Genome-wide association study of nevirapine hypersensitivity in a sub-Saharan African HIV-infected population

Daniel F. Carr1*,†, Stephane Bourgeois2†, Mas Chaponda1,3, Louise Y. Takeshita4, Andrew P. Morris1,5, Elena M. Cornejo Castro1, Ana Alfirevic1, Andrew R. Jones4, Daniel J. Rigden4, Sam Haldenby6, Saye Khoo1, David G. Lalloo7, Robert S. Heyderman3,8, Collet Dandara9, Elizabeth Kampira9, Joep J. van Oosterhout3,10, Francis Ssali11, Paula Munderi12, Giuseppe Novelli13, Paola Borgiani13, Matthew R. Nelson14, Arthur Holden15, Panos Deloukas2,16,17 and Munir Pirmohamed1

1Department of Molecular and Clinical Pharmacology, University of Liverpool, Liverpool, UK; 2William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK; 3Malawi-Liverpool-Wellcome Trust Clinical Research Programme, College of Medicine, University of Malawi, Malawi; 4Institute of Integrative Biology, University of Liverpool, Liverpool, UK; 5Department of Biostatistics, University of Liverpool, Liverpool, UK; 6Centre for Genomic Research, University of Liverpool, Liverpool, UK; 7Liverpool School of Tropical Medicine, Liverpool, UK; 8Division of Infection and Immunity, University College London, London, UK; 9Division of Human Genetics, University of Cape Town, Cape Town, South Africa; 10Dignitas International, Zomba, Malawi; 11Joint Clinical Research Centre, Headquarters, Kampala, Uganda; 12UVRI/MRC Uganda Research Unit on AIDS, Entebbe, Uganda; 13Department of Biomedicine and Prevention, University of Rome ‘Tor Vergata’, Rome, Italy; 14GlaxoSmithKline, Research Triangle Park, NC, USA; 15SAEC Consortium, Ltd, Chicago, IL, USA; 16Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK; 17Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, 21589, Saudi Arabia

*Corresponding author. Tel: +44-151-795-5392; E-mail: d.carr@liv.ac.uk
†Authors contributed equally.

Received 15 August 2016; returned 10 October 2016; revised 13 October 2016; accepted 20 November 2016

Background: The antiretroviral nevirapine is associated with hypersensitivity reactions in 6%–10% of patients, including hepatotoxicity, maculopapular exanthema, Stevens–Johnson syndrome (SJS) and toxic epidermal necrolysis (TEN).

Objectives: To undertake a genome-wide association study (GWAS) to identify genetic predisposing factors for the different clinical phenotypes associated with nevirapine hypersensitivity.

Methods: A GWAS was undertaken in a discovery cohort of 151 nevirapine-hypersensitive and 182 tolerant, HIV-infected Malawian adults. Replication of signals was determined in a cohort of 116 cases and 68 controls obtained from Malawi, Uganda and Mozambique. Interaction with ERAP genes was determined in patients positive for HLA-C*04:01. In silico docking studies were also performed for HLA-C*04:01.

Results: Fifteen SNPs demonstrated nominal significance (P < 1 × 10−3) with one or more of the hypersensitivity phenotypes. The most promising signal was seen in SJS/TEN, where rs5010528 (HLA-C locus) approached genome-wide significance (P < 8.5 × 10−8) and was below HLA-wide significance (P < 2.5 × 10−6) in the meta-analysis of discovery and replication cohorts [OR 4.84 (95% CI 2.71–8.61)]. rs5010528 is a strong proxy for HLA-C*04:01 carriage: in silico docking showed that two residues (35 and 123) in the B pocket were the most likely nevirapine interactors. There was no interaction between HLA-C*04:01 and ERAP1, but there is a potential protective effect with ERAP2 (P = 0.019, OR 0.43 (95% CI 0.21–0.87)).

Conclusions: HLA-C*04:01 predisposes to nevirapine-induced SJS/TEN in sub-Saharan Africans, but not to other hypersensitivity phenotypes. This is likely to be mediated via binding to the B pocket of the HLA-C peptide. Whether this risk is modulated by ERAP2 variants requires further study.
Introduction

Nevirapine, an NNRTI used for HIV infection is effective as part of combination antiretroviral therapy, but causes hypersensitivity in 6%–10% of patients. This can manifest in various ways, ranging from nevirapine-induced rash (NIR) (i.e. a maculopapular exanthema without any systemic manifestations), hypersensitivity syndrome (HSS) to severe blistering skin reactions such as Stevens–Johnson syndrome (SJS) and toxic epidermal necrolysis (TEN). Extra-cutaneous involvement typically manifests as hepatotoxicity.

Identification of the genetic risk factors for nevirapine hypersensitivity has focused on candidate gene approaches. Nevirapine is primarily metabolized by the hepatic cytochrome P450s 2B6 (CYP2B6) and 3A4 (CYP3A4). The exon 4 variant in CYP2B6 (c.516G>T), which encodes a non-synonymous amino acid substitution (Gln172His) (rs3745274), leads to loss of function, with the variant T allele resulting in higher nevirapine plasma concentrations in both Caucasian and sub-Saharan adult patients. The associations with CYP2B6 polymorphisms are rather confusing with the CYP2B6 c.516G>T SNP associated with nevirapine-induced cutaneous adverse events in black and white populations but not with nevirapine-induced hepatotoxicity.

The association with HLA alleles is even more complex, with HLA-DRB1*01:01 (Caucasian, Chinese and Black), HLA-C*04 (Thai, Chinese and Black), HLA-C*08 (Japanese) and HLA-B*35:05 (Thai) acting as predisposing alleles for nevirapine hypersensitivity. Our own previous study within a subset of patients from the Malawian HIV population described in this paper identified an association between HLA-C*04:01 and nevirapine-induced SJS.

In this study, in order to overcome some of the issues associated with candidate gene analysis, we have undertaken a genome-wide association study (GWAS) in a Malawian HIV cohort of nevirapine-exposed patients in order to identify genetic biomarkers of nevirapine hypersensitivity in an unbiased manner. We have also investigated whether there is any interaction between HLA-C*04:01 in SJS/TEN patients and the endoplasmic reticulum aminopeptidase genes (ERAP1 and ERAP2), which have been shown to modulate the risk of various immune diseases, in particular ankylosing spondylitis.

Methods

Patients

Discovery cohort

Antiretroviral-naive patients (n = 1117) were prospectively recruited as previously described (Figure 1) from the Queen Elizabeth Central Hospital (QECH), Blantyre, Malawi, between March 2007 and December 2008. All were self-reported black African, over the age of 16, and had no baseline jaundice. CD4+ counts and liver function tests were monitored at 0, 2, 6, 10, 14, 18 and 22 weeks. Fifty-seven patients from this prospective cohort had nevirapine-induced hypersensitivity fulfilling the criteria of one or more of the following phenotypes:

- NIR: widespread maculopapular exanthema with no systemic manifestations but which worsened on treatment continuation.
- HSS: widespread rash with systemic manifestations (i.e. fever, cough or abnormal liver function tests). This is also known as DRESS (drug reaction with eosinophilia and systemic symptoms).

We recruited a number of patients with nevirapine hypersensitivity, with different phenotypes, from a number of centres (Table 1) to replicate our findings:

- Thirty nevirapine-hypersensitive patients and matched (age and gender) HIV-positive nevirapine-treated controls from Malawi. All controls and eight of the cases were from the original study but not included in the initial GWAS due to DNA quantity restraints. The other 22 cases presenting with the hypersensitivity phenotype according to the above criteria were identified from the QECH after the conclusion of the initial recruitment phase (December 2008).
- Thirty-two nevirapine-hypersensitive cases and age- and gender-matched controls identified in Uganda from the DART study cohort.
- Twenty-seven pregnant female patients with nevirapine-induced hepatitis and 10 nevirapine-tolerant pregnant controls from Mozambique. Cases were defined as previously stated, and included patients who discontinued nevirapine due to increased liver enzymes (grade 3/4). Controls were excluded if ALT/AST levels exceeded the median value observed in the case cohort.
- Twenty-seven female patients with nevirapine-induced SJS/TEN from Mozambique. In this instance SJS/TEN was defined as development of exanthema and blistering starting mainly on the trunk, involving >10% of the body surface with mucosal involvement.

DNA extraction

Genomic DNA was extracted from whole blood for the discovery cohort and replication cohorts as previously described.

Ethics

Full ethics approval for the study was received from the Liverpool School of Tropical Medicine Research Ethics Committee (Liverpool, UK), the College of Medicine Research and Ethics Committee, University of Malawi (Blantyre, Malawi) and the Uganda National Council for Science and Technology. All patients gave their written informed consent and those who met the criteria for a case had nevirapine withdrawn in accordance with Malawian National Treatment Guidelines. Local ethics approval was obtained for the DART study as previously described with subsequent ethics approval for a pharmacogenetic sub-study also obtained.

Genomic DNA was extracted from whole blood for the discovery cohort and replication cohorts as previously described.
Discovery cohort genotyping and sample quality control (QC)
A total of 352 samples were genotyped for 1,048,713 variants using the HumanOmni1-Quad_v1 chip (Illumina). Variants were excluded from analysis if their minor allele frequency (MAF) was <1%, the call rate was <99% for an MAF between 1% and 3%, the call rate was <98% for an MAF of >3%, or if Hardy–Weinberg expectations were not satisfied ($P < 10^{-4}$).

Individuals were excluded if the sample call rate was <95%, the assigned gender contradicted genetic information from the X chromosome heterozygosity, or if they appeared to be duplicates, or related to other individuals in the study (as measured by identity by state using PLINK). Multidimensional scaling analysis of genotype data was undertaken by merging the data with HapMap3 cohort data and using the mds function in PLINK in order to determine population stratification (Figure S1, available as Supplementary data at JAC Online).

Discovery cohort imputation
Imputation of genotypes, after phasing of each chromosome using ShapeIt, was carried out using IMPUTE V2.3.1, 1000G phase 1 integrated v3 macGT1 reference panel haplotypes (March 2012). After imputation, SNPs with an information measure (info score) <0.8 were discarded, and a threshold of 0.5 was applied on genotype uncertainty. Imputed variants with an MAF <1% were then excluded.

Discovery cohort association analysis
Univariate logistic regression analysis of non-genetic covariates (age, gender, BMI, CD4+ cell count) was undertaken for each hypersensitivity phenotype. Statistically significant variables ($P < 0.05$) were included in the subsequent logistic regressions to test for the association of each hypersensitivity phenotype with each SNP passing QC. All statistical analyses were undertaken using PLINK and R. Given prior associations between nevirapine hypersensitivity and HLA allele associations, it was felt reasonable to specify a Bonferroni-corrected HLA-wide significance threshold of $P < 2.5 \times 10^{-5}$, based on the presumption that there are usually <200 effective HLA allele tests.

Replication cohort genotyping, QC and association analysis
SNPs determined to have a nominally significant association with a nevirapine hypersensitivity phenotype ($P < 1 \times 10^{-5}$) in the discovery cohort were subsequently typed in the replication cohort using either the Sequenom MassArray iPLEX platform (Sequenom Inc., San Diego, CA, USA) or custom TaqMan real-time PCR SNP genotyping assays (Life Technologies, Paisley, UK) according to the manufacturer’s protocols. SNPs were excluded if they failed to meet the genotype QC thresholds as outlined for the discovery cohort or if assay design software parameters prohibited their inclusion.

Logistic regression analysis of the replication cohort, including and excluding CD4+ count as a covariate, where appropriate (as determined in the discovery cohort), was carried out. Meta-analysis of combined discovery and replication cohorts was undertaken using a fixed-effects model with inverse-variant effect size weighting in GWAMA.

Imputation of HLA allele type and MHC locus
Imputation of HLA-C allele type from the discovery cohort SNP array data was undertaken using HLA*IMP:02.

HLA-C and ERAP gene–gene interactions
In cases and tolerant controls positive for carriage of the rs5010528 G allele, which was used as a proxy for HLA-C*04:01, we investigated both ERAP1 (rs10050860 and rs30187) and ERAP2 (rs2248374, rs2549782) SNPs
which have previously been shown to interact with HLA-mediated immune diseases. Association of ERAP1 and ERAP2 SNPs with SJS/TEN risk was determined in the HLA-C*04:01-positive cohort (cases and controls) by logistic regression with CD4+ cell count as a covariate using PLINK. A Bonferroni adjustment for multiple testing was applied with a significance threshold of $P = 0.125$.

**Targeted sequencing of MHC region**

Sixteen genomic DNA samples from nevirapine-induced SJS/TEN and 16 age- and gender-matched tolerant controls were carried forward for MHC-targeted sequencing. The methodology is detailed in the Supplementary data.

**Allelotyping**

HLA allelotyping was performed from raw FASTQ data files using Omixon Target v1.81 HLA Typing software and utilized the HLA database version 3.15.0 (Omixon Ltd, Budapest, Hungary).

**In silico docking**

In order to predict possible modes of interaction between nevirapine and HLA-C*04:01, in silico docking was undertaken. The methodology is detailed in the Supplementary data.

**Results**

**Discovery cohort**

A total of 333 samples (151 cases and 182 controls) out of 352 passed QC. Of the 19 excluded samples, 9 failed heterozygosity checks (outliers by $> 3$ SD), 8 failed identity checks and 2 failed the call rate threshold. Multi-dimensional analysis for population stratification (Figure S1) demonstrated no population outliers. In total, 817 728 SNPs passed QC and were carried over for imputation with the 1000 genomes panel. Imputation produced a dataset of 1 421 8511 variants. Cohort characteristics are shown in Table 1. We considered five nevirapine-induced hypersensitivity phenotypes for analysis—NIR, HSS, SJS/TEN, DILI (Table 1)—and also combined these different phenotypes into an overall hypersensitivity group.

Univariate logistic regression analysis showed CD4+ cell count to be a statistically significant variable for NIR ($P = 0.016$), SJS ($P = 0.003$) and all hypersensitivity cases ($P = 0.002$). Therefore, we included CD4+ cell count as a covariate in the SNP logistic regression model for these three phenotypes. Multidimensional scaling (MDS) variables were not included as covariates in the logistic regression since the population stratification analysis suggested that the cohort was homogeneous and genomic control was unnecessary (Figure S1). From the genome-wide logistic regression analyses, we identified 15 SNPs with $P < 5\times 10^{-8}$, with at least one of the five different phenotypes analysed (Figure 2; summarized in Table 2). No variant reached genome-wide significance.

**Replication cohort**

Of the 15 SNPs considered for replication, one (rs150223496) could not be typed due to proximal sequence constraints of the Sequenom assay design process, and QC failure for TaqMan genotyping Hardy-Weinberg equilibrium (HWE), (using data from the Illumina array), which have previously been shown to interact with HLA-mediated immune diseases. Association of ERAP1 and ERAP2 SNPs with SJS/TEN risk was determined in the HLA-C*04:01-positive cohort (cases and controls) by logistic regression with CD4+ cell count as a covariate using PLINK. A Bonferroni adjustment for multiple testing was applied with a significance threshold of $P = 0.125$.

**Targeted sequencing of MHC region**

Sixteen genomic DNA samples from nevirapine-induced SJS/TEN and 16 age- and gender-matched tolerant controls were carried forward for MHC-targeted sequencing. The methodology is detailed in the Supplementary data.

**Allelotyping**

HLA allelotyping was performed from raw FASTQ data files using Omixon Target v1.81 HLA Typing software and utilized the HLA database version 3.15.0 (Omixon Ltd, Budapest, Hungary).

**In silico docking**

In order to predict possible modes of interaction between nevirapine and HLA-C*04:01, in silico docking was undertaken. The methodology is detailed in the Supplementary data.

**Results**

**Discovery cohort**

A total of 333 samples (151 cases and 182 controls) out of 352 passed QC. Of the 19 excluded samples, 9 failed heterozygosity checks (outliers by $> 3$ SD), 8 failed identity checks and 2 failed the call rate threshold. Multi-dimensional analysis for population stratification (Figure S1) demonstrated no population outliers. In total, 817 728 SNPs passed QC and were carried over for imputation with the 1000 genomes panel. Imputation produced a dataset of 1 421 8511 variants. Cohort characteristics are shown in Table 1. We considered five nevirapine-induced hypersensitivity phenotypes for analysis—NIR, HSS, SJS/TEN, DILI (Table 1)—and also combined these different phenotypes into an overall hypersensitivity group.

Univariate logistic regression analysis showed CD4+ cell count to be a statistically significant variable for NIR ($P = 0.016$), SJS ($P = 0.003$) and all hypersensitivity cases ($P = 0.002$). Therefore, we included CD4+ cell count as a covariate in the SNP logistic regression model for these three phenotypes. Multidimensional scaling (MDS) variables were not included as covariates in the logistic regression since the population stratification analysis suggested that the cohort was homogeneous and genomic control was unnecessary (Figure S1). From the genome-wide logistic regression analyses, we identified 15 SNPs with $P < 5\times 10^{-8}$, with at least one of the five different phenotypes analysed (Figure 2; summarized in Table 2). No variant reached genome-wide significance.

**Replication cohort**

Of the 15 SNPs considered for replication, one (rs150223496) could not be typed due to proximal sequence constraints of the Sequenom assay design process, and QC failure for TaqMan genotyping Hardy-Weinberg equilibrium (HWE)
Figure 2. Manhattan plots of association for logistic regression SNP analysis for the five defined nevirapine hypersensitivity phenotypes. The phenotypes, all hypersensitivity, NIR and SJS/TEN show P values from logistic regression incorporating CD4+ cell count as a covariate. For HSS and DILI no covariates were incorporated. The broken line indicates genome-wide significance (P = 5 × 10^{-8}).

Table 2. Top SNPs identified in association with a nevirapine hypersensitivity phenotype within the main cohort analysis

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>SNP</th>
<th>Chr</th>
<th>Position (GRCh37.p13)</th>
<th>Reference allele</th>
<th>Associated allele</th>
<th>Gene</th>
<th>Typed/imputed</th>
<th>Logistic regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P value</td>
</tr>
<tr>
<td>All hypersensitivity (n = 151)</td>
<td>rs34213790</td>
<td>6</td>
<td>31318579</td>
<td>G</td>
<td>A</td>
<td>3’ of HLA-B</td>
<td>imputed</td>
<td>1.15 x 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>rs11764223</td>
<td>7</td>
<td>137742056</td>
<td>T</td>
<td>G</td>
<td>AKR1D1</td>
<td>imputed</td>
<td>2.99 x 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>rs11988543</td>
<td>8</td>
<td>68596362</td>
<td>T</td>
<td>A</td>
<td>CPA6</td>
<td>imputed</td>
<td>2.69 x 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>rs9527426</td>
<td>13</td>
<td>56208216</td>
<td>T</td>
<td>C</td>
<td>MIR5007</td>
<td>typed</td>
<td>9.86 x 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>rs35990155</td>
<td>23</td>
<td>11827017</td>
<td>G</td>
<td>T</td>
<td>KIAA1210</td>
<td>imputed</td>
<td>6.33 x 10^{-7}</td>
</tr>
<tr>
<td>Rash (n = 56)</td>
<td>rs10815440</td>
<td>9</td>
<td>661057</td>
<td>G</td>
<td>A</td>
<td>KANK1</td>
<td>typed</td>
<td>8.86 x 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>rs115848367</td>
<td>10</td>
<td>128045499</td>
<td>A</td>
<td>C</td>
<td>ADAM12</td>
<td>imputed</td>
<td>8.83 x 10^{-7}</td>
</tr>
<tr>
<td></td>
<td>rs74373347</td>
<td>12</td>
<td>68919428</td>
<td>A</td>
<td>G</td>
<td>OR8S1</td>
<td>imputed</td>
<td>1.08 x 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>rs6511720</td>
<td>19</td>
<td>11202306</td>
<td>G</td>
<td>T</td>
<td>LDLR</td>
<td>typed</td>
<td>5.49 x 10^{-6}</td>
</tr>
<tr>
<td>SJS/TEN (n = 51)</td>
<td>rs5010528</td>
<td>6</td>
<td>31241032</td>
<td>A</td>
<td>G</td>
<td>HLA-C</td>
<td>typed</td>
<td>4.13 x 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>rs150223496</td>
<td>15</td>
<td>76421885</td>
<td>G</td>
<td>A</td>
<td>C15orf27</td>
<td>imputed</td>
<td>8.20 x 10^{-6}</td>
</tr>
<tr>
<td>DILI (n = 21)</td>
<td>rs147773805</td>
<td>1</td>
<td>107741393</td>
<td>A</td>
<td>T</td>
<td>NTNG1</td>
<td>imputed</td>
<td>4.47 x 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>rs114693001</td>
<td>9</td>
<td>76074363</td>
<td>A</td>
<td>T</td>
<td>intergenic</td>
<td>imputed</td>
<td>8.82 x 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>rs142213069</td>
<td>13</td>
<td>77962182</td>
<td>G</td>
<td>C</td>
<td>S’ of MYCBP2</td>
<td>imputed</td>
<td>6.64 x 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>rs6139258</td>
<td>20</td>
<td>3958616</td>
<td>T</td>
<td>C</td>
<td>RNF24</td>
<td>typed</td>
<td>6.64 x 10^{-6}</td>
</tr>
</tbody>
</table>

P values and ORs were determined by logistic regression with CD4+ cell count as a covariate (except for HSS and DILI).
Thus, 14 SNPs were carried forward for analysis. A total of 59 Malawian samples (30 cases and 29 controls) and 63 Ugandan samples (32 cases and 31 controls) passed QC (call rate >90%) as described in Table 1. Due to sample constraints, SNP signals identified in the discovery cohort for the ‘all hypersensitivity’ phenotype were not typed in the samples from Mozambique in the replication cohort.

For nevirapine-induced DILI (Table 3), combining the discovery and replication cohorts for the SNP rs6139258 in the RNF24 locus strengthened the association \( P = 5.7 \times 10^{-7}, OR 13.62 \) (95% CI 4.90–37.84) (Table 4). A single SNP (rs5010528 in the HLA-C locus) showed an association with SJS/TEN in the replication cohort (38 cases, 59 controls) \( P = 0.006; OR 5.12 \) (95% CI 1.60–16.42). Combining the discovery and replication cohorts strengthened the
GWAS of nevirapine hypersensitivity

**Figure 3.** *In silico* docking of HLA-C*04:01 and nevirapine. (a) Specific peptide residues present in the different HLA-C alleles previously identified in the Malawian cohort (n = 116). *G* shows continuity with the reference peptide (C*04:01) and *~* identifies peptides not sequenced in an allelotype. *~* denotes putative nevirapine interaction and # denotes residue substituted by SNP (rs1050409) identified in GWAS. Residue reference numbers are as defined by the *in silico* model. (b) HLA-C*04:01/nevirapine docking mode of conformation with the highest predicted affinity (lowest score) as produced using the PyMOL software. Key interacting residues are highlighted. (c) A LIGPLOT + 2D schematic representation of the interaction of nevirapine with HLA-C*04:01 PBD residues in the highest affinity predicted docking conformation mode. The broken line indicates a hydrogen bond and radiating lines indicate hydrophobic interactions. (d) HLA-C*04:01/nevirapine (orange) docking mode of conformation for the 20 highest predicted affinities (lowest scores) as produced using the PyMOL software.

overall association, which approached genome-wide significance [P = 8.5 × 10^{-8}, OR 4.84 (95% CI 2.71–8.61)]. No positive signals were identified for the ‘all hypersensitivity’ phenotype (62 cases, 59 controls).

**HLA-C allelotype imputation**

Overall allelotype imputation from the SNP array data using HLA*IMP demonstrated 71.5% concordance with the HLA typing obtained for 116 of our patients using the sequence-based methodology. However, the ability of the imputation to correctly call HLA-C*04:01 alleles was 90%.

Within the 116 patients for which HLA allele typing and SNP array genotype data were available, HLA-C*04:01 allele carriage co-occurred with the rs5010528 G allele in 112/116 cases (96.5%). For the imputed HLA allelotype data, C*04:01 co-occurred with rs5010528 G in 303/333 cases (91%).

The initial discovery logistic regression analysis demonstrated two non-synonymous SNPs in the HLA-C locus associated with SJS/TEN that were in absolute linkage disequilibrium (LD) with rs5010528 (Table 5). The first SNP (rs146911342) encodes a valine-to-methionine amino acid substitution at residue 327 and the second (rs1050409) encodes an alanine to glutamic acid at residue 73 (close to the peptide binding domain), which was also associated with SJS/TEN [P = 4.1 × 10^{-5}, OR 4.75 (95% CI 2.45–9.23)]. Both are key defining residues of the HLA-C*04 allelotype as defined in the HLA-IMGT database and within the Malawian cohort (Figure 3). Verification of either SNP in our discovery or replication cohorts via other genotyping methodologies was not possible due to sequence constraints in assay design (the proximal nucleotide sequence for primer design was not sufficiently unique or contained a restrictive number of other genetic variants).

However, targeted sequencing of the HLA locus in 16 SJS/TEN and 16 tolerant controls (Table 5) suggested that both the
Table 5. Association of nevirapine-induced SJS/TEN and imputed SNPs of the HLA-C locus from the main cohort and the targeted sequencing cohort

<table>
<thead>
<tr>
<th>SNP</th>
<th>bp (GRCh37.p13)</th>
<th>A1/A2</th>
<th>Amino acid substitution</th>
<th>typed/imputed</th>
<th>SJS/TEN MAF (n=51)</th>
<th>control MAF (n=182)</th>
<th>P</th>
<th>OR (95% CI)</th>
<th>SJS/TEN MAF (n=16)</th>
<th>control MAF (n=16)</th>
<th>P</th>
<th>OR (95% CI)</th>
<th>LD with rs5010528 (Df)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs146911342</td>
<td>31:237:779</td>
<td>C/T</td>
<td>p.V327M</td>
<td>imputed</td>
<td>0.36</td>
<td>0.15</td>
<td>4.13 × 10⁻⁶</td>
<td>4.75 (2.45–9.23)</td>
<td>0.34</td>
<td>0.13</td>
<td>0.013</td>
<td>34.14 (2.07–561.6)</td>
<td>1</td>
</tr>
<tr>
<td>rs41562714</td>
<td>31:238:538</td>
<td>G/C</td>
<td></td>
<td>imputed</td>
<td>0.36</td>
<td>0.15</td>
<td>4.13 × 10⁻⁶</td>
<td>4.75 (2.45–9.23)</td>
<td>0.34</td>
<td>0.13</td>
<td>0.013</td>
<td>34.14 (2.07–561.6)</td>
<td>1</td>
</tr>
<tr>
<td>rs1050409</td>
<td>31:239:501</td>
<td>G/T</td>
<td>p.A73E</td>
<td>imputed</td>
<td>0.36</td>
<td>0.15</td>
<td>4.13 × 10⁻⁶</td>
<td>4.75 (2.45–9.23)</td>
<td>0.34</td>
<td>0.13</td>
<td>0.013</td>
<td>34.14 (2.07–561.6)</td>
<td>1</td>
</tr>
<tr>
<td>rs41553018</td>
<td>31:239:742</td>
<td>G/C</td>
<td></td>
<td>imputed</td>
<td>0.36</td>
<td>0.15</td>
<td>3.71 × 10⁻⁶</td>
<td>4.75 (2.45–9.23)</td>
<td>0.34</td>
<td>0.13</td>
<td>0.013</td>
<td>34.14 (2.07–561.6)</td>
<td>1</td>
</tr>
<tr>
<td>rs4361609</td>
<td>31:240:635</td>
<td>G/C</td>
<td></td>
<td>imputed</td>
<td>0.36</td>
<td>0.15</td>
<td>4.13 × 10⁻⁶</td>
<td>4.75 (2.45–9.23)</td>
<td>0.34</td>
<td>0.13</td>
<td>0.013</td>
<td>34.14 (2.07–561.6)</td>
<td>1</td>
</tr>
<tr>
<td>rs5010528</td>
<td>31:241:032</td>
<td>A/G</td>
<td></td>
<td>typed</td>
<td>0.36</td>
<td>0.15</td>
<td>4.13 × 10⁻⁶</td>
<td>4.75 (2.45–9.23)</td>
<td>0.34</td>
<td>0.13</td>
<td>0.013</td>
<td>34.14 (2.07–561.6)</td>
<td>–</td>
</tr>
<tr>
<td>rs58019823</td>
<td>31:241:215</td>
<td>A/G</td>
<td></td>
<td>imputed</td>
<td>0.36</td>
<td>0.15</td>
<td>4.13 × 10⁻⁶</td>
<td>4.75 (2.45–9.23)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>NA</td>
</tr>
<tr>
<td>rs2524087</td>
<td>31:241:294</td>
<td>C/G</td>
<td></td>
<td>imputed</td>
<td>0.37</td>
<td>0.15</td>
<td>3.71 × 10⁻⁶</td>
<td>4.73 (2.45–9.14)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>NA</td>
</tr>
<tr>
<td>rs59103503</td>
<td>31:287:944</td>
<td>G/T</td>
<td></td>
<td>imputed</td>
<td>0.37</td>
<td>0.15</td>
<td>2.75 × 10⁻⁶</td>
<td>5.00 (2.55–9.80)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>NA</td>
</tr>
<tr>
<td>rs77641320</td>
<td>31:298:229</td>
<td>C/G</td>
<td></td>
<td>imputed</td>
<td>0.35</td>
<td>0.14</td>
<td>1.64 × 10⁻⁶</td>
<td>5.53 (2.75–11.12)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>NA</td>
</tr>
<tr>
<td>rs9468965</td>
<td>31:300:247</td>
<td>T/A</td>
<td></td>
<td>imputed</td>
<td>0.44</td>
<td>0.19</td>
<td>2.24 × 10⁻⁶</td>
<td>4.61 (2.24–8.69)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>NA</td>
</tr>
<tr>
<td>rs35364987</td>
<td>31:309:423</td>
<td>T/C</td>
<td></td>
<td>imputed</td>
<td>0.49</td>
<td>0.27</td>
<td>8.01 × 10⁻⁶</td>
<td>3.76 (2.10–6.74)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>NA</td>
</tr>
<tr>
<td>rs35359945</td>
<td>31:311:374</td>
<td>T/G</td>
<td></td>
<td>imputed</td>
<td>0.49</td>
<td>0.27</td>
<td>9.83 × 10⁻⁶</td>
<td>3.72 (2.08–6.67)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>NA</td>
</tr>
<tr>
<td>rs35278939</td>
<td>31:319:780</td>
<td>G/A</td>
<td></td>
<td>imputed</td>
<td>0.47</td>
<td>0.23</td>
<td>8.01 × 10⁻⁶</td>
<td>4.04 (2.19–7.47)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>NA</td>
</tr>
</tbody>
</table>

The list comprises all SNPs with a P value of < 1 × 10⁻⁵ in the main cohort analysis. Statistical significance and OR (95% CI) for the targeted sequencing cohort is determined by logistic regression with CD4+ cell count as covariate. NA denotes LD indeterminable in targeted sequencing cohort as SNP typing not available.

SNP was imputed in the discovery cohort analysis but not detected and/or called in the sequencing data.
imputed non-synonymous SNPs may be in absolute LD with the original discovery cohort signal SNP (rs5010528), again demonstrating a significant association with nevirapine-induced SJS/TEN. Allele-type inference from the targeted sequencing SNP data also confirmed that both non-synonymous SNPs were in 100% co-occurrence with HLA-C*04:01 (data not shown). Our data do not show an association between any of the other HLA gene loci and other nevirapine-induced hypersensitivity phenotypes.

**HLA-C*04:01 and ERAP1 and ERAP2 SNP interactions**

Given the previously reported interactions between ERAP genes and HLA class I-mediated diseases, in particular ankylosing spondylitis, we determined whether there was an interaction with the carriage of HLA-C*04:01 using the rs5010528 G allele as a proxy SNP (Table S1). There was no significant association (P > 0.05) between the ERAP1 variants and SJS/TEN risk in carriers of HLA-C*04:01. However, both ERAP2 variants showed a nominal association with SJS/TEN risk (P = 0.019, OR 0.43 (95% CI 0.21–0.87)), though this did not pass the Bonferroni threshold for multiple testing (P = 0.0125).

**In silico docking**

In light of the association observed between nevirapine-induced SJS/TEN and an SNP (rs1050409) encoding an amino acid substitution at residue 73 of HLA-C (p.A73E), in silico docking was undertaken to determine the possible effect of the residue substitution on nevirapine binding. The data suggest that none of the predicted modes of nevirapine docking conformation interact with residue 73, which appears to be on the periphery of the peptide-binding domain (Figure 3). The lowest scoring (predicted highest affinity) mode highlights an interaction between nevirapine and residues 33 and 123 in the B pocket (Figure 3). In HLA-C*04:01, residues 33 and 123 are serine and phenylalanine respectively (Figure 3). The majority of other allelotypes do not possess these particular residues (with the exception of C*04:07 and C*14:02). Docking of the metabolite 12-hydroxy-nevirapine was also undertaken, since it has also been suggested as potentially responsible for nevirapine-induced adverse drug reactions; these were in general agreement with those for nevirapine, in that docking seems to take place around the B pocket (e.g. near residues 33 and 123), but with more variability in the different modes predicted than for nevirapine. None of the predicted modes interacted with residue 73 (data not shown). Taken together, the docking results suggest that binding of either nevirapine or 12-hydroxy-nevirapine around the centre of the peptide-binding regions is likely to be important in the mechanism of the immune-mediated reaction.

**Discussion**

The investigation of genetic factors predisposing to serious adverse drug reactions is challenging because of their rarity. Despite this, we have assembled one of the largest cohorts of patients with clinically well-characterized nevirapine hypersensitivity, including SJS/TEN. GWAS analysis of our Malawian discovery cohort (n = 333) identified 15 polymorphisms having a suggestive association with nevirapine hypersensitivity (Table 2). Subsequent analysis of these variants in our replication cohort suggested that three of the SNPs may be potential risk factors (Table 3): rs34213790 3’ of the HLA-B gene locus with all hypersensitivity phenotypes; rs5010528 in the HLA-C gene locus with SJS/TEN, and rs6139258 in the RNF24 gene locus with DILI. The weakest of the above three association signals, SNP rs34213790, is unlikely to be an independent marker of nevirapine hypersensitivity in general, and its association may be due to a haplotype effect between HLA-C*04:01 (rs5010528; see below) and B allelotypes.

SNP rs6139258 in the RNF24 gene locus only marginally failed to pass the Bonferroni threshold of significance in the replication cohort. Very little is known regarding the function of RNF24. However, it is known that it is a protein that interacts with transient receptor potential cation channel 6 (TRPC6), a receptor-activated channel, expressed in liver cells, which plays a role in cellular calcium homeostasis. TRPC6 has been suggested to play a role in hepatoma cell-line proliferation, possibly via a cyclin D-modulated mechanism. Thus RNF24 may have some biological plausibility in the pathogenesis of nevirapine-induced liver injury, and merits further investigation in additional patients with nevirapine-induced DILI, and functional work to uncover the possible mechanisms (if any) of the association.

The most compelling of the three signals, rs5010528, gave an OR of 4.84 for nevirapine-induced SJS/TEN, and was replicated in patients from three countries (Malawi, Uganda and Mozambique) at the Bonferroni threshold (P < 0.05), approaching genome-wide significance in the combined analysis (P = 8.5 × 10^-8) (Table 4). SNP rs5010528 is located within the HLA-C gene locus. High co-occurrence of rs5010528 with HLA-C*04:01 was observed (96.5%) in 116 patients within this study who had previously been HLA typed by sequence-based methods. The association between nevirapine and rs5010528 (as a proxy for C*04:01) can be considered statistically significant when applying an HLA-wide significance threshold of P < 2.5 × 10^-6. Additionally, HLA-C allelotypes imputed from the SNP array data also showed a high co-occurrence (91%) in the main study cohort, suggesting rs5010528 may be a good proxy for HLA-C*04:01. Thus, the GWAS data appear to confirm our previous finding associating HLA-C*04:01 with nevirapine-induced SJS/TEN. Of note, the association of rs5010528 with other hypersensitivity phenotypes was not as strong, suggesting that the risk conferred by rs5010528 and thus HLA-C*04:01 is specific for nevirapine-induced SJS/TEN. The reason for this is unclear, and will require further investigation.

In terms of clinical utility, rs5010528 appears to have little potential as a pre-emptive genetic test. Indeed, based on a prevalence of SJS/TEN in our prospective cohort of 1.07% and assuming a dominant mode of inheritance, the positive (PPV) and negative (NPV) predictive values were 2.8% and 42.4% respectively. For the RNF24 variant (rs6139258) the PPV, based on a prevalence of DILI of 0.63%, is also very low (0.2%).

Only one previous GWAS investigating nevirapine hypersensitivity has been reported, but in a smaller Thai population (72 cases, 77 controls). Patients had a wide variety of rashes, with only 11 grade 4 cases (6.9%), which would be equivalent to our cases with SJS/TEN. The SNP rs9461684 in the HLA-C locus was significantly associated with nevirapine rash, but no HLA alleлотype imputation or HLA sequencing was carried out. In our data, rs9461684 is in high LD with our top SNP, rs5010528 (D’ = 1.0, r2 = 0.972). The discrepancy may be a result of the different LD patterns in the
different ethnic groups studied, as well as the much lower numbers of patients with serious skin reactions in the Thai study.

From the imputed SNP data of the discovery cohort and targeted resequencing data (Table 5), it is clear that rs5010528 is in LD with a functional non-synonymous SNP (rs1050409) that leads to an alanine-to-glutamic acid substitution at residue 73, which lies near to the peptide-binding domain of the HLA-C protein. However, in silico modelling suggested that this residue does not interact with nevirapine in any docking conformation. Two other residues of HLA-C*04:01 (33 and 123 in the model) appear to be the key interactors in the majority of the predicted modes of nevirapine docking (Figure 3), including the most favoured. However, it should be noted that this is a predictive model and further analysis of the HLA-C/nevirapine complex is needed to further elucidate the potential for docking. The association signal at residue 73 (rs1050409) is likely to be a proxy for the 33 and 123 residues also present in HLA-C*04:01. However, this work has provided the first evidence that nevirapine binds to the B pocket of HLA-C*04:01.

ERAP gene variants interact in a protective manner in HLA-mediated diseases such as ankylosing spondylitis in individuals who carry the risk HLA alleles. ERAP1 and ERAP2 are enzymes involved in antigenic peptide precursor trimming prior to loading into HLA class I molecules (and may thus alter the peptide) and may potentially also alter the expression of the risk HLA class I allele. To our knowledge, this is one of the first examinations of whether there is interaction between drug-induced HLA disease and the ERAP genes. We were, however, unable to detect an interaction between ERAP1 variants and HLA-C*04:01 in African patients with SJS/TEN. However, a nominal association (P = 0.019) was observed for both ERAP2 SNPs (Table S1). A limitation of our analysis is the small sample size, particularly given the much larger numbers that have been studied in ankylosing spondylitis. Nevertheless, the possibility of an association with ERAP is intriguing, and needs further investigation not only with nevirapine-induced hypersensitivity, but also with other HLA-related adverse drug reactions.

In identifying an SNP in the HLA-C locus that appears to be a proxy for the HLA-C*04:01 allele, as a risk factor for nevirapine-induced SJS/TEN, this study has added further weight to existing evidence. The data generated also suggest that, in sub-Saharan African HIV patients, no other strong, significant genetic risk factors for nevirapine hypersensitivity exist that could be utilized as clinical predictive markers. However, the data are valuable in terms of the mechanistic insights they provide. Additionally, in silico analysis has identified two putative HLA-C peptide residues that are predicted to be key for the binding of nevirapine, which warrant further investigation as to their role in the pathogenesis of SJS/TEN. Further work is also needed to determine the reasons for organ-specific toxicities in different patients.

Acknowledgements
We thank the patients and staff of the ART clinic of Queen Elizabeth Central Hospital, Blantyre, in particular Mr S. Kaunda, Clinical Officer. We also thank Dr Christiane Hertz-Fowler, Dr Margaret Hughes, Dr Lisa Olahan and Dr Anita Lucaci from the Centre for Genomic Research for undertaking the library preparation and sequencing of the MHC region in the cohort of 32 DNA samples.

Funding
The initial GWAS was funded by the International Serious Adverse Events Consortium (iSAEC). The iSAEC is a non-profit organization dedicated to identifying and validating DNA variants useful in predicting the risk of drug-related serious adverse events. The Consortium brings together the pharmaceutical industry, regulatory authorities and academic centres to address clinical and scientific issues associated with the genomics of drug-related serious adverse events. The iSAEC’s current funding members include: Abbott, Amgen, AstraZeneca, Daiichi Sankyo, GlaxoSmithKline, Merck, Novartis, Pfizer, Takeda and the Wellcome Trust. Mas Chapa was funded by a 3 year Wellcome Trust training fellowship WT078857MA administered through the University of Liverpool. Malawi-Liverpool-Wellcome Trust Clinical Research Programme is funded through a Core Programme Grant award from the Wellcome Trust. Munir Pirmohamed is a National Institute for Health Research Senior Investigator, and also wishes to thank the MRC Centre for Drug Safety Science for support.

The DART study was supported by the UK Medical Research Council (grant number G0600344), the UK Department for International Development and the Rockefeller Foundation.

Andrew P. Morris is a Wellcome Trust Senior Research Fellow in Basic Biomedical Science (grant number WT098017). Louise Y. Takeshita is funded by a PhD fellowship from CNPq (National Council for Scientific and Technological Development, Brazil). Panos Deloukas’ work forms part of the research themes contributing to the translational research portfolio of Barts Cardiovascular Biomedical Research Unit which is supported and funded by the National Institute for Health Research.

Transparency declarations
All authors: no conflicts of interest to declare.
GlaxoSmithKline, Gilead Sciences, Boehringer-Ingelheim and AbbVie donated drugs for the DART study.

Supplementary data
Figure S1 and Table S1 are available as supplementary data at JAC Online.

References
1 van Leth F, Phanuphak P, Ruxungham K et al. Comparison of first-line antiretroviral therapy with regimens including nevirapine, efavirenz, or both drugs, plus stavudine and lamivudine: a randomised open-label trial, the 2NN Study. Lancet 2004; 363: 1253–63.
GWAS of nevirapine hypersensitivity


27. Purcell S, Neale B, Todd-Brown K et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81: 559–75.


