HEPATOLOGY

HEPATOLOGY, VOL. 73, NO. 1, 2021

*Human Leukocyte Antigen B*14:01* and *B*35:01* Are Associated With Trimethoprim-Sulfamethoxazole Induced Liver Injury

Yi-Ju Li [1,2] Elizabeth J. Phillips [3] Andrew Dellinger [2] Paola Nicoletti,[4] Ryan Schutte,[5] Danmeng Li,[5] David A. Ostrov,[5] Robert J. Fontana,[6] Paul B. Watkins,[7] Andrew Stolz,[8] Ann K Daly,[9] Guruprasad P Aithal,[10,11] Huiman Barnhart,[1,12] Naga Chalasani,[13] and the Drug-induced Liver Injury Network

BACKGROUND AND AIMS: Trimethoprim (TMP)–sulfamethoxazole (SMX) is an important cause of idiosyncratic drug-induced liver injury (DILI), but its genetic risk factors are not well understood. This study investigated the relationship between variants in the human leukocyte antigen (HLA) class 1 and 2 genes and well-characterized cases of TMP-SMX DILI.

APPROACH AND RESULTS: European American and African American persons with TMP-SMX DILI were compared with respective population controls. HLA sequencing was performed by Illumina MiSeq (Illumina, San Diego, CA) for cases. The HLA genotype imputation with attribute bagging program was used to impute HLA alleles for controls. The allele frequency difference between case patients and controls was tested by Fisher's exact tests for each ethnic group. For European Americans, multivariable logistic regression with Firth penalization was used to test the HLA allelic effect after adjusting for age and the top two principal components. Molecular docking was performed to assess HLA binding with TMP and SMX. The European American subset had 51 case patients and 12,156 controls, whereas the African American subset had 10 case patients and 5,439 controls. Four HLA alleles were significantly associated in the European American subset, with *HLA-B*14:01* ranking at the top (odds ratio, 9.20; 95% confidence interval, 3.16, 22.35; P = 0.0003) after covariate adjustment. All carriers of *HLA-B*14:01* with TMP-SMX DILI possessed *HLA-C*08:02*, another significant allele (P = 0.0026). This pattern was supported by *HLA-B*14:01–HLA-C*08:02* haplotype association ($P = 1.33 \times 10^{-5}$). For the African American patients, *HLA-B*35:01* had 2.8-fold higher frequency in case patients than in controls, with 5 of 10 patients carrying this allele. Molecular docking showed cysteine at position 67 in *HLA-B*14:01* and phenylalanine at position 67 in *HLA-B*35:01* to be the predictive binding sites for SMX metabolites.

CONCLUSIONS: *HLA-B*14:01* is associated with TMP-SMX DILI in European Americans, and *HLA-B*35:01* may be a potential genetic risk factor for African Americans. (Hepatology 2021;73:268-281).

Trimethoprim (TMP)–sulfamethoxazole (SMX) is a fixed-combination synthetic antimicrobial that is commonly used to treat various bacterial infections and as a prophylaxis against opportunistic infections.(1) TMP-SMX use has been associated with life-threatening and presumed immune-mediated idiosyncratic adverse drug reactions such as

Abbreviations: AA, amino acid; AF, allele frequency; CF, carriage frequency; CI, confidence interval; Cys^{67}, cysteine at position 67; dbGaP, Database of Genotypes and Phenotypes; DILI, drug-induced liver injury; DILIN, DILI Network; eMERGE, Electronic Medical Records and Genomics; FDR, false discovery rate; GWAS, genome-wide association study; HIV, human immunodeficiency virus; HLA, human leukocyte antigen; iDILIC, International DILI Consortium; MHC, major histocompatibility complex; NCBI, National Center for Biotechnology Information; OR, odds ratio; PAGE, Population Architecture Using Genomics and Epidemiology; PC, principal component; PDB, Protein Data Bank; Phe^{67}, phenylalanine at position 67; SJS, Stevens-Johnson syndrome; SMX, sulfamethoxazole; SNP, single-nucleotide polymorphism; TEN, toxic epidermal necrolysis; Thr^{163}, threonine at position 163; TMP, trimethoprim; Trp^{97}, tryptophan at position 97.

Received December 9, 2019; accepted March 12, 2020.

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.31258/suppinfo.

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 the following grants (to the Drug-induced Liver Injury Network [https://dilin.dcri.duke.edu/]): U01-DK065176 and U24-DK065176 (Duke University), U01-DK065211 (Indiana University), U01-DK065201 (University of North Carolina), U01-DK065184

Stevens-Johnson syndrome (SJS), toxic epidermal necrolysis (TEN), drug reactions with eosinophilia and systemic symptoms, severe drug-induced liver injury (DILI), and blood dyscrasias.[2] TMP-SMX is among the top five leading causes of DILI in the United States.[3] The risk factors for liver injury associated with TMP-SMX are largely unknown, other than that they are more common among African Americans and those infected with human immunodeficiency virus (HIV).[1,4] The pattern of liver injury associated with TMP-SMX is typically cholestatic or mixed, although severe hepatocellular injury and acute liver failure have been reported as well.[5]

Immunogenetic factors have been increasingly shown to play an important role in the pathogenesis of severe immune-mediated adverse drug reactions, including DILI.[6] Previous studies have identified several genetic variants, especially in the human leukocyte antigen (HLA) region associated with any-cause DILI or liver injury due to specific drugs.[7-10] Variants in the HLA region have also been linked to a few TMP-SMX–induced adverse reactions. A Turkish study found a significantly higher frequency of *HLA-A*30*, *HLA-A*30–HLA-B*13-Cw6* haplotype, and *HLA-B*55* among 42 patients with TMP-SMX–induced fixed drug eruption as compared with 2,378 healthy blood donors.[11] In another study consisting of 43 Thai patients with TMP-SMX–induced SJS and TEN and 91 TMP-SMX–tolerant controls, *HLA-B*15:02*, *HLA-C*06:02*, and *HLA-C*08:01* were

(University of Michigan), U01-DK065193 (University of Connecticut), U01-DK065238 (University of California San Francisco/California Pacific Medical Center), U01-DK083023 (University of Texas-Southwestern), U01-DK083020 (University of Southern California), U01-DK082992 (Mayo Clinic), U01-DK083027 (Thomas Jefferson/Albert Einstein Medical Center, University of Pennsylvania), and U01-DK100928 (Icahn School of Medicine at Mount Sinai). Funding support for the cohort and genotyping of Electronic Medical Records and Genomics samples was provided by the following: Group Health Seattle: U01-AG006781 and U01-HG004438, with genotyping performed at Johns Hopkins University; the Mayo Clinic: U01-HG004599 and U01-HG004424, with genotyping performed at The Broad Institute; Marshfield Clinic: U01-HG004608 and U01-HG004438, with genotyping performed at Johns Hopkins University; Northwestern University: U01-HG04603 and U01-HG004424, with the genotyping performed at The Broad Institute; Vanderbilt University: U01-HG004603 and U01-HG004424, with the genotyping performed at The Broad Institute. Phenotype data collection for the Population Architecture Using Genomics and Epidemiology data set was supported by The Andrea and Charles Bronfman Philanthropies. Funding support for genotyping, which was performed at The Center for Inherited Disease Research, was provided by the NIH (U01-HG007417).

View this article online at wileyonlinelibrary.com.
DOI 10.1002/hep.31258

Potential conflict of interest: Dr. Phillips consults for Biocryst and received royalties from UpToDate. Dr. Nicoletti consults for Chiesi and is employed by Sema4. Dr. Fontana consults for Sanofi and received grants from AbbVie, Gilead, and Bristol-Myers Squibb. Dr. Aithal consults for and advises GlaxoSmithKline and Pfizer. He consults for Fractyl and AstraZeneca. He advises Amryt. Dr. Chalasani consults for and received grants from Galectin. He consults for NuSirt, AbbVie, Afimmune, Allergan, Madrigal, Siemens, Foresite, Zydus, and La Jolla. He received grants from Exact Sciences and Intercept.

ARTICLE INFORMATION:

From the [1]Department of Biostatistics and Bioinformatics, Duke University Medical Center, Duke University, Durham, NC; [2]Molecular Physiology Institute, Duke University Medical Center, Duke University, Durham, NC; [3]Department of Medicine, Vanderbilt University Medical School and School of Medicine Vanderbilt University, Nashville, TN; [4]Department of Genetics and Genomic Science, Icahn School of Medicine at Mount Sinai, Mount Sinai Health System, New York, NY; [5]Department of Pathology, Immunology and Laboratory Medicine, College of Medicine, University of Florida, Gainesville, FL; [6]Department of Internal Medicine, University of Michigan, Ann Arbor, MI; [7]Division of Pharmacotherapy and Experimental Therapeutics, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC; [8]Division of Gastroenterology and Hepatology, Keck School of Medicine, University of Southern California, Los Angeles, CA; [9]Institute of Translational and Clinical Research, Newcastle University, Newcastle Upon Tyne, United Kingdom; [10]Nottingham Digestive Diseases Centre, Nottingham University Hospital National Health Service Trust and University of Nottingham, Nottingham, United Kingdom; [11]National Institute for Health Research Nottingham Biomedical Research Centre, Nottingham University Hospital National Health Service Trust and University of Nottingham, Nottingham, United Kingdom; [12]Clinical Research Institute, Duke University Medical Center, Duke University, Durham, NC; [13]Division of Gastroenterology and Hepatology, School of Medicine, Indiana University, Indianapolis, IN.

ADDRESS CORRESPONDENCE AND REPRINT REQUESTS TO:

Naga Chalasani, M.D.
Division of Gastroenterology and Hepatology
Indiana University School of Medicine
Emerson Hall 305, 545 Barnhill Drive
Indianapolis, IN 46202
E-mail: nchalasa@iu.edu
Tel.: +1-317-278-0414

found to be significantly associated with severe cutaneous adverse reactions.[12] To our knowledge, there have been no studies to date that have attempted to elucidate the genetic basis underlying liver injury caused by TMP-SMX using genome-wide association studies (GWASs) or targeted HLA allelic association analysis.

The main objectives of this study were to investigate the HLA allele association with TMP-SMX induced liver injury in a well-defined cohort of European Americans and African Americans enrolled in the DILI Network (DILIN) Prospective and Retrospective studies.

Participants and Methods

STUDY PARTICIPANTS

Individuals who were ≥2 years of age with suspected DILI and who met predefined eligibility criteria were enrolled into the DILIN Prospective (NCT00345930) and Retrospective (NCT00360646) studies. The designs of these two studies have been published.[13,14] In brief, written informed consent was obtained from individuals with acute liver injury suspected to be due to prescription medications, over-the-counter medications, or herbal and dietary supplements before enrollment at multiple centers throughout the United States. Our study followed the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the institutional review board of each study site. On recruitment, participants underwent systematic evaluation for competing causes, and all cases were adjudicated for causality by structured expert consensus. The causal relationship between suspected medication(s) and the liver injury episode was categorized as definite (>95% likelihood), highly likely (76%-95% likelihood), probable (51%-75% likelihood), possible (26%-50% likelihood), and unlikely (≤25% likelihood).[14] Cases included in this study were adjudicated to have definite, highly likely, or probable DILI due to TMP-SMX, referred to as high-confidence DILI cases. For this study's purposes, instead of self-reported race and ethnicity, we used the genetic ancestry inferred by genome-wide single-nucleotide polymorphism (SNP) data from DILIN and the 1000 Genomes Project[15] (https://www.internationalgenome.org/) (Supporting Information). Of 1,916 patients with DILI, 86 (4.5%) were suspected to have DILI due to TMP-SMX, with 72 patients classified as high-confidence case patients. Based on genetically inferred ethnicity, these patients were then grouped as 51 European Americans, 10 African Americans, 6 Hispanics, 4 Asians, and 1 other (Supporting Information, Section 1; Supporting Table S2). In this study, the 51 European American and 10 African American case patients with DILI were analyzed. Participants in this paper have been included in other publications from the DILIN.[16]

STUDY DESIGN AND HLA DATA

We used a case-control study design to identify HLA alleles that may play a role in the risk of developing TMP-SMX DILI in European Americans and African Americans, respectively. Because of the lack of control data in the DILIN, considering the low incidence rate of DILI in the population, we used large population cohorts as controls to overcome the potential bias of including potential patients with DILI events. With the approval from the National Center for Biotechnology Information (NCBI) Database of Genotypes and Phenotypes (dbGaP) (https://www.ncbi.nlm.nih.gov/gap/), we obtained the GWAS data from Electronic Medical Records and Genomics (eMERGE)-1: Genome-Wide Association Studies of Network Phenotypes (phs000360.v3.p1) and Population Architecture Using Genomics and Epidemiology (PAGE): the Charles Bronfman Institute for Personalized Medicine BioMe BioBank (phs000925.v1.p1). The eMERGE-1 GWAS data set consisted of 17,045 participants who were predominantly Americans with European ancestry, whereas the PAGE GWAS data set consisted of 12,932 participants with diverse racial backgrounds. Genotyping platforms were the Human 660W Quad (Illumina, San Diego, CA) for eMERGE-1 and MEGA (Illumina) for PAGE. Similarly, we used genome-wide SNP data to infer ethnicity for each data set. This led to 12,156 European Americans from eMERGE-1 and 5,439 African Americans from PAGE to serve as controls.

For case patients with TMP-SMX DILI, HLA alleles were determined from the deep sequencing of HLA class 1 and 2 genes using Illumina MiSeq (details in Supporting Information). For population controls, we used the GWAS SNPs to impute HLA alleles in HLA class 1 and 2 genes for European

Americans and African Americans, respectively. Specifically, four-digit HLA alleles were imputed on the basis of the SNPs in the major histocompatibility complex (MHC) region in chromosome 6 using the HLA genotype imputation with attribute bagging (HIBAG) program.[17] All analyses were conducted for European Americans (51 case patients vs. 12,156 controls) and African Americans (10 case patients vs. 5,439 controls), respectively.

STATISTICAL ANALYSIS

Descriptive statistics for demographic and patient characteristics were computed for 51 European American and 10 African American case patients with TMP-SMX DILI, respectively. We presented the frequency (percentage) for categorical variables and the median (interquartile range) for continuous variables. Allele frequencies (AFs) and carriage frequencies (CFs; proportion of patients carrying the allele) of all four-digit HLA alleles were computed for case patients and controls, respectively, for each ethnic group. To evaluate the accuracy of imputed HLA alleles in controls, we compared the AFs of imputed HLA alleles to population AFs available from BeTheMatch (https://bioinformatics.bethematchclinical.org)[18] and Allele Frequency Net Database (http://www.allelefrequencies.net/)[19] for European Americans and African Americans, respectively. We set an arbitrary threshold to exclude imputed HLA alleles that have an absolute AF difference from population AFs greater than 0.05. Because HLA alleles were determined by all variants in the coding sequence of the gene, this reduced variants to one marker with multiple alleles for each HLA gene (e.g., *HLA-A**, *HLA-B**, and so on). Therefore, association tests were performed for a single allele at a time. Considering that rare alleles will have low statistical power in a single-allele association test, we excluded ultra-rare HLA alleles that had an allele count of either <2 in case patients with TMP-SMX DILI or <10 in controls.

For allelic association tests, each four-digit HLA allele was coded as 0, 1, or 2 on the basis of the number of the allele an individual carried. Fisher's exact test was used to test AF differences between case patients and controls for each HLA allele. For the European American subset, multivariable logistic regression with Firth penalization for rare events (Firth logistic regression) was used to test the association between each HLA allele and DILI with covariate adjustment. For covariate selection, univariate Firth logistic regression was performed for age, sex, and principal component (PC)1 to PC10 individually first. We selected the top three variables meeting $P < 0.05$ as covariates. Furthermore, because eMERGE-1 released age data in decades, we categorized age at three levels (1, <40 years; 2, 40-69 years; 3, >69 years). False discovery rate (FDR) by Benjamini-Hochberg[20] was computed to correct multiple testing. Significant HLA alleles were determined on the basis of an FDR < 0.15 from the multivariable Firth logistic regression analysis. To delineate the relationship among the significant HLA alleles, we conducted conditional analyses by adjusting the most significant HLA allele, along with other covariates. For the African American data set, because of having 10 case patients only, we did not pursue multivariable regression analysis. Instead, we conducted Fisher's exact test only and highlighted alleles meeting $P < 0.05$.

Several secondary analyses were performed to understand the top HLA alleles further. First, haplotype association was tested for two-gene and three-gene combinations using the score statistics in the haplo.stats program (Comprehensive R Archive Network, R Foundation, Vienna, Austria).[21] Second, using the NetMHCpan-2.8 algorithm in MHCcluster (http://www.cbs.dtu.dk/services/MHCcluster/; Technical University of Denmark, Lyngby, Denmark),[22] HLA alleles in the gene of interest were clustered on the basis of their predicted peptide-binding specificity. We then tested the association between each allelic cluster and TMP-SMX DILI. That is, HLA alleles in the same cluster were coded as the same marker, and then the same association analyses as described above were followed. Third, we checked whether the HLA allele of interest was present in persons with high-confidence DILI due to non-antibacterial sulfonamides enrolled in the DILIN studies. Fourth, to evaluate whether the HLA allele of interest was specific to the subset with TMP-SMX, we compared CFs of the HLA allele of interest among DILIN drug groups that had at least 10 high-confidence DILI cases. Finally, we tested whether the HLA allele of interest was associated with the any-cause DILI for the corresponding ethnicity.

REPLICATION OF TOP HLA ALLELES

We were able to identify 5 European American patients with TMP-SMX and 3 patients with TMP-induced liver injury enrolled in the International DILI Consortium (iDILIC) to evaluate our findings in DILIN. Causality assessment was performed as reported.[23] The cases were previously analyzed as a part of the broad GWAS on susceptibility to DILI.[24]

AMINO ACID ASSOCIATION ANALYSIS

We first obtained the amino acid (AA) sequence alignment of all available HLA alleles in the same HLA gene from the Immuno Polymorphism Database (https://www.ebi.ac.uk/ipd/imgt/hla/align.html; European Molecular Biology Laboratory–European Bioinformatics Institute, Hinxton, United Kingdom).[25] We then identified a set of polymorphic AA positions within the binding site. Using the HLA allele of interest as the reference allele, we converted the genotypes of the HLA allele to AA residue pairs of all polymorphic AA positions on the basis of whether the reference AA residue was present. In other words, assuming K polymorphic AA positions identified for an HLA gene, HLA genotype data of each participant was recoded as a vector of K markers with 0, 1, or 2 on the basis of the number of reference AA residues present at the AA position. The same association tests, as described above for HLA alleles, were performed for each polymorphic AA position for European Americans and African Americans, respectively.

MOLECULAR DOCKING

Because of the lack of crystal structure of HLA-B*14:01, an atomic model of *HLA-B*14:01* was generated on the basis of the crystal structure of *HLA-B*14:02* (99.6% identical, Protein Data Bank [PDB] code 3BVN)[26] using SWISS-MODEL (Protein Structure Bioinformatics Group, Basel, Switzerland).[27] The *HLA-B*14:01* model and *HLA-B*35:01* crystal structure, PDB code 6BJ8,[28] were used for molecular docking. AutoDock Tools (Scripps Research Institute, La Jolla, CA) was used for molecular docking of TMP and SMX with AutoDock Vina[29] (e.g., add H, generate charges for each atom to be used in scoring). Scoring grids were 20 × 20 × 20 Å and centered on three sites within the antigen-binding clefts of *HLA-B*14:01* and *HLA-B*35:01*, including sites corresponding to the Cα of the first peptide position, Cα of the fifth peptide position (the middle of the antigen-binding cleft), and Cα of the terminal peptide position. TMP and SMX were docked with exhaustiveness set to 10. The top eight scoring orientations were output and compared. Sulfonamide was manually positioned in the modelled antigen-binding cleft of *HLA-B*14:01* with PyMOL (version 1.8; Schrödinger, LLC, New York, NY), which was used to generate molecular graphics.

Results

From September 2004 through January 2019, the DILIN had enrolled a total of 2,256 participants with suspected DILI, and 1,916 of them had undergone causality adjudication and had their HLA genes sequenced (Fig. 1). There were 86 participants with suspected DILI in whom TMP-SMX was the primary implicated agent. On adjudication, 72 were deemed to have high-confidence DILI events (causality scores of definite, highly likely, or probable), and these participants included 51 European Americans and 10 African Americans. Table 1 describes the summary statistics for selected clinical characteristics of European Americans and African Americans, respectively. The median latency was longer in European Americans (25 days) than in African Americans (14.5 days). Although the pattern of liver injury was evenly distributed among hepatocellular, cholestatic, and mixed categories in European Americans, it was predominantly hepatocellular in African Americans. Eleven patients (22%) of the European American subset had peripheral eosinophilia, but this condition did not appear in the African American subset. Although a cutaneous rash was commonly reported (European Americans, 47%; African Americans, 40%), there were no reported instances of SJS in our patients with TMP-SMX DILI.

We first examined the quality of imputed HLA alleles in controls: the eMERGE-1 and PAGE data. We identified four class 2 alleles in European Americans and seven class 2 alleles in African Americans with AFs deviating from population AFs by more than 0.05 (Supporting Table S3). These

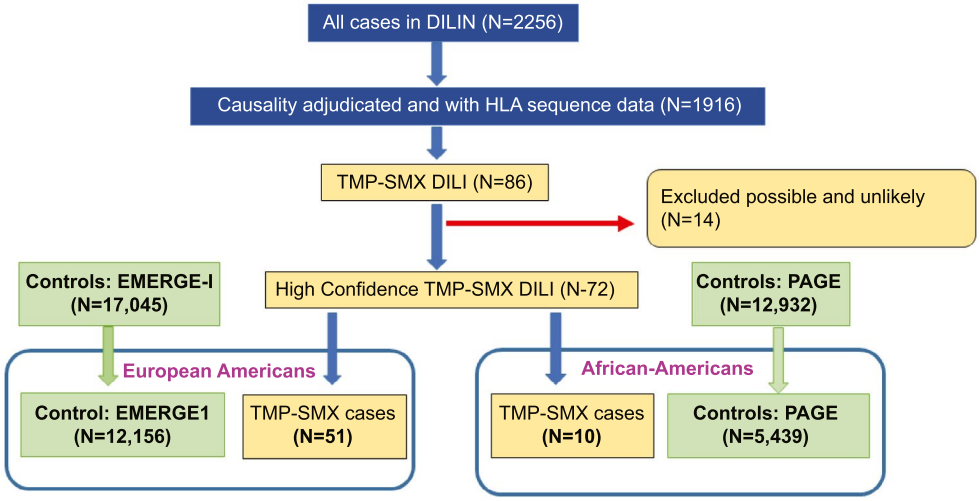

FIG. 1. Consort diagram to outline the sample sizes used in this study. Abbreviations: EMERGE-1, Electronic Medical Records and Genomics1; DILI, drug-induced liver injury; DILIN, DILI Network; HLA, human leukocyte antigen; PAGE, Population Architecture Using Genomics and Epidemiology; SMX, sulfamethoxazole; TMP, trimethoprim.

alleles were excluded from the association analysis. In addition, 114 ultra-rare HLA alleles in European Americans and 130 ultra-rare HLA alleles in African Americans were excluded. This led to 85 alleles in European Americans and 26 alleles in African Americans remaining for interrogation in this study.

HLA ALLELIC ASSOCIATION

Table 2 summarizes the top HLA alleles associated with TMP-SMX DILI for European Americans and African Americans, respectively. In European Americans, four class 1 HLA alleles were considered significant, with *HLA-B*14:01* ranking at the top after adjusting for covariates and multiple testing. The AF of *HLA-B*14:01* in case patients was about 5.5-fold higher than in controls (AF = 0.049 vs. AF = 0.009, Fisher $P = 0.002$). The effect size of *HLA-B*14:01* on TMP-SMX DILI was estimated, with an odds ratio (OR) of 9.20 (95% confidence interval [CI], 3.16, 22.35; adjusted $P = 0.0003$; FDR = 0.024 based on 85 alleles tested). The other three top HLA alleles were *HLA-A*34:02*, *HLA-C*08:02*, and *HLA-B*27:02*. The OR estimates were unstable, with large CIs for both *HLA-A*34:02* (95% CI = 5.17, 320.43) and *HLA-B*27:02* (95% CI = 2.50, 48.04), which is likely due to low AFs in both case patients and controls. On the other hand, *HLA-C*08:02*, a relatively common allele (AF = 0.03 in controls), had the most stable effect-size estimate (narrower CI) among the four top alleles (OR [95% CI], 3.78 [1.66, 7.62]; adjusted $P = 0.0026$). However, once we adjusted for the effect of *HLA-B*14:01*, *HLA-C*08:02* was no longer significant ($P = 0.230$). This was due to the fact that all five patients carrying *HLA-B*14:01* also carried *HLA-C*08:02* (Supporting Table S4).

Given the rarity of the *HLA-B*14:01* allele in the European American general population (population AF = 0.008), we did not expect any carrier in the 5 case patients with TMP-SMX DILI from the iDILIC. Instead, we found that 1 case patient with TMP-SMX DILI carried the *HLA-B*14:01* allele, a higher AF than expected (AF = 0.1). Interestingly, this case patient with *HLA-B*14:01* also carried *HLA-C*08:02*. None of the 5 case patients with TMP-SMX DILI from the iDILIC carried *HLA-A*34:02* or *HLA-B*27:02*. Further, among 3 case patients with TMP DILI in the iDILIC, there were no carriers of *HLA-B*14:01* or *HLA-B*14:02*.

For the African American subset, three HLA alleles showed a nominally significant AF difference (Fisher $P < 0.05$) (Table 2). Among them, *HLA-B*35:01* ranked at the top, with a 2.8-fold higher AF in case patients than in controls (AF = 0.25 vs. AF = 0.087, $P = 0.026$). More specifically, 5 out

TABLE 1. Selected Clinical and Laboratory Characteristics for Patients With TMP-SMX DILI by Ethnicity

	European Americans (N = 51)	African Americans (N = 10)
Age (years), mean (SD)	48 (20)	45 (18)
Women, n (%)	29 (57)	4 (40)
BMI (kg/m^2), mean (SD)	25.8 (6.5)	34.0 (13.5)
Diabetes mellitus, n (%)	7 (14)	5 (50)
Latency (days), median (IQR)	25 (10, 34)	14.5 (9, 20)
Jaundice n (%)	31 (61)	4 (40)
Liver biochemistries, peak values		
ALT (U/L), mean (SD)	1,015 (1,719.7)	798 (671.0)
Alk P (U/L), mean(SD)	481 (294.6)	485 (685.8)
Total bilirubin (mg/dL), mean (SD)	10.2 (10.0)	10.4 (11.1)
INR, mean (SD)	1.4 (0.8)	1.2 (0.25)
Peripheral eosinophilia (>500/μL), n (%)	11 (22)	0 (0)
SJS, n (%)	0 (0)	0 (0)
Pattern of liver injury, n (%)		
Hepatocellular	13 (26)	6 (60)
Cholestatic	18 (35)	3 (30)
Mixed	20 (39)	1 (10)
Severity of Liver Injury, n (%)		
Mild	8 (15.7)	2 (20)
Moderate	15 (29.4)	0 (0)
Moderate-hospitalized	16 (31.4)	7 (70)
Severe	10 (19.6)	1 (10)
Death, n (%)	2 (3.9)	0 (0)
Liver transplantation, n (%)	0 (0)	0 (0)
Chronic DILI, n (%)	8 (15.7)	2 (20)

Abbreviations: Alk P, serum alkaline phosphatase; ALT, serum alanine aminotransferase; BMI, body mass index; DILI, drug-induced liver injury; INR, international normalized ratio; IQR, interquartile range (25%, 75%); SJS, Stevens-Johnson syndrome; SMX, sulfamethoxazole; TMP, trimethoprim.

of 10 case patients with TMP-SMX DILI carried *HLA-B*35:01*. Both *HLA-A*68:02* and *HLA-DPB1*04:01* had four allele counts from 10 African American patients. *HLA-A*68:02* was present in 3 patients with TMP-SMX DILI, with 1 patient carrying homozygous *HLA-A*68:02*, whereas *HLA-DPB1*04:01* was present in 4 patients with heterozygous *HLA-DPB1*04:01*.

HLA-B*14:01 HAPLOTYPE ASSOCIATION

Haplotypes of *HLA-B*14:01* were assessed for the following combinations of HLA class 1 genes: A-B, B-C, and A-C-B in the European American subset. The combinations of *HLA-A*32:01*, *HLA-C*08:02*, and *HLA-B*14:01* were the most promising *HLA-B*14:01* haplotypes to show a significant association with DILI (Supporting Fig. S3). The *HLA-C*08:02–HLA-B*14:01* haplotype (frequency = 0.05, $P = 2.79 \times 10^{-6}$) is more common than the other two haplotypes (A-C and A-C-B combinations, frequency = 0.02) in case patients with TMP-SMX DILI, which can be explained by all 5 carriers of *HLA-B*14:01* carrying *HLA-C*08:02* but only 2 carriers of *HLA-B*14:01* having *HLA-A*32:01* (Supporting Table S4).

HLA ALLELIC CLUSTER ASSOCIATION

Figure 2A,B depicts the HLA allelic clusters by peptide binding for European Americans and African Americans, respectively. For European Americans, the *HLA-B*14:01* and *HLA-B*14:02* cluster was the only cluster that showed a significant association with TMP-SMX DILI (adjusted $P = 0.0028$). In African Americans, the cluster of *HLA-B*35:01*, *HLA-B*53:01*, and *HLA-B*51:01* had a significantly higher cluster frequency in case patients than in controls (frequency = 0.55 vs. frequency = 0.23, Fisher $P = 0.002$) (Table 3).

AA ASSOCIATION

In the European American subset, a total of 45 AA positions in the binding sites of HLA-B were aligned, whereas 28 AA positions were polymorphic. *HLA-B*14:01* and *HLA-B*14:02* had identical AA residues in all sites except position 11, which is located at the carboxy terminal end of the first β strand on the floor of the antigen-binding cleft, forming contact with the α1 helix (not oriented toward the peptide). Among all polymorphic AA residues in *HLA-B*14:01*, AA residues at positions 67 (cysteine at position 67 [Cys^{67}]), 97 (tryptophan at position 97 [Trp^{97}]), and 163 (threonine at position 163 [Thr^{163}]) showed significant association with DILI (Table 4). Trp^{97} was present only in *HLA-B*14:01* and *HLA-B*14:02*, with an estimated effect size of the OR of 3.8 (95% CI, 1.7, 7.6) (adjusted $P = 0.003$). Cys^{67} was present in *HLA-B*14:01*, *HLA-B*14:02*, and nine other HLA-B* alleles, with an OR (95% CI) of 2.2 (1.2, 3.7) (adjusted $P = 0.008$), whereas Thr^{163} was present in *HLA-B*14:01*, *HLA-B*14:02*

TABLE 2. HLA Alleles Associated With TMP-SMX DILI by Ethnic Group

		Fisher's Exact Test			Firth Logistic Regression*			Conditional Analysis on B*14:01*	
Ethnicity	Allele	Case Patient AF	Control AF	Fisher *P*	OR (95% CI)	Adjusted *P*	FDR†	OR (95% CI)	Adjusted *P*
European American	B*14:01	0.049	0.009	0.002	9.20 (3.16, 22.35)	0.0003	0.024	—	—
	A*34:02	0.020	0.001	0.002	47.52 (5.17, 320.43)	0.0012	0.050	27.25 (3.85, 150.33)	0.002
	C*08:02	0.078	0.032	0.016	3.78 (1.66, 7.62)	0.0026	0.074	2.11(0.57, 5.56)	0.230
	B*27:02	0.020	0.002	0.024	13.53 (2.50, 48.04)	0.0056	0.119	14.94 (2.74, 53.53)	0.004
African American	B*35:01	0.25	0.09	0.026	—	—	—	—	—
	DPB1*04:01	0.20	0.06	0.034	—	—	—	—	—
	A*68:02	0.20	0.05	0.038	—	—	—	—	—

*Firth logistic regression adjusted for age, PC2, and PC3; conditional analysis: Firth logistic regression adjusted for age, PC2, PC3, and *HLA-B*14:01*.
†FDR by Benjamini-Hochberg method to correct multiple testing of 85 alleles.
Abbreviations: AF, allele frequency; CI, confidence interval; DILI, drug-induced liver injury; FDR, false discover rate; HLA, human leukocyte antigen; OR, odds ratio; PC, principal component; SMX, sulfamethoxazole; TMP, trimethoprim.

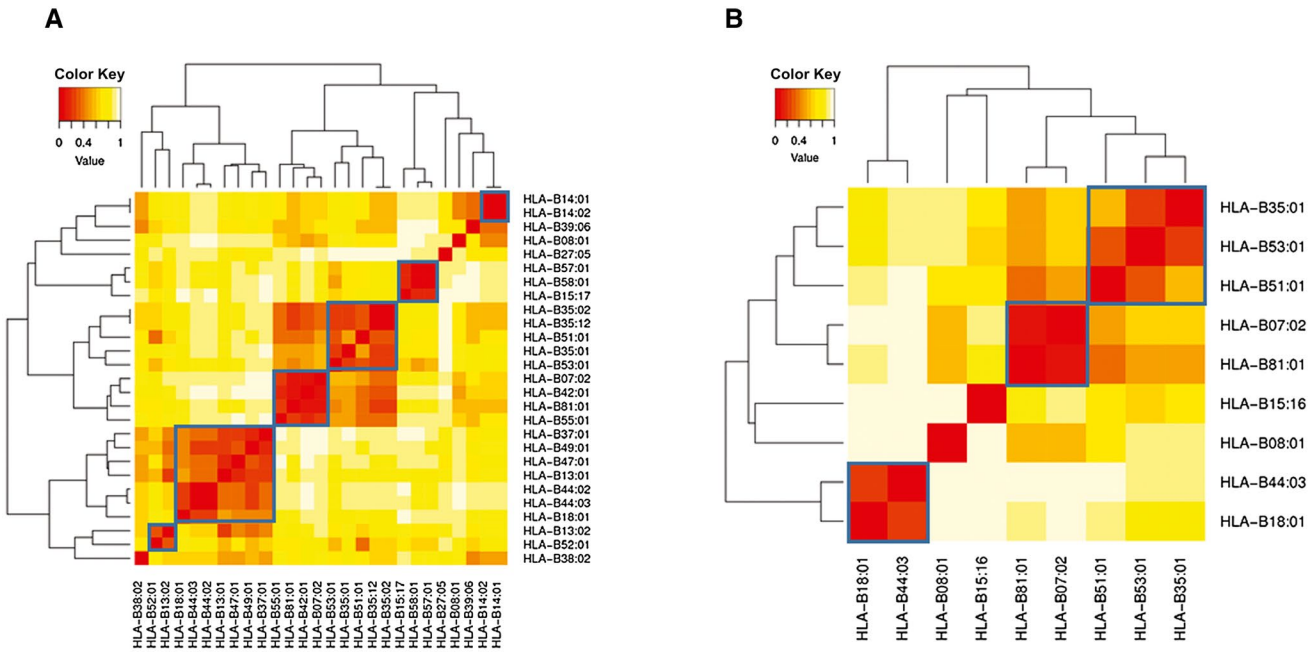

FIG. 2. HLA-B allele cluster from predicted peptide-binding specificity for (A) European Americans and (B) African Americans. The color key was based on the distance between allele binding specificity. The green color (near 0) indicates high similarity on binding specificity between alleles, and white (near 1) indicates completely different binding specificity. Allele clusters tested for association with TMP-SMX DILI are marked by rectangles. Abbreviations: HLA-B, human leukocyte antigen B; SMX, sulfamethoxazole; TMP, trimethoprim.

and 15 other B* alleles, with an OR (95% CI) of 1.9 (1.2, 3.1) (adjusted $P = 0.007$).

In the African American subset, among the same 45 AA residues in the binding sites, there were 26 polymorphic AA residues in the available HLA-B alleles. The AA residues at positions 63 (asparagine at position 63 [Asn^{63}]) and 67 (phenylalanine at position 67 [Phe^{67}]) in *HLA-B*35:01* reached the nominal significance threshold (Table 4). Phe^{67}, present in *HLA-B*35:01*, *HLA-B*08:01*, and *HLA-B*51:01*,

TABLE 3. Association Results Between HLA-B Allelic Cluster and TMP-SMX DILI by Ethnic Group

Ethnicity	Alleles in the Cluster	Cluster Freq: Case Patients	Cluster Freq: Controls	Fisher *P**	Adjusted *P**
European American	B*14:01, B*14:02	0.078	0.032	0.018	0.0028
	B*57:01, B*58:01, B*15:17	0.049	0.046	0.810	0.799
	B*35:02, B*35:12, B*51:01, B*35:01, B*53:01	0.115	0.115	1	0.604
	B*07:02, B*42:01, B*81:01, B*55:01	0.137	0.158	0.683	0.891
	B*37:01, B*49:01, B*47:01, B*13:01, B*44:02, B*44:03, B*18:01	0.245	0.217	0.472	0.537
	B*13:02, B*52:01	0.020	0.029	1	0.874
African American	B*35:01, B*53:01, B*51:01	**0.55**	**0.23**	**0.002**	—
	B*07:02, B*81:01	0.1	0.1	1	—
	B*44:03, B*18:01	0.15	0.08	0.224	—

For the African American group, only Fisher's exact test was performed.
*Fisher *P*: *P* values from Fisher's exact tests; adjusted *P*: *P* values from Firth logistic regression with covariate adjustment of age, PC2, and PC3.
Abbreviations: DILI, drug-induced liver injury; Freq, frequency; HLA-B, human leukocyte antigen B; PC, principal component; SMX, sulfamethoxazole; TMP, trimethoprim.

TABLE 4. AAs Showing Significant Association With TMP-SMX DILI by Ethnic Group

Ethnicity/ Reference HLA Allele	AA Position (Residue)	B* Alleles Sharing the Same Residue	Fisher's Exact Test			Firth Logistic Regression*	
			Case Patient AA Freq	Control AA Freq	*P*	OR (95% CI)	Adjusted *P*
European American/ HLA-B*14:01	67 (Cys)	B*14:01, B*14:02, B*15:10, B*15:18, B*27:02, B*27:05, B*38:01, B*39:01, B*39:06, B*39:24, B*73:01	0.18	0.11	0.057	2.2 (1.2, 3.7)	0.008
	97 (Trp)	B*14:01, B*14:02	0.08	0.03	0.02	3.8 (1.7,7.6)	0.003
	163 (Thr)	B*14:01, B*14:02, B*08:01, B*18:01, B*18:02, B*18:13, B*18:18, B*37:01, B*38:01, B*39:01, B*39:06, B*39:24, B*41:01, B*41:02, B*42:01, B*42:02, B*55:01	0.41	0.27	0.002	1.9 (1.2,3.1)	0.007
African American/ HLA-B*35:01	63 (Asn)	B*14:01, B*14:02, B*07:02, B*08:01, B*15:10, B*18:01, B*38:01, B*39:01, B*39:03, B*39:06, B*39:10, B*42:01, B*42:02, B*51:01	0.85	0.58	0.02	—	—
	67 (Phe)	B*08:01, B*51:01, B*53:01, B*78:01	0.55	0.25	0.007	—	—

*Firth logistic regression included covariate adjustment of age, PC2, and PC3.
Abbreviations: AA, amino acid; Asn, asparagine; CI, confidence interval; Cys, cysteine; Freq, frequency; HLA, human leukocyte antigen; OR, odds ratio; PC, principal component; Phe, phenylalanine; Thr, threonine; Trp, tryptophan.

ranked at the top, with a 2-fold higher frequency in case patients than in controls (frequency = 0.55 vs. frequency = 0.25, $P = 0.007$).

MOLECULAR DOCKING OF SMX OR TMP WITH B*14:01 AND B*35:01

For *HLA-B*14:01*, molecular docking predictions of non-covalent interactions (H bonds and van der Waal contacts) showed that neither SMX nor TMP was predicted to bind with high affinity to the antigen-binding cleft of *HLA-B*14:01*. The estimated DG values (Gibbs free energy) were −6.7 kcal/mol for SMX and −6.4 kcal/mol for TMP (less than −10 kcal/mol suggests high-affinity binding, −6.5 to −9.0 kcal/mol suggests moderate-affinity binding, and greater than −6.5 kcal/mol suggests low-affinity binding). However, drug

metabolites and peptides may form intermolecular contacts with HLA-B*14:01 by covalent interactions. It is notable that *HLA-B*14:01* exhibited a chemically reactive side chain located at a site in the antigen-binding cleft critical for peptide binding: a Cys^{67}, the residue detected in our AA analysis. As the cysteine sulfhydryl group is often in a reactive thiolate form in cells, free unpaired cysteine residues have the potential to form covalent bonds with reactive groups (e.g., with peptides that contain cysteine, or other molecules with reactive groups). SMX is known to be metabolized into reactive forms, such as sulfonamide, that have the potential to bind Cys^{67} in the antigen-binding cleft of *HLA-B*14:01*. These covalent bonds may partially explain the association between TMP-SMX and *HLA-B*14:01*, as metabolite binding at this site would result in altered peptide presentation to T cells. On the basis of the location of the reactive group in sulfonamide and the free sulfhydryl of Cys^{67} in the *HLA-B*14:01* model, SMX was accommodated in a site composed of residues associated with TMP-SMX DILI (Cys^{67}, Trp^{97}, and Thr^{163}; Fig. 3A,B). These data suggested functional interactions between HLA antigen-binding clefts and reactive SMX metabolites.

For *HLA-B*35:01*, molecular docking predicted SMX to bind better than TMP (estimated G = −7.6 kcal/mol for SMX vs. estimated G = −6.9 kcal/mol for TMP) (Fig. 3C,D). SMX was predicted to form hydrogen-bond interactions with

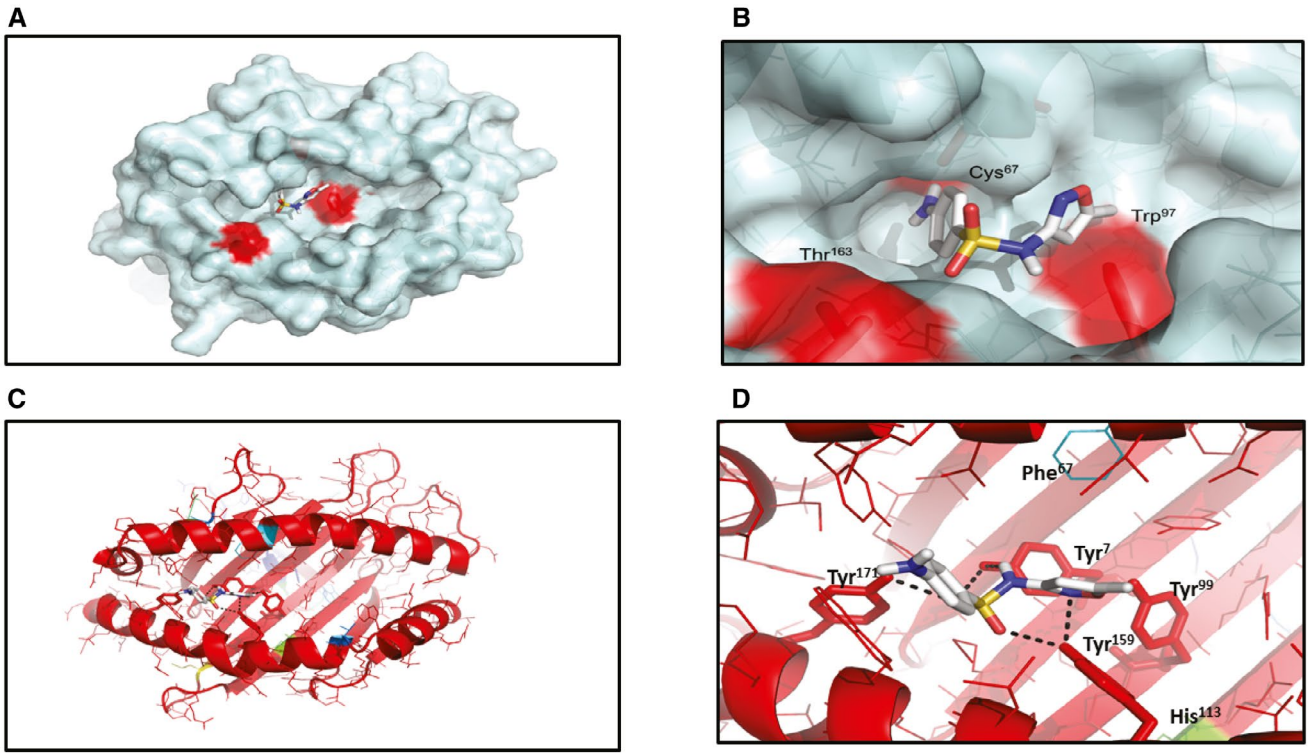

FIG. 3. Results of molecular docking. (A,B) *HLA-B*14:01* exhibited an unpaired Cys^{67} with the potential to bind reactive SMX metabolite sulfonamide. The molecular surface of a model of *HLA-B*14:01* is shown in light blue. Sulfonamides are shown as sticks: white for carbon, blue for nitrogen, and red for oxygen. Cys^{67}, Trp^{97}, and Thr^{163} are shown in red. (C,D) Molecular docking predicted interaction between SMX and the antigen-binding cleft of *HLA-B*35:01*. The crystal structure of *HLA-B*35:01* is shown (PDB code 6BJ8), colored on the basis of the sequence similarity to *HLA-B*15:02*. Blosum62 similarity values are as follows: blue, 40-50; cyan, 50-60; green, 60-70; yellow, 70-80; orange, 80-90; and red, 90-100. Molecular docking (AutoDock Vina, ΔG = −7.6 kcal/mol) predicted SMX interactions with residues shared by *HLA-B*35:01* and *HLA-B*15:02* are shown as sticks. Predicted H-bond interactions are shown as black, dashed lines. The position 67 (Phe^{67}), which we detected in our AA analysis, was not predicted to form contact with SMX by molecular docking. Abbreviations: AA, amino acid; Blosum62, Blocks Substitution Matrix 62; Cys^{67}, cysteine at position 67; His^{113}, histidine at position 113; HLA-B, human leukocyte antigen B; PDB, Protein Data Bank; Phe^{67}, phenylalanine at position 67; SMX, sulfamethoxazole; Thr^{163}, threonine at position 163; Trp^{97}, tryptophan at position 97; Tyr^{171}, tyrosine at position 171.

residues in the antigen-binding clefts of *HLA-B*35:01* that are identical in *HLA-B*15:02*, a previously identified HLA allele associated with TMP-SMX–induced severe cutaneous adverse drug reactions, such as SJS/TEN in a Thai population.[12]

HLA-B*14:01 AND HLA-B*35:01 IN PATIENTS WITH DILI BECAUSE OF OTHER AGENTS

In the European American subset, among 20 drugs with at least 10 DILI cases, only six drug subsets, including TMP-SMX, have *HLA-B*14:01* present (Supporting Table S5). *HLA-B*14:01* was the most prevalent in TMP-SMX, with 5 of 51 patients carrying *HLA-B*14:01* (CF = 0.098). Isoniazid and minocycline ranked second, with 2 of 32 patients carrying *HLA-B*14:01* (CF = 0.063). We did not find *HLA-B*14:01* in 7 patients with DILI because of non-antibacterial sulfonamides (3 patients in the celecoxib subset, 3 patients in the sulfasalazine subset, and 1 patient in the zonisamide subset of patients with DILI) (Supporting Table S6). We further tested whether *HLA-B*14:01* is associated with DILI regardless of the causal drugs by performing multivariable logistic regression with age, sex, and PC1 through PC10 adjusted. On the basis of 903 case patients with high-confidence DILI and 12,156 controls, the association between *HLA-B*14:01* and DILI did not reach the significance level (OR [95% CI], 1.58 [0.96, 2.50]; $P = 0.073$).

For *HLA-B*35:01* in the African American subset, three drug subsets, including TMP-SMX, had at least 10 DILI cases. *HLA-B*35:01* was the most prevalent in the TMP-SMX cases, with 5 of 10 patients carrying *HLA-B*35:01* (CF = 0.5). The other two drugs were isoniazid, with 2 of 19 patients with DILI carrying *HLA-B*35:01* (CF = 0.105), and amoxicillin with clavulanic acid, with 1 of 10 patients carrying *HLA-B*35:01* (CF = 0.1). Among 2 African American patients with DILI caused by darunavir, a non-antibacterial sulfonamide, 1 patient carried *HLA-B*35:01* (Supporting Table S6). We also tested whether *HLA-B*35:01* was associated with overall DILI, regardless of the causal drugs, in 124 case patients and 5,439 controls by adjusting age and PC1; *HLA-B*35:01* did not have a significant effect on overall DILI (OR [95% CI], 1.06 [0.35, 2.42]; $P > 0.99$).

PATIENT CHARACTERISTICS FOR CARRIERS AND NONCARRIERS OF HLA-B*14:01 AND HLA-B*35:01

We observed a slightly longer latency of liver injury (median, 29 vs. 21) and higher proportion of peripheral eosinophilia (40% vs. 16%) in patients with TMP-SMX DILI carrying *HLA-B*14:01* than in those without the allele, but the differences were not statistically significant. No significant differences were found for liver biochemistries at onset and at peak, pattern of liver injury, recovery of liver biochemistries, or severity and outcomes, either (Supporting Table S7). Similarly, the characteristics and the outcomes of African Americans with and without *HLA-B*35:01* were not significantly different from each other.

Discussion

Our study found that *HLA-B*14:01* was associated with TMP-SMX–related DILI in European Americans, with the risk being 9.2-fold higher among individuals carrying the allele than in those without the allele. The *HLA-B*14:01* allele was about 5.5-fold more frequent in European American patients with TMP-SMX DILI than in controls. This large frequency difference was also present in 1 of 5 validation case patients with TMP-SMX DILI carrying *HLA-B*14:01* from the iDILIC cohort. For the African American subset, *HLA-B*35:01* ranked at the top as a potential risk factor for TMP-SMX DILI. Although only 10 African American patients with TMP-SMX DILI were available in our data set, half of them carried *HLA-B*35:01*, a much higher AF and CF than in the control group. Although replication studies with a larger sample size are needed, our results are promising and consistent with previous genetic studies of drug-specific DILI that identified HLA class 1 and 2 genes[8,30,31] as genetic risk factors for DILI.

In the European American subset, *HLA-A*34:02*, *HLA-B*27:02*, and *HLA-C*08:02* were the other top alleles that showed a significant association with TMP-SMX DILI. However, we considered *HLA-B*14:01* as the most promising risk factor for the following reasons. First, *HLA-B*14:01* remained significant after correcting for multiple testing. Second,

both *HLA-A*34:02* and *HLA-B*27:02* were present in only 2 patients with TMP-SMX DILI. Together with their low frequency in controls, the estimated effect sizes (ORs) were unstable, as shown by the large 95% CIs. Therefore, further replication is definitely needed. Third, although *HLA-C*08:02* is the most common allele among these four top alleles, *HLA-C*08:02* is known to be in strong linkage disequilibrium with B*14:01/B*14:02.[32] In our data, all 8 carriers of *HLA-C*08:02* carried either *HLA-B*14:01* (5 patients) or *HLA-B*14:02* (3 patients). Because *HLA-B*14:02* did not show association with TMP-SMX DILI, the association signal of *HLA-C*08:02* is more likely dependent on *HLA-B*14:01*. As we demonstrated in the conditional analysis, after adjusting for the effect of *HLA-B*14:01*, the significant effect of *HLA-C*08:02* on TMP-SMX DILI was diminished.

*HLA-B*14:01* and *HLA-B*14:02* are highly correlated in terms of their predicted peptide-binding specificity. Despite this similarity, *HLA-B*14:01* has a much lower frequency than *HLA-B*14:02* in the general population (e.g., AF = 0.008 vs. AF = 0.03 in European Americans). Interestingly, within the subset of patients with TMP-SMX DILI, we observed an opposite phenomenon, in which *HLA-B*14:01* had a higher frequency than *HLA-B*14:02* (AF = 0.049 vs AF = 0.029). Further, the *HLA-B*14:02* frequency was similar between case patients and controls. This implied that the effect of *HLA-B*14:01* was independent of *HLA-B*14:02*. To date, we have not found any reports linking *HLA-B*14:01* to DILI in general or to drug-specific cases. Among all causal drugs in the DILIN cohort, *HLA-B*14:01* was more prevalent in the subset with TMP-SMX than in other drug-specific subsets or the subset with DILI caused by non-antibacterial sulfonamide-like drugs. Therefore, if the effect of *HLA-B*14:01* on DILI risk we observed in this study is correct, it is specific to TMP-SMX only (and likely to SMX).

In the literature, the *HLA-B*14:01* allele has been linked to protection against HIV progression and with lower plasma viral levels in patients with HIV.[32,33] Different AA residues at a single AA position can also a play role (protective or risk-associated) in disease outcomes, as illustrated in HIV-related studies.[33] In the current study, the main identified risk allele, *HLA-B*14:01*, has an unpaired Cys^{67}. This provides clues not only to a shared epitope with other HLA alleles that could in theory be implicated in risk of DILI but also to the structure of the reactive metabolites and specific groups of SMX, such as sulfonamide RS(O)NR'_2, that might be expected to covalently bind with host proteins.

In the smaller group of African American persons with TMP-SMX–related DILI, *HLA-B*35:01* appeared to be an important HLA risk allele. *HLA-B*35:01* was reported to be associated with DILI caused by *Polygonum multiflorum*, a Chinese herbal medicine,[34] in Chinese patients. However, the relationship between *HLA-B*35:01* and DILI with other drugs has not been reported. *HLA-B*35:01* is known to share peptide-binding specificities with *HLA-B*51:01* and *HLA-B*53:01*. Our allelic cluster analysis also showed that the cluster of these three alleles had stronger association with TMP-SMX DILI than *HLA-B*35:01* alone. The two significant AA residues in positions 63 (Asn^{63}) and 67 (Phe^{67}) identified are also present in *HLA-B*51:01* and *HLA-B*53:01*. Therefore, it is likely that these three alleles are equally important, but this conclusion deserves further study in large cohorts of African Americans with DILI. On the other hand, molecular docking led us to *HLA-B*15:01*, which shares peptide-binding preferences nearly identical to those of *HLA-B*35:01*. Because *HLA-B*15:01* was reported to associate with SJS/TEN due to TMP-SMX in the Thai population, the similarity in the antigen-binding cleft between these two alleles offers potential clues to explain the role of *HLA-B*35:01* in TMP-SMX DILI in African Americans.

Despite these comprehensive analyses suggesting *HLA-B*14:01* and *HLA-B*35:01* as risk factors for TMP-SMX–induced DILI in European Americans and African Americans, respectively, the limitations in this study need to be considered. First, the numbers of cases were limited, a common problem in most drug-specific genetic association studies. Using our European American subset for a post hoc power calculation, given 51 case patients and 12,156 controls and assuming AF of the risk allele at 0.009 (the same AF as HLA-B*14:01) and complete linkage disequilibrium between the risk allele and the testing allele, we had 80% power to detect genetic relative risk at 5.9 for an allele at the same AF. This calculation implies that our sample size was sufficient to detect *HLA-B*14:01* but will require more samples to detect other HLA risk alleles with smaller effect sizes. In addition, because of this sample size limitation, several approaches were used to enhance the analysis quality,

including using large population control data sets for comparison, Fisher's exact tests to accommodate the low counts, and logistic regression models with Firth penalization to reduce the potential bias of parameter estimates. With these strategies, we expected a reduced chance of biased findings. Second, although HLA genotype data in patients with DILI were obtained from advanced sequencing techniques, the HLA data in population controls (eMERGE-1 and PAGE) were obtained from imputation of SNPs in the MHC region. This may raise concerns if HLA calls are inaccurate in the control group. The imputation quality by various programs has been compared before. Karnes et al.[35] reported a 97.6% concordance rate for European Americans and a 92.9% concordance rate for African Americans using HIBAG. Particularly, alleles in class 1 genes had concordance rates over 98%. Therefore, with the large sample size of the control groups used in this study, the impact of minor imputation errors on the result should be small, particularly for the class 1 genes. Third, binding interactions between SMX and TMP were predicted, rather than being experimentally determined. Moreover, predictions of SMX and TMP contact with *HLA-B*14:01* were based on an atomic model (as opposed to a solved crystal structure). However, *HLA-B*14:01* is nearly identical to the crystal structure of HLA-*B*14:02*, with only one difference in α1 and α2 domains, residues 1-180, that form the antigen-binding cleft: serine at position 11 versus alanine at position 11. Because the side chain at this position is not oriented toward peptide ligands (in class 1 HLA molecules, the side chain at position 11 is oriented toward the α1 helix), this polymorphic difference between *HLA-B*14:01* and *–B*14:02* is unlikely to influence reduce confidence in modelling the antigen-binding cleft of *HLA-B*14:01*. Overall, the results from our molecular docking may help us to understand interactions with SMX and its metabolites.

Given the rarity of liver injury due to TMP-SMZ and the relatively small effect size for the associated HLA alleles, our observations are not necessarily immediately clinically actionable from the perspective of pre–prescription screening to avoid a clinical event. However, our observations may help to both inform the immunopathogenesis of DILI related to TMP-SMX and also help in the risk stratification and early diagnosis of such patients. Indeed, in the future when integrated health care systems (e.g., Geisenger Medical Group) undertake whole-exome or whole-genome sequencing of all their patients, our genetic association findings may assist in the thorough evaluation of risk-benefit ratios and help promote safe prescription practices.

Acknowledgment:: We thank all patients and study coordinators who participated in the DILIN study. We also thank the NCBI/dbGaP (http://www.ncbi.nlm.nih.gov/gap) for approving our application to obtain control data sets. The data of control participants without DILI were from eMERGE-1 and PAGE GWAS studies from dbGaP. The eMERGE network is a consortium of five participating sites (Group Health Seattle, Marshfield Clinic, Mayo Clinic, Northwestern University, and Vanderbilt University) funded by the National Human Genome Research Institute. The eMERGE GWAS data sets were accessed using dbGaP accession number phs000360.v3.p. The PAGE data set was from the Charles Bronfman Institute for Personalized Medicine BioMe BioBank at the Icahn School of Medicine at Mount Sinai (New York). The PAGE GWAS data were accessed using dbGaP accession number phs000925.v1.p1.

Author Contributions: Study concept and design: Y.J.L., E.P., H.B., N.C.; acquisition of data: Y.J.L., E.P., P.N., A.K.D., G.P.A., H.B., N.C.; analysis and interpretation of data: Y.J.L., E.P., A.D., D.O., R.S., D.L., N.C.; drafting of the manuscript: Y.J.L., E.P., N.C.; critical revision of the manuscript for important intellectual content: Y.J.L., E.P., A.D., P.N., D.O., R.I.F., P.B.W., A.S., A.K.D., H.B., N.C.; statistical analysis: Y.J.L., A.D., P.N.; obtained funding: R.I.F., P.B.W., A.S., H.B., N.C.

REFERENCES

1) Sulfamethoxazole-Trimethoprim. LiverTox: Clinical and Research Information on Drug-Induced Liver Injury. Bethesda, MD: National Institute of Diabetes and Digestive and Kidney Diseases; 2012. https://www.accessdata.fda.gov/drugsatfda_docs/label/2013/017377s068s073lbl.pdf. Accessed April 2020.
2) U.S. Food & Drug Admission. FDA approved drug: Bactrim. https://www.accessdata.fda.gov/drugsatfda_docs/label/2013/017377s068s073lbl.pdf. Accessed April 2020.
3) Chalasani N, Bonkovsky HL, Fontana R, Lee W, Stolz A, Talwalkar J, et al.; U.S. Drug Induced Liver Injury Network. Features and outcomes of 899 patients with drug-induced liver injury: the DILIN prospective study. Gastroenterology 2015;148:1340-1352.e7.
4) Chalasani N, Reddy KRK, Fontana RJ, Barnhart H, Gu J, Hayashi PH, et al. Idiosyncratic drug induced liver injury in African-Americans is associated with greater morbidity

and mortality compared to Caucasians. Am J Gastroenterol 2017;112:1382-1388.
5) Björnsson E, Jerlstad P, Bergqvist A, Olsson R. Fulminant drug-induced hepatic failure leading to death or liver transplantation in Sweden. Scand J Gastroenterol 2005;40:1095-1101.
6) Pavlos R, Mallal S, Phillips E. HLA and pharmacogenetics of drug hypersensitivity. Pharmacogenomics 2012;13:1285-1306.
7) Urban TJ, Daly AK, Aithal GP. Genetic basis of drug-induced liver injury: present and future. Semin Liver Dis 2014;34:123-133.
8) Urban TJ, Nicoletti P, Chalasani N, Serrano J, Stolz A, Daly AK, et al.; Drug-Induced Liver Injury Network (DILIN); Pharmacogenetics of Drug-Induced Liver Injury group (DILIGEN); International Serious Adverse Events Consortium (iSAEC). Minocycline hepatotoxicity: Clinical characterization and identification of HLA-B*35:02 as a risk factor. J Hepatol 2017;67:137-144.
9) Nicoletti P, Werk AN, Sawle A, Shen Y, Urban TJ, Coulthard SA, et al.; International Drug-induced Liver Injury Consortium. HLA-DRB1*16: 01-DQB1*05: 02 is a novel genetic risk factor for flupirtine-induced liver injury. Pharmacogenet Genomics 2016;26:218-224.
10) **Nicoletti P**, **Aithal GP**, Bjornsson ES, Andrade RJ, Sawle A, Arrese M, et al.; International Drug-Induced Liver Injury Consortium; Drug-Induced Liver Injury Network Investigators; International Serious Adverse Events Consortium. Association of liver injury from specific drugs, or groups of drugs, with polymorphisms in HLA and other genes in a genome-wide association study. Gastroenterology 2017;152:1078-1089.
11) Özkaya-Bayazit E, Akar U. Fixed drug eruption induced by trimethoprim-sulfamethoxazole: evidence for a link to HLA-A30 B13 Cw6 haplotype. J Am Acad Dermatol 2001;45:712-717.
12) Kongpan T, Mahasirimongkol S, Konyoung P, Kanjanawart S, Chumworathayi P, Wichukchinda N, et al. Candidate HLA genes for prediction of co-trimoxazole-induced severe cutaneous reactions. Pharmacogenet Genomics 2015;25:402-411.
13) Rochon J, Protiva P, Seeff LB, Fontana RJ, Liangpunsakul S, Watkins PB, et al. Reliability of the Roussel Uclaf Causality Assessment Method for assessing causality in drug-induced liver injury. Hepatology 2008;48:1175-1183.
14) Fontana RJ, Watkins PB, Bonkovsky HL, Chalasani N, Davern T, Serrano J, et al. Drug-Induced Liver Injury Network (DILIN) prospective study: rationale, design and conduct. Drug Saf 2009;32:55-68.
15) Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al.; 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 2015;526:68-74.
16) Chalasani N, Fontana RJ, Bonkovsky HL, Watkins PB, Davern T, Serrano J, et al. Causes, clinical features, and outcomes from a prospective study of drug-induced liver injury in the United States. Gastroenterology 2008;135:1924-1934, 1934.e1-4.
17) Zheng X, Shen J, Cox C, Wakefield JC, Ehm MG, Nelson MR, et al. HIBAG–HLA genotype imputation with attribute bagging. Pharmacogenomics J 2014;14:192-200.
18) Maiers M, Gragert L, Klitz W. High-resolution HLA alleles and haplotypes in the United States population. Hum Immunol 2007;68:779-788.
19) **Gonzalez-Galarza FF**, **Takeshita LY**, Santos EJ, Kempson F, Maia MH, da Silva AL, et al. update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. Nucleic Acids Res 2015;43:D784-D788.
20) Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 1995;57:289-300.
21) Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002;70:425-434.
22) Hoof I, Peters B, Sidney J, Pedersen LE, Sette A, Lund O, et al. NetMHCpan, a method for MHC class I binding prediction beyond humans. Immunogenetics 2009;61:1-13.
23) Aithal GP, Watkins PB, Andrade RJ, Larrey D, Molokhia M, Takikawa H, et al. Case definition and phenotype standardization in drug-induced liver injury. Clin Pharmacol Ther 2011;89:806-815.
24) **Cirulli ET**, **Nicoletti P**, Abramson K, Andrade RJ, Bjornsson ES, Chalasani N, et al.; Drug-Induced Liver Injury Network (DILIN) investigators; International DILI consortium (iDILIC). A missense variant in PTPN22 is a risk factor for drug-induced liver injury. Gastroenterology 2019;156:1707-1716.e2.
25) Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG. The IPD and IMGT/HLA database: allele variant databases. Nucleic Acids Res 2015;43:D423-D431.
26) Kumar P, Vahedi-Faridi A, Saenger W, Merino E, Lopez de Castro JA, Uchanska-Ziegler B, et al. Structural basis for T cell alloreactivity among three HLA-B14 and HLA-B27 antigens. J Biol Chem 2009;284:29784-29797.
27) Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure homology modeling using SWISS-MODEL workspace. Nat Protoc 2009;4:1-13.
28) Sibener LV, Fernandes RA, Kolawole EM, Carbone CB, Liu F, McAffee D, et al. Isolation of a structural mechanism for uncoupling T cell receptor signaling from peptide-MHC binding. Cell 2018;174:672-687.e27.
29) Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 2009;30:2785-2791.
30) Lucena MI, Molokhia M, Shen Y, Urban TJ, Aithal GP, Andrade RJ, et al.; Spanish DILI Registry; EUDRAGENE; DILIN; DILIGEN; International SAEC. Susceptibility to amoxicillin-clavulanate-induced liver injury is influenced by multiple HLA class I and II alleles. Gastroenterology 2011;141:338-347.
31) Singer JB, Lewitzky S, Leroy E, Yang F, Zhao X, Klickstein L, et al. A genome-wide study identifies HLA alleles associated with lumiracoxib-related liver injury. Nat Genet 2010;42:711-714.
32) Valenzuela-Ponce H, Alva-Hernandez S, Garrido-Rodriguez D, Soto-Nava M, Garcia-Téllez T, Escamilla-Gomez T, et al.; Mesoamerican HIV Project Group. Novel HLA class I associations with HIV-1 control in a unique genetically admixed population. Sci Rep 2018;8:6111. https://doi.org/10.1038/s41598-018-23849-7. https://www.nature.com/srep/.
33) Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PI, Walker BD, et al.; International HIV Controllers Study. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science 2010;330:1551-1557.
34) Li C, Rao T, Chen X, Zou Z, Wei A, Tang J, et al. HLA-B*35:01 allele is a potential biomarker for predicting polygonum multiflorum-induced liver injury in humans. Hepatology 2019;70:346-357.
35) Karnes JH, Shaffer CM, Bastarache L, Gaudieri S, Glazer AM, Steiner HE, et al. Comparison of HLA allelic imputation programs. PLoS One 2017;12:e0172444. https://doi.org/10.1371/journal.pone.0172444. https://journals.plos.org/plosone/.

Author names in bold designate shared co-first authorship.

Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.31258/suppinfo.