Integrative genomic strategies applied to a lymphoblast cell line model reveal specific transcriptomic signatures associated with clozapine response.

Short title: in vitro genomic analyses of clozapine response.

Contributors: SAJ de With#1,2, APS On#1,3,4*, T Wang1, SL Pulit5, E Strengman1, J Viana6, J Mill7, S de Jong8, RA Ophoff1**,9,10.

Affiliations:

1UCLA Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, California, USA
2Brain Center Rudolf Magnus, Department of Psychiatry, University Medical Center Utrecht, The Netherlands
3Department of Genetics, University Medical Center Groningen, The Netherlands*
4Department of Psychiatry, University Medical Center Groningen, The Netherlands*
5Department of Medical Genetics, University Medical Center Utrecht, The Netherlands
6Institute of Psychiatry, SGDP Research Centre, King’s College London, United Kingdom
7University of Exeter Medical School, University of Exeter, Exeter, United Kingdom.
8MRC Social, Developmental and Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London, UK
9Department of Human Genetics, University of California Los Angeles, Los Angeles, CA, USA.
10Department of Psychiatry, Erasmus Medical Center, Erasmus University, Rotterdam, The Netherlands.

#These authors contributed equally
*Current affiliation

**Correspondence to RAO at ophoff@ucla.edu

UCLA Center for Neurobehavioral Genetics
Semel Institute for Neuroscience and Human Behavior
Gonda Center, room 4357B
695 Charles E. Young Drive South
Los Angeles, California 90095-176
United States of America
Abstract
Clozapine is an important antipsychotic drug. However, its use is often accompanied by metabolic adverse effects and, in rare instances, agranulocytosis. The molecular mechanisms underlying these adverse events are unclear. To gain more insights into the response to clozapine at the molecular level, we exposed lymphoblastoid cell lines (LCLs) to increasing concentrations of clozapine and measured genome-wide gene expression and DNA methylation profiles. We observed robust and significant changes in gene expression levels due to clozapine (n = 463 genes at FDR < 0.05) affecting cholesterol and cell cycle pathways. At the level of DNA methylation, we find significant changes upstream of the LDL receptor, in addition to global enrichments of regulatory, immune and developmental pathways. By integrating these data with human tissue gene expression levels obtained from the Genotype-Tissue Expression project (GTEx), we identified specific tissues, including liver and several tissues involved in immune, endocrine and metabolic functions, that clozapine treatment may disproportionally affect. Notably, differentially expressed genes were not enriched for genome-wide disease risk of schizophrenia or for known psychotropic drug targets. However, we did observe a nominally significant association of genetic signals related to total cholesterol and low-density lipoprotein levels. Together, these results shed light on the biological mechanisms through which clozapine functions. The observed associations with cholesterol pathways, its genetic architecture and specific tissue effects may be indicative of the metabolic adverse effects observed in clozapine users. LCLs may thus serve as a useful tool to study these molecular mechanisms further.
Introduction

Antipsychotic drugs (APs) play an important role in the treatment of psychotic disorders such as schizophrenia (SCZ). Clozapine is one of the most effective antipsychotic drugs (AP) (Kane et al., 1988; Leucht et al., 2013; Taylor, 2017). However, the decision to prescribe clozapine is complicated by its potential to induce severe adverse effects (Leucht et al., 2013). The most severe adverse effect, with a prevalence of <1%, is clozapine-induced agranulocytosis, a dramatic reduction of white blood cells (Andersohn et al., 2007). More common adverse effects include weight gain, dyslipidemia and type 2 diabetes. These adverse metabolic effects are, addition to the chance of developing agranulocytosis, the primary reasons for patient noncompliance and discontinuation of treatment (Cohen, 2014; Weiden et al., 2004).

The biological mechanisms underpinning the effect of clozapine, as well as its adverse effects, remain elusive. A twin study estimated that the heritability of APs-induced weight gain is approximately 60% (Gebhardt et al., 2010), suggesting a substantial role for genetic factors. Candidate genes studies of clozapine-induced adverse effects have yielded ambiguous results and lack consistent replication (reviewed by Chowdhury et al., 2011; Lett et al., 2012; Müller et al., 2013; Roerig et al., 2011; Yan et al., 2013). Two genome-wide association studies (GWAS) investigating antipsychotic-induced metabolic adverse effects have yielded inconclusive findings, primarily due to insufficient sample sizes and the potential polygenic nature of this trait (Adkins et al., 2011; Malhotra et al., 2012).

Intermediate molecular phenotypes, such as gene expression studies in specific cell lines or tissues, may improve our understanding of the molecular function of clozapine. Previous studies have found that atypical antipsychotic drugs may induce cholesterol metabolism through transcription factors such as sterol regulatory element binding proteins (SREBP1 and 2) (Ferno et al., 2011), suggesting that drug-induced cholesterol metabolism is related to these adverse
metabolic effects. However, such findings were not consistently replicated when profiling whole blood (Ferno et al., 2011; Harrison et al., 2016; Vik-Mo et al., 2008). It is possible that changes in DNA methylation could mediate changes in gene expression, such as the AP-induced hypomethylation of the FAR2 gene leading to insulin resistance (Burghardt et al., 2016). However, such studies are currently limited and those performed provide inconsistent results (Burghardt et al., 2016; Houtepen et al., 2016; Kinoshita et al., 2017; Melas et al., 2012; Ota et al., 2014; Rukova et al., 2014; Stapel et al., 2017; Swathy et al., 2017; Swathy and Banerjee, 2017).

A major obstacle towards understanding clozapine-induced metabolic effects is that clozapine therapy is a relatively rare (~6%) treatment plan in patients diagnosed with schizophrenia (Burghardt et al., 2016; Stroup et al., 2016), of which then only 1% develop CIA. Such factors challenge our ability to adequately sample a large and controlled prospective cohort of patients thereby limiting progress in understanding both metabolic and hematological adverse effects. To augment the lack of available in vivo data, we implemented an in vitro lymphoblast cell line (LCL) model to study the effects of drug exposure at the molecular level. Cell-based models have been successfully employed for pharmacogenomic studies, including LCL models to study clozapine function (de With et al., 2015; Morag et al., 2010; Welsh et al., 2009; Wen et al., 2012). Here, we exposed LCLs to increasing doses of clozapine and collected both expression and methylation profiles. Through integrative genomic analyses, we aim to extrapolate in vitro molecular signatures of clozapine towards relevance of in vivo function and study clozapine response without the need to assemble a large cohort of patients. We identified significant changes in transcriptomic signatures associated with cholesterol and cell cycle pathways. By then integrating these molecular profiles with genome-wide association study (GWAS) summary statistics of different traits and diseases, we show that clozapine-associated genes also overlap with genetic signals related to total cholesterol and low-density lipoprotein (LDL) levels but not
schizophrenia genetic risk. Clozapine-associated genes are furthermore related to specific human tissues, such as liver and those involved in immune, endocrine and metabolic functioning.

**Material and Methods**

**Lymphoblast cell lines**

We used lymphoblast cell lines (LCLs) from four unrelated samples, all part of the collection of Utah residents of Northern and Western European ancestry (HapMap CEPH/CEU phase 1)(Consortium and †The International HapMap Consortium, 2003). We obtained LCLs from the Coriell Institute for Medical Research (Camden, NJ, USA) and maintained the cell lines as previously described(Consortium and †The International HapMap Consortium, 2003; de With et al., 2015). To study methylation changes after exposure to clozapine, we performed a separate experiment using six LCLs (HapMap, CEPH/CEU phase 1) consisting of two parent-offspring trios.

**In vitro experimental design and clozapine exposure**

We exposed LCLs to different clozapine concentrations, with clinical concentration set at 2µM(Baumann et al., 2004). Clozapine, purchased from Sigma Aldrich, was dissolved in culture medium with dimethyl sulfoxide (DMSO), with a maximum concentration of 0.1%. Cell lines were exposed for 24 hours to clinical concentration (Supplementary Methods), 10x, 50x and 100x clinical concentration (20µM-100 µM-200 µM clozapine) and vehicle (DMSO); each concentration was measured in 4 cell lines, after which RNA was obtained for gene expression analysis (Supplemental Figure 1A). To study DNA methylation changes in response to clozapine, we performed an experiment similar to the gene expression study; LCLs were exposed to different concentrations of clozapine (vehicle (DMSO), 1x, 20x, 40x and 60 times clinical concentration) and exposure times were 24h and 96h (Supplemental Figure 1B). We measured cell viability using the TC10 automated cell counter.
Sample processing and data collection

After desired exposure time, we lysed cells and performed RNA and DNA collections using column-based extraction methods from Qiagen according to manufacturer’s instructions (Supplementary Methods). Gene expression profiling was carried out using Illumina® HumanHT-12 v4 Expression BeadChip technology. DNA methylation assays were performed with Illumina® Infinium HumanMethylation450 Beadchip arrays.

Data preprocessing and normalization

We processed raw gene expression values using the “Lumi” R-package (Du et al., 2008). We log2 transformed and quantile normalized the raw data, keeping only expressed gene transcripts (detection p < 0.01) for further analysis (22,926 probes). We processed DNA methylation values using the “WateRmelon” Bioconductor package (Pidsley et al., 2013), removing probes that were known to cross-hybridize, probes containing SNPs in target CpG regions, probes with detection p-value greater than 0.01 in 5% of samples, and probes with beadcounts > 3 (n=86,068 probes in total) (Chen et al., 2013; Price et al., 2013). We normalized the data using the dasen function and computed β-values, defined as the ratio of the methylated probe intensity and the overall intensity (sum of methylated and unmethylated probe intensities), to measure methylation levels. To limit the effect of heteroskedasticity, we included only variable probes with β-values between 0.2 – 0.8 in our analyses (165,014 probes) (Du et al., 2010).

Statistical analyses

To detect clozapine-induced molecular changes, for each probe, we tested for association between gene expression (1) or DNA methylation levels (2) with increasing clozapine concentrations using the following linear models implemented in R using the “Limma” package (Smyth, n.d.):
where Y is the normalized gene expression (g) or DNA methylation levels (m) for an individual probe, $b_0$ the intercept, $b_1$ the effect of clozapine concentration, $b_2$ the effect of RIN, and $\epsilon$ the residual variation of the model. We ran each model for each probe per individual i and subsequently performed a meta-analysis across all individuals by combining p-values using Stouffer’s method with directionality of the effect sizes taken into account. We included RNA-integrity number (RIN) as a covariate in the gene expression model and applied a Bonferroni correction to correct for multiple testing, resulting in a significance threshold of $p < 2.18 \times 10^{-6}$ for the gene expression analysis ($n = 22,926$ probes) and $p < 3.03 \times 10^{-7}$ for the DNA methylation analysis ($n = 165,014$ probes).

**Gene ontology analysis**

We performed functional gene ontology analysis using DAVID (Database for Annotation, Visualization and Integrated Discovery, version 6.8, interrogated February 2018) (Huang et al., 2009a, 2009b), with default settings (Supplementary Methods).

**Functional enrichment analysis of DNA methylation data**

We used Genomic Regions Enrichment of Annotations Tools (GREAT, v3.0) to predict the biological function of the top methylation probes associated with clozapine exposure. GREAT links both proximal and distal genomic CpG sites with their putative target genes and implements both a gene-based test and a region-based test using the hypergeometric and binomial test, respectively, which allows us to assess enrichment of genomic regions in biological annotations of pathway databases (Supplementary Methods) (McLean et al., 2010). The statistical outputs of GREAT for both gene-based and region-based tests were subsequently adjusted for multiple testing using Bonferroni correction.
Weighted gene co-expression network analysis

We performed a gene expression network analysis using weighted gene co-expression network analysis (WGCNA) in R. Briefly, WCGNA identifies distinct modules using the shared variation in gene expression based on pairwise correlation. To account for the biases related to differing probe numbers between genes assayed on the array, we provided as input the mean probe expression of genes residing within nominally significant differentially expressed genes (p<0.05), considering 5,708 probes within 4,897 genes (Horvath, 2011; Langfelder and Horvath, 2008; Zhang and Horvath, 2005). To assign biological function to each WCGNA module, we performed gene ontology analysis using DAVID we performed.

Additional DNA methylation analyses

We ran a candidate gene study for CpG sites in close proximity of genes with evidence of clozapine-induced differential gene expression. Methylation probes within the gene body, in the 3’ and 5’ untranslated regions and up to 1,500 nucleotides upstream of the transcription start site of the 463 ‘top genes’ (p<0.05) were selected for a post-hoc analysis (n = 1,004). These results can be found in Table 1 of the Supplemental Material.

GTEx cross tissue analysis

To translate the in vitro effects of clozapine to in vivo human biology, we investigated how clozapine genes behave across 22 human tissues represented in the GTEx data set. We downloaded gene level quantifications (version 6, date: April 24, 2019) from the GTEx Project web portal (Melé et al., 2015) and transformed gene expression values using a log2 transformation (1 + RPKM value). We then tested 1) if clozapine-associated genes have higher or lower average expression within each tissue and 2) if between tissue distance is different for clozapine-associated genes compared to the expected based on chance. To calculate between tissue distance, we used the top half most variable genes across GTEx samples that were also significantly detected in our in vitro assay (n=7,025). We then performed multidimensional scaling...
using the isoMDS() function in the MASS R package (v7.3) (Melé et al., 2015; Venables and Ripley, 2002) and calculated between tissue Euclidean distances using the dist() function in R (v.3.3.3) and the median gene expression values for each tissue for clozapine associated genes (N=463). To establish whether clozapine-associated genes significantly deviated from the expected, we established a null distribution by repeating this procedure using the Euclidean distance for samplings of similarly-sized sets of genes. Within each sampling, genes were sampled with a probability that matched the distribution of average gene expression of clozapine-associated genes in our LCL experiment. This was done to account for differentially expressed genes having higher gene expression levels. Our null distribution therefore better captured the expected effect. Null distributions were built separately for the expected average gene expression within each tissue and for the expected Euclid distance for each tissue pair. P-values were then established by calculating the proportion of samplings that were higher or lower than the statistic of the observed clozapine-associated genes and corrected for multiple testing by Bonferroni correction. Analyses were conducted separately for genes upregulated and downregulated after clozapine exposure.

**MAGMA gene-set analysis**

To investigate if clozapine-associated genes were enriched for schizophrenia or cardiovascular-related genetic association signals, we performed gene-set analysis using MAGMA (multimarker analysis of genomic annotation) (de Leeuw et al., 2015). We tested the following three gene-sets, each with three different significance thresholds:

1. All genes differentially expressed, either up- or downregulated;
2. Upregulated differentially expressed genes;
3. Downregulated differentially expressed genes.

We obtained GWAS summary statistics for schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014), body mass index (Yengo et al., n.d.), coronary artery
disease (Nikpay et al., 2015), type 2 diabetes (Scott et al., 2017), total cholesterol, triglycerides, high-density-lipoprotein, and low-density-lipoprotein (Scott et al., 2017; Surakka et al., 2015). For each GWAS, we computed aggregate gene-level test statistics using a 10kb window around the transcription start and end site of each gene; a total of 11,533 genes were available for analysis. We then tested for association between GWAS gene level Z-scores and experimental gene-sets (defined above). We estimated linkage disequilibrium (LD) using the 1000 Genomes European reference panel (1000 Genomes Project Consortium et al., 2015; Scott et al., 2017; Surakka et al., 2015) and ran MAGMA using a two-sided competitive test by accounting for probes tested in our experiment while also correcting for gene size, SNP density, minor allele count and LD.

**Antipsychotic drug gene-sets and enrichment of schizophrenia heritability**

As a complementary analysis, we investigated whether manually curated antipsychotic drug gene-sets overall are associated with schizophrenia heritability. We performed gene-set analysis on antipsychotic drug targets that were previously reported to be enriched for schizophrenia heritability [55]. Drug target list were curated using drug-gene interaction databases (Roth et al., 2000; Wagner et al., 2016). We identified 50 sets of drugs with ATC (Anatomical Therapeutic Chemical) code N05A, a drug class to which all antipsychotic drugs belong, including clozapine. Using the schizophrenia GWAS as input to MAGMA, we ran gene-set analysis (1) per individual N05A antipsychotic drug gene-set, (2) for all N05A drug target genes combined, and for (3) all N05A drug target genes combined excluding clozapine.

**LD Score Regression estimation of heritability and genetic correlations**

Pre-computed LD scores from the 1000 Genomes European reference panel provided through the LD Score Regression (LDSR) github and GWAS summary statistics file processed and reformatted to sumstats format were used as input to (LDSR) to estimate SNP-based heritability (SNP-\(h^2\)) (B. K. Bulik-Sullivan et al., 2015). Genetic correlations between traits were estimated using cross-trait LDSR via the --rg flag (B. Bulik-Sullivan et al., 2015).
Results

Clozapine exposure induces widespread gene expression changes

We exposed lymphoblast cell lines to different concentrations of clozapine and tested for subsequent changes in gene expression levels (Supplemental Figure 1A). In total, we tested 22,926 gene expression probes, of which 5,708 showed nominal significance (p<0.05) and 518 probes exceeded a Bonferroni-corrected p < 2.18*10^-6. These 518 probes consisted of 234 up-regulated probes and 284 down-regulated probes, representing a total of 463 unique genes. The top 10 up-regulated and down-regulated probes are shown in Table 1. See Supplement for the complete list of differentially expressed genes.

<table>
<thead>
<tr>
<th>Up-regulated genes</th>
<th>P-value</th>
<th>Down-regulated genes</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>STMN1 (Stathmin 1)</td>
<td>5.16*10^-16</td>
<td>LSS (Lanosterol synthase)</td>
<td>6.14*10^-17</td>
</tr>
<tr>
<td>HIST2H2AC (Histone cluster 2 H2A family member C)</td>
<td>1.98*10^-15</td>
<td>MAL (T-cell differentiation protein)</td>
<td>1.10*10^-16</td>
</tr>
<tr>
<td>RPS7 (Ribosomal protein S7)</td>
<td>2.52*10^-15</td>
<td>MIR1974 (MicroRNA 1974)</td>
<td>5.56*10^-16</td>
</tr>
<tr>
<td>HIST1H2BJ (Histone cluster 1 H2B family member J)</td>
<td>4.42*10^-15</td>
<td>LDLR (Low density lipoprotein receptor)</td>
<td>3.97*10^-14</td>
</tr>
<tr>
<td>HIST2H2AA3 (Histone cluster 1 H2A family member A3)</td>
<td>1.97*10^-14</td>
<td>DHCR7 (7-dehydrocholesterol reductase)</td>
<td>4.26*10^-14</td>
</tr>
<tr>
<td>RPS15 (Ribosomal protein S15)</td>
<td>2.17*10^-14</td>
<td>PASK (PAS Domain Kinase)</td>
<td>5.09*10^-14</td>
</tr>
<tr>
<td>HIST2H2AA4 (Histone cluster 2 H2A family member A4)</td>
<td>2.78*10^-14</td>
<td>RGS1</td>
<td>2.31*10^-13</td>
</tr>
<tr>
<td>HIST1H2AC (Histone cluster 1 H2A family member C)</td>
<td>2.79*10^-14</td>
<td>LYPD6B (Regular Of G Protein Signaling 1)</td>
<td>4.38*10^-13</td>
</tr>
<tr>
<td>AURKA (Aurora Kinase A)</td>
<td>5.19*10^-14</td>
<td>TPP1 (LY6/PLAUR Domain Containing 6B)</td>
<td>6.77*10^-13</td>
</tr>
<tr>
<td>SGOL1 (Shugoshin 1)</td>
<td>2.09*10^-13</td>
<td>RENBP (Tripeptidyl Peptidase 1)</td>
<td>7.54*10^-13</td>
</tr>
</tbody>
</table>

Table 1. Top 10 up- and downregulated genes after clozapine exposure. Gene symbols are shown with corresponding gene name and p-value of differential gene expression analysis. A full list of genes is in Supplementary Table 1.

Gene ontology enrichment analysis of up-regulated transcripts showed significant functional enrichment for cholesterol metabolism (13 genes, p = 4.15*10^-15) and steroid biosynthesis (11 genes, p = 1.01*10^-8) (Table 2), while analysis of down-regulated transcripts were enriched for cell
division processes and related annotations, such as mitosis (44 genes, \( p = 1.87 \times 10^{-39} \)), chromosome (49 genes \( p = 3.03 \times 10^{-35} \)) and nucleosome (16 genes, \( p = 3.12 \times 10^{-14} \)) and other cell cycle pathways (Table 2).

Gene expression network analysis after clozapine exposure

WGCNA network analysis of clozapine-induced differentially expressed genes (N=4,987) yielded 15 co-expression modules, ranging in size from n=61 to 1,791 genes. Five gene co-expression modules were altered upon clozapine exposure: M14 (upregulated, 61 genes), M10 (upregulated, 155 genes), M9 (upregulated, 158 genes), M3 (upregulated, 579) and M1 (downregulated, 1,791 genes), which were nominally significant. The M14 co-expression module was enriched for genes involved in cholesterol metabolism (13 genes, \( p = 4.8 \times 10^{-15} \)), the M1 co-expression module was enriched for genes involved in cell cycle (133 genes, \( p = 7.3 \times 10^{-32} \)) and the M9 and M3 co-expression module was enriched for mitochondrial genes (15 genes, \( p = 6.7 \times 10^{-6} \) and 46 genes \( p = 3.8 \times 10^{-7} \) respectively). The M10 co-expression module was enriched for genes involved in the nucleosome (9 genes, \( p = 1.4 \times 10^{-7} \)).

Minimal changes in DNA methylation at single CpG sites after clozapine exposure

We then performed DNA methylation profiling to assess whether these effects were due to epigenetic changes. For statistical analysis, we applied an analytical approach similar to our gene expression analyses. Targeted analysis on the 1,004 CpG sites near the 463 differentially-expressed genes revealed 3 probes exhibiting significant changes in DNA methylation (\( p < 1.08 \times 10^{-4} \)), including a probe upstream of the low-density lipoprotein receptor (LDL-R) gene (cg2971501, \( p=4.75 \times 10^{-5} \)) after 24h; one probe upstream of the cyclin F (CCNF) gene showed a significant change in DNA methylation after 96h (Table S3). Beyond these examples, global methylation differences were not observed after 24h or 96h at the level of individual probes.
<table>
<thead>
<tr>
<th>Functional annotation cluster</th>
<th>Enrichment score</th>
<th>Best P-value</th>
<th>Best fold enrichment</th>
<th>Average P-value</th>
<th>Average fold enrichment</th>
<th>Number of pathways</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upregulated pathways</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cholesterol/lipid/steroid biosynthesis</td>
<td>8.44</td>
<td>4.15 x 10^-15</td>
<td>46.92</td>
<td>1.19 x 10^-3</td>
<td>18.46</td>
<td>19</td>
</tr>
<tr>
<td>Cholesterol/Steroid biosynthesis</td>
<td>4.77</td>
<td>1.01 x 10^-8</td>
<td>68.45</td>
<td>4.66 x 10^-4</td>
<td>54.32</td>
<td>3</td>
</tr>
<tr>
<td><strong>Down-regulated pathways</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mitosis/cell division</td>
<td>31.26</td>
<td>1.87 x 10^-37</td>
<td>14.77</td>
<td>1.12 x 10^-21</td>
<td>10.69</td>
<td>5</td>
</tr>
<tr>
<td>Chromosome/centromere</td>
<td>16.47</td>
<td>3.03 x 10^-35</td>
<td>21.32</td>
<td>7.76 x 10^-8</td>
<td>15.79</td>
<td>9</td>
</tr>
<tr>
<td>Histone/nucleosome</td>
<td>6.93</td>
<td>3.12 x 10^-14</td>
<td>20.72</td>
<td>1.73 x 10^-3</td>
<td>9.13</td>
<td>18</td>
</tr>
<tr>
<td>Spindle</td>
<td>6.02</td>
<td>4.95 x 10^-12</td>
<td>12.89</td>
<td>7.45 x 10^-4</td>
<td>10.13</td>
<td>6</td>
</tr>
<tr>
<td>Nucleotide/ATP-binding</td>
<td>5.72</td>
<td>1.46 x 10^-13</td>
<td>2.71</td>
<td>7.78 x 10^-5</td>
<td>2.47</td>
<td>5</td>
</tr>
<tr>
<td>Microtubule/kinesin</td>
<td>4.60</td>
<td>1.90 x 10^-4</td>
<td>16.92</td>
<td>4.99 x 10^-2</td>
<td>8.56</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 2. Main functional annotations associated with clozapine exposure based on gene expression. Output of pathway enrichment analyses using DAVID is shown. The enrichment score is used as the main metric of importance. It is defined as the geometric mean of all enrichment p-values of each annotation term within the group. It is expressed as the minus log of the p-value, an enrichment score of 1.3 is nominally significant. This table shows clusters with an enrichment score > 4.6, corresponding to a p-value < 0.01. Fold enrichment is a measure to express the enrichment of this particular group of genes in comparison with the genes in the human genome. A list of all pathways is available in supplemental information.
DNA methylation is affected at the pathway level

To investigate if our top associated DNAm probes aggregated to changes at the pathway level, we performed pathway enrichment analysis using GREAT, which incorporates functional annotation from various databases to predict cis-regulatory function of genomic regions of interest. When considering probes exhibiting FDR < 10%, we observed enrichment for protein binding and regulation of cellular processes after 24h of clozapine treatment. When considering the top 1000 probes, we also observed enrichment of immune-related functions, such as the Major Histocompatibility Complex (MHC) class II protein complex and Graft-versus-host disease after 24h. We found significant enrichment for estradiol regulation and various embryonic developmental processes (Figure 1 and Table S5) when considering the top 1000 probes after 96h.

Figure 1. Clozapine-associated DNAm probes concentrate to specific biological annotations. Genomic Regions Enrichment of Annotations Tool (GREAT) was used to assign biological context to genomic regions associated to clozapine exposure. GREAT uses the annotation of nearby genes and regulatory elements to performed enrichment analysis across known functional databases. The x-axis shows the -log 10 p-value of the region-based analysis (binomial test), while the y-axis shows the enrichment for the gene-based analysis (hypergeometric overlap test). Results are shown for (A) the top probes after 24 hours (FDR < 0.10, n=177), (B) the top 1,000 probes after 24 hours, and (C) the top 1,000 probes after 96 hours. Each dot represents an annotation. Significant annotations, after Bonferroni correction, are color-coded according to the test used.
DNA methylation changes are time-dependent

To examine if clozapine affected each time point similarly, we correlated effect sizes at each CpG probe between time points. Across all probes, we found a significant negative correlation between the Z-score of the association between DNA methylation levels and clozapine exposure at 24h and 96h (Pearson r = -0.32, P<2.2x10^{-16}). This correlation was preserved among probes at FDR < 10% at 24h (n=177, Pearson r = -0.38, P=1.3x10^{-7}) but not among the top probes at 96 hours (n=177, Pearson r = -0.12, P = 0.10), highlighting possible time-dependent effects after clozapine exposure (Figure 2).

Figure 2. Clozapine-induced DNA methylation changes are time-dependent. Correlation analysis was performed between DNAm probe associations after 24 and 96 hours of clozapine exposure. The Pearson correlation was computed for (A) all probes tested, (B) the top probes associated after 24hour exposure (FDR < 0.10, n = 177 probes), and (C) the top 177 probes after 96 hours of drug exposure.

Clozapine transcriptomic profiles highlight multi-tissue effects in GTEx

We then asked whether our in vitro derived transcriptomic signatures could be used to help translate the function of clozapine in humans. For this purpose, we used gene expression data from the GTEx Project, including LCLs (n = 22 GTEx tissues, Figure 3A). First, we overlapped preferentially-expressed genes of each GTEx tissue, as previously reported (Melé et al., 2015), with the clozapine-associated genes detected in our assay. Preferentially expressed genes in GTEx-LCLs exhibited the most overlap with the genes identified by our experiment followed by
whole blood (Supplementary Figure 2). We then investigated if the mean tissue expression of clozapine-associated genes was significantly different from the mean expression of all the genes tested in our experiment. We found that genes downregulated by clozapine have lower expression in all tissues except testis and LCLs (Figure 3B and Table S6). Downregulated genes are enriched for cell cycle processes and their higher expression in the testis and LCLs likely represent the proliferative nature of these tissues. Genes upregulated by clozapine have significantly higher average expression in liver, muscle, lung, and fibroblasts, and lower average expression in testis tissue. We then asked whether clozapine-associated genes have different between-tissue distances as a proxy to investigate possible functional links with other tissues. Distance is calculated using the Euclidean distance measure. A lower value indicates that two tissues are more similar while a higher distance value indicates that these tissues are more dissimilar. Of the 406 tissue pairs, we found that genes upregulated by clozapine have significantly deviating between-tissue distances across 31 pairs of tissues (Figure 3C and Table S7). Spleen tissue stands out with having the most different distance with other tissues (N=20). We also found multiple differences for adipose, lung, and breast tissue pairs. The distance between cervix uteri and ovary tissue, adrenal gland and thyroid tissue, muscle and nerve tissue are significantly more dissimilar as well for upregulated clozapine genes compared to all genes detected in our assay. For downregulated clozapine genes, we detected significantly dissimilar distances for 27 tissue-pairs, all involving testis tissue (Figure 3C).
Figure 3. Clozapine-associated genes show tissue-specific expression patterns in GTEx. (A) Using gene expression data from the GTEx project, multidimensional scaling (MDS) was performed to visualize tissues and their relationships in a parsimonious way. The top half most variable genes (n=7,025), that were also significantly detected in our experiment, across a random subset of 2,000 GTEx samples were used as input in the MDS analysis. (B) A forest plot visualizing mean tissue gene expression of clozapine-associated genes across several tissues. Clozapine genes were divided into groups that are up regulated and down regulated after drug exposure. The expected mean gene expression, based on 10,000 weighted samplings, is shown as well (dotted vertical line). Within each tissue, the observed mean tissue gene expression (x-axis) is normalized by subtracting the expected mean gene expression. P-values indicate whether the mean gene expression of clozapine up regulated genes significantly deviate from the expected mean expression within a tissue. P-values are corrected for the number of tissues tested. Downregulated genes were not tested. See Supplementary Tables for results of all tissues (C) Between tissue distance across all GTEx tissue pairs. Each point represents one tissue pair. The y-axis shows the Euclidean distance between tissues computed using only clozapine-associated genes. The x-axis shows the expected mean Euclidean distance across 50,000 weighted samplings. Tissue pairs for which the between tissue distance, based on clozapine genes only, significantly deviates from mean expected distance (Padjusted <0.05), are color-coded. P-values are adjusted for multiple testing by Bonferroni correction (n test = 465 pairs x tests = 930).
Clozapine transcriptome signatures are not enriched for schizophrenia disease risk

Previous studies have found associations between schizophrenia genetic susceptibility and antipsychotic drug targets (Gaspar and Breen, 2017; Skene et al., 2018). To investigate whether clozapine-induced in vitro gene expression signatures are also associated with genetic susceptibility of schizophrenia, we conducted gene-set enrichment analysis using all differentially expressed genes (n=463). We did not observe a significant enrichment (p = 0.91). We then considered upregulated genes, and did not observe significant enrichment (p = 0.74), nor for downregulated genes (p = 0.64). Gene sets defined according to <1% and <5% FDR did not change these findings (Table S8).

To further understand these findings, we investigated whether SCZ genetic susceptibility aggregates to antipsychotic drug target genes with and without clozapine targets. As SCZ genetic risk has been associated with antipsychotic drug target genes, it could be that this signal is primarily driven by non-clozapine genes. To examine this, we used drug target gene lists from drug-gene interaction databases as previously reported (Roth et al., 2000; Wagner et al., 2016). We were able to extract 53 drug target gene-sets belonging to the N05A class of drugs with antipsychotic actions, including clozapine. Across drug target gene-sets, we mapped all targets to 104 unique genes of which 41 were detectable in our gene expression data but none overlapped with differentially expressed genes identified. We found no evidence for SCZ risk to be enriched in antipsychotic drug target genes overall (n = 96 genes, p = 0.96) nor with clozapine targets excluded (n = 52, genes, p = 0.60). We did observe a strong concordance between the p-values of individual drug gene-sets tested in our analysis and the p-values reported by the previous study (Gaspar and Breen, 2017) (n = 50 drug gene sets, rho = 0.85, p = 1.49 \times 10^{-15})

indicating our analysis framework was able to reproduce previous findings at the level of individual drug target gene-sets. Our analysis however does not observe an association between schizophrenia genetic risk and clozapine-associated genes nor antipsychotic drug target genes.
Genes upregulated after clozapine exposure show an association with GWAS signal of total cholesterol and low-density lipoprotein.

Next, we explored whether differentially expressed genes were enriched for genetic signal of cholesterol- and cardiovascular disease-related traits. Using summary statistics of large GWASs for each trait, we observe significant SNP-h$^2$ and a rich correlation structure between these traits (Figure 4A and 4B). We then integrated the observed SNP-h$^2$ with our detected clozapine transcriptomic signatures. While no association remained significant after correction of multiple testing (72 tests, p < 6.9$^{-4}$), we did observe an increasing association between total cholesterol and LDL heritability and clozapine gene-sets across more stringent thresholds of differentially expressed genes (i.e. <FDR 5%, <FDR 1%, and Bonferroni correction (Figure 4C and Table S9).

Such trends were not observed for HDL or any of the other traits tested, including SCZ.
Figure 4. Clozapine gene-set analysis across cardiovascular traits. (A) SNP-based heritability estimates on the observed scale. (B) Genetic correlations between traits used in the analysis. The magnitude of the genetic correlation is only color-coded for estimates with p-value < 0.01. (C) MAGMA effect sizes (β) of GWAS trait associations with clozapine target genes are shown for down- and unregulated genes identified to be differentially expressed at FDR 5%, FDR 1%, and Bonferroni correction thresholds. Asterisks denote nominal significance of *P<0.05, **P<0.01. SCZ = schizophrenia, BMI = body mass index, CAD = coronary artery disease, Type2DB = type 2 diabetes, TC = total cholesterol, TG = triglycerides, HDL = high-density-lipoprotein, LDL = low-density-lipoprotein.
We used an in vitro cell system of LCLs to study molecular effects of clozapine exposure to better understand the mechanisms involved in adverse effects of psychotropic drug use and make several important observations. First, genome-wide gene expression profiling of cells exposed to clozapine showed strong activation of cholesterol metabolism and deactivation of genes involved in cell cycle processes. These findings align with our observation of changed levels of DNAm upstream of the LDL-R and CCNF genes after clozapine exposure. Second, DNAm analyses suggest that effects of clozapine are time-dependent indicating that time of drug exposure is an important experimental variable to take into account when studying clozapine. Third, integration of in vitro transcriptomic signatures with human tissues in GTEx highlight liver tissue and several immune and endocrine tissues as possible downstream effectors of clozapine exposure. Finally, genetic analysis of the results suggests that clozapine-response in LCLs is independent from schizophrenia disease risk and depleted from known drug targets of antipsychotic drugs, while it is likely linked to the genetic architecture involved in cholesterol and low-density lipoprotein levels.

In vitro gene expression changes after exposure to clozapine have been reported before for various cell lines. Here, we for the first time applied a genome-wide analysis of gene expression in lymphoblastoid cell lines in response to clozapine exposure. The central role of cholesterol metabolism in the response to clozapine is concordant with previously reported studies performed in other cell types and for different antipsychotic drugs (Ferno et al., 2011; Foley and Mackinnon, 2014). Even though most of these were based on candidate genes rather than by genome-wide analyses, the results consistently implicate cholesterol metabolism in response to antipsychotic drugs (Choi et al., 2009; Fernø et al., 2009, 2005; Hu et al., 2010; Laressergues et al., 2012, 2011, 2010; Liu et al., 2009; Raeder et al., 2006; Vik-Mo et al., 2009; Yang et al., 2007, 2009). At the gene network level, clozapine-associated genes group to clear gene clusters. In addition to
cholesterol metabolism, we observed mitochondrial and nucleosome pathways to be upregulated after clozapine exposure, while cell cycle processes were downregulated, for example.

A small number of studies have identified a link between genetic markers in genes involved in cholesterol metabolism and adverse effects of antipsychotic treatment, helping to further substantiate these findings (Chowdhury et al., 2011; Yang et al., 2016, 2015). Additional studies suggest that alterations in cholesterol metabolism by antipsychotic drugs may contribute to the beneficial effects in treating psychosis. Cholesterol is extremely important in brain development and in sustaining neuronal connections and myelination (Dietschy, 2009). Furthermore, Le Hellard et al. described an association between genes important in cholesterol metabolism (SREBP1 and SREBP2) and schizophrenia (Dietschy, 2009; Le Hellard et al., 2010), which was confirmed in a schizophrenia GWAS (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014).

In addition to significantly upregulated genes, we also identified 284 genes that are downregulated in the presence of clozapine, with significant enrichment for genes involved in cell cycle pathways. A small number of studies have shown differing gene expression levels in cell cycle genes in patients with schizophrenia, compared to healthy controls (Gassó et al., 2017; Lin et al., 2016; Okazaki et al., 2016; Wang et al., 2010; Yang et al., 2017). A similar connection with antipsychotic drug effects has not been reported before. It is possible that genes involved in cell cycle play a role in the etiology of schizophrenia, making it a possible target of antipsychotic drugs as clozapine. Conversely, the described effect may be a direct consequence of clozapine toxicity. We have shown before that clozapine, in high concentrations, has a direct effect on viability of lymphoblastoid cell lines (de With et al., 2015). Downregulation of genes involved in cell cycle processes could thus be a direct (toxic) effect of clozapine. The toxic effects of clozapine and its metabolites have been associated with clozapine-induced agranulocytosis (de With et al., 2015;
Lahdelma et al., 2010; Pereira and Dean, 2006; Williams et al., 1997). We note that for this study, we used LCLs, which are blood-derived cells, but from a different progenitor cell and with different cellular functions than neutrophils. The question remains whether findings from in vitro LCL can be directly extrapolated to in vivo effects of clozapine on neutrophils.

While we found strong effects of clozapine exposure at the gene expression level, we did not observe similar global effects at the level of DNA methylation. We did find 3 differently methylated CpGs located within genes implicated by our gene expression analysis, including the LDL receptor gene. These effects, however, were not observed after 96h of clozapine exposure. Additionally, we observed overall enrichment of regulatory and immune pathways in the top associated DNAm probes. Although previous studies have indicated that antipsychotic drugs may induce changes in DNA methylation (Burghardt et al., 2016; Houtepen et al., 2016; Kinoshita et al., 2017; Rukova et al., 2014), we did not find evidence for immediate large effects on DNA methylation based on CpG sites assayed in our experiments. Possibly, subtle changes in DNA methylation patterns play a regulatory role in the observed gene expression changes but a larger sample size is needed to decipher these changes (Jones, 2012). In addition, a wider range of clozapine concentrations in vitro may provide finer experimental resolution to observe subtle epigenetic changes and identify key regulatory drivers. DNA methylation is one type of epigenetic regulation and assaying other regulatory mechanisms, such as histone modification or RNA regulation, may provide further insights into the regulatory dynamics that drive widespread changes in gene expression (Allis and Jenuwein, 2016). Lastly, we observed DNA methylation changes to be time-dependent. While the clinical meaning of this remains speculative, our findings do suggest that the existing literature should be evaluated in the context of the duration of drug exposure. In addition, future studies could gain more insights into the function of clozapine when modeling drug exposure time as a variable in their analyses.
To further explore our gene expression findings, we investigated their association with genetic susceptibility of schizophrenia. We did not observe significant enrichment of schizophrenia heritability across clozapine-associated genes. Two previous studies did report an association between schizophrenia heritability and antipsychotic target genes (Gaspar and Breen, 2017; Skene et al., 2018). Their approach, however, differed from ours. While they used lists of antipsychotic target genes originating from pharmacological databases, we used a list of experimentally derived gene sets after \textit{in vitro} exposure to clozapine in LCLs. The genes we found to be differentially expressed did not overlap with antipsychotic drug target genes used in these two previous studies, which may explain the discrepancy in findings. If we assume that the clozapine-induced gene expression differences are a proxy for adverse effects, the lack of evidence of the enrichment analysis suggests that the genetic architecture of disease susceptibility is likely independent from susceptibility to adverse effects. We explored this further by examining the enrichment of genetic signal linked to cholesterol and cardiovascular disease-related traits such as body mass index, coronary artery disease, type 2 diabetes, and triglyceride levels. We observed a nominally significant increased enrichment for total cholesterol and LDL genetic signal in upregulated genes in response to clozapine exposure. While these findings did not survive multiple testing correction and should be interpreted with caution, the observed cholesterol and LDL heritability enrichment increased as we narrowed down on the most strongly associated clozapine genes suggesting that this association warrants further investigation. Mapping heritability from large-scale population studies to \textit{in vitro} experimental systems can then serve as a powerful approach to study biological pathways through integration of polygenic disease risk (Ori et al., 2019). Metabolic adverse effects are often observed in patients using clozapine (and antipsychotics in general) and improving our understanding of the mechanisms that underlie these adverse effects is an imperative area of research.
An important advantage of in vitro experimental models is the controlled laboratory environment, which not only decreases the signal-to-noise ratio in the collected data but also allows for precise manipulation of the model in follow up work. While our current results are consistent with previous findings in other cell types (Choi et al., 2009; Fernø et al., 2009, 2005; Hu et al., 2010; Lauressergues et al., 2012, 2011, 2010; Liu et al., 2009; Raeder et al., 2006; Vik-Mo et al., 2009; Yang et al., 2007, 2009), it remains unclear whether LCLs are an appropriate cell type to capture the molecular changes that are most relevant for studying adverse effects of antipsychotics in patients. To gain insight into the transferability of the in vitro clozapine-response in LCLs to functions of human tissues, we investigated how clozapine-associated genes are expressed across GTEx tissues. As gene expression patterns have been shown to have significant degree of sharing across human tissues (Buil et al., n.d.; GTEx Consortium et al., 2017) effects discovered in one tissue may thus be informative for other tissues. We observed that preferentially expressed genes in GTEx-LCL tissue overlap the most with clozapine genes identified in LCLs in vitro, with whole blood ranked as second highest tissue. Preferentially expressed genes however, represent only a subset of the clozapine genes. Using all associated genes, we demonstrated that upregulated clozapine genes have significantly different average expression in liver, muscle, lung, and testis tissue. This suggests multi-tissue downstream effects through clozapine treatment, which fits its clinical presentation of a diverse set of adverse effects that have been reported (Iqbal et al., 2003). Our finding of higher expression in the liver is in line with the observed cholesterol gene ontology signature of upregulated clozapine genes in our assay. The liver is a central player in the function of cholesterol in the body and clozapine-related hepatotoxicity has furthermore been reported in cases of treatment with the drug (Keane et al., 2009; Kellner et al., 1993). We observe no differences for any of the GTEx brain tissues.

On top of within tissue effects, we also performed a more explorative analysis to examine clozapine gene expression-based similarity between tissues and found specific tissue pairs that
were significantly different in tissue similarities. Here, we highlight some interesting observations. We observe a decrease in similarity for tissues involved in immune (spleen), endocrine (testis, ovary, adrenal and thyroid gland) and metabolic (adipose) functioning. The spleen and testis in particular stand out with a significant change in dissimilarity with many tissues. The spleen is the largest secondary lymphoid organ in the body and hosts a wide range of immunological functions, including the storage of leukocytes (Lewis et al., 2019). While it remains speculative if splenic dysfunction could for example lead to agranulocytosis, our findings do solicit for more research on the mechanism between clozapine and splenic function. An intriguing, and perhaps also surprising, finding is the implication of the testis among analysis of genes downregulated by clozapine. The testis has two primary functions; to produce sperm and to produce hormones, in particular testosterone, which is a sex steroid synthesized from cholesterol (Eacker et al., 2008). Genes downregulated by clozapine are enriched for cell division processes and cell proliferation is an important function of testicular cells (Sohni et al., 2019). While little is currently known about how clozapine or other antipsychotics may affect testicular function in adult humans, our finding does again point to cholesterol-related biology. In addition to the testis, we also observe deviation in between-tissue similarity for several other endocrine tissues. As endocrine abnormalities are known causes of human obesity, the identified transcriptomic profile of clozapine response together with downstream tissue effects may point to new avenues to study clozapine-induced weight gain (Baptista, 1999). Together, these findings suggest that clozapine disproportionately affects the function of specific tissues. Our results furthermore demonstrate how experimental molecular signatures can be integrated with external genomic datasets, such as the GTEx project, to help translate in vitro findings to human biology.

We used LCLs as an in vitro model for studying molecular effects of clozapine exposure in an effort to improve our understanding of antipsychotic-induced adverse effects. Genome-wide gene expression profiling demonstrated a robust up-regulation of cholesterol metabolism and down-
regulation of cell cycle pathways, with only limited changes in DNA methylation profiles. We did not find evidence that genes up- or down-regulated during clozapine exposure were enriched for genetic variation associated with schizophrenia. On the other hand, the observed enrichment signal with the genetic basis of total cholesterol and LDL levels and multi-tissue involved across immune, endocrine, and metabolic functions may provide important leads linked to antipsychotic drug induced metabolic adverse effects. The necessary challenge of large-scale, systematic prospective patient cohort studies of adverse effects of clozapine and other AP remains, while in vitro studies such as ours provide only glimpses of what may be relevant.
Acknowledgements

The authors would like to thank Drs. Barbara Franke, Jeffrey Glennon, Jan Buitelaar, and Wouter Staal for their input in the early stages of the project and the GTEx Project and all used GWAS studies for making their data publicly available.

Funding

This work was partly supported by funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) Pediatric European Risperidone Studies (PERS) under grant agreement n°241959. RAO is supported by NIH grants R01 DA028526, MH090553, and MH078075. SDW was supported by a personal grant from Stichting de Drie Lichten. SDJ was part funded by NARSAD Young Investigator Grant (YI 60373). The GTEx Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS.

Competing interest

The authors declare no conflict of interest.
572 References


Consortium †the International Hapmap, †The International HapMap Consortium. 2003. The International HapMap Project. Nature. doi:10.1038/nature02168


GTEx Consortium, Laboratory, Data Analysis & Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund, NIH/NCI, NIH/NHGRI, NIH/NIMH, NIH/NIDA, Biospecimen Collection Source Site—NDRI, Biospecimen Collection Source Site—RPCI, Biospecimen Core Resource—VARI, Brain Bank Repository—University of Miami Brain Endowment Bank, Leidos Biomedical—Project Management, ELSI Study, Genome Browser Data Integration & Visualization—EBI, Genome Browser Data Integration & Visualization—UCSC Genomics Institute, University of California Santa Cruz, Lead analysts:, Laboratory, Data Analysis & Coordinating Center (LDACC):, NIH program management:, Biospecimen collection:, Pathology:, eQTL manuscript working group:, Battle A, Brown CD, Engelhardt BE, Montgomery SB. 2017. Genetic effects on gene expression across human tissues. Nature 550:204–213.


Ota VK, Noto C, Gadelha A, Santoro ML, Spindola LM, Gouveia ES, Stilhano RS, Ortiz BB, Silva


Yang Z, Li M, Hu X, Xiang B, Deng W, Wang Q, Wang Y, Zhao L, Ma X, Sham PC, Northoff G, Li T. 2017. Rare damaging variants in DNA repair and cell cycle pathways are associated with hippocampal and cognitive dysfunction: a combined genetic imaging study in first-episode...
treatment-naive patients with schizophrenia. Transl Psychiatry 7:e1028.


