Chapter 8

Genome-wide association study biomarkers in bipolar disorder

Gustavo C. Medeiros and Fernando S. Goes
Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, MD, United States

8.1 Introduction

As most human traits—including those relating to cognition, emotions, and behavior—show prominent evidence of heritability, genetic biomarkers have the potential to improve prevention and diagnosis of psychiatric disorders, particularly, of the more highly heritable conditions such as bipolar disorder (BD) (Polderman et al., 2015). Moreover, compared to other types of biomarkers, genetic biomarkers have a number of specific advantages: they can be obtained from virtually any tissue source; have a temporal primacy that facilitates causal inferences; and can be potentially informative throughout the lifespan. Nevertheless, although twin and family studies, some dating from almost a century ago, have indicated that BD is among the most heritable psychiatric disorders (Smoller & Finn, 2003), the specific genetic risk factors involved in BD heritability have remained largely unknown until the recent emergence of large-scale genome-wide association studies (GWAS). In contrast to prior molecular genetic studies, GWAS have allowed for a broad evaluation of genetic variation, which has been necessary to address the complex etiology of psychiatric disorders, where a large number of small-effect variables are likely to be involved (Tam et al., 2019). GWAS have now identified many common variants associated with BD, laying the foundations to discover causally grounded pathophysiology, and to potentially identify novel therapeutic targets.

In this chapter we will provide a brief introduction to GWAS, summarize their main findings, and discuss the areas of greatest translational potential in the years ahead.

Biomarkers in Bipolar Disorders. DOI: https://doi.org/10.1016/B978-0-12-821398-8.00016-3

8.2 Generic epidemiology: heritability and family studies

The primary starting point for the investigation of genetic markers as risk factors is the observation that family members, by virtue of their greater likelihood to sharing identical stretches of DNA, find themselves at increased risk of developing a specific disorder. Among the different family studies, the most genetically informative are twin studies, which compare the risk of illness in monozygotic twins—who essentially share 100% of their DNA—to dizygotic twins—who share only 50% of their DNA. As both types of twins are assumed to share an identical, or almost identical, environment, comparing the rates disorders between monozygotic and dizygotic twins provides an estimate of the overall phenotypic heritability (h^2), a parameter that indicates the degree to which phenotypic variability can be explained by genetic differences (Visscher, Hill, & Wray, 2008). However, high heritability does not itself imply the presence of genes of high effect (more typical of mendelian disorders) and, in many disorders (including BD), the heritability is likely to be composed of a very large number of genetic variants, across the frequency spectrum and mostly characterized by small-effect sizes (Visscher et al., 2008).

Twin and family studies in BD have found heritability estimates to be between 60% and 80%, with significantly increased risk in first-degree relatives (relative risk $\sim 4-8$) (Smoller & Finn, 2003). More recent twin studies using large national registries have found broadly consistent estimates compared to the earlier investigations, with the added advantage of being more representative of the overall population (Lichtenstein et al., 2009; Song et al., 2014). In addition, due to their much larger sample sizes, such registry-based family studies have been able to provide a more comprehensive picture of BD heritability, and of coheritability of BD with other disorders. The results have revealed the familial transmission of both a specific vulnerability toward BD, as well as a more generalized increased risk for other types of psychopathology (Song et al., 2014), reflecting a shared pattern of common variant risk that has also emerged from GWAS (Anttila et al., 2018; Mullins et al., 2020) (described below).

8.3 Genome-wide association studies

The human genome is composed of $\sim$3.2 billion base pairs within each set of 22 autosomal and two sex chromosomes. Recent studies of tens of thousands of individuals have found evidence for variation, mostly single base substitution or small-scale insertion or deletions (indel), found at least once every seven base pairs (Taliun et al., 2021). When analyzed in a sample or population, the majority of these variants are classified as rare (defined as being present in less than 0.5% of a sample or population), with approximately half of all variants being seen only once in sample as large as

$\sim$50,000 s individuals ("singleton variants"). In contrast, most variants found in a single individual represent common variation and are termed *polymorphisms*. Depending on the population of ancestry, individuals have $\sim$3.5–4 million variants, of which the vast majority ($>$95%) are classified as a common (minor allele frequency $>$ 0.5%) (Consortium, 1000 Genomes Project et al., 2012). Most polymorphisms are found in the noncoding region (intergenic) and represent evolutionarily ancient variation shared across ethnicities that have risen to high frequencies largely due to chance events ("population drift") rather than balancing selection.

GWAS microarrays assay anywhere from 1 to 2.5 million genotypes. However, because of the correlation across nearby markers (i.e., linkage disequilibrium -LD), nearby genotypes that are not directly genotype can nevertheless be accurately inferred, or imputed, using information from a population-specific reference genotype data source. Importantly, imputation can provide highly reliable predictions of nearby markers, thereby, allowing metaanalyses across studies originally genotyped on different micro-arrays. Following imputation and quality control, a GWAS metaanalysis will usually have approximately 8–10 million imputed markers. Subsequently each individual marker is tested, usually in case-control framework, with covariates to correct for ancestry and/or batch effects. To account for multiple testing, an association is considered to be significant if its p-value is less than 5×10^{-8}, reflecting a correction for approximately one million independent common variants. Although stringent, this p-value threshold has been essential to ensure a high degree of subsequent replication in independent samples.

GWAS have now been performed on essentially all biomedical phenotypes and/or traits and, almost invariably, when sample sizes reach a certain threshold, genome-wide significant loci have been discovered (Claussnitzer et al., 2020). With a few exceptions, samples sizes for an initial discovery need to be large (at least several thousands to tens of thousands) since the effect sizes are usually modest, with most significant GWAS findings having odds-ratios (ORs) in the $\sim$1.05–1.30 range.

8.4 Recent results of genome-wide association studies in bipolar disorder

Initial GWAS of BD began to be published in 2007 (Consortium, 2007; Sklar et al., 2008) but robust associations only emerged as GWAS datasets were combined in larger meta-analyses. The first such meta-analysis by the Psychiatric Genomics Consortium (PGC) BD workgroup (PGC BD 1) included 7481 participants with BD and 9250 controls (Sklar et al., 2011), and identified two genome-wide significant loci: (1) a locus associated with the *CACNA1C* gene, which encodes a subunit of voltage-gated calcium channels, and (2) a locus related to the *ODZ4* gene (also known as teneurin

transmembrane protein 4), which encodes a brain-expressed transmembrane protein involved in oligodendrocyte differentiation.

The second study of the PGC BD workgroup (PGC BD 2) was published in 2019, and was able to identify 30 loci in a sample of 20,352 cases and 31,358 controls (Stahl et al., 2019). The third and most recent analysis by the group (PGC BD 3) found associations between BD and 64 loci in a meta-analysis of 41,917 BD cases and 371,549 controls (Mullins et al., 2020). These 64 genome-wide significant loci identified in PGC BD 3 (shown in Fig. 8.1) included the two loci initially discovered in PGC BD 1, as well as 28 of the 30 loci from the PGC BD 2. Among the genome-wide significant findings in PGC BD 3, the mean OR was 1.07 (range 1.05−1.15), and the mean risk allele frequency was 0.47 (0.06−0.93). These findings reinforce that GWAS in BD have small effects and, mostly, involve very common risk alleles.

In PGC BD 3 meta-analysis, the strongest associations were: (1) a locus on chromosome 3 encompassing the *TRANK1* gene, which encodes a protein of unclear function expressed in the brain and other tissues; (2) an intergenic locus on chromosome 6; (3) a locus on chromosome 12 enclosing the *CACNA1C* gene, originally identified in PGC 1 BD and PGC 2 BD analyses; and (4) a locus in the HLA region, albeit independent of the schizophrenia-associated locus that has been fine-mapped to the complement gene *C4*.

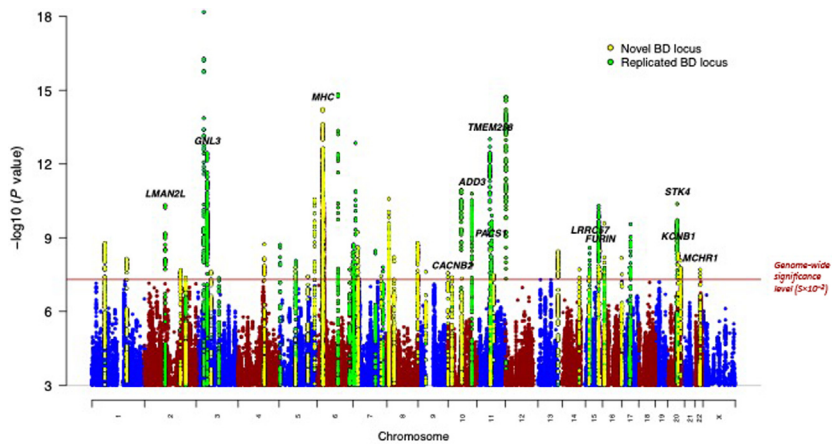

FIGURE 8.1 Manhattan plot showing the results of the most recent genome-wide association study meta-analysis of the Psychiatric Genomics Consortium Bipolar Disorder Working Group. The *X* axis represents markers across the genomes, with the *Y* axis representing the -log *p*-value of each markers association. Sixty-four (loci) cross the level of genome-wide significance. *From Mullins, N., Forstner, A. J., O'Connell, K. S., Coombes, B., I. Coleman, J. R., Qiao, Z., … Andreassen, O. A. (2020). Genome-wide association study of over 40,000 bipolar disorder cases provides novel biological insights [Preprint].* Psychiatry and Clinical Psychology. https://doi.org/10.1101/2020.09.17.20187054.

Other significant loci were found in genes with known functions in neurons including the sodium voltage-gated channel *SCN2A*, the postsynaptic density protein *SHANK2*, and the glutamate receptor subunit *GRIN2A*. The replication of the initial associated locus in the calcium voltage-gated gene *CACNA1C*, and the identification of novel locus related to the calcium voltage-gated channel gene *CACNB2* further strengthened the hypothesis that neuronal voltage-gated calcium channels may play a role in the etiology of BD. Subsequent analyses found strong enrichment for genes expressed in the brain, particularly those related to calcium-signaling, synaptic functioning, and neurogenesis. Not unexpectedly, several genes associated with BD in PGC BD 3, particularly those related to neuronal function, were also seen in the GWAS of other psychiatric disorders, particularly schizophrenia (Consortium, Ripke, Walters, & O'Donovan, 2020) and major depression (Howard et al., 2019).

8.5 Fine mapping: from loci to genes

GWAS can point to a locus, or *region* of the chromosome, but it does not usually have the resolution to pinpoint the specific causal marker(s) responsible for the statistical association. Determining the precise variant(s) usually requires additional statistical analysis and/or further functional experimentation through a strategy broadly knows as *fine mapping* (Broekema, Bakker, & Jonkers, 2020). The major challenge of fine mapping is the correlation of nearby genetic markers from LD. Since LD is strongly determined by demographic history, the inclusion of diverse populations, each with their own somewhat distinct pattern of LD, can facilitate fine mapping as regions of interest for further study can be restricted to smaller regions that show overlap in LD across various populations (Pereira, Mutesa, Tindana, & Ramsay, 2021). Unfortunately, the vast majority of GWAS have been so far conducted in subjects of European Ancestry, a population with relatively high levels of LD, hindering our ability to fine-map GWAS loci. This is an important premise for the current impetus to increase the ancestral diversity of GWAS (Martin et al., 2019).

A complementary approach to help prioritizing causal variants in a locus is to overlay and statistically "integrate" genetic data with nearby molecular features such as gene expression and epigenetic features (e.g., methylation, histone functioning, and chromatic accessibility). Such studies, broadly known as quantitative trait loci (QTL) mapping studies, are usually more limited in size compared to GWAS because they require access to the relevant tissue of interest (e.g., postmortem brain samples). The most common type of QTL mapping studies involve the use of gene-expression measures (*e*xpression or *e*QTL studies) in blood or brain tissues and, when combined with GWAS association results, are termed transcript-wide association studies (TWAS) (Wainberg et al., 2019). TWAS have the potential advantages of

focusing on the gene (rather than a locus) as the level of analysis, which is both more interpretable and requires a less stringent level of correction of multiple testing. In the PGC BD 3 study, for example, a TWAS analysis using a brain gene-expression dataset led to the identification of 77 significantly associated *genes* compared to 64 genome-wide associated *loci* in the GWAS parent study. Evidence for a significant QTL association suggests that the associated loci may be at least partially mediated by the relevant molecular features, and aids in prioritizing further analyses or experiments needed to strengthen causal inference.

8.6 Common variant heritability and coheritability

Due to their small-effect sizes, genome-wide significant loci account for only a very small proportion of the overall heritability ($\sim$1%–2%). However, using mixed models initially described by Yang and colleagues, it is possible to estimate an SNP-based overall measures of heritability (h^2_{SNP}), which predicts the proportion of heritability attributable to all common variants (Yang et al., 2010). SNP-based heritability is expected to be less than twin-based measures of heritability because the former represents only the effect of additive common variant-based heritability, in contrast to the twin-based phenotypic estimates that can include the effects of rare variants and nonadditive interactions. Using the PGC BD 3 data, h^2_{SNP} was estimated to be 18.6% (SE = 0.008), indicating that approximately one-quarter to one-third of the phenotypic heritability (estimated to be $\sim$60%–80%) may ultimately be discovered by common variants (see "Missing Heritability" below). Interestingly, the PGC BD 3 SNP-based heritability estimate was less than that estimated by the PGC BD 1 (27%, SE = 0.03) and PGC BD 2 (24%, SE = 0.01), potentially reflecting the greater phenotypic heterogeneity of more recent samples that, for example, include a greater proportion of BD-type II cases. Compared with BD-type I, samples with BD-type II were found to have significantly lower estimates of h^2_{SNP} (11.6% vs 20.9%).

Measures of common variant heritability can additionally provide a means to test for the degree of shared common variant risk, or coheritability, across psychiatric and medical conditions. In BD, there is strong evidence for common variant-based coheritability with schizophrenia (0.69 $\pm$ 0.02 s.e.), and major depressive disorder (0.48 $\pm$ 0.03 s.e.). A lower, but still significant, coheritability has been observed with attention-deficit hyperactivity disorder (0.21 $\pm$ 0.04 s.e.), anorexia nervosa (0.20 $\pm$ 0.03 s.e.), and autism spectrum disorder (0.21 $\pm$ 0.042 s.e.). Although these genetic correlations may provide important nosological insights, such interpretations must be made with caution, since they may partially reflect the influence of potential confounds, or of shared but unaccounted phenotypes, such as substance abuse or neuroticism.

8.7 Polygenic risk scores

GWAS results may be used to calculate polygenic risk scores (PRS), which are aggregate measures of genetic risk that are derived by creating a summary of the overall number of associated risk alleles (Fig. 8.2).

PRS utilize association thresholds that are usually much more liberal than the genome-wide significance thresholds used for individual locus identification. Although alleles associated at subthreshold levels of association are individually unlikely to represent true positives, when analyzed collectively, they show strong evidence for aggregate association in independent replication samples. PRS are derived by initially using the results of a GWAS meta-analysis as a *training dataset* to identify risk alleles and their effect sizes. Subsequently, individual level risk scores are calculated in an independent *testing dataset* using a weighted sum of the risk alleles identified in the training dataset (Baselmans, Yengo, van Rheenen, & Wray, 2021). A polygenic association is calculated by comparing the weighted sum of risk alleles in cases versus controls, with the strength of the overall polygenic association usually represented by a pseudo-R^2 statistic, representing the phenotyping variability explained by the polygenic scores in the independent case-control sample. As pseudo-R^2 values are affected the proportion of cases in the GWAS, a calibrated version of the pseudo-R^2, known as the liability pseudo-R^2, is usually reported to allow comparability across studies. In the most recent PGC BD GWAS meta-analysis, the PRS has been found to have pseudo-R^2 scores of $\sim$ 4.5%, representing an intermediate effect compared to the pseudo-R^2 values from the latest PGC schizophrenia PRS and the PGC unipolar depression PRS, which have explained $\sim$8% and $\sim$ 2%−3% of the variance, respectively.

Although the magnitude of the effect of these PRS remains modest, the expectation is that, with increased sample sizes, the explanatory power of the PRS may approximate their maximal potential as determined by the genome-wide common variant heritability (h^2_{SNP}). An important limitation of PRS, however, is their poor utility or "portability" across ancestries. In PGC BD 3, for example, PRS explained $\sim$4.75% of the phenotypic variance in European cohorts but its performance was lower in individuals of East Asian (1.9%−2.3%) or African American ancestry (0.4%−1.2%). Although novel methods that leverage functional information as well as ancestry-specific LD patterns may mitigate the loss of predictive power across ancestries, the simplest and most full proof solution is to increase representation of non-European ancestry participants in current and future studies. This represents a major focus of ongoing research, for both scientific and health equity reasons (Wand et al., 2021).

As sample sizes increase and become more geographically representative, PRS may have a role in predictive models of clinically relevant phenotypes. In other common complex disorders, PRS have been found to have

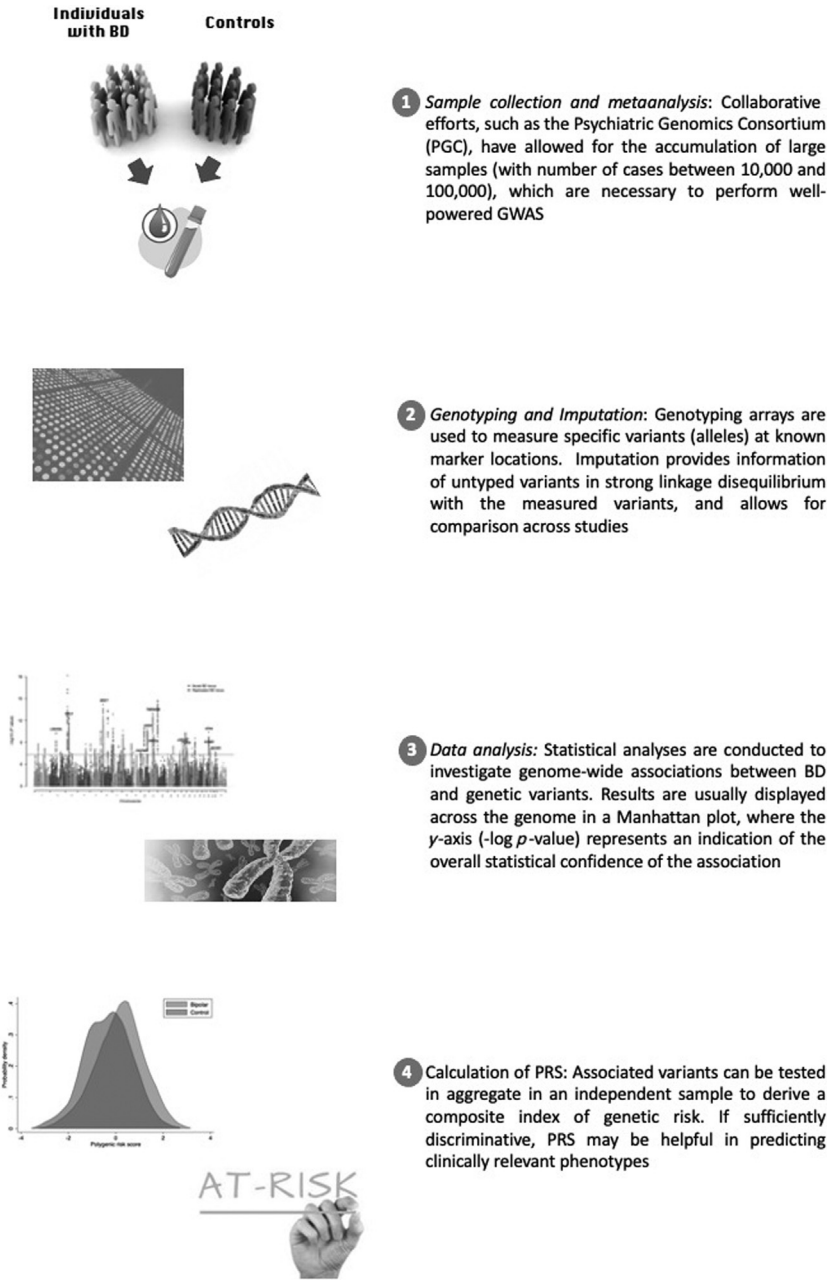

FIGURE 8.2 A broad schematic of the major steps involved in the generation of genome-wide association studies and the derivation of polygenic risk scores.

incremental predictive ability that may aid clinical decision making, such as earlier treatment for the prevention of coronary heart disease (Khera et al., 2018) and earlier screening guidelines in common cancers (Mars et al., 2020). In psychiatric disorders, there is preliminary data suggesting that certain clinical decisions—such as treatment response to antipsychotics (Zhang et al., 2018) or prediction of bipolarity in high-risk subjects with depression (Liebers, Pirooznia, Ganna, Bipolar Genome Study BiGS, & Goes, 2020)—could benefit from the use of PRS; however, the state of the evidence remains premature for clinical use and should ultimately be validated using randomized clinical trials.

8.8 Missing heritability

As noted above, while genome-wide significant loci currently only account for a very limited proportion of heritability, modeling of the overall degree of common variant heritability (h^2_{SNP}) indicates that, as GWAS sample sizes increase, the proportion of heritability explained by genome-wide significant loci will increase, eventually approximating the upper bound limits estimated by common variant heritability ($h^2_{SNP} \sim 18.6\%$). Yet, even under optimistic scenarios, this upper bound limit is much lower than the estimates of heritability from twin samples (estimated at $\sim 60\% - 80\%$), leading to an explanatory gap that has been dubbed as the "missing heritability."

Although what underlies the major components of missing heritability remains a matter of debate (Zuk et al., 2014), recent evidence from easily measured phenotypes—such as height—has provided support for the role of low-frequency variants. However, rare variation is poorly correlated with common variants and, therefore is not "tagged" by the common variants measured in GWAS. Instead, they must be measured directed using large-scale sequencing studies, which are beginning to emerge in BD. While they have initially focused on the more interpretable exonic regions of the genome (Jia et al., 2021; Palmer et al., 2021), rare-variant studies will eventually transition to whole genome sequencing as decreasing costs make large-scale sequencing studies more feasible.

8.9 Genome-wide association studies and treatment of bipolar disorder (pharmacogenetics)

Arguably, the most translatable genetic biomarkers are those related to the efficacy and/or tolerance of pharmacological treatments (Bousman et al., 2021). The use of genetics to inform pharmacotherapy, known as pharmacogenetics, has had greater success identifying potential markers associated with drug levels (pharmacokinetics) than markers associated with therapeutic drug functioning (pharmacodynamics). Compared to case-control GWAS, pharmacogenetic studies have generally been smaller in scope, and focused

on a small number of genetic markers rather than a comprehensive genome-wide assay. More recently, there has been an increased application of the more rigorous randomized controlled trial methodology to compare a pharmacogenetically informed treatment with treatment as usual (Greden et al., 2019; Jürgens et al., 2020, p. 6). However, these studies, with sample sizes in the hundreds, have not shown significant improvement in outcomes with pharmacogenetic testing, suggesting the need for larger sample sizes and/or more comprehensive pharmacogenetic assays. For antidepressant response, there is now emerging evidence that clinical phenotypes—such as response or remission rates—are likely to be polygenic and, therefore more appropriately tested with genome-wide assays and polygenic modeling (Pain et al., 2020).

In BD, more specifically, there have been initial attempts to identify common variants associated with response to lithium and other mood stabilizers (Ho et al., 2020; Hou et al., 2016; Song et al., 2016). However, these initial efforts represent secondary analyses from previously collected samples and remain as yet unreplicated. Moreover, similarly to disorder-related GWAS, the effect sizes of individual loci in pharmacogenetic studies are modest and insufficient for clinical translation, echoing the likely necessity of polygenic models in prognostic or predictive studies.

8.10 Challenges and opportunities for the next era of genome-wide association studies in bipolar disorder

The current GWAS era has led to the identification of common variants associated with almost all major psychiatric disorders. However, the small-effect sizes and the challenge of linking a locus to a clearly identifiable downstream biological function have led to the realization that the discovery of a genome-wide association represents only the beginning of a potentially long path of discovery. The ultimate objective of GWAS is the identification of a sufficiently actionable disease-related biological target that will inform therapeutic hypotheses, and lead to novel drug discovery. Thus far, most of the examples of genetically informed novel drug discovery have emerged from rare-variant studies (Spreafico, Soriaga, Grosse, Virgin, & Telenti, 2020), although there is an expectation that, as GWAS sample sizes increase and fine-mapping efforts continue, there will be more direct links to actionable biology (King, Davis, & Degner, 2019). Particularly informative is the potential for an "allelic series" of associated variants ranging from common to rare variants that provide the statistical backing of common variant studies with the more precise mechanistic insights of rare-variant studies. In schizophrenia, for example, there is now strong genome-wide evidence for an association with common *and* rare variants in the genes *GRIN2A* and *SP4*. Rare-variant studies of BD have been smaller in scale than those of schizophrenia and, while there is emerging evidence for convergence to broad gene-sets

such as the postsynaptic density, there is not as yet evidence for convergence with the recent GWAS meta-analysis results (Goes et al., 2019; Palmer et al., 2021).

In addition to the impetus to increase samples sizes for further discovery of novel loci, there is a strong emphasis on greater ethnic representativeness to leverage the genetic diversity of global populations, ensuring that novel pathophysiological insights have the potential to improve the health of all populations. Such ethnically diverse studies are being increasingly carried out in lower- and middle-income countries (Stevenson et al., 2019), as well as in high-income countries (Bigdeli et al., 2021), where ethnic minorities are underrepresented in research studies despite suffering from a disproportioned disease burden (Akinhanmi et al., 2018).

Finally, as sample size requirements become larger and larger, there is an important need to reconsider the appropriateness and sufficiency of phenotypic characterization (Cai et al., 2020). The common phenotypic overlap of BD with many of the major psychiatric syndromes (e.g., psychosis, anxiety, mood instability, substance abuse) cautions against increasing reliance on "light" phenotypic approaches that may be more associated with comorbid conditions rather than the primary diagnosis of BD (Baldessarini, 2000).

8.11 Conclusions

Following the disappointment of the genome-wide linkage and candidate gene eras, the transition to the GWAS era has been marked by an unprecedented focus on large-scale and highly reproducible science that has led to robust loci discovery in all major psychiatric syndromes, including BD. GWAS, arguably, represent the most important "entry point" into the underlying biology of mental disorders. However, for its translation potential to be realized, there is now a pressing need to link novel loci discovery with functional insights and clinically relevant phenotypes, which ultimately may lead to novel drug discovery and improved clinical outcomes.

References

Akinhanmi, M. O., Biernacka, J. M., Strakowski, S. M., McElroy, S. L., Balls Berry, J. E., Merikangas, K. R., ... Frye, M. A. (2018). Racial disparities in bipolar disorder treatment and research: A call to action. *Bipolar Disorders*, *20*(6), 506–514. Available from https://doi.org/10.1111/bdi.12638.

Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., Duncan, L., ... Malik, R. (2018). Analysis of shared heritability in common disorders of the brain. *Science (New York, N.Y.)*, *360*(6395).

Baldessarini, R. J. (2000). A plea for integrity of the bipolar disorder concept. *Bipolar Disorders*, *2*(1), 3–7. Available from https://doi.org/10.1034/j.1399-5618.2000.020102.x.

Baselmans, B. M. L., Yengo, L., van Rheenen, W., & Wray, N. R. (2021). Risk in relatives, heritability, SNP-based heritability, and genetic correlations in psychiatric disorders: A review.

Biological Psychiatry, *89*(1), 11−19. Available from https://doi.org/10.1016/j.biopsych.2020.05.034.

Bigdeli, T. B., Fanous, A. H., Li, Y., Rajeevan, N., Sayward, F., Genovese, G., … Harvey, P. D. (2021). Genome-wide association studies of schizophrenia and bipolar disorder in a diverse cohort of United States veterans. *Schizophrenia Bulletin*, *47*(2), 517−529. Available from https://doi.org/10.1093/schbul/sbaa133.

Bousman, C. A., Bengesser, S. A., Aitchison, K. J., Amare, A. T., Aschauer, H., Baune, B. T., … Müller, D. J. (2021). Review and consensus on pharmacogenomic testing in psychiatry. *Pharmacopsychiatry*, *54*(1), 5−17. Available from https://doi.org/10.1055/a-1288-1061.

Broekema, R. V., Bakker, O. B., & Jonkers, I. H. (2020). A practical view of fine-mapping and gene prioritization in the post-genome-wide association era. *Open Biology*, *10*(1), 190221. Available from https://doi.org/10.1098/rsob.190221.

Cai, N., Revez, J. A., Adams, M. J., Andlauer, T. F. M., Breen, G., Byrne, E. M., … Flint, J. (2020). Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nature Genetics*, *52*(4), 437−447. Available from https://doi.org/10.1038/s41588-020-0594-5.

Claussnitzer, M., Cho, J. H., Collins, R., Cox, N. J., Dermitzakis, E. T., Hurles, M. E., … McCarthy, M. I. (2020). A brief history of human disease genetics. *Nature*, *577*(7789), 179−189. Available from https://doi.org/10.1038/s41586-019-1879-7.

Consortium, T. S. W. G. of the P. G., Ripke, S., Walters, J. T., & O'Donovan, M. C. (2020). Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. *MedRxiv*, 2020.09.12.20192922. Available from https://doi.org/10.1101/2020.09.12.20192922.

Consortium, W. T. C. C. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, *447*(7145), 661−678.

Consortium, 1000 Genomes Project., Abecasis, G. R., Auton, A., Brooks, L. D., DePristo, M. A., Durbin, R. M., … McVean, G. A. (1000). An integrated map of genetic variation from 1,092 human genomes. *Nature*, *491*(7422), 56−65. Available from https://doi.org/10.1038/nature11632.

Goes, F. S., Pirooznia, M., Tehan, M., Zandi, P. P., McGrath, J., Wolyniec, P., … Pulver, A. E. (2019). De novo variation in bipolar disorder. *Molecular Psychiatry*. Available from https://doi.org/10.1038/s41380-019-0611-1.

Greden, J. F., Parikh, S. V., Rothschild, A. J., Thase, M. E., Dunlop, B. W., DeBattista, C., … Dechairo, B. (2019). Impact of pharmacogenomics on clinical outcomes in major depressive disorder in the GUIDED trial: A large, patient- and rater-blinded, randomized, controlled study. *Journal of Psychiatric Research*, *111*, 59−67. Available from https://doi.org/10.1016/j.jpsychires.2019.01.003.

Ho, A. M.-C., Coombes, B. J., Nguyen, T. T. L., Liu, D., McElroy, S. L., Singh, B., … Biernacka, J. M. (2020). Mood-stabilizing antiepileptic treatment response in bipolar disorder: A genome-wide association study. *Clinical Pharmacology & Therapeutics*, *108*(6), 1233−1242. Available from https://doi.org/10.1002/cpt.1982.

Hou, L., Heilbronner, U., Degenhardt, F., Adli, M., Akiyama, K., Akula, N., … Schulze, T. G. (2016). Genetic variants associated with response to lithium treatment in bipolar disorder: A genome-wide association study. *The Lancet*, *387*(10023). Available from https://doi.org/10.1016/S0140-6736(16)00143-4.

Howard, D. M., Adams, M. J., Clarke, T.-K., Hafferty, J. D., Gibson, J., Shirali, M., … McIntosh, A. M. (2019). Genome-wide *meta*-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, *22*(3), 343−343. Available from https://doi.org/10.1038/s41593-018-0326-7.

Jia, X., Goes, F. S., Locke, A. E., Palmer, D., Wang, W., Cohen-Woods, S., … Scott, L. J. (2021). Investigating rare pathogenic/likely pathogenic exonic variation in bipolar disorder. *Molecular Psychiatry*. Available from https://doi.org/10.1038/s41380-020-01006-9.

Jürgens, G., Andersen, S. E., Rasmussen, H. B., Werge, T., Jensen, H. D., Kaas-Hansen, B. S., & Nordentoft, M. (2020). Effect of routine cytochrome P450 2D6 and 2C19 genotyping on antipsychotic drug persistence in patients with schizophrenia: A randomized clinicaltTrial. *JAMA Network Open*, *3*(12), e2027909. Available from https://doi.org/10.1001/jamanetworkopen.2020.27909.

Khera, A. V., Chaffin, M., Aragam, K. G., Haas, M. E., Roselli, C., Choi, S. H., … Kathiresan, S. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics*, *50*(9), 1219−1224. Available from https://doi.org/10.1038/s41588-018-0183-z.

King, E. A., Davis, J. W., & Degner, J. F. (2019). Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLOS Genetics*, *15*(12), e1008489. Available from https://doi.org/10.1371/journal.pgen.1008489.

Lichtenstein, P., Yip, B. H., Bjork, C., Pawitan, Y., Cannon, T. D., Sullivan, P. F., & Hultman, C. M. (2009). Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: A population-based study. *Lancet*, *373*(9659), 234−239. Available from https://doi.org/10.1016/S0140-6736(09)60072-6.

Liebers, D. T., Pirooznia, M., Ganna, A., Bipolar Genome Study (BiGS)., & Goes, F. S. (2020). Discriminating bipolar depression from major depressive disorder with polygenic risk scores. *Psychological Medicine*, 1−8. Available from https://doi.org/10.1017/S003329172000015X.

Mars, N., Koskela, J. T., Ripatti, P., Kiiskinen, T. T. J., Havulinna, A. S., Lindbohm, J. V., … Ripatti, S. (2020). Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. *Nature Medicine*, *26*(4), 549−557. Available from https://doi.org/10.1038/s41591-020-0800-0.

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., & Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics*, *51*(4), 584−591. Available from https://doi.org/10.1038/s41588-019-0379-x.

Mullins, N., Forstner, A. J., O'Connell, K. S., Coombes, B., I. Coleman, J. R., Qiao, Z., … Andreassen, O. A. (2020). Genome-wide association study of over 40,000 bipolar disorder cases provides novel biological insights [Preprint]. *Psychiatry and Clinical Psychology*. Available from https://doi.org/10.1101/2020.09.17.20187054.

Pain, O., Hodgson, K., Trubetskoy, V., Ripke, S., Marshe, V. S., Adams, M. J., … Lewis, C. M. (2020). Antidepressant response in major depressive disorder: A genome-wide association study. *MedRxiv*, 2020.12.11.20245035. Available from https://doi.org/10.1101/2020.12.11.20245035.

Palmer, D. S., Howrigan, D. P., Chapman, S. B., Adolfsson, R., Bass, N., Blackwood, D., … Neale, B. M. (2021). Exome sequencing in bipolar disorder reveals shared risk gene *AKAP11* with schizophrenia. *MedRxiv*, 2021.03.09.21252930. Available from https://doi.org/10.1101/2021.03.09.21252930.

Pereira, L., Mutesa, L., Tindana, P., & Ramsay, M. (2021). African genetic diversity and adaptation inform a precision medicine agenda. *Nature Reviews. Genetics*, 1−23. Available from https://doi.org/10.1038/s41576-020-00306-8.

Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). *Meta*-analysis of the heritability of human traits

based on fifty years of twin studies. *Nature Genetics*, *47*(7), 702–709. Available from https://doi.org/10.1038/ng.3285.

Sklar, P., Smoller, J. W., Fan, J., Ferreira, M. A., Perlis, R. H., Chambert, K., … Purcell, S. M. (2008). Whole-genome association study of bipolar disorder. *Molecular Psychiatry*, *13*(6), 558–569. Available from https://doi.org/10.1038/sj.mp.4002151.

Sklar, P., Ripke, S., Scott, L. J., Andreassen, O. A., Cichon, S., Craddock, N., … Psychiatric GWAS Consortium Bipolar Disorder Working Group. (2011). Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nature Genetics*, *43*(10), 977–983. Available from https://doi.org/10.1038/ng.943.

Smoller, J. W., & Finn, C. T. (2003). Family, twin, and adoption studies of bipolar disorder. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics*, *123C*(1), 48–58. Available from https://doi.org/10.1002/ajmg.c.20013.

Song, J., Bergen, S. E., Di Florio, A., Karlsson, R., Charney, A., Ruderfer, D. M., … Belliveau, R. A. (2016). Genome-wide association study identifies SESTD1 as a novel risk gene for lithium-responsive bipolar disorder. *Molecular Psychiatry*, *21*(9), 1290–1297. Available from https://doi.org/10.1038/mp.2015.165.

Song, J., Bergen, S. E., Kuja-Halkola, R., Larsson, H., Landen, M., & Lichtenstein, P. (2014). Bipolar disorder and its relation to major psychiatric disorders: A family-based study in the Swedish population. *Bipolar Disorders*, *17*(2), 184–193. Available from https://doi.org/10.1111/bdi.12242.

Spreafico, R., Soriaga, L. B., Grosse, J., Virgin, H. W., & Telenti, A. (2020). Advances in genomics for drug development. *Genes*, *11*(8). Available from https://doi.org/10.3390/genes11080942.

Stahl, E. A., Breen, G., Forstner, A. J., McQuillin, A., Ripke, S., Trubetskoy, V., … Sklar, P. (2019). Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nature Genetics*, *51*(5), 793–803. Available from https://doi.org/10.1038/s41588-019-0397-8.

Stevenson, A., Akena, D., Stroud, R. E., Atwoli, L., Campbell, M. M., Chibnik, L. B., … Koenen, K. C. (2019). Neuropsychiatric genetics of African populations-psychosis (NeuroGAP-Psychosis): A case-control study protocol and GWAS in Ethiopia, Kenya, South Africa and Uganda. *BMJ Open*, *9*(2). Available from https://doi.org/10.1136/bmjopen-2018-025469.

Taliun, D., Harris, D. N., Kessler, M. D., Carlson, J., Szpiech, Z. A., Torres, R., … Abecasis, G. R. (2021). Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature*, *590*(7845), 290–299. Available from https://doi.org/10.1038/s41586-021-03205-y.

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews. Genetics*, *20*(8), 467–484.

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era—concepts and misconceptions. *Nature Reviews Genetics*, *9*(4), 255–266. Available from https://doi.org/10.1038/nrg2322.

Wainberg, M., Sinnott-Armstrong, N., Mancuso, N., Barbeira, A. N., Knowles, D. A., Golan, D., … Kundaje, A. (2019). Opportunities and challenges for transcriptome-wide association studies. *Nature Genetics*, *51*(4), 592–599. Available from https://doi.org/10.1038/s41588-019-0385-z.

Wand, H., Lambert, S. A., Tamburro, C., Iacocca, M. A., O'Sullivan, J. W., Sillari, C., … Wojcik, G. L. (2021). Improving reporting standards for polygenic scores in risk prediction studies. *Nature*, *591*(7849), 211–219. Available from https://doi.org/10.1038/s41586-021-03243-6.

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., ... Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, *42*(7), 565–569. Available from https://doi.org/10.1038/ng.608.

Zhang, J.-P., Robinson, D., Yu, J., Gallego, J., Fleischhacker, W. W., Kahn, R. S., ... Lencz, T. (2018). Schizophrenia polygenic risk score as a predictor of antipsychotic efficacy in first-episode psychosis. *American Journal of Psychiatry*, *176*(1), 21–28. Available from https://doi.org/10.1176/appi.ajp.2018.17121363.

Zuk, O., Schaffner, S. F., Samocha, K., Do, R., Hechter, E., Kathiresan, S., ... Lander, E. S. (2014). Searching for missing heritability: Designing rare variant association studies. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(4), E455–E464. Available from https://doi.org/10.1073/pnas.1322563111.