10

From Common Variant to Function

State-of-the-Art Approaches

Ellen Hu and Hyejung Won

Abstract

Genome-wide association studies (GWASs) have been integral to our understanding of the polygenic architecture of psychiatric disorders, yet distilling disease biology from GWAS remains a challenge because GWAS-identified genomic regions often contain dozens of variants with highly correlated structure. The majority of variants within these loci are located within the noncoding genome, so their functional consequences are not immediately apparent. Moreover, to characterize thousands of such variants requires a highly scalable experimental approach. Such systematic interrogation of the functional consequences of variants is enabled by population-scale molecular assays (e.g., quantitative trait loci) or scalable genomic perturbation assays (e.g., massively parallel reporter assays or CRISPR screens). This review describes available genomic resources and cutting-edge experimental approaches that have been adopted to infer functional consequences of risk variants. Additionally, it outlines gaps in defining causal variants, linking variants to target genes, interrogating combinatorial effects of variants within the polygenic background, and considering the context-specific nature of variants. Extended efforts to fill these gaps will enable more comprehensive interpretation of GWAS and ultimately reveal the fundamental biological context behind polygenic psychiatric disorders.

Introduction

Despite the success of genome-wide association studies (GWASs) in identifying common variants associated with psychiatric disorders, specific properties of common risk variants pose a unique challenge for interpreting their functions. The traditional strategy to link (*de novo*) rare variants to function is to identify their impact on protein function. Since rare variants are often located in

From "Exploring and Exploiting Genetic Risk for Psychiatric Disorders," edited by Joshua A. Gordon and Elisabeth B. Binder. Strüngmann Forum Reports, vol. 31, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262547383 (paperback) 9780262377423 (pdf)

protein-coding sequences, this strategy is often straightforward. Nonetheless, two unique features of common variants distinguish them from rare variants, making the scientific approach to linking them to function extremely challenging. First, common variants exhibit a correlation structure with nearby variants due to linkage disequilibrium. Thus, a typical GWAS locus contains dozens to hundreds of variants with strong association signals, making it difficult to identify which variants are playing a causal role in disease etiology. Second, over 90% of common variants are in the noncoding genome (Watanabe et al. 2019b). Unlike variants located in protein-coding genes, we often do not know the function of noncoding variants, as they are likely involved in complex, cell-type, and context-dependent gene regulatory activities. Strategies that have been used to decipher the impact of rare variants on psychiatric disorders are thus not directly applicable to common variants where we do not expect direct impact on single protein function by coding sequence changes but effects on the regulatory fine-tuning of gene expression.

What Is the Definition of Causal Variants?

A challenge in the field is a lack of consensus on the definition of causal common variants. Computational fine mapping was initially used to identify putative causal or "credible" variants, which refer to the smallest set of variants that can explain the association statistics of a given GWAS locus including many correlated variants. We have previously shown that different fine-mapping algorithms can predict different sets of credible variants (Mah and Won 2020). This lack of reproducibility raises questions on how accurately any given fine-mapping algorithm can predict the most likely causal variants.

To address this issue, experimental approaches that measure the functional consequences of noncoding variants are now being increasingly used. For example, effects on gene regulatory activity can be assessed using massively parallel reporter assays (MPRA) and CRISPR editing. These assays measure the functional impact on gene regulation of a large number of risk variants simultaneously, based on the hypothesis that causal variants likely act through altered gene expression. MPRA adopts the design of a traditional luciferase assay, which has been used to study the regulatory activity of noncoding elements. A luciferase assay examines the ability of small sections of DNA that contain the variants of interest to modulate the expression of the reporter gene, luciferase, which can be read out via luminescence. Whereas luminescence of luciferase is measured as readouts for regulatory activity in luciferase assays, random barcode expression is measured as readouts for regulatory activity in MPRA via RNA sequencing, greatly increasing its throughput (Figure 10.1a). MPRA permits rapid testing of regulatory effects of many variants in a single experiment and identifies variants with differential allelic regulatory effects (McAfee et al. 2022a).

From "Exploring and Exploiting Genetic Risk for Psychiatric Disorders," edited by Joshua A. Gordon and Elisabeth B. Binder. Strüngmann Forum Reports, vol. 31, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262547383 (paperback) 9780262377423 (pdf)

CRISPR screens utilize nuclease-dead Cas9 (dCas9) fused with a repressor domain (KRAB; Yeo et al. 2018). After introducing guide-RNAs (gRNA) that target dCas9-KRAB to loci containing the variants of interest, gene expression profiles of nearby genes are measured by RNA sequencing (Figure 10.1b). If a certain gene is up- or downregulated after introducing a gRNA, we can conclude that the targeted locus can contain a regulatory variant that affects expression of that gene. Therefore, regulatory regions carrying the variants of interest and many potential target genes can be interrogated in a single assay. While CRISPR screens also allow the testing of allele-specific differences, the scale of such experiments is far below those which just target putative regulatory regions that contain the variant.

Joint use of MPRA and CRISPR screens can comprehensively illuminate the variant function, at least with regard to gene regulatory functions. When used together, MPRA can identify variants with allelic regulatory activity at a scale beyond the capacity of CRISPR, whereas CRISPR screens can supplement the endogenous regulatory context and target genes of the variants.

Note again that the MPRA and CRISPR screens are powerful experimental strategies that depend on a prevailing hypothesis in the field; namely, that "functional regulatory variants" are "causal variants." It remains unknown, however, whether all functional regulatory variants with significant association signals are causal variants or whether causal variants are always engaged in gene regulatory function. Other options include indirect effects on gene regulation via noncoding RNAs or regulation of alternative splicing.

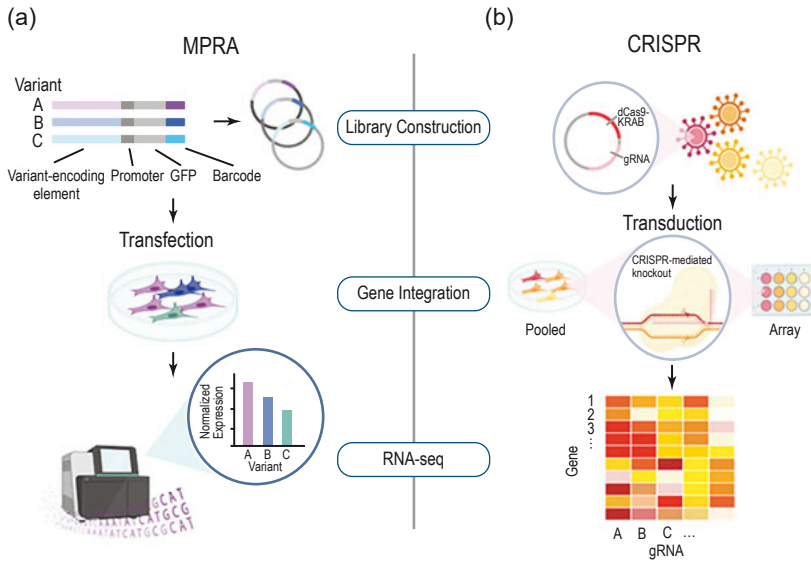

Figure 10.1 Schematic workflow of experimental assays (MPRA and CRISPR) for large-scale functional validation of genetic variants.

From "Exploring and Exploiting Genetic Risk for Psychiatric Disorders," edited by Joshua A. Gordon and Elisabeth B. Binder. Strüngmann Forum Reports, vol. 31, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262547383 (paperback) 9780262377423 (pdf)

Development of additional experimental and computational strategies, as well as a standardized benchmark to define causal variants against, will be pivotal to unravel the role of variants in disease etiology. Alternative approaches are also discussed below.

Next we discuss how multi-omic data sets generated and made available through consortia-level efforts have enabled functional annotation of noncoding variants.

Population-Scale Molecular Assays to Link Variants to Genes

Population-scale RNA sequencing data combined with genotype provides a resource for expression quantitative trait loci (eQTL) that can correlate variants with gene expression levels. Because eQTLs can translate variant information to quantified gene expression, they have played a pivotal role in addressing the functional consequence of variants associated with human traits and diseases. Accordingly, compiling large eQTL resources has been the focus of many groups and consortia (Table 10.1). The NIH-funded GTEx Consortium was a pioneering effort that gathered RNA sequencing data from 838 individuals across 49 tissues that include multiple brain regions, such as the cortex, cerebellum, hippocampus, striatum, amygdala, and substantia nigra (GTEx Consortium 2020). The PsychENCODE Consortium generated a more comprehensive catalog of brain eQTLs by amassing dorsolateral prefrontal cortex (DLPFC) RNA sequencing data acquired from 1,387 individuals (Wang et al. 2018). Meta-analysis of multi-ancestry eQTLs further expanded this resource by profiling RNA sequencing data from 2,119 individuals, with focus on diverse ancestries for more non-European representation (Zeng et al. 2022).

Table 10.1 Existing resources of brain expression quantitative trait loci (eQTL).

Resources	Number of Individuals	Brain Regions or Cell Types
GTEx Consortium (2020)	≤209	Amygdala, cortex, cerebellum, hippocampus, hypothalamus, nucleus accumbens, putamen, substantia nigra
BrainSeq Consortium (2015)	738	Dorsolateral prefrontal cortex, hippocampus
PsychENCODE Consortium		
Wang et al. (2018)	1,387	Dorsolateral prefrontal cortex
Zeng et al. (2022)	2,119	Dorsolateral prefrontal cortex
de Klein et al. (2023)	6,518	Amygdala, basal ganglia, cortex, cerebellum, hippocampus, hypothalamus
Bryois et al. (2022)	196	Astrocytes, endothelial cells, excitatory/inhibitory neurons, microglia, oligodendrocytes, oligodendrocyte precursor cells, pericytes

From "Exploring and Exploiting Genetic Risk for Psychiatric Disorders," edited by Joshua A. Gordon and Elisabeth B. Binder. Strüngmann Forum Reports, vol. 31, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262547383 (paperback) 9780262377423 (pdf)

More efforts are being undertaken to create a more comprehensive eQTL catalog that covers seven brain regions (de Klein et al. 2023) or different cell types (Bryois et al. 2022). Indeed, eQTL catalogs are now transitioning from sequencing bulk tissue to obtaining data with cell-type resolution using single-cell sequencing approaches.

While eQTL resources have been central to functionally interpreting GWAS-identified variants, currently available eQTL resources cannot explain all GWAS loci. According to current data, many genetic variants associated with psychiatric disorders do not exhibit detectable effects on gene expression in eQTL studies. For example, colocalization analysis between multi-ancestry brain eQTLs and joint GWAS of schizophrenia and bipolar disorder assigned only 20 out of 144 loci to genes (Zeng et al. 2022). One way to address this gap is to increase the scale of eQTL studies. However, increasing sample size to detect additional eQTLs in steady-state adult brain samples using bulk sequencing is approaching saturation (Wang et al. 2018; de Klein et al. 2023), so other strategies that target unexplored aspects of regulatory mechanisms may be a more effective target of focus. Below, we discuss some of the missing QTL resources that may be pivotal to closing the gap in the QTL approach to GWAS functional interpretation.

Quantitative Trait Loci for Different Molecular Assays

Changes in transcriptional regulation are only one kind of potential functional consequence of common or noncoding genetic variation. RNA species besides polyadenylated RNA, such as long noncoding RNA (lncRNA) and microRNA, have also been shown to play a critical role in gene regulation. However, single nucleotide polymorphism (SNP) associations with lncRNA and microRNA have not yet been evaluated comprehensively at the genome-wide level. QTLs that target these RNA species may identify unexplored mechanisms underlying psychiatric disorders. Promisingly, transcriptomic profiling of postmortem brain samples with psychiatric disorders has shown that many lncRNAs are dysregulated in psychiatric conditions (Gandal et al. 2018b). In addition, other noncoding RNAs, such as circular RNAs, should also be explored.

Another important transcriptional control occurs at the level of alternative splicing. Postmortem brain samples with schizophrenia show widespread dysregulation of splice isoforms (Gandal et al. 2018b), and processes involved with RNA splicing are associated with various psychiatric disorders (Sey et al. 2020), indicating that RNA splicing can provide imperative insights into disease biology. Variants associated with RNA splicing or isoform usage in the human brain have been identified by the detection of splice-QTLs and splicing-isoform QTLs (Li et al. 2016; Wang et al. 2018).

Chromatin QTLs (cQTLs) propose another functional explanation for the association between risk loci and disease mechanism. Specifically, analyzing variant effects on profiles of open-chromatin (from ATAC sequencing)

From "Exploring and Exploiting Genetic Risk for Psychiatric Disorders," edited by Joshua A. Gordon and Elisabeth B. Binder. Strüngmann Forum Reports, vol. 31, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262547383 (paperback) 9780262377423 (pdf)

gives chromatin accessibility QTLs (caQTLs), while that for specific chromatin marks (from ChIP-seq) gives histone QTLs (haQTLs). PsychENCODE pioneered the large-scale identification of brain caQTLs (Bryois et al. 2018) and haQTLs (Wang et al. 2018). Overlapping cQTLs with eQTLs can provide mechanistic insight into (a) how chromatin changes facilitate downstream effects on gene expression and (b) on the level of impact of the genetic variant.

Cell Type-Specific Quantitative Trait Loci

Another key limitation in the current eQTL resources is attributed to bulk RNA sequencing, where expression quantification is heavily affected by cellular heterogeneity. Since gene expression varies with cell type, QTL resources obtained from the bulk brain homogenate may lack cell type-specific regulatory variants that could have a strong impact on psychiatric disorders. Hence, identifying cell type-specific QTLs can help prioritize cell types with the largest impact on disease development.

A significant advancement in the identification of cell type-specific QTLs could be achieved by single-cell (sc)RNA sequencing, which solves the cellular heterogeneity issue with bulk sequencing by measuring gene expression at cellular resolution (Bryois et al. 2022). QTLs can be combined with scRNA sequencing to catch otherwise undetectable functional effects of variants that are specific to particular cell types. Large-scale scRNA sequencing data, however, is very expensive and suffers a low signal-to-noise ratio due to the sparsity of the data set; hence conducting such studies on large samples with robust expression profiles represents a practical challenge. When such a large-scale scRNA sequencing resource is not available, computational deconvolution methods by which cell type composition is inferred from bulk data can provide a potential alternative to the existing RNA sequencing data sets generated from the bulk brain samples. While current deconvolution methods are only modestly accurate (Kim-Hellmuth et al. 2020), once improved, they may provide a resourceful way to extract cell type-specific signatures from abundant existing bulk sequencing data.

In addition to cell type-specific QTLs, SNPs may affect the disease biology by regulating cellular proportions within tissue; in the case of brain tissue, this could have a large impact. The PsychENCODE Consortium has identified cell fraction QTLs, which indicate SNPs associated with variance in cell-type proportions, by deconvolving bulk RNA sequencing data with scRNA sequencing signatures as a reference (Wang et al. 2018). Fraction QTLs were found to explain a large portion of gene expression variability in the brain samples, implying that changes in cell composition can shape bulk expression profiles. This result suggests the possibility that variant impact on cell-type composition could provide additional insights beyond variant function within a cell type. Such fraction QTLs may emerge from genetic variants related to cell-type differentiation during development and may also moderate cell-type proportion

From "Exploring and Exploiting Genetic Risk for Psychiatric Disorders," edited by Joshua A. Gordon and Elisabeth B. Binder. Strüngmann Forum Reports, vol. 31, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262547383 (paperback) 9780262377423 (pdf)

changes over age. Taken together, cell type-specific QTLs and fraction QTLs can elucidate unexplored facets of variant function in psychiatric disorders.

Quantitative Trait Loci that Span Developmental Stages

Many psychiatric disorders are thought to have neurodevelopmental origins. Both gene expression and the regulatory landscape are highly variable across developmental stages (Kang et al. 2011; Li et al. 2018), supporting the idea that psychiatric disorders need to be characterized across critical stages of brain development. Enhancers in the developing human brain were found to be enriched in variants associated with psychiatric disorders, illustrating the importance of the regulatory landscape during neurogenesis to the etiology of psychiatric disorders (de la Torre-Ubieta et al. 2018; Spiess and Won 2020; Won et al. 2016).

To expand QTL resources across developmental stages, eQTLs have been compiled from fetal brains (Walker et al. 2019; Werling et al. 2020). Though many of these eQTLs are not entirely distinct to the fetal brain, their enrichment in psychiatric risk variants implies that mechanisms contributing to neuropsychiatric disease start as early as the fetal stage (Werling et al. 2020). The existence of fetal-specific eQTLs and identification of temporal-dominant eQTLs (prenatal- or postnatal-dominant; Werling et al. 2020) suggests that eQTLs exert varied effects on expression across developmental stages. Therefore, large-scale expression data across multiple developmental stages will be necessary to understand a crucial window for the development of psychiatric disorders.

Context-Specific Quantitative Trait Loci

Regulatory variants may require an external stimulus, such as a drug or environmental factor, to become activated. Response eQTLs and caQTLs have been identified in macrophages upon immune activation, demonstrating that some variants affect gene regulation only upon stimulation (Alasoo et al. 2018). Context-specific profiling of the human brain is sparse due to a lack of large data sets with environmental exposures and gene expression. Further exploration of context-specific QTLs may help explain why some GWAS loci have been linked to target genes but were not associated with changes in expression (Javierre et al. 2016); the loci may not have been studied under conditions necessary to activate variants, so experimenting with various stimuli may reveal the true variant function on gene expression.

Are Quantitative Trait Loci Gold Standards to Link GWAS to Function?

While refined QTL studies, such as those noted above, are needed to elucidate the full picture of variant effects on gene regulation, the importance of eQTL studies for linking GWAS hits to biological function has not been definitively

From "Exploring and Exploiting Genetic Risk for Psychiatric Disorders," edited by Joshua A. Gordon and Elisabeth B. Binder. Strüngmann Forum Reports, vol. 31, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262547383 (paperback) 9780262377423 (pdf)

established. For example, recent studies have suggested that eQTL studies and GWAS may identify different sets of variants driven by selective pressure (Mostafavi et al. 2022). Supporting this finding, over 60% of schizophrenia risk variants whose regulatory functions are validated by MPRA do not exhibit eQTL signals (McAfee et al. 2022b). Furthermore, MPRA-validated variants do not overlap with eQTLs mapped to genes with higher mutational constraints and richer functional annotations, compared to those that do overlap with eQTLs. The observed discrepancy can be partially explained by selective pressure: the episomal design of MPRA may not be subject to the selective pressure that depletes eQTLs around mutation-intolerant genes. Since eQTLs remain a dominant strategy to link GWAS to function, there is a need to develop and adopt additional genomic approaches for linking variants to genes that are immune to linkage disequilibrium structure and selective pressure. In the next section, we review how other types of functional genomic data sets offer such alternatives to investigate variant function and disease mechanism.

Functional Genomic Approaches to Annotate Variants

Studying chromatin architecture (e.g., chromatin accessibility, histone modifications, chromatin interaction profiles) provides a complementary opportunity to link GWAS variants to genes. For example, chromatin architecture may be well-suited for linking variants to genes that are under selective pressure. Mutation-intolerant genes associated with various human diseases (including schizophrenia; see Pardiñas et al. 2018) were shown to be depleted of eQTLs because deleterious regulatory variants for these genes may occur at low frequency in the population (Mostafavi et al. 2022). In contrast, mutation-intolerant genes were shown to have more complex regulatory architecture (e.g., enhancer-gene interactions) to buffer them from deleterious regulatory variants (Wang and Goldstein 2020). For those genes, chromatin architecture may allow for more comprehensive annotation of disease-associated variants.

Here, we will first use chromatin interaction profiles as an example to show how chromatin architecture has facilitated functional annotation of GWAS. Thereafter, we show how integration of multi-omics data sets can further improve variant-gene relationships.

Chromatin Interaction Profiles

Chromatin interaction maps, defined by chromosome conformation capture techniques (e.g., Hi-C), can supply insight on functional consequences of noncoding variants by linking them to the target genes. For example, if a variant A forms a loop with a gene B, it is postulated that the variant A may regulate gene B given their physical proximity (Figure 10.2).

From "Exploring and Exploiting Genetic Risk for Psychiatric Disorders," edited by Joshua A. Gordon and Elisabeth B. Binder. Strüngmann Forum Reports, vol. 31, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262547383 (paperback) 9780262377423 (pdf)

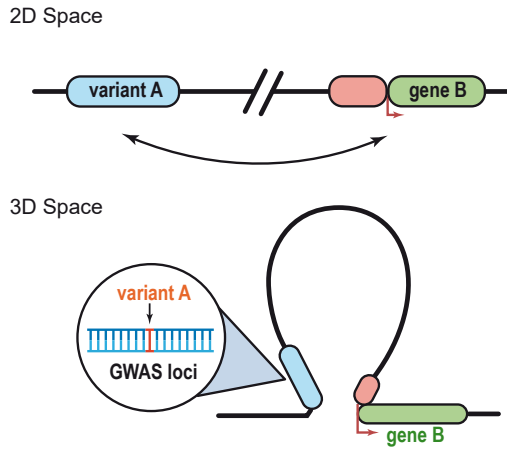

Figure 10.2 Three-dimensional proximity of variant A to gene B in a loop configuration suggests that variant A holds a regulatory role in the expression of gene B (e.g., via an enhancer-promoter interaction).

Chromatin interaction profiles have shown that over 80% of noncoding variants physically interact with distal genes (Sey et al. 2020). This underscores the importance of distal regulatory relationships in mapping variants to genes. Notably, chromatin interaction maps have been built across brain development and in different cell types (Pratt and Won 2022). These maps potentially facilitate the identification of genes associated with psychiatric disorders in a developmental stage and cell type-specific manner (Table 10.2).

The genome-wide nature of Hi-C allows unbiased characterization of interacting targets of all variants. This, together with the genome-wide gene mapping tool MAGMA, led to the development of H-MAGMA, which converts the search space from variants of unknown function to genes of well-characterized function leveraging chromatin interaction maps (de Leeuw et al. 2020; Sey et al. 2020). H-MAGMA aggregates variant-level association statistics to

Table 10.2 Existing resources of brain chromatin interaction profiles.

Reference	Brain region	Cell types	Developmental Stage
Won et al. (2016)	Cortex	Germinal zone, cortical plate	Fetal
Song et al. (2020)	Cortex	Radial glia, intermediate progenitors, excitatory and inhibitory neurons	Fetal
Nott et al. (2019)	Cortex	Neuron, microglia, oligodendrocyte	Pediatric
Wang et al. (2018)	Cortex	Brain homogenate	Adult
Hu et al. (2021)	Cortex	Neuron, glia	Adult
Sey et al. (2022)	Midbrain	Dopaminergic neuron	Adult

From "Exploring and Exploiting Genetic Risk for Psychiatric Disorders," edited by Joshua A. Gordon and Elisabeth B. Binder. Strüngmann Forum Reports, vol. 31, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262547383 (paperback) 9780262377423 (pdf)

gene-level summary in a genome-wide fashion, allowing characterization of subthreshold loci or relatively low-powered GWAS.

Integrative Multi-omic Approaches

Advances in multi-omic experimental approaches and computational integrative analytic tools allow us to refine regulatory relationships between variants and genes. For example, chromatin interaction profiles can be combined with other epigenomic profiling to capture a specific type of chromatin interaction. HiChIP, a hybrid of Hi-C and ChIP-seq, not only identifies likely target genes of genetic variants on the basis of enhancer-promoter interactions, it also functionally annotates variants (i.e., whether variants reside in enhancers) (Mumbach et al. 2017). Similar to the idea of HiChIP, the activity-by-contact (ABC) model has been developed to combine enhancer activity and chromatin contact frequency to infer variant function (Fulco et al. 2019). This model proposes that the magnitude of variant effect and its target gene can be better predicted by taking both enhancer activity and chromatin contact frequency into account. The resulting ABC score was shown to outperform alternative models for predicting enhancer-promoter relationships. In addition, single-cell multi-omic profiling that integrates scRNA sequencing with scATAC sequencing allows us to infer relationships between chromatin accessibility and gene expression, which can be further extended to refine variant-gene expression relationships (Ma et al. 2020). Together, simultaneous dissection of multilayered molecular phenotypes can synergistically bridge the gap between variants and genes.

Beyond Cataloging

Assigning variants to genes they regulate is the first step to understand mechanisms by which common variants contribute risk for psychiatric disorders. The next step is to study the function of risk genes, which is in itself a challenge, especially considering the large number of genes identified from GWAS (e.g., H-MAGMA identified >5,000 genes associated with schizophrenia; Sey et al. 2020). Furthermore, different genomic approaches can give different sets of risk genes, resulting in a growing list of risk genes reported from different studies. So far, the field lacks a gold standard set of risk genes that can be benchmarked against risk genes identified from different genomic approaches. Transcriptomic data sets can be used to create such benchmarks by checking whether GWAS risk genes are dysregulated in psychiatric conditions. However, the majority of transcriptomic data sets are obtained from postmortem brain tissue, which does not accurately represent the onset of disease. Indeed, differential expression signatures from postmortem brain samples of schizophrenia showed weak to moderate association with schizophrenia heritability. This

From "Exploring and Exploiting Genetic Risk for Psychiatric Disorders," edited by Joshua A. Gordon and Elisabeth B. Binder. Strüngmann Forum Reports, vol. 31, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262547383 (paperback) 9780262377423 (pdf)

suggests that the accumulation of environmental exposures over the life course can dilute the effects of genetic risk factors on transcriptomic signatures (Yu et al. 2021). Human-induced pluripotent stem cells (hiPSCs) derived from individuals with psychiatric disorders could fill this gap by providing transcriptomic signatures in the early phase of the disease directly impacted by genetic risk factors. While hiPSC-based disease models have limitations, such as reduced cell maturity and inter-lab variance due to nonstandardized differentiation protocols, they still offer an accessible way to explore molecular landscapes of polygenic disorders that cannot be recapitulated by animal models. To implement transcriptomic signatures from hiPSC-derived models as a benchmark for GWAS risk genes, a standardized differentiation protocol as well as a large collection of hiPSCs with psychiatric disorders will be integral.

Once GWAS risk genes are prioritized, their function can be interrogated using traditional toolkits developed for studying monogenic disorders (e.g., knocking out a gene and measuring cellular and behavioral phenotypes). However, given the polygenic architecture of psychiatric disorders, their functional characterization demands a systems genetics approach. One such approach, Perturb-seq, provides a scalable method to test the functional impact of dozens of genes in a single assay by combining CRISPR gene perturbation with scRNA sequencing (Jin et al. 2020) (Figure 10.3). Single-cell expression profiling after CRISPR-mediated gene editing enables us to identify cell type-specific alterations in the transcriptomic architecture upon perturbation of psychiatric risk genes. Furthermore, *in vivo* application in animal models of Perturb-seq across developmental epochs or in response to specific external stimuli would allow temporal frame- and context-specific dissection of gene function.

Conclusion

A large compendium of common variants has been identified via GWAS to be associated with psychiatric disorders, proving their central role in psychiatric genetics. Common variants present multilayered challenges to functional interrogation due to their correlation structure, enrichment in the noncoding

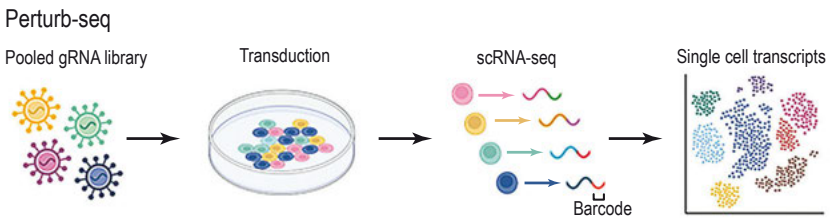

Figure 10.3 Perturb-seq merges a pooled CRISPR screen and scRNA sequencing for high-throughput identification of perturbation effects by cell type.

From "Exploring and Exploiting Genetic Risk for Psychiatric Disorders," edited by Joshua A. Gordon and Elisabeth B. Binder. Strüngmann Forum Reports, vol. 31, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262547383 (paperback) 9780262377423 (pdf)

genome, and polygenic architecture. Such problems warrant tailored solutions for thorough investigation of polygenic disease etiology. One critical challenge lies in identifying causal variants and interrogating variant function, which can be advanced by high-throughput experimental strategies such as MPRA and CRISPR screens. However, the field needs to reach a consensus on how to define causal variants. Another critical challenge is to link variants to targetable genes. Population-scale molecular assays such as eQTLs have been playing a central role, whereas emerging evidence suggests that currently available eQTL resources alone may not lead to the complete understanding of GWAS. Here, we discussed complementary genomic resources to fill this gap (e.g., QTL resources that span multiple molecular assays, developmental stages, and cell types; multi-omic data sets that combine chromatin accessibility and interactions in a cell type-specific fashion). Finally, variant function needs to be extended to biological underpinnings that go beyond cataloging a new gene list. In linking variants to biology, we need to profile joint effects of the variants rather than characterizing individual variant effects to truly embrace the polygenic architecture of psychiatric disorders.

From "Exploring and Exploiting Genetic Risk for Psychiatric Disorders," edited by Joshua A. Gordon and Elisabeth B. Binder. Strüngmann Forum Reports, vol. 31, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262547383 (paperback) 9780262377423 (pdf)