Genome-wide association study of treatment-resistance in depression and meta-analysis of three independent samples

Chiara Fabbri, Siegfried Kasper, Alexander Kautzky, Lucie Bartova, Markus Dold, Joseph Zohar, Daniel Souery, Stuart Montgomery, Diego Albani, Ilaria Raimondi, Dimitris Dikeos, Dan Rujescu, Rudolf Uher, Cathryn M. Lewis, Julien Mendlewicz and Alessandro Serretti

Background
Treatment-resistant depression (TRD) is the most problematic outcome of depression in terms of functional impairment, suicidal thoughts and decline in physical health.

Aims
To investigate the genetic predictors of TRD using a genome-wide approach to contribute to the development of precision medicine.

Method
A sample recruited by the European Group for the Study of Resistant Depression (GSRD) including 1148 patients with major depressive disorder (MDD) was characterised for the occurrence of TRD (lack of response to at least two adequate antidepressant treatments) and genotyped using the Infinium PsychArray. Three clinically relevant patient groups were considered: TRD, responders and non-responders to the first antidepressant trial, thus outcomes were based on comparisons of these groups. Genetic analyses were performed at the variant, gene and gene-set (i.e. functionally related genes) level. Additive regression models of the outcomes and relevant covariates were used in the GSRD participants and in a fixed-effect meta-analysis performed between GSRD, STAR*D (n = 1316) and GENDEP (n = 761) participants.

Results
No individual polymorphism or gene was associated with TRD, although some suggestive signals showed enrichment in cytoskeleton regulation, transcription modulation and calcium signalling. Two gene sets (GO:0043949 and GO:0000183) were associated with TRD versus response and TRD versus response and non-response to the first treatment in the GSRD participants and in the meta-analysis, respectively (corrected \( P = 0.030 \) and \( P = 0.027 \)).

Conclusions
The identified gene sets are involved in cyclic adenosine mono-phosphate mediated signal and chromatin silencing, two processes previously implicated in antidepressant action. They represent possible biomarkers to implement personalised antidepressant treatments and targets for new antidepressants.

Declaration of interest
D.S. has received grant/research support from GlaxoSmithKline and Lundbeck; has served as a consultant or on advisory boards for AstraZeneca, Bristol-Myers Squibb, Eli Lilly, Janssen and Lundbeck. S.M. has been a consultant or served on advisory boards for: AstraZeneca, Bristol-Myers Squibb, Forest, Johnson & Johnson, Leo, Lundbeck, Medelink, Neurim, Pierre Fabre, Richter. S.K. has received grant/research support from Eli Lilly, Lundbeck, Bristol-Myers Squibb, GlaxoSmithKline, Organon, Seprocar and Servier; has served as a consultant or on advisory boards for AstraZeneca, Bristol-Myers Squibb, GlaxoSmithKline, Eli Lilly, Lundbeck, Pfizer, Organon, Schwabe, Seprocar, Servier, Janssen and Novartis; and has served on speakers’ bureaus for AstraZeneca, Eli Lilly, Lundbeck, Schwabe, Seprocar, Servier, Pierre Fabre, Janssen and Neurapharm. J.Z. has received grant/research support from Lundbeck, Servier, Brainsway and Pfizer, has served as a consultant or on advisory boards for Servier, Pfizer, Abbott, Lilly, Actelion, AstraZeneca and Roche and has served on speakers’ bureaus for Lundbeck, Roch, Lilly, Servier, Pfizer and Abbott. J.M. is a member of the Board of the Lundbeck International Neuroscience Foundation and of Advisory Board of Servier. A.S. is or has been consultant/speaker for: Abbott, AbbVie, Angelini, AstraZeneca, Clinical Data, Boehringer, Bristol Myers Squibb, Eli Lilly, GlaxoSmithKline, Innovapharma, Italfarmaco, Janssen, Lundbeck, Naurex, Pfizer, Polifarma, Sanofi and Servier. C.M.L. receives research support from RGA UK Services Limited.

Keywords
Treatment-resistant depression; antidepressants; polymorphism; GWAS; pathway.

Copyright and usage
© The Royal College of Psychiatrists 2018.
regulating actin cytoskeleton, although the finding did not survive multiple-testing correction.4 The present paper reports a GWAS of common variants and a meta-analysis at the polymorphism, gene and gene-set (pathway) level, with the aim of contributing to filling the gap in our knowledge about TRD genetics.

Method

Participants

European group for the study of resistant depression (GSRD)

The GSRD participants were recruited within a multicentre, cross-sectional study including 1346 adults who were in- and out-patients with MDD according to DSM-IV-TR criteria.5 The GSRD has been active for more than 20 years in the field of clinical and genetic modulators of TRD. Diagnosis was confirmed using the Mini International Neuropsychiatric Interview (MINI).6 Inclusion criteria were: (a) having received at least one antidepressant during the current MDD episode (≥4 weeks at adequate dose); (b) Montgomery–Åsberg Depression Rating Scale (MADRS)7 score >22 at the onset of the current MDD episode. Exclusion criteria were (a) any other primary psychiatric disorder than MDD, (b) any substance disorder (except nicotine and caffeine) in the previous 6 months, and (c) any condition that could interfere with the ability to give informed consent or with the assessments required by the study (for example linguistic barrier). Depressive symptom severity was assessed using the MADRS at study inclusion and at the onset of the current MDD episode. Information on previous and current antidepressant and other pharmacological treatments during the current MDD episode was collected as well as clinical–demographic characteristics. Antidepressant treatment was naturalistic according to best-clinical practice principles.

All procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. All procedures involved in the study were approved by the local ethics committees of each participating centre (coordinating centre approval number: B406201213479). Further details can be found elsewhere.8 Written informed consent was obtained from all patients included in this study.

Sequenced treatment alternatives to relieve depression (STAR*D)

A total of 4041 participants with MDD (DSM-IV criteria)9 were enrolled from primary care or psychiatric out-patient clinics. Of these, 2876 patients met the inclusion criterion of having at least moderate depression severity and 1948 were genotyped. Severity of depression was assessed using the 16-item Quick Inventory of Depressive Symptomatology-Clinician Rated (QIDS-C16)10 at baseline and then every 2 weeks until week 12. At level 1, all patients received citalopram. Participants without sufficient symptomatic benefit were eligible for randomisation to level 2 treatments, received citalopram. Participants without sufficient symptomatic improvement by the study (for example linguistic barrier). Depressive symptom severity was assessed using the MADRS, Hamilton Rating Scale for Depression (HRSD–17)13 and other measures. Other information about the GENDEP study can be found elsewhere.14

Phenotypes

TRD was defined according to the most common definition of lack of response to at least two adequate antidepressant treatments.15 In the participants in the GSRD sample response was defined as a MADRS score <22 and a score decrease of at least 50% compared with the onset of the current MDD episode. In the participants in the STAR*D sample a comparable definition of response was used (QIDS-C16 <13 (equivalent to MADRS of 22) and score decrease of at least 50% compared with baseline), whereas in GENDEP a decrease of at least 50% in MADRS score compared with baseline was used as the only criterion according to the available definition of response in this sample.14

The same phenotypes were investigated in the GSRD participants and in the meta-analysis of GSRD, STAR*D and GENDEP participants: (a) TRD versus response to the first antidepressant treatment; (b) per cent symptom improvement in TRD and responders to the first antidepressant treatment; (c) TRD versus others (responders and non-responders to the first antidepressant treatment). Phenotypes (a) and (b) are expected to provide more specific information on the genetic risk factors involved in TRD, because non-responders may develop TRD or not. Phenotype (c) was investigated to provide comparative results.

In the GSRD sample response to the antidepressants prescribed during the current MDD episode was assessed using a cross-sectional design, in the STAR*D sample longitudinal data referred to level 1 and level 2 were used to create the phenotypes and in the GENDEP sample prospective data collected during the 12-week trial were combined with the retrospective information on previous antidepressant treatments of the current episode to determine the phenotypes.14 In STAR*D and GENDEP missing QIDS-C16 and MADRS evaluations were handled as previously described.16

Genotyping and imputation

In the GSRD participants genotyping was performed using the Illumina Infinium PsychArray 24 BeadChip (Illumina, Inc., San Diego). Genome-wide data available in STAR*D were obtained using Affymetrix Human Mapping 500 K Array Set or Affymetrix Genome-Wide Human SNP Array 5.0 (Affymetrix, San Francisco, California), whereas in GENDEP, Illumina Human610-quad bead chip (Illumina, Inc., San Diego) was used.

Pre-imputation quality control was performed according to the following criteria: (a) variants with a missing rate ≥5%; (b) monomorphic variants; (c) participants with genotyping rate <97%; (d) participants with gender discrepancies; (e) participants with abnormal heterozygosity; (f) related participants (identity by descent (IBD) >0.1875); (g) population outliers according to Eigensoft analysis of linkage-disequilibrium-pruned genetic data;17 (h) participants from ethnic groups other than White.

Genotypes were imputed using the Haplotype Reference Consortium (HRC) r1.1 2016 data as reference panel and Minimac3. Post-imputation quality control was performed prunning variants according to the following criteria: (a) poor imputation quality ($R^2$ estimated squared correlation between imputed genotypes and true genotypes) <0.30;18 (b) minor allele frequency (MAF) <0.01.

Statistical analysis

Linear or logistic regression models were applied as appropriate to study the role of single variants in GSRD alone and in all samples through a fixed-effects meta-analysis, using the continuous and binary outcomes described in the Phenotypes section. The first
ten population principal components were used as covariates since they explained the 70% of variance in population structure. Other covariates were the variables showing an impact on outcomes (centre of recruitment and baseline symptom severity in all groups, age in STAR*D and GENDEP). We decided not to apply mixed linear models because we did not expect spurious associations caused by population structure and relatedness. Indeed, only participants who were White were included in this study, related participants were excluded (IBD >0.1875) and population principal components were used as covariates. Furthermore, the genomic inflation factor was calculated in order to assess if the observed results may be affected by population stratification, cryptic relatedness or genotyping errors. Heterogeneity across studies was measured using Cochran’s Q statistic and I² statistic. A standard genome-wide significance threshold of 5 × 10⁻⁸ was used while a suggestive significance threshold was set at a P-value of 5 × 10⁻⁶.19 plink 1.9 was used for these analyses assuming an additive model. At the genome-wide threshold of significance, the meta-analysis of TRD versus others provides a power of 0.80 to identify a variant with a MAF difference of 2.4% between patients with TRD and those who are non-TRD (MAF of 6.4% and 4%, respectively), corresponding to an odds ratio (OR) = 1.59, and the comparison of patients with TRD versus responders the identifiable MAF difference is 2.7% (MAF of 6.7% and 4%, respectively), corresponding to an OR = 1.72.

The independent variants with suggestive level of association were annotated using FUMA (http://fuma.ctglab.nl/), including expression quantitative trait loci (which investigates if the expression of a gene is associated with allelic variation at a single nucleotide polymorphism (SNP) of interest) and chromatin state across 127 tissue/cell type. Enrichment in gene ontology (GO) functional categories of genes harbouring variants with a suggestive P-value was also considered. Briefly, this test evaluated the proportion of overlapping genes between those with suggestive association signals and those in GO functional categories using a hypergeometric test.20

The hypothesis that psychiatric traits are highly polygenic has received increasing support in recent years and multimarker tests show higher power than single-variant analysis. Thus, gene- and gene-set (i.e. sets of functionally related genes or pathways) analyses were performed using MAGMA (https://ctg.cncr.nl/software/magma) in the GSRD group and in all samples using a fixed-effects meta-analysis.21 These analyses are based on multiple regression models that included as predictors the variants in a gene or gene set and the same covariates considered in the variant-level analysis. MAGMA performs both a self-contained (i.e. testing if a gene set is more associated with the trait than it would be expected by chance) and a competitive gene-set analysis, the latter is more conservative and it was applied in this study as it reflects the difference in association between genes in the analysed gene set and genes outside it (i.e. in the rest of the genome). Genotyped variants with MAF <0.01 were included in the gene- and gene-set analyses and different weights were assigned to polymorphisms according to their MAF (rare variants were defined as function of sample size22). The same covariates used for SNP-level analysis were included. For MAGMA gene-level analysis, the false discovery rate (FDR) correction was applied,23 and for gene-set analysis 10 000 permutations were performed. The analysed pathways were downloaded from http://software.broadinstitute.org/gsea/downloads.jsp (Biocarta, KEGG, Gene Ontology, Reactome, microRNA targets and transcription factor targets, v. 6.1).

## Results

After quality control, 1148, 1316 and 761 participants (total = 3223) were included in the analysis of TRD versus others, from GSRD, STAR*D and GENDEP, respectively, and 759, 1119 and 336 participants (total = 2214), respectively, for the analyses including only patients with TRD and responders. The clinical-demographic characteristics and number of variants available in the analysed participants are shown in supplementary Table 1 available at https://doi.org/10.1192/bjp.2018.256. All included patients were of White ethnicity, the proportion of patients with TRD broadly varied across studies (60% in STAR*D, 42% in GSRD and 14% in GENDEP).

## Single-variant analysis

### GSRD participants

In total, 7 605 870 variants were available in this group after quality control and there was no evidence of genomic inflation (lambda values were ~1, QQ plots and lambda values are shown in supplementary Fig. 1). No variant was associated with TRD or symptom improvement at the genome-wide level of significance. The intergenic SNP rs7665833 was the closest to the significance threshold (P = 1.05 × 10⁻⁶, phenotype: patients with TRD versus others). The characteristics of variants with suggestive level of association are reported in supplementary Table 2. Genes within the regions harbouring variants with suggestive P-values showed enrichment in GO sets regulating intermediate filament cytoskeleton (supplementary Table 3). These genes were not differentially expressed across 30 general tissue types (supplementary Fig. 2).

### Meta-analysis

Approximately 7 000 000 variants overlapped across the three groups and were included in the meta-analysis (see Fig. 1 for the exact number of SNPs included in each analysis). There was no evidence of genomic inflation (lambda values ~1, QQ plots and lambda values are shown in supplementary Fig. 1). There were no loci associated with the analysed phenotypes, but two intergenic SNPs (rs12160925 and rs12160621) in complete linkage-disequilibrium were close to the significance threshold for association with symptom improvement (P = 9.14 × 10⁻⁵). These variants are located upstream of the seizure related 6 homolog like (SEZ6L) gene (supplementary Fig. 3). Other suggestive findings are shown in supplementary Table 4 and Manhattan plots are shown in Fig. 1. Genes in the regions of variants with suggestive P-values showed enrichment particularly in gene sets involved in transcription regulation, apoptosis, calcium signalling, synaptic transmission, second messenger cascades, secretion and response to hormones such as steroids (supplementary Table 5). When considering the phenotype TRD versus response, these genes showed a significant higher expression in the brain cerebellar hemisphere, thyroid and pituitary across 30 tissue types (P = 3.15 × 10⁻⁵, 3.15 × 10⁻⁴ and 5.01 × 10⁻⁴, respectively, supplementary Fig. 4). For the phenotype per cent improvement in patients with TRD and responders, genes of interest showed a significant higher expression in the hypothalamus (P = 3.15 × 10⁻⁵, supplementary Fig. 4), whereas there was no differential expression when considering patients with TRD versus others.

### Genes and gene sets

#### GSRD group

No gene was associated with the phenotypes of interest after FDR correction (genes with nominal P < 5 × 10⁻⁵ are reported in supplementary Table 6). The GO:0043949 gene set (regulation of cyclic adenosine monophosphate (cAMP) mediated signal) was associated with TRD versus response (nominal comparative P = 1.82 × 10⁻⁵, corrected comparative P = 0.030, Table 1). The most...
significant genes in this functional category were CRTC3 (CREB regulated transcription coactivator 3, \( P = 0.0024 \)) and PDE10A (phosphodiesterase 10A, \( P = 0.021 \)). All the other gene sets showed a corrected comparative \( P > 0.50 \) (results not shown).

**Meta-analysis**

No significant genes were identified, genes with \( P < 5 \times 10^{-4} \) are reported in supplementary Table 7. The gene ontology term GO:0000183, involved in the repression of transcription of ribosomal DNA by altering the structure of chromatin, was associated with TRD versus others (comparative corrected \( P = 0.027 \), Table 1). This gene set included 29 genes and the top ones were HIST1H4E (\( P = 0.015 \)), BEND3 (\( P = 0.017 \)) and SIRT2 (\( P = 0.021 \)). Gene sets with corrected comparative \( P < 0.50 \) are shown in supplementary Table 8.

**Discussion**

This GWAS aimed to identify common variants, genes and gene sets associated with TRD in order to contribute to the development of personalised treatments of MDD and reduce the heavy personal and socioeconomic burden of this disease. At variant level, no polymorphism was associated with the phenotypes of interest in the GSRD participants or in the meta-analysis, but several suggestive findings were identified.

**Key findings and interpretation**

In the meta-analysis, the intergenic variants rs12160925 and rs12160621, in complete linkage-disequilibrium, were close to the genome-wide significant threshold (\( P = 9.14 \times 10^{-8} \)). These variants

---

**Table 1** Significant results of the gene-set analysis

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Gene set</th>
<th>Description</th>
<th>( n ) of included genes</th>
<th>Comparative ( P )</th>
<th>Comparative ( P ) corrected</th>
<th>Top genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRD versus response (GSRD)</td>
<td>GO:0043949</td>
<td>Regulation of cAMP mediated signal</td>
<td>23</td>
<td>( 1.82 \times 10^{-6} )</td>
<td>0.030</td>
<td>CRTC3, PDE10A</td>
</tr>
<tr>
<td>TRD versus others (meta-analysis)</td>
<td>GO:0000183</td>
<td>Chromatin silencing</td>
<td>29</td>
<td>( 1.67 \times 10^{-5} )</td>
<td>0.027</td>
<td>HIST1H4E, BEND3, SIRT2</td>
</tr>
</tbody>
</table>

TRD, treatment-resistant depression; GSRD, Group for the Study of Resistant Depression; GO, gene ontology; cAMP, cyclic adenosine monophosphate.

a. Meta-analysis of participants in the three samples.
are located ~30 kilobase pairs upstream of the seizure related 6 homolog like (SEZ6L) and showed relevant linkage-disequilibrium ($R^2 = 0.7$) with an intronic variant of this gene (rs113368973). SEZ6L is involved in the modulation of excitatory synaptic transmission and it is important for the achievement of balance between elongation and branching during dendritic arborisation. Variants in this gene have been associated with bipolar disorder and differential levels of the coded protein were detected in the cerebrospinal fluid between patients with mood disorders and healthy controls. Other suggestive loci are within genes having a previously reported link with mood disorders and/or antidepressant action, such as CACNA1C and NEDD4.

Interestingly, regions harbouring suggestive variants were enriched in gene sets regulating calcium signalling and related pathways (apoptosis and synaptic transmission, particularly glutamatergic), neural projection development and hormone signalling (including response to steroids). All these processes are known to mediate antidepressant effects, suggesting that at least part of the suggestive variants may play a role in TRD although they did not reach the genome-wide significance threshold. Enrichment was also identified in gene sets involved in cytoskeleton regulation and regulation of second messenger cascades, in line with previous findings. Another encouraging finding was the observation of a significant higher expression of the genes of interest in some brain regions compared with other tissues (supplementary Figure 4).

We did not identify individual genes associated with TRD, although the functional gene set GO:0043949 was associated with TRD in the GSRD participants and GO:0000183 in the meta-analysis. The first is involved in the regulation of cAMP signalling, a pathway that in the brain is activated by neurotransmitters (for example through adrenergic receptors), hormones or chemokines. Through the activation of a heterotrimeric G protein, it stimulates adenyl cyclase and increases the cellular concentration of cAMP. The subsequent signalling cascade is known to control the activation of CREB (cAMP responsive element binding protein) and the transcription of target genes such as brain-derived neurotrophic factor. This pathway is involved in numerous neuronal biological processes, including cell survival, synaptic structure and synaptic plasticity, and it mediates antidepressant action. Posttranscriptional signal regulation has recently received attention for the potential development of antidepressants with new mechanisms of action. For example, PDE10A was one of the top genes identified in GO:0043949 and it has been reported as a potential target for new antidepressants. Our findings suggest the hypothesis that antidepressants acting via this alternative route may be effective in TRD.

Another potential target may be CRTC3 that was reported to be essential for the regulation of CREB-stimulated transcription of corticotropin releasing factor following a stressful stimulus. The effect of the GO:0043949 gene set was not confirmed in the meta-analysis (nominal comparative $P = 0.037$ and 0.046 for symptom improvement and TRD versus response, respectively, not surviving multiple-testing correction), and GO:0000183 was associated with TRD risk in the meta-analysis only. This GO gene set regulates chromatin silencing, and pathways related to the modulation of chromatin have been previously associated with antidepressant response in humans. This association may be mediated through the modulation of gene expression related to neurotransmission and neuroplasticity. Downregulation of histone deacetylase in the hippocampus was demonstrated to have antidepressant-like effect in mice through chromatin remodelling and consequent modulation of gene expression. Consistently, histone deacetylase (HDAC) inhibitors show antidepressant effects. SIRT2 (sin3 R2) was one of the top genes in GO:0000183, it codes for a class III NAD+ dependent HDAC that is oppositely regulated by stress and antidepressants. Very interestingly, SIRT2 inhibition is able to reverse anhedonia in different animal models by modulation of the glutamate and serotonin system in the prefrontal cortex.

**Limitations**

The present results should also be interpreted taking into account the limitations of this study. The power to detect associations at variant level was limited (odds ratios of at least ~1.60 were identifiable with adequate power), and this may explain the lack of genome-wide significant findings. The three samples of participants were recruited based on different protocols, hence they were heterogeneous for some clinical–demographic characteristics such as treatment and the available phenotypes were comparable but not exactly defined in the same way, partly explaining the different proportion of patients with TRD across samples.

**Implications**

In conclusion, available sample sizes are still limited to identify individual variants associated with TRD risk, but multimarker tests at gene-set level were again demonstrated to provide meaningful results. The gene sets reported by this study underlined the relevance of posttranscriptional regulation and chromatin remodelling as potential targets for the development of antidepressants with alternative mechanisms of action and potential benefit in TRD.

---

**Supplementary material**

Supplementary material is available online at https://doi.org/10.1192/bip.2018.25e.

**Funding**

STAR+D was supported by NIMH Contract No. N31MH00003 to the University of Texas Southwestern Medical Center. The ClinicalTrials.gov identifier is NCT0001528. The GENOMED project was supported by a European Commission Framework 6 grant (contract reference: LSHB-CT-2003-503428). The Medical Research Council, United Kingdom, and GlaxoSmithKline (GSK) provided support for genotyping. High performance computing facilities were funded with capital equipment grants from the GSTT Charity (TR130505) and Maudsley Charity (980). The collection of the GSRD (Group for the Study of Resistant Depression) sample analyzed in this study was supported by an unrestricted grant from Lundbeck. Lundbeck had no further role in the study design, in the collection, analysis, and interpretation of data, in the writing of the paper, and in the decision to submit the paper for publication. Dr Rudolf Uher is
Acknowledgements

We thank the NIMH for allowing the analyses of their data for the STAR*D participants. We also thank the authors of previous publications in this data-set, and foremost, we thank the patients and their families who enrolled in the study. Data and biomaterials were obtained from the limited access data-sets distributed from the NH-supported ‘Sequenced Treatment Alternatives to Relieve Depression’ (STAR*D). We thank Intromics (Copenhagen, Denmark) for genotype calling in the GSRD sample. All authors were actively involved in the design of the study, the analytical method of the study, the selection and review of all scientific content.

References


https://doi.org/10.1192/bjp.2018.256 Published online by Cambridge University Press