and DNA copy number variants between BALB/cJ and BALB/cByJ substrains. *Behavior Genetics*, *40*(2), 201–210. http://dx.doi.org/10.1007/s10519-009-9325-5.

Vu, T. H., Coccaro, E. F., Eichler, E. E., & Girirajan, S. (2011). Genomic architecture of aggression: Rare copy number variants in intermittent explosive disorder. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *156B*(7), 808–816. http://dx.doi.org/10.1002/ajmg.b.31225.

Washington, N. L., Haendel, M. A., Mungall, C. J., Ashburner, M., Westerfield, M., & Lewis, S. E. (2009). Linking human diseases to animal models using ontology-based phenotype annotation. *PLoS Biology*, 7(11), e1000247. http://dx.doi.org/10.1371/journal.pbio.1000247.

Williams, R. T., Lim, J. E., Harr, B., Wing, C., Walters, R., Distler, M. G., et al. (2009). A common and unstable copy number variant is associated with differences in Glo1 expression and anxiety-like behavior. *PloS One*, *4*(3), e4649. http://dx.doi.org/10.1371/journal.pone.0004649.

Yalcin, B., Wong, K., Agam, A., Goodson, M., Keane, T. M., Gan, X., et al. (2011). Sequence-based characterization of structural variation in the mouse genome. *Nature*, *477*(7364), 326–329. http://dx.doi.org/10.1038/nature10432.

Yalcin, B., Wong, K., Bhomra, A., Goodson, M., Keane, T. M., Adams, D. J., et al. (2012). The fine-scale architecture of structural variants in 17 mouse genomes. *Genome Biology*, *13*(3), R18. http://dx.doi.org/10.1186/gb-2012-13-3-r18.