HLA-DQB1 6672G>C (rs113332494) is associated with clozapine-induced neutropenia and agranulocytosis in individuals of European ancestry

Bettina Konte1, James T. R. Walters2, Dan Rujescu1, Sophie E. Legge2, Antonio F. Pardiñas2, Dan Cohen3, Munir Pirmohamed4, Jari Tiihonen5,6, Annette M. Hartmann1, Jan P. Bogers6, Jan van der Weide8,9, Karen van der Weide8, Anu Putkonen5, Eila Repo-Tiihonen5, Tero Hallikainen5, Ed Silva6, Oddur Ingimarsson10,11, Engilbert Sigurdsson10,11, James L. Kennedy12,13,14, Patrick F. Sullivan15,16,17, Marcella Rietschel18, Gerome Breen19, Hreinn Stefansson20, Kari Stefansson20, David A. Collier21,22, Michael C. O'Donovan23 and Ina Giegling1

Abstract

The atypical antipsychotic clozapine is the only effective medication for treatment-resistant schizophrenia. However, it can also induce serious adverse drug reactions, including agranulocytosis and neutropenia. The mechanism by which it does so is largely unknown, but there is evidence for contributing genetic factors. Several studies identified HLA-DQB1 variants and especially a polymorphism located in HLA-DQB1 (6672G>C, rs113332494) as associated with clozapine-induced agranulocytosis and neutropenia. We analysed the risk allele distribution of SNP rs113332494 in a sample of 1396 controls and 178 neutropenia cases of which 60 developed agranulocytosis. Absolute neutrophil counts of 500/mm³ and 1500/mm³ were used for defining agranulocytosis and neutropenia cases, respectively. We also performed association analyses and analysed local ancestry patterns in individuals of European ancestry, seeking replication and extension of earlier findings. HLA-DQB1 (6672G>C, rs113332494) was associated with neutropenia (OR = 6.20, P = 2.20E-06) and agranulocytosis (OR = 10.49, P = 1.83E-06) in individuals of European ancestry. The association signal strengthened after including local ancestry estimates (neutropenia: OR = 10.38, P = 6.05E-08; agranulocytosis: OR = 16.31, P = 1.39E-06), with effect sizes being considerably larger for agranulocytosis. Using local ancestry estimates for prediction, the sensitivity of rs113332494 increased from 11.28 to 55.64% for neutropenia and from 16.67 to 53.70% for agranulocytosis. Our study further strengthens the evidence implicating HLA-DQB1 in agranulocytosis and neutropenia, suggesting components of the immune system as contributing to this serious adverse drug reaction. Using local ancestry estimates might help in identifying risk variants and improve prediction of haematological adverse effects.

Introduction

Clozapine is the most effective antipsychotic medication, particularly for treatment resistant patients where clozapine is the only evidence based option1–3. Compared to alternative antipsychotics it offers superior symptom control, longer duration of treatment, reduced hospitalisation, improved cognition, better work and social function, a higher quality of life and reductions in violence and...
Benign ethnic neutropenia (BEN) is characterized by exposure to clozapine as well as studies identifying risk factors. Severe and rapid course of agranulocytosis can occur after the neutrophil. The possibility of an underlying immunological mechanism is indicated by the finding that agranulocytosis seems to be the result of a toxic event involving an unstable intermediate metabolite, a nitrenium ion. Binding of this ion to neutrophil proteins could lead to a disturbance of neutrophil function or hapten formation, thereby triggering the immune destruction of neutrophils. Neutropenia and agranulocytosis describe absolute neutrophil counts (ANCs) below 1500 and 500 cells/mm³, respectively. In 1975, treatment of 2260 patients with clozapine was associated with agranulocytosis in 16 patients, 8 of whom died, leading to clozapine withdrawal from the market in many countries. Clozapine was still prescribed in German-speaking countries and Finland and reintroduced in other countries in 1990 under strict conditions which required regular blood count monitoring. The incidence of neutropenia and agranulocytosis has recently been reported as 3.8% (95% CI: 2.7–5.2%) and 0.9% (95% CI: 0.7–1.1%), respectively.

The mechanisms underlying clozapine-induced hematological adverse effects are largely unknown. There is evidence for toxic, immunological and genetic mechanisms indicating a multi-factorial pathogenesis. Agranulocytosis seems to be an idiosyncratic reaction, is dose-independent and is influenced by genetic or environmental factors. Neutropenia has been proposed to be a direct toxic effect of the parent drug, agranulocytosis seems to be the result of a toxic event involving an unstable intermediate metabolite, a nitrenium ion. Binding of this ion to neutrophil proteins could lead to a disturbance of neutrophil function or hapten formation, thereby triggering the immune destruction of the neutrophil. The possibility of an underlying immunological mechanism is indicated by findings that a more severe and rapid course of agranulocytosis can occur after re-exposure to clozapine as well as studies identifying risk variants located in the major histocompatibility complex (MHC). Benign ethnic neutropenia (BEN) is characterized by low neutrophil counts not associated with an increased risk of infections and shows an increased prevalence in individuals of African or Middle Eastern ancestry, potentially due to the contribution of genetic variants that are rare elsewhere. These ancestry dependent differences, and the fact that other adverse drug reactions also have genetic mediators, present a strong argument that genetic factors explain, at least partly, inter-individual variability and differences in bioactivation, detoxification and immunological response.

Very few genetic markers have been clearly associated with clozapine induced agranulocytosis or neutropenia. Identified genetic risk factors include specific human leucocyte antigen (HLA) variants and haplotypes. The HLA region was first implicated in small studies that identified the HLA allele DQB1*05:02 as risk factor for agranulocytosis in patients of European ancestry. Evidence strengthened when Athanasiou et al. analysed 74 candidate genes in a cohort recruited from sites within the USA, Russia and South Africa in order to replicate the findings in a second cohort of non-Jewish schizophrenia patients of German descent. Only variants in HLA-DQB1 were significantly associated with agranulocytosis in both cohorts. A subsequent refinement analysis revealed association to a single SNP (HLA-DQB1 6672G>C, rs113332494, OR = 16.9), which is in strong linkage disequilibrium (LD) with DQB1*05:02. Although two other smaller studies were not able to replicate these findings, substantial evidence for the importance of the HLA-DQB1 region has also come from the largest genome-wide association studies (GWAS) conducted to date. The Clozapine-Induced Agranulocytosis Consortium (CIAC) analysed imputed HLA alleles and amino acid changes in 162 patients and 4319 controls. Clozapine treated cases (ANC ≤ 1000) were compared to two populations: patients with schizophrenia treated with clozapine for at least 1 year without developing neutropenia (ANC > 1500) and population controls. Two genome-wide associated variants were identified, including a single non-synonymous variant in HLA-DQB1 (126Q) (OR = 0.19, P = 4.7E–14), for which the most common HLA allele at high-risk is HLA-DQB1*05:02. Moreover, rs113332494 was genotyped in a British sample of 60 clozapine-treated cases and 305 clozapine-treated controls and again, significantly associated with neutropenia (OR = 15.6, P = 0.015). The association strengthened (OR = 38.1, P = 0.0079) using a stricter ANC threshold of 1000 mm⁻³. The variant rs113332494 is located in the HMC region on chromosome 6, a region with extensive LD spanning more than 3 Mb. The region is highly polymorphic and this variability is maintained by positive natural selection. Population-specific recombination sites might contribute to the high diversity of haplotypes and it has been shown that the region exhibits an excess of local ancestry switches in non-European populations. A local ancestry switch is defined as a genomic region exhibiting a change in ancestry, such that while most haplotypes across the broader region originate from one ancestral population, local haplotypes in a subregion might be derived from a different population as a result of recombination. Interestingly, an increase in East Asian ancestry haplotypes in the HLA-DQB1 region was found.
in a neutropenia case–control sample of British ancestry. The switch was seen in cases and controls, but nevertheless, local ancestry switches could be additional relevant factors influencing disease risk.

In order to replicate and refine previous findings and clarify the influence of the marker on the risk of neutropenia and agranulocytosis, we analysed the risk allele distribution, global and local population structure and association of the marker rs113332494 with neutropenia and agranulocytosis in 1396 controls and 178 neutropenia cases of which 60 developed agranulocytosis. To the best of our knowledge, this is the first study which determines the predictive power of rs113332494 incorporating local ancestry information.

Methods

Sample description
The CRESTAR sample comprised 1576 clozapine treated individuals, 1396 controls and 180 neutropenia cases of which 61 developed agranulocytosis (Supplementary Tables S1 and S2). Any psychiatric diagnosis was allowed; the primary psychiatric diagnosis was schizophrenia. ANC thresholds of 500/mm³ or 1500/mm³ were used for defining agranulocytosis and neutropenia cases, respectively. A small number of cases (N = 22, 3 agranulocytosis and 19 neutropenia cases) were classified by clinical judgement due to a precipitous drop in ANC while being intensively monitored. A detailed description of all contributing groups and inclusion criteria can be found in Supplementary Material.

Overlap to other studies
Overlaps to CIAC and CLOZUK were examined by identity-by-descent estimations. Using genome-wide data we identified a small number of individuals (eight controls, seven cases) of European ancestry as duplicates in CIAC. These individuals were not excluded, as the marker, which is the focus of the present study, was not imputed at adequate quality in the GWAS conducted by that consortium. Overlap to CLOZUK cases was checked by identifying duplicates and first-degree relatives using ~1000 markers. One neutropenia and one agranulocytosis case were excluded from association and prediction analyses as they were found to be part of the CLOZUK study. See Supplementary Material for details and a sensitivity analysis excluding the individuals overlapping with CIAC.

Genotyping/Imputation
Individuals were genotyped on Illumina HumanOmniExpress-12v1 and standard quality checks were applied (see Supplementary Methods). Pre-phasing was performed with SHAPEIT2 and imputation with IMPUTE2. We used the reference panel 1000 Genomes Project Phase 1 (August 2012), provided by the SHAPEIT2 authors, for both steps.

Population structure
Population structure was analyzed by EIGENSTRAT (Supplementary Fig. S1) and an admixture analysis in a supervised setting using African (AFR), East Asian (EAS) and European (EUR) 1000 Genomes Project (1 KG) super populations as reference using ADMIXTURE v1.3.0. African (N = 48) and European (N = 1006) subsamples were defined as individuals with at least 80% AFR or EUR ancestry fractions. All other individuals were defined as admixed (N = 522).

Statistical analysis
The statistical analysis included 1574 individuals (1396 controls, 178 neutropenia cases of whom 60 developed agranulocytosis) once the two duplicate individuals from CLOZUK were excluded (see above).

Risk allele distribution of marker rs113332494 was examined based on best guess genotypes derived from PLINK v1.933,34 using the following parameters: --hard-call-threshold 0.3, --geno 0.1, --maf 0.01, --hwe 1E-05. Genotypes with posterior probabilities below 0.7 were set to missing.

Association analysis was performed on individuals of European ancestry using logistic regression in PLINK v1.933,34 applying an additive model corrected for relevant principal components inferred by EIGENSTRAT v6.0.135. Supplementary Fig. S2 shows the scatterplot of first two principal components.

The main association analyses were conducted on neutropenia (ANC ≤ 1500) and agranulocytosis (ANC ≤ 500). To investigate the contribution of agranulocytosis cases, we excluded all cases with ANC ≥ 500 and tested a range of ANC thresholds (500–1500 in steps of 100). We also analysed different ANC thresholds excluding neutropenia cases for a range of ANC thresholds in steps of 100 to determine how rs113332494 behaves as a predictor.

Estimating local ancestry haplotypes across the extended MHC region (25–35 Mb) was performed with ELAI using the 1KG super populations EUR, EAS and AFR as reference. We set the number of upper clusters, lower clusters and mixture generations to 3, 15 and 200 as recommended by the software author. Inference was performed with 20 EM steps and averaged over ten replicates. The contribution of each of the three ancestral populations is given as an allele dosage estimate ranging from 0 to 2, corresponding respectively to no or complete contribution of the reference haplotype. Means for the region of interest were averaged across 1 kb up- and downstream from rs113332494. Mean dosage estimates of a single ancestral population and normed differences of two populations [POP1/(POP1 + POP2) − POP2/(POP1 + POP2)] were compared between groups. Permutation tests were used to derive the significance of
group comparisons. We conducted 10,000 permutations shuffling group membership while keeping the group size constant.

Best guess genotypes with maximum posterior probability > 0.7 were used to determine the predictive power of rs113332494. The risks in the high- and low-risk group as well as positive and negative predictive value were calculated with respect to the assumed risk of agranulocytosis (0.9%) and neutropenia (3.8%). Individuals were classified as “high risk” when carrying the risk allele G, additionally local ancestry information was used in a second model.

Power calculations were performed with Quanto using a log-additive model. Power was derived using the estimated allele frequencies and case-control ratios setting the prevalence for neutropenia and agranulocytosis to 3.8 and 0.9%. SNP annotations, conservation scores and LD plots were derived with SNiPA. For annotation of marker effects see https://snipa.helmholtz-muenchen.de/snipa3/index.php?task=supplement. Further details are provided in Supplementary Methods.

Results
Risk allele distribution

We examined rs113332494 risk allele (G) carriers and the distribution of risk alleles in individuals of different ancestries (N = 1574). The study sample was mainly comprised of individuals of European (N = 1004, 133 cases) ancestry, but also included individuals of African (N = 48, 15 cases) and admixed (N = 522, 30 cases) ancestry. Eight controls had maximum genotype posterior probabilities below the threshold (0.7) and were set to missing. We found 68 (4.3%) risk allele carriers in the 1566 remaining individuals, 52 (3.7% of 1396) controls and 16 (9.0% of 178) cases of which 9 (15.0% of 60) developed agranulocytosis. Fourteen cases were confirmed by neutrophil counts, whereas two cases (one neutropenia, one agranulocytosis) were confirmed by clinical judgement (blind to genotype). Of the 68 risk allele carriers, 66 had a diagnosis of schizophrenia, the other 2 had diagnoses of bipolar affective disorder and psychosis. We found 45 (15 cases) risk allele carriers of European, 6 (1 case) of African and 17 (0 cases) of mixed ancestry (Supplementary Fig. S3, Supplementary Table S3). The risk allele frequency for all 1566 individuals was 2.2%, which is in line with all non-Asian 1KG phase 1 populations (2.4%). Finnish (1.8%) and EUR (2.3%) frequencies also agree with the estimates of the reference populations (FIN = 1.6%, EUR = 2.5%). The risk allele in the European sample was three times more frequent in cases than in controls. In contrast, the risk allele was more common in African ancestry controls compared with cases. We found no admixed ancestry cases carrying the risk allele. All subsequent analyses were conducted in individuals of European ancestry only.

Local ancestry of the HLA-DQB1 region

We estimated local ancestry patterns in the extended MHC region and investigated the contribution of African, East Asian and European haplotypes in the region surrounding rs113332494 in individuals of European ancestry (N = 1004). In general, European dosage estimates dominated, but decreased in the HLA-DQB1 region. Cases had higher African and lower East Asian local ancestry estimates than controls (Supplementary Fig. S4). Additional stratification by rs113332494 revealed two striking patterns. First, risk allele carriers exhibited considerably higher East Asian than European ancestry (EAS = 1.34,
EUR = 0.44) (Fig. 1A, Supplementary Fig. S5A, C, E), whereas in non-carriers EUR local haplotypes dominated (EAS = 0.64, EUR = 0.89) (Fig. 1B, Supplementary Fig. S5B, D, F). Second, in the group of non-risk allele carriers, cases exhibited higher African and lower East Asian dosage estimates compared to controls (AFR controls = 0.45, AFR neutropenia = 0.60; EAS controls = 0.67, EAS neutropenia = 0.45) (Fig. 1B, Supplementary Fig. S5B, D, F). Patterns were similar for neutropenia and agranulocytosis cases.

The significance of group differences assessed by permutation testing is presented in Table 1. Risk allele carriers had significantly higher East Asian ancestry estimates than non-risk allele carriers (P < 1.00E−04). In addition, we found significant differences of East Asian (P < 1.00E−04) and African (P = 1.20E−03) ancestry in non-risk allele carriers between neutropenia cases and controls. The difference of AFR and EAS estimates also differed significantly between neutropenia cases and controls regardless of risk allele status, but this effect was driven by the difference in non-risk allele carriers, as there was no significant difference in risk allele carriers. Results were similar when restricting to agranulocytosis cases (Supplementary Tables S4 and S5).

### Association analyses

We performed an association analysis of the marker HLA-DQB1 6672G>C (rs113332494) in a European sample comprised of 871 controls, 133 neutropenia and 54 agranulocytosis cases. The risk allele was more frequent in cases (neutropenia = 5.4%, agranulocytosis = 8.0%, controls = 1.8%) and the variant was significantly associated with neutropenia (OR = 6.20, 95% CI: 2.91–13.21, P = 2.20E−06) and agranulocytosis (OR = 10.49, 95% CI: 3.99–27.56, P = 1.83E−06) (Table 2).

Including AFR and EAS dosage estimates as additional covariates in association analyses gave higher odds ratio estimates (neutropenia: OR = 10.38, 95% CI: 4.45–24.22, P = 6.05E−08; agranulocytosis: OR = 16.31, 95% CI: 5.25–50.68, P = 1.39E−06), with the P value in the neutropenia analysis almost reaching genome-wide significance (Table 2). We had 100% power to detect effects in the size of the estimated odds ratios given the estimated effect size (Supplementary Table S6).

### Linear effect of ANC

The odds ratio estimates were higher for agranulocytosis than for neutropenia. To determine if the effect on neutropenia is driven by inclusion of those with agranulocytosis, we excluded all cases with ANC ≥ 500 and tested a range of ANC thresholds (500–1500 in steps of 100) (Supplementary Table S7). The association diminished when excluding all agranulocytosis cases but was still significant, odds ratios decreased slightly with higher ANC thresholds (Supplementary Table S7, Supplementary Fig. S6). To determine how rs113332494 behaves as a predictor we also analysed different ANC thresholds excluding neutropenia cases for a range of ANC thresholds in steps of 100 (Supplementary Table S8). The odds ratios nearly followed a linear trend, where estimates decreased with higher ANC thresholds (Fig. 2).

### Predictive test characteristics

Predictive power was established by comparing affection to risk allele status, where heterozygote and homozygote risk allele carriers were combined. The HLA-DQB1 variant had high specificity, but low sensitivity in predicting risk for agranulocytosis and neutropenia (Table 3). Sensitivity was higher in agranulocytosis (17%) cases in comparison to those with neutropenia (11%). Classifying all individuals carrying the rs113332494 risk allele as “high

### Table 1 Local ancestry estimates in individuals of European ancestry for ancestral populations AFR, EAS and EUR.

| Ancestral population | Non-risk vs risk allele carriers | Non-risk allele carriers: controls vs cases |
|----------------------|----------------------------------|--------------------------------------------|
|                      | Dosage | Difference | P value | dosage | difference | P value |
| AFR                  | 0.470/0.213 | 0.257 | <1.00E−04 | 0.452/0.600 | −0.148 | 1.20E−03 |
| EAS                  | 0.641/1.344 | −0.703 | <1.00E−04 | 0.668/0.452 | 0.216 | <1.00E−04 |
| EUR                  | 0.889/0.443 | 0.446 | <1.00E−04 | 0.881/0.948 | −0.067 | 8.31E−02 |
| AFR–EAS              | −0.154/−0.726 | 0.573 | <1.00E−04 | −0.193/0.141 | −0.334 | <1.00E−04 |
| AFR–EUR              | −0.308/−0.351 | 0.043 | <4.00E−01 | −0.322/0.224 | −0.098 | 5.77E–02 |
| EAS–EUR              | −0.162/0.504 | −0.666 | <1.00E−04 | −0.138/0.353 | 0.216 | 9.00E−04 |

Dosage means for single ancestral populations and normed differences for two populations were compared between risk (N = 45) and non-risk allele (N = 956) carriers as well as between controls (N = 838) and neutropenia cases (N = 118) not carrying the risk allele. P values were determined by permutation tests (N = 10,000). Significant results are in bold.
risk” and classifying non-risk allele carriers with respect to
the difference in African and East Asian ancestry esti-
mates (positive: higher AFR = "high risk", negative: lower
AFR = “low risk”) resulted in a substantial increase in
sensitivity (neutropenia = 56%, agranulocytosis = 54%),
but also a decrease in specificity (Table 3).

Discussion
Our study shows an association between rs113332494
(HLA-DQB1 6672G>C) and clozapine-induced neu-
topenia and agranulocytosis, further adding to the evi-
dence of the importance of this genetic variant. The
estimated odds ratio for agranulocytosis was higher (OR
= 10.49, \(P = 1.83 \times 10^{-06}\)) than for neutropenia (OR = 6.20,

\[P = 2.20 \times 10^{-06}\]). Exclusion of individuals with agrano-
lucytosis revealed a residual, but diminished association
signal, in those with neutropenia. Additional association
analyses including cases with varying ANC thresholds
(500–1500) showed that effect sizes increased linearly
with stricter ANC thresholds. Most likely, this indicates
that neutropenia cases are composed of two different
populations: (1) individuals who would have developed
agranulocytosis if clozapine had not been stopped and (2)
individuals who show decreased ANC because of random
fluctuation or other factors not related to clozapine-
induced agranulocytosis. In this scenario, it might be the
case that the number of “true” agranulocytosis cases
increases as ANC counts decline.

Given the complex recombination pattern of the MHC
region, it is unclear if rs113332494, which is intronic, is
itself the causal variant rather than simply indicating an
indirect association with another variant with which
it is in LD. Annotations revealed three SNPs in
strong LD (\(r^2 > 0.8\)) to rs113332494 in the European
population: rs41542812 (\(r^2 = 0.95\), direct transcript
effect), rs113664950 (\(r^2 = 0.89\), putative regulatory/tran-
script effect) and rs114262759 (\(r^2 = 0.84\), putative tran-
script effect). The missense variant rs41542812 is a
glutamine to histidine polymorphism at position 158.
Rs113664950 is annotated as expressed promoter (FAN-
TOM5) for HLA-DQB1 and promoter-associated distal DHS
(enhancer) for PBX2, TAP1, HLA-DRB9, HLA-DQA1,
HLA-DPA1, HLA-DQB1, HLA-DQA1, PBX2, HLA-DRB5,
HLA-DOA and HLA-DPA1 (ENCODE). Interestingly, this
variant is also associated with bare lymphocyte syndrome
type I. Given the high LD between the investigated variant
and rs113664950, an association of rs113332494 with this
syndrome seems possible. One of twenty one variants in
middle to high LD (0.6 < \(r^2 < 0.8\)) showed a direct regu-
latory effect: rs35139945 (\(r^2 = 0.74\), which is a cis-eQTL
variant for HLA-DQA1, HLA-DRB5 and HLA-DRB1.

Among other tissues differentially regulated, gene

### Table 2 Results of association analyses of rs113332494 for neutropenia and agranulocytosis including 1004 and 925
individuals of European ancestry.

| Phenotype     | \(N\) cases | \(N\) controls | Info | Freq cases | freq controls | Global ancestry | Global and local ancestry |
|---------------|-------------|----------------|------|------------|---------------|------------------|--------------------------|
|               | OR 95% CI   | \(P\) value   | OR 95% CI | \(P\) value |
| Neutropenia   | 6.20        | \(2.20 \times 10^{-06}\) | 10.38 | \(6.05 \times 10^{-08}\) |
| Agranulocytosis | 10.49      | \(1.83 \times 10^{-06}\) | 16.31 | \(1.39 \times 10^{-06}\) |

Estimates are given for risk allele G. Global ancestry: association results corrected for principal components (PC1–PC7). Global and local ancestry: association results corrected for principal components (PC1–PC7) and local ancestry estimates for ancestral populations AFR and EAS. Significant results are in bold. Freq frequency of risk allele G, OR odds ratio for risk allele G, 95% CI 95% confidence interval for odds ratio.

Fig. 2 Odds ratio estimates for varying ANC thresholds (500–1500). Odds ratios were determined in association analyses of European individuals corrected for global (principal components PC1–PC7) and additionally for local ancestry (dosage estimates for ancestral populations AFR and EAS). Regression lines were determined by weighted least squares regression where the reciprocals of the standard errors were used as weights.
LD decreases in the downstream direction of HLA-DQB1 overlapping HLA-DQA1 \((r^2 > 0.7)\) and the HLA-DRB1-DRB6 (pseudo)-DRBS region \((r^2 > 0.4)\) in the European population (Fig. 3), whereas in African and Asian populations, the regions including high to low LD variants are more sharply limited surrounding the region of interest (Supplementary Fig. S7). Which variant(s) in this region is/are causal and whether variations in this region contribute to the risk in specific populations only remain subject to further investigation.

There is evidence that besides global population patterns, local haplotypes might be a relevant feature in the HLA-DQB1 region. An excess of ancestry switches was reported
for several populations and European ancestry was shown
to diminish across the HLA-DQB1 region in a British study
on neutropenia. We determined the local haplotype struc-
ture in a sample without overlapping cases and also found
decreasing European contribution to the region surround-
ing the marker rs113332494. The analysis revealed a sub-
stantial increase of East Asian local haplotypes in risk allele
carriers, whereas in non-risk allele carriers European allele
dosage estimates dominated. This pattern was present in
cases and controls and thus did not provide additional
information for risk detection or prediction. Interestingly,
we also found significantly lower EAS and higher AFR
dosage estimates in cases compared to controls not carrying
the risk allele. Furthermore, including the local haplotype
information substantially strengthened the association with
both neutropenia (OR = 10.38, P = 6.05E−08) and agran-
ulocytosis (OR = 16.31, P = 1.39E−06). However, local
ancestry inferences should not be interpreted as conclusive
assignments of haplotypes to specific populations40, as the
resolution of the reference panel might not be enough to
resolve the observed pattern of genetic diversity. In addition,
itis not clear how the input sample or the choice of
markers influences local ancestry estimates and if the effect
is additive as was assumed in this study. Nevertheless, if
further evaluated and replicated, these results could con-
tribute to improve prediction of haematological side effects.

Individuals of African ancestry have a higher prevalence
of BEN14. This study included individuals with >80%
global European ancestry patterns. Thus, one might
speculate that higher local African ancestry in those with
low ANC might be related to global African ancestry
patterns, and specifically, a higher propensity for BEN. How-
however, local and global African ancestry do not cor-
relate in non-risk allele carriers (controls: Kendall’s tau-b
= 0.003, P = 0.89, neutropenia: Kendall’s tau-b = 0.024, P
= 0.74), suggesting that the local effect is not driven by
uncontrolled effects of global ancestry. Moreover, among
all non-risk allele carriers, we found only one individual,
who was actually a control, who was homozygous for the
C-allele at SNP rs2814778, which is actually thought to be
the causal genotype for BEN, and is a better predictor of
BEN than global ancestry15. Thus, our finding appears to
reflect local, not genome-wide, effects of African ancestry
beyond those indexed by rs113332494. The observation of
a local effect of African ancestry independent of
rs113332494 suggests that there are either additional
causal alleles within the region that are enriched in people
of African ancestry, and/or that rs113332494 is not causal
but is instead in LD with the causal variant.

To be of clinical use, a predictive pharmacogenetic test
should be sufficient to identify individuals for which the
risk of developing adverse effects is low enough to make
blood monitoring either unnecessary or reduce the fre-
quency by which it is performed. This corresponds to a
low risk of developing adverse effects in the group of non-
risk allele carriers. Verbelen et al. analysed the char-
acteristics of a useful predictive variant and showed that
the key factor is high sensitivity, especially in combination
with a low frequency of risk allele carriers. Sensitivity or
the true positive rate for a test denotes the proportion of
individuals developing a side effect who are positive for
the test, high sensitivity implying that the number of
people falsely classified as being not at risk is low. In the
context of high sensitivity, a low-risk allele frequency also
implies that the absolute numbers of people correctly
classified as being not at risk is larger, and thus blood
monitoring can be relaxed for a large number of individ-
uals. Assuming an acceptable risk of 0.13% (corre-
sponding to the risk of chlorpromazine which does not
require extensive blood monitoring in the UK) for agran-
ulocytosis in the group of non-risk allele carriers and an
agranulocytosis risk of 0.9%, sensitivity must be at least
85.7%11 in order to relax monitoring in those that are not
positive for the test (i.e. do not have the high-risk allele).
As in earlier studies, where sensitivity was found to be
21.5%21 and 5%25, we failed to reach the required threshold.
However, using local ancestry estimates for prediction,
the sensitivity of rs113332494 increased from 11.28 to
55.64% for neutropenia and from 16.67 to 53.70% for
agranulocytosis, thus coming closer to the goal of using
variants as a predictive tool to give patients the chance of
going an effective treatment as soon as possible and with
minimal disruption.

While our results show the importance of the marker
under study, they should be viewed in light of some
limitations. First of all, the sample size is arguably small,
but this is unavoidable given the low prevalence of agra-
ulocytosis. For this reason, the project allowed the
inclusion of patients of all ancestries. However, the anal-
ysis of the risk allele distribution showed that we do not
have enough samples to conduct an association analysis in
individuals of African, Asian or mixed ancestry and thus is
limited to individuals of European ancestry. Second, the
current study is limited to one single SNP. The variant
was not imputed with enough quality in other GWAs, but
has been replicated in independent samples and is thus of
special interest. Therefore, we conducted additional ana-
yses for refinement of findings, while genome-wide ana-
lyses are underway. Third, different groups recruited
patients under different conditions. While following
standardized guidelines, some patients were recruited in
hospitals during treatment, others were recruited from
registers where the available information on clinical and
demographic features was limited (see Supplementary
Information). For example, treatment compliance was
monitored and good compliance was defined as clinical
impression >80% of dosage and/or consistently docu-
mented plasma levels. However, plasma levels could not
be included in the statistical analysis, as this information was not available for all study subjects. Last, the neutropenia sample might also include individuals who would have developed agranulocytosis if clozapine had not been stopped, but this follows the monitoring guidelines in countries of contributing groups and is the situation encountered in clinical practice.

In summary, we were able to replicate previous findings for an association of rs113332494 (HLA-DQB1 6672G>C) with clozapine-induced neutropenia and agranulocytosis in individuals of European ancestry. We showed that local haplotype structure might be an important factor when considering regions containing an excess of ancestry switches, as is the case for the MHC region, and that these patterns might be relevant for determining risk. These findings could lead to novel insights into the pathogenesis of neutropenia and agranulocytosis. Using local ancestry estimates for prediction increased the sensitivity of rs113332494. Further studies are required to clarify the role of the marker rs113332494 and to analyse the influence of global and local ancestry patterns.

Acknowledgements
The authors thank all participants, clinicians and researchers for making this study possible. This CRESTAR project has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement #279227. M.P. wishes to thank the MRC Centre for Drug Safety Science (grant code MR/L007678/1) for support. A.F.P. acknowledges support by a “Springboard” award (ref SBF005 \1/083) from the Academy of Medical Sciences. J.T.R.W. and M.C.O. acknowledge support by a collaborative research grant from Takeda (Takeda played no part in the conception, design, implementation or interpretation of this study). Genome-wide data of CIAC used for identifying overlapping individuals were obtained from NIMH Repository & Genomics Resource, a centralized national biorepository for genetic studies of psychiatric disorders (www.nimhgenetics.org). The collection of data and biomaterials for this study was supported by funding provided by R01 MH080403 (PI: P.F.S., Edwin J.C.G. van den Oord) from the US National Institute of Mental Health via the American Recovery and Reinvestment Act of 2009. Dr. Denis Fourches and Dr. Alexander Tropsha acknowledge the support from National Science Foundation grant ABI 10-567. To Dr. Denis Fourches and Dr. Alexander Tropsha acknowledge the support from National Science Foundation grant ABI 10-567. Dr. Stanton Gerson (Department of Hematology, The University of Texas Galveston) and Dr. Nancy Berliner (Department of Severe Mental Illness, Mental Health Care Organization North-Holland North, Heerhugowaard, The Netherlands) acknowledge the support from National Science Foundation grant ABI 10-567. Dr. Stanton Gerson (Department of Hematology, The University of Texas Galveston) and Dr. Nancy Berliner (Department of Severe Mental Illness, Mental Health Care Organization North-Holland North, Heerhugowaard, The Netherlands) acknowledge the support from National Science Foundation grant ABI 10-567. Dr. Alexandre Robert acknowledges support by a collaborative research grant from Takeda (Takeda played no part in the conception, design, implementation or interpretation of this study).

Author details
1Department of Psychiatry, Psychotherapy and Psychosomatics, Martin-Luther-University Halle-Wittenberg, Halle, Germany. 2MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK. 3Department of Psychiatry, University of Liverpool, Liverpool, England. 4Department of Forensic Psychiatry, University of Eastern Finland, Niuvanniemi Hospital, Kuopio, Finland. 5Department of Clinical Neuroscience, Karolinska Institutet, and Center for Psychiatry Research, Stockholm City Council, Stockholm, Sweden. 6Mental Health Services Rivierduinen, Oegstgeest, The Netherlands. 7Department of Clinical Chemistry, St Jansdal Hospital, harderwijk, The Netherlands. 8Psychiatric Hospital GZ Centraal, Dependance Meerkanten, Ermelo, The Netherlands. 9Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland. 10Landsdómskóli University Hospital, Mental Health Services, Hringbraut, 101 Reykjavik, Iceland. 11CAMPBELL Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON MST 1R8, Canada. 12Department of Psychiatry, University of Toronto, Toronto, ON MST 1R8, Canada. 13Institute of Medical Science, University of Toronto, Toronto, ON M5S 1A8, Canada. 14Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm SE-17177, Sweden. 15Department of Genetics, University of North Carolina, Chapel Hill, NC 27599-7264, USA. 16Department of Psychiatry, University of North Carolina, Chapel Hill, NC 27599-7160, USA. 17Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany. 18King’s College London, London, UK. 19deCODE Genetics/Amgen, Reykjavik, Iceland. 20Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London, UK. 21Eli Lilly and Company, Bracknell, UK

Author contributions
D.R., M.C.O. and I.G. were responsible for conception and design of the study. B.K. performed the statistical analyses and drafting of the first manuscript version. All authors contributed to the collection and interpretation of data and writing of the manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL.

Conflict of interest
D.A.C. is a full-time employee and stockholder of Eli Lilly and Company.

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

Supplementary information
The online version contains supplementary material available at https://doi.org/10.1038/s41398-021-01322-w.

Received: 14 July 2020 Revised: 21 February 2021 Accepted: 11 March 2021

Published online: 12 April 2021

References
1. Kane, J., Hongfeldt, G., Singer, J. & Meltzer, H. Clozapine for the treatment-resistant schizophrenic: A double-blind comparison with chlorpromazine. Arch. Gen. Psychiatry 45, 789–796 (1988).
2. Meltzer, H. Y. Treatment-resistant schizophrenia—the role of clozapine. Curr. Med. Res. Opin. 14, 1–20 (1997).
3. Säkkinen, H., McCarty, N., Goldschlager, R. & Kisely, S. Clozapine v. first- and second-generation antipsychotics in treatment-refractory schizophrenia: systematic review and meta-analysis. Br. J. Psychiatry 209, 385–392 (2016).
4. Fogley, C., Taylor, D., Dickens, G. & Picchioni, M. A systematic review of the evidence of clozapine’s anti-aggressive effects. Int. J. Neuropsychopharmacol. 15, 1351–1371 (2012).
5. Meltzer, H. Y. Clozapine: balancing safety with superior antipsychotic efficacy. Clin. Schizophr. Relat. Psychoses 6, 134–144 (2012).
6. Kim, Y.-J., No, S.-H. & Lee, S.-Y. Successful early clozapine trial in the treatment of first-episode schizophrenia: a case report. Clin. Psychopharmacol. Neurosci. 11, 168–169 (2013).
7. Idanpaan-Heikkilä, J., Alavä, E., Okninuora, M. & Palva, I. Letter: clozapine and agranulocytosis. Lancet 2, 611 (1975).
8. Naber, D., Haasen, C. & Perro, C. Clozapine: the first atypical antipsychotic. In Atypical Antipsychotics (eds Ellenbroek, B. A. & Cools, A. R.) 145–162 (Birkhäuser Basel, 2000).
9. Myles, N. et al. Meta-analysis examining the epidemiology of clozapine-associated neutropenia. Acta Psychiatr. Scand. 138, 101–109 (2018).
Athanasiou, M. C. et al. Candidate gene analysis identifies HLA-DQB1 and HLA-B alleles as risk factors for clozapine-induced agranulocytosis in the USA experience. *Drug Saf* 7, 139–158 (1997).

11. Gillman, K. Paradoxical pattern of haematological risk with clozapine. *Br. J. Psychiatry* 177, 88 (2000).

12. Yunis, J. J., Lieberman, J. & Yunis, E. J. Major histocompatibility complex associations with clozapine-induced agranulocytosis. The USA experience. *Drug Saf* 7(Suppl 1), 7–9 (1992).

13. Marchesi, C., Paini, M., Tamborini, S., Ampollini, P. & Maggini, C. Recurrence of clozapine-induced agranulocytosis. *J. Clin. Psychopharmacol.* 25, 276–277 (2005).

14. Haddy, T. B., Rana, S. R. & Castro, O. Benign ethnic neutropenia: what is a normal absolute neutrophil count? *J. Lab. Clin. Med.* 133, 15–22 (1999).

15. Legge, S. E. et al. A genome-wide association study in individuals of African ancestry reveals the importance of the Duffy-null genotype in the assessment of clozapine-related neutropenia. *Mol. Psychiatry* 24, 328–337 (2019).

16. Wei, C.-Y., Lee, M.-T. & Chen, Y.-T. Pharmacogenomics of adverse drug reactions: implementing personalized medicine. *Hum. Mol. Genet.* 21, R58–R65 (2012).

17. Pouget, J. G., Shams, T. A., Tiwari, A. K. & Müller, D. J. Pharmacogenomics and outcome with antipsychotic drugs. *Dialogues Clin. Neurosci.* 16, 555–566 (2014).

18. Lieberman, J. A. et al. HLA-B38, DRA, DQw3 and clozapine-induced agranulocytosis in Jewish patients with schizophrenia. *Arch. Gen. Psychiatry* 47, 945–948 (1990).

19. Yunis, J. J. et al. HLA associations in clozapine-induced agranulocytosis. *Blood* 86, 1177–1183 (1995).

20. Detting, M., Schaub, R. T., Mueller-Oerlinghausen, B., Roots, I. & Cascoconti, I. Further evidence of human leukocyte antigen-encoded susceptibility to clozapine-induced agranulocytosis independent of ancestry. *Pharmacogenet. Genom.* 11, 135–141 (2001).

21. Athanasiou, M. C. et al. Candidate gene analysis identifies a polymorphism in HLA-DQA1 associated with clozapine-induced agranulocytosis. *J. Clin. Psychiatry* 72, 458–463 (2011).

22. Goldstein, J. I. et al. Clozapine-induced agranulocytosis is associated with rare HLA-DQA1 and HLA-B alleles. *Nat. Commun.* 5, 4757 (2014).

23. Detting, M., Cascoconti, I., Oppen-Rhein, C. & Schaub, R. Clozapine-induced agranulocytosis in schizophrenic Caucasians: confirming clues for associations with human leukocyte class I and II antigens. *Pharmacogenomics* J. 7, 325–332 (2007).

24. van der Weide, K. et al. Genetic risk factors for clozapine-induced neutropenia and agranulocytosis in a Dutch psychiatric population. *Pharmacogenomics J.* 17, 471–478 (2017).

25. 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).

26. Shiner, D., Adeyemo, A., Ramos, E., Chen, G. & Rotimi, C. N. Mapping of disease-associated variants in admixed populations. *Genome Biol.* 12, 223 (2011).

27. Lam, T. H., Shen, M., Chia, J.-M., Chan, S. H. & Ren, E. C. Population-specific recombination sites within the human MHC region. *Heredity* 111, 131–138 (2013).

28. Ordoñez, G. et al. Genomewide admixture study in Mexican Mestizos with multiple sclerosis. *Clin. Neurol. Neurosurg.* 130, 55–60 (2015).

29. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G. R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* 44, 955–959 (2012).

30. Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3* 1, 457–470 (2011).

31. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* 5, e1000529 (2009).

32. Zhou, H., Alexander, D. & Lange, K. A quasi-Newton acceleration for high-dimensional optimization algorithms. *Stat. Comput.* 21, 261–273 (2011).

33. Chang, C. C. et al. Second-generation PLINK rising to the challenge of larger and richer datasets. *GigaScience* 4 https://doi.org/10.1186/s13742-015-0047-8 (2015).

34. Purcell, S. & Chang, C. PLoS 1.9. www.cog-genomics.org/plink/1.9/.

35. Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38, 904–909 (2006).

36. Guan, Y. Detecting structure of haplotypes and local ancestry. *Genetics* 196, 625–642 (2014).

37. Gauderman, W. J. & Morrison, J. QUANTO 1.1: a Computer Program for Power and Sample Size Calculations for Genetic-epidemiology Studies. http://hydra.usc.edu/gxe (2006).

38. Arnold, M., Raffler, J., Pfeuffer, A., Suhe, K. & Kastenmüller, G. SNIPA: an interactive, genetic variant-centered annotation browser. *Bioinformatics* 31, 1334–1336 (2015).

39. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* 45, 580–585 (2013).

40. Lawson, D. J., van Dorp, L. & Falush, D. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nat. Commun.* 9, 3258 (2018).

41. VerbeLEN, M., Coller, D. A., Cohen, D., MacCabe, J. H. & Lewis, C. M. Establishing the characteristics of an effective pharmacogenetic test for clozapine-induced agranulocytosis. *Pharmacogenomics J.* 15, 461–466 (2015).