

ARTICLE

Open Access

# Identification of a novel polymorphism associated with reduced clozapine concentration in schizophrenia patients—a genome-wide association study adjusting for smoking habits

Robert Løvsletten Smith<sup>1</sup>, Kevin O'Connell<sup>2</sup>, Lavinia Athanasiu<sup>2</sup>, Srdjan Djurovic<sup>3,4</sup>, Marianne Kristiansen Kringen<sup>1,5</sup>, Ole A. Andreassen<sup>2</sup> and Espen Molden<sup>1,6</sup>

## Abstract

Clozapine (CLZ) is the superior antipsychotic drug for treatment of schizophrenia, but exhibits an extensive interpatient pharmacokinetic variability. Here, we conducted a genome-wide association study (GWAS) of CLZ serum concentration adjusting for known smoking habits, which is a major nongenetic factor reducing CLZ levels. The study included 484 patients with 10,283 steady-state serum concentrations of CLZ and *N*-desmethylclozapine, prescribed dosing, co-medications and known smoking habits ( $n = 422$ ; 9284 serum samples) from a therapeutic drug monitoring (TDM) service. The GWAS analyses were performed with and without smoking habits as covariate, where possible hits were assessed in relation to the target CLZ concentration range applied in the TDM service (300–2500 nmol/L). The smoking-independent analysis of *N*-desmethylclozapine serum concentration and the CLZ-to-*N*-desmethylclozapine ratio replicated the previously identified locus on chromosome 4. After adjusting for smoking habits in patients confirmed as 'smokers' (61%) or 'nonsmokers' (39%), a novel variant (*rs28379954*; minor *T*>*C* allele frequency 4.1%; 7.6% *CT* carriers in the population) within the gene encoding the nuclear factor 1 B-type (*NF1B*) was significantly associated with reduced CLZ serum concentration ( $p = 1.68 \times 10^{-8}$ ,  $\beta = -0.376$ ; explained variance 7.63%). There was no significant association between *rs28379954* and *N*-desmethylclozapine concentration in the GWAS analysis ( $p = 5.63 \times 10^{-5}$ ). The fraction of CLZ TDM samples below 300 nmol/L was significantly higher in carriers vs. noncarriers of the *rs28379954* minor *C* allele [12.0% (95% CI: 9.4–14.7) vs. 6.2% (95% CI: 5.7–6.8),  $p < 0.001$ ]. We identified a novel variant in the *NF1B* gene associated with reduced CLZ levels and increased risk of subtherapeutic serum concentrations. This warrants testing of clinical relevance of screening for this gene variant, and also experimental studies to investigate the biological mechanisms of *NF1B* involvement in CLZ pharmacokinetics.

## Introduction

Clozapine (CLZ) is the most efficient antipsychotic drug in treatment of schizophrenia<sup>1</sup>. However, CLZ is only

indicated in treatment-resistant schizophrenia (TRS) due to risk of neutropenia or agranulocytosis<sup>2,3</sup>, which is reported to occur in about 3% and 0.5–1% of the patients, respectively<sup>4</sup>. Therefore, close monitoring of neutrophil granulocyte counts is mandatory during use of CLZ<sup>5</sup>. The suppression of granulocyte levels, as well as risk of other adverse effects such as seizures, cardiometabolics (weight gain and type 2 diabetes), sedation, and hypersalivation, have been associated with CLZ and/or metabolite levels<sup>6,7</sup>.

Correspondence: Espen Molden (espen.molden@farmasi.uio.no)

<sup>1</sup>Center for Psychopharmacology, Diakonhjemmet Hospital, Oslo, Norway

<sup>2</sup>CoE NORMENT, Division of Mental Health and Addiction, Oslo University

Hospital, and Institute of Clinical Medicine, University of Oslo, Oslo, Norway

Full list of author information is available at the end of the article

These authors contributed equally: Robert Løvsletten Smith, Kevin O'Connell

© The Author(s) 2020



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

CLZ is subjected to complex metabolism involving biotransformation via multiple pathways<sup>8,9</sup>. Several cytochrome P450 (CYP) enzymes, including CYP1A2, CYP3A4, CYP2D6, and CYP2C19, are in vitro shown to be involved in the oxidative metabolism of CLZ to the major metabolites *N*-desmethylclozapine and clozapine-*N*-oxide<sup>9</sup>. In addition, CLZ undergoes metabolism to reactive nitrenium ion metabolites, which are believed to be important for the toxicity of CLZ treatment<sup>10,11</sup>. Less is known about the enzyme(s) mediating the formation of nitrenium ion metabolites of CLZ, but the myeloperoxidase system in granulocytes seem to play a role<sup>9,12</sup>.

Previous studies have shown that interindividual variability in CLZ concentration is significantly associated with age, sex, and smoking habits<sup>13,14</sup>, where the latter is quantitatively most important. However, these factors only explain parts of the variability, which has motivated the search for pharmacogenetic determinants of CLZ pharmacokinetics. Most studies so far have been unsuccessful in identifying genetic polymorphisms associated with individual variability in CLZ pharmacokinetics<sup>15</sup>, but recently a genome-wide association study (GWAS) in a British patient population ( $n = 2989$ ) identified a single locus (15:74817689–75404506) between the *CYP1A1* and *CYP1A2* genes associated with a reduction in CLZ serum levels<sup>16</sup>. Further, a locus on chromosome 4 (4:69542100–70387482), containing the *UGT2B10* gene, was associated with both *N*-desmethylclozapine serum concentration and the CLZ-to-*N*-desmethylclozapine metabolic ratio. For the *N*-desmethylclozapine serum concentration and CLZ-to-*N*-desmethylclozapine metabolic ratio two additional loci were identified<sup>16</sup>, i.e., at chr2:234611523–234676118 and 10:96098093–96974830, respectively. In contrast to polygenic complex traits, such as schizophrenia<sup>17</sup>, these results support the hypothesis that the genetic architecture of CLZ metabolism is less polygenic and may constitute a smaller number of variants with larger effect sizes<sup>16,18</sup>. This is typical for the pharmacogenetics of drug-metabolizing enzymes, where a limited number of variant allele diplotypes have large effect sizes and separate the population into distinct metabolizer phenotypes<sup>19,20</sup>. In addition, CLZ related adverse effects are associated with few genetic variants with large effect sizes<sup>21,22</sup>.

Tobacco smoke reduces the serum concentration of CLZ by ~30–50% due to polycyclic aromatic hydrocarbons inducing gene transcription of drug-metabolizing enzymes, e.g., CYP1A enzymes<sup>23–25</sup>. Thus, information on smoking habits is crucial when studying factors that may influence the CLZ serum concentrations. The previous British GWAS did not account for the individual patient's smoking habits<sup>16</sup>, but used a proxy measure (polygenic risk scores for smoking), which may imply that significant variants were not

captured in the analysis. The aim of the present study was to perform a GWAS on serum concentrations of CLZ and its main metabolite *N*-desmethylclozapine in a Norwegian population of schizophrenia patients with known smoking habits.

## Materials and methods

### Study population

Patients with serum concentration measurements of CLZ and the metabolite *N*-desmethylclozapine, and bio-banked blood samples for genotyping, were included from the therapeutic drug monitoring (TDM) service at the Center for Psychopharmacology, Diakonhjemmet Hospital, Oslo, Norway, from March 2005 to May 2019. Inclusion criteria were: (i) blood sampling for serum concentration analysis of CLZ at steady-state conditions and 10–30 h after the last drug intake, and (ii) prescribed daily dose of CLZ. For the GWAS analysis adjusting for smoking only, a further inclusion criterion was (iii) information about confirmed smoking habits written by the physician on the TDM requisition form (cigarette smoking 'yes' or 'no'). The time intervals between last dose intake and blood sampling for serum concentration analysis, and the psychotropic co-medication profiles, were obtained from the requisition forms. Exclusion criteria were (i) concurrent use of interacting drugs, i.e., the CYP1A2/2C19/3A4 inhibitor fluvoxamine, or the CYP3A4/2C19 inducers phenobarbital, phenytoin, and carbamazepine, and (ii) concentration of CLZ or *N*-desmethylclozapine below the lower limit of quantification (LLOQ; i.e., 20 nmol/L). All serum concentration measurements from each of the patients fulfilling the criteria were included in the statistical analyses, as described below.

The study was approved by the Regional Committee for Medical and Health Research Ethics and the Hospital Investigational Review Board. As the study was solely based on historical data, the project did not have the potential to cause any harm or burden, informed patient consent was not required.

### Clinical characteristics

The study population consisted of patients with TRS as determined by TDM of CLZ per se, since use of this drug is solely indicated for TRS<sup>3</sup>. In Norway, TDM analyses are reimbursed by the governmental health service, and it is used as a tool for clinical follow-up in psychiatry. The analyses are typically requested as a routine control to monitor e.g., adherence and drug–drug interactions, or to assess adverse effects or lack of effect, in relation to a therapeutic concentration reference range. All results from the analyses are interpreted by the TDM laboratory against the therapeutic reference range as basis for the reports sent to the physicians.

The therapeutic reference ranges used in our TDM service are based on international standards as defined by the AGNP

working group<sup>26</sup>. However, they are adjusted according to the concentrations measured at our laboratory in a Norwegian patient population during therapeutic dosing. In the case of CLZ, the applied therapeutic reference range at our laboratory is 300–2500 nmol/L, while the AGNP reference range is 1071–1835 nmol/L. The AGNP definition of 1071 nmol/L (=350 ng/ml) as the lower therapeutic concentration boundary is based on several studies correlating clinical response and serum concentration of CLZ, e.g., the study by Perry et al. showing that only 22% of the patients with concentration <350 ng/ml were responders (i.e.,  $\geq 20\%$  reduction in symptom score), while 64% of the patients above the limit were responders<sup>27</sup>. This was supported by a follow-up study where the majority of nonresponders achieved clinical effects when the serum CLZ levels were raised above 350 ng/ml<sup>28</sup>. However, in the present study, we use a more conservative lower serum concentration boundary, i.e., 300 nmol/L, since it is applied in our clinical TDM service based on Norwegian data. Use of 300 nmol/L as the lower boundary also increases the likelihood of poor clinical efficacy of CLZ when serum concentrations are below this threshold. In addition, a CLZ concentration of 300 nmol/L corresponds with a brain dopamine-2 (D2) receptor occupancy of about 20–40%, which is regarded as critical for clinical effect<sup>29</sup>.

The TDM service at the Diakonhjemmet Hospital in Oslo analyses serum samples from psychiatric patients mainly located in the South-Eastern Norway Regional Health Authority, and ~25% of the submitted samples are from hospitalized patients.

### Genotyping and imputation

DNA was extracted from whole blood and genotyped with the Human Omni Express-24 v.1.1 array (Illumina Inc., San Diego, CA, USA), at deCODE Genetics (Reykjavik, Iceland), in accordance with the standard Illumina protocol. Standard pre-imputation quality control was performed using PLINK v1.9<sup>30,31</sup>. Finally, chromosome-wide haplotypes were phased with Eagle2<sup>32,33</sup> and missing variants were imputed with Minimac3<sup>34</sup> using the first release of the haplotype reference consortium reference set. After imputation, we removed variants if their minor allele frequency was <1% or if they showed departure from Hardy–Weinberg equilibrium ( $P$ -value  $< 1 \times 10^{-6}$ ). Furthermore, we removed related individuals (pairwise Identity-By-Descent  $\hat{\pi} > 0.2$  according to PLINK v1.9<sup>30,31</sup>) and those individuals with high rates of genotyping missingness (>5%). The final dataset included 484 unrelated individuals and ~6.2 million variants for analysis.

### GWAS of serum concentrations with and without adjusting for smoking habits

We applied a regression modeling framework to combine serum concentration measurements of each patient

from multiple time points into a single phenotype, as described previously<sup>16</sup>. The R package *ftdistrplus* v1.0-14 was used to determine the best-fitting distribution for each phenotype outcome variable (dose-adjusted serum concentrations of CLZ and *N*-desmethylclozapine, and the CLZ-to-*N*-desmethylclozapine metabolic ratio). Thereafter, the R package *gamlss* v5.1-2 was used to specify a random-effects model for each phenotype outcome variable. The fixed effect covariates in the model included known predictors of CLZ serum concentration (CLZ dose, age at sampling and time between sampling and last dose intake). In addition, a random effect was added to model the distribution of the outcome in each individual, controlling for these covariates. The inclusion of this random effect for each individual permits unbalanced data patterns (differing number of serum measurements) in the model<sup>35</sup>. For each model, the coefficients of this random effect for each individual were extracted. These coefficients correspond to the variation within the serum concentrations for each individual independent of the included covariates, and were used as the outcome phenotypes for the GWAS analyses.

We conducted a GWAS for each outcome phenotype (variation in serum concentrations of CLZ and *N*-desmethylclozapine, and the CLZ-to-*N*-desmethylclozapine metabolic ratio) using linear regression analyses implemented in PLINK v1.9<sup>30,31</sup>, controlling for participant age, sex, the first 20 genetic principal components, and genotyping batch. We also performed a GWAS for each outcome phenotype in the sample with known smoking habits ( $n = 422$ ) by adding this variable as a covariate in the regression models.

Details describing the methods used for gene loci definitions and annotations, estimating the single-nucleotide polymorphism (SNP) proportion of variance explained, and follow-up GWAS meta-analyses with the British sample including testing the effect of smoking PRS as proxy for smoking and body mass index, are included in the Supplementary file (Supplementary Methods; Supplementary Figs. 1–4 and Supplementary Tables 1 and 2). Also included is a methodological description of the targeted Taqman genotyping used for verification of SNPs significantly associated with CLZ concentration in the GWAS.

### Follow-up analyses of SNPs identified in the GWAS

For SNPs significantly associated with CLZ concentration in the GWAS, the various phenotypes were compared in carriers vs. noncarriers of the identified variant(s) using linear mixed model analysis. These analyses were performed for both smokers and nonsmokers adjusting for sex, age, and sampling time as covariates. The Pearson Chi-square test was used to evaluate the potential differences in proportion of patients with TDM measurements

of CLZ either below or above the defined target concentration range (300–2500 nmol/L; c.f. information provided in the beginning of the “Methods” section).

### Statistical tools

The statistical analyses were performed using R statistics and PLINK v1.9<sup>30,31</sup> for the GWAS and SPSS® (version 25.0) for the assessments of the significant GWAS hits, as described above.

## Results

### Sample characteristics

The study population comprised 484 CLZ-treated patients with a total of 10,283 CLZ and *N*-desmethylclozapine serum concentrations measurements (Table 1), with median age of 37 years (IQR: 29–47) and 38.4% females ( $n = 186$ ). The median prescribed daily dose of CLZ was 400 mg (IQR: 250–500 mg), whereas the median serum concentration of CLZ and *N*-desmethylclozapine were 1167 and 770 nmol/L, respectively. Among the patients with known smoking habits ( $n = 422$ ; 9284 serum samples), 258 were confirmed smokers (61.1%) and 164 were confirmed nonsmokers (38.9%). The median dose-adjusted CLZ serum concentrations (C/D ratio) were 2.46 nmol/L/mg (IQR: 1.67–3.64) and 4.63 nmol/L/mg (IQR: 2.91–6.49) in smokers and nonsmokers, respectively, whereas the median CLZ-to-*N*-desmethylclozapine ratios were 1.43 (IQR: 1.17–1.71) in smokers and 1.69 (IQR: 1.40–2.06) in nonsmokers.

### GWAS of serum concentrations adjusting for smoking habits

The GWAS of CLZ serum concentration, adjusting for smoking habits ( $n = 422$ ), identified a single novel genome-wide significant locus at chromosome 9 (Fig. 1a, Table 2 and Supplementary Fig. 5a), represented by lead SNP (*rs28379954*,  $p = 1.677 \times 10^{-8}$ , 9:14163907). Data from the 1000 Genomes phase 3-genome browser shows that this SNP lies within an LD-sparse region (Supplementary Fig. 6).

**Table 1 Population characteristics.**

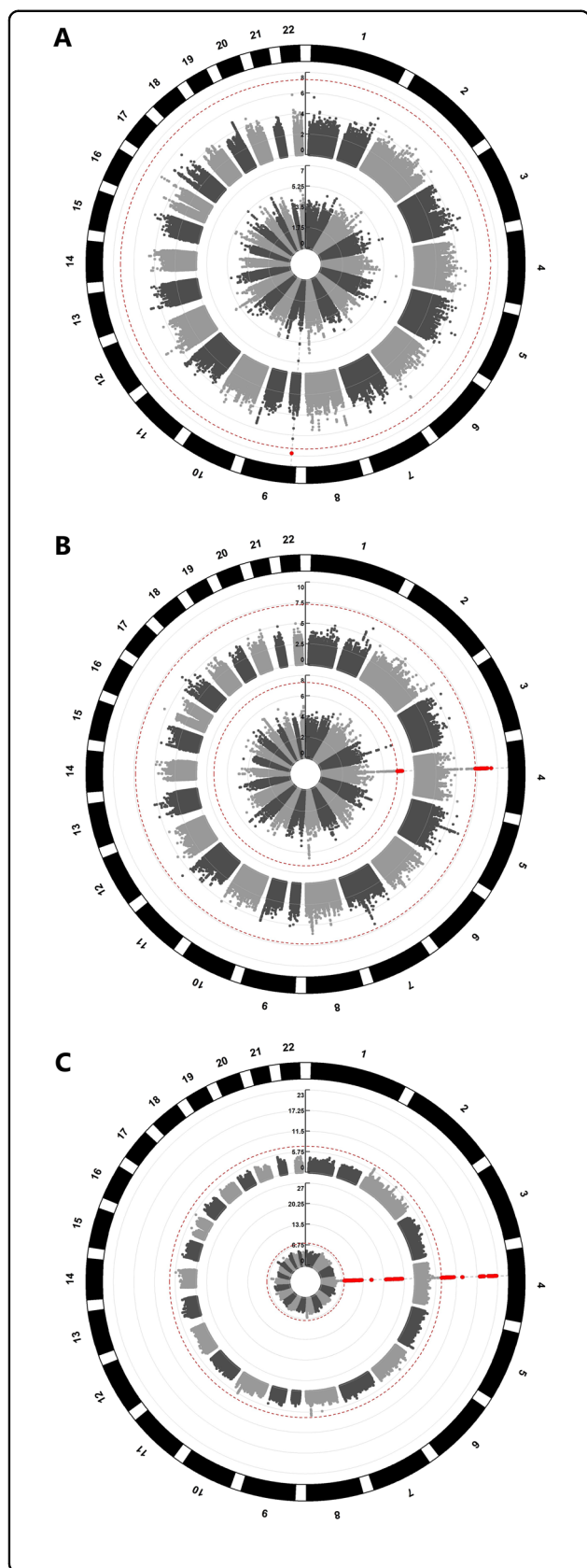
Variables	Value
No. of patients, $n$ (measurements)	484 (10,283)
Male/female, $n$	298/186
Age, years; median (IQR)	37 (29, 47)
No. of measurements per patient; median (IQR)	13 (5, 28)
Clozapine concentration, nmol/L; median (IQR)	1167 (687, 1786)
<i>N</i> -desmethylclozapine concentration, nmol/L; median (IQR)	770 (469, 1135)
Clozapine dose, mg (IQR)	400 (250, 500)
Withdrawal time (time between dose and sampling), $h$ , median (IQR)	13.1 (12.2, 14.5)
Smoking habits, $n$ (%)	422 (87.2)
Smoker/nonsmoker (% females)	258 (31.4)/164 (48.2)

However, four other SNPs, in LD with the lead SNP ( $r^2 > 0.2$ ) and within 250 kb, showed nominal significance ( $p < 1.000 \times 10^{-4}$ ): *rs10481501* C>T (9:14172871,  $p = 4.506 \times 10^{-7}$ , LD  $r^2 = 0.554$ ), *rs10961406* A>G (9:14182043,  $p = 3.856 \times 10^{-5}$ , LD  $r^2 = 0.451$ ), *rs10120632* G>C (9:14226937,  $p = 3.892 \times 10^{-5}$ , LD  $r^2 = 0.329$ ) and *rs12156582* C>T (9:14163383,  $p = 4.546 \times 10^{-6}$ , LD  $r^2 = 0.411$ ; Supplementary Fig. 7).

The *rs28379954* variant is located within an intron of the nuclear factor 1 B-type (*NF1B*) gene and the minor C allele (C allele frequency = 0.041; explained variance 7.63%) was associated with reduced CLZ serum concentration (beta = -0.3755; Table 2 and Fig. 1a). The *NF1B* genotypes from GWAS and targeted Taqman analyses of *rs28379954* T>C showed 97% consistency, supporting the significant association with CLZ serum concentration. Functional analyses do not suggest this SNP, or the four others in LD ( $r^2 > 0.2$ ), to be deleterious (CADD scores: <12.37) or likely to have regulatory functionality (RegulomeDB scores = 5–7). Furthermore, none of these SNPs were associated with any other phenotype in the GWAS catalog and none were identified as potential eQTLs in the GTEx portal.

GWAS analyses of *N*-desmethylclozapine serum concentration and the CLZ-to-*N*-desmethylclozapine ratio as phenotypes, adjusting for smoking habits, revealed the same significant loci as those identified in the GWAS analyses by Pardiñas et al.<sup>16</sup> (Fig. 1b, c, Table 2, and Supplementary Fig. 5b, c). We did not replicate the additional locus associated with *N*-desmethylclozapine serum concentration in the study by Pardiñas et al.<sup>16</sup> (2:234611523–234676118). Of the 242 SNPs within this region, the most significant association to *N*-desmethylclozapine concentration was identified for *rs17862876* ( $p = 0.059$ , beta = -0.227). Furthermore, we did not replicate the additional locus associated with the CLZ-to-*N*-desmethylclozapine ratio (10:96098093–96974830). Of the 2098 SNPs within this region, the most significant association to the CLZ-to-*N*-desmethylclozapine ratio was identified for *rs66634733* ( $p = 1.065 \times 10^{-4}$ , beta = 0.80). Functional annotation of the lead SNPs from these GWAS did not predict any deleterious effects on protein structure/function and did not suggest any regulatory functionality (Table 2).

Among the patients with smoking habits data ( $n = 422$ ), 32 patients (7.6%) were in the GWAS identified as *rs28379954* minor C allele carriers. In Fig. 2a, the linear regressions of absolute, unadjusted CLZ serum concentration vs. prescribed daily dose are shown according to *rs28379954* genotype. As compared with the *rs28379954* TT genotype carriers (smokers: 2.31 (95% CI: 2.19–2.44); nonsmokers: 3.39 (95% CI: 3.15–3.64)), the C allele carriers (smokers: 0.81 (95% CI: 0.56–1.06);



**Fig. 1** Manhattan plots showing clozapine concentration, *N*-desmethylclozapine concentration, and metabolic ratio associations. Circular Manhattan plots for (A) clozapine serum concentration, (B) *N*-desmethylclozapine serum concentration, and (C) Clozapine-*N*-desmethylclozapine metabolic ratio. The inner plot shows the results from the initial GWAS without controlling for smoking habits. The outer plot shows the results of the GWAS after controlling for smoking habits. The  $-\log_{10}$  transformed  $p$ -values for each SNP are indicated on the y-axes and chromosomal positions along the exterior of the plot. The dashed red line represents the threshold for significant associations ( $p < 5 \times 10^{-8}$ ). SNPs surpassing this threshold are indicated in red.

nonsmokers: 0.77 (95% CI: 0.37–1.16)) had a significantly lower linear slope (Fig. 2a). This was reflected by a significantly higher proportion of CLZ serum concentrations below the therapeutic range ( $<300$  nmol/L) in *C* vs. *TT* allele-carrying smokers (16.0% vs. 7.2%,  $p < 0.001$ ) and nonsmokers (8.0% vs. 4.9%,  $p = 0.017$ ), despite similar daily dosing ( $p > 0.05$ ). Regarding the pharmacokinetic effect of *rs28379954*, *C* allele carriers had an ~42% lower CLZ dose-adjusted serum concentration than *TT* genotype carriers in smokers ( $p = 0.004$ ), whereas the *C* allele-carrying nonsmokers had 33% lower CLZ dose-adjusted serum concentration ratio than the *TT* genotype carrying nonsmokers ( $p = 0.031$ ; Fig. 2b), after adjusting for age, sex, and sampling time as covariates. The dose-adjusted serum concentration of *N*-desmethylclozapine and CLZ-to-*N*-desmethylclozapine ratio were not associated with the *rs28379954* *C*>*T* polymorphism in the GWAS analyses (Table 2). Due to reduced substrate (CLZ) availability, the *rs28379954* *C* allele carriers still had 30% and 31% lower dose-adjusted serum concentration of *N*-desmethylclozapine compared with the smoking ( $p = 0.029$ ) and nonsmoking ( $p = 0.016$ ) *TT* genotype carriers, respectively (Fig. 2c). In smoking *rs28379954* *TT* allele carriers, the CLZ-to-*N*-desmethylclozapine ratio was 19% ( $p = 0.014$ ) higher than in smoking *C* allele carriers (Fig. 2d), while no significant change in CLZ-to-desmethylclozapine ratio was observed among *C* allele-carrying nonsmokers ( $p = 0.265$ ; Fig. 2d).

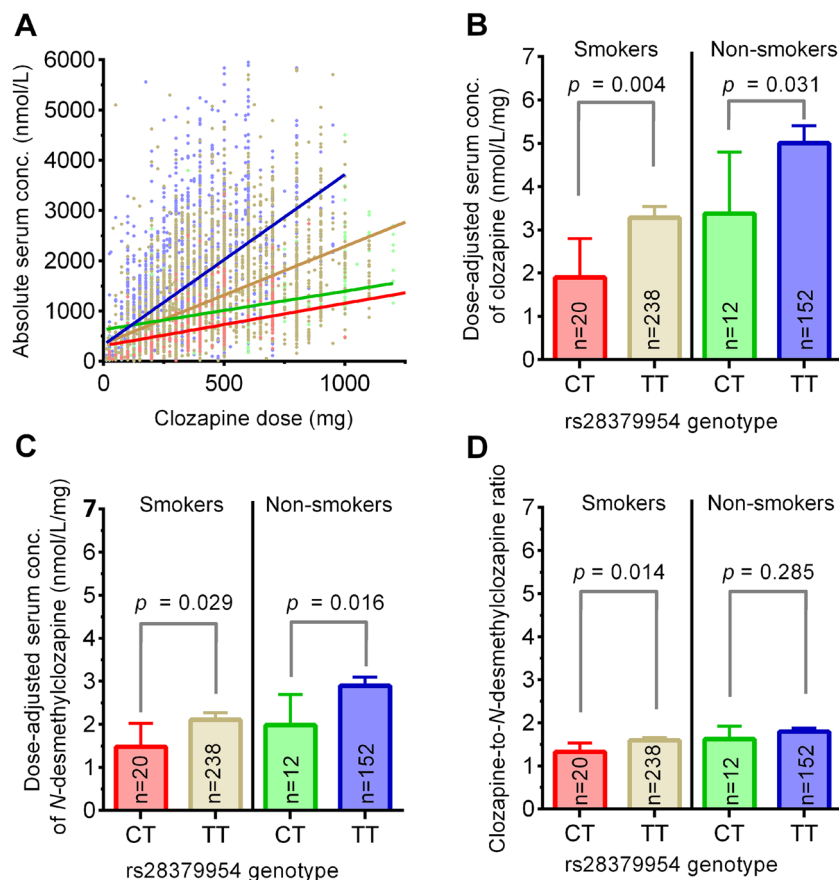
#### GWAS of serum concentrations without adjusting for smoking habits

Without adjusting for smoking habits, we identified no genome-wide significant loci in our GWAS of CLZ serum concentration (Fig. 1a and Supplementary Fig. 8a). Further investigation of the previously implicated *CYP1A* locus on chromosome 15 (15:74817689–75404506)<sup>16</sup> showed that the lead SNP *rs2472297* had a call rate of 64.2% and was therefore excluded from our GWAS analyses. However, 773 SNPs were present within this locus, and the most significant association was observed for *rs11635266* *T*>*C* (15:75386649) at  $p = 0.002$  (beta = 0.080).

**Table 2 Association statistics and effect sizes of the lead SNPs for each phenotype and genomic locus from the GWASs, without or with controlling for smoking habits.**

Phenotype and locus	Lead SNP	Position: alleles	Effect allele	GWAS p	Beta	SE	Annotation	Nearest gene	Distance to gene	CADD	RDB	Min Chr state	Common Chr state
Statistics without controlling for smoking habits:													
<i>N</i> -desmethylclozapine:													
4:69601886–70138176	rs1670747	69610466: G/A	A	1.616 × 10 <sup>-8</sup>	-0.183	0.032	Intergenic	<i>RP11-1267H10.4</i>	32249	2.713	NA	2	15
Metabolic Ratio:													
4:69535335–70387482	rs1513559	69655555: A/G	G	4.819 × 10 <sup>-27</sup>	0.246	0.021	Intergenic	<i>CTD-2005D20.1</i>	2362	0.004	5	9	15
Statistics controlling for smoking habits:													
Clozapine:													
9:14163907–14163907	rs28379954	14163907: T/C	C	1.677 × 10 <sup>-8</sup>	-0.376	0.065	Intronic	<i>NFB</i>	0	8.433	7	5	15
<i>N</i> -desmethylclozapine:													
4:69601886–70138176	rs1670747	69626729: G/A	A	5.805 × 10 <sup>-10</sup>	-0.215	0.034	Intergenic	<i>RP11-1267H10.4</i>	32249	2.713	NA	2	15
Metabolic ratio:													
4:69544277–70340670	rs292638	69662832: C/G	G	2.067 × 10 <sup>-23</sup>	0.247	0.023	Intergenic	<i>CTD-2005D20.1</i>	9639	2.223	5	5	15

CADD = Combined Annotation-Dependent depletion score, which predict how deleterious the SNP effect is on protein structure/function (higher scores indicate more deleterious); RegulomeDB (RDB) scores predict likelihood of regulatory functionality (lower scores indicate higher likelihood); min Chr State = minimum chromatin state across 127 tissue types (lower scores indicate more open chromatin); common Chr State = most common chromatin state in 127 tissue types.



**Fig. 2** The impact of *NFIB* *rs28379954* minor C allele carriers on clozapine and *N*-desmethylclozapine concentration, as well as clozapine-to-*N*-desmethylclozapine ratio, in relation to smoking habits. **a** The figure shows all serum concentration measurements (dots) versus clozapine prescribed daily dose. The regression lines for the various subpopulations are calculated using a linear mixed model analyses without any covariates (blue, nonsmoking *rs28379954* TT carriers; golden, smoking TT carriers; green, nonsmoking CT carriers; red, smoking CT carriers). The figures show the impact of *rs28379954* minor C allele carriers on dose-adjusted serum concentration of clozapine (C/D ratio; **b**), *N*-desmethylclozapine (C/D ratio; **c**) and clozapine-to-*N*-desmethylclozapine ratio (**d**) in relation to smoking habits. The ratios were analyzed by linear mixed model including sex, age, and sampling time (time between last dose and blood sampling) as covariates. The numbers of patients within each subpopulation are shown in the various bars.

Our analysis of *N*-desmethylclozapine serum concentration and the CLZ-to-*N*-desmethylclozapine ratio replicated the significant locus on chromosome 4 (Fig. 1b, c, Table 2 and Supplementary Fig. 8b, c) reported by Pardiñas et al.<sup>16</sup>, regardless of adjusting for smoking habits ( $p < 1.62 \times 10^{-8}$ ). On the other hand, we did not replicate the additional loci previously associated with *N*-desmethylclozapine serum concentration (2:234611523–234676118) or with the CLZ-to-*N*-desmethylclozapine ratio (10:96098093–96974830), respectively.

Meta-analyses of our smoking-independent GWAS and the previously published British GWAS<sup>16</sup> did not identify any novel significant loci, while all previously identified loci were retained (Supplementary Figs. 1 and 2, Supplementary Table 1).

## Discussion

The present study replicated the previously identified locus on chromosome 4 for both *N*-desmethylclozapine and the CLZ-to-*N*-desmethylclozapine metabolic ratio, independent of smoking habits<sup>9,24</sup>. When adjusting for known smoking habits, which have a major effect on CLZ metabolism<sup>13,14</sup>, we identified a novel locus in the *NFIB* gene (*rs28379954* T>C) associated with a significant reduction in CLZ serum concentrations.

According to our findings in the GWAS analysis adjusting for smoking habits, heterozygous carriers of the minor *rs28379954* C allele are associated with a substantial decrease in CLZ serum concentration, i.e., an estimated reduction of 37.6% in the dose-adjusted serum concentration of CLZ compared with homozygous

carriers of the wild type allele (*rs28379954 TT*). This effect was only identified after adjusting for individually ascertained smoking habits, and not when controlling for a polygenic score proxy measure of smoking, as performed in the recent GWAS by Pardiñas et al.<sup>16</sup> Smokers comprised 61.1% of the included patient population, a typical prevalence for a sample of individuals with schizophrenia<sup>36</sup>. These results suggest that patients heterozygous for the minor *C* allele, representing 7.6% of the current study population, are at risk of subtherapeutic CLZ serum concentrations at standard recommended CLZ dosing. Indeed, the patients carrying the minor *C* allele had significant lower absolute CLZ serum concentration regardless of their smoking habits; however, smoking was shown to reinforce this effect. Since smokers are already prone to lower CLZ serum concentrations, smokers carrying the minor *C* allele may represent a subpopulation particularly vulnerable to suboptimal concentration and treatment failure. This is supported by the more than threefold higher proportion of CLZ concentration measurements below the therapeutic reference in smoking *C* carriers vs. nonsmoking *TT* carriers in our population. Thus, genotyping of *rs28379954* may potentially aid individualized dosing of CLZ in clinical practice to ensure optimal treatment response.

The *rs28379954 T>C* polymorphism was not associated with the pharmacokinetic variability of either the dose-adjusted serum concentration of *N*-desmethylclozapine or the CLZ-to-*N*-desmethylclozapine metabolic ratio in the GWAS analyses, suggesting that the mechanisms behind the reduced CLZ levels may be related to an increased formation of other CLZ metabolites than *N*-desmethylclozapine. The SNP *rs28379954* is located in an intronic region of the *NFIB* gene, which encodes a transcription factor (NFIB) that is expressed in several tissues, such as lungs, liver, kidneys, and brain<sup>37,38</sup>. NFIB plays a role during embryonic development by initiating tissue differentiation in the fetus<sup>38</sup>, but has not previously been reported to regulate the expression or activity of drug-metabolizing enzymes.

From the findings of the present study, it may be hypothesized that the *rs28379954 T>C* polymorphism affects the transcription of genes involved in CLZ metabolism. Based on observations of a quantitatively higher and statistically stronger difference in CLZ levels in smoking than nonsmoking *CT* vs. *TT* carriers (Fig. 2b), the most relevant hypothesis is that *NFIB* modulates the expression of the ligand-activated transcription factor aryl hydrocarbon receptor (AhR), which is induced by smoking and important for CLZ metabolism. However, experimental studies are required to elucidate the functional impact of the *NFIB* gene on CLZ metabolism. In addition, it should be investigated if *rs28379954* genotype is associated with formation of metabolites with potential impact on CLZ toxicity.

Recently, Pardiñas et al. showed an association between the SNP *rs2472297* in *CYP1A1/1A2* locus and CLZ serum concentration in a British cohort<sup>16</sup>, which was not replicated in the present study. In their study, however, smoking habits data were not available, and polygenic risk scores for smoking were therefore used as a proxy measure of smoking habits instead. Notably, cigarette smoking has been shown to reduce CLZ levels by ~30–50% by inducing *CYP1A1/1A2*<sup>23–25</sup>, and up to 60% of individuals with schizophrenia are smokers<sup>36</sup>. Given the large effect of smoking on CLZ metabolism, and the high prevalence of smoking in people with schizophrenia, the differences in accounting for smoking habits may explain the discrepancy between the studies. It should be noted that in the present study, the SNP *rs2472297* in the *CYP1A1/1A2* locus had a call rate of 64.2% and was therefore excluded from GWAS analyses. However, no other SNPs within the previously defined *CYP1A1/1A2* locus (15:74817689–75404506) showed association signals, without and after adjusting for smoking habits, respectively. These results suggest that caution is required when polygenic risk score is utilized as a proxy for smoking.

Similarly to Pardiñas et al.<sup>16</sup>, we found that the genetic variants in the *UGT2B10* locus are associated with the *N*-desmethylclozapine serum concentration, and hence the CLZ-to-*N*-desmethylclozapine metabolic ratio, irrespective of smoking habits. In particular, the intergenic minor alleles of the lead SNPs *rs1670747* (4:69601886–70138176) and *rs2926038* (4:69544277–70340670) are responsible for decreasing *N*-desmethylclozapine levels and increasing the CLZ-to-*N*-desmethylclozapine ratio by ~21% and 25%, respectively. *UGT2B10* is involved in *N*-glucuronidation of a variety of psychotropic drugs including CLZ, and likely *N*-desmethylclozapine<sup>39</sup>. These variants might be of potential relevance to genotype before starting CLZ treatment, since increasing *N*-desmethylclozapine level has been associated with positive effects of CLZ treatment<sup>6,40</sup>, but further studies are required to elucidate if *UGT2B10* genotype has an impact on the clinical response of CLZ.

Our results add support to the hypothesis that the genetic architecture of CLZ metabolism is driven by few variants with large effect sizes<sup>16,18</sup>, in line with typical pharmacogenetics findings<sup>19,20</sup>. Despite our sample being approximately sixfold smaller than the British GWAS<sup>16</sup>, we robustly replicate their findings that the *UGT2B10* locus on chromosome 4 is associated with *N*-desmethylclozapine serum concentration, and hence the CLZ-to-*N*-desmethylclozapine metabolic ratio. These results, in addition to the 97% overlap in *rs28379954 T>C* genotypes by GWAS and targeted Taqman analyses, increase our confidence that the novel locus we identified to be associated with CLZ serum concentration (*rs28379954*), after adjusting for smoking habits, is a true positive and not a



spurious association due to winner's curse. Additional studies are still required to replicate this association in independent cohorts of CLZ-treated schizophrenia patients with known smoking habits.

The use of TDM data for research purposes is associated with some methodological limitations, where potential nonadherence is a relevant issue creating 'nonbiological' variability in the concentration phenotype. Unknown co-medications may also alter CLZ metabolism and concentration, but in the current material, co-prescription of the most relevant interacting drugs was accounted for, i.e., phenytoin, phenobarbital, carbamazepine, and fluvoxamine, which all change the CLZ concentration several-fold<sup>41,42</sup>. In addition, information on comorbidity, weight, and potential consumption of dietary supplements potentially affecting concentration phenotypes was not available. The amount of cigarette consumption could not be obtained, but the enzyme-inducing effect of smoking is little affected by the degree of cigarette consumption<sup>25</sup>.

In conclusion, the present GWAS on CLZ pharmacokinetics, adjusting for known smoking habits, identified a novel variant in the *NFIB* gene (*rs28379954*) associated with a significant reduction in CLZ levels. These results suggest that patients carrying the minor *C* allele are at risk of subtherapeutic CLZ serum concentrations, particularly in smokers. Thus, genotyping of *rs28379954* may potentially aid individualized dosing of CLZ in clinical practice.

#### Acknowledgements

This project received funding from the South-Eastern Norway Regional Health Authority (Grant number 2016097, 2013-123), the NIH (NS057198, EB00790) and the Research Council of Norway (229129, 213837, 223273, and 226971). The authors thank Francesco Bettella, PhD, for sharing his helpful knowledge.

#### Author details

<sup>1</sup>Center for Psychopharmacology, Diakonhjemmet Hospital, Oslo, Norway.

<sup>2</sup>CoE NORMENT, Division of Mental Health and Addiction, Oslo University Hospital, and Institute of Clinical Medicine, University of Oslo, Oslo, Norway.

<sup>3</sup>Department of Medical Genetics, Oslo University Hospital, Oslo, Norway.

<sup>4</sup>NORMENT, Department of Clinical Science, University of Bergen, Bergen, Norway.

<sup>5</sup>Department of Life Sciences and Health, Oslo Metropolitan University, Oslo, Norway. <sup>6</sup>Department of Pharmacy, Section for Pharmacology and Pharmaceutical Biosciences, University of Oslo, Oslo, Norway

#### Conflict of interest

O.A.A. has received a speaker's honorarium from Lundbeck and is a consultant for HealthLytix Inc. E.M. received speaker's honorarium from Lilly, Lundbeck and Otsuka. The other authors have no conflicts of interest to declare.

#### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Supplementary Information** accompanies this paper at (<https://doi.org/10.1038/s41398-020-00888-1>).

Received: 20 December 2019 Revised: 9 June 2020 Accepted: 9 June 2020  
Published online: 19 June 2020

#### References

- Leucht, S. et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet* **382**, 951–962 (2013).
- Joober, R. & Boksa, P. Clozapine: a distinct, poorly understood and under-used molecule. *J. Psychiatry Neurosci.* **35**, 147–149 (2010).
- FDA. *U.S. Food and Drug Administration, Clozaril, label approval, NDA no. 019758* (FDA, 2014).
- Atkin, K. et al. Neutropenia and agranulocytosis in patients receiving clozapine in the UK and Ireland. *Br. J. Psychiatry* **169**, 483–488 (1996).
- Alvir, J. M., Lieberman, J. A., Safferman, A. Z., Schwimmer, J. L. & Schaaf, J. A. Clozapine-induced agranulocytosis. Incidence and risk factors in the United States. *N. Engl. J. Med.* **329**, 162–167 (1993).
- Smith, R. L. et al. Correlation between serum concentrations of N-desmethylclozapine and granulocyte levels in patients with schizophrenia: a retrospective observational study. *CNS Drugs* **31**, 991–997 (2017).
- Citrome L, McEvoy J. P., Saklad S. R. A guide to the management of clozapine-related tolerability and safety concerns. *Clin Schizophr Relat Psychoses.* **10**, 163–177 (2016).
- Dain, J. G., Nicoletti, J. & Ballard, F. Biotransformation of clozapine in humans. *Drug Metab. Dispos.* **25**, 603–609 (1997).
- Dragovic, S., Gunness, P., Ingelman-Sundberg, M., Vermeulen, N. P. & Commandeur, J. N. Characterization of human cytochrome P450s involved in the bioactivation of clozapine. *Drug Metab. Dispos.* **41**, 651–658 (2013).
- Wicinski, M. & Weclawicz, M. M. Clozapine-induced agranulocytosis/granulocytopenia: mechanisms and monitoring. *Curr. Opin. Hematol.* **25**, 22–28 (2018).
- Pirmohamed, M. & Park, K. Mechanism of clozapine-induced agranulocytosis: current status of research and implications for drug development. *CNS Drugs* **7**, 139–158 (1997).
- Pirmohamed, M., Williams, D., Madden, S., Templeton, E. & Park, B. K. Metabolism and bioactivation of clozapine by human liver in vitro. *J. Pharmacol. Exp. Ther.* **272**, 984–990 (1995).
- Rostami-Hodjegan, A. et al. Influence of dose, cigarette smoking, age, sex, and metabolic activity on plasma clozapine concentrations: a predictive model and nomograms to aid clozapine dose adjustment and to assess compliance in individual patients. *J. Clin. Psychopharmacol.* **24**, 70–78 (2004).
- Couchman, L., Bowskill, S. V., Handley, S., Patel, M. X. & Flanagan, R. J. Plasma clozapine and norclozapine in relation to prescribed dose and other factors in patients aged <18 years: data from a therapeutic drug monitoring service, 1994–2010. *Early Inter. Psychiatry* **7**, 122–130 (2013).
- Krivoy, A., Gaughran, F., Weizman, A., Breen, G. & MacCabe, J. H. Gene polymorphisms potentially related to the pharmacokinetics of clozapine: a systematic review. *Int. Clin. Psychopharmacol.* **31**, 179–184 (2016).
- Pardinas, A. F. et al. Pharmacogenomic variants and drug interactions identified through the genetic analysis of clozapine metabolism. *Am. J. Psychiatry* **176**, 477–486 (2019).
- Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
- Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J. & Richards, J. B. Genetic architecture: the shape of the genetic contribution to human traits and disease. *Nat. Rev. Genet.* **19**, 110–124 (2018).
- Zanger, U. M. & Schwab, M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol. Ther.* **138**, 103–141 (2013).
- Sim, S. C. & Ingelman-Sundberg, M. The Human Cytochrome P450 (CYP) Allele Nomenclature website: a peer-reviewed database of CYP variants and their associated effects. *Hum. Genomics.* **4**, 278–281 (2010).
- Goldstein, J. I. et al. Clozapine-induced agranulocytosis is associated with rare HLA-DQB1 and HLA-B alleles. *Nat. Commun.* **5**, 4757 (2014).
- Legge, S. E. et al. Genome-wide common and rare variant analysis provides novel insights into clozapine-associated neutropenia. *Mol. Psychiatry* **22**, 1502–1508 (2017).
- Nebert, D. W., Dalton, T. P., Okey, A. B. & Gonzalez, F. J. Role of aryl hydrocarbon receptor-mediated induction of the CYP1 enzymes in environmental toxicity and cancer. *J. Biol. Chem.* **279**, 23847–23850 (2004).
- Smith, R. L., Kyles, L., Haslemo, T., Andreassen, O. A. & Molden, E. Reduction in N-desmethylclozapine level is determined by daily dose but not serum concentration of valproic acid-indications of a presystemic interaction mechanism. *Ther. Drug Monit.* **41**, 503–508 (2019).

25. Haslemo, T., Eikeseth, P. H., Tanum, L., Molden, E. & Refsum, H. The effect of variable cigarette consumption on the interaction with clozapine and olanzapine. *Eur. J. Clin. Pharmacol.* **62**, 1049–1053 (2006).
26. Hiemke, C. et al. AGNP consensus guidelines for therapeutic drug monitoring in psychiatry: update 2011. *Pharmacopsychiatry* **44**, 195–235 (2011).
27. Perry, P. J., Miller, D. D., Arndt, S. V. & Cadoret, R. J. Clozapine and nortriptyline plasma concentrations and clinical response of treatment-refractory schizophrenic patients. *Am. J. Psychiatry* **148**, 231–235 (1991).
28. Miller, D. D., Fleming, F., Holman, T. L. & Perry, P. J. Plasma clozapine concentrations as a predictor of clinical response: a follow-up study. *J. Clin. Psychiatry* **55**, 117–121 (1994).
29. Seeman, P. Clozapine, a fast-off-D2 antipsychotic. *ACS Chem. Neurosci.* **5**, 24–29 (2014).
30. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
31. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
32. Loh, P. R. et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).
33. Loh, P. R., Palamara, P. F. & Price, A. L. Fast and accurate long-range phasing in a UK Biobank cohort. *Nat. Genet.* **48**, 811–816 (2016).
34. Das, S. et al. Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
35. Detry, M. A. & Ma, Y. Analyzing repeated measurements using mixed models. *JAMA* **315**, 407–408 (2016).
36. de Leon, J. & Diaz, F. J. A meta-analysis of worldwide studies demonstrates an association between schizophrenia and tobacco smoking behaviors. *Schizophr. Res.* **76**, 135–157 (2005).
37. Steele-Perkins, G. et al. The transcription factor gene Nfib is essential for both lung maturation and brain development. *Mol. Cell Biol.* **25**, 685–698 (2005).
38. Chaudhry, A. Z., Lyons, G. E. & Gronostajski, R. M. Expression patterns of the four nuclear factor I genes during mouse embryogenesis indicate a potential role in development. *Dev. Dyn.* **208**, 313–325 (1997).
39. Lu, D., Xie, Q. & Wu, B. N-glucuronidation catalyzed by UGT1A4 and UGT2B10 in human liver microsomes: assay optimization and substrate identification. *J. Pharm. Biomed. Anal.* **145**, 692–703 (2017).
40. Sur, C. et al. N-desmethylclozapine, an allosteric agonist at muscarinic 1 receptor, potentiates N-methyl-D-aspartate receptor activity. *Proc. Natl Acad. Sci. USA* **100**, 13674–13679 (2003).
41. Raitasuo, V., Lehtovaara, R. & Huttunen, M. O. Carbamazepine and plasma levels of clozapine. *Am. J. Psychiatry* **150**, 169 (1993).
42. Perucca, E. Clinically relevant drug interactions with antiepileptic drugs. *Br. J. Clin. Pharmacol.* **61**, 246–255 (2006).