REVIEW

Genetic Determinants of Depression: Recent Findings and Future Directions

Erin C. Dunn, ScD, MPH, Ruth C. Brown, PhD, Yael Dai, BA, Jonathan Rosand, MD, MSc, Nicole R. Nugent, PhD, Ananda B. Amstadter, PhD, and Jordan W. Smoller, MD, ScD

Learning Objectives: After participating in this activity, learners should be better able to:
1. Evaluate current evidence regarding the genetic determinants of depression
2. Assess findings from studies of gene-environment interaction
3. Identify challenges to gene discovery in depression

Abstract: Depression is one of the most prevalent, disabling, and costly mental health conditions in the United States and also worldwide. One promising avenue for preventing depression and informing its clinical treatment lies in uncovering the genetic and environmental determinants of the disorder as well as their interaction (G×E). The overarching goal of this review article is to translate recent findings from studies of genetic association and G×E related to depression, particularly for readers without in-depth knowledge of genetics or genetic methods. The review is organized into three major sections. In the first, we summarize what is currently known about the genetic determinants of depression, focusing on findings from genome-wide association studies (GWAS). In the second section, we review findings from studies of G×E, which seek to simultaneously examine the role of genes and exposure to specific environments or experiences in the etiology of depression. In the third section, we describe the challenges to genetic discovery in depression and promising strategies for future progress.

Keywords: copy-number variant, depression, genetics, gene-environment interaction, genome-wide association study, genome-wide environment interaction study, rare variants

Depression is one of the most prevalent, disabling, and costly mental health conditions in the United States, with lifetime prevalence estimates of 11.7% among adolescents[1] and 16.6% among adults.[2] It is projected to be *the* leading cause of disease burden worldwide by 2030.[3] Although the impact of depression can be minimized or prevented through early detection, treatment, and ongoing care, numerous individual and structural barriers—including stigma, lack of health insurance, and other barriers to accessing mental health services—prevent many from seeking help. Indeed, only slightly more than half of all people who experience depression seek treatment, and those who do tend to drop out prematurely or receive poor quality care.[4,5] Existing treatments for depression are modestly effective; only about one-fifth of adults receiving cognitive-behavioral therapy or psychodynamic therapy alone,[6] and one-third of adults receiving antidepressant medication alone,[7,8] will experience remission after an initial course of treatment. In children and adolescents, the efficacy of existing treatments is also limited.[9–11] Moreover, nearly three-quarters of people with depression will experience a relapse at some point in their lives.[12] These findings underscore the urgent need to prioritize prevention alongside treatment.

A deeper understanding of the etiology of depression, including its genetic and environmental determinants as well as their interplay (e.g., gene-environment interaction (G×E)), will have implications for preventing depression and informing its clinical treatment. Numerous environmental risk factors for depression have been established, including poverty,[13,14] negative family relationships and parental divorce,[15,16] child

From the Department of Psychiatry, Harvard Medical School (Drs. Dunn and Smoller); Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA (Drs. Dunn, Rosand, and Smoller, and Ms. Dai); Stanley Center for Psychiatric Research (Drs. Dunn and Smoller) and Program in Medical and Population Genetics (Dr. Rosand), Broad Institute of Harvard and MIT, Cambridge, MA; Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University (Drs. Brown and Amstadter); Department of Neurology, Massachusetts General Hospital, Boston, MA (Dr. Rosand); Department of Psychiatry and Human Behavior, Alpert Brown Medical School (Dr. Nugent); Center on the Developing Child, Harvard University (Dr. Smoller).

Supported, in part, by the Harvard University Center on the Developing Child (Dr. Dunn) and by National Institute of Mental Health grant nos. K01MH087240 (Dr. Nugent) and K24MH094614 (Dr. Smoller).

Original manuscript received 1 February 2014; revised manuscript received 26 May 2014, accepted for publication subject to revision 25 June 2014; revised manuscript received 18 July 2014.

Correspondence: Erin C. Dunn, Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, Simches Research Building, 6th Floor, 185 Cambridge St., Boston, MA 02114. Email: edunn2@mgh.harvard.edu

CME *Harvard Review of Psychiatry* offers CME for readers who complete questions about featured articles. Questions can be accessed from the *Harvard Review of Psychiatry* website (www.harvardreviewofpsychiatry.org) by clicking the CME tab. Please read the featured article and then log into the website for this educational offering. If you are already online, click here to go directly to the CME page for further information.

DOI: 10.1097/HRP.0000000000000054

maltreatment,[17,18] and other stressful life events more generally.[19,20] Although the risk of depression is elevated in the immediate aftermath of experiencing these environmental adversities, the effects of adversity can persist over the life course.[21,22]

A robust literature implicates genetic factors in the etiology of depression and other psychiatric disorders. Depression is known to run in families; people with major depressive disorder (MDD) are three times more likely than those without the disorder to have a first-degree relative who also has depression.[23] Twin studies, which allow for the simultaneous quantification of genetic and environmental influences, suggest that depression is moderately heritable. Specifically, twin studies have estimated that approximately 40% of the variation in the population risk of depression is attributable to genetic variation.[24]

In recent years, the combination of advances in our understanding of human genomic variation (e.g., Human Genome Project, HapMap Project, 1,000 Genomes Project) and cost-effective genotyping techniques have led to extraordinary growth in molecular genetic studies of depression and other "complex" psychiatric phenotypes. These studies typically examine whether specific alleles (i.e., alternative forms of DNA sequence at a specific locus) or genotypes (i.e., the combination of alleles at a given locus) are associated with the phenotype of interest. Until recently, genetic studies of depression focused largely on candidate genes—that is, genes that are hypothesized to be implicated in the neurobiology of depression. Some of the most commonly studied candidate genes have been those regulating serotonin (5-HT) and dopamine (DA) neurotransmission, given the suspected involvement of these neurotransmitters in the pathophysiology of depression and their role as targets of antidepressant drugs.[25–27] Unfortunately, most candidate gene studies have been underpowered, and replication of findings has been rare. More recently, the availability of DNA microarrays has enabled genome-wide association studies (GWAS) that do not rely on prior hypotheses. The GWAS approach allows for a hypothesis-free analysis of a million or more variants across the entire genome. The ultimate goal of GWAS is to improve diagnosis, prevention, and treatment through a nuanced understanding of the genetic underpinnings of the disease.

In this article we review recent findings from genetic association studies and G×E studies related to depression, and outline areas for future research. Several excellent reviews of this literature aimed at the genetic research community have already been published (see, e.g., references 28–33). We aim to provide a review for a broad audience, who may be unfamiliar with genetic concepts and methods. We have organized this review into three major sections. In the first, we describe recent findings from GWAS of depression. We begin with GWAS, rather than older methods (i.e., linkage and candidate gene association studies), since these older methods have already been extensively covered by prior reviews. We also do not review studies on genetic markers of antidepressant treatment response, or pharmacogenomics,[34] since our focus is on the genetic determinants of illness risk. In the second section, we review findings from G×E studies, which aim to simultaneously examine the respective roles of genetic variants and environmental exposures in the etiology of depression. As described below, G×E studies have the potential to help identify genetic variants associated with both the risk of, and resilience against, depression—which are revealed only in specific subgroups of the population that have experienced a given environment. In the third section, we address the challenges that face genetic studies of depression and describe emerging strategies that may be useful for overcoming these challenges. We encourage readers who may be unfamiliar with basic genetic concepts to refer to two articles by Attia and colleagues[35,36] and the resources listed in Text Box 1.

Text Box 1
Resources to Learn More About Concepts and Findings from Genetics and Genomics

- National Human Genome Research Institute. Talking glossary of genetic terms. www.genome.gov/Glossary (detailed glossary of genetic terms and concepts)
- National Coalition for Health Professional Education in Genetics. www.nchpeg.org (provides health professionals with online training and continuing education series on topics related to human genetics)
- NIH Pharmacogenetics Research Network. www.pgrn.org (network of scientists focused on research identifying genetic influences on medication response)

FINDINGS FROM GENOME-WIDE ASSOCIATION STUDIES

GWAS have been one of the most widely used methods for identifying risk loci in the past decade.[37–39] In a typical GWAS, one million or more common variants known as single nucleotide polymorphisms (SNPs) are examined for their association to disease. Common variants are generally defined as those alleles that are carried by at least 5% of the population. GWAS are typically conducted using a case-control design in which allele frequencies are compared between cases (with a disease) and controls (without the disease). Compared to candidate gene studies, GWAS provide a hypothesis free, or "unbiased," approach to detecting susceptibility loci. To account for the large number of tests conducted, however, the threshold for declaring genome-wide significance in a GWAS is a p-value of less than 5×10^{-8}, which is equivalent to a p-value of .05 that has been corrected for a million independent tests ($p < .00000005$).[40] Because common variant effects are typically modest, large samples (in the order of 10,000 or more cases and controls) are usually needed to have sufficient power to detect such effects at this statistical threshold.

According to the National Human Genome Research Institute GWAS catalog, more than 2000 GWAS have been published to date.[41] A total of 14 GWAS have been conducted for either MDD or depressive symptoms. In addition, one GWAS focusing on age at onset of MDD has been conducted. These 15 studies were identified by conducting a systematic search of PubMed for articles published before October 2013. We searched the PubMed database using the following MESH terms: (depression OR depressive disorder OR depressive disorder, major OR depressive disorder, treatment-resistant) AND (genome-wide association study). We also searched for articles by examining the reference pages of review articles, meta-analyses, and other empirical articles published since 2005. As shown in Table 1, all of these studies were based on samples of European ancestry and represent a combination of population- and clinic-based samples.

The first GWAS of depression was published in 2009 and included 1738 cases and 1802 controls. Although no SNPs reached genome-wide significance, 11 of the top 200 SNPs were found in a 167 kilobase (kb) region overlapping the gene *PCLO* (piccolo presynaptic cytomatrix protein), which is involved in establishing active synaptic zones and synaptic vesicle tracking.[42] In several subsequent studies,[43,48,57] investigators found mixed evidence regarding the association of *PCLO* SNPs and MDD.

In the first study to report a genome-wide significant association for depression, Kohli and colleagues[49] found support for a recessive effect of a SNP (rs1545843) in the gene *SLC6A15* (solute carrier family 6, neutral amino acid transporter, member 15), which is involved in transporting neutral amino acids. The authors provided additional evidence in support of this association by demonstrating that risk alleles were correlated with reduced *SLC6A15* expression in hippocampal tissue (taken from individuals undergoing surgery for epilepsy) and reduced hippocampal volume and neuronal integrity (as determined by neuroimaging). Mice susceptible to chronic stress were also found to have reduced hippocampal *SLC6A15* expression. This locus, however, has not emerged as a prominent finding in subsequent depression GWAS (described below).

As in the case of other complex traits,[58,59] one of the major lessons from these early GWAS of depression was that the effect of most SNPs is small in magnitude (allelic odd ratios of around 1.3 or less) and that considerably larger samples would therefore be needed to identify genetic loci associated with depression. To enhance the power of psychiatric GWAS studies, the Psychiatric Genomics Consortium was established in 2007 as an international collaborative effort to define the spectrum of risk variants across psychiatric disorders (http://www.med.unc.edu/pgc/). One of the consortium's major goals is to conduct mega-analyses for MDD as well as autism, attention-deficit/hyperactivity disorder, bipolar disorder, and schizophrenia.[60–62] A mega-analysis pools individual-level phenotype and genotype data from across many studies; this approach differs from a meta-analysis, where the summary statistics produced by each study are analyzed. In 2012, the consortium published the results of a GWAS mega-analysis of MDD comprising 9240 cases and 9519 controls across nine primary samples, all of European ancestry.[52] Although this sample was the largest to date, no SNP reached genome-wide significance. The most significant SNPs in the discovery sample were rs11579964 ($p = 1.0 \times 10^{-7}$), a variant closest to several genes (*CNIH4*, *NVL*, *WDR26*), and rs7647854 ($p = 6.5 \times 10^{-7}$), a variant closest to *C3orf70* and *EHHADH*. These findings were not supported, however, in a large, independent replication sample.

GWAS of depressive symptoms have also been largely unrevealing. The first GWAS of depressive symptoms did not find any SNPs reaching genome-wide significance.[54] One modestly associated ($p = 1.59 \times 10^{-6}$) SNP (rs7582472) did show evidence of replication in two independent cohorts. However, this SNP was more than 300kb away from two genes, and neither gene showed significant association to depression in a gene-based analysis. A second study of depressed mood, while finding no genome-wide significant SNP, did find in the meta-analysis that an intronic SNP (rs12912233) in *RORA* (retinoidrelated orphan receptor alpha gene) was modestly associated ($p = 6.3 \times 10^{-7}$).[44] Although this result is interesting because another *RORA* SNP has been linked through GWAS to posttraumatic stress disorder,[63] it awaits replication. In the largest study—a meta-analysis comprising 17 population-based studies (n = 34,549 individuals) as the discovery sample—no SNP reached genome-wide significance.[55] The strongest association was for rs8020095 ($p = 1.05 \times 10^{-7}$), located in the gene *GPHN*. When the discovery and replication samples were combined into one meta-analysis of 22 studies with 51,258 respondents, one region (indexed by the SNP rs40465) was associated with depressive symptoms at genome-wide levels of significance.[55] This variant is in a *gene desert*, an area of the genome where there are long regions without protein-coding sequences and whose biological function is unknown.

Another major lesson from depression GWAS has been that popular candidate genes have generally not shown evidence of association. Prior to the GWAS era, meta-analyses of candidate gene studies concluded that there was nominally significant evidence (at $p < 0.05$) for six candidate genes in depression: *APOE*, *DRD4*, *GNB3*, *MTHFR*, *SLC6A3*, and *SLC6A4*.[64,65] To date, however, none of these genes, or any of the more than 100 frequently examined candidate genes, has shown evidence of significant association in the published GWAS of depression. Replication of candidate genes in GWAS is challenging since several widely studied candidate gene markers, including the serotonin transporter 5-HTTLPR variable tandem number repeat, are not directly captured by a typical GWAS platform. Some groups have developed techniques to impute or derive best-guess estimates of these genetic markers using available SNP data,[66,67] though these efforts have not yet been widely adopted. Nonetheless, the evidence for many candidate genes has not been compelling.

Table 1

Published GWAS of MDD, Depressive Symptoms, or Age at Onset of MDD

Study	Discovery sample	Definition of depression	GWAS discovery results	Replication sample	GWAS replication results	Other genetic findings
Sullivan et al. (2009)[42]	1738 cases (mainly drawn from primary care screening and outpatient mental health clinics) 1802 controls (drawn from twin registry; selected to be at low risk of depression) European ancestry	Lifetime diagnosis of DSM-IV MDD based on CIDI	No SNP reached GWS Of the top 200 SNPs, 11 were within a 167 kb region overlapping *PCLO* Top hits: rs2715148 ($p = 7.7 \times 10^{-7}$) rs2522833 ($p = 1.2 \times 10^{-6}$) Secondary analyses revealed findings were largely driven by women and participants with recurrent, early onset MDD	5 independent samples: 6979 cases 5893 controls	No SNP reached significance after correction The replication sample most similar to the discovery had a nominal effect: rs2522833 6.4×10^{-8}	Examined 75 candidate genes and found significant effect for *NOS1* (p = .0006)
Rietschel et al. (2010)[43]	597 cases (hospital based) 1295 controls (population based) European ancestry	Lifetime MDD based on SCI, medical records, or family history	No SNP reached GWS Top 3 hits: rs2765501 ($p = 1.66 \times 10^{-7}$) in *CD5L* rs7713917 ($p = 5.87 \times 10^{-5}$) located 20 kb upstream of *HOMER1* rs9943849 ($p = 6.22 \times 10^{-5}$) located 14 kb upstream of *CPM* Did not replicate previous findings for *PCLO*	409 cases 541 controls	The two top hits showed a trend toward significance in replication sample: rs7713917 ($p = 7.61 \times 10^{-3}$) rs9943849 ($p = 1.59 \times 10^{-2}$) Meta-analysis with discovery and replication samples: rs9943849 ($p = 3.24 \times 10^{-6}$) located upstream of *CPM* rs7713917 ($p = 1.48 \times 10^{-6}$) located in *HOMER1*	Examined genes that contained SNPs with p-values <.001 and found the most significant pathway related to cell signaling (included *GRM5* and *HOMER1*) After correcting for gene size, found 6 of 22 genes that had a probability >10% of being selected by chance (*ACTN2, CYP17A1, DLC1, GRM5, MAP2K4, MAP24K, RSU1*) Conducted fMRI follow-up analyses with top variants to examine several intermediate phenotypes and found some differences by genotype group in working memory
Terracciano et al. (2010)[44]	3972 participants European ancestry	Depression inventory from the revised NEO-PI	No SNP reached GWS Top hit: rs349475 ($p = 2.4 \times 10^{-7}$) in *CDH18*	839 community-dwelling respondents from the Baltimore Longitudinal Study of Aging European ancestry	No SNP reached GWS Top hit: rs4885589 ($p = 2.4 \times 10^{-4}$)	Conducted meta-analysis in combined sample (n = 4811) Top hit in meta-analysis: rs12912233 ($p = 6.3 \times 10^{-7}$) (an intronic SNP in *RORA*) Described rs17864092 ($p = 5.5 \times 10^{-6}$) in *GRM8* as most biologically plausible top hit

Table 1

Continued

Study	Discovery sample	Definition of depression	GWAS discovery results	Replication sample	GWAS replication results	Other genetic findings
Muglia et al. (2010)[45]	926 recurrent MDD cases (clinical sample) 866 controls (age and gender matched from community sample) European ancestry	Recurrent MDD (at least 2 separate episodes of severe intensity) as assessed by SCAN Cases allowed to have comorbid anxiety (no OCD or PTSD); 26% did	No SNP reached GWS	492 recurrent cases (clinical sample) 1052 controls (from community survey) European ancestry	No SNP reached GWS in replication sample Little agreement between discovery and replication sample with respect to the most significant SNP associations identified Conducted meta-analysis: no SNP reached GWS; most significant SNP was for rs4238010 (nearest gene was *CCND2*; $p = 5.8 \times 10^{-6}$)	Gene-based analysis found the most significant results were similar to the SNP-level analysis; strongest adjusted gene-based association ($p = .009$) was for *SMG7* Examined whether top 104 genes ($p < 10^{-4}$) from meta-analysis were found in a previous study to associate with mood disorders or schizophrenia; none of these genes was associated with depression in previous studies; some genes (*ADCY9*, *GRM7*) were related to bipolar disorder/schizophrenia Candidate gene analysis with 15 candidates found none was GWS; strongest effect in *GRM7* (rs1485171; $p = .0001$) Functional annotation analysis (i.e., generated pathways) of the most significant ($p < .0005$) SNPs from meta-analysis found this subset of genes was enriched ($p < .05$) in four pathways: (1) synaptic long-term depression, (2) cAMP-mediated signaling, (3) G-protein-coupled receptor signaling, and (4) glutamate receptor signaling

Continued on next page

Table 1

Continued

Study	Discovery sample	Definition of depression	GWAS discovery results	Replication sample	GWAS replication results	Other genetic findings
Lewis et al. (2010)[46]	1636 cases (from 3 clinic-based groups) 1594 controls (primary care screening; screened negative for depression or anxiety) European ancestry	Recurrent depression of at least moderate severity defined by SCAN	No SNP reached GWS Four genotyped SNPs showed suggestive evidence: 2 SNPs in *BICC1*: rs9416742 ($p = 1.3 \times 10^{-7}$) and rs999845 ($p = 3.1 \times 10^{-7}$); rs2698195 ($p = 3.54 \times 10–6$) located closest to *KIAA1841*; rs8050326 ($p = 4.07 \times 10–6$) located closest to (over 100 kb from) *IRF8* After imputing SNPs in the *BICC1* region, found GWS for 6 SNPs in strong linkage disequilibrium; strongest evidence at rs7903712 ($p = 5.7 \times 10^{-9}$)	Conducted meta-analysis 1418 cases (recurrent depression from clinical sample) 1918 controls (from population-based study)	No SNP attained GWS in meta-analysis Top hit: rs606149 ($p = 2.57 \times 10^{-6}$) near *LOC647167* No replication for *BICC1*	Analyzed haplotypes in discovery at 2 SNPs in *BICCI* (rs7903712 and rs9416742) and found association ($p = 9.04 \times 10^{-8}$) Conducted sex-specific analyses for top 20 SNPs in discovery and found GWS association for women with rs9416742 in *BICC1* ($p = 1.8 \times 10^{-8}$); in women, 4 additional SNPs had suggestive evidence (rs8067196, rs2930553, rs13079811, rs987390); in men, rs6989226 (near *TUSC3*) had suggestive evidence ($p = 1.81 \times 10^{-6}$) Analyzed 84 candidate genes and found strongest evidence for rs13050655 in *PDE9A* ($p = 3.58 \times 10^{-5}$)
Shi et al. (2011)[47]	1020 cases (with recurrent, early-onset MDD; clinical sample) 1636 controls (no lifetime MDD; population based) European ancestry	MDD based on DIGS and consensus by 2 independent reviewers	No SNP reached GWS Top hit: rs17077540 ($p = 1.83 \times 10^{-7}$)	Compared results to other meta-analyses described in Shyn et al. (2011)[51]	No SNP reached GWS Strongest support for rs17144465 ($p = 8.38 \times 10^{-7}$)	Examined 41 candidate genes using single SNP tests and found the lowest p-value ($p = 6.7 \times 10^{-5}$) for *CACNA1C* Also conducted an aggregate test (of whether p-values in gene were more significant than expected by chance) and found no significant findings
Aragam et al. (2011)[48]	1726 population-based cases 1630 population-based controls European ancestry	MDD diagnosis based on CIDI	No SNP reached GWS Top hit: rs1558477 ($p = 2.63 \times 10^{-7}$) in *ADCYAP1R1* 4 SNPs in *PCLO* had marginal effects: rs2715148 ($p = 1.38 \times 10^{-6}$) rs 2522833 ($p = 2.46 \times 10^{-6}$) rs2522840 ($p = 4.38 \times 10^{-6}$) rs2107828 ($p = 1.48 \times 10^{-5}$)	None		Also conducted subgroup analyses stratified by sex Best SNP for males: rs9352774 ($p = 2.26 \times 10^{-6}$) in *LGSN* Best SNP for females: rs2715148 ($p = 5.64 \times 10^{-7}$) in *PCLO* Also tested for interactions by gender and found best SNP in rs12692709 ($p = 5.75 \times 10^{-6}$)

Table 1

Continued

Study	Discovery sample	Definition of depression	GWAS discovery results	Replication sample	GWAS replication results	Other genetic findings
Kohli et al. (2011)[49]	353 cases (inpatients in tertiary clinic) 366 matched controls (no lifetime MDD; community sample) European ancestry	Met DSM-IV criteria for first depressive episode or recurrent depressive disorder and had HAM-D score ≥14	One SNP reached GWS: rs1545843 ($p = 5.53 \times 10^{-8}$)	6 independent samples of different racial/ethnic background	Nominally significant association in 4 of the 5 initial replication samples with recessive model GWS replication for rs1545843 ($p = 4.37 \times 10^{-8}$) in meta-analysis, after adjustment for multiple testing Further replicated findings for rs154843 in second replication sample (p = .008)	Further validated findings in analyses of (1) associations to pre-mortem human hippocampus and lymphoblastoid cell line expression profiles, (2) imaging, and (3) hippocampal expression in mouse model of chronic social stress
Shyn et al. (2011)[50]	1221 cases (from STAR*D; outpatient clinics) 1636 controls (population based; no lifetime history of MDD) European ancestry	Diagnosis of MDD by clinician rating and HAM-D score ≥14 by independent raters	No SNP reached GWS in analyses of genotyped SNPs Top hit: rs12462886 ($p = 1.73 \times 10^{-6}$) located in gene desert	Meta-analysis with 2 additional datasets (GenRED and GAIN) (n = 3957 cases; n = 4328 controls) Examined broad (all cases) and narrow phenotype (recurrent depression with onset before age 31)	No SNP (imputed or genotyped) reached GWS Strongest evidence in broad meta-analysis was for intronic SNPs in *ATP6V1B2* (rs1106634; $p = 6.78 \times 10^{-7}$), *SP4* (rs17144465; $p = 7.68 \times 10^{-7}$), and *GRM7* (rs9870680; $p = 1.11 \times 10^{-6}$) Best SNP in narrow analysis was in stratified analysis of males only (rs11710109; $p = 5.64 \times 10^{-8}$)	Examined 41 candidate genes in discovery, but none were supported; best finding was for rs3788477, a SNP intronic to *SYN3* ($p = 1.64 \times 10^{-4}$); no other SNP achieved $p < 10^{-3}$ Also examined in meta-analysis; aggregate analysis did not suggest an excess of low p-values among these candidates
Wray et al. (2012)[51]	2431 cases 3673 controls (population based; family members without disease; drawn from five different sites) European ancestry	Lifetime diagnosis of MDD by CIDI or other interview instrument	No SNP reached GWS Top hits in total sample: rs3732293 ($p = 1.5 \times 10^{-6}$) rs17226852 ($p = 1.5 \times 10^{-6}$)	Compared results to other published studies and meta-analyses	No SNPs reached GWS	Tested 183 candidate genes and found none reached significance after correction in the discovery sample, other published studies, or meta-analysis

Continued on next page

Table 1

Continued

Study	Discovery sample	Definition of depression	GWAS discovery results	Replication sample	GWAS replication results	Other genetic findings
Psychiatric GWAS Consortium (2012)[52]	9240 cases and 9519 controls (mostly population based; no lifetime history of depression) Came from 9 primary samples European ancestry	Lifetime MDD established using structured diagnostic instruments from direct interviews or clinician-administered checklists	No SNP reached GWS in mega-analysis Top hits: rs11579964 ($p = 1.0 \times 10^{-7}$) rs7647854 ($p = 6.5 \times 10^{-7}$) Top 201 SNPs and 1655 in linkage disequilibrium with those did not overlap with literature from the National Human Genome Research Institute catalog, transcripts expressed in brain samples, or prior PGC analyses; several SNPs were near (20 kb) genes studied in MDD (e.g., *ADCY9*, *PDC1M5*) or other psychiatric disorders (e.g., *GRM7*, *HTR7*, *RELN*) No SNP reached GWS on X chromosome Most significant SNP across all analyses was rs12837650 in female-only analysis ($p = 5.6 \times 10^{-6}$)	6783 cases and 50,695 controls (7 independent samples from discovery) European ancestry	Tried to replicate 554 SNPs with $p < .001$ Did not find SNPs that replicated in the same direction as discovery analysis more frequently than chance No SNP achieved GWS for a joint analysis of the discovery and replication samples; top hit was for rs1969253 ($p = 4.8 \times 10^{-6}$) located in *DVL3*	Conducted a cross-disorder meta-analysis and a set of secondary analyses (by sex; recurrent, early age at onset; and subtype) Direction of effects was generally consistent between discovery and replication for analyses restricted to women and for recurrent MDD, but no SNP reached GWS Only in MDD/bipolar disorder cross-disorder analysis did 15 SNPs exceed GWS; top hit was for rs2535629 ($p = 5.9 \times 10^{-9}$) Conducted a polygene analysis using discovery-phase samples and found SNPs explained 0.6% of the variance in case-control status ($p < 10^{-6}$)
Power et al. (2012)[53]	Time-to-event analysis for age at onset 1480 cases and 1584 controls, both from the UK cohorts in the RADIANT study (see Lewis et al. [2010])[46] Additional analyses used all RADIANT participants (n = 2746) European ancestry	Age at onset to recurrent depression of at least moderate severity defined by SCAN	No SNP achieved GWS in any analysis Top hit (in case-control analysis): rs2273289 in *PLOD1* ($p = 1.29 \times 10^{-7}$)	2 clinical cohorts based in Germany	None of the previously identified suggestive loci replicated	Also performed a GCTA analysis and found that 55% of the variance in age at onset was explained by common SNPs Sex-specific analyses found suggestive evidence for 36 SNPs

Table 1
Continued

Study	Discovery sample	Definition of depression	GWAS discovery results	Replication sample	GWAS replication results	Other genetic findings
Luciano et al. (2013)[54]	5 population-based cohorts (n = 4525) European ancestry	Depressive symptoms measured by BDI or HADS	No SNP reached GWS in meta-analysis Top hits: 5 SNPs with $p < 6.09 \times 10^{-6}$, including rs2141848 in *FAM190A*, rs4888786 in *WWOX*, and rs10410977 in *RAVER1*	1 population-based German cohort using the POMS, and the Netherlands twin register using BDI	One SNP (rs7582472), whose closest gene was 300 kb away, replicated in the German (p = .01) and Netherlands cohorts (p = .006) None of the other 4 SNPs replicated in either sample	Performed a gene-based test; did not find GWS results; best gene was 1.9×10^{-5} (*GRAP*)
Hek et al. (2013)[55]	17 population-based studies (n = 34,549 individuals) European ancestry	Depressive symptoms measured by 10-, 11-, or 20-item CESD	No SNP reached GWS Top hit: rs8020095 ($p = 1.05 \times 10^{-7}$)	5 population-based studies (n = 16,709) Focused on 7 SNPs	No SNP reached significance after correction; best SNP was rs161645 ($p = 9.19 \times 10^{-3}$)	Performed combined meta-analysis of 22 studies (n = 51,258) and found rs40465 reached GWS ($p = 4.78 \times 10^{-8}$) Conducted pathway analysis with 104 genes to identify and classify biological processes among SNPs with p-values $<10^{-4}$; found neurotransmitter secretion ($p = 9.94 \times 10^{-3}$), vitamin transport (p = .014), and synaptic transmission (p = .037) processes were overrepresented among top SNPs Examined 17 candidate SNPs based on previous findings and found none that replicated
Power et al. (2013)[56]	805 case-control pairs matched first on ancestry and second on exposure to stressful life events from the RADIANT study (see Lewis et al. [2010])[46] European ancestry	Recurrent depression of at least moderate severity defined by SCAN	No SNPs achieved GWS or suggestive evidence ($p < 5 \times 10^{-6}$)	None		

BDI, Beck Depression Inventory; CESD, Center for Epidemiological Studies of Depression Scale; CIDI, Composite International Diagnostic Interview; DIGS, Diagnostic Interview for Genetic Studies; GAIN, Genetic Association Information Network; GCTA, genome-wide complex trait analysis; GenRED, Genetics of Recurrent Early-Onset Depression; GWAS, genome-wide association studies; GWS, genome-wide significance; HADS, Hospital Anxiety and Depression Scale; HAM-D, Hamilton Depression Rating Scale; MDD, major depressive disorder; NEO-PI, NEO Personality Inventory; OCD, obsessive-compulsive disorder; POMS, Profile of Mood States; PTSD, posttraumatic stress disorder; SCAN, Schedules for Clinical Assessment in Neuropsychiatry; SNP, single-nucleotide polymorphism; STAR*D, Sequenced Treatments to Relieve Depression.

Another interesting observation from GWAS has been the failure to consider the role of environment. As we describe below, we believe that GWAS may be limited by not taking into account how genetic influences on depression may vary among individuals with certain environmental exposures. One exception is a study by Powers and colleagues,[56] who used propensity-score matching to conduct a GWAS among case-control pairs matched on exposure to recent stressful life events. Although they did not formally test for G×E, the use of propensity-score matching enabled them to reduce sample heterogeneity and to compare cases to controls with similar levels of exposure. In their analysis, no SNPs were genome-wide significant or even suggestive ($p < 5 \times 10^{-6}$); this finding was likely due to the very small sample size (n = 805).

FINDINGS FROM GENE-ENVIRONNENT INTERACTION STUDIES

The long-standing recognition that both genes ("nature") and environments ("nurture") contribute to the etiology of depression has motivated a great deal of interest in studying gene-environment interactions. G×E studies examine the degree to which genetic variants modify the association between environmental factors and depression (or similarly, the extent to which environmental factors modify the association between genes and depression).[68–70] Typically, G×E studies have assumed a *diathesis-stress* model, where a genetic liability, also referred to as a diathesis, interacts with a stressful life event to give rise to depression. In this model, genes either exacerbate or buffer the effects of stress.[71] More recently, the concept of G×E has been expanded to incorporate positive aspects of the environment, such as social support, psychosocial interventions, and other protective factors that reduce the risk of disease.[72,73] Emerging work has focused on differential susceptibility to the environment[74,75] or on the extent to which genetic variation makes individuals more likely to respond adversely to negative environments but more positively to salutary environments.

Research on G×E in depression was essentially launched with a 2003 publication in *Science*. In that study Caspi and colleagues[76] used data from a 26-year longitudinal study in New Zealand to test whether a functional length polymorphism in the promoter region (5-HTTLPR) of the serotonin transporter gene (*SLC6A4*) interacted with stressful life events to increase the risk of depression. Results of the Caspi study suggested that individuals with at least one short (s) allele (i.e., the "s/s" or "s/l" genotype of the biallelic coded version) had more depression in response to stressful life events when compared to subjects who were not s allele carriers. This result held regardless of how depression was measured, whether by level of depressive symptoms, depression diagnosis, incident depression, or suicidality. They also found that s allele carriers, compared to those without an s allele, had a greater probability of experiencing depression resulting from exposure to probable or severe childhood maltreatment. The Caspi article has become one of the most influential studies in the field, having been cited more than 5000 times.

Since the publication of Caspi and colleagues' seminal research, numerous replication attempts have been made. Most of these have also focused on 5-HTTLPR, although other genetic variants have been studied, including variants in *BDNF* (brain-derived neurotropic factor), *MAOA* (monoamine oxidase A), *FKBP5* (FK506 binding protein 51), *CRHR1* (corticotropin-releasing hormone receptor 1), *COMT* (catechol-O-methyltransferase), and *CREB1* (also known as cAMP or responsive element-binding protein 1). Many replication attempts have focused on stressful life events (either recent or in childhood) or on child maltreatment—namely, physical abuse, sexual abuse, or neglect. All of these "candidate" environments are appropriate to study in G×E research. Child maltreatment, for example, is one of the most potent environmental stressors in the etiology and course of depression and other types of psychopathology. Extant studies suggest child maltreatment at least doubles the risk of internalizing problems, including depression.[18,20,21,77,78]

The large number of empirical studies trying to replicate Caspi's G×E findings for depression have been summarized in several reviews focusing on G×E with 5-HTTLPR (see, e.g., references 70 and 79–86). These reviews ultimately fueled a heated debate on the plausibility of the Caspi findings. Including somewhat similar individual studies, review articles have drawn opposing conclusions about the support for G×E effects, with some studies finding consistent G×E effects and others failing to detect them.[82,87] Meta-analyses have provided a quantitative summary of these studies but have also reached opposing conclusions. Specifically, the results of two meta-analyses,[80,84] which found evidence against a consistent G×E effect, differed from a third meta-analysis,[81] which concluded that the evidence was strong to support the 5-HTTLPR G×E. These conflicting results may be explained by differences in the selection of studies for inclusion in the meta-analyses.[88,89] For example, the meta-analyses that used the most stringent inclusion criteria[80,84] failed to support the G×E association.[90] Some commentators have noted that an inverse relationship exists between the power of the replication studies and support for the 5-HTTLPR association—precisely the opposite of what one would expect if the association is valid.[90] Moreover, the most direct replication attempt of the Caspi findings, which was not included in any prior meta-analysis, found no evidence in support of the G×E effect on depression. This longitudinal, birth cohort study followed a similar population (New Zealand residents) for a similar length of time (30 years) and used comparable phenotypic measures.[91] The authors observed no interaction between stressful life events and 5-HTTLPR genotype, even after conducting 104 different regression models.[91]

By contrast, some have argued that support for the 5-HTTLPR G×E has been more consistent when childhood maltreatment is the exposure variable[81,82,87] or when direct interview assessments (as opposed to self-report questionnaires)

have been used.[82,87] These findings are important since there has been substantial variability in the characteristics of study populations, the measurements of depression and environmental exposures, and the analytic methods used across empirical studies to test for G×E in depression.[70] Others have also tried to place these individual G×E studies in the context of the broader literature examining genetic variability and stress sensitivity in depression. They have appealed to the more consistent findings from animal studies showing that loss-of-function mutations in the serotonin gene have been associated with depressive-like behavior in rodents and that genetic variation in the serotonin transporter gene has been linked to depression among nonhuman primates.[87] Proponents have noted that the results are more convincing when considered alongside both experimental imaging studies showing 5-HTTLPR variation in amygdala activity and treatment-response studies showing 5-HTTLPR variation in antidepressant treatment response.[87,92] Overall, the validity of the influential 5-HTTLPR G×E finding remains unclear.

G×E studies focusing on other candidate genes, however, have found more consistent results. For example, studies examining *FKBP5* and *CHRH1* have shown that variants in these genes moderate the effect of exposure to child maltreatment, childhood adversities, or negative life events on adult depression.[93–96] These genes are interesting candidates because they regulate the stress response via the hypothalamic-pituitary-adrenal axis.[97] Additional replications of these candidates would be helpful to further evaluate their role in shaping risk for depression. Evidence for other candidates, such as *BDNF*, has been mixed. For instance, a recent review found stronger evidence to support interactions with the *BDNF* Val66Met polymorphism and stressful life events compared to childhood adversity.[98] As we later discuss, genome-wide approaches to G×E remain an important, but relatively unexplored, area.

CURRENT AND FUTURE DIRECTIONS FOR RESEARCH

The limited success of GWAS for depression is in contrast to other psychiatric disorders, where established risk variants are accumulating through GWAS. For example, at the time of writing, more than 100 loci have been associated with schizophrenia and bipolar disorder at stringent levels of statistical significance.[99–104] Although no individual risk loci have been identified for depression, we know that such variants will be found with adequate sample sizes. For example, it is now possible through genome-wide complex trait analysis to estimate the common variant contribution to depression using genome-wide SNP data (these estimates are sometimes referred to as *SNP chip heritability*).[105] These methods have yielded estimates of the common variant contribution to depression, ranging from a high of 32%[106] to a low of 21%.[107] It should be noted that these estimates are lower bound because SNP chip heritability reflects only the effect of common variation that is captured on genotyping arrays.

Text Box 2
Possible Explanations for the Lack of Success of GWAS and G×E Studies for Depression, and Strategies to Increase Gene Finding

Explanations	Strategies to Address
Depression has a different genetic architecture	• Increase sample size to improve power to detect associations of individually small effect loci • Aggregate genetic signals into pathways or gene sets • Examine rare variants and other types of structural variants (e.g., copy-number variants) in addition to common variants
Previous GWAS did not consider the role of environment	• Conduct GEWIS • Test for G×E using candidate genes from GWAS
Depression is highly heterogeneous	• Examine depressive symptoms (quantitative phenotype) rather than only diagnoses of depression • Use novel analytic methods (e.g., factor analysis, latent class analysis) to identify and refine distinct subtypes • Focus on intermediate phenotypes or endophenotypes, consistent with RDoC

GEWIS, genome-environment wide interaction studies; GWAS, genome-wide association studies; G×E, gene-environment interaction; RDoC, National Institute of Mental Health Research Domain Criteria Initiative.

Thus, the field faces two major questions: what explains the lack success of GWAS and G×E studies for depression, and how can we best move forward? As described below (and summarized in Text Box 2), there are several likely explanations for the limited progress to date and several strategies that may help overcome these challenges.

Genetic Architecture and the Need for Larger Studies

The genetic architecture of depression is likely to be highly complex. Genetic architecture refers to the number of genetic loci associated with a phenotype, the effect size of each locus, and the manner in which these loci behave (e.g., whether they have additive or multiplicative effects). While all psychiatric disorders are thought to be polygenic, or influenced by multiple genes, the genetic basis of depression may reflect an even larger number of loci of individually small effect. Results from studies that have calculated polygenic risk scores (capturing aggregate effects of loci across the genome) support such a hypothesis.[52,108] It is therefore likely that much larger samples than those examined to date will be needed to detect these individually small effects. Simulations suggest that to have comparable power to GWAS of schizophrenia or bipolar disorder, studies of depression will need to have sample sizes as much as five times larger than the sample sizes required for those disorders.[51] Experience with GWAS for other psychiatric disorders has established that once a critical

sample size threshold is crossed, a larger and larger sample size yields more and more loci.

If depression is driven by many thousands of loci of weak effect, another strategy may be to combine genetic signals across many SNPs into functionally defined gene sets or pathways. Pathway approaches can be considerably more powerful than single-variant analyses, as the aggregation of weak signals from multiple causal variants may yield statistically significant evidence in support of a given gene or pathway.[109,110] Thus far, investigators have primarily examined pathways related to specific biological functions (e.g., axon guidance, cell functioning) as defined by human-curated bioinformatics resources, such as the Kyoto Encyclopedia of Genes and Genomes[111] or Gene Ontology.[112] Recent studies of candidate gene pathways have found evidence that genes involved in glutamatergic synaptic neurotransmission,[113] among others,[114] were significantly associated with depression. Evidence in support of gene sets or pathways also comes from several GWAS that we have described above (see Table 2). These studies found significant support for some pathways.[45,55] One of the major drawbacks of gene-set analyses is that they require predefined sets of genes. Gene sets defined by current annotation databases, such as the Kyoto Encyclopedia or Gene Ontology, vary in their completeness; some pathways are more complete than others. Moreover, databases also vary in how they define gene sets. Thus, a given gene may belong to one pathway in one database and a second pathway in another. Although these challenges are substantial, we think that greater use of pathway-type analyses is needed.

Understudied Components of the Genetic Architecture of Depression

A related consideration is that GWAS are designed to capture common, but not rare, genetic variation. Rare variants can include genetic single-nucleotide variations (SNVs; present in <1% of the population) and rare copy-number variations (CNVs; that is, structural variations in DNA sequence that involve the duplication or deletion of thousands or more than a million base pairs). Such variants have now been shown to play a role in autism,[115,116] schizophrenia,[117,118] and bipolar disorder,[119] but to date these components of the genetic architecture of depression have been largely unexplored.

Fortunately, advances in sequencing technology now provide an opportunity to address the role of rare SNVs. In recent years, the cost of direct DNA sequencing has dropped dramatically, and technologic advances have facilitated the development of "high-throughput" sequencing.[120,121] To date, these "next-generation sequencing technologies" have been largely applied to study variants in exons, which are the protein-coding regions of the genome, collectively known as the *exome*. Exons comprise about 30 megabases of DNA or 1% of the total genome. Although no exome-sequencing studies of depression have been reported at the time of writing, such studies are under way. Next-generation sequencing technologies can also be applied to the entire genome (*whole genome sequencing*), enabling researchers to explore a full range of genetic variants in both coding and noncoding regions of the genome.

The major strength of sequencing is that it captures variants that have been previously uncharacterized by candidate gene and GWAS methods and thus may provide new insights into the genetic underpinnings of depression. Like all techniques, however, sequencing approaches face a number of challenges. For example, despite enormous reductions in the cost of sequencing, well-powered studies remain very expensive. Whole-genome sequencing costs at least US$1000 per genome, whereas exome sequencing costs only several hundred dollars. Exome sequencing also assesses polymorphisms that, by definition, are rare and thus occur much less frequently than common variants. To have sufficient statistical power to identify an association between these rare variants and depression, very large sample sizes—in the order of 10,000 or more cases—are needed. In addition, rare-variant association methods are still largely under development.

Structural variation, including CNVs, is also a potential source of depression risk loci. CNVs can be inherited or spontaneous (de novo). De novo CNVs—those that are present in offspring but not in either parent—have been shown to be important risk factors for several neuropsychiatric disorders—namely autism,[115,116] schizophrenia,[117,118] and bipolar disorder.[119] After conducting a systematic literature search of PubMed for articles published by December 2013 using the MESH terms for depression that were described above and the phrase "copy number var*," we identified four studies that provide preliminary evidence implicating CNVs in depression.[122–125] In the largest of these studies, Glessner and colleagues[124] found 12 CNV regions that were exclusive to cases with MDD. The region with the highest frequency in cases was a locus on chromosome 5 (5q35.1) that overlapped the genes *SLIT3*, *CCDC99*, and *DOCK2*. The finding of a CNV overlapping the gene *SLIT3* is interesting since *SLIT3* is known to play a role in axon development and neurodevelopmental disorders.

One of the major strengths of studying CNVs is that the methods for association testing are similar, by and large, to examining common variants. Simultaneous examination of SNPs and CNVs in large samples may identify whether CNVs play a significant role in depression and what their importance is relative to common variants. One of the major drawbacks of association testing with CNVs is that catalogs of these variants do not exist with the same number or specificity as they do for SNPs. For example, the location, size, and boundary of CNVs in these publicly available resources have been relatively imprecise. Opportunities for misclassification of variants is consequently much higher for CNVs than for SNPs.[126] Efforts are now under way to provide a more comprehensive catalog of CNVs (see, e.g., http://www.sanger.ac.uk/research/areas/humangenetics/cnv/). Moreover, until recently no genotyping array that could detect both SNPs and CNVs was commercially available. With the advent of

the *PsychChip*, a customized genotyping chip for psychiatric phenotypes, investigators will soon be able to simultaneously examine multiple genetic variants, including SNPs, CNVs, and rare variants. The importance of rare variants to depression risk remains to be seen. Large-scale studies will be needed to clarify their contribution.

Accounting for the Role of Gene-Environment Interaction

As noted previously, existing studies have not systematically addressed the possibility that a substantial proportion of the risk of depression is attributable to nonadditive effects, including G×E. Moreover, G×E studies to date have focused on a limited set of candidate genes and have typically been underpowered, creating a risk of both false-positive and false-negative results. It is well established that environmental factors, including exposure to stressful life events and child maltreatment, are important risk factors for depression, but we still know little about whether these environmental effects are moderated by genetic variation and, if so, which genetic variants are relevant.

One approach to filling this gap may come from genome-environment wide interaction studies (GEWIS), pronounced *gee whiz*.[127,128] In a GEWIS, investigators test for statistical interaction or G×E, with the *G* defined as the genetic loci (e.g., SNPs) included in a GWAS and the *E* defined as a known environmental exposure. Unlike candidate gene G×E, GEWIS offers the opportunity to conduct a genetically unbiased search—that is, one in which prior genetic or biologic hypotheses are not required. In one type of GEWIS, investigators could focus on loci for which a main effect of a genetic variant has been established by GWAS. In this scenario, loci that have been identified by GWAS become candidates for G×E analysis but with the advantage over traditional candidate gene studies that the locus is already known to influence the phenotype of interest.

To our knowledge, no GEWIS of depression has been published to date. Although research on GEWIS of depression and other psychiatric phenotypes is lacking, a small but emerging body of research on other complex phenotypes suggests GEWIS can yield important new gains. For example, studies have identified significant genome-wide G×E interactions in cancer,[129,130] diabetes[131] and insulin resistance,[132] Parkinson's disease,[133] pulmonary function,[134] and nonsyndromic cleft palate.[135] Interest in GEWIS is growing, but several challenges to conducting this type of study remain.[127] The first is identifying the best methods to test for genome-wide G×E. Several methodological approaches have been developed (see, e.g., the reviews by Winham & Biernacka[136] and Gauderman et al.[137]), though without any consensus as to which is the best. Selection of a specific analytic method depends largely on whether the goal is to leverage G×E to discover novel loci, or to characterize the joint effect of genetic variants and environmental factors.[138]

The second challenge is that the "environment" is somewhat indeterminate; it is unbounded in a way that the genome is not. Both children and adults are exposed to a range of experiences across the multiple social and physical contexts in which they are embedded (e.g., families, school, neighborhoods, workplaces); all of these experiences and exposures can contribute to health.[139] One way to start is to focus on well-defined measures of environment where robust and consistent evidence supports a relationship between the exposure and depression. Such a list of measures could include in utero exposures (e.g., viruses, toxins, alcohol and drugs), social deprivation (e.g., poverty, child maltreatment), and enrichment (e.g., psychosocial interventions and treatments). However, even if we select the same environment, such as child maltreatment, we are still faced with multiple different types of maltreatment, multiple ages at which the maltreatment occurred, and multiple ways to measure maltreatment (e.g., self-report, administrative records, clinical interview).

The final, and perhaps the biggest, challenge is balancing the trade-off between large samples and precise measures of environmental exposure. Large samples are needed to detect G×E (larger even than those needed in standard GWAS). Large samples, however, often lack the depth and breadth that are necessary to capture data on environmental or phenotype measures. Although smaller samples frequently have rich and repeated measures, they are underpowered to establish robust associations. Smaller samples can be combined to increase statistical power, but challenges will arise in trying to harmonize measures of environment across these data sets. In other words, efforts to ensure an adequate sample size for each unique combination of risk factors and G×E strata can lead to watered-down environmental measures that lack any meaningful variability; a classic example would be an instance where respondents are simply classified as "exposed" or "non-exposed." Longitudinal birth cohort studies, which can include prospective measures of environmental exposures along with detailed phenotype data and genome-wide data, may be one promising avenue for conducting GEWIS in the future. Moreover, the growing interest in the concept of the exposome, in environment-wide association studies (EWAS), and in ways to systematically identify relevant environmental factors (see, e.g., Wild[140] and Patel et al.[141]) could yield new insights to guide GEWIS in the future.

The Phenotypic Complexity of Depression

Another obstacle to identifying susceptibility loci is that depression is a heterogeneous phenotype. Indeed, it is possible to meet DSM-IV or DSM-5 diagnostic criteria for a major depressive episode through at least 227 different symptom combinations.[142] As currently described by DSM-5, MDD can manifest with or without (1) anxious distress, (2) mixed features, (3) melancholic features, (4) atypical features, (5) mood-congruent psychotic features, (6) mood-incongruent psychotic features, (7) catatonia, (8) peripartum onset, and (9) a seasonal pattern.[143] These subtypes of MDD could reflect different genetic contributions. Consistent with such a

hypothesis, studies suggest that depression with a history of child maltreatment has a different onset, course, and response to treatment than a depression that arises among individuals without a history of abuse.[144,145] Recent twin studies have also suggested that genetic liability to MDD reflects not *one* but *three* distinct symptom dimensions (psychomotor/cognitive, mood, and neurovegetative).[146] Thus, GWAS that simply examine "depressed" cases versus controls may decrease the ratio of signal to noise by combining multiple disorder subtypes that vary in their genetic etiology. In light of evidence suggesting that there is no truly categorical threshold for depression caseness[147] and that different lifetime prevalence estimates of depression are found when comparing cross-sectional retrospective reports to cumulative evaluations based on multiple interviews,[148] it is reasonable to posit that misclassification of individuals as cases or controls may be undermining the power of typical case-control GWAS.

We think several strategies can help reduce the heterogeneity in depression. First, examining the full range of variation in depression (e.g., depressive symptoms) rather than dichotomizing the phenotype (cases and controls) could be a statistically more powerful approach to identify variants associated with depression.[149] This approach would be consistent with evidence that the diagnostic threshold for MDD has been artificially imposed on a continuity of depression risk.[147] Second, more data-driven approaches to examine shared features or subtypes of depression through use of latent class analysis[150] may also prove helpful. Prior studies applying such methods in both adolescents and adults have found distinct subtypes that differ based on severity, symptoms, and episode length.[151,152] Examination of these subtypes in a genetic association study may help to identify variants that are common across, or unique to, specific subtypes. Third, another strategy would be to continue efforts to examine phenotypes thought to be more proximal to a genetic substrate than are clinically defined categories.[153] Putative *intermediate phenotypes* or *endophenotypes* that are related to depression include emotion-based attention biases,[154,155] impaired reward function,[156] and deficits in domains of executive functioning, such as learning and memory.[157] Investigation of endophenotypes is consistent with the National Institute on Mental Health Research Domain Criteria Initiative,[158–161] which aims to provide a bottom-up characterization of psychopathology incorporating genetics, neural circuitry, and behavioral phenotypes. Endophenotypes have not yet been the subject of large-scale studies that might fully evaluate their relative power. One exception is the ENIGMA consortium, where a GWAS meta-analysis of structural magnetic resonance imaging phenotypes yielded a genome-wide significant association with hippocampal volume,[162] one of the best-established biomarkers of depression risk. Nonetheless, this result required sample sizes in the thousands, challenging the view that endophenotype-based studies will be more powerful than studies of MDD itself.

CONCLUSIONS

Research on the genetic underpinnings of depression is at an exciting yet challenging crossroad. On the one hand, genotyping technologies have allowed for the characterization of individual and population-based genetic variation and have provided analytic tools to examine the individual and joint effects of genetic and environmental determinants. On the other hand, GWAS of depression have yet to see the same success that has been achieved with other psychiatric or medical disorders. Studies of G×E have thus far failed to provide clarity but have fueled plenty of debate. Some argue that positive findings reflect chance results among small, underpowered studies,[84] whereas others see consistencies when focusing on studies that are methodologically comparable.[81,82,87]

We have reviewed some of the potential explanations for the lack of success to date for GWAS and G×E studies of depression. Given the established heritability of depression, there is every reason to expect that increasingly well-powered studies will indeed identify risk loci. The genetic and phenotypic complexity of depression, however, may mean that such successes will require samples on the order of tens of thousands of participants. Efforts to parse the heterogeneity of depression and validate phenotypic subtypes may also be essential to facilitate gene identification. Further, as we have noted, potentially important areas to uncover the genetic basis of depression—specifically, rare variation and G×E—remain relatively unexplored on a large scale. It remains to be seen how much of the "missing heritability" of depression will be revealed thorough studies of these components.

Although the path forward to detect genetic risk loci for depression remains challenging, what is certain is that a deeper understanding of the etiology of depression is needed. Existing treatments for depression are based on decades-old biology, and genetic discoveries have already begun to identify promising targets for novel therapies in other disorders. Given the enormous burden of depression, identifying its genetic underpinnings may be essential to preventing the onset of this disorder and improving the lives of those who already suffer.

The authors thank Caitlin Clements, Patience Gallagher, Stephanie Kravitz, and Preetha Palasuberniam for their assistance in conducting the literature review for this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health.

REFERENCES

1. Merikangas KR, He S, Burstein M, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry 2011;49:980–9.
2. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005;62:593–602.

3. Mathers C, Ma Fat D, Boerma JT. The global burden of disease: 2004 update. Geneva, Switzerland: World Health Organization, 2008.
4. Wang PA, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC. Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005;62:629–40.
5. Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003;289:3095–105.
6. Driessen E, Van HL, Don FJ, et al. The efficacy of cognitive behavioral therapy and psychdynamic therapy in the outpatient treatment of major depression: a randomized clinical trial. Am J Psychiatry 2013;170:1041–50.
7. Rush AJ, Trivedi M, Fava M. Depression, IV: STAR*D treatment trial for depression. Am J Psychiatry 2003;160:237.
8. Warden D, Rush AJ, Trivedi MH, Fava M, Wisniewski SR. The STAR*D project results: a comprehensive review of findings. Curr Psychiatry Rep 2007;9:449–59.
9. Cox GR, Callahan P, Churchill R, et al. Psychological therapies versus antidepressant medication, alone and in combination for depression in children and adolescents. Cochrane Database Syst Rev 2012;11:CD008324.
10. Weisz JR, McCarty CA, Valeri SM. Effects of psychotherapy for depression in children and adolescents: a meta-analysis. Psychol Bull 2006;132:132–49.
11. Weisz JR, Kuppens S, Eckshtain D, Ugueto AM, Hawley KM, Jensen-Doss A. Performance of evidence-based youth psychotherapies compared with usual clinical care: a multilevel meta-analysis. JAMA Psychiatry 2013;70:750–61.
12. Kessler RC, Zhao S, Blazer DG. Prevalence, course, and correlates of minor and MDD in the National Comorbidity Survey. J Affect Disord 1997;45:19–30.
13. Brooks-Gunn J, Duncan GJ. The effects of poverty on children. Future Child 1997;7:55–71.
14. McLeod JD, Shanahan MJ. Trajectories of poverty and children's mental health. J Health Soc Behav 1996;37:207–20.
15. Repetti RL, Taylor SE, Seeman TE. Risky families: family social environments and the mental and physical health of offspring. Psychol Bull 2002;128:330–66.
16. Gilman SE, Kawachi I, Fitzmaurice GM, Buka SL. Family disruption in childhood and risk of adult depression. Am J Psychiatry 2003;160:939–46.
17. Slopen N, Koenen KC, Kubzansky LD. Cumulative adversity in childhood and emergent risk factors for long-term health. J Pediatr 2014;164:631–8.
18. Widom CS, DuMont K, Czaja SJ. A prospective investigation of major depressive disorder and comorbidity in abused and neglected children grown up. Arch Gen Psychiatry 2007;64:49–56.
19. Kessler RC. The effects of stressful life events on depression. Annu Rev Psychol 1997;48:191–214.
20. Hammen C. Stress and depression. Annu Rev Clin Psychol 2005;1:293–319.
21. Dunn EC, Gilman SE, Slopen N, Willett JB, Molnar BE. The impact of exposure to interpersonal violence on gender differences in adolescent-onset major depression: results from the National Comorbidity Survey Replication (NCS-R). Depress Anxiety 2012;29:392–9.
22. Dunn EC, McLaughlin KA, Slopen N, Rosand J, Smoller JW. Developmental timing of child maltreatment and symptoms of depression and suicidality in young adulthood: results from the National Longitudinal Study of Adolescent Health. Depress Anxiety 2013;30:955–64.
23. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta analysis. Am J Psychiatry 2000;157:1552–62.
24. Rice F, Harold G, Thapar A. The genetic aetiology of childhood depression: a review. J Child Psychol Psychiatry 2002;43:65–79.
25. Dunlop BW, Nemeroff CB. The role of dopamine in the pathophysiology of depression. Arch Gen Psychiatry 2007;64:327–37.
26. Owens MJ, Nemeroff CB. Role of serotonin in the pathophysiology of depression: focus on the serotonin transporter. Clin Chem 1994;40:288–95.
27. Thase ME, ed. Neurobiological aspects of depression. 2nd ed. New York: Guilford, 2009.
28. Levinson DF. The genetics of depression: a review. Biol Psychiatry 2006;60:84–92.
29. Lohoff FW. Overview of the genetics of major depressive disorder. Curr Psychiatry Rep 2010;12:539–46.
30. Hettema JM. Genetics of depression. Focus 2010;8:316–22.
31. Shyn SI, Hamilton SP. The genetics of major depression: moving beyond the monoamine hypothesis. Psychiatr Clin North Am 2010;33:125–40.
32. Cohen-Woods S, Craig IW, McGuffin P. The current state of play on the molecular genetics of depression. Psychol Med 2012;43:673–87.
33. Saveanu RV, Nemeroff CB. Etiology of depression: genetic and environmental factors. Psychiatr Clin North Am 2012;35:51–71.
34. Perlis RH. Pharmacogenomic testing and personalized treatment of depression. Clin Chem 2014;60:53–9.
35. Attia J, Ioannidis JP, Thakkinstian A, et al. How to use an article about genetic association: background concepts. JAMA 2009;301:74–81.
36. Attia J, Ioannidis JP, Thakkinstian A, et al. How to use an article about genetic association: are the results of the study valid? JAMA 2009;301:191–7.
37. Corvin A, Craddock N, Sullivan PF. Genome-wide association studies: a primer. Psychol Med 2010;40:1063–77.
38. Manolio TA. Genomewide association studies and assessment of the risk of disease. N Engl J Med 2010;36:166–76.
39. Balding DJ. A tutorial on statistical methods for population association studies. Nat Rev Genet 2006;7:781–91.
40. Pearson TA, Manolio TA. How to interpret a genome-wide association study. JAMA 2008;299:1335–44.
41. National Human Genome Research Institute. A catalog of published genome-wide association studies. 2013. At http://www.genome.gov/gwastudies
42. Sullivan PF, de Geus EJ, Willemsen G, et al. Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. Mol Psychiatry 2009;14:359–75.
43. Rietschel M, Mattheisen M, Frank J, et al. Genome-wide association-, replication-, and neuroimaging study implicates HOMER1 in the etiology of major depression. Biol Psychiatry 2010;68:578–85.
44. Terracciano A, Tanaka T, Sutin AR, et al. Genome-wide association scan of trait depression. Biol Psychiatry 2010;68:811–7.
45. Muglia P, Tozzi F, Galwey NW, et al. Genome-wide association study of recurrent major depressive disorder in two European case-control cohorts. Mol Psychiatry 2010;15:589–601.
46. Lewis CM, Ng MY, Butler AW, et al. Genome-wide association study of major recurrent depression in the U.K. population. Am J Psychiatry 2010;167:949–57.
47. Shi J, Potash JB, Knowles JA, et al. Genome-wide association study of recurrent early-onset major depressive disorder. Mol Psychiatry 2011;16:193–201.
48. Aragam N, Wang KS, Pan Y. Genome-wide association analysis of gender differences in major depressive disorder in the Netherlands NESDA and NTR population-based samples. J Affect Disord 2011;133:516–21.

49. Kohli MA, Lucae S, Saemann PG, et al. The neuronal transporter gene SLC6A15 confers risk to major depression. Neuron 2011;70:252–65.
50. Shyn SI, Shi J, Kraft JB, et al. Novel loci for major depression identified by genome-wide association study of Sequenced Treatment Alternatives to Relieve Depression and meta-analysis of three studies. Mol Psychiatry 2011;16:202–15.
51. Wray NR, Pergadia ML, Blackwood DH, et al. Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. Mol Psychiatry 2012;17:36–48.
52. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke S, Wray NR, Lewis CM, et al. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry 2013;18:497–511.
53. Power RA, Keers R, Ng MY, et al. Dissecting the genetic heterogeneity of depression through age at onset. Am J Med Genet B Neuropsychiatr Genet 2012;159B:859–68.
54. Luciano M, Huffman JE, Arias-Vasquez A, et al. Genome-wide association uncovers shared genetic effects among personality traits and mood states. Am J Med Genet B Neuropsychiatr Genet 2012;159B:684–95.
55. Hek K, Demirkan A, Lahti J, et al. A genome-wide association study of depressive symptoms. Biol Psychiatry 2013;73:667–78.
56. Power RA, Cohen-Woods S, Ng MY, et al. Genome-wide association analysis accounting for environmental factors through propensity-score matching: application to stressful live events in major depressive disorder. Am J Med Genet B Neuropsychiatr Genet 2013;162B:521–9.
57. Hek K, Mulder CL, Luijendijk HJ, et al. The PCLO gene and depressive disorders: replication in a population-based study. Hum Mol Genet 2010;15:731–4.
58. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty, and challenges. Nat Rev Genet 2008;9:356–69.
59. Donnelly P. Progress and challenges in genome-wide association studies in humans. Nature 2008;456:728–31.
60. Psychiatric GWAS Consortium Coordinating Committee. Genomewide association studies: history, rationale, and prospects for psychiatric disorders. Am J Psychiatry 2009;166:540–56.
61. Sullivan PF. The Psychiatric GWAS Consortium: big science comes to psychiatry. Neuron 2010;68:182–6.
62. Psychiatric GWAS Consortium Steering Committee. A framework for interpreting genome-wide association studies of psychiatric disorders. Mol Psychiatry 2009;14:10–7.
63. Logue MW, Baldwin C, Guffanti G, et al. A genome-wide association study of post-traumatic stress disorder identifies the retinoid-related orphan receptor alpha (RORA) gene as a significant risk locus. Mol Psychiatry 2013;18:937–42.
64. Lopez-Leon S, Janssens AC, Gonzalez-Suloeta Ladd AM, et al. Meta-analyses of genetic studies on major depressive disorder. Mol Psychiatry 2008;13:772–85.
65. Lopez-Leon S, Croes EA, Sayed-Tabatabaei FA, Claes SJ, Van Broeckhoven C, van Duijn CM. The dopamine D4 receptor gene 48-base-pair-repeat polymorphism and mood disorders: a meta analysis. Biol Psychiatry 2005;57:999–1003.
66. Haenisch B, Herms S, Mattheisen M, et al. Genome-wide association data provide further support for an association between 5-HTTLPR and major depressive disorder. J Affect Disord 2013;146:438–40.
67. Vinkhuyzen AA, Dumenil T, Ryan L, et al. Identification of tag haplotypes for 5HTTLPR for different genome-wide SNP platforms. Mol Psychiatry 2011;16:1073–5.
68. Moffitt TE, Caspi A, Rutter M. Strategy for investigating interactions between measured genes and measured environments. Arch Gen Psychiatry 2005;62:473–81.
69. Moffitt TE, Caspi A, Rutter M. Measured gene-environment interactions in psychopathology: concepts, research strategies, and implications for research, intervention, and public understanding of genetics. Perspect Psychol Sci 2006;1:5–27.
70. Dunn EC, Uddin M, Subramanian SV, Smoller JW, Galea S, Koenen KC. Gene-environment interaction (G×E) research in youth depression: a systematic review with recommendations for future research. J Child Psychol Psychiatry 2011;52:1223–38.
71. Monroe SM, Simons AD. Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. Psychol Bull 1991;110:406–25.
72. Bakermans-Kranenburg MJ, Van IMH, Pijlman FT, Mesman J, Juffer F. Experimental evidence for differential susceptibility: dopamine D4 receptor polymorphism (DRD4 VNTR) moderates intervention effects on toddlers' externalizing behavior in a randomized controlled trial. Dev Psychol 2008;44:293–300.
73. Brody GH, Beach SR, Philibert RA, Chen YF, Murry VM. Prevention effects moderate the association of 5-HTTLPR and youth risk behavior initiation: gene x environment hypotheses tested via a randomized prevention design. Child Dev 2009;80:645–61.
74. Belsky J, Pleuss M. Beyond diathesis stress: differential susceptibility to environmental influences. Psychol Bull 2009;135:885–908.
75. Ellis BJ, Boyce WT. Biological sensitivity to context. Curr Dir Psychol Sci 2008;17:183–7.
76. Caspi A, Sugden K, Moffitt TE, et al. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. Science 2003;301:386–9.
77. Chapman DP, Whitfield CL, Felitti VJ, Dube SR, Edwards VJ, Anda RF. Adverse childhood experiences and the risk of depressive disorders in adulthood. J Affect Disord 2004;82:217–25.
78. McLaughlin KA, Greif Green J, Gruber MJ, Sampson NA, Zaslavsky AM, Kessler RC. Childhood adversities and adult psychiatric disorders in the National Comorbidity Survey Replication II: associations with persistence of DSM-IV disorders. Arch Gen Psychiatry 2010;67:124–32.
79. Nugent NR, Tyrka AR, Carpenter LL, Price LH. Gene-environment interactions: early life stress and risk for depressive and anxiety disorders. Psychopharmacology (Berl) 2011;214:175–96.
80. Risch N, Herrell R, Lehner T, et al. Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta analysis. JAMA 2009;301:2462–71.
81. 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. Arch Gen Psychiatry 2011;68:444–54.
82. Uher R, McGuffin P. The moderation by the serotonin transporter gene of environmental adversity in the aetiology of mental illness: review and methodological analysis. Mol Psychiatry 2008;13:131–46.
83. Uher R, McGuffin P. The moderation by the serotonin transporter gene of environmental adversity in the etiology of depression: 2009 update. Mol Psychiatry 2010;15:18–22.
84. Munafo MR, Durrant C, Lewis G, Flint J. Gene X environment interactions at the serotonin transporter locus. Biol Psychiatry 2009;65:211–9.
85. Brown GW, Harris TO. Depression and the serotonin transporter 5-HTTLPR polymorphism: a review and a hypothesis concerning gene-environment interaction. J Affect Disord 2008;111:1–12.
86. Wankerl M, Wust S, Otte C. Current developments and controversies: does the serotonin transporter gene–linked polymorphic

region (5-HTTLPR) modulate the association between stress and depression? Curr Opin Psychiatry 2010;23:582–7.
87. Caspi A, Hariri AR, Holmes A, Uher R, Moffitt TE. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. Am J Psychiatry 2010;167:1–19.
88. Kaufman J, Gelernter J, Kaffman A, Caspi A, Moffitt T. Arguable assumptions, debatable conclusions. Biol Psychiatry 2010; 67:e19–20.
89. Munafo MR, Durrant C, Lewis G, Flint J. Defining replication: a response to Kaufman and colleagues. Biol Psychiatry 2010; 67:e21–3.
90. Duncan LE, Pollastri AR, Smoller JW. Mind the gap: why many geneticists and psychological scientists have discrepant views about gene-environment interaction (G×E) research. Am Psychol 2014;69:249–68.
91. Fergusson DM, Horwood LJ, Miller AL, Kennedy MA. Life stress, 5-HTTLPR and mental disorder: findings from a 30-year longitudinal study. Br J Psychiatry 2011;198:129–35.
92. Munafo MR. The serotonin transporter gene and depression. Depress Anxiety 2012;29:915–7.
93. Appel K, Schwahn C, Mahler J, et al. Moderation of adult depression by a polymorphism in the FKBP5 gene and childhood physical abuse in the general population. Neuropsychopharmacology 2011;36:1982–91.
94. Bradley RG, Binder EB, Epstein MP, et al. Influence of child abuse on adult depression: moderation by the corticotropin-releasing hormone receptor gene. Arch Gen Psychiatry 2008; 65:190–200.
95. Lavebratt C, Aberg E, Sjoholm LK, Forsell Y. Variations in FKBP5 and BDNF genes are suggestively associated with depression in a Swedish population-based cohort. J Affect Disord 2010;125:249–55.
96. Zimmermann P, Brucki T, Nocon A, et al. Interaction of FKBP5 gene variants and adverse life events in predicting depression onset: results from a 10-year prospective study. Am J Psychiatry 2011;168:1107–16.
97. Gillespie CF, Phifer J, Bradley B, Ressler KJ. Risk and resilience: genetic and environmental influences on development of the stress response. Depress Anxiety 2009;26:984–92.
98. Hosang G, Shiles C, Tansey KE, McGuffin P, Uher R. Interaction between stress and the BDNF Val66Met polymorphism in depression: a systematic review and meta-analysis. BMC Med 2014;12:7.
99. Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nat Genet 2011;43:977–83.
100. Ferreira MA, O'Donovan MC, Meng Y, et al. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. Nat Genet 2008;40: 1056–8.
101. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. Nat Genet 2011;43:969–76.
102. Nyegaard M, Demontis D, Foldager L, Hedemand A, Flint TJ, et al. CACNA1C (rs1006737) is associated with schizophrenia. Mol Psychiatry 2010;15:119–21.
103. Hamshere ML, Walters JT, Smith R, Richards AL, Green E, et al. Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the Schizophrenia PGC. Mol Psychiatry 2013;18:708–12.
104. Ripke S, O'Dushlaine C, Chambert K, et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat Genet 2013;45:1150–9.
105. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011; 88:76–82.
106. Lubke GH, Hottenga JJ, Walters R, et al. Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. Biol Psychiatry 2012;72:707–9.
107. Lee SH, Ripke S, Neale BM, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nat Genet 2013;45:984–94.
108. Demirkan A, Penninx BW, Hek K, et al. Genetic risk profiles for depression and anxiety in adult and elderly cohorts. Mol Psychiatry 2011;16:773–83.
109. Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. Nat Rev Genet 2010;11: 843–54.
110. Holmans P. Statistical methods for pathway analysis of genome-wide data for association with complex genetic traits. Adv Genet 2010;72:141–79.
111. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28:27–30.
112. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25–9.
113. Lee PH, Perlis RH, Jung JY, et al. Multi-locus genome-wide association analysis supports the role of glutamatergic synaptic transmission in the etiology of major depressive disorder. Transl Psychiatry 2012;2:e184.
114. Kao CF, Jia P, Zhao Z, Kuo PH. Enriched pathways for major depressive disorder identified from a genome-wide association study. Int J Neuropsychopharmacol 2012;15:1401–11.
115. Sanders SJ, Ercan-Sencicek AG, Hus V, et al. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. Neuron 2011;70:863–85.
116. Sebat J, Lakshmi B, Malhotra D, et al. Strong association of de novo copy number mutations with autism. Science 2007;316: 445–9.
117. Stefansson H, Rujescu D, Cichon S, et al. Large recurrent microdeletions associated with schizophrenia. Nature 2008;455:232–6.
118. Purcell SM, Moran JL, Fromer M, et al. A polygenic burden of rare disruptive mutations in schizophrenia. Nature 2014;506: 185–90.
119. Malhotra D, McCarthy S, Michaelson JJ, et al. High frequencies of de novo CNVs in bipolar disorder and schizophrenia. Neuron 2011;72:951–63.
120. Bras J, Guerreiro R, Hardy J. Use of next-generation sequencing and other whole-genome strategies to dissect neurological disease. Nat Rev Neurosci 2012;12:453–64.
121. Rizzo JM, Buck MJ. Key principles and clinical applications of "next-generation" DNA sequencing. Cancer Prev Res (Phila) 2012;5:887–900.
122. Rucker JJ, Breen G, Pinto D, et al. Genome-wide association analysis of copy number variation in recurrent depressive disorder. Mol Psychiatry 2013;18:183–9.
123. Degenhardt F, Priebe L, Herms S, et al. Association between copy number variants in 16p11.2 and major depressive disorder in a German case-control sample. Am J Med Genet B Neuropsychiatr Genet 2012;159B:263–73.
124. Glessner JT, Wang K, Sleiman PM, et al. Duplication of the SLIT3 locus on 5q35.1 predisposes to major depressive disorder. PLoS One 2010;5:e15463.
125. O'Duschlaine C, Ripke S, Ruderfer DM, et al. Rare copy number variation in treatment-resistant major depressive disorder. Biol Psychiatry 2014;76:536–41.
126. Wain LV, Armour JA, Tobin MD. Genomic copy number variation, human health, and disease. Lancet 2009;374:340–50.

127. Thomas D. Gene-environment-wide association studies: emerging approaches. Nat Rev Genet 2010;11:259–72.
128. Thomas D. Methods for investigating gene-environment interactions in candidate pathway and genome-wide association studies. Annu Rev Public Health 2010;31:21–36.
129. Wu C, Kraft P, Zhai K, et al. Genome-wide association analyses of esophageal squamous cell carcinoma in Chinese identify multiple susceptibility loci and gene-environment interactions. Nat Genet 2012;44:1090–7.
130. Seigert S, Hampe J, Schafmayer C, et al. Genome-wide investigation of gene-environment interactions in colorectal cancer. Hum Genet 2013;132:219–31.
131. Cornelis MC, Tchetgen EJ, Liang L, et al. Gene-environment interactions in genome-wide association studies: a comparative study of tests applied to empirical studies of type 2 diabetes. Am J Epidemiol 2012;175:191–202.
132. Manning AK, Hivert M-F, Scott RA, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. Nat Genet 2012;44:659–69.
133. Hamza TH, Chen H, Hill-Burns EM, et al. Genome-wide gene-environment study identifies glutamate receptor gene GRIN2A as a Parkinson's disease modifier gene via interaction with coffee. PLoS Genet 2011;7:e1002237.
134. Hancock DB, Artigas MS, Gharib SA, et al. Genome-wide joint meta analysis of SNP and SNP-by-smoking interaction identifies novel loci for pulmonary function. Plos Genet 2012;8: e1003098.
135. Beaty TH, Ruczinski I, Murray JC, et al. Evidence for gene-environment interaction in a genome wide study of nonsyndromic cleft palate. Genet Epidemiol 2011;35:469–78.
136. Winham SJ, Biernacka JM. Gene-environment interactions in genome-wide association studies: current approaches and new directions. J Child Psychol Psychiatry 2013;54:1120–34.
137. Gauderman WJ, Zhang P, Morrison JL, Lewinger JP. Finding novel genes by testing G×E interactions in a genome-wide association study. Genet Epidemiol 2013;37:603–13.
138. Hutter CM, Mechanic LE, Chatterjee N, Kraft P, Gillanders EM; NCI Gene-Environment Think Tank. Gene-environment interactions in cancer epidemiology: a National Cancer Institute Think Tank report. Genet Epidemiol 2013;37:643–57.
139. Dunn EC, Masyn KE, Yudron M, Jones SM, Subramanian SV. Translating multilevel theory into multilevel research: challenge and opportunities for understanding the social determinants of psychiatric disorders. Soc Psychiatry Psychiatr Epidemiol 2014;49:859–72.
140. Wild CP. The exposome: from concept to utility. Int J Epidemiol 2012;41:24–32.
141. Patel CJ, Bhattacharya J, Butte AJ. An environment-wide association study (EWAS) on type 2 diabetes mellitus. PLoS One 2010;5:e10746.
142. Galatzer-Levy IR, Bryant RA. 626,120 ways to have posttraumatic stress disorder. Perspect Psychol Sci 2013;8:651–62.
143. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: American Psychiatric Publishing, 2013.
144. Teicher MH, Samson JA. Childhood maltreatment and psychopathology: a case for ecophenotypic variants as clinically and neurobiologically distinct subtypes. Am J Psychiatry 2014;170:1114–33.
145. Nemeroff CB, Heim CM, Thase ME, et al. Differential responses to psychotherapy versus pharmacotherapy in patients with chronic forms of major depression and childhood trauma. Proc Natl Acad Sci 2003;100:14293–6.
146. Kendler KS, Aggen SH, Neale MC. Evidence for multiple genetic factors underlying DSM-IV criteria for major depression. JAMA Psychiatry 2013;70:599–607.
147. Kendler KS, Gardner CO. Boundaries of major depression: an evaluation of DSM-IV criteria. Am J Psychiatry 1998; 155:172–7.
148. Takayanagi Y, Spira AP, Roth KB, Gallo JJ, Eaton WW, Mojtabai R. Accuracy of reports of lifetime mental and physical disorders: results from the Baltimore Epidemiological Catchment Area Study. JAMA Psychiatry 2014;71: 272–80.
149. Plomin R, Haworth CMA, Davis OSP. Common disorders are quantitative traits. Nat Rev Genet 2009;10:872–8.
150. van Loo HM, de Jonge P, Romeijn JW, Kessler RC, Schoevers RA. Data-driven subtypes of major depressive disorder: a systematic review. BMC Med 2012;10:156.
151. Kendler KS, Eaves LJ, Walters EE, Neale MC, Heath AC, Kessler RC. The identification and validation of distinct depressive syndromes in a population-based sample of female twins. Arch Gen Psychiatry 1996;53:391–9.
152. Lamers F, Burstein M, He J, Avenevoli S, Angst J, Merikangas KR. Structure of major depressive disorder in adolescents and adults in the US general population. Br J Psychiatry 2012;201: 143–50.
153. Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry 2003;160:636–45.
154. Mathews A, MacLeod C. Cognitive vulnerability to emotional disorders. Annu Rev Clin Psychol 2005;1:167–95.
155. Gotlib IH, Joormann J. Cognition and depression: current status and future directions. Annu Rev Clin Psychol 2010;6:285–312.
156. Pizzagalli DA, Holmes AJ, Dillon DG, et al. Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder. Am J Psychiatry 2009;166:702–10.
157. Hasler G, Drevets WC, Manji HK, Charney DS. Discovering endophenotypes for major depression. Neuropsychopharmacology 2004;29:1765–81.
158. Sanislow CA, Pine DS, Quinn JP, et al. Developing constructs for psychopathology research: research domain criteria. J Abnorm Psychol 2010;4:631–9.
159. Morris SE, Cuthbert BN. Research domain criteria: cognitive systems, neural circuits, and dimensions of behavior. Dialogues Clin Neurosci 2012;14:29–37.
160. Craddock N, Owen MJ. The Kraepelinian dichotomy—going, going . . . but still not gone. Br J Psychiatry 2010;196:92–5.
161. Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry 2010;167:748–51.
162. Stein JL, Medland SE, Vasquez AA, et al. Identification of common variants associated with human hippocampal and intracranial volumes. Nat Genet 2012;44:552–61.