Association of Liver Injury From Specific Drugs, or Groups of Drugs, With Polymorphisms in HLA and Other Genes in a Genome-Wide Association Study


BACKGROUND & AIMS: We performed a genome-wide association study (GWAS) to identify genetic risk factors for drug-induced liver injury (DILI) from licensed drugs without previously reported genetic risk factors. METHODS: We performed a GWAS of 862 persons with DILI and 10,588 population-matched controls. The second set of cases was recruited before May 2009 in Europe (n = 137) and the United States (n = 274). The second set of cases were identified from May 2009 through May 2013 from international collaborative studies performed in Europe, the United States, and South America. For the GWAS, we included only cases with patients of European ancestry associated with a particular drug (but not fluvoxacin or amoxicillin-clavulanate). We used DNA samples from all subjects to analyze HLA genes and single nucleotide polymorphisms. After the discovery analysis was concluded, we validated our findings using data from 283 European patients with diagnosis of DILI associated with various drugs. RESULTS: We associated DILI with rs114577328 (a proxy for A*33:01 allele; odds ratio [OR], 2.7; 95% confidence interval

See Covering the Cover synopsis on page 915.

BACKGROUND & AIMS: We performed a genome-wide association study (GWAS) to identify genetic risk factors for drug-induced liver injury (DILI) from licensed drugs without previously reported genetic risk factors. METHODS: We performed a GWAS of 862 persons with DILI and 10,588 population-matched controls. The first set of cases was recruited before May 2009 in Europe (n = 137) and the United States (n = 274). The second set of cases were identified from May 2009 through May 2013 from international collaborative studies performed in Europe, the United States, and South America. For the GWAS, we included only cases with patients of European ancestry associated with a particular drug (but not fluvoxacin or amoxicillin-clavulanate). We used DNA samples from all subjects to analyze HLA genes and single nucleotide polymorphisms. After the discovery analysis was concluded, we validated our findings using data from 283 European patients with diagnosis of DILI associated with various drugs. RESULTS: We associated DILI with rs114577328 (a proxy for A*33:01 allele; odds ratio [OR], 2.7; 95% confidence interval

1Department of Systems Biology, Columbia University, New York, New York;2National Institute for Health Research, Nottingham Digestive Diseases Biomedical Research Unit, Nottingham University Hospital, National Health Service Trust, and University of Nottingham, Nottingham, United Kingdom;3Department of Internal Medicine, Landspitali University Hospital, Reykjavik, Iceland;4Unidad de Gestión Clínica Digestivo, Instituto de Investigación Biomédica de Málaga, Hospital Universitario Virgen de la Victoria, Universidad de Málaga, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas, Málaga, Spain;5Departamento de Gastroenterología, Escuela de Medicina, Pontificia Universidad Católica de Chile, Santiago, Chile;6Duke University, Durham, North Carolina;7Université de Toulouse, Toulouse, France;8Department of Internal Medicine, University of North Carolina-Chapel Hill, Chapel Hill, North Carolina;9Servicio de Gastroenterología y Hepatología, Universidad Nacional de Rosario, Rosario, Argentina;10Universidad de Valladolid, Valladolid, Spain;11Institute of Experimental and Clinical Pharmacology, University Hospital Schleswig-Holstein, Kiel, Germany;12Division of Gastroenterology and Hepatology, Indiana University School of Medicine, Indianapolis, Indiana;13University Hospital, Verona, Italy;14Institute of Cellular Medicine, Newcastle University, Newcastle Upon Tyne, United Kingdom;15Broad Institute of Harvard and Massachusetts Institute of Technology, Boston, Massachusetts;16Medical Research Institute, University of Dundee, Ninewells Hospital, Dundee, United Kingdom;17University of Michigan, Ann Arbor, Michigan;18Department of Medical Sciences and Science for Life Laboratory, Uppsala University, Uppsala, Sweden;19Facultad de Medicina, Universidad de la Republica, Montevideo, Uruguay;20Fundació Institut Català de Farmacologia, Hospital Universitari Vall d’Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain;21Department of Clinical Pharmacology and Toxicology, University Hospital and University of Zurich, Zurich, Switzerland;22Clinical Research Unit for Pulmonary Diseases, Helsinki University Central Hospital, Helsinki, Finland;23Liver Unit, Centre Hospitalier Universitaire, St Elis Hospital, Montpellier, France;24Division of Pharmacopidemiology and Clinical Pharmacology, Utrecht University, Utrecht, Netherlands;25School of Medicine and Public Health, University of Newcastle, New South Wales, Australia;26Department of Primary Care and Public Health Sciences, King’s College, London, United Kingdom;27Department of Molecular and Clinical Pharmacology, University of Liverpool, Liverpool, United Kingdom;28Centre for Liver Disease Research, School of Medicine, The University of Queensland, Brisbane, Queensland, Australia;29Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders, Shanghai Jiao Tong University, Shanghai, China;30National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland;31University of Southern California, Los Angeles, California;32University of North Carolina Institute for Drug Safety Sciences, Eshelman School of Pharmacy, Chapel Hill, North Carolina;33Target Sciences, GlaxoSmithKlein, King of Prussia, Pennsylvania; and34UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, North Carolina
Hepatotoxicity is the second most common cause of drug attrition during development, and post-marketing withdrawal, and idiosyncratic drug-induced liver injury (DILI) accounts for 11%–17% of cases of acute liver failure in the United States and Europe. The typical incidence of DILI varies from approximately 1% with the anti-tumor necrosis factor agents to 0.04% with some widely used antimicrobials such as amoxicillin-clavulanate. During the past 15 years, increasing progress on identifying genetic risk factors for DILI has been made. In particular, associations with HLA class I and II alleles have been reported for DILI caused by a range of drugs, though a particular HLA genotype does not appear to be relevant to all forms of idiosyncratic DILI.

Previously, a genome-wide association study (GWAS) involving cohorts of DILI cases related to one particular drug only resulted in identification of 1 or more drug-specific HLA risk alleles. A large study involving 783 DILI cases due to a range of different drugs also resulted in a genome-wide significant HLA signal, but this association was abolished once 296 cases of DILI due to fluvoxacinil and amoxicillin-clavulanate were excluded. This partly reflects the fact that amoxicillin-clavulanate is a very common cause of DILI worldwide and fluvoxacinil is an equally common cause in a number of Northern European countries. Therefore, DNA collections from DILI cases will generally be highly enriched in cases relating to these 2 drugs, making detection of associations related to other compounds more difficult.

We have expanded our previous study of DILI caused by a range of different drugs, and after excluding cases relating to amoxicillin-clavulanate and fluvoxacinil, we have more than doubled the number of cases, with additions from Europe, Australia, South America, and the United States. We now report that HLA-A*33:01 is associated with risk of DILI, particularly due to terbinafine, fenofibrate, and ticlopidine, and especially with a cholestatic or mixed phenotype. We have also found novel non–major histocompatibility complex (MHC)-related signals apparently shared across a range of different drugs; an intronic single nucleotide polymorphism (SNP) in the lipopolysaccharide-responsive vesicle trafficking, beach and anchor containing (LRBA) gene is associated with hepatocellular DILI and an intergenic SNP on chromosome 2, rs72631567, with DILI generally. An additional drug-specific genome-wide significant signal that could not be confirmed is also reported.

Materials and Methods

Drug-Induced Liver Injury Discovery Cohort

The cases in the study were from 2 separate recruitment phases. Phase 1 consists of 411 cases included in a previous study (from the Drug-Induced Liver Injury Network [DILIN], DILIGEN, and Eudragene) and phase 2 more recently recruited cases (n = 451) of which a small subset was included in a recent report.

Phase 1 cases. These cases included 411 DILI cases not due to amoxicillin-clavulanate or fluvoxacinil, with a defined casual drug and with causality score greater than possible (Roussel Uclaf Causality Assessment Method score ≥3) recruited in Europe (n = 137) or the United States (n = 274) before 2009. Clinical characteristics of these cases and methods used for genotyping have been described in detail previously. Additional exome chip analysis (Illumina Infinium HumanCoreExome BeadChip) was performed on 150 of these 411 cases at the Broad Institute, Boston, MA.

Phase 2 case recruitment: International Drug-Induced Liver Injury Consortium. The International Drug-Induced Liver Injury Consortium (iDILIC) cases were recruited between May 2009 and May 2013 as part of an international collaborative study involving recruitment centers in the United Kingdom (Newcastle, Nottingham, Liverpool, London, Dundee), Sweden (Uppsala and Gothenburg), Spain (Malaga and Barcelona), France (Montpellier), The Netherlands (Utrecht), Germany (Kiel), Australia (Brisbane), Switzerland (Zurich), Finland (Helsinki), Argentina (Rosario), Uruguay (Montevideo), and Chile (Santiago). All participants provided

*Authors share co-first authorship; §Authors share co-senior authorship.

Abbreviations used in this paper: AF, allele frequency; CM, cholestatic-mixed; DILI, drug-induced liver injury; DILIN, Drug-Induced Liver Injury Network; GWAS, genome-wide association study; HC, hepatocellular; iDILIC, International Drug-Induced Liver Injury Consortium; LRBA, lipopolysaccharide-responsive vesicle trafficking, beach and anchor containing; MHC, major histocompatibility complex; OR, odd ratio; SNP, single nucleotide polymorphism.

© 2017 by the AGA Institute
0016-5085/$36.00
http://dx.doi.org/10.1053/j.gastro.2016.12.016
written informed consent and each study had been approved by the appropriate national or institutional ethical review boards. For the GWAS, only cases of European ancestry where there were at least 2 cases due to a particular drug available (when phase 1 cases from Europe and the United States were also considered) and where the DILI was not due to either fluoxacinil or amoxicillin-clavulanate were included (n = 339). Clinical inclusion criteria for all cases were those described by Aithal et al.\textsuperscript{15}

**Phase 2 case recruitment: Drug-Induced Liver Injury Network.** Details of the US-based DILIN prospective study including institutional review board approval information have been described previously.\textsuperscript{16} A total of 112 eligible new cases of European ancestry and 18 years and older were included in the current GWAS. These new cases were selected from the larger DILIN sample collection, such that only cases relating to drugs also included among the iDILIC cases were represented. Laboratory inclusion criteria were as described previously.\textsuperscript{16} Patients were excluded if there was known or suspected acetaminophen overdose, history of bone marrow or liver transplantation before DILI onset, or history of immune-related liver disease, such as autoimmune hepatitis.

**Additional Cases Used for Confirmation of Associations**

After the discovery analysis was concluded, we enrolled an additional 283 European patients with diagnosis of DILI across multiple causal drugs (6 from iDILIC and 277 from DILIN networks recruited subsequent to the GWAS). The causal drug distribution is reported in Supplementary Table 1A. An additional 12 statin DILI samples from the Spanish iDILIC network and 3 UK-DILICEN cases were recruited later in the study to confirm the class-specific association (Supplementary Table 1B).

Of the 283, we used 272 DILI cases to directly type SNPs associated across multiple drugs or specific for drugs and drug classes and 11 DILI cases for HLA typing to confirm HLA associations. An additional Chinese terbinafine DILI sample was also HLA typed.

**Causality Assessment**

The iDILIC cases were evaluated by application of the Council for International Organizations of Medical Science scale, also called the Roussel Uclaf Causality Assessment Method\textsuperscript{15} and by expert review by a panel of 3 hepatologists. The pattern of liver injury was classified according to the International Consensus Meeting Criteria.\textsuperscript{17} Only cases having at least possible causality (score \(\geq 3\)) were included in the study. For all cases in DILIN, causality assessment was by expert consensus, as described previously.\textsuperscript{16}

**Controls**

Because DILI has a very low prevalence, we used general population samples as study controls. We selected 10588 European ancestry controls from multiple available sources; Welcome Trust Case Control Consortium (http://www.wtccc.org.uk), the Population Reference Sample,\textsuperscript{18} PGX40001,\textsuperscript{19} and Spanish Bladder cancer cohort (phs000346.v1) from dbGAP.\textsuperscript{20} In order to increase the case-control ratio for Italian, Spanish and Swedish, we added samples from Hypergenes cohort (http://www.hypergenes.eu/dissemination.html#pub), the National Spanish DNA Bank (http://www.bancoadn.org/), Italian Penicillin Tolerant Controls, and the Swedish Twin Registry (http://ki.se/en/research/the-swedish-twin-registry).

**Genotyping**

DNA preparation from phase 2 cases. For iDILIC cases, DNA was prepared as described previously.\textsuperscript{1} DILIN DNA was extracted from lymphocytes and stored at the National Institute of Diabetes and Digestive and Kidney Diseases biosample repository at Rutgers University, Piscataway, NJ.

**Genome-wide analysis.** Genome-wide genotyping of the phase 2 and 150 phase 1 cases was performed by the Broad Institute, Boston, MA by Illumina Infinium HumanCoreExome BeadChip. iDILIC and DILIN cases were genotyped in 2 separate batches. A total of 505,740 markers shared across the batches passed quality control and no samples were excluded for low-quality profile (see Supplementary Materials and Methods).

Details on the genotype data available for each control collection are reported in Supplementary Table 2.

**Imputation.** SNP imputation was performed in batches dividing the cohorts according to genotyping platforms. Imputation methods are described in detail in the Supplementary Material. For HLA genotypes, 4-digit HLA alleles were inferred using HIBAC.\textsuperscript{21}

**Single nucleotide polymorphism genotyping.** The top associated imputed SNPs were validated by SNP genotyping in subsets of iDILIC cases and in the overall DILIN cohort (see Supplementary Material). The SNPs were further confirmed in additional cases using TaqMan predesigned and custom SNP genotyping assays (ThermoFisher Scientific, Waltham, MA) in accordance with the manufacturer’s recommendations.

**HLA genotyping.** High-resolution genotyping of HLA-A, B, C, DRB1, DQA1, and DQB1 was performed on selected cases by Histogenetics (Ossining, NY). Sequencing data files were analyzed using Histogenetics’ proprietary analysis software (Histomatcher and Histomagic) for HLA genotype calling. Allele assignments are based on IMGT/HLA Database release version 2.21.0, dated April 2008 (http://www.ebi.ac.uk/imgt/hla/).

**Statistical Analysis**

The effect of population structure was assessed through principal components analysis using the smartPCA program from the EIGENSTRAT package (version 3.0).\textsuperscript{22} Single marker and haplotype association analyses and heterogeneity test analyses were carried out by PLINK.\textsuperscript{23} The statistical association of each marker, HLA alleles and SNPs, was determined in a logistic regression framework with scores for the first 7 principal components as covariates under an additive model using PLINK. We used the same statistical test for subpopulation analyses, using 2, 7, and 10 most significant principal components as covariates in Italian, Spanish, and North European populations, respectively. We set the genome-wide traditional significance \(P\)-value threshold to \(5.0 \times 10^{-8}\) to correct for multiple testing.\textsuperscript{24} When we obtained genome-wide significant signals, we tested for independent effects from the neighboring variants by including the most associated variants as a covariate and then testing the significance of others in the region. We also tested interaction effects among them by including interaction terms in the logistic regression. Differences in clinical characteristics among sample groups were tested by Fisher’s
Results

Clinical Characteristics of the Cases

Clinical details of the DILI cases included in the main GWAS are summarized in Table 1. A variety of different causative drugs were represented but the most common was diclofenac with 67 cases, followed by nitrofurantoin with 64 cases. A few drugs, including azathioprine, isoniazid, fenofibrate, and diclofenac had significantly disproportionate number of cases in 1 of the 2 recruitment phases. Details of all the causative drugs are shown in Supplementary Table 3.

Overall Analysis

The discovery cohort included 862 European ancestry DILI cases (411 from phase 1 and 451 from phase 2) and 10,588 controls. Principal components analysis showed that all cases (including those from South America) clustered within 3 major groups (Italian, Spanish, and Northern European) and matched with the population controls (Supplementary Figure 1A). Consistent with the previous study, phase 1 cases were predominantly Northwest European. The most significant genome-wide associated SNPs were rs72631567 on chromosome 2 (odds ratio [OR], 2.0; 95% confidence interval [CI], 1.6–2.5; \( P = 9.7 \times 10^{-9} \)) and rs114577328 in the MHC region of chromosome 6 (OR, 2.7; 95% CI, 1.9–3.8; \( P = 2.4 \times 10^{-8} \)) (Figure 1A, Table 2, and Supplementary Figures 2 and 3). Data for both SNPs had been obtained by imputation in cases and controls and subsequently validated by SNP typing (see Supplementary Methods). The associations were consistent among geographic clusters and study phases (Supplementary Table 4) and not due to artifact(s) of population structure, missing genotypes rate (Supplementary Table 5) or variability in imputation quality among populations or genotyping platforms (see Supplementary Methods).

For the chromosome 2 SNP rs72631567, breakdown by drug showed that 10 unrelated drug causes had an OR >2.0 with at least 2 carriers (Supplementary Table 6). Ciprofloxacin-related cases showed the strongest association (\( n = 21; \text{OR}, 7.4; 95\% \text{CI}, 17.3–161; \text{P} = 4.0 \times 10^{-6} \)).

The chromosome 6 SNP rs114577328 is the SNP proxy of an uncommon HLA class I allele, HLA-A*33:01. Indeed, this SNP was in near-perfect linkage disequilibrium with A*33:01 (\( r^2 = 0.98 \)). From the imputed HLA allele assignments, a strong association with DILI for this allele is confirmed (\( \text{OR}, 2.6; 95\% \text{CI}, 1.8–3.7; \text{P} = 8.0 \times 10^{-8} \)), Supplementary Figure 4). Including rs114577328 or A*33:01 as a covariate removed any association in the MHC region, indicating that there is only 1 MHC association signal.

Table 1. Clinical Details of the Drug-Induced Liver Injury Cases Included in the Genome-Wide Association Study

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Phase 1 (n = 411)</th>
<th>Phase 2 (n = 451)</th>
<th>Combined (n = 862)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical information</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age, y</td>
<td>51</td>
<td>54</td>
<td>53</td>
</tr>
<tr>
<td>% Female, %</td>
<td>63.0</td>
<td>60.5</td>
<td>61.8</td>
</tr>
<tr>
<td>Mean alanine aminotransferase, IU/L</td>
<td>895.1</td>
<td>757.7</td>
<td>822.2</td>
</tr>
<tr>
<td>Mean alkaline phosphatase, IU/L</td>
<td>388.2</td>
<td>282.5</td>
<td>330.6</td>
</tr>
<tr>
<td>Mean latency, d</td>
<td>201.7</td>
<td>177.4</td>
<td>188.2</td>
</tr>
<tr>
<td>Injury type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cholestatic</td>
<td>76</td>
<td>87</td>
<td>163</td>
</tr>
<tr>
<td>Hepatocellular</td>
<td>202</td>
<td>272</td>
<td>474</td>
</tr>
<tr>
<td>Mixed</td>
<td>69</td>
<td>91</td>
<td>160</td>
</tr>
<tr>
<td>Not available(^a)</td>
<td>64</td>
<td>1</td>
<td>65</td>
</tr>
<tr>
<td>Genotype chip</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illumina 1 M</td>
<td>261</td>
<td></td>
<td>261</td>
</tr>
<tr>
<td>Illumina 1M/Illumina Infinium HumanCoreExome BeadChip</td>
<td>150</td>
<td></td>
<td>150</td>
</tr>
<tr>
<td>Illumina Infinium HumanCoreExome BeadChip</td>
<td>447</td>
<td></td>
<td>447</td>
</tr>
<tr>
<td>Illumina HumanOmniExpress BeadChip</td>
<td>4</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Country of birth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>274</td>
<td>112</td>
<td>386</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>71</td>
<td>79</td>
<td>150</td>
</tr>
<tr>
<td>Spain</td>
<td>16</td>
<td>95</td>
<td>111</td>
</tr>
<tr>
<td>Sweden</td>
<td>81</td>
<td></td>
<td>81</td>
</tr>
<tr>
<td>France</td>
<td>30</td>
<td>7</td>
<td>37</td>
</tr>
<tr>
<td>Germany</td>
<td></td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Italy</td>
<td>16</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>56</td>
<td>60</td>
</tr>
</tbody>
</table>

\(^a\)Because of the retrospective nature of the phase 1 study, minimal clinical information needed to establish the type of injury were not available for a subset of initial nonsteroidal anti-inflammatory drug–induced liver injury cases from DILIGEN because of missing alkaline phosphatase and upper limit of normal values.
The A*33:01 association appears independent of the chromosome 2 signal, because rs72631567, when conditioned on A*33:01, showed an almost unchanged effect size (ORrs72631567, 1.7; 95% CI, 1.25–2.2; P = .0006). There was no statistically significant interaction effect between the 2 signals (P = .5).

Breakdown by drug showed DILI due to terbinafine was most strongly associated with the HLA-A*33:01 signal (OR, 40.5; 95% CI, 12.5–131.4; P = 6.7 × 10⁻¹⁰) and a similarly strong association was seen with rs114577328 (OR, 58.7; 95% CI, 18.3–188.2; P = 7.3 × 10⁻¹²; Figure 1B and Supplementary Figure 6). As summarized in Table 3, in addition to terbinafine cases, cases due to 6 additional drugs showed an association with A*33:01 with P < .01. The largest case subset related to terbinafine, but we found that A*33:01 was also a risk factor for ticlopidine (OR, 163.1; 95% CI, 16.2–1642.0; P = .00002), methyldopa (OR, 97.8; 95% CI, 12.3–743.0; P = .00001) and fenofibrate DILI (OR, 58.7; 95% CI, 12.3–279.8; P = 3.2 × 10⁻⁷). Indeed, although fewer positive carriers were observed, A*33:01 also seems to be a common risk factor for enalapril (OR, 34.8), sertraline (OR, 29), and erythromycin DILI (OR, 10.2). An erythromycin case was positive for A*33:03, an allele rare in European population controls (allele frequency [AF], 0.002) which belongs to the A*33 group. Overall, we found that 87% (n = 36) of the A*33:01-positive carriers were also positive for HLA-B*14:02 and HLA-C*08:02. The haplotype showed a larger OR than A*33:01 as single marker in terbinafine (ORhaplotype, 49.2; P = 9.54 × 10⁻¹¹), ticlopidine (ORhaplotype, 201; P = 7.2 × 10⁻⁶), fenofibrate (ORhaplotype, 68.5; P = 1.1 × 10⁻⁷), and erythromycin (ORhaplotype, 13.1; P = .002) DILI, but not with DILI as a phenotype (ORhaplotype, 2.7; P = 1.6 × 10⁻⁷; Supplementary Table 7).

We verified the imputed A*33:01 genotype by sequence-based HLA typing in 35 cases related to the main A*33:01-associated drugs (Supplementary Table 8). The A*33:01 predictions were confirmed in all cases except that one methyldopa case was negative for this allele (false positive) and an additional terbinafine case was a carrier (false negative). This validation result suggests that methyldopa might not share the HLA risk factor. The validation confirmed that all the A*33:01-positive terbinafine cases carried the complete HLA A*33:01-B*14:02-C*08:02 haplotype, increasing the strength of the haplotype association in the terbinafine DILI cases (ORhaplotype, 70; P = 8.7 × 10⁻¹³) and in the overall analyses (ORhaplotype, 2.8; P = 5.1 × 10⁻⁸). We also typed the A*33:01 proxy SNP across DILIN cases to confirm imputed genotypes. We found only one new carrier of the minor allele, not related to the major-A*33:01 associated drugs.

**Analysis by Type of Injury and Causative Drugs**

We further investigated the association of genotypes with particular patterns of DILI by grouping the cases into hepatocellular (HC) and cholestatic/mixed (CM) pattern. The chromosome 2 association described was similar in the
2 phenotypic categories (direct comparison between CM cases vs HC cases; logistic $P = .5$), although the effect was marginally stronger in the CM cases (Table 2). The association with rs114577328 was genome-wide significant only in the CM cases ($n = 323$; OR, 5.3; 95% CI, 3.4–8.2; $P = 4.5 \times 10^{-14}$; Figure 2B and Supplementary Figure 6) and similarly with A*33:01 (OR, 5.1; 95% CI, 3.3–7.9; $P = 4.2 \times 10^{-13}$; Figure 2B and Supplementary Figure 6). Conditional analysis on the variant and HLA allele indicated only one genetic association was present in the region, as shown for the main analysis (Supplementary Figure 7). There was no association between the proxy SNP or A*33:01 in the HC cases ($n = 474$; OR for A*33:01, 1.5; 95% CI, 0.82–2.6; $P = .19$; Table 2). The A*33:01-B*14:02-C*08:02 haplotype showed an OR of 5.6; $P = 2.5 \times 10^{-13}$ in CM cases.

The CM only terbinafine-specific OR increased 2-fold compared with the value for all terbinafine cases (OR, 88.1, 95% CI, 19.3–402.4; $P = 7.5 \times 10^{-9}$) because all the A*33:01 carriers belonged to this injury type. Following the injury correlation pattern established for terbinafine, A*33:01 appeared to be a stronger risk factor for CM injury than for HC injury also for fenofibrate, ticlopidine, enalapril, and erythromycin-related DILI. This was not the case for injury due to sertraline and methyldopa. These top 7 drugs account for 51% ($n = 21$) of all A*33:01 positive cases (Supplementary Table 9). Sixteen other drugs account for the rest of the carriers showing slight enrichment in CM phenotypes, which showed a marginal association with A*33:01 (OR, 2.6; 95% CI, 1.4–4.9; $P = .003$) (Supplementary Tables 9 and 10).

We detected a new HC-specific genome-wide significant signal on chromosome 4 (Figure 2B). The signal lies within the LRBA gene with the imputed variant rs28521457, located in an intronic region, the most significant SNP (OR, 2.1; 95% CI, 1.6–2.7, $P = 4.8 \times 10^{-5}$) (Table 2 and Figure 2B). The AF for this SNP in the CM cases (AF, 0.04) was comparable to that in controls with no evidence of association with this phenotype. The risk allele was carried by >4% of the HC cases in cases due to a total of 45 drugs, but in general, there were no drug-specific signals (Supplementary Table 11).

We also investigated associations with particular causative drugs or specific therapeutic classes where a group including more than 40 samples was available. Detail on the groups studied is summarized in Supplementary Table 12. Genome-wide significance was detected only for one group examined, the statins, with no significant signals for the other drug classes (Supplementary Figure 8 and Supplementary Table 13). In the case of the statins, rs116561224, a common intergenic SNP on chromosome 18, was genome-wide significant (OR, 5.4; 95% CI, 3.0–9.5; $P = 7.1 \times 10^{-9}$; Figure 2C and Supplementary Figure 9) with the signal mainly driven by simvastatin (Supplementary Table 14).

### Confirmation of Associations

The European cohort used to confirm the associations ($n = 283$) had a wider range of causal drugs, mostly different from the discovery cohort (Supplementary

### Table 2. Association Effect Size of rs72631567, A*33:01, and rs28521457 Across Different Liver Injury Patterns

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Associated variant</th>
<th>OR</th>
<th>95% CI</th>
<th>$P$ value</th>
<th>AF cases</th>
<th>AF controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entire DILI cohort</td>
<td>rs72631567</td>
<td>2.0</td>
<td>1.6–2.5</td>
<td>$9.7 \times 10^{-9}$</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>A*33:01</td>
<td>2.6</td>
<td>1.8–3.7</td>
<td>$7.0 \times 10^{-8}$</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>rs28521457</td>
<td>1.5</td>
<td>1.3–1.9</td>
<td>$5.1 \times 10^{-5}$</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>Cholestatic and mixed DILI</td>
<td>rs72631567</td>
<td>2.4</td>
<td>1.7–3.4</td>
<td>$9.5 \times 10^{-7}$</td>
<td>0.06</td>
<td>0.03</td>
</tr>
<tr>
<td>cohort</td>
<td>A*33:01</td>
<td>5.0</td>
<td>3.3–7.9</td>
<td>$4.2 \times 10^{-13}$</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>rs28521457</td>
<td>1.0</td>
<td>0.7–1.5</td>
<td>$9$</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Hepatocellular DILI cohort</td>
<td>rs72631567</td>
<td>1.6</td>
<td>1.2–2.3</td>
<td>$2.5 \times 10^{-3}$</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>A*33:01</td>
<td>1.5</td>
<td>0.8–2.6</td>
<td>$19$</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>rs28521457</td>
<td>2.1</td>
<td>1.6–2.7</td>
<td>$4.8 \times 10^{-9}$</td>
<td>0.08</td>
<td>0.040</td>
</tr>
</tbody>
</table>

### Table 3. Association Effect Size of A*33:01 Signal for the Causal Drugs Enriched in A*33:01-Positive Carriers

<table>
<thead>
<tr>
<th>Drug</th>
<th>No. of cases tested</th>
<th>No. of A*33:01 alleles in cases</th>
<th>OR</th>
<th>95% CI</th>
<th>$P$ value</th>
<th>CF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ticlopidine</td>
<td>5</td>
<td>4</td>
<td>163.1</td>
<td>16.2–1642</td>
<td>.00002</td>
<td>0.8</td>
</tr>
<tr>
<td>Methyldopa</td>
<td>4</td>
<td>2</td>
<td>97.8</td>
<td>12.8–743.8</td>
<td>.00001</td>
<td>0.5</td>
</tr>
<tr>
<td>Fenofibrate</td>
<td>7</td>
<td>4</td>
<td>58.7</td>
<td>12.3–279.8</td>
<td>3.2 $\times 10^{-7}$</td>
<td>0.43</td>
</tr>
<tr>
<td>Terbinafine</td>
<td>14</td>
<td>6</td>
<td>40.5</td>
<td>12.5–131.4</td>
<td>6.7 $\times 10^{-10}$</td>
<td>0.43</td>
</tr>
<tr>
<td>Enalapril</td>
<td>4</td>
<td>2</td>
<td>34.8</td>
<td>3.9–302.9</td>
<td>.001</td>
<td>0.5</td>
</tr>
<tr>
<td>Sertraline</td>
<td>5</td>
<td>2</td>
<td>29</td>
<td>4–207.2</td>
<td>.0008</td>
<td>0.4</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>10</td>
<td>2</td>
<td>10.2</td>
<td>2–51.7</td>
<td>.005</td>
<td>0.2</td>
</tr>
</tbody>
</table>

CF, carrier frequency.

*Causal drug involved.*
Table 1). Later in time, we had access to 15 additional cases relating specifically to the statin cohort.

The A*33:01 association was further investigated in the additional cases by directly genotyping rs114577328 in 272 cases and by direct HLA typing on 11 additional samples who developed DILI due to drugs for which we had detected an enrichment in A*33:01 alleles in the discovery cases. Overall, the rs114577328 carriers were enriched in cases from drugs previously associated with the allele (Supplementary Table 15). Eight of all 23 additional cases relating to drugs previously associated with A*33:01 were shown to carry this allele or the proxy SNP (AF, 0.17) compared with an expected population frequency of 0.01. We specifically confirmed the association of A*33:01/rs114577328 with terbinafine having a carrier frequency of 0.63 (5 of 8 terbinafine-related cases across both the injury types) and with sertraline at a carrier frequency of 0.75 (3 of 4 sertraline-related cases) (Supplementary Table 15). Although fenofibrate had a high carriage rate for A*33:01 in the discovery cohort, none of the 7 additional cases carried this allele or the proxy SNP. Few additional cases were available for other A*33:01-related drugs to confirm the association.

Interestingly, a terbinafine DILI case from Finland was positive for A*33:05, a very rare allele in the general European population (AF, 0.0001, USA NMDP European Caucasian in http://www.allelefrequencies.net/; n = 1,242,890) and Finnish populations. An additional terbinafine DILI case of Chinese origin was positive for A*33:03. In total, 10 of the 24 additional cases (23 European cases and 1 Chinese case)

Figure 2. Manhattan plot displaying the association results for (A) cholestatic/mixed only cases (n = 323); (B) hepatocellular only cases (n = 474 cases); (C) statin cases (n = 59). SNPs in green have a significance level <5 × 10⁻⁶ and red have a significance level <5 × 10⁻⁸.
were carriers of an A*33 allele, in line with expectations based on the effects observed in the discovery sample.

We further genotyped rs72631567 and rs28521457 in 272 additional European cases. The rs72631567 and rs28521457 variants were found at AFs comparable to those for controls (AF_{rs72631567} = 0.022 and AF_{rs28521457} in HC only, 0.025) and so the association was not confirmed. However, rs72631567 carriers were slightly enriched in ciprofloxacin, atorvastatin, and mercaptopurine-induced DILI cases, as in the discovery cohort, with ORs in the same direction in both cohorts (Supplementary Table 16). Similarly, rs28521457 carriers seemed to be more common in the same subgroup of causal drugs in both cohorts (Supplementary Table 17). This suggested a limited replication of the signal for these drugs.

We also attempted to confirm the rs116561224 signal for statins. The number of additional cases available for this purpose was small (n = 29, Supplementary Table 1B) with only 4 simvastatin cases. None of the statin cases were positive for rs116561224 so the signal could not be confirmed.

### Discussion

Our previous studies have been successful in identifying genetic risk factors for both fluoxacinil and amoxicillin-clavulanate DILI. However, our most recent GWAS did not identify any risk factors that were common for DILI in general or specific genetic risk factors for DILI due to individual drugs, which accounted for a smaller number of cases of DILI. The current study included 451 additional cases of DILI due to a wide variety of causative drugs, including at least 10 DILI cases relating to each of 22 different drugs. This increase in numbers and the exclusion of the amoxicillin-clavulanate and fluoxcinil cases together with use of improved imputation methods has enabled the detection and confirmation of a novel genome-wide significant signal relating to a relatively rare HLA class I allele A*33:01. Although 3 other interesting signals were detected in the course of the study, an intergenic signal on chromosome 2, an intrinsic SNP in LRBA in HC cases only and a signal on chromosome 18 for statins, the failure to confirm these signals is a limitation. There are some indications that, as observed for HLA-A*33:01, the chromosome 2 and LRBA signals are shared across multiple unrelated drugs instead of being non–drug-specific risk variants. As supporting evidence, the chromosome 2 signal has been consistently associated in both replication and discovery cohorts with DILI due to ciprofloxacin, atorvastatin, and mercaptopurine. There remains a possibility that replication could be achieved in a larger study involving a different mix of causative drugs, but the degree of heterogeneity in drugs originally associated with the signals also increases the risk that these were chance observations.

Interestingly, unlike previously recognized HLA associations for DILI, A*33:01 also appears to be a risk factor for DILI due to several, structurally unrelated drugs. Our results also suggest that a haplotype comprising A*33:01, B*14:02, and C*08:02 may participate in concert to confer risk for DILI, as opposed to A*33:01 alone. However, because these alleles are so highly correlated, our current sample size does not allow us to distinguish between these possible explanations by genetic association evidence alone. This conceivable hypothesis could be further verified in a larger study, or by experiments with recombinant HLA proteins.

In the case of terbinafine where the A*33:01 association showed genome-wide significance for cases relating to this drug only, information on the underlying mechanism for hepatotoxicity is limited, N-dealkylation leads to the formation of an aldehyde metabolite, TBF-A, and this metabolite shows reactivity with glutathione. It has been proposed that the glutathione-adduct is transported across the canalicular membrane and concentrated in the bile, where it can cause damage to biliary epithelial cells. There are limited data from the various case reports on an underlying inflammatory mechanism, but it has been demonstrated that treatment of monocytes with terbinafine results in the release of the proinflammatory cytokines interleukin 8 and tumor necrosis factor—α. Metabolism of terbinafine is complex involving several different cytochromes P450. However, there was no evidence from the GWAS for a role for either CYP genes or innate immunity genes in the terbinafine DILI cases studied.

The other drugs showing the most convincing associations with A*33:01 were fenofibrate, ticlopidine, and sertraline. Failure to see individual genome-wide significant associations with these drugs is likely to be due to fewer cases being available than for terbinafine. The A*33:01 association was seen for 3 of 7 cases due to fenofibrate, all with CM DILI. The literature on fenofibrate DILI is quite limited, but it appears that this drug is extensively metabolized, mainly by CYP3AA4, and there is a report of a drug interaction resulting in cholestatic injury together with other isolated reports of idiosyncratic cholestatic DILI.

Ticlopidine-related DILI has been well studied previously, including 2 studies investigating genetic risk factors in Japanese individuals. Cholestatic liver injury also predominates in this form of DILI. Ticlopidine is subject to extensive metabolism by several cytochrome P450 isoforms and carboxyesterase. A study in rats suggests that adducts are formed after metabolism by cytochrome P450 with evidence for toxicity after biliary excretion of glutathione-conjugated metabolites via MRP2-facilitated transport. In previous studies, 22 Japanese patients with ticlopidine DILI showed an association with an HLA haplotype including A*33:03 (OR, 13). In line with current observations, the association was strongest with cholestatic cases with 12 of 14 cases positive for A*33:03. It should be noted that A*33:03 is relatively common in Japan with approximately 10%–15% of individuals carrying this allele.

The observations on the HLA association for Japanese ticlopidine DILI cases were followed up by a report that those carrying a −2320T>C polymorphism in CYP2B6 were more susceptible to ticlopidine DILI due to high CYP2B6 expression (OR, 2; P = .04). The CYP2B6 polymorphism (rs7254579) is less frequent in Europeans than in Asians.
(minor AF_{\text{cau}} 0.29; minor AF_{\text{Asian}} 0.45) and its low effect size limited our ability to replicate the association in this small European ticlopidine DILI population, but the nonsignificant effect is in the same direction as the previous study (OR, 2.8: 95% CI, 0.7–10.18; P = .11).

In contrast with these 3 drug examples, sertraline is associated predominantly with HC DILI. In line with this phenotypic association, there is evidence that sertraline can cause mitochondrial damage and induce endoplasmic reticulum stress in liver cells. There are parallels with a previous example of an HLA risk factor (DRB1*15:01) that is associated with predominantly CM DILI with amoxicillin-clavulanate, but HC DILI with lumiracoxib.

In line with the Japanese report of a role for A*33:03 in ticlopidine DILI and a case report showing an association between A*33 and tiopronin-induced cholestasis in a Chinese patient, we found 2 DILI cases positive for A*33:03 after direct HLA typing. One of these was a Chinese terbinafine DILI case and the second a European-American with erythromycin DILI. In our imputed GWAS data set, A*33:03 was carried by 8 DILI cases due to a range of drugs in our cohort and showed an apparent risk effect for CM DILI but this was not genome-wide significant (OR, 2.1; 95% CI, 1.02–4.6; P = .04). Another A*33 allele, A*33:05, was also represented in a terbinafine case from Finland. There is very strong homology between these 3 HLA-A*33 alleles at the protein sequence level with A*33:03 differing at only 2 positions from A*33:01 (Tyr instead of His at position 171 and Lys instead of Arg at position 186) and A*33:05 differing at only 1 position (Arg at 54 in place of Gln). In particular, all 3 alleles conserve the key residues for specific peptide binding within the B and F pockets. This is in contrast to a related HLA allele A*31:01, which is associated with carbamazepine-induced skin rash, but does not appear to be a risk factor for DILI, where the B pocket sequence, though homologous, is not conserved.

As mentioned, the association of a common HLA allele with DILI due to chemically unrelated compounds had been observed previously for DRB1*15:01 with amoxicillin-clavulanate and lumiracoxib and for DRB1*07:01 with DILI from lapatinib and ximelagatran. The association of A*33:01 with DILI in general and secondary to a number of structurally dissimilar compounds is consistent with these observations. Together with recent findings from in vitro studies on T-cell responses to fluoroxxacin and amoxicillin-clavulanate, these observations support the hypothesis that either the parent drug or metabolites bind covalently to cellular or circulating proteins to form adducts in a mechanism that is probably slightly different from the direct drug effect seen with hypersensitivity reactions to abacavir. Adduct formation may then allow binding to the peptide binding groove of HLA molecules, leading to activation and differentiation of T cells with a consequent adaptive immune response-mediated liver injury. Evidence that the majority of the drugs showing the A*33:01 association undergo hepatic metabolism and biliary excretion may explain the stronger association of A*33:01 with CM DILI and could indicate that, unlike in the case of fluoroxxacin and amoxicillin-clavulanate, metabolites contribute to the toxicity mechanism. Further investigation of potential interaction of both the various drugs and their metabolites with the A*33:01 gene product by molecular modeling and in vitro studies on T cells as undertaken previously for fluoroxxacin would be of interest.

The novel association of HC DILI with LRBA is interesting because this gene is a biologically plausible DILI candidate. LRBA deficiency due to rare mutations is associated with primary immunodeficiency of variable severity with a particular feature of decreased regulatory T cell levels, other immunodeficiencies, and inherited autoimmune disease. Patients with mutations in LRBA leading to immunodeficiency have been demonstrated to show loss of cytotoxic T lymphocyte antigen–4. Studies in a mouse model suggest that low CTLA4 is a risk factor for DILI. Unlike the HLA-A*33:01 association, no genome-wide significant associations for single drugs were detected with the LRBA SNP, and there were no obvious features in common between cases positive for the variant other than the HC phenotype. It remains possible that this association could be replicated if a larger cohort were available.

The intergenic signal on chromosome 2 is from a region 800 kb upstream from SOX11, is independent of A*33:01 and associated with an almost 2-fold risk of DILI, with the top SNP showing a frequency of 0.02 in Europeans. This risk factor seems to be shared across unrelated drugs, among which ciprofloxacin showed the strongest association. The ENCODE project suggests there are no regulatory elements in this region so the basis for the signal is unclear. Neither rs72631567 nor any of its linkage disequilibrium SNPs (r^2 > 0.5) are known expression quantitative trait loci variants (http://www.gtexportal.org/). The failure to confirm this association and the absence of any apparent biological basis suggests the observed significance could have been a chance finding.

Most data for individual drug classes that were comparatively well represented in our cohort were entirely negative, but the finding of a signal for statins that was driven by several class members was entirely novel. Similar to the more general signal seen on chromosome 2, the chromosome 18 is intergenic with the closest known gene, cadherin 19, located approximately 300,000 bp downstream. Although functionally such a protein could be of relevance to the liver injury process, any biologic significance seems tenuous. The failure to confirm the signal in additional cases could be due to the availability of only a small cohort of additional cases, which reflects the rarity of this form of DILI.

In conclusion, this study has detected a novel HLA association (HLA-A*33:01) in cases of DILI due to a number of different drugs, together with several novel non-HLA signals. Overall sensitivity and specificity of the A*33:01 allele as a predictor of DILI is low, but our findings may be important for future drug treatment in cases of DILI due to one of the drugs for which the A*33:01 association is relevant. Follow-up studies are required to further explore the intergenic signal on chromosome 2, the biologically
interesting signal in LRBA, and the rs116561224 signal for statins in larger cohorts.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of Gastroenterology at www.gastrojournal.org, and at http://dx.doi.org/10.1053/j.gastro.2016.12.016.

References


30. Vickers AE, Sinclair JR, Zollinger M, et al. Multiple cytochrome P-450s involved in the metabolism of


33. Kim MJ, Jeong ES, Park JS, et al. Multiple cytochrome P450 isoforms are involved in the generation of a pharmacologically active thiol metabolite, whereas paraxonase 1 and carboxylesterase 1 catalyze the formation of a thiol metabolite isomer from ticlopidine. Drug Metab Dispos 2014;42:141–152.


42. Kurosaki M, Takagi H, Mori M. HLA-A33/B44/DR6 is highly related to intrahepatic cholestasis induced by tiotprion. Dig Dis Sci 2000;45:1103–1108.

technical, study management, and assistance with recruitment on the iDILIC study; and to Daniele Cusi (Hypergenes), Patrik K. Magnusson (Swedish Twin Registry), and Javier Martin (Spanish DNA bank) for provision of control data. Contributors to sample collection via iDILIC, the Spanish DILI registry, EUDRAGENE, DILIN, DILIGEN and iSAEC are listed in the Supplementary Material.

Drs Nicoletti, Aithal, Urban, and Daly contributed equally to this manuscript.

Conflicts of interest
These authors disclose the following: Dr Nelson is an employee of GlaxoSmithKline. Drs Chalasani, Fontana, and Watkins report consulting agreements and research grants with several pharmaceutical companies but none represent as potential conflicts for this paper. The Drug-Induced Liver Injury Network causality committee considers potential conflicts while assigning cases for adjudication to individual investigators. The remaining authors disclose no conflicts.

Funding
The genome-wide association study, HLA genotyping, and International Drug-Induced Liver Injury Consortium case enrollment and sample collection were funded by the International Serious Adverse Events Consortium with (phase 2) membership support from Abbott, Amgen, Daiichi-Sankyo, GlaxoSmithKline, Merck, Novartis, Pfizer, Roche, Sanofi-Aventis, Takeda, and the Wellcome Trust. This is a summary of independent research party (the DILIGEN and International Drug-Induced Liver Injury Consortium sample collection) funded by the National Institute for Health Research (NIHR) Nottingham Digestive Diseases Biomedical Research Unit at the Nottingham University Hospitals National Health Service (NHS) Trust and University of Nottingham. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health. The Drug-Induced Liver Injury Network (https://dilin.dcri.duke.edu) is supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health (NIH) as a Cooperative Agreement (U01s) under grants: U01-DK065176 (Duke University), U01-DK065201 (University of North Carolina [UNC]), U01-DK065184 (University of Michigan), U01-DK065211 (Indiana University), U01DK065193 (University of Connecticut), U01-DK065238 (University of California-San Francisco/California Pacific Medical Center), U01-DK083023 (University of Texas Southwestern Medical Center), U01-DK083027 (Thomas Jefferson Hospital/University of Pennsylvania), U01-DK082982 (Mayo Clinic), U01-DK083020 (University of Southern California), U01-DK100928 (Icahn School of Medicine). Additional funding is provided by Clinical and Translational Science Award grants UL1 RR025761 (Indiana University), UL1 TR001111 (UNC), and UL1 RR024986 (University of Michigan). The EUDRAGENE collaboration received support from the EC 5th Framework program (QLRI-CT-2002-02757). The Spanish Drug-Induced Liver Injury Registry is partly funded by the Spanish Medicine Agency, Fondo Europeo de Desarrollo Regional--FEDER (P10-CTS-6470, FIS PI12/00378). CIBERehd is funded by Instituto de Salud Carlos III. The Swedish case collection (SWEDEGENE) has received support from the Swedish Medical Products Agency, the Swedish Society of Medicine (2008-21619), Swedish Research Council (Medicine 521-2011-2440), and Swedish Heart and Lung Foundation (20120557). Mariam Molokhia was supported by the NIHR Biomedical Research Centre at Guy’s and St Thomas’ NHS Foundation Trust and King’s College London. Luisa Ibáñez was supported by Instituto de Salud Carlos III (EC08/00250). Marco Arrese was supported by CONICYT PIA/Basal PFB12. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.