HHS Public Access

Author manuscript

Brain Imaging Behav. Author manuscript; available in PMC 2018 August 01.

Published in final edited form as:

Brain Imaging Behav. 2018 August ; 12(4): 1211–1219. doi:10.1007/s11682-017-9780-1.

Opportunities for an Enhanced Integration of Neuroscience and Genomics

Ashlee A. Moore[1,2], **Chelsea Sawyers**[1,3], **Daniel E. Adkins**[1,4,5,6], and **Anna R. Docherty**[1,4,5]

[1]Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University School of Medicine, Richmond, VA 23220 USA

[2]Center for Clinical and Translational Research, Virginia Commonwealth University, Richmond, VA, 23220 USA

[3]Department of Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA, 23220 USA

[4]University Neuropsychiatric Institute, University of Utah School of Medicine, Salt Lake City, UT, 84110 USA

[5]Department of Psychiatry, University of Utah School of Medicine, Salt Lake City, UT, 84110 USA

[6]Department of Sociology, University of Utah, Salt Lake City, UT, 84110 USA

Abstract

Neuroimaging and genetics are two rapidly expanding fields of research. Thoughtful integration of these areas is critical for ongoing large-scale research into the genetic mechanisms underlying brain structure, function, and development. Neuroimaging genetics has been slow to evolve relative to psychiatric genetics research, and some may be unaware that new statistical methods allow for the genomic analysis of more modestly-sized imaging samples. We present a broad overview of the extant imaging genetics literature, provide an interpretation of the major problems surrounding the integration of neuroimaging and genetics, discuss the influence and impact of genetics consortia, and suggest statistical genetic analyses that expand the repertoire of imaging researchers amassing rich behavioral data in modestly-sized samples. Specific attention is paid to the creative use of polygenic risk scoring in imaging genetic analyses, with primers on the most current risk scoring applications.

Keywords

Imaging; Neuroimaging; MRI; Genetic; Genomic; GWAS; Polygenic

Introduction

The following update on imaging genomics methodology is aimed specifically at imaging researchers who are currently performing small-scale genetic studies ($N < 200$), and/or those

*Correspondence concerning this article should be addressed to Anna R. Docherty, University Neuropsychiatric Institute, University of Utah School of Medicine, 501 Chipeta Way, Salt Lake City, UT, 84110 USA, Phone: +1 763 516 367, anna.docherty@utah.edu.

who are interested in expanding their research program by including new genetic methodology. This manuscript briefly reviews the historical context of research at the intersection of neuroimaging and genetics/genomics, draws attention to current challenges facing the field of imaging genetics, and outlines how new genetic methodologies may bypass some of the historically difficult challenges faced by researchers attempting to integrate these two fields. Specifically, the methodologies suggested to overcome the current challenges in imaging genetics include, (1) the consortium approach, and (2) polygenic risk scoring. A detailed overview of these methods is provided. Bold terms throughout this manuscript can be found in the glossary (Table 1).

Genetics to Genomics

Twin and family studies are important platforms for quantifying the genetic influence on a trait of interest (e.g., **biometrical genetic studies**; Neale and Cardon 1992). Twin studies exploit the known genetic and environmental similarities and differences of monozygotic (MZ) and dizygotic (DZ) twins in order to determine the magnitude of genetic and environmental factors contributing to variance in a trait. Additive genetic (A) influences represent the effect of all **polymorphic allele**s, and can be determined by accounting for the fact that MZ twins share 100% of their polymorphic alleles, whereas DZ twins share, on average, 50%. Environmental effects can be grouped into two types: common/shared environment (C) and unique/unshared environment (E). In this type of biometrical model, both MZ and DZ twin pairs share 100% of their shared/common environment, and 0% of their unshared/unique environment. Common environment is presumed to make members of the same family more alike, whereas unique environmental experiences are presumed to create differences among family members.

The first study using this biometrical methodology to examine the **heritability** of brain structure involved computerized tomography (CT) imaging (Reveley et al. 1984), and the inclusion of magnetic resonance imaging (MRI) data appeared about a decade later (e.g., Bartley et al. 1997). Almost all neuroanatomical measures studied thus far appear to be under substantial genetic control (40–97%), with areas of the prefrontal cortex evidencing the highest heritabilities (90–95%; for a review see Peper et al. 2007). Only a few exceptions to high heritabilities have been noted, such as a robust influence of environment on cerebellar volume in children (Wallace et al. 2006; genetic factors appear to be more influential on the cerebellum later in life, e.g.; Posthuma et al. 2000). Some of the most influential, large-scale, biometrical genetic imaging studies include the Vietnam Era Twin Study of Aging (VETSA; e.g., Eyler et al. 2011; Kremen et al. 2010; Panizzon et al. 2009; Schmitt et al. 2007), and the NIMH Longitudinal Structural MRI Study of Human Brain Development (e.g., Giedd et al. 2015; Lenroot and Giedd 2008; Schmidtt et al. 2014; Wallace et al. 2010). These twin studies, along with many other more modestly sized twin samples continue to inform **genomic research** today, and continue to provide insights into the genetics of brain morphology and function. Some of this recent research highlights the unique genetic factors influencing cortical thickness, gyrification, and subcortical volumes (Chen et al. 2012; Panizzon et al. 2009; Rimol et al. 2010), as well as the shared genetic covariance of these neuroanatomical features with cognitive ability (Docherty et al. 2015).

Most of the early gene identifying efforts in the field of neuroimaging focused on associations between **phenotype**s and single (or a few) candidate **gene**s. In candidate gene studies, a phenotype, either quantitative or case-control, is examined in the context of either (1) the number of alleles (0, 1, or 2) at a particular **single nucleotide polymorphisms (SNP)**, or (2) the carrier status of an allele (binary variable) at a particular SNP. This research in neuroimaging generally focused on SNPs in genes that were thought to have biological plausibility in influencing neural processes. For example, some studies examined the COMT Val158Met SNP (rs4680), involved in pre-frontal dopamine levels (e.g., Cerasa et al. 2008; Egan et al. 2001; Honea et al. 2009; Taylor et al. 2007), and the serotonin-transporter-linked polymorphic region (5-HTTLPR) in the serotonin transporter gene (SLC6A4; e.g., Dutt et al. 2009; Hariri et al. 2002). However the replication rate of these associations was often ill-fated (e.g., Barnes et al. 2009; Gonzalez-Castro et al. 2016; Nickl-Jockschat et al. 2015; Wang et al. 2013). This trend of unsuccessful candidate gene replications is not limited to imaging genetics studies, and problems with candidate gene studies have been well documented in many areas of psychiatric genetics (e.g., Duncan and Keller 2011; Karg et al. 2011). The smaller samples needed for sufficient statistical power in candidate gene studies, as opposed to more rigorous genomic methods, puts the field of neuroimaging at special risk of excessive un-replicated candidate gene findings.

In other fields of psychiatric genetics, candidate gene studies have largely been replaced by genomic analyses. This is especially true for traits that are complex, and/or **polygenic** in nature. Several larger imaging studies have expanded their protocols to include genomic data metrics. **Genome wide association (GWA) studies** use statistical processes similar those described above for candidate gene studies, but examining many (500,000 or more), instead of one, polymorphic variants across the genome. However, the large number of tests produces a stringent multiple testing significance threshold ($p < 5 \times 10^{-8}$). The first neuroimaging GWA study examined the blood oxygen level dependent (BOLD) signal during an item recognition task in schizophrenia cases and controls and was unable to find any **genome-wide significant associations** with 138 subjects (Potkin et al. 2009). Null findings in a genomic study of this size is not surprising, especially when one considers the first genome-wide significant finding for schizophrenia only emerged once > 9000 cases as > 12,000 controls had been amassed (The Schizophrenia Psychiatric Genome-Wide Association Study Consortium 2011). Although a few imaging GWA studies of modest size have produced nominally significant genome-wide results (e.g., Hashimoto et al. 2014; Hass et al. 2013; Ousdal et al. 2012), these results have rarely been replicated.

In addition to traditional genomic methods (such as GWA studies) neurogenetic researchers are starting to take advantage of pooled data in the form of consortia and/or meta-analyses. This large-scale genomic research, complementing family-based studies, can provide new potential genomic models for future examination, potentially aiding in our understanding of the mechanisms underlying complex human phenotypes.

Challenges in Imaging Genetics

Genetic neuroimaging research has principally focused on either twin and family studies of relatively modest size ($N < 500$), or relatively small case-control, candidate gene studies (N

< 60 per cell). Neuroimaging research has notable strengths relative to most genomics research, in that rich clinical and cognitive data are often available, and case-control samples are carefully screened for major confounding factors such as head injury or current drug use. However, the high cost of imaging often prohibits large samples, and thus imaging sample sizes are often modest, at best. While sample size is not a traditional constraint of imaging studies, a small N quickly becomes prohibitive of large-scale genomic questions. Because GWA studies have a stringent multiple-testing burden (genome-wide significant *p*-value of $< 5 \times 10^{-8}$), there is a lack of statistical power to detect genome-wide effects in currently available imaging samples. This is especially true given that effect sizes emerging from large meta-analytic GWA studies of neuroanatomy are much smaller than originally expected (e.g., Adams et al. 2016; Stein et al. 2012). It is therefore easy to see why a candidate gene approach may be appealing to imaging researchers, whose research is inherently costly, and as it has proven useful in certain types of neurogenetic animal models (for a helpful review, see Gordon 2016). However, most candidate gene successes in humans have involved traits with a simple **Mendelian** basis or well understood pathways in which genes interact to directly cause a phenotype.

It is now widely agreed that genetic studies of behavioral and/or psychiatric traits in humans require genomic, polygenic, and/or pedigree-based information. The concept of a candidate gene for a psychiatric disorder conjures for many the idea of a “dark era” of psychiatric genetics prior to the GWA era. The overload of underpowered and unreplicated studies was an unfortunate setback for the field of genetics, and it is now widely accepted that *candidate gene studies of most psychiatric phenotypes reflect obsolete methodology*, create noise in the literature, and seldom replicate. Large-scale GWA studies have confirmed that almost all of our widely measured psychiatric phenotypes are massively polygenic (Jarick et al. 2014; Mattheisen et al. 2015), with hundreds, or even thousands, of **genomic loci** each accounting for a very small proportion of the phenotypic variance. Unfortunately, identifying single genetic loci of robust influence has been difficult for anything but rare neurodevelopmental disorders such as fragile X, Rett syndrome, and Barth syndrome (Amir et al. 1999; Johnston et al. 1997; van Loo and Martens 2007; Verkerk et al. 1991). There is accumulating evidence that autism and schizophrenia may be influenced both by many of genes of small effect, as well as rare **de novo mutation**s of larger effect, but nevertheless these de novo effects are small (e.g., Brandler et al. 2016; Corominas et al. 2014; Fromer et al. 2016).

Despite the polygenicity of most psychiatric conditions, there is growing evidence that some single genes (such as BDNF or FTO) do impact health factors relevant to psychiatry, with methodologically sound functional neuroimaging work to support such theories of association (Glaser et al. 2014; Gluskin and Mickey 2016, Hsu et al. 2013; Love et al. 2012; Mickey et al. 2011). In this context, phenotyping methodology takes a front seat and constructs tapping more general domains of human experience (for example, paradigms measuring **Research Domain Criteria [RDoC]** constructs, as opposed to criteria for case-control status) can provide quantitative assessments more sensitive to genetic variance. More thorough phenotyping in these cases would allow for a more accurate portrayal of the degree of influence that certain genes have on human neurophysiology and behavioral outcomes.

Based on extant imaging genetics literature, as well as the methodological challenges noted above, we argue that the primary reason imaging genetics has generally lagged behind other fields in adopting relevant genetic methodology is the prohibitively small sample sizes often observed in such studies. However, as we explain in the following sections, techniques are emerging that allow more modestly-sized samples, such as those often seen in neuroimaging genetics, to adopt genetic methodology that is more relevant to understanding the true underlying genetic architecture of human complex traits. These include (1) participation in large-scale consortia, and (2) polygenic risk scoring.

The Consortium Approach

Over the last decade the field of genetics has rapidly shifted to the consortia paradigm, in which numerous samples are combined to produce large-scale genomic datasets (Hagg et al. 2015; Schizophrenia Psychiatric Genome-Wide Association Study Consortium et al. 2011; Wellcome Trust Case Control Consortium 2007). Consortia are well designed to address genetic questions that the specific studies within the consortia would likely not, individually, have the statistical power to address.

The neuroimaging genetics community has also moved toward consortium methodology. The Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA; Thompson et al. 2014) and Cohorts for Heart and Aging Research in Genetic Epidemiology (CHARGE; Psaty et al. 2009) are two examples that show promise. Much like early schizophrenia GWA work by the Psychiatric Genetics Consortium (PGC; Hamshere et al. 2013; Schizophrenia Psychiatric Genome-Wide Association Study Consortium et al. 2011), ENIGMA and CHARGE leverage meta-analytics to gather the largest pool of data to date and continually add new samples, resulting in genetic and imaging data for > 12,000 participants (current estimates) for each of the two consortia. Researchers in a variety of specialties (imaging, genetics, psychiatry, neurology, etc.) contribute data from a range of studies focused on specific phenotypic areas (e.g., psychiatric disorders) and broader imaging genetics topics. These topics include (but are not limited to) GWA, **copy number variant**s, **genome-wide complex trait analysis (GCTA)**, diffusion tensor imaging (DTI), subcortical volumes, and cortical thickness.

Thus far, three GWA publications by ENIGMA have replicated fifteen SNP associations with specific subcortical structure or intracranial volumes in healthy samples, using CHARGE as a replication sample (Adams et al. 2016; Hibar et al. 2015; Stein et al. 2012). Between the three ENIGMA studies, hippocampal volume has been associated with two SNPs, putamen with four, caudate nucleus with one, and intracranial volume with seven. Additionally, CHARGE has published two imaging GWA studies, replicating 18 SNPs at 2 loci associated with hippocampal volume (Bis et al. 2012) and 2 SNPs associated with total intracranial volume (Ikram et al. 2012). Overall, these SNPs have been found in or near genes expressed during early development, and are involved in axon guidance pathways, cellular apoptosis, synapse formation, dendrite growth promotion, neurotransmitter release, and cell membrane and cytoskeletal organization (Adams et al. 2016; Bis et al. 2012; Hibar et al. 2015; Ikram et al. 2012; Stein et al. 2012).

These genomic loci associations, with their stringent significance threshold and replication, are less likely to represent spurious findings than would candidate gene studies. In the context of neuroimaging, it is almost certainly the case that genomic information from large meta-analyses and consortia will be necessary to follow up on the gains made by twin and family studies. A thorough integration of genomic and biometrical models is on the horizon, and is likely to be a significant boon for our understanding of the shared genetics of brain morphology and risk for psychopathology.

The Polygenic Risk Scoring Approach

With careful consideration of variable selection in the context of larger consortia efforts, studies of smaller size can approach similar questions in a slightly different manner. This can facilitate the maximization of power within smaller samples, and provide the extant literature with necessary replication of findings. A method commonly used in smaller studies ($N < 200$) to maximize power is the calculation of a *polygenic risk score* (PRS). This is also sometimes referred to as a *genetic profile score* or a *risk profile score*. Whereas a GWA study hunts for significant "hits," a PRS incorporates the effects of all SNPs across the genome by using a single aggregated quantitative metric, based on previously published, publically available GWA data.

Traditional PRS methods use summary statistics available from large psychiatric case-control GWA studies to derive a quantitative measure (the PRS) for each individual in an independent dataset. Thus, the PRS can be evaluated as to whether or not it predicts the same, or a related, phenotype(s) in a much smaller independent sample. Methods for calculating PRSs vary, and two of the most popular methods (traditional and Bayesian) are discussed below. In psychiatric genetics, this method of scoring has been used successfully to understand the aggregated genetic risk of schizophrenia, its contribution to the risk of bipolar disorder, general cognitive ability, and other phenotypes, as well as the elevated polygenic risk of family members (International Schizophrenia Consortium et al. 2009; Lencz et al. 2014). Traditional PRS methods have been utilized across multiple disorders and domains of personality, and across ancestries (de Moor et al. 2012; Docherty et al. 2016a; Genetics of Personality Consortium et al. 2015; for additional review of this area, see Docherty et al. 2016b).

The choice of SNP variants for a PRS can be driven either from previous GWA literature (e.g., using many "top hits" associated with the trait) or via molecular pathways with enrichment for disorders (for example, gene sets consisting of loci found in/near genes of significance within alcohol metabolism pathways). As a proof of concept, one study was able to replicate the significant single SNP findings from 13 GWA studies as well as replicate their own novel discoveries (Peng et al. 2010). While the enriched pathway-based method is one option for SNP selection, it should be noted that great care ought to be taken when deciding how best to apply gene set analyses to a specific sample (Mooney and Wilmot 2015). Given the high degree of polygenicity of psychiatric traits, we recommend using large swathes of the genome to derive PRSs.

In one example to illustrate our point, a recent study by Franke and colleagues (2016) built upon findings of the ENIGMA consortium by incorporating them into the examination of the shared genetic covariance of structural brain regions in schizophrenia via PRS analyses. One advantage of these analyses is that they examined a disorder with already well-established structural brain differences (Boos et al. 2007; Fusar-Poli et al. 2013; Haijma et al. 2013; Hulshoff Pol et al. 2002; Thermenos et al. 2013; van Erp et al. 2016; van Haren et al. 2012; Wright et al. 2000), and used only already well-established genetic variants associated with case-control status (Schizophrenia Psychiatric Genome-Wide Association Study Consortium 2011; Schizophrenia Working Group of the PGC 2014). However, despite the significant volumetric differences previously observed in schizophrenia versus control samples, there was no significant genetic overlap observed between the two phenotypes. The genetic data in this analysis were limited to only a modest number top hits from previous GWA studies by the PGC, meaning that these polygenic analyses did not use all variants from the initial PGC analysis, and thus the combined genetic variance accounted for by these hits was likely to be quite small. The findings reported were proof of concept that the method works, and we would argue, suggestive that deeper genotyping (using all common variants measured, for example, or a combination of common and rare variants) and deeper quantitative phenotyping (at the level of behavior or symptom dimension) may be more fruitful.

Finally, genomic research of non-imaging phenotypes to date indicates that there is substantial overlap of genetic influence across major psychiatric phenotypes (Cross-Disorder Group of the Psychiatric Genomics Consortium 2013; Cross-Disorder Group of the Psychiatric Genomics Consortium et al. 2013; Mühleisen et al. 2012). PRSs for psychiatric disorders differentially predict cognitive, personality, and health factor phenotypes (Docherty et al. 2017; Krapohl et al. 2016; Lo et al. 2017). We would argue that this provides a portal to innovative and impactful small-scale imaging research. In an attempt to hasten this research paradigm shift, we highlight relevant PRS methods below.

Traditional polygenic risk scoring

Traditional PRS methods, first described by Purcell and colleagues (International Schizophrenia Consortium 2009), use summary statistics available from large GWA studies (also known as the “discovery sample”) to derive a quantitative measure, the PRS, for each individual in an independent dataset. Typically, a list of SNPs common to both the discovery and independent samples is constructed. Next, **clumping** chooses only the most significant SNP in a specified genomic region (e.g., 500 kb). This clumping step (also known as LD pruning) is performed because nearby SNPs are often inherited together due to the process of **linkage disequilibrium (LD)**, and therefore a group of significant SNPs in a small genomic region are likely tagging the influence of a single causally influential SNP. After clumping, SNPs are chosen that have previously reached a certain *p*-value cutoff in the discovery GWA study, and usually scores from several different *p*-value thresholds are examined in a single study (for example, using all variants with $p < .001$, $p < .01$, $p < .1$, then $p < 1$). Each SNP list is then used to calculate a PRS. Here, the known effect size of each SNP based on the discovery GWA summary statistics (for example, the log of the odds ratio in case-control studies), is multiplied by the number of alleles an individual possesses

for that SNP. These weighted effect sizes are summed across all included SNPs, thereby deriving the PRS for each individual.

This method can be used to predict liability toward, or resilience to, the development of a disorder. However, the variance accounted for in a trait will tend to decrease when the number of loci decrease, which is directly related to the *p*-value threshold chosen. Another consideration in using this PRS method is that all variant lists must be pruned for LD based on separate samples of identical ancestry. Some of these established samples for LD pruning are small, and thus significantly limit studies of non-European ancestry. Researchers also need to ensure that their discovery and independent samples include unrelated individuals, so that variance accounted for is not contaminated by shared environmental effects (Wray et al. 2014). Additionally, Dudbridge (2013), suggests that researchers collecting new discovery and independent data ensure that the samples are of equal sample size until the independent sample reaches $N = 2{,}000$, after which all subsequent individuals should be assigned to the discovery sample to maximize power. These are a few of the major considerations when constructing a PRS, although a comprehensive review of limitations in PRS methodology can be found elsewhere (see Wray et al. 2013).

PRSice (Euesden et al. 2015) is a useful program for conducting traditional PRS analyses. PRSice is fully automatic, and allows for the calculation, application, evaluation, and plotting of results. This software is written in R, with wrappers for PLINK2 and bash data management scripts. An innovative feature of PRSice is that it can calculate PRS at multiple *p*-value thresholds, and identify the most precise (i.e., predictive) threshold. Furthermore, it allows the user to specify various parameters, including the removal or inclusion of SNPs in LD, input of ancestry dimensions to control for population structure, and genotyped vs. imputed data. PRSice is freely available for download at https://PRSice.info.

Bayesian polygenic risk scoring

Bayesian PRS methodology was developed in order to maximize the number of SNPs used and therefore also maximize predictive ability of the PRS. Typically, a list of SNPs common to the discovery GWAS, independent sample, and reference sample (later used for determining LD structure) is constructed. Next, genomic LD structure is estimated using the reference sample, and Bayesian methodology (e.g., Markov chain Monte Carlo) is used to estimate the posterior mean effect sizes of each variant based on discovery GWAS summary statistics. These effect sizes are estimated using LD structure and genomic architecture, which is inferred via two parameters: (1) the phenotype heritability based on the variants measured in the discovery GWAS and (2) the fraction of variants with non-zero effect sizes. Much like *p*-value thresholding in traditional PRS methodology, it is recommended that several different fractions of non-zero-effect variants be estimated. Finally, all estimated effects for each allele are summed across the genome for each individual, creating the PRS.

The Bayesian approach is beginning to be more widely used, primarily because simulation studies have shown that traditional methods, via LD pruning and *p*-value thresholding, often fall short of predicting the same proportion of phenotypic variance due to genetic factors (heritability) that is found in twin studies of the same phenotype (Vilhjalmsson et al. 2015). Based on these findings, Bayesian methodology may be preferential given it accounts for

LD and does not rely on *p*-value thresholds, thereby maximizing the number of variants used in constructing the PRS and increasing the proportion of variance accounted for. As with traditional PRS methodology, researchers should ensure that all samples (in this case, discovery, independent, and reference) are made up of unrelated individuals and therefore reflect truly independent samples. Furthermore, Vilhjalmsson and colleagues (2015) suggest that LD structure reference panels should contain at least 1000 individuals. However, they note an exception that decreases the need for three samples: If the independent sample contains more than 1000 individuals of the same ancestry as the discovery GWAS, then the independent sample may also be used as the LD reference panel.

LDpred is a software program developed by researchers at Harvard's Broad Institute (Vilhjalmsson et al. 2015) that is used to conduct Bayesian PRS analyses. LDpred is unique in that it can be applied to traits and diseases with a wide range of genetic architectures, and the exact architecture does not need to be known a priori. LDpred is fully automatic, adjusting for LD using a sliding window along the genome (2 Mb on average, although this setting can be changed). Simulations have suggested that the increase in predictive ability of LDpred over and above that of traditional PRS methods is due primarily to the fact that LDpred accounts for LD structure (Vilhjalmsson et al. 2015). LDpred is freely available for download at https://github.com/bvilhjal/ldpred. Finally, Pak Sham's lab has also very recently released in pre-print a new penalized-regression method for PRS called "lassosum" (details and application can be found at http://biorxiv.org/content/early/2017/03/22/058214).

Concluding Remarks

We have attempted to provide an overview of imaging genetics, and to touch on new methods for researchers working with modestly-sized imaging samples who may have the advantage of analyzing deeper phenotypic data. Resources for these analyses are open source (i.e., free) by design, largely utilizing R or Python-based programming. The data science is progressing daily, as can be witnessed on pre-print online sources, and the field is ripe with opportunity for those with more restrictive samples to not only replicate and verify findings of the larger consortia, but also to apply the consortia results to novel and well-powered analyses of rich clinical and imaging phenotypes using polygenic risk scoring methods. Our hope is that imaging research may contribute increasingly nuanced analyses of phenotype data, as we move toward to a true integration of neuroimaging and genetics.

Acknowledgments

Funding This work was funded by the National Institutes of Health (MH093731 to ARD; MH020030 and MH111229 to AAM), and the Brain & Behavior Research Foundation (ARD).

References

Adams HH, Hibar DP, Chouraki V, Stein JL, Nyquist PA, Renteria ME, et al. Novel genetic loci underlying human intracranial volume identified through genome-wide association. Nature Neuroscience. 2016; 19:1569–1582. [PubMed: 27694991]

Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nature Genetics. 1999; 23:185–188. [PubMed: 10508514]

Barnes A, Isohanni M, Barnett JH, Pietilainen O, Veijola J, Miettunun J, et al. No association of COMT (Val158Met) genotype with brain structure differences between men and women. PLoS One. 2009; 7:e33964.

Bartley AJ, Jones DW, Weinberger DR. Genetic variability of human brain size and cortical gyral patterns. Brain. 1997; 120:257–269. [PubMed: 9117373]

Bis JC, DeCarli C, Smith AV, van der Lijn F, Crivello F, Fornage M, et al. Common variants at 12q14 and 12q24 are associated with hippocampal volume. Nature Genetics. 2012; 44:545–551. [PubMed: 22504421]

Boos HB, Aleman A, Cahn W, Pol Hulshoff H, Kahn RS. Brain volumes in relatives of patients with schizophrenia: a meta-analysis. Archives of General Psychiatry. 2007; 64:297–304. [PubMed: 17339518]

Brandler WM, Antaki D, Gujral M, Noor A, Rosanio G, Chapman TR, et al. Frequency and complexity of de novo structural mutation in autism. American Journal of Human Genetics. 2016; 98:667–679. [PubMed: 27018473]

Cerasa A, Gioia MC, Labate A, Liguori M, Lanza P, Quattrone A. Impact of catechol-O-methyltransferase val(108/158) met genotype on hippocampal and prefrontal gray matter volume. Neuroreport. 2008; 19:405–408. [PubMed: 18287936]

Chen CH, Gutierrez ED, Thompson W, Panizzon MS, Jernigan TL, Eyler LT, et al. Heriarchical genetic organization of human cortical surface area. Science. 2012; 335:1634–1636. [PubMed: 22461613]

Corominas R, Yang X, Lin GN, Kang S, Shen Y, Ghamsari L, et al. Protein interaction network of alternatively spliced isoforms from brain links genetic risk factors for autism. Nature Communications. 2014; 5:3650.

Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet. 2013; 381:1371–1379. [PubMed: 23453885]

Cross-Disorder Group of the Psychiatric Genomics Consortium. Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nature Genetics. 2013; 45:984–994. [PubMed: 23933821]

de Moor MH, Costa PT, Terracciano A, Krueger RF, de Geus EJ, Toshiko T, et al. Meta-analysis of genome-wide association studies for personality. Molecular Psychiatry. 2012; 17:337–349. [PubMed: 21173776]

Docherty AR, Hagler DJ Jr, Panizzon MS, Neale MC, Eyler LT, Fennema-Notestine C, et al. Does degree of gyrification underlie the phenotypic and genetic associations between cortical surface area and cognitive ability? NeuroImage. 2015; 106:154–160. [PubMed: 25433211]

Docherty AR, Moscati A, Peterson R, Edwards AC, Adkins DE, Bacanu SA, et al. SNP-based heritability estimates of the personality dimensions and polygenic prediction of both neuroticism and major depression: findings from CONVERGE. Translational Psychiatry. 2016a; 6:e926. [PubMed: 27779626]

Docherty AR, Moscati AA, Fanous AH. Cross-disorder psychiatric genomics. Current Behavioral Neuroscience Reports. 2016b; 3:256–263. [PubMed: 28042526]

DochertyAR, , MoscatiA, , DickD, , SavageJE, , SalvatoreJE, , CookeM. , et al. Polygenic prediction of the phenome, across ancestry, in emerging adulthood2017

Dudbridge F. Power and predictive accuracy of polygenic risk scores. PLoS Genetics. 2013; 9:e1003348. [PubMed: 23555274]

Duncan LE, Keller MC. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. American Journal of Psychiatry. 2011; 168:1041–1049. [PubMed: 21890791]

Dutt A, McDonald C, Dempster E, Prata D, Shaikh M, Williams I, et al. The effect of COMT, BDNF, 5-HTT, NRG1 and DTNBP1 genes on hippocampal and lateral ventricular volume in psychosis. Psychological Medicine. 2009; 39:1783–1797. [PubMed: 19573260]

Egan MF, Goldberg TE, Kolachana BS, Callicott JH, Mazzanti CM, Straub RE, et al. Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. Proceedings of the National Academy of Sciences. 2001; 98:6917–6922.

Euesden J, Lewis CM, O'Reilly PF. PRSice: polygenic risk score software. Bioinformatics. 2015; 31:1466–1468. [PubMed: 25550326]

Eyler LT, Prom-Wormley E, Fennema-Notestine C, Panizzon MS, Neale MC, Jernigan TL, et al. Genetic patterns of correlation among subcortical volumes in humans: results from a magnetic resonance imaging twin study. Human Brain Mapping. 2011; 32:641–653. [PubMed: 20572207]

Franke B, Stein JL, Ripke S, Anttila V, Hibar DP, van Hulzen KJ, et al. Genetic influences on schizophrenia and subcortical brain volumes: large-scale proof of concept. Nature Neuroscience. 2016; 19:420–431. [PubMed: 26854805]

Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nature Neuroscience. 2016; 19:1442–1453. [PubMed: 27668389]

Fusar-Poli P, Smieskova R, Kempton MJ, Ho BC, Andreasen NC, Borgwardt S. Progressive brain changes in schizophrenia related to antipsychotic treatment? A meta-analysis of longitudinal MRI studies. Neuroscience and Biobehavioral Reviews. 2013; 37:1680–1691. [PubMed: 23769814]

Genetics of Personality Consortium. de Moor MH, van den Berg SM, Verweij KJ, Krueger RF, Luciano M, et al. Meta-analysis of genome-wide association studies for neuroticism, and the polygenic association with major depressive disorder. JAMA Psychiatry. 2015; 72:642–650. [PubMed: 25993607]

Giedd JN, Raznaham A, Alexander-Bloch A, Schmitt E, Gogtay N, Rappaport J. Child psychiatry branch of the National Institute of Mental Health longitudinal structural magnetic resonance imaging study of human brain development. Neuropsychopharmacology. 2015; 40:43–49. [PubMed: 25195638]

Glaser YG, Zubieta JK, Hsu DT, Villafuerte S, Mickey BJ, Trucco EM, et al. Indirect effect of corticotropin-releasing hormone receptor 1 gene variation on negative emotionality and alcohol use via right ventrolateral prefrontal cortex. The Journal of Neuroscience. 2014; 34:4099–4107. [PubMed: 24623788]

Gluskin BS, Mickey BJ. Genetic variation and dopamine D2 receptor availability: a systematic review and meta-analysis of human in vivo molecular imaging studies. Translational Psychiatry. 2016; 6:e747. [PubMed: 26926883]

Gonzalez-Castro TB, Hernandez-Diaz Y, Juarez-Rojop IE, Lopez-Narvaez ML, Tovilla- Zarate CA, Fresan A. The role of a catechol-O-methyltransferase (COMT) Val158Met genetic polymorphism in schizophrenia: a systematic review and updated meta-analysis on 32,816 subjects. Neuromolecular Medicine. 2016; 18:216–231. [PubMed: 27020768]

Gordon JA. On being a circuit psychiatrist. Nature Neuroscience. 2016; 19:1385–1386. [PubMed: 27786177]

Hagg S, Ganna A, Van der Laan SW, Esko T, Pers TH, Locke AE, et al. Gene-based meta-analysis of genome-wide association studies implicates new loci involved in obesity. Human Molecular Genetics. 2015; 24:6849–6860. [PubMed: 26376864]

Haijma SV, Van Haren N, Cahn W, Koolschijn PC, Hulshoff Pol HE, Kahn RS. Brain volumes in schizophrenia: a meta-analysis in over 18 000 subjects. Schizophrenia Bulletin. 2013; 39:1129–1138. [PubMed: 23042112]

Hamshere ML, Walters JT, Smith R, Richards AL, Green E, Grozeva D, et al. Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the schizophrenia PGC. Molecular Psychiatry. 2013; 18:708–712. [PubMed: 22614287]

Hariri AR, Mattay VS, Tessitore A, Kolachana B, Fera F, Goldman D, et al. Serotonin transporter genetic variation and the response of the human amygdala. Science. 2002; 297:400–403. [PubMed: 12130784]

Hashimoto R, Ikeda M, Yamashita F, Ohi K, Yamamori H, Yasuda Y, et al. Common variants at 1p36 are associated with superior frontal gyrus volume. Translational Psychiatry. 2014; 4:e472. [PubMed: 25335168]

Hass J, Walton E, Kirsten H, Liu J, Priebe L, Wolf C, et al. A genme-wide association study suggests novel loci associated with a schizophrenia related brain-based phenotype. PLoS One. 2013; 8:e64872. [PubMed: 23805179]

Hibar DP, Stein JL, Renteria ME, Arias-Vasquez A, Desrivieres S, Jahanshad N, et al. Common genetic variants influence human subcortical brain structures. Nature. 2015; 520:224–229. [PubMed: 25607358]

Honea R, Verchinski BA, Pezawas L, Kolachana BS, Callicott JH, Mattay VS, et al. Impact of interacting functional variants in COMT on regional gray matter volume in human brain. NeuroImage. 2009; 45:44–51. [PubMed: 19071221]

Hsu DT, Sanford BJ, Meyers KK, Love TM, Hazlett KE, Wang H, et al. Response of the mu-opioid system to social rejection and acceptance. Molecular Psychiatry. 2013; 18:1211–1217. [PubMed: 23958960]

Hulshoff Pol HE, Schnack HG, Bertens MG, van Haren NE, van der Tweel I, Staal WG, et al. Volume changes in gray matter in patients with schizophrenia. The American Journal of Psychiatry. 2002; 159:244–250. [PubMed: 11823266]

Ikram MA, Fornage M, Smith AV, Seshadri S, Schmidt R, Debette S, et al. Common variants at 6q22 and 17q21 are associated with intracranial volume. Nature Genetics. 2012; 44:539–544. [PubMed: 22504418]

International Schizophrenia Consortium. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009; 460:748–752. [PubMed: 19571811]

Jarick I, Volckmar AL, Putter C, Pechlivanis S, Nguyen TT, Dauvermann MR, et al. Genome-wide analysis of rare copy number variations reveals PARK2 as a candidate gene for attention-deficit/hyperactivity disorder. Molecular Psychiatry. 2014; 19:115–121. [PubMed: 23164820]

Johnston J, Kelley RI, Feigenbaum A, Cox GF, Iyer GS, Funanage VL, et al. Mutation characterization and genotype-phenotype correlation in barth syndrome. American Journal of Human Genetics. 1997; 61:1053–1058. [PubMed: 9345098]

Karg K, Burmeister M, Shedden K, Sen S. The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. Archives of General Psychiatry. 2011; 68:444–454. [PubMed: 21199959]

Krapohl E, Euesden J, Zabaneh D, Pingault JB, Rimfeld K, von Stumm S, et al. Phenome-wide analysis of genome-wide polygenic scores. Molecular Psychiatry. 2016; 21:1188–1193. [PubMed: 26303664]

Kremen WS, Prom-Wormley E, Panizzon MS, Eyler LT, Fischl B, Neale MC, et al. Genetic and environmental influences on the size of specific brain regions in midlife: the VETSA MRI study. NeuroImage. 2010; 49:1213–1223. [PubMed: 19786105]

Lencz T, Knowles E, Davies G, Guha S, Liewald DC, Starr JM, et al. Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: a report from the cognitive genomics consorTium (COGENT). Molecular Psychiatry. 2014; 19:168–174. [PubMed: 24342994]

Lenroot RK, Giedd JN. The changing impact of genes and environment on brain development during childhood and adolescence: Initial findings from a neuroimaging study of pediatric twins. Developmental Psychopathology. 2008; 20:1161–1175.

Lo MT, Hinds DA, Tung JY, Franz C, Fan CC, Wang Y, et al. Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. Nature Genetics. 2017; 49:152–156. [PubMed: 27918536]

Love TM, Enoch MA, Hodgkinson CA, Pecina M, Mickey B, Koeppe RA, et al. Oxytocin gene polymorphisms influence human dopaminergic function in a sex- dependent manner. Biological Psychiatry. 2012; 72:198–206. [PubMed: 22418012]

Mattheisen M, Samuels JF, Wang Y, Greenberg BD, Fyer AJ, McCracken JT, et al. Genome-wide association study in obsessive-compulsive disorder: results from the OCGAS. Molecular Psychiatry. 2015; 20:337–344. [PubMed: 24821223]

Mickey BJ, Zhou Z, Heitzeg MM, Heinz E, Hodgkinson CA, Hsu DT, et al. Emotion processing, major depression, and functional genetic variation of neuropeptide Y. Archives of General Psychiatry. 2011; 68:158–166. [PubMed: 21300944]

Mooney MA, Wilmot B. Gene set analysis: a step-by-step guide. American Journal of Medical Genetics Part B, Neuropsychiatric Genetics. 2015; 168:517–527.

NealeMC, , CardonLR. Methodology for genetic studies of twins and familiesDordecht: Kluwer Academic Publishers; 1992

Nickl-Jockschat T, Janouschek H, Eickhoff SB, Eickhoff CR. Lack of meta-analytic evidence for an impact of COMT Val158Met genotype on brain activation during working memory tasks. Biological Psychiatry. 2015; 78:e43–e46. [PubMed: 25861704]

Ousdal OT, Anand Brown A, Jensen J, Nakstad PH, Melle I, Agartz I, et al. Associations between variants near a monoaminergic pathways gene (PHOX2B) and amygdala reactivity: a genome-wide functional imaging study. Twin Research and Human Genetics. 2012; 15:273–285. [PubMed: 22856363]

Panizzon MS, Fennema-Notestine C, Eyler LT, Jernigan TL, Prom-Wormley E, Neale M, et al. Distinct genetic influences on cortical surface area and cortical thickness. Cerebral Cortex. 2009; 19:2728–2735. [PubMed: 19299253]

Peng G, Luo L, Siu H, Zhu Y, Hu P, Hong S, et al. Gene and pathway-based second-wave analysis of genome-wide association studies. European Journal of Human Genetics. 2010; 18:111–117. [PubMed: 19584899]

Peper JS, Brouwer RM, Boomsma DI, Kahn RS, Hulshoff Pol HE. Genetic influences on human brain structure: a review of brain imaging studies in twins. Human Brain Mapping. 2007; 28:464–473. [PubMed: 17415783]

Posthuma D, de Geus EJC, Neale MC, Pol HEH, Baare WEC, Kahn RS, et al. Multivariate genetic analysis of brain structure in an extended twin design. Behavior Genetics. 2000; 30:311–319. [PubMed: 11206086]

Potkin SG, Turner JA, Guffanti G, Lakatos A, Fallon JH, Nguyen DD, et al. A genome-wide association study of schizophrenia using brain activation as a quantitative phenotype. Schizophrenia Bulletin. 2009; 35:96–108. [PubMed: 19023125]

Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, et al. Cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium: design of prospective meta-analyses of genome-wide association studies from five cohorts. Circulation: Cardiovascular Genetics. 2009; 2:73–80. [PubMed: 20031568]

Reveley AM, Reveley MA, Chitkara B, Clifford C. The genetic basis of cerebral ventricular volume. Psychiatry Research. 1984; 13:261–266. [PubMed: 6597463]

Rimol LM, Panizzon MS, Fennema-Notestine C, Eyler LT, Fischl B, Franz CE, et al. Cortical thickness is influenced by regionally specific genetic factors. Biological Psychiatry. 2010; 67:493–499. [PubMed: 19963208]

Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. Nature Genetics. 2011; 43:969–976. [PubMed: 21926974]

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

Schmidtt JE, Neale MC, Fassassi B, Perez J, Lenroot RK, Wells EM, et al. The dynamic role of genetics on cortical patterning during childhood and adolescence. Proceedings of the National Academy of Science. 2014; 111:6774–6749.

Schmitt JE, Wallace GL, Rosenthal MA, Molloy EA, Ordaz S, Lenroot R, et al. A multivariate analysis of neuroanatomic relationships in a genetically informative pediatric sample. Neuroimage. 2007; 35:70–82. [PubMed: 17208460]

Stein JL, Medland SE, Vasquez AA, Hibar DP, Senstad RE, Winkler AM, et al. Identification of common variants associated with human hippocampal and intracranial volumes. Nature Genetics. 2012; 44:552–561. [PubMed: 22504417]

Taylor WD, Zuchner S, Payne ME, Messer DF, Doty TJ, MacFall JR, et al. The COMT Val158Met polymorphism and temporal lobe morphometry in healthy adults. Psychiatry Research. 2007; 155:173–177. [PubMed: 17521892]

The Schizophrenia Psychiatric Genome-Wide Association Study Consortium. Genome-wide association study identifies five new schizophrenia loci. Nature Genetics. 2011; 43:969–976. [PubMed: 21926974]

Thermenos HW, Keshavan MS, Juelich RJ, Molokotos E, Whitfield-Gabrieli S, Brent BK, et al. A review of neuroimaging studies of young relatives of individuals with schizophrenia: a developmental perspective from schizotaxia to schizophrenia. American Journal of Medical Genetics, Part B, Neuropsychiatric Genetics. 2013; 162:604–635.

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

van Erp TG, Hibar DP, Rasmussen JM, Glahn DC, Pearlson GD, Andreassen OA, et al. Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. Molecular Psychiatry. 2016; 21:585. [PubMed: 26283641]

van Haren NE, Rijsdijk F, Schnack HG, Picchioni MM, Toulopoulou T, Weisbrod M, et al. The genetic and environmental determinants of the association between brain abnormalities and schizophrenia: The schizophrenia twins and relatives consortium. Biological Psychiatry. 2012; 71:915–921. [PubMed: 22341827]

van Loo KM, Martens GJ. Genetic and environmental factors in complex neurodevelopmental disorders. Current Genomics. 2007; 8:429–444. [PubMed: 19412416]

Verkerk AJ, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DP, Pizzuti A, et al. Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell. 1991; 65:905–914. [PubMed: 1710175]

Vilhjalmsson BJ, Yang J, Finucane HK, Gusev A, Lindstrom S, Ripke S, et al. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. American Journal of Human Genetics. 2015; 97:576–592. [PubMed: 26430803]

Wallace GL, Eric SJ, Lenroot R, Viding E, Ordaz S, Rosenthal MA, et al. A pediatric twin study of brain morphometry. Journal of Child Psychology and Psychiatry. 2006; 47:987–993. [PubMed: 17073977]

Wallace GL, Lee NR, Prom-Wormley EC, Medland SE, Lenroot RK, Clasen LS, et al. A bivariate twin study of regional brain volumes and verbal and nonverbal intellectual skills during childhood and adolescence. Behavior Genetics. 2010; 40:125–134. [PubMed: 20112131]

Wang Y, Li J, Chen C, Chen C, Zhu B, Moysis RK, et al. COMT rs4680 met is not always the ‘smart allele’: val allele is associated with better working memory and larger hippocampal volume in healthy Chinese. Genes, Brain and Behavior. 2013; 12:323–329.

Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007; 447:661–678. [PubMed: 17554300]

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

Wray NR, Lee SH, Mehta D, Vinkhuyzen AA, Dudbridge F, Middeldorp CM. Research review: polygenic methods and their application to psychiatric traits. Journal of Child Psychology and Psychiatry, and Allied Disciplines. 2014; 55:1069–1087.

Wright IC, Rabe-Hesketh S, Woodruff PW, David AS, Murray RM, Bullmore ET. Meta-analysis of regional brain volumes in schizophrenia. The American Journal of Psychiatry. 2000; 157:16–25. [PubMed: 10618008]

Table 1

Glossary of Genetic Terminology

Term	Definition
Biometrical genetic study	A study that seeks to quantify the influence of genes and environment on a trait or behavior, usually in the form of twin and family studies using statistical methodology such as structural equation modeling.
Polymorphic allele	A specific position (locus) in the genome that possesses more than one variant in the population. A monomorphic allele, on the other hand, is an allele that has only one observable form in the population.
Heritability	The proportion of a trait's phenotypic variance that is due to genetic variance in the population.
Genomic research	Methods that involve the examination of a large set of genetic variants across the genome, such as genome-wide association (GWA) and genome-wide complex trait analysis (GCTA).
Phenotype	An observable characteristic or set of observable characteristics (in the case of a disease or disorder).
Gene	A single unit of genetic information that codes the amino acid sequence of a polypeptide.
Single nucleotide polymorphism (SNP)	A variation of a single base pair at a specific genomic location.
Polygenic	Influenced by a multiple genes.
Genome-wide association (GWA) study	An examination of a large number of SNPs across the genome to see if any are associated with a specific trait.
Genome-wide significant associations	A statistical test meeting a stringent significance criterion based on a multiple-testing correction that accounts for the high number of associations tested in genomic research. Typically $p < 5 \times 10^{-8}$.
Mendelian	A pattern of biological inheritance following the genetic laws proposed by Gregor Mendel.
Genomic locus/loci	A specific position on a chromosome (singular: locus, plural: loci).
De novo mutation	A genetic mutation that appears in a child due to a mutation in the germ cell of a parent or in the fertilized egg.
Research domain criteria (RDoC)	A project initiated by the National Institutes of Mental Health (NIMH) that aims to examine psychiatric disorders in terms of more basic constructs, such as genes and neural circuitry.
Copy number variant (CNV)	A section of the genome that is repeated a variable number of times within the population. The number of copies can range from missing to multiple copies.
Genome-wide complex trait analysis (GCTA)	A method of estimating heritability using unrelated individuals, by directly estimating the degree of relationship for each pair of individuals in the dataset (via measured genetic variants).
Clumping	A step in constructing a traditional polygenic risk score (PRS), whereby only the most significant SNP in a specified genomic region (e.g., 500 kb) is retained for inclusion in the PRS.
Linkage disequilibrium (LD)	The non-random association of alleles at different genetic loci, insofar that alleles closer together on a chromosome are more likely to be inherited together.