Genome-wide association studies on Northern Italy isolated populations provide further support concerning genetic susceptibility for major depressive disorder

Vincenzo Dattilo[a], Sheila Ulivi[b], Alessandra Minelli[a,c], Martina La Bianca[b], Edoardo Giacopuzzi[d,e], Marco Bortolomasi[f], Stefano Bignotti[g], Massimo Gennarelli[a,c], Paolo Gasparini[b,h] and Maria Pina Concas[b]

[a]Genetics Unit, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy; [b]Institute for Maternal and Child Health-IRCCS Burlo Garofolo, Trieste, Italy; [c]Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy; [d]Wellcome Centre for Human Genetics, Oxford University, Oxford, UK; [e]NIHR Oxford Biomedical Research Centre, Oxford, UK; [f]Psychiatric Hospital "Villa Santa Chiara", Verona, Italy; [g]Unit of Psychiatry, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy; [h]Department of Medicine, Surgery and Health Science, University of Trieste, Trieste, Italy

ABSTRACT

Objectives: Major depressive disorder (MDD) is a psychiatric disorder with pathogenesis influenced by both genetic and environmental factors. To date, the molecular-level understanding of its aetiology remains unclear. Thus, we aimed to identify genetic variants and susceptibility genes for MDD with a genome-wide association study (GWAS) approach.

Methods: We performed a meta-analysis of GWASs and a gene-based analysis on two Northern Italy isolated populations (cases/controls $n = 166/472$ and 33/320), followed by replication and polygenic risk score (PRS) analyses in Italian independent samples (cases $n = 464$, controls $n = 339$).

Results: We identified two novel MDD-associated genes, *KCNQ5* (lead SNP rs867262, $p = 3.82 \times 10^{-9}$) and *CTNNA2* (rs6729523, $p = 1.25 \times 10^{-8}$). The gene-based analysis revealed another six genes ($p < 2.703 \times 10^{-6}$): *GRM7*, *CTNT4*, *SNRK*, *SRGAP3*, *TRAPPC9*, and *FHIT*. No replication of the genome-wide significant SNPs was found in the independent cohort, even if 14 SNPs around *CTNNA2* showed association with MDD and related phenotypes at the nominal level of p (<0.05). Furthermore, the PRS model developed in the discovery cohort discriminated cases and controls in the replication cohort.

Conclusions: Our work suggests new possible genes associated with MDD, and the PRS analysis confirms the polygenic nature of this disorder. Future studies are required to better understand the role of these findings in MDD.

KEYWORDS
Major depressive disorder; genetic isolates; genome-wide association study; *KCNQ5*; *CTNNA2*

Introduction

Major Depressive Disorder (MDD) is the most common psychiatric illness consisting of the third leading cause of years lived with disability worldwide (WHO 2021). This disorder affects not only adolescents and adults but also children. It is a complex disorder influenced by both genetic and environmental factors. The genetic contribution to the disorder arises from parental and twin-based studies that estimated a heritability between 30 and 50% (Kendall et al. 2021). The first approach used in the genetic studies of MDD was the candidate-gene association. This method aims to identify possible associations between the disorder and genes selected *a priori* based on their biological function, with a putative role in the pathophysiological mechanisms of the disorder (Shadrina et al. 2018). Although candidate-gene studies have revealed more than 100 candidates as suspected risk genes, the results are often conflicting. The discordances are mainly due to a lack of statistical power deriving from small sample sizes and, in some cases, incomplete knowledge of the biological function of the genes or neglected interactions between gene and environment (Verbeek et al. 2014). More recently, genome-wide association studies (GWAS) have been used to identify variant loci associated with MDD without *a priori* selection of genes. To date, GWAS on MDD

endeavoured in identifying individual associated loci and in replicating significant findings. The first MDD GWASs, conducted on small sample sizes, failed in discovering common variants significantly associated with the disorder (Kendall et al. 2021). Since 2016, large-scale GWASs of major depression, conducted on increased sample size with minimal phenotyping, have allowed the identification of genome-wide significant loci (Hyde et al. 2016; Okbay et al. 2016; Direk et al. 2017). In 2018, a GWAS conducted on 322,580 UK Biobank participants revealed 16 independent loci associated with broad depression phenotypes (Howard et al. 2018). The most recent successes of GWASs come from two meta-analyses conducted by the PGC MDD working group with an increasing sample size (Wray et al. 2018; Howard et al. 2019). Wray et al. conducted the analyses on 130,664 MDD cases and 330,470 controls by pooling cohorts with different phenotype definitions and described 44 associated genetic loci. Howard et al. used three combined datasets with a minimal phenotyping (246,363 cases and 561,190 controls) in their analyses, yielding 102 genome-wide significant variants, 87 of which were replicated in an independent sample. Recently, bivariate genome-wide association analyses on multiple psychiatric phenotypes revealed eight novels independently replicated depression loci (Amare et al. 2020). Collectively, these GWAS confirmed that the difficulty in identifying risk loci for MDD mainly lies in the high prevalence of the disorder within the population as well as in its polygenic architecture, thus resulting in a combined effect of many genetic variants with individually small effect sizes (Wray et al. 2014). In addition, the genetic heterogeneity of MDD is also accompanied by a broad symptomatologic spectrum and sex differences in the disorder prevalence (Kendler et al. 2001). Other factors, such as ethnicity, onset time, recurrence, severity, and childhood traumatic history are all sources of heterogeneity in the disorder (Zhang and Rong 2019). While a reasonable strategy to overcome the difficulty in these studies is to increase sample size at the expense of phenotype accuracy, another is to study more accurate phenotypes in smaller cohorts (Mullins and Lewis 2017). The CONVERGE (China, Oxford, and Virginia Commonwealth University Experimental Research on Genetic Epidemiology) Consortium has collected a homogeneous sample by restricting the phenotype to recurrent severe depression in Han Chinese women (Cai et al. 2015). The authors identified two loci followed by replication in an independent Chinese sample. An alternative strategy in genetic association studies consists in leveraging the unique characteristics of isolated populations, defined as subpopulations derived from a small number of individuals, isolated because of a founding event and stayed so for many generations. Isolates show higher phenotypic and genetic homogeneity compared to the outbred population, thus allowing a reduced population size in the observation of genetic variability (Charlesworth 2009), mainly for complex disorders, such as MDD (Amin, Belonogova, et al. 2017; Amin, Jovanova, et al. 2017; Amin et al. 2018).

In the present study, we aimed to identify variants/genes implicated in MDD in two Italian genetic isolate populations. We performed GWAS meta-analysis and gene-based analyses in two genetic isolates of Northern Italy. A replication study and polygenic risk score (PRS) analysis were carried out in an independent cohort.

Materials and methods

Discovery step

Samples

The present study involved samples of two genetically isolated cohorts from the Italian Network of Genetic Isolates (INGI), forming a discovery cohort of 199 MDD patients and 792 controls. The first one was made of 638 samples (166 cases and 472 controls) coming from six geographically isolated villages in the Italian Friuli-Venezia Giulia (FVG) region (Clauzetto, Erto, Illegio, Resia, Sauris, and San Martino del Carso). The second cohort included 353 samples (33 cases and 320 controls) referring to two Piedmontese populations in close geographical proximity to the Val Borbera (VBI), namely Trino Vercellese and Val di Susa. The geographic isolates met the criteria defining 'genetic isolates' as separate geographical locations with high rates of endogamy, language barriers, few surnames, few founders, low rates of emigration and immigration, and for which genetic homogeneity was already shown. Indeed, the genetic isolation and the small effective population size of the FVG village populations were manifested by higher levels of genomic homozygosity and elevated linkage disequilibrium (Esko et al. 2013). Regarding the VBI, Colonna and colleagues' findings reveal that the valley has features characteristic of a genetic isolate, such as reduced genetic heterogeneity and reduced effective population size (Colonna et al. 2013). The FVG participants received a structured diagnostic interview using the Composite International Diagnostic Interview (CIDI) to assess current and lifetime diagnoses according to the

Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition—Text Revision (DSM-IV-TR) and the Hamilton Depression Rating Scale (HAM-D) (Hamilton 1960) for assessing depressive symptoms. The VBI participants filled in the self-report Millon Clinical Multiaxial Inventory-III (MCMI-III) (Millon and Davis 1997) composed of 175 true-false questions to assess both Axis I and Axis II disorders according to DSM-IV. Moreover, they were evaluated with a non-structured interview schedule for the presence of current and/or lifetime psychiatric disorders, psychiatric drug treatments, access to psychiatric services, and first-degree family history of psychiatric disorders. In addition, all subjects completed the Symptom Checklist-90-R (SCL-90-R) (Prunas et al. 2012) to evaluate a broad spectrum of psychological symptoms and psychological distress. Subjects who obtained a score lower than 24/30 on the Mini-Mental State Examination (M.M.S.E.), or had cognitive impairment, dementia, or other severe neurological disorders were excluded. All socio-demographic and clinical characteristics of the two genetic isolates are represented in Table 1.

Genotyping

All samples have been genotyped with Illumina 370 K/700K high-density SNP array (Illumina Inc., San Diego, CA, USA). Genotypes were called with Illumina GenomeStudio. Each batch was processed according to standard quality control (QC) procedures with the following criteria for inclusion: sample call rate $\geq$0.95, gender check, SNP call rate $\geq$0.95, Hardy-Weinberg Equilibrium (HWE) p-value $>1 \times 10^{-6}$, and minor allele frequency (MAF) $\geq$0.01. Genotype imputation was conducted using IMPUTE2 (Howie et al. 2009) considering as reference a custom panel generated merging the 1000 Genomes phase 3 (Altshuler et al. 2012) and whole-genome sequences of INGI samples (Cocca et al. 2020). After imputation, SNPs with MAF $<$ 0.01 and Info Score $<$ 0.4 were discarded from the statistical analyses.

Genome-wide association (GWA) analysis and meta-analysis

GWAS was conducted in FVG and VBI separately, using mixed linear models [ABEL R packages (Aulchenko et al. 2007)]. Genomic kinship matrices were used as random effects to take into account relatedness. An additive model was used, adjusted by sex and age. Meta-analysis was performed using an inverse-variance method (METAL) (Willer et al. 2010). We used LD-Score regression (Bulik-Sullivan et al. 2015) and the HapMap3 reference panel to estimate the LD-score regression intercept for the results of both GWAS and the meta-analysis. To better understand the effect of the SNPs on the trait, we transformed the effect size to the Odds Ratio (OR) using LMOR (Lloyd-Jones et al. 2018). Variants with a significant p-value ($<$0.05) for heterogeneity Cochran Q were excluded from the meta-analysis and only concordant variants in the two populations were examined. SNPs with p-value $< 5 \times 10^{-8}$ were considered genome-wide significant. SNPs were annotated with the Variant Effect Predictor tool (VEP, https://www.ensembl.org/info/docs/tools/vep/index.html) (McLaren et al. 2016) to determine the closest genes and to obtain functional characteristics.

Gene-based analysis

Starting from the results of the meta-analysis, gene-based association analyses were performed with MAGMA v1.08 (de Leeuw et al. 2015), implemented in the FUMAGWAS web tool. The genome-wide significance threshold was set based on the number of genes tested, at a p-value $= 2.703 \times 10^{-6}$ (0.05/18499). The analyses were run using the 1000 Genome Phase3 reference panel.

Genotype-tissue expression (GTEx) analysis

Human gene expression levels of the genes highlighted by meta-analysis and gene-based analysis were obtained using the GTEx Database, release v8 (GTEX Consortium 2015).

Replication step

Samples

The replication cohort has been made of 463 MDD patients and 339 controls. Patients with at least moderate to severe depression, who met the Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV) classification system criteria, were voluntarily enrolled in the study. All of them had been referred to Villa S. Chiara Psychiatric Hospital in Verona or the Psychiatry Rehabilitation Unit of IRCCS Centro Fatebenefratelli 'S. Giovanni di Dio' in Brescia. Diagnosis of unipolar depression was confirmed using the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) diagnostic scale. The exclusion criteria were as follows: (a) mental retardation or cognitive disorder; (b) a lifetime history of schizophrenic, schizoaffective, or bipolar disorder; (c) personality disorder, substance abuse, alcohol abuse or dependency, obsessive-compulsive disorder, or post-traumatic stress disorder as the

Table 1. Demographic and clinical characteristics of the subjects belonging to the discovery cohorts, stratified in control and MDD patient groups.

	FVG			VBI		
Characteristic	Controls ($N = 472$)	MDD ($N = 166$)	p-Value	Controls ($N = 320$)	MDD ($N = 33$)	p-Value
Age (years), mean (s.d.)	46.1 (14.0)	49.0 (13.7)	0.02	57.5 (15.2)	57.9 (14.7)	0.87
Gender (F), n (%)	212 (44.9)	130 (78.3)	<0.001	181 (56.6)	29 (87.9)	<0.001
Co-occurrence with anxiety disorders, n (%)	–	68 (41.0)	–	–	20 (60.6)	–
Depression scale score* mean (*SD*)	–	10.04 (8.42)	–	–	1.06 (0.66)	–

FVG: Friuli Venezia Giulia; VBI: Val Borbera; MDD: major depression disorder; s.d.: standard deviation; F: female; p-Value refers to the difference in controls and MDD patients obtained by t-test for continuous variables and Chi-square test for dichotomic variables.
*HAM-D for FVG, SCL-90 depression subscale for VBI.

primary diagnosis; and (d) comorbidity with an eating disorder.

The control sample consisted of unrelated healthy volunteers who were screened for DSM-IV Axis I disorders by expert psychologists using the Mini-International Neuropsychiatric Interview (MINI) (Sheehan et al. 1998). Only healthy volunteers without a history of drug or alcohol abuse or dependence and without a personal or first-degree family history of psychiatric disorders were enrolled in the study. Both patients and controls were Caucasians of Italian descent for at least two generations, residing in north Italy and unrelated to other participants. In both groups, subjects who obtained a score lower than 24/30 on the Mini-Mental State Examination (M.M.S.E.), or had cognitive impairment, dementia, or other severe neurological disorders were excluded. All socio-demographic and clinical characteristics for all groups are shown in Table 2.

Genotyping

Samples were genotyped on three different arrays: Genome-Wide Human SNP 6.0 array (Affymetrix), Illumina Infinium PsychArray-24 BeadChip, and Infinium multi-ethnic genotyping array (MEGA) (Illumina). Each sample batch was processed [Plink v1.9 (Chang et al. 2015)] excluding SNPs with call rate < 0.95, MAF < 0.05, and a Hardy-Weinberg equilibrium p-value $<1 \times 10^{-6}$ and removing samples with a call rate < 0.99, sex discrepancy, unusual heterozygosity (<0.20 or >0.40) and cryptic relatedness (p_hat $>$ 0.20). The batches were then imputed [Minimac3 v.2.0.1 (Das et al. 2016), Shapeit v.2.r837 (Delaneau et al. 2011), 1000G EUR reference panel]. After imputation, SNPs with $R^2 \geq 0.5$ and genotype probability $\geq$0.9 were selected from each batch and merged into a combined dataset. The second round of QC filtering was then applied to remove SNPs with missing genotypes >5%, HWE test p-value $<1 \times 10^{-6}$, and MAF $<$ 1%. We then created two separate datasets, one including only MDD cases and one including both cases and controls.

Replication analysis

We searched in the replication sample the SNPs found as genome-wide significant in the discovery step, but, unfortunately, these were discarded from the analysis due to the applied QC. Thus, we investigated association signals for SNPs located in a 500 kb range around the genes showing significant association at the genome-wide level in the meta-analysis (p-value 5×10^{-8}). Overall, 1554 SNPs were tested for each phenotype, except MDD for which 1437 SNPs were analysed. A significant threshold of the p-value of 3.22×10^{-5} was fixed (Bonferroni correction: 0.05/1554 SNPs). The association of SNPs with case/control status was tested in the dataset including both cases and controls, while the association with age of onset, MADRS (Montgomery-Asberg Depression rating scale) score, co-occurrence of anxiety disorders, and co-occurrence psychotic symptoms was tested in the dataset of MDD cases. We used Plink v.1.9 logistic/linear regression corrected for sex, age, and the first five Principal Components (PCs) and an additive model. Then we used summary statistics from the replication cohort to perform a gene-based analysis with the same method described for the discovery cohort (MAGMA v1.08).

PRS analysis

Standardised PRS on the replication cohorts were computed using PRSice2 (Choi and O'Reilly 2019) with the—score std option, based on the summary statistics obtained from the discovery cohort, including sex, age, and the first five PCs as covariates. Clumping was performed on a 500 kb window with—clump-r^2 set to 0.5. The association between PRS and the specific phenotype was tested using logistic or linear regression and an empirical p-value was computed using 5000 permutations as implemented in PRSice2. The best model was selected for each analysis based on p-value thresholding. The association with case/control

Table 2. Demographic and clinical characteristics of the subjects belonging to the replication cohort, stratified in control and MDD patient groups.

Characteristic	Controls ($N = 339$)	MDD ($N = 464$)	p-Value
Age (years), mean (s.d.)	50.8 (15.4)	56.3 (13.7)	<0.001
Gender (F), *n* (%)	190 (47.6)	316 (68.1)	<0.001
TRD, *n* (%)	–	72 (15.5)	–
Co-occurrence with anxiety disorders, *n* (%)	–	196 (42.2)	–
Co-occurrence with Psychotic symptoms, *n* (%)	–	135 (29.1)	–
MADRS, mean (s.d.)	–	31.6 (5.7)	–
Age of onset, mean (s.d.)	–	38.1 (13.4)	–

MDD: major depressive disorder; s.d.: standard deviation; F: female; MADRS: Montgomery-Asberg Depression rating scale. p-Value refers to the difference in controls and MDD patients obtained by t-test for continuous variables and Chi-square test for dichotomic variables.

status was tested in the case-control dataset, while associations with the age of onset, co-occurrence of anxiety disorder, co-occurrence psychotic symptoms, and MADRS score were tested using the MDD-only dataset.

The entire workflow of the study is shown in Figure 1.

Results

GWAS meta-analysis results

Figure S1 shows Manhattan and QQ plots (Turner 2018) of the results of the meta-analysis. The effect of population stratification was negligible, as confirmed by the values of the genomic inflation factor (lambda = 0.9995 in the meta-analysis, and 1.0165 and 1.0075 in the FVG and VBI GWAS, respectively). LD-score regression intercepts were 0.9936 for the meta-analysis, 1.0049 for FVG, and 1.0021 for VBI. Genome-wide significant results of the meta-analysis are displayed in Table 3, while all the results with p-value $< 1 \times 10^{-5}$ are shown in Table S1. Their VEP annotation is available in Table S2. Eight SNPs reached the statistically significant p-value $< 5 \times 10^{-8}$, identifying two independent regions. Seven SNPs in linkage disequilibrium (LD, min $r^2 = 0.93$, lead SNP rs867262, p-value 3.82×10^{-9}) are located in chromosome 6q13. The SNPs are intronic in the *KCNQ5* (potassium voltage-gated channel subfamily Q member 5) gene (see the regional plot in Figure 2(a)). The minor allele G of the top SNP rs867262 shows a risk effect with an OR of 4.11. The second region is identified by a single genome-wide significant SNP (rs6729523, p-value 1.25×10^{-8}) in chromosome 2p12, within the *CTNNA2* (catenin alpha 2) gene (Figure 2(b)). This signal is supported by two other SNPs in LD with rs6729523 and with suggestive p-value (Table S1 and Figure 2(b)). The minor allele G of rs6729523 shows a risk effect with an OR of 5.14. All the identified SNPs showed info scores ≥0.847 in FVG and ≥0.879 in VBI as displayed in Table S1. GTEx analysis revealed that both genes are expressed in the brain (Figures S2(a,b)).

Replication study

The results of the replication study were summarised in Table S3. After multiple test corrections, no significant associations for single SNPs were identified in the replication cohort, even if 14 SNPs around the *CTNNA2* gene identified from the discovery cohort showed association with MDD, age of onset, and MADRS at the nominal level of the p-value (<0.05).

PRS analysis

The PRS analysis based on discovery cohort summary statistics showed that the combined effect of the detected SNPs can discriminate cases and controls in the case-control replication dataset (p-value 1.43×10^{-4}, empirical p-value 3.3×10^{-3}). The distribution of PRS for MDD phenotype is reported in Figure S3 and complete results for all the PRS associations tested are given in Table S4.

Gene-based analysis

In the gene-based analysis, *KCNQ5* and *CTNNA2* showed significant p-values at the nominal level (p-value of 0.015 and 0.0012, respectively). Moreover, this study highlighted six significant genes (p-value $< 2.703 \times 10^{-6}$, Table 4): *GRM7* (glutamate metabotropic receptor 7), *CNTN4* (contactin 4), *SNRK* (SNF related kinase), *SRGAP3* (SLIT-ROBO Rho GTPase activating protein 3), *TRAPPC9* (trafficking protein particle complex 9) and *FHIT* (fragile histidine triad diadenosine triphosphatase). All these genes showed expression in cerebral tissue as revealed by GTEx analysis (Figures S2(c–h)). No significant results were found in the gene-based analysis of the replication cohort (Table S5).

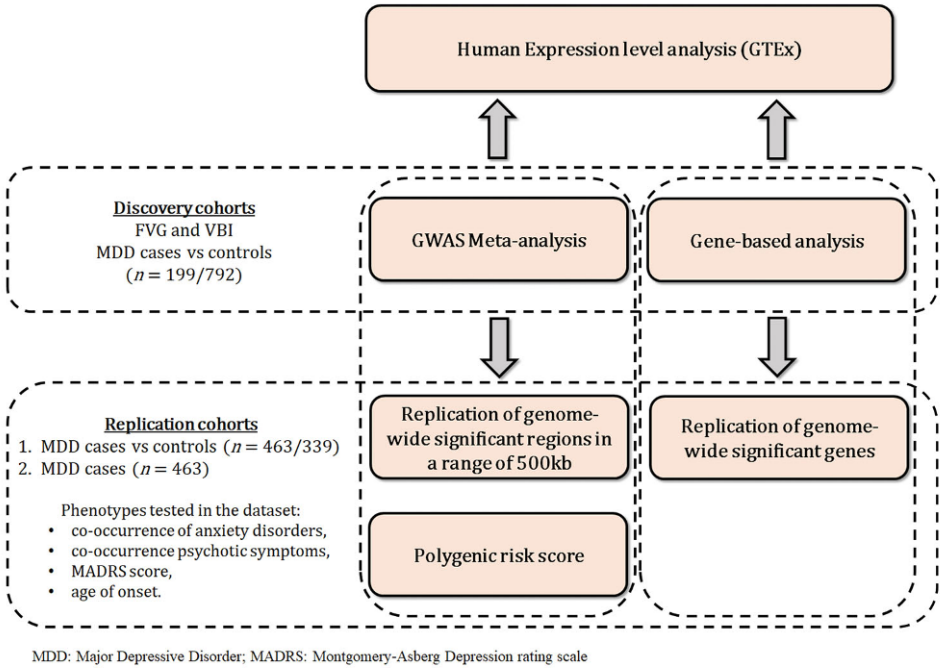

Figure 1. Research flow of the present study.

Table 3. SNPs with a genome-wide significant association (p-value $< 5 \times 10^{-8}$) with major depressive disorder (MDD) in discovery cohorts as obtained by meta-analysis.

SNP	Chr	Pos	EA/OA	EAF	OR (95% CI)	p-Value	Gene	Consequence
rs867262	6	73578179	G/A	0.04	4.11 (2.57, 6.57)	3.82×10^{-9}	*KCNQ5*	intron
rs9360607	6	73580412	C/T	0.03	4.10 (2.57, 6.56)	3.85×10^{-9}	*KCNQ5*	intron
rs9446777	6	73581051	G/A	0.04	4.10 (2.57, 6.55)	3.97×10^{-9}	*KCNQ5*	intron
rs9442860	6	73585145	A/T	0.03	4.08 (2.54, 6.56)	6.39×10^{-9}	*KCNQ5*	intron
rs9442861	6	73588327	A/G	0.03	4.06 (2.53, 6.53)	7.25×10^{-9}	*KCNQ5*	intron
rs9293915	6	73589605	T/C	0.03	4.02 (2.50, 6.47)	1.00×10^{-8}	*KCNQ5*	intron
rs6729523	2	80750869	G/A	0.02	5.14 (2.93, 9.03)	1.25×10^{-8}	*CTNNA2*	intron

SNP: SNP name; Chr: chromosome; Pos: position (base pair); EA: effect allele; OA: other allele; EAF: effect allele frequency; OR: odds ratio; CI: 95% OR confidential interval; FVG: Friuli Venezia Giulia; VBI: Val Borbera.
The seven SNPs in chromosome 6 are in linkage disequilibrium (min $r^2 = 0.93$). All the SNPs showed info score $\geq$0.847 in FVG and $\geq$0.879 in VBI as displayed in Table S1.

Discussion

To identify new genes/variants involved in MDD, we performed a GWAS meta-analysis on two Italian genetic isolate populations, followed by a gene-based analysis and replication study. Figure 3 summarises the results we obtained.

Although the replication step did not allow us to replicate the top SNPs, the results of PRS suggest that the genetic associations detected in the discovery cohorts are also associated with MDD in the replication cohort.

Considering the discovery phase, the study allowed us to identify, by means of SNP-based GWAS meta-analysis, two interesting genes: *KCNQ5* and *CTNNA2*. Even if we failed to replicate a strong signal around these genes in our replication cohort, the presence around *CTNNA2* of a block of SNPs associated at the nominal level of p-value to MDD, age of onset and MADRS further supports the relevance of this gene. Interestingly, both genes are expressed in the brain and encode for proteins involved, although with different mechanisms, in synaptic plasticity and in the formation as well as maintenance of neuronal circuitry (Schaffer et al. 2018; Baculis et al. 2020).

KCNQ5 (OMIM 607357) gene encodes for a voltage-gated potassium channel, important for the regulation of the current modulating the neuronal excitability (Lehman et al. 2017). It belongs to the KV7 family of voltage-gated potassium channels, comprising five KCNQ members (KCNQ1-5) (Brown and Passmore 2009). Four *KCNQ* genes (*KCNQ2-5*) are expressed in the central nervous system both on RNA and protein level (Brown and Passmore 2009) and are therefore excellent candidate susceptibility genes for a wide range of neuronal disorders. Moreover, the expression of ion-channel subunits has been reported to be

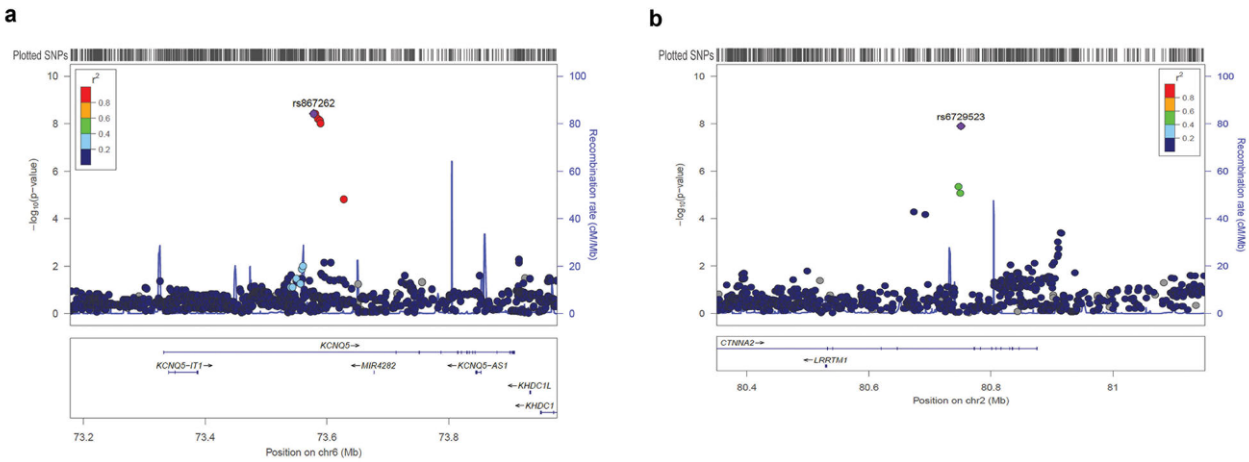

Figure 2. Regional association plot for significant signals in discovery cohort on chromosome 6 (a) and chromosome 2 (b). The minus logarithm of SNP association *p*-value is shown on the y-axis and the SNP position (with gene annotation) on the x-axis. For each SNP, the strength of LD with the lead SNP is colour coded on its r^2. Plots were produced in LocusZoom (Pruim et al. 2010).

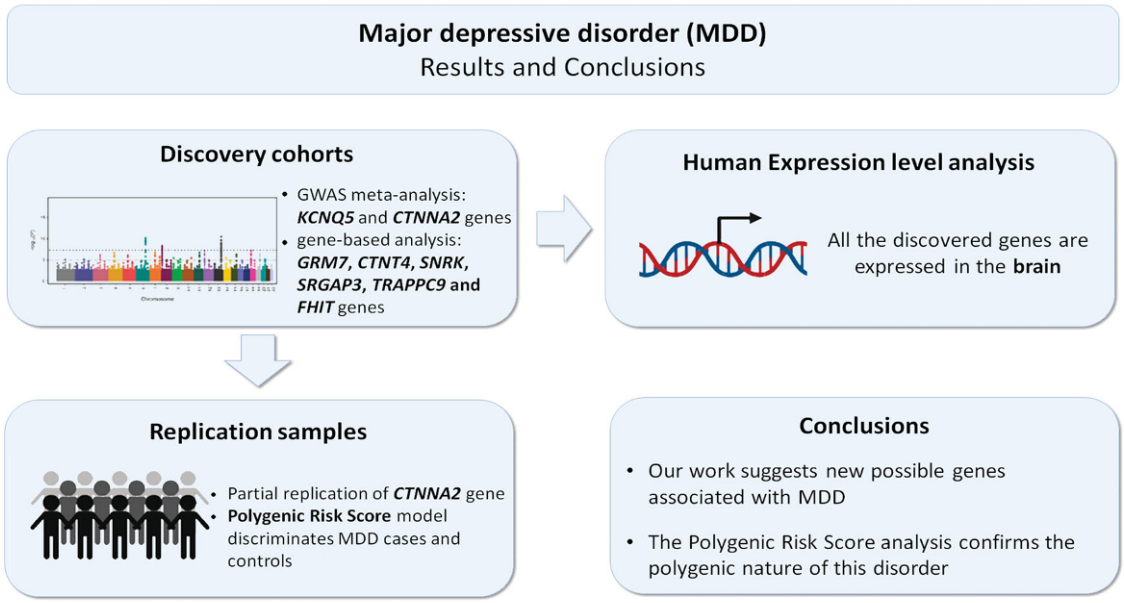

Figure 3. Summary of the obtained results.

modulated in mice exposed to chronic psychotropic drugs as well as electroconvulsive treatment (ECT) (Duncan et al. 2008; Hjaeresen et al. 2008). KV7.5 can form heterotetrameric channels with KV7.3 (Schroeder et al. 2000). Some genetic variants of *KCNQ3* leading to a functional impairing of the channel complex formed by KV7.3 and KV7.5 are reported in patients affected by MDD or other psychiatric and neurodevelopmental diseases (Gilling et al. 2013). A dominant-negative *Kcnq5* mutation in mice has been shown to alter synaptic activity in the hippocampus, where the channel is highly expressed (Tzingounis et al. 2010; Fidzinski et al. 2015). Recently, transcriptomic analyses revealed dysregulation in the expression of *KCNQ5* in patients with neurological and psychiatric diseases including MDD (Baird et al. 2021; Verma and Shakya 2021). On these bases, *KCNQ5* could be a suggestive susceptibility gene for MDD and some typologies of neurological disorders, and due to the considerable overlap in aetiologies also for other psychiatric disorders including anxiety. Given the significance of ion channels in neuronal activity, neurotransmission, plasticity, and the formation of neuronal circuitry, which regulates major processes relevant to psychiatric disorders, the development of drugs modulating the ion-channel activity could counteract anhedonia and anergy that are often resistant to standard treatments in MDD (Smolin et al. 2012). Interestingly, KCNQ

Table 4. Significant gene-based results ($p < 2.703 \times 10^{-6}$) for MDD in discovery cohort as obtained.

Gene name	Chromosome	Start position of gene	End position of gene	Number of SNPS in gene	Number of model parameters	*N*	*Z* statistic	*p*-Value
GRM7	3	6811688	7783215	1651	144	991	5.1381	1.39×10^{-7}
CNTN4	3	2140497	3099645	2193	242	991	5.1346	1.41×10^{-7}
SNRK	3	43328004	43466256	94	18	991	4.9381	3.94×10^{-7}
SRGAP3	3	9022275	9404737	664	104	991	4.8999	4.80×10^{-7}
TRAPPC9	8	140742586	141468678	1202	126	991	4.837	6.59×10^{-7}
FHIT	3	59735036	61237133	3174	281	991	4.5959	2.15×10^{-6}

GRM7: glutamate metabotropic receptor 7; CNTN4: contactin 4; SNRK: SNF-related kinase; SRGAP3: SLIT-ROBO Rho GTPase-activating protein 3; TRAPPC9: trafficking protein particle complex subunit 9; FHIT: fragile histidine triad diadenosine triphosphatase.

channel openers, including FDA-approved drug retigabine (ezogabine), have been proven to normalise the connectivity of brain circuitry reversing the depressive symptomatology through the potentiation of active resilience mechanisms (Friedman et al. 2016; Tan et al. 2018).

CTNNA2 (OMIM 114025) is a large gene of ~1Mb, conserved across species, and is highly expressed in the central nervous system, especially in the prefrontal cortex, temporal lobe, cingulate cortex, hypothalamus, and amygdala (Terracciano et al. 2011). *CTNNA2* encodes for the catenin cadherin-associated protein alpha 2, a cell-adhesion protein that has been shown to regulate synaptic plasticity, acting as a linker between cadherins and the actin cytoskeleton and as such is important for maintaining the stability of dendritic spines and synaptic contact (Abe et al. 2004; Smith et al. 2005). The deletion of the homologous gene (*Catna2*) in knockout mice produced alterations in brain morphogenesis (Uemura and Takeichi 2006) and led to abnormal behaviours (Park et al. 2002), conditions reverted by the transgenic restoration of *Catna2* expression (Uemura and Takeichi 2006). Similarly, biallelic loss of human *CTNNA2*, deriving from truncating mutations, led to cortical neuronal migration defects (Schaffer et al. 2018). This evidence indicates a potential role for *CTNNA2* in the regulation of personality features. Along the same line, several GWASs revealed the association of *CTNNA2* with psychiatric disorders (Scott et al. 2009; Adkins et al. 2011; Goodbourn et al. 2014; Ehlers et al. 2016; Ryu et al. 2018) and related phenotypes (McGue et al. 2013; Johnston et al. 2019). Consistently, a GWAS on a primate model of infant inhibited temperament (Fox et al. 2021) showed a genome-wide significant hit near *CTNNA2*, a gene that maintained an association also after gene-level enrichment permutation analyses based on published human association studies for neuroticism (Nagel et al. 2018) and depression (Coleman et al. 2020), thus suggesting a shared evolutionarily-conserved mechanism. Furthermore, *CTNNA2* was previously reported to be proximal (within 250 kb) and in reasonably high linkage disequilibrium ($r^2 \geq 0.5$) with markers associated with MDD (Sullivan et al. 2009). Finally, copy number variations (CNVs) were found to be located within the *CTNNA2* gene in 5 MDD cases of a multigenerational pedigree derived from a Dagestan remote highland isolate containing an aggregation of MDD and suicides (Bulayeva et al. 2012). All this evidence may suggest a pleiotropic action of *CTNNA2*, supporting the concept that genes harbouring common variants associated with personality traits might harbour other variants associated with more extreme psychiatric phenotypes. For this reason, the development of drugs modulating the expression or function of CTNNA2 protein might be useful for a potential new therapeutic approach in psychiatric disorders including depression (Eszlari et al. 2021).

The gene-based analysis revealed six significantly associated genes, all involved in different physiological mechanisms that usually appear to be deregulated in MDD and affect human brain development. Among these genes, *GMR7*, *CNTN4*, and *SRGAP3* are all located on chromosome 3, precisely in a well-defined region that has been identified to be in genetic linkage with MDD in two independent linkage studies (Breen et al. 2011; Pergadia et al. 2011), with a maximum LOD score nearest the *GMR7* gene. This evidence suggests that this gene is one of the most accredited genes that can explain the linkage of such a broad region with MDD.

GRM7 (OMIM 604101) encodes for the metabotropic glutamate receptors mGlu7. It is highly expressed in the amygdala and hippocampus, and its modulation in these brain regions has been demonstrated to affect behavioural models of anxiety disorders or depression (Matson and Cervantes 2014). Results from GWA and candidate-gene studies on MDD revealed a significant association between *GRM7* SNPs and MDD (Li et al. 2016) or recurrent MDD (Muglia et al. 2010). The involvement of *GRM7* in MDD has been further proposed by transcriptome studies on human postmortem brain samples (Chang et al. 2014) and animal models (Sansig et al. 2001; Bushell et al. 2002; Cryan

et al. 2003; Callaerts-Vegh et al. 2006; Mitsukawa et al. 2006). In addition, pharmacogenetics investigations of this gene and antidepressant response have been conducted as well (Wierońska et al. 2007; Fabbri et al. 2013; Sun et al. 2019), thus highlighting its potential role as a therapeutic target in MDD.

The *CNTN4* (OMIM 607280) gene encodes for a neuronal cell adhesion immunoglobulin, probably involved in axon growth, guidance, and fasciculation in the central nervous system (CNS) (Zhang et al. 2008). Associations between *CNTN4* and MDD (Breen et al. 2011; Pergadia et al. 2011) or other psychiatric and neurodevelopmental disorders (Glessner et al. 2009; Cottrell et al. 2011; Guo et al. 2012) have been reported, suggesting an important role of this gene in normal and abnormal development of CNS.

The *SRGAP3* (OMIM 606525) gene encodes for a GAP protein, highly expressed in foetal and adult brain tissue, including the cortex and the hippocampus (Waltereit et al. 2008; Bacon et al. 2009). In a case report on a family with a history of schizophrenia and MDD, the authors identified a ~134 kb duplication spanning exons 2–4 of the *SRGAP3* gene that segregates with psychotic illness (Wilson et al. 2011). Recently, a GWAS on 1758 European adults with a ruminative response style (Eszlari et al. 2019), often associated with MDD (Nolen-Hoeksema and Watkins 2011), showed that top SNPs for rumination on chromosome 3 were associated with expression levels of *SRGAP3* in the dorsolateral prefrontal cortex (Eszlari et al. 2019). Therefore, the disruption/alteration of SRGAP3 may be considered, at least in part, etiologic for mental disorders.

The *FHIT* (OMIM 601153) gene encodes for a hydrolase involved in purine metabolism, expressed in multiple brain regions. Intronic variants in *FHIT* have been reported to be associated with a broad depression phenotype in two different GWAS meta-analyses (Direk et al. 2017; Amare et al. 2020). Moreover, Howard et al. (2019) identified, through a meta-analysis of 807,553 individuals from three large GWAS of depression, another intronic SNP of *FHIT* among 102 independent variants significantly associated with MDD.

The *TRAPPC9* (Trafficking Protein Particle Complex 9) (OMIM 611966) gene encodes for a protein that likely plays a role in NF-kappa-B signalling. *TRAPPC9* is highly expressed in the postmitotic neurons of the cerebral cortex and exerts an essential role in human brain development (Hu et al. 2005). Genetic variants of *TRAPPC9* have been described as causative of autosomal-recessive intellectual disability in families with different ethnic backgrounds (Mir et al. 2009; Mochida et al. 2009; Philippe et al. 2009; Marangi et al. 2013; Abbasi et al. 2017; Mortreux et al. 2018), also in comorbidity with ASD (Hnoonual et al. 2019; Wilton et al. 2020). Deficient mice developed a wide range of behavioural deficits and exhibited brain structural abnormalities (Ke et al. 2020).

The last gene found in the gene-based analysis, *SNRK* (OMIM 612760), encodes for SNF-related serine/threonine-protein kinase, involved in neuronal apoptosis (Thirugnanam and Ramchandran 2020). Although there are no studies referring to its involvement in psychiatric disorders, an association of *SNRK* with bipolar disorder and schizophrenia on the GWAS Catalog has been reported.

A weakness of our study is that we were unable to replicate the two genes found in the meta-analysis (*KCNQ5* and *CTNNA2*) by applying the gene-based approach, although their nominal *p*-values were <0.05. In the gene-based analysis, genetic data are aggregated to test the association of all markers within each gene with the phenotype. Specifically, MAGMA gene-based analysis uses multiple regression models to account for LD and catch the markers' combined effects. Thus, in addition to the impact of individual SNPs, other factors are involved, which could explain the lack of association.

Overall, our work highlighted eight genes as candidates for MDD. The relatively high number of identified genes is in line with the complex nature of MDD, in which the accumulation of the effect of individual variants and environmental factors contribute to the formation of the final phenotype.

The same complexity is observed in the results of GWAS for neuroticism, an important risk factor for psychiatric traits, including depression. Indeed, many researchers highlighted the difficulty to find and replicate signals associated with this phenotype (de Moor et al. 2015; Smith et al. 2016). On the other hand, many studies evidence the polygenic nature of neuroticism finding several associated genes (Nagel et al. 2018), although, to our knowledge, no overlapping genes with our results are present.

The study of isolated populations confirms its usefulness for identifying genes associated with complex traits, and in our case it allowed us to find seven variants in two genes associated with MDD in the SNPs-based GWAS. The lack of replication represents a weakness of our study, and a probable reason could be attributed to the different compositions of the replication sample compared to the discovery. Indeed, the replication sample is more heterogeneous both for

genetic and environmental aspects. On the other hand, the PRS analysis confirms that the accumulation of the effects of the variants discovered in our genetic isolation can predict MDD also in the outbreeding independent population. This finding is very important because it means that valid signals are involved in the meta-analysis results and confirms, although the difficulty to replicate each single gene, the polygenic nature of this disorder. However, given the multifactorial nature of this trait, other features (such as the interaction with other genes, environmental factors, etc.) should be taken into account to better understand the contribution of the identified genes. Nonetheless, our results contribute to elucidate the genetic bases of MDD in the perspective of developing personalised therapeutic strategies and preventive approaches for this disorder.

Acknowledgments

The authors wish to thank the staff and the study participants. The authors thank Veronika Collovati for the English revision of the manuscript.

Ethical approval

The study on the discovery cohorts was approved by the Ethics Committee of the IRCCS-Burlo Garofolo of Trieste (2007 242/07) and the Ethical Committee of the 'San Raffaele' Hospital and of the Piemonte Region. The study on the replication cohort was approved by the Local Ethics Committees (CEIOC IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia N: 50/2008 and Ethics Committee of the province of Verona N: 4997/09.11.01).

Consent

Participants of both cohorts were informed about the study and those who provided written consent were included, according to the recommendations of the declaration of Helsinki and the Italian DL no 675 of the 31 December 1996.

Author contributions

SU, AM, MG, PG, and MPC contributed to the conception and design of the study. VD, MPC, and EG performed the statistical analysis. SU and AM collected and reviewed the clinical data. MLB collected the genetic data. MB, SB, and AM recruited and evaluated the MDD patients of the replication sample. AM recruited and evaluated the controls of the replication sample. VD, AM, MPC, and SU wrote the manuscript. All authors read and approved the final manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the Italian Ministry of Health (GR-2010-2310715) to SU, Italian Ministry of Health (RC 01/21) to MPC, and supported by the NIHR Oxford Biomedical Research Centre.

ORCID

Vincenzo Dattilo http://orcid.org/0000-0001-8493-6188
Sheila Ulivi http://orcid.org/0000-0003-3606-835X
Alessandra Minelli http://orcid.org/0000-0002-9463-7808
Edoardo Giacopuzzi http://orcid.org/0000-0002-5692-366X
Massimo Gennarelli http://orcid.org/0000-0003-4192-945X
Paolo Gasparini http://orcid.org/0000-0002-0859-0856
Maria Pina Concas http://orcid.org/0000-0003-3598-2537

Data availability statement

A subset of the data is already available on the European Genome-phenome Archive (EGA) at the following links: FVG cohort: BAM files https://www.ebi.ac.uk/ega/studies/EGAS00001000252; sample list, vcf files https://www.ebi.ac.uk/ega/studies/EGAS00001001597; https://www.ebi.ac.uk/ega/datasets/EGAD00001002729; VBI cohort: BAM files https://www.ebi.ac.uk/ega/studies/EGAS00001000398; https://www.ebi.ac.uk/ega/studies/EGAS00001000458; A vcf file including all the INGI variants (SNPs and INDELs) with information on allele frequencies in the whole dataset and each cohort, has been submitted to the European Variation Archive (EVA), study accession number: PRJEB33648. The data is accessible at the following link: https://www.ebi.ac.uk/ena/data/view/PRJEB33648.

References

Abbasi AA, Blaesius K, Hu H, Latif Z, Picker-Minh S, Khan MN, Farooq S, Khan MA, Kaindl AM. 2017. Identification of a novel homozygous TRAPPC9 gene mutation causing non-syndromic intellectual disability, speech disorder, and secondary microcephaly. Am J Med Genet B Neuropsychiatr Genet. 174(8):839–845.

Abe K, Chisaka O, Van Roy F, Takeichi M. 2004. Stability of dendritic spines and synaptic contacts is controlled by alpha N-catenin. Nat Neurosci. 7(4):357–363.

Adkins DE, Aberg K, McClay JL, Bukszár J, Zhao Z, Jia P, Stroup TS, Perkins D, McEvoy JP, Lieberman JA, et al. 2011. Genomewide pharmacogenomic study of metabolic side effects to antipsychotic drugs. Mol Psychiatry. 16(3):321–332.

Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, Flicek P, Gabriel SB, et al. 2012. An integrated map of genetic variation from 1,092 human genomes. Nature. 491:56–65.

Amare AT, Vaez A, Hsu Y-H, Direk N, Kamali Z, Howard DM, McIntosh AM, Tiemeier H, Bültmann U, Snieder H, et al. 2020. Bivariate genome-wide association analyses of the broad depression phenotype combined with major depressive disorder, bipolar disorder or schizophrenia reveal eight novel genetic loci for depression. Mol Psychiatry. 25(7):1420–1429.

Amin N, Belonogova NM, Jovanova O, Brouwer RWW, van Rooij JGJ, van den Hout MCGN, Svishcheva GR, Kraaij R, Zorkoltseva IV, Kirichenko AV, et al. 2017. Nonsynonymous Variation in NKPD1 increases depressive symptoms in European populations. Biol Psychiatry. 81(8):702–707.

Amin N, Jovanova O, Adams HHH, Dehghan A, Kavousi M, Vernooij MW, Peeters RP, de Vrij FMS, van der Lee SJ, van Rooij JGJ, et al. 2017. Exome-sequencing in a large population-based study reveals a rare Asn396Ser variant in the LIPG gene associated with depressive symptoms. Mol Psychiatry. 22(4):537–543.

Amin N, de Vrij FMS, Baghdadi M, Brouwer RWW, van Rooij JGJ, Jovanova O, Uitterlinden AG, Hofman A, Janssen HLA, Darwish Murad S, et al. 2018. A rare missense variant in RCL1 segregates with depression in extended families. Mol Psychiatry. 23(5):1120–1126.

Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. 2007. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 23(10):1294–1296.

Bacon C, Endris V, Rappold G. 2009. Dynamic expression of the Slit-Robo GTPase activating protein genes during development of the murine nervous system. J Comp Neurol. 513(2):224–236.

Baculis BC, Zhang J, Chung HJ. 2020. The role of Kv7 channels in neural plasticity and behavior. Front Physiol. 11: 568667.

Baird DA, Liu JZ, Zheng J, Sieberts SK, Perumal T, Elsworth B, Richardson TG, Chen C-Y, Carrasquillo MM, Allen M, et al. 2021. Identifying drug targets for neurological and psychiatric disease via genetics and the brain transcriptome. PLOS Genet. 17(1):e1009224.

Breen G, Webb BT, Butler AW, van den Oord EJCG, Tozzi F, Craddock N, Gill M, Korszun A, Maier W, Middleton L, et al. 2011. A genome-wide significant linkage for severe depression on chromosome 3: The Depression Network study. Am J Psychiatry. 168(8):840–847.

Brown DA, Passmore GM. 2009. Neural KCNQ (Kv7) channels. Br J Pharmacol. 156(8):1185–1195.

Bulayeva K, Lencz T, Glatt S, Takumi T, Gurgenova F, Kawakami H, Bulayev O. 2012. [Mapping genes related to early onset major depressive disorder in Dagestan genetic isolates]. Turk Psikiyatri Derg. 23:161–170.

Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Patterson N, Daly MJ, Price AL, Neale BM. 2015. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 47(3): 291–295.

Bushell TJ, Sansig G, Collett VJ, van der Putten H, Collingridge GL. 2002. Altered short-term synaptic plasticity in mice lacking the metabotropic glutamate receptor mGlu7. ScientificWorldJournal. 2:730–737.

Cai N, Bigdeli TB, Kretzschmar W, Lei Y, Liang J, Song L, Hu J, Li Q, Jin W, Hu Z, et al. 2015. Sparse whole-genome sequencing identifies two loci for major depressive disorder. Nature. 523:588–591.

Callaerts-Vegh Z, Beckers T, Ball SM, Baeyens F, Callaerts PF, Cryan JF, Molnar E, D'Hooge R. 2006. Concomitant deficits in working memory and fear extinction are functionally dissociated from reduced anxiety in metabotropic glutamate receptor 7-deficient mice. J Neurosci. 26(24): 6573–6582.

Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 4:7.

Chang LC, Jamain S, Lin CW, Rujescu D, Tseng GC, Sibille E. 2014. A conserved BDNF, glutamate- and GABA-enriched gene module related to human depression identified by coexpression meta-analysis and DNA variant genome-wide association studies. PLoS One. 9(3):e90980.

Charlesworth B. 2009. Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. Nat Rev Genet. 10(3):195–205.

Choi SW, O'Reilly PF. 2019. PRSice-2: Polygenic Risk Score software for biobank-scale data. Gigascience. 8(7):giz082.

Cocca M, Barbieri C, Concas MP, Robino A, Brumat M, Gandin I, Trudu M, Sala CF, Vuckovic D, Girotto G, et al. 2020. A bird's-eye view of Italian genomic variation through whole-genome sequencing. Eur J Hum Genet. 28(4):435–444.

Coleman JRI, Gaspar HA, Bryois J, Byrne EM, Forstner AJ, Holmans PA, de Leeuw CA, Mattheisen M, McQuillin A, Whitehead Pavlides JM, et al. 2020. The genetics of the mood disorder spectrum: genome-wide association analyses of more than 185,000 cases and 439,000 controls. Biol Psychiatry. 88(2):169–184.

Colonna V, Pistis G, Bomba L, Mona S, Matullo G, Boano R, Sala C, Viganò F, Torroni A, Achilli A, et al. 2013. Small effective population size and genetic homogeneity in the Val Borbera isolate. Eur J Hum Genet. 21(1):89–94.

Cottrell CE, Bir N, Varga E, Alvarez CE, Bouyain S, Zernzach R, Thrush DL, Evans J, Trimarchi M, Butter EM, et al. 2011. Contactin 4 as an autism susceptibility locus. Autism Res. 4(3):189–199.

Cryan JF, Kelly PH, Neijt HC, Sansig G, Flor PJ, Van Der Putten H. 2003. Antidepressant and anxiolytic-like effects in mice lacking the group III metabotropic glutamate receptor mGluR7. Eur J Neurosci. 17(11):2409–2417.

Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, Vrieze SI, Chew EY, Levy S, McGue M, et al. 2016. Next-generation genotype imputation service and methods. Nat Genet. 48(10):1284–1287.

de Leeuw CA, Mooij JM, Heskes T, Posthuma D. 2015. MAGMA: generalized gene-set analysis of GWAS data. PLOS Comput Biol. 11:e1004219.

de Moor MHM, Van Den Berg SM, Verweij KJH, Krueger RF, Luciano M, Arias Vasquez A, Matteson LK, Derringer J, Esko T, Amin N, et al. 2015. Meta-analysis of genome-wide association studies for neuroticism, and the polygenic association with major depressive disorder. JAMA Psychiatry. 72:642–650.

Delaneau O, Marchini J, Zagury JF. 2011. A linear complexity phasing method for thousands of genomes. Nat Methods. 9(2):179–181.

Direk N, Williams S, Smith JA, Ripke S, Air T, Amare AT, Amin N, Baune BT, Bennett DA, Blackwood DH, et al. 2017. An analysis of two genome-wide association meta-analyses

identifies a new locus for broad depression phenotype. Biol Psychiatry. 82(5):322–329.
Duncan CE, Chetcuti AF, Schofield PR. 2008. Coregulation of genes in the mouse brain following treatment with clozapine, haloperidol, or olanzapine implicates altered potassium channel subunit expression in the mechanism of antipsychotic drug action. Psychiatr Genet. 18(5):226–239.
Ehlers CL, Gizer IR, Bizon C, Slutske W, Peng Q, Schork NJ, Wilhelmsen KC. 2016. Single nucleotide polymorphisms in the REG-CTNNA2 region of chromosome 2 and NEIL3 associated with impulsivity in a Native American sample. Genes Brain Behav. 15(6):568–577.
Esko T, Mezzavilla M, Nelis M, Borel C, Debniak T, Jakkula E, Julia A, Karachanak S, Khrunin A, Kisfali P, et al. 2013. Genetic characterization of northeastern Italian population isolates in the context of broader European genetic diversity. Eur J Hum Genet. 21(6):659–665.
Eszlari N, Bagyura Z, Millinghoffer A, Nagy T, Juhasz G, Antal P, Merkely B, Bagdy G. 2021. Catenin alpha 2 may be a biomarker or potential drug target in psychiatric disorders with perseverative negative thinking. Pharmaceuticals. 14: 850.
Eszlari N, Millinghoffer A, Petschner P, Gonda X, Baksa D, Pulay AJ, Réthelyi JM, Breen G, Deakin JFW, Anta P, et al. 2019. Genome-wide association analysis reveals KCTD12 and miR-383-binding genes in the background of rumination. Transl Psychiatry. 9(1):119.
Fabbri C, Drago A, Serretti A. 2013. Early antidepressant efficacy modulation by glutamatergic gene variants in the STAR*D. Eur Neuropsychopharmacol. 23(7):612–621.
Fidzinski P, Korotkova T, Heidenreich M, Maier N, Schuetze S, Kobler O, Zuschratter W, Schmitz D, Ponomarenko A, Jentsch JT. 2015. KCNQ5 K + channels control hippocampal synaptic inhibition and fast network oscillations. Nat Commun. 6:6254.
Fox AS, Harris RA, Rosso LD, Raveendran M, Kamboj S, Kinnally EL, Capitanio JP, Rogers J. 2021. Infant inhibited temperament in primates predicts adult behavior, is heritable, and is associated with anxiety-relevant genetic variation. Mol Psychiatry. 26(11):6609–6618.
Friedman AK, Juarez B, Ku SM, Zhang H, Calizo RC, Walsh JJ, Chaudhury D, Zhang S, Hawkins A, Dietz DM, et al. 2016. KCNQ channel openers reverse depressive symptoms via an active resilience mechanism. Nat Commun. 7:11671.
Gilling M, Rasmussen HB, Calloe K, Sequeira AF, Baretto M, Oliveira G, Almeida J, Lauritsen MB, Ullmann R, Boonen SE, et al. 2013. Dysfunction of the heteromeric KV7.3/KV7.5 potassium channel is associated with autism spectrum disorders. Front Genet. 4:54.
Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S, Zhang H, Estes A, Brune CW, Bradfield JP, et al. 2009. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. Nature. 459(7246):569–572.
GTEX Consortium. 2015. The genotype-tissue expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. 348:648.
Goodbourn PT, Bosten JM, Bargary G, Hogg RE, Lawrance-Owen AJ, Mollon JD. 2014. Variants in the 1q21 risk region are associated with a visual endophenotype of autism and schizophrenia. Genes Brain Behav. 13(2):144–151.
Guo H, Xun G, Peng Y, Xiang X, Xiong Z, Zhang L, He Y, Xu X, Liu Y, Lu L, et al. 2012. Disruption of Contactin 4 in two subjects with autism in Chinese population. Gene. 505(2): 201–205.
Hamilton M. 1960. A rating scale for depression. J Neurol Neurosurg Psychiatry. 23:56–62.
Hjaeresen M-L, Hageman I, Wortwein G, Plenge P, Jørgensen MB. 2008. Chronic electroconvulsive stimulation but not chronic restraint stress modulates mRNA expression of voltage-dependent potassium channels Kv7.2 and Kv11.1 in the rat piriform cortex. Brain Res. 1217:179–184.
Hnoonual A, Graidist P, Kritsaneepaiboon S, Limprasert P. 2019. Novel compound heterozygous mutations in the TRAPPC9 gene in two siblings with autism and intellectual disability. Front Genet. 10:61.
Howard DM, Adams MJ, Clarke T-K, Hafferty JD, Gibson J, Shirali M, Coleman JRI, Hagenaars SP, Ward J, Wigmore EM, et al. 2019. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci. 22(3):343–352.
Howard DM, Adams MJ, Shirali M, Clarke TK, Marioni RE, Davies G, Coleman JRI, Alloza C, Shen X, Barbu MC, et al. 2018. Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. Nat Commun. 9:1–10.
Howie BN, Donnelly P, Marchini J. 2009. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLOS Genet. 5(6): e1000529.
Hu W-H, Pendergast JS, Mo X-M, Brambilla R, Bracchi-Ricard V, Li F, Walters WM, Blits B, He L, Schaal SM, et al. 2005. NIBP, a novel NIK and IKK(beta)-binding protein that enhances NF-(kappa)B activation. J Biol Chem. 280(32): 29233–29241.
Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR, Tung JY, Hinds DA, Perlis RH, Winslow AR, et al. 2016. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. Nat Genet. 48(9):1031–1036.
Johnston KJA, Adams MJ, Nicholl BI, Ward J, Strawbridge RJ, Ferguson A, McIntosh AM, Bailey MES, Smith DJ. 2019. Genome-wide association study of multisite chronic pain in UK Biobank. PLOS Genet. 15(6):e1008164.
Ke Y, Weng M, Chhetri G, Usman M, Li Y, Yu Q, Ding Y, Wang Z, Wang X, Sultana P, et al. 2020. Ppc9 deficiency in mice impairs learning and memory by causing imbalance of dopamine d1 and d2 neurons. Sci Adv. 6(47):eabb7781.
Kendall KM, Van AE, Andlauer TFM, Choi KW, Luykx JJ, Schulte EC, Lu Y. 2021. The genetic basis of major depression. Psychol Med. 51:2217–2230.
Kendler KS, Gardner CO, Neale MC, Prescott CA. 2001. Genetic risk factors for major depression in men and women: similar or different heritabilities and same or partly distinct genes? Psychol Med. 31(4):605–616.
Lehman A, Thouta S, Mancini GMS, Naidu S, van Slegtenhorst M, McWalter K, Person R, Mwenifumbo J, Salvarinova R, Guella I, et al. 2017. Loss-of-function and gain-of-function mutations in KCNQ5 cause intellectual disability or epileptic encephalopathy. Am J Hum Genet. 101(1):65–74.
Li W, Ju K, Li Z, He K, Chen J, Wang Q, Yang B, An L, Feng G, Sun W, et al. 2016. Significant association of GRM7 and GRM8 genes with schizophrenia and major depressive

disorder in the Han Chinese population. Eur Neuropsychopharmacol. 26(1):136–146.
Lloyd-Jones LR, Robinson MR, Yang J, Visscher PM. 2018. Transformation of summary statistics from linear mixed model association on all-or-none traits to odds ratio. Genetics. 208(4):1397–1408.
Marangi G, Leuzzi V, Manti F, Lattante S, Orteschi D, Pecile V, Neri G, Zollino M. 2013. TRAPPC9-related autosomal recessive intellectual disability: report of a new mutation and clinical phenotype. Eur J Hum Genet. 21(2):229–232.
Matson JL, Cervantes PE. 2014. Commonly studied comorbid psychopathologies among persons with autism spectrum disorder. Res Dev Disabil. 35(5):952–962.
McGue M, Zhang Y, Miller MB, Basu S, Vrieze S, Hicks B, Malone S, Oetting WS, Iacono WG. 2013. A genome-wide association study of behavioral disinhibition. Behav Genet. 43(5):363–373.
McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F. 2016. The Ensembl variant effect predictor. Genome Biol. 17(1):122.
Millon T, Davis RD. 1997. The MCMI-III: present and future directions. J Pers Assess. 68(1):69–85.
Mir A, Kaufman L, Noor A, Motazacker MM, Jamil T, Azam M, Kahrizi K, Rafiq MA, Weksberg R, Nasr T, et al. 2009. Identification of Mutations in TRAPPC9, which encodes the NIK- and IKK-beta-binding protein, in nonsyndromic autosomal-recessive mental retardation. Am J Hum Genet. 85(6):909–915.
Mitsukawa K, Mombereau C, Lötscher E, Uzunov DP, van der Putten H, Flor PJ, Cryan JF. 2006. Metabotropic glutamate receptor subtype 7 ablation causes dysregulation of the HPA axis and increases hippocampal BDNF protein levels: implications for stress-related psychiatric disorders. Neuropsychopharmacology. 31(6):1112–1122.
Mochida GH, Mahajnah M, Hill AD, Basel-Vanagaite L, Gleason D, Hill RS, Bodell A, Crosier M, Straussberg R, Walsh CA, et al. 2009. A truncating mutation of TRAPPC9 is associated with autosomal-recessive intellectual disability and postnatal microcephaly. Am J Hum Genet. 85(6): 897–902.
Mortreux J, Busa T, Germain DP, Nadeau G, Puechberty J, Coubes C, Gatinois V, Cacciagli P, Duffourd Y, Pinard J-M, et al. 2018. The role of CNVs in the etiology of rare autosomal recessive disorders: the example of TRAPPC9-associated intellectual disability. Eur J Hum Genet. 26(1): 143–148.
Muglia P, Tozzi F, Galwey NW, Francks C, Upmanyu R, Kong XQ, Antoniades A, Domenici E, Perry J, Rothen S, et al. 2010. Genome-wide association study of recurrent major depressive disorder in two European case-control cohorts. Mol Psychiatry. 15(6):589–601.
Mullins N, Lewis CM. 2017. Genetics of depression: progress at last. Curr Psychiatry Rep. 19:1–7.
Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J, Savage JE, Hammerschlag AR, Skene NG, Muñoz-Manchado AB, et al. 2018. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. Nat Genet. 50(7):920–927.
Nolen-Hoeksema S, Watkins ER. 2011. A heuristic for developing transdiagnostic models of psychopathology: explaining multifinality and divergent trajectories. Perspect Psychol Sci. 6(6):589–609.
Okbay A, Baselmans BML, De Neve J-E, Turley P, Nivard MG, Fontana MA, Meddens SFW, Linnér RK, Rietveld CA, Derringer J, et al. 2016. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. Nat Genet. 48(6):624–633.
Park C, Falls W, Finger JH, Longo-Guess CM, Ackerman SL. 2002. Deletion in Catna2, encoding alpha N-catenin, causes cerebellar and hippocampal lamination defects and impaired startle modulation. Nat Genet. 31(3): 279–284.
Pergadia ML, Glowinski AL, Wray NR, Agrawal A, Saccone SF, Loukola A, Broms U, Korhonen T, Penninx BWJH, Grant JD, et al. 2011. A 3p26-3p25 genetic linkage finding for DSM-IV major depression in heavy smoking families. Am J Psychiatry. 168(8):848–852.
Philippe O, Rio M, Carioux A, Plaza J-M, Guigue P, Molinari F, Boddaert N, Bole-Feysot C, Nitschke P, Smahi A, et al. 2009. Combination of linkage mapping and microarray-expression analysis identifies NF-kappaB signaling defect as a cause of autosomal-recessive mental retardation. Am J Hum Genet. 85(6):903–908.
Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. 2010. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 26(18):2336–2337.
Prunas A, Sarno I, Preti E, Madeddu F, Perugini M. 2012. Psychometric properties of the Italian version of the SCL-90-R: a study on a large community sample. Eur Psychiatry. 27(8):591–597.
Ryu E, Nassan M, Jenkins GD, Armasu SM, Andreazza A, McElroy SL, Vawter MP, Frye MA, Biernacka JM. 2018. A genome-wide search for bipolar disorder risk loci modified by mitochondrial genome variation. Mol Neuropsychiatry. 3(3):125–134.
Sansig G, Bushell TJ, Clarke VRJ, Rozov A, Burnashev N, Portet C, Gasparini F, Schmutz M, Klebs K, Shigemoto R, et al. 2001. Increased seizure susceptibility in mice lacking metabotropic glutamate receptor 7. J Neurosci. 21(22): 8734–8745.
Schaffer AE, Breuss MW, Caglayan AO, Al-Sanaa N, Al-Abdulwahed HY, Kaymakçalan H, Yılmaz C, Zaki MS, Rosti RO, Copeland B, et al. 2018. Biallelic loss of human CTNNA2, encoding αN-catenin, leads to ARP2/3 complex overactivity and disordered cortical neuronal migration. Nat Genet. 50(8):1093–1101.
Schroeder BC, Hechenberger M, Weinreich F, Kubisch C, Jentsch TJ. 2000. KCNQ5, a novel potassium channel broadly expressed in brain, mediates M-type currents. J Biol Chem. 275(31):24089–24095.
Scott LJ, Muglia P, Kong XQ, Guan W, Flickinger M, Upmanyu R, Tozzi F, Li JZ, Burmeister M, Absher D, et al. 2009. Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry. Proc Natl Acad Sci USA. 106(18):7501–7506.
Shadrina M, Bondarenko EA, Slominsky PA. 2018. Genetics factors in major depression disease. Front Psychiatry. 9: 334.
Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC. 1998. The

Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry. 59:22–33.
Smith A, Bourdeau I, Wang J, Bondy CA. 2005. Expression of catenin family members CTNNA1, CTNNA2, CTNNB1 and JUP in the primate prefrontal cortex and hippocampus. Brain Res Mol Brain Res. 135(1–2):225–231.
Smith DJ, Escott-Price V, Davies G, Bailey MES, Colodro-Conde L, Ward J, Vedernikov A, Marioni R, Cullen B, Lyall D, et al. 2016. Genome-wide analysis of over 106 000 individuals identifies 9 neuroticism-associated loci. Mol Psychiatry. 21(6):749–757.
Smolin B, Karry R, Gal-Ben-Ari S, Ben-Shachar D. 2012. Differential expression of genes encoding neuronal ion-channel subunits in major depression, bipolar disorder and schizophrenia: implications for pathophysiology. Int J Neuropsychopharmacol. 15(7):869–882.
Sullivan PF, de Geus EJC, Willemsen G, James MR, Smit JH, Zandbelt T, Arolt V, Baune BT, Blackwood D, Cichon S, et al. 2009. Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. Mol Psychiatry. 14(4):359–375.
Sun Q, Yuan F, Yuan R, Ren D, Zhu Y, Bi Y, Hu J, Guo Z, Xu F, Niu W, et al. 2019. GRIK4 and GRM7 gene may be potential indicator of venlafaxine treatment responses in Chinese of Han ethnicity. Medicine. 98(19):e15456.
Tan A, Costi S, Morris LS, Van Dam NT, Kautz M, Whitton AE, et al. 2018. Effects of the KCNQ channel opener ezogabine on functional connectivity of the ventral striatum and clinical symptoms in patients with major depressive disorder. Mol Psychiatry. 25(6):1323–1333.
Terracciano A, Esko T, Sutin AR, de Moor MHM, Meirelles O, Zhu G, Tanaka T, Giegling I, Nutile T, Realo A, et al. 2011. Meta-analysis of genome-wide association studies identifies common variants in CTNNA2 associated with excitement-seeking. Transl Psychiatry. 1:e49.
Thirugnanam K, Ramchandran R. 2020. SNRK: a metabolic regulator with multifaceted role in development and disease. Vessel Plus. 4:26.
Turner SD. 2018. qqman: An R package for visualizing GWAS results using Q-Q and Manhattan plots. J Open Source Softw. 3(25):731
Tzingounis AV, Heidenreich M, Kharkovets T, Spitzmaul G, Jensen HS, Nicoll RA, Jentsch TJ. 2010. The KCNQ5 potassium channel mediates a component of the afterhyperpolarization current in mouse hippocampus. Proc Natl Acad Sci USA. 107(22):10232–10237.
Uemura M, Takeichi M. 2006. Alpha N-catenin deficiency causes defects in axon migration and nuclear organization in restricted regions of the mouse brain. Dev Dyn. 235(9): 2559–2566.
Verbeek EC, Bevova MR, Hoogendijk WJG, Heutink P. 2014. The genetics of MDD – a review of challenges and opportunities.
Verma P, Shakya M. 2021. Transcriptomics and sequencing analysis of gene expression profiling for major depressive disorder. Indian J Psychiatry. 63(6):549–553.
Waltereit R, Kautt S, Bartsch D. 2008. Expression of MEGAP mRNA during embryonic development. Gene Expr Patterns. 8(5):307–310.
[WHO] World Health Organization. 2021. Depression.
Wierońska JM, Kłak K, Pałucha A, Brański P, Pilc A. 2007. Citalopram influences mGlu7, but not mGlu4 receptors' expression in the rat brain hippocampus and cortex. Brain Res. 1184:88–95.
Willer CJ, Li Y, Abecasis GR. 2010. METAL: Fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 26(17):2190–2191.
Wilson NKA, Lee Y, Long R, Hermetz K, Rudd MK, Miller R, Rapoport JL, Addington AM.,. 2011. A novel microduplication in the neurodevelopmental gene SRGAP3 that segregates with psychotic illness in the family of a COS proband. Case Rep Genet. 2011:1–5.
Wilton KM, Gunderson LB, Hasadsri L, Wood CP, Schimmenti LA. 2020. Profound intellectual disability caused by homozygous TRAPPC9 pathogenic variant in a man from Malta. Mol Genet Genomic Med. 8(5):e1211.
Wray NR, Lee SH, Mehta D, Vinkhuyzen AAE, Dudbridge F, Middeldorp CM. 2014. Research review: polygenic methods and their application to psychiatric traits. J Child Psychol Psychiatry. 55(10):1068–1087.
Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, Adams MJ, Agerbo E, Air TM, Andlauer TMF, et al. 2018. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet. 50(5):668–681.
Zhang C, Rong H. 2019. Genetic advance in depressive disorder. Adv Exp Med Biol. 1180:19–57.
Zhang Y, Yeh J, Richardson PM, Bo X. 2008. Cell adhesion molecules of the immunoglobulin superfamily in axonal regeneration and neural repair. Restor Neurol Neurosci. 26(2–3):81–96.