Epigenetic Signatures in Antidepressant Treatment Response: a Methylome-wide Association Study in the EMC Trial

Jan Engelmann
Lea Zillich
Josef Frank
Central Institute for Mental Health https://orcid.org/0000-0003-4867-9465
Stefanie Wagner
University Medical Center Mainz
Metin Cetin
David Herzog
Marianne Müller
University Medical Center, Johannes Gutenberg University Mainz https://orcid.org/0000-0002-0269-6131
André Tadic
Jerome Foo
Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg
Lea Sirignano
Central Institute of Mental Health https://orcid.org/0000-0002-7989-5833
Dieter Braus
Norbert Dahmen
Sabrina Sordon
Matthias Riemenschneider
Christian Spaniol
Gilles Gasparoni
Marcella Rietschel
University of Mannheim https://orcid.org/0000-0002-5236-6149
Stephanie Witt
University Medical Centre Mannheim https://orcid.org/0000-0002-1571-1468
Klaus Lieb
University Medical Center Mainz
Fabian Streit (✉ fabian.streit@zi-mannheim.de)
University of Heidelberg https://orcid.org/0000-0003-1080-4339

Article

Keywords:

Posted Date: January 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1253852/v1

Version of Record: A version of this preprint was published at Translational Psychiatry on July 7th, 2022. See the published version at https://doi.org/10.1038/s41398-022-02032-7.

Abstract

Although the currently available antidepressants are well established in the treatment of major depressive disorder (MDD), there is strong variability in the response of individual patients. Reliable predictors to guide treatment decisions before or in an early stage of treatment are needed. DNA-methylation has been proven a useful biomarker in different clinical conditions, but its importance for mechanisms of antidepressant response has not yet been determined. 80 MDD patients were selected out of >500 participants from the Early Medication Change (EMC) cohort with available genetic material based on their antidepressant response after four weeks and stratified into clear responders and age- and sex-matched non-responders (N=40, each). Early improvement after two weeks was analyzed as a secondary outcome. DNA-methylation was determined using the Illumina EPIC BeadChip. Epigenome-wide association studies were performed and differentially methylated regions (DMRs) identified using the comb-p algorithm. Enrichment was tested for hallmark gene-sets and in genome-wide association studies of depression and antidepressant response. No epigenome-wide significant differentially methylated positions were found for treatment response or early improvement. Twenty DMRs were associated with response; the strongest in an enhancer region in *SORBS2*, which has been related to cardiovascular diseases and type II diabetes. Another DMR was located in *CYP2C18*, a gene previously linked to antidepressant response. Results pointed towards differential methylation in genes associated with cardiac function, neuroticism, and depression. Linking differential methylation to antidepressant treatment response is an emerging topic and represents a step towards personalized medicine, potentially facilitating the prediction of patients' response before treatment.

Introduction

Major depressive disorder (MDD) is one of the most common, burdensome, and costly mental disorders worldwide [1,2]. Although currently available pharmacological treatments of MDD are well established and safe, there is a strong variability in antidepressant treatment response and considerable number of depressed patients do not respond to the first antidepressant administered, requiring optimization of antidepressant pharmacotherapy [3-5]. Due to the unpredictable treatment outcome, there is a vital need to identify reliable predictors of antidepressant response to guide treatment decisions. In clinical studies, early improvement, defined as a decrease in depressive symptomatology after two weeks, is considered the most consistent clinical predictor of antidepressant response [6]. However, research has not identified any clinical and biological predictor of sufficient clinical utility to inform the selection of a specific antidepressant agent for an individual depressed patient to date [7,8].

MDD is moderately heritable, with heritability estimates from twin studies ranging between 30% and 40% [9]. In a recent genome-wide association study (GWAS), which investigate the association of common genetic variants with depression, 102 independent genome—wide significant variants contributing to disorder risk were identified, and the phenotypic variance explained by all investigated SNPs (i.e. SNP-heritability) was estimated to be 8.9% [10]. During the last few decades, scientific knowledge about the genetic background of depression has increased steadily and pharmacogenetic

approaches have broadly been investigated to identify genetic variation contributing to individual treatment response in order to improve response prediction. A recent GWAS of antidepressant treatment response in 5,151 depressed patients did not yield any genome-wide significant finding, but showed that genetic variation explained around 13% of variance in the total meta-analysis and 20-40% of variance within each cohort [11]. It has to be noted that the included samples showed high heterogeneity, and because of the relatively small sample size compared to other GWAS in psychiatric genetics, analyses were not stratified for diagnosis, drug, or drug dosage.

In addition to genetic variation, epigenetic alterations, specifically DNA-methylation, i.e. the addition of a methyl-group to a cytosine nucleobase at the 5' position in CpG dinucleotides, can influence gene expression and may induce a wide range of changes at the cellular and systems function level [12,13]. The investigation of epigenetic variation represents a promising approach to not only investigate the biological mechanisms underlying depression, but especially the response to antidepressant treatments. Epigenome-wide association studies (EWAS) have investigated the association of DNA-methylation with depression [14-17]), as well as with antidepressant use [18]. Only one of the mentioned studies [17] yielded epigenome-wide significant results, although the two identified CpG-sites were not annotated to nearby genes, which makes functional interpretation difficult. Previous epigenetic investigation of antidepressant treatment response focused mainly on well-described candidate genes like *BDNF, NR3C1,* and *FKBP5* (reviews: [19-21]), but results are inconclusive. Additionally, candidate gene approaches are based on already existing hypotheses, and limited to a small selection out of the ~20,000 genes in the human genome. EWAS represents a promising approach to identify new biological mechanisms underlying individual differences in the context of antidepressant response, and identified methylation signatures might guide clinical decisions in the future, by predicting therapy outcomes.

So far, only two studies have investigated genome-wide differences in baseline DNA methylation between responders and non-responders to pharmacological antidepressant treatment. Ju and colleagues revealed several CpG-sites differentially methylated between responders and non-responders to eight weeks of escitalopram treatment, which were also associated with gene expression differences between both groups [22]. The study highlighted a differential methylated position (DMP) located in the *CHN2* gene, which was most significantly associated with mRNA expression and was replicable in an external cohort with a similar treatment [22]. The second study by Martinez-Pinteno et al. [23] identified 21 differentially methylated DMPs between responders and non-responders to eight weeks of fluoxetine treatment in a cohort of depressed children and adolescents. The *Ras Homolog Family Member J* (*RHOJ*) gene, encoding signaling molecules in the regulation of cytoskeletal organization, showed four significantly hypermethylated CpG-sites in non-responders [23]. Due to the inconsistent findings of the mentioned studies, further investigations of DNA methylation signatures between later responders and non-responders are needed.

The primary aim of this study was to identify epigenetic signatures associated with antidepressant treatment response, by testing baseline differential methylation between responders and non-responders after four weeks of antidepressant treatment. Patients were enrolled from a large well-characterized

antidepressant trial, allowing us to carefully select clear responders and sex- and age-matched non-responders. As a secondary aim, we explored the DNA-methylation signatures of early improvement in the same sample.

Patients And Methods

Sample

This investigation is a secondary analysis of 80 MDD patients, who have participated in the “Randomized clinical trial comparing an early medication change (EMC) strategy with treatment as usual (TAU) in patients with Major Depressive Disorder (MDD) – the EMC trial” (ClinicalTrials.gov NCT00974155). A total of 889 depressed patients were enrolled between 2009 and 2014 in this trial. Genetic material at baseline was available for 560 patients, of which the 40 most clear responders and sex- and aged-matched non-responders were selected. The selection was based on their treatment response, measured with the Hamilton Rating Scale for Depression – 17 items ($HAMD_{17}$) to antidepressant study medication after four weeks. In addition, the course of depression severity was considered in weekly intervals from baseline to week 4. The non-responders showed no improvement in depressive symptomatology despite four weeks of antidepressant treatment; in the group of responders, depressive symptomatology decreased steadily and patients showed a complete remission after four weeks (see also *Study Procedures*). Details of the EMC study protocol have been described previously [24-26] and the treatment algorithm can be openly accessed by *https://trialsjournal.biomedcentral.com/articles/10.1186/1745-6215-11-21*. In summary, the EMC trial was a multi-center, randomized, controlled clinical trial investigating whether patients with non-improvement after 14 days of escitalopram treatment take advantage to an early medication change (EMC: change to venlafaxine from day 14 onwards followed by an augmentation with lithium after again non-response at day 28) compared to patients treated according to current guideline recommendations (TAU: continuing escitalopram for two more weeks and switching to venlafaxine at day 28). Study procedures were approved by the local ethics committee of the Landesärztekammer Rheinland-Pfalz and are compliant with the Code of Ethics of the World Medical Association (Declaration of Helsinki) in its current version.

Study Procedures

Diagnoses were based on the German Version of the Mini International Neuropsychiatric Interview (M.I.N.I. ; [27]) and the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II;[28]). The socio-demographic and clinical characteristics, such as previously diagnosed cardiovascular or metabolic diseases, were assessed relying on patients’ self-reports. Depression severity was measured weekly from baseline to day 56 by the HAMD [29] by trained and blind raters [30]. Morning blood samples were obtained weekly. Antidepressant premedication was – if necessary – washed out after inclusion and before baseline visit. The antidepressant treatment according to study protocol was 20 mg escitalopram from baseline to day 14, followed by a predefined treatment algorithm. Other medications

were administered to treat depression-associated symptoms (e.g., insomnia) or adverse drug reactions (e.g., agitation or anxiety) with short-acting hypnotics (zolpidem or zopiclone), low potency antipsychotic drug pipamperone, histamine-receptor antagonist promethazine in standard doses as well as benzodiazepines in a dose-equivalent up to 15 mg diazepam per day was allowed. The main outcome parameters were: : a) response, defined as decrease of 50% during four weeks of treatment, and b) early improvement, defined as a decrease of depression severity of at least 20% from baseline to day 14 [6].

DNA-methylation

DNA was extracted from whole blood using the QIAamp DNA Blood Midi Kit from Qiagen (Qiagen, Hilden, Germany). The genomic DNA samples were stored at -20°C. Responders and non-responders were matched based on age and gender and the DNA from the matched samples were randomized and pipetted on processing plates. 500ng genomic DNA were bisulfite converted using the EZ-96 DNA methylation gold kit (Zymo research, Irvine, USA). Epigenome-wide methylation levels were determined using the Illumina HumanMethylationEPIC Beadchip and Illumina HiScan array scanning systems (Illumina, San Diego, CA).

Data preprocessing and quality control

The R statistical environment, version 3.6.1, was used for all data preprocessing and analysis steps. We used an updated version of the CPACOR-pipeline to extract methylation data from raw intensity data and performed quality control [31]. Thresholds for sample removal were: (i) DNA quality was not sufficient (missing rate > 0.10) or (ii) a discrepancy between methylation-based and phenotypic sex emerged. Thresholds for probe removal were: (i) the call-rate was insufficient (< 0.95), (ii) SNPs with a minor allele frequency > 0.10 were located in the probe sequence, (iii) the probes were located on the X or Y chromosome. After quality control all 80 samples remained. After filtering, 706,677 out of 843,232 sites were available for analysis.

Statistical Analysis

Differences in clinical and sociodemographic characteristics between responders and non-responders were calculated by t-tests for independent variables or Chi^2-tests, depending on the level of measurement.

Methylation values were log-transformed (base2) and included as dependent variables in the association analyses [32]. Principal component analysis was performed to extract signals of the internal control probes of the EPIC array and the resulting first ten principal components were included in all analyses to control for batch effects and technical quality. Additionally, the chip number and position on the chip were included. Cell-type heterogeneity was accounted for by estimating the cell counts based on the methylation data [33]. This approach results in six estimates, which roughly sum up to one. To avoid

multicollinearity in the EWAS, variance inflation factors were calculated for each cell count estimate. The estimated granulocyte count was subsequently removed from further analyses. For two participants, data on smoking was not available, and their smoking status was therefore estimated based on a validated set of sites [34].

Epigenome-wide association analysis. Tests of single site methylation differences between responders and non-responders were performed with linear models, adjusting for sex, age, smoking, standardized cell counts, and the first ten principal components of the internal control probes. Additionally, all analyses were run with early improvement after two weeks as a secondary outcome. Correction for multiple testing was applied using the Benjamini-Hochberg (FDR) correction and the resulting values are reported as *q*-values. CpG-sites were annotated using the manufacturer's manifest (http://webdata.illumina.com.s3-website-us-east-1.amazonaws.com/downloads/productfiles/methylationEPIC/infinium-methylationepic-v-1-0-b4-manifest-file-csv.zip; downloaded on 10th of August 2018).

Differentially methylated regions (DMRs). The comb-p algorithm was applied to identify DMRs. Comb-p accounts for autocorrelation between tests of adjacent methylation sites and combines these sites, in a given window, to regions of enrichment [35]. In the present study the settings were: Seed-p value < 0.01, minimum of 2 probes, sliding window 500 bp. Correction for multiple testing was applied using the Šidák correction as implemented in comb-p.

Gene-Set Enrichment Analysis. missMethyl [36] was used for functional analysis to test differentially methylated CpG-sites overrepresented in.hallmark gene-sets Sites with a threshold of $p_{nominal} < 0.001$ were included and the Hallmark gene-set collection (MSigDB Version 7.1), which consists of 50 gene-sets representing specific well-designed biological states or processes, was used as reference [37]. missMethyl controls for several potential confounders, such as probe number bias, which is the increased likelihood of a gene being differentially methylated, if more probes cover the gene, and multi-gene bias, since probes can be annotated to more than one gene.

GWAS-Enrichment-Analysis. Gene-sets consisted of the genes to which CpG-sites with an uncorrected p-value < 0.001 in the EWAS were annotated to. Two gene-sets were created, one for early improvement and one for treatment response. Gene-set enrichment was tested in results of the two recent genome-wide association studies described in the introduction: one of antidepressant treatment response ($N_{remission}$ = 1,852, $N_{non\text{-}remission}$ = 3,299) [11] and one of major depressive disorder including PGC and UKB samples (N_{cases} = 246,363, $N_{controls}$ = 561,190) [10]. This test was performed using Multi-marker Analysis of GenoMic Annotation (MAGMA) [38].

Results

The course of depression severity in responders and non-responders is shown in Figure 1. Mean age (±SD) was 41.5 (±11.1) years, 58% of patients were women and depression severity at baseline was 22.3

(±4.1) points ($HAMD_{17}$). As individuals were matched, there were no differences in age or sex distribution (all $p \geq 0.94$). For details of the clinical and sociodemographic characteristics see Table 1.

Table 1

Clinical and Sociodemographic Data. Continuous measures are presented as mean (standard deviation) and categorical measures as frequency (percent).

	Total (N=80)	Responders (N=40)	Non-Responders (N=40)	p-value (group comparison)
Age – yrs (SD)	41.55 (11.09)	41.45 (10.83)	41,65 (11.49)	0.94[a]
Female	46 (57.5%)	23 (57.5%)	23 (57.5%)	
Male – (%)	34 (42.5%)	17 (42.5%)	17 (42.5%)	1.0[b]
Age at onset - yrs (SD)	33.48 (12.02)	36.00 (12.13)	30.95 (11.50)	0.06[a]
Duration of current episode – wks (SD)	30.38 (41.94)	28.15 (40.87)	32.60 (43.39)	0.64[a]
1st episode	28 (35%)	19 (48%)	9 (23%)	
recurrent – (%)	52 (65%)	21 (52%)	31 (77%)	0.034[b]
Hamilton Scores				
Baseline (SD)	22.33 (4.15)	22.83 (3.88)	21.82 (4.4)	0.28[a]
Day 14 (SD)	12.69 (8.83)	4.57 (2.26)	20.8 (4.21)	0.00[a]
Day 28 (SD)	12.61 (10.50)	2.63 (1.76)	22.60 (3.97)	0.00[a]
Smokers – (%)				0.35[b]
yes	28 (36%)	16 (42%)	12 (30%)	
no	50 (64%)	22 (58%)	28 (70%)	
Cardiovascular Disease – (%)				0.0504[b]
yes	16 (20%)	4 (10%)	12 (30%)	
no	64 (80%)	36 (90%)	28 (70%)	
Metabolic Disease – (%)				0.735[b]
yes	10 (12.5%)	4 (10%)	6 (15%)	
no	70 (87.5%)	36 (90%)	34 (85%)	
Early Improvement – (%)				
yes	36 (45%)	39 (97.5%)	5 (12.5%)	
no	44 (55%)	1 (2.5%)	35 (87.5%)	

[a] t test; [b] χ^2 test; SD = standard deviation, wks = weeks; yrs = years

Epigenome-wide association study. No epigenome-wide significant differentially methylated positions emerged for either treatment response or early improvement after controlling for multiple testing. The strongest association with response was observed with hypermethylation of cg02107110 in *WDR47* (β =

0.17, $p = 9.59*10^{-7}$, $q = 0.57$). For early improvement, the strongest association was observed for cg04568295, which was annotated to *SIRT7* and *MAFG* (β = 0.11, $p = 1.58*10^{-6}$, $q = 0.86$). Regression coefficients for the 100 DMPs showing the strongest association can be found in Supplementary Table S1 for response and S2 for early improvement.

Differentially methylated regions (DMRs).

The DMR analysis identified twenty differentially methylated regions associated with treatment response and eleven with early improvement, Table 2 lists the DMRs and Figure 2 depicts the Manhattan plot of this analysis; DMRs are highlighted; results for early improvement are listed in Supplementary Table S3. The DMR showing the strongest association for both response to treatment and early improvement was annotated to *Sorbin And SH3 Domain Containing 2 (SORBS2),* a protein coding gene. The DMR consists of eight CpG-sites, hypermethylated in the responder group, seven of which are part of an enhancer region of *SORBS2*, pointing towards a potential functional relevance.

Table 2

Differentially Methylated Regions Associated with Response After Four Weeks of Treatment

Chr	Start	Ende	N probes	P	Sidak P	Gene	Direction
4	186732837	186733061	8	5.84E-16	1.75E-12	*SORBS2*	+
6	32016214	32016427	8	2.49E-12	8.26E-09	*TNXB*	+
1	174844397	174844561	5	1.19E-09	5.11E-06	*RABGAP1L*	+
8	1713005	1713013	3	2.74E-08	2.42E-03	*LOC101927752;CLN8*	+
10	96442621	96442675	3	4.59E-08	6.01E-04	*CYP2C18*	+
3	23244051	23244131	6	4.84E-08	4.28E-04	*UBE2E2-AS1;UBE2E2*	+
19	996220	996374	2	5.30E-08	2.43E-04		+
12	25801455	25801622	5	5.41E-08	2.29E-04	*LMNTD1*	+
22	50528213	50528299	4	6.20E-08	5.10E-04	*MOV10L1*	+
3	52099522	52099562	3	7.73E-08	1.37E-03	*LINC00696*	-
22	50585229	50585401	4	1.58E-07	6.49E-04	*MOV10L1*	+
1	108023366	108023487	5	2.65E-07	1.55E-03	*NTNG1*	+
3	48632568	48632724	4	5.57E-07	2.52E-03	*COL7A1*	+
2	128366514	128366595	2	5.90E-07	5.13E-03	*MYO7B*	+
3	49459909	49460112	6	8.09E-07	2.81E-03	*AMT;NICN1*	+
20	30225681	30225852	4	8.45E-07	3.49E-03	*COX4I2*	+
6	116886276	116886350	2	9.22E-07	8.76E-03		+
6	32121355	32121523	8	1.44E-06	6.05E-03	*PPT2;LOC100507547*	+
17	36997563	36997732	4	3.75E-06	1.56E-02	*C17orf98*	+
10	45719880	45720041	2	5.02E-06	2.18E-02		-

Chr = chromosome, - = hypomethylation of DMR in responders, + = hypermethylation of DMR in responders

Gene-Set Enrichment Analysis. Results from the EWAS were most strongly overrepresented in the Hallmark gene-sets “apical surface” (p = .001, q = .078) and “myogenesis“ (p = 0.006, q = 0.142), although none of the terms remained significant after multiple testing correction. Results of the gene-set enrichment analysis can be found in Supplementary Table S4 for response and S5 for early improvement.

GWAS-enrichment analysis. No significant enrichment of genes implicated by GWAS of MDD and antidepressant treatment response was observed (all $p \geq 0.12$). Detailed results are listed in Table 3.

Table 3

Results of GWAS-enrichment analyses.

Outcome	GWAS	N Genes	Beta	SE	P
Response	ADR	560	0.01	0.036	0.388
	Depression	557	0.027	0.041	0.259
Early Improvement	ADR	451	0.028	0.041	0.125
	Depression	448	0.06	0.049	0.12

ADR = antidepressant treatment response, SE = standard error.

Exploratory Analysis. Genetic Variation in *SORBS2* has repeatedly been associated with cardiovascular and metabolic diseases. Therefore, we performed follow-up analyses to investigate whether there was an association between previously diagnosed cardiovascular diseases, such as high blood pressure or arrhythmia, metabolic diseases, such as diabetes or obesity, and treatment response. Descriptively, patients who responded to antidepressant therapy were less likely to have a previous diagnosis of a cardiovascular disease, but a chi-square test was not significant ($X^2(1) = 3.83$, p = 0.0504). No association between metabolic diseases and treatment response was observed.

When we included previously diagnosed cardiovascular diseases in the EWAS regression model as a covariate, associations between treatment response and methylation of the DMR in *SORBS2* remained significant. Also, in a separate analysis, previously diagnosed cardiovascular diseases did not predict methylation in *SORBS2* (all p > 0.46).

Discussion

The aim of the present study was to identify differential methylation signatures before treatment initiation associated with antidepressant treatment outcome in 80 MDD patients, who were part of a large randomized controlled trial. We focused our analyses on antidepressant response after four weeks and

additionally investigated early improvement after two weeks of treatment. For both outcomes several differentially methylated regions at baseline were observed, which may point to possible underlying mechanisms of differential response to antidepressant pharmacotherapy.

While the epigenome-wide association study did not yield findings remaining significant after correction for multiple testing on the single site level, the region based analyses highlighted several CpG-sites as potentially relevant in antidepressant treatment response. The most strongly associated CpG-site for response was observed with hypermethylation of cg02107110 in the *WDR47* gene. WDR47 is a microtubule-associated protein and plays a role in neuronal regulation, brain development and brain connectivity [39]. Regarding early improvement, the strongest associated CpG-site was found in cg04568295. DNA-methylation at this site has been associated with HbA1c-levels in type 1 diabetes [40].

The first epigenome-wide association study of antidepressant response by Ju and colleagues identified three DMPs before treatment between later responders (N = 82) and non-responders (N = 95). One DMP located in the *CHN2* gene could be replicated in a second cohort receiving the same antidepressant treatment [22]. However, the three CpG-sites highlighted in *CHN2* were not available for analysis in the present study after quality control had been performed. It also has to be taken into account that the reported results of Ju and colleagues were not corrected for major drivers of differential methylation such as smoking [34]. The second study by Martinez-Pinteno and colleagues investigated baseline differences in DNA methylation between responders and non-responders (N = 11, each) and reported 21 significantly differential methylated CpG-sites associated with response to fluoxetine in adolescents. Within the two genes *RHOJ* and *OR2L13* (*Olfactory Receptor family 2 subfamily L member 13*), four and three DMPs were found between responders and non-responders [23]. These results were not replicated for antidepressant response in adults in the present study.

The significant DMRs were found in genes associated with a variety of domains, such as psychiatric, skeletal, immunological and metabolic traits. For example, genetic variation in *RABGAP1L* has been associated with the psychiatric traits ease of getting up in the morning [41], depressive affect [42] and neuroticism [43], but also metabolic traits such as BMI [44]. The strongest association was observed for a region in *SORBS2 (ARGBP2).* This DMR was differentially methylated between responders and non-responders, as well as between patients who showed an early improvement after two weeks of treatment, and those who did not. *SORBS2* is a protein-coding gene, which encodes the sorbin and SH3 domain containing 2 protein. Interestingly, the identified DMR is in an enhancer region, which provides evidence for a potential functional mechanism. Genetic variation in *SORBS2* has previously been implicated in cardiovascular diseases [45], type II diabetes [46], and educational attainment [47] in European ancestry populations. A pharmacogenetic GWAS suggested an association of *SORBS2* in response to lithium treatment [48] and in the recently published EWAS by Zhu and colleagues, *SORB2* was found as significant DMR associated with lifetime history of MDD in monozygotic discordant twins [15]. In addition, a review by Gharipour et al. highlighted *SORBS2* as one of three overlapping genes between mood disorders and obesity and formulated the hypothesis that hypermethylation in *SORBS2* might play a role in the co-occurence of both syndromes due to inflammation processes [49]. Based on the findings

on *SORBS2* in cardiovascular and metabolic diseases, we performed additional exploratory analysis, including previously diagnosed cardiac and metabolic comorbidities of our MDD patients. Even after taking the comorbidities into account, the association between response and differentially methylation of *SORBS2* remained significant. Mechanistically, the *SORBS2* splice variant *neural Abelson-related gene-binding protein 2* (*nArgBP2*) could be of particular interest, because it is specifically expressed in neurons. The nArgBP2 protein is enriched at dendritic spines where it acts as a cytoskeletal adaptor protein [50]. Due to the specificity of nArgBP2 to excitatory synaptic inputs, dysregulation results in an excitatory/inhibitory imbalance that could contribute to the disease course in mood disorders [51] and also to the response to antidepressants.

Another DMR in *CYP2C18* identified in the present study is of particular interest as *CYP2C18* belongs to the cytochrome P450 super family, which is involved in metabolism of many drugs, and genetic variation in this gene has previously been associated with escitalopram treatment response [52]. The pharmacogenetic study by Braten and colleagues investigated novel CYP2C-haplotypes to improve genetic prediction of escitalopram metabolism. The presence of the two SNPs (i.e., rs2860840 (C>T: CYP2C18, 3'UTR) and rs11188059 (G>A: CYP2C18, intron5) was associated with a significantly lower serum concentration of escitalopram [54].

None of the gene-sets were significantly enriched for genetic variation identified in recent GWAS. This could be because the cut-off of $p < 0.001$ resulted in a relatively large gene-set, but a restriction to genes implied by DMR analysis did not yield significant findings either. Also, the GWAS of antidepressant treatment response, which is the most relevant for the present analysis, is relatively small in comparison to other GWAS on psychiatric phenotypes and therefore limited in statistical power [11].

Overall, there is a wide correlation between the two investigated outcome parameters (response and early improvement) and the majority of differentially methylated CpG-sites were implicated in both outcomes. This can be simply explained from the patient sample investigated here in which 97.5% (N = 39) of responders at week 4 also showed an early improvement to antidepressant treatment at week 2. However, this is also in line with convergent evidence in literature from our own studies as well as a number of additional investigations that early improvement, defined as a 20% decrease of depressive symptomatology within the first two weeks, is one of the most consistent clinical predictors of later response to antidepressants [6, 53, 54], and that biological predictors of antidepressant treatment response may be identified already in the early course of treatment.

The major strength of our study is the selection of the investigated patients from the large well-characterized EMC trial, which enabled to choose clear responders and non-responders and to match the two groups according to several criteria, such as age and sex. In addition, we were able to control for potentially influencing factors such as smoking and to include concomitant cardiovascular and metabolic diseases in our analysis.

Several limitations apply to the present study. The sample size is relatively small compared to case-control EWAS of MDD and therefore our results need to be confirmed in larger well-characterized MDD

samples. Secondly, gene expression patterns could not be investigated in our sample, as the respective biomaterial is not available, limiting the possibility to draw conclusions about functional mechanisms. Thirdly, DNA methylation was assessed at baseline. While from a prediction perspective, it is important to identify pretreatment biomarkers of later therapy response, a longitudinal assessment could provide important insights into methylation changes associated with antidepressant treatment outcomes. Furthermore, methylation was assessed in peripheral blood samples and may potentially not reflect methylation in the brain of depressed patients.

In conclusion, we identified differential methylated regions associated with pharmacological antidepressant response in a well-characterized MDD study sample before treatment initiation. The DMR showing the strongest association was annotated to *SORBS2*, which has previously been described as an overlapping gene between mood disorders and obesity. *SORBS2* may therefore be a potential target gene enabling better understanding of mood disorders and additionally antidepressant treatment response, but confirmation in larger samples is needed. In summary, our results provide further evidence for the role of DNA methylation in patients' response to antidepressant treatment. Exploring DNA methylation in larger and clinically well-characterized samples may lead enable the stratification into different response subtypes.

Declarations

Author Contributions

SHW, MRiet, FS, KL, MRiem and MBM planned the investigation. JE and DPH performed the DNA extraction. SS, MRiem and CS performed the methylation analysis. SHW, MRiet, LZ were responsible for generating genome-wide methylation data. LZ, JF, JCF and FS developed the analysis plan. LZ and JF performed all statistical analyses. LZ, JE, SW, JCF, LS, MR, FS, DPH and SHW reviewed the literature for the paper. LZ, JE, FS and SHW drafted the manuscript. All authors contributed, revised, and edited the final manuscript critically. All authors agreed to the publication of the final version of the manuscript.

Funding and Disclosures

The EMC trial was funded by the German Federal Ministry for Education and Research (BMBF grant n°: 01 KG 0906); the herein presented additional investigations are not part of the funding. The analysis was funded through the ERA-NET NEURON project "EMBED-impact of Early life MetaBolic and psychosocial strEss on susceptibility to mental Disorders; from converging epigenetic signatures to novel targets for therapeutic intervention" [01EW1904]. The BMBF had no role in the conception of the study design, in the writing of the manuscript or the decision to submit the manuscript for publication. J. Engelmann is supported by the Mainz Research School of Translational Biomedicine (TransMed) with clinical scientist fellowship. A. Tadic is designated as inventor of the European patent number 12171541.1–2404 'Method for predicting response or non-response to a mono-aminergic antidepressant'. He has received during the last five years consultancy fees from Janssen, Novartis and ROVI. K. Lieb is designated as inventor of the European patent number 12171541.1–2404 'Method for predicting response or non-response to a mono-aminergic antidepressant'. D.F. Braus has received in the last 5 years fees for unrestricted educational

lectures from Lilly, Janssen, Bayer, Lundbeck, Servier, Shire, TAD Pharma, Lundbeck, Berlin-Chemie, Takeda, Rovi, Biogen, MSD. None of the authors has relevant financial interests in this manuscript.

Data Availability

Raw data and summary statistics for all analyses are available on request.

References

1. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2224–60.
2. Smith K. Mental health: a world of depression. Nature. 2014;515(7526):181.
3. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. Lancet. 2018;391(10128):1357–66.
4. Iniesta R, Malki K, Maier W, Rietschel M, Mors O, Hauser J, et al. Combining clinical variables to optimize prediction of antidepressant treatment outcomes. J Psychiatr Res. 2016;78:94–102.
5. Uher R, Perlis RH, Henigsberg N, Zobel A, Rietschel M, Mors O, et al. Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms. Psychol Med. 2012;42(5):967–80.
6. Stassen HH, Angst J, Hell D, Scharfetter C, Szegedi A. Is there a common resilience mechanism underlying antidepressant drug response? Evidence from 2848 patients. J Clin Psychiatry. 2007;68(8):1195–205.
7. Kessler RC. The potential of predictive analytics to provide clinical decision support in depression treatment planning. Curr Opin Psychiatry. 2018;31(1):32–39.
8. Kraus C, Kadriu B, Lanzenberger R, Zarate CA, Jr., Kasper S. Prognosis and Improved Outcomes in Major Depression: A Review. Focus (Am Psychiatr Publ). 2020;18(2):220–35.
9. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. Am J Psychiatry. 2000;157(10):1552–62.
10. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci. 2019;22(3):343–52.
11. Pain O, Hodgson K, Trubetskoy V, Ripke S, Marshe VS, Adams MJ, et al. Identifying the Common Genetic Basis of Antidepressant Response. Biol Psychiatry (Global Open Science). 2021.
12. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13(7):484–92.

13. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008;454(7205):766–70.
14. Story Jovanova O, Nedeljkovic I, Spieler D, Walker RM, Liu C, Luciano M, et al. DNA Methylation Signatures of Depressive Symptoms in Middle-aged and Elderly Persons: Meta-analysis of Multiethnic Epigenome-wide Studies. JAMA Psychiatry. 2018;75(9):949–59.
15. Zhu Y, Strachan E, Fowler E, Bacus T, Roy-Byrne P, Zhao J. Genome-wide profiling of DNA methylome and transcriptome in peripheral blood monocytes for major depression: A Monozygotic Discordant Twin Study. Transl Psychiatry. 2019;9(1):215.
16. Clark SL, Hattab MW, Chan RF, Shabalin AA, Han LKM, Zhao M, et al. A methylation study of long-term depression risk. Mol Psychiatry. 2020;25(6):1334–43.
17. Howard DM, Pain O, Arathimos R, Barbu MC, Amador C, Walker RM, et al. Methylome-wide association study of early life stressors and adult mental health. Hum Mol Genet. 2021.
18. Barbu MC, Huider F, Campbell A, Amador C, Adams MJ, Lynall ME, et al. Methylome-wide association study of antidepressant use in Generation Scotland and the Netherlands Twin Register implicates the innate immune system. Mol Psychiatry. 2021:2020.10.06.20207621.
19. Menke A, Binder EB. Epigenetic alterations in depression and antidepressant treatment. Dialogues Clin Neurosci. 2014;16(3):395–404.
20. Chen D, Meng L, Pei F, Zheng Y, Leng J. A review of DNA methylation in depression. J Clin Neurosci. 2017;43:39–46.
21. Webb LM, Phillips KE, Ho MC, Veldic M, Blacker CJ. The Relationship between DNA Methylation and Antidepressant Medications: A Systematic Review. Int J Mol Sci. 2020;21(3).
22. Ju C, Fiori LM, Belzeaux R, Theroux JF, Chen GG, Aouabed Z, et al. Integrated genome-wide methylation and expression analyses reveal functional predictors of response to antidepressants. Transl Psychiatry. 2019;9(1):254.
23. Martinez-Pinteño A, Rodriguez N, Blázquez A, Plana MT, Varela E, Gassó P, et al. DNA Methylation of Fluoxetine Response in Child and Adolescence: Preliminary Results. Pharmacogenomics and personalized medicine. 2021;14:459–67.
24. Tadic A, Gorbulev S, Dahmen N, Hiemke C, Braus DF, Roschke J, et al. Rationale and design of the randomised clinical trial comparing early medication change (EMC) strategy with treatment as usual (TAU) in patients with major depressive disorder--the EMC trial. Trials. 2010;11:21.
25. Tadic A, Wachtlin D, Berger M, Braus DF, van Calker D, Dahmen N, et al. Randomized controlled study of early medication change for non-improvers to antidepressant therapy in major depression–The EMC trial. Eur Neuropsychopharmacol. 2016;26(4):705–16.
26. Tadic A, Wagner S, Gorbulev S, Dahmen N, Hiemke C, Braus DF, et al. Peripheral blood and neuropsychological markers for the onset of action of antidepressant drugs in patients with Major Depressive Disorder. BMC Psychiatry. 2011;11:16.
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. Münster RD. Wittchen, H.-U., Zaudig, M. & Fydrich, T. (1997). SKID Strukturiertes Klinisches Interview für DSM-IV. Achse I und II. Göttingen: Hogrefe, DM 158,-. Hiller, W., Zaudig, M. & Mombour, W. (1997). IDCL Internationale Diagnosen Checklisten für DSM-IV und ICD-10. Göttingen: Hogrefe, DM 198,- bzw. DM 239. Z Klin Psychol Psychother (Gott). 1999;28(1):68-70.
29. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry. 1960;23:56–62.
30. Wagner S, Helmreich I, Lieb K, Tadic A. Standardized rater training for the Hamilton Depression Rating Scale (HAMD(1)(7)) and the Inventory of Depressive Symptoms (IDS(C30)). Psychopathology. 2011;44(1):68–70.
31. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, et al. Erratum to: A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. Genome Biol. 2016;17:73.
32. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinform. 2010;11:587.
33. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinform. 2012;13:86.
34. Zeilinger S, Kuhnel B, Klopp N, Baurecht H, Kleinschmidt A, Gieger C, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. PLoS ONE. 2013;8(5):e63812.
35. Pedersen BS, Schwartz DA, Yang IV, Kechris KJ. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. Bioinformatics. 2012;28(22):2986–88.
36. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. Bioinformatics. 2016;32(2):286–8.
37. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov Jill P, Tamayo P. The Molecular Signatures Database Hallmark Gene Set Collection. Cell Systems. 2015;1(6):417–25.
38. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol. 2015;11(4):e1004219.
39. Kannan M, Bayam E, Wagner C, Rinaldi B, Kretz PF, Tilly P, et al. WD40-repeat 47, a microtubule-associated protein, is essential for brain development and autophagy. Proc Natl Acad Sci U S A. 2017;114(44):E9308-E17.
40. Chen Z, Miao F, Braffett BH, Lachin JM, Zhang L, Wu X, et al. DNA methylation mediates development of HbA1c-associated complications in type 1 diabetes. Nat Metab. 2020;2(8):744–62.
41. Jansen PR, Watanabe K, Stringer S, Skene N, Bryois J, Hammerschlag AR, et al. Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. Nat Genet. 2019;51(3):394–403.
42. Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J, et al. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. Nat Genet. 2018;50(7):920–27.

43. Nagel M, Watanabe K, Stringer S, Posthuma D, van der Sluis S. Item-level analyses reveal genetic heterogeneity in neuroticism. Nat Commun. 2018;9(1):905.
44. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum Mol Genet. 2019;28(1):166–74.
45. Koyama S, Ito K, Terao C, Akiyama M, Horikoshi M, Momozawa Y, et al. Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease. Nat Genet. 2020;52(11):1169–77.
46. Vujkovic M, Keaton JM, Lynch JA, Miller DR, Zhou J, Tcheandjieu C, et al. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. Nat Genet. 2020;52(7):680–91.
47. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet. 2018;50(8):1112–21.
48. McCarthy MJ, Leckband SG, Kelsoe JR. Pharmacogenetics of lithium response in bipolar disorder. Pharmacogenomics. 2010;11(10):1439–65.
49. Gharipour M, Barekatain M, Sung J, Emami N, Sadeghian L, Dianatkhah M, et al. The Epigenetic Overlap between Obesity and Mood Disorders: A Systematic Review. Int J Mol Sci. 2020;21(18).
50. Lee SE, Kim Y, Han JK, Park H, Lee U, Na M, et al. nArgBP2 regulates excitatory synapse formation by controlling dendritic spine morphology. Proc Natl Acad Sci U S A. 2016;113(24):6749–54.
51. Lee SE, Kim JA, Chang S. nArgBP2-SAPAP-SHANK, the core postsynaptic triad associated with psychiatric disorders. Exp Mol Med. 2018;50(4):1–9.
52. Braten LS, Haslemo T, Jukic MM, Ivanov M, Ingelman-Sundberg M, Molden E, et al. A Novel CYP2C-Haplotype Associated With Ultrarapid Metabolism of Escitalopram. Clin Pharmacol Ther. 2021;110(3):786–93.
53. Wagner S, Engel A, Engelmann J, Herzog D, Dreimuller N, Muller MB, et al. Early improvement as a resilience signal predicting later remission to antidepressant treatment in patients with Major Depressive Disorder: Systematic review and meta-analysis. J Psychiatr Res. 2017;94:96–106.
54. Szegedi A, Jansen WT, van Willigenburg AP, van der Meulen E, Stassen HH, Thase ME. Early improvement in the first 2 weeks as a predictor of treatment outcome in patients with major depressive disorder: a meta-analysis including 6562 patients. J Clin Psychiatry. 2009;70(3):344–53.

Figures

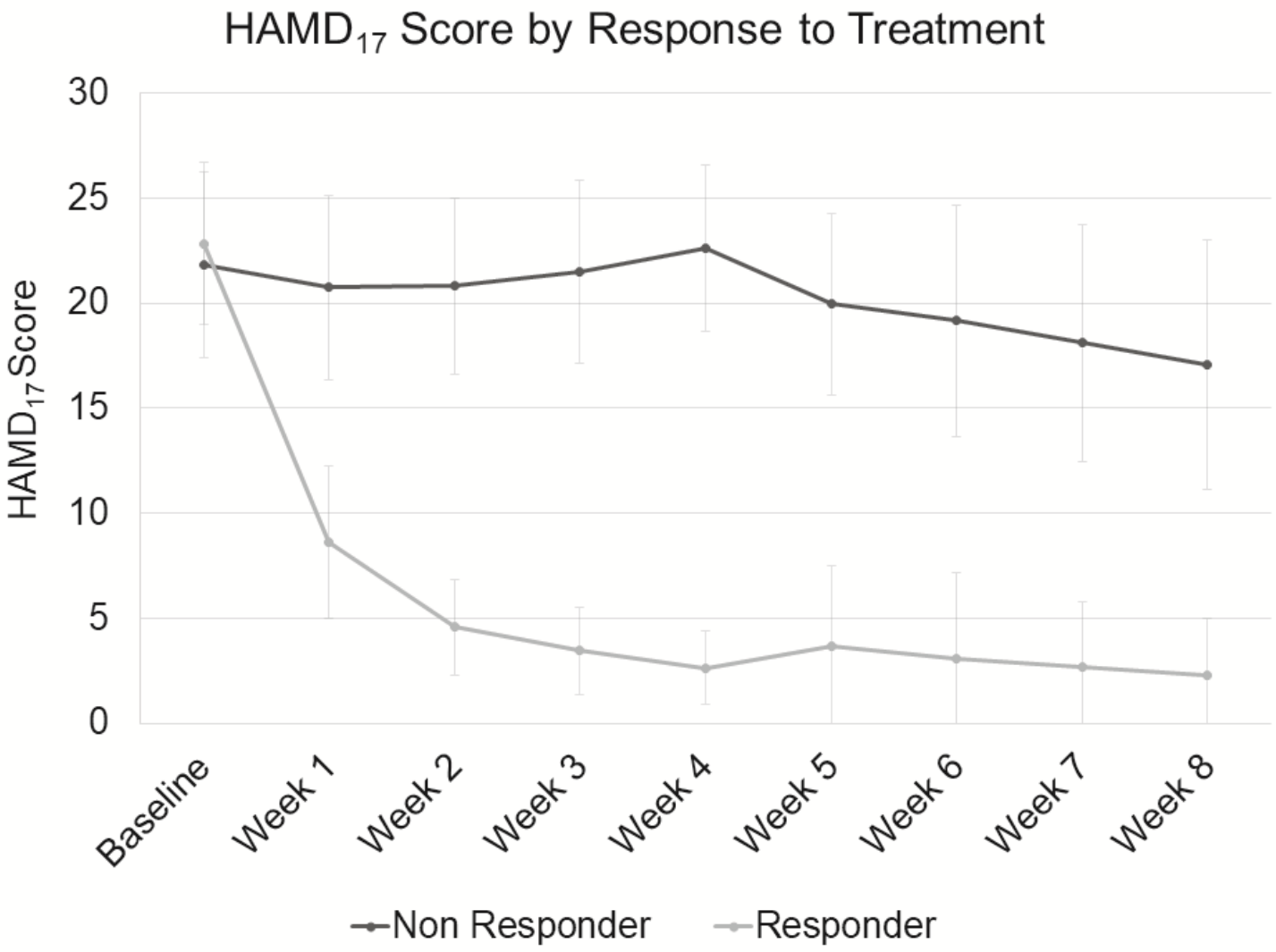

Figure 1

Mean Hamilton Rating Scale for Depression – 17 items ($HAMD_{17}$) Score by response to treatment, light gray represents responder, dark gray non-responder. Error bars represent standard deviations.

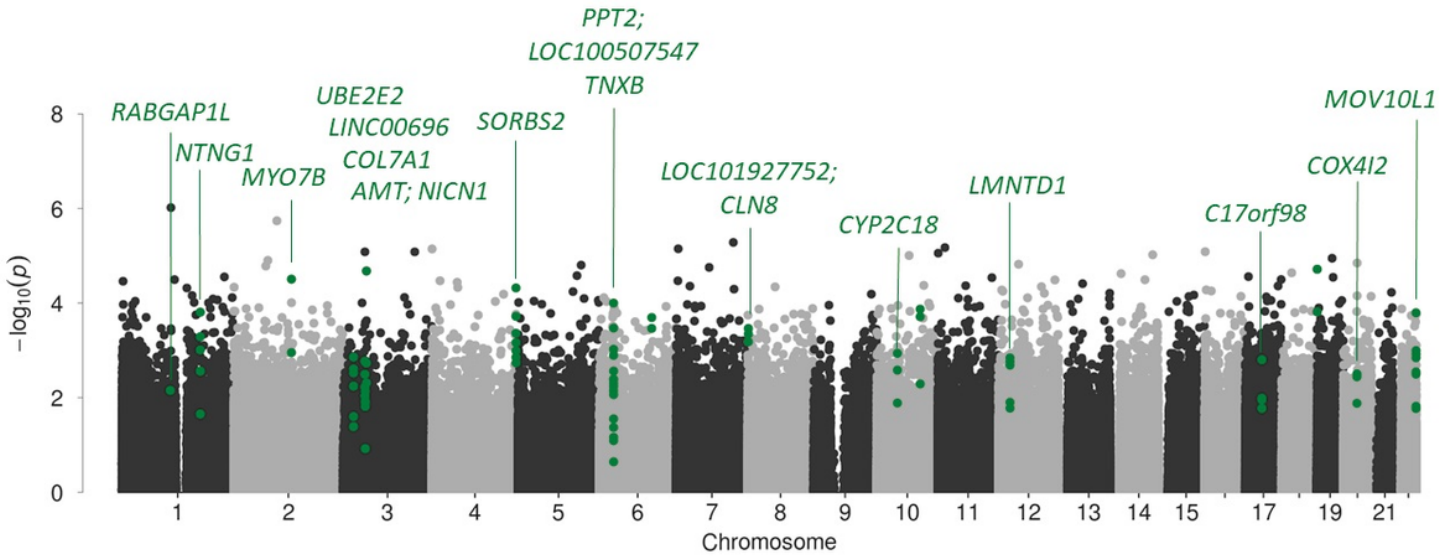

Figure 2

Differentially methylated CpG-sites and regions (green) associated with treatment response after four weeks.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- EMCMethylationSupplement.xlsx