

Genetic risk factors in drug-induced liver injury due to isoniazid-containing anti-tuberculosis drug regimens

Paola Nicoletti¹, Harshad Devarbhavi², Ashish Goel³, Radha Venkatesan⁴, Chundamannil E. Eapen³, Jane I. Grove⁵, Samreen Zafer¹, Einar Bjornsson^{6,7}, M. Isabel Lucena⁷, Raul J. Andrade⁷, Munir Pirmohamed⁸, Mia Wadelius⁹, Dominique Larrey¹⁰, Anke-Hilse Maitland-van der Zee^{11,12}, Luisa Ibanez¹³, Paul B Watkins¹⁴, *Ann K. Daly¹⁵ and *Guruprasad P. Aithal⁵

¹ Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, USA

² Department of Gastroenterology, St Johns' Medical College Hospital, Bangalore, India

³ Christian Medical College, Vellore, India

⁴ Department of Molecular Genetics, Madras Diabetes Research Foundation, Chennai, India

⁵ NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and University of Nottingham, UK.

⁶ Division of Gastroenterology and Hepatology, Department of Internal Medicine, The National University Hospital of Iceland, Reykjavik, Iceland

⁷ Faculty of Medicine, University of Iceland, Reykjavik, Iceland

⁸ UGC Digestivo y Servicio de Farmacología Clínica, Instituto de Investigación Biomédica de Málaga (IBIMA), Hospital Universitario Virgen de la Victoria, Universidad de Málaga, Málaga, Spain and Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), Madrid, Spain

⁹ Department of Pharmacology and Therapeutics, University of Liverpool, Liverpool University Hospitals, and Liverpool Health Partners, Liverpool, UK

¹⁰ Department of Medical Sciences and Science for Life Laboratory, Uppsala University, Uppsala, Sweden,

¹¹ Liver Unit, CHU St Eloi Hospital, Montpellier, France

¹² Department of Respiratory Medicine, Academic Medical Center (AMC), University of Amsterdam, Amsterdam, Netherlands

¹³ Department of Pharmaceutical Sciences, Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht University, Utrecht, Netherlands

¹⁴ Fundació Institut Català de Farmacologia, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain

¹⁵ University of North Carolina Institute for Drug Safety Sciences, Eshelman School of Pharmacy, North Carolina, USA

¹⁶ Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, UK

*joint senior authors

Correspondence to

Guruprasad P. Aithal

Nottingham Digestive Diseases Centre,

School of Medicine, University of Nottingham,

Nottingham UK.

Email address: Guru.Aithal@nottingham.ac.uk

Key words: drug-induced liver injury, HLA genes, adverse drug reactions, genetic polymorphisms, N-acetyltransferase 2, isoniazid

Electronic word count: main text 4485 words

2 figures and 5 tables

CONFLICTS OF INTEREST

MP has received grants from the MRC Clinical Pharmacology Training Scheme (joint funding by MRC and Roche, UCB, Eli Lilly and Novartis), grants for a Joint PhD studentship funded by EPSRC and Astra Zeneca, and educational grants from Bristol Myers Squibb and UCB, but all these grants were in relation to other work outside of the subject area covered in this paper. PN is employee of Sema⁴. GPA has served as an advisory board member for Medicines and Healthcare Products Regulatory Agency, Department of Health and Social Care, UK, Pfizer and Glaxo SmithKline; he has been a consultant to Amryt Pharmaceuticals and Astra Zeneca. All other authors report no competing interests for this work.

FUNDING

The genome-wide association study and iDILIC case enrolment and sample collection in Europe and India was 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 partly (the DILIGEN and iDILIC sample collection) funded by the National Institute for Health Research (NIHR) Nottingham Digestive Diseases Biomedical Research Unit and NIHR Nottingham Biomedical Research Centre [BRC-1215-20003] at the Nottingham University Hospitals NHS Trust and University of Nottingham. GPA is the gastrointestinal and liver disorder theme lead for the NIHR Nottingham BRC (Reference no: BRC-1215-20003). The Spanish DILI Registry (RJA, MIL) is partly funded by the Spanish Medicine Agency, Fondo Europeo de Desarrollo Regional - FEDER (P10-CTS-6470, FIS PI12/00378, PI16/01748). CIBERehd is funded by Instituto de Salud Carlos III.

The Swedish case collection (SWEDEGENE) (PH, MW) has received support from the Swedish Medical Products Agency, the Swedish Society of Medicine (2008-21619), Swedish Research Council (Medicine 521-2011-2440 and 521-2014-3370), and Swedish Heart and Lung Foundation (20120557). The Swedish Twin Registry which provided control data is managed by Karolinska Institutet and receives funding through the Swedish Research Council under the grant no 2017-00641. MP is a NIHR Emeritus Senior Investigator. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Drug-induced liver injury (DILI) is a complication of treatment with anti-tuberculosis (TB) drugs, especially in isoniazid-containing regimens. To investigate genetic risk factors, we performed a genome-wide association study (GWAS) involving anti-TB DILI cases (55 Indian, 70 European) and controls (1199 Indian, 10397 European). Most cases were treated with a standard anti-TB drug regimen; all received isoniazid. We imputed single nucleotide polymorphism and HLA genotypes and performed trans-ethnic meta-analysis on GWAS and candidate gene genotypes. GWAS found one significant association (rs117491755) in Europeans only. For HLA, *HLA-B*52:01* was significant (meta-analysis odds ratio (OR) 2.67; 95%CI 1.63-4.37; $P=9.4 \times 10^{-5}$). For N-acetyltransferase 2 (NAT2), *NAT2*5* frequency was lower in cases (OR 0.69; 95%CI 0.57-0.83, $P=0.01$). *NAT2*6* and *NAT2*7* were more common, with homozygotes for *NAT2*6* and/or *NAT2*7* enriched among cases (OR 1.89; 95%CI 0.84-4.22; $P=0.004$). We conclude HLA genotype makes a small contribution to TB drug-related DILI and that the NAT2 contribution is complex, but consistent with previous reports when differences in the metabolic effect of *NAT2*5* compared with those of *NAT2*6* and *NAT2*7* are considered.

Introduction

Up to 20% of patients receiving isoniazid either as mono or combination therapy for tuberculosis (TB) may develop transient asymptomatic elevation of liver enzymes but this elevation usually resolves without drug discontinuation.^{1, 2} In a recent large prospective cohort study based in China, 5.4% of patients on anti-TB combination therapy developed drug-induced liver injury (DILI)³ as defined by the International DILI Expert Working Group.⁴ In 16% of these cases, DILI was accompanied by other symptoms of hepatotoxicity including jaundice and 5.3% developed acute liver failure.³ Recurrence of DILI upon retreatment (called positive rechallenge) has been reported to occur in approximately 9%-25% with at least one of the anti-TB drugs.^{1, 2, 5, 6} The incidence of DILI when isoniazid (INH) is combined with rifampicin appears higher than for INH alone, with addition of pyrazinamide increasing the risk further.⁷

A number of drug-specific and host related factors influence the susceptibility of a patient to DILI with anti-tuberculosis drugs.⁸ Potential mechanisms related to INH DILI have been the most widely investigated. Acetylhydrazine is generally considered to be a key INH metabolite contributing to INH-induced DILI and is produced by N-acetyltransferase 2 (NAT2). It can undergo further metabolism by cytochrome P450 to a toxic metabolite or by NAT2 to the less toxic diacetylhydrazine. It has been suggested that fast acetylators (those with NAT2 activity within the normal range) will form diacetylhydrazine efficiently and therefore levels of both acetylhydrazine and toxic P450 metabolites will be low.⁹ It has also been proposed that slow acetylators who lack NAT2 activity may form higher concentrations of the toxic metabolite hydrazine by cleavage of the amide bond on INH to form isonicotinic acid.¹⁰ There are also data to suggest formation of a reactive metabolite directly from INH oxidation which may contribute to liver toxicity through the formation of protein adducts¹¹ inducing an inappropriate immune response.¹²

Because of the important contribution by acetylation to INH metabolism and the existence of common loss of function polymorphisms in *NAT2*, the gene encoding the acetylation enzyme, there have been a large number of studies examining these variants as DILI risk factors. At least four different meta-analyses including large numbers of cases have concluded that slow acetylators have an increased risk of TB drug DILI with an overall odds ratio (OR) varying from 1.59 to 6.42, although the risk varied depending

on the precise genotypic definition of slow acetylation and the population studied.¹³⁻¹⁶ Although many of these studies involved a limited number of patients experiencing only mild liver injury, the *NAT2* association has also been observed in a study involving moderate to severe DILI cases only.¹⁷ A clinical trial based in Japan with differential dosing with INH on the basis of *NAT2* genotype found a lower incidence of DILI when slow acetylators were given a lower dosing regimen.¹⁸ A genome-wide association study (GWAS) involving patients from Thailand recently described a genome-wide significant signal for *NAT2*, suggesting a stronger risk for DILI development in those positive for slow acetylator alleles in line with many of the earlier candidate gene studies.¹⁹ However, two GWAS which included small numbers of European cases with anti-TB DILI did not find any genome-wide significant association with the *NAT2* slow acetylator genotype.^{20, 21}

In addition to *NAT2* variants, other candidate genetic risk factors for anti-TB DILI have also been investigated, with reports suggesting that genotypes for genes relevant to anti-TB drug disposition and oxidative stress such as *CYP2E1*,²² *SOD2*,²³ *GST* isoforms,²³ carboxyesterase (*CES*) isoforms²⁴ and *PXR* (*NR1I2*)²⁵ may modulate risk. Human leukocyte antigen (HLA) genotype is a strong risk factor for a number of forms of DILI²¹ and it has been suggested that HLA class II genotype is relevant to risk of DILI due to anti-TB drugs in an Indian population.²⁶ However, further studies using either candidate gene or genome-wide approaches have failed to confirm a role for HLA genotype in susceptibility to anti-TB drug-related DILI in Europeans.^{20, 21, 27} A recent study of DILI due to anti-TB drugs in an Ethiopian population who were also HIV-positive and receiving anti-HIV treatment found an interesting association with the class I HLA B*57 alleles,²⁸ but this form of DILI showed distinct phenotypic differences from that normally associated with anti-TB drugs. A non-HLA immunogenetic risk factor for DILI generally in the gene *PTPN22* has also been identified recently but its relevance to anti-TB drug related DILI has not been investigated in detail to date.²⁹

The aims of the current study were to perform a GWAS together with additional candidate gene studies on a newly recruited group of Indian patients with moderate to severe anti-TB drug DILI and on an enlarged European cohort which includes some cases studied previously.^{17, 20, 21}

Methods

Study design

This study combined in a trans-ethnic meta-analysis framework results from GWAS conducted separately in subjects with European and Indian ancestry. The study was conducted according to the Declaration of Helsinki (Hong Kong Amendment) and Good Clinical Practice (European guidelines). All participants provided written informed consent and each study was approved by the appropriate local (Department of Gastroenterology, St John's Medical College Hospital, Bangalore, India and Christian Medical College, Vellore, India), national or institutional ethical review boards as reported previously.^{21, 29}

Indian cohort

Patients (n=55) who developed DILI after exposure to INH in combination with rifampicin, pyrazinamide and ethambutol (all 4 drugs for first 2 months and INH and rifampicin for further 4 months), 105 patients treated with these drugs without DILI development and 104 healthy South Indian adults of mixed ancestries were enrolled from Aug 2009 to Feb 2014 at St John's Medical College, Bengaluru and Christian Medical College, Vellore, South India. Roussel Uclaf Causality Assessment Method (RUCAM) was used for case adjudication as described previously.⁴ To further increase study power, we added to the control set a total of 990 ethnically matched samples comprising 356 from 1000G project and 634 controls of Indian descent from Charles Bronfman Institute for Personalized Medicine BioMe BioBank (phs000925.v1.p1) identified by principal component analysis (see Supplementary Materials).

European cohort

We analyzed 70 European ancestry DILI cases exposed at least to INH alone or in combination with one or more of rifampicin, pyrazinamide and ethambutol, collected by the DILIGEN, iDILIC and DILIN consortia. Of these cases, 43 were treated with INH alone as standard TB prophylaxis and 27 had been treated with any combination of INH and one or more additional anti-TB drugs with 11 of those exposed to all four drugs. These cases all form part of a large cohort of DILI cases previously analyzed by GWAS,²⁹ with some also included in earlier studies as summarised in Table S1. A subgroup of 12 UK cases (DILIGEN study) had been included in an earlier study

involving direct *NAT2* genotyping only¹⁷ but subsequently also underwent genome-wide genotyping.²⁰ European ancestry controls (n=10397) were used.²⁹

Clinical characterisation of DILI

Criteria used for case definition, categorisation of DILI pattern as well as grading of severity of DILI were harmonised across all cohorts using previous guidelines.⁴ Causality assessment was done by RUCAM score for the DILIGEN and iDILIC cohorts and by both structured expert opinion and RUCAM score for the DILIN cohort, as previously reported.^{20, 21}

DNA preparation and genotyping

For the Indian cases and controls, DNA was isolated from whole blood samples using the standard phenol-chloroform method. DNA samples were genotyped in one batch using the Illumina Human Core Exome-24 BeadChip by Department of Molecular Genetics, Madras Diabetes Research Foundation, India. DNA isolation and genotyping of European cases from the DILIGEN, iDILIC and DILIN studies was as described previously.^{20, 21}

Population structure and imputation

Quality control (QC) checks on the initial genotype data were performed as described previously.²⁹ To assess the extent of population structure of the study cohorts, and derive eigenvectors to account for confounding, we applied principal component analysis on each cohort separately, using the smartPCA program from the EIGENSTRAT package (version 3.0)³⁰ on the overlapping single nucleotide polymorphisms (SNPs) (minor allele frequency, MAF>0.01) across the range of genotyping arrays used for typing cases and controls. We used 1000 Genome Project samples as the reference panel to select cases and controls of Caucasian and South Indian ancestries. SNP imputation was performed in batches dividing the samples according to ethnicity and genotyping platforms. For each batch imputation was carried out using Michigan Imputation Server as described previously.²⁹

GWAS analysis and meta-analysis

We tested for association of each SNP with DILI, separately in Indian and European GWAS, in a logistic regression framework, under an additive genetic model, with adjustment for the principal components from smartPCA to account for population structure using PLINK v 1.07.³¹ No other additional covariates were included in the model since we did not have clinical information for controls. Association summary statistics from the two cohorts were combined using effective sample size weighted Z-score fixed-effects meta-analysis, implemented in METAL.³² Allelic odds ratios (ORs) across the two cohorts were obtained through inverse-variance weighting of effect sizes, with heterogeneity assessed with Cochran's Q statistic,³³ implemented in METAL. We reported only those SNPs that attained, in addition to genome-wide significance, nominal evidence of association ($P < 0.05$) with the same direction of effect on DILI in both GWAS phases (internal validation). Genome-wide significance with clinical outcomes was defined using a common threshold of $P < 5 \times 10^{-8}$. Since the phenotypes studied are rare, the number of cases analyzed was limited. All detailed analyses and Manhattan plots were produced with R (version 3.0.2, The R Project for Statistical Computing, <http://www.r-project.org>). Regional plots were drawn by LocusZoom as described previously.²¹

HLA analysis

For each cohort, HLA alleles were inferred using HIBAG³⁴ using the reference predictor panels specific for the genotyping chip and ancestry provided in the software webpage. To impute Indian samples we used the provided Asian reference data.³⁴ In total, we imputed 217 HLA alleles in the overall European cohort and 192 in the Indian cohort. We set the MHC-region-wide significance p-value threshold for the HLA allele association to 2.5×10^{-4} to correct for multiple testing (Bonferroni correction for 200 predicted HLA alleles). Association test for each HLA allele and meta-analysis were carried out as reported above. Haplotype analysis was performed by Plink 1.07 including the most significantly associated HLA alleles.

NAT2 genotypes

We predicted *NAT2* alleles using genotypes for rs1801280, rs1799930 and rs1799931 from the GWAS (*NAT**5, *6 and *7 alleles respectively) by haplo.stats (<https://cran.r-project.org/web/packages/haplo.stats/index.html>) and extracted the best haplotype predictions based on posterior probability. We tested for association between DILI and

allele by logistic regression with adjustment for the principal components in Plink. NAT2 genotypes were recorded and samples were divided into different acetylator status groups as originally proposed³⁵ and refined more recently:³⁶ (a) rapid (*4/*4); (b) intermediate (*4/*5,*4/*6,*4/*7); (c) slow (*5/*5,*5/*6,*5/*7); (d) ultraslow (*6/*6,*6/*7,*7/*7). We then tested for association between DILI and acetylator status groups and genotypes in a multivariate regression model including principal components axes and acetylator status groups or genotypes as binomial variable or as categorical variable having 44 genotype or rapid/intermediate as baseline groups. The meta-analysis p-value was calculated using METAL with the default approach that combines p-value and direction of effect, weighted according to sample size.³² Significance was defined using a Bonferroni threshold of $P < 0.01$ considering 4 comparisons.

Candidate gene analysis

For analysis of additional candidate genes, we selected four genes previously proposed to have a role in the pathogenesis of INH-related DILI, including CYP2E1, CES2, CES1 and PXR/NR1I2. We extracted all variants belonging to each gene from GnomAD (<https://gnomad.broadinstitute.org/>). We performed an association analysis by Plink 1.07.³¹ Significance was defined using a Bonferroni threshold of $P < 0.003$ considering 16 multiple comparisons.

Results

Cases and controls

Clinical characteristics of the 55 DILI cases from India and 70 European cases are summarised in Table 1. The total Indian DILI cohort was enriched in cases with severe liver injury with 22% of all patients progressing to acute liver failure, transplant or death. The majority of the European DILI cases met the definition of moderate to severe DILI used in our previous European DILI studies.²¹ Nine European patients (13%) suffered liver failure, underwent liver transplant or died (Table 1).

As controls for the Indian cases, we used 1199 ethnically matched individuals (105 patients treated for TB by similar regimens who did not develop DILI and 104 healthy adults recruited for this study, 356 from 1000G project phase 3 dataset and 634 from BioMe dataset (Figure S1)). As European controls, we used 10397 ethnically matched individuals (Figure S2).

A schematic summary of the overall study is provided in Figure 1.

GWAS analysis

The case-control GWAS in the European cohort showed one marker that passed the significance threshold (rs117491755, OR=4.37, 95%CI [2.702-7.061], $P=1.8 \times 10^{-9}$; $AF_{\text{cases}} = 0.143$; $AF_{\text{controls}} = 0.037$) (Figure 2A). rs117491755 is an intronic SNP in *ASTN2*. For the Indian cohort, there were no genome-wide significant associations between DILI and imputed or genotyped variants and rs117491755 did not pass QC checks in the imputed dataset (Figure 2B). Trans-ethnic meta-analysis between the two ethnic groups analysed 4,900,532 shared markers between the two cohorts. None of the variants in the meta analysis reached genome-wide significance (Figure 2C).

HLA analysis

In the European cohort, we identified two major histocompatibility complex (MHC) significant signals, *HLA-C*12:02* (OR=6.43; 95%CI [2.53-16.37] ; $P = 9.4 \times 10^{-5}$, Table 2 and Table S1) and *HLA-B*52:01* (OR=6.39; 95%CI [2.25-16.29] ; $P = 1.0 \times 10^{-4}$, Table 2). All carriers presented with both alleles. The *B*52:01-C*12:02* haplotype conferred a significant increase in DILI risk of almost 7-fold ($P = 7.8 \times 10^{-5}$, Table S2), although the overall haplotype frequency is low even in the cases. The alleles were mainly

associated with DILI cases due to drug combinations compared with INH alone (Table S3). Slightly different HLA results were observed for Indian cases. These cases showed an enrichment in class II HLA risk alleles with *HLA-DQA1*03:01* (OR= 2.60, 95%CI[1.54-4.38]; P = 0.0003, Table 2) as the most significantly associated allele, though not passing the MHC multiple correction threshold. Indian cases also showed a nominal enrichment of *HLA-B*52:01* and *HLA-C*12:02* alleles compared with controls (Table S2). In line with the data from Europeans, the haplotype conferred a 1.45 fold increase in risk of DILI (P = 0.04, Tables 2 and S2). In meta-analysis *HLA-B*52:01* showed a MHC-significant association (OR=2.67; 95%CI 1.63-4.37; P = 9.4×10^{-5} , Table 3).

Candidate gene analysis

NAT2 genotypes

The frequency of the *NAT2* alleles in both cohorts are summarised in Table 3. *NAT2* allele frequencies for our cohorts were similar to frequencies reported previously with Indians showing a high frequency of the *NAT2*6* allele in the control group compared to Europeans whereas in the European controls *NAT2*5* was the most common allele.³⁷ We tested association between each allele and DILI correcting for population stratification (Table 4). We found that *NAT2*5* was the most underrepresented *NAT2* allele in both European and Indian cases, passing the Bonferroni correction for the number of the predicted *NAT2* alleles ($P_{\text{Meta-analysis}} = 0.01$, Table 4). *NAT2*6* was significantly enriched in Indian cases (P = 0.01). Severe Indian cases (n = 12) showed a slightly higher frequency for *NAT2*6* (AF = 0.54) and lower frequency of *NAT2*5* (AF = 0.11) compared to the overall cases, with *NAT2*5* statistically significant as a protective factor compared to controls (OR = 0.30; 95%CI 0.08 - 0.94; P = 0.05). Among Europeans, INH alone cases (n = 43) showed a significantly decreased frequency for *NAT2*5* (AF = 0.34, P = 0.04) (Table 4). The group of European severe cases (n=9) was too small to detect significant trends.

We then assigned *NAT2* genotypes for each individual and classified them in four acetylator status groups. We evaluated if there was an enrichment of any of the groups in our cases compared to controls. We found that ultraslow acetylators (those carrying homozygous or compound heterozygous genotypes for *NAT2*6* and *NAT2*7*) were

significantly enriched in cases compared with controls ($P_{\text{Meta-analysis}} = 0.004$, Table 5). This enrichment was not significant in a multivariable regression model for which we combined rapid and intermediate acetylators as the baseline group ($P_{\text{European}} = 0.08$, and $P_{\text{Indian}} = 0.1$). Examining the enrichment of single genotype groups, the only set significantly associated with DILI by meta-analysis was *NAT2*6/NAT2*7* ($P = 0.002$) which was also significant in Europeans alone ($P = 0.01$, Table S4).

Other candidate gene analysis

We extracted variants located in selected candidate genes relevant to INH metabolism (CYP2E1, CES2, CES1 and PXR/NR1I2) from GnomAD (<https://gnomad.broadinstitute.org>) for a total of 16 imputed/genotyped SNPs in European (CYP2E1 n = 9; CES2 n = 2; CES1 n = 0; PXR/NR1I2 n = 5) 14 in Indians (CYP2E1 n = 5; CES2 n = 3; CES1 n = 4; PXR/NR1I2 n = 2). Since imputation is based on ethnicities and genotyping platforms, the SNPs available for analysis in the candidate genes were different for the two groups. No variants were significantly associated with DILI (Table S5).

Frequency of rs2476601 in *PTPN22* that has been previously associated with DILI due to several different drugs,²⁹ was not increased in European DILI cases ($AF_{\text{cases}} = 0.10$; $P = 0.40$) and INH alone cases ($AF_{\text{cases}} = 0.12$; $P = 0.15$) but was marginally increased in the Indian cases ($OR = 3.8$ 95%CI [1.06-13.83] $P = 0.04$ $AF_{\text{cases}} = 0.03$ and $AF_{\text{controls}} = 0.01$).

Discussion

Despite a relatively large number of published studies, the genetic basis for susceptibility to DILI due to anti-TB drugs including INH remains poorly understood compared with DILI caused by certain other drugs such as flucloxacillin and amoxicillin-clavulanate. There are a number of reasons for this including: (a) the complexity of the phenotype (both mild and more serious cases of DILI due to anti-TB drugs are common); (b) the fact that the standard treatment typically involves a combination of four different drugs; (c) both TB as a disease and DILI induced by anti-TB drug treatment are more common in developing countries where relevant genetic polymorphisms may show differing frequencies than in Europeans where DILI has been studied more extensively, making worldwide comparisons difficult. Furthermore, in some countries where concomitant infection with TB and HIV is more common, assessment of causality as to whether the DILI is due to the anti-HIV drugs, anti-TB drugs, or both, is often unclear.

Similar to the findings in two earlier GWAS involving Europeans,^{20, 21} we failed to detect genome-wide significance when we undertook a GWAS in the Indian population. This is in contrast to the recent GWAS performed on a Thai population, where genome-wide significance was seen for *NAT2*,¹⁹ but is more consistent with a separate GWAS performed in Ethiopians which also failed to observe genome-wide significant signals,³⁸ although the Ethiopian patients were also undergoing HIV treatments that can cause DILI. In our enlarged European population which includes the previously studied cases^{20, 21} as well as 8 new cases,²⁹ we did see one genome-wide significant signal in an intronic SNP in *ASTN2*. This gene product appears to affect synaptic strength by trafficking and degradation of surface proteins.³⁹ The relevance to DILI is not immediately clear and this signal was not seen in the Indian population. Without a positive replication, we have to consider this signal as either a false-positive or putative, until further patients have been studied.

While the current study and several previous reports failed to detect strong HLA associations, we found some evidence that a rare *HLA-C*12:02 -B*52:01* haplotype, which has been recently reported to be a risk factor for Crohn's disease in Asians, might contribute to risk of anti-TB drug DILI in some individuals.⁴⁰ The effect of the haplotype seems to be consistent across both cohorts and the association passes the

Bonferroni correction based on the number of imputed alleles. Meta-analysis showed that the effect of the *B*52:01* allele alone was more significant than for *C*12:02*. The number of cases positive for the risk allele was low but importantly both cohorts showed this association and that there is an already reported association for the haplotype with an autoimmune disease.⁴⁰ In view of the small number of cases positive for the "at risk" haplotype and the absence of a signal in the INH only cases, it is possible that the HLA signal may reflect DILI induced by one of the other anti-TB drugs, possibly pyrazinamide.

A previous HLA class II typing study in an Indian population reported that absence of *HLA-DQA1*01:02*, and presence of *HLA-DQB1*02:01* were risk factors for DILI due to anti-TB drugs.²⁶ However, in the current study the most significant findings for class II were increased frequencies of *HLA-DQA1*01:03* and *HLA-DQA1*03:01*, although we did not observe this in Europeans. We also saw no association with *B*57* alleles but as discussed previously, we consider that this particular association may relate to a combination of anti-HIV and anti-TB drug treatment²⁸ which we would not expect to see confirmed in the current study.

The *PTPN22* variant rs2476601 has recently been found to be an additional risk factor for some forms of DILI showing HLA associations.²⁹ We therefore also evaluated the role of this variant in our patients with anti-TB-related DILI, but did not find a significant association. However, the allele frequency of rs2476601 in South Asian populations is consistently lower compared to European populations (1% vs 10%, as reported in GnomAD) so our ability to detect any association was also limited.

In view of very limited genome-wide or near genome-wide significant signals in the GWAS, we proceeded with additional candidate gene analysis using the GWAS data, focussing on *NAT2*, given the extensive literature which has demonstrated that *NAT2* genotype and phenotype is a risk factor for INH-induced DILI. Three meta analyses on *NAT2* as a DILI risk factor which include studies published up to 2017 together with two recent large studies appear to be the most informative to use for comparison with the current study.^{13-15, 19, 41} In general, most previous studies on *NAT2* genotype as a risk factor have examined all slow acetylators in comparison with either homozygous wild-types (who are often now classified as the fast acetylator group without inclusion of heterozygotes) or both homozygous wild-types and those heterozygous for one

variant allele only (the traditional fast acetylator group). In some studies, heterozygotes are designated as intermediate acetylators and were analysed separately.

There is however increasing data available which indicates that *NAT2**5, which is common in Europeans, South Asians and Africans, but not in East Asians, is not a true "slow acetylator" allele with the gene product retaining some enzyme activity while the enzymes encoded by both *NAT2**6 and *NAT2**7 are associated with no activity.^{36, 42, 43} In line with this, a recent study of South African Zulus examined levels of INH and certain metabolites in relation to *NAT2* genotype in patients undergoing treatment with INH and did not find a significant difference in drug and metabolite levels when comparing *NAT2**5 homozygotes or heterozygotes with those homozygous for two rapid acetylator alleles. *NAT2**6 and *NAT2**7 alleles were not detected in this population.⁴⁴ One relatively small study of phenotype-genotype relationships in healthy Swedish volunteers used INH for phenotype determination and reported a higher metabolic ratio for *NAT2**5 homozygotes compared with *NAT2**6 homozygotes.⁴⁵

The *NAT2* genotype distribution among the DILI cases in the current study indicates a protective effect for the *NAT2**5 allele with *NAT2**4 "neutral" whereas an increased risk was seen for *NAT2**6 with the combined "ultraslow" group also showing a statistically significant increased risk. This is in agreement with a meta-analysis reporting an increased risk for *NAT2**6 and *NAT2**7.⁴³ A recent study performed in Singapore also suggests that the risk for INH-related DILI is from *NAT2**6 and *7 only.⁴¹ This could also explain the recent genome-wide significance reported for *NAT2* variants in Thailand¹⁹ since *NAT2**5 is rarely seen in this East Asian population. One of the earliest reports on *NAT2* genotype as a risk factor for INH-related DILI was performed in Taiwan.⁴⁶ This study found a small number of individuals positive for *NAT2**5 but hepatotoxicity was seen almost entirely in those carrying at least one *NAT2**6 or *NAT2**7 allele. The biological basis for this complex association with *NAT2* genotype is not completely clear but *6 and *7 carriers may be at increased risk of toxicity due to higher levels of the parent drug undergoing metabolism by alternative routes to toxic intermediates such as hydrazine or possibly by accumulation of the acetylhydrazine metabolite which also may be converted to hydrazine.⁹ INH-related DILI was reported to be more common among East Asians compared with white Europeans and African-Americans in early population studies.^{47, 48} This could reflect the higher frequency of the *NAT2**6 and *NAT2**7 alleles in these populations⁴⁶

compared with those reported for Europeans,⁴⁵ despite the overall average higher acetylation activity seen in East Asians. It also remains possible that the recent GWAS findings reported for a Thai population¹⁹ showing significance for *NAT2* are not directly comparable to the current study as the liver enzyme elevation thresholds for participation in that study were considerably lower than in the current study.

We also studied four additional genes potentially relevant to INH disposition in detail in both cohorts to see if any evidence for trends towards genome-wide significance could be detected. These were chosen on the basis of direct relevance to the INH metabolic pathway⁹ and either encode enzymes (*CYP2E1*, *CES1* and *CES2*) or transcription regulators with a role in regulation of gene expression (*PXR/NR1I2* which regulates *CES* expression). Our findings were entirely negative. We believe this is not too surprising and is generally consistent with reports in the existing literature of no significance or small effects.^{24, 49} Larger studies might enable the detection of smaller effects than was feasible in the current study.

A limitation of this study and most others on DILI due to anti-TB drugs is that in addition to INH, the other drugs used in treatment, especially pyrazinamide, can also cause DILI. All cases in the Indian cohort were related to combination anti-TB drug therapy while a significant proportion of DILI cases in the European cohort were attributable to INH monotherapy so the two cohorts are not identical in terms of drug treatment. This is an important limitation but the results obtained for the two cohorts, especially for *NAT2* genotype and to some extent for *HLA* genotype, are still comparable. Alternative regimens not involving INH show slightly lower incidence of hepatotoxicity but DILI may still occur⁵⁰ and although it may be possible to determine which drug causes DILI in a particular patient, this requires a series of individual drug rechallenges which is difficult to perform and was not feasible in the current study.

In conclusion, we have obtained some evidence that certain *NAT2* alleles increase the risk of DILI in patients receiving INH-containing anti-TB drug regimens. The two cohorts we studied were large compared with most previous studies with a well-defined phenotype but were still small compared with GWAS on more common diseases where relatively small but significant effects can be detected by use of very large cohorts. There is a need for additional studies involving either larger cases numbers or additional

meta analyses to better understand the underlying risk factors for DILI due to INH and other anti-TB drugs.

STUDY HIGHLIGHTS

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

Anti-tuberculosis (TB) drugs including isoniazid are a common cause of drug-induced liver injury (DILI). Previous reports suggest *NAT2* genotype and some HLA alleles are risk factors but not all studies agree on this.

WHAT QUESTION DID THIS STUDY ADDRESS?

We aimed to identify novel genetic risk factors for DILI due to anti-TB drugs including isoniazid in European and Indian cases and consolidate understanding on relevance of HLA and *NAT2* genotypes to risk of DILI.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

The study provides further support for *NAT2**6 and *NAT2**7 variants of *NAT2* as risk factors for development of anti-TB drug-related DILI and for *NAT2**5 being protective. There may also be increased risk in those carrying the *HLA-B**52:01 allele.

HOW THIS MIGHT CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

The study provides further support for *NAT2* and HLA contributions to risk of DILI from TB drugs, adding to knowledge that may lead to genetic tests capable of identifying those patients at risk.

Acknowledgements

We are extremely grateful to Arthur Holden of iSAEC for his support. We are grateful to Julia Patch, Julian Leathart and Julian Arbuckle for technical help, 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. MP thanks the MRC Centre for Drug Safety Science for infrastructure support. We thank the Drug Induced Liver Injury Network (DILIN) for providing its data included in this paper which is now publicly available in dbGAP. The DILIN was not involved in data analyses, manuscript preparation or its review. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the UK Department of Health.

Author contributions: G.P.A. and A.K.D. designed the research; G.P.A., H.D., A.G., C.E.E., E.B., M.W., D.L., L.I., P.B.W., A.-H.M.-Z., R.J.A., M.I.L., M.P., R.V., J.I.G. and A.K.D. performed the research; P.N., S.Z., G.P.A and A.K.D. analyzed the data; P.N., G.P.A. and A.K.D. wrote the manuscript.

References

1. Sharma, S.K. *et al.* Safety of 3 different reintroduction regimens of antituberculosis drugs after development of antituberculosis treatment-induced hepatotoxicity. *Clin Infect Dis* **50**, 833-9 (2010).
2. Tahaoglu, K. *et al.* The management of anti-tuberculosis drug-induced hepatotoxicity. *Int J Tuberc Lung Dis* **5**, 65-9 (2001).
3. Liang, L., Huadong, Y., Du, J.Y., Susu, J., Hu, Y. & Aithal, G.P. Prevalence and risk factors of first line anti-tuberculosis drug induced liver injury: Large cohort study involving 4652 patients *J Hepatol* **70**, E62-E3 (2019).
4. Aithal, G.P. *et al.* Case definition and phenotype standardization in drug-induced liver injury. *Clin Pharmacol Ther* **89**, 806-15 (2011).
5. Mitchell, J.R. *et al.* Isoniazid liver injury: clinical spectrum, pathology and probable pathogenesis. *Annals of Internal Medicine* **84**, 181-92 (1976).
6. Singanayagam, A. *et al.* A comparison between two strategies for monitoring hepatic function during antituberculous therapy. *Am J Respir Crit Care Med* **185**, 653-9 (2012).
7. Steele, M.A., Burk, R.F. & DesPrez, R.M. Toxic hepatitis with isoniazid and rifampin. A meta-analysis. *Chest* **99**, 465-71 (1991).
8. Ramappa, V. & Aithal, G.P. Hepatotoxicity Related to Anti-tuberculosis Drugs: Mechanisms and Management. *J Clin Exp Hepatol* **3**, 37-49 (2013).
9. Wang, P., Pradhan, K., Zhong, X.B. & Ma, X. Isoniazid metabolism and hepatotoxicity. *Acta Pharm Sin B* **6**, 384-92 (2016).
10. Sarich, T.C., Adams, S.P., Petricca, G. & Wright, J.M. Inhibition of isoniazid-induced hepatotoxicity in rabbits by pretreatment with an amidase inhibitor. *J Pharmacol Exp Ther* **289**, 695-702 (1999).
11. Mitchell, J.R. *et al.* Increased incidence of isoniazid hepatitis in rapid acetylators: possible relation to hydranize metabolites. *Clin Pharmacol Ther* **18**, 70-9 (1975).
12. Usui, T. *et al.* Detection of Drug-Responsive T-Lymphocytes in a Case of Fatal Antituberculosis Drug-Related Liver Injury. *Chem Res Toxicol* **29**, 1793-5 (2016).
13. Zhang, M. *et al.* The association between the NAT2 genetic polymorphisms and risk of DILI during anti-TB treatment: a systematic review and meta-analysis. *Br J Clin Pharmacol* **84**, 2747-60 (2018).
14. Khan, S. *et al.* Pharmacogenetic association between NAT2 gene polymorphisms and isoniazid induced hepatotoxicity: trial sequence meta-analysis as evidence. *Biosci Rep* **39**, BSR20180845 (2019).
15. Richardson, M., Kirkham, J., Dwan, K., Sloan, D.J., Davies, G. & Jorgensen, A.L. NAT2 variants and toxicity related to anti-tuberculosis agents: a systematic review and meta-analysis. *Int J Tuberc Lung Dis* **23**, 293-305 (2019).
16. Cai, Y., Yi, J., Zhou, C. & Shen, X. Pharmacogenetic study of drug-metabolising enzyme polymorphisms on the risk of anti-tuberculosis drug-induced liver injury: a meta-analysis. *PLoS One* **7**, e47769 (2012).
17. Ng, C.S. *et al.* N-acetyltransferase 2 (NAT2) genotype as a risk factor for development of drug-induced liver injury relating to antituberculosis drug treatment in a mixed-ethnicity patient group. *Eur J Clin Pharmacol* **70**, 1079-86 (2014).

18. Azuma, J. *et al.* NAT2 genotype guided regimen reduces isoniazid-induced liver injury and early treatment failure in the 6-month four-drug standard treatment of tuberculosis: a randomized controlled trial for pharmacogenetics-based therapy. *Eur J Clin Pharmacol* **69**, 1091-101 (2013).
19. Suvichapanich, S. *et al.* Genomewide Association Study Confirming the Association of NAT2 with Susceptibility to Antituberculosis Drug-Induced Liver Injury in Thai Patients. *Antimicrob Agents Chemother* **63**, e02692-18 (2019).
20. Urban, T.J. *et al.* Limited contribution of common genetic variants to risk for liver injury due to a variety of drugs. *Pharmacogenet Genomics* **22**, 784-95 (2012).
21. Nicoletti, P. *et al.* 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* **152**, 1078-89 (2017).
22. Vuilleumier, N. *et al.* CYP2E1 genotype and isoniazid-induced hepatotoxicity in patients treated for latent tuberculosis. *Eur J Clin Pharmacol* **62**, 423-9 (2006).
23. Huang, Y.S. *et al.* Genetic polymorphisms of manganese superoxide dismutase, NAD(P)H:quinone oxidoreductase, glutathione S-transferase M1 and T1, and the susceptibility to drug-induced liver injury. *J Hepatol* **47**, 128-34 (2007).
24. Yamada, S. *et al.* Genetic variation in carboxylesterase genes and susceptibility to isoniazid-induced hepatotoxicity. *Pharmacogenomics J* **10**, 524-36 (2010).
25. Wang, Y. *et al.* Association of PXR and CAR Polymorphisms and Antituberculosis Drug-Induced Hepatotoxicity. *Sci Rep* **9**, 2217 (2019).
26. Sharma, S.K., Balamurugan, A., Saha, P.K., Pandey, R.M. & Mehra, N.K. Evaluation of clinical and immunogenetic risk factors for the development of hepatotoxicity during antituberculosis treatment. *Am J Respir Crit Care Med* **166**, 916-9 (2002).
27. Leiro-Fernandez, V., Valverde, D., Vazquez-Gallardo, R., Constenla-Carames, L., Del Campo-Perez, V. & Fernandez-Villar, A. HLA-DQ B1*0201 and A1*0102 Alleles Are Not Responsible for Antituberculosis Drug-Induced Hepatotoxicity Risk in Spanish Population. *Front Med (Lausanne)* **3**, 34 (2016).
28. Petros, Z., Kishikawa, J., Makonnen, E., Yimer, G., Habtewold, A. & Aklillu, E. HLA-B(*)57 Allele Is Associated with Concomitant Anti-tuberculosis and Antiretroviral Drugs Induced Liver Toxicity in Ethiopians. *Front Pharmacol* **8**, 90 (2017).
29. Cirulli, E.T. *et al.* A Missense Variant in PTPN22 is a Risk Factor for Drug-induced Liver Injury. *Gastroenterology* **156**, 1707-16.e2 (2019).
30. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. & Reich, D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-9 (2006).
31. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
32. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).

33. Ioannidis, J.P., Patsopoulos, N.A. & Evangelou, E. Heterogeneity in meta-analyses of genome-wide association investigations. *PLoS One* **2**, e841 (2007).
34. Zheng, X. *et al.* HIBAG--HLA genotype imputation with attribute bagging. *Pharmacogenomics J* **14**, 192-200 (2014).
35. Deguchi, T. Sequences and expression of alleles of polymorphic arylamine N-acetyltransferase of human liver. *J Biol Chem* **267**, 18140-7 (1992).
36. Selinski, S., Blaszkewicz, M., Ickstadt, K., Hengstler, J.G. & Golka, K. Refinement of the prediction of N-acetyltransferase 2 (NAT2) phenotypes with respect to enzyme activity and urinary bladder cancer risk. *Arch Toxicol* **87**, 2129-39 (2013).
37. Sabbagh, A., Darlu, P., Crouau-Roy, B. & Poloni, E.S. Arylamine N-acetyltransferase 2 (NAT2) genetic diversity and traditional subsistence: a worldwide population survey. *PLoS One* **6**, e18507 (2011).
38. Petros, Z. *et al.* Genome-Wide Association and Replication Study of Hepatotoxicity Induced by Antiretrovirals Alone or with Concomitant Anti-Tuberculosis Drugs. *Omics* **21**, 207-16 (2017).
39. Behesti, H. *et al.* ASTN2 modulates synaptic strength by trafficking and degradation of surface proteins. *Proc Natl Acad Sci U S A* **115**, 9717-26 (2018).
40. Han, B. *et al.* Amino acid position 37 of HLA-DRbeta1 affects susceptibility to Crohn's disease in Asians. *Hum Mol Genet* **27**, 3901-10 (2018).
41. Chan, S.L. *et al.* Association and clinical utility of NAT2 in the prediction of isoniazid-induced liver injury in Singaporean patients. *PLoS One* **12**, e0186200 (2017).
42. Doll, M.A. & Hein, D.W. Genetic heterogeneity among slow acetylator N-acetyltransferase 2 phenotypes in cryopreserved human hepatocytes. *Arch Toxicol* **91**, 2655-61 (2017).
43. Suvichapanich, S. *et al.* NAT2 ultra-slow acetylator and risk of anti-tuberculosis drug-induced liver injury: a genotype-based meta-analysis. *Pharmacogenet Genomics* **28**, 167-76 (2018).
44. Mthiyane, T. *et al.* N-Acetyltransferase 2 Genotypes among Zulu-Speaking South Africans and Isoniazid and N-Acetyl-Isoniazid Pharmacokinetics during Antituberculosis Treatment. *Antimicrob Agents Chemother* **64**, e02376-19 (2020).
45. Smith, C.A., Wadelius, M., Gough, A.C., Harrison, D.J., Wolf, C.R. & Rane, A. A simplified assay for the arylamine N-acetyltransferase 2 polymorphism validated by phenotyping with isoniazid. *J Med Genet* **34**, 758-60 (1997).
46. Huang, Y.S. *et al.* Polymorphism of the N-acetyltransferase 2 gene as a susceptibility risk factor for antituberculosis drug-induced hepatitis. *Hepatology* **35**, 883-9 (2002).
47. Black, M., Mitchell, J.R., Zimmerman, H.J., Ishak, K.G. & Epler, G.R. Isoniazid-associated hepatitis in 114 patients. *Gastroenterology* **69**, 289-302 (1975).
48. Kopanoff, D.E., Snider, D.E., Jr. & Caras, G.J. Isoniazid-related hepatitis: a U.S. Public Health Service cooperative surveillance study. *Am Rev Respir Dis* **117**, 991-1001 (1978).
49. Wang, F.J., Wang, Y., Niu, T., Lu, W.X., Sandford, A.J. & He, J.Q. Update meta-analysis of the CYP2E1 RsaI/PstI and DraI polymorphisms and risk of

- antituberculosis drug-induced hepatotoxicity: evidence from 26 studies. *J Clin Pharm Ther* **41**, 334-40 (2016).
50. Tweed, C.D. *et al.* Liver toxicity associated with tuberculosis chemotherapy in the REMoxTB study. *BMC Med* **16**, 46 (2018).

Table 1 Clinical phenotype for the Indian and European cases

Groups	Indian cohort	European cohort
Clinical characteristics		
Number of cases	55	70
Gender (F/M)	28/27	37/33
Age (years) (Mean ,SD)	40 (16.1)	55.1 (13.9)
Time to onset from first drug exposure (days) (Mean, SD)	46.8 (55.1)	55.2 (45.8)
Pattern of DILI		
Cholestatic	18% (10)	5.7% (4)
Hepatocellular	59% (32)	77.1% (54)
Mixed	22% (12)	10.0% (7)
Unknown	1% (1)	7.1% (5)
Causal drug		
Isoniazid and rifampicin (IR)		1.4% (1)
Isoniazid, rifampicin, pyrazinamide (IRP)		21.4% (15)
Isoniazid, rifampicin, pyrazinamide and ethambutol (IRPE)	100% (55)	15.7% (11)
Isoniazid (I)		61.43% (43)
Severity		
Mild	0	11% (8)
Moderate	78% (43)	73% (51)
Severe/Fatal	22% (12)	13% (9)
Not Reported	0	3% (2)
Causality Score (CIOMS/RUCAM)		
3-5 (possible)	6% (3)	18% (13)
6-8 (probable)	60% (33)	60% (42)
>8 (highly probable)	34% (19)	21% (15)
Genotyping platform		
1M Illumina Duo (% ,N)		39% (27)
Infinium Core Exome (% ,N)	100% (55)	50% (35)
Multi-Ethnic Genotyping Array Consortium (% ,N)		11% (8)

Table 2 The most significant HLA associations for European and Indian cohorts

Markers	OR	95%CI	P	AF Cases	AF Controls	AF reference dataset*
European associations						
<i>HLA-C*12:02</i>	6.43	2.526-16.37	0.00009	0.04	0.006	0.009
<i>HLA-B*52:01</i>	6.40	2.511-16.29	0.0001	0.04	0.007	0.009
<i>HLA-DRB1*15:02</i>	6.36	2.489-16.25	0.0001	0.04	0.006	0.007
Indian associations						
<i>HLA-DQA1*03:01</i>	2.60	1.53-4.38	0.00035	0.15	0.06	0.09
<i>HLA-DPB1*01:01</i>	3.24	1.58-6.61	0.0013	0.09	0.02	0.02
<i>HLA-DPB1*03:01</i>	3.60	1.58-8.19	0.0023	0.07	0.02	0.05
<i>HLA-DRB1*04:06</i>	8.48	2.02-35.52	0.00346	0.02	0.00	0.00
<i>HLA-DQA1*01:03</i>	1.85	1.21-2.84	0.0047	0.27	0.14	0.14

OR= Odds ratio of a multivariate regression model correcting for population stratification; 95%CI= confident interval of the Odd Ratio; AF Cases=allele frequency in cases; AF Controls =allele frequency in controls ; AF reference dataset = allele frequency calculated based on the number of carriers estimated from all cohorts belonging to a particular geographic region reported in www.allelefrequencies.net; P =multinomial p-value

Table 3 HLA alleles meta-analysis association results

Marker	Direction of effect	OR	95%CI	PV_m	HetPV
HLA-B*52:01	Concordant	2.67	1.63-4.37	9.4x10 ⁻⁵	0.03
HLA-C*12:02	Concordant	2.31	1.41-3.75	0.0008	0.01
HLA-DQA1*01:03	Concordant	1.75	1.24-2.45	0.0013	0.66

OR= Odds ratio of a multivariate regression model correcting for population stratification; 95%CI= confidence interval of the Odds Ratio; AF Cases=allele frequency in cases; AF Controls = allele frequency in controls; AF reference dataset = allele frequency calculated based on the number of carriers estimated from all cohorts belonging to a particular geographic region reported in www.allelefrequencies.net; PV=multinomial p value; PV_m=meta-analysis pvalue; HetPV= Heterogeneity p value

Table 4 Frequency of NAT2 alleles in case and controls of European and Indian cohorts

	Europeans (70 cases)							Indians (55 cases)					Meta-analysis				
Allele	Control freq	Case freq	INH only	INH comb	OR	95%CI	P	Control freq	Case freq	OR	95%CI	P	Effect dir	OR	95% CI	P _m	Het P
<i>NAT2*4</i>	0.23	0.24	0.29	0.14	1.06	0.71-1.58	0.75	0.22	0.16	0.66	0.39-1.11	0.12	N/P	0.77	0.61-0.98	0.81	0.1
<i>NAT2*5</i>	0.46	0.37	0.34	0.41	0.7	0.49-0.98	0.04	0.33	0.25	0.68	0.43-1.07	0.1	C	0.69	0.57-0.83	0.01	0.43
<i>NAT2*6</i>	0.29	0.35	0.31	0.41	1.3	0.91-1.85	0.14	0.37	0.5	1.77	1.18-2.65	0.01	C	1.42	0.97-2.08	0.03	0.05
<i>NAT2*7</i>	0.02	0.04	0.04	0.04	1.88	0.81-4.34	0.14	0.08	0.08	1.09	0.55-2.17	0.79	C	1.21	0.61-2.38	0.14	0.82

Freq=allele frequency; INH=Isoniazid; INH comb = INH combination; P=multinomial p value; OR= odds ratio of a multivariate regression model correcting for population stratification; P_m =meta-analysis p value; HetP= Heterogeneity p value. Effect dir=effect direction; N/P=null/positive; C=concordant

Table 5 Frequency of NAT2 acetylator status genotypes in European and Indian case/control cohorts and their association in a multivariate regression model

	Europeans (70 cases)					Indians (55 cases)					Meta-analysis				
Acetylator status	Control freq	Case freq	OR	95%CI	P	Control freq	Case freq	OR	95%CI	P	Effect dir	OR	95% CI	P _m	Het P
Rapid	0.05	0.06	1.18	0.44-3.28	0.73	0.06	0.02	0.29	0.03-2.20	0.23	D	0.38	0.22-0.65	0.1	0.75
Intermediate	0.35	0.36	1.02	0.62-1.66	0.93	0.34	0.29	0.8	0.43-1.45	0.45	D	0.9	0.64-1.26	0.95	0.21
Slow	0.49	0.4	0.6	0.41-1.08	0.1	0.41	0.4	0.93	0.53-1.64	0.82	C	0.68	0.52-0.88	0.87	0.46
Ultra Slow	0.1	0.19	2.03	1.10-3.72	0.02	0.19	0.29	1.78	0.98-3.39	0.06	C	1.89	0.84-4.22	0.004	0.3

We considered (a) in rapid group *4/*4 genotypes; (b) in intermediate group *4/*5,*4/*6,*4/*7 genotypes; (c) in slow group *5/*5,*5/*6,*5/*7 genotypes; (d) in ultra slow *6/*6,*7/*6,*7/*7 genotypes.

Freq=allele frequency; P=multinomial p value; OR= odds ratio of a multivariate regression model correcting for population stratification; P_m=meta-analysis p value; HetP= Heterogeneity p value. Effect dir=effect direction; D=discordant; C=concordant

Figure legends

Figure 1 Schematic summary of the study including details on the number of cases and controls and analysis steps undertaken.

Figure 2 Manhattan plots displaying the association results of the genetic association analyses. (A) Manhattan plot showing the summary statistics for the European GWAS; (B) Manhattan plot showing the summary statistics for the Indian GWAS (C) Manhattan plot showing the summary statistics for the meta analysis. The overall significance level was set at 5×10^{-8} . SNPs in green have a significance level less than 5×10^{-6} .

SUPPLEMENTAL INFORMATION

Supplemental Material

Genome-wide genotyping of 55
Indian anti-TB DILI cases & 1199
Indian controls

Genome-wide genotyping on 70
European anti-TB DILI cases & 10397
population controls

After imputation of additional SNPs, assess association of individual SNPs with DILI by logistic regression, using additive model and adjusting for population structure. Perform separately in both cohorts followed by meta analysis.

Impute HLA genotypes from genome-wide genotyping data for cases and controls. 217 alleles were imputed in Europeans and 192 in Indians. Assess association of individual alleles with DILI. Set p for significance at less than 2.5×10^{-4} . Perform separately in both cohorts followed by meta analysis.

Extract NAT2 genotypes from imputed GWAS data. Assess association of individual SNPs and acetylator genotypes with DILI in cases and controls. Perform separately in both cohorts followed by meta analysis.

Extract individual SNPs in additional candidate genes from imputed GWAS data. Assess association of SNPs individually with DILI in cases and controls. Set significance of $p < 0.003$ for 16 multiple comparisons. Perform separately in both cohorts.

