Overview

Recent Molecular Genetic Studies and Methodological Issues in Major Depression Research

Shih-Jen Tsai, M.D.¹²*

Major depressive disorder (MDD) is a genetically complex mental disorder involving numerous susceptible genes whose expression may be modulated by environmental factors. In the past decade, case-control association approaches have been the mainstream in MDD genetic studies. But whole genome gene-expression microarray and genome-wide association studies have begun to emerge in recent years. Genetic association studies have suggested several genes related to MDD, although the results to date are inconclusive and no single genetic variation has been identified to increase the risk of depression substantially. Recent approaches based on gene-environment interaction have demonstrated more consistent findings. Therefore, the usefulness from the study results of case-control studies is limited by present knowledge of MDD pathophysiology. Genome-wide association studies are not constrained by our limited knowledge. Although the results of several recent genome-wide association studies did not reach the desired level of statistical significance, these studies do support interesting candidate genes and genomic regions for further study. In short, the field of genetic studies in general has been disappointing because to find common gene variants of large effect in MDD pathogenesis has been unsuccessful. However, the field of psychiatric genetics is rapidly growing, and several new technological advances (e.g. whole-genome sequencing) will be soon available for large-scale studies. These approaches provide exciting new avenues to identify new candidate genes for MDD. A better understanding of the MDD-related genes may potentially lead to developing effective prevention and treatment of this disease.

Key words: genetics, association study, genome-wide association study, epigenetics

Introduction

Major depressive disorder (MDD) is a heterogeneous, highly prevalent, and moderately heritable mental disorder. Estimated heritability of MDD based on monozygotic vs. dizygotic twin concordance differences exhibits a modest heritable contribution of 37% [1]. Adoption studies also have suggested both an important genetic impact of parental depression, and a significant environmental impact of maternal depression in mediating depression among adopted adolescents [2].

Although genetic factors play a role in MDD, identifying specific genes involved has proved challenging. In the past decade, molecular-genetic technologies have made great advances in genome-wide searches for disease-causing genes with the linkage disequilibrium (LD) approach and genome-wide studies of gene expression and chromatin modifications that reflect the epigenetic response of the genome to environmental exposure. The technological revolution has shifted MDD genetic study from case-control association studies (focused on hypothesis-driven candidates) to genome-wide searches (employed an unbiased exploration using linked polymorphic markers that span the whole genome in search of previously unimplicated loci). Here, the author intends to provide an overview of recent genetic studies of MDD with different approaches to advance knowledge of the genetic bases of MDD, and to explore methodological issues in MDD genetic studies.

Candidate Gene Association Studies

Candidate gene association studies test genetic polymorphisms in genes involved in disease pathogenesis for statistical association. The classic design of a genetic association analysis is a case-control study in which the frequency of a possible risk allele of a candidate gene in unrelated individuals affected by the same disease (case group) is compared to the frequency observed in healthy individuals of the same ethnic group (control group). If the risk allele/genotype analyzed is found more frequently in the case group than in controls, an association exists between the genetic variant and the disease. This means that the presence of this risk allele increases the risk of, or susceptibility to, the disease. But researchers need to interpret positive association finding with great caution since many reports have problems with the study design and/or statistical analysis. In addition, population stratification, insufficient sample power, genotyping problems, and phenotype heterogeneity could cause false positive or negative findings in genetic association studies. Given these limitations, candidate gene association research is still the most popular method to investigate genetic factors in MDD due to its simple technique. Methodological aspects of the candidate gene association studies have progressed to use increasingly larger case control samples or meta-analytic summaries to increase the sample size.

Using key words “depressi*” AND ((genetic association) OR polymorphism*); filters activated: humans” in a PubMed search, I found about three hundreds of MDD genetic association studies. Among them, 36 reports are from Taiwan (Table 1) [3-38]. In 2008, Lopez-Leon et al. reviewed MDD association studies reported in 183 papers that studied 393 polymorphisms in 102 genes [39]. They found 22 of these polymorphisms were investigated by three or more studies allowing a meta-analysis, and found that significant evidence for association for APOE (apolipo-
<table>
<thead>
<tr>
<th>Reference number</th>
<th>Polymorphism /gene</th>
<th>Result</th>
<th>Authors</th>
<th>Publication year</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>T102C /HTR2A</td>
<td>No association</td>
<td>Tsai SJ et al.</td>
<td>1999</td>
</tr>
<tr>
<td>4</td>
<td>A218C /TPH1</td>
<td>Association</td>
<td>Tsai SJ et al.</td>
<td>1999</td>
</tr>
<tr>
<td>5</td>
<td>(AAT)n repeat /CNR1</td>
<td>No association</td>
<td>Tsai SJ et al.</td>
<td>2001</td>
</tr>
<tr>
<td>6</td>
<td>C825T /GNB3</td>
<td>No association</td>
<td>Lin CN et al.</td>
<td>2001</td>
</tr>
<tr>
<td>7</td>
<td>2 bp deletion /CHRNA7</td>
<td>Modest association</td>
<td>Lai IC et al.</td>
<td>2001</td>
</tr>
<tr>
<td>8</td>
<td>C267T /HTR6</td>
<td>No association</td>
<td>Wu WH et al.</td>
<td>2001</td>
</tr>
<tr>
<td>9</td>
<td>I/D polymorphism /ACE</td>
<td>No association</td>
<td>Hong CJ et al.</td>
<td>2002</td>
</tr>
<tr>
<td>10</td>
<td>5-HTTVNTR /SLC6A4</td>
<td>No association</td>
<td>Yen FC et al.</td>
<td>2003</td>
</tr>
<tr>
<td>11</td>
<td>Val66Met /BDNF</td>
<td>No association</td>
<td>Hong CJ et al.</td>
<td>2003</td>
</tr>
<tr>
<td>12</td>
<td>2 SNPs /ESR1</td>
<td>No association</td>
<td>Tsai SJ et al.</td>
<td>2003</td>
</tr>
<tr>
<td>13</td>
<td>C-511T /IL1B</td>
<td>No association</td>
<td>Yu YW et al.</td>
<td>2003</td>
</tr>
<tr>
<td>14</td>
<td>C276T /NOS1</td>
<td>No association</td>
<td>Yu YW et al.</td>
<td>2003</td>
</tr>
<tr>
<td>15</td>
<td>Val66Met /BDNF</td>
<td>No association</td>
<td>Tsai SJ et al.</td>
<td>2003</td>
</tr>
<tr>
<td>16</td>
<td>Val66Met /BDNF</td>
<td>No association</td>
<td>Tsai SJ et al.</td>
<td>2003</td>
</tr>
<tr>
<td>17</td>
<td>A-161T /HTR1B</td>
<td>No association</td>
<td>Tsai SJ et al.</td>
<td>2004</td>
</tr>
<tr>
<td>18</td>
<td>uVNTR /MAOA</td>
<td>No association</td>
<td>Yu YW et al.</td>
<td>2005</td>
</tr>
<tr>
<td>19</td>
<td>C-634G /IL6</td>
<td>No association</td>
<td>Hong CJ et al.</td>
<td>2005</td>
</tr>
<tr>
<td>20</td>
<td>C197T6 /ADORA2A2</td>
<td>No association</td>
<td>Tsai SJ et al.</td>
<td>2006</td>
</tr>
<tr>
<td>21</td>
<td>Val66Met /BDNF</td>
<td>Met66 allele was associated with geriatric MDD</td>
<td>Hwang JP et al.</td>
<td>2006</td>
</tr>
<tr>
<td>22</td>
<td>epsilon /APOE</td>
<td>A lower frequency of epsilon2 allele in MDD</td>
<td>Fan PL et al.</td>
<td>2006</td>
</tr>
<tr>
<td>23</td>
<td>epsilon /APOE</td>
<td>No association</td>
<td>Hwang JP et al.</td>
<td>2006</td>
</tr>
<tr>
<td>25</td>
<td>5 SNPs /NGFR</td>
<td>No association</td>
<td>Gau YT et al.</td>
<td>2008</td>
</tr>
<tr>
<td>26</td>
<td>4 SNPs /GSK3B</td>
<td>No association</td>
<td>Tsai SJ et al.</td>
<td>2008</td>
</tr>
<tr>
<td>27</td>
<td>6 SNPs /SERPINE1</td>
<td>rs2227684-G and rs7242-T alleles were more frequent in MDD</td>
<td>Tsai SJ et al.</td>
<td>2008</td>
</tr>
<tr>
<td>28</td>
<td>3 SNPs /S100A10</td>
<td>No association</td>
<td>Tzang RF et al.</td>
<td>2008</td>
</tr>
<tr>
<td>29</td>
<td>uVNTR, rs1137070/ MAOA</td>
<td>Association in female MDD</td>
<td>Huang SY et al.</td>
<td>2009</td>
</tr>
<tr>
<td>30</td>
<td>5 SNPs /TPH2</td>
<td>rs17110747-G homozygote was more frequent in MDD</td>
<td>Tsai SJ et al.</td>
<td>2009</td>
</tr>
<tr>
<td>31</td>
<td>6 SNPs /TPH2</td>
<td>2755A allele was associated with female MDD</td>
<td>Lin YM et al.</td>
<td>2009</td>
</tr>
<tr>
<td>32</td>
<td>C-511T /IL1B</td>
<td>No association</td>
<td>Hwang JP et al.</td>
<td>2009</td>
</tr>
<tr>
<td>33</td>
<td>7 SNPs /VEGFA</td>
<td>No association</td>
<td>Tsai SJ et al.</td>
<td>2009</td>
</tr>
<tr>
<td>34</td>
<td>5 SNPs /KCNK2</td>
<td>rs6685629 was associated with MDD</td>
<td>Liou YJ et al.</td>
<td>2009</td>
</tr>
<tr>
<td>35</td>
<td>Val66Met /BDNF &amp; 4 SNPs /NTRK2</td>
<td>BDNF and NTRK2 interacted in MDD susceptibility</td>
<td>Lin E et al.</td>
<td>2009</td>
</tr>
<tr>
<td>36</td>
<td>6 SNPs /PAWR</td>
<td>rs8176874 was associated with MDD</td>
<td>Liou YJ et al.</td>
<td>2011</td>
</tr>
<tr>
<td>37</td>
<td>17 SNPs /SLC6A3</td>
<td>Modest association</td>
<td>Huang CC et al.</td>
<td>2011</td>
</tr>
<tr>
<td>38</td>
<td>uVNTR /MAOA</td>
<td>MAOA 4R allele was more frequent in male MDD</td>
<td>Lung FW et al.</td>
<td>2011</td>
</tr>
</tbody>
</table>

SNP, single-nucleotide polymorphism; Bp, base pairs. uVNTR, upstream variable-number tandem repeat; I/D, insertion-deletion.
protein E), \textit{DRD4} (D4 subtype of the dopamine receptor), \textit{GNB3} (guanine nucleotide binding protein (G protein), beta polypeptide 3), \textit{MTHFR} (methylene tetrahydrofolate reductase), \textit{SLC6A3} (gene encoding the dopamine transporter) and \textit{SLC6A4} (gene encoding the serotonin transporter).

Several candidate genes deserve mentioning, as they have been suggested repeatedly to be implicated in MDD genetic association studies.

**Serotonin transporter (5-HTT)**

Following serotonin release, the serotonin transporter is the major site of serotonin reuptake into the presynaptic neuron. By regulating the reuptake of the released serotonin, the serotonin transporter is central to fine-tune the serotonergic neurotransmission. The \textit{SLC6A4} gene encoding serotonin transporter has been one of the genes most studied in MDD because serotonin transporter is the major therapeutic target of selective serotonin reuptake inhibitors (SSRIs) and serotonin-and-norepinephrine reuptake inhibitors (SNRIs). Among the 5-HTT genetic variants, a 44-base pairs (bp) insertion-deletion in the promoter region (\textit{5-HTTLPR}) polymorphism has been extensively studied for association with MDD. The \textit{5-HTTLPR} short variant suppresses transcriptional activity of the promoter [40]. A meta-analysis in 2010 indicated evidence of a small but statistically significant association between the \textit{5-HTTLPR} polymorphism and MDD (odds ratio [OR] 1.08, 95% confidence interval [CI] 1.03-1.12) [41]. A recent study also found that \textit{5-HTTLPR} is functionally trisomic [42]. The presence of an A-to-G single nucleotide polymorphism further subdivides the long allele into La or Lg, which is equivalent to the S allele in reduced functional activity. Using the trisomic classification, Segal et al. in 2009 compared the frequency of \textit{5-HTTLPR} alleles in 94 MDD patients who had history of attempted suicide with 94 controls who were free of any psychiatric disorder, showing no distributional difference between those two groups [43].

**Tryptophan hydroxylase-1 (TPH1)**

\textit{TPH1} encodes tryptophan hydroxylase-1, which is the rate-limiting enzyme in serotonin synthesis. The \textit{TPH1} A218C (rs1800532) polymorphism has been the most extensively studied \textit{TPH1} polymorphism in MDD. In 1997, Mann et al. reported that the \textit{TPH1} A218C polymorphism is associated with suicidal behaviors in patients with major depression [44]. Our study replicated their finding that this \textit{TPH1} polymorphism is associated with major depression and the association was most significant in patients who attempted suicide [4]. But a meta-analysis across eight MDD samples (1,812 cases and 2,223 controls) has not shown any association for any rs1800532 genotype and MDD [45].

**Brain-derived neurotrophic factor (BDNF)**

Brain-derived neurotrophic factor is a member of the neurotrophin family and plays an important role in neuronal survival and brain plasticity. Thus, \textit{BDNF} is considered an attractive candidate gene for the study of healthy and diseased brain function and behaviors. Evidence from animal and clinical studies has implicated decreased central BDNF activity in the pathogenesis of MDD; therefore, activating the BDNF-dependent pathway plays an important role in the mechanism of antidepressant therapeutic action [46].

In 2003, we studied the association between the functional \textit{BDNF} Val66Met polymorphism and MDD outpatients [15] and inpatients [16], and demonstrated negative findings. The negative association between \textit{BDNF} Val66Met polymor-
phism and MDD is supported by most of the following studies, suggesting that the BDNF Val66Met polymorphism is not a major contributing factor to MDD susceptibility. In 2006, we tested the association between the BDNF Val66Met polymorphism and geriatric MDD, and found that a highly significant excess of the Met66 allele is present in the geriatric MDD patients [21]. The depolarization-induced secretion of BDNF was reduced in Met66 BDNF-transfected neurons compared with Val66 BDNF analogs [47]. Since subjects carrying the Met66/Met66 genotype may have lower BDNF activity than Val66 carriers, the increased Met66/Met66 frequency in geriatric MDD patients is consistent with the hypothesis that MDD is associated with reduced BDNF activity in the brain [46]. A comparison study with 245 elderly depressed white subjects and 94 elderly white normal controls replicated our findings indicating geriatric depressed subjects are more likely to be Met66 allele carriers than the normal controls [48]. Verhagen et al. recently did a meta-analysis of the BDNF Val66Met polymorphism in MDD of 14 studies involving 2,812 MDD cases and 10,843 non-depressed controls [49]. It revealed that the BDNF Val66Met polymorphism is not significantly associated with MDD in the total sample; however, the gender stratified analyses revealed significant effects in both the allelic and genotypic analyses in men (OR [Met66], 95% CI; 1.27 [1.10-1.47]) ; OR (Met66/Met66), 95% CI; 1.67 [1.19-2.36]) but not in women.

**Gene-environment interaction and gene-gene interaction**

Epidemiologic studies suggest that MDD results from multiple interdependent environmental factors (such as childhood maltreatment, including psychological, physical, and sexual abuse and neglect) and genetic risk factors that are alone neither necessary nor sufficient for the disorder to develop [1]. Gene-environment interaction (G×E) studies suggest that by moderating the effects of the environment, genetic variation explains why some individuals are vulnerable and others are resistant to the effects of adversity.

Caspi et al. were the first investigators to show that childhood maltreatment and later stressful life events can predict the onset of depressive symptoms only in genetically predisposed individuals with a short allele of the 5-HTTLPR polymorphism, while the long allele carriers were found to be more resilient to depression after exposing to adverse life events [50]. In a recent meta-analysis of 54 studies, 5-HTTLPR has been found to moderate the relationship between stress and depression, with the 5-HTTLPR short allele associated with an increased risk of developing depression under stress (p = 0.00002) [51]. Significant G×E interactions on MDD susceptibility have also been reported for BDNF, FKBP5 and CRHR1 genetic variants [52].

Because MDD may be related to multiple genes, several studies have investigated gene-gene (G×G) interaction in MDD. For example, in genetic study of geriatric depression, we found that BDNF Val66Met polymorphism and NTRK2 (BDNF receptor gene) genetic variants interact in the 2-locus, 3-locus, 4-locus, and 5-locus gene-gene interaction models using a generalized multifactor dimensionality reduction (GMDR) method [35]. The results suggest that the BDNF and NTRK2 genes may contribute to the risk of geriatric depression in an interactive manner.

**Linkage Studies**

Linkage studies are genetic analyses of pedigrees and are well-suited for studying disorders
with a clear pattern of Mendelian inheritance and disorders driven by genes of large effect. While both conditions are less likely to operate in depression, there have been few reports of linkage studies in MDD.

Supportive findings of many MDD linkage studies include regions on chromosomes 1, 3, 4, 6, 8, 11, 12, 15, and 18 [53]. A major limitation of family-based linkage studies is an inherent difficulty to ascertain sizable numbers of families with multiple affected relatives sharing restrictive phenotype characteristics that will allow a large enough sample power to detect variants with a small effect size. Further, large chromosomal regions shared among family members constrain the narrowing down of a linkage signal sufficiently to identify a causative gene.

**Genome-wide Scans and DNA Microarray**

Technical improvements in genotyping in the past decade have led to markedly increased genotyping capacity at reduced cost. Combined with the identification of millions of common polymorphisms through the HapMap project, high-throughput genotyping approaches and bioinformatics analysis have made genome-wide association study (GWAS) a feasible study design to identify genetic loci and ultimately candidate genes associated with complex psychiatric disorders. This method is based on genotyping arrays or microarrays that allow the variability of the human genome to be traced for assessing the hypothesis of common disease-common variant without the need to conduct a hypothesis-driven study of the etiology of the disease. Limitations of this approach are the immense amount of data, costs, and issues regarding multiple testings. Because of the unprecedented amount of hypothesis testing (and the resultant inflation of type I error), the standard in the field has been to set “genome-wide significance” at $p \leq 1.0 \times 10^{-8}$, which is highly restrictive and difficult to attain.

For MDD, there are now five published GWAS studies. In 2009, Sullivan et al. first used a sample of 1,738 MDD cases and 1,802 controls from the Netherlands [54]. Among the 435,291 single-nucleotide polymorphisms (SNPs), the authors found the top signal to be at rs1558477 (trend-test $p$-value of $1 \times 10^{-6}$), 12.4 kb downstream of ADCYAP1R1 gene. Additional GWAS of MDD have followed. Disappointingly, there were no genome-wide significant findings in these GWAS studies. Although the results of each individual GWAS studies did not reach the desired level of statistical significance, they do suggest interesting candidate genes that may be worthwhile following up in future studies. It is becoming increasingly clear that individual genetic susceptibility factors for MDD are likely to have only minor effects, and large sample size will be necessary to identify them.

Microarray technology examines the expression levels of thousands of gene transcripts in postmortem brains of suicide completers. This technology has opened new avenues for discovering MDD pathogenesis and may aid in identifying candidate genes for a further case-control association study. Using this technique, Sequeira et al. demonstrated lower expression of the spermine/spermidine N1-acetyltransferase (SSAT) gene, the rate-limiting enzyme in the catabolism of polyamines, in the brain of suicide completers and further confirmed the role of SSAT in suicide behaviors by genetic case-control analysis [55].
**Epigenetics**

Epigenetics comprises mechanisms of gene expression modifications that do not alter the genomic code itself. The two major epigenetic mechanisms are DNA methylation and histone modification. Epigenetic mechanisms may be involved in the reprogramming of gene expression in response to stressful stimuli and have long-lasting effects within mature neurons. Studies in animals have shown that epigenetic mechanisms with histone modifications or DNA methylation affect diverse pathways leading to depression-like behaviors [56]. Human studies have revealed the GABA$_A$ receptor promoter [57] and the BDNF promoter hypermethylation [58] in the brain of depressed suicide completers. Recent study using genome-wide DNA methylation scan in postmortem frontal cortex of 39 MDD samples and 26 controls identified that 224 candidate regions with DNA methylation differences are greater than 10% [59]. These regions are highly enriched for neuronal growth and development genes. The above-mentioned findings suggested that epigenetic pharmaceuticals, such as DNA methylation inhibitors, are promising and deserves further attention in developing depression treatment.

**Conclusion**

MDD is a complex and multifactorial mental disorder with important genetic and nongenetic contributory factors. Although many genetic association studies have investigated MDD, at present no identified gene are likely to be found as the specific cause of MDD. Researchers have reported many genetic polymorphisms are associated with MDD, but their studies cannot be duplicated later. Here, the author would like to conclude this overview with six comments on methodological issues for future MDD genetic studies to improve future consistent findings between studies.

- Many MDD genetic association studies to date address only hypothesis-driven candidate genes that mostly revolve around known mechanisms of antidepressant drugs. Novel ground-breaking hypotheses for MDD are needed to produce new sets of candidate genes. Furthermore, most genetic studies have devoted disproportionate focus to being on few highly reported functional polymorphisms (DNA sequence variations that alter the expression and/or functioning of the gene product), neglecting comprehensive coverage of the coding and promoter regions of the gene of interest.

- One strategy to investigate gene effects in etiologically complex diseases is through deconstructing complexity into parts that are likely to be less etiologically heterogeneous, such as intermediate phenotypes [60]. Heritable intermediate phenotypes that are disease-associated have been termed endophenotypes. Several promising endophenotypes proposed for MDD genetic studies include neuroticism, the dexamethasone suppression test and the neuroimaging [60].

- We must not overlook the role of the environment factors in studying genetics. The most consistent MDD genetic findings are the GxE studies (e.g. 5-HTTLPR polymorphism and adverse life events). It should become possible to study gene-environment interactions by conducting larger-scale molecular studies on samples drawn from longitudinal cohorts and by measuring these variables in samples collected for genetic studies. More intensive studies to measure factors like childhood trauma and life stress are needed.
Although candidate gene case-control studies have value in MDD genetic studies, they are naturally limited to the study of genes that have some contributing function to the development of MDD. Previous genetic studies have focused on monoaminergic systems, particularly the serotonergic system, and recent studies have focused on genes in neurotrophic systems. But many genes with unknown function may contribute to causing MDD. Genome-wide studies of gene expression and genetic variation are now possible and are not constrained by our limited knowledge. Although these unbiased approaches are at the infancy stage, they have already provided some new avenues to identify new candidate genes for MDD genetic studies and to confirm the findings in genome-wide studies.

Genome-wide association studies using indirect mapping have been successful in localizing genes associated with complex diseases when common variants are the underlying cause of disease etiology. A strong LD between tagSNPs and underlying causal variants makes it feasible to use indirect mapping. To detect associations with rare variants, indirect LD mapping will be low-powered due to weak correlations between common tagSNPs and rare causal variants. But for associations with rare variants, to perform direct mapping and to first identify rare variants within a sample are necessary. Sequencing candidate genes or entire genomes are the optimal way to identify rare variants. Compared with present clinical genetic testing, whole genome sequencing greatly expands the breadth of testing from genes associated with MDD to the whole genome and, potentially, all the information that the genome contains about MDD.

MDD is a complex disorder and the predisposition towards MDD consists of numerous genetic factors and non-genetic factors. Genes interact not only among each other, but also environmental factors may interplay with SB. Using statistical methods that account for multi-factorial etiology of MDD would be important in future research. For example, the GMDR method, which extends the MDR method, is a non-parametric data-mining approach applied to continuous and dichotomous phenotypes. Additionally, GMDR permits adjusting discrete and quantitative covariates, and has been used in various population-based studies with unbalanced groups of case and control subjects.

**Summary**

Many genes are likely responsible for MDD susceptibility. A better understanding of the genes involved in MDD and their interaction with non-genetic factors such age, gender, and stress, may help develop more effective prevention, diagnosis and management of MDD.

**Acknowledgements**

Fundings for this report were provided by grant NSC 101-2314-B-075-040 from the National Science Council, Taiwan and by grant VGHUST102-G1-2-1 from the Taipei Veterans General Hospital.

**References**


22. Pan PL, Chen CD, Kao WT, Shu BC, Lung FW: Protective effect of the apo epsilon2 allele in major


