

The GWAS Risk Genes for Depression May Be Actively Involved in Alzheimer's Disease

Hua Ni^{a,1}, Min Xu^{b,c,1}, Gui-Lai Zhan^{a,1}, Yu Fan^b, Hejiang Zhou^b, Hong-Yan Jiang^d, Wei-Hong Lu^e, Liwen Tan^f, Deng-Feng Zhang^{b,*}, Yong-Gang Yao^{b,c,g,h,*} and Chen Zhang^{e,*}

^aCenter for Disease Control and Prevention, Shanghai Xuhui Mental Health Center, Shanghai, China

^bKey Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences and Yunnan Province, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

^cKunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan, China

^dDepartment of Psychiatry, the First Affiliated Hospital of Kunming Medical University, Kunming, Yunnan, China

^eDepartment of Psychiatry, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China

^fMental Health Institute of the Second Xiangya Hospital, Central South University, Changsha, China

^gCAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China

^hKIZ/CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming, Yunnan, China

Handling Associate Editor: Yan-Jiang Wang

Accepted 30 May 2018

Abstract. Depression is one of the most frequent psychiatric symptoms observed in people during the development of Alzheimer's disease (AD). We hypothesized that genetic factors conferring risk of depression might affect AD development. In this study, we screened 31 genes, which were located in 19 risk loci for major depressive disorder (MDD) identified by two recent large genome-wide association studies (GWAS), in AD patients at the genomic and transcriptomic levels. Association analysis of common variants was performed by using summary statistics of the International Genomics of Alzheimer's Project (IGAP), and association analysis of rare variants was conducted by sequencing the entire coding region of the 31 MDD risk genes in 107 Han Chinese patients with early-onset and/or familial AD. We also quantified the mRNA expression alterations of these MDD risk genes in brain tissues of AD patients and AD mouse models, followed by protein-protein interaction network prediction to show their potential effects in AD pathways. We found that common and rare variants of *L3MBTL2* were significantly associated with AD. mRNA expression levels of 18 MDD risk genes, in particular *SORCS3* and *OAT*, were differentially expressed in AD brain tissues. 13 MDD risk genes were predicted to physically interact with core AD genes. The involvement of *HACE1*, *NEGR1*, and *SLC6A15* in AD was supported by convergent lines of evidence. Taken together, our results showed that MDD risk genes might play an active role in AD pathology and supported the notion that depression might be the "common cold" of psychiatry.

Keywords: Alzheimer's disease, depression, genome-wide association studies, genomics, transcriptomics

¹These authors contributed equally to this work.

*Correspondence to: Dr. Deng-Feng Zhang and Yong-Gang Yao, Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences and Yunnan Province, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China. Tel.: +86 871 68125441/86 871 6812

5440; E-mails: zhangdengfeng@mail.kiz.ac.cn (D.-F. Zhang), yaoyg@mail.kiz.ac.cn (Y.-G. Yao) and Dr. Chen Zhang, Department of Psychiatry, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai 200030, China. E-mail: zhangchen645@163.com.

INTRODUCTION

Alzheimer's disease (AD), the most common neurodegenerative disease, is becoming a serious global health issue [1, 2]. It is generally considered as a cognitive disorder, resulting from accumulation of extracellular amyloid- β ($A\beta$) plaques and intracellular neurofibrillary tangles, neuron loss, cerebral atrophy, and neuroinflammation [1, 2]. Almost all AD patients have neuropsychiatric symptoms during the course of the disease [3]. Depression is one of the most frequent psychiatric symptoms observed in the development of mild cognitive impairment (MCI) and AD [3, 4]. The symptoms of depression, technically referred to as major depressive disorder (MDD), are clinically heterogeneous, with a lifetime prevalence of approximately 16% [5, 6]. Depression and AD are common conditions in older age, with impaired cognition [5]. However, the relationship between depression and AD is elusive. It is unclear whether the comorbidity of MDD in AD is the cause or just a byproduct. Besides the comorbidity, it has been reported that a history of depression might increase risk of developing AD, and people with both MCI and a history of depression would progress to AD at a much higher rate [7–10]. Other studies also suggested that depression may be a prodromal symptom of AD [11–13]. These reports indicated a potentially causal role of depression in the course of AD.

Genetic factors account for $\sim 37\%$ of liability to MDD and depressive symptoms [14], while AD has a high heritability ($\sim 79\%$) [1, 15]. It is reasonable to hypothesize that genetic factors conferring risk of depression might also underlie the genetic basis of AD, with shared risk alleles for both MDD and AD. Indeed, our previous findings have shown that the MDD-risk allele *CFH* Y402H [16] also affects AD risk [17]. Moreover, the *MEF2C* gene, identified by the most recent and largest genome-wide association study (GWAS) for MDD [18], was a well-established AD risk gene in the largest GWAS for AD [19]. There might be more depression risk genes involved in AD, yet this kind of investigation was highly dependent on the reliability of recognized MDD risk genes. Numerous GWASs for MDD have been conducted in recent years, but no genome-wide significant risk locus has been identified due to clinical heterogeneity of the disorder [20–22]. Recently, two large scale GWASs for MDD had successfully identified several robust risk loci: the CONVERGE performed a GWAS for MDD in a phenotypically homogeneous sample of

female Han Chinese and identified two loci exceeding the genome-wide significance [23]; the largest meta-analysis combining datasets of 23andMe and MDD GWAS from the Psychiatric Genomics Consortium (PGC) identified 17 single nucleotide polymorphisms (SNPs) showing the genome-wide significance [18]. The MDD risk gene list of these two studies [18, 23] offered a start point for us to test a hypothesis of common genetic basis of MDD and AD. In this study, we analyzed 31 genes that were located in the reported MDD risk loci [18, 23] in AD at the genomic and transcriptomic levels to test this hypothesis.

MATERIALS AND METHODS

Assignment of MDD risk genes

Though there were several GWASs for MDD [20–22], genome-wide hits were only observed in the most recent 23andMe-based GWAS in populations of European ancestry [18] and the CONVERGE study in Han Chinese populations [23]. The 23andMe-based GWAS meta-analysis combined three data sets: Discovery phase, 75,607 cases and 231,747 controls; PGC MDD, 9,240 cases and 9,519 controls; Replication stage, 45,773 cases and 106,354 controls, resulted in the largest sample size in MDD genetic research [18]. A total of 17 independent genome-wide significant SNPs from 15 loci, covering 21 surrounding genes (defined by a distance of <300 kb window within the GWAS loci), were identified and validated in the original study [18] (Supplementary Table 1). The CONVERGE Consortium performed whole genome sequencing of 5,303 MDD cases and 5,337 controls at a lower coverage and identified two loci reaching a genome-wide significance, which could be validated in 3,231 cases and 3,186 controls [23]. The two loci contained 10 surrounding genes in ~ 500 kb of the signal as shown in the original study [23] (Supplementary Table 1). These 19 SNPs represented the most robust MDD risk loci ever identified, albeit it was yet unknown which gene is the causal gene targeted by the GWAS hits and more independent validations should be performed to confirm these risk loci. In this study, the 19 proxy MDD GWAS SNPs, as well as the 31 genes (including 21 for the 23andMe-based GWAS [18] and 10 for the CONVERGE [23]) within the 19 GWAS loci (as defined by the original studies [18, 23]) were subjected to comprehensive analyses in AD datasets (Fig. 1 and Supplementary Table 1).

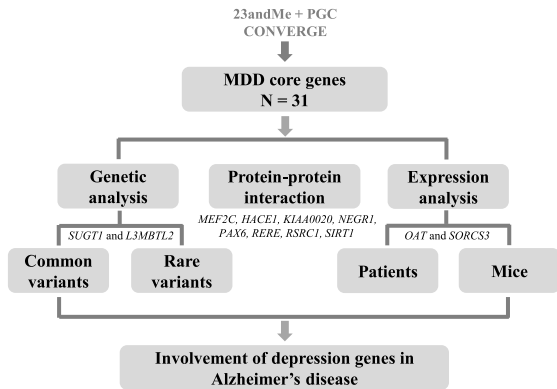


Fig. 1. Workflow and summary of the current study. We aimed to systematically investigate whether candidate genes located in the major depressive disorder (MDD)-associated loci were involved in Alzheimer's disease (AD). Genome-wide significant hits from the largest meta-analysis combining 23andMe and MDD GWAS from the Psychiatric Genomics Consortium (PGC) in populations of European ancestry [18], as well as the CONVERGE study in Chinese population [23], were used to identify potential MDD candidate genes. We initially checked whether common variants within these MDD-associated GWAS loci were associated with AD by using summary data of the International Genomics of Alzheimer's Project (IGAP) [19]. In addition to common variant association, we examined the entire coding regions of the 31 MDD candidate genes in 107 Han Chinese patients with early-onset and/or familial AD for rare variants. Transcriptomic profiling were performed using GEO data as processed in our recent report [33]. Protein-protein interaction network was constructed using online tool STRING (<https://string-db.org/>).

Association of common variants in MDD risk genes with AD using publicly available GWAS data

Involvement of MDD risk genes in AD was initially evaluated by using the publicly available International Genomics of Alzheimer's Project (IGAP) GWAS data [19], which focused on the association of common variants with AD susceptibility. The IGAP data contained 7,055,881 SNPs in 17,008 AD cases and 37,154 controls [19]. Summary statistics were downloaded from http://web.pasteur-lille.fr/en/recherche/u744/igap/igap_download.php. Single-site association results of all variants within 10 kb of the MDD risk genes were retrieved. The single-site association results were subjected to the gene-based test using the online tool Versatile Gene-based Association Study (VEGAS2 v02, <https://vegas2.qimrberghofer.edu.au/>) [24, 25]. Detailed methodology of the gene-based test was described in the original study [24]. Briefly, this test considers a full set of markers within a gene and the linkage disequilibrium information between markers by using simulations [24].

Identification of coding variants in MDD risk genes in AD patients using next generation sequencing

The coding and flanking regions (UTRs and exons) of the 31 MDD risk genes were analyzed in 107 Han Chinese patients with AD by using next generation sequencing, to identify potentially causal variants. We focused on these AD patients that had an early onset age <55 years old and/or a positive family history (46.7% females, age 64.6 ± 10.29 years, 37.4% APOE $\epsilon 4$ carriers), as these patients might be genetically informative compared with sporadic or late-onset AD. The targeted regions were captured by Nimblegene SeqCap Kit. Processed final libraries for each individual were pooled and sequenced on Illumina HiSeq2500 or 4000 (150-bp paired-end, Illumina, San Diego, CA, USA). Low quality raw reads were removed using Trimmomatic-0.32 [26]. Quality-filtered sequenced reads were aligned to the human genome reference assembly (build GRCh37/hg19) using Burrows-Wheeler Aligner (BWA) [27]. Picard Tools (<http://broadinstitute.github.io/picard/>) were used to flag duplicated reads. SNP call was performed through the canonical pipeline recommended by the Best Practice Variant Detection with the GATK (Genome Analysis Toolkit, <https://www.broadinstitute.org/gatk/guide/best-practices>) [28]. Variant Quality Score Recalibration (VQSR) from the GATK package was used to filter spurious variants due to sequencing errors and/or mapping artifacts. The ANNOVAR was used to annotate variants, which were assigned into different functional categories according to their locations and expected effects on the encoded gene products [29]. Missense, nonsense, frameshift, and splice site variants were defined as functional variants. We used the exome data of 160 Chinese individuals (40.6% females, age 52.6 ± 16.5 years) showing no signs of memory loss and no familial history of neurodegenerative disorders [30] as a control sample. To increase the sample size of controls, the whole genome data of Han Chinese in Beijing (CHB, N=103) and Southern Han Chinese (CHS, N=105) from the 1000 Genome Project phase 3 [31] were pooled with the exome data of 160 Chinese individuals [30] as the population control (N=368). Allele frequency data of 4327 East Asians from the ExAC (Exome Aggregation Consortium, <http://exac.broadinstitute.org/>) [32] was used as another control to validate the results in Chinese.

Expression changes of the MDD risk genes in AD brain samples and mouse models using our previously processed data

In addition to genomic analysis, we analyzed the mRNA expression profiles of the MDD risk genes in four AD-affected brain regions (frontal cortex tissues of 413 patient and 284 controls; temporal cortex tissues of 158 patients and 174 controls; hippocampus tissues of 74 patients and 65 controls; entorhinal cortex tissues of 39 patients and 39 controls), following the approach described in our recent study [33]. In brief, raw microarray data for each brain region was retrieved from the NCBI Gene Expression Omnibus database (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) and renormalized as one combined dataset with an enlarged sample size. Details of data sources and processing were described in our previous study (Ref. [33] and references therein), and the expression data collection and renormalized expression profiles can be accessed through our database <http://www.alzdata.org> [33].

In addition to the differential expression analysis in brain tissues of AD patients, we investigated the expression alterations of these MDD risk genes in brain tissues of mouse AD models [34]. In brief, the transgenic mouse models with human mutant genes responsible for familial AD, which showed the presence of AD pathological features, were used for the genome-wide microarrays [34]. Hippocampus and cortex tissues were tested by using the MouseRef8 v2 (Illumina) microarray platform. Microarray data was processed and shared by the Mouse Dementia Network, available at Mouseac (<http://www.mouseac.org>). More details about this dataset were described in the original study [34]. Correlation between mRNA expression levels of genes of interest and the quantified level of pathology (A β plaques and tau burden) was measured based on the processed data, using the Pearson correlation test, as had been described in our recent study [33].

Protein-protein interaction network of the MDD risk genes with AD core genes

To evaluate the involvement of these MDD risk genes in the molecular network of AD, we performed the protein-protein interaction (PPI) analysis using proteins of the 31 MDD risk genes, together with 43 known AD genes (identified by GWAS and linkage studies [35]) using the online tool STRING (<https://string-db.org/>).

Convergent functional genomics (CFG) analysis

To cross-validate the involvement of target genes in AD at different levels, we used the convergent functional genomic (CFG) approach [36, 37] that integrates multiple lines of AD-related evidence, as we had done recently [33]. A gene was defined as AD-related if it: 1) showed a significant association with AD at the gene-based test using the IGAP data [19]; 2) had rare variants that were significantly associated with AD in 107 Han Chinese patients with AD; 3) was differentially expressed in any brain tissue of AD patients according to the AlzData.org [33]; 4) was correlated with AD pathology in AD mice based on the Mouseac (<http://www.mouseac.org>) [34]; or 5) was involved in the PPI network that was formed by the AD core genes. For each line of evidence, one point was assigned if the above fact was observed; otherwise zero point. This integration system led to CFG scores ranging from 0 to 5 points.

RESULTS

Associations of common and rare variants of the MDD risk genes with AD

Among the 19 genome-wide significant MDD-related SNPs (Supplementary Table 1), two SNPs (*L3MBTL2* rs2179744 and intergenic rs4543289) showed positive associations with AD in IGAP dataset [19]. It is reasonable that other SNPs within the 31 MDD risk genes, rather than the 19 GWAS hits, might be associated with AD. We searched for other SNPs within the 31 MDD risk genes and found that 21 out of the 31 genes had SNPs showing nominally significant associations ($p < 0.05$) with AD (Table 1). Among them, three genes, *SUGT1* (Gene-based $p = 1.97 \times 10^{-3}$), *L3MBTL2* (Gene-based $p = 9.60 \times 10^{-3}$), and the known AD gene *MEF2C* [19] (Gene-based $p = 1.19 \times 10^{-2}$), were significantly associated with AD at the gene level (Table 1).

Besides associations of common variants with AD, we analyzed the entire coding region of the 31 genes in 107 Han Chinese patients with early-onset and/or familial AD using targeted sequencing, to investigate whether there was an enrichment of rare coding variants in these MDD risk genes in AD patients (Supplementary Table 2). Six genes, *TMEM161B*, *L3MBTL2*, *SLC6A15*, *KIAA0020*, *LHPP*, and *MYPN*, had a missense variant with nominally significance in AD (Table 2). Two of them (*L3MBTL2* and *SLC6A15*)

Table 1
List of MDD risk genes and association analysis of common variants in the MDD GWAS loci with AD

Chromosome	Gene	No. of SNPs	Start	Stop	Gene-based <i>p</i> -value	Top SNP	Top SNP <i>p</i> -value
13	<i>SUGT1</i>	83	53226830	53262433	0.00197	rs4885767	0.00079
22	<i>L3MBTL2</i>	80	41601312	41627275	0.00960	rs9611520	0.00014
5	<i>MEF2C</i>	288	88014057	88199922	0.01190	rs3850651	0.00028
15	<i>MEIS2</i>	392	37183221	37393500	0.05030	rs4924109	0.00348
9	<i>KIAA0020</i>	80	2804154	2844130	0.05889	rs10217194	0.00709
1	<i>RERE</i>	583	8412463	8877699	0.08392	rs112053331	0.00090
1	<i>NEGR1</i>	739	71868624	72748277	0.10689	rs522395	0.00011
9	<i>RFX3</i>	574	3218296	3526001	0.16583	rs62526377	0.00487
10	<i>FAM53B</i>	227	126307862	126432930	0.22478	rs11245345	0.02208
10	<i>SORCS3</i>	675	106400858	107024993	0.29970	rs112898269	0.00344
6	<i>LIN28B</i>	128	105404922	105531207	0.37962	rs10457125	0.01275
13	<i>OLFM4</i>	72	53602875	53626196	0.39960	rs3803259	0.03261
9	<i>PAX5</i>	671	36833271	37034476	0.40160	rs138898773	0.00800
10	<i>CTNNA3</i>	971	67672275	69455949	0.41658	rs116950805	0.00372
5	<i>TMEM161B</i>	106	87485449	87564696	0.43157	rs115873146	0.00750
12	<i>SLC6A15</i>	60	85253266	85306608	0.47453	rs79265864	0.12470
3	<i>MLF1</i>	106	158288952	158324249	0.51648	rs7617983	0.11220
10	<i>METTL10</i>	50	126447405	126480439	0.58042	rs11245363	0.07227
10	<i>LHPP</i>	743	126150340	126302710	0.60140	rs60015709	0.01357
15	<i>TMCO5A</i>	31	38227456	38243623	0.69431	rs112187315	0.17920
6	<i>HACE1</i>	264	105175967	105307794	0.77323	rs79147057	0.02067
10	<i>MYPN</i>	387	69865873	69971773	0.79321	rs76451279	0.00867
10	<i>HERC4</i>	244	69681655	69835103	0.80120	rs17454621	0.08052
10	<i>SIRT1</i>	79	69644426	69678147	0.82118	rs10997864	0.04805
10	<i>FAM175B</i>	44	126490353	126525239	0.82717	rs2277263	0.12360
10	<i>DNAJC12</i>	56	69556426	69597937	0.84615	rs41299238	0.13750
11	<i>PAX6</i>	41	31806339	31839509	0.89710	rs113859447	0.22190
3	<i>RSRC1</i>	448	157827840	158262624	0.89710	rs144747720	0.00288
10	<i>OAT</i>	66	126085871	126107545	0.90909	rs140350519	0.08203
2	<i>VRK2</i>	434	58134785	58387055	0.98402	rs140500963	0.01353

Note: No. of SNPs, number of SNPs included in the gene-based test. Start/Stop, range of target gene (position based on human genome reference assembly (build GRCh37/hg19)). The gene-based *p*-value was calculated by the online tool Versatile Gene-based Association Study (VEGAS2 v02, <https://vegas2.qimrberghofer.edu.au/>) [24, 25]. Top SNP and Top SNP *p*-value were retrieved from the summary statistics of the International Genomics of Alzheimer's Project (IGAP) [19]. Significant *p*-values (<0.05) were marked in bold. One (*MIR759*) out of the 31 genes had no data available.

showed consistent associations when compared with both our in-house control [30] and the ExAC reference control [32]. Notably, *L3MBTL2*, which was also associated with AD in the common variant analysis, had a rare missense variant rs3804097 (p. I7V) that was enriched in AD (MAF=4.7%) compared with our in-house controls (MAF=1.6%, $p = 1.63 \times 10^{-2}$, OR = 2.99) and the ExAC reference control (MAF=2.7%, $p = 8.60 \times 10^{-2}$, OR = 1.78).

Differential gene expression of MDD genes in brain tissues of AD patients and mice

To further investigate the involvement of MDD risk genes in AD development, we analyzed the mRNA expression profiles of the MDD genes in four brain tissues (hippocampus, entorhinal cortex, frontal cortex, and temporal cortex) of AD patients

and controls ([33] and references therein). Eighteen genes showed a nominally significant ($p < 0.05$) differential expression in AD brain tissues compared with that of controls (Table 3). In particular, five genes (*MEF2C*, *SORCS3*, *OAT*, *DNAJC12*, and *SIRT1*) were consistently altered in all four brain regions, with *SIRT1* being upregulated and the other four genes being downregulated in AD brain tissues (Table 3). We then checked the differentially expressed pattern of these 18 genes in hippocampus tissues of AD mouse models [34]. Seven genes (*FAM53B*, *HACE1*, *NEGR1*, *SORCS3*, *OAT*, *DNAJC12*, and *SLC6A15*) were significantly correlated with AD pathology burden (number of amyloid plaques or tau tangles) (Table 3). Of note, mRNA expression levels of *SUGT1* and *L3MBTL2*, which showed a genetic association with AD, had no expression alteration in AD brain tissues.

Table 2
Significant coding variants in MDD risk genes in 107 AD patients

Position	SNP	Allele	Gene	Mutation	AC/AN		<i>p</i> -value	OR	AC/AN ExAC	<i>p</i> -value	OR
					AD	CN					
chr5:87493525	rs75451667	C/T	<i>TMEM161B</i>	p. V383I [#]	41/212	101/736	0.049	1.51	1352/7864	0.407	1.15
chr22:41601385	rs3804097	A/G	<i>L3MBTL2</i>	p. I7V	10/212	12/736	0.016	2.99	213/7864	0.086	1.78
chr12:85257227	rs3782369	A/C	<i>SLC6A15</i>	p. I603M[#]	3/212	0/366	0.049	12.25	26/7858	0.040	4.32
chr9:2828765	rs2173904	C/G	<i>KIAA0020</i>	p. R289P	129/212	380/736	0.019	1.45	4757/7852	1.000	1.01
chr10:126172863	rs6597801	A/G	<i>LHPP</i>	p.Q94R	205/212	682/736	0.038	2.32	7429/7866	0.219	1.72
chr10:69934258	rs3814182	C/G	<i>MYPN</i>	p.S803R [#]	87/212	251/736	0.073	1.34	2837/7816	0.169	1.22

Note: Position was based on human genome reference assembly (build GRCh37/hg19); Allele, reference allele/alternative allele; AC, allele count; AN, allele number; OR, odds ratio; ExAC, Exome Aggregation Consortium (<http://exac.broadinstitute.org/>) [32]; AD, 107 subjects with Alzheimer's disease; CN, 368 subjects combining 160 in-house non-dementia individuals [30] and 208 Chinese from the 1000 Genome project [31]. Significant *p*-values (<0.05) were marked in bold. The ANNOVAR was used to annotate variants, which were assigned into different functional categories according to their locations and expected effect on the encoded gene products [29]. Missense, nonsense, frameshift, and splice site variants were defined as functional variants. [#]Coding variants predicted to be damaging according to ANNOVAR annotation. All rare coding variants were listed in Supplementary Table 2.

Table 3
mRNA expression changes of MDD risk genes in brain tissues of AD patients and AD mouse models

Gene	Human AD brain						Mouse AD model					
	Entorhinal Cortex		Hippocampus		Temporal Cortex		Frontal Cortex		A β		Tau	
	log2FC	p-value	log2FC	p-value	log2FC	p-value	log2FC	p-value	r	p-value	r	p-value
<i>CTNNA3</i>	0.17	0.422	-0.02	0.877	-0.05	0.847	-0.13	0.209	0.419463	0.0063336	0.077534	0.7835861
<i>DNAJC12</i>	-0.49	0.003	-0.38	0.001	-0.68	2.1E-05	-0.16	0.015	-0.35167	0.0241533	-0.37774	0.16509865
<i>FAM175B</i>	-0.13	0.095	-0.08	0.270	-0.17	0.088	-0.07	0.062	NA	NA	NA	NA
<i>FAM53B</i>	0.30	0.018	0.13	0.302	0.07	0.681	0.17	0.022	0.594041	4.23E-05	0.060471	0.83048573
<i>HACE1</i>	-0.41	0.002	-0.12	0.108	-0.28	0.097	NA	NA	-0.33535	0.0320821	-0.4933	0.06167796
<i>HERC4</i>	0.06	0.550	0.07	0.427	-0.15	0.234	0.13	0.055	-0.5922	4.53E-05	0.074053	0.79310244
<i>KIAA0020</i>	0.14	0.131	0.01	0.921	-0.05	0.653	0.10	0.034	NA	NA	NA	NA
<i>L3MBTL2</i>	-0.06	0.431	0.06	0.449	-0.02	0.833	NA	NA	-0.22469	0.157844013	-0.45847	0.08565248
<i>LHPP</i>	-0.05	0.781	0.04	0.723	0.33	0.025	0.23	0.009	NA	NA	NA	NA
<i>LIN28B</i>	-0.33	0.011	-0.43	0.008	0.08	0.688	NA	NA	NA	NA	NA	NA
<i>MEF2C</i>	-0.72	0.002	-0.56	1.6E-05	-0.91	6.3E-06	-0.25	0.010	-0.18712	0.241395885	-0.19811	0.47908788
<i>MEIS2</i>	0.11	0.271	0.20	0.056	0.34	0.028	0.06	0.324	-0.27327	0.083855988	-0.87303	2.15E-05
<i>METTL10</i>	0.01	0.909	-0.10	0.127	-0.12	0.381	-0.19	0.009	NA	NA	NA	NA
<i>MLF1</i>	0.13	0.251	0.11	0.252	-0.15	0.241	-0.09	0.150	-0.03942	0.806703421	0.115341	0.6822973
<i>MYPN</i>	0.29	0.090	-0.15	0.257	0.07	0.777	NA	NA	NA	NA	NA	NA
<i>NEGR1</i>	-0.51	0.002	-0.18	0.099	-0.49	0.001	NA	NA	-0.406	0.0084445	-0.51538	0.0492703
<i>OAT</i>	-0.36	0.015	-0.22	0.005	-0.60	0.001	-0.25	0.006	0.727227	7.23E-08	0.53886	0.0382033
<i>OLFM4</i>	-0.67	0.001	0.09	0.611	-0.29	0.146	-0.27	0.075	-0.04742	0.76845086	0.215175	0.44120421
<i>PAX5</i>	-0.09	0.285	-0.04	0.699	0.11	0.561	-0.05	0.473	NA	NA	NA	NA
<i>PAX6</i>	0.16	0.192	0.33	1.6E-04	0.68	6.5E-07	0.21	0.008	NA	NA	NA	NA
<i>RERE</i>	0.03	0.817	0.14	0.116	-0.21	0.228	0.11	0.037	NA	NA	NA	NA
<i>RFX3</i>	0.25	0.002	0.21	0.010	-0.20	0.041	0.15	0.019	0.212547	0.182144629	-0.00283	0.99201833
<i>RSRC1</i>	-0.07	0.522	-0.10	0.298	-0.46	0.003	-0.12	0.196	NA	NA	NA	NA
<i>SIRT1</i>	0.36	0.014	0.24	0.010	0.21	0.035	0.21	0.005	0.058888	0.714570112	-0.3476	0.20425918
<i>SLC6A15</i>	-0.22	0.130	-0.10	0.288	-0.61	1.9E-04	-0.17	0.087	-0.28047	0.075699482	-0.54521	0.0355581
<i>SORCS3</i>	-0.38	0.038	-0.25	0.031	-0.77	1.2E-04	-0.31	1.9E-05	-0.76766	4.73E-09	-0.62775	0.0122258
<i>SUGT1</i>	0.13	0.129	-0.01	0.883	0.09	0.408	NA	NA	-0.24314	0.125564623	-0.09847	0.72699952
<i>TMCO5A</i>	-0.07	0.636	-0.03	0.833	0.01	0.969	NA	NA	NA	NA	NA	NA
<i>TMEM161B</i>	0.16	0.059	0.04	0.585	-0.12	0.274	NA	NA	NA	NA	NA	NA
<i>VRK2</i>	0.28	0.063	0.01	0.903	0.07	0.687	0.05	0.626	0.498012	0.0009218	0.556909	0.0310465

Note: log2FC, log2 of fold change of gene expression level compared with that of control; r, Pearson correlation; Expression microarray data of the MDD genes in four brain tissues (hippocampus, entorhinal cortex, frontal cortex, and temporal cortex) in AD patients were retrieved from the NCBI Gene Expression Omnibus database (GEO, <http://www.ncbi.nlm.nih.gov/geo/>). Details were described in our previous paper [33] and the expression data collection and renormalized expression profiles can be accessed through our database <http://www.alzdata.org> [33]. Expression data of AD mouse models were taken from the mouseac.org [34]. The number of amyloid- β (A β) plaques and the tau (Tau) burden were quantified and the correlation between mRNA expression levels of genes of interest, and the quantified level of pathology was measured based on the processed data, using the *Pearson* correlation test. Significant *p*-values (<0.05) were marked in bold. One (*MIR759*) out of the 31 genes was not available.

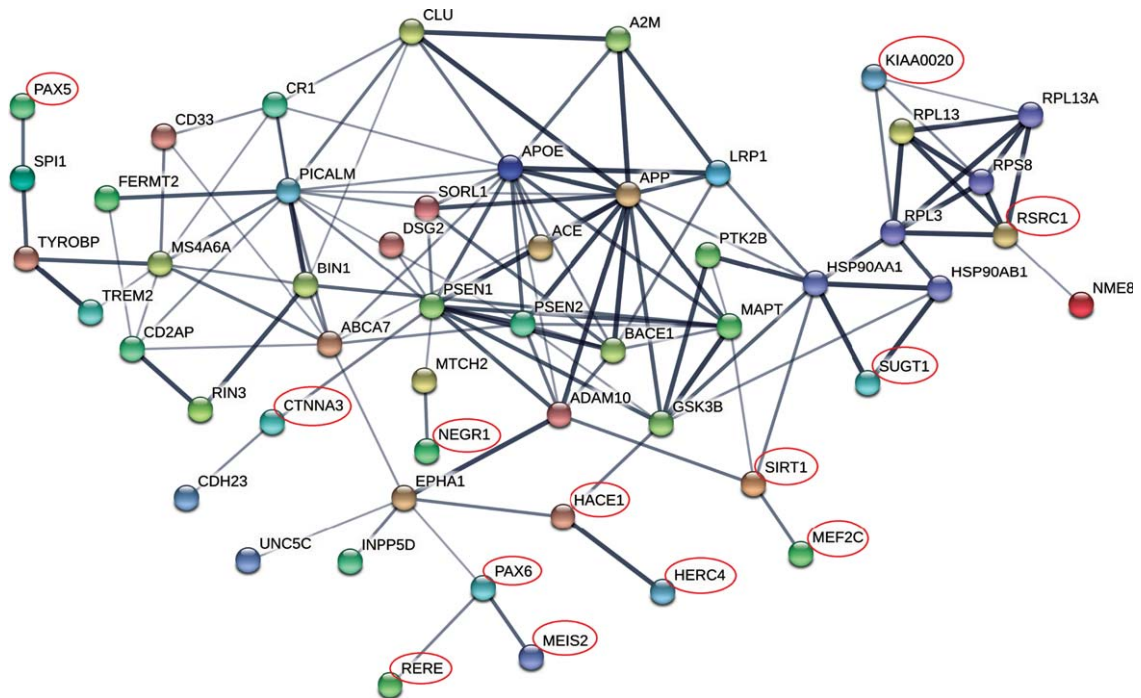


Fig. 2. Protein-protein interaction network of MDD risk genes and AD core genes. Protein-protein interaction (PPI) network was constructed using online tool STRING (<https://string-db.org/>). AD GWAS genes from the IGAP list [19] and reported causal genes as reviewed by Guerreiro et al. [35], were defined as core AD genes. PPI interactions were defined using resources from text-mining, experiments, databases, co-expression, and co-occurrence, as the default setting by the software. Line thickness indicates the strength of confidence. MDD risk genes were marked by red circle.

Protein-protein interaction of MDD genes with core AD genes

The AD GWAS genes from the IGAP study [19] and reported causal genes, as reviewed by Guerreiro et al. [35], were defined as core AD genes. The 31 MDD risk genes and the 43 core AD genes formed a network with significantly more interactions than expected (PPI enrichment $p = 2.0 \times 10^{-13}$) (Fig. 2). In total, 13 genes (PAX5, CTNNA3, KIAA0020, RSRC1, NEGR1, SIRT1, HACE1, HERC4, PAX6, RERE, MEIS2, SUGT1 and MEF2C), including the two genetically risk genes SUGT1 and MEF2C, were integrated into the AD network, suggesting an involvement of these genes in AD pathology (Fig. 2). In particular, the SIRT1 gene interacted with MAPT, ADAM10, and MEF2C, whereas HACE1 and PAX6 interacted with EPHA1 and GSK3B and were linked with RERE, MEIS2, and HERC4. The PAX5 interacted with the most recently recognized AD gene SPI1 [38] (Fig. 2). Notably, eight (MEF2C, HACE1, KIAA0020, NEGR1, PAX6, RERE, RSRC1, and SIRT1) of the 13 genes recognized from the PPI network were differentially expressed in AD brain

(Table 3), suggesting functional roles of these genes in AD pathology.

Cross-validated targets using the CFG approach

Though we observed an apparent involvement of several MDD risk genes in AD at different levels, targets cross-validated by convergent lines of evidence might be more reliable. In our CFG analysis (Supplementary Table 3), we observed four genes (HACE1, MEF2C, NEGR1, and SLC6A15) being weighted with a score of 3, and 13 genes (CTNNA3, DNAJC12, FAM53B, HERC4, KIAA0020, L3MBTL2, LHPP, MEIS2, OAT, PAX6, SIRT1, SORCS3, and SUGT1) with a score of 2. These 17 genes were cross-validated by at least two independent lines of evidence, further supporting their active involvement in AD.

DISCUSSION

People with AD usually experience both cognitive and psychiatric symptoms [3, 4, 8]. Behavioral or psychiatric symptoms were thought to be more

challenging effects of the disease [3, 11, 13, 39]. Understanding the relationship and cause of psychiatric symptoms in AD might be essential to improve the quality of life for people suffered. Depression, with high lifetime prevalence, is a common condition in older age. It has been recognized to be one of the most frequent psychiatric comorbidities in AD [3, 8, 40]. It is yet unclear whether depression in AD is causal or just a byproduct. Previous studies indicated that depression may be a prodromal feature and predict risk of AD [7, 11–13, 39, 40]. Therefore, genetic factors contributing to depression might affect risk to AD onset and development. We tested the hypothesis that MDD risk alleles might also contribute to AD risk in this study.

In fact, a previous study by Gibson et al. [41] assessed the genetic correlation of the MDD GWAS findings with neurodegenerative disease using linkage disequilibrium score regression. They observed no overlapping polygenic architecture between lifetime MDD and AD at the whole-genome level and suggested that the genetic overlap might be restricted to some specific genes [41]. Our previous findings showed that the MDD-risk allele *CFH* p.Y402H conferred AD risk [16, 17]. The *MEF2C* gene was identified as a risk gene of MDD by the most reliable MDD GWAS [18]; this gene was a previously established AD-risk gene [19, 35]. It is thus valuable to check systematically whether the top MDD risk loci were involved in AD. Taking advantage of the MDD risk gene list from two recently MDD GWAS [18, 23] with a large sample size, we analyzed these MDD GWAS hits in AD at the genomic variation, mRNA expression and PPI levels, to further explore the genetic correlation between depression and AD.

We initially checked whether the MDD GWAS SNPs were associated with AD using publicly available GWAS data [19]. Among the 19 MDD GWAS hits, only one SNP (*L3MBTL2* rs2179744) within the genic region showed a positive association with AD; whereas 21 out of the 31 genes had nominally significant non-GWAS hit SNPs. Note that an intergenic MDD GWAS SNP rs4543289 also showed a positive association with AD, and further studies are needed to identify the underlying genes. Three genes (*SUGT1*, *L3MBTL2*, and *MEF2C*) were significantly associated with AD at the gene-based level. *MEF2C* was a recognized AD risk gene [19, 35] and was highly expressed in brain tissues according to the search at <http://www.alzdata.org> [33]. It is a transcription factor that functions in immune response and regulates synapse numbers and function [42]; both biological

processes would contribute to the course of MDD and AD. The *L3MBTL2* gene showed robust evidence for genetic association with AD in both common and rare variant analyses. Intriguingly, a recent large genome-wide analyses for personality traits also identified an association between *L3MBTL2* SNPs and neuroticism [43]. The Schizophrenia Working Group of the PGC also identified association of *L3MBTL2* SNPs with schizophrenia in the largest GWAS of schizophrenia [44]. These observations strongly suggested that *L3MBTL2* might be a genetic risk factor common for neuropsychiatric disorders via its functions in neural development. However, data from the International Mouse Phenotyping Consortium (IMPC, <http://www.mousephenotype.org/>) showed that mice knocking out *L3MBTL2* have decreased lumbar vertebrae number and increased sacral vertebrae number, but no abnormal behavior/neurological or nerve system phenotypes. Though it is highly expressed in neuron (cf. single cell expression data at our AlzData.org database [33]), its function in neural system remained unclear. Further functional assay is needed to characterize the underlying mechanism of *L3MBTL2* in neuropsychiatric disorders.

We found that the *SUGT1* gene was involved in the AD network and was genetically associated with AD. Previous report has shown that *SUGT1* is highly expressed in temporal cortex and significantly decreased in the temporal, angular, and posterior cingulate cortex in AD brains as compared to aged controls [45]. Its decrease might be a result of neuron degeneration in AD, yet it is still unclear how it contributes to MDD and AD. Collectively, our results showed that *SUGT1* and *L3MBTL2*, together with previously identified *CFH* [16, 17] and *MEF2C* [18, 19], might play a role in the neural system and serve as reliable common targets in MDD and AD.

Though only limited number of the MDD risk genes was found to be associated with AD genetic risk, 18 MDD risk genes were differentially expressed in AD brain tissues. In particularly, five genes (*MEF2C*, *SORCS3*, *OAT*, *DNAJC12*, and *SIRT1*) were consistently altered in all four brain regions of patients according to the AlzData [33]. What's more, mRNA expression levels of *SORCS3* and *OAT* were also significantly correlated with pathology burden in hippocampus of AD mice. These two genes might be highly involved in AD pathology. The *OAT* gene encodes the mitochondrial enzyme ornithine aminotransferase, which is a key enzyme in the synthesis of major excitatory and inhibitory neurotransmitters glutamate and GABA

[46]. Its activities in brains of Huntington's disease were found to be reduced compared to age-matched control brains [46]. However, no such changes were observed in AD or schizophrenia [46]. Though there was no doubt that *OAT* plays an essential role in neuron function, its involvement in MDD and AD needs further investigation. *SORCS3* belongs to the sorting and signaling receptor family central in control of neuronal viability and function [47]. It is highly expressed in hippocampus and localizes to the postsynaptic density [47]. *SORCS3*-deficient mice (<http://www.informatics.jax.org/marker/phenotypes/MGI:1913923>) had deficits in behavioral activities and spatial learning [47]. Intriguingly, a previous study reported that SNPs of *SORCS1*, a family member of *SORCS3*, were associated with AD susceptibility [48]. Overexpression of *SORCS1* reduced γ -secretase activity and A β level, while suppression of *SORCS1* increased γ -secretase processing of A β PP and the level of A β [48]. Consistently, another important sorting related receptor *SORL1* was an established AD top risk gene [19]. These observations suggested that *SORCS3*, together with other sorting related receptors, might be involved in both MDD and AD.

Besides the genomic and transcriptomic evidence, several MDD risk genes and AD core genes formed a network with high interactions than expected. Eight genes (*MEF2C*, *HACE1*, *KIAA0020*, *NEGR1*, *PAX6*, *RERE*, *RSRC1*, and *SIRT1*) in the PPI network were differentially expressed in AD brain (Fig. 2), suggesting potentially functional involvement of these genes in AD pathology. The genetically risk gene *SUGT1* was also integrated into the AD network. Notably, the key MDD gene *SIRT1* [23] interacted with AD core genes *MAPT* and *ADAM10*. In addition, though no single gene showed convergent lines of supporting evidence, our CFG analysis found that four genes (*HACE1*, *MEF2C*, *NEGR1*, and *SLC6A15*) (Supplementary Table 3) were cross-validated by data at different levels, including the well-known hit *MEF2C*. The other three genes (*HACE1*, *NEGR1*, and *SLC6A15*) were all highly expressed in neural system and might be core genes shared between MDD and AD.

By focusing on MDD risk genes locating in the susceptibility loci identified by GWAS, we found that: 1) two genes (*SUGT1* and *L3MBTL2*) showing nominally significant associations with AD; 2) *SORCS3* and *OAT* were significantly correlated with AD pathology in both human AD brain tissues and AD mouse models; and 3) three new targets *HACE1*,

NEGR1, and *SLC6A15* showing convergent lines of evidence to be involved in AD. However, there were several limitations in this study. First, the distilling of causal MDD risk genes is very challenging and the current list of candidate genes is incomplete. Due to the huge heterogeneity of depression, it is difficult to map the most reliable risk genes as shown by most of previous MDD GWASs [20–22]. Age at onset of depression was suggested to be bimodal at around 20 years old and 70–80 years old [49]. It is reasonable to speculate that there might be different genetic bases for early-life and late-life depressions. Genetic correlation between AD and depression might be dependent on the onset age of depression, as AD mainly occurred in the elderly. However, there were some observations suggesting potentially similar biological processes underlying early-onset and late-onset depressions: 1) most of the MDD-risk genes shared by AD showed a stable expression level in prefrontal cortex from early adulthood (20 years old) to late-life (80 years old) (Supplementary Figure 1); 2) both early-life [10] and late-life depression [50] were indicated to increase risk of dementia. As the MDD-risk genes from the 23andMe-based GWAS meta-analysis were identified in subjects of all age groups [18], whereas the CONVERGE findings were based on subjects younger than 60 years old [23], the current MDD risk gene list might not be specific to early-onset and/or late-onset depressions, and refined MDD gene list are needed to understand the shared genetic basis between late-onset depression and AD. Second, the sample size to identify rare coding variants in these MDD risk genes was relatively small, and more samples will offer enhanced statistical power. Third, there was no well-designed functional assay to characterize these genes in the context of AD and depression in this study. In addition, considering the polygenic nature of depression [18, 23], it will be promising to test the correlation between polygenic risk score of MDD with depression and cognitive symptoms in AD patients in the future.

In short, our results were generally consistent with the report by Gibson et al. [41] that the comorbidity of MDD and AD might not be largely driven by genetic factors. However, our data showed that some MDD risk genes located in the GWAS regions might be involved in AD pathogenesis. It is currently unclear whether depression is a consequence of cognitive decline and neuronal loss affected by AD, or depression at early adulthood contributes to neuronal damage that leads to cognitive decline. Considering the report that *APOE* $\epsilon 4$, the major AD

risk allele [19], has no contribution to depression [51], it is more possible that depression might be the upstream of AD or a byproduct event. Further research is needed to understand other non-genetic causes (e.g., common environmental factors) and underlying biological mechanisms that account for the complex relationship between neuropsychiatric disorders. This knowledge might help to reduce the psychiatric burden that AD patients and their caregivers suffered.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant nos. 31730037 to YGY, 81560230 to HYJ, 81500940 to HZ, and 31601039 to DFZ), the Bureau of Frontier Sciences and Education, Chinese Academy of Sciences (CAS, Grant no. QYZDJ-SSW-SMC005 to YGY), and the West Light Foundation of the CAS (to DFZ). We thank the International Genomics of Alzheimer's Project (IGAP) for providing summary results data for these analyses. The investigators within IGAP contributed to the design and implementation of IGAP and/or provided data but did not participate in analysis or writing of this report. IGAP was made possible by the generous participation of the control subjects, the patients, and their families. The i-Select chips were funded by the French National Foundation on Alzheimer's disease and related disorders. EADI was supported by the LABEX (laboratory of excellence program investment for the future) DISTALZ grant, Inserm, Institut Pasteur de Lille, Université de Lille 2 and the Lille University Hospital. GERAD was supported by the Medical Research Council (Grant no. 503480), Alzheimer's Research UK (Grant no. 503176), the Wellcome Trust (Grant no. 082604/2/07/Z) and German Federal Ministry of Education and Research (BMBF): Competence Network Dementia (CND) Grant no. 01GI0102, 01GI0711, 01GI0420. CHARGE was partly supported by the NIH/NIA grant R01 AG033193 and the NIA AG081220 and AGES contract N01-AG-12100, the NHLBI grant R01 HL105756, the Icelandic Heart Association, and the Erasmus Medical Center and Erasmus University. ADGC was supported by the NIH/NIA grants: U01 AG032984, U24 AG021886, U01 AG016976, and the Alzheimer's Association grant ADGC-10-196728.

Authors' disclosures available online (<https://www.j-alz.com/manuscript-disclosures/18-0276r1>).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: <http://dx.doi.org/10.3233/JAD-180276>.

REFERENCES

- [1] Alzheimer's Association (2016) 2016 Alzheimer's disease facts and figures. *Alzheimers Dement* **12**, 459-509.
- [2] Alzheimer's Association (2015) 2015 Alzheimer's disease facts and figures. *Alzheimers Dement* **11**, 332-384.
- [3] Lyketsos CG, Lopez O, Jones B, Fitzpatrick AL, Breitner J, DeKosky S (2002) Prevalence of neuropsychiatric symptoms in dementia and mild cognitive impairment: Results from the cardiovascular health study. *JAMA* **288**, 1475-1483.
- [4] Lee HB, Lyketsos CG (2003) Depression in Alzheimer's disease: Heterogeneity and related issues. *Biol Psychiatry* **54**, 353-362.
- [5] Gureje O, Kola L, Afolabi E (2007) Epidemiology of major depressive disorder in elderly Nigerians in the Ibadan Study of Ageing: A community-based survey. *Lancet* **370**, 957-964.
- [6] Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS (2003) The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *JAMA* **289**, 3095-3105.
- [7] Taragano FE, Allegri RF, Krupitzki H, Sarasola DR, Serrano CM, Lon L, Lyketsos CG (2009) Mild behavioral impairment and risk of dementia: A prospective cohort study of 358 patients. *J Clin Psychiatry* **70**, 584-592.
- [8] Spira AP, Rebok GW, Stone KL, Kramer JH, Yaffe K (2012) Depressive symptoms in oldest-old women: Risk of mild cognitive impairment and dementia. *Am J Geriatr Psychiatry* **20**, 1006-1015.
- [9] Janzing JG, Naarding P, Eling P (2005) Depressive symptoms predict slow cognitive decline in mild dementia. *Dement Geriatr Cogn Disord* **20**, 77-81.
- [10] Speck CE, Kukull WA, Brenner DE, Bowen JD, McCormick WC, Teri L, Pfanschmidt ML, Thompson JD, Larson EB (1995) History of depression as a risk factor for Alzheimer's disease. *Epidemiology* **6**, 366-369.
- [11] Wilson RS, Barnes LL, Mendes de Leon CF, Aggarwal NT, Schneider JS, Bach J, Pilat J, Beckett LA, Arnold SE, Evans DA, Bennett DA (2002) Depressive symptoms, cognitive decline, and risk of AD in older persons. *Neurology* **59**, 364-370.
- [12] Gatz JL, Tyas SL, St John P, Montgomery P (2005) Do depressive symptoms predict Alzheimer's disease and dementia? *J Gerontol A Biol Sci Med Sci* **60**, 744-747.
- [13] Mirza SS, Wolters FJ, Swanson SA, Koudstaal PJ, Hofman A, Tiemeier H, Ikram MA (2016) 10-year trajectories of depressive symptoms and risk of dementia: A population-based study. *Lancet Psychiatry* **3**, 628-635.
- [14] Sullivan PF, Neale MC, Kendler KS (2000) Genetic epidemiology of major depression: Review and meta-analysis. *Am J Psychiatry* **157**, 1552-1562.
- [15] Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, Fiske A, Pedersen NL (2006) Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry* **63**, 168-174.

- [16] Zhang C, Zhang DF, Wu ZG, Peng DH, Chen J, Ni J, Tang W, Xu L, Yao YG, Fang YR (2016) Complement factor H and susceptibility to major depressive disorder in Han Chinese. *Br J Psychiatry* **208**, 446-452.
- [17] Zhang DF, Li J, Wu H, Cui Y, Bi R, Zhou HJ, Wang HZ, Zhang C, Wang D, Kong QP, Li T, Fang Y, Jiang T, Yao YG (2016) CFH variants affect structural and functional brain changes and genetic risk of Alzheimer's disease. *Neuropsychopharmacology* **41**, 1034-1045.
- [18] Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR, Tung JY, Hinds DA, Perlis RH, Winslow AR (2016) Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat Genet* **48**, 1031-1036.
- [19] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thornton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin CF, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau MT, Choi SH, Reitz C, Pasquier F, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston JA, Evans D, Lovestone S, Letenneur L, Moron FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fievet N, Huettelmann MW, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGuinness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogeava E, Gallacher J, St George-Hyslop P, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossu P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Deniz Naranjo MC, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Caffarra P, Nacmias B, Gilbert JR, Mayhaus M, Lannefelt L, Hakonarson H, Pichler S, Carrasquillo MM, Ingelsson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E, Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM, Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley TH Jr, Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Bruijn RF, Passmore P, Montine TJ, Bettens K, Rotter JJ, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell JF, Nalls MA, Ritchie K, Lunetta KL, Kauwe JS, Boerwinkle E, Riemenschneider M, Boada M, Hiltunen M, Martin ER, Schmidt R, Rujescu D, Wang LS, Dartigues JF, Mayeux R, Tzourio C, Hofman A, Nothen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA, van Duijn CM, Van Broeckhoven C, Moskvin V, Seshadri S, Williams J, Schellenberg GD, Amouyel P (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-1458.
- [20] Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, Breen G, Byrne EM, Blackwood DH, Boomsma DI, Cichon S, Heath AC, Holsboer F, Lucae S, Madden PA, Martin NG, McGuffin P, Muglia P, Nothen MM, Penninx BP, Pergadia ML, Potash JB, Rietschel M, Lin D, Muller-Myhsok B, Shi J, Steinberg S, Grabe HJ, Lichtenstein P, Magnusson P, Perlis RH, Preisig M, Smoller JW, Stefansson K, Uher R, Kutalik Z, Tansey KE, Teumer A, Viktorin A, Barnes MR, Bettecken T, Binder EB, Breuer R, Castro VM, Churchill SE, Coryell WH, Craddock N, Craig IW, Czamara D, De Geus EJ, Degenhardt F, Farmer AE, Fava M, Frank J, Gainer VS, Gallagher PJ, Gordon SD, Goryachev S, Gross M, Guipponi M, Henders AK, Herms S, Hickie IB, Hoefels S, Hoogendijk W, Hottenga JJ, Iosifescu DV, Ising M, Jones I, Jones L, Jung-Ying T, Knowles JA, Kohane IS, Kohli MA, Korszun A, Landen M, Lawson WB, Lewis G, Macintyre D, Maier W, Mattheisen M, McGrath PJ, McIntosh A, McLean A, Middeldorp CM, Middleton L, Montgomery GM, Murphy SN, Nauck M, Nolen WA, Nyholt DR, O'Donovan M, Oskarsson H, Pedersen N, Scheftner WA, Schulz A, Schulze TG, Shyn SI, Sigurdsson E, Slager SL, Smit JH, Stefansson H, Steffens M, Thorgeirsson T, Tozzi F, Treutlein J, Uhr M, van den Oord EJ, Van Grootheest G, Volzke H, Weiburg JB, Willemsen G, Zitman FG, Neale B, Daly M, Levinson DF, Sullivan PF (2013) A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry* **18**, 497-511.
- [21] Muglia P, Tozzi F, Galwey NW, Francks C, Upmanyu R, Kong XQ, Antoniadis A, Domenici E, Perry J, Rothen S, Vandeleur CL, Mooser V, Waeber G, Vollenweider P, Preisig M, Lucae S, Muller-Myhsok B, Holsboer F, Middleton LT, Roses AD (2010) Genome-wide association study of recurrent major depressive disorder in two European case-control cohorts. *Mol Psychiatry* **15**, 589-601.
- [22] Shi J, Potash JB, Knowles JA, Weissman MM, Coryell W, Scheftner WA, Lawson WB, DePaulo JR Jr, Gejman PV, Sanders AR, Johnson JK, Adams P, Chaudhury S, Jancic D, Evgrafov O, Zvinyatskovskiy A, Ertman N, Gladis M, Neimanas K, Goodell M, Hale N, Ney N, Verma R, Mirel D, Holmans P, Levinson DF (2011) Genome-wide association study of recurrent early-onset major depressive disorder. *Mol Psychiatry* **16**, 193-201.
- [23] CONVERGE consortium (2015) Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* **523**, 588-591.
- [24] Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, Hayward NK, Montgomery GW, Visscher PM, Martin NG, Macgregor S (2010) A versatile gene-based test for genome-wide association studies. *Am J Hum Genet* **87**, 139-145.
- [25] Mishra A, Macgregor S (2015) VEGAS2: Software for more flexible gene-based testing. *Twin Res Hum Genet* **18**, 86-91.
- [26] Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120.
- [27] Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760.
- [28] Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA (2013) From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* **43**, 11 10 11-33.
- [29] Wang K, Li M, Hakonarson H (2010) ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164.
- [30] Wang D, Fan Y, Malhi M, Bi R, Wu Y, Xu M, Yu XF, Long H, Li YY, Zhang DF, Yao YG (2018) Missense variants in HIF1A and LACC1 contribute to leprosy risk in Han Chinese. *Am J Hum Genet* **102**, 794-805.
- [31] The 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR (2015) A global reference for human genetic variation. *Nature* **526**, 68-74.

- [32] Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, DeFlaux N, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G, Tusie-Luna MT, Weisburd B, Won HH, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S, Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarrroll S, McCarthy MI, McGovern D, McPherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J, Tsuang MT, Watkins HC, Wilson JG, Daly MJ, MacArthur DG (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285-291.
- [33] Xu M, Zhang DF, Luo R, Wu Y, Zhou H, Kong LL, Bi R, Yao YG (2018) A systematic integrated analysis of brain expression profiles reveals YAP1 and other prioritized hub genes as important upstream regulators in Alzheimer's disease. *Alzheimers Dement* **14**, 215-229.
- [34] Matarin M, Salih DA, Yasvoina M, Cummings DM, Guelfi S, Liu W, Nahaboo Solim MA, Moens TG, Paublete RM, Ali SS, Perona M, Desai R, Smith KJ, Latcham J, Fullelove M, Richardson JC, Hardy J, Edwards FA (2015) A genome-wide gene-expression analysis and database in transgenic mice during development of amyloid or tau pathology. *Cell Rep* **10**, 633-644.
- [35] Guerreiro R, Bras J, Hardy J (2013) SnapShot: Genetics of Alzheimer's disease. *Cell* **155**, 968-968 e961.
- [36] Niculescu AB 3rd, Segal DS, Kuczenski R, Barrett T, Hauger RL, Kelsoe JR (2000) Identifying a series of candidate genes for mania and psychosis: A convergent functional genomics approach. *Physiol Genomics* **4**, 83-91.
- [37] Niculescu AB, Le-Niculescu H (2010) Convergent Functional Genomics: What we have learned and can learn about genes, pathways, and mechanisms. *Neuropsychopharmacology* **35**, 355-356.
- [38] Huang KL, Marcora E, Pimenova AA, Di Narzo AF, Kapoor M, Jin SC, Harari O, Bertelsen S, Fairfax BP, Czajkowski J, Chouraki V, Grenier-Boley B, Bellenguez C, Deming Y, McKenzie A, Raj T, Renton AE, Budde J, Smith A, Fitzpatrick A, Bis JC, DeStefano A, Adams HHH, Ikram MA, van der Lee S, Del-Aguila JL, Fernandez MV, Ibanez L, Sims R, Escott-Price V, Mayeux R, Haines JL, Farrer LA, Pericak-Vance MA, Lambert JC, van Duijn C, Launer L, Seshadri S, Williams J, Amouyel P, Schellenberg GD, Zhang B, Borecki I, Kauwe JSK, Cruchaga C, Hao K, Goate AM (2017) A common haplotype lowers PU.1 expression in myeloid cells and delays onset of Alzheimer's disease. *Nat Neurosci* **20**, 1052-1061.
- [39] Kaup AR, Byers AL, Falvey C, Simonsick EM, Satterfield S, Ayonayon HN, Smagula SF, Rubin SM, Yaffe K (2016) Trajectories of depressive symptoms in older adults and risk of dementia. *JAMA Psychiatry* **73**, 525-531.
- [40] Saczynski JS, Beiser A, Seshadri S, Auerbach S, Wolf PA, Au R (2010) Depressive symptoms and risk of dementia: The Framingham Heart Study. *Neurology* **75**, 35-41.
- [41] Gibson J, Russ TC, Adams MJ, Clarke TK, Howard DM, Hall LS, Fernandez-Pujals AM, Wigmore EM, Hayward C, Davies G, Murray AD, Smith BH, Porteous DJ, Deary IJ, McIntosh AM (2017) Assessing the presence of shared genetic architecture between Alzheimer's disease and major depressive disorder using genome-wide association data. *Transl Psychiatry* **7**, e1094.
- [42] Barbosa AC, Kim MS, Ertunc M, Adachi M, Nelson ED, McAnally J, Richardson JA, Kavalali ET, Monteggia LM, Bassel-Duby R, Olson EN (2008) MEF2C, a transcription factor that facilitates learning and memory by negative regulation of synapse numbers and function. *Proc Natl Acad Sci U S A* **105**, 9391-9396.
- [43] Lo MT, Hinds DA, Tung JY, Franz C, Fan CC, Wang Y, Smeland OB, Schork A, Holland D, Kauppi K, Sanyal N, Escott-Price V, Smith DJ, O'Donovan M, Stefansson H, Bjornsdottir G, Thorgeirsson TE, Stefansson K, McEvoy LK, Dale AM, Andreassen OA, Chen CH (2017) Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. *Nat Genet* **49**, 152-156.
- [44] Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427.
- [45] Spiechowicz M, Bernstein HG, Dobrowolny H, Lesniak W, Mawrin C, Bogerts B, Kuznicki J, Filipek A (2006) Density of Sgt1-immunopositive neurons is decreased in the cerebral cortex of Alzheimer's disease brain. *Neurochem Int* **49**, 487-493.
- [46] Wong PT, McGeer PL, Rossor M, McGeer EG (1982) Ornithine aminotransferase in Huntington's disease. *Brain Res* **231**, 466-471.
- [47] Breiderhoff T, Christiansen GB, Pallesen LT, Vaegter C, Nykjaer A, Holm MM, Glerup S, Willnow TE (2013) Sortilin-related receptor SORCS3 is a postsynaptic modulator of synaptic depression and fear extinction. *PLoS One* **8**, e75006.
- [48] Reitz C, Tokuhira S, Clark LN, Conrad C, Vonsattel JP, Hazrati LN, Palotas A, Lantigua R, Medrano M, I ZJ-V, Vardarajan B, Simkin I, Haines JL, Pericak-Vance MA, Farrer LA, Lee JH, Rogaeva E, George-Hyslop PS, Mayeux R (2011) SORCS1 alters amyloid precursor protein processing and variants may increase Alzheimer's disease risk. *Ann Neurol* **69**, 47-64.
- [49] Pedersen CB, Mors O, Bertelsen A, Waltoft BL, Agerbo E, McGrath JJ, Mortensen PB, Eaton WW (2014) A comprehensive nationwide study of the incidence rate and lifetime risk for treated mental disorders. *JAMA Psychiatry* **71**, 573-581.
- [50] Li G, Wang LY, Shofer JB, Thompson ML, Peskind ER, McCormick W, Bowen JD, Crane PK, Larson EB (2011) Temporal relationship between depression and dementia: Findings from a large community-based 15-year follow-up study. *Arch Gen Psychiatry* **68**, 970-977.
- [51] Mauricio M, O'Hara R, Yesavage JA, Friedman L, Kraemer HC, Van De Water M, Murphy GM Jr (2000) A longitudinal study of apolipoprotein-E genotype and depressive symptoms in community-dwelling older adults. *Am J Geriatr Psychiatry* **8**, 196-200.

Supplementary Materials

Supplementary Table 1. Candidate genes in the MDD-associated GWAS loci and their potential association with AD based on the IGAP data

rs ID	Genes in the GWAS loci	MDD-GWAS p-value	AD-GWAS p-value
rs10514299	<i>TMEM161B, MEF2C</i>	9.99×10^{-16}	0.8776
rs1518395	<i>VRK2</i>	4.32×10^{-12}	0.9664
rs2179744	<i>L3MBTL2</i>	6.03×10^{-11}	0.0004
rs11209948	<i>NEGR1</i>	8.38×10^{-11}	0.2800
rs454214	<i>TMEM161B, MEF2C</i>	1.09×10^{-9}	0.0213
rs301806	<i>RERE</i>	1.90×10^{-9}	0.3574
rs1475120	<i>HACE1, LIN28B</i>	4.17×10^{-9}	0.2659
rs10786831	<i>SORCS3</i>	8.11×10^{-9}	0.2988
rs12552	<i>OLFM4, SUGT1, MIR759</i>	8.16×10^{-9}	0.0777
rs6476606	<i>PAX5</i>	1.20×10^{-8}	0.9275
rs8025231	<i>MEIS2, TMC05A, PAX6, PAX5</i>	1.23×10^{-8}	0.8981
rs12065553	/	1.32×10^{-8}	0.2481
rs1656369	<i>RSRC1, MLF1</i>	1.34×10^{-8}	0.9078
rs4543289	/	1.36×10^{-8}	0.0039
rs2125716	<i>SLC6A15</i>	3.05×10^{-8}	0.9882
rs2422321	<i>NEGR1</i>	3.18×10^{-8}	0.8054
rs7044150	<i>KIAA0020, RFX3</i>	4.31×10^{-8}	0.2996
rs12415800	<i>OAT, LHPP, FAM53B, METTL10, FAM175B</i>	2.37×10^{-10}	/
rs35936514	<i>CTNNA3, DNAJC12, SIRT1, HERC4, MYPN</i>	6.43×10^{-12}	0.8101

Note: GWAS results for MDD were retrieved from the 23andMe-based MDD GWAS [1] and the CONVERGE study [2]. GWAS results for AD were retrieved from summary statistics of the International Genomics of Alzheimer's Project (IGAP) [3]. Significant p-values (<0.05) were marked in bold.

Supplementary Table 2. Rare coding variants in MDD risk genes in 107 Chinese patients with AD.

Gene	Position	SNP	Allele	Mutation	AC/AN		p-value	OR
					AD	CN		
<i>VRK2</i>	chr2:58386773	rs192754138	G/A	p.R468H	1/214	1/576	0.469	2.700
	chr2:58373581	rs369008157	T/C	p.I385T	2/214	1/526	0.202	4.953
	chr2:58316786	rs35966666	A/G	p.I157M [#]	2/214	6/736	1.000	1.148
<i>L3MBTL2</i>	chr22:41601385	rs3804097	A/G	p.I7V	10/212	12/736	0.016	2.987
	chr22:41605809	rs150194965	G/C	p.S45T [#]	3/212	7/736	0.702	1.495
	chr22:41620735	rs151151722	G/A	p.R394Q [#]	2/212	0/160	0.508	3.812
	chr22:41615519	.	G/A	p.A233T [#]	2/212	0/160	0.508	3.812
	chr22:41616786	.	A/G	p.H256R [#]	1/212	0/160	1.000	2.277
	chr22:41617247	rs2277846	C/T	p.R300W [#]	3/212	4/736	0.190	2.627
<i>NEGR1</i>	chr1:72076799	.	C/A	p.G233V [#]	1/212	0/160	1.000	2.277
<i>RERE</i>	chr1:8415639	.	G/T	p.L1503M [#]	3/140	1/136	0.622	2.956
	chr1:8418541	rs200525338	C/T	p.A1352T [#]	1/212	5/526	0.679	0.494
	chr1:8418655	.	T/G	p.I1314L [#]	1/212	0/160	1.000	2.277
	chr1:8422837	rs200477297	T/C	p.N603S [#]	1/214	4/526	1.000	0.613
	chr1:8418898	.	C/T	p.E1233K [#]	1/210	0/160	1.000	2.298
<i>SORCS3</i>	chr10:106401650	.	G/T	p.A189S [#]	1/206	1/320	1.000	1.556
	chr10:106401666	.	C/T	p.S194L [#]	1/208	0/160	1.000	2.320
	chr10:107023094	rs146780166	G/A	p.E1211K [#]	1/212	0/366	0.367	5.199
	chr10:107016619	.	T/C	p.M1127T [#]	1/212	0/160	1.000	2.277
	chr10:106899177	.	G/A	p.S412N [#]	1/212	0/160	1.000	2.277
	chr10:107005317	.	C/G	p.D962E	1/212	0/160	1.000	2.277
<i>OLFM4</i>	chr13:53624380	rs77938256	A/C	p.N336T [#]	1/212	5/736	1.000	0.693
	chr13:53624425	.	T/A	p.V351D [#]	1/212	0/160	1.000	2.277
<i>PAX5</i>	chr9:37034023	rs139701864	A/T	p.D2E [#]	4/206	7/736	0.269	2.062
	chr9:36882052	rs543302698	G/A	p.P321 [#] S	1/212	0/160	1.000	2.277
<i>TMC05A</i>	chr15:38230419	.	G/C	p.E126 [#] Q	1/212	0/160	1.000	2.277
	chr15:38233869	rs2937991	T/C	spliceSite	212/212	736/736	1.000	0.289
<i>RSRC1</i>	chr3:157840012	.	G/A	p.G40E	1/214	0/160	1.000	2.255
	chr3:157839961	.	G/A	p.R23Q [#]	1/214	0/160	1.000	2.255
<i>MLF1</i>	chr3:158320674	.	C/A	p.T191N	1/212	0/160	1.000	2.277
	chr3:158320598	rs201104179	G/A	p.D166N [#]	1/214	0/366	0.369	5.150
<i>SLC6A15</i>	chr12:85266484	rs12424429	G/A	p.A400V [#]	4/212	13/736	1.000	1.070
	chr12:85277540	rs139320132	T/C	p.H285R	2/210	1/366	0.302	3.510
	chr12:85277727	.	C/T	p.G223S [#]	1/212	0/160	1.000	2.277
	chr12:85257227	rs3782369	A/C	p.I603M [#]	3/212	0/366	0.049	12.246
<i>KIAA0020</i>	chr9:2811557	rs2270889	C/T	p.R480Q [#]	1/212	4/526	1.000	0.618

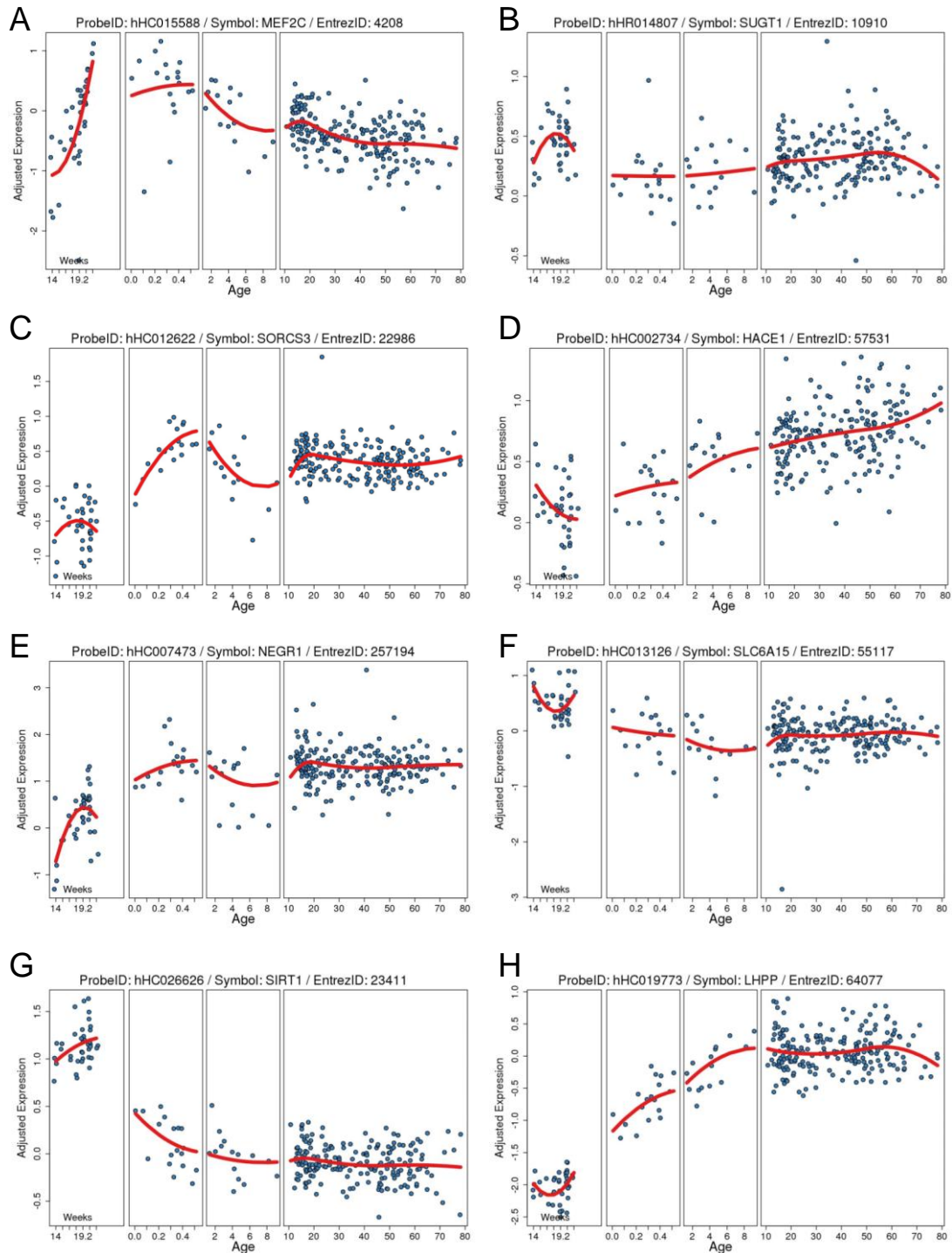
	chr9:2827077	.	C/G	p.R344T [#]	1/212	0/160	1.000	2.277
	chr9:2812238	.	C/A	p.G465V [#]	1/212	1/320	1.000	1.512
	chr9:2828745	rs201465798	C/G	p.E296Q [#]	1/212	1/576	0.466	2.725
	chr9:2810343	.	C/T	spliceSite	1/212	0/160	1.000	2.277
	chr9:2838470	rs10968457	C/T	p.S13N	9/212	33/736	1.000	0.944
<i>RFX3</i>	chr9:3225130	.	A/G	p.L721P [#]	1/212	1/320	1.000	1.512
	chr9:3271018	rs41314211	G/A	p.T396I [#]	1/212	0/160	1.000	2.277
<i>OAT</i>	chr10:126089508	.	T/C	p.I354V	1/212	0/160	1.000	2.277
	chr10:126097420	rs199957428	A/G	p.V105A [#]	1/212	2/576	1.000	1.360
<i>LHPP</i>	chr10:126150523	rs75426652	C/A	p.T31K	2/200	8/718	1.000	0.896
	chr10:126172794	.	T/C	p.L71P [#]	1/212	0/160	1.000	2.277
	chr10:126176996	.	C/T	p.R107 [#] C	1/212	0/160	1.000	2.277
	chr10:126177053	rs200710743	G/A	p.A126T [#]	1/212	0/160	1.000	2.277
<i>FAM53B</i>	chr10:126370915	.	C/G	p.C56S [#]	1/212	0/160	1.000	2.277
	chr10:126395213	rs200550066	G/A	p.R24C	2/210	5/736	0.655	1.406
<i>METTL10</i>	chr10:126463309	.	G/A	p.P112S [#]	1/212	0/160	1.000	2.277
	chr10:126477664	.	T/C	p.D80G [#]	1/212	0/160	1.000	2.277
<i>FAM175B</i>	chr10:126523250	.	T/G	p.L320V	1/212	0/160	1.000	2.277
<i>CTNNA3</i>	chr10:67680239	.	A/G	p.V846A [#]	1/212	0/160	1.000	2.277
	chr10:68040262	rs190239785	A/G	p.I617T [#]	1/212	4/736	1.000	0.867
<i>DNAJC12</i>	chr10:69571259	rs201856212	G/A	p.T107M	1/212	1/576	0.466	2.725
	chr10:69571365	rs569240271	G/A	p.R72*	1/212	0/160	1.000	2.277
<i>SIRT1</i>	chr10:69676240	rs114182972	G/A	p.A712T [#]	1/212	2/366	1.000	0.863
	chr10:69651268	.	G/A	p.E300K [#]	1/212	0/160	1.000	2.277
<i>HERC4</i>	chr10:69699400	.	T/C	p.Q847R [#]	1/212	0/160	1.000	2.277
	chr10:69682825	rs202149588	C/T	p.G1013R [#]	1/212	0/366	0.367	5.199
	chr10:69793910	.	T/C	p.Y166C [#]	1/212	0/160	1.000	2.277
<i>MYPN</i>	chr10:69957212	rs372824930	C/T	p.R1088C [#]	1/212	0/160	1.000	2.277
	chr10:69933986	.	C/T	p.P713S	1/212	0/160	1.000	2.277
	chr10:69908157	rs11596653	T/C	p.V393A	1/212	1/320	1.000	1.512
	chr10:69881590	rs201203517	C/T	p.P132L	1/212	2/736	0.532	1.739
	chr10:69934278	.	G/A	p.R810H	2/212	4/320	1.000	0.752
	chr10:69881606	.	G/C	p.R137S	1/212	0/160	1.000	2.277
	chr10:69959320	rs138313730	C/A	p.L1161I [#]	2/212	1/366	0.558	3.476

Note: Position was based on human genome reference assembly (build GRCh37/hg19); Allele, reference allele / alternative allele; AC, allele count; AN, allele number; OR, odds ratio; ExAC, Exome Aggregation Consortium (<http://exac.broadinstitute.org/>); AD, 107 subjects with Alzheimer's disease; CN, 368 subjects combining 160 in-house non-dementia individuals [4] and 208 Chinese from the 1000 Genome project [5]. The ANNOVAR was used to annotate variants, which were assigned into different functional categories according to their locations and expected effect on the encoded gene products [6]. [#] Coding variants predicted to be damaging according to ANNOVAR annotation.

Supplementary Table 3. Convergent lines of evidence for the MDD GWAS genes

Gene	Common variant	Coding variant	Patient	Mouse	PPI	CFG score
<i>HACE1</i>	0	0	1	1	1	3
<i>MEF2C</i>	1	0	1	0	1	3
<i>NEGR1</i>	0	0	1	1	1	3
<i>SLC6A15</i>	0	1	1	1	0	3
<i>CTNNA3</i>	0	0	0	1	1	2
<i>DNAJC12</i>	0	0	1	1	0	2
<i>FAM53B</i>	0	0	1	1	0	2
<i>HERC4</i>	0	0	0	1	1	2
<i>KIAA0020</i>	0	1	0	0	1	2
<i>L3MBTL2</i>	1	1	0	0	0	2
<i>LHPP</i>	0	1	1	0	0	2
<i>MEIS2</i>	0	0	0	1	1	2
<i>OAT</i>	0	0	1	1	0	2
<i>PAX6</i>	0	0	1	0	1	2
<i>SIRT1</i>	0	0	1	0	1	2
<i>SORCS3</i>	0	0	1	1	0	2
<i>SUGT1</i>	1	0	0	0	1	2
<i>LIN28B</i>	0	0	1	0	0	1
<i>MYPN</i>	0	1	0	0	0	1
<i>OLFM4</i>	0	0	1	0	0	1
<i>PAX5</i>	0	0	0	0	1	1
<i>RERE</i>	0	0	0	0	1	1
<i>RFX3</i>	0	0	1	0	0	1
<i>RSRC1</i>	0	0	0	0	1	1
<i>TMEM161B</i>	0	1	0	0	0	1
<i>VRK2</i>	0	0	0	1	0	1
<i>FAM175B</i>	0	0	0	0	0	0
<i>METTL10</i>	0	0	0	0	0	0
<i>MLF1</i>	0	0	0	0	0	0
<i>TMC05A</i>	0	0	0	0	0	0

Note: A gene was defined as AD-associated if it 1) showed a significant association with AD at the gene-based test using the IGAP data [3]; 2) had significantly AD-associated rare variants in 107 Han Chinese patients; 3) was differentially expressed in any brain tissue of AD patients [7]; 4) was correlated with AD pathology in AD mice [8]; 5) was involved in the PPI network as evaluated by the online tool STRING. For each line of evidence, one point was assigned as the convergent functional genomics (CFG) score if the above fact was observed; otherwise zero point. The higher CFG score, the stronger association with AD.



Supplementary Fig. 1. Expression pattern of MDD-risk genes in the prefrontal cortex across human lifespan. Data retrieved from BrainCloud (<http://braincloud.jhmi.edu/>) [9], an online application for exploring the temporal dynamics of transcription in the human prefrontal cortex across the lifespan. Most of the MDD-risk genes *MEF2C* (A), *SUGT1* (B), *SORCS3* (C), *HACE1* (D), *NEGR1* (E), *SLC6A15* (F), *SIRT1* (G), and *LHPP* (H) showed a stable mRNA expression level in prefrontal cortex from early adulthood (20 years old) to late-life (80 years old).

REFERENCES

- [1] Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR, Tung JY, Hinds DA, Perlis RH, Winslow AR (2016) Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat Genet* **48**, 1031-1036.
- [2] CONVERGE consortium (2015) Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* **523**, 588-591.
- [3] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thornton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin CF, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau MT, Choi SH, Reitz C, Pasquier F, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston JA, Evans D, Lovestone S, Letenneur L, Moron FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fievet N, Huentelman MW, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGuinness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogaeva E, Gallacher J, St George-Hyslop P, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossu P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Deniz Naranjo MC, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Caffarra P, Nacmias B, Gilbert JR, Mayhaus M, Lannefelt L, Hakonarson H, Pichler S, Carrasquillo MM, Ingelsson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E,

- Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM, Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley TH, Jr., Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Bruijn RF, Passmore P, Montine TJ, Bettens K, Rotter JI, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell JF, Nalls MA, Ritchie K, Lunetta KL, Kauwe JS, Boerwinkle E, Riemenschneider M, Boada M, Hiltunen M, Martin ER, Schmidt R, Rujescu D, Wang LS, Dartigues JF, Mayeux R, Tzourio C, Hofman A, Nothen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA, van Duijn CM, Van Broeckhoven C, Moskvina V, Seshadri S, Williams J, Schellenberg GD, Amouyel P (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-1458.
- [4] Wang D, Fan Y, Malhi M, Bi R, Wu Y, Xu M, Yu XF, Long H, Li YY, Zhang DF, Yao YG (2018) Missense variants in HIF1A and LACC1 contribute to leprosy risk in Han Chinese. *Am J Hum Genet* **102**, 794-805.
- [5] The 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR (2015) A global reference for human genetic variation. *Nature* **526**, 68-74.
- [6] Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164.
- [7] Xu M, Zhang DF, Luo R, Wu Y, Zhou H, Kong LL, Bi R, Yao YG (2018) A systematic integrated analysis of brain expression profiles reveals YAP1 and other prioritized hub

- genes as important upstream regulators in Alzheimer's disease. *Alzheimers Dement* **14**, 215-229.
- [8] Matarin M, Salih DA, Yasvoina M, Cummings DM, Guelfi S, Liu W, Nahaboo Solim MA, Moens TG, Paublete RM, Ali SS, Perona M, Desai R, Smith KJ, Latcham J, Fulleylove M, Richardson JC, Hardy J, Edwards FA (2015) A genome-wide gene-expression analysis and database in transgenic mice during development of amyloid or tau pathology. *Cell Rep* **10**, 633-644.
- [9] Colantuoni C, Lipska BK, Ye T, Hyde TM, Tao R, Leek JT, Colantuoni EA, Elkahloun AG, Herman MM, Weinberger DR, Kleinman JE (2011) Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* **478**, 519-523.