

Pharmacogenomics of antidepressant drugs: perspectives for the personalization of treatment in depression

Citation for published version (APA):

Fabbri, C. (2018). *Pharmacogenomics of antidepressant drugs: perspectives for the personalization of treatment in depression: Genetic biomarkers of drug treatment outcomes in depression*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20181114cf>

Document status and date:

Published: 01/01/2018

DOI:

[10.26481/dis.20181114cf](https://doi.org/10.26481/dis.20181114cf)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

**PHARMACOGENOMICS OF ANTIDEPRESSANT DRUGS:
PERSPECTIVES FOR THE PERSONALIZATION OF
TREATMENT IN DEPRESSION**

Genetic biomarkers of drug treatment outcomes in depression

To obtain the degree of Doctor at the Maastricht University,
on the authority of the Rector Magnificus, Prof. Dr. Rianne M. Letschert,
in accordance with the decision of the Board of Deans,

To be defended in public on
Wednesday 14th of November 2018 at 14:00
by Chiara Fabbri

Supervisor: Prof. dr. Koen Schruers

Co-supervisor: Prof. dr. Alessandro Serretti, University of Bologna

Assessment Committee:

- Prof. dr. B.P.F. Rutten (Chairman);
- Prof. dr. S. Claes, Katholieke Universiteit Leuven, België;
- Prof. dr. F. P.M.L. Peeters;
- Prof. dr. G. Perna, Hermanas Hospitalarias, Villa San Benedetto Menni Hospital, Como, Italy and Miami University, Miami, United States of America;
- Prof. dr. W. Riedel.

Cover image: *Garden of Melancholia* by Mike Worrall

Contents

1. General Introduction	4
Part 1: Candidate genes involved in antidepressant pharmacokinetics and pharmacodynamics	19
2. Effect of cytochrome CYP2C19 metabolizing activity on antidepressant response and side effects: meta-analysis of data from genome-wide association studies.....	19
3. Pleiotropic genes in psychiatry: calcium channels and the stress-related FKBP5 gene in antidepressant response and resistance.....	38
Part 2: Genome-wide association studies of antidepressant response	59
4. New insights on the pharmacogenomics of antidepressant response from the GENDEP and STAR*D studies: rare variant analysis and high-density imputation.....	59
Part 3: Multi-marker tests as innovative approach to antidepressant pharmacogenomics	82
5. PPP3CC gene: a putative modulator of antidepressant response through the B-cell receptor signaling pathway.....	82
6. Neuronal cell adhesion genes and antidepressant response in three independent samples.....	107
7. Genetics of long-term treatment outcome in bipolar disorder.....	132
8. Summary	155
9. General Discussion	158
10. Valorization chapter	174
Acknowledgments	178

Chapter 1

General Introduction

Background

The high life-time prevalence (~13%) and high morbidity of major depressive disorder (MDD), in conjunction with unsatisfactory response rates to antidepressant medications (complete remission in only ~1/3 of patients), constitute an enormously heavy personal and socio-economic burden (Otte et al., 2016). In 2015 depressive disorders became the third cause of years lived with disability on a global scale (GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, 2016). Impaired functioning and reduced quality of life in MDD is comparable to that of observed in severe physical disorders such as arthritis and heart disease (Buist-Bouwman et al., 2006). Treatment-resistant depression (TRD) is usually defined as lack of response to at least two treatments (30% to 40% of major depression episodes) and affected patients continuously suffer from symptoms and their sequelae (Schosser et al., 2012).

Over 30 different antidepressant medications are available and several antidepressant combinations or augmentation with other drugs are suggested by guidelines in case of TRD. The broad variability in therapeutic options reflects the hypothesis that different biological dysfunctions may be implicated in different patients, but no biomarker is currently available to guide clinicians' choice of treatment.

Common genetic variants were estimated to explain 42% of individual differences in antidepressant response (Tansey et al., 2013). Despite this relevant contribution, the complex interactions among the variants involved and the contribution of rare variants prevented the identification of valid and reproducible genetic markers that could be translated in clinical applications.

The identification of valid genetic markers of treatment response and TRD is expected to substantially reduce the time to identify the best treatment strategy since clinicians could implement second level treatments from the beginning in patients at risk of TRD and know if a patient may be resistant to some antidepressant classes. In other words, genetic biomarkers of antidepressant efficacy would facilitate the development of new evidence-based treatments guided by genetic polymorphisms of the individual patient. A number of pharmacogenetic tests have been already commercialized and they can be bought on the Internet, they all combine a number of polymorphisms involved in antidepressant pharmacodynamics and pharmacokinetics (Fabbri et al., 2016). The major issue of these tests is the lack of demonstration of cost/effectiveness independently from the producing companies and this explains the reason why they are not currently

reimbursed by any national health system. Indeed the complete list of the polymorphisms included and the algorithms used to classify drugs in recommended or not recommended are not publicly available. The inconsistent findings reported for the greatest part of polymorphisms in candidate genes included in commercial pharmacogenetic tests makes doubtful their validity in predicting antidepressant response (Niitsu et al., 2013), at least until the contrary is demonstrated by independent investigators. In any case, the complex polygenic nature of treatment response in depression requires also the use of approaches complementary to candidate gene studies, such as genome-wide association studies (GWAS) and genome sequencing, since multiple and probably partly unknown polymorphisms within pathways involved in antidepressant action are hypothesized to modulate the phenotype.

Treatment response to antidepressant medications and other psychotropic medications is known to be different between MDD and bipolar disorder (BD), in terms of clinical effects and underlying biological mechanisms (Citrome, 2014). Anyway, a similar polygenic architecture of treatment efficacy has been hypothesized in BD (Fabbri and Serretti, 2016), thus the same issues and the same possible strategies to deal with them can be assumed. Fewer pharmacogenetic studies have investigated treatment response in BD compared to MDD, and the most part of them were focused on the prophylactic efficacy of lithium (Song et al., 2016; Hou et al., 2016; Squassina et al., 2011), leaving largely unclear the genetic variants involved in response to other medications. Pharmacogenetic studies of antidepressant efficacy included both unipolar and bipolar depressed patients in many cases (Niitsu et al., 2013), despite the limitation of this approach as outlined above. Thus, there is no evidence about which polymorphisms/genes/pathways may be specifically involved in antidepressant efficacy in BD compared to MDD.

This dissertation aimed to discuss new strategies to study the pharmacogenetics of antidepressant medications in MDD, while one study (**chapter 7**) illustrated an alternative approach applied to study the genetics of treatment outcome in bipolar patients treated in a naturalistic setting (combinations of mood stabilizers, antidepressants and antipsychotics), since the similar genetic architecture of treatment efficacy in MDD and BD.

The genetics of antidepressant response

The hypothesis that antidepressant response may be a hereditary trait was formulated thanks to the observation that this phenotype clustered in families (Franchini et al., 1998). Accordingly, it is part of common good clinical practice to inquire antidepressant response in relatives before choosing which antidepressant

to prescribe to a patient. More recently, the variance in response explained by common polymorphisms was estimated to be around 42% using Genome-Wide Complex Trait Analysis (GCTA) that quantify the proportion of phenotypic variance explained by the genetic relatedness matrix with a mixed-effect model (Tansey et al., 2013). This estimation did not take into account rare and unknown polymorphisms, the contribution of which is still unclear. Previous studies had two types of approach to the study of the polymorphisms involved in antidepressant response: candidate gene studies and genome-wide association studies (GWAS), or combinations of these two.

Candidate genes: antidepressant pharmacodynamics and pharmacokinetics

Candidate gene studies examined the effect of polymorphisms in genes that are involved in the mechanisms of antidepressant action (pharmacodynamics) or antidepressant metabolism and transport (pharmacokinetics) according to the findings of pre-clinical studies. This approach has been applied since the 1990s and one of the first studied genes was the serotonin transporter gene (SLC6A4) because its product is the main target of selective serotonin reuptake inhibitors (SSRIs). Some functional polymorphisms within SLC6A4 have been object of multiple studies and meta-analyses. A 44 bp insertion/deletion (5-HTTLPR) in the promoter of the gene can carry the 16-repeat long (L) allele, which is associated with a level of basal gene expression that is twice compared to the 14-repeat short (S) allele (Heils et al., 1996). The S allele was associated with several psychiatric disorders with affective symptomatology and with pathological behaviors and personality traits related to anxiety, impulsivity, and stress (Serretti et al., 2006). The greatest part of studies reported that the L allele or LL genotype were associated with better antidepressant outcome, especially in Caucasians and during SSRI treatment, while results in Asians were more controversial (Porcelli et al., 2012). A single nucleotide polymorphism (SNP), namely rs25531, was reported within the 5-HTTLPR sequence and it was hypothesized to influence the functional effect of 5-HTTLPR itself (Hu et al., 2006). Accordingly, the rs25531 G variant in conjunction with the L allele may result in a reduced expression of SLC6A4, equivalent to that conferred by the S allele (Hu et al., 2006), but studies considering the haplotype 5-HTTLPR/rs25531 mostly reported negative findings (Fabbri et al., 2013). Polymorphisms in the SLC6A4 gene were without doubt the most studied using the candidate gene approach, but other genes involved in monoaminergic neurotransmission (HTR1A, HTR2A, HTR6, TPH1, COMT, SLC6A2), neuroplasticity (BDNF), second messenger cascades (GNB3) and

corticosteroid signaling (FKBP5) were object of particular interest because their functions are linked to antidepressant pharmacodynamics (Niitsu et al., 2013).

Variants in genes coding for several subtypes of serotonin receptors (HTR1A, HTR2A, HTR6) were studied as modulators of antidepressant response since the coded receptors are targets of several antidepressant medications (e.g. trazodone, mirtazapine, vortioxetine) and some antipsychotics with antidepressant properties (e.g. lurasidone, olanzapine, quetiapine). Among serotonin receptor genes, the most promising results were obtained for HTR2A (serotonin receptor 2A), indeed rs6311, rs6313 and rs7997012 in this gene were associated with antidepressant efficacy by the most recent meta-analyses (Lin et al., 2014; Niitsu et al., 2013). A role of this subtype of serotonin receptor in antidepressant action is suggested also by pre-clinical studies. In HTR2A *-/-* mutants SSRIs failed to elicit a progressive increase in the firing rate of serotonergic neurons with decreased ability to stimulate hippocampal plasticity and to produce antidepressant-like activities (Qesseveur et al., 2016).

TPH1 (tryptophan hydroxylase 1) and COMT (catechol-O-methyltransferase) are involved in the main steps of serotonin synthesis and monoamines metabolism, respectively, with a modulatory effect on the levels of monoamines and their metabolites in the central nervous system (CNS). According to the monoaminergic hypothesis, depression is caused by a decreased monoaminergic function in the CNS and the currently available antidepressant molecules work on the basis of this theory (Porcelli et al., 2011a). Despite their potential role, pharmacogenetic studies that investigated polymorphisms in TPH1 and COMT yielded mainly no evidence of association with antidepressant efficacy (Niitsu et al., 2013). SLC6A2 codes for the norepinephrine transporter that is one of the main targets of serotonin and noradrenaline reuptake inhibitors (SNRIs), tricyclic antidepressants (in particular desipramine, nortriptyline, trimipramine) and some antipsychotics with antidepressant properties (quetiapine). Results pertaining variants in this gene were contradictory and the most recent meta-analysis suggested no relevant effects (Niitsu et al., 2013).

BDNF (brain derived neurotrophic factor) gene codes for one of the main known neurotrophic factors and it plays a pivotal function in promoting neuronal growth and survival. Several gene expression studies demonstrated that antidepressant response is associated with an increase in BDNF expression in peripheral blood cells (Fabbri et al., 2016). The BDNF Valine66Methionine (Val66Met) SNP (rs6265) codes for the replacement of an evolutionarily conserved valine with a methionine at position 66 in the BDNF protein and it has been the most studied polymorphisms in this gene because of its functional consequences. *In vitro*

analysis of BDNF Val66Met has demonstrated that the variant Met protein is less efficiently targeted by the secretory pathway than the Val protein, resulting in decreased secretion of the protein. In humans, the rs6265 Met allele has been associated with reduced hippocampal volume, reduced cognitive performance, and psychiatric disorders including mood disorders and anxiety despite none of these associations has been consistently replicated (Glatt and Lee, 2016). Pharmacogenetic studies mainly reported better response in the heterozygote genotype, particularly in Asian populations (Niitsu et al., 2013). The result can be explained by animal models showing that although BDNF exerts an antidepressant effect, BDNF overexpression may increase stress-induced plasticity in the amygdala and increase anxiety-like behaviors (Govindarajan et al., 2006).

GNB3 (guanine nucleotide binding protein beta polypeptide 3) is involved in the generation of second messengers in response to a number of signals, such as neurotransmitters and hormones, that are hypothesized to modulate mechanisms involved in neuronal plasticity. The T allele of the rs5443 (C825T) polymorphism leads to a splice variant that results in altered activity of the protein (Ruiz-Velasco and Ikeda, 2003). Pharmacogenetic studies mostly reported a better antidepressant response or higher symptom improvement in T-allele carriers; however, controversial results have been also reported (Niitsu et al., 2013). The association between this variant and antidepressant response/remission was confirmed in subjects of Asian ancestry by the most recent meta-analyses (Niitsu et al., 2013; Hu et al., 2015).

FKBP5 (FK506 binding protein 51) encodes for a co-chaperone of hsp90 that modulates glucocorticoid receptor (GR) sensitivity. When FKBP5 is bound to the receptor complex, cortisol binds with lower affinity and nuclear translocation of the receptor occurs to a less extent (Wochnik et al., 2005). Several polymorphisms in this gene (rs3800373, rs1360780, rs9296158, rs9470080) were associated with enhanced expression of FKBP5 that results in increased GR resistance and decreased efficiency of the negative feedback of the stress hormone axis (Binder et al., 2004). Meta-analytical findings suggested that rs1360780 and rs3800373 may be the most promising polymorphisms in this gene for association with antidepressant response (Niitsu et al., 2013), as confirmed by a recent pilot study showing that rs1360780 may influence the effect of an algorithm-based antidepressant treatment (Stamm et al., 2016).

Finally, genes responsible for antidepressant metabolism and transport are worth of some description. Cytochrome P450 (CYP) enzymes are the major enzymes responsible for Phase I reactions in the metabolism of several substances, including antidepressant medications. CYP2D6, CYP2C19 and CYP2C9 are the isoforms

most studied in the prediction of antidepressant response. As a consequence of functional polymorphisms in these genes, the coded enzymes show different level of metabolizing activity, defining distinct functional groups. The wild type genotype results in the extensive metabolizer (EM) phenotype, which is characterized by the presence of two active alleles, whereas intermediate metabolizer (IM) is characterized by the presence of one wild-type allele plus a partially or totally defective allele and it is expected to be between EMs and poor metabolizers (PMs). PMs show a combination of two partially or totally defective alleles. The ultrarapid metabolizer (UM) category is a consequence of multiple copies of the gene or functional polymorphisms associated with increased gene expression (Porcelli et al., 2011b). There is quite consolidated evidence supporting that CYP P450 level of activity is associated with the pharmacokinetics of the metabolized antidepressants (e.g. concentration, plasma half-life, oral clearance), while the association with clinical efficacy is more controversial (Porcelli et al., 2011b). Several reasons can explain these contradictory (and often negative) findings: 1) several antidepressants are metabolized by different CYP P450 isoforms and the interaction among functional variants in the corresponding genes were not taken into account; 2) peripheral measures of drug metabolism such as plasma concentration may not reflect CNS drug concentration in a linear way, and small effect sizes may have not been identified in previous studies that were usually performed on quite small samples; 3) the inhibitory/stimulating effects of drugs, foods and smoke may have confounded the results. A meta-analysis of existing studies could address part of these limitations.

ABCB1 codes for P-glycoprotein (P-gp), an ATP-dependent drug efflux pump that limits the uptake and accumulation of some lipophilic drugs into the brain, including a number of antidepressants (e.g. amitriptyline, nortriptyline, citalopram, venlafaxine, sertraline, trimipramine). rs2032582 and rs1045642 are two particularly interesting SNPs within ABCB1 because they alter P-gp expression and/or function (Eichelbaum et al., 2004). The role of rs2032582 was confirmed at meta-analytical level (Niitsu et al., 2013) and this SNP together with rs2235015 showed potential usefulness in adjusting antidepressant dose or guide antidepressant switch (Breitenstein et al., 2014).

As previously reported, the pharmacogenetic tests currently available for the prediction of antidepressant treatment outcomes include a combination of polymorphisms identified by candidate gene association studies. The most commonly included variants are within SLC6A4, serotonin receptor genes, BDNF, CYP2D6, CYP2C19 and COMT (Fabbri et al., 2016), but none of these tests has been validated outside the producing companies.

Genome-wide association studies (GWAS)

The introduction of chip-based microarray technology capable of interrogating thousands of polymorphisms within the whole genome has given fresh momentum to genetic association studies and ushered new design studies, namely genome-wide association studies (GWAS). Such advances have brought in a new term, pharmacogenomics, which reflects the shift from single genes to virtually the whole genome, accurately genotyping from 500,000 to over four million SNPs in a cost-effective manner. GWAS do not need *a priori* hypothesis, they overcome the problem that the mechanisms of antidepressant action are not completely known. Further, the most significantly associated polymorphisms in behavioral GWAS were found in noncoding sequence (Tak and Farnham, 2015) and the disruption in a gene or pathway function is due to the cumulative effect of multiple variants that can be studied only using genome-wide genotyping or sequencing. Thus, on one hand GWAS overcome some of the main limitations of candidate gene studies, on the other hand they have their own limitations: 1) microarrays include a limited number of pre-selected SNPs while over 39 millions were identified in the human genome (McCarthy et al., 2016); 2) statistical methods for the analysis of multi-marker genetic effects in large databases are not easy to develop and implement; 3) genuine but small effect sizes may not survive the strict correction for genome-wide significance in underpowered sample sizes; 4) GWAS are focused on the analysis of common variants, while the effect of rare and unknown polymorphisms should not be overlooked. Regarding the latter issue, rare variants in specific pathways may be responsible for relevant effects in common disorders as suggested by the observation that the heritability of several common traits remains partly unexplained considering common variants only (Kaakinen et al., 2017). Previous GWAS investigating antidepressant response have been penalized by the limitations listed above. As a consequence, very few findings survived multiple-testing correction and no one was replicated across independent samples. Meta-analyses of GWAS in Caucasian samples provided no genome-wide signals (Tansey et al., 2012) or an association with an intronic region of the myosin X (MYO10) gene that was not confirmed after genotyping leading to the conclusion that the finding was only an imputation artifact (GENDEP Investigators et al., 2013). The GWAS on the largest Asian sample did not report any genome-wide significant finding either, but a meta-analysis of this sample with a Caucasian cohort found some polymorphisms close to the significance threshold, in particular one SNP in the 5' upstream of the NRG1 gene (neuregulin-1, that is involved in brain development and it was associated with mental disorders) (Biernacka et al., 2015). A multi-stage GWAS in Korean patients reported a significant signal in

AUTS2, a gene that was repeatedly implicated in neurodevelopmental disorders including autism but also schizoaffective or bipolar affective disorder (Myung et al., 2015). Another GWAS in a Korean population did not provide genome-wide hits in single polymorphisms, but it replicated the role of the inorganic cation transmembrane transporter activity pathway in two independent samples (Cocchi et al., 2016). The results of this study are in line with the hypothesis that pathway analysis provides better power and greater chances to obtain independent replication compared to the analysis of single variants, as most extensively discussed in the next section.

The analysis of rare variants was highly overlooked by previous GWAS. Only two studies investigated rare variants using genome-wide arrays, and they reported negative findings (Tansey et al., 2014; O'Dushlaine et al., 2014). Few and not replicated data were obtained using exome sequencing (Tammiste et al., 2013).

Multi-marker tests using GWAS or sequencing data

Multi-marker tests or aggregated approaches test the effect of many variants in a gene or pathway instead of single variants. The power increase in an aggregated approach can be attributed to the reduction in the number of tests performed (variants from a genomic region of interest are collapsed into one score) and the capture of the cumulative effects of a number of variants (the disruption of a gene or pathway functioning is the result of cumulative effects of variants within it) (Li et al., 2017). Multi-marker tests were also reported to balance heterogeneity within and between samples, for example the effects of ancestry stratification or different genotyping rates are weakened when variants are grouped and analyzed in pathways instead of individually. Aggregated tests requires the use of specific methodological approaches to take into account the non-independence among variants in linkage disequilibrium (LD), the weight to assign to variants with different frequency, functional consequences or other characteristics, the higher random chance that genes or gene sets including a higher number of variants provide significant results if there is not a comparator term in the analysis. For the latter reason, gene size and gene density are usually included as covariates in gene analysis and the significance of a gene set should be compared to that of genes outside the set to obtain more robust results (de Leeuw et al., 2015). The weight to assign to each variant in a gene or gene set is another interesting point. The most common method takes into account the minor allele frequency (MAF) of each variant to assign more influence to SNPs with lower MAF (Ionita-Laza et al., 2013; de Leeuw et al., 2015), but other SNPs characteristics may be used to increase the power of aggregated tests. For example, the use of prior GWAS

results, functional annotation and conservation scores have been recently demonstrated to improve the prediction of multi-marker tests (Shi et al., 2016).

Previous pathway analysis suggested the role of pathways involved in mechanisms of neurogenesis and neural plasticity and inflammation/extracellular matrix remodeling in antidepressant efficacy (Hunter et al., 2013; Ising et al., 2009), but negative findings were also reported (Tansey et al., 2012). Other issues in addition to those listed above can affect the results of pathway analysis, one of them is the pathway database used as reference. For example, some pathway databases are curated manually by experts on the base of existing literature, others include *in silico* data, some databases integrate data from different sources, some types of signaling cascades (e.g. stress activated) are included in few databases only (Chowdhury and Sarkar, 2015).

An alternative approach to study the aggregated effect of a gene set on a phenotype is enrichment analysis. This method has an exploratory nature: instead of studying the effect of a gene set previously identified as a pathway in existing databases, it aims to identify if a group of genes of interest is enriched with members of a known pathway or group of related genes more frequently than what is expected by chance. In this case the gene set of interest includes the genes associated with the phenotype or the genes where associated polymorphisms are located. Pathway enrichment suggests the possible cellular and molecular mechanisms that mediate the effect of associated genes/polymorphisms on the phenotype. Different datasets (e.g. predicted protein interaction data, physical and genetic interaction data, pathway and molecular interaction data) can be integrated to determine relationships among query genes, assigning weights based on how well query genes interact among each other (e.g. at level of protein interaction) or reproduce GO (Gene Ontology) co-annotation patterns (Warde-Farley et al., 2010). The identified functional networks can be a very useful starting point for further analysis (e.g. polygenic risk scores, pathway analysis, machine learning models). GWAS and multi-marker tests performed on genome-wide data have not produced prototypes of clinical applications yet, differently from candidate gene studies.

Aim and Outline of this dissertation

The overall aim of this dissertation was to contribute to the identification of genetic markers of antidepressant efficacy that may help in the process leading to clinical applications.

Part 1 of this dissertation identifies strong candidate genes for involvement in antidepressant efficacy on the basis of previous literature and it applies innovative approaches to their study. Genome-wide data were used to extract functional

categorization of CYP2C19 (cytochrome P450 family 2 subfamily C member 19) and perform the first meta-analysis of CYP2C19 metabolizing activity in antidepressant treatment outcomes (*chapter 2*). The use of the candidate gene approach was integrated with the analysis of genome-wide data to identify the biological pathways that may modulate the risk of treatment-resistant depression (*chapter 3*).

Part 2 of this dissertation discusses new strategies to exploit genome-wide data in the study of antidepressant response. More specifically, we combined the study of common and rare variants, we maximized the coverage of genetic variation using high-density imputation, and we used for the first time a pathway meta-analysis to identify the biological mechanisms that may be involved across independent samples (*chapter 4*).

Part 3 of this dissertation will focus on multi-marker tests in the prediction of antidepressant efficacy. Emphasis here will lie on the use of pathway analysis to identify the signaling cascades involved in antidepressant action (*chapter 5* and *chapter 6*) and enrichment analysis to identify networks that may mediate the effect of associated genes with treatment response in patients with bipolar disorder (*chapter 7*).

Chapter 8 will summarize key findings of each study.

Chapter 9 will discuss the innovations of each study compared to previous literature and possible implications for treatment strategies and direction of future research.

Chapter 10 is about valorization and it will discuss the contribution of this dissertation to future research, the socio-economic value of the reported findings and their possible further exploitation.

References

Biernacka, J.M., Sangkuhl, K., Jenkins, G., Whaley, R.M., Barman, P., Batzler, A., Altman, R.B., Arolt, V., Brockmüller, J., Chen, C.H., et al. (2015). The International SSRI Pharmacogenomics Consortium (ISPC): a genome-wide association study of antidepressant treatment response. *Transl. Psychiatry* 5, e553.

Binder, E.B., Salyakina, D., Lichtner, P., Wochnik, G.M., Ising, M., Pütz, B., Papiol, S., Seaman, S., Lucae, S., Kohli, M.A., et al. (2004). Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. *Nat. Genet.* 36, 1319–1325.

Breitenstein, B., Scheuer, S., Pfister, H., Uhr, M., Lucae, S., Holsboer, F., Ising, M., and Brückl, T.M. (2014). The clinical application of ABCB1 genotyping in antidepressant treatment: a pilot study. *CNS Spectr.* 19, 165–175.

Buist-Bouwman, M.A., De Graaf, R., Vollebergh, W. a. M., Alonso, J., Bruffaerts, R., Ormel, J., and ESEMEd/MHEDEA 2000 Investigators (2006). Functional disability of mental disorders and comparison with physical disorders: a study among the general population of six European countries. *Acta Psychiatr. Scand.* 113, 492–500.

Chowdhury, S., and Sarkar, R.R. (2015). Comparison of human cell signaling pathway databases--evolution, drawbacks and challenges. *Database J. Biol. Databases Curation* 2015.

Citrome, L. (2014). Treatment of bipolar depression: making sensible decisions. *CNS Spectr.* 19 Suppl 1, 4-11; quiz 1-3, 12.

Cocchi, E., Fabbri, C., Han, C., Lee, S.-J., Patkar, A.A., Masand, P.S., Pae, C.-U., and Serretti, A. (2016). Genome-wide association study of antidepressant response: involvement of the inorganic cation transmembrane transporter activity pathway. *BMC Psychiatry* 16, 106.

Eichelbaum, M., Fromm, M.F., and Schwab, M. (2004). Clinical aspects of the MDR1 (ABCB1) gene polymorphism. *Ther. Drug Monit.* 26, 180–185.

Fabbri, C., and Serretti, A. (2016). Genetics of long-term treatment outcome in bipolar disorder. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 65, 17–24.

Fabbri, C., Di Girolamo, G., and Serretti, A. (2013). Pharmacogenetics of antidepressant drugs: An update after almost 20 years of research. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 162, 487–520.

Fabbri, C., Crisafulli, C., Calabrò, M., Spina, E., and Serretti, A. (2016). Progress and prospects in pharmacogenetics of antidepressant drugs. *Expert Opin. Drug Metab. Toxicol.* 1–12.

Franchini, L., Serretti, A., Gasperini, M., and Smeraldi, E. (1998). Familial concordance of fluvoxamine response as a tool for differentiating mood disorder pedigrees. *J. Psychiatr. Res.* 32, 255–259.

GBD 2015 Disease and Injury Incidence and Prevalence Collaborators (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Lond. Engl.* 388, 1545–1602.

GENDEP Investigators, MARS Investigators, and STAR*D Investigators (2013). Common genetic variation and antidepressant efficacy in major depressive disorder: a meta-analysis of three genome-wide pharmacogenetic studies. *Am. J. Psychiatry* 170, 207–217.

Glatt, C.E., and Lee, F.S. (2016). Common Polymorphisms in the Age of Research Domain Criteria (RDoC): Integration and Translation. *Biol. Psychiatry* 79, 25–31.

Govindarajan, A., Rao, B.S.S., Nair, D., Trinh, M., Mawjee, N., Tonegawa, S., and Chattarji, S. (2006). Transgenic brain-derived neurotrophic factor expression causes both anxiogenic and antidepressant effects. *Proc. Natl. Acad. Sci. U. S. A.* 103, 13208–13213.

Heils, A., Teufel, A., Petri, S., Stöber, G., Riederer, P., Bengel, D., and Lesch, K.P. (1996). Allelic variation of human serotonin transporter gene expression. *J. Neurochem.* 66, 2621–2624.

Hou, L., Heilbronner, U., Degenhardt, F., Adli, M., Akiyama, K., Akula, N., Arda, R., Arias, B., Backlund, L., Banzato, C.E.M., et al. (2016). Genetic variants associated with response to lithium treatment in bipolar disorder: a genome-wide association study. *Lancet Lond. Engl.* 387, 1085–1093.

Hu, Q., Zhang, S.-Y., Liu, F., Zhang, X.J., Cui, G.-C., Yu, E.-Q., Xu, X.-F., Li, P., Xiao, J.-Q., Wei, D.-M., et al. (2015). Influence of GNB3 C825T polymorphism on the efficacy of antidepressants in the treatment of major depressive disorder: A meta-analysis. *J. Affect. Disord.* 172, 103–109.

Hu, X.-Z., Lipsky, R.H., Zhu, G., Akhtar, L.A., Taubman, J., Greenberg, B.D., Xu, K., Arnold, P.D., Richter, M.A., Kennedy, J.L., et al. (2006). Serotonin transporter promoter gain-of-function genotypes are linked to obsessive-compulsive disorder. *Am. J. Hum. Genet.* 78, 815–826.

Hunter, A.M., Leuchter, A.F., Power, R.A., Muthén, B., McGrath, P.J., Lewis, C.M., Cook, I.A., Garriock, H.A., McGuffin, P., Uher, R., et al. (2013). A genome-wide association study of a sustained pattern of antidepressant response. *J. Psychiatr. Res.* 47, 1157–1165.

Ionita-Laza, I., Lee, S., Makarov, V., Buxbaum, J.D., and Lin, X. (2013). Sequence kernel association tests for the combined effect of rare and common variants. *Am. J. Hum. Genet.* 92, 841–853.

Ising, M., Lucae, S., Binder, E.B., Bettecken, T., Uhr, M., Ripke, S., Kohli, M.A., Hennings, J.M., Horstmann, S., Kloiber, S., et al. (2009). A genomewide association study points to multiple loci that predict antidepressant drug treatment outcome in depression. *Arch. Gen. Psychiatry* 66, 966–975.

Kaakinen, M., Mägi, R., Fischer, K., Heikkinen, J., Järvelin, M.-R., Morris, A.P., and Prokopenko, I. (2017). A rare-variant test for high-dimensional data. *Eur. J. Hum. Genet* 25, 988-994.

de Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* 11, e1004219.

Li, C.-I., Samuels, D.C., Zhao, Y.-Y., Shyr, Y., and Guo, Y. (2017). Power and sample size calculations for high-throughput sequencing-based experiments. *Brief. Bioinform.* [in press].

Lin, J.-Y., Jiang, M.-Y., Kan, Z.-M., and Chu, Y. (2014). Influence of 5-HT_{2A} genetic polymorphisms on the efficacy of antidepressants in the treatment of major depressive disorder: a meta-analysis. *J. Affect. Disord.* 168, 430–438.

McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang, H.M., Fuchsberger, C., Danecek, P., Sharp, K., et al. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* 48, 1279–1283.

Myung, W., Kim, J., Lim, S.-W., Shim, S., Won, H.-H., Kim, S., Kim, S., Lee, M.-S., Chang, H.S., Kim, J.-W., et al. (2015). A genome-wide association study of antidepressant response in Koreans. *Transl. Psychiatry* 5, e633.

Niitsu, T., Fabbri, C., Bentini, F., and Serretti, A. (2013). Pharmacogenetics in major depression: a comprehensive meta-analysis. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 45, 183–194.

O’Dushlaine, C., Ripke, S., Ruderfer, D.M., Hamilton, S.P., Fava, M., Iosifescu, D.V., Kohane, I.S., Churchill, S.E., Castro, V.M., Clements, C.C., et al. (2014). Rare copy number variation in treatment-resistant major depressive disorder. *Biol. Psychiatry* 76, 536–541.

Otte, C., Gold, S.M., Penninx, B.W., Pariante, C.M., Etkin, A., Fava, M., Mohr, D.C., and Schatzberg, A.F. (2016). Major depressive disorder. *Nat. Rev. Dis. Primer* 2, 16065.

Porcelli, S., Drago, A., Fabbri, C., and Serretti, A. (2011a). Mechanisms of antidepressant action: an integrated dopaminergic perspective. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 35, 1532–1543.

Porcelli, S., Fabbri, C., Spina, E., Serretti, A., and De Ronchi, D. (2011b). Genetic polymorphisms of cytochrome P450 enzymes and antidepressant metabolism. *Expert Opin. Drug Metab. Toxicol.* 7, 1101–1115.

Porcelli, S., Fabbri, C., and Serretti, A. (2012). Meta-analysis of serotonin transporter gene promoter polymorphism (5-HTTLPR) association with antidepressant efficacy. *Eur. Neuropsychopharmacol. J. Eur. Coll. Neuropsychopharmacol.* 22, 239–258.

Qesseveur, G., Petit, A.C., Nguyen, H.T., Dahan, L., Colle, R., Rotenberg, S., Seif, I., Robert, P., David, D., Guilloux, J.-P., et al. (2016). Genetic dysfunction of serotonin 2A receptor hampers response to antidepressant drugs: A translational approach. *Neuropharmacology* 105, 142–153.

Ruiz-Velasco, V., and Ikeda, S.R. (2003). A splice variant of the G protein beta 3-subunit implicated in disease states does not modulate ion channels. *Physiol. Genomics* 13, 85–95.

Schosser, A., Serretti, A., Souery, D., Mendlewicz, J., Zohar, J., Montgomery, S., and Kasper, S. (2012). European Group for the Study of Resistant Depression (GSRD)--where have we gone so far: review of clinical and genetic findings. *Eur. Neuropsychopharmacol. J. Eur. Coll. Neuropsychopharmacol.* 22, 453–468.

Serretti, A., Calati, R., Mandelli, L., and De Ronchi, D. (2006). Serotonin transporter gene variants and behavior: a comprehensive review. *Curr. Drug Targets* 7, 1659–1669.

Shi, J., Park, J.-H., Duan, J., Berndt, S.T., Moy, W., Yu, K., Song, L., Wheeler, W., Hua, X., Silverman, D., et al. (2016). Winner's Curse Correction and Variable Thresholding Improve Performance of Polygenic Risk Modeling Based on Genome-Wide Association Study Summary-Level Data. *PLoS Genet.* 12, e1006493.

Song, J., Bergen, S.E., Di Florio, A., Karlsson, R., Charney, A., Ruderfer, D.M., Stahl, E.A., Members of the International Cohort Collection for Bipolar Disorder (ICCBD), Chambert, K.D., Moran, J.L., et al. (2016). Genome-wide association study identifies *SESTD1* as a novel risk gene for lithium-responsive bipolar disorder. *Mol. Psychiatry* 21, 1290–1297.

Squassina, A., Manchia, M., Borg, J., Congiu, D., Costa, M., Georgitsi, M., Chillotti, C., Ardaù, R., Mitropoulos, K., Severino, G., et al. (2011). Evidence for association of an *ACCN1* gene variant with response to lithium treatment in Sardinian patients with bipolar disorder. *Pharmacogenomics* 12, 1559–1569.

Stamm, T.J., Rampp, C., Wiethoff, K., Stingl, J., Mössner, R., O Malley, G., Ricken, R., Seemüller, F., Keck, M., Fisher, R., et al. (2016). The *FKBP5* polymorphism rs1360780 influences the effect of an algorithm-based antidepressant treatment and is associated with remission in patients with major depression. *J. Psychopharmacol. Oxf. Engl.* 30, 40–47.

Tak, Y.G., and Farnham, P.J. (2015). Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics Chromatin* 8, 57.

Tammiste, A., Jiang, T., Fischer, K., Mägi, R., Krjutškov, K., Pettai, K., Esko, T., Li, Y., Tansey, K.E., Carroll, L.S., et al. (2013). Whole-exome sequencing identifies a polymorphism in the *BMP5* gene associated with SSRI treatment response in major depression. *J. Psychopharmacol. Oxf. Engl.* 27, 915–920.

Tansey, K.E., Guipponi, M., Perroud, N., Bondolfi, G., Domenici, E., Evans, D., Hall, S.K., Hauser, J., Henigsberg, N., Hu, X., et al. (2012). Genetic predictors of response to serotonergic and noradrenergic antidepressants in major depressive disorder: a genome-wide analysis of individual-level data and a meta-analysis. *PLoS Med.* 9, e1001326.

Tansey, K.E., Guipponi, M., Hu, X., Domenici, E., Lewis, G., Malafosse, A., Wendland, J.R., Lewis, C.M., McGuffin, P., and Uher, R. (2013). Contribution of Common Genetic Variants to Antidepressant Response. *Biol. Psychiatry* 73, 679–682.

Tansey, K.E., Rucker, J.J.H., Kavanagh, D.H., Guipponi, M., Perroud, N., Bondolfi, G., Domenici, E., Evans, D.M., Hauser, J., Henigsberg, N., et al. (2014). Copy number variants and therapeutic response to antidepressant medication in major depressive disorder. *Pharmacogenomics J.* 14, 395–399.

Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C.T., et al. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* 38, W214-220.

Wohnik, G.M., Rüegg, J., Abel, G.A., Schmidt, U., Holsboer, F., and Rein, T. (2005). FK506-binding proteins 51 and 52 differentially regulate dynein interaction and nuclear translocation of the glucocorticoid receptor in mammalian cells. *J. Biol. Chem.* 280, 4609–4616.

PART 1:

Candidate genes involved in antidepressant pharmacokinetics and pharmacodynamics

Chapter 2

Effect of cytochrome CYP2C19 metabolizing activity on antidepressant response and side effects: meta-analysis of data from genome-wide association studies

Chiara Fabbri, Katherine E. Tansey, Roy H. Perlis, Joanna Hauser, Neven Henigsberg, Wolfgang Maier, Ole Mors, Anna Placentino, Marcella Rietschel, Daniel Souery, Gerome Breen, Charles Curtis, Sang-Hyuk Lee, Stephen Newhouse, Hamel Patel, Michael O'Donovan, Glyn Lewis, Gregory Jenkins, Richard M. Weinshilboum, Anne Farmer, Katherine J. Aitchison, Ian Craig, Peter McGuffin, Koen Schruers, Joanna M. Biernacka, Rudolf Uher, Cathryn M. Lewis

Published: Fabbri et al. Effect of cytochrome CYP2C19 metabolizing activity on antidepressant response and side effects: Meta-analysis of data from genome-wide association studies. *European Neuropsychopharmacology*, 2018 28(8):945-954.

Abstract

Cytochrome (CYP) P450 enzymes have a primary role in antidepressant metabolism and variants in these polymorphic genes are targets for pharmacogenetic investigation. This is the first meta-analysis to investigate how *CYP2C19* polymorphisms predict citalopram/escitalopram efficacy and side effects.

CYP2C19 phenotypes comprise poor metabolizers (PM), intermediate and intermediate+ metabolizers (IM; IM+), extensive and extensive+ metabolizers (EM [wild type]; EM+) and ultra-rapid metabolizers (UM) defined by the two most common *CYP2C19* functional polymorphisms (rs4244285 and rs12248560). These polymorphisms were genotyped or imputed from genome-wide data in four samples treated with citalopram or escitalopram (GENDEP, STAR*D, GenPod, PGRN-AMPS). Treatment efficacy was defined as percentage symptom improvement and remission. Side effect data in GENDEP, PGRN-AMPS and STAR*D were analysed at weeks 2-4, 6 and 9. A fixed-effects meta-analysis was performed using EM as the reference group.

Analysis of 2558 patients for efficacy and 2037 patients for side effects showed that PMs had higher symptom improvement (SMD=0.43, CI=0.19-0.66) and higher remission rates (OR=1.55, CI=1.23-1.96) compared to EMs. At weeks 2-4, PMs showed higher risk of gastro-intestinal side effects (OR=1.26, CI=1.08-1.47), neurological side effects (OR=1.28, CI=1.07-1.53) and sexual side effects (OR=1.52, CI=1.23-1.87; week 6 values similar) but no difference for cardiovascular and sleep side effects. No difference was seen at week 9 or in total side effect burden.

Despite the size of the whole sample, PMs are relatively rare in the Caucasian population and only 51 were available. *CYP2C19* polymorphisms may provide helpful information for guiding citalopram/escitalopram treatment.

Keywords: *CYP2C19*, gene, polymorphism, citalopram, escitalopram, response, side effects, meta-analysis

1. Introduction

Major depressive disorder (MDD) is a leading cause of disability-adjusted life years worldwide (GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, 2016). Although anti-depressant drugs can be an effective therapy, remission rates are disappointing, largely as a consequence of high variability in efficacy among individuals combined with early discontinuation or poor compliance due to side effects (Hodgson et al., 2012; Crawford et al., 2014). Genetic variants are considered key modulators of antidepressant efficacy and side effects (Cacabelos et al., 2012). Common variants were estimated to explain approximately 42% of inter-individual variability in antidepressant response (Tansey et al., 2013), confirming the role of genetic polymorphisms as promising markers to provide personalized treatments.

Previous pharmacogenetic studies for antidepressant efficacy and side effects have focused on genes involved in antidepressant mechanisms of action (pharmacodynamics) or in antidepressant transport/metabolism (pharmacokinetics), including the cytochrome P450 genes (CYP450) (Fabbri and Serretti, 2015). These CYP450 genes are included in commercial pharmacogenetic tests (e.g. GeneSight Psychotropic, Genecept Assay™, YouScript Psychotropic (GTR: Genetic Testing Registry, 2017)). They form promising targets for personalizing antidepressant treatment, since they are responsible for antidepressant drug metabolism and their polymorphisms define phenotypic groups with different level of metabolic activity (Porcelli et al., 2011). An association between CYP450 metabolizer status (CYP450 phenotypes) and metabolite plasma levels has been consistently reported for antidepressants, but the association of CYP450 phenotypes with antidepressant efficacy and side effects is more controversial (Porcelli et al., 2011).

CYP2C19 is the primary CYP450 isoform responsible for the metabolism of citalopram and escitalopram, two commonly prescribed SSRIs (selective serotonin reuptake inhibitors) (Hicks et al., 2015). Elevated drug concentrations have been observed in *CYP2C19* poor metabolizers (PMs), which may increase the risk of adverse drug reactions, while *CYP2C19* ultrarapid metabolizers (UMs) may have lower exposure to these drugs leading to treatment failure. *CYP2C19*-adjusted doses for citalopram and escitalopram have been estimated, but these were based on observed differences in drug pharmacokinetics, not differences in clinical outcomes of efficacy and side effects (Hicks et al., 2015).

Inconsistent associations between *CYP2C19* phenotypes and citalopram/escitalopram outcomes have been observed, and several factors may

have led to the contradictory results (Peters et al., 2008; Mrazek et al., 2011; Hodgson et al., 2014; Hodgson et al., 2015):

1) Only a weak correlation exists between SSRI dose and efficacy and drug plasma levels may not be associated with either efficacy or side effects (Jakubovski et al., 2016; Hodgson et al., 2014; Hodgson et al., 2015);

2) Pharmacodynamic mechanisms may modulate the association between *CYP2C19* phenotypes and citalopram/escitalopram efficacy and some side effects, weakening the association between pharmacokinetic parameters and treatment outcomes (Jukić et al., 2016);

3) *CYP2C19* PM phenotypes are rare, and studies may have lacked power to detect a pharmacogenetic association with this phenotype.

In this study, we present the first meta-analysis to investigate association between *CYP2C19* phenotypes and citalopram/escitalopram efficacy and side effects. This large study aimed to identify a link between *CYP2C19* phenotypes and treatment outcomes and to determine whether dose adjustments based on *CYP2C19* phenotypes should be part of personalized medicine for antidepressant treatment.

2. Materials and Methods

2.1. Samples

2.1.1. GENDEP

The Genome-Based Therapeutic Drugs for Depression (GENDEP) project was a 12-week partially randomized open-label pharmacogenetic study with two active treatment arms. 867 patients with unipolar depression (ICD-10 or DSM-IV criteria) aged 19–72 years were recruited at nine European centres. Eligible participants were allocated to flexible-dosage treatment with either escitalopram (10–30 mg daily) or nortriptyline. Only 499 patients treated with escitalopram were included in the current meta-analysis. Severity of depression was assessed weekly by the Montgomery-Asberg Depression Rating Scale (MADRS) (Montgomery and Asberg, 1979), Hamilton Rating Scale for Depression (HRSD–17) (Hamilton, 1967) and other measures. Side effects were assessed at baseline and then weekly using the Antidepressant Side-Effect Checklist (ASEC) and UKU Side Effect Rating Scale, with good agreement between them. The ASEC data were analysed for this study, since they have lower rates of missing data (Uher et al., 2009). Detailed information about the GENDEP study has been previously reported (Uher et al., 2010).

2.1.2 STAR*D

The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study was a NIMH-funded study to determine the effectiveness of different treatments

for patients with MDD who have not responded to the first antidepressant treatment. Non-psychotic MDD (DSM-IV criteria) patients with age between 18 and 75 years were enrolled from primary care or psychiatric outpatient clinics. Severity of depression was assessed using the 16-item Quick Inventory of Depressive Symptomatology-Clinician Rated (QIDS-C16) (Trivedi et al., 2004) at baseline, weeks 2, 4, 6, 9, and 12. Side effects were measured at the same time points using the Patient-Rated Inventory of Side Effects (PRISE). This study uses data from level 1, where all patients received citalopram. Detailed description of the study design and population are reported elsewhere (Rush et al., 2004).

2.1.3. PGRN-AMPS

The Pharmacogenomic Research Network Antidepressant Medication Pharmacogenomic Study (PGRN-AMPS) included 529 participants with nonpsychotic MDD recruited at inpatient and outpatient practices of the Department of Psychiatry and Psychology, Mayo Clinic, Rochester, Minnesota. Participants were offered an eight-week course of treatment with either citalopram or escitalopram and depressive symptoms were rated using QIDS-C16 as in STAR*D. Side effects were assessed using the PRISE scale at weeks 4 and 8. Further details were reported elsewhere (Ji et al., 2013).

2.1.4. GenPod

The GENetic and clinical Predictors Of treatment response in Depression (GenPod) was a multi-centre randomized clinical trial of 601 patients recruited in primary care who had an ICD-10 diagnosis of major depression of at least moderate severity as assessed by the Clinical Interview Schedule-Revised (CIS-R) (Lewis et al., 1992) and the Beck Depression Inventory (BDI) (Beck et al., 1961). Individuals were randomly allocated to either reboxetine (4 mg twice daily) or citalopram (20 mg/day). 240 patients of European ancestry and treated with citalopram were included in this meta-analysis. Further details about this study can be found elsewhere (Thomas et al., 2008).

2.2. Outcomes

2.2.1. Treatment efficacy

Treatment efficacy was measured by percentage symptom improvement and by remission at study endpoint. Continuous measures, such as percentage improvement, capture more information and have higher power than cutoff-based dichotomous measures, however remission has a particular clinical relevance since it is associated with MDD prognosis (Streiner, 2002; Gaynes et al., 2009).

The percentage symptom improvement was corrected for possible confounding variables (age, baseline severity, and center for multi-center studies) and then standardized to allow comparability across studies.

Remission was defined as a binary variable according to standard definitions (HRSD-17 ≤ 7 in GENDEP; QIDS-C16 ≤ 5 in STAR*D and PGRN-AMPS; BDI < 10 in GenPod). In GENDEP symptom improvement was calculated using the MADRS scale similarly to previous studies (Uher et al., 2010) while HRSD-17 was used to define remission given the stronger consensus about the threshold to identify remission on this scale in contrast to MADRS, where different definitions of remission have been reported (Li et al., 2016; Jacobsen et al., 2015).

HRSD-17 and QIDS-C16 missing values at follow-up were imputed using the best unbiased estimate from a mixed-effect linear regression model, with fixed linear and quadratic effects of time and random effects of individual and center of recruitment, following previously reported methods (Uher et al., 2010).

2.2.2. Side effects

Measures of side effects were available in GENDEP, STAR*D and PGRN-AMPS. In GENDEP we chose to use the ASEC because data was more complete than the UKU (Uher et al., 2009). In STAR*D and PGRN-AMPS side effects were assessed using the PRISE scale. Both scales use a rating of severity for each side effect (coded 0-3 in ASEC, and 0-2 in PRISE) which was dichotomized (0=absent, 1=present) for the meta-analysis. Side effects were grouped in categories that were assessed in both samples: gastro-intestinal (dry mouth, diarrhea, constipation, nausea or vomiting), cardiovascular (palpitations, dizziness or feeling light-headed on standing), central nervous system (headache, tremor, feeling like the room is spinning), sleep (insomnia, drowsiness or oversleeping) and sexual (loss of desire, trouble achieving orgasm, trouble with erection). These categories were analysed as dichotomous variables (presence of at least one side effect in each category). To assess the overall severity of side effects across both studies, we summed the number of side effects reported, and dichotomized at the 3rd quartile of the distribution in each sample. Study retention at week 4 was compared among *CYP2C19* phenotypes since patients who did not benefit from treatment or had troubling side effects are expected to be lost from follow-up early in the study.

Antidepressant-induced side effects are more frequent at the beginning of treatment and then decrease (Uher et al., 2009). We therefore meta-analysed side effects at weeks 2-4 (no assessment was performed at week 2 in PGRN-AMPS), week 6 and weeks 8-9 (no assessment was performed at week 8 in STAR*D while in GENDEP we used week 8 data because of lower missing rate compared to week 9).

In GENDEP side effects were common at baseline in medication-free patients (Uher et al., 2009). We therefore performed a sensitivity analysis excluding side effects there were present also at baseline in drug-free GENDEP patients.

2.3. Genotyping and definition of *CYP2C19* phenotypes

CYP2C19 phenotypes comprise poor metabolizers (PM), intermediate and intermediate+ metabolizers (IM; IM+), extensive and extensive+ metabolizers (EM [wild type]; EM+) and ultra-rapid metabolizers (UM) defined by the two most common *CYP2C19* functional polymorphisms (rs4244285 and rs12248560) which capture the *CYP2C19* *1, *2 and *17 functional alleles (Supplementary Table 1) (Hodgson et al., 2014). These polymorphisms were directly genotyped in GENDEP using the AmpliChip CYP450 test (Hodgson et al., 2014) and they were imputed in the other samples using the Haplotype Reference Consortium (HRC version r1.1 2016) panel as reference and Minimac3. Pre-imputation quality control was performed according to standard criteria (variants with missing rate \geq 5%; monomorphic variants; subjects with genotyping rate $<$ 97%; subjects with gender discrepancies; subjects with abnormal heterozygosity; related subjects (identity by descent (IBD) $>$ 0.1875 (Anderson et al., 2010)); population outliers according to Eigensoft analysis of linkage-disequilibrium-pruned genetic data (Price et al., 2006); and non-white subjects). Imputation quality was assessed using R^2 (Li et al., 2010) and comparing imputed and genotyped *CYP2C19* phenotypes in GENDEP.

2.4. Statistical analysis

Individual-level phenotypes and genotypes were available for all studies. A fixed-effects meta-analysis was performed with the R package “Netmeta” (<https://cran.r-project.org/web/packages/netmeta/index.html>). This package has been created for performing network meta-analysis and it was useful for this study since multiple groups needed to be compared to the reference group even if there were not indirect comparisons (i.e. all the studies provided data for each of the considered *CYP2C19* phenotypes). Phenotypic groups were compared using the wild-type EM as the reference group. A random-effects meta-analysis was carried out for completeness and comparison of findings. Standardized mean difference (SMD) or odds ratio (OR) with 95% confidence intervals (CI) were calculated. Heterogeneity across studies was assessed using I^2 and Cochran’s Q (Higgins et al., 2003).

This meta-analysis provided 80% power to identify an effect size (SMD) of $d=0.40$ when comparing PMs (the smallest group, $n=51$) with EMs (the reference group, $n=1049$) for a continuous outcome and $OR=2.21$ for a binary outcome, at a significance level of 0.05 (Faul et al., 2007).

We estimated that a corrected p value of 0.008 would account for the six independent tests that were carried out (improvement and response were correlated and considered as one test; gastro-intestinal side effects, cardiovascular side effects, sleep side effects, sexual side effects, and CNS side effects were

considered as independent outcomes). Side effects at different weeks are not independent and *CYP2C19* metabolic groups are not considered independent (they all derive from two functional SNPs in the gene), and they have specific functional meaning.

3. Results

A description of the clinical-demographic characteristics of the included samples is provided in Supplementary Table 2. There was no difference in mean citalopram or escitalopram dose by *CYP2C19* phenotypes at study endpoint in GENDEP, STAR*D and PGRN-AMPS (dose information was not available in GenPod). The distribution of phenotypic groups in the analysed samples is reported in Supplementary Table 3A. Imputation quality was high in all samples for both polymorphisms (R^2 between 0.95 and 0.99 (Li et al., 2010)). GENDEP participants had 97.6% consistency between genotyped and imputed SNPs (Supplementary Table 3B).

3.1. Treatment efficacy

In total, 2558 patients were included in the meta-analysis. The distribution of efficacy outcomes across *CYP2C19* phenotypes was reported in Supplementary Table 4. Compared to EMs, PMs had higher symptom improvement scores (SMD=0.43, CI=0.19-0.66, $p=0.00037$) and higher remission rates (OR=1.55, CI=1.23-1.96, $p=0.00025$), with low or absent heterogeneity (I^2 was 11.5% and 0%, respectively). Other *CYP2C19* phenotypes did not show different outcomes compared to EMs (Figure 1). Results did not change using a random-effects model.

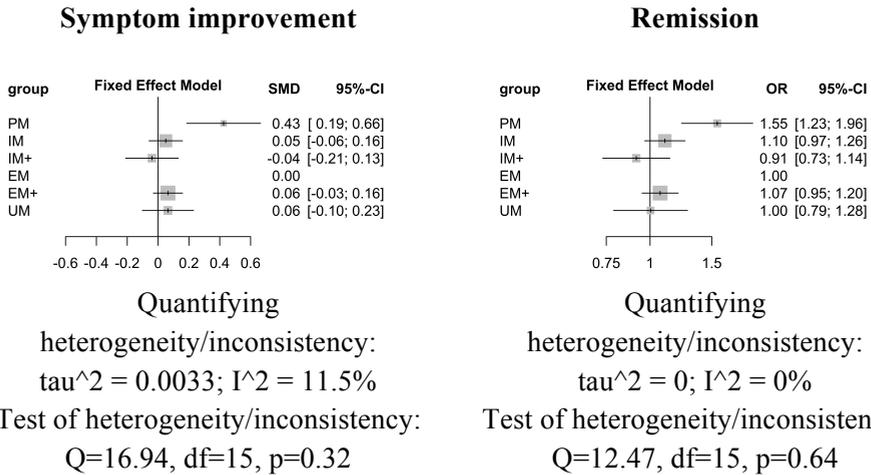
3.2. Treatment side effects

Across STAR*D, GENDEP and PGRN-AMPS 2037 patients were included in the analysis. The distribution of side effects across *CYP2C19* phenotypes was reported in Supplementary Table 5. At weeks 2-4, PMs showed higher risk of gastrointestinal side effects (OR=1.26, CI=1.08-1.47, $p=0.0033$), of CNS side effects (OR=1.28, CI=1.07-1.53, $p=0.0068$) and of sexual side effects (OR=1.52, CI=1.23-1.87, $p=0.0001$) (Figure 2).

Considering a corrected p threshold of 0.008, all these side effects were significantly more frequent in PMs. At week 6, PMs showed higher risk of sexual side effects (OR=1.64, CI=1.23-2.17, $p=0.0007$) but no higher risk of other side effects. For all these comparisons heterogeneity was low (I^2 range 0%-24%). No difference was seen at week 8-9 for any side effect, except a weak non-significant trend for sexual side effects; no difference in total side effects burden was observed at any time point. *CYP2C19* IM+ group was the only phenotype to show higher

risk of drop out at week 4 (OR=1.80, 95% CI=1.08-3.00, p=0.024), but this association did not survive multiple-testing correction.

Figure 1: meta-analysis results for improvement and remission. PM=poor metabolizers; IM=intermediate metabolizers; IM+= intermediate metabolizers plus; EM=extensive metabolizers; EM+= extensive metabolizers+; UM=ultrarapid metabolizers. EM was taken as reference group. SMD=standardized mean difference. CI=confidence interval.



Other *CYP2C19* phenotypic groups did not show relevant differences compared to EMs, except lower risk of cardiovascular side effects and sleep side effects in EM+ at weeks 2-4 (OR=0.77, CI=0.64-0.92, p=0.0048) and 6 (OR=0.84, CI=0.75-0.95, p=0.0039), respectively, and higher risk of CNS side effects at week 8 in UMs (OR=1.26, 95% CI=1.04-1.53, p=0.019), but the latter did not survive multiple-testing correction.

The use of a random-effects model did not change the results.

Excluding those side effects there were already present at baseline in drug-free patients in GENDEP, results did not change, except that PMs showed higher risk of gastro-intestinal side effects also at week 6 (OR=1.47, CI=1.13-1.92, p=0.004). In addition, the trend of higher sexual side effects in PMs at weeks 8-9 was not observed, the lower risk of cardiovascular side effects in EM+ at weeks 2-4 became a non-significant trend (OR=0.82, CI=0.67-0.99) and there was a non-significant trend of higher gastro-intestinal side effects in PMs at weeks 8-9 (OR=1.35, CI=1.01-1.81).

4. Discussion

This study shows that *CYP2C19* PMs had higher symptom improvement and higher remission probability compared to EMs during treatment with citalopram or escitalopram (Figure 1). The observed SMD of 0.43 in symptom improvement between PMs and EMs is statistically considered close to a medium effect size (0.50) (Faraone, 2008). Statistical outcomes cannot be equated with clinical relevance and a clinical relevance cutoff of SMD=0.24 was proposed based on the effect size observed for antidepressant drugs (SMD=0.31, CI=0.27-0.35) and psychotherapy (SMD=0.25, CI= 0.14-0.36) in depression (Cuijpers et al., 2014). Other *CYP2C19* phenotypes, including UMs, showed no differences in efficacy outcomes compared to EMs. In addition to increased treatment efficacy, PMs showed higher risk of gastro-intestinal, CNS and sexual side effects early in treatment (particularly during the first 2-4 weeks), but not later in treatment (weeks 8-9) (Figure 2). At week 4, PMs did not show a higher burden of total side effects and had no higher risk of drop-out. These results suggest that although side effects are more common in PMs in the first weeks of treatment, they were not more troubling than in other *CYP2C19* groups. No difference in antidepressant dose was found across *CYP2C19* phenotypic groups. These findings are consistent with a previous STAR*D study that investigated remission and tolerance to citalopram (Mrazek et al., 2011), where tolerance represents a measure of side effect level. Tolerance was defined as continuation of citalopram treatment after the completion of Level 1 of the STAR*D trial.

Previous studies in GENDEP and STAR*D failed to establish association between *CYP2C19* metabolizer status (PM vs. EM) and response, side effects or study retention (Peters et al., 2008; Hodgson et al., 2015; Hodgson et al., 2014), but individual studies would have limited power given the low number of subjects with PM phenotype (~2% of all patients analysed), particularly in GENDEP which has only six PM subjects. A previous analysis of *CYP2C19* in GENDEP used different definitions of side-effect, investigating each ASEC item and the sum of ASEC items (Hodgson et al., 2015) No difference in treatment efficacy or side effects was identified between UMs and EMs, except for a non-significant higher risk of CNS side effects only at weeks 8-9 (Figure 2) that was probably the effect of random noise.

The only phenotypic group that showed lower risk of side effects was EM+ (lower risk of sleep side effects at week 6 and of cardiovascular side effects at weeks 2-4), suggesting that weak differences may depend on metabolic level but the UM group may have not provided enough power to observe them (~4-5% of patients were UMs in the analysed samples).

In addition to pharmacokinetic mechanisms, pharmacodynamic mechanisms may be involved in the association between *CYP2C19* and antidepressant response, since *CYP2C19* activity was reported to influence central neurotransmitters and neurotrophins relevant to antidepressant mechanisms of action (Jukić et al., 2016). Our results conflict with the recommendation, based on pharmacokinetic parameters, of a 50% reduction in the starting dose of citalopram/escitalopram in *CYP2C19* PMs (Hicks et al., 2015), since we showed that a standard dose was associated with greater efficacy without higher drop-out rates or higher total burden of side effects. Antidepressant treatment with citalopram/escitalopram may be particularly indicated in *CYP2C19* PMs given the efficacy profile, if appropriate clinical support and monitoring is provided and the patient is informed of potential side effects at the beginning of the treatment. Effective plasma (and brain) drug concentrations may be reached in a higher proportion of PMs than other phenotypes, at the price of more frequent early side effects. The good tolerability profile of citalopram/escitalopram implies that these side effects are usually not troubling, which may not be true for other antidepressants, such as tricyclic antidepressants (TCAs) or venlafaxine (Cipriani et al., 2012; Cipriani et al., 2009). It should be noted that TCAs and venlafaxine have specific profiles of efficacy and they represent valid alternatives to SSRIs as currently reported in clinical guidelines, but it should not be assumed that the current results referred to *CYP2C19* PMs can be applied to antidepressants different from citalopram and escitalopram.

The limitations and strengths of this study should be considered. This was the first meta-analysis to investigate the role of *CYP2C19* phenotypes in citalopram/escitalopram efficacy/side effects, individual level data were available in all samples and the total sample size was the largest ever used for investigating this topic. On the other hand, PMs are rare in the Caucasian population resulting in limited power to identify differences involving this group even in this sample of 2558 patients.

Side effect assessment was not available in all samples, and at weeks 6 and 8-9 part of patients dropped from the study and side effects data could not be imputed because it would be unreliable. At weeks 6 and 8-9, respectively, side effects were available in 84.4% and 73.6% of the initial sample in STAR*D, while in 85.9% and 83.3% of the initial sample in GENDEP. In PGRN-AMPS 0.87% of patients initially included had side effect data at week 4 and 80% at week 8. Our findings suggest that *CYP2C19* PMs may benefit from standard doses of citalopram/escitalopram, with a higher response than other phenotypes. No conclusions could be drawn for UMs since which showed no significant

differences in outcomes compared to EMs, and the study was probably under-powered to detect weak effects. EM+ was the only group showing lower risk of some side effects compared to EMs. We observed no to low heterogeneity among studies for both efficacy and side effects. For the former group all samples showed similar better outcome in PMs compared to EMs except GENPOD, which included only three PM patients explaining the marginal effect on heterogeneity. Finally, the possible confounding effect of *CYP2C19* enhancers/inhibitors was not assessed, but a previous analysis in GENDEP concluded that the exclusion of subjects with concomitant use of enhancers/inhibitors did not change the pattern of results (Hodgson et al., 2014).

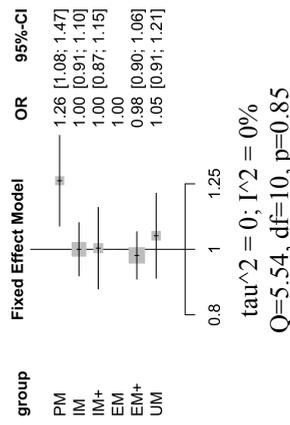
In conclusion, this meta-analysis shows good efficacy in *CYP2C19* poor metabolisers with citalopram/escitalopram, contrasting previous pharmacokinetic findings (Hicks et al., 2015). Our results show better treatment outcomes in PMs treated with standard doses with no relevant impact on late side effects (after the 6th week of treatment). Careful information for patients and monitoring of side effects during the early phase of treatment are recommended. Other *CYP2C19* phenotypes, including UMs, did not show differences in efficacy or side-effect outcomes compared to EMs. An interesting implication of this study is the possibility to derive *CYP2C19* metabolic groups from standard genome-wide data with a good level of quality.

Supplementary materials:

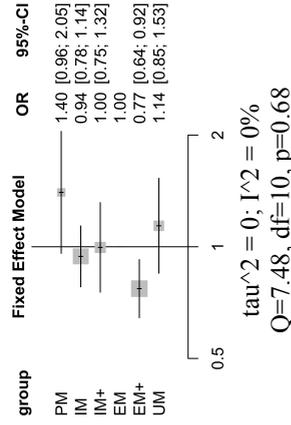
Supplementary materials are available for the online version of this publication.

Figure 2: meta-analysis results for side effects. PM=poor metabolizers; IM=intermediate metabolizers; IM+= intermediate metabolizers plus; EM=extensive metabolizers; EM+= extensive metabolizers+; UM=ultrarapid metabolizers. EM was taken as reference group. SMD=standardized mean difference. CI=confidence interval. For each comparison heterogeneity is quantified using τ^2 , I^2 and assessed using Q test.

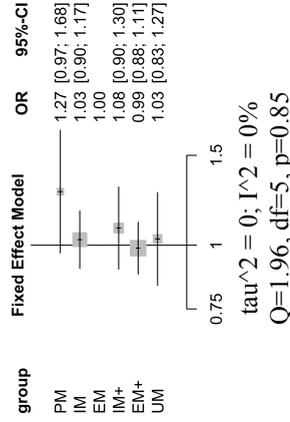
Gastro-intestinal side effects at weeks 2-4



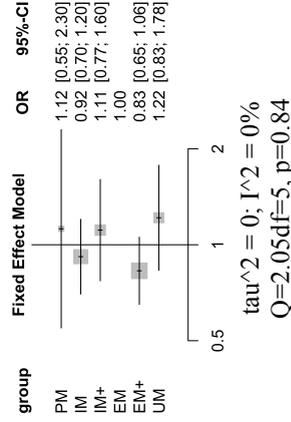
Cardiovascular side effects at weeks 2-4



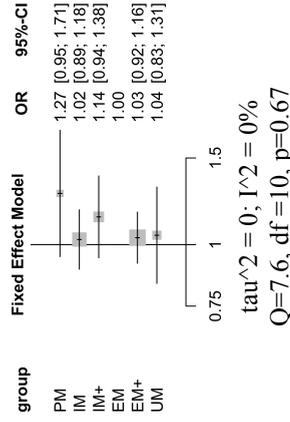
Gastro-intestinal side effects at week 6



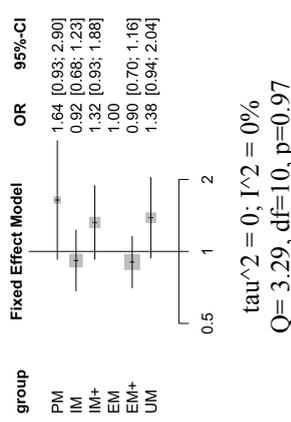
Cardiovascular side effects at week 6



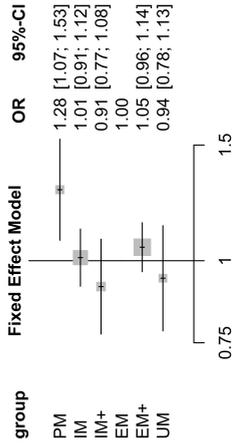
Gastro-intestinal side effects at weeks 8-9



Cardiovascular side effects at weeks 8-9

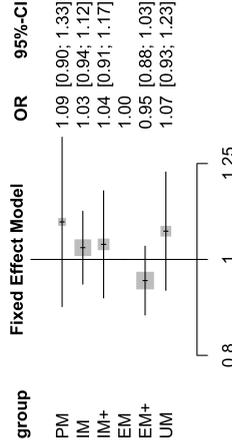


CNS side effects at weeks 2-4



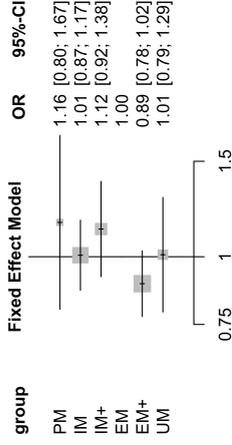
tau² = 0.0011; I² = 5.2%
Q=10.55, df=10, p=0.39

Sleep side effects at weeks 2-4



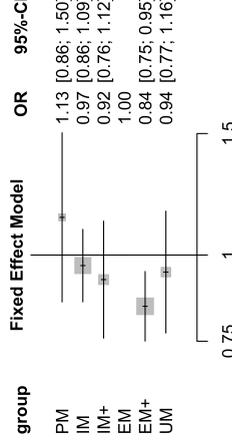
tau² = 0; I² = 0%
Q=4.24, df=10, p=0.94

CNS side effects at week 6



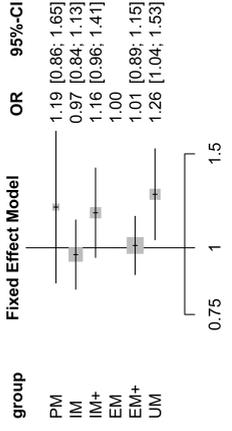
tau² = 0; I² = 0%
Q=0.68, df=5, p=0.98

Sleep side effects at week 6



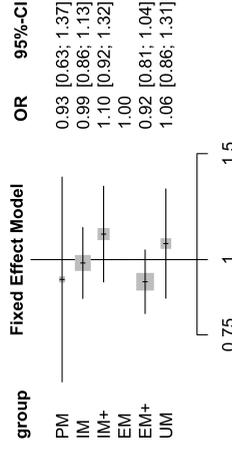
tau² = 0; I² = 0%
Q=3.24, df=5, p=0.66

CNS side effects at weeks 8-9



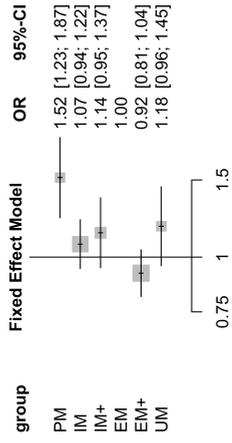
tau² = 0.0018; I² = 4.8%
Q=10.51, df=10, p=0.40

Sleep side effects at weeks 8-9



tau² = 0; I² = 0%
Q=4.57, df=10, p=0.92

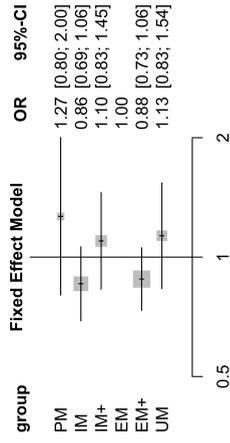
Sexual side effects at weeks 2-4



$\tau^2 = 0$; $I^2 = 0\%$

$Q = 6.89$, $df = 10$, $p = 0.74$

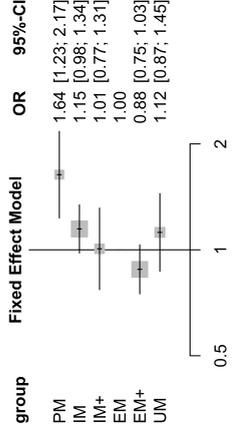
Total side effects at week 2 >=3rd quartile



$\tau^2 = 0.0039$, $I^2 = 5.2\%$

$Q = 5.27$, $df = 5$, $p = 0.38$

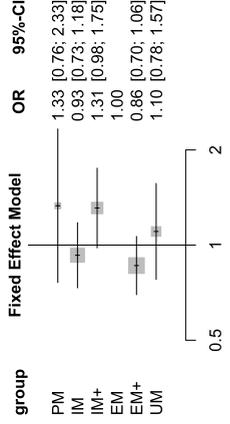
Sexual side effects at week 6



$\tau^2 = 0.0152$, $I^2 = 24\%$

$Q = 6.58$, $df = 5$, $p = 0.25$

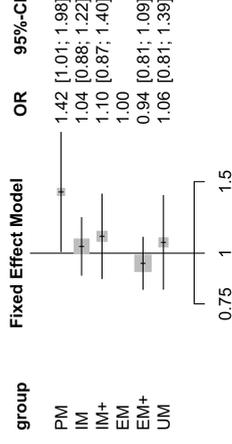
Total side effects at week 6 >=3rd quartile



$\tau^2 = 0$, $I^2 = 0\%$

$Q = 3.61$, $df = 5$, $p = 0.61$

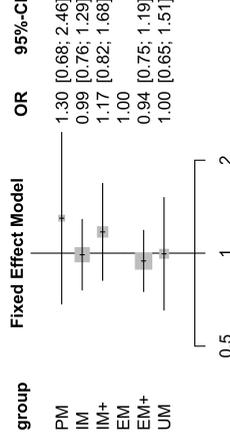
Sexual side effects at weeks 8-9



$\tau^2 = 0$; $I^2 = 0\%$

$Q = 9.95$, $df = 10$, $p = 0.45$

Total side effects at weeks 8-9 >=3rd quartile



$\tau^2 = 0$; $I^2 = 0\%$

$Q = 2.5$, $df = 5$, $p = 0.78$

References

- Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, and Zondervan, KT (2010). Data quality control in genetic case-control association studies. *Nat Protoc* 5: 1564–1573.
- Beck AT, Ward CH, Mendelson M, Mock J, and Erbaugh J (1961). An inventory for measuring depression. *Arch Gen Psychiatry* 4: 561–571.
- Cacabelos R, Martinez-Bouza R, Carril JC, Fernandez-Novoa L, Lombardi V, Carrera I *et al* (2012). Genomics and pharmacogenomics of brain disorders. *Curr Pharm Biotechnol* 13: 674–725.
- Cipriani A, Santilli C, Furukawa TA, Signoretti A, Nakagawa A, McGuire H *et al* (2009). Escitalopram versus other antidepressive agents for depression. *Cochrane Database Syst Rev* CD006532.
- Cipriani A, Purgato M, Furukawa TA, Trespidi C, Imperadore G, Signoretti A *et al* (2012). Citalopram versus other anti-depressive agents for depression. *Cochrane Database Syst Rev* CD006534.
- Crawford AA, Lewis S, Nutt D, Peters TJ, Cowen P, O'Donovan MC *et al* (2014). Adverse effects from antidepressant treatment: randomised controlled trial of 601 depressed individuals. *Psychopharmacology (Berl)* 231: 2921–2931.
- Cuijpers P, Turner EH, Koole SL, van Dijke A, and Smit F (2014). What is the threshold for a clinically relevant effect? The case of major depressive disorders. *Depress Anxiety* 31: 374–378
- Fabbri C, and Serretti A (2015). Pharmacogenetics of Major Depressive Disorder: Top Genes and Pathways Toward Clinical Applications. *Curr Psychiatry Rep* 17:50.
- Faraone SV (2008). Interpreting estimates of treatment effects: implications for managed care. *Pharmacy & Therapeutics* 33: 700–711.
- Faul F, Erdfelder E, Lang A-G, and Buchner A (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 39: 175–191.
- Gaynes BN, Warden D, Trivedi MH, Wisniewski SR, Fava M, and Rush AJ (2009). What did STAR*D teach us? Results from a large-scale, practical, clinical trial for patients with depression. *Psychiatr Serv Wash DC* 60: 1439–1445.
- Disease and Injury Incidence and Prevalence Collaborators (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Lond Engl* 388: 1545–1602.

Hamilton M (1967). Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol* 6: 278–296.

Hicks JK, Bishop JR, Sangkuhl K, Müller DJ, Ji Y, Leckband SG *et al* (2015). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2D6 and CYP2C19 Genotypes and Dosing of Selective Serotonin Reuptake Inhibitors. *Clin Pharmacol Ther* 98: 127–134.

Higgins JPT, Thompson SG, Deeks JJ, and Altman DG (2003). Measuring inconsistency in meta-analyses. *BMJ* 327: 557–560.

Hodgson K, Mufti SJ, Uher R, and McGuffin P (2012). Genome-wide approaches to antidepressant treatment: working towards understanding and predicting response. *Genome Med* 4: 52.

Hodgson K, Tansey K, Dernovsek MZ, Hauser J, Henigsberg N, Maier W *et al* (2014). Genetic differences in cytochrome P450 enzymes and antidepressant treatment response. *J Psychopharmacol Oxf Engl* 28: 133–141.

Hodgson K, Tansey KE, Uher R, Dernovšek MZ, Mors O, Hauser J *et al* (2015). Exploring the role of drug-metabolising enzymes in antidepressant side effects. *Psychopharmacology (Berl)* 232: 2609–2617.

Jacobsen PL, Mahableshwarkar AR, Serenko M, Chan S, and Trivedi MH (2015). A randomized, double-blind, placebo-controlled study of the efficacy and safety of vortioxetine 10 mg and 20 mg in adults with major depressive disorder. *J Clin Psychiatry* 76: 575–582.

Jakubovski E, Varigonda AL, Freemantle N, Taylor MJ, and Bloch MH (2016). Systematic Review and Meta-Analysis: Dose-Response Relationship of Selective Serotonin Reuptake Inhibitors in Major Depressive Disorder. *Am J Psychiatry* 173: 174–183.

Ji Y, Biernacka JM, Hebring S, Chai Y, Jenkins GD, Batzler A *et al* (2013). Pharmacogenomics of selective serotonin reuptake inhibitor treatment for major depressive disorder: genome-wide associations and functional genomics. *Pharmacogenomics J* 13: 456–463.

Jukić MM, Opel N, Ström J, Carrillo-Roa T, Miksys S, Novalen M *et al* (2016). Elevated CYP2C19 expression is associated with depressive symptoms and hippocampal homeostasis impairment. *Mol Psychiatry* 22, 1155-1163.

Lewis G, Pelosi AJ, Araya R, and Dunn G (1992). Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychol Med* 22: 465–486.

Li H, Gu N, Zhang H, Wang G, Tan Q, Yang F *et al* (2016). Efficacy and safety of quetiapine extended release monotherapy in bipolar depression: a multi-center, randomized, double-blind, placebo-controlled trial. *Psychopharmacology (Berl)* 233: 1289–

1297.

Li Y, Willer CJ, Ding J, Scheet P, and Abecasis GR (2010). MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 34: 816–834.

Montgomery SA, and Asberg M (1979). A new depression scale designed to be sensitive to change. *Br J Psychiatry J Ment Sci* 134: 382–389.

Mrazek DA, Biernacka JM, O’Kane DJ, Black JL, Cunningham JM, Drews MS *et al* (2011). CYP2C19 variation and citalopram response. *Pharmacogenet Genomics* 21: 1–9.

Peters EJ, Slager SL, Kraft JB, Jenkins GD, Reinalda MS, McGrath PJ *et al* (2008). Pharmacokinetic genes do not influence response or tolerance to citalopram in the STAR*D sample. *PLoS One* 3: e1872.

Porcelli S, Fabbri C, Spina E, Serretti A, and De Ronchi D (2011). Genetic polymorphisms of cytochrome P450 enzymes and antidepressant metabolism. *Expert Opin Drug Metab Toxicol* 7: 1101–1115.

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, and Reich D (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904–909.

Rush AJ, Fava M, Wisniewski SR, Lavori PW, Trivedi MH, Sackeim HA *et al* (2004). Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. *Control Clin Trials* 25: 119–142.

Streiner DL (2002). Breaking up is hard to do: the heartbreak of dichotomizing continuous data. *Can J Psychiatry* 47: 262–266.

Tansey KE, Guipponi M, Hu X, Domenici E, Lewis G, Malafosse A *et al* (2013). Contribution of Common Genetic Variants to Antidepressant Response. *Biol Psychiatry* 73: 679–682.

Thomas L, Mulligan J, Mason V, Tallon D, Wiles N, Cowen P *et al* (2008). GENetic and clinical predictors of treatment response in depression: the GenPod randomised trial protocol. *Trials* 9: 29.

Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T *et al* (2004). The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med* 34: 73–82.

Uher R, Farmer A, Henigsberg N, Rietschel M, Mors O, Maier W *et al* (2009). Adverse reactions to antidepressants. *Br J Psychiatry* 195: 202–210.

Uher R, Perroud N, Ng MY, Hauser J, Henigsberg N, Maier W *et al* (2010). Genome-wide pharmacogenetics of antidepressant response in the GENDEP project. *Am J Psychiatry* 167: 555-64.

Chapter 3

Pleiotropic genes in psychiatry: calcium channels and the stress-related FKBP5 gene in antidepressant response and resistance

Chiara Fabbri, Filippo Corponi, Diego Albani, Ilaria Raimondi, Gianluigi Forloni, Koen Schruers, Siegfried Kasper, Alexander Kautzky, Joseph Zohar, Daniel Souery, Stuart Montgomery, Carlotta Pia Cristalli, Vilma Mantovani, Julien Mendlewicz, Alessandro Serretti

Published: Fabbri et al. Pleiotropic genes in psychiatry: calcium channels and the stress-related FKBP5 gene in antidepressant response and resistance. *Neuropsychopharmacology & Biological Psychiatry*, 2018; 81:203-210.

Abstract

A candidate gene and a genome-wide approach were combined to study the pharmacogenetics of antidepressant response and resistance. Investigated genes were selected on the basis of pleiotropic effect across psychiatric phenotypes in previous genome-wide association studies and involvement in antidepressant response.

Three samples with major depressive disorder (total=671) were genotyped for 44 SNPs in 8 candidate genes (*CACNA1C*, *CACNB2*, *ANK3*, *GRM7*, *TCF4*, *ITIH3*, *SYNE1*, *FKBP5*). Phenotypes were response/remission after 4 weeks of treatment and treatment-resistant depression (TRD). Genome-wide data from STAR*D were used to replicate findings for response/remission (n=1409) and TRD (n=620). Pathways including the most promising candidate genes were investigated in STAR*D for involvement in TRD.

FKBP5 polymorphisms showed replicated but nominal associations with response, remission or TRD. *CACNA1C* rs1006737 and rs10848635 were the only polymorphisms that survived multiple-testing correction. In STAR*D the best pathway associated with TRD included *CACNA1C* (GO:0006942, permuted p=0.15). Machine learning models showed that independent SNPs in this pathway predicted TRD with a mean sensitivity of 0.83 and specificity of 0.56 after 10-fold cross validation repeated 100 times.

FKBP5 polymorphisms appear good candidates for inclusion in antidepressant pharmacogenetic tests. Pathways including the *CACNA1C* gene may be involved in TRD and they may provide the base for developing multi-marker predictors of TRD.

Keywords: antidepressants; treatment-resistant depression; gene; GWAS; pharmacogenetics

1. Introduction

Major depressive disorder (MDD) is a high-prevalence disease associated with a heavy personal and socio-economic burden, partly as a result of the lack of tailored treatment options (1). Common single nucleotide polymorphisms (SNPs) were estimated to account for 42% of the variance in antidepressant response (2), confirming the hypothesis that genetic polymorphisms may be used as effective markers to provide personalized treatments.

Candidate gene and genome-wide analyses can provide a complementary strategy, since the former can be applied to clarify the role of SNPs with high pre-test probability of association with the trait and the latter is useful to study the joined effects of a number of SNPs in a gene or a set of genes (3).

The present study aimed to apply this complementary strategy to the study of eight genes that are optimal candidates for involvement in antidepressant response on the base of their biological function and previous literature. The genes of interest are involved in the regulation of neurotransmission (*CACNA1C*, *CACNB2*, *ANK3*), neural differentiation, synaptic plasticity, adhesion processes and structural organization (*GRM7*, *TCF4*, *ITIH3*, *SYNE1*) and glucocorticoid signaling (*FKBP5*).

CACNA1C, *CACNB2*, *ANK3*, *TCF4*, *ITIH3* and *SYNE1* have been associated with multiple psychiatric phenotypes in Psychiatric Genomics Consortium (PGC) genome-wide association studies (GWAS) (4–7). PGC was able to collect well-powered samples for genome-wide analysis, while GWAS investigating antidepressant efficacy mostly failed to identify significant signals because of limited power (8,9). Shared genetic factors were hypothesized to contribute to the pathogenesis of different psychiatric traits (i.e. pleiotropy) (7), thus we decided to investigate the role of the aforementioned genes in antidepressant efficacy.

CACNA1C, *CACNB2* and *ANK3* play a role in the regulation of neurotransmission and polymorphisms in these genes were associated with vulnerability to a number of different neuropsychiatric conditions, suggesting that abnormalities in voltage-sensitive ion channels: 1) may represent a continuum of shared neurobiological vulnerability across several neuropsychiatric diseases and 2) may provide a basis for common psychiatric susceptibility (10). *CACNA1C* encodes for the α -1C subunit of the L-type voltage-dependent calcium channel, it transiently increases calcium-mediated membrane depolarization and modulates intracellular signaling, gene transcription and synaptic plasticity (10). Genetic variations in *CACNA1C* (particularly rs1006737, but also rs72552065, rs2370419, rs2370411, rs7297582, rs4765913, rs4765905) have been associated with MDD, schizophrenia, autism spectrum disorders (ASD), as well as changes in brain function and structure in

healthy subjects (10,11). *CACNB2* codes for the auxiliary Cav β subunit, which modulates calcium channel activity. Five mental disorders (ASD, attention deficit-hyperactivity disorder (ADHD), bipolar disorder, MDD, schizophrenia) have been associated with polymorphisms in *CACNB2* (11,12). Ankyrin G (*ANK3*) is an adaptor protein expressed in the axonal initial segments and the nodes of Ranvier, and it has been shown to regulate the assembly of voltage-gated sodium channels (13). Variants in *ANK3* have been associated with bipolar disorder (14), depression, suicide risk, stress response (15) and schizophrenia (16).

ITIH3 and *SYNE1* are involved in adhesion processes and structural organization. *ITIH3* codes for a serine protease inhibitor that is involved in extracellular matrix stabilization. *ITIH3* rs2535629 showed the strongest association signal with susceptibility to five psychiatric disorders (ASD, ADHD, bipolar disorder, MDD and schizophrenia) in PGC samples (11). *SYNE1* is part of the complex linking the nucleoskeleton to the cytoskeleton and cellular membrane. In the brain, it is hypothesized to modulate synaptic plasticity and endocytosis of glutamate receptors (17). There is strong evidence that polymorphisms in this gene (particularly rs9371601) are involved in bipolar disorder and MDD (18).

TCF4 is a member of the basic helix–loop–helix (bHLH) family of transcription factors that have an important role in developmental processes. In the CNS *TCF4* is involved in the differentiation of glial cells, but it is also plays a role in neuronal differentiation and survival (19). Variants in *TCF4* have been strongly associated with schizophrenia (20,21), neurophysiologic traits in schizophrenia and bipolar disorder (22). A reduced expression of the *TCF4* gene was recently detected in MDD (23).

GRM7 (metabotropic glutamate receptor 7) and *FKBP5* (FK506 Binding Protein 5) were also selected because of previous convincing evidence that they are involved in MDD pathogenesis and antidepressant response. *GRM7* is one of the strongest genes reported by GWAS meta-analysis for association with MDD (24). *GRM7* has also been associated with bipolar disorder (25) and to the mechanisms of action of mood stabilizers and antidepressants (26). *GRM7* variants may modulate early antidepressant efficacy (27) and antipsychotic response (28). Very interesting results suggested that *FKBP5* may be involved in antidepressant efficacy and preliminary evidence that polymorphisms in this gene may be useful to guide treatment was provided (29). Polymorphisms (rs3800373, rs1360780, rs9296158, rs9470080) associated with enhanced expression of *FKBP5* lead to an increased glucocorticoid receptor (GR) resistance and decreased efficiency of the negative feedback of the stress hormone axis, thus they may increase the risk of MDD, bipolar disorder and post-traumatic stress disorder (30). Previous studies suggested

a possible role of rs1360780, rs3800373, rs352428 and rs4713916 in antidepressant treatment outcome (31–36).

The present paper aimed to study the role of the aforementioned genes in antidepressant response and treatment-resistance. The effect of these genes was studied applying a candidate gene approach in four independent samples. Subsequently, genome-wide data were used to study the pathways that may mediate their role in antidepressant action.

2. Materials and Methods

2.1 Samples

For all three original samples ethical approval was obtained from local research ethics committees. The clinical-demographic characteristics of samples are described in Supplementary Table 1. ES1 and ES2 were collected thanks to the European Group for the Study of Resistant Depression (GSRD) (37).

2.1.1 European Sample 1 (ES1)

357 patients with a diagnosis of major depressive episode (DSM-IV criteria) were recruited in the context of an European multicenter project. Eligible patients were treated with antidepressants in a naturalistic setting. Depressive symptoms were evaluated according to Hamilton Depression Rating Scale (HDRS 21-item version) at week 4. Detailed description of the whole sample has been provided elsewhere (38) and other information, including inclusion/exclusion criteria, is provided in Supplementary Methods.

2.1.2 European Sample 2 (ES2)

218 MDD patients were recruited in the context of an European multicenter project. Inclusion and exclusion criteria have been described previously in detail (39). Briefly, patients met DSM-IV-TR criteria for major depressive episode defined as moderate or severe (Montgomery Asberg Depression Rating Scale (MADRS) total score at baseline > 22). Patients entered a two stage trial after the failure of at least one adequate antidepressant treatment (retrospectively assessed), firstly receiving a 6 week venlafaxine treatment and then, in case of non response, a 6 week escitalopram treatment. Depressive symptoms were evaluated according to MADRS at baseline and biweekly until week 12. More detailed information about this sample is provided in Supplementary Methods.

2.1.3 Italian Sample (ITAS)

96 patients with diagnosis of non-psychotic MDD (DSM-IV-TR criteria) and with a baseline HDRS score ≥ 13 were included. Inclusion and exclusion criteria were detailed elsewhere (40). Eligible patients were treated with antidepressants in a naturalistic setting. Patients were evaluated for depressive symptomatology (HDRS

21-items) by trained psychiatrists at baseline and weekly until week 8. More detailed information about this sample is provided in Supplementary Methods.

2.1.4 STAR*D sample

Sequenced Treatment Alternatives to Relieve Depression (STAR*D) sample was used for replication purposes and for multi-marker tests (see Statistical analysis). Further descriptions of the study design and study population are detailed elsewhere (41). In brief, non-psychotic MDD (DSM-IV criteria) patients were enrolled from primary care or psychiatric outpatient clinics. Severity of depression was assessed using the 16-item Quick Inventory of Depressive Symptomatology-Clinician Rated (QIDS-C) (42) at baseline, weeks 2, 4, 6, 9, and 12. All patients received citalopram in level 1. Patients who did not show response in level 1 were encouraged to enter level 2. In level 2, there were seven different treatment options, and cognitive behavioral therapy (CBT) was included as the psychotherapy option. There were three combination options (either bupropion-SR, buspirone or CBT added to citalopram), and four switch options (either sertraline, bupropion-SR, venlafaxine-XR, or CBT). Depressive symptoms were rated as in level 1.

2.2 Outcomes under investigation

In all original samples response and remission at week 4 or 6 were investigated according to standard definitions (response was defined as a decrease of at least 50% in the HDRS-21 or the MADRS, while remission was defined as HDRS \leq 7 or MADRS $<$ 10).

In the European samples, treatment resistance depression (TRD) was also investigated. In ES1, two definitions of treatment resistance have been considered: (1) non-response to at least two adequate consecutive antidepressant treatments administered during the last episode (wide definition, TRD-W (38)); (2) non-response to at least two adequate consecutive antidepressant treatments of different classes administered during the last episode (different classes definition, TRD-DC (43,44)). In ES2, TRD was defined as: (1) non-response to escitalopram (TRDA) and (2) non-remission to escitalopram (TRDB). TRDA and TRDB are equivalent to TRD-DC since the two antidepressants prescribed in ES 2 belong to different classes (SNRIs and SSRIs).

In STAR*D response and remission were investigated in level 1 and in level 2, in order to provide comparable phenotypes to the original samples, including TRD (response and remission in level 2). Response was defined as a decrease of at least 50% in the QIDS-C scale and remission as a QIDS-C score \leq 5 at endpoint.

2.4 SNPs selection and genotyping

The list of genotyped SNPs and their characteristics is shown in Supplementary Table 2. Criteria for SNP selection and methods used for genotyping are provided in Supplementary Methods.

Genome-wide data were available in STAR*D, quality control and imputation were performed according to the procedures described in Supplementary Methods.

2.5 Statistical analysis

In the original samples the effect of individual markers (alleles and genotypes) on phenotypes was tested through logistic regression models. Covariates were selected according to their impact on outcomes as previously reported (3,45). Odds ratios (OR) with 95% confidence intervals (CI) were estimated for the effects of high-risk genotypes/alleles. Markers showing $p < 0.05$ for association with binary phenotypes were further investigated through a repeated-measures mixed-effect linear regression model. In detail, the time x SNP interaction and baseline severity score were used as fixed effects and the random effect of the subject was included to study the variation of symptom severity over time. Baseline severity was included in the model because it significantly affected the AIC (Akaike information criterion) and BIC (Bayesian information criterion) parameters. The significance of each model was evaluated through its comparison to a null model (not including the SNP x time interaction term) using ANOVA. The R cran lme4 package was used (<https://cran.r-project.org/package=lme4>).

STAR*D was used for replication at individual SNP level, but pathway analysis was also carried out. At SNP level, the index variants (i.e. those genotyped in the original samples, see Supplementary Table 2) and those in linkage disequilibrium (LD) with them ($R^2 \geq 0.30$) were extracted from imputed data. Their possible effect on phenotypes was tested through logistic regression models (46). Covariates were age, baseline severity of depressive symptoms, recruitment center and ancestry-informative principal components. Pathway effects were tested using MAGMA (47). All the pathways including the genes of interest were selected from GSEA Broad Institute database version 6 (48). MAGMA performs both a self-contained and a competitive gene-set analysis, the latter is more conservative and it was used in this study. This measure reflects if genes in the pathway are more associated with the outcome than genes outside the pathway. The same covariates used for SNP analysis were included in pathway analysis. 10,000 permutations were performed to correct for multiple testing. SNP analysis was performed both in level 1 and level 2, while pathway analysis was carried out to investigate the possible biological mechanisms mediating the effect of top candidate genes on TRD (i.e. response/remission were investigated only in STAR*D level 2).

Finally, the top pathway(s) found in STAR*D were investigated through machine learning models using the R cran Caret package (<https://cran.r-project.org/package=caret>). The aim of this analysis was to provide an exploratory estimation of SNPs' predictive performance in a pathway of interest since no previous study estimated this for antidepressant response or TRD. Briefly, independent variants associated with the outcome of interest were selected and the number of risk alleles for each variant was used as predictor. Near-zero variance predictors were removed and 10-fold cross validation was repeated 100 times in order to estimate the mean accuracy of the model. Five machine learning models (neural networks, recursive partitioning, learning vector quantization, gradient boosted machine and random forests) were tested and compared in terms of accuracy in order to select the one with the best performance. Further details are provided in Supplementary Methods.

2.6. Power analysis and significance level

In single marker analysis, ES1 provided a power of 0.80 to detect risk alleles with $OR \geq 1.96$ setting alpha value to 0.05 two-tailed. Setting the same parameters, risk alleles with $OR \geq 2.35$ are detectable in ES2 and with $OR \geq 3.67$ in ITAS. Replicated SNPs at $p=0.05$ were reported among potentially relevant results because the investigated candidates have previous strong support in literature (previous GWAS) and they were investigated in four independent samples in order to reduce the risk of false-positive findings. We suggest that results involving strong candidate genes validated across independent samples are worth reporting without strict multiple-testing correction, but only SNPs surviving a Bonferroni corrected $p=0.001$ (0.05/44 SNPs, phenotypes were considered as non-independent) were referred as significant results. For pathway analysis 10.000 permutations were performed to correct for multiple-testing and for machine learning models 10-fold cross validation was repeated 100 times.

3. Results

A detailed description of the clinical-demographic characteristics of samples is provided in Supplementary Table 1. SNPs with $p < 0.05$ in at least two original samples were reported in Table 1. These SNPs were located within the *CACNA1C*, *CACNB2*, *ANK3* and *FKBP5* genes. In the following paragraphs an alternative perspective on results was provided, describing in each sample the findings that showed consistent effect on two or more of the investigated phenotypes. A comprehensive overview of all results is available in Supplementary Table 3.

3.1 Results in the European sample 1 (ES1)

rs1006737 and rs10848635 in the *CACNA1C* gene showed trends of association with at least two clinical outcomes and significant associations with response. In detail, AA genotype and A allele were associated with higher response for both rs1006737 and rs10848635, while we found nominal associations with remission and TRD-W. rs3617 in the *ITIH3* gene showed nominal associations with response, TRD-W and TRD-DC, with the CC genotype and C allele having a protective effect compared to the A allele and AA genotype. *FKBP5* rs3800373 CC genotype showed a non-significant higher risk of non-response. Nominal associations were found between single phenotypes and other SNPs in the *ANK3*, *CACNB2*, *TCF4* and *FKBP5* genes (Supplementary Table 3).

3.2 Results in the European Sample 2 (ES2)

SNPs showing nominal associations with clinical outcomes were concentrated in the *FKBP5* gene, particularly rs3800373 and rs1360780 yielded significant results for two or more phenotypes. Anyway, these results did not survive multiple-testing correction. rs3800373 AA genotype and A allele showed nominal evidence of better response and remission. rs1360780 CC genotype and C allele were associated with a trend of better response and remission to venlafaxine and lower risk of TRDA. Other *FKBP5* SNPs showed consistent but only nominal allelic associations with more than one investigated phenotype, in detail rs9470080 (protective effect of the C allele compared to T allele), rs9296157 (protective effect of the G allele compared to C allele) and rs4713916 (protective effect of the G allele compared to A allele). The nominal effect of rs3800373 and rs1360780 on venlafaxine efficacy was confirmed by repeated-time analysis (rs3800373: $p=0.03$ for the allelic analysis; rs1360780 $p=0.01$ and $p=0.0059$ for the allelic and genotypic analysis, respectively). The results of repeated-time analysis did not outline other SNP x time effects (Supplementary Table 4). Nominal associations with individual phenotypes were found in the *CACNA1C*, *CACNB2* and *SYNE1* genes (Supplementary Table 3).

3.3 Results in the Italian Sample (ITAS)

No association survived multiple-testing correction, but nominal findings were concentrated in the *FKBP5* gene and they showed consistent direction with the results found in ES2. rs1360780 CC genotype and C allele were nominally associated with better response and remission compared to the T allele. rs3800373 AA and rs9470080 CC were nominally associated with higher response and remission. rs2799573 C allele (*CACNB2* gene) showed a trend of protective effect compared to the T allele. Repeated-time analysis (Supplementary Table 4) supported the effect of *FKBP5* rs1360780 ($p=0.014$ and $p=0.001$ for the genotypic and allelic analysis, respectively) and *FKBP5* rs3800373 ($p=0.003$ in the allelic

analysis) with p values very close to the significance threshold. Nominal associations with individual phenotypes were found in *ANK3*, *CACNA1C* and *CACNB2* genes (Supplementary Table 3).

3.4 Replication results in STAR*D

All candidate SNPs were available in STAR*D and 219 SNPs showed LD ($R^2 \geq 0.30$) with at least one of them, thus a total of 263 SNPs were analyzed. Considering the SNPs with nominal associations replicated in at least two original samples, interesting results were found for *CACNA1C* rs11062157 ($R^2=0.65$ with rs10848635) for association with response in level 2 ($p=0.044$) and *FKBP5* rs9368882 ($R^2=0.67$ with rs9470080) for association with remission in level 1 ($p=0.035$) as shown in Table 1. An overview of findings is provided in Supplementary Table 5.

950 pathways including the best candidate genes (*ANK3*, *CACNA1C*, *CACNB2*, *FKBP5*) found in the original samples were investigated for association with response and remission in level 2 (i.e. TRD). No pathway survived after multiple-testing correction and the best result was found for the Gene Ontology term regulation of striated muscle contraction (GO:0006942) (permuted comparative $p=0.15$ for association with response). It should be noted that this pathway includes *CACNA1C* and several other genes coding for ion channels as well as other genes that are expressed in the central nervous system and are relevant for neural plasticity. Results referred to GO:0006942 and other pathways are reported in Supplementary Table 6. Despite GO:0006942 was not associated with response after multiple-testing correction, it was the top pathway and 33 genes on a total of 72 genes in this pathway were previously linked to neurogenesis, neural plasticity, long term potentiation (LTP) and/or antidepressant action (Supplementary Table 7). Thus the predictive performance of SNPs in this pathway was tested using machine learning models. 63 SNPs were included after clumping variants according to linkage disequilibrium and exclude SNPs with near-zero variance (see Supplementary Methods). The list of the selected SNPs and their gene annotation is shown in Supplementary Table 8. Neural networks and gradient boosted machine showed the best accuracy among the models tested (Supplementary Figure 1). For the former the best predictive performance was found using one hidden layer and weights decay=0.1. 10 fold cross-validation, repeated 100 times, showed a mean accuracy of 0.73, sensitivity of 0.83, specificity of 0.56, PPV (positive predictive value) of 0.77 and NPV (negative predictive value) of 0.65 in the prediction of TRD. The best observed performance showed accuracy=0.82, sensitivity=0.90, specificity=0.69, PPV=0.83 and NPV=0.79. Gradient boosted machine results were obtained using 150 boosting iterations, maximum number of

interacting predictors (tree depth) of 3, minimum terminal node size of 3 and shrinkage=0.1. 10 fold cross-validation, repeated 100 times, showed a mean accuracy of 0.73, sensitivity of 0.83, specificity of 0.55, PPV of 0.76 NPV of 0.65 in the prediction of TRD. The best observed performance showed accuracy=0.84, sensitivity=0.93, specificity=0.68, PPV=0.83 and NPV=0.85. The 30 SNPs of GO:0006942 pathway showing the highest importance in predicting TRD using gradient boosted machine are shown in Supplementary Figure 2 (results obtained using neural networks were similar). These SNPs are located within 21 different genes and 16 of these genes have been previously linked to neurogenesis, neuroplasticity and/or antidepressant action (Supplementary Table 7). The addition of clinical-demographic variables previously associated with TRD (see Supplementary Methods) did not improve substantially the results (gradient boosted machine showed the same mean accuracy, specificity and PPV with slightly higher mean sensitivity (0.84) and NPV (0.66)).

Table 1. Results of logistic regression in European Sample 1 (ES1), European Sample 2 (ES2), Italian Sample (ITAS) and STAR*D. Genotypic (a) and allelic (b) analysis are reported for SNPs showing $p < 0.05$ in at least two samples. For genotypic analysis the major homozygote was used as reference phenotype. R2 is a measure of linkage disequilibrium. CI=95% confidence interval; ES1=European Sample 1; ES2=European Sample 2; ITAS=Italian Sample; OR=odds ratio.

(a)

Gene	SNP	Sample	Phenotype	Risk Genotype	Statistics
<i>ANK3</i>	rs1049862	ES1	Remission	GA	p=0.03 OR=0.46 (CI 0.22;0.95)
		ITAS	Remission	GA	p=0.04 OR=0.32 (CI 0.10;0.95)
<i>CACNA1C</i>	rs2283326	ES1	TRD-DC	GG	p=0.03 OR=1.83 (CI 1.08;3.11)
		ITAS	Remission	GG	p=0.036 OR=2.71 (CI 1.09;7.07)
	rs10848635	ES1	Response	AT	p=0.007 OR=0.40 (CI 0.20;0.77)
			Response	TT	p=1.39e-05 OR=0.21(CI 0.10;0.42)
			Remission	TT	p=0.03 OR=0.39(CI 0.17;0.92)

			TRD-W	TT	p=0.042 OR=2.35 (CI 1.07;5.64)
		ES2	TRDA	TT	p=0.03 OR=0.36 (CI 0.14;0.89)
	rs11062157 (R2=0.65 with rs10848635)	STAR*D	Response in level 2	AA	p=0.044, OR=1.31 (CI 1.008; 1.70)
<i>FKBP5</i>	rs9470080	ES2	TRDB	TT	p=0.0469 OR=3.96 (CI 1.06;16.65)
			Response	CT	p=0.01 OR=0.27 (CI 0.09;0.79)
		Remission	CT	p=0.03 OR=0.25 (CI 0.06;0.84)	
	rs9368882 (R2=0.67 with rs9470080)	STAR*D	Remission Level 1	TT	p=0.035, OR=0.82 (CI 0.69;0.98)
	rs3800373	ES1	Response	CC	p=0.046 OR= 0.11(CI 0.005; 0.68)
			ES2	TRDA	CC
		TRDB		CA	p=0.04 OR=2.93 (CI 1.07;8.55)
		TRDB		CC	p=0.02 OR=6.44 (CI 1.51;35.34)
		ITAS	Response	CA	p=0.01 OR=0.28 (CI 0.09;0.73)
			Remission	CA	p=0.03 OR=0.27 (CI 0.07;0.83)
		rs1360780	ES2	Response	TT
	TRDA			TT	p=0.01 OR=4.27 (CI 1.40;13.04)
	TRDB			TT	p=0.02 OR=4.76 (CI 1.36;19.77)
	ITAS		Response	CT	p=0.002 OR=0.22 (CI 0.08;0.55)
			Remission	CT	p=0.003 OR=0.17 (CI 0.05;0.52)

(b)

Gene	SNP	Sample	Phenotype	Risk Allele	Statistics
<i>CACNA1C</i>	rs2283326	ES1	TRD-DC	G	p=0.002 OR=1.83 (CI 1.26;2.66)
		ITAS	Remission	G	p=0.003 OR=2.71 (CI 1.42;5.30)
	rs1006737	ES1	Response	G	p=5.25e-05 OR=0.50 (CI 0.35;0.70)
			Remission	G	p=0.02 OR=0.61 (CI 0.40;0.93)
		ES2	TRDA	G	p=0.04 OR=0.58 (CI 0.35; 0.97)
	rs10848635	ES1	Response	T	p=0.007 OR=0.47 (CI 0.34;0.65)
			Remission	T	p=0.021 OR=0.61 (CI 0.41;0.93)
		ES2	TRDA	T	p=0.01 OR= 0.53 (CI 0.32;0.88)
	<i>CACNB2</i>	rs2799573	ES1	Response	T
ITAS			Response	T	p=0.01 OR=3.32 (CI 1.48;8.23)
			Remission	T	p=0.049 OR=2.57 (CI 1.05;7.03)
<i>FKBP5</i>	rs3800373	ES1	Response	C	p=0.03 OR=0.50 (CI 0.26;0.93)
		ES2	Response	C	p=0.02 OR= 0.52 (CI 0.29; 0.91)
			TRDA	C	p=0.005 OR=2.35 (CI 1.29;4.27)
			TRDB	C	p=0.002 OR= 3.07 (CI 1.52;6.42)
		ITAS	Response	C	p=0.03 OR=0.45 (CI 0.21;0.91)
	rs1360780	ES2	Response	T	p=0.02 OR=0.60 (CI 0.38;0.93)
			TRDA	T	p=0.01 OR=1.95 (CI 1.16;3.26)
			TRDB	T	p=0.01 OR= 2.05 (CI 1.17;3.60)
		ITAS	Response	T	p=0.009 OR=0.40 (CI 0.20;0.78)
			Remission	T	p=0.04 OR=0.45 (CI 0.20;0.95)

4. Discussion

The present study combined a candidate gene approach and a genome-wide approach to investigate the role of eight strong candidate genes in antidepressant response. Polymorphisms in these genes have been recognized to have a pleiotropic effect across different psychiatric phenotypes in previous GWAS (4–7). SNPs showing nominally replicated or significant associations were concentrated in the *FKBP5* and *CACNA1C* genes. *FKBP5* rs3800373 AA genotype and A allele, rs1360780 CC/CT genotypes and C allele and rs9470080 CC genotype showed consistent but nominal associations with better treatment outcomes in two or three original samples, including the risk of TRD. A SNP in high LD with rs9470080 (rs9368882) had nominal association with remission in STAR*D Level 1. rs3800373 C allele, rs1360780 T allele and rs9470080 T allele have been associated with enhanced cortisol-induced expression of *FKBP5* leading to an increased GR resistance and decreased efficiency of the negative feedback of the stress hormone axis. These alleles may increase the risk of MDD, bipolar disorder and stress-related disorders (30). Previous studies mainly reported a protective effect of rs1360780 T or TT when early antidepressant response was considered (Binder et al., 2004; Kirchheiner et al., 2008) or they reported negative findings (32,51–53). In mixed ethnic groups, a better response was reported in CC subjects (49). For *FKBP5* rs3800373, C allele carriers were reported to have a trend of better response to antidepressants during the first 2-3 weeks of treatment (30,33), while other studies found no association between this SNP and clinical outcomes (31,51). Previous associations between rs3800373 C / rs1360780 T and antidepressant response were found during the early phase of treatment (2-3 weeks), while the present results contributed in clarifying the role of these SNPs later in treatment and in TRD patients. In TRD hypothalamic-pituitary-adrenal (HPA) axis activity is reset at a higher level, although feedback remains intact (54,55) while increased resistance of the HPA axis to the feedback of corticosteroids was reported in broad MDD (56). These observations suggest that TRD patients may suffer from specific detrimental effects of *FKBP5* polymorphisms that reduce the efficiency of the HPA negative feedback.

CACNA1C encodes for the α -1C subunit of the L-type voltage-dependent calcium channel. Transient increases of calcium intracellular concentration through that channel appears to trigger changes in intracellular signaling, gene transcription, and synaptic plasticity (10). rs2283326, rs1006737 and rs10848635 were found to affect antidepressant treatment outcomes in at least two of the analyzed samples, but only rs1006737 and rs10848635 showed significant effects after multiple-testing correction. rs2283326 was associated also with TRD-DC. In STAR*D, a

SNP in high LD with rs10848635 (rs11062157) was nominally associated with response in Level 2. However, the direction of association was not consistent across the three samples. Among these three SNPs, rs1006737 was previously associated with schizophrenia (57) and bipolar disorder (58), with greater risk of disease for the A allele compared to the G allele. On the contrary, the G allele was shown to increase the risk of ASD (59). In our study, the A allele was associated with a better outcome of antidepressant treatment compared to the G allele in ES1, while the opposite was found in ES2. Despite *CACNA1C* SNPs were not previously associated with antidepressant response individually, a previous GWAS supported the hypothesis that the inorganic cation transmembrane transporter activity pathway (GO:0022890) is associated with antidepressant response and it includes *CACNA1A*, *CACNA1C*, and *CACNB2* genes (60). In line with those results, the present study suggested that the GO:0006942 pathway (that includes *CACNA1C* and other ion channels) may be a useful predictor of TRD. The regulation of calcium currents through L-type calcium channels was demonstrated to play a critical role for the expression of a variety of forms of neuroplasticity (61). A number of genes in this pathway (33 genes on a total of 72 genes) were linked to long-term potentiation, neural survival, neurogenesis and neuroplasticity, but also to MDD and antidepressant efficacy as detailed in Supplementary Table 7. Machine learning is an interesting approach to study the contribution of genetic variants to complex traits such as antidepressant response or TRD. It has been previously applied using clinical-demographic predictors of these phenotypes (62,63) and combinations of few candidate polymorphisms and clinical variables (64), but not to more numerous sets of polymorphisms. To the best of our knowledge, this is the first study to test the predictive properties of a set of polymorphisms using genome-wide data from a candidate pathway. Results suggested good sensitivity and PPV in the detection of TRD but not so good specificity and NPV. The addition of clinical-demographic predictors (gender, marital status, current post-traumatic stress disorder, psychotic symptoms at screening) did not improve substantially the prediction of TRD, probably as a result of limited size of the investigated sample in relation to the number of tested predictors.

Other meaningful polymorphisms with trends of association in more than one sample were in the *ANK3* and *CACNB2* genes. *ANK3* rs1049862 AA genotype may predict better remission accordingly to the results obtained in ES1 and ITAS. *ANK3* was recently proposed as a mediator of aging, stress response, depressive symptoms and antidepressant effects (65). rs2799573 *CACNB2* showed nominal

associations with clinical outcomes in two samples but the direction of association was not consistent.

The finding that some SNPs in the *CACNA1C* and *CACNB2* genes showed inconsistent direction of effect in different samples needs further discussion. Different explanations can be hypothesized: 1) these results may be false positives; 2) the risk allele/genotype may depend from unknown confounding factors (the modulating effect of other polymorphisms or non-genetic factors). When the latter involves genetic confounders is also called flip-flop phenomenon and it is due to multilocus effects and variation in interlocus correlation. In other words, flip-flop associations can occur when the investigated variant is correlated, through interactive effects or linkage disequilibrium, with a causal variant at another locus, as previously described (66).

Polymorphisms in the other genes of interest (*GRM7*, *TCF4*, *ITIH3*, *SYNE1*) did not show significant or replicated findings across samples.

The limitations and strengths of this study should be considered. The relatively limited size of three original samples, particularly the Italian one, represents a limitation of this study. It was not possible to combine the original samples because of different trial designs. On the other hand, the investigated genes were selected among the best candidates according to recent genome-wide association studies thus providing a higher pre-test probability compared to other genes. Further, replication across independent samples was used in order to control for false positive findings. The analyzed samples were heterogeneous in terms of treatment and other clinical variables, but this may be seen as a strength since the obtained results may be applicable across different clinical contexts. Treatment adherence was monitored by inquiry during follow-up visits but no dosage of antidepressant plasma level was determined. In ES1 and ES2 there was a very small percent of non-Caucasian subjects, but previous analyses demonstrated that the exclusion of these subjects did not affect the results (42). In STAR*D, only Caucasian subjects were selected. Finally, machine learning models had good sensitivity and PPV in the detection of TRD but not so good specificity and NPV, and no replication in an independent sample was performed. On the other hand, pathway analysis was not previously performed in STAR*D level 2 and this was the first study to use genome-wide variants from a candidate pathway to predict TRD. The addition of clinical predictors to the model did not substantially affect the prediction of TRD risk. This suggests that genetic variants may provide good prediction by themselves with a plateau effect or our sample (n=620) may not be large enough to evaluate a high number of predictors. On the other hand, the number of clinical predictors included was clearly limited compared to the genetic

ones, since this study was primarily a pharmacogenetic study. Alternatively, the investigated genes may themselves be modulating in a complex way the clinical predictors resulting in a reduced impact on TRD when analyzed together. In conclusion, the present results are in line with the hypothesis that *FKBP5* and *CACNA1C* polymorphisms may play a role in antidepressant response and TRD. Polymorphisms in these genes are good candidates for pharmacogenetic tests aimed to predict antidepressant response and TRD. Accordingly, a previous study has suggested the clinical utility of *FKBP5* polymorphisms for guiding antidepressant treatment (29). Pathways including *CACNA1C* and other genes relevant for neurotransmission, long-term potentiation, neurogenesis and neuroplasticity may provide the basis for developing multi-marker predictors of TRD. This study may pave the way to the identification of sets of genetic predictors in specific pathways able to predict the risk of TRD. It is reasonable to hypothesize a certain degree of variability in the genetic variants involved in TRD across different patients, but the involved pathways are expected to be more stable. Validated genetic markers of TRD could have a pivotal role in the implementation of personalized antidepressant treatments.

Supplementary materials:

Supplementary materials are available for the online version of this publication.

References

1. JUDD LL, AKISKAL HS, ZELLER PJ et al. Psychosocial disability during the long-term course of unipolar major depressive disorder. *Arch Gen Psychiatry* 2000;**57**:375–380.
2. TANSEY KE, GUIPPONI M, HU X et al. Contribution of Common Genetic Variants to Antidepressant Response. *Biol Psychiatry* 2013;**73**:679–682.
3. FABBRI C, MARSANO A, ALBANI D et al. PPP3CC gene: a putative modulator of antidepressant response through the B-cell receptor signaling pathway. *Pharmacogenomics J* 2014;**14**:463–472.
4. MÜHLEISEN TW, LEBER M, SCHULZE TG et al. Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat Commun* 2014;**5**:3339.
5. SCHIZOPHRENIA WORKING GROUP OF THE PSYCHIATRIC GENOMICS CONSORTIUM. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014;**511**:421–427.
6. GREEN EK, GROZEVA D, FORTY L et al. Association at SYNE1 in both bipolar disorder and recurrent major depression. *Mol Psychiatry* 2013;**18**:614–617.
7. CROSS-DISORDER GROUP OF THE PSYCHIATRIC GENOMICS CONSORTIUM. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet (London, England)* 2013;**381**:1371–1379.
8. GENDEP INVESTIGATORS R, MARS INVESTIGATORS KE, STAR*D INVESTIGATORS N et al. Common genetic variation and antidepressant efficacy in major depressive disorder: a meta-analysis of three genome-wide pharmacogenetic

- studies. *Am J Psychiatry* 2013;**170**:207–217.
9. BIERNACKA JM, SANGKUH K, JENKINS G et al. The International SSRI Pharmacogenomics Consortium (ISPC): a genome-wide association study of antidepressant treatment response. *Transl Psychiatry* 2015;**5**:e553.
 10. BHAT S, DAO DT, TERRILLION CE et al. CACNA1C (Cav1.2) in the pathophysiology of psychiatric disease. *Prog Neurobiol* 2012;**99**:1–14.
 11. CROSS-DISORDER GROUP OF THE PSYCHIATRIC GENOMICS CONSORTIUM. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 2013;**381**:1371–1379.
 12. SOLDATOV NM. CACNB2: An Emerging Pharmacological Target for Hypertension, Heart Failure, Arrhythmia and Mental Disorders. *Curr Mol Pharmacol* 2015;**8**:32–42.
 13. SHINOZAKI G, POTASH JB. New Developments in the Genetics of Bipolar Disorder. *Curr Psychiatry Rep* 2014;**16**:493.
 14. HAYASHI A, LE GAL K, SÖDERSTEN K, VIZLIN-HODZIC D, ÅGREN H, FUNA K. Calcium-dependent intracellular signal pathways in primary cultured adipocytes and ANK3 gene variation in patients with bipolar disorder and healthy controls. *Mol Psychiatry* 2015;**20**:931–940.
 15. RANGARAJU S, LEVEY DF, NHO K et al. Mood, stress and longevity: convergence on ANK3. *Mol Psychiatry* 2016;**21**:1037–1049.
 16. NIE F, WANG X, ZHAO P et al. Genetic analysis of SNPs in *CACNA1C* and *ANK3* gene with schizophrenia: A comprehensive meta-analysis. *Am J Med Genet Part B Neuropsychiatr Genet* 2015;**168**:637–648.
 17. COTTRELL JR, BOROK E, HORVATH TL, NEDIVI E. CPG2. *Neuron* 2004;**44**:677–690.
 18. GREEN EK, GROZEVA D, FORTY L et al. Association at SYNE1 in both bipolar disorder and recurrent major depression. *Mol Psychiatry* 2013;**18**:614–617.
 19. NAVARRETE K, PEDROSO I, DE JONG S et al. *TCF4* (*e2-2*; *ITF2*): A schizophrenia-associated gene with pleiotropic effects on human disease. *Am J Med Genet Part B Neuropsychiatr Genet* 2013;**162**:1–16.
 20. FORREST MP, HILL MJ, QUANTOCK AJ, MARTIN-RENDON E, BLAKE DJ. The emerging roles of TCF4 in disease and development. *Trends Mol Med* 2014;**20**:322–331.
 21. RIPKE S, SANDERS AR, KENDLER KS et al. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011;**43**:969–976.
 22. HALL M-H, LEVY DL, SALISBURY DF et al. Neurophysiologic effect of GWAS derived schizophrenia and bipolar risk variants. *Am J Med Genet Part B Neuropsychiatr Genet* 2014;**165**:9–18.
 23. MOSSAKOWSKA-WÓJCIK J, ORZECZOWSKA A, TALAROWSKA M, SZEMRAJ J, GALECKI P. The importance of TCF4 gene in the etiology of recurrent depressive disorders. *Prog Neuro-Psychopharmacology Biol Psychiatry* Published Online First: 21 March 2017. doi:10.1016/j.pnpbp.2017.03.014
 24. SHYN SI, SHI J, KRAFT JB et al. Novel loci for major depression identified by genome-wide association study of Sequenced Treatment Alternatives to Relieve Depression and meta-analysis of three studies. *Mol Psychiatry* 2011;**16**:202–215.
 25. KANDASWAMY R, MCQUILLIN A, CURTIS D, GURLING H. Allelic association, DNA resequencing and copy number variation at the metabotropic glutamate receptor *GRM7* gene locus in bipolar disorder. *Am J Med Genet Part B Neuropsychiatr*

- Genet 2014;**165**:365–372.
26. SHYN SI, HAMILTON SP. The genetics of major depression: moving beyond the monoamine hypothesis. *Psychiatr Clin North Am* 2010;**33**:125–140.
 27. FABBRI C, DRAGO A, SERRETTI A. Early antidepressant efficacy modulation by glutamatergic gene variants in the STAR*D. *Eur Neuropsychopharmacol* 2013;**23**:612–621.
 28. STEVENSON JM, REILLY JL, HARRIS MSH et al. Antipsychotic pharmacogenomics in first episode psychosis: a role for glutamate genes. *Transl Psychiatry* 2016;**6**:e739.
 29. STAMM TJ, RAMPP C, WIETHOFF K et al. The FKBP5 polymorphism rs1360780 influences the effect of an algorithm-based antidepressant treatment and is associated with remission in patients with major depression. *J Psychopharmacol* 2016;**30**:40–47.
 30. BINDER EB. The role of FKBP5, a co-chaperone of the glucocorticoid receptor in the pathogenesis and therapy of affective and anxiety disorders. *Psychoneuroendocrinology* 2009;**34**:S186–S195.
 31. ELLSWORTH KA, MOON I, ECKLOFF BW et al. FKBP5 genetic variation. *Pharmacogenet Genomics* 2013;**23**:156–166.
 32. LEKMAN M, LAJE G, CHARNEY D et al. The FKBP5-Gene in Depression and Treatment Response—an Association Study in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Cohort. *Biol Psychiatry* 2008;**63**:1103–1110.
 33. KIRCHHEINER J, LORCH R, LEBEDEVA E et al. Genetic variants in *FKBP5* affecting response to antidepressant drug treatment. *Pharmacogenomics* 2008;**9**:841–846.
 34. HORSTMANN S, LUCAE S, MENKE A et al. Polymorphisms in *GRIK4*, *HTR2A*, and *FKBP5* Show Interactive Effects in Predicting Remission to Antidepressant Treatment. *Neuropsychopharmacology* 2010;**35**:727–740.
 35. STAMM TJ, RAMPP C, WIETHOFF K et al. The FKBP5 polymorphism rs1360780 influences the effect of an algorithm-based antidepressant treatment and is associated with remission in patients with major depression. *J Psychopharmacol* 2016;**30**:40–47.
 36. APPEL K, SCHWAHN C, MAHLER J et al. Moderation of adult depression by a polymorphism in the FKBP5 gene and childhood physical abuse in the general population. *Neuropsychopharmacology* 2011;**36**:1982–1991.
 37. SCHOSSER A, SERRETTI A, SOUERY D et al. European Group for the Study of Resistant Depression (GSRD)--where have we gone so far: review of clinical and genetic findings. *Eur Neuropsychopharmacol* 2012;**22**:453–468.
 38. SOUERY D, OSWALD P, MASSAT I et al. Clinical factors associated with treatment resistance in major depressive disorder: results from a European multicenter study. *J Clin Psychiatry* 2007;**68**:1062–1070.
 39. SOUERY D, CALATI R, PAPAGEORGIOU K et al. What to expect from a third step in treatment resistant depression: A prospective open study on escitalopram. *World J Biol Psychiatry* 2015;**16**:472–482.
 40. FABBRI C, MARSANO A, BALESTRI M, DE RONCHI D, SERRETTI A. Clinical features and drug induced side effects in early versus late antidepressant responders. *J Psychiatr Res* 2013;**47**:1309–1318.
 41. HOWLAND RH. Sequenced Treatment Alternatives to Relieve Depression (STAR*D). Part 1: study design. *J Psychosoc Nurs Ment Health Serv* 2008;**46**:21–

- 24.
42. TRIVEDI MH, RUSH AJ, IBRAHIM HM et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med* 2004;**34**:73–82.
 43. THASE ME. The need for clinically relevant research on treatment-resistant depression. *J Clin Psychiatry* 2001;**62**:221–224.
 44. CHMP. Note for guidance on clinical investigation of medicinal products in the treatment of depression. 2002.
 45. FABBRI C, CRISAFULLI C, CALATI R et al. Neuroplasticity and second messenger pathways in antidepressant efficacy: pharmacogenetic results from a prospective trial investigating treatment resistance. *Eur Arch Psychiatry Clin Neurosci* Published Online First: 4 March 2017. doi:10.1007/s00406-017-0766-1
 46. PURCELL S, NEALE B, TODD-BROWN K et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet* 2007;**81**:559–575.
 47. DE LEEUW CA, MOOIJ JM, HESKES T, POSTHUMA D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Comput Biol* 2015;**11**:e1004219.
 48. LIBERZON A, BIRGER C, THORVALDSDÓTTIR H, GHANDI M, MESIROV JP, TAMAYO P. The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Syst* 2015;**1**:417–425.
 49. NIITSU T, FABBRI C, BENTINI F, SERRETTI A. Pharmacogenetics in major depression: A comprehensive meta-analysis. *Prog Neuro-Psychopharmacology Biol Psychiatry* 2013;**45**:183–194.
 50. BINDER EB, SALYAKINA D, LICHTNER P et al. Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. *Nat Genet* 2004;**36**:1319–1325.
 51. SARGINSON JE, LAZZERONI LC, RYAN HS, SCHATZBERG AF, MURPHY GM. FKBP5 polymorphisms and antidepressant response in geriatric depression. *Am J Med Genet Part B Neuropsychiatr Genet* 2010;**153B**:554–560.
 52. TSAI S-J, HONG C-J, CHEN T-J, YU YW-Y. Lack of supporting evidence for a genetic association of theFKBP5 polymorphism and response to antidepressant treatment. *Am J Med Genet Part B Neuropsychiatr Genet* 2007;**144B**:1097–1098.
 53. PAPIOL S, ARIAS B, GAST? C, GUTI?RREZ B, CATAL?N R, FA?AN?S L. Genetic variability at HPA axis in major depression and clinical response to antidepressant treatment. *J Affect Disord* 2007;**104**:83–90.
 54. JURUENA MF, CLEARE AJ, PAPADOPOULOS AS, POON L, LIGHTMAN S, PARIANTE CM. The prednisolone suppression test in depression: dose-response and changes with antidepressant treatment. *Psychoneuroendocrinology* 2010;**35**:1486–1491.
 55. JURUENA MF, PARIANTE CM, PAPADOPOULOS AS, POON L, LIGHTMAN S, CLEARE AJ. Prednisolone suppression test in depression: prospective study of the role of HPA axis dysfunction in treatment resistance. *Br J Psychiatry* 2009;**194**:342–349.
 56. HERBERT J. Cortisol and depression: three questions for psychiatry. *Psychol Med* 2013;**43**:449–469.
 57. PORCELLI S, LEE S-J, HAN C, PATKAR AA, SERRETTI A, PAE C-U. CACNA1C gene and schizophrenia. *Psychiatr Genet* 2015;**25**:163–167.
 58. OU X, CRANE DE, MACINTOSH BJ et al. CACNA1C rs1006737 genotype and

- bipolar disorder: Focus on intermediate phenotypes and cardiovascular comorbidity. *Neurosci Biobehav Rev* 2015;**55**:198–210.
59. LI J, ZHAO L, YOU Y et al. Schizophrenia Related Variants in CACNA1C also Confer Risk of Autism. *PLoS One* 2015;**10**:e0133247.
 60. COCCHI E, FABBRI C, HAN C et al. Genome-wide association study of antidepressant response: involvement of the inorganic cation transmembrane transporter activity pathway. *BMC Psychiatry* 2016;**16**:106.
 61. FRANK CA. How voltage-gated calcium channels gate forms of homeostatic synaptic plasticity. *Front Cell Neurosci* 2014;**8**:40.
 62. KAUTZKY A, BALDINGER-MELICH P, KRANZ GS et al. A New Prediction Model for Evaluating Treatment-Resistant Depression. *J Clin Psychiatry* 2017;**78**:215–222.
 63. INIESTA R, MALKI K, MAIER W et al. Combining clinical variables to optimize prediction of antidepressant treatment outcomes. *J Psychiatr Res* 2016;**78**:94–102.
 64. KAUTZKY A, BALDINGER P, SOUERY D et al. The combined effect of genetic polymorphisms and clinical parameters on treatment outcome in treatment-resistant depression. *Eur Neuropsychopharmacol* 2015;**25**:441–453.
 65. RANGARAJU S, LEVEY DF, NHO K et al. Mood, stress and longevity: convergence on ANK3. *Mol Psychiatry* 2016;**21**:1037–1049.
 66. LIN P-I, VANCE JM, PERICAK-VANCE MA, MARTIN ER. No gene is an island: the flip-flop phenomenon. *Am J Hum Genet* 2007;**80**:531–538.

PART 2:

Genome-wide association studies of antidepressant response

Chapter 4

New insights on the pharmacogenomics of antidepressant response from the GENDEP and STAR*D studies: rare variant analysis and high-density imputation

Chiara Fabbri, Katherine E. Tansey, Roy H. Perlis, Joanna Hauser, Neven Henigsberg, Wolfgang Maier, Ole Mors, Anna Placentino, Marcella Rietschel, Daniel Souery, Gerome Breen, Charles Curtis, Lee Sang-Hyuk, Stephen Newhouse, Hamel Patel, Michel Guipponi, Nader Perroud, Guido Bondolfi, Michael O'Donovan, Glyn Lewis, Joanna M. Biernacka, Richard M. Weinshilboum, Anne Farmer, Katherine J. Aitchison, Ian Craig, Peter McGuffin, Rudolf Uher, Cathryn M. Lewis

Published: Fabbri et al. New insights on the pharmacogenomics of antidepressant response from the GENDEP and STAR*D studies: rare variant analysis and high-density imputation. *The pharmacogenomics Journal* 2018; 18:413-421.

Abstract

Genome-wide association studies have generally failed to identify polymorphisms associated with antidepressant response. Possible reasons include limited coverage of genetic variants that this study tried to address by exome genotyping and dense imputation.

A meta-analysis of Genome-Based Therapeutic Drugs for Depression (GENDEP) and Sequenced Treatment Alternatives to Relieve Depression (STAR*D) studies was performed at SNP, gene and pathway level. Coverage of genetic variants was increased compared to previous studies by adding exome genotypes to previously available genome-wide data and using the Haplotype Reference Consortium panel for imputation. Standard quality control was applied. Phenotypes were symptom improvement and remission after 12 weeks of antidepressant treatment. Significant findings were investigated in NEWMEDS consortium samples and Pharmacogenomic Research Network Antidepressant Medication Pharmacogenomic Study (PGRN-AMPS) for replication.

7,062,950 SNPs were analysed in GENDEP (n=738) and STAR*D (n=1409). rs116692768 (p=1.80e-08, *ITGA9* (integrin alpha 9)) and rs76191705 (p=2.59e-08, *NRXN3* (neurexin 3)) were significantly associated with symptom improvement during citalopram/escitalopram treatment. At gene level, no consistent effect was found. At pathway level, the Gene Ontology terms GO:0005694 (chromosome) and GO:0044427 (chromosomal part) were associated with improvement (corrected p=0.007 and 0.045, respectively). The association between rs116692768 and symptom improvement was replicated in PGRN-AMPS (p=0.047), while rs76191705 was not. The two SNPs did not replicate in NEWMEDS.

ITGA9 codes for a membrane receptor for neurotrophins and *NRXN3* is a transmembrane neuronal adhesion receptor involved in synaptic differentiation. Despite their meaningful biological rationale for being involved in antidepressant effect, replication was partial. Further studies may help in clarifying their role.

Keywords: antidepressant, major depression, GWAS, gene, pathway, integrin, neurexin

1. Introduction

Major depressive disorder (MDD) became one of the five leading diseases contributing to disability-adjusted life years (DALYs) in 2010 in the US (1). MDD is associated with a huge increase in suicide risk (2), poor quality of life (comparable to that observed in severe physical disorders such as arthritis and heart disease (3)) and health expenditure (direct costs alone amount to 42 billion dollars per year in Europe (4)).

Despite the availability of antidepressant drugs belonging to different classes, high inter-individual variability is observed in response. The lack of reliable and reproducible markers of treatment outcome contributes to unsatisfactory response and remission rates as well as to side effect burden, poor treatment adherence and early treatment discontinuation (5). Following the observation that antidepressant response clusters in families, genetic variants were considered promising biomarkers to tailor antidepressant treatments and improve the prognosis of MDD (6, 7). Genome-wide association studies (GWAS) were a promising tool to identify the polymorphisms involved in antidepressant response after the overall contradictory and non-replicated findings of candidate gene studies (8). But GWAS results fell below expectations, with no genome-wide significant signal ($p < 5e-08$) that was replicated in different samples (9-15). Possible reasons for these disappointing results include: 1) limited coverage of genetic variants (e.g. ~ 500 K common polymorphisms were originally analyzed in STAR*D, GENDEP and MARS studies and ~ 1.2 millions in the meta-analysis of these studies thanks to imputation, while ~ 40 millions polymorphisms are known to date thanks to sequencing studies (9-11) (12, 16); 2) limited sample size; 3) sample heterogeneity (e.g. different subtypes of depression and severity, different antidepressants); 4) analysis of common variants (minor allele frequency (MAF) > 0.01) alone. Further, previous GWAS meta-analyses focused on single marker analysis and pathway analysis was performed in single samples and not as meta-analysis among different samples (12, 17).

Considering these limitations, the current study aimed to:

- 1) Increase the coverage of genetic variants by analysing exonic polymorphisms and dense imputation;
- 2) Test for association with antidepressant response at the level of SNPs, genes and pathways including both common and rare variants;
- 3) Reduce heterogeneity in treatment by analysing patients treated with the same antidepressant, as performed previously (e.g. (10, 17)).

2. Materials and methods

2.1 Samples

2.1.1 GENDEP

The Genome-Based Therapeutic Drugs for Depression (GENDEP) project was a 12-week partially randomized open-label pharmacogenetic study with two active treatment arms. 867 patients with unipolar depression (ICD-10 or DSM-IV criteria) aged 19–72 years were recruited at nine European centres. Eligible participants were allocated to flexible-dosage treatment with either escitalopram (10–30 mg daily, 499 subjects) or nortriptyline (50–150 mg daily, 368 subjects). Severity of depression was assessed weekly by the Montgomery-Asberg Depression Rating Scale (MADRS) (18), Hamilton Rating Scale for Depression (HRSD–17) (19) and Beck Depression Inventory (BDI) (20). Detailed information about the GENDEP study has been previously reported (10). 736 subjects had no missing phenotypes and covariates after quality control (see quality control details in paragraph 2.3. Genotyping and imputation) and they were included in this study.

2.1.2 STAR*D

The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study was a NIMH-funded study aimed to determine the effectiveness of different treatments for patients with MDD who have not responded to the first antidepressant treatment. Non-psychotic MDD (DSM-IV criteria) patients with age between 18 and 75 years were enrolled from primary care or psychiatric outpatient clinics. Severity of depression was assessed using the 16-item Quick Inventory of Depressive Symptomatology–Clinician Rated (QIDS-C16) (21) at baseline, weeks 2, 4, 6, 9, and 12, while HRSD–17 was administered at each level entry and exit. All patients received citalopram in level 1 and the present study is based on level 1 data. 1953 patients were included in the genetic study. Detailed description of the study design and population are reported elsewhere (22). 1409 subjects had no missing phenotypes and covariates after quality control (see quality control details in paragraph 2.3. Genotyping and imputation) and they were included in this study.

2.1.3 Replication samples

NEWMEDS consortium (<http://www.newmeds-europe.com>) (23) samples other than GENDEP (17) and PGRN-AMPS (Pharmacogenomic Research Network Antidepressant Medication Pharmacogenomic Study) sample (14) were used for replication of significant findings obtained in the GENDEP-STAR*D meta-analysis.

As part of the NEWMEDS consortium (for further details see (17)), three studies conducted by academic institutions (GENDEP, see paragraph 2.1.1; GENPOD, a randomized controlled trial of two active antidepressants, n=601 (24); and GODS,

a treatment cohort of severe depression, n=131 (25)) and two studies by pharmaceutical industry members of the European Federation of Pharmaceutical Industries and Associations (active comparator arms from randomized controlled trials by Pfizer, n=355, and GlaxoSmithKline, n=191) were combined. All included patients were diagnosed with MDD and treated for 6 to 12 weeks with either an antidepressant that acts primarily through blocking the reuptake of serotonin (SSRIs: escitalopram, citalopram, paroxetine, sertraline, fluoxetine) or an antidepressant that acts primarily through blocking the reuptake of norepinephrine (NRIs: nortriptyline, reboxetine), see (17) for details. 980 subjects had no missing phenotypes and covariates after quality control (see quality control details in paragraph 2.3.Genotyping and imputation) and they were included in this study.

PGRN-AMPS included 529 participants with nonpsychotic MDD who were recruited at the Mayo Clinic in Rochester, Minnesota primarily through the inpatient and outpatient practices of the Department of Psychiatry and Psychology. Participants were offered an eight-week course of treatment with either citalopram or escitalopram and depressive symptoms were rated using QIDS-C16 as in the STAR*D study in addition to the HRSD-17. For further details see (14). 492 subjects had no missing phenotypes and covariates after quality control (see quality control details in paragraph 2.3.Genotyping and imputation) and they were included in this study.

2.2 Outcomes

The primary outcome of this study was depressive symptom improvement after 12 weeks of antidepressant treatment. Continuous measures, such as percentage improvement, capture more information and have higher power than cutoff-based dichotomous measures, such as remission, however remission was associated with MDD prognosis (26) (27). The percentage change in scores between baseline and 12 weeks was used to measure symptom improvement, using MADRS in the GENDEP study, and QIDS-C16 in STAR*D, as in previous studies in these samples (10, 28).

As a secondary outcome, we investigated symptom remission after 12 weeks of antidepressant treatment. According to standard criteria, remission was defined as $HRSD-17 \leq 7$ (29) and $QIDS-C16 \leq 5$ (28) in GENDEP and STAR*D, respectively. HRSD-17 was used to define remission in the GENDEP given the stronger consensus about the threshold to identify remission on this scale in contrast to MADRS, where different definitions of remission have been reported (≤ 12 (30), ≤ 10 (31), ≤ 8 (32)).

Each outcome measure (percentage change, remission) was analyzed separately in GENDEP and STAR*D, and then a meta-analysis performed. Two analyses were

performed, initially using all samples, and then including only escitalopram-treated patients from GENDEP, since escitalopram is the active isomer of citalopram (33), the antidepressant used in STAR*D level 1.

Missing data were handled as in previous studies on the investigated samples (9, 10). When at least one post-baseline assessment was available, the percentage improvement at 12 weeks was estimated as the best unbiased estimate of mixed effect linear models. Participants without any post-baseline measurement were excluded from the analyses. Since specific antidepressant response is associated with depression severity (34), a minimum depression severity score of 14 on the HRSD-17 was an inclusion criterion in the STAR*D study (all the subjects that we included satisfied this criteria) but not in GENDEP thus a sensitivity analysis excluding GENDEP subjects with HRSD-17<14 was performed for the validation of significant findings.

2.3.Genotyping and imputation

Genome-wide data available in STAR*D were obtained using Affymetrix Human Mapping 500K Array Set in 969 subjects and Affymetrix Genome-Wide Human SNP Array 5.0 (Affymetrix, South San Francisco, California) in the remaining 979 samples, while in GENDEP Illumina Human610-quad bead chip (Illumina, Inc., San Diego) was used (9, 10). In STAR*D the two groups genotyped by different arrays were balanced by ethnic grouping, gender and proportions of responders and non-responders. Twelve samples were genotyped on both the 500K and 5.0 Arrays, and > 99% concordance was found across these platforms (9). Further genotyping in both samples was performed by the Illumina Infinium Exome-24 v1.0 BeadChip that includes ~ 250K variants. Pre-imputation quality control was performed according to the following criteria: 1) variants with missing rate $\geq 5\%$; 2) monomorphic variants; 3) subjects with genotyping rate < 97%; 4) subjects with gender discrepancies; 5) subjects with abnormal heterozygosity; 6) related subjects (identity by descent (IBD) >0.1875 (35)); 7) population outliers according to Eigensoft analysis of linkage-disequilibrium-pruned genetic data (36, 37); 8) GWAS discordant subjects (referred to exome data only) and 9) non-white subjects (referred to STAR*D only since all subjects included in the GENDEP are of Caucasian ethnicity). Hardy-Weinberg equilibrium was not used as an exclusion criterion for markers (as previously done in the same datasets (17)), since departures from Hardy-Weinberg equilibrium are expected in a case-only study (38).

Data were imputed using Minimac3 as provided by the Michigan imputation Server (<https://imputationserver.sph.umich.edu/start.html>). Post-imputation quality control was performed pruning variants according to the following criteria: 1) poor

imputation quality ($R^2 < 0.30$ (39, 40); 2) minor allele frequency (MAF) < 0.01 (see further details in paragraph 2.4 Statistical analysis).

For a flow chart describing pre- and post-imputation quality control on each dataset see Supplementary Figure 1. Since exome array data were available only in 1015 subjects from the STAR*D sample (after quality control), the imputation of these data was performed separately from the imputation of genome-wide array data (1470 subjects).

2.4 Statistical analysis

We performed a fixed-effects meta-analysis to test the association between single polymorphisms and phenotypes using PLINK (41). Heterogeneity measures (Cochrane's Q statistic and the I^2 heterogeneity index) were calculated. We tested linear or logistic regression models including the ancestry-informative principal components, recruitment centre, age, and baseline severity as covariates in line with previous publications on these samples (10, 12, 42). The same covariates were used for gene and pathway analysis.

We tested the association between genes and phenotypes as well as pathways and phenotypes using MAGMA (43). MAGMA performs both a self-contained and a competitive gene-set analysis, the latter is more conservative and it was applied in this study since it reflects the difference in association between genes in the pathway and genes outside the pathway. Only for the replication analyses self-contained analysis results were also reported since the replication nature of these tests. Both rare (MAF <0.01) and common variants were included in gene and pathway analysis, but only genotyped rare variants were retained while imputed rare variants were excluded. Indeed imputation quality of rare variants using the HRC panel was found to be better than using 1000 Genomes data ($R^2=0.64$ vs $R^2=0.36$ at MAF=0.1% when using a 1M SNP chip) but we preferred to be conservative (44). Thus we chose a relatively conservative approach, including only subjects genotyped on both genome-wide and exome array in the gene and pathway analysis. In the gene- and pathway-level meta-analysis different weights were assigned to polymorphisms according to their MAF, thus higher weight was assigned to rare variants as implemented in MAGMA (43).

For pathway analysis in each dataset and their meta-analysis the reported results refer to a competitive gene-set analysis, which uses a conditional model to correct for confounding due to gene size, gene density and (if applicable) differences in sample size per gene (43). The analysed pathways were downloaded from <http://software.broadinstitute.org/gsea/downloads.jsp> (Biocarta, KEGG, Gene Ontology, Reactome, microRNA targets and transcription factor targets).

An attempt to replicate significant results was performed in NEWMEDS (omitting GENDEP) and PGRN-AMPS. In NEWMEDS phenotypes and covariates were as described in a previous study (17). Briefly, percent symptom improvement at end-point was adjusted for covariates (age, gender, baseline severity, ancestry-informative principal components and centre in case of multi-centric studies) and z-score transformed. Samples were genotyped on Illumina Human610-Quad BeadChips or Illumina Human660W-Quad BeadChips.

For PGRN-AMPS details about phenotypes and covariates were described elsewhere (14). Percent symptom improvement at end-point was used as phenotype and covariates were the first four population principal components and age as in the original GWAS (14). Samples were genotyped on Illumina Human610-Quad BeadChips (Illumina, San Diego, CA).

In both replication samples we performed genotype imputation using the same method applied in GENDEP and STAR*D; pre- and post-imputation quality control were performed according to the same criteria. For replication of individual SNP results, the index SNP and those in linkage disequilibrium ($R^2 \geq 0.30$) were considered.

2.5 Multiple-testing correction and power analysis

For individual SNP analysis, a genome-wide significance threshold was set at $p=5e-08$. A suggestive significance threshold was set at a p value of $5e-06$, which is two orders of magnitude below the genome-wide significance level and approximately corresponds to a level at which one association per genome-wide analysis is expected by chance (45). A nominal p value of 0.05 was set for replication of significant findings in NEWMEDS and PGRN-AMPS since no genome-wide analysis was performed in these samples.

For MAGMA gene level analysis, the False Discovery Rate (FDR) correction was applied. For MAGMA pathway analysis, 50,000 permutations were performed to correct for multiple testing.

For a continuous outcome in the whole sample ($n=2145$) and setting $\alpha=5e-08$, we had 80% power to identify a SNP with an effect size (heritability) of 0.018, while in citalopram-escitalopram treated sample ($n=1739$) we had 80% power to identify an effect size of 0.022 (46).

For a dichotomous phenotype in the whole sample and setting $\alpha=5e-08$, we had 80% power to identify a risk allele with $MAF=0.06$ and $RR=1.50$, while in citalopram-escitalopram treated sample we had 80% power to identify a risk allele with $MAF=0.07$ and $RR=1.50$ (47).

3. Results

The clinical-demographic characteristics of the samples are reported in Supplementary Table 1. 2145 subjects (1409 and 736 from STAR*D and GENDEP, respectively) were included in the SNP-level meta-analysis, 1828 of them were treated with citalopram or escitalopram. 1739 (1003 and 736 from STAR*D and GENDEP, respectively) subjects were included in the gene and pathway meta-analysis, 1422 of them were treated with citalopram or escitalopram.

3.1 SNP analysis results

The GENDEP and STAR*D meta-analysis included 7,062,950 SNPs and showed no evidence of genomic inflation (λ values were ≤ 1.01 , QQ-plots are shown in Supplementary Figure 2).

In the whole sample, no SNP reached the genome-wide significance threshold for association with symptom improvement or remission. Eighty SNPs, from 17 genomic regions, reached suggestive level of association ($p < 5e-06$) (Supplementary Table 2).

In STAR*D and escGENDEP (GENDEP escitalopram) meta-analysis, rs116692768 (MAF=0.033, beta STAR*D=0.14, beta escGENDEP=-0.20, $p=1.87e-08$) and rs76191705 (MAF=0.012, beta STAR*D=-0.26, beta escGENDEP=-0.23, $p=2.39e-08$) were significantly associated with symptom improvement (Figure 1; Table 1). These SNPs are located within introns of the *ITGA9* (integrin alpha 9) and *NRXN3* (neurexin 3) genes. Another *NRXN3* SNP (rs79302561) was close to the significance threshold ($p=6.76e-08$). There was no heterogeneity as measured by I^2 for these three SNPs, thus fixed-effects and random-effects p values corresponded. Each SNP was imputed, with R^2 values of over 0.6 for each SNP in each cohort (Table 1). These SNPs retained significance after the exclusion of GENDEP patients with baseline HRSD-17<14 ($p=2.69e-08$ and $3.43e-08$, respectively). Regional association plots for *ITGA9* and *NRXN3* genes are reported in Figure 2. An overview of SNPs with suggestive level of association ($p < 5e-06$) with symptom improvement in STAR*D and escGENDEP is reported in Supplementary Table 3.

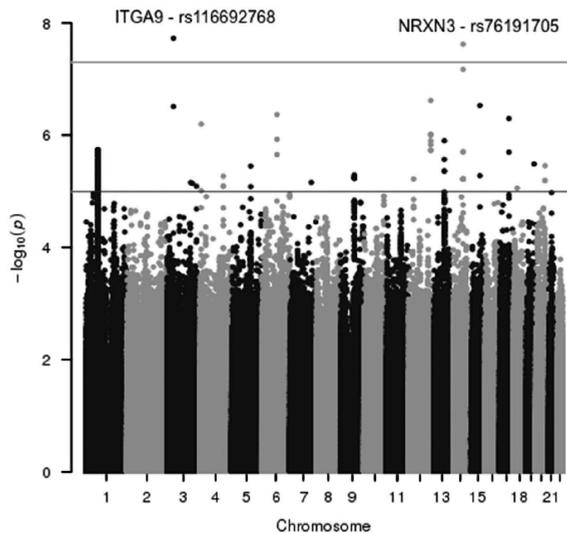
In the analysis of remission, no SNPs reached significance in the full meta-analysis or the meta-analysis with escGENDEP, SNPs with suggestive level of association ($p < 5e-06$) are reported in Supplementary Tables 4-5.

Table 1: SNPs showing genome-wide association with symptom improvement in GENDEP escitalopram-treated sample and STAR*D meta-analysis. Chr=chromosome. Pos=position (GRCh37). NMD=non-sense mediated decay. Q=Cochrane’s Q statistic. I=I2 heterogeneity index. Rsq=measure of imputation quality. MAF=minor allele frequency. Random and fixed p values corresponded for these SNPs.

SNP	Chr	Base pair	Gene	P
rs116692768	3	37540482	<i>ITGA9</i> (intron)	1.87e-08
rs76191705	14	79718462	<i>NRXN3</i> (intron)	2.39e-08

SNP	I	Rsq STARD/ escGENDEP	Beta STARD/ escGENDEP	MAF STARD /escGENDEP
rs116692768	0	0.75 / 0.78	-0.14 / -0.20	0.035 / 0.023
rs76191705	0	0.61 / 0.96	-0.26 / -0.23	0.010 / 0.016

Figure 1: Manhattan plot referred to symptom improvement in STAR*D and GENDEP escitalopram-treated subsample meta-analysis. The two lines indicate the threshold for suggestive findings and genome-wide significant findings.



3.2 Gene analysis results

The gene-level analysis included 17,996 genes with 3,185,134 SNPs, a schematic representation of this analysis was reported in Supplementary Figure 3A.

In the whole sample meta-analysis, the olfactory receptor family 4 subfamily K member 2 (*OR4K2*) gene was associated with symptom improvement after multiple-testing correction (nominal $p=2.13e-06$, corrected $p=0.038$ (FDR)). *OR4K2* included 4 rare genotyped SNPs in both datasets but the overall effect of *OR4K2* rare alleles on improvement was in the opposite direction between the samples. GENDEP subjects carrying rare alleles (rs199718838 A, rs116972349 A, rs151057533 C and rs147651981 T, n=8 subjects) showed lower mean symptom improvement ($18.58\pm 53.43\%$) compared to common alleles carriers ($59.78\pm 24.49\%$), while in STAR*D the opposite was found (fourteen subjects carrying rare alleles (rs199718838 A, rs116972349 A, rs142549715 A and rs150417989 G) had a mean improvement of $70.30\pm 22.23\%$ compared to a mean improvement of $50.33\pm 32.98\%$ in common allele carriers).

No gene survived multiple-testing correction in the analysis of remission (Supplementary Table 6).

In the meta-analysis of STAR*D and escGENDEP, no gene was associated with symptom improvement or remission (Supplementary Table 7). Several genes with nominal $p\leq 0.0005$ overlapped with those found in the whole sample, such as *POUIF1*, *PAG1*, *PKM*, *RPUSD3* and *PARP6*.

3.3 Pathway analysis results

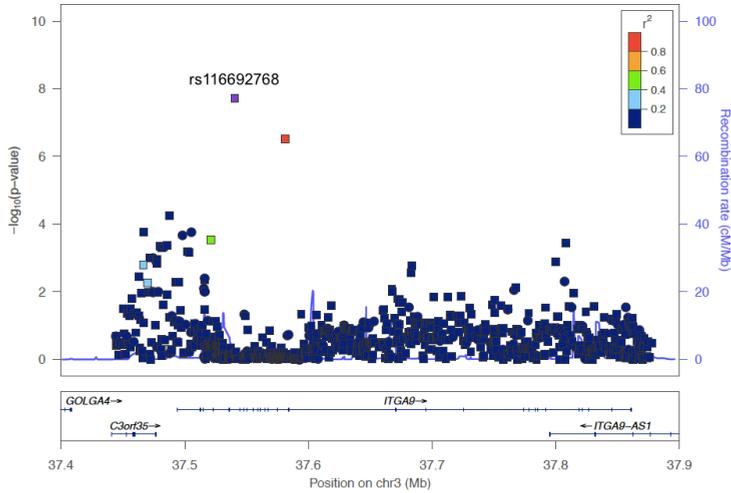
17,996 genes including 3,185,134 SNPs were included in this analysis, a schematic representation of this analysis was reported in Supplementary Figure 3B.

In the whole sample, the Gene Ontology term corresponding to the chromosome pathway (GO:0005694) and the chromosomal part (GO:0044427) pathway were associated with symptom improvement (corrected $p=0.007$ and 0.045 , respectively). No pathway was associated with remission.

In STAR*D and escGENDEP, the steroid hormone receptor signaling pathway (GO:0030518) was very close to the significance threshold for association with remission (corrected $p=0.055$).

An overview of results is reported in Table 2 (including all pathways with $p<0.10$) and the functional role of the variants in each of these pathways is shown in Supplementary Figure 4. Interestingly, the chromosome pathway (GO:0005694) was the richest in rare missense variants (6.6% versus 2.5-3% in the other pathways).

Figure 2: regional association plots referred to ITGA9 (A) and NRXN3 (B) and symptom improvement in STAR*D and GENDEP escitalopram-treated subsample. Plots were obtained using LocusZoom (locuszoom.org).. Imputed SNP are plotted as squares and genotyped (in at least one sample) SNPs as circles. **A**



B

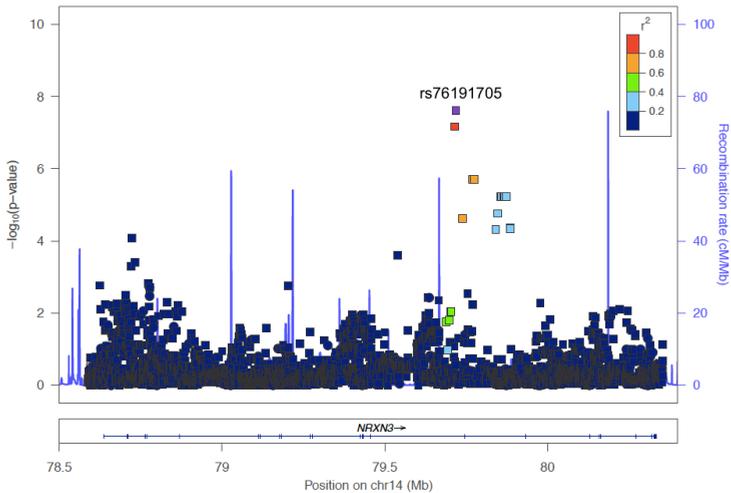


Table 2: results of pathway meta-analysis in the whole sample and STAR*D – GENDEP escitalopram subgroup (CIT-ESCI). Results with corrected (permutated) $p < 0.10$ are reported. For each pathway “Top genes” are those showing nominal $p \leq 0.10$ for association with the phenotype.

Pathway	Sample Phenotype	N genes	Top genes	P
Chromosome pathway (GO:0005694)	Whole sample - improvement	116	NPM2, PAM, BUB1, UPF1, POLG2, NDC80, RFC4, HMGB1, FOXC1, TIPIN, SMC2, TOP1, H2AFY, OIP5	0.0069
Chromosomal part (GO:0044427)		92	NPM2, PAM, BUB1, UPF1, NDC80, RFC4, FOXC1, SMC2, TIPIN, OIP5, H2AFY	0.045
Targets of JUN		247	CCNA2, GPBP1, VGF, GNB4, RAI1, RUNDC3A, AHI1, UBE2H, ALS2, ABHD16A, SIK1, SLC18A2, PDP1, PCSK1, IRF2BPL, ELAVL1, ELOVL5, ANAPC10, TIPRL, B3GALT2, HOXC10, RPS29, ZFAND2B, SYNCRIP, MAP3K13, LGR5, RAB25, TSC22D2, PACRGL, LENG9, GLI1, TGIF2, GPR3, RCE1, ATG5, TRPC1, HHIP, SYT11, RBM18	0.092
Targets of CREB1		250	CCNA2, GPBP1, VGF, GNB4, RAI1, RUNDC3A, AHI1, UBE2H, ALS2, ABHD16A, SIK1, SLC18A2, PDP1, PCSK1, IRF2BPL, ELAVL1, ELOVL5, ANAPC10, TIPRL, HOXC10, RPS29, ZFAND2B, MAP3K13, RBMS2, LGR5, RAB25, TSC22D2, ZNF687, TUBB2B, PACRGL, GLI1, TGIF2, GPR3, RCE1, ATG5, HHIP, SYT11, RBM18	0.094
Steroid hormone	CIT-ESCI –	18	UBR5	0.070

receptor signaling pathway (GO:0030518)	improvement			
Intracellular receptor mediated signaling pathway (GO:0030522)		19	UBR5	0.087
Steroid hormone receptor signaling pathway (GO:0030518)	CIT-ESCI – remission	18	YWHAH, MED13, DAXX	0.055
Intracellular receptor mediated signaling pathway (GO:0030522)		19	YWHAH, MED13, DAXX	0.079

3.4 Replication samples

In both replication samples, only the genome-wide significant findings of STAR*D–escGENDEP meta-analysis and independent SNPs showing $R^2 \geq 0.30$ with them were analysed. Secondly, pathways reported in Table 2 were also investigated. In NEWMEDS replication analyses were performed in the whole sample or only in subjects treated with some antidepressants according to results in the discovery samples. 980 and 492 subjects were included after quality control in NEWMEDS and PGRN-AMPS, respectively.

In NEWMEDS the two SNPs associated with symptom improvement in the STAR*D – escGENDEP meta-analysis were not replicated and weak nominal associations were found for two SNPs (s7152916 and rs7152941) in LD with *NRXN3* rs76191705 in the subsample treated with SSRIs (n=751), while trends of association were found in the subsample treated with citalopram/escitalopram (n=370) (Supplementary Table 8A). In this sample only three independent SNPs in LD with the index SNPs were available.

Over 90% of SNPs included in the pathways showing significant associations (or trends, see Table 2) in the GENDEP-STAR*D meta-analysis were available also in NEWMEDS and only these pathways were analysed in NEWMEDS. In the subsample treated with citalopram or escitalopram (n=370), the intracellular receptor mediated signaling pathway and the steroid hormone receptor signaling pathway were associated with improvement (permutated comparative $p=0.0028$ and $p=0.017$, respectively, self- $p=0.0089$ and 0.014 (self- p values are also reported because only six pathways were analysed here and for a replication purpose, see paragraph 2.4 Statistical analysis for details about comparative and self- p values)), other pathways reported in Table 2 did not show association with phenotypes.

In PGRN-AMPS only *ITGA9* rs116692768 and *NRXN3* rs76191705 were investigated because no SNP showed $LD \geq 0.30$ with them. rs116692768 was associated with symptom improvement in the same direction to STAR*D – escGENDEP meta-analysis ($p=0.047$, Supplementary Table 8B; no multiple testing correction was applied here since only two SNPs were investigated for replication purpose), while rs76191705 showed no association with this phenotype ($p=0.34$). In PGRN-AMPS only the intracellular receptor mediated signaling pathway and the steroid hormone receptor signaling pathway were analysed because all patients were treated with citalopram or escitalopram in this sample. Over 90% of SNPs included in these pathways in STAR*D – escGENDEP meta-analysis was available in PGRN-AMPS, but no association with symptom improvement was found (nominal self- p values were 0.10 and 0.11, respectively).

4. Discussion

4.1 Main findings

The present study has exploited existing pharmacogenetic samples (GENDEP and STAR*D), through: 1) the increase of genetic variant coverage (exome genotyping, high-density imputation) and 2) SNP-, gene- and pathway-level meta-analysis.

In the SNP-level meta-analysis, rs116692768 and rs76191705 were associated with symptom improvement during citalopram/escitalopram treatment (Table 1), rs116692768 was nominally replicated in PGRN-AMPS but none of the two SNPs was replicated in the remaining NEWMEDS samples. The effect sizes of these SNPs may appear relatively larger compared to those reported for other complex traits, consistently with the findings of other GWAS (the median effect size of pharmacogenomics variants was reported to be twice as large as that observed for associations with complex disease risk (48)). These SNPs are located within *ITGA9* (integrin subunit alpha 9) *NRXN3* (neurexin 3), respectively. Both these SNPs are intronic, but the latter is a NMD (non-sense mediated decay) transcript variant (i.e. it is a target of NMD). NMD is a post-transcriptional surveillance process that recognizes and degrades mRNAs containing premature termination codons but it also targets 3–10% of normal transcripts in humans, thereby serving as a widespread gene regulatory mechanism (49-51). rs79302561 T allele was reported to act as an enhancer of gene expression in several cell types (Ensembl GRCh37 release 84). *ITGA9* rs143661452 lies within both *ITGA9* and *ITGA9-AS1* (antisense RNA). Antisense RNAs are involved in multiple regulatory processes in eukaryotes, such as transcriptional interference and RNA masking (52). No regulatory role of rs143661452 is known so far.

The proteins coded by *ITGA9* and *NRXN3* show interesting functions in relation to antidepressant action. Integrins are heterodimeric (one alpha and one beta subunit) transmembrane proteins that connect the extracellular environment to intracellular signaling. In the CNS they are involved in the control of synaptic plasticity, long-term potentiation (LTP), cell adhesion and migration (53, 54). Polymorphisms in another beta isoform (*ITGB3*) and expression level of this gene have been associated with antidepressant response in humans (42, 55). Neurexins are type I transmembrane neuronal adhesion receptors and their interaction with neuroligins is sufficient to trigger postsynaptic and presynaptic differentiation (56).

Non-genome-wide significant SNPs showing suggestive p-values ($<5e-06$) were mainly intergenic but several of them lie in genes previously reported to be involved in mood disorders or antidepressant response (particularly *SORCS2*, *GRIN2D*, *CTNND2*, *CSGALNACT1*, *DISC1*, *TSNAX-DISC1*) (57-66).

The gene-level meta-analysis did not show any consistent finding. Indeed the olfactory receptor family 4 subfamily K member 2 (*OR4K2*) gene reached the significance threshold (Supplementary Table 6) but the effect of rare alleles on symptom improvement had opposite direction between GENDEP and STAR*D. Olfactory receptor genes have been shown to be over-represented among homozygous loss of function (HLOF) genes and segregating polymorphisms of functional and non-functional copies of olfactory genes are common (67). The olfactory receptor family is the most polymorphic family of genes in humans after the major histocompatibility complex and the phenotypic consequences of such genetic variability are not completely clear but likely not dramatic given their large diffusion (68).

Our pathway meta-analysis in GENDEP-STAR*D showed that the chromosome pathway (GO:0005694) and the chromosomal part (GO:0044427) pathway were associated with symptom improvement. These pathways were not associated with improvement in NEWMEDS, but GO:0044427 and particularly GO:0005694 were much richer in rare missense variants compared to the other significant pathways (Supplementary Figure 3) and this may have limited the replicability of the finding since no exome genotyping was performed in NEWMEDS. Interestingly, the chromosomal part pathway has been previously associated with antidepressant response by a genome-wide gene expression study (69). Among the top genes of the chromosome and chromosomal part pathways (Table 2), some are involved in the differentiation of neural stem cells into neurons (*UPFI* (Regulator Of Nonsense Transcripts Homolog), *HMGB1* (High Mobility Group Box 1) and neural development (*FOXCI* (Forkhead Box C1) (70-72)). Thus these genes may play a role in the neurogenesis process that is known to mediate the effect of

antidepressants. Another member of these pathways is *PAM* (Peptidylglycine Alpha-Amidating Monooxygenase) that acts as a regulator of amygdala excitability and synaptic plasticity (73) and consistently it plays a role in emotional responses regulation (74).

The meta-analysis of escGENDEP and STAR*D showed that the steroid hormone receptor signaling pathway and the intracellular receptor mediated signaling pathway (that largely overlap between each other) were close to the significance threshold after permutations. Impaired glucocorticoid receptor (GR) function has been suggested to be causal for HPA axis hyperactivity in MDD that results in impaired neurogenesis and reduction of hippocampal volume. Antidepressants modulate the expression of GR, its translocation to the nucleus and also the transcription of its target genes (75). The steroid hormone receptor signaling pathway and the intracellular receptor mediated signaling pathway were associated with improvement in NEWMEDS citalopram-escitalopram treated subsample but not in PGRN-AMPS. *YWHAH* was one of the top genes of this pathway found in both discovery sample and NEWMEDS. *YWHAH* codes for the η subtype of the 14-3-3 protein family, it is expressed mainly in the brain, and it is a positive regulator in the glucocorticoid signal pathway by blocking the degradation of the GR. Variants in this genes have been associated with bipolar disorder and schizophrenia (76).

4.2 Limitations

The limitations of the present study should be considered. First, the samples included in this study (particularly NEWMEDS samples used for replication) were heterogeneous from several points of view, for example baseline severity, scales used to assess depressive symptoms, time points of evaluation, antidepressant treatment and dose, setting of recruitment. Heterogeneity across samples and small size of replication samples have limited the power to replicate our findings. In GENDEP and part of STAR*D exome genotypes were available (Illumina Infinium Exome-24 v1.0 BeadChip) but these were not genotyped in the replication samples, making it difficult to replicate the gene and pathway analysis findings. Antidepressant response is known to be a heterogeneous phenotype that is affected by a number of genetic and non-genetic variables. In this study we considered age, baseline severity, recruitment centre and ancestry-informative principal components as covariates and we investigated genes and pathways as analysis units in addition to individual SNPs, but there are presumably a number of factors that we did not take into account, such as the effect of the environment. MDD was demonstrated to be genetically heterogeneous, since independent samples showed relatively low genetic correlation compared to schizophrenia and

bipolar disorder (77). The significant SNPs that we identified explained a limited percent of variance in symptom improvement (rs116692768 explained 3.5% and rs76191705 2.7% in escGENDEP, while both variants explained ~ 1.4% of phenotypic variance in STAR*D), that is not expected to be clinically significant. Indeed it has been suggested that a variant explaining 6.3% of variance may be considered as clinically meaningful (78). Placebo response is another issue not taken into account by antidepressant pharmacogenomics trials, despite placebo response was reported to significantly contribute to the antidepressant effect and to have a relevant genetic component (79). Finally, the replication of findings was weak or absent and no validation of the results was obtained through the use of complementary investigation strategies such as gene expression studies.

4.3 Conclusions

The increase in genetic variant coverage seems useful to identify new variants that may influence antidepressant efficacy, but the difficulty in signal replication across different samples is still a problem. Adequate sample size represents a primary issue to allow replication, but also the use of standardized criteria for patient inclusion, treatment and evaluation. The identification of more homogenous groups of patients appears a critical issue in MDD and antidepressant pharmacogenomics. Results of pathway analysis and more in general gene set analysis may be easier to replicate across different samples than individual SNPs because sources of heterogeneity or bias (e.g. genotyping or imputation errors) are expected to have a lower influence and effect sizes are expected to be higher than for individual SNPs. LD score regression and polygenic risk scoring are recent complementary approaches that are expected to provide a relevant contribution to the study of the polygenic complexity of antidepressant response and the sharing of genetic variants with correlated traits, such as MDD, response to other psychotropic drugs or placebo response.

Supplementary materials:

Supplementary materials are available for the online version of this publication.

References

1. Murray CJ, Atkinson C, Bhalla K, Birbeck G, Burstein R, Chou D, et al. The state of US health, 1990-2010: burden of diseases, injuries, and risk factors. *Jama*. 2013;310(6):591-608.
2. Bradvik L, Mattisson C, Bogren M, Nettelbladt P. Long-term suicide risk of depression in the Lundby cohort 1947-1997--severity and gender. *Acta psychiatrica Scandinavica*. 2008;117(3):185-91.
3. Buist-Bouwman MA, De Graaf R, Vollebergh WA, Alonso J, Bruffaerts R, Ormel J, et al. Functional disability of mental disorders and comparison with physical disorders: a

- study among the general population of six European countries. *Acta psychiatrica Scandinavica*. 2006;113(6):492-500.
4. Sobocki P, Jonsson B, Angst J, Rehnberg C. Cost of depression in Europe. *The journal of mental health policy and economics*. 2006;9(2):87-98.
 5. Fabbri C, Di Girolamo G, Serretti A. Pharmacogenetics of antidepressant drugs: an update after almost 20 years of research. *American journal of medical genetics Part B, Neuropsychiatric genetics*. 2013;162B(6):487-520.
 6. O'Reilly RL, Bogue L, Singh SM. Pharmacogenetic response to antidepressants in a multicaser family with affective disorder. *Biological psychiatry*. 1994;36(7):467-71.
 7. Tansey KE, Guipponi M, Hu X, Domenici E, Lewis G, Malafosse A, et al. Contribution of common genetic variants to antidepressant response. *Biological psychiatry*. 2013;73(7):679-82.
 8. Perlis RH. Pharmacogenomic testing and personalized treatment of depression. *Clinical chemistry*. 2014;60(1):53-9.
 9. Garriock HA, Kraft JB, Shyn SI, Peters EJ, Yokoyama JS, Jenkins GD, et al. A genome-wide association study of citalopram response in major depressive disorder. *Biological psychiatry*. 2010;67(2):133-8.
 10. Uher R, Perroud N, Ng MY, Hauser J, Henigsberg N, Maier W, et al. Genome-wide pharmacogenetics of antidepressant response in the GENDEP project. *The American journal of psychiatry*. 2010;167(5):555-64.
 11. Ising M, Lucae S, Binder EB, Bettecken T, Uhr M, Ripke S, et al. A genome-wide association study points to multiple loci that predict antidepressant drug treatment outcome in depression. *Archives of general psychiatry*. 2009;66(9):966-75.
 12. GENDEP Investigators, MARS Investigators, STAR*D Investigators. Common genetic variation and antidepressant efficacy in major depressive disorder: a meta-analysis of three genome-wide pharmacogenetic studies. *The American journal of psychiatry*. 2013;170(2):207-17.
 13. Biernacka JM, Sangkuhl K, Jenkins G, Whaley RM, Barman P, Batzler A, et al. The International SSRI Pharmacogenomics Consortium (ISPC): a genome-wide association study of antidepressant treatment response. *Translational psychiatry*. 2015;5:e553.
 14. Ji Y, Biernacka JM, Hebbiring S, Chai Y, Jenkins GD, Batzler A, et al. Pharmacogenomics of selective serotonin reuptake inhibitor treatment for major depressive disorder: genome-wide associations and functional genomics. *The pharmacogenomics journal*. 2013;13(5):456-63.
 15. Li QS, Tian C, Seabrook GR, Drevets WC, Narayan VA. Analysis of 23andMe antidepressant efficacy survey data: implication of circadian rhythm and neuroplasticity in bupropion response. *Translational psychiatry*. 2016;6(9):e889.
 16. McCarthy S, Das S, Kretzschmar W, Durbin R, Abecasis G, Marchini J. A reference panel of 64,976 haplotypes for genotype imputation 2015 [cited 2016 28th June].
 17. Tansey KE, Guipponi M, Perroud N, Bondolfi G, Domenici E, Evans D, et al. Genetic predictors of response to serotonergic and noradrenergic antidepressants in major depressive disorder: a genome-wide analysis of individual-level data and a meta-analysis. *PLoS medicine*. 2012;9(10):e1001326.
 18. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *The British journal of psychiatry*. 1979;134:382-9.
 19. Hamilton M. Development of a rating scale for primary depressive illness. *The British journal of social and clinical psychology*. 1967;6(4):278-96.

20. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Archives of general psychiatry*. 1961;4:561-71.
21. Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychological medicine*. 2004;34(1):73-82.
22. Rush AJ, Fava M, Wisniewski SR, Lavori PW, Trivedi MH, Sackeim HA, et al. Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. *Controlled clinical trials*. 2004;25(1):119-42.
23. Hughes B. Novel consortium to address shortfall in innovative medicines for psychiatric disorders. *Nature reviews Drug discovery*. 2009;8(7):523-4.
24. Thomas L, Mulligan J, Mason V, Tallon D, Wiles N, Cowen P, et al. GENetic and clinical predictors of treatment response in depression: the GenPod randomised trial protocol. *Trials*. 2008;9:29.
25. Bondolfi G, Aubry JM, Golaz J, Gex-Fabry M, Gervasoni N, Bertschy G. A stepwise drug treatment algorithm to obtain complete remission in depression: a Geneva study. *Swiss medical weekly*. 2006;136(5-6):78-85.
26. Gaynes BN, Warden D, Trivedi MH, Wisniewski SR, Fava M, Rush AJ. What did STAR*D teach us? Results from a large-scale, practical, clinical trial for patients with depression. *Psychiatric services*. 2009;60(11):1439-45.
27. Streiner DL. Breaking up is hard to do: the heartbreak of dichotomizing continuous data. *Canadian journal of psychiatry*. 2002;47(3):262-6.
28. Mrazek DA, Biernacka JM, McAlpine DE, Benitez J, Karpyak VM, Williams MD, et al. Treatment outcomes of depression: the pharmacogenomic research network antidepressant medication pharmacogenomic study. *Journal of clinical psychopharmacology*. 2014;34(3):313-7.
29. Frank E, Prien RF, Jarrett RB, Keller MB, Kupfer DJ, Lavori PW, et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder. Remission, recovery, relapse, and recurrence. *Archives of general psychiatry*. 1991;48(9):851-5.
30. Li H, Gu N, Zhang H, Wang G, Tan Q, Yang F, et al. Efficacy and safety of quetiapine extended release monotherapy in bipolar depression: a multi-center, randomized, double-blind, placebo-controlled trial. *Psychopharmacology*. 2016;233(7):1289-97.
31. Jacobsen PL, Mahableshwarkar AR, Serenko M, Chan S, Trivedi MH. A randomized, double-blind, placebo-controlled study of the efficacy and safety of vortioxetine 10 mg and 20 mg in adults with major depressive disorder. *The Journal of clinical psychiatry*. 2015;76(5):575-82.
32. Moller HJ, Demyttenaere K, Olausson B, Szamosi J, Wilson E, Hosford D, et al. Two Phase III randomised double-blind studies of fixed-dose TC-5214 (dextropropylamine) adjunct to ongoing antidepressant therapy in patients with major depressive disorder and an inadequate response to prior antidepressant therapy. *The world journal of biological psychiatry*. 2015;16(7):483-501.
33. Svensson S, Mansfield PR. Escitalopram: superior to citalopram or a chiral chimera? *Psychotherapy and psychosomatics*. 2004;73(1):10-6.
34. Fournier JC, DeRubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, Shelton RC, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *Jama*. 2010;303(1):47-53.

35. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. *Nature protocols*. 2010;5(9):1564-73.
36. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS genetics*. 2006;2(12):e190.
37. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*. 2006;38(8):904-9.
38. Wittke-Thompson JK, Pluzhnikov A, Cox NJ. Rational inferences about departures from Hardy-Weinberg equilibrium. *American journal of human genetics*. 2005;76(6):967-86.
39. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic epidemiology*. 2010;34(8):816-34.
40. Pistis G, Porcu E, Vrieze SI, Sidore C, Steri M, Danjou F, et al. Rare variant genotype imputation with thousands of study-specific whole-genome sequences: implications for cost-effective study designs. *European journal of human genetics*. 2015;23(7):975-83.
41. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics*. 2007;81(3):559-75.
42. Fabbri C, Crisafulli C, Gurwitz D, Stingl J, Calati R, Albani D, et al. Neuronal cell adhesion genes and antidepressant response in three independent samples. *The pharmacogenomics journal*. 2015;15(6):538-48.
43. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS computational biology*. 2015;11(4):e1004219.
44. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nature genetics*. 2016;48(10):1279-83.
45. Dudbridge F, Gusnanto A. Estimation of significance thresholds for genomewide association scans. *Genetic epidemiology*. 2008;32(3):227-34.
46. Feng S, Wang S, Chen CC, Lan L. GWAPower: a statistical power calculation software for genome-wide association studies with quantitative traits. *BMC genetics*. 2011;12:12.
47. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nature genetics*. 2006;38(2):209-13.
48. Maranville JC, Cox NJ. Pharmacogenomic variants have larger effect sizes than genetic variants associated with other dichotomous complex traits. *The pharmacogenomics journal*. 2016;16(4):388-92.
49. Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutates genomic noise. *Nature genetics*. 2004;36(10):1073-8.
50. Yepiskoposyan H, Aeschimann F, Nilsson D, Okoniewski M, Muhlemann O. Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. *Rna*. 2011;17(12):2108-18.

51. Tani H, Imamachi N, Salam KA, Mizutani R, Ijiri K, Irie T, et al. Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. *RNA biology*. 2012;9(11):1370-9.
52. Werner A, Berdal A. Natural antisense transcripts: sound or silence? *Physiological genomics*. 2005;23(2):125-31.
53. Smith AC, Scofield MD, Kalivas PW. The tetrapartite synapse: Extracellular matrix remodeling contributes to corticoaccumbens plasticity underlying drug addiction. *Brain research*. 2015;1628(Pt A):29-39.
54. Lynch G, Kramar EA, Gall CM. Protein synthesis and consolidation of memory-related synaptic changes. *Brain research*. 2015;1621:62-72.
55. Probst-Schendzielorz K, Scholl C, Efimkina O, Ersfeld E, Viviani R, Serretti A, et al. CHL1, ITGB3 and SLC6A4 gene expression and antidepressant drug response: results from the Munich Antidepressant Response Signature (MARS) study. *Pharmacogenomics*. 2015;16(7):689-701.
56. Bot N, Schweizer C, Ben Halima S, Fraering PC. Processing of the synaptic cell adhesion molecule neurexin-3beta by Alzheimer disease alpha- and gamma-secretases. *The Journal of biological chemistry*. 2011;286(4):2762-73.
57. Christoforou A, McGhee KA, Morris SW, Thomson PA, Anderson S, McLean A, et al. Convergence of linkage, association and GWAS findings for a candidate region for bipolar disorder and schizophrenia on chromosome 4p. *Molecular psychiatry*. 2011;16(3):240-2.
58. Fabbri C, Serretti A. Genetics of long-term treatment outcome in bipolar disorder. *Progress in neuro-psychopharmacology & biological psychiatry*. 2016;65:17-24.
59. Dean B, Gibbons AS, Boer S, Uezato A, Meador-Woodruff J, Scarr E, et al. Changes in cortical N-methyl-D-aspartate receptors and post-synaptic density protein 95 in schizophrenia, mood disorders and suicide. *The Australian and New Zealand journal of psychiatry*. 2016;50(3):275-83.
60. Nivard MG, Mbarek H, Hottenga JJ, Smit JH, Jansen R, Penninx BW, et al. Further confirmation of the association between anxiety and CTNND2: replication in humans. *Genes, brain, and behavior*. 2014;13(2):195-201.
61. Hunter AM, Leuchter AF, Power RA, Muthen B, McGrath PJ, Lewis CM, et al. A genome-wide association study of a sustained pattern of antidepressant response. *Journal of psychiatric research*. 2013;47(9):1157-65.
62. Hashimoto R, Numakawa T, Ohnishi T, Kumamaru E, Yagasaki Y, Ishimoto T, et al. Impact of the DISC1 Ser704Cys polymorphism on risk for major depression, brain morphology and ERK signaling. *Human molecular genetics*. 2006;15(20):3024-33.
63. Thomson PA, Parla JS, McRae AF, Kramer M, Ramakrishnan K, Yao J, et al. 708 Common and 2010 rare DISC1 locus variants identified in 1542 subjects: analysis for association with psychiatric disorder and cognitive traits. *Molecular psychiatry*. 2014;19(6):668-75.
64. Arias B, Fabbri C, Serretti A, Drago A, Mitjans M, Gasto C, et al. DISC1-TSNAX and DAOA genes in major depression and citalopram efficacy. *Journal of affective disorders*. 2014;168:91-7.
65. Schosser A, Gaysina D, Cohen-Woods S, Chow PC, Martucci L, Craddock N, et al. Association of DISC1 and TSNAX genes and affective disorders in the depression case-control (DeCC) and bipolar affective case-control (BACCS) studies. *Molecular psychiatry*. 2010;15(8):844-9.

66. Okuda A, Kishi T, Okochi T, Ikeda M, Kitajima T, Tsunoka T, et al. Translin-associated factor X gene (TSNAX) may be associated with female major depressive disorder in the Japanese population. *Neuromolecular medicine*. 2010;12(1):78-85.
67. Kaiser VB, Svinti V, Prendergast JG, Chau YY, Campbell A, Patarcic I, et al. Homozygous loss-of-function variants in European cosmopolitan and isolate populations. *Human molecular genetics*. 2015;24(19):5464-74.
68. Menashe I, Man O, Lancet D, Gilad Y. Different noses for different people. *Nature genetics*. 2003;34(2):143-4.
69. Belzeaux R, Bergon A, Jeanjean V, Loriod B, Formisano-Treziny C, Verrier L, et al. Responder and nonresponder patients exhibit different peripheral transcriptional signatures during major depressive episode. *Translational psychiatry*. 2012;2:e185.
70. Alrahbeni T, Sartor F, Anderson J, Miedzybrodzka Z, McCaig C, Muller B. Full UPF3B function is critical for neuronal differentiation of neural stem cells. *Molecular brain*. 2015;8:33.
71. Meneghini V, Bortolotto V, Francese MT, Dellarole A, Carraro L, Terzieva S, et al. High-mobility group box-1 protein and beta-amyloid oligomers promote neuronal differentiation of adult hippocampal neural progenitors via receptor for advanced glycation end products/nuclear factor-kappaB axis: relevance for Alzheimer's disease. *The Journal of neuroscience*. 2013;33(14):6047-59.
72. Colleoni S, Galli C, Gaspar JA, Meganathan K, Jagtap S, Hescheler J, et al. Development of a neural teratogenicity test based on human embryonic stem cells: response to retinoic acid exposure. *Toxicological sciences*. 2011;124(2):370-7.
73. Gaier ED, Eipper BA, Mains RE. Pam heterozygous mice reveal essential role for Cu in amygdalar behavioral and synaptic function. *Annals of the New York Academy of Sciences*. 2014;1314:15-23.
74. Gaier ED, Rodriguiz RM, Ma XM, Sivaramakrishnan S, Bousquet-Moore D, Wetsel WC, et al. Haploinsufficiency in peptidylglycine alpha-amidating monooxygenase leads to altered synaptic transmission in the amygdala and impaired emotional responses. *The Journal of neuroscience*. 2010;30(41):13656-69.
75. Anacker C, Zunszain PA, Carvalho LA, Pariante CM. The glucocorticoid receptor: pivot of depression and of antidepressant treatment? *Psychoneuroendocrinology*. 2011;36(3):415-25.
76. Grover D, Verma R, Goes FS, Mahon PL, Gershon ES, McMahon FJ, et al. Family-based association of YWHAH in psychotic bipolar disorder. *American journal of medical genetics Part B, Neuropsychiatric genetics*. 2009;150B(7):977-83.
77. Gratten J, Wray NR, Keller MC, Visscher PM. Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nat Neurosci*. 2014;17(6):782-90.
78. Uher R, Tansey KE, Malki K, Perlis RH. Biomarkers predicting treatment outcome in depression: what is clinically significant? *Pharmacogenomics*. 2012;13(2):233-40.
79. Holmes RD, Tiwari AK, Kennedy JL. Mechanisms of the placebo effect in pain and psychiatric disorders. *The pharmacogenomics journal*. 2016;16(6):491-500.

PART 3:

Multi-marker tests as innovative approach to antidepressant pharmacogenomics

Chapter 5

PPP3CC gene: a putative modulator of antidepressant response through the B-cell receptor signaling pathway

Chiara Fabbri, Agnese Marsano, Diego Albani, Armando Chierchia, Raffaella Calati, Antonio Drago, Concetta Crisafulli, Marco Calabrò, Siegfried Kasper, Rupert Lanzenberger, Joseph Zohar, Alzbeta Juven-Wetzler, Daniel Souery, Stuart Montgomery, Julien Mendlewicz, Alessandro Serretti

Published: Fabbri et al., PPP3CC gene: a putative modulator of antidepressant response through the B-cell receptor signaling pathway. *The Pharmacogenomics Journal* 2014; 14: 463-472.

Abstract

Antidepressant pharmacogenetics represent a stimulating but often discouraging field. The present study propose a combination of several methodologies across 3 independent samples.

Genes belonging to monoamine, neuroplasticity, circadian rhythm, and transcription factor pathways were investigated in two samples (n=369 and 88) with diagnosis of major depression who were treated with antidepressants. Phenotypes were response, remission, and treatment-resistant depression. Logistic regression including appropriate covariates was performed. Genes associated with outcomes were investigated in the STAR*D genome-wide study (n=1861). Top genes were further studied through a pathway analysis.

In both original samples markers associated with outcomes were concentrated in the PPP3CC gene. Other interesting findings were particularly in the HTR2A gene in one original sample and the STAR*D. The B cell receptor signaling pathway resulted the putative mediator of PPP3CC effect on antidepressant response (p=0.03).

Among innovative candidates, PPP3CC – involved in the regulation of immune system and synaptic plasticity – seems promising for further investigation.

Key words: gene, antidepressant, pharmacogenetics, PPP3CC, B lymphocyte, pathway

1. Introduction

Major depressive disorder (MDD) is a high-prevalence disorder with heavy impact on individual well-being and health care systems (1). Current pharmacological treatments provide a complete remission of symptoms only in about 30% of subjects (2), partly due to the high variability in efficacy and lack of predictors of response. Antidepressant response has a relevant genetic component, since common SNPs were estimated to explain 0.428 of variance to SSRIs response and 0.420 of variance to mixed antidepressant response (3).

Previous pharmacogenetic studies have been conducted according to the candidate gene approach or the genome-wide approach. Both of them show relevant limitations (4) and have been mainly focused on the analysis of individual polymorphisms, while the functional unit of genome is represented by genes. Thus, a gene-based analysis and the integration of findings obtained through the candidate gene and genome-wide approaches may represent a suitable way to contrast the limitations of both methodologies. Pathway analysis provides the opportunity to deepen the functional role of candidate genes under the view of the antidepressant mechanisms of action.

The most confirmed theories of MDD pathophysiology suggest the involvement of monoamines, neuroplasticity (5) and circadian rhythm alterations (6). The serotonin (5-HT) 2A receptor (HTR2A) gene is supposed to affect both the risk of MDD and antidepressant response (7): notably, 5-HT_{2A} receptor blockade leads to rapid antidepressant effect in rodents (8) and the antidepressant potential of several drugs is hypothesized to involve 5-HT_{2A} antagonism (9). Previous pharmacogenetic studies focused on HTR2A suggest a multilocus model, since several studies reported different SNPs in this gene among the top candidates (10-12). The catechol-O-methyl transferase (COMT) is involved in the metabolism of monoamines, and previous studies have been largely focused on one single variant, i.e. rs4680 (Val108/158Met) (7). Recently the variability of the gene has been more extensively investigated (13-15).

Among neurotrophic factors, brain derived neurotrophic factor (BDNF) is without doubt the most investigated gene. Similarly to COMT, studies have been almost all focused on one single SNP, i.e. rs6265 (196G/A or Val66Met), with inconsistent findings (4). Preliminary findings were provided for other BDNF SNPs (16-18). The inhibitory phosphorylation of glycogen synthase kinase 3 B (GSK3B) occurs in the context of the signaling cascades in response to 5-HT, 5-HT₁ receptor agonists, dopamine, lithium, and antidepressants (19). It is involved in the control of gene expression, cell behavior, cell adhesion and cell polarity, and plays a major role in neurodevelopment, regulation of neuroplasticity, and cell survival (20). Further,

GSK-3 inhibitors have antidepressant effects in animal models of depression (21). The gene has been poorly studied, and preliminary findings pertain rs334558 (22). PLA2G4A (phospholipase A2, group IVA) is a calcium-dependent arachidonic acid-selective cytosolic phospholipase, which is found in post-synaptic sites in the brain (23). The released arachidonic acid and its metabolites can modulate signal transduction, transcriptional regulation, neuronal activity, apoptosis and a number of other neuronal processes (24). Further, PLA2G4A can be linked to various G-protein-coupled receptors including HTR2A/2C (25, 26) and ionotropic N-methyl-D-aspartate (NMDA) receptor (27). Animal studies suggested that the gene is involved in antidepressant mechanisms of action (28), and it has been hypothesized a risk factor for MDD (29), but no previous pharmacogenetic study has been conducted. Protein phosphatase 3, catalytic subunit, gamma isozyme (PPP3CC) gene is a calcium-dependent, calmodulin-stimulated protein phosphatase. PPP3CC may have a role in the calmodulin activation of calcineurin, a neuron-enriched phosphatase that regulates synaptic plasticity, and antagonizes the effects of the cyclic AMP activated protein/kinase A. The kinase/phosphatase dynamic balance seems to be critical for transition to long-term cellular responses in neurons (30). PPP3CC is also involved in B cell receptor signaling, suggesting another possible mechanism of association with antidepressant action since the role of the immune system in the pathogenesis of major depression (31). Further, MDD treatment outcome is influenced by the antidepressant induced modulation of cytokines (32) and of B cells proliferation (33). The gene has not been previously investigated as predictor of antidepressant efficacy. The sialyltransferase 8B (ST8SIA2) gene may be involved in the production of polysialic acid, a modulator of the adhesive properties of neural cell adhesion molecule (NCAM1) that is involved in neuronal plasticity. The gene has previously shown association with bipolar spectrum disorders (34).

Besides genes involved in neuroplasticity, genes involved in the regulation of the circadian rhythm represent interesting candidates. The RAR-related orphan receptor A (RORA) gene is a member of the nuclear hormone-receptor superfamily, and has emerged as an important component of mammalian circadian rhythms (35). A GWAS suggested the gene as a predictor of trait depression (36) and a suggestive genome-wide association with citalopram response has been reported (37). The vasoactive intestinal peptide receptor 2 (VIPR2) gene contribute to circadian rhythms regulation, since KO mice express disrupted behavioral and metabolic rhythms and show altered suprachiasmatic nuclei neuronal activity and clock gene expression (38). The gene has been associated with the risk of bipolar

disorder and MDD (39), but no data exist about its possible involvement in antidepressant effect.

Given that the effectors of antidepressant action are hypothesized to finally act on gene expression modulation, functional variants in transcription factors may be relevant. Zinc finger protein 804A (ZNF804A) alleles are linked to changes in neural activity and connectivity in healthy subjects (40), as well as neuroanatomical changes in both white and grey matter in several brain regions (41). The gene has been associated with the risk of bipolar disorder and especially psychosis (42, 43), and it is considered a cross-diagnostic candidate. Finally, Sp4 transcription factor (SP4) is a zinc-finger brain-specific transcription factor that have been associated with MDD in two GWAS meta-analyses (44, 45). SP4 mutant mice showed decreased granule cell density in the hippocampal dentate gyrus (46) and the gene may play a role in glutamate-induced neurotoxicity (47, 48).

Given the provided picture, the aim of the present study is to investigate the reported 11 candidate genes in two original samples and replicate findings in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) GWAS. Further, a pathway analysis in the STAR*D has been used to deepen the mechanisms by which the top candidates may be involved in antidepressant effect.

2. Materials and Methods

2.1. Samples

2.1.1. European sample

285 MDD patients and 84 bipolar disorder (BD) patients (total sample size=369) were recruited in the context of the European multicenter project ‘Patterns of treatment resistance and switching strategies in unipolar affective disorder’. The study protocol was approved by the ethical committees of all participating centers. Detailed description of the whole sample has been previously reported (49). Depressive symptoms were evaluated according to Hamilton Depression Rating Scale (HDRS, 21-item version) at baseline and after 4 weeks of treatment.

2.1.2. Italian sample

88 patients with diagnosis of non psychotic MDD (DSM-IV criteria) and with a score ≥ 13 on HDRS (21-item version) were included. Exclusion criteria were detailed elsewhere (50). Eligible patients were treated with antidepressants in a naturalistic setting (Supplementary Table 2). Patients were evaluated for depressive symptomatology (21-item HDRS) by trained psychiatrists at baseline and weekly until week 8.

Patients were outpatients recruited at the Department of Biomedical and NeuroMotor Sciences, Bologna University. Ethical approval was obtained from local research ethic committee.

2.1.3. Sequenced Treatment Alternatives to Relieve Depression (STAR*D)

Detailed descriptions of the study design and study population are detailed elsewhere (51). In brief, non psychotic MDD (DSM-IV criteria) patients were enrolled from primary care or psychiatric outpatient clinics and a current 17-item Hamilton Depression Rating score of ≥ 14 by independent raters was obtained. Severity of depression was assessed using the 16-item Quick Inventory of Depressive Symptomatology-Clinician Rated (QIDS-C) (52) at baseline, weeks 2, 4, 6, 9, and 12. All patients received citalopram in level 1. Patients who failed to reach response, were included in level 2, in which different substrategies were defined by participant acceptability (Supplementary Table 2).

2.2. Phenotypes under investigation

The main investigated phenotypes were response and remission at weeks 4 and 8 in the Italian sample and at week 4 in the European sample and in the STAR*D, according to standard criteria (decrease of at least 50% in the HDRS-21 or the QIDS-C, score ≤ 7 of the HDRS-21 or ≤ 5 of the QIDS-C, respectively).

In the European sample treatment-resistance (TRD) was also investigated. Two definitions of treatment resistance have been considered in the analyses: 1) non response to at least 2 adequate consecutive antidepressant treatments administered during the last episode (49) (wide definition, TRD-W); 2) non response to at least 2 adequate consecutive antidepressant treatments of different classes administered during the last episode (53, 54) (different classes definition, TRD-DC).

2.3. Aims of the study

The primary aim of the study was to investigate the possible association between the selected genes and antidepressant response and remission in all the available samples. The secondary aim was to investigate the possible association between the selected genes and TRD in the European sample, and try to replicated these findings in the STAR*D level 2 (which included patients who did not respond to Level 1 treatment).

In the European sample, the primary analysis was focused on the MDD subgroup, while a secondary analysis was performed in the whole sample (MDD+BD).

2.4. Genotyping

Genomic DNA was purified with an automated workstation (Maxwell, Promega) and checked for quality and quantity by a small scale spectrophotometer (Nanodrop, Thermo Scientific). The genotyping was performed using a Sequenom MassArray platform (Sequenom, California, USA) in conjunction with the iPLEX

assay (<http://www.sequenom.com>). Genotyping was then performed according to the manufacturer's standard protocols. MassArrayTyper V.4.0 3.4 was used to read the extended mass and genotype calls. Forward and reverse primers' sequences are available upon request. Genetic SNPs were chosen among those 1) with a reported prevalence of at least 5% for the variant allele among Caucasians (data from <http://hapmap.ncbi.nlm.nih.gov/>, $R^2=0.08$ and $MAF=0.05$), and 2) with availability of a validated assay in our laboratory. We also considered variants not investigated before. The list of genotyped SNPs is shown in Supplementary Table 1.

2.5. Statistical analysis

The effect of individual markers (alleles and genotypes) on phenotypes was tested through logistic regression models in all samples. Covariates were selected according to their impact on outcomes, while gender and age were used as covariates in all the analyses, and ancestry in the STAR*D since ancestry stratification has been previously reported in this sample (55). Briefly, a complete agglomerative clustering was applied, based on a multidimensional scaling of a matrix of pairwise identity-by-state (IBS) values between samples, and clusters were defined on the base of the pairwise population concordance test ($PCC < 0.0001$, according to (56)). Identity-by-descent (IBD) analysis was used to identify related subjects ($IBD > 0.1875$ (57)) in the STAR*D. Odds ratios (OR) with 95% confidence intervals (CI) were estimated for the effects of high-risk genotypes/alleles. In the European and Italian samples Haploview 3.2 was used to generate a linkage disequilibrium map and the R software (cran.r-project.org/) "haplo.stat" package was used to perform haplotype analysis. Permutation (20.000 permutations) was used to estimate the global significance of the results for haplotype analyses to confirm the asymptotic p values when $p < 0.05$.

In the STAR*D available SNPs within the genes showing evidence of association ($p < 0.05$) in the two original samples were extracted from the genome-wide data according to gene physical positions (Genome Build 36.3). Secondly, the genes of interest were imputed using IMPUTE2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) and 1000 Genomes data (NCBI Build 36 (dbSNP b126)) as reference panel. An info value threshold ≥ 0.8 was applied in order to prune poorly imputed SNPs. Finally, top candidate genes emerging from the analyses in the original samples were further investigated through a pathway analysis in the STAR*D. For genes of interest the respective molecular pathways were identified according to the KEGG database (www.genome.jp/kegg/pathway.html). Genes belonging to the pathways under analysis were imputed according to the method reported above. Variations showing

$p < 0.05$ and $p < 0.01$ in the pathway under analysis were tested for a significant different distribution (Fisher exact test) compared to a random pathway. Each random pathway was matched with the index pathway in terms of number of SNPs within it and intragenic position of the SNPs but with random distribution within the genome. $10e-5$ permutations were run.

In single marker analysis, the whole European sample provided a power of 0.80 to detect risk alleles with $OR \geq 1.98$ setting alpha value to the exploratory value of 0.05 two-tailed and considering an explained variance of 0.02. Setting the same parameters, risk alleles with $OR \geq 3.7$ are detectable in the Italian sample and with $OR \geq 1.37$ in the STAR*D. Alpha value was set at the conservative value of 0.05 because the use of replication samples reduce the risk of false positive findings. In the case of haplotype and pathway analyses, $20e-4$ and $10e-5$ permutations were run, respectively, to control for false positive findings.

3. Results

Clinical demographic features and treatment of the analyzed samples are reported in Supplementary Table 2. In the European sample suicidal risk impacted on all clinical outcomes, and anxiety disorder comorbidity on response and remission, (data not shown), thus they were used as covariates. In the Italian sample and STAR*D baseline severity affected outcomes and it was used as covariate. A representation of gene selection process, performed analyses and results in the three samples is reported in Figure 2.

3.1. Results in the European sample

Results with $p < 0.05$ are shown in Table 1, while an overview of all results is shown in Supplementary Table 3. SNPs associated with clinical outcomes were concentrated in the PPP3CC gene: rs10108011 was associated with response ($p=0.005$), TRD-W ($p=0.03$), and TRD-DC ($p=0.04$) in the MDD subsample; rs7430 was associated with response ($p=0.0002$), TRD-W ($p=0.01$), and TRD-DC ($p=0.008$), especially in the MDD subsample; rs2249098 was associated with response ($p=0.002$) especially in the MDD subsample again. Some SNPs in ST8SIA2 were associated with outcomes in MDD subjects: rs3784723 with response (genotypic $p=0.006$, allelic $p=0.007$) and remission (genotypic $p=0.029$, allelic $p=0.03$), rs8035760 with response (genotypic $p=0.04$, allelic $p=0.008$), remission (genotypic $p=0.04$, allelic $p=0.02$). Further, rs11030101 in the BDNF gene showed a weak evidence of association with response (genotypic $p=0.037$; allelic $p=0.032$) in the MDD subsample, while rs11030104 alleles showed a selective effect on the risk of TRDC ($p=0.038$). Other weak associations were

found for ZNF804A rs7603001-rs1344706 and PLA2G4A rs10489407 in the whole sample.

Haplotypic analysis did not retrieve any significant result (data not shown).

Table 1: week 4 (w4) response (A.), treatment-resistant depression wide definition (TRD-W) and treatment-resistant depression different classes definition (TRD-DC) (B.) in the European sample. Only results with $p < 0.05$ are shown. All the results reported here were obtained in the MDD subsample. NS=non significant result. CI=95% confidence interval.

A

Gene	SNP	w4 response: genotypic	w4 response: allelic	w4 remission: genotypic	w4 remission: allelic
PPP3CC	rs10108011	Z=-2.81, p=0.005; OR=0.37, CI=0.18-0.73	NS	NS	NS
	rs7430	Z=-3.69, p=0.0002; OR=0.23, CI=0.11-0.50	NS	NS	NS
	rs2249098	Z=-3.08, p=0.002; OR=0.30, CI=0.14-0.64	NS	NS	NS
BDNF	rs11030101	Z=-2.09, p=0.037; OR=0.44, CI=0.20- 0.94	Z=-2.14, p=0.032; OR=0.65, CI=0.44-0.96	NS	NS
ST8SIA2	rs3784723	Z=-2.74, p=0.006; OR=0.15, CI=0.04-0.57	Z= -2.70, p=0.007; OR=0.17, CI=0.04-0.60	Z= -2.19, p=0.029; OR=0.17, CI=0.03-0.94	Z=-2.17, p=0.03; OR=0.19, CI=0.04-0.99
	rs8035760	Z=-2.02, p=0.04; OR= 0.17, CI=0.02-0.87	Z=-2.66, p=0.008; OR=0.47, CI=0.27-0.82	Z=-2.10, p=0.04; OR=0.15, CI=0.02-0.95	Z=-2.36, p=0.02; OR=0.39, CI=0.18-0.88
PLA2G4A	rs10489407	NS	Z=2.19, p=0.028; OR=1.79, CI=1.08- 3.07	NS	NS

B

Gene	SNP	TRD-W: genotypic	TRD-W: allelic	TRD-DC: genotypic	TRD-DC: allelic
PPP3CC	rs10108011	Z=2.14, p=0.03; OR=2.52, CI=1.09-6.00	NS	Z=2.03, p=0.04; OR=2.60, CI=1.05- 6.68	NS
	rs7430	Z=2.55, p=0.01; OR=3.36,	NS	Z=2.64, p=0.008;	NS

		CI=1.35-8.78		OR=4.02, CI=1.47- 11.70	
BDNF	rs11030104	NS	NS	NS	Z=-2.08, p=0.038; OR=0.48, CI=0.24-0.95

3.2. Results in the Italian sample

Results with $p < 0.05$ are shown in Table 2. Significant findings were mainly in the PPP3CC gene: rs11780915 was associated with response (genotypic $p=0.022$; allelic $p=0.0096$) and remission (genotypic $p=0.04$), rs10108011 with response (genotypic $p=0.002$; allelic $p=0.0003$) and remission (genotypic $p=0.007$; allelic $p=0.004$), rs7430 and rs2249098 with response (genotypic $p=0.025$ and allelic $p=0.01$; genotypic $p=0.022$ and allelic $p=0.0097$, respectively). rs2037547 and rs1381841 in GSK3B were weakly associated with response ($p=0.039$; $p=0.026$, respectively). rs643627 in HTR2A and RORA rs11630262 were associated with remission ($p=0.03$, $p=0.037$, respectively). The rs11780915-rs10108011-rs7430-rs2249098 G-G-G-C haplotype of the PPP3CC gene was associated with non response (permutated $p=0.046$). Haplotypes in other genes did not provide any evidence of association with outcomes (data not shown).

Table 2: week 4 response (w4 resp.) and week 4 remission (w4 rem.) (A.), week 8 response (w8 resp.) and week 8 remission (w8 rem.) (B.) in the Italian sample. Only SNPs with $p < 0.05$ are shown. NS=non significant result. CI=95% confidence interval.

A

Gene	SNP	w4 resp: genotypic	w4 resp: allelic	w4 rem: genotypic
PPP3CC	rs10108011	Z=2.11, $p=0.03$; OR=3.60, CI=1.10-11.79	Z=2.35, $p=0.019$; OR=2.08, CI=1.13-3.84	NS
	rs11780915	NS	NS	Z=-2.05, $p=0.04$; OR=0.22, CI=0.05-0.94
HTR2A	rs643627	NS	NS	Z=2.16, $p=0.03$; OR=3.20, CI=1.11-9.17

B

Gene	SNP	W8 resp: genotypic	W8 resp: Allelic	W8 rem: genotypic	W8 rem: allelic
PPP3CC	rs11780915	Z=2.29, $p=0.02$; OR=5.69,	Z=2.59, $p=0.0096$; OR=2.29,	NS	NS

		CI=1.29-25.18	CI=1.22- 4.29		
	rs10108011	Z=3.05, p=0.002; OR=7.17, CI=2.03-25.38	Z=3.57, OR=3.13, p=0.0003; CI=1.67-5.85	Z=2.67, p=0.007; OR=6.03, CI=1.62- 22.49	Z=2.87, p=0.004; OR=2.55, CI=1.34-4.83
	rs7430	Z=2.24, p=0.03; OR=4.00, CI=1.20-13.46	Z=2.52, p=0.011; OR=2.17, CI=1.19-3.98	NS	NS
	rs2249098	Z=2.29, p=0.02; OR=2.09, CI=0.71-6.22	Z=2.59, p=0.0097; OR=2.25, CI=1.22- 4.15	NS	NS
GSK3B	rs2037547	Z=2.06, p=0.039; OR=3.42, CI=1.06-10.98	NS	NS	NS
	rs1381841	Z=2.23, p=0.026; OR=4.79, CI=1.21-19	NS	NS	NS
RORA	rs11630262	NS	NS	NS	Z=2.08, p=0.037; OR=3.30, CI=1.07- 10.15

3.3. Results in the STAR*D sample

Results of single marker analyses in Level 1 are shown in Table 3 (genotyped SNPs) and Table 4 (imputed SNPs). Among genotyped SNPs (n=194), rs1923888 in HTR2A resulted the top SNP (p response=0.006 and p remission=0.0036). This SNP was only 4226 bp far from rs643627, that was associated with remission in the Italian sample (see above). The HTR2A gene showed evidence of association with remission also in Level 2 (rs1928038 p=0.006, Supplementary Table 3). After imputation, 2,202 SNPs were available. In Level 1 rs7333412 in HTR2A resulted the top SNP (p=3.91e-04) and a cluster of SNPs showed p values of size e-04 in the downstream/first intron region of the gene (Figure 1). On the other hand, in Level 2 only two SNPs within ST8SIA2 showed a p value < 0.005 (Table 4).

Despite no individual SNP in the PPP3CC gene showed interesting results, the gene was further analyzed through a pathway analysis, since several SNPs within it were associated with outcomes in both the original samples. Among the KEGG pathways including PPP3CC (Supplementary Table 4), the B cell receptor signaling pathway was associated with remission (permuted p=0.03) and showed a trend of association with response (permuted p=0.059), as reported in Table 5.

Interactions among components of the pathway are shown in Supplementary Figure 1.

Table 3: week 4 response (A.) and week 4 remission (B.) in STAR*D level 1, SNPs with $p < 0.05$ are shown. 194 SNPs in total were available. The number of SNPs available in each gene is reported in brackets. In PPP3CC 9 SNPs, and in RORA 9 SNPs were retrieved, respectively. Additive, dominant, recessive and deviance from dominance models were tested. Distance from the top SNPs in the original samples is reported (column “Chr position”) in base pairs (bp). HWE=Hardy-Weinberg equilibrium. CI=95% confidence interval.

A

Gene	Polymorphism	Model	Statistics	Chr position	P HWE
HTR2A (24 SNPs)	rs977003	Additive	OR=0.84, CI=0.73- 0.97, p=0.016	47415001, 13610 bp from rs643627	0.02
	rs1923888	Dominant	OR=0.76, CI=0.62- 0.93, p=0.006	47424385, 4226 bp from rs643627	0.61
	rs1745837	Dominant	OR=0.76, CI=0.62-0.93, p=0.007	47424812, 3800 bp from rs643627	0.50
	rs2296972	Dominant	OR=0.76, CI=0.62- 0.94, p=0.009	47428471, 140 bp from rs643627	0.29
GSK3B (10 SNPs)	rs4340737	Additive	OR=1.35, CI=1.02-1.77, p=0.035	119622014, 56788 bp from rs1381841	0.68
	rs4624596	Additive	OR=1.32, CI=1.01- 1.73, p=0.042	119571541, 107261 bp from rs1381841	0.33
ZNF804A (15 SNPs)	rs7590852	Dominant	OR=1.55, CI=1.02-2.34, p=0.040	185766912, 96 bp from rs7603001	0.22
	rs725617	Dominant	OR=1.54, CI=1.02-2.33, p=0.042	185778262, 11350 bp from rs7603001	0.05
BDNF (6 SNPs)	rs7127507	Domdev	OR= 1.96, CI=1.26-3.04, p=0.0028	27714884, 34140 bp from rs11030101	0.16
	rs11030119	Domdev	OR= 1.62, CI=1.009-2.56, p= 0.046	27728102, 47358 bp from rs11030101	0.82
PLA2G4 A (57 SNPs)	rs12720662	Additive	OR=1.27, CI=1.05-1.54, p=0.015	186934092, 109979 bp from rs10489407	0.008

B

Gene	Polymorphism	Model	Statistics	Chr position	P HWE
HTR2A (24 SNPs)	rs977003	Recessive	OR=0.67, CI=0.50-0.90, p=0.0068	47415001, 13610 bp from rs643627	0.02
	rs985934	Domdev	OR=1.42, CI= 1.07-1.88, p=0.014	47455725, 27114 bp from rs643627	0.12
	rs985933	Domdev	OR=1.39, CI=1.05-1.84, p=0.02	47455863, 27252 bp from rs643627	0.09
	rs1923888	Dominant	OR=0.67, CI=0.51-0.88, p=0.0036	47424385, 4226 bp from rs643627	0.61
	rs9316232	Additive	OR=0.80, CI=0.65-0.98, p=0.029	47426722, 1889 bp from rs643627	0.92
	rs9567739	Dominant	OR=0.71, CI=0.53- 0.94, p=0.018	47424944, 3667 bp from rs643627	0.44
	rs1745837	Dominant	OR=0.70, CI=0.53-0.91, p=0.009	47424812, 3800 bp from rs643627	0.50
	rs2296972	Dominant	OR=0.69, CI=0.53-0.91, p=0.008	47428471, 140 bp from rs643627	0.29
ZNF804A (15 SNPs)	rs1344706	Domdev	OR= 0.56, CI= 0.35- 0.88, p=0.013	185778428, 11612 bp from rs7603001	0.15
PLA2G4 A (57 SNPs)	rs10752991	Additive	OR=1.28, CI=1.01- 1.61, p=0.037	186972935, 38843 bp from rs10489407	0.62
ST8SIA2 (64 SNPs)	rs3784732	Dominant	OR=1.38, CI=1.02-1.87, p=0.036	92986916, 18843 bp from rs8035760	0.93

Table 4: A. Results after imputation of the genes of interest in the STAR*D level 1. Phenotype: remission at week 4. In level 1 no SNP reached the threshold when considering response as outcome. B. Results after imputation of the genes of interest in the STAR*D level 2. Phenotype: response at week 4. In level 2 no SNP reached the threshold when considering remission as outcome. Only SNPs with IMPUTE2 info ≥ 0.8 were considered (see number of available SNPs). Only results with $p < 0.005$ are shown. For GSK3B (217 SNPs), VIPR2 (45 SNPs), PPP3CC (80 SNPs), and PLA2G4A (87) no SNP reached this p threshold. CI=95% confidence interval.

A

Gene (n° of available SNPs)	SNP	Info	Test	MAF remitters	MAF non remitters	Statistics
BDNF (99)	rs7124442	0.9475	Recessive	0.3737	0.3166	OR=0.78, CI=0.64-0.94, p= 0.0048
	rs10767658	0.9717		0.3736	0.3165	OR=0.78, CI=0.64-0.94, p=0.0048
	rs1519480	0.9729		0.3738	0.3167	OR= 0.78, CI=0.64-0.94, p= 0.0048
	rs10767657	0.9712		0.3736	0.3165	OR= 0.78, CI=0.64-0.94, p= 0.0048
	rs925946	0.9710		0.3736	0.3165	OR= 0.78, CI=0.64-0.94, p= 0.0048
ZNF80A4 (497)	rs17508713	0.8216	Additive	0.2230	0.1755	OR=1.35, CI=1.08-1.69, p=0.0027
	rs10202700	0.9211		0.2076	0.1611	OR= 1.36, CI=1.08-1.72, p=0.0043
	rs6715977	0.9212		0.2076	0.1611	OR=1.36, CI=1.08-1.72, p=0.0043
	rs12693382	0.8465		0.2196	0.1745	OR=1.33, CI=1.06-1.67, p=0.0049
HTR2A (145)	rs7333412	0.8349	Additive	0.3273	0.2674	OR=1.33, CI=1.09-1.62, p=3.91e-04
	rs7324017	0.8591		0.3278	0.2675	OR=1.34, CI=1.10-1.63, p=4.31e-04
	rs1923882	0.9737		0.3083	0.2480	OR=1.35, CI=1.11-1.65, p=4.92e-04
	rs3125	0.8880		0.2552	0.2085	OR=1.30, CI=1.05-1.61, p=0.0024
	rs3803189	0.8858		0.2551	0.2085	OR=1.30, CI=1.05-1.61, p=0.0024
	rs55948462	0.9067		0.2556	0.2089	OR=1.30, CI=1.05-1.61, p=0.0027
	rs61948314	0.8087		0.2357	0.1945	OR=1.28, CI=1.03-1.59, p=0.0041

RORA (872)	SNP	Info	Test	MAF in remitters	MAF non remitters	Statistics
	rs62002750	0.9515	Recessive	0.2772	0.3431	OR=0.73, CI=0.60-0.90, p=0.0023
	rs17270446	0.9779	Recessive	0.2768	0.3427	OR=0.73, CI=0.60-0.90, p=0.0024
	rs7162615	0.9432	Recessive	0.2443	0.3092	OR=0.72, CI=0.58-0.89, p=0.0034
	rs2445927	0.9770	recessive	0.2363	0.3007	OR=0.72, CI=0.58-0.89, p= 0.0038
	rs12913922	0.9566	Additive	0.2367	0.3010	OR=1.39, CI=1.12-1.72, p=0.0047
	rs62004363	0.9568	recessive	0.2370	0.3013	OR=0.72, CI=0.58-0.89, p= 0.0039
	rs236220	0.8785	Additive	0.2773	0.3405	OR=0.74, CI=0.61-0.91, p=0.0048
	rs7171681	0.9571	Additive	0.2374	0.3017	OR=1.39, CI=1.12-1.72, p=0.0049
	rs341383	0.9852	Recessive	0.2430	0.3064	OR= 0.73, CI=0.59-0.90, p= 0.0048
	rs341374	0.9696	Recessive	0.2456	0.3090	OR= 0.73, CI=0.59-0.90, p= 0.0049
	rs341373	0.9687	Recessive	0.2457	0.3091	OR=0.73, CI=0.59-0.90, p= 0.0049

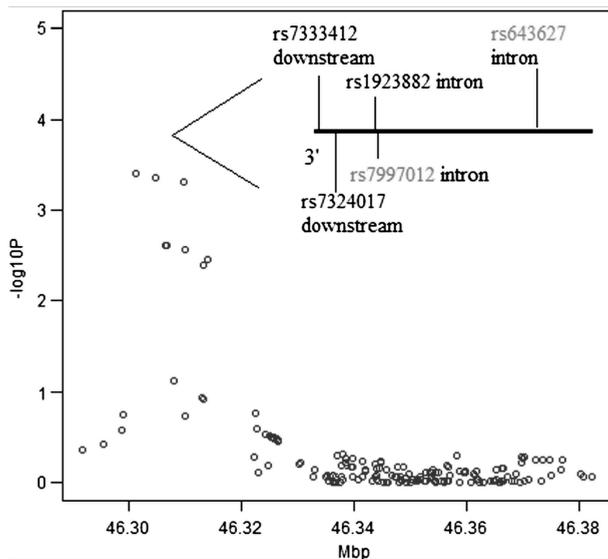
B

Gene (n° of available SNPs)	SNP	Info	Test	MAF in remitters	MAF non remitters	Statistics
ST8SIA2 (72)	rs11074067	0.99	Dominant	0.5423	0.4800	OR=0.49, CI=0.32-0.78, p=0.0023
	rs12592946	0.94		0.5393	0.4816	OR=0.52, CI=0.34-0.82, p=0.0027

Table 5: results of the PPP3CC pathway analysis in the STAR*D Level 1.

Phenotype	Pathway	n SNPs with p<0.01 (%)	n SNPs with p>0.01 (%)	Top genes: n SNPs, % SNPs with p<0.01	Perm. p
Response (week 4)	B cell receptor signaling pathway	36 (0.02)	1295 (0.97)	PIK3API: 88, 19% CD79B: 8, 12% KRAS: 50, 6% PIK3CA: 51, 3% BLNK: 82, 2%	0.059
	Random pathway	25 (0.01)	1306 (0.98)	-	
Remission (week 4)	B cell receptor signaling pathway	32 (0.02)	1301 (0.97)	PIK3API: 88, 19% PLCG2: 199, 4% CD22: 32, 3% PPP3CA: 165, 2%	0.03
	Random pathway	20 (0.01)	1313 (0.98)	-	

Figure 1: Effect of HTR2A imputed SNPs on week 4 remission in STAR*D. rs643627 was associated with remission in the Italian sample while rs7997012 was associated with treatment outcome in previous studies.



4. Discussion

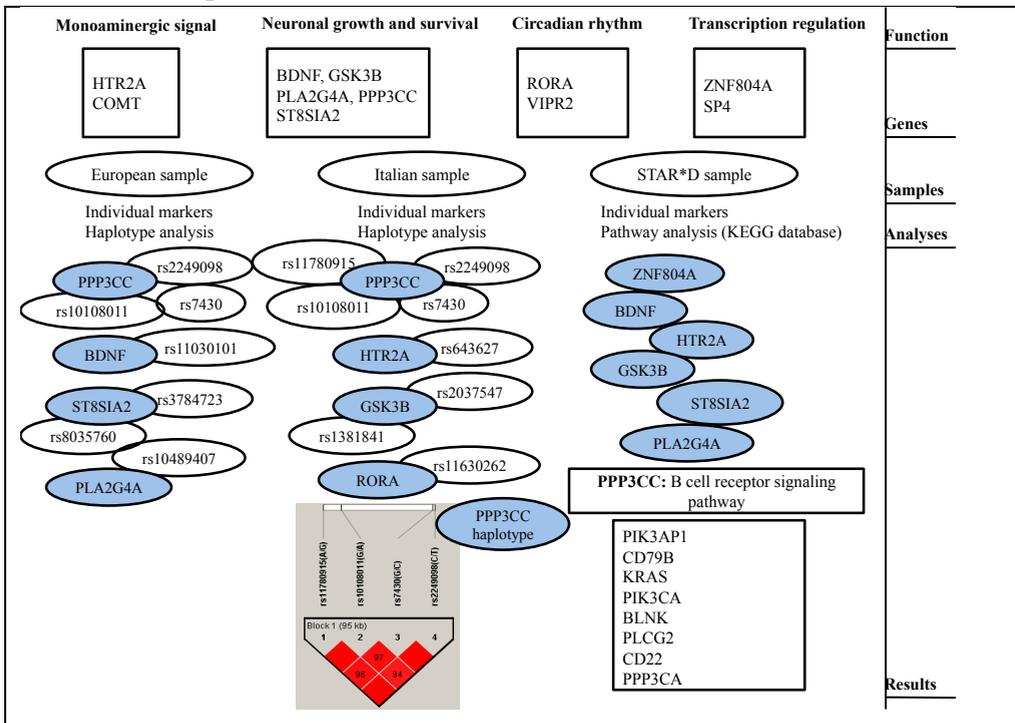
The present study investigated the effect of 11 genes belonging to monoamine, neuroplasticity, circadian rhythm, and transcription factor pathways on antidepressant efficacy in three independent samples. Our findings suggest an association between PPP3CC (gene and B cell receptor signaling pathway) and

antidepressant efficacy in all the analyzed samples, and they represent the first demonstration of PPP3CC involvement in antidepressant treatment outcome. In detail, PPP3CC rs10108011, rs7430, rs2249098, and rs11780915 were associated with clinical outcomes in the original samples. PPP3CC encodes the calcineurin- γ -catalytic subunit, and calcineurin is a calcium-dependent serine/threonine protein phosphatase, widely expressed in the CNS. Although we did not find any association between PPP3CC individual SNPs and antidepressant outcomes in the STAR*D, pathway analysis suggested that the B cell receptor signaling pathway may mediate the effect of PPP3CC on remission (permutated $p=0.03$). This finding is consistent with recent studies that demonstrated the role of immune system in the pathogenesis of major depression (31). Further, MDD treatment outcome is influenced by the antidepressant induced modulation of cytokines (32). Increased percentage of B cells was demonstrated in peripheral blood of MDD subjects, particularly in subjects suffering of melancholic forms (58), and long term treatment with SSRIs may modulate B cells proliferation (59). Furthermore, alteration in B-cells proliferation induced by chronic stress exposure was restored after treatment with fluoxetine (33). Previous pharmacogenetic studies support the involvement of immune system in antidepressant efficacy. In detail, the top candidate genes are interleukin 1 beta (IL1B) and interleukin 11 (IL11) (4). Among their functions, IL11 is shown to stimulate the T-cell-dependent development of immunoglobulin-producing B cells and IL1B stimulates B-cell maturation and proliferation.

In addition to its role in immunity, PPP3CC is involved in the regulation of dopaminergic signal transduction and in the induction of some forms of NMDA-dependent synaptic plasticity (60). Recent evidence supports PPP3CC as a susceptibility gene for affective and cognitive disorders. Indeed, altered calcineurin signaling may contribute to bipolar disorder (61) and schizophrenia (62) pathogenesis and the inhibition of calcineurin can induce depressive-like behaviors (63). Nevertheless, no previous study focused on role of calcineurin in antidepressant response has been conducted in humans. To the best of our knowledge, only a previous study conducted on mice has demonstrated that chronic antidepressant administration lead to an increase of calcineurin levels in the hippocampus (64).

Figure 2: summary of the gene selection process, analyses and results in the three samples. Only genes and SNPs with $p<0.05$ are shown. In the STAR*D sample available SNPs did not overlap with genotyped SNPs in the European and Italian samples; thus only genes are shown and STAR*D results are referred to genotyped

SNPs and not imputed ones. In STAR*D, genes belonging to the B-cell receptor signaling pathway and including the highest proportion of SNPs associated with outcomes are reported.



While PPP3CC is an innovative candidate gene in antidepressant response, our study confirmed the involvement of the known candidate HTR2A. Despite previous results were not completely consistent (65), the gene was among the strongest candidates both in the STAR*D (10, 11) and GENDEP (12) studies. According to our data, in the Italian sample HTR2A rs643627 is associated with remission, and in the STAR*D GWAS rs17069005 (located only 4493 bp from rs643627) is a modulator of the same phenotype. The imputation of STAR*D data allowed to identify a region of association in the downstream of the gene, including rs7333412 (Figure 1). The intronic rs1923888 was near (12400 bp) to the previous top finding of the STAR*D (rs7997012, according to two studies based on independent genotyping (10, 11)). The HTR2A rs643627 variation has been reported to be protective against suicidal behavior (66), but the present study was the first one to investigated this polymorphism as a modulator of antidepressant efficacy. Our results and those previously reported in literature suggest that HTR2A may modulate antidepressant response through a multilocus model involving different regions of the gene (downstream and intronic regions).

In addition to HTR2A, the well supported candidate BDNF gene received a confirmation of contribution to antidepressant efficacy. The present study did not support a role of the extensively studied rs6265, but suggested the involvement of rs11030101 and rs11030104. A recent study showed an association between rs11030101 and response to electroconvulsive therapy (67), supporting the contribution of the SNP to antidepressant efficacy.

Other associations with clinical outcomes were found in GSK3B, RORA, ZNF804A, ST8SIA2, and PLA2G4A genes. Few preliminary but promising data are available about the role of GSK3B in antidepressant efficacy (4), thus it represents a primary candidate for further investigation. The RORA gene is implicated in circadian rhythm control and is expressed in CNS regions involved in the pathogenesis of MDD, particularly cerebral cortex, thalamus and hypothalamus and it shows neuroprotective functions towards oxidative stress (68). Furthermore, this gene has been previously associated with response and remission by a candidate gene study in the STAR*D sample (37). ZNF804A encodes a zinc finger protein and other than being a modulator of schizophrenia and bipolar disorder risk (69), it seems to be involved in fronto-frontal connectivity during emotion perception (70), suggesting a possible role of ZNF804A in antidepressants response. ST8SIA2 is responsible for the post-translational glycosylation of the neuronal cell adhesion molecule (NCAM), which in its polysialylated form plays an important role in processes such as synaptogenesis and neuroplasticity (71). A weak evidence of association was found for PLA2G4A, that catalyzes the hydrolysis of membrane phospholipids to release arachidonic acid, a lipid that acts as cellular hormone. Chronic fluoxetine treatment was demonstrated to increase brain PLA2G4A gene expression post-transcriptionally by increasing mRNA stabilization (28).

Our study has some limitations that have to be stated. Firstly the heterogeneity among the samples under analysis in terms of treatment and scales of evaluation. On the other hand, the replication of results across different samples supports their potential applicability in different clinical context. Further, setting alpha value to 0.05 for individual SNPs analyses entails a relatively higher risk of false positive findings, but this risk is controlled through the use of multiple samples and selection of genes with good pre-test probability of association with outcomes of interest. Indeed, all the selected genes have a biological rationale for being involved in antidepressant mechanisms of action and some of them (e.g. HTR2A and BDNF) show also a number of previous encouraging pharmacogenetic results. Regarding HTR2A, the clustering of associated SNPs (both in the present and previous studies) in some regions of the gene (Figure 1) further reduce the risk of

type I error. The increase in type I error risk due to the use of multiple phenotypes is partially balanced by the correlation existing among them (e.g. response is related to remission). The use of multiple genetic models also increase the risk of false positives, but the inheritance pattern of the causal alleles is unknown for the analyzed SNPs. Other limitations consist in the small size of the Italian sample and in the retrospective design of the European study. Finally, only limited availability of the variants genotyped in the two original samples was found in the STAR*D GWAS and this could be only partially overcome through imputation. Imputation itself provide reliable results but should be verified through genotyping, thus results on imputed SNPs should be considered only preliminary. Despite the limited correspondence of SNPs genotyped in the original samples and STAR*D, the focus of this study is on genes and not individual SNPs, since multilocus models are supposed to modulate antidepressant effects. Several genetic signals were replicated across two of the three samples (Figure 2), despite partial overlap was observed probably due to the limitations reported above (particularly limited statistical power and clinical/genetic heterogeneity among samples).

In conclusion, the present study is the first to suggest PPP3CC gene as a new promising modulator of antidepressant response, possibly through the B cell receptor signaling pathway. The contribution of the known candidates HTR2A and BDNF found support in the present study and other new candidate genes (GSK3B, RORA, ST8SIA2, and ZNF804A) should receive attention.

Supplementary materials:

Supplementary materials are available for the online version of this publication.

References

1. Judd LL, Akiskal HS, Zeller PJ, Paulus M, Leon AC, Maser JD, et al. Psychosocial disability during the long-term course of unipolar major depressive disorder. *Archives of general psychiatry* 2000;**57**:375-80.
2. Kemp AH, Gordon E, Rush AJ, Williams LM. Improving the prediction of treatment response in depression: integration of clinical, cognitive, psychophysiological, neuroimaging, and genetic measures. *CNS spectrums* 2008;**13**:1066-86; quiz 87-8.
3. Tansey KE, Guipponi M, Hu X, Domenici E, Lewis G, Malafosse A, et al. Contribution of common genetic variants to antidepressant response. *Biological psychiatry* 2013;**73**:679-82.
4. Fabbri C, Di Girolamo G, Serretti A. Pharmacogenetics of antidepressant drugs: an update after almost 20 years of research. *Am J Med Genet B Neuropsychiatr Genet* 2013;**162B**:487-520.
5. Racagni G, Popoli M. Cellular and molecular mechanisms in the long-term action of antidepressants. *Dialogues in clinical neuroscience* 2008;**10**:385-400.

6. Courtet P, Olie E. Circadian dimension and severity of depression. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* 2012;**22 Suppl 3**:S476-81.
7. Porcelli S, Drago A, Fabbri C, Gibiino S, Calati R, Serretti A. Pharmacogenetics of antidepressant response. *J Psychiatry Neurosci* 2011;**36**:87-113.
8. Pandey DK, Mahesh R, Kumar AA, Rao VS, Arjun M, Rajkumar R. A novel 5-HT(2A) receptor antagonist exhibits antidepressant-like effects in a battery of rodent behavioural assays: approaching early-onset antidepressants. *Pharmacology, biochemistry, and behavior* 2010;**94**:363-73.
9. Fountoulakis KN, Kelsoe JR, Akiskal H. Receptor targets for antidepressant therapy in bipolar disorder: an overview. *Journal of affective disorders* 2012;**138**:222-38.
10. McMahon FJ, Buervenich S, Charney D, Lipsky R, Rush AJ, Wilson AF, et al. Variation in the gene encoding the serotonin 2A receptor is associated with outcome of antidepressant treatment. *American journal of human genetics* 2006;**78**:804-14.
11. Peters EJ, Slager SL, Jenkins GD, Reinalda MS, Garriock HA, Shyn SI, et al. Resequencing of serotonin-related genes and association of tagging SNPs to citalopram response. *Pharmacogenet Genomics* 2009;**19**:1-10.
12. Uher R, Huezo-Diaz P, Perroud N, Smith R, Rietschel M, Mors O, et al. Genetic predictors of response to antidepressants in the GENDEP project. *Pharmacogenomics J* 2009;**9**:225-33.
13. Kocabas NA, Faghel C, Barreto M, Kasper S, Linotte S, Mendlewicz J, et al. The impact of catechol-O-methyltransferase SNPs and haplotypes on treatment response phenotypes in major depressive disorder: a case-control association study. *International clinical psychopharmacology* 2010;**25**:218-27.
14. Perlis RH, Fijal B, Adams DH, Sutton VK, Trivedi MH, Houston JP. Variation in catechol-O-methyltransferase is associated with duloxetine response in a clinical trial for major depressive disorder. *Biological psychiatry* 2009;**65**:785-91.
15. Ji Y, Biernacka J, Snyder K, Drews M, Pelleymounter LL, Colby C, et al. Catechol O-methyltransferase pharmacogenomics and selective serotonin reuptake inhibitor response. *Pharmacogenomics J* 2012;**12**:78-85.
16. Licinio J, Dong C, Wong ML. Novel sequence variations in the brain-derived neurotrophic factor gene and association with major depression and antidepressant treatment response. *Archives of general psychiatry* 2009;**66**:488-97.
17. Domschke K, Lawford B, Laje G, Berger K, Young R, Morris P, et al. Brain-derived neurotrophic factor (BDNF) gene: no major impact on antidepressant treatment response. *The international journal of neuropsychopharmacology / official scientific journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)* 2010;**13**:93-101.
18. Kocabas NA, Antonijevic I, Faghel C, Forray C, Kasper S, Lecrubier Y, et al. Brain-derived neurotrophic factor gene polymorphisms: influence on treatment response phenotypes of major depressive disorder. *International clinical psychopharmacology* 2011;**26**:1-10.
19. Beaulieu JM, Gainetdinov RR, Caron MG. Akt/GSK3 signaling in the action of psychotropic drugs. *Annual review of pharmacology and toxicology* 2009;**49**:327-47.
20. Grimes CA, Jope RS. The multifaceted roles of glycogen synthase kinase 3beta in cellular signaling. *Progress in neurobiology* 2001;**65**:391-426.
21. Maes M, Fisar Z, Medina M, Scapagnini G, Nowak G, Berk M. New drug targets in depression: inflammatory, cell-mediated immune, oxidative and nitrosative stress,

- mitochondrial, antioxidant, and neuroprogressive pathways. And new drug candidates-- Nrf2 activators and GSK-3 inhibitors. *Inflammopharmacology* 2012;**20**:127-50.
22. Tsai SJ, Liou YJ, Hong CJ, Yu YW, Chen TJ. Glycogen synthase kinase-3beta gene is associated with antidepressant treatment response in Chinese major depressive disorder. *Pharmacogenomics J* 2008;**8**:384-90.
 23. Murakami M, Kambe T, Shimbara S, Kudo I. Functional coupling between various phospholipase A2s and cyclooxygenases in immediate and delayed prostanoid biosynthetic pathways. *The Journal of biological chemistry* 1999;**274**:3103-15.
 24. O'Banion MK. Cyclooxygenase-2: molecular biology, pharmacology, and neurobiology. *Critical reviews in neurobiology* 1999;**13**:45-82.
 25. Garcia MC, Kim HY. Mobilization of arachidonate and docosahexaenoate by stimulation of the 5-HT2A receptor in rat C6 glioma cells. *Brain research* 1997;**768**:43-8.
 26. Qu Y, Villacreses N, Murphy DL, Rapoport SI. 5-HT2A/2C receptor signaling via phospholipase A2 and arachidonic acid is attenuated in mice lacking the serotonin reuptake transporter. *Psychopharmacology* 2005;**180**:12-20.
 27. Basselin M, Chang L, Bell JM, Rapoport SI. Chronic lithium chloride administration attenuates brain NMDA receptor-initiated signaling via arachidonic acid in unanesthetized rats. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* 2006;**31**:1659-74.
 28. Rao JS, Ertley RN, Lee HJ, Rapoport SI, Bazinet RP. Chronic fluoxetine upregulates activity, protein and mRNA levels of cytosolic phospholipase A2 in rat frontal cortex. *Pharmacogenomics J* 2006;**6**:413-20.
 29. Pae CU, Yu HS, Kim JJ, Lee CU, Lee SJ, Lee KU, et al. BanI polymorphism of the cytosolic phospholipase A2 gene and mood disorders in the Korean population. *Neuropsychobiology* 2004;**49**:185-8.
 30. Xia Z, Storm DR. The role of calmodulin as a signal integrator for synaptic plasticity. *Nature reviews Neuroscience* 2005;**6**:267-76.
 31. Krishnadas R, Cavanagh J. Depression: an inflammatory illness? *Journal of neurology, neurosurgery, and psychiatry* 2012;**83**:495-502.
 32. Janssen DG, Caniato RN, Verster JC, Baune BT. A psychoneuroimmunological review on cytokines involved in antidepressant treatment response. *Human psychopharmacology* 2010;**25**:201-15.
 33. Edgar VA, Cremaschi GA, Sterin-Borda L, Genaro AM. Altered expression of autonomic neurotransmitter receptors and proliferative responses in lymphocytes from a chronic mild stress model of depression: effects of fluoxetine. *Brain, behavior, and immunity* 2002;**16**:333-50.
 34. McAuley EZ, Scimone A, Tiwari Y, Agahi G, Mowry BJ, Holliday EG, et al. Identification of sialyltransferase 8B as a generalized susceptibility gene for psychotic and mood disorders on chromosome 15q25-26. *PloS one* 2012;**7**:e38172.
 35. Ko CH, Takahashi JS. Molecular components of the mammalian circadian clock. *Human molecular genetics* 2006;**15 Spec No 2**:R271-7.
 36. Terracciano A, Tanaka T, Sutin AR, Sanna S, Deiana B, Lai S, et al. Genome-wide association scan of trait depression. *Biological psychiatry* 2010;**68**:811-7.
 37. Garriock HA, Kraft JB, Shyn SI, Peters EJ, Yokoyama JS, Jenkins GD, et al. A genomewide association study of citalopram response in major depressive disorder. *Biological psychiatry* 2010;**67**:133-8.

38. Hughes AT, Guilding C, Piggins HD. Neuropeptide signaling differentially affects phase maintenance and rhythm generation in SCN and extra-SCN circadian oscillators. *PLoS one* 2011;**6**:e18926.
39. Soria V, Martinez-Amoros E, Escaramis G, Valero J, Perez-Egea R, Garcia C, et al. Differential association of circadian genes with mood disorders: CRY1 and NPAS2 are associated with unipolar major depression and CLOCK and VIP with bipolar disorder. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* 2010;**35**:1279-89.
40. Esslinger C, Kirsch P, Haddad L, Mier D, Sauer C, Erk S, et al. Cognitive state and connectivity effects of the genome-wide significant psychosis variant in ZNF804A. *NeuroImage* 2011;**54**:2514-23.
41. Lencz T, Szeszko PR, DeRosse P, Burdick KE, Bromet EJ, Bilder RM, et al. A schizophrenia risk gene, ZNF804A, influences neuroanatomical and neurocognitive phenotypes. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* 2010;**35**:2284-91.
42. Lett TA, Zai CC, Tiwari AK, Shaikh SA, Likhodi O, Kennedy JL, et al. ANK3, CACNA1C and ZNF804A gene variants in bipolar disorders and psychosis subphenotype. *The world journal of biological psychiatry : the official journal of the World Federation of Societies of Biological Psychiatry* 2011;**12**:392-7.
43. Williams HJ, Craddock N, Russo G, Hamshere ML, Moskvina V, Dwyer S, et al. Most genome-wide significant susceptibility loci for schizophrenia and bipolar disorder reported to date cross-traditional diagnostic boundaries. *Human molecular genetics* 2011;**20**:387-91.
44. Shi J, Potash JB, Knowles JA, Weissman MM, Coryell W, Scheftner WA, et al. Genome-wide association study of recurrent early-onset major depressive disorder. *Molecular psychiatry* 2011;**16**:193-201.
45. Shyn SI, Shi J, Kraft JB, Potash JB, Knowles JA, Weissman MM, et al. Novel loci for major depression identified by genome-wide association study of Sequenced Treatment Alternatives to Relieve Depression and meta-analysis of three studies. *Molecular psychiatry* 2011;**16**:202-15.
46. Zhou X, Qyang Y, Kelsoe JR, Masliah E, Geyer MA. Impaired postnatal development of hippocampal dentate gyrus in Sp4 null mutant mice. *Genes, brain, and behavior* 2007;**6**:269-76.
47. Mao X, Moerman-Herzog AM, Wang W, Barger SW. Differential transcriptional control of the superoxide dismutase-2 kappaB element in neurons and astrocytes. *The Journal of biological chemistry* 2006;**281**:35863-72.
48. Mao X, Yang SH, Simpkins JW, Barger SW. Glutamate receptor activation evokes calpain-mediated degradation of Sp3 and Sp4, the prominent Sp-family transcription factors in neurons. *Journal of neurochemistry* 2007;**100**:1300-14.
49. Souery D, Oswald P, Massat I, Bailer U, Bollen J, Demyttenaere K, et al. Clinical factors associated with treatment resistance in major depressive disorder: results from a European multicenter study. *J Clin Psychiatry* 2007;**68**:1062-70.
50. Fabbri C, Marsano A, Balestri M, De Ronchi D, Serretti A. Clinical features and drug induced side effects in early versus late antidepressant responders. *Journal of psychiatric research* 2013;**47**:1309-18.
51. Howland RH. Sequenced Treatment Alternatives to Relieve Depression (STAR*D). Part 1: study design. *Journal of psychosocial nursing and mental health services* 2008;**46**:21-4.

52. Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychological medicine* 2004;**34**:73-82.
53. CHMP. Note for guidance on clinical investigation of medicinal products in the treatment of depression 2002. 2002; Available from: <http://www.emea.europa.eu/pdfs/human/ewp/051897en.pdf>.
54. Thase ME. The need for clinically relevant research on treatment-resistant depression. *J Clin Psychiatry* 2001;**62**:221-4.
55. Fabbri C, Drago A, Serretti A. Early antidepressant efficacy modulation by glutamatergic gene variants in the STAR(D). *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* 2012; **23**:612-21
56. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* 2007;**81**:559-75.
57. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. *Nature protocols* 2010;**5**:1564-73.
58. Maes M, Stevens WJ, DeClerck LS, Bridts CH, Peeters D, Schotte C, et al. A significantly increased number and percentage of B cells in depressed subjects: results of flow cytometric measurements. *Journal of affective disorders* 1992;**24**:127-34.
59. Hernandez ME, Martinez-Fong D, Perez-Tapia M, Estrada-Garcia I, Estrada-Parra S, Pavon L. Evaluation of the effect of selective serotonin-reuptake inhibitors on lymphocyte subsets in patients with a major depressive disorder. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* 2010;**20**:88-95.
60. Gerber DJ, Hall D, Miyakawa T, Demars S, Gogos JA, Karayiorgou M, et al. Evidence for association of schizophrenia with genetic variation in the 8p21.3 gene, PPP3CC, encoding the calcineurin gamma subunit. *Proceedings of the National Academy of Sciences of the United States of America* 2003;**100**:8993-8.
61. Mathieu F, Miot S, Etain B, El Khoury MA, Chevalier F, Bellivier F, et al. Association between the PPP3CC gene, coding for the calcineurin gamma catalytic subunit, and bipolar disorder. *Behav Brain Funct* 2008;**4**:2.
62. Horiuchi Y, Ishiguro H, Koga M, Inada T, Iwata N, Ozaki N, et al. Support for association of the PPP3CC gene with schizophrenia. *Molecular psychiatry* 2007;**12**:891-3.
63. Yu JJ, Zhang Y, Wang Y, Wen ZY, Liu XH, Qin J, et al. Inhibition of calcineurin in the prefrontal cortex induced depressive-like behavior through mTOR signaling pathway. *Psychopharmacology* 2013;**225**:361-72.
64. Crozatier C, Farley S, Mansuy IM, Dumas S, Giros B, Tzavara ET. Calcineurin (protein phosphatase 2B) is involved in the mechanisms of action of antidepressants. *Neuroscience* 2007;**144**:1470-6.
65. Lucae S, Ising M, Horstmann S, Baune BT, Arolt V, Muller-Myhsok B, et al. HTR2A gene variation is involved in antidepressant treatment response. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* 2010;**20**:65-8.

66. Giegling I, Hartmann AM, Moller HJ, Rujescu D. Anger- and aggression-related traits are associated with polymorphisms in the 5-HT-2A gene. *Journal of affective disorders* 2006;**96**:75-81.
67. Viikki ML, Jarventausta K, Leinonen E, Huuhka M, Mononen N, Lehtimäki T, et al. BDNF polymorphism rs11030101 is associated with the efficacy of electroconvulsive therapy in treatment-resistant depression. *Psychiatric genetics* 2013;**23**:134-6.
68. Miller MW, Wolf EJ, Logue MW, Baldwin CT. The retinoid-related orphan receptor alpha (RORA) gene and fear-related psychopathology. *Journal of affective disorders* 2013;**151**:702-8.
69. Lee KW, Woon PS, Teo YY, Sim K. Genome wide association studies (GWAS) and copy number variation (CNV) studies of the major psychoses: what have we learnt? *Neuroscience and biobehavioral reviews* 2012;**36**:556-71.
70. Fornito A, Bullmore ET. Connectomic intermediate phenotypes for psychiatric disorders. *Frontiers in psychiatry* 2012;**3**:32.
71. Hildebrandt H, Muhlenhoff M, Weinhold B, Gerardy-Schahn R. Dissecting polysialic acid and NCAM functions in brain development. *Journal of neurochemistry* 2007;**103 Suppl 1**:56-64.

Chapter 6

Neuronal cell adhesion genes and antidepressant response in three independent samples

Chiara Fabbri, Concetta Crisafulli, David Gurwitz, Julia Stingl, Raffaella Calati, Diego Albani, Gianluigi Forloni, Marco Calabrò, Rosalba Martines, Siegfried Kasper, Joseph Zohar, Alzbeta Juven-Wetzler, Daniel Souery, Stuart Montgomery, Julien Mendlewicz, Giovanni De Girolamo, Alessandro Serretti

Published: Fabbri et al., Neuronal cell adhesion genes and antidepressant response in three independent samples. *The pharmacogenomics Journal* 2015;15:538-48.

Abstract

Drug-effect phenotypes in human lymphoblastoid cell lines (LCLs) recently allowed to identify *CHLI*, *GAP43*, *ITGB3* as new candidates for involvement in the antidepressant effect. *CHLI* and *ITGB3* code for adhesion molecules, while *GAP43* codes for a neuron-specific cytosolic protein expressed in neuronal growth cones; all 3 gene products are involved in synaptic plasticity.

16 polymorphisms in these genes were genotyped in two samples (n=369 and 90) with diagnosis of major depressive episode who were treated with antidepressants in a naturalistic setting. Phenotypes were response, remission, and treatment-resistant depression. Logistic regression including appropriate covariates was performed. Genes associated with outcomes were investigated in the STAR*D genome-wide study (n=1861) as both individual genes and through a pathway analysis (Reactome and String databases).

Gene-based analysis suggested *CHLI* rs4003413, *GAP43* rs283393 and rs9860828, *ITGB3* rs3809865 as the top candidates due to their replication across the largest original sample and the STAR*D cohort. *GAP43* molecular pathway was associated with both response and remission in the STAR*D, with *ELAVL4* representing the gene with the highest percent of SNPs associated with outcomes. Other promising genes emerging from the pathway analysis were *ITGB1* and *NRP1*.

The present study was the first to analyze cell adhesion genes and their molecular pathways in antidepressant response. Genes and biomarkers involved in neuronal adhesion should be considered by further studies aimed to identify predictors of antidepressant response.

Key words: *ITGB3*, *CHLI*, *GAP43*, neuronal adhesion, gene, antidepressant, major depression, treatment-resistance, pharmacogenetics, polymorphism

1. Introduction

Major depressive disorder (MDD) is a high-prevalence disorder that entails a heavy personal and socioeconomic burden in developed countries [1], mainly due to the high variability in antidepressant efficacy and lack of predictors of response in the individual patient. Antidepressant response has a relevant genetic component, since it shows a familiar clustering [2] and common SNPs only were estimated to explain 0.42 of variance in antidepressant response [3].

As we previously suggested [4], a gene-based analysis with the integration of findings obtained through the candidate gene and genome-wide approaches may represent a useful converging strategy in order to overcome the limitations of each individual methodology. Further, the study of individual polymorphisms and genes should be enhanced by the study of groups of genes and molecular pathways. Indeed, variations in pathways involved in antidepressant mechanisms of action are supposed to affect antidepressant response in an interrelated manner and probably with a “threshold” effect.

Taking into account these issues, the present study was focused on in the investigation of the effect on antidepressant response of three genes (*CHLI*, *ITGB3*, *GAP43*) and their molecular pathways that are involved in neuronal adhesion and synaptogenesis.

CHLI (cell adhesion molecule with homology to L1CAM) is a cell adhesion protein that plays a role in nervous system development and in synaptic plasticity, with involvement in mechanisms regulating structural reorganization and signal transduction implicated in learning and memory [5]. *CHLI* knockout mice exhibit reduced novelty seeking and impaired social behavior [6]. At neuroanatomical level they show misguided circuitry connecting the limbic system with the cerebral cortex [7, 8]. Moreover, *CHLI* expression is reduced in the hippocampus of mice exposed to early postnatal stress [9], a known risk factor for mood disorders. Consistently, the gene was implicated in psychiatric disorders pathogenesis, in detail bipolar disorder [10] and schizophrenia [11, 12]. A microarray study in lymphoblastoid cell lines (LCLs) exposed to paroxetine found *CHLI* as the most notable genome-wide transcriptome difference between LCLs displaying high versus low paroxetine sensitivity [13]. A following study demonstrated that *CHLI* lower expression in paroxetine-sensitive LCLs corresponded to increased expression of miR-151-3p, a microRNA that targets *CHLI* [14].

ITGB3 (Integrin Beta 3) is an adhesion protein implicated in synaptogenesis that regulates excitatory synaptic strength [15] and hippocampal AMPA receptors expression [16]. It has been implicated in schizophrenia [17] and autism [18], diseases that are both related to deficient brain plasticity. Importantly, *ITGB3* was

shown to be essential for the action of the serotonin transporter, the target of SSRI and TCA antidepressant drugs, in both blood platelets [19] and brain synaptosomes [20].

GAP43 (growth associated protein 43) is also named as ‘plasticity’ protein because it is expressed at high levels in neuronal growth cones during development and axonal regeneration [21, 22]. *GAP43* deficient mice express extreme stress sensitivity and autistic-like features [23] and the gene was demonstrated to be low-expressed in schizophrenia [24].

As well as for *CHL1*, the modulation of *GAP43* expression has been implicated in LCLs sensitivity to paroxetine [13], while *ITGB3* was shown to be upregulated in LCLs of patients treated with paroxetine compared to controls [25]. Thus, these genes were suggested as involved in antidepressant mechanisms of action. Neither *CHL1*, *ITGB3* or *GAP43* polymorphisms have been previously studied as putative genetic markers of antidepressant efficacy.

The present study aimed to investigate 16 polymorphisms harbored by *CHL1*, *GAP43* and *ITGB3* genes in two independent samples of MDD patients treated with antidepressant drugs and replicate the findings in the STAR*D genome-wide dataset. Both individual gene analysis and pathway analysis were performed.

2. Materials and Methods

2.1. Samples

2.1.1. European sample

285 MDD patients and 84 bipolar disorder (BD) patients (total sample size=369) were recruited in the context of the European multicenter project ‘Patterns of treatment resistance and switching strategies in unipolar affective disorder’. The study protocol was approved by the ethical committees of all participating centers. Detailed description of the whole sample has been previously reported [26]. Briefly, patients met DSM-IV criteria (Mini-international Neuropsychiatric Interview version 5.0.0 Modified for GSRD [27]) for major depressive episode defined as primary (i.e. not secondary to other Axis I disorders), but no minimal Hamilton Depression Rating Scale (HDRS) cut-off score at baseline was set as inclusion criteria. Eligible patients were treated with antidepressants in a naturalistic setting (Supplementary Table 1). Depressive symptoms were evaluated according to Hamilton Depression Rating Scale (HDRS, 21-item version) at baseline and after 4 weeks of treatment.

2.1.2. Italian sample

90 patients with diagnosis of non psychotic MDD (DSM-IV criteria) and with a score ≥ 13 on HDRS (21-item version) were included. Exclusion criteria were detailed elsewhere [28]. Eligible patients were treated with antidepressants in a

naturalistic setting (Supplementary Table 1). Patients were evaluated for depressive symptomatology (21-item HDRS) by trained psychiatrists at baseline and weekly until week 8.

Patients were outpatients recruited at the Department of Biomedical and NeuroMotor Sciences, Bologna University. Ethical approval was obtained from local research ethic committee.

2.1.3. Sequenced Treatment Alternatives to Relieve Depression (STAR*D)

Detailed descriptions of the study design and study population are detailed elsewhere [29]. In brief, non psychotic MDD (DSM-IV criteria) patients were enrolled from primary care or psychiatric outpatient clinics and a current 17-item Hamilton Depression Rating score of ≥ 14 by independent raters was obtained. Severity of depression was assessed using the 16-item Quick Inventory of Depressive Symptomatology-Clinician Rated (QIDS-C) [30] at baseline, weeks 2, 4, 6, 9, and 12. All patients received citalopram in level 1.

2.2. Phenotypes under investigation

The main investigated phenotypes were response and remission at weeks 4 and 8 in the Italian sample and at week 4 in the European sample and in the STAR*D, according to standard criteria (decrease of at least 50% in the HDRS-21 or the QIDS-C, score ≤ 7 of the HDRS-21 or ≤ 5 of the QIDS-C, respectively).

In the European sample treatment-resistance (TRD) was also investigated. Two definitions of treatment resistance have been considered in the analyses: 1) non response to at least 2 adequate consecutive antidepressant treatments administered during the last episode [26] (wide definition, TRD-W); 2) non response to at least 2 adequate consecutive antidepressant treatments of different classes administered during the last episode [31, 32] (different classes definition, TRD-DC).

2.3. Aims of the study

The primary aim of the study was to investigate the possible association between *CHL1*, *ITGB3* and *GAP43* polymorphisms and antidepressant response and remission in all the available samples. The secondary aim was to investigate the possible association between the selected genes and TRD in the European sample. In the European sample, the primary analysis was focused on the MDD subgroup, while a secondary analysis was performed in the whole sample (MDD+BD).

2.4. SNP selection and genotyping

Genetic SNPs were chosen according to the following criteria: 1) a reported prevalence of at least 5% for the variant allele among Caucasians (MAF=0.05 according to <http://hapmap.ncbi.nlm.nih.gov/> CEU population); 2) tagging approach ($R^2 \geq 0.08$); 3) position within the candidate gene boundaries; and 4) availability of a validated assay in our laboratory. We also considered variants not

investigated before. The list of genotyped SNPs is shown in Supplementary Table 2.

Genomic DNA was purified with an automated workstation (Maxwell, Promega) and checked for quality and quantity by a small scale spectrophotometer (Nanodrop, Thermo Scientific). Genotyping was performed using RFLP, ASO and melting curve analysis on Applied BioSystems 7500 Real-Time PCR System. Genotyping was performed according to the manufacturer's standard protocols. Forward and reverse primers' sequences are available upon request.

2.5. Statistical analysis

The effect of individual SNPs on phenotypes was tested through logistic regression models in all samples. Covariates were selected according to their impact on outcomes, while gender and age were used as covariates in all the analyses, and ancestry in the STAR*D since ancestry stratification has been previously reported in this sample [33]. Briefly, a complete agglomerative clustering was applied, based on a multidimensional scaling of a matrix of pairwise identity-by-state (IBS) values between samples, and clusters were defined on the base of the pairwise population concordance test ($PCC < 0.0001$, according to [34]). Identity-by-descent (IBD) analysis was used to identify related patients ($IBD > 0.1875$ [35]) in the STAR*D. Odds ratios (OR) with 95% confidence intervals (CI) were estimated for the effects of high-risk genotypes/alleles. In the European and Italian samples Haploview 3.2 was used to generate a linkage disequilibrium map.

In the STAR*D polymorphisms in the genes of interest were extracted from the genome-wide data according to gene physical position (Genome Build 36.3) and they were analyzed before and after imputation. Imputation was performed using IMPUTE2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) and 1000 Genomes data (NCBI Build 36 (dbSNP b126)) as reference panel. An info value threshold ≥ 0.8 was applied in order to prune poorly imputed SNPs.

Finally, the genes of interest was further investigated through a pathway analysis in the STAR*D. Reactome (www.reactome.org/) or String Interaction Network (<http://string-db.org>) were used to identify molecules that interact with *CHL1*, *ITGB3* and *GAP43*. Genes belonging to each pathway under analysis were imputed according to the method reported above. Variations showing $p < 0.05$ and $p < 0.01$ in each pathway under analysis were tested for a significant different distribution (Fisher exact test) compared to a random pathway. Each random pathway was matched with the index pathway in terms of number of SNPs within it and intragenic position of the SNPs but with random distribution within the genome. 10^4 permutations were run.

In single marker analysis, the whole European sample provided a power of 0.80 to detect risk alleles with $OR \geq 1.98$ setting alpha value to the exploratory value of 0.05 two-tailed and considering an explained variance of 0.02. Setting the same parameters, risk alleles with $OR \geq 3.7$ are detectable in the Italian sample and with $OR \geq 1.37$ in the STAR*D. Alpha value was set at the conservative value of 0.05 because the use of replication samples reduce the risk of false positive findings.

3. Results

Clinical demographic features and treatment of the analyzed samples are reported in Supplementary Table 1. In the European sample suicidal risk impacted on all clinical outcomes, and anxiety disorder comorbidity on response and remission (data not shown), thus they were used as covariates. In the Italian sample and STAR*D baseline severity affected phenotypes thus it was used as covariate. A complete list of the SNPs analyzed in the European, Italian and STAR*D (before and after imputation) is shown in Supplementary Table 2. A summary of pharmacogenetic results in the three samples is shown in Figure 1.

3.1. Results in the European sample

CHL1 rs2133402 G allele was associated with remission especially in the MDD subsample ($p=0.014$) but weakly also in the whole sample ($p=0.043$). *CHL1* rs4003413 C allele predicted better remission in the MDD subsample only (C vs. A $p=0.04$; AC vs. AA $p=0.013$). For both rs2133402 and rs4003413 there was a linear relationship between the number of G and C alleles, respectively, and remission ($t=2.90$, $p=0.004$; $t=2.14$, $p=0.033$, respectively). No SNP showed evidence of association with TRDW or TRDC.

GAP43 rs9860828 was associated with response (genotypic $p=0.03$; allelic $p=0.0088$), TRDW (allelic $p=0.02$) and TRDC (allelic $p=0.02$) in the MDD group. rs2028248 and rs283393 within the same gene were associated with remission (allelic $p=0.03$ for both of them).

GAP43 rs6803101 was associated with remission (allelic $p=0.001$; genotypic $p=0.0008$), TRDW (allelic $p=0.004$) and TRDC (allelic $p=0.018$) in the whole sample.

ITGB3 rs15908 (better outcome in AA and A carriers) and rs8074094 (better outcome in C carriers) were associated with remission (allelic $p=0.03$ and $p=0.007$, for the former and latter SNP respectively) in the MDD sample and whole sample, respectively.

ITGB3 rs3809865 was associated with response and remission in the whole sample ($p=0.003$ and $p=0.008$, respectively).

Significant results in the MDD sample (primary analysis) were reported in Table 1.

No evidence of linkage disequilibrium (LD) among the *CHL1*, *GAP43* and *ITGB3* SNPs was found (LD plots are shown in Supplementary Figure 1).

3.2. Results in the Italian sample

A weak correlation between *CHL1* rs2133402 and response at week 8 was found (genotypic $p=0.039$ and allelic $p=0.015$), but in the opposite direction compared to the European sample. No effect of this polymorphism was found for outcomes at week 4.

GAP43 rs6803101 G allele was a predictor of remission at week 8 (allelic $p=0.021$).

No SNP within the *ITGB3* gene was associated with any of the investigated phenotypes.

Significant results were reported in Table 1.

No evidence of linkage disequilibrium (LD) among the *CHL1*, *GAP43* and *ITGB3* SNPs was found (LD plots are shown in Supplementary Figure 1).

Table 1: results with $p<0.05$ in the MDD European (**A.**) and Italian (**B.**) samples. For each SNP both genotypic and allelic tests were performed. For genotypic test, the homozygote taken as reference was compared to the heterozygote and the opposite homozygote. TRD-W=treatment-resistance wide definition. TRD-DC=treatment-resistance different classes definition. w4=week 4. NS=non significant. CI=95% confidence interval.

A

Gene - SNP	w4 response	w4 remission	TRD-W	TRD-DC
CHL1 - rs4003413	NS	ACvsAA: $z=2.48$, $p=0.013$ (OR=3.80, CI=1.38-11.70) CvsA: $z=2.05$, $p=0.04$ (OR=1.93, CI=1.03-3.62)	NS	NS
CHL1 - rs2133402	NS	TvsG: $z=-2.46$, $p=0.014$ (OR=0.33, CI=0.14-0.82)	NS	NS
GAP43 - rs2028248	NS	CTvsCC: $z=-2.78$, $p=0.005$ (OR=0.15, 95% CI=0.04-0.56) TTvsCC: $z=-2.61$, $p=0.009$ (OR=0.24, 95% CI=0.08-0.72) TvsC: $z=-2.11$, $p=0.03$ (OR=0.52, 95%CI=0.28-0.97)	NS	NS

GAP43 - rs283393	NS	CTvsCC: z=2.31, p=0.02 (OR= 3.33, CI=1.16 -9.14) TvsC: z=2.19, p=0.03 (OR=2.06, CI=1.06-3.88)	NS	NS
GAP43 - rs9860828	AGvsAA: z=- 2.13, p=0.03 (OR=0.40, CI=0.16-0.89) GvsA: z=- 2.62, p=0.0088 (OR=0.34, CI=0.14-0.73)	NS	GvsA: z=2.27, p=0.02 (OR=3.16, CI=1.23-9.24)	GvsA: z=2.31,p=0.02 (OR=3.48, CI=1.26- 10.73)
ITGB3 - rs15908	NS	ACvsAA: z=-2.29, p=0.02 (OR=0.17, CI=0.04-0.78) CCvsAA: z=-2.60, p=0.009 (OR=0.16, CI=0.04-0.66) CvsA: z=-2.14, p=0.03 (OR=0.49, CI=0.26-0.96)	NS	NS

B

Gene - SNP	w8 response	w8 remission
CHL1 rs2133402	- TTvsGG: z=2.07, p=0.039 (OR=9.98, CI=1.50-200.19) TvsG: z=2.43, p=0.015 (OR=2.93, CI=1.30-7.01)	NS
GAP43 rs6803101	- NS	GvsA: z=2.29, p=0.021 (OR=2.46, CI=1.15-5.39)

3.3. Results in the STAR*D sample

Before imputation, the most interesting results in the *CHL1* gene were rs13091420 (3098 bp from rs4003413), rs17264626 (8731 bp from rs4003413) and rs931678 (3199 bp from rs2133402) that showed evidence of association with week 4 remission. Within the *GAP43* gene, rs12488667 showed an effect on response (p=0.003) and interestingly this SNP is located at only 889 bp from rs9860828, while rs4831199 and rs11926976 were only a bit farther (1328 bp and 8611 bp, respectively) from rs9860828 and both had evidence of impact on phenotypes. rs283393 that was associated with remission in the European sample had an influence on the same phenotype in the STAR*D (p=0.024). rs11655536 was the

only SNP within *ITGB3* that show a (weak) evidence of association with outcomes ($p=0.047$) and it is about 30 Kbp from rs8074094.

After imputation, two *GAP43* regions that are located ~9000 bp around rs12490536 and ~18 Kbp from rs9860828 represented the best predictors of response in this gene (Figure 2). A *CHLI* region extended from 915 bp to 7037 bp from rs4003413 (which boundaries were rs11714146 and rs1846448) showed evidence of association with outcomes, as well as rs11707773, rs11709854, and rs13061475 that lie from 6 Kbp to 15 Kbp from rs2133402 (Figure 2). Within the *ITGB3* gene rs7214096 was associated with both response and remission and is ~ 5 Kbp from rs15908; rs3809865 and a cluster of SNPs from 8 Kbp to 2 Kbp around it showed association with outcomes. Finally, rs8078614, rs7217214, rs55845779, rs11649785, rs8064871, and rs8069732 were among the top *ITGB3* SNPs and lie from 3 Kbp to 8 Kbp from rs8074094 (Figure 2).

Results before and after imputation of the genes of interest are reported in Table 2. The Reactome and String molecular pathways harboring the genes of interest are listed in Supplementary Table 3 and Table 3 summarizes the results of the pathway analysis. The *GAP43* pathway (Supplementary Figure 2) was associated with week 4 response and remission since a higher proportion of SNPs with $p<0.05$ within the index pathway than in the random pathway (permutated $p=0.017$ and 0.005 , respectively) was found. The top genes in the *GAP43* pathway were *ELAVL4*, *ZDHHC7* and *GRIN2B* (Table 3). The other pathways did not show evidence of association with outcomes, despite the *CHLI* interaction pathway (REACT_22292.1) showed a nominal significant p value and harbored the *ITGB1* gene which 35.56% of SNPs (32/90) had $p<0.01$.

Table 2: results in the STAR*D before (A.) and after imputation (B.) of the genes of interest. Only results with $p<0.05$ were reported in A. and with $p<0.01$ in B.. The number of available SNPs in each gene before and after imputation was reported in parenthesis. CI=95% confidence interval.

A

Gene	SNP	Phenotype	Test	Statistics
CHLI (32 SNPs)	rs17354487	W4 response	Additive	OR=0.73, CI=0.56-0.98, $p=0.03$
	rs17016919	W4 response	Dominant	OR=0.79, CI=0.64-0.98, $p=0.037$
	rs17028711	W4 remission	Recessive	OR=0.17, CI=0.04-0.71, $p=0.015$
	rs17017010	W4 remission	Dominant	OR=0.72, CI=0.54-0.95, $p=0.019$
	rs17264626	W4 remission	Additive	OR=0.69, CI=0.50-0.95, $p=0.02$

	rs13091420	W4 remission	Dominant	OR=0.73, CI=0.54-0.98, p=0.035
	rs931678	W4 remission	Recessive	OR=0.27, CI=0.08-0.93, p=0.038
	rs13091217	W4 remission	Dominant	OR=0.74, CI=0.55-0.99, p=0.043
GAP43 (52 SNPs)	rs12488667	W4 response	Additive	OR=1.76, CI=1.20-2.56, p=0.003
	rs4831199	W4 response	Dominant	OR=0.77, CI=0.61-0.98, p=0.033
	rs11926976	W4 remission	Dominant	OR=0.63, CI=0.46-0.86, p=0.003
	rs283393	W4 remission	Additive	OR=1.27, CI=1.03-1.56, p=0.024
ITGB3 (34 SNPs)	rs11655536	W4 remission	Additive	OR=0.79, CI=0.62-0.99, p=0.047

B

Gene	SNP	Phenotype	Info	Test	Statistics
CHL1	rs3933036	W4 response	0.93	Recessive	OR=0.59, CI=0.42-0.83, p=0.0006
		W4 remission		Recessive	OR=0.52, CI=0.31-0.86, p=0.0002
	rs7629803	W4 response	0.95	Recessive	OR=0.60, CI=0.43-0.84, p=0.0008
		W4 remission		Recessive	OR=0.53, CI=0.32-0.87, p=0.0008
	rs4541385	W4 response	0.94	Recessive	OR=0.60, CI=0.43-0.84, p=0.0008
		W4 remission		Recessive	OR=0.53, CI=0.32-0.88, p=0.0003
	rs3933034	W4 response	0.94	Recessive	OR=0.60, CI=0.43-0.84, p=0.0008
		W4 remission		Recessive	OR=0.53, CI=0.32-0.88, p=0.0003
	rs13061475	W4 response	0.90	Additive	OR=0.77, CI= 0.62-0.95, p=0.009
	rs4685598	W4 response	0.89	Dominant	OR=2.05, CI= 1.15-3.65, p=0.006
	rs11718181	W4 response	0.86	Recessive	OR=0.57, CI=0.36-0.91, p=0.007
	rs11715973	W4 response	0.89	Recessive	OR=0.57, CI=0.36-0.90, p=0.008
	rs17019589	W4 response	0.91	Recessive	OR=0.58, CI=0.37-0.89, p=0.008
	rs3901552	W4 response	0.92	Recessive	OR=0.59, CI=0.37-0.93, p=0.009
	rs11712380	W4 response	0.89	Recessive	OR=0.59, CI=0.38-0.91, p=0.0097
	rs11707773	W4 remission	0.83	Additive	OR=1.29, CI=1.03-1.62, p=0.003
rs11709854	W4 remission	0.90	Additive	OR=1.30, CI= 1.04-1.64 p=0.003	

	rs1846448	W4 remission	0.94	Additive	OR=1.30, CI=1.03-1.64, p=0.004
	rs6766301	W4 remission	0.96	Additive	OR=1.30, CI=1.03-1.64, p=0.004
	rs11714146	W4 remission	0.94	Additive	OR=1.29, CI=1.02-1.63, p=0.005
	rs58003001	W4 remission	0.96	Additive	OR=1.30, CI=1.03-1.64, p=0.005
	rs62228294	W4 remission	0.96	Additive	OR=1.30, CI=1.03-1.64, p=0.005
	rs12494871	W4 remission	0.96	Additive	OR=1.29, CI=1.03-1.63, p=0.005
	rs402356	W4 remission	0.88	Dominant	OR=0.59, CI=0.38-0.89, p=0.0008
	rs17264828	W4 remission	0.98	Additive	OR=1.28, CI=1.02-1.62, p=0.007
	rs9871112	W4 remission	0.87	Recessive	OR=1.54, CI=1.05-2.25, p=0.001
	rs17012663	W4 remission	0.88	Recessive	OR=1.72, CI=1.13-2.63, p=0.002
	rs331874	W4 remission	0.93	Recessive	OR=0.63, CI=0.42-0.93, p=0.004
	rs425366	W4 remission	0.97	Recessive	OR=0.63, CI=0.42-0.95, p=0.004
	rs332450	W4 remission	0.91	Recessive	OR=0.63, CI=0.42-0.94, p=0.004
	rs1909382	W4 remission	0.95	Recessive	OR=0.64, CI=0.43-0.97, p=0.004
	rs162731	W4 remission	0.90	Recessive	OR=0.63, CI=0.43-0.95, p=0.004
	rs9990181	W4 remission	0.99	Recessive	OR=0.61, CI=0.40-0.93, p=0.005
	rs191362	W4 remission	0.93	Recessive	OR=0.64, CI=0.43-0.95, p=0.005
	rs331893	W4 remission	0.88	Recessive	OR=0.67, CI=0.45-0.99, p=0.007
	rs331895	W4 remission	0.90	Recessive	OR=0.66, CI=0.44-0.98, p=0.007
	rs87018	W4 remission	0.89	Recessive	OR=0.67, CI=0.45-0.99, p=0.007
	rs331894	W4 remission	0.90	Recessive	OR=0.66, CI=0.45-0.98, p=0.008
GAP43 (123 SNPs)	rs12490536	W4 response	0.98	Additive	OR=0.79,9CI=0.67-0.95, p=0.006
	rs10934305	W4 response	0.95	Additive	OR=0.80, CI=0.67-0.96, p=0.008
	rs283373	W4 response	0.99	Recessive	OR=0.66, CI=0.47-0.94, p=0.008
ITGB3 (176 SNPs)	rs11079766	W4 response	0.98	Additive	OR=0.72, CI=0.61-0.85, p=5.18e-05
		W4 remission		Additive	OR=0.70, CI=0.55-0.88, p=0.002

rs11079767	W4 response	0.98	Additive	OR=0.72, CI=0.61-0.85, p=5.25e-05
	W4 remission		Additive	OR=0.70, CI=0.55-0.88, p=0.002
rs11658244	W4 response	0.98	Additive	OR=0.72, CI=0.61-0.85, p=5.48e-05
	W4 remission		Additive	OR=0.70, CI=0.55-0.88, p=0.002
rs11079768	W4 response	0.98	Additive	OR=0.72, CI=0.61-0.85, p=5.65e-05
	W4 remission		Additive	OR=0.69, CI=0.55-0.87, p=0.001
rs57044056	W4 response	0.98	Additive	OR=0.72, CI=0.61-0.86, p=5.69e-05
	W4 remission		Additive	OR=0.70, CI=0.56-0.88, p=0.002
rs59827675	W4 response	0.98	Additive	OR=0.72, CI=0.61-0.86, p=5.75e-05
	W4 remission		Additive	OR=0.70, CI=0.56-0.88, p=0.002
rs7220842	W4 response	0.99	Additive	OR=0.72, CI=0.61-0.86, p=5.86e-05
	W4 remission		Additive	OR=0.70, CI=0.56-0.88, p=0.002
rs7216442	W4 response	0.99	Additive	OR=0.72, CI=0.61-0.86, p=5.91e-05
	W4 remission		Additive	OR=0.70, CI=0.56-0.88, p=0.002
rs12451895	W4 response	0.99	Additive	OR=0.72, CI=0.61-0.86, p=5.95e-05
	W4 remission		Additive	OR=0.70, CI=0.56-0.88, p=0.002
rs12449371	W4 response	0.98	Additive	OR=0.72, CI=0.61-0.86, p=5.95e-05
	W4 remission		Additive	OR=0.70, CI=0.56-0.88, p=0.002
rs12452359	W4 response	0.99	Additive	OR=0.73, CI=0.61-0.86, p=6.18e-05
	W4 remission		Additive	OR=0.69, CI=0.55-0.87, p=0.001
rs1806687	W4 response	0.90	Dominant	OR=0.68, CI=0.55-0.85, p=0.0001
rs62074393	W4 response	0.88	Dominant	OR=0.65, CI=0.51-0.83, p=0.0002
	W4 remission		Dominant	OR=0.56, CI=0.39-0.79, p=0.0007
rs8078614	W4 response	0.94	Dominant	OR=0.70, CI=0.56-0.87, p=0.0002
	W4 remission		Additive	OR=0.71, CI=0.56-0.91, p=0.007
rs7217214	W4 response	0.94	Dominant	OR=0.70, CI=0.56-0.87, p=0.0002

		W4 remission		Additive	OR=0.70, CI=0.55-0.89, p=0.006
rs55845779	W4 response	0.94		Dominant	OR=0.70, CI=0.56-0.87, p=0.0002
	W4 remission			Additive	OR=0.70, CI=0.55-0.89, p=0.006
rs11649785	W4 response	0.93		Dominant	OR=0.70, CI=0.56-0.87, p=0.0002
	W4 remission			Additive	OR=0.70, CI=0.55-0.89, p=0.006
rs1165373303	W4 response	0.86		Dominant	OR=0.66, CI=0.51-0.85, p=0.0003
rs884696	W4 response	0.90		Additive	OR=0.73, CI=0.61-0.88, p=3.40e-04
rs62074394	W4 response	0.88		Dominant	OR=0.66, CI=0.51-0.86, p=0.0004
	W4 remission			Additive	OR=0.61, CI=0.44-0.84, p=0.002
rs62074395	W4 response	0.91		Dominant	OR=0.68, CI=0.53-0.87, p=0.0005
	W4 remission			Additive	OR=0.65, CI=0.48-0.88, p=0.007
rs8077238	W4 response	0.91		Dominant	OR=0.71, CI=0.56-0.89, p=0.0007
	W4 remission			Additive	OR=0.67, CI=0.51-0.88, p=0.004
rs62074396	W4 response	0.93		Dominant	OR=0.68, CI=0.53-0.87, p=0.0007
	W4 remission			Additive	OR=0.66, CI=0.49-0.89, p=0.008
rs8073179	W4 response	0.93		Dominant	OR=0.68, CI=0.53-0.87, p=0.0007
	W4 remission			Additive	OR=0.65, CI=0.48-0.88, p=0.007
rs11653733	W4 response	0.86		Additive	OR=0.70, CI=0.56-0.88, p=0.0008
	W4 remission			Dominant	OR=0.57, CI=0.40-0.82, p=0.0006
rs4629025	W4 response	0.99		Recessive	OR=0.60, CI=0.44-0.82, p=0.001
rs8064871	W4 response	0.96		Dominant	OR=0.70, CI=0.54-0.89, p=0.0016
	W4 remission			Additive	OR=0.64, CI=0.48-0.87, p=0.006
rs62074428	W4 response	0.95		Dominant	OR=0.70, CI=0.55-0.90, p=0.0019
	W4 remission			Additive	OR=0.65, CI=0.48-0.87, p=0.007
rs5918	W4 response	0.94		Dominant	OR=0.70, CI=0.55-0.90, p=0.0019
	W4 remission			Additive	OR=0.65, CI=0.48-0.88, p=0.007

rs17218711	W4 response	0.93	Dominant	OR=0.71, CI=0.55-0.91, p=0.002
	W4 remission		Additive	OR=0.65, CI=0.48-0.88, p=0.007
rs7214096	W4 response	0.91	Dominant	OR=0.71, CI=0.55-0.91, p=0.002
	W4 remission		Additive	OR=0.65, CI=0.48-0.88, p=0.007
rs11650022	W4 response	0.89	Dominant	OR=0.72, CI=0.56-0.92, p=0.002
	W4 remission		Additive	OR=0.66, CI=0.49-0.89, p=0.008
rs60544493	W4 response	0.86	Recessive	OR=0.27, CI=0.09-0.84, p=0.003
rs17605348	W4 response	0.85	Recessive	OR=0.28, CI=0.09-0.85, p=0.003
rs58097871	W4 response	0.85	Recessive	OR=0.29, CI=0.10-0.87, p=0.003
rs8069732	W4 response	1.00	Additive	OR=0.74, CI=0.59-0.91, p=0.004
	W4 remission		Additive	OR=0.63, CI=0.46-0.86, p=0.006
rs7223956	W4 response	0.81	Additive	OR=1.32, CI=1.06-1.66, p=0.005
	W4 remission		Additive	OR=1.57, CI=1.14-2.16, p=0.003
rs11652097	W4 response	0.99	Additive	OR=0.83, CI=0.72-0.96, p=0.0099
rs9303534	W4 remission	0.86	Additive	OR=1.36, CI=1.10-1.70, p=0.004
rs7224753	W4 remission	0.87	Additive	OR=1.34, CI=1.08-1.66, p=0.005
rs12600865	W4 remission	0.95	Additive	OR=1.33, CI=1.07-1.64, p=0.006
rs11658426	W4 remission	0.95	Additive	OR=1.33, CI=1.07-1.64, p=0.006
rs3809865	W4 remission	0.96	Additive	OR=1.33, CI=1.08-1.64, p=0.006
rs35435669	W4 remission	0.89	Additive	OR=1.33, CI=1.07-1.64, p=0.006
rs4968314	W4 remission	0.89	Additive	OR=1.33, CI=1.07-1.64, p=0.006
rs7218813	W4 remission	0.89	Additive	OR=1.33, CI=1.07-1.64, p=0.006
rs6504836	W4 remission	0.89	Additive	OR=1.33, CI=1.07-1.64, p=0.006
rs4968316	W4 remission	0.89	Additive	OR=1.33, CI=1.07-1.64, p=0.006
rs2271803	W4 remission	0.89	Additive	OR=1.33, CI=1.07-1.64, p=0.006
rs7221232	W4 remission	0.89	Additive	OR=1.33, CI=1.07-1.64, p=0.006

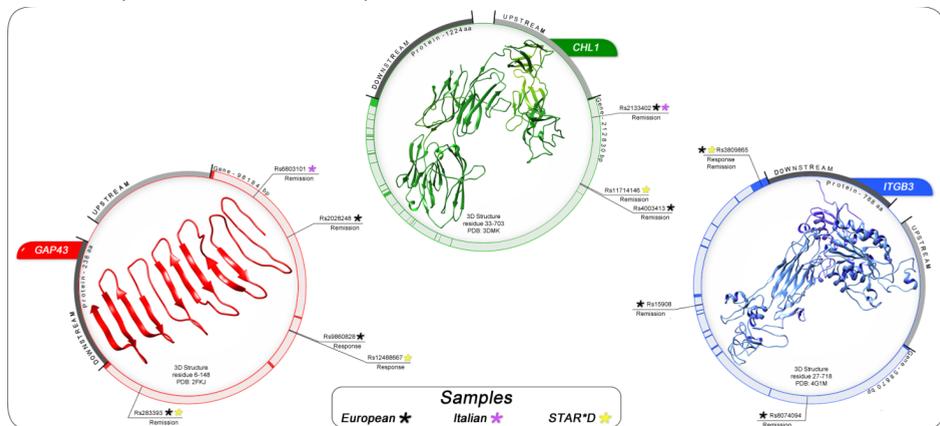
rs8075410	W4 remission	0.95	Additive	OR=1.33, CI=1.07-1.64, p=0.006
rs9893410	W4 remission	0.95	Additive	OR=1.33, CI=1.07-1.64, p=0.006
rs35161461	W4 remission	0.93	Additive	OR=1.32, CI=1.07-1.63, p=0.006
rs3760372	W4 remission	0.87	Additive	OR=1.32, CI=1.07-1.62, p=0.006
rs1969267	W4 remission	0.87	Additive	OR=1.32, CI=1.07-1.62, p=0.006
rs35983484	W4 remission	0.84	Additive	OR=1.31, CI=1.06-1.62, p=0.007
rs9894860	W4 remission	0.82	Additive	OR=1.28, CI=1.05-1.56, p=0.007
rs34589418	W4 remission	0.84	Additive	OR=1.30, CI=1.06-1.59, p=0.008
rs6504833	W4 remission	0.85	Additive	OR=1.29, CI=1.06-1.57, p=0.008
rs7225700	W4 remission	0.85	Additive	OR=1.29, CI=1.06-1.57, p=0.008
rs8068200	W4 remission	0.86	Additive	OR=1.29, CI=1.06-1.57, p=0.008
rs2317677	W4 remission	0.85	Additive	OR=1.29, CI=1.06-1.57, p=0.008
rs11650072	W4 remission	0.85	Additive	OR=1.29, CI=1.06-1.57, p=0.008
rs9899121	W4 remission	0.84	Additive	OR=1.28, CI=1.06-1.56, p=0.008

Table 3: results of the pathway analysis. P values were permuted only if they were < 0.05. CI=95% confidence interval.

Gene/pathway	Results for p<0.05	Results for p<0.01	Top genes
CHL1 / REACT_22292.1	Response: p=0.71 (OR=0.97, CI=0.82-1.14) Remission: p=0.13 (OR=0.87, CI=0.74-1.04)	Response: p=0.015 (OR=0.67, CI=0.49-0.93); permuted p=0.32 Remission: p=0.16 (OR=0.80, CI=0.58-1.09)	Response and p<0.01: ITGB1 32/90 SNPs (35.56%) NRP1 39/511 SNPs (7.63%)
GAP43 / String 9.1 database	Response: p=9.02e-06 (OR=0.65, CI=0.53-0.79); permuted p=0.017* Remission: p=3.35e-06 (OR=0.64, CI=0.53-0.77); permuted p=0.005*	Response: p=0.006 (OR=0.56, CI=0.36-0.85); permuted p=0.83 Remission: p=0.0087 (OR=0.60, CI=0.40-0.88); permuted p=0.45	Response and p<0.05: ELAVL4 17/30 SNPs (56.67%) ZDHHC7 24/112 (21.43%) GRIN2B 235/1280 SNPs (18.36%) Remission and p<0.05: ELAVL4 17/30 (56.67%) ZDHHC7 24/112 (21.43%)

			GRIN2B 242/1280 (18.91%)
ITGB3 / REACT_12519.1	Response: p=0.18 (OR=0.83, CI=0.63-1.09) Remission: p=0.78 (OR=0.95, CI=0.72-1.26)	Response: p=0.36 (OR=1.31, CI=0.76-2.29) Remission: p=0.51 (OR=1.24, CI=0.71-2.17)	Response and p<0.05: PLCG1 18/36 SNPs (50.00%) YES1 66/164 SNPs (40.24%)
ITGB3 / REACT_15381.1	Response: p=2.70e-15 (OR=3.58, CI=2.53-5.16) Remission: -	Response: p=6.56e-05 (OR=4.62, CI=2.00-12.43) Remission: -	-
ITGB3 / REACT_22272.1	Response: p=0.15 (OR=1.12, CI=0.96-1.32) Remission: p=0.57, (OR=1.05, CI=0.89-1.23)	Response: p=0.21 (OR=0.81, CI=0.59-1.12) Remission: p=0.70, (OR=0.93, CI=0.68-1.27)	Response and p<0.05: ITGB1 54/90 SNPs (60.00%) NRP1 139/511 SNPs (27.20%) RAC1 14/71 SNPs (19.72%) ITGB3 18/166 SNPs (10.84%)
ITGB3 / REACT_16394.2.2	Response: p=0.002 (OR=1.25, CI=1.08-1.45) Remission: -	Response: p=3.57e-09 (OR=2.70, CI=1.90-3.90) Remission: -	-

Figure 1: the SNPs that were found associated with the outcomes and their positions were reported. Data on genes and the position of SNPs were obtained from NCBI GRCh38 assembly. The 3D structure was obtained through Chimera software (Pettersen et al. 2004).



4. Discussion

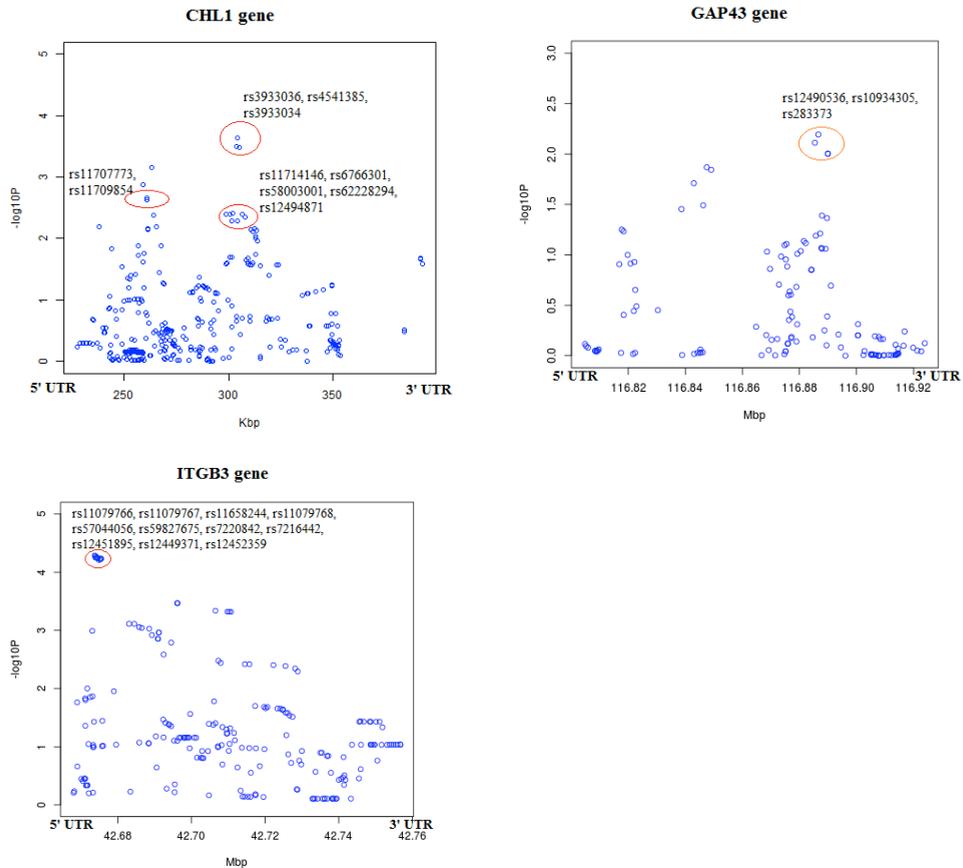
4.1. Main findings

The present study investigated the effect of polymorphisms in three genes (*CHLI*, *GAP43*, *ITGB3*) involved in neural adhesion and synaptic plasticity on antidepressant efficacy. These genes were selected on the basis of previous microarray studies in human lymphoblastoid cell lines (LCLs) that suggested them as a transcriptome biomarkers of paroxetine sensitivity [13] or chronic *in vivo* paroxetine exposure [25] and the lack of data about their role in antidepressant response in humans. For our purpose, 16 polymorphisms were genotyped in two independent original samples of MDD patients and the STAR*D genome-wide dataset was used as replication sample.

CHLI gene codes for a cell adhesion protein implicated in neurodevelopment and neuroplasticity [5]. Gene-based analysis of *CHLI* suggested rs4003413 as the best candidate, since its replication across the European sample and the STAR*D (in the STAR*D rs4003413 was not available, but rs11714146 was at only ~ 900 bp from it and it was associated with outcomes together with a surrounding cluster of SNPs). *CHLI* rs2133402 had an effect on outcomes in both original samples, but it showed a weak effect in the Italian one that was opposite to that identified in the European sample. Both rs4003413 and rs2133402 are located in the intron V of *CHLI*, thus it is plausible that they may be in LD with exonic or regulatory SNPs maybe in the near last exon of the gene or 3' UTR. Despite *CHLI* interaction pathway showed only nominal association with response in the STAR*D, it outlined another interesting candidate belonging to the integrin family, i.e. the *ITGB1* gene (Table 3).

GAP43 is expressed at high levels in neuronal growth cones during development and axonal regeneration [21, 22]. Within the *GAP43* gene rs283393 was associated with remission in both the European and STAR*D sample, while rs9860828 was found a predictor of response, TRDW and TRDC in the former sample and rs12488667 (only ~ 800 bp from rs9860828) was associated with response in the latter sample. All these three SNPs are intronic. The analysis of *GAP43* interaction pathway demonstrated an effect on both response and remission, and the pathway genes with the highest proportion of SNPs associated with these phenotypes were *ELAVL4*, *ZDHHC7* and *GRIN2B* (Table 3). *ELAVL4* (ELAV like neuron-specific RNA binding protein 4) may play a role in neuron-specific RNA processing [36] and protects *CDKN1A* (cyclin-dependent kinase inhibitor 1A (p21)) mRNA from decay.

Figure 2: graphical representation of results obtained for CHL1, GAP43 and ITGB3 after imputation in STAR*D.



In vitro and *in vivo* (human and mice) experiments showed that hippocampal p21 levels are correlated with antidepressant administration and they are considered a marker of antidepressant-induced hippocampal neurogenesis [37, 38]. *ZDHHC7* (zinc finger, DHHC-type containing 7) codes for a palmitoyltransferase that has GABA gamma subunits (GABRG1, GABRG2 and GABRG3) among its targets and thereby regulate their synaptic clustering and/or cell surface stability. Further, *ZDHHC7* palmitoylates sex steroid hormone receptors affecting the rapid intracellular signaling by sex hormones via ERK and AKT kinases [39]; sex steroid hormones are known to modulate synaptic plasticity [40], mood, and antidepressant response [41]. *GRIN2B* (glutamate receptor ionotropic N-methyl D-aspartate 2) can be hypothesized to be involved in antidepressant action since the increasing evidence supporting the involvement of the glutamatergic system in MDD and antidepressant action [33]. Interestingly, increased *GRIN2B* expression

in the locus coeruleus (LC) was demonstrated in MDD, suggesting disrupted glutamatergic-noradrenergic interactions at this level that may be responsible for the rapid antidepressant effects of glutamate antagonists [42].

ITGB3 regulates excitatory synaptic strength [15] and hippocampal AMPA receptors expression [16]. Within this gene rs3809865 was associated with remission in the European sample and the finding was replicated in the STAR*D together with a cluster of SNPs from 8 Kbp to 2 Kbp around this SNP. Some of the SNPs in this cluster around rs3809865 were in the downstream region of *ITGB3* (rs6504836, rs4968316, rs2271803, rs7221232, rs8075410, rs9893410) and may thus have regulatory roles. None of the *ITGB3* pathways showed evidence of association with phenotypes, but the top genes interacting with *ITGB3* were *PLCG1*, *YES1*, *ITGB1*, *NRP1*, and *RAC1* (Table 3). *PLCG1* (phospholipase C, gamma 1) mediates the production of the second messenger molecules diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP3) and thus it plays an important role in the intracellular transduction of receptor-mediated tyrosine kinase activators. *PLCG1* was suggested as a predictor of lithium response [43], and animal experiments demonstrated the activation of phospholipase C gamma signaling pathways in mouse brain after exposure to antidepressants [44]. *NRP1* (neuropilin 1) is involved in several signaling pathways that participate to the regulation of cell migration and attraction and its role in neuronal guidance is mediated by its interaction with class 3A semaphorin (Sema3A) [45]. It has been established that Sema3A signaling through NRP-1 triggers the sequential inhibition of phosphatidylinositol 3-kinase (PI3K) and protein kinase B (Akt) leading to activation of glycogen synthase kinase-3beta (GSK-3b) [46]. Specific inhibitors of GSK-3 have antidepressant behavioral activity [47]. *NRP1* gene was found upregulated in depressed patients [48] and *NRP1* signaling partners (e.g. *VEGF*) have been shown to be regulated by stress, antidepressants, and depressive disorder [49, 50]. *RAC1* (RAS-related C3 botulinum substrate 1) is a GTPase which belongs to the RAS superfamily of small GTP-binding proteins and it acts as critical modulator of synaptic structure. In mice *RAC1* modulates depression-like behaviors and its downregulation has strong anti-stress efficacy [51]. *YES1* (v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1) codes for a non-receptor protein tyrosine kinase that is involved in the regulation of several cell processes, among which cell growth and survival, apoptosis, cell-cell adhesion, cytoskeleton remodeling, and differentiation. It has not been investigated in relation to MDD and antidepressant response by previous studies.

4.2. Limitations

The present results should be considered together with the limitations of study. First the heterogeneity among the samples under analysis in terms of treatment and scales of evaluation represents a limitation. On the other hand, the replication of results across different samples supports their potential applicability in different clinical context. Other limitations consist in the small size of the Italian sample and in the retrospective design of the European study. No correction for multiple-testing was performed in SNP-based analysis, even if the replication of results across three independent samples is expected to balance the risk of false positive findings. Finally, only limited availability of the variants genotyped in the two original samples was found in the STAR*D GWAS and this could be only partially overcome through imputation. Imputation itself provide reliable results but should be verified through genotyping, thus results on imputed SNPs should be considered only preliminary. Despite the limited correspondence of SNPs genotyped in the original samples and STAR*D, the focus of the present study is on genes and not on individual SNPs, since multilocus models are supposed to modulate the antidepressant effect.

4.3. Conclusion

Our study concluded that *GAP43* gene and its pathway (in particular, *ELAVL4*, *ZDHHC7* and *GRIN2B* genes), *CHL1* with the interacting *ITGB1* gene, and *ITGB3* with the interacting *NRPI* gene represent promising candidate genes implicated in antidepressant response. Their molecular pathways should be carefully taken into account in further investigations.

Supplementary materials:

Supplementary materials are available for the online version of this publication.

References

1. Judd LL, Akiskal HS, Zeller PJ, Paulus M, Leon AC, Maser JD, et al.: Psychosocial disability during the long-term course of unipolar major depressive disorder. *Arch Gen Psychiatry* 2000; 57: 375-80.
2. Franchini L, Serretti A, Gasperini M, Smeraldi E: Familial concordance of fluvoxamine response as a tool for differentiating mood disorder pedigrees. *J Psychiatr Res* 1998; 32: 255-9.
3. Tansey KE, Guipponi M, Hu X, Domenici E, Lewis G, Malafosse A, et al.: Contribution of common genetic variants to antidepressant response. *Biol Psychiatry* 2013; 73: 679-82.
4. Fabbri C, Marsano A, Albani D, Chierchia A, Calati R, Drago A, et al.: PPP3CC gene: a putative modulator of antidepressant response through the B-cell receptor signaling pathway. *Pharmacogenomics J* 2014; 14:463-72.

5. Maness PF, Schachner M: Neural recognition molecules of the immunoglobulin superfamily: signaling transducers of axon guidance and neuronal migration. *Nat Neurosci* 2007; 10: 19-26.
6. Morellini F, Lepsveridze E, Kahler B, Dityatev A, Schachner M: Reduced reactivity to novelty, impaired social behavior, and enhanced basal synaptic excitatory activity in perforant path projections to the dentate gyrus in young adult mice deficient in the neural cell adhesion molecule CHL1. *Mol Cell Neurosci* 2007; 34: 121-36.
7. Demyanenko GP, Siesser PF, Wright AG, Brennaman LH, Bartsch U, Schachner M, et al.: L1 and CHL1 Cooperate in Thalamocortical Axon Targeting. *Cereb Cortex* 2011; 21: 401-12.
8. Demyanenko GP, Halberstadt AI, Rao RS, Maness PF: CHL1 cooperates with PAK1-3 to regulate morphological differentiation of embryonic cortical neurons. *Neuroscience* 2010; 165: 107-15.
9. Desarnaud F, Jakovcevski M, Morellini F, Schachner M: Stress downregulates hippocampal expression of the adhesion molecules NCAM and CHL1 in mice by mechanisms independent of DNA methylation of their promoters. *Cell Adh Migr* 2008; 2: 38-44.
10. Noor A, Lionel AC, Cohen-Woods S, Moghimi N, Rucker J, Fennell A, et al.: Copy number variant study of bipolar disorder in Canadian and UK populations implicates synaptic genes. *Am J Med Genet B Neuropsychiatr Genet* 2014; 165B: 303-13.
11. Tam GW, van de Lagemaat LN, Redon R, Strathdee KE, Croning MD, Malloy MP, et al.: Confirmed rare copy number variants implicate novel genes in schizophrenia. *Biochem Soc Trans* 2010; 38: 445-51.
12. Chu TT, Liu Y: An integrated genomic analysis of gene-function correlation on schizophrenia susceptibility genes. *J Hum Genet* 2010; 55: 285-92.
13. Morag A, Pasmanik-Chor M, Oron-Karni V, Rehavi M, Stingl JC, Gurwitz D: Genome-wide expression profiling of human lymphoblastoid cell lines identifies CHL1 as a putative SSRI antidepressant response biomarker. *Pharmacogenomics* 2011; 12: 171-84.
14. Oved K, Morag A, Pasmanik-Chor M, Oron-Karni V, Shomron N, Rehavi M, et al.: Genome-wide miRNA expression profiling of human lymphoblastoid cell lines identifies tentative SSRI antidepressant response biomarkers. *Pharmacogenomics* 2012; 13: 1129-39.
15. Cingolani LA, Goda Y: Differential involvement of beta3 integrin in pre- and postsynaptic forms of adaptation to chronic activity deprivation. *Neuron Glia Biol* 2008; 4: 179-87.
16. Pozo K, Cingolani LA, Bassani S, Laurent F, Passafaro M, Goda Y: beta3 integrin interacts directly with GluA2 AMPA receptor subunit and regulates AMPA receptor expression in hippocampal neurons. *Proc Natl Acad Sci U S A* 2012; 109: 1323-8.
17. Wang KS, Liu X, Arana TB, Thompson N, Weisman H, Devargas C, et al.: Genetic association analysis of ITGB3 polymorphisms with age at onset of schizophrenia. *J Mol Neurosci* 2013; 51: 446-53.
18. Napolioni V, Lombardi F, Sacco R, Curatolo P, Manzi B, Alessandrelli R, et al.: Family-based association study of ITGB3 in autism spectrum disorder and its endophenotypes. *Eur J Hum Genet* 2011; 19: 353-9.

19. Carneiro AM, Cook EH, Murphy DL, Blakely RD: Interactions between integrin alphaIIb beta3 and the serotonin transporter regulate serotonin transport and platelet aggregation in mice and humans. *J Clin Invest* 2008; 118: 1544-52.
20. Whyte A, Jessen T, Varney S, Carneiro AM: Serotonin transporter and integrin beta 3 genes interact to modulate serotonin uptake in mouse brain. *Neurochem Int* 2014; 73: 122-6.
21. Verhaagen J, Oestreicher AB, Grillo M, Khew-Goodall YS, Gispén WH, Margolis FL: Neuroplasticity in the olfactory system: differential effects of central and peripheral lesions of the primary olfactory pathway on the expression of B-50/GAP43 and the olfactory marker protein. *J Neurosci Res* 1990; 26: 31-44.
22. Leu B, Koch E, Schmidt JT: GAP43 phosphorylation is critical for growth and branching of retinotectal arbors in zebrafish. *Dev Neurobiol* 2010; 70: 897-911.
23. Zaccaria KJ, Lagace DC, Eisch AJ, McCasland JS: Resistance to change and vulnerability to stress: autistic-like features of GAP43-deficient mice. *Genes Brain Behav* 2010; 9: 985-96.
24. Fung SJ, Sivagnanasundaram S, Weickert CS: Lack of change in markers of presynaptic terminal abundance alongside subtle reductions in markers of presynaptic terminal plasticity in prefrontal cortex of schizophrenia patients. *Biol Psychiatry* 2011; 69: 71-9.
25. Oved K, Morag A, Pasmanik-Chor M, Rehavi M, Shomron N, Gurwitz D: Genome-wide expression profiling of human lymphoblastoid cell lines implicates integrin beta-3 in the mode of action of antidepressants. *Transl Psychiatry* 2013; 3: e313.
26. Souery D, Oswald P, Massat I, Bailer U, Bollen J, Demyttenaere K, et al.: Clinical factors associated with treatment resistance in major depressive disorder: results from a European multicenter study. *J Clin Psychiatry* 2007; 68: 1062-70.
27. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al.: 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* 1998; 59 Suppl 20: 22-33; quiz 34-57.
28. Fabbri C, Marsano A, Balestri M, De Ronchi D, Serretti A: Clinical features and drug induced side effects in early versus late antidepressant responders. *J Psychiatr Res* 2013; 47: 1309-18.
29. Howland RH: Sequenced Treatment Alternatives to Relieve Depression (STAR*D). Part 1: study design. *J Psychosoc Nurs Ment Health Serv* 2008; 46: 21-4.
30. Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al.: The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med* 2004; 34: 73-82.
31. CHMP. *Note for guidance on clinical investigation of medicinal products in the treatment of depression 2002*. 2002; Available from: <http://www.emea.europa.eu/pdfs/human/ewp/051897en.pdf>.
32. Thase ME: The need for clinically relevant research on treatment-resistant depression. *J Clin Psychiatry* 2001; 62: 221-4.

33. Fabbri C, Drago A, Serretti A: Early antidepressant efficacy modulation by glutamatergic gene variants in the STAR(D). *Eur Neuropsychopharmacol* 2012; 23:612-21.
34. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al.: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; 81: 559-75.
35. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT: Data quality control in genetic case-control association studies. *Nat Protoc* 2010; 5: 1564-73.
36. Bronicki LM, Jasmin BJ: Emerging complexity of the HuD/ELAV14 gene; implications for neuronal development, function, and dysfunction. *RNA* 2013; 19: 1019-37.
37. Pechnick RN, Zonis S, Wawrowsky K, Cosgayon R, Farrokhi C, Lacayo L, et al.: Antidepressants stimulate hippocampal neurogenesis by inhibiting p21 expression in the subgranular zone of the hippocampus. *PLoS One* 2011; 6: e27290.
38. Epp JR, Beasley CL, Galea LA: Increased hippocampal neurogenesis and p21 expression in depression: dependent on antidepressants, sex, age, and antipsychotic exposure. *Neuropsychopharmacology* 2013; 38: 2297-306.
39. Pedram A, Razandi M, Deschenes RJ, Levin ER: DHHC-7 and -21 are palmitoylacyltransferases for sex steroid receptors. *Mol Biol Cell* 2012; 23: 188-99.
40. Ooishi Y, Kawato S, Hojo Y, Hatanaka Y, Higo S, Murakami G, et al.: Modulation of synaptic plasticity in the hippocampus by hippocampus-derived estrogen and androgen. *J Steroid Biochem Mol Biol* 2012; 131: 37-51.
41. Parker G, Brotchie H: Gender differences in depression. *Int Rev Psychiatry* 2010; 22: 429-36.
42. Chandley MJ, Szebeni A, Szebeni K, Crawford JD, Stockmeier CA, Turecki G, et al.: Elevated gene expression of glutamate receptors in noradrenergic neurons from the locus coeruleus in major depression. *Int J Neuropsychopharmacol* 2014: 1-10.
43. Ikeda A, Kato T: Biological predictors of lithium response in bipolar disorder. *Psychiatry Clin Neurosci* 2003; 57: 243-50.
44. Rantamaki T, Hendolin P, Kankaanpaa A, Mijatovic J, Piepponen P, Domenici E, et al.: Pharmacologically diverse antidepressants rapidly activate brain-derived neurotrophic factor receptor TrkB and induce phospholipase-Cgamma signaling pathways in mouse brain. *Neuropsychopharmacology* 2007; 32: 2152-62.
45. Kolodkin AL, Leventgood DV, Rowe EG, Tai YT, Giger RJ, Ginty DD: Neuropilin is a semaphorin III receptor. *Cell* 1997; 90: 753-62.
46. Pasterkamp RJ, Giger RJ: Semaphorin function in neural plasticity and disease. *Curr Opin Neurobiol* 2009; 19: 263-74.
47. Gould TD, Picchini AM, Einat H, Manji HK: Targeting glycogen synthase kinase-3 in the CNS: implications for the development of new treatments for mood disorders. *Curr Drug Targets* 2006; 7: 1399-409.
48. Goswami DB, Jernigan CS, Chandran A, Iyo AH, May WL, Austin MC, et al.: Gene expression analysis of novel genes in the prefrontal cortex of major depressive disorder subjects. *Prog Neuropsychopharmacol Biol Psychiatry* 2013; 43: 126-33.

49. Warner-Schmidt JL,Duman RS: VEGF is an essential mediator of the neurogenic and behavioral actions of antidepressants. *Proc Natl Acad Sci U S A* 2007; 104: 4647-52.
50. Warner-Schmidt JL,Duman RS: VEGF as a potential target for therapeutic intervention in depression. *Curr Opin Pharmacol* 2008; 8: 14-9.
51. Golden SA, Christoffel DJ, Heshmati M, Hodes GE, Magida J, Davis K, et al.: Epigenetic regulation of RAC1 induces synaptic remodeling in stress disorders and depression. *Nat Med* 2013; 19: 337-44.

Chapter 7

Genetics of long-term treatment outcome in bipolar disorder

Chiara Fabbri, Alessandro Serretti

Published: Fabbri and Serretti, Genetics of long-term treatment outcome in bipolar disorder. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 2016; 65:17-24.

Abstract

Bipolar disorder (BD) shows one of strongest genetic predisposition among psychiatric disorders and the identification of reliable genetic predictors of treatment response could significantly improve the prognosis of the disease.

The present study investigated genetic predictors of long-term treatment-outcome in 723 patients with BD type I from the STEP-BD (Systematic Treatment Enhancement Program for Bipolar Disorder) genome-wide dataset. BD I patients with > 6 months of follow-up and without any treatment restriction (reflecting a natural setting scenario) were included. Phenotypes were the total and depressive episode rates and the occurrence of one or more (hypo)manic/mixed episode during follow-up. Quality control of genome-wide data was performed according to standard criteria and linear/logistic regression models were used as appropriate under an additive hypothesis. Top genes were further analyzed through a pathway analysis.

Genes previously involved in the susceptibility to BD (*DFNB31*, *SORCS2*, *NRXN1*, *CNTNAP2*, *GRIN2A*, *GRM4*, *GRIN2B*), antidepressant action (*DEPTOR*, *CHRNA7*, *NRXN1*), and mood stabilizer or antipsychotic action (*NTRK2*, *CHRNA7*, *NRXN1*) may affect long-term treatment outcome of BD. Promising findings without previous strong evidence were *TRAF3IP2-AS1*, *NFYC*, *RNLS*, *KCNJ2*, *RASGRF1*, *NTF3* genes. Pathway analysis supported particularly the involvement of molecules mediating the positive regulation of MAPK cascade and learning/memory processes.

Further studies focused on the outlined genes may be helpful to provide validated markers of BD treatment outcome.

Keywords: bipolar disorder; treatment response; gene; pharmacogenomics; pathway analysis

1. Introduction

Bipolar disorder (BD) is a chronic disease characterized by the alternation of periods of (hypo)mania and depression, resulting in high personal and socio-economic burden in terms of poor quality of life, increased rates of suicide, direct and indirect costs (Whiteford et al. 2013). An adequate treatment may allow for long-term remission and good functioning in the majority of patients, but the lack of biological markers for guiding drug choice makes often difficult the identification of the most effective treatment. Limited treatment efficacy and side effects often due to polypharmacy contribute to treatment non-adherence that occurs at a rate between 12% and 64% among individuals with BD (Leclerc et al. 2013).

BD shows one of highest genetic predisposition among psychiatric disorders and the heritability index is estimated to be 0.85 (McGuffin et al. 2003). Genetics accounts for 20% to 95% of variability in CNS drug disposition and pharmacodynamics (Cacabelos et al. 2012), supporting the hypothesis that treatment efficacy in BD may significantly be affected by genetic variants.

Previous pharmacogenetic studies were mainly focused on the investigation of predictors of lithium response and they applied both candidate gene and genome-wide association approaches.

Candidate gene studies mainly investigated genes pertaining to the monoaminergic system (especially *SLC6A4* and genes coding for dopaminergic receptors), intracellular second messengers (especially *INPP1*, *GSK3 β* , and *CREB1*), and neurotrophin system (especially *BDNF*) (Rybakowski 2013).

Genome-wide association studies (GWAS) represent a fundamental turning point since they allow the genotyping of several hundreds of thousands polymorphisms throughout the whole genome, providing the opportunity to uncover the multiple variants with small effect size that are supposed to be involved. Two independent groups of researchers identified the chromosomal region 18q23 as connected with lithium response (Ewald et al. 1999; Turecki et al. 2001), but no specific genes were identified within this region. In the STEP-BD (Systematic Treatment Enhancement Program for Bipolar Disorder) study, the risk for recurrence among patients treated with lithium was associated with a region on chromosome 4q32 spanning the *GRIA2* gene, coding for the glutamate AMPA receptor (Perlis et al. 2009). Results from a GWAS on a small Italian sample indicated the *ACCNI* gene as a potential candidate for response to lithium (Squassina et al. 2011). The gene codes for a cation channel with high affinity for sodium that may play a role in neurotransmission. Thanks to the ConLiGen initiative (Schulze et al. 2010) a sample including more than 1200 patients characterized for response to lithium

was collected and first results suggested *SLC4A10* gene as the top finding (despite it did not reach the genome-wide significance threshold) (Schulze 2012). The gene codes for solute carrier family 4, sodium bicarbonate transporter, member 10, which belongs to a family of sodium-coupled bicarbonate transporters. *SLC4A10* is highly expressed in the hippocampus and cerebral cortex. Interestingly, it has been proposed as a susceptibility gene for recurrent major depression (Schosser et al. 2011). A recent study on Chinese Han bipolar patients reported two SNPs in high linkage disequilibrium in the *GADLI* gene as correlated with lithium efficacy with impressive p values ($10e-37$) (Chen et al. 2014). Nevertheless, the risk alleles found by this study are rare in persons of European ancestry and following evidence did not confirm this finding (Consortium on Lithium et al. 2014).

The pharmacogenetics of mood stabilizers different from lithium was much less studied. Some candidate gene studies exist and provided negative findings (Yun et al. 2008; Wang et al. 2013), while some preliminary results indicated that *DRD2/ANKK1* polymorphisms may be associated with dextromethorphan augmentation to valproate (Lee et al. 2012) and the -116C/G SNP in the *XBPI* gene may correlate with valproate response (Masui et al. 2006; Kim et al. 2009).

Previous studies mainly investigated short-term treatment response and lithium. Despite lithium is considered as a first choice mood stabilizer for the treatment of BD according to current guidelines, polypharmacy is actually the most frequent scenario in the clinical practice (Sachs et al. 2014). Thus, the identification of genetic markers of treatment outcome in a real clinical setting may provide helpful information to identify patients at higher risk of poor outcome who should receive additional clinical attention.

In addition to the investigation of individual genes and polymorphisms, the analysis of molecular pathways that harbor genes of interest represents a recent and exciting challenge for pharmacogenetics. Pathway analysis allows to integrate various "omics" data such as gene expression and GWAS, and several considerations support its application. Indeed, genes are not expected to work alone, but in a complex network of interactions; further, complex phenotypes are supposed to be caused by the dysregulation of multiple targets in connected pathways and/or different genes in the same pathways. Under a statistical point of view, pathway analysis can balance the heterogeneity of genome-wide data (e.g. due to population stratification or differential rates of genotyping error between the groups under analysis) and provide a focused set of genes for validation (e.g. by sequencing) (Fabbri et al. 2013). The identification of pathways associated with treatment response in BD can help to understand the biological mechanisms involved and potentially to develop new treatment strategies.

Given the aforementioned issues, the present study aimed to investigate the genetic predictors of long-term treatment efficacy (more than 6 months) in the STEP-BD genome-wide study, including all treatment regimens in order to reflect a real clinical practice scenario. Depressive and (hypo)manic recurrences were both separately and jointly considered. Both individual polymorphisms/genes and molecular pathways were analyzed.

2. Methods

2.1. Sample

Systematic treatment enhancement program for bipolar disorder (STEP-BD) is the largest clinical prospective trial including BD patients to date (Sachs et al. 2003). STEP-BD was a prospective study, designed to develop and expand knowledge on the management and treatment of BD and evaluate the longitudinal outcome of the disease. Included patients were required to meet DSM-IV criteria for bipolar I disorder, bipolar II disorder, cyclothymia, bipolar disorder not otherwise specified (NOS), or schizoaffective manic or bipolar subtypes. Patients received pharmacological interventions as clinically indicated by the principles of evidenced-based medicine in a naturalistic setting and visits occurred according to clinical demand. Further details on study design and sample are provided elsewhere (Sachs et al. 2003).

Genotyping was performed at the Center for Genotyping and Analysis of the Broad Institute and was performed using the Affymetrix GeneChip Human 500K Mapping Array Set (Sklar et al. 2008).

2.2. Phenotypes

Prospective clinical information at each visit was registered using the clinical monitoring form (CMF), that included the assessment of manic and depressive symptoms with clinical diagnosis of acute episodes of each polarity according to DSM-IV criteria. More in detail, recovery was defined as two or fewer syndromal features of a mood episode for at least 8 weeks, and recurrence was defined as meeting full DSM criteria for a mood episode on any single subsequent visit. The presence of subsyndromal mood symptoms during follow-up was not considered recurrence (Perlis et al. 2009). In order to provide information about the long-term outcome of treatment, mood episodes occurring within follow-up were considered after excluding patients with follow-up duration < 6 months, accordingly to previous studies (Post et al. 2012). Patients were included independently from any ongoing pharmacological treatment. Considering the naturalistic design of the study, a treatment was considered as ongoing when it was prescribed to the patient for more than 50% of the available visits.

Three phenotypes were considered: 1) the total number of episodes / total number of visits ratio; 2) the number of depressive episodes / total number of visits ratio; and 3) the occurrence of at least one hypomanic, manic or mixed episode during follow-up. Thus 1) and 2) are continuous variables (representing the rate of visits in which the patient was labeled as being in an acute phase) and 3) is a binary variable. For 1) and 2) the use of the ratio with total number of visits as denominator allowed to control for the wide range in the number of visits among patients. The total number of visits was chosen instead of the duration of follow-up because the time interval between visits extensively varied according to clinical needs. For 3) a binary variable was chosen because the continuous variable similar to 1) and 2) was flattened toward the unit (no or few episodes). Given that 3) did not take into account the number of visits, only for this phenotype patients with follow-up duration $>$ mean + 2SD were excluded.

2.3. Quality control

SNP pruning for quality control was based on call rates (genotype rate of at least 95% for each SNP and genotype availability in at least 90% of subjects), departure from Hardy-Weinberg equilibrium ($p < 0.001$), and minor allele frequency (MAF < 0.05). The result of quality control is shown by the QQ plot for each phenotype (Supplementary Figure 1). Lambda (genotypic inflation factor) values were 1.015, 1.018 and 0.999 for the phenotypes describing total, depressive and (hypo)manic episodes, respectively. A complete agglomerative clustering was applied, based on a multidimensional scaling of a matrix of pairwise identity-by-state (IBS) values between samples, as implemented in PLINK (Purcell et al. 2007). Ancestry clusters were defined on the base of the pairwise population concordance test (PCC < 0.0001 , according to (Purcell et al. 2007)). Identity-by-descent (IBD) analysis was used to identify related subjects (IBD > 0.1875 , (Anderson et al. 2010)).

2.4. Statistical analysis

The effect of individual markers on phenotypes was tested through linear or logistic regression models under the hypothesis of an additive genetic effect. Covariates were age, gender, ancestry (see paragraph 2.3) and monotherapy vs. polytherapy. Monotherapy was defined as being in monotherapy for more than 50% of considered visits. PLINK served for these analyses (pengu.mgh.harvard.edu/~purcell/plink/).

Genes containing SNPs with $p < 10e-4$ were analyzed through the GeneMANIA plugin (www.genemania.org/) in Cytoscape (www.cytoscape.org/) in order to retrieve possible enrichment in specific molecular pathways. The threshold of $p < 10e-4$ was chosen in order to allow the detection of relatively mild signals but

balancing the risk of false positives. A threshold of $p < 10e-05$ (that was applied in previous studies (Song and Lee 2013)) was considered excessively conservative since only about 70 SNPs and 40 genes could be included in the analysis if this threshold was applied. Benjamini–Hochberg FDR multiple testing correction was applied. When doing the enrichment analysis, only GO terms with at least 10 annotations in humans and no more than 300 were considered in order to control the size of the multiple testing correction (Zuberi et al. 2013). GO terms with FDR < 0.05 and containing at least 10 genes were further investigated through imputation and pathway analysis. In detail, genes pertaining to each identified GO pathway were extracted from the genome-wide database according to their physical positions (Genome Build 36.3). Secondly, the genes of interest were imputed using IMPUTE2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) and 1000 Genomes data (NCBI Build 36 (dbSNP b126)) as reference panel. Imputed SNPs were pruned according to linkage disequilibrium ($r^2 \geq 0.8$), and poor imputation quality ($\text{info} < 0.8$). Variations showing $p < 0.05$ and $p < 0.01$ in each pathway of interest were tested for a significant different distribution (Fisher exact test) compared to a random pathway. Each random pathway was matched with the index pathway in terms of number of SNPs within it and intragenic position of the SNPs but with random distribution within the genome. $10e04$ permutations were run.

For individual marker analysis a p threshold of $10e-07$ was considered according to the main standard reported in literature (Broer et al. 2013). In case of significant ($p < 10e-07$) or suggestive findings ($p < 10e-06$), results were validated through $10e04$ permutations. Concerning pathway analysis, Benjamini–Hochberg FDR multiple testing correction and permutation were used for the first and second step of the analysis, respectively.

3. Results

329806 SNPs were available after quality control. 723 patients satisfied inclusion criteria for phenotypes 1) and 2) and their clinical-demographic characteristics are reported in Supplementary Table 1. For phenotype 3) (i.e. (hypo)manic/mixed recurrence) 706 patients were included. Monotherapy was associated with a lower frequency of total acute phases compared to polytherapy ($p = 8.80e-05$), as well as with a lower frequency of depressive episodes ($p = 9.31e-05$). No related subjects were identified according to the IBD.

3.1. SNP-based analysis

The rate of episodes of each polarity was associated with rs6568686 in the *TRAF3IP2-ASI* gene, rs1359582 in the *RNLS* gene, rs10513249 that lays 36 Kbp

far from the *DFNB31* gene, rs10489167 in the *NFYC* gene and rs6993270 in the *DEPTOR* gene. All SNPs with $p < 10e-06$ were reported in Table 1. Among them a signal was found in the *SORCS2* gene (rs16840900). Permuted p values were all $\leq 9e-06$.

The rate of depressive episodes showed evidence of association with rs10513249 that is located 36 Kbp from the *DFNB31* gene and with a couple of intergenic SNPs (rs2190547 and rs41368245) at 17q24.3 which nearest gene is *KCNJ2* (about 700 Kbp far). Several other SNPs at 17q24.3 showed $p < 10e-06$. All the SNPs with $p < 10e-06$ were reported in Table 1 and they are located in (or near) the *DEPTOR*, *C6orf163*, *TRAF3IP2-ASI*, and *TSPAN5* genes. Permuted p values were all $\leq 7e-06$.

Manhattan plots representing the results obtained for phenotypes 1) and 2) are shown in Figure 1, while regional association plots are shown in Supplementary Figure 2.

No marker was associated with (hypo)manic/mixed recurrence during follow-up.

Table 1: results of linear regression testing genetic polymorphisms associated with: **A)** the total number of episodes during follow-up; and **B)** the number of depressive episodes during follow-up. Only SNPs with $p < 10e-06$ were shown. No SNP with $p < 10e-06$ was found for the phenotype describing manic recurrence. Results are referred to an additive test. For intergenic SNPs the nearest gene was reported. Perm. P = p value after 100,000 permutations.

A

Gene	SNP	Beta	Stat	P	Perm. P
TRAF3IP2-ASI	rs6568686	0.11	5.58	3.66e-08	2e-06
RNLS	rs1359582	0.09	5.35	1.28e-07	1e-06
DFNB31 (36Kbp)	rs10513249	0.18	5.09	4.75e-07	1.8e-05
NFYC	rs10489167	0.16	5.06	5.53e-07	1.3e-05
DEPTOR	rs6993270	0.14	4.96	9.24e-07	1.7e-05
RNLS	rs10736359	0.08	4.82	1.82e-06	1e-05
DEPTOR	rs6999724	0.13	4.69	3.44e-06	2.4e-05
DEPTOR	rs7818752	0.13	4.69	3.44e-06	2.4e-05
SORCS2	rs16840900	0.09	4.66	3.94e-06	2.4e-05
ETS1 (770Kbp)	rs7123770	0.15	4.54	7.01e-06	8.74e-05
NRXN1	rs10187465	-0.06	-4.46	9.99e-06	9e-06

B

Gene	SNP	Beta	Stat	P	Perm. P
DFNB31 (36Kbp)	rs10513249	0.18	5.83	9.35e-09	3e-06

KCNJ2 (753Kbp)	rs2190547	0.08	4.96	9.49e-07	4e-06
KCNJ2 (734Kbp)	rs41368245	0.08	4.95	9.86e-07	5e-06
KCNJ2 (749Kbp)	rs9900281	0.08	4.94	1.03e-06	4e-06
KCNJ2 (752Kbp)	rs17823861	0.08	4.92	1.14e-06	5e-06
KCNJ2 (745Kbp)	rs4570897	0.08	4.91	1.18e-06	5e-06
KCNJ2 (745Kbp)	rs9901138	0.08	4.90	1.23e-06	5e-06
DEPTOR	rs6993270	0.12	4.88	1.39e-06	3.5e-05
C6orf163 (14Kbp)	rs9362426	0.08	4.87	1.43e-06	6e-06
C6orf163 (4Kbp)	rs1203156	0.08	4.86	1.50e-06	6e-06
DEPTOR	rs6999724	0.12	4.85	1.61e-06	6e-06
DEPTOR	rs7818752	0.12	4.85	1.61e-06	3.3e-05
TRAF3IP2- AS1	rs6568686	0.08	4.79	2.08e-06	1.1e-05
ETS1 (770Kbp)	rs7123770	0.14	4.74	2.68e-06	7.07e-05
C6orf163 (13Kbp)	rs1884320	0.08	4.64	4.44e-06	2e-05
KCNJ2 (802Kbp)	rs8068956	0.07	4.47	9.27e-06	3e-05
TSPAN5 (89Kbp)	rs1528010	0.06	4.47	9.31e-06	2.5e-05

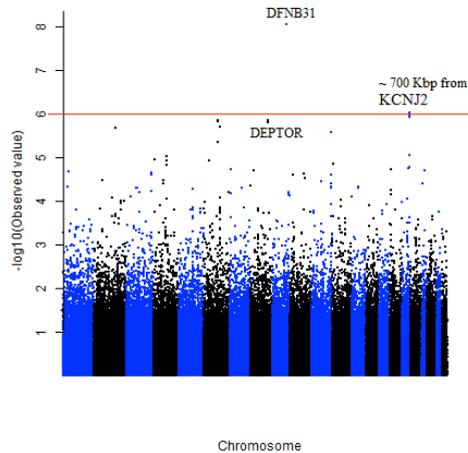
3.2. Pathway analysis

The molecular pathways that showed evidence of functional enrichment are summarized in Table 2.

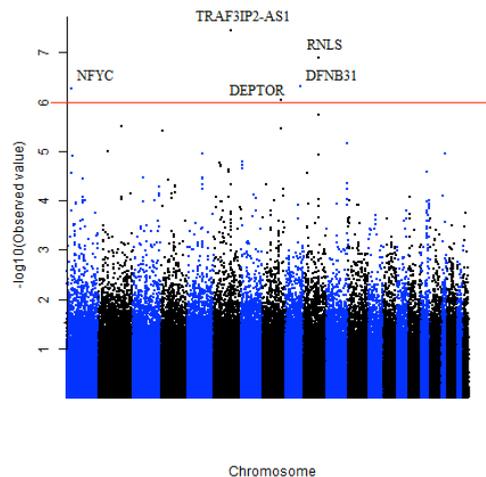
Regarding total episode rate, pathways with evidence of enrichment are involved in behavior, learning and memory, cognition and brain development (Table 2). Pathways including at least 9 genes were further studied as described in paragraph 2.4., but none of them reached the threshold for significance. The pathway GO:0007611 (involved in learning and memory processes) showed a trend of significance ($p=0.09$ and $p=0.07$, see Supplementary Table 2), with *RASGRF1* and *NTRK2* genes emerging as the top ones in terms of percent of associated SNPs. A number of pathways were found to be enriched for the phenotype describing the rate of depressive episodes.

Figure 1: manhattan plots showing the results obtained for the phenotypes total episode (A) and depressive (B) episode frequency during follow-up.

A



B



They pertain to the regulation of MAPK cascade, protein kinase activity, neuron projection, cell morphogenesis and chemotaxis, stem cell differentiation, cytoskeleton reorganization, cell migration, glutamate receptor activity (Table 2). The pathway GO:0043410 (positive regulation of MAPK cascade) showed nominal significance (nominal $p= 0.0006$) that was lost after permutation

(permuted $p=0.55$). The most interesting genes in terms of percent of associated SNPs were *GRM4*, *CHRNA7*, *LRRK2*, and *NTF3* (Supplementary Table 2).

No pathway was enriched when considering phenotype 3).

Table 2: results of pathway analysis for **A)** the total number of episodes during follow-up; and **B)** the number of depressive episodes during follow-up. No pathway showed $FDR < 0.05$ for the phenotype assessing (hypo)manic/mixed recurrence.

A

GO ID	Description	FDR	Genes in the sample
GO:0007610	Behavior	3.41e-08	ZIC1, USP46, TAS2R1, SHANK1, RASGRF1, PLK2, PLCB1, NTRK2, LRRK2, JAM3, GRIN2B, GRIN2A, FYN, F3, EFNB2, DLG4, CNTNAP2, CDH13, CAMK1D
GO:0007611	Learning memory or	1.42e-05	SHANK1, RASGRF1, PLK2, PLCB1, NTRK2, GRIN2B, GRIN2A, FYN, DLG4
GO:0050890	Cognition	9.96e-05	SHANK1, RASGRF1, PLK2, PLCB1, NTRK2, GRIN2B, GRIN2A, FYN, DLG4
GO:0021537	Telencephalon development	0.01	SALL1, PLCB1, NTRK2, LRRK2, LEF1, EPHB2, CNTNAP2
GO:0007420	Brain development	0.04	ZIC1, SALL1, PLCB1, PCDH18, NTRK2, LRRK2, LEF1, FOXA2, EPHB2, CTNNA2, CNTNAP2

B

GO ID	Description	FDR	Genes in the sample
GO:0043410	Positive regulation of MAPK cascade	2.28e-04	TGFB2, SDCBP, PDGFRA, NTF3, MAP2K6, LRRK2, KIT, GRM4, CHRNA7, ARRB1
GO:0007610	Behavior	6.59e-04	PLK2, PDGFRA, NTF3, LRRK2, FYN, F7, F3, CNTNAP2, CHRNA7
GO:0032147	Activation of protein kinase activity	0.003	TGFB2, NTF3, MAP3K1, MAP2K6, LRRK2, KIT, GRM4, CHRNA7
GO:0043005	Neuron projection	0.003	TGFB2, PLK2, MAGI2, LRRK2,

			KIF1B, GRM7, EPHA4, CNTNAP2
GO:0000187	Activation of MAPK activity	0.003	NTF3, MAP2K6, LRRK2, KIT, GRM4, CHRNA7
GO:0022604	Regulation of cell morphogenesis	0.003	TGFB2, PTPRD, LRRK2, LEF1, KIT, FOXA2, EPHA4, CAPZB
GO:0060326	Cell chemotaxis	0.003	TGFB2, PDGFRA, LEF1, KIT, F7, CCR6, CCL20
GO:0048863	Stem cell differentiation	0.004	TGFB3, TGFB2, PDGFRA, LEF1, KIT, FOXA2, CTR9
GO:0030425	Dendrite	0.004	PLK2, MAGI2, LRRK2, GRM7, EPHA4, CNTNAP2
GO:0032970	Regulation of actin filament-based process	0.004	PPFIA1, PDGFRA, PDE4D, NTF3, MTPN, KCNJ2, HAX1
GO:0097458	Neuron part	0.01	TGFB2, PLK2, MAGI2, LRRK2, KIF1B, GRM7, EPHA4, CNTNAP2
GO:0010769	Regulation of cell morphogenesis involved in differentiation	0.01	TGFB2, PTPRD, LRRK2, LEF1, FOXA2, EPHA4
GO:0014068	Positive regulation of phosphatidylinositol 3-kinase signaling	0.01	TGFB2, PDGFRA, KIT, HAX1
GO:0051241	Negative regulation of multicellular organismal process	0.01	TGFB2, PDGFRA, PDE4D, LRRK2, LEF1, CHRNA7, ARRB1
GO:0048864	Stem cell development	0.02	TGFB3, TGFB2, LEF1, KIT, CTR9
GO:0050927	Positive regulation of positive chemotaxis	0.02	NTF3, F7, F3
GO:0044708	Single-organism behavior	0.02	PLK2, LRRK2, FYN, CNTNAP2, CHRNA7
GO:0002092	Positive regulation of receptor internalization	0.02	NTF3, MAGI2, ARRB1
GO:0031623	Receptor internalization	0.02	NTF3, MAGI2, DNM2, ARRB1
GO:0016311	Dephosphorylation	0.03	TGFB2, PTPRG, PTPRD, PTPN5, MAGI2, DLG3, DLG2
GO:0071902	Positive regulation of protein serine/threonine kinase activity	0.03	NTF3, MAP2K6, LRRK2, KIT, GRM4, CHRNA7

GO:2000249	Regulation of actin cytoskeleton reorganization	0.03	PDGFRA, NTF3, HAX1
GO:0007611	Learning or memory	0.03	PLK2, FYN, CNTNAP2, CHRNA7
GO:0031532	Actin cytoskeleton reorganization	0.03	PDGFRA, NTF3, KIT, HAX1
GO:0008066	Glutamate receptor activity	0.03	GRM7, GRM4, GRM3
GO:0030335	Positive regulation of cell migration	0.03	TGFB2, PDGFRA, NTF3, LEF1, F7, F3
GO:0048008	Platelet-derived growth factor receptor signaling pathway	0.04	PDGFRA, F7, F3
GO:0042698	Ovulation cycle	0.04	TGFB2, PDGFRA, KIT
GO:0019199	Transmembrane receptor protein kinase activity	0.04	TGFB3, PDGFRA, KIT, EPHA4
GO:0051272	Positive regulation of cellular component movement	0.04	TGFB2, PDGFRA, NTF3, LEF1, F7, F3
GO:0048066	Developmental pigmentation	0.04	MITF, LEF1, KIT
GO:2000736	Regulation of stem cell differentiation	0.04	TGFB2, PDGFRA, LEF1, FOXA2,
GO:0050890	Cognition	0.04	PLK2, FYN, CNTNAP2, CHRNA7
GO:0001837	Epithelial to mesenchymal transition	0.04	TGFB3, TGFB2, LEF1, FOXA2
GO:0051896	Regulation of protein kinase B signaling	0.04	MAGI2, HAX1, F7, F3

4. Discussion

The present study investigated the genetic predictors of long-term (more than 6 months) treatment efficacy in BD patients from the STEP-BD genome-wide study. Total, depressive, and (hypo)manic recurrence were examined by polymorphism-, gene- and pathway-based approaches.

The rate of total episode recurrence was correlated with SNPs in the *TRAF3IP2-ASI*, *RNLS*, *DFNB31*, *NFYC*, and *DEPTOR* genes, and several of them appear promising on the basis of previous literature.

TRAF3IP2-ASI is an antisense non-coding RNA located in the region 6q21, it has no apparent open reading frame and most likely represents a structural RNA gene that is transcribed but not translated. The gene is probably involved in the regulation of *C6orf4* (a near gene) expression, and these two genes were cloned from a subregion at 6q21 containing a replicative senescence gene, a tumor suppressor gene and a gene involved in hereditary schizophrenia (Cao et al. 1997; Morelli et al. 2000). Very few knowledge about this gene is available for hypothesizing the mechanism(s) behind its possible involvement in treatment outcome of BD, but its location seems a promising clue.

RNLS (renalase, FAD-dependent amine) codes for a FAD-dependent amine oxidase that is involved in the degradation of catecholamines. High concentrations of catecholamines activate plasma renalase and promote its secretion and synthesis, that results in the reduction of blood pressure and heart rate, suggesting a role in cardiovascular risk. This gene has been correlated with the risk of type I diabetes by several independent studies (e.g. (Hisanaga-Oishi et al. 2014; Winkler et al. 2014)) but no previous evidence linked it to BD.

The *DFNB31* gene is thought to function in the organization and stabilization of stereocilia and actin elongation and cytoskeletal assembly. In neurons the coded protein is called whirlin, which binds to most members of the Usher protein complex. The Usher protein complex is an effector of β -catenin, and also affects neuronal morphogenesis and structural plasticity (Piazza et al. 1987). Some polymorphisms (rs942518, rs16929770, and rs10982256) within this gene were associated with BD susceptibility by a GWAS (Baum et al. 2008), a GWAS meta-analysis (Baum et al. 2008) and a confirmation was provided by an independent candidate gene study (Ollila et al. 2009). rs942518 and rs16929770 lay at only 31 Kbp and 29 Kbp, respectively, from the *DFNB31* marker outlined by the present study (rs10513249, see Table 1).

NFYC codes for one subunit of a trimeric complex forming a highly conserved transcription factor that binds with high specificity to CCAAT motifs in the promoters of a variety of genes. For example, the gene is involved in the control of cell cycle blocks and signaling pathways (Benatti et al. 2011), mineralocorticoid receptor expression (Murai-Takeda et al. 2010), lipid metabolism, and possibly neuronal plasticity as suggested in animal models (Milton et al. 2013).

DEPTOR (DEP domain containing MTOR-interacting protein) is part of the mammalian target of rapamycin complex 1 (mTORC1) which contains also MTOR, MLST8, RPTOR, and AKT1S1/PRAS40. Lithium augmentation to ketamine treatment was associated with the activation of the mTOR/BDNF signaling pathways in the rat prefrontal cortex (PFC) and restoration of dendritic

spine density in the medial PFC with concomitant evidence of antidepressant-like behavior (Chiu et al. 2014). Several molecules with antidepressant properties (ketamine (Abelaira et al. 2014), lurasidone (Luoni et al. 2014), GLYX-13 (Lu et al. 2014)) but also commonly prescribed antidepressants (escitalopram and paroxetine (Park et al. 2014)) were demonstrated to regulate the mTOR pathway. Besides the mTOR signaling pathway demonstrated an antidepressant response to various drugs, this seems to be more associated with antidepressant N-methyl-d-aspartate (NMDA) receptor antagonists, such as ketamine (Abelaira et al. 2014). Molecules mediating mTOR signaling thus represent promising candidates for deepening the knowledge of mechanisms of antidepressant action and identify targets for innovative drugs.

SORCS2 and *NRXN1* are among our top gene findings and despite they did not reach statistical significance they appear interesting candidates. *SORCS2* encodes one family member of vacuolar protein sorting 10 (VPS10) domain-containing receptor proteins. This have been implicated in BD risk by a GWAS (Baum et al. 2008) and two candidate gene studies (Ollila et al. 2009; Takata et al. 2011). *NRXN1* codes for neurexin-1-alpha and it functions as cell adhesion molecule and receptor. Molecular pathways involved in cell adhesion were found to contribute to the risk of schizophrenia and BD and the most replicated (by three GWAS) individual genes were *NRXN1* and *CNTNAP2* (O'Dushlaine et al. 2011). *CNTNAP2* is included in several of the enriched pathway in the present study (Table 2). Further, *NRXN1* polymorphisms were related with antidepressant response in major depression (Tansey et al. 2014) and antipsychotic response in schizophrenia (Lett et al. 2011; Jenkins et al. 2014).

The rate of depressive episodes during follow-up showed a stronger association with the *DFNB31* rs10513249 polymorphism compared to the total episode rate. Further, some markers ~ 700 Kbp from the *KCNJ2* gene showed association with this phenotype (Table 1). *KCNJ2* codes for a inward rectifier potassium channel which mutations have been implicated in the pathogenesis of the Andersen-Tawil syndrome (ATS), that includes major depression among its clinical manifestations (Chan et al. 2010). The SNPs near to this gene that have been identified by this study are located at 17q24.3, a region that has been associated with BD by linkage studies (Ewald et al. 2005; Tomas et al. 2006). Finally, recent evidence from GWAS implicates an entire class of genes related to the structure and regulation of ion channels in the etiology of BD, among which several genes coding for potassium channels (*KCNQ2*, *KCNQ3*, *KCNN3*, *KCNA4*, *KCNA1*, *KCNAB1*, *KCNS3*). Accumulating evidence suggests that existing treatments for BP may potentially exert their therapeutic action via ion channel regulation and potassium

channels are considered attractive targets for novel therapeutics for BD (Judy and Zandi 2013).

Pathway analysis showed enrichment signals in pathways mediating behavior, learning/memory, cognition, brain development, MAPK cascade activation, cell morphogenesis and chemotaxis, stem cell differentiation, cytoskeleton reorganization, cell migration, and glutamate receptor activity (Table 2). Some genes (*JAM3*, *CNTNAP2*, *PTPRD*) that are part of our enriched pathways are also members of the cell adhesion molecule pathway that has been previously associated with BD susceptibility (O'Dushlaine et al. 2011). None of the enriched pathway showed a different distribution of SNPs with $p < 0.05$ or $p < 0.01$ compared to a random pathway, even if positive regulation of MAPK cascade showed a nominal significant p value and learning/memory pathway had a p value approaching nominal significance (Supplementary Table 2). The most interesting candidate genes emerging from these pathways are *RASGRF1*, *NTRK2*, *GRIN2B*, *GRIN2A*, *GRM4*, *CHRNA7*, *LRRK2*, and *NTF3*.

RASGRF1 could mediate forms of synaptic plasticity and might participate in the regulation of neuronal excitability and neurite outgrowth (Zhu et al. 2013; Barman et al. 2014) by several signaling pathway among which GluN2B signaling to p38 MAPK (Kochlamazashvili et al. 2010). The ablation of *RASGRF1* was found to be able to restore impaired synaptic plasticity in brain slices showing deficit in neural cell adhesion molecules (Kochlamazashvili et al. 2010).

NTRK2 (neurotrophic tyrosine kinase receptor type 2) is a specific receptor for neurotrophin brain-derived neurotrophic factor (BDNF). The *BDNF/NTRK2* pathway might be involved in the pathogenesis of BD and the therapeutic mechanisms of lithium (Shaltiel et al. 2007). Genetic studies consistently suggested that the *NTRK2* gene may be associated with BD susceptibility (Smith et al. 2009) and lithium response (Bremer et al. 2007).

GRIN2B and *GRIN2A* encode for two subunits of the N-methyl-d-aspartate (NMDA) glutamate receptor. Abnormalities in glutamatergic function and signaling pathways through the NMDA receptors are involved in the pathophysiology of BD and mood stabilizers are supposed to exert their neuroprotective effects through reducing NMDA receptor-induced excitotoxicity (Hashimoto et al. 2002). Genetic and pharmacogenetic studies indicated that variations in *GRIN2B* may be implicated in BD risk (Fallin et al. 2005; Avramopoulos et al. 2007; Lorenzi et al. 2010) and relapse in BD (Dalvie et al. 2010). *GRIN2A* polymorphisms have been also associated with BD (Cherlyn et al. 2010), as well as *GRM4* (metabotropic glutamate receptor 4) (Fallin et al. 2005).

CHRNA7 (cholinergic receptor nicotinic alpha 7) is involved in cognitive function and synaptic plasticity. A dysfunction of *CHRNA7*-dependent signaling was reported in bipolar disorder (Thomsen et al. 2011) and variants of this gene were associated with several neuropsychiatric and cognitive phenotypes among which BD and sustained attention performance in patients with BD (Ancin et al. 2010; Ancin et al. 2011; Soler-Alfonso et al. 2014). Targeting cholinergic receptor nicotinic alpha 7 ($\alpha 7$ nAChR) is an attractive strategy for the development of new antidepressants. Indeed, $\alpha 7$ nAChR agonists show antidepressant-like activity in rats when administered alone or in combination with a sub-active dose of citalopram (Andreasen et al. 2012) and may be promising for treating cognitive symptoms (Pichat et al. 2007) and psychosis (Saur et al. 2013).

Variants in *LRRK2* (leucine-rich repeat kinase 2) were associated with Parkinson disease but also with premorbid mood disorders in affected patients (Shanker et al. 2011).

NTF3 (neurotrophin 3) controls survival and differentiation of mammalian neurons, it is closely related to both nerve growth factor and brain-derived neurotrophic factor. Polymorphisms in *NTF3* were hypothesized to have a modest effect on schizophrenia susceptibility (Hattori et al. 2002) but no effect was demonstrated on BD risk by one study (Tadokoro et al. 2004).

Some limitations of the present study should be considered. Firstly, no selection of patients was performed on the basis of treatment/treatment class, resulting in potential bias due to stratification. On the other hand, our approach is more representative of the real clinical practice and other GWAS followed the same principle (e.g. the MARS study (Ising et al. 2009)). Given our approach, the present findings could represent genetic markers of disease course/severity rather than markers of treatment response. Follow-up duration also lacked of standardization and widely varied among included patients (Supplementary Table 1); for this reason we constructed phenotypes considering the rate of visits labeled as acute episodes on the total number of visits. Anyway, this was not possible for the phenotype assessing (hypo)manic/mixed recurrences, thus outliers for follow-up duration were excluded in order to reduce the heterogeneity deriving from extreme values of this variable. Lambda (genotypic inflation factor) values for total and depressive episode recurrence showed a tendency toward inflation, despite similar lambda values were reported by previous GWAS (Gaudet et al. 2010; Drago et al. 2011). Sample size (723 patients) is relatively small compared to other recent GWAS, even if this difference is lower or inverted when considering only GWAS focused on treatment-response phenotypes (Perlis et al. 2009; Squassina et

al. 2011; Chen et al. 2014). No independent sample is currently available for replication in the present study, but multiple-testing corrections were applied.

5. Conclusion

The present study provides evidence that variants in genes previously involved in the susceptibility to BD (*DFNB31*, *SORCS2*, *NRXN1*, *CNTNAP2*, *GRIN2A*, *GRM4*, *GRIN2B*), antidepressant action (*DEPTOR*, *CHRNA7*, *NRXN1*) and mood stabilizer or antipsychotic action (*NTRK2*, *CHRNA7*, *NRXN1*) may affect long-term treatment outcome of BD under naturalistic treatment. Promising findings without previous strong evidence were *TRAF3IP2-AS1*, *NFYC*, *RNLS*, *KCNJ2*, *RASGRF1*, *NTF3* genes. Pathway analysis supported in particular the involvement of molecules mediating regulation of MAPK cascade and learning/memory processes.

Supplementary materials:

Supplementary materials are available for the online version of this publication.

References

- Abelaira HM, Reus GZ, Neotti MV and Quevedo J. The role of mTOR in depression and antidepressant responses. *Life Sci* 2014; 101: 10-14.
- Ancin I, Barabash A, Vazquez-Alvarez B, Santos JL, Sanchez-Morla E, Martinez JL, et al. Evidence for association of the non-duplicated region of *CHRNA7* gene with bipolar disorder but not with Schizophrenia. *Psychiatr Genet* 2010; 20: 289-297.
- Ancin I, Cabranes JA, Santos JL, Sanchez-Morla E, Vazquez-Alvarez B, Rodriguez-Moya L, et al. *CHRNA7* haplotypes are associated with impaired attention in euthymic bipolar disorder. *J Affect Disord* 2011; 133: 340-345.
- Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP and Zondervan KT. Data quality control in genetic case-control association studies. *Nat Protoc* 2010; 5: 1564-1573.
- Andreasen JT, Redrobe JP and Nielsen EO. Combined alpha7 nicotinic acetylcholine receptor agonism and partial serotonin transporter inhibition produce antidepressant-like effects in the mouse forced swim and tail suspension tests: a comparison of SSR180711 and PNU-282987. *Pharmacol Biochem Behav* 2012; 100: 624-629.
- Avramopoulos D, Lasseter VK, Fallin MD, Wolynec PS, McGrath JA, Nestadt G, et al. Stage II follow-up on a linkage scan for bipolar disorder in the Ashkenazim provides suggestive evidence for chromosome 12p and the *GRIN2B* gene. *Genet Med* 2007; 9: 745-751.
- Barman A, Assmann A, Richter S, Soch J, Schutze H, Wustenberg T, et al. Genetic variation of the *RASGRF1* regulatory region affects human hippocampus-dependent memory. *Front Hum Neurosci* 2014; 8: 260.
- Baum AE, Akula N, Cabanero M, Cardona I, Corona W, Klemens B, et al. A genome-wide association study implicates diacylglycerol kinase eta (*DGKH*) and several other genes in the etiology of bipolar disorder. *Mol Psychiatry* 2008; 13: 197-207.

- Baum AE, Hamshere M, Green E, Cichon S, Rietschel M, Noethen MM, et al. Meta-analysis of two genome-wide association studies of bipolar disorder reveals important points of agreement. *Mol Psychiatry* 2008; 13: 466-467.
- Benatti P, Dolfini D, Vigano A, Ravo M, Weisz A and Imbriano C. Specific inhibition of NF-Y subunits triggers different cell proliferation defects. *Nucleic Acids Res* 2011; 39: 5356-5368.
- Bremer T, Diamond C, McKinney R, Shehktman T, Barrett TB, Herold C, et al. The pharmacogenetics of lithium response depends upon clinical co-morbidity. *Mol Diagn Ther* 2007; 11: 161-170.
- Broer L, Lill CM, Schuur M, Amin N, Roehr JT, Bertram L, et al. Distinguishing true from false positives in genomic studies: p values. *Eur J Epidemiol* 2013; 28: 131-138.
- Cacabelos R, Martinez-Bouza R, Carril JC, Fernandez-Novoa L, Lombardi V, Carrera I, et al. Genomics and pharmacogenomics of brain disorders. *Curr Pharm Biotechnol* 2012; 13: 674-725.
- Cao Q, Martinez M, Zhang J, Sanders AR, Badner JA, Cravchik A, et al. Suggestive evidence for a schizophrenia susceptibility locus on chromosome 6q and a confirmation in an independent series of pedigrees. *Genomics* 1997; 43: 1-8.
- Chan HF, Chen ML, Su JJ, Ko LC, Lin CH and Wu RM. A novel neuropsychiatric phenotype of KCNJ2 mutation in one Taiwanese family with Andersen-Tawil syndrome. *J Hum Genet* 2010; 55: 186-188.
- Chen CH, Lee CS, Lee MT, Ouyang WC, Chen CC, Chong MY, et al. Variant GADL1 and response to lithium therapy in bipolar I disorder. *N Engl J Med* 2014; 370: 119-128.
- Cherlyn SY, Woon PS, Liu JJ, Ong WY, Tsai GC and Sim K. Genetic association studies of glutamate, GABA and related genes in schizophrenia and bipolar disorder: a decade of advance. *Neurosci Biobehav Rev* 2010; 34: 958-977.
- Chiu C, Scheuing L, Liu G, Liao H, Linares GR, Lin D, et al. The Mood Stabilizer Lithium Potentiates the Antidepressant-like Effects and Ameliorates Oxidative Stress Induced by Acute Ketamine in a Mouse Model of Stress. *Int J Neuropsychopharmacol* 2014; 18(6).
- Consortium on Lithium G, Hou L, Heilbronner U, Rietschel M, Kato T, Kuo PH, et al. Variant GADL1 and response to lithium in bipolar I disorder. *N Engl J Med* 2014; 370: 1857-1859.
- Dalvie S, Horn N, Nossek C, van der Merwe L, Stein DJ and Ramesar R. Psychosis and relapse in bipolar disorder are related to GRM3, DAOA, and GRIN2B genotype. *Afr J Psychiatry (Johannesbg)* 2010; 13: 297-301.
- Drago A, Crisafulli C and Serretti A. The genetics of antipsychotic induced tremors: a genome-wide pathway analysis on the STEP-BD SCP sample. *Am J Med Genet B Neuropsychiatr Genet* 2011; 156B: 975-986.
- Ewald H, Wang AG, Vang M, Mors O, Nyegaard M and Kruse TA. A haplotype-based study of lithium responding patients with bipolar affective disorder on the Faroe Islands. *Psychiatr Genet* 1999; 9: 23-34.
- Ewald H, Wikman FP, Teruel BM, Buttenschon HN, Torralba M, Als TD, et al. A genome-wide search for risk genes using homozygosity mapping and microarrays with 1,494 single-nucleotide polymorphisms in 22 eastern Cuban families with bipolar disorder. *Am J Med Genet B Neuropsychiatr Genet* 2005; 133B: 25-30.

- Fabbri C, Di Girolamo G and Serretti A. Pharmacogenetics of antidepressant drugs: an update after almost 20 years of research. *Am J Med Genet B Neuropsychiatr Genet* 2013; 162B: 487-520.
- Fallin MD, Lasseter VK, Avramopoulos D, Nicodemus KK, Wolyniec PS, McGrath JA, et al. Bipolar I disorder and schizophrenia: a 440-single-nucleotide polymorphism screen of 64 candidate genes among Ashkenazi Jewish case-parent trios. *Am J Hum Genet* 2005; 77: 918-936.
- Gaudet MM, Kirchoff T, Green T, Vijai J, Korn JM, Guiducci C, et al. Common genetic variants and modification of penetrance of BRCA2-associated breast cancer. *PLoS Genet* 2010; 6: e1001183.
- Hashimoto R, Hough C, Nakazawa T, Yamamoto T and Chuang DM. Lithium protection against glutamate excitotoxicity in rat cerebral cortical neurons: involvement of NMDA receptor inhibition possibly by decreasing NR2B tyrosine phosphorylation. *J Neurochem* 2002; 80: 589-597.
- Hattori M, Kunugi H, Akahane A, Tanaka H, Ishida S, Hirose T, et al. Novel polymorphisms in the promoter region of the neurotrophin-3 gene and their associations with schizophrenia. *Am J Med Genet* 2002; 114: 304-309.
- Hisanaga-Oishi Y, Nishiwaki-Ueda Y, Nojima K and Ueda H. Analysis of the expression of candidate genes for type 1 diabetes susceptibility in T cells. *Endocr J* 2014; 61: 577-588.
- Ising M, Lucae S, Binder EB, Bettecken T, Uhr M, Ripke S, et al. A genomewide association study points to multiple loci that predict antidepressant drug treatment outcome in depression. *Arch Gen Psychiatry* 2009; 66: 966-975.
- Jenkins A, Apud JA, Zhang F, Decot H, Weinberger DR and Law AJ. Identification of candidate single-nucleotide polymorphisms in NRXN1 related to antipsychotic treatment response in patients with schizophrenia. *Neuropsychopharmacology* 2014; 39: 2170-2178.
- Judy JT and Zandi PP. A review of potassium channels in bipolar disorder. *Front Genet* 2013; 4: 105.
- Kim B, Kim CY, Lee MJ and Joo YH. Preliminary evidence on the association between XBP1-116C/G polymorphism and response to prophylactic treatment with valproate in bipolar disorders. *Psychiatry Res* 2009; 168: 209-212.
- Kochlamazashvili G, Senkov O, Grebenyuk S, Robinson C, Xiao MF, Stummeyer K, et al. Neural cell adhesion molecule-associated polysialic acid regulates synaptic plasticity and learning by restraining the signaling through GluN2B-containing NMDA receptors. *J Neurosci* 2010; 30: 4171-4183.
- Leclerc E, Mansur RB and Brietzke E. Determinants of adherence to treatment in bipolar disorder: a comprehensive review. *J Affect Disord* 2013; 149: 247-252.
- Lee SY, Chen SL, Chang YH, Chen SH, Chu CH, Huang SY, et al. The DRD2/ANKK1 gene is associated with response to add-on dextromethorphan treatment in bipolar disorder. *J Affect Disord* 2012; 138: 295-300.
- Lett TA, Tiwari AK, Meltzer HY, Lieberman JA, Potkin SG, Voineskos AN, et al. The putative functional rs1045881 marker of neurexin-1 in schizophrenia and clozapine response. *Schizophr Res* 2011; 132: 121-124.
- Lorenzi C, Delmonte D, Pirovano A, Marino E, Bongiorno F, Catalano M, et al. Searching susceptibility loci for bipolar disorder: a sib pair study on chromosome 12. *Neuropsychobiology* 2010; 61: 10-18.

- Lu Y, Wang C, Xue Z, Li C, Zhang J, Zhao X, et al. PI3K/AKT/mTOR signaling-mediated neuropeptide VGF in the hippocampus of mice involve in the rapid onset antidepressant-like effects of GLYX-13. *Int J Neuropsychopharmacol* 2014; 18(5).
- Luoni A, Macchi F, Papp M, Molteni R and Riva MA. Lurasidone exerts antidepressant properties in the chronic mild stress model through the regulation of synaptic and neuroplastic mechanisms in the rat prefrontal cortex. *Int J Neuropsychopharmacol* 2014; 18(4).
- Masui T, Hashimoto R, Kusumi I, Suzuki K, Tanaka T, Nakagawa S, et al. A possible association between the -116C/G single nucleotide polymorphism of the XBP1 gene and lithium prophylaxis in bipolar disorder. *Int J Neuropsychopharmacol* 2006; 9: 83-88.
- McGuffin P, Rijsdijk F, Andrew M, Sham P, Katz R and Cardno A. The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Arch Gen Psychiatry* 2003; 60: 497-502.
- Milton AC, Packard AV, Clary L and Okkema PG. The NF-Y complex negatively regulates *Caenorhabditis elegans* *tbx-2* expression. *Dev Biol* 2013; 382: 38-47.
- Morelli C, Magnanini C, Mungall AJ, Negrini M and Barbanti-Brodano G. Cloning and characterization of two overlapping genes in a subregion at 6q21 involved in replicative senescence and schizophrenia. *Gene* 2000; 252: 217-225.
- Murai-Takeda A, Shibata H, Kurihara I, Kobayashi S, Yokota K, Suda N, et al. NF- κ B functions as a corepressor of agonist-bound mineralocorticoid receptor. *J Biol Chem* 2010; 285: 8084-8093.
- O'Dushlaine C, Kenny E, Heron E, Donohoe G, Gill M, Morris D, et al. Molecular pathways involved in neuronal cell adhesion and membrane scaffolding contribute to schizophrenia and bipolar disorder susceptibility. *Mol Psychiatry* 2011; 16: 286-292.
- Ollila HM, Soronen P, Silander K, Palo OM, Kieseppa T, Kaunisto MA, et al. Findings from bipolar disorder genome-wide association studies replicate in a Finnish bipolar family-cohort. *Mol Psychiatry* 2009; 14: 351-353.
- Park SW, Lee JG, Seo MK, Lee CH, Cho HY, Lee BJ, et al. Differential effects of antidepressant drugs on mTOR signalling in rat hippocampal neurons. *Int J Neuropsychopharmacol* 2014; 17: 1831-1846.
- Perlis RH, Smoller JW, Ferreira MA, McQuillin A, Bass N, Lawrence J, et al. A genomewide association study of response to lithium for prevention of recurrence in bipolar disorder. *Am J Psychiatry* 2009; 166: 718-725.
- Piazza L, Fishman GA, Kaplan RD, Horowitz AL, Hinds WA and Mafee MF. Magnetic resonance imaging of central nervous system defects in Usher's syndrome. *Retina* 1987; 7: 241-245.
- Pichat P, Bergis OE, Terranova JP, Urani A, Duarte C, Santucci V, et al. SSR180711, a novel selective $\alpha 7$ nicotinic receptor partial agonist: (II) efficacy in experimental models predictive of activity against cognitive symptoms of schizophrenia. *Neuropsychopharmacology* 2007; 32: 17-34.
- Post RM, Leverich GS, Altshuler LL, Frye MA, Suppes T, McElroy SL, et al. Relationship of prior antidepressant exposure to long-term prospective outcome in bipolar I disorder outpatients. *J Clin Psychiatry* 2012; 73: 924-930.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; 81: 559-575.

- Rybakowski JK. Genetic influences on response to mood stabilizers in bipolar disorder: current status of knowledge. *CNS Drugs* 2013; 27: 165-173.
- Sachs GS, Peters AT, Sylvia L and Grunze H. Polypharmacy and bipolar disorder: what's personality got to do with it? *Int J Neuropsychopharmacol* 2014; 17: 1053-1061.
- Sachs GS, Thase ME, Otto MW, Bauer M, Miklowitz D, Wisniewski SR, et al. Rationale, design, and methods of the systematic treatment enhancement program for bipolar disorder (STEP-BD). *Biol Psychiatry* 2003; 53: 1028-1042.
- Saur T, DeMarco SE, Ortiz A, Sliwoski GR, Hao L, Wang X, et al. A genome-wide RNAi screen in *Caenorhabditis elegans* identifies the nicotinic acetylcholine receptor subunit ACR-7 as an antipsychotic drug target. *PLoS Genet* 2013; 9: e1003313.
- Schosser A, Gaysina D, Cohen-Woods S, Domenici E, Perry J, Tozzi F, et al. A follow-up case-control association study of tractable (druggable) genes in recurrent major depression. *Am J Med Genet B Neuropsychiatr Genet* 2011; 156B: 640-650.
- Schulze T (2012). The Consortium on Lithium Genetics (ConLiGen) genome-wide association studies of lithium response phenotypes in bipolar disorder. *CINP Congress*. Stockholm: 36.
- Schulze TG, Alda M, Adli M, Akula N, Ardu R, Bui ET, et al. The International Consortium on Lithium Genetics (ConLiGen): an initiative by the NIMH and IGSLI to study the genetic basis of response to lithium treatment. *Neuropsychobiology* 2010; 62: 72-78.
- Shaltiel G, Chen G and Manji HK. Neurotrophic signaling cascades in the pathophysiology and treatment of bipolar disorder. *Curr Opin Pharmacol* 2007; 7: 22-26.
- Shanker V, Groves M, Heiman G, Palmese C, Saunders-Pullman R, Ozelius L, et al. Mood and cognition in leucine-rich repeat kinase 2 G2019S Parkinson's disease. *Mov Disord* 2011; 26: 1875-1880.
- Sklar P, Smoller JW, Fan J, Ferreira MA, Perlis RH, Chambert K, et al. Whole-genome association study of bipolar disorder. *Mol Psychiatry* 2008; 13: 558-569.
- Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W, et al. Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry* 2009; 14: 755-763.
- Soler-Alfonso C, Carvalho CM, Ge J, Roney EK, Bader PI, Kolodziejaska KE, et al. CHRNA7 triplication associated with cognitive impairment and neuropsychiatric phenotypes in a three-generation pedigree. *Eur J Hum Genet* 2014; 22: 1071-1076.
- Song GG and Lee YH. Pathway analysis of genome-wide association studies for Parkinson's disease. *Mol Biol Rep* 2013; 40: 2599-2607.
- Squassina A, Manchia M, Borg J, Congiu D, Costa M, Georgitsi M, et al. Evidence for association of an ACCN1 gene variant with response to lithium treatment in Sardinian patients with bipolar disorder. *Pharmacogenomics* 2011; 12: 1559-1569.
- Tadokoro K, Hashimoto R, Tatsumi M, Kamijima K and Kunugi H. Analysis of enhancer activity of a dinucleotide repeat polymorphism in the neurotrophin-3 gene and its association with bipolar disorder. *Neuropsychobiology* 2004; 50: 206-210.
- Takata A, Kawasaki H, Iwayama Y, Yamada K, Gotoh L, Mitsuyasu H, et al. Nominal association between a polymorphism in DGKH and bipolar disorder detected in a meta-analysis of East Asian case-control samples. *Psychiatry Clin Neurosci* 2011; 65: 280-285.
- Tansey KE, Rucker JJ, Kavanagh DH, Guipponi M, Perroud N, Bondolfi G, et al. Copy number variants and therapeutic response to antidepressant medication in major depressive disorder. *Pharmacogenomics J* 2014; 14: 395-399.

- Thomsen MS, Weyn A and Mikkelsen JD. Hippocampal alpha7 nicotinic acetylcholine receptor levels in patients with schizophrenia, bipolar disorder, or major depressive disorder. *Bipolar Disord* 2011; 13: 701-707.
- Tomas C, Canellas F, Rodriguez V, Picornell A, Lafau O, Nadal M, et al. Genetic linkage study for bipolar disorders on chromosomes 17 and 18 in families with a high expression of mental illness from the Balearic Islands. *Psychiatr Genet* 2006; 16: 145-151.
- Turecki G, Grof P, Grof E, D'Souza V, Lebus L, Marineau C, et al. Mapping susceptibility genes for bipolar disorder: a pharmacogenetic approach based on excellent response to lithium. *Mol Psychiatry* 2001; 6: 570-578.
- Wang Z, Fan J, Gao K, Li Z, Yi Z, Wang L, et al. Neurotrophic tyrosine kinase receptor type 2 (NTRK2) gene associated with treatment response to mood stabilizers in patients with bipolar I disorder. *J Mol Neurosci* 2013; 50: 305-310.
- Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet* 2013; 382: 1575-1586.
- Winkler C, Krumsiek J, Buettner F, Angermuller C, Giannopoulou EZ, Theis FJ, et al. Feature ranking of type 1 diabetes susceptibility genes improves prediction of type 1 diabetes. *Diabetologia* 2014; 57: 2521-2529.
- Yun DH, Pae CU, Drago A, Mandelli L, De Ronchi D, Patkar AA, et al. Effect of the dysbindin gene on antimanic agents in patients with bipolar I disorder. *Psychiatry Investig* 2008; 5: 102-105.
- Zhu Q, Wang L, Xiao Z, Xiao F, Luo J, Zhang X, et al. Decreased expression of Ras-GRF1 in the brain tissue of the intractable epilepsy patients and experimental rats. *Brain Res* 2013; 1493: 99-109.
- Zuberi K, Franz M, Rodriguez H, Montojo J, Lopes CT, Bader GD, et al. GeneMANIA prediction server 2013 update. *Nucleic Acids Res* 2013; 41: W115-122.

Chapter 8

Summary

The aim of this dissertation was to contribute to the clarification of the genetic component of antidepressant response in order to facilitate the development of clinical applications. The included studies aimed to contribute to this aim using the following strategies:

1) The study of strong candidate genes for involvement in antidepressant metabolism and antidepressant mechanisms of action through innovative approaches. The first meta-analysis of CYP2C19 (cytochrome P450 family 2 subfamily C member 19) metabolizing activity in antidepressant treatment outcomes was performed (*chapter 2*). A candidate gene approach was integrated with the analysis of genome-wide data to identify the biological pathways that may modulate the risk of treatment-resistant depression (*chapter 3*).

2) The improvement of the exploitation of existing genome-wide data. More specifically, the study of common and rare variants was combined, genetic variation covering was maximized thanks to the use of high-density imputation, and for the first time a pathway meta-analysis was performed in order to identify common biological mechanisms that may be involved across independent samples (*chapter 4*).

3) The use of multi-marker tests to improve the prediction of treatment efficacy. Pathway analysis was carried out to identify the signaling cascades involved in antidepressant action (*chapter 5* and *chapter 6*) and a genome-wide study of treatment outcome in bipolar patients was used to select genes to be included in a gene-set enrichment analysis (*chapter 7*).

Candidate genes involved in antidepressant pharmacokinetics and pharmacodynamics: Part 1

CYP2C19 is the main isoenzyme responsible for citalopram and escitalopram metabolism. The first study included in this dissertation used genotyped or imputed data to identify CYP2C19 metabolizing groups on the base of the two most common functional polymorphisms in this gene (rs4244285 and rs12248560) and it investigated their association with citalopram/escitalopram clinical outcomes in four pharmacogenomic samples using a meta-analysis. Poor metabolizers (PMs) had higher symptom improvement and remission rates but in the first weeks of treatment they showed higher risk of several side effects compared to extensive metabolizers (EMs). This study suggested that CYP2C19 polymorphisms may provide helpful information for guiding citalopram/escitalopram treatment (*chapter 2*). In *chapter 3* genes that were previously associated with several

psychiatric traits by genome-wide association studies and which are involved in biological processes highly linked to antidepressant action were studied in four independent samples. Association tests were performed at SNP-level but also at pathway-level using genome-wide data. This was the first study to apply pathway analysis to the investigation of resistance to antidepressants and to use genetic variants from a candidate pathway to predict resistance using machine learning. At SNP-level, variants in the stress-related FKBP5 gene were the most consistently replicated for association with antidepressant response and resistance across the investigated samples. The best pathway associated with antidepressant resistance included the CACNA1C gene (GO:0006942). Using neural networks and gradient boosted machine independent SNPs in this pathway predicted resistance with a mean sensitivity of 0.83 and specificity of 0.56 after 10-fold cross validation repeated 100 times. This study suggested that FKBP5 polymorphisms should be considered for clinical applications, while pathways including CACNA1C may be helpful to develop multi-marker predictors of resistance to antidepressants.

Genome-wide association studies of antidepressant response: Part 2

In *chapter 4* we addressed some of the limitations that may have interfered with the success of previous genome-wide associations studies in the identification of polymorphisms associated with antidepressant response. These limitations included limited coverage of genetic variants and limited use of pathway analysis that this study tried to address by exome genotyping, dense imputation and use of pathway meta-analysis to investigate common pathways across independent samples. The study identified some new SNPs that may be associated with depressive symptom improvement in the ITGA9 (integrin alpha 9) and NRXN3 (neurexin 3) genes. ITGA9 codes for a membrane receptor for neurotrophins and NRXN3 is a transmembrane neuronal adhesion receptor involved in synaptic differentiation. Some significant pathways emerged from our meta-analysis (GO:0005694 (chromosome) and GO:0044427 (chromosomal part)), suggesting that variants in molecules mediating neurogenesis may have particular relevance.

Multi-marker tests as innovative approach to antidepressant pharmacogenomics: part 3

Chapter 5 and *chapter 6* aimed to apply pathway analysis to identify the mechanisms that may mediate the contribution of candidate genes to antidepressant action. Variants in the PPP3CC (protein phosphatase 3 catalytic subunit gamma) gene were associated with antidepressant efficacy across different samples and pathway analysis suggested that its effect on the outcome may be mediated by the

B-cell receptor signaling pathway, consistently with several inflammatory abnormalities involving B cells previously reported (*chapter 5*). GAP43 gene codes for a neuron-specific cytosolic protein expressed in neuronal growth cones. GAP43 pathway is involved in neural and synaptic plasticity and it was associated with response and remission to antidepressants (*chapter 6*). In *chapter 7* a GWAS of treatment outcome in bipolar patients was carried out and associated genes were investigated for functional interaction (enrichment) within known biological processes. This exploratory approach aimed to identify the possible functional link(s) among the genes of interest and consequently identify the corresponding biological/molecular mechanisms involved. The risk of depressive recurrence was associated with several genetic networks mediating neurogenesis and neuroplasticity, such as MAPK cascade, neuron projection, cell morphogenesis and chemotaxis, stem cell differentiation, cytoskeleton reorganization, cell migration, glutamate receptor activity.

Chapter 9

General Discussion

The aim of this dissertation was to contribute to the identification of genetic markers of antidepressant treatment outcomes in major depressive disorder (MDD). This disease produces an enormous burden on a global scale, being the third cause of years lived with disability (GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, 2016). This huge personal and social burden is partly due to the fact that valid biomarkers to guide treatment choice are still lacking. As outlined in *chapter 1*, there is convincing evidence that genetic variants contribute to a significant percent of variance in antidepressant response (Tansey et al., 2013), but the identification of the polymorphisms explaining this proportion of variance has encountered a number of obstacles. Some strategies for trying to address them were described and implemented in this dissertation.

Treatment response to antidepressant medications and other psychotropic medications is known to be different between MDD and bipolar disorder (BD), in terms of clinical effects and underlying biological mechanisms (Citrome, 2014). Anyway, a similar polygenic architecture of treatment efficacy has been hypothesized in BD (Fabbri and Serretti, 2016), thus the same issues and the same possible strategies to deal with them can be assumed.

An important consideration is that neither a candidate gene approach nor a classical genome-wide association study (GWAS) was successful in the identification of consistently replicated variants of treatment outcome in MDD and BD. Despite some tentative clinical applications have been produced (e.g. GeneSight and CNSDose tests (Peterson et al., 2017)), their use is currently not recommended by any clinical guideline because they have unclear cost/benefit ratio. Some of the main issues behind these unsatisfying results are the complex polygenic nature of treatment response, the small effect size of each single polymorphism involved and the heterogeneity of MDD. The latter is suggested by GWAS demonstrating a high cross-sample correlation in the polygenic component associated with schizophrenia and bipolar disorder, but low cross-sample correlation (around 50%) in MDD suggesting higher genetic heterogeneity in this disorder (Gratten et al., 2014). This dissertation mainly aimed to implement and test strategies to dissect the polygenic component of treatment response in mood disorders and increase power of detecting positive associations, since this was considered more feasible in the available samples than the study of the sources of genetic heterogeneity. The latter indeed would require samples of larger size in order to provide adequate power for the study of more homogeneous subpopulations.

Lack of power is one of the main obstacles in the implementation of successful GWAS of antidepressant response. For example a sample size ~ 2000 subjects provides adequate power to identify individual polymorphisms associated with a binary trait with heritability ~ 40% (Visscher et al., 2014), but all the GWAS that investigated antidepressant response had smaller sample sizes, except one that identified one genome-wide significant locus (Li et al., 2016). Similar issues affected GWAS investigating treatment response in BD (Song et al., 2016; Hou et al., 2016; Squassina et al., 2011); further, the last group of studies were focused on lithium efficacy, while response to other psychotropic drugs was largely overlooked. GWAS investigating other psychiatric traits showed an increase in the number of genome-wide significant hits as the sample size increased, the most notable example is the Psychiatric Genomics Consortium (PGC) Schizophrenia sample (Ripke et al., 2014). As we discussed in the Introduction paragraph, aggregated or multi-marker tests are a strategy to improve power and it seems particularly suited to the study of response to psychotropic drugs since the known polygenic nature of the trait and the fact that almost all the available samples are probably underpowered to detect associations with single polymorphisms. Pathway analysis and gene set enrichment were the aggregated approaches applied in order to investigate the polygenic component of treatment efficacy.

As reported previously, the small effect size of polymorphisms associated with treatment response makes particularly difficult to identify them in underpowered GWAS, because a strict multiple-testing correction is needed. An interesting strategy to address this issue is to start from some candidate genes which association with treatment effect is suggested by their biological role and previous genetic studies, then test which sets of genes interacting with the selected ones may mediate the phenotype of interest and focus the analysis on variants in the identified gene sets. Polygenic risk scores calculated for the variants in a gene set of interest are a possible strategy to study the cumulative effect of a gene set on treatment outcome.

Finally, the improvement in the covering of genetic variability is another issue addressed in the present dissertation. Poor covering of known genetic variants was provided in previous GWAS of antidepressant response, particularly in terms of rare variants (GENDEP Investigators et al., 2013), resulting in the risk of overlooking relevant polymorphisms.

Start from strong candidate genes and move forward: how previous knowledge can guide targeted analysis in GWAS

No *a priori* hypothesis is needed in GWAS and this can be interpreted as a considerable advantage for the investigation of treatment response in mood disorders, since the genes involved and the biological mechanisms mediating response are only partially known. The results of previous GWAS showed that this advantage could turn in a disadvantage. As previously discussed, the broad application of the genome-wide multiple-testing correction prevents the identification of genuine but small effects in underpowered sample sizes. The restriction of the focus on candidate genes supported by the existing literature can allow more fruitful exploitations of existing genome-wide data. In **chapter 2** we demonstrated that genome-wide data could be used to extract functional polymorphisms needed to determine CYP2C19 (cytochrome P450 family 2 subfamily C member 19) metabolizing activity and successfully study its association with antidepressant efficacy and side effects. CYP2C19 has been demonstrated to affect the metabolism of several antidepressants and particularly citalopram and escitalopram. Despite CYP2C19 metabolizing activity has a quite consolidated association with citalopram/escitalopram pharmacokinetics (e.g. drug plasma levels), contradictory findings were reported in terms of association with response and side effects (Porcelli et al., 2011). If plasma concentration of the drug reflects CYP2C19 metabolizing activity but no clear association could be found for clinical outcomes, this suggests a missing piece of knowledge: the proportion of the drug reaching the brain (or drug concentration in the brain) is unknown. It is reasonable to hypothesize a correlation between plasma and brain drug concentration (Paulzen et al., 2015), but transport through the blood-brain barrier via the P-glycoprotein (O'Brien et al., 2013) and possible central metabolism (CYP2C19 gene is expressed also in the brain (Jukić et al., 2016)) are expected to be responsible for different brain drug concentration. Thus, the effect of CYP2C19 metabolizing level on drug concentration in the brain is expected to be decreased of an unknown amount compared to the effect on plasma drug concentration. Previous studies investigating the association between CYP2C19 metabolizing activity and antidepressant outcomes were underpowered to detect this effect, since the group with most reduced enzyme activity (i.e. poor metabolizers or PMs) is observed only in ~2% of the Caucasian population. Indeed the largest sample studied for determining this particular association was almost the only one to report a positive finding (Mrazek et al., 2011). Thus, the use of a meta-analysis was an optimal approach to maximize power and clarify the effect of CYP2C19 metabolizing level on the clinical outcomes of citalopram/escitalopram treatment.

The results suggested that CYP2C19 PMs showed higher treatment efficacy but also higher risk of side effects in the first weeks of treatment, with small effect sizes according to the usually accepted statistical classification (Chen et al., 2010). Under a clinical point of view, the identified effect sizes were significant since a clinical relevance cutoff of $SMD=0.24$ was proposed based on the effect size observed for antidepressant drugs ($SMD=0.31$, $CI=0.27-0.35$) and psychotherapy ($SMD=0.25$, $CI= 0.14-0.36$) in depression (Cuijpers et al., 2014). Another interesting point emerging from the study was the possibility to obtain CYP2C19 metabolizing activity from genome-wide data using imputation with a good level of accuracy, as demonstrated by the comparison of the classification obtained using genotyped and imputed data in one of the investigated samples. Thus no need of additional genotyping to determine CYP2C19 functionality is needed in patients with genome-wide genotyping.

Genes with previous evidence of involvement in the mechanisms of antidepressant action are also interesting in order to develop targeted analysis of genome-wide data. In *chapter 3* we demonstrated that the study of strong candidate genes according to previous literature could provide promising findings when expanded to interacting genes. This study was the first to use machine learning to investigate the contribution of SNPs in a candidate gene pathway to antidepressant efficacy. Indeed machine learning has been previously applied using clinical-demographic predictors (Kautzky et al., 2017; Iniesta et al., 2016) and combinations of few candidate polymorphisms and clinical variables (Kautzky et al., 2015), but not to more numerous sets of polymorphisms. Genes in the identified pathway showed an overlap with another pathway previously associated with antidepressant response that involves inorganic cation signaling (Cocchi et al., 2016). Results suggested good sensitivity and PPV (positive predictive value) in the detection of treatment-resistant depression (TRD) but not so good specificity and NPV (negative predictive value) and replication in an independent sample was not performed. Interestingly, the 30 SNPs contributing more to the model were mostly located in genes that have been previously linked to neurogenesis, neuroplasticity and/or antidepressant action. These results may pave the way to the identification of sets of genetic predictors in specific pathways able to predict the risk of TRD. It is reasonable to hypothesize a certain degree of variability in the genetic variants involved in TRD across different patients, but the involved pathways are expected to be more stable and the demonstration of this hypothesis is one of the lines of our future research plans.

Exome genotyping and high-density imputation to improve the covering of genetic variants

Previous meta-analyses of GWAS of antidepressant response included from 500,000 to 1.3 million common variants (Tansey et al., 2012; GENDEP Investigators et al., 2013), that are not more than 3.3% of the known common variants in the human genome (McCarthy et al., 2016). Rare variants were overlooked since the common disease – common variant theory inspired previous GWAS of antidepressant response (Reich and Lander, 2001). On the other hand, part of the so called missing heritability was attributed to the putative effect of rare variants (Cirulli and Goldstein, 2010). In *chapter 4* we performed a GWAS meta-analysis that combined genotyping by an exome array and a genome-wide array in order to include both common and rare variants in the analysis. In addition, for the first time in the study of antidepressant pharmacogenomics we performed imputation of genotypes using the Haplotype Reference Consortium data that represents the largest database of human common and rare genetic variants (McCarthy et al., 2016). This approach allowed the inclusion of over 7 million variants in the analysis and the identification of two new genome-wide significant loci, rs116692768 and rs76191705, in the ITGA9 (integrin alpha 9) and NRXN3 (neurexin 3) genes, respectively. Both the loci were obtained thanks to high-density imputation and they showed good imputation quality. The proteins coded by ITGA9 and NRXN3 show interesting functions in relation to antidepressant action. Integrins are heterodimeric transmembrane proteins that connect the extracellular environment to intracellular signaling. In the brain they are involved in the control of synaptic plasticity, long-term potentiation (LTP), cell adhesion and migration (Smith et al., 2015; Lynch et al., 2015). Polymorphisms in another beta isoform (ITGB3) and expression level of this gene have been associated with antidepressant response in humans (Probst-Schendzielorz et al., 2015; Fabbri et al., 2015). Neurexins are type I transmembrane neuronal adhesion receptors and their interaction with neuroligins is sufficient to trigger postsynaptic and presynaptic differentiation (Bot et al., 2011). The inclusion of rare exomic variants in the analysis provided the information needed to identify the association of the chromosome pathway (GO:0005694) with depressive symptom improvement. Indeed rare missense variants were 6.6% of all variants in this pathway, while missense variants were ~2% in the most part of the other investigated pathways. Abnormalities in gene expression levels in this pathway were previously associated with antidepressant efficacy in MDD (Belzeaux et al., 2012).

Pathway analysis: a way to investigate the molecular mechanisms implicated in treatment response

In this dissertation, several approaches were applied to perform pathway analysis. In *chapters 3, 5* and *6* pathway analysis was used to identify the putative biological mechanisms that mediate the association of candidate genes with antidepressant efficacy or antidepressant resistance. We showed that this approach provided consistent findings with the previous knowledge about the mechanisms of antidepressant action and with previous similar studies, but it added information about the role of the investigated candidate genes. In particular, *chapter 5* demonstrated that the association between the PPP3CC (protein phosphatase 3 catalytic subunit gamma) gene and antidepressant remission may be mediated by the involvement of this gene in the B-cell receptor signaling pathway. Several inflammatory abnormalities have been reported in MDD and they included also B cell proliferation and signaling. Indeed an increased percentage of B cells was demonstrated in the peripheral blood of patients with MDD and long-term treatment with SSRIs was demonstrated to restore alterations in B cells proliferation found in MDD (Maes et al., 1992; Hernandez et al., 2010; Edgar et al., 2002). A previous study reported interesting trends of association for other pathways involved in inflammation, in particular the KEGG antigen processing and presentation pathway and the KEGG tumor necrosis factor pathway (Hunter et al., 2013). Genes involved in extracellular matrix remodeling (e.g. ADAMTSL1, CD36, PON2, APOB, and PIK3R1) and thus in the release of inflammatory factors were members of another pathway associated with antidepressant efficacy in a previous GWAS (Ising et al., 2009). Similarly, in *chapter 6* we showed that the effect of the GAP43 (growth-associated protein 43) gene on antidepressant efficacy may involve the interaction of other genes in the same pathway that mediates different forms of neural plasticity. This pathway includes the ELAVL4, ZDHHC7, and GRIN2B genes. GRIN2B (glutamate receptor ionotropic N-methyl d-aspartate 2) is a known member of the long-term potentiation (LTP) pathway that has been associated with antidepressant efficacy in a previous GWAS (Hunter et al., 2013). ELAVL4 (ELAV-like neuron-specific RNA-binding protein 4) is involved in the processing of mRNAs that regulate hippocampal neurogenesis in response to antidepressants. ZDHHC7 (zinc finger, DHHC-type containing 7) regulates GABA subunits synaptic clustering and cell surface stability but it also affects intracellular signaling by sex hormones via ERK and AKT kinases which are known to modulate synaptic plasticity, mood, and antidepressant response (Fabbri et al., 2016). Several genes involved in GABAergic and glutamatergic neurotransmission but also neuroplasticity are members of a pathway previously

associated with antidepressant efficacy (Ising et al., 2009), particularly ERBB2, NRG1, and HOMER1. Genes involved in cell adhesion processes (e.g. FN1, EFNA5, and EPHA5) and modulation of microtubules and neurofilaments were also implicated in synaptic plasticity and they are involved in pathways associated with antidepressant response (O'Dushlaine et al., 2014).

In **chapter 4** we applied a different approach to pathway analysis: we did not investigate the known pathways that may mediate the effects of some candidate genes, but we applied pathway analysis with no *a priori* hypothesis including all the pathways from validated databases. Two GWAS were meta-analyzed in order to provide better statistical power and this represented the first application of pathway meta-analysis to the study of antidepressant response. Results involved chromatin remodeling (chromosome pathway (GO:0005694) and the chromosomal part (GO:0044427)), referring to the known role of neurogenesis and neural differentiation in antidepressant action. Indeed some of the top genes in these pathways are known regulators of the differentiation of neural stem cells into neurons (UPF1 (Regulator Of Nonsense Transcripts Homolog), HMGB1 (High Mobility Group Box 1) and neural development (FOXC1 (Forkhead Box C1; Alrabheni et al., 2015; Meneghini et al., 2013; Colleoni et al., 2011). Another promising pathway identified through this meta-analysis was the steroid hormone receptor signaling pathway (GO:0030518). The association between this pathway and antidepressant efficacy can be interpreted within the inflammation and HPA (hypothalamic–pituitary–adrenal) axis dysfunctions demonstrated in MDD and restored during antidepressant treatment (Anacker et al., 2011). The association of this pathway this depressive symptom improvement was replicated in an independent sample, demonstrating that pathway analysis improves power to detect and replicate positive associations.

In **chapter 7** a more exploratory strategy was applied to study the interactions among genes involved in treatment response. This study started from a classical genome-wide analysis of common variants involved in treatment response in BD; then polymorphisms with $p < 10e-4$ were annotated and the corresponding genes were investigated for possible enrichment with Gene Ontology (GO) terms in order to identify the biological functions these genes play a role in. This analysis was performed separately for treatment efficacy in preventing depressive episodes, (hypo)manic episodes and overall episodes in order to detect possible different biological mechanisms involved. Genes including SNPs associated with overall risk of recurrence were enriched with GO processes mediating behavior, learning and memory, cognition and brain development. Genes including SNPs associated with depressive recurrence were enriched with GO processes including the

regulation of MAPK cascade, protein kinase activity, neuron projection, cell morphogenesis and chemotaxis, stem cell differentiation, cytoskeleton reorganization, cell migration, glutamate receptor activity. These processes are linked with neurogenesis and neuroplasticity that have been extensively associated with treatment outcome in mood disorders (Crupi et al., 2011). The MAPK cascade is activated by the BDNF receptor TrkB that is essential for the transcription of neuronal plasticity genes (Arango-Lievano et al., 2015). Several genes (JAM3, CNTNAP2, PTPRD) that are part of the identified networks are also members of the cell adhesion molecule pathway that has been previously associated with BD susceptibility (O'Dushlaine et al., 2011). Other interesting candidate genes emerging from the identified networks were RASGRF1, NTRK2, GRIN2B, GRIN2A, GRM4, CHRNA7 and NTF3. RASGRF1 (Ras Protein Specific Guanine Nucleotide Releasing Factor 1), NTRK2 (Neurotrophic Receptor Tyrosine Kinase 2) and NTF3 (neurotrophin 3) are directly involved in the regulation of synaptic plasticity (Barman et al., 2014; Harward et al., 2016; Guo et al., 2017). GRIN2B and GRIN2A encode for two subunits of the N-methyl-d-aspartate (NMDA) glutamate receptor that is known to be involved in the pathophysiology of BD, as well as in antidepressant and mood stabilizers efficacy (Miller et al., 2017; Hashimoto et al., 2002). CHRNA7 (cholinergic receptor nicotinic alpha 7) is involved in cognitive function and synaptic plasticity and it is an attractive target for the development of new antidepressants with efficacy on cognitive symptoms (Andreasen et al., 2012; Pichat et al., 2007).

The enriched networks can be the starting point for further investigation, such as pathway analysis and polygenic risk scores (PRS). PRS are calculated as the sum of associated alleles, weighted by effect sizes, for polymorphisms with p-values less than pre-defined thresholds. PRS can be calculated at genome-wide level but this approach has often not provided brilliant results in the prediction of psychotropic drug response, probably because of the difficulty in separating genuine small effects from false positives when the p threshold is relaxed (Fabbri, 2017). Calculating PRS only including variants from a gene set with previous evidence of association with treatment outcome has been demonstrated a valid strategy. For example, pathway-level regional heritability was demonstrated to be significant in MDD (Zeng et al., 2016).

Strengths and limitations

The studies included in this dissertation provided some positive achievements that are expected to contribute to the identification of the genetic factors modulating treatment outcome in mood disorders. Indeed meta-analytical or independently

replicated evidence was provided for polymorphisms in several genes possibly involved in antidepressant action (CYP2C19, FKBP5, ITGA9, NRXN3, PPP3CC, CHL1) but also some pathways (chromosome pathway, chromosomal part pathway, steroid hormone receptor signaling pathway). Several successful strategies were implemented to achieve these results and overcome some of the limitations of previous GWAS, such as increased covering of common and rare genetic variants, more extensive use of meta-analysis also at pathway level, different approaches to gene set analysis, combination of the candidate gene approach and the genome-wide approach in order to exploit genome-wide data in a hypothesis-driven way.

On the other hand, the limitations of the studies included in this dissertation should also be considered. Though the strategies described above, statistical power was an issue that limited the identification of positive findings or their replication. Some findings were not replicated and they should be considered preliminary, particularly the results of machine learning in the prediction of antidepressant resistance described in *chapter 3*. Further, machine learning provided specificity not suitable for clinical application, despite sensitivity and PPV were good. The study of rare exonic variants was improved compared to previous studies, but it was still limited to the use of an exome array (*chapter 4*) while exome sequencing would be the state of the art technology to provide a complete covering of exonic variants. Further, rare imputed variants were excluded from the analyses because of higher risk of imputation artifacts compared to common variants. We were not able to effectively implement methods to reduce clinical and genetic heterogeneity of the analyzed samples. In *chapter 4*, stratifying the analyses by antidepressant provided some significant results in the subgroup of patients treated with citalopram/escitalopram, but other strategies were not implemented. Finally, we did not test possible clinical implementations of our pharmacogenetic findings.

Clinical implications and additional future directions

Despite we did not test possible clinical applications of our findings, the studies included in this dissertation suggest some interesting clinical implications and future lines of research.

The results described in *chapter 2* suggest that CYP2C19 poor metabolizers (PMs) show higher treatment efficacy using standard doses of citalopram or escitalopram, at the price of more common side effects in the first weeks of treatment but no higher total burden of side effects and no higher risk of dropout. Thus these results support that PMs should not receive a dose of citalopram or escitalopram reduced by 50% as previously suggested by pharmacokinetic studies (Hicks et al., 2015).

Accurate information of patients and careful monitoring of side effects during the first weeks of treatment are expected to minimize the burden of side effects in CYP2C19 PMs treated with standard citalopram/escitalopram doses while symptom improvement and remission rates could be improved.

The genetic predictors of antidepressant resistance identified using machine learning (*chapter 3*) represent the result of the first attempt of this kind. It is reasonable to hypothesize that future clinical application of antidepressant pharmacogenomics will involve polygenic predictors throughout the genome or specific molecular pathways, since the complex polygenic nature of the trait. Pathways consistently associated with antidepressant efficacy could become the starting point to study predictive models based on the variants within them. Indeed pathways involved in antidepressant efficacy are expected to be replicable through different samples, while the effects of single polymorphisms within them could be different from subject to subject because of the interaction with other genetic variants (or even more intriguing, environmental exposures). Machine learning could be the most suitable approach to take into account this variability.

The results of gene set enrichment obtained in *chapter 7* suggested several possible biological mechanisms involved in treatment outcome in BD, but the enriched GO processes should be investigated more in depth in order to confirm their role and identify possible clinical applications.

Concluding remarks

The studies included in this dissertation aimed to contribute to the identification of genetic predictors of treatment efficacy in mood disorders, with focus on antidepressant efficacy. This field of research has provided quite frustrating results in the last two decades, since no clinical application has been included in routine clinical practice yet. This dissertation tried to demonstrate that advances in the field are possible and can be implemented. We should not give up in the attempt to provide new evidence-based treatment to patients with mood disorders and genetic variants were demonstrated to be the most suitable biomarker for guiding clinicians in the choice of treatment.

References

Alrahbeni, T., Sartor, F., Anderson, J., Miedzybrodzka, Z., McCaig, C., and Müller, B. (2015). Full UPF3B function is critical for neuronal differentiation of neural stem cells. *Mol. Brain* 8, 33.

Anacker, C., Zunszain, P.A., Carvalho, L.A., and Pariante, C.M. (2011). The glucocorticoid receptor: pivot of depression and of antidepressant treatment? *Psychoneuroendocrinology* 36, 415–425.

Andreasen, J.T., Redrobe, J.P., and Nielsen, E.Ø. (2012). Combined $\alpha 7$ nicotinic acetylcholine receptor agonism and partial serotonin transporter inhibition produce antidepressant-like effects in the mouse forced swim and tail suspension tests: a comparison of SSR180711 and PNU-282987. *Pharmacol. Biochem. Behav.* 100, 624–629.

Arango-Lievano, M., Lambert, W.M., Bath, K.G., Garabedian, M.J., Chao, M.V., and Jeanneteau, F. (2015). Neurotrophic-priming of glucocorticoid receptor signaling is essential for neuronal plasticity to stress and antidepressant treatment. *Proc. Natl. Acad. Sci. U. S. A.* 112, 15737–15742.

Barman, A., Assmann, A., Richter, S., Soch, J., Schütze, H., Wüstenberg, T., Deibele, A., Klein, M., Richter, A., Behnisch, G., et al. (2014). Genetic variation of the RASGRF1 regulatory region affects human hippocampus-dependent memory. *Front. Hum. Neurosci.* 8, 260.

Belzeaux, R., Bergon, A., Jeanjean, V., Lloriod, B., Formisano-Tréziny, C., Verrier, L., Loundou, A., Baumstarck-Barrau, K., Boyer, L., Gall, V., et al. (2012). Responder and nonresponder patients exhibit different peripheral transcriptional signatures during major depressive episode. *Transl. Psychiatry* 2, e185.

Bot, N., Schweizer, C., Ben Halima, S., and Fraering, P.C. (2011). Processing of the synaptic cell adhesion molecule neurexin-3beta by Alzheimer disease alpha- and gamma-secretases. *J. Biol. Chem.* 286, 2762–2773.

Chen, H., Cohen, P., and Chen, S. (2010). How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies. *Commun. Stat. - Simul. Comput.* 39, 860–864.

Cirulli, E.T., and Goldstein, D.B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat. Rev. Genet.* 11, 415–425.

Citrome, L. (2014). Treatment of bipolar depression: making sensible decisions. *CNS Spectr.* 19 Suppl 1, 4-11; quiz 1-3, 12.

Cocchi, E., Fabbri, C., Han, C., Lee, S.-J., Patkar, A.A., Masand, P.S., Pae, C.-U., and Serretti, A. (2016). Genome-wide association study of antidepressant response: involvement of the inorganic cation transmembrane transporter activity pathway. *BMC Psychiatry* 16, 106.

Colleoni, S., Galli, C., Gaspar, J.A., Meganathan, K., Jagtap, S., Hescheler, J., Sachinidis, A., and Lazzari, G. (2011). Development of a neural teratogenicity test based on human embryonic stem cells: response to retinoic acid exposure. *Toxicol. Sci. Off. J. Soc. Toxicol.* 124, 370–377.

Crupi, R., Marino, A., and Cuzzocrea, S. (2011). New therapeutic strategy for mood disorders. *Curr. Med. Chem.* 18, 4284–4298.

Cuijpers, P., Turner, E.H., Koole, S.L., van Dijke, A., and Smit, F. (2014). What is the threshold for a clinically relevant effect? The case of major depressive disorders. *Depress. Anxiety* 31, 374–378.

Edgar, V.A., Cremaschi, G.A., Sterin-Borda, L., and Genaro, A.M. (2002). Altered expression of autonomic neurotransmitter receptors and proliferative responses in lymphocytes from a chronic mild stress model of depression: effects of fluoxetine. *Brain. Behav. Immun.* 16, 333–350.

Fabbri, C. (2017). Is a polygenic predictor of antidepressant response a possibility? *Pharmacogenomics* 18, 749–752.

Fabbri, C., and Serretti, A. (2016). Genetics of long-term treatment outcome in bipolar disorder. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 65, 17–24.

Fabbri, C., Crisafulli, C., Gurwitz, D., Stingl, J., Calati, R., Albani, D., Forloni, G., Calabrò, M., Martines, R., Kasper, S., et al. (2015). Neuronal cell adhesion genes and antidepressant response in three independent samples. *Pharmacogenomics J.* 15, 538–548.

Fabbri, C., Crisafulli, C., Calabrò, M., Spina, E., and Serretti, A. (2016). Progress and prospects in pharmacogenetics of antidepressant drugs. *Expert Opin. Drug Metab. Toxicol.* 12, 1157–1168.

GBD 2015 Disease and Injury Incidence and Prevalence Collaborators (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Lond. Engl.* 388, 1545–1602.

GENDEP Investigators, MARS Investigators, and STAR*D Investigators (2013). Common genetic variation and antidepressant efficacy in major depressive disorder: a meta-analysis of three genome-wide pharmacogenetic studies. *Am. J. Psychiatry* 170, 207–217.

Gratten, J., Wray, N.R., Keller, M.C., and Visscher, P.M. (2014). Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nat. Neurosci.* 17, 782–790.

Guo, Y., Su, Z.-J., Chen, Y.-K., and Chai, Z. (2017). Brain-derived neurotrophic factor/neurotrophin 3 regulate axon initial segment location and affect neuronal excitability in cultured hippocampal neurons. *J. Neurochem.* 142, 260–271.

Harward, S.C., Hedrick, N.G., Hall, C.E., Parra-Bueno, P., Milner, T.A., Pan, E., Laviv, T., Hempstead, B.L., Yasuda, R., and McNamara, J.O. (2016). Autocrine BDNF-TrkB signalling within a single dendritic spine. *Nature* 538, 99–103.

Hashimoto, R., Hough, C., Nakazawa, T., Yamamoto, T., and Chuang, D.-M. (2002). Lithium protection against glutamate excitotoxicity in rat cerebral cortical neurons:

involvement of NMDA receptor inhibition possibly by decreasing NR2B tyrosine phosphorylation. *J. Neurochem.* 80, 589–597.

Hernandez, M.E., Martinez-Fong, D., Perez-Tapia, M., Estrada-Garcia, I., Estrada-Parra, S., and Pavón, L. (2010). Evaluation of the effect of selective serotonin-reuptake inhibitors on lymphocyte subsets in patients with a major depressive disorder. *Eur. Neuropsychopharmacol. J. Eur. Coll. Neuropsychopharmacol.* 20, 88–95.

Hicks, J.K., Bishop, J.R., Sangkuhl, K., Müller, D.J., Ji, Y., Leckband, S.G., Leeder, J.S., Graham, R.L., Chiulli, D.L., LLerena, A., et al. (2015). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2D6 and CYP2C19 Genotypes and Dosing of Selective Serotonin Reuptake Inhibitors. *Clin. Pharmacol. Ther.* 98, 127–134.

Hou, L., Heilbronner, U., Degenhardt, F., Adli, M., Akiyama, K., Akula, N., Arda, R., Arias, B., Backlund, L., Banzato, C.E.M., et al. (2016). Genetic variants associated with response to lithium treatment in bipolar disorder: a genome-wide association study. *Lancet Lond. Engl.* 387, 1085–1093.

Hunter, A.M., Leuchter, A.F., Power, R.A., Muthén, B., McGrath, P.J., Lewis, C.M., Cook, I.A., Garriock, H.A., McGuffin, P., Uher, R., et al. (2013). A genome-wide association study of a sustained pattern of antidepressant response. *J. Psychiatr. Res.* 47, 1157–1165.

Iniesta, R., Malki, K., Maier, W., Rietschel, M., Mors, O., Hauser, J., Henigsberg, N., Dernovsek, M.Z., Souery, D., Stahl, D., et al. (2016). Combining clinical variables to optimize prediction of antidepressant treatment outcomes. *J. Psychiatr. Res.* 78, 94–102.

Ising, M., Lucae, S., Binder, E.B., Bettecken, T., Uhr, M., Ripke, S., Kohli, M.A., Hennings, J.M., Horstmann, S., Kloiber, S., et al. (2009). A genomewide association study points to multiple loci that predict antidepressant drug treatment outcome in depression. *Arch. Gen. Psychiatry* 66, 966–975.

Jukić, M.M., Opel, N., Ström, J., Carrillo-Roa, T., Miksys, S., Novalen, M., Renblom, A., Sim, S.C., Peñas-Lledó, E.M., Courtet, P., et al. (2016). Elevated CYP2C19 expression is associated with depressive symptoms and hippocampal homeostasis impairment. *Mol. Psychiatry* 22, 1155-1163.

Kautzky, A., Baldinger, P., Souery, D., Montgomery, S., Mendlewicz, J., Zohar, J., Serretti, A., Lanzenberger, R., and Kasper, S. (2015). The combined effect of genetic polymorphisms and clinical parameters on treatment outcome in treatment-resistant depression. *Eur. Neuropsychopharmacol. J. Eur. Coll. Neuropsychopharmacol.* 25, 441–453.

Kautzky, A., Baldinger-Melich, P., Kranz, G.S., Vanicek, T., Souery, D., Montgomery, S., Mendlewicz, J., Zohar, J., Serretti, A., Lanzenberger, R., et al. (2017). A New Prediction Model for Evaluating Treatment-Resistant Depression. *J. Clin. Psychiatry* 78, 215–222.

Li, Q.S., Tian, C., Seabrook, G.R., Drevets, W.C., and Narayan, V.A. (2016). Analysis of 23andMe antidepressant efficacy survey data: implication of circadian rhythm and neuroplasticity in bupropion response. *Transl. Psychiatry* 6, e889.

Lynch, G., Kramár, E.A., and Gall, C.M. (2015). Protein synthesis and consolidation of memory-related synaptic changes. *Brain Res.* 1621, 62–72.

Maes, M., Stevens, W.J., DeClerck, L.S., Bridts, C.H., Peeters, D., Schotte, C., and Cosyns, P. (1992). A significantly increased number and percentage of B cells in depressed subjects: results of flow cytometric measurements. *J. Affect. Disord.* 24, 127–134.

McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang, H.M., Fuchsberger, C., Danecek, P., Sharp, K., et al. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* 48, 1279–1283.

Meneghini, V., Bortolotto, V., Francese, M.T., Dellarole, A., Carraro, L., Terzieva, S., and Grilli, M. (2013). High-mobility group box-1 protein and β -amyloid oligomers promote neuronal differentiation of adult hippocampal neural progenitors via receptor for advanced glycation end products/nuclear factor- κ B axis: relevance for Alzheimer's disease. *J. Neurosci. Off. J. Soc. Neurosci.* 33, 6047–6059.

Miller, O.H., Bruns, A., Ben Ammar, I., Mueggler, T., and Hall, B.J. (2017). Synaptic Regulation of a Thalamocortical Circuit Controls Depression-Related Behavior. *Cell Rep.* 20, 1867–1880.

Mrazek, D.A., Biernacka, J.M., O'Kane, D.J., Black, J.L., Cunningham, J.M., Drews, M.S., Snyder, K.A., Stevens, S.R., Rush, A.J., and Weinshilboum, R.M. (2011). CYP2C19 variation and citalopram response. *Pharmacogenet. Genomics* 21, 1–9.

O'Brien, F.E., O'Connor, R.M., Clarke, G., Dinan, T.G., Griffin, B.T., and Cryan, J.F. (2013). P-glycoprotein inhibition increases the brain distribution and antidepressant-like activity of escitalopram in rodents. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* 38, 2209–2219.

O'Dushlaine, C., Kenny, E., Heron, E., Donohoe, G., Gill, M., Morris, D., International Schizophrenia Consortium, and Corvin, A. (2011). Molecular pathways involved in neuronal cell adhesion and membrane scaffolding contribute to schizophrenia and bipolar disorder susceptibility. *Mol. Psychiatry* 16, 286–292.

O'Dushlaine, C., Ripke, S., Ruderfer, D.M., Hamilton, S.P., Fava, M., Iosifescu, D.V., Kohane, I.S., Churchill, S.E., Castro, V.M., Clements, C.C., et al. (2014). Rare copy number variation in treatment-resistant major depressive disorder. *Biol. Psychiatry* 76, 536–541.

Paulzen, M., Groppe, S., Tauber, S.C., Veselinovic, T., Hiemke, C., and Gründer, G. (2015). Venlafaxine and O-desmethylvenlafaxine concentrations in plasma and cerebrospinal fluid. *J. Clin. Psychiatry* 76, 25–31.

Peterson, K., Dieperink, E., Anderson, J., Boundy, E., Ferguson, L., and Helfand, M. (2017). Rapid evidence review of the comparative effectiveness, harms, and cost-effectiveness of pharmacogenomics-guided antidepressant treatment versus usual care for major depressive disorder. *Psychopharmacology (Berl.)* 234, 1649–1661.

Pichat, P., Bergis, O.E., Terranova, J.-P., Urani, A., Duarte, C., Santucci, V., Gueudet, C., Voltz, C., Steinberg, R., Stemmelin, J., et al. (2007). SSR180711, a novel selective alpha7 nicotinic receptor partial agonist: (II) efficacy in experimental models predictive of activity against cognitive symptoms of schizophrenia. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* 32, 17–34.

Porcelli, S., Fabbri, C., Spina, E., Serretti, A., and De Ronchi, D. (2011). Genetic polymorphisms of cytochrome P450 enzymes and antidepressant metabolism. *Expert Opin. Drug Metab. Toxicol.* 7, 1101–1115.

Probst-Schendzielorz, K., Scholl, C., Efimkina, O., Ersfeld, E., Viviani, R., Serretti, A., Fabbri, C., Gurwitz, D., Lucae, S., Ising, M., et al. (2015). CHL1, ITGB3 and SLC6A4 gene expression and antidepressant drug response: results from the Munich Antidepressant Response Signature (MARS) study. *Pharmacogenomics* 16, 689–701.

Reich, D.E., and Lander, E.S. (2001). On the allelic spectrum of human disease. *Trends Genet.* 17, 502–510.

Ripke, S., Neale, B.M., Corvin, A., Walters, J.T.R., Farh, K.-H., Holmans, P.A., Lee, P., Bulik-Sullivan, B., Collier, D.A., Huang, H., et al. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511, 421–427.

Smith, A.C.W., Scofield, M.D., and Kalivas, P.W. (2015). The tetrapartite synapse: Extracellular matrix remodeling contributes to corticoaccumbens plasticity underlying drug addiction. *Brain Res.* 1628, 29–39.

Song, J., Bergen, S.E., Di Florio, A., Karlsson, R., Charney, A., Ruderfer, D.M., Stahl, E.A., Members of the International Cohort Collection for Bipolar Disorder (ICCBD), Chambert, K.D., Moran, J.L., et al. (2016). Genome-wide association study identifies SESTD1 as a novel risk gene for lithium-responsive bipolar disorder. *Mol. Psychiatry* 21, 1290–1297.

Squassina, A., Manchia, M., Borg, J., Congiu, D., Costa, M., Georgitsi, M., Chillotti, C., Ardau, R., Mitropoulos, K., Severino, G., et al. (2011). Evidence for association of an ACCN1 gene variant with response to lithium treatment in Sardinian patients with bipolar disorder. *Pharmacogenomics* 12, 1559–1569.

Tansey, K.E., Guipponi, M., Perroud, N., Bondolfi, G., Domenici, E., Evans, D., Hall, S.K., Hauser, J., Henigsberg, N., Hu, X., et al. (2012). Genetic predictors of response to serotonergic and noradrenergic antidepressants in major depressive disorder: a genome-wide analysis of individual-level data and a meta-analysis. *PLoS Med.* 9, e1001326.

Tansey, K.E., Guipponi, M., Hu, X., Domenici, E., Lewis, G., Malafosse, A., Wendland, J.R., Lewis, C.M., McGuffin, P., and Uher, R. (2013). Contribution of Common Genetic Variants to Antidepressant Response. *Biol. Psychiatry* 73, 679–682.

Visscher, P.M., Hemani, G., Vinkhuyzen, A.A.E., Chen, G.-B., Lee, S.H., Wray, N.R., Goddard, M.E., and Yang, J. (2014). Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. *PLoS Genet.* 10, e1004269.

Zeng, Y., Navarro, P., Fernandez-Pujals, A.M., Hall, L.S., Clarke, T.-K., Thomson, P.A., Smith, B.H., Hocking, L.J., Padmanabhan, S., Hayward, C., et al. (2016). A Combined Pathway and Regional Heritability Analysis Indicates NETRIN1 Pathway Is Associated With Major Depressive Disorder. *Biol. Psychiatry.* 81, 336-346.

Chapter 10

Valorization chapter

The importance of collaboration: open source data and expertise sharing should be a pillar of future research

The studies included in the present dissertation provided several new approaches to the study of pharmacogenetics in depression and they valorized the importance of data and expertise sharing as a successful model for future research.

Pharmacogenetics is a research field mostly represented by scientists with a non-clinical background, such as bioinformatics and statistical genetics. A clinical-oriented point of view is therefore precious in psychiatric pharmacogenetics, since the importance of a good knowledge of psychopharmacology to guide hypothesis formulation and testing. *Chapter 2* of this dissertation provides an example of this concept: the knowledge of the pharmacokinetics of antidepressant drugs was the preliminary step for the realization of the study which was born as a spontaneous idea during the discussion between me (a clinical scientist) and a professor in statistical genetics (Cathryn Lewis). Cytochrome P450 (CYP450) genes show a high complexity because of their polymorphic nature and the interpretation of the consequences of the different possible allelic combinations is not straightforward (<https://www.pharmvar.org>). The direct collaboration with experts in complementary fields allowed the combination of psychopharmacology, genetics and biostatistics in the study, which was the first to demonstrate that CYP2C19 enzymatic level can be accurately estimated using imputed genome-wide data (bioinformatics and genetic relevant finding) and that CYP2C19 poor metabolizers do not probably need any dose reduction when treated with citalopram or escitalopram (clinically relevant finding). This provides an example of complementary expertise sharing that should guide future studies in order to avoid the compartmentalization of science. Genetics and pharmacogenetics should indeed not be the ground of non-clinical scientists only.

Another key point that should be taken into account in research is the possibility of accessing previously collected databases that are often available to qualified researchers and can provide very valuable resources

for primary analyses or replication. We entered in a scientific era more and more often characterized by the sharing of not only results but also raw data. The initiatives of the US NIMH Center for Collaborative Genomic Studies provide a collection of over 150,000 well characterized patient and control samples from a wide-range of mental disorders (<https://www.nimhgenetics.org>). Genomic and clinical data from NIMH repositories were included in all the studies of this dissertation for replication purpose, for providing complementary evidence or as part of the primary analysis, demonstrating that open access data can provide exceptional extra-value and the importance of not being focused on only in-house data (which is also much more expensive and time consuming). Under the same spirit, data from my ongoing projects will be included in international consortiums, such as the Psychiatric Genetics Consortium (PGC). Future research should not ignore the resources already available, and at the same time sharing our own data should become a routine procedure.

Social and economic considerations

Depressive disorders are the third cause of disability considering all non-communicable diseases (Disease and Injury Incidence and Prevalence Collaborators, 2016). The economic burden of major depression increased from \$173.2 billion to \$210.5 billion in the period 2005-2010 in the US and in Europe it was estimated to be ~92 billion of euro in 2010 (Greenberg et al., 2015; Olesen et al., 2012). Treatment-resistant depression (TRD) was demonstrated to be a relevant factor in contributing to depression-related costs (Mrazek et al., 2014). For this reason, several studies included in the present dissertation assessed the genetic factors involved in TRD risk (*chapters 3, 5, and 6*).

In this scenario, the identification of biomarkers able to predict antidepressant response and guide treatment prescription is of undoubted value. This dissertation identified genetic markers of antidepressant efficacy and side effects at individual gene level and at pathway level. At individual gene level, there is encouraging evidence that combinatorial approaches including relevant polymorphisms can provide clinically meaningful information to guide antidepressant treatment, as showed for CYP2C19 and

FKBP5 genes in *chapter 2* and *chapter 3*. Pharmacogenetic tests including variants in these genes have preliminary evidence of improving remission rates compared to standard care (Stamm et al., 2016; Bradley et al., 2018). Combinatorial approaches applied at pathway level (i.e. a set of functionally related genes) encompass the exciting opportunity to explain a higher proportion of variance in antidepressant response and reflect the final functional balance of a biological system relevant to antidepressant action. The application of predictive modeling at pathway level was tested in *chapter 3* and it was one of the first studies of this kind in psychiatric pharmacogenetics. Many academic centers are planning or implementing start up efforts that aim to integrate genomic information and clinical information from electronic health records to optimize treatments using predictive modeling (Kalinin et al., 2018; Kannry and Williams, 2013). A deep learning approach may represent a game-changing advance compared to currently available pharmacogenetic tests that are based on the results of candidate gene studies (Fabbri et al., 2018).

Starting point for future research

The studies included in this dissertation were the starting point for planning further projects that are now in the implementation phase. Indeed, the study of treatment-resistant depression (TRD) and the idea of expanding genomic coverage to rare exomic variants were considered particularly promising since the paucity of previously published results and the points discussed in the previous section. The studies described in *chapters 3, 4, 5* and *6* express the innovativeness and relevance of these concepts for improving the potential impact of pharmacogenetics at clinical level. This idea was shared by stakeholders who gave credit to the potential socio-economic impact of this line of research and funded new projects that are going to extend the studies included in this dissertation (“Fondazione Umberto Veronesi” and European Union (Marie Skłodowska-Curie fellowship)). Moving the focus of analysis from individual variants to pathways was another key point in *chapters 3-7* of this dissertation and it is also the main methodological frame guiding the genetic analysis of the new projects. One important advance of the new analyses will be the computing of polygenic scores at both common and rare variant level, and their calculation using all available

variants in the genome/exome but also variants in specific pathways only. Indeed part of the so-called missing heritability may be due to the effect of rare (and possibly unknown) variants, as suggested in *chapter 4*.

Ideally, prediction algorithms should include comprehensive genetic information (common and rare variants), but also clinical-demographic variables, a point that will be addressed in ongoing projects. Indeed the combination of these predictors is expected to fit the hypothesized mechanisms underling TRD and maximize the chances to develop valid and reproducible tests suitable for clinical application.

References

- Bradley P., et al., 2018. *Psychiatr Res* 96, 100–107.
- Dold M. et al., 2018. *Int. J. Neuropsychopharmacol.* 21, 539–549.
- Fabbri C. et al., 2018. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 86, 36–44.
- GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, 2016. *Lancet* 388, 1545–1602.
- Greenberg P.E., et al., 2015. *J Clin Psychiatry* 76, 155–162.
- Hornberger J. et al., 2015. *Am J Manag Care* 21, e357-365.
- Kalinin A.A. et al., 2018. *Pharmacogenomics* 19, 629–650.
- Kannry J.L. et al., 2013. *Genet. Med.* 15, 757–760.
- Mrazek D.A. et al., 2014. *Psychiatr Serv* 65, 977–987.
- Olesen J. et al., 2012. *Eur. J. Neurol.* 19, 155–162.
- Stamm T.J. et al. 2016. *J. Psychopharmacol.* 30, 40–47.

Aknowledgments

I want to thank Prof. dr. Koen Schruers and Prof. dr. Alessandro Serretti for supervising me in the preration of this thesis and the whole assessment committee for giving me the opportunity to defend it at Maastricht University. Prof. dr. Alessandro Serretti also taught me the basics of genetic research and supervised my work since 2010, when I was still a medical student and I started to write my first research papers.

I also want to thank Prof. dr. Cathryn M. Lewis for supervising part of the work included in this thesis and for supporting me and my projects.

I cannot forget about my family that was a fundamental support during my studies.