Published in final edited form as:

Bioessays. 2014 June ; 36(6): 606–616. doi:10.1002/bies.201300147.

Molecular network analysis enhances understanding of the biology of mental disorders

Kay S. Grennan[±], Chao Chen[±], Elliot S. Gershon, and Chunyu Liu^{*}

Department of Psychiatry, University of Illinois at Chicago, Chicago, IL 60607, USA

Summary

We provide an introduction to network theory, evidence to support a connection between molecular network structure and neuropsychiatric disease, and examples of how network approaches can expand our knowledge of the molecular bases of these diseases. Without systematic methods to derive their biological meanings and inter-relatedness, the many molecular changes associated with neuropsychiatric disease, including genetic variants, gene expression changes and protein differences, present an impenetrably complex set of findings. Network approaches can potentially help integrate and reconcile these findings, as well as provide new insights into the molecular architecture of neuropsychiatric diseases. Network approaches to neuropsychiatric disease are still in their infancy, and we discuss what might be done to improve their prospects.

Keywords

mental disorders; molecular networks; neuropsychiatric disease; systems biology

Introduction

Decades of research have been devoted to identifying the molecular bases of neuropsychiatric diseases. The heritabilities of these diseases have been established [1]. The diseases and many of their traits have been shown to be associated with numerous alterations at the molecular level; these alterations range from genetic variations to cell signaling changes. However, the relationships among these changes have proven difficult to piece together, and attempts to predict disease phenotypes from genotypes have been only modestly successful.

Systems biology is a holistic approach to studying biological systems. Rather than studying disease-associated molecular changes in isolation, systems biology approaches try to identify changes within their larger context. Networks are an important tool for modeling the molecular interactions that make up these complex systems. These approaches have been widely used in studies of complex diseases, particularly cancer. Cancer research has led other research areas in the utilization of advanced technologies that generate a huge amount

^{*}Corresponding author: Chunyu Liu, liucy@uic.edu.

[±]Co-first authors

of data in functional genomics and proteomics, including genome-wide studies of genetic variation, gene expression, protein-protein interactions and so on. The National Cancer Institute has realized the importance of data integration and established dozens of Centers of Cancer Systems Biology. One way to integrate these data is to find the relationships among these molecules using network approaches. As a result, most of the network data available today are cancer-related.

Neuropsychiatric diseases share some important features with cancer, such as their multi-genic or polygenic nature, heterogeneous causes and complex phenotypes. Some systems biology tools such as network approaches, which have been successfully used in cancer, could be adapted for neuropsychiatric disease studies. However, network approaches have not yet been fully embraced by researchers in these diseases. Therefore, we discuss how network approaches can help integrate and reconcile previous findings, and how these approaches can pave the way for understanding the complex molecular mechanisms underlying these diseases.

First, we introduce the reader to some basic concepts in network theory. We briefly review the kind of data that is used for network construction, and how neuropsychiatric disorders have already been shown to be associated with molecular network structure. We describe approaches to characterizing molecules by exploiting network structure and approaches to identifying molecular interactions relevant to neuropsychiatric disorders. We demonstrate how molecular interactions can be connected back to their regulators using network approaches. We further discuss the potential utilization of network characterization in the classification and treatment of neuropsychiatric disorders, and matters that should be addressed so network approaches to identifying the molecular bases of neuropsychiatric disorders can move forward.

Introduction to molecular networks

Basic concepts

In a molecular network, the **nodes** or vertices represent molecules, which can be genes, RNA, proteins and metabolites, and the **edges** represent relationships between the nodes, which can be direct physical interactions or other indirect relationships, such as a regulatory relationship (Fig. 1A). In a protein-protein interaction network, the nodes are proteins and the edges physical interactions between them. In a gene regulation network, the nodes are regulators (transcription factors) and effectors (the target genes), and the edges represent transcription regulation. Other networks include gene co-expression networks, metabolic networks, cell signaling networks and so on. These are all examples of *molecular* networks: the nodes are molecules and the edges represent some kind of relationship between two molecules.

A example of a PPI network from psychiatry is that of Camargo et al. [2], who used a yeast two-hybrid (Y2H) assay [3] to identify protein binding partners of the DISC1 protein in human fetal brain. *DISC1*, or disrupted-in-schizophrenia-1, had previously been identified as a schizophrenia (SZ) risk gene because a balanced translocation disrupting it co-segregated with psychiatric illness in a Scottish family [4]. Its function, however, was not well-

characterized, and it was not among the top associations in SZ genome-wide association studies (GWASs). Camargo et al. did two rounds of Y2H assays. First, they identified 286 proteins that interact directly with DISC1. Next, they identified proteins that interact with a subset of those DISC1-interacting proteins. Finally, they identified any previously reported interactions involving the identified proteins. All these interactions were used to construct a PPI network. More proteins in the network than would be expected by chance were involved in cytoskeletal organization and biogenesis, mRNA/protein synthesis, cell cycle/division, intracellular transport and signal transduction processes. This functional bias of the DISC1 PPI network implicated DISC1 in those functions.

This implication leads us to a fundamental property of molecular networks: modularity. A **module** is a sub-network of highly interconnected nodes (Fig. 1B) that is relatively sparsely connected to the larger network (Fig. 1C). Molecules in the same module are more likely to interact and/or to play roles in the same function --a function that may be compromised in disease [5]. Conversely, molecules that are associated with a disease or a set of similar diseases are more likely to interact [6]. Either way, the principle --referred to as guilt-by-association -- has been generally well-supported, including evidence from co-expression networks [7], though it is not necessarily universal (see Bogdanov et al. [8] below, and [9]).

Modularity is pertinent to disease studies for two major reasons: first, working with modules, rather than individual genes, reduces the dimensionality of the data; for example, modules from a co-expression network can be tested for association with disease, rather than every individual gene, which reduces the multiple-testing burden. Second, modules provide a set of genes that are highly likely to be functionally related in one or more contexts, which can be exploited in numerous ways, making a module of genes a more biologically meaningful unit in relation to disease.

Another fundamental property of complex biological networks, related to their modularity, is that relatively few of their nodes are highly connected to other nodes, while most of their nodes only connect to a few other nodes; the highly-connected nodes are referred to as **hubs** (Fig. 1B). They approximate scale-free networks, meaning that their degree distribution -- where degree is the number of edges that a node touches -- approximately follows a power law. In the literature, they are often referred to as actually being scale-free. However, this and other properties thought to be universal properties of biological networks -- based on observations of early networks -- are being re-evaluated as more data are collected and network accuracy improves [10, 11]. These assumptions affect network analysis: for example, when scale-freeness is used as a criterion in selecting one network from a range of possible networks (e.g. [12]).

How molecular networks are constructed: up from the bottom and down from the top

The Camargo et al. paper can be thought of as a “bottom-up” approach to defining molecular networks, where pair-wise molecular interactions are identified and pieced together to form a network. Previously reported interactions can be retrieved, individually or already assembled into networks, from databases, including curated projects such as Kyoto Encyclopedia of Genes and Genome (KEGG) (www.genome.jp/kegg), GeneMANIA

(www.genemania.org), BioGRID (thebiogrid.org/), and PINA (cbg.garvan.unsw.edu.au/pina/) or commercial tools such as Ingenuity's Knowledge Base (www.ingenuity.com).

However, networks can also be constructed in a top-down manner, using genome-wide data, a.k.a. “omics” data. This method is unbiased relative to a bottom-up approach, which depends on which interactions have already been identified and reported. A top-down example is our study of post-mortem brain tissue genome-wide expression data[13]: we defined a gene co-expression network based on the pair-wise Pearson correlations between expression levels of genes, and then identified modules of co-expressed genes via a clustering algorithm, per the Weighted Gene Co-expression Network Analysis method of Zhang & Horvath [12]. Top-down networks are statistically inferred from the data, using one of a number of different network models, including hierarchical clustering [14, 15] and random forest [16]. Principal component analysis (PCA) and independent component analysis (ICA), amongst others, are powerful tools for retrieving underlying biological features from complex networks.

Top-down network approaches for other data types have lagged behind those for co-expression [17], because genome-wide, high-throughput assays for collecting these data have developed more slowly. Such assays exist for protein-protein [18–23], protein-DNA [24], protein-RNA interactions [25–30]; however, unlike expression microarray assays, they are still not easily employed for specific tissues like brain or to capture population variation (see Box 1). Currently, epigenetic markers, except for DNA methylation, generally appear as discrete variables in networks, i.e. presence or absence of a particular histone mark. As these technologies improve, the narrow focus of network construction should broaden to include many non-genetic variables.

Neuropsychiatric diseases are associated with molecular network structure

There are two primary network approaches to studying neuropsychiatric diseases. The first focuses on mapping known disease genes, whether GWAS results, candidate genes, differently-expressed genes, or rare disease-associated variants, onto previously identified molecular networks. The second is based on networks built from the top-down with genome-wide data, and is used to identify case-control differences in networks.

The first approach is particularly important because it explains some of the genetic heterogeneity observed in neuropsychiatric diseases. For example, a PPI network was constructed from the protein products of genes associated with major depression (MD) via GWAS [31], showing that these disease-associated genes are functionally-related. The same general design has been used for autism (AUT) [32, 33], SZ [34, 35] and BD [33]. Ben-David & Shifman [36] performed a weighted gene co-expression network analysis (WGCNA) on post-mortem brain tissue from neuropsychiatrically normal adult patients, and identified three modules containing genes associated with neuronal functions. They found two of those neuronal modules were enriched with both rare mutations and common variants associated with AUT. This same general principle has also been tested and observed in SZ [37, 38] and in other AUT studies [39–42]. It has recently been applied to fetal and early post-natal expression data from multiple brain layers; co-expression network analyses of

these spatiotemporally-specific data suggest cortical glutamatergic neurons during fetal development may be relevant to AUT pathophysiology [41, 42].

The second network approach demonstrates genome-wide network organization. WGCNA has been used to demonstrate that the brain transcriptome is organized in a network manner: functional modules correspond to brain cell types including neurons, oligodendrocytes, astrocytes and microglia [43]. These functional modules may be up- or down-regulated during cortical development [41]. Module topologies – i.e. module members and edges – have been found to be fairly consistent in studies of SZ [13], BD [13] and AUT [44]. This means that there appears to be little change in the pair-wise correlations of gene expression levels within modules. However, there were case-control differences: entire co-expression modules were up- or down-regulated in these SZ, BD and AUT studies. In contrast, in Alzheimer's disease (AD) it was the module topologies that were altered: loss of modules and gain of modules were observed [45].

Practical applications of the network approach to the study of neuropsychiatric disease

Network approaches can help characterize and prioritize genes for disease studies

Gene annotation—Classically, genes have been annotated by experimentation or by sequence similarity to known genes. Networks can be efficiently used to annotate genes and molecules, either functionally or for disease association. Shared membership in a module is a useful annotation in and of itself. As for more standard annotations, Bogdanov et al. [46] described four basic kinds of annotation methods using networks: direct, module-directed, probabilistic and pattern-based approaches. Direct methods propagate annotation information from neighbor to neighbor. Module-directed methods use the annotations of other module genes to infer the function of an unannotated module gene, similar to what Camargo et al. [2] did qualitatively. Probabilistic methods are based on Markov Random Fields and are designed to assess annotations for all network nodes at once. All three are based on the assumption that network neighbors are likely to share annotations.

Bogdanov et al. [46] have found that this assumption holds truer for genes with annotations from the “biological process” or “cellular localization” domains of the Gene Ontology (GO) than for the lower-order “molecular functions” domain. Pattern-based annotation approaches are an alternative intended to address this disparity. They are based on the hypothesis that genes with similar functions will have similar patterns of annotations in their immediate network neighborhoods; for these methods, the distances between the nodes within the network do not matter, and neither does sparse annotation of a node's immediate neighbors. Bogdanov & Singh [8] proposed and demonstrated such a method.

Prioritizing candidate genes—Candidate genes are generally identified based on the strength of their linkage or association signals in genetic studies, on their biological feasibility, or on their similarity to or interaction with known disease genes. However, the resulting gene lists can be very long. For example, linkage regions can harbor hundreds of genes. Even an associated single nucleotide polymorphism (SNP) can implicate multiple

adjacent genes. Network-based ranking algorithms can aid in prioritizing one gene over another for biological validation by adding another layer of information to these classical approaches, specifically, the network context of the gene under consideration.

The simpler gene prioritization algorithms are based on the direct protein interaction partners of a candidate gene's protein product. Franke et al. [47] scored and ranked genes based on how many known disease genes with which they interacted in the PPI network, while Lage et al.'s [48] method differed in that it also incorporated interactions with genes known to be associated with other, similar phenotypes. Information flow methods are not limited to directly interacting nodes; they can take into account a candidate gene's proximity to all known disease genes ([49, 50]). Using a Google PageRank-related algorithm, Lee et al. [51] ranked GWAS results by combining a "guilt-by-association" score and an association score calculated from GWAS data.

Doncheva et al. [52] reviewed gene selection methods and showed that PPI networks are among the most powerful data sources for prioritizing candidate genes. They also observed that creating tissue-specific PPI networks – by filtering out PPIs that include products of genes not expressed in a given tissue – improved prioritization (e.g. [53]), as did direct incorporation of disease-differential gene expression data (e.g. [54]). This is unsurprising, given that disease genes are more likely to have tissue-specific functions [5]. It supports the importance of developing brain-specific networks for neuropsychiatric disease studies

An alternative, qualitative, method for candidate gene prioritization is to prioritize intra-modular hub genes from disease-associated modules. For example, a recent WGCNA analysis [55], comparing mouse and human transcriptional organization, identified a human-specific brain module associated with various measures of AD progression; the functions of three of the module's four hub genes were unknown, and could be prioritized for biological validation. Langfelder et al. [56] evaluated the effectiveness of this approach relative to other methods.

Molecular interactions are associated with neuropsychiatric diseases

As well as characterizing individual molecules, network analyses can be used to characterize the interactions among molecules and how they are associated with disease. One method is exemplified by the studies described above, in which WGCNA was used to identify co-expression modules associated with neuropsychiatric diseases AUT [44], SZ [13], and BD [13]. The disease associations are based on differential expression of the modules' 'eigengenes' in cases relative to controls; the eigengene represents the overall expression profile of a module. Such an approach not only greatly reduces the dimensionality of the data and lessens the multiple testing correction burden, but also tests for case-control differences at the level of gene-co-expression regulation, rather than at the level of individual genes.

Once a module has been identified, it can be further functionally characterized bioinformatically according to the known functions of molecules in the modules or experimentally by genetically manipulating the components of the module. One can test whether genes from a particular pathway, functional class, or other set of genes occur in a

module more frequently than expected by chance, i.e. if the module is “enriched” with a class of genes (see [57] and <http://david.abcc.ncifcrf.gov/> for more). For example, in Chen et al. [13], one of the two modules differentially regulated in SZ and BD was enriched with genes involved with neuron protection functions, and the second module was enriched with neuron differentiation and development genes, as well as genes previously found to be associated with SZ via GWASs.

However, these enrichment categories can be general or have many members, meaning that their over-representation may not always be useful. In any case, enrichment of a module with disease-associated genes does not exactly validate either the module’s existence or association with disease, but could be said to provide what Poultney et al. [58] call a “sanity check.” Interpreting the lack of enrichment poses a problem, as well: it could mean that the module is not valid, that the module is not a heritable phenotype, or merely that the genes are sparsely annotated.

A more fundamental problem is determining whether the observed disease-associated module changes are at the root of disease etiology, or simply downstream effects of the primary disease mechanisms or even of chronic drug therapy. It is even possible that genetic variation does not underlie observed module changes. These issues, though, can be addressed by integrating different data types, e.g. genetic, epigenetic, co-expression, protein, metabolite and environmental data.

Molecular network analysis connects genes to their regulators, using multi-dimensional data

In the graphs discussed so far, the edges are lines without arrows, representing un-directed relationships. In molecular networks, directionality is interpreted as causation, which can be considered a Holy Grail of biology. To predict phenotype from genotype, to differentiate causes from consequences, we must be able to orient network edges (Fig. 2).

Key to this orientation is the integration of multiple data types, i.e. rather than only using gene co-expression data or PPI data to infer a network, other data, like genetic variants, are needed. Using multiple data types can make up for missing or problematic data in any one type of data, and converging evidence from two or more data types make findings more robust [59]. Also, modules from different data types do not tend to overlap [60], so the complete biological model is unlikely to be accurately and adequately represented without the full complement of levels of biological regulation [61]. Microarray data cannot capture the DNA-binding activity of many transcription factors, which depends on post-transcriptional modifications. However, inferring causation is the most important reason for integrating genomics, epigenomics, transcriptomics, proteomics, and/or metabolomics data into a single network.

For example, in their late-onset AD study mentioned above, Zhang et al. [45] used expression quantitative trait loci (eQTLs) as prior probabilities to infer directed co-expression modules. Because they could infer directionality, this Bayesian approach allowed them to identify the TYROBP protein a key causal regulator in an immune and microglia-specific co-expression module (Fig. 3). One drawback of this approach is that Bayesian

analyses cannot infer feedback loops or similar non-linear motifs; their importance are discussed below.

Non-linear features of molecular networks are associated with complex phenotypes

Multiple network properties can disrupt the linear relationship between genotype and phenotype. These properties are difficult to investigate using reductionist methods. Some examples are epistasis, pleiotropy, and recurrent non-linear motifs such as feedback loops. They may contribute to the complexity of the genetic architecture of neuropsychiatric disorders.

Epistasis—Epistasis is generally defined as the phenomena of one gene masking the effect of another. It can be more narrowly defined in two ways, statistically and biologically. Statistically, epistasis is the non-additive effects of two different loci on a phenotype. It can only be detected if there is variation at both loci and is often referred to as gene-gene interaction. Biologically, epistasis can be described as interaction among molecules that contributes to phenotype.

The frequency of epistasis has been hotly debated over the last decade or so, as has been its role in the genetic architecture of complex phenotypes (e.g. [62–64]). Fueling the debate is the fact that the frequency and effect of epistasis has been difficult to assess. Detection of epistasis statistically in genome-wide data sets is hampered by the sheer number of tests that would be required: in a set of one million SNPs, that would be 5×10^{11} possible pair-wise interactions, without even considering three-way tests. Correcting for multiple testing would reduce the tests' statistical power drastically, while the computational power requirement would be enormous.

However, various studies have shown epistasis to be an inherent and ubiquitous property of networks [65]. Since the perturbation of network might be associated with neuropsychiatric disorders, a prediction that epistasis might play a role in those disorders is reasonable, and some interactions have been reported. Judy et al. [66] circumvented the multiple testing correction problem by testing SNPs from 14 previously identified BD risk genes for gene-gene interaction; they detected statistical epistasis between *ANK3* and *KCNQ2*, already known to interact physically. Qin et al. [67] only tested SNPs in *GRIN1* and *GRIN2B* for interaction in SZ in a Chinese population: they found one, namely between a *GRIN1* SNP and a *GRIN2B* SNP. Ma et al.'s [68] AUT association study was restricted to GABA receptor subunit genes, and identified an interaction. There have been many epistatic interactions reported for AD: Combarros et al. [69] used synergy factor analysis to evaluate 100 previously reported interactions affecting AD risk. This logistic regression analysis-based method only replicated 27 interactions. This suggests that, while epistasis likely plays a role in the genetic architecture of AD, and possibly in neuropsychiatric disease in general, methods for detecting it (see [70] for a review) still need improvement.

As epistasis is a product of molecular network interactions, incorporating prior biological interaction data into detection methods could be one such improvement. Emily et al. [71] tested only gene-gene pairs for which the protein products were already known to interact; using that prior biological information allowed them to identify a previously unrecognized

significant statistical interaction that increased BD risk. Further progress may result from combining statistical methods for detecting epistasis with biological network knowledge, similar to Lee et al.'s [72] approach in yeast. Members of a disease-associated module would be tested for SNPs with non-additive effects on phenotype, reducing the test space for statistical interactions and thus decreasing the chances of false negatives. Any set of genes that appear to have a non-additive effect on phenotype can then be followed up with biological experimentation; designing these experiments would be aided by the fact that the gene pair would have come from the same functional module. This unbiased method would include in the interaction tests genes with small or no main effects, a function that is not accomplished by all methods for detecting epistasis.

Pleiotropy—Like epistasis, pleiotropy has been shown to be an inherent and ubiquitous property of molecular networks [65]. There have been many different definitions of the term since it was introduced over one hundred years ago [73]; here, we will define it broadly, as the phenomenon in which one locus, genetic variant or gene, affects multiple phenotypes. Pleiotropy complicates the relationship between genotype to phenotype to such an extent that Featherstone & Brodie [74] proposed using targeted attacks to disrupt gene expression networks to facilitate the study of gene function.

Also like epistasis, the role of pleiotropy in the genetic architecture of complex disease has been a matter of some dispute (e.g. [75–77]). However, there is evidence that it plays a role in neuropsychiatric disease: for example, many common risk loci are associated with multiple neuropsychiatric disorders [78–80]. Rare syndromic copy number variations (CNVs) affecting multiple organ systems have also been associated with multiple neuropsychiatric disorders, e.g. the 22q1.1 deletion with both SZ and AUT [81, 82], and the Timothy Syndrome *CACNA1C* mutations with both AUT [83] and BD [84]. Other CNVs are associated with a range of neuropsychiatric disorders, e.g. the 15q13.3 deletion with SZ, AUT, intellectual disability and generalized epilepsy [81], and 16p11.2 CNVs with SZ, AUT and intellectual disability [81].

Furthermore, regulatory elements are by definition pleiotropic, and a number of them have been associated with neuropsychiatric diseases. Variants in the transcription factor gene *TCF4* are associated with SZ and with two other conditions, Fuchs' corneal dystrophy, and Pitt-Hopkins Syndrome [85]. MicroRNAs (miRNAs) are non-coding RNAs that regulate gene expression. Each can target multiple genes, possibly hundreds [86], and a number of them have been found to be associated with neuropsychiatric diseases or therapies [87]; in particular, there are multiple lines of evidence that miR-137, an miRNA enriched in neurons, is associated with SZ (e.g. [88, 89]). Zhang et al. [90] discussed “eQTL hotspots”, i.e. eQTLs that were associated, not simply with a single gene expression trait, but a group of them. They found that the group of traits can correspond to a co-expression module.

Given these observations, it appears likely that pleiotropic genes would play a role in the genetic architecture of neuropsychiatric diseases. Placing these genes and their interactions into their network context should help us understand the genetic relationships among different neuropsychiatric diagnoses.

Non-linear motifs—Non-linear motifs within a network, such as negative feedback, positive feedback, and oscillatory loops [91] can inhibit or amplify signals in a way that cannot be predicted without knowing their spatial and temporal network context [92]. Feed-forward loops have been shown to be enriched in the ENCODE human gene regulatory network, particularly between levels of the regulatory hierarchy [93]. The mammalian clock gene circuit, which controls circadian rhythms [94] illustrates how complex these non-linear motifs can be (Fig. 3): it is an oscillatory circuit, made up of multiple delayed feedback loops [94]. Epistatic interactions among genes from this circuit are associated with BD risk [95].

Robustness—Complex molecular networks tend to be robust, meaning that perturbations, whether genetic or environmental, can be buffered and ultimately have no effect on phenotype. Robustness can be considered an emergent property of complex networks [96], emergent properties being those that are not attributable to individual components of the network, but instead arise from the network as a whole.

Many of the network features already discussed contribute to robustness. Modularity is one: a module's relatively sparse connectedness to the rest of the network can effectively contain any damage to that module, and prevent it from affecting the rest of the network. An approximately scale-free network is more robust to random attacks [97], since, while damaging a network hub node would fragment the network, the low incidence of highly-connected nodes means there is a low probability of a random attack hitting one of these nodes. Negative feedback loops are associated with increased robustness to the effects of noise and other external perturbations [91].

Pleiotropy and epistasis also contribute to robustness. A property that exemplifies both is phenotypic capacitance [98]: phenotypic capacitors are single genes that buffer genotypic variation in many other genes. Heat shock protein 90 (HSP90) is the best-studied example: as a chaperone molecule that aids in protein folding, it compensates for mutations that would have otherwise rendered a protein unable to fold correctly. Those mutations can then accumulate, essentially hidden from selection pressure, which is why they are often referred to as cryptic genetic variation. Levy & Siegal [99] identified over 300 gene products in yeast that act as capacitors, i.e. when knocked out, they increased phenotypic variability.

Our interest in these compensatory mechanisms is two-fold. First, understanding how disease-associated molecular changes produce a disease phenotype despite these mechanisms will aid our understanding of disease etiology. Second, we need to understand these mechanisms if we are going to successfully design drug therapies that can circumvent them [91]. In addition, knowledge of these buffering mechanisms is important when designing animal or cell models, particularly in controlling for differences in genetic background among individuals.

Molecular network data can help us refine neuropsychiatric disease classification

A major obstacle to connecting genotype to neuropsychiatric phenotype, in both gene-based and network-based studies, is that many neuropsychiatric phenotypes are not well-defined [100, 101]. There are few objective diagnostic criteria, making the diagnostic process quite

subjective. The clinical traits of diseases overlap substantially [102]. The disease presentation can vary greatly among individuals. Multiple neuropsychiatric diagnoses in individuals are common [103, 80], suggesting that they do not represent genuinely independent, biologically-based phenotypes.

For decades, it has been thought that endophenotypes of neuropsychiatric diseases would help circumvent some of these issues (e.g. [104]). However, most of the behavioral and imaging endophenotypes have not been found to have stronger heritabilities than classical disease diagnoses [105, 106]. Endophenotypes also do not necessarily differentiate among neuropsychiatric diseases [107]. If, however, they are found to have stronger modular associations than other phenotypic classifications, they might prove their worth.

We suggest that molecular network characterization in brain and in disease states may ultimately improve neuropsychiatric disease classification and treatment. Identification of disease-associated alterations in the interactions between molecules and in the functional integrity of networks of these interactions, i.e. “networked biomarkers,” can aid in the development of predictive models for both genetic risk and for drug therapy success, by connecting disorders to their biological bases. Several statistical methods have been developed to identify “differential networks” for disease, for example, differential dependency network (DDN) analysis [108].

Issues to consider as network-based studies of neuropsychiatric disorders move forward

Molecular interactions, pathways and networks are context-specific

Whether two molecules interact strongly depends on their biological context, so molecular networks in general, and modules in particular, are context-specific. They can be tissue-specific [23, 109], allele-specific [45, 110, 93], splicing isoform-specific [111], cell-type-specific [43] or brain layer-specific [41, 42]. They also vary over time. So time has to be incorporated into network models [112], whether the interval of interest is the moments between activation of a gene and its subsequent repression via a negative feedback loop [92], a cell cycle, a circadian cycle, or developmental variations over the life cycle. For example, the clock circuit mentioned above changes with age [113], possibly via alterations in DNA methylation [114].

Ideally, all these context variables should be considered in molecular network models, whether built from the top down or bottom up, but, right now, they are not. For example, most PPI networks provided by network tools like GeneMANIA (www.genemania.org) or the STRING database (string-db.org) are not tissue-specific or cell-type-specific. Moreover, one cannot know how the individual interactions were identified without accessing to the original references. Tissue-specific PPI networks have been constructed by filtering the PPIs for protein products of genes known to be expressed in a particular tissue [23]: the HIPPIE (Human Integrated PPI rEference) database [115], and TissueNet [116] include such a filter. Lopes et al. [23] found that such networks tended to be substantially more fragmented than those constructed with unfiltered PPIs. These are not perfect substitutions for a true brain-specific PPI network, since post-translational modifications cannot be taken into account.

Another context variable that must eventually be incorporated into network models is the environment. While many neuropsychiatric diseases are highly heritable, environmental factors do contribute to disease risk, to the extent that an environmental stressor may be required for diagnosis, as for PTSD. The herpes simplex virus may affect AD risk [117] and a GRIN2 variant has been shown to reduce Parkinson's via an interaction with coffee [118]. In addition, the brain has a tightly-regulated developmental program in which timely environmental inputs play a significant role post-natally in activity-dependent neurodevelopment [119]; those inputs will have to be included as well. Ultimately, environmental factors must be considered as network nodes, with their modifying effects as edges.

Conclusions and outlook

Multiple lines of evidence support an association between neuropsychiatric disease and alterations to molecular networks in brain. We have described methods for exploiting network structure to aid in investigation at the level of individual molecules and for characterizing entire networks. These methods are important both for creating more realistic models of biological function and for reducing the dimensionality of the massive amount of data that is beginning to accumulate due to advances in high-throughput assays.

Ultimately, the validity and utility of the network methods described here can only be demonstrated by the construction of a predictive model, such as the identification of a disease biomarker or a successful drug target. Some predictive network-based models of treatment response have been developed for ovarian [120] and breast cancers [121], but nothing similar has yet been identified for neuropsychiatric disease. If we successfully address the issues described above, we may be able to move forward from the static network maps depicted here to a dynamic functional systems model, one with the power to reliably predict phenotypes and therapeutic success [122].

It is possible that a very large-scale predictive model will be available in the not-too-distant future: the goal of the European Commission's Human Brain Project is to simulate the human brain in a computer over the next ten years (www.humanbrain.project.eu). The Brain Initiative (www.nih.gov/science/brain), proposed by the Obama administration, is another large project to map cellular and molecular networks in brain, though its actual content has yet to be defined. These new government-supported consortia will drive the production of massive amounts of data, improved computational and statistical methods, and a better understanding of brain networks, benefiting neuropsychiatric disease studies.

Abbreviations

AUT	autism
BD	bipolar disorder
eQTL	expression quantitative trait locus
GWAS	genome-wide association study

SNP	single nucleotide polymorphism
SZ	schizophrenia
WGCNA	weighted gene co-expression network analysis

Reference List

1. Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet.* 2012; 13:537–551. [PubMed: 22777127]
2. Camargo LM, Collura V, Rain JC, Mizuguchi K, et al. Disrupted in Schizophrenia 1 Interactome: evidence for the close connectivity of risk genes and a potential synaptic basis for schizophrenia. *Mol Psychiatry.* 2007; 12:74–86. [PubMed: 17043677]
3. Fields S, Song O. A novel genetic system to detect protein-protein interactions. *Nature.* 1989; 340:245–246. [PubMed: 2547163]
4. Millar JK, Wilson-Annan JC, Anderson S, Christie S, et al. Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum Mol Genet.* 2000; 9:1415–1423. [PubMed: 10814723]
5. Goh KI, Cusick ME, Valle D, Childs B, et al. The human disease network. *Proc Natl Acad Sci USA.* 2007; 104:8685–8690. [PubMed: 17502601]
6. Feldman I, Rzhetsky A, Vitkup D. Network properties of genes harboring inherited disease mutations. *Proc Natl Acad Sci USA.* 2008; 105:4323–4328. [PubMed: 18326631]
7. Wolfe CJ, Kohane IS, Butte AJ. Systematic survey reveals general applicability of "guilt-by-association" within gene coexpression networks. *BMC Bioinformatics.* 2005; 6:227. [PubMed: 16162296]
8. Bogdanov P, Singh AK. Molecular function prediction using neighborhood features. *IEEE/ACM Trans Comput Biol Bioinform.* 2010; 7:208–217. [PubMed: 20431141]
9. Gillis J, Pavlidis P. "Guilt by association" is the exception rather than the rule in gene networks. *PLoS Comput Biol.* 2012; 8:e1002444. [PubMed: 22479173]
10. Milenkovic, T.; Przulj, N. Topological Characteristics of Molecular Networks. In: Koyuturk, Mehmet; Subramaniam, Shankar; Green, E., editors. *Functional Coherence of Molecular Networks in Bioinformatics.* New York: Springer; 2012. p. 15-48.
11. Lima-Mendez G, van Helden J. The powerful law of the power law and other myths in network biology. *Mol Biosyst.* 2009; 5:1482–1493. [PubMed: 20023717]
12. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol.* 2005; 4:Article17. [PubMed: 16646834]
13. Chen C, Cheng L, Grennan K, Pibiri F, et al. Two gene co-expression modules differentiate psychotics and controls. *Mol Psychiatry.* 2013; 18:1308–1314. [PubMed: 23147385]
14. Draisma HH, Reijmers TH, Meulman JJ, van der Greef J, et al. Hierarchical clustering analysis of blood plasma lipidomics profiles from mono- and dizygotic twin families. *Eur J Hum Genet.* 2013; 21:95–101. [PubMed: 22713803]
15. Chen G, Sullivan PF, Kosorok MR. Biclustering with heterogeneous variance. *Proc Natl Acad Sci USA.* 2013; 110:12253–12258. [PubMed: 23836637]
16. Diaz-Uriarte R, Alvarez de Andres S. Gene selection and classification of microarray data using random forest. *BMC Bioinformatics.* 2006; 7:3. [PubMed: 16398926]
17. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science.* 1995; 270:467–470. [PubMed: 7569999]
18. Weimann M, Grossmann A, Woodsmith J, Ozkan Z, et al. A Y2H-seq approach defines the human protein methyltransferase interactome. *Nat Methods.* 2013; 10:339–342. [PubMed: 23455924]
19. Waaijers S, Koorman T, Kerver J, Boxem M. Identification of human protein interaction domains using an ORFeome-based yeast two-hybrid fragment library. *J Proteome Res.* 2013; 12:3181–3192. [PubMed: 23718855]

20. Fujimori S, Hirai N, Ohashi H, Masuoka K, et al. Next-generation sequencing coupled with a cell-free display technology for high-throughput production of reliable interactome data. *Sci Rep*. 2012; 2:691. [PubMed: 23056904]
21. Yu H, Tardivo L, Tam S, Weiner E, et al. Next-generation sequencing to generate interactome datasets. *Nat Methods*. 2011; 8:478–480. [PubMed: 21516116]
22. Ewing RM, Chu P, Elisma F, Li H, et al. Large-scale mapping of human protein-protein interactions by mass spectrometry. *Mol Syst Biol*. 2007; 3:89. [PubMed: 17353931]
23. Lopes TJ, Schaefer M, Shoemaker J, Matsuoka Y, et al. Tissue-specific subnetworks and characteristics of publicly available human protein interaction databases. *Bioinformatics*. 2011; 27:2414–2421. [PubMed: 21798963]
24. Bernstein BE, Birney E, Dunham I, Green ED, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012; 489:57–74. [PubMed: 22955616]
25. Murigneux V, Sauliere J, Roest CH, Le HH. Transcriptome-wide identification of RNA binding sites by CLIP-seq. *Methods*. 2013; 63:32–40. [PubMed: 23545196]
26. Zhang C, Darnell RB. Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HTS-CLIP data. *Nat Biotechnol*. 2011; 29:607–614. [PubMed: 21633356]
27. Hafner M, Landthaler M, Burger L, Khorshid M, et al. PAR-CLIP--a method to identify transcriptome-wide the binding sites of RNA binding proteins. *J Vis Exp*. 2010; pii:2034. [PubMed: 20644507]
28. Konig J, Zarnack K, Rot G, Curk T, et al. iCLIP--transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution. *J Vis Exp*. 2011; pii:2638. [PubMed: 21559008]
29. Helwak A, Kudla G, Dudnakova T, Tollervey D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell*. 2013; 153:654–665. [PubMed: 23622248]
30. Li JH, Liu S, Zhou H, Qu LH, et al. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res*. 2014; 42:D92–D97. [PubMed: 24297251]
31. Jia P, Kao CF, Kuo PH, Zhao Z. A comprehensive network and pathway analysis of candidate genes in major depressive disorder. *BMC Syst Biol*. 2011; 5(Suppl 3):S12. [PubMed: 22784618]
32. Noh HJ, Ponting CP, Boulding HC, Meader S, et al. Network topologies and convergent aetiologies arising from deletions and duplications observed in individuals with autism. *PLoS Genet*. 2013; 9:e1003523. [PubMed: 23754953]
33. Lee SA, Tsao TT, Yang KC, Lin H, et al. Construction and analysis of the protein-protein interaction networks for schizophrenia, bipolar disorder, and major depression. *BMC Bioinformatics*. 2011; 12(Suppl 13):S20. [PubMed: 22373040]
34. Luo X, Huang L, Jia P, Li M, et al. Protein-protein interaction and pathway analyses of top schizophrenia genes reveal schizophrenia susceptibility genes converge on common molecular networks and enrichment of nucleosome (chromatin) assembly genes in schizophrenia susceptibility Loci. *Schizophr Bull*. 2014; 40:39–49. [PubMed: 23671194]
35. Sun J, Jia P, Fanous AH, van den Oord EJ, et al. Schizophrenia gene networks and pathways and their applications for novel candidate gene selection. *PLoS One*. 2010; 5:e11351. [PubMed: 20613869]
36. Ben-David E, Shifman S. Networks of neuronal genes affected by common and rare variants in autism spectrum disorders. *PLoS Genet*. 2012; 8:e1002556. [PubMed: 22412387]
37. Gulsuner S, Walsh T, Watts AC, Lee MK, et al. Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell*. 2013; 154:518–529. [PubMed: 23911319]
38. Gilman SR, Chang J, Xu B, Bawa TS, et al. Diverse types of genetic variation converge on functional gene networks involved in schizophrenia. *Nat Neurosci*. 2012; 15:1723–1728. [PubMed: 23143521]
39. Gilman SR, Iossifov I, Levy D, Ronemus M, et al. Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron*. 2011; 70:898–907. [PubMed: 21658583]
40. Sakai Y, Shaw CA, Dawson BC, Dugas DV, et al. Protein interactome reveals converging molecular pathways among autism disorders. *Sci Transl Med*. 2011; 3:86ra49.

41. Parikshak NN, Luo R, Zhang A, Won H, et al. Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell*. 2013; 155:1008–1021. [PubMed: 24267887]
42. Willsey AJ, Sanders SJ, Li M, Dong S, et al. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell*. 2013; 155:997–1007. [PubMed: 24267886]
43. Oldham MC, Horvath S, Geschwind DH. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proc Natl Acad Sci USA*. 2006; 103:17973–17978. [PubMed: 17101986]
44. Voineagu I, Wang X, Johnston P, Lowe JK, et al. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature*. 2011; 474:380–384. [PubMed: 21614001]
45. Zhang B, Gaiteri C, Bodea LG, Wang Z, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell*. 2013; 153:707–720. [PubMed: 23622250]
46. Bogdanov P, Macropol K, Singh AK. Function Annotation in Gene Networks. *Functional Coherence in Molecular Networks in Bioinformatics*. 2012:49–68.
47. Franke L, van Bakel H, Fokkens L, de Jong ED, et al. Reconstruction of a functional human gene network, with an application for prioritizing positional candidate genes. *Am J Hum Genet*. 2006; 78:1011–1025. [PubMed: 16685651]
48. Lage K, Karlberg EO, Storling ZM, Olason PI, et al. A human phenome-interactome network of protein complexes implicated in genetic disorders. *Nat Biotechnol*. 2007; 25:309–316. [PubMed: 17344885]
49. Kohler S, Bauer S, Horn D, Robinson PN. Walking the interactome for prioritization of candidate disease genes. *Am J Hum Genet*. 2008; 82:949–958. [PubMed: 18371930]
50. Vanunu O, Magger O, Ruppin E, Shlomi T, et al. Associating genes and protein complexes with disease via network propagation. *PLoS Comput Biol*. 2010; 6:e1000641. [PubMed: 20090828]
51. Lee I, Blom UM, Wang PI, Shim JE, et al. Prioritizing candidate disease genes by network-based boosting of genome-wide association data. *Genome Res*. 2011; 21:1109–1121. [PubMed: 21536720]
52. Doncheva NT, Kacprowski T, Albrecht M. Recent approaches to the prioritization of candidate disease genes. *Wiley Interdiscip Rev Syst Biol Med*. 2012; 4:429–442. [PubMed: 22689539]
53. Magger O, Waldman YY, Ruppin E, Sharan R. Enhancing the prioritization of disease-causing genes through tissue specific protein interaction networks. *PLoS Comput Biol*. 2012; 8:e1002690. [PubMed: 23028288]
54. Karni S, Soreq H, Sharan R. A network-based method for predicting disease-causing genes. *J Comput Biol*. 2009; 16:181–189. [PubMed: 19193144]
55. Miller JA, Horvath S, Geschwind DH. Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways. *Proc Natl Acad Sci USA*. 2010; 107:12698–12703. [PubMed: 20616000]
56. Langfelder P, Mischel PS, Horvath S. When is hub gene selection better than standard meta-analysis? *PLoS One*. 2013; 8:e61505. [PubMed: 23613865]
57. Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput Biol*. 2012; 8:e1002375. [PubMed: 22383865]
58. Poultney CS, Greenfield A, Bonneau R. Integrated inference and analysis of regulatory networks from multi-level measurements. *Methods Cell Biol*. 2012; 110:19–56. [PubMed: 22482944]
59. Holzinger ER, Ritchie MD. Integrating heterogeneous high-throughput data for meta-dimensional pharmacogenomics and disease-related studies. *Pharmacogenomics*. 2012; 13:213–222. [PubMed: 22256870]
60. Costanzo M, Baryshnikova A, Bellay J, Kim Y, et al. The genetic landscape of a cell. *Science*. 2010; 327:425–431. [PubMed: 20093466]
61. Ames RM, Macpherson JJ, Pinney JW, Lovell SC, et al. Modular biological function is most effectively captured by combining molecular interaction data types. *PLoS One*. 2013; 8:e62670. [PubMed: 23658761]

62. Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci USA*. 2012; 109:1193–1198. [PubMed: 22223662]
63. Hill WG, Goddard ME, Visscher PM. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet*. 2008; 4:e1000008. [PubMed: 18454194]
64. Moore JH. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. *Hum Hered*. 2003; 56:73–82. [PubMed: 14614241]
65. Tyler AL, Asselbergs FW, Williams SM, Moore JH. Shadows of complexity: what biological networks reveal about epistasis and pleiotropy. *BioEssays*. 2009; 31:220–227. [PubMed: 19204994]
66. Judy JT, Seifuddin F, Pirooznia M, Mahon PB, et al. Converging evidence for epistasis between ANK3 and potassium channel gene KCNQ2 in bipolar disorder. *Front Genet*. 2013; 4:87. [PubMed: 23730306]
67. Qin S, Zhao X, Pan Y, Liu J, et al. An association study of the N-methyl-D-aspartate receptor NR1 subunit gene (GRIN1) and NR2B subunit gene (GRIN2B) in schizophrenia with universal DNA microarray. *Eur J Hum Genet*. 2005; 13:807–814. [PubMed: 15841096]
68. Ma DQ, Whitehead PL, Menold MM, Martin ER, et al. Identification of significant association and gene-gene interaction of GABA receptor subunit genes in autism. *Am J Hum Genet*. 2005; 77:377–388. [PubMed: 16080114]
69. Combarros O, Cortina-Borja M, Smith AD, Lehmann DJ. Epistasis in sporadic Alzheimer's disease. *Neurobiol Aging*. 2009; 30:1333–1349. [PubMed: 18206267]
70. Cordell HJ. Detecting gene-gene interactions that underlie human diseases. *Nat Rev Genet*. 2009; 10:392–404. [PubMed: 19434077]
71. Emily M, Mailund T, Hein J, Schausser L, et al. Using biological networks to search for interacting loci in genome-wide association studies. *Eur J Hum Genet*. 2009; 17:1231–1240. [PubMed: 19277065]
72. Lee I, Lehner B, Vavouri T, Shin J, et al. Predicting genetic modifier loci using functional gene networks. *Genome Res*. 2010; 20:1143–1153. [PubMed: 20538624]
73. Stearns FW. One hundred years of pleiotropy: a retrospective. *Genetics*. 2010; 186:767–773. [PubMed: 21062962]
74. Featherstone DE, Broadie K. Wrestling with pleiotropy: genomic and topological analysis of the yeast gene expression network. *BioEssays*. 2002; 24:267–274. [PubMed: 11891763]
75. Wagner GP, Zhang J. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nat Rev Genet*. 2011; 12:204–213. [PubMed: 21331091]
76. Hill WG, Zhang XS. On the pleiotropic structure of the genotype-phenotype map and the evolvability of complex organisms. *Genetics*. 2012; 190:1131–1137. [PubMed: 22214609]
77. Paaby AB, Rockman MV. Pleiotropy: what do you mean? Reply to Zhang and Wagner. *Trends Genet*. 2013; 29:384. [PubMed: 23746964]
78. Smoller JW, Craddock N, Kendler K, Lee PH, et al. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet*. 2013; 381:1371–1379. [PubMed: 23453885]
79. Lee SH, Ripke S, Neale BM, Faraone SV, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet*. 2013; 45:984–994. [PubMed: 23933821]
80. Rzhetsky A, Wajngurt D, Park N, Zheng T. Probing genetic overlap among complex human phenotypes. *Proc Natl Acad Sci USA*. 2007; 104:11694–11699. [PubMed: 17609372]
81. Vorstman JA, Breetvelt EJ, Thode KI, Chow EW, et al. Expression of autism spectrum and schizophrenia in patients with a 22q11.2 deletion. *Schizophr Res*. 2013; 143:55–59. [PubMed: 23153825]
82. Jonas RK, Montojo CA, Bearden CE. The 22q11.2 deletion syndrome as a window into complex neuropsychiatric disorders over the lifespan. *Biol Psychiatry*. 2013; 75:351–360. [PubMed: 23992925]
83. Splawski I, Timothy KW, Decher N, Kumar P, et al. Severe arrhythmia disorder caused by cardiac L-type calcium channel mutations. *Proc Natl Acad Sci USA*. 2005; 102:8089–8096. [PubMed: 15863612]

84. Gershon ES, Grennan K, Busnello J, Badner JA, et al. A rare mutation of CACNA1C in a patient with bipolar disorder, and decreased gene expression associated with a bipolar-associated common SNP of CACNA1C in brain. *Mol Psychiatry*. 2013 in press.
85. Navarrete K, Pedrosa I, de Jong S, Stefansson H, et al. TCF4 (e2-2; ITF2): a schizophrenia-associated gene with pleiotropic effects on human disease. *Am J Med Genet B Neuropsychiatr Genet*. 2013; 162B:1–16. [PubMed: 23129290]
86. Selbach M, Schwanhaussner B, Thierfelder N, Fang Z, et al. Widespread changes in protein synthesis induced by microRNAs. *Nature*. 2008; 455:58–63. [PubMed: 18668040]
87. Grennan, K.; Zhang, F.; Liu, CY. MicroRNAs and their potential roles in schizophrenia and bipolar disorder. In: Yue; Junming; Lu, editors. *Lu miRNAs and Human Diseases*. Trivandrum: Research Signpost; 2012. p. 131-154.
88. Guella I, Sequeira A, Rollins B, Morgan L, et al. Analysis of miR-137 expression and rs1625579 in dorsolateral prefrontal cortex. *J Psychiatr Res*. 2013; 47:1215–1221. [PubMed: 23786914]
89. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet*. 2011; 43:969–976. [PubMed: 21926974]
90. Zhang W, Zhu J, Schadt EE, Liu JS. A Bayesian partition method for detecting pleiotropic and epistatic eQTL modules. *PLoS Comput Biol*. 2010; 6:e1000642. [PubMed: 20090830]
91. Miguez DG. Network nonlinearities in drug treatment. *Interdiscip Sci*. 2013; 5:85–94. [PubMed: 23740389]
92. Isalan M. Gene networks and liar paradoxes. *BioEssays*. 2009; 31:1110–1115. [PubMed: 19722183]
93. Gerstein MB, Kundaje A, Hariharan M, Landt SG, et al. Architecture of the human regulatory network derived from ENCODE data. *Nature*. 2012; 489:91–100. [PubMed: 22955619]
94. Zhang EE, Kay SA. Clocks not winding down: unravelling circadian networks. *Nat Rev Mol Cell Biol*. 2010; 11:764–776. [PubMed: 20966970]
95. Shi J, Wittke-Thompson JK, Badner JA, Hattori E, et al. Clock genes may influence bipolar disorder susceptibility and dysfunctional circadian rhythm. *Am J Med Genet B Neuropsychiatr Genet*. 2008; 147B:1047–1055. [PubMed: 18228528]
96. Kitano H. Biological robustness. *Nat Rev Genet*. 2004; 5:826–837. [PubMed: 15520792]
97. Albert R, Jeong H, Barabasi AL. Error and attack tolerance of complex networks. *Nature*. 2000; 406:378–382. [PubMed: 10935628]
98. Bergman A, Siegal ML. Evolutionary capacitance as a general feature of complex gene networks. *Nature*. 2003; 424:549–552. [PubMed: 12891357]
99. Levy SF, Siegal ML. Network hubs buffer environmental variation in *Saccharomyces cerevisiae*. *PLoS Biol*. 2008; 6:e264. [PubMed: 18986213]
100. Alawieh A, Zaraket FA, Li JL, Mondello S, et al. Systems biology, bioinformatics, and biomarkers in neuropsychiatry. *Front Neurosci*. 2012; 6:187. [PubMed: 23269912]
101. Craddock N, Kendler K, Neale M, Nurnberger J, et al. Dissecting the phenotype in genome-wide association studies of psychiatric illness. *Br J Psychiatry*. 2009; 195:97–99. [PubMed: 19648536]
102. Lyalina S, Percha B, Lependu P, Iyer SV, et al. Identifying phenotypic signatures of neuropsychiatric disorders from electronic medical records. *J Am Med Inform Assoc*. 2013; 20:e297–e305. [PubMed: 23956017]
103. Fava GA, Tossani E, Bech P, Berrocal C, et al. Emerging clinical trends and perspectives on comorbid patterns of mental disorders in research. *Int J Methods Psychiatr Res*. 2014; 23(Suppl 1):92–101. [PubMed: 24375537]
104. Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry*. 2003; 160:636–645. [PubMed: 12668349]
105. Greenwood TA, Swerdlow NR, Gur RE, Cadenhead KS, et al. Genome-wide linkage analyses of 12 endophenotypes for schizophrenia from the Consortium on the Genetics of Schizophrenia. *Am J Psychiatry*. 2013; 170:521–532. [PubMed: 23511790]

106. Geng X, Prom-Wormley EC, Perez J, Kubarych T, et al. White matter heritability using diffusion tensor imaging in neonatal brains. *Twin Res Hum Genet.* 2012; 15:336–350. [PubMed: 22856369]
107. Tamminga CA, Ivleva EI, Keshavan MS, Pearlson GD, et al. Clinical phenotypes of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *Am J Psychiatry.* 2013; 170:1263–1274. [PubMed: 23846857]
108. Zhang B, Li H, Riggins RB, Zhan M, et al. Differential dependency network analysis to identify condition-specific topological changes in biological networks. *Bioinformatics.* 2009; 25:526–532. [PubMed: 19112081]
109. Dobrin R, Zhu J, Molony C, Argman C, et al. Multi-tissue coexpression networks reveal unexpected subnetworks associated with disease. *Genome Biol.* 2009; 10:R55. [PubMed: 19463160]
110. Xu L, Barker B, Gu Z. Dynamic epistasis for different alleles of the same gene. *Proc Natl Acad Sci USA.* 2012; 109:10420–10425. [PubMed: 22689976]
111. Foley KS, Young PW. An analysis of splicing, actin-binding properties, heterodimerization and molecular interactions of the non-muscle alpha-actinins. *Biochem J.* 2013; 452:477–488. [PubMed: 23557398]
112. Przytycka TM, Singh M, Slonim DK. Toward the dynamic interactome: it's about time. *Brief Bioinform.* 2010; 11:15–29. [PubMed: 20061351]
113. Kunieda T, Minamino T, Katsuno T, Tateno K, et al. Cellular senescence impairs circadian expression of clock genes in vitro and in vivo. *Circ Res.* 2006; 98:532–539. [PubMed: 16424366]
114. Zhang L, Lin QL, Lu L, Yang CC, et al. Tissue-specific modification of clock methylation in aging mice. *Eur Rev Med Pharmacol Sci.* 2013; 17:1874–1880. [PubMed: 23877850]
115. Schaefer MH, Lopes TJ, Mah N, Shoemaker JE, et al. Adding protein context to the human protein-protein interaction network to reveal meaningful interactions. *PLoS Comput Biol.* 2013; 9:e1002860. [PubMed: 23300433]
116. Barshir R, Basha O, Eluk A, Smoly IY, et al. The TissueNet database of human tissue protein-protein interactions. *Nucleic Acids Res.* 2013; 41:D841–D844. [PubMed: 23193266]
117. Bearer EL. HSV, axonal transport and Alzheimer's disease: in vitro and in vivo evidence for causal relationships. *Future Virol.* 2012; 7:885–899. [PubMed: 23335944]
118. Hamza TH, Chen H, Hill-Burns EM, Rhodes SL, et al. Genome-wide gene-environment study identifies glutamate receptor gene GRIN2A as a Parkinson's disease modifier gene via interaction with coffee. *PLoS Genet.* 2011; 7:e1002237. [PubMed: 21876681]
119. Purves, D.; Augustine, GJ.; Fitzpatrick, D.; Hall, WC., et al. *Neuroscience.* 4th edition. Sinauer Associates, Inc; 2008.
120. Zhang W, Ota T, Shridhar V, Chien J, et al. Network-based survival analysis reveals subnetwork signatures for predicting outcomes of ovarian cancer treatment. *PLoS Comput Biol.* 2013; 9:e1002975. [PubMed: 23555212]
121. Taylor IW, Linding R, Warde-Farley D, Liu Y, et al. Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat Biotechnol.* 2009; 27:199–204. [PubMed: 19182785]
122. Bader JS. Grand network convergence. *Genome Biol.* 2011; 12:306. [PubMed: 21682935]

Figure 1.

Illustration of network concepts. **A:** Dots indicate **node**s in the network, which represent objects, and the lines, or **edges**, represent pair-wise relationships between the objects. **B:** A **module** is a sub-network of highly interconnected nodes in the network, and an **intramodular hub** is the most connected node or nodes in a module. **C:** Biological networks comprise many modules, as well as sparsely-connected nodes. Purple lines indicate relationships between nodes belonging to different modules, and cyan lines indicate relationships between non-modular nodes.

A)

B)

Figure 2.
Undirected versus directed graphs. **A:** An undirected graph, **B:** A directed graph

Figure 3.

The mammalian clock gene circuit. Arrows indicate activation; T-ended lines indicate repression. Red nodes and edges represent activation relationships (retinoid-related orphan receptors (RORs) and D-box-binding protein-thyrotroph embryonic factor-hepatic leukaemia factor (DBP-TEF-HLF) are transcriptional activators). Blue nodes and edges represent repression relationships (REV-ERBs and E4 promoter-binding protein 4 (E4BP4) are transcriptional repressors). After Figure 2 of [94].