Published in final edited form as:

Nat Genet. 2014 October ; 46(10): 1131–1134. doi:10.1038/ng.3093.

HLA-DQA1-HLA-DRB1 variants confer susceptibility to pancreatitis induced by thiopurine immunosuppressants

A full list of authors and affiliations appears at the end of the article.

These authors contributed equally to this work.

Abstract

Pancreatitis occurs in approximately 4% of patients treated with the thiopurines azathioprine or mercaptopurine. Its development is unpredictable and almost always leads to drug withdrawal. We identified inflammatory bowel disease patients who had developed pancreatitis within three months of starting these drugs from 168 sites around the world. After detailed case adjudication, we performed a genome wide association study on 172 cases, and 2035 inflammatory bowel disease controls. Strong evidence of association was identified within the class II HLA region with the most significant association at rs2647087 (Odds ratio 2.59, 95% CI 2.07 – 3.26 $P=2 \times 10^{-16}$). These findings were replicated in an independent set of 78 cases and 472 IBD controls matched for drug exposure. Fine mapping of the HLA region identified association with the HLA-DQA1*02:01-HLA-DRB1*07:01 haplotype. Patients heterozygous at rs2647087 have a 9% risk of developing pancreatitis after administration of a thiopurine, while homozygous patients have a 17% risk.

Almost one million prescriptions for the thiopurines azathioprine and mercaptopurine were dispensed in England alone in 2011¹. Mercaptopurine (MP) and its prodrug azathioprine (AZA) are the most commonly prescribed immunosuppressive agents used to maintain corticosteroid-free remission and to prevent postoperative recurrence in patients with inflammatory bowel disease (IBD)^{2–4}. Outside of gastroenterology they are widely used as anti-rejection agents after solid organ transplantation and as steroid sparing agents for conditions such as rheumatoid arthritis. Despite their widespread use, it has been estimated

Correspondence should be addressed to Dr Graham Heap (grahamheap@nhs.net) or Dr Tariq Ahmad (tariq.ahmad1@nhs.net).

⁴⁶A complete list of all consortia members is displayed in the supplementary information

Author Contributions

G.A.H, K.S, T.K.D, D.A.vH, P.C.D, Ar.H and T.A participated in the conception, design and coordination of the study. C.M.B was the project manager and coordinated recruitment. G.A.H, M.W and T.A analysed the data and participated in writing the paper. K.P. performed DNA preparation. The International Serious Adverse Events Consortium advised on the GWA study, performed genotyping and participated in writing the paper. A.C, T.C, P.D, E.G, P.M.I, J.O.L, J.M, Z.M, T.R.O, D.R, G.A.H, A.S, and T.A performed adjudication of the cases. M.C, J.B.S. and M.B. performed the SSP based HLA genotyping. J.R. and J.P.V. performed the *in silico* structural analysis. K.S, J.M.A, V.A, P.B, S.B, A.C, S.J.C, T.C, F.R.C, M.D'A, T.K.D, R.N.F, T.H.F, D.R.G, E.G, J.H, Al.H, P.M.I, G.J, A.K, I.C.L, J.C.L, C.L, R.L-T, J.O.L, J.Man, J.Maw, Z.M, M.P, T.R.O, G.R-S, R.K.R, D.R, J.S, M.S.S, G.C.S, M.T, E.V.T, A.W, G.W, R.K.W, S.Z contributed a substantial number of samples in line with ICMJE criteria. All authors assisted in the writing, reviewing and approval of the manuscript.

URLs

qqman R Script available from <https://raw.githubusercontent.com/stephenturner/qqman/master/qqman.r>

For raw data access please visit <https://dataportal.saeconsortium.org/>

European Genome-Phenome Archive <https://www.ebi.ac.uk/ega/home>

Competing Interests Statement

The International Serious Adverse Events Consortium funded this study. The authors declare no competing financial interests.

that 17% of patients taking these medications develop side effects that necessitate drug withdrawal 5. Acute pancreatitis after thiopurine therapy is a well recognised, idiosyncratic, dose-independent adverse drug reaction with an incidence of approximately 4 to 7% in patients with IBD 5,6. The pathogenesis of thiopurine-induced pancreatitis is unknown and the clinical picture is poorly described. Development of pancreatitis, which can be life threatening, precludes the patient from continuing on conventional thiopurine-based therapy and necessitates the use of other agents, which may be less effective or more costly.

Recent studies have confirmed that some rare adverse responses to drug therapy are associated with clinically useful genetic variants that can be identified using a small number of rigorously characterised cases by genome wide association study methodologies. For example the HLA class I allele HLA-B*57:01 has been shown to be a major determinant of flucloxacillin-induced cholestatic hepatitis with an odds ratio of 80, using a cohort of only 51 patients 7. This same HLA allele was earlier found to be associated with abacavir hypersensitivity. In Europe, the USA and Australia HLA-B*57:01 testing is now mandatory before prescribing abacavir 8. We aimed to characterise the clinical features of thiopurine-induced pancreatitis and identify genetic markers that might predict the development of this serious adverse drug reaction.

In total we recruited 441 patients to the study in two recruitment rounds, the initial GWAS round and a case-control replication cohort. Eight individuals submitted to the study failed to meet the eligibility criteria and were removed leaving 433 patients in the combined cohort. Each case was reviewed by an expert panel of gastroenterologists and a causal link between AZA/MP and pancreatitis development assessed using a modified version of the Liverpool Causality Assessment Tool 9. Definite cases required the development of recurrent pancreatitis upon thiopurine re-challenge. Cases classified as probable demonstrated a clear temporal relationship with thiopurine administration, with no other identifiable risk factors for pancreatitis, including the concomitant use of other drugs recognised as causing pancreatitis. Only definite and probable cases were taken forward for analyses. Details of case recruitment, the adjudication process and results are available in the Methods section and Supplementary Information. The number of cases in each adjudication group is displayed in Supplementary Table 1. A summary of the clinical characteristics of the recruited patients is shown in Supplementary Table 2.

We first conducted a genome wide association study with 217 of the 248 patients with thiopurine-induced pancreatitis and 2035 previously genotyped control Crohn's disease and ulcerative colitis cases 10,11, matched for disease (Crohn's disease or Ulcerative Colitis) but unselected for thiopurine exposure or pancreatitis development. Of the initial 217 patients, we restricted the analyses to 177 patients of European descent. After genotyping quality control procedures (see Methods) 172 cases remained. In order to account for the different genotyping platforms used within the control cohort and the case cohort, we first performed an imputation analysis and used genotype dosages from all individuals (0% missingness) for the analysis. The case cohort and the Crohn's disease and ulcerative colitis control cohorts were imputed separately to the 1000 Genomes European phase1 version 3 (20101123) panel using minimac 12,13. We restricted our analysis to the 2,819,700 SNPs that had an imputation $R^2 > 0.95$ across all three cohorts. The genomic inflation factor was 1.03.

We identified an association with pancreatitis development at rs2647047 (Odds Ratio 2.26, $P=1.9 \times 10^{-12}$) (Supplementary Table 3). As the association appeared to be within the Class II HLA region, we performed dedicated HLA imputation using SNP2HLA into the T1DGC reference panel of 5,224 individuals that have had classical HLA alleles typed as well as SNPs and indels by immunochip 14. This enabled us to refine the identified association with thiopurine-induced pancreatitis to the SNPs rs2647087 and rs7745656 within the Class II HLA region (Odds Ratio 2.59, 95% CI 2.07 – 3.26, $P=2 \times 10^{-16}$) (Figure 1). This association was robust to principal component correction (Odds ratio 2.62, $p=2 \times 10^{-15}$ after adjustment for 20 PCs). The HLA alleles HLA-DQA1*02:01 and HLA-DRB1*07:01 were also significantly associated with pancreatitis development and were partially tagged by rs2647087 (R^2 0.49). Significant HLA allele associations are displayed in Table 1, while complete results are displayed in Supplementary Table 4. In order to validate the HLA imputation, we performed classical HLA SSP based PCR genotyping in 160 case samples. The resulting correlation ($R^2 = 0.93$ for all HLA-DRB1 alleles) was in line with published observations 14. There was a 99% concordance between the imputed and directly genotyped HLA-DRB1*07:01 alleles. Full results for imputed and SSP PCR based genotyping are shown in Supplementary Table 5. Including rs2647087 as a covariate removed all the associations in the HLA regions indicating that there is only one HLA association signal (Supplementary Figure 1). rs2647087 remained borderline significant after conditioning the association on HLA-DRB1*07:01 ($P=2.0 \times 10^{-5}$). This disparity, together with the low R^2 between rs2647087 and HLA-DRB1*07:01 indicates that the causal variant may be an unidentified rare allele that is being tagged by these two alleles. The well described association of HLA-B*57:01 with Abacavir hypersensitivity requires only one copy of the allele for the development of hypersensitivity. In this study, however, there was no evidence of a dominant effect ($P=0.37$).

Including only the definite cases ($n=37$) increased the odds ratio marginally (Odds Ratio 3.18), but due to the much smaller sample size ($n = 37$ European cases) the association did not achieve genome-wide significance (HLA-DQA1*02:01-HLA-DRB1*07:01 $P=5.88 \times 10^{-8}$). To investigate if smoking, a recognised risk factor for pancreatitis, was associated with rs2647087 genotype, the cases were subdivided by smoking status at the time of their pancreatitis. rs2647087 had an odds ratio of 2.99 (95% CI 2.18 – 4.11) in patients who were smoking at the time of their pancreatitis and 2.19 (95% CI 1.58 – 2.96) in patients who were not ($p=0.17$), providing no evidence for an interaction between smoking and genotype. rs2647087 remained the most significantly associated marker after including smoking as a covariate in the association test (Odds Ratio 2.50 $P=3.4 \times 10^{-13}$). Thiopurine methyltransferase (*TPMT*) genotype (TPMT*3A, *3C, *2, *4 & *8) was not associated with pancreatitis development at any of the five most common loss of function loci ($P=0.99$) that predispose patients to bone marrow suppression with thiopurine treatment.

We next attempted to replicate our finding in an independent study of 78 cases and 472 IBD controls treated with thiopurines for at least 12 months without the development of pancreatitis. From the four variants chosen for replication, all of which are in complete linkage disequilibrium with rs2647087, the SNP rs6935723 had the highest genotype success rate (98.2%) and demonstrated robust replication of the association, Odds Ratio 2.21, $P=4 \times 10^{-6}$ (Supplementary Tables 6 and 7). There was insufficient statistical power

within the replication cohort to attempt to replicate the putative associations observed on chromosomes 3 and 10. The combined results of the GWAS and replication cohort are displayed in Table 2.

HLA-DRB1*07:01 has previously been described to be associated with ileal Crohn's disease with a relative risk of 1.6 15. To ensure the association was not due to a disparity in the anatomical location of Crohn's disease between cases and controls (i.e. ileal vs. colonic disease), the association was performed with disease location as a covariate in the regression analysis (rs2647087, Odds Ratio 2.71, 95% CI 2.07 – 3.26, $P=8.67 \times 10^{-13}$). In addition we performed the association test using only cases and controls with ulcerative colitis. This yielded consistent evidence of association despite the smaller number of cases (rs6935723, $n=38$, Odds Ratio 2.23, 95% CI 1.43 – 3.49, $P=0.0003$). A shared genetic association between ileal Crohn's disease and thiopurine-induced pancreatitis might direct clinicians to other therapies including surgery for patients with this pattern of disease.

The prevalence of pancreatitis in patients treated with thiopurines is approximately 4-7%. Our results suggest that heterozygotes at rs2647087 will be approximately 2.5 times more likely to get pancreatitis and risk allele homozygotes approximately 5 times more likely to get pancreatitis than common allele homozygotes. In a clinical setting this means that for every 1000 patients tested for rs2647087, 77 risk allele homozygotes will be identified and these individuals will an approximate risk of 17% for the development of pancreatitis (based on a relative risk of 4.31, see Methods). Patients heterozygous at rs2647087 would have a 9% risk of developing pancreatitis. If AZA/MP are subsequently avoided in all homozygote risk allele individuals, this equates to an overall number needed to test of 76 patients to prevent 1 case of pancreatitis.

Pancreatitis was first described in renal transplant recipients treated with AZA in the early 1960s 16. It wasn't until 1972, however, that AZA was implicated as the causal agent when a patient with Crohn's disease was re-challenged with AZA and developed pancreatitis for a second time 17. Candidate gene case-control studies conducted prior to this study have suggested an association between a polymorphism in *ITPA* 18 and thiopurine-induced pancreatitis, however, we were unable to replicate this association (rs1127354, $P=0.177$).

To date six pharmacogenetic genome wide association studies for adverse drug reactions have demonstrated association with the class II HLA region 19. The mechanism for these associations has not yet been elucidated, however, the clinical features described here, with an average time from starting the drug to development of pancreatitis of 23.8 days (95% CI 21.2 – 26.4), would be consistent with a delayed immunological or T cell mediated reaction. The extended HLA DRB1*07-DQA1*02 haplotype identified here has previously been described as associated with drug-induced liver injury to both the tyrosine kinase inhibitor Lapatinib and the anticoagulant Ximelagatran although the mechanism of action is unknown 20,21.

Recent work investigating the HLA-B*57:01 association with abacavir has described how the molecule interacts with the HLA binding pocket to alter the antigen presentation repertoire. This alteration can result in novel "self" antigen presentation or novel

presentation of constitutive self-peptides 22. Class II HLA associations are less well established, but Ximelagatran has been demonstrated to competitively bind preferentially to the HLA DRB1*07:01 haplotype 20. We have constructed a computational docking model (Supplementary Figure 2), which predicts that MP is able to bind the HLA DRB1*07:01 molecule; the effects of this on T cell activation, however, are at this stage unknown.

We have undertaken the first large scale clinical and genetic analyses of thiopurine-induced pancreatitis and identified an association with a common variant in the Class II HLA region. This is the first description of any genetic association with a drug-induced pancreatitis. Although in this study the indication for thiopurine use was IBD, these results are likely to be generalisable to many other patient populations where AZA/MP are used. Information on the risk of pancreatitis development may form part of a future panel of genetic tests that could aid clinicians and patients when deciding upon treatment options.

Online Methods

Patient Recruitment and selection

Patients were identified and recruited from 115 NHS hospital UK research sites and 53 international sites. Cases were identified through electronic searches of patient databases, pathology services, directly through gastroenterology clinics and by direct advertising to patients. The protocol was approved by the National Research Ethics Committee South West, Exeter (11/SW/0222) and by all local research and development offices.

Inclusion criteria for patient recruitment included all of the following:

- a) Onset of acute severe abdominal pain within three months of starting mercaptopurine or azathioprine treatment for ulcerative colitis or Crohn's disease.
- b) Greater than or equal to a two-fold rise in amylase or lipase as defined by the research site's local laboratory.
- c) Medical opinion implicating thiopurine therapy and subsequent drug withdrawal.

At the patient recruitment visit, a case report form was completed that detailed demographic, clinical and drug history after informed consent was obtained. Two 6ml EDTA blood samples were taken at the same visit for DNA extraction (BD Vacutainer, USA).

To assess patient eligibility, an adjudication panel assessed causality from case report forms using an adapted version of the validated Liverpool Adverse Drug Reaction Causality Assessment Tool displayed in Supplementary Figure 3 9. All cases were reviewed independently by four medically trained experts and assigned a causality category based on the assessment tool. Confounding medications or conditions (listed below) classified the patient as a possible cause of thiopurine-induced pancreatitis:

- Gallstones
- Alcohol

- Hyperlipidaemia (in particular hypertriglyceridaemia)
- Concomitant administration of other medications from Badalov et al 23
- Infection (e.g. Viruses-mumps, coxsackie, hepatitis B, CMV, varicella-zoster, HSV)
- Post-ERCP
- Ischaemia
- Trauma

The collective results from each panel member were collated and the panel discussed discrepant cases before a final adjudication decision was reached. A set of guidelines was drawn up through the course of these expert meetings to ensure consistent decision-making in borderline cases. These guidelines are displayed in the Supplementary Information and help illustrate the robust phenotype assessment undertaken for all cases. Co-administration of a thiopurine with any drug known to cause drug-induced pancreatitis (Class 1a or class 1b as defined by Badalov et al 23) within the three months prior to the development of pancreatitis classified the patient as having had “possible” thiopurine-induced pancreatitis. Only patients classified as “definite” or “probable” cases of thiopurine-induced pancreatitis were taken forward to subsequent clinical and genetic analyses.

DNA Extraction and Genotyping

DNA was extracted from EDTA stabilised blood using the Qiagen Autopure LS with Puregene chemistry. 248 samples were genotyped on the Illumina Infinium HumanCoreExome beadchip (Illumina, USA), which contains 264,909 tagging SNP markers, and 244,593 exome focused markers by the Broad Institute (Boston, USA).

472 IBD patients who had been screened for thiopurine use without development of pancreatitis from the Royal Devon and Exeter NHS Trust were used as a control cohort for replication analysis. These 472 samples, together with the 87 thiopurine induced pancreatitis cases were genotyped at four SNPs (rs2647087, rs6935723, rs2647089 and rs7745656) using the KASP™ genotyping assay by LGC Limited, UK. Analysis was limited to the 78 samples that self identified as “White”. The genotyping success rates were > 97% for all SNPs.

Genetic and Statistical Analysis

The study was open for data collection from March 2012 to December 2013. Multivariate linear regression analysis was used to calculate risk variables. Logistic regression was used for categorical variables. Severe pancreatitis was defined as single organ failure or greater. Smoking status was calculated based on number of pack years (number of packs of 20 cigarettes per day multiplied by years smoked). For study entry a raised amylase or lipase was characterised as a greater than or equal to a two fold increase above the upper limit of normal for the research site’s local laboratory. Relative risks for genotypes were calculated from odds ratios based on the formula 24:

$$\text{Relative risk} = \text{odds ratio} / (1 - \text{baseline risk} + (\text{baseline risk} \times \text{odds ratio}))$$

All statistical analysis was undertaken in R (Version 3.0.2) and Stata (Version 13). Manhattan and QQ plots were created with the qqman R package.

Quality control of SNPs and samples before imputation

Genotyping was performed on 217 cases assigned as having “definite” or “probable” thiopurine-induced pancreatitis using the HumanCoreExome SNP Chip. Genotypes were called using Gencall 25. We excluded SNPs with a HWE $P < 0.0001$ and a genotype success rate < 0.99 . We excluded indels. Exclusion criteria for case samples were genotyping success rate < 0.98 and a heterozygosity rate $> 4SD$ (5 samples removed). We used zCall to improve calling of low frequency variants 26. After running zCall, SNPs were excluded if they had a HWE $P < 0.0001$, MAF < 0.01 or if they were duplicated. This left 254,457 autosomal SNPs for imputation. The control patients with Crohn’s disease and ulcerative colitis were obtained from the International IBD Genetics Consortium as part of the Wellcome Trust Case Control Consortium (WTCCC 1 for Crohn’s disease and WTCCC 2 for Ulcerative Colitis) 27,28. There were 1748 CD control samples genotyped on the Affymetrix 500K SNP chip and 2361 UC samples genotyped on the Affymetrix 6 SNP chip available for this analysis. Preliminary QC had already been performed on the 1748 CD and 2361 UC samples 27,28. From these two control cohorts, we excluded SNPs with a genotyping success rate < 0.99 , MAF < 0.01 and a HWE $P < 0.0001$ – this left 396,255 (Crohn’s disease) and 727,195 (Ulcerative Colitis) autosomal SNPs. To exclude ethnic outliers we performed principal components analysis using GCTA 29. To generate the principal components we used a set of 79,974 SNPs that were imputed with $R^2 > 0.99$ in the cases (see below) and directly genotyped in the two control cohorts (Supplementary Figure 4). 40 cases and 66 control samples were excluded for being $> 4SD$ for the first or second principal components. The case exclusions were consistent with self-reported non-European ancestry. We used KING 30 to test for cryptic relatedness between samples. If a case and control pair of samples had a kinship coefficient > 0.2 we excluded the control sample, otherwise we excluded one of the pair of samples at random if both were in the same cohort. Six control samples were excluded because of relatedness to a pancreatitis case patient. Nine Crohn’s disease control samples were excluded because they were diagnosed with pancreatitis and also included in the case cohort. After exclusions this left 172 “probable” and “definite” cases, 80% of which had Crohn’s disease. To match the ratio of Crohn’s disease to ulcerative colitis patients in the control group to that in the case cohort we used all the 1669 Crohn’s disease patients that passed QC and a random 366 samples from the ulcerative colitis cohort in subsequent association analyses.

Genome-wide and HLA Imputation and association analyses

We used minimac 13 to impute into the European phase1 version 3 (20101123) panel SNPs and indels reference panel. 76% of the 9,412,474 SNPs with MAF $> 1\%$ frequency was imputed at $RSQ > 0.6$ in the cases: 75% in the Crohn’s disease (CD) controls and 82% in the ulcerative colitis (UC) controls. Each of the three case and control cohorts used a different SNP genotyping chip. This has been reported to lead to spurious associations 31. Therefore,

to avoid excessive false positive associations we focused subsequent association analyses on a very conservative subset of 2,819,700 SNPs that had an $RSQ > 0.95$ in all three cohorts. For dedicated imputation of the HLA region we used SNP2HLA 14 and imputed into the T1DGC reference panel of 5,224 individuals that have had classical HLA alleles typed as well as SNPs and indels by the immunochip. 8321 of the 8961 variants in the T1DGC panel were captured with an INFO score > 0.8 . mach2dat 32 was used to perform the association analyses for the genome-wide analysis, and PLINK 33 was used to perform the association analyses for the HLA imputed analysis. A standard test for dominance deviation analysis was performed whereby we tested if the genotype coded 0/1/0 improved upon a regression model which included a term for the genotype coded 0/1/2.

In-silico peptide binding

Computational docking was performed to probe the binding of mercaptopurine to HLA-DRB1*07:01. As a crystal structure of HLA-DRB1*07:01 is not available, the HLA-DRB1*07:01 was homology modelled from the crystal structure of HLA-DRB1*04:01 (PDB code 4MCY 34). For comparison, docking runs were also performed with HLA-DRB1*01:01 (PDB code 4AH2), HLA-DRB1*04:01 (PDB code 4MCY), HLA-DRB3*01:01 (PDB code 2Q6W), HLA-DP2 (PDB code 3LQZ) and HLA-DQ8 (PDB code 2NNA). AutoDock Tools 1.5.6 35 was used to assign hydrogens, Gasteiger charges and rotatable bonds to the compounds. Each docking run was performed in the absence of peptide within the binding cleft and utilised the AutoDock Vina software 36 to search a docking grid that encompassed the entire peptide-binding cleft.

Data access

Phenotype and genotype data for cases is freely available upon request from the iSAEC Data Access Committee for users who comply with the Consortium's Data Release and IP Policy. Data will be available within 12 months of genotype completion. Raw genotype data is freely available to researchers upon request. For further data access details please contact saec@c2b2.columbia.edu.

Genotype data for the WTCCC Ulcerative colitis and Crohn's disease cases are available from the European Genome-Phenome Archive.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Graham A Heap^{#1,2}, Michael N Weedon^{#2}, Claire M Bewshea², Abhey Singh¹, Mian Chen³, Jack B Satchwell³, Julian P Vivian⁴, Kenji So¹, Patrick C Dubois⁵, Jane M Andrews⁶, Vito Annese⁷, Peter Bampton⁸, Martin Barnardo³, Sally Bell⁹, Andy Cole¹⁰, Susan J Connor¹¹, Tom Creed¹², Fraser R Cummings¹³, Mauro D'Amato¹⁴, Tawfique K Daneshmend¹, Richard N Fedorak¹⁵, Timothy H Florin¹⁶, Daniel R Gaya¹⁷, Emma Greig¹⁸, Jonas Halfvarson¹⁹, Alisa Hart²⁰, Peter M Irving²¹, Gareth Jones²², Amir Karban²³, Ian C Lawrance²⁴, James C Lee²⁵, Charlie Lees²², Raffi

Lev-Tzion²⁶, James O Lindsay²⁷, John Mansfield²⁸, Joel Mawdsley²⁹, Zia Mazhar³⁰, Miles Parkes²⁵, Kirstie Parnell², Timothy R Orchard³¹, Graham Radford-Smith^{32,33}, Richard K Russell³⁴, David Reffitt³⁵, Jack Satsangi²², Mark S Silverberg³⁶, Giacomo C Sturniolo³⁷, Mark Tremelling³⁸, Epameinondas V Tsianos³⁹, David A van Heel⁴⁰, Alissa Walsh⁴¹, Gill Watermeyer⁴², Rinse K Weersma⁴³, Sebastian Zeissig⁴⁴, Jamie Rossjohn⁴, Arthur L Holden⁴⁵, International Serious Adverse Events Consortium⁴⁶, IBD Pharmacogenetics Study Group⁴⁶, and Tariq Ahmad^{1,2}

Affiliations

¹IBD Pharmacogenetics, Royal Devon and Exeter Hospital, Exeter, United Kingdom
²Precision Medicine Exeter, University of Exeter, Exeter, United Kingdom ³Oxford Transplant Centre, Oxford University Hospital NHS Trust, Oxford, UK ⁴Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia ⁵King's College Hospital, London, United Kingdom ⁶IBD Service, Department of Gastroenterology & University of Adelaide at Royal Adelaide Hospital, Adelaide, SA, Australia ⁷Division of Gastroenterology, AOU Careggi, Florence, Italy ⁸Flinders Medical Centre, Flinders University of South Australia, Adelaide, SA, Australia ⁹Department of Gastroenterology, St Vincent's Hospital, Fitzroy, VIC, Australia ¹⁰Gastroenterology & Hepatology, Royal Derby Hospital, Derby, United Kingdom ¹¹Department of Gastroenterology & Hepatology, Liverpool Hospital, Sydney, NSW, Australia ¹²Joint Clinical Research Unit, University Hospitals Bristol NHS Foundation Trust, Bristol, United Kingdom ¹³Department of Gastroenterology, University Hospital Southampton NHS Foundation Trust, Southampton, UK ¹⁴Department of Biosciences and Nutrition, Karolinska Institute, Stockholm, Sweden ¹⁵Division of Gastroenterology, University of Alberta, Edmonton, AB, Canada ¹⁶The University of Queensland School of Medicine, South Brisbane, QLD, Australia ¹⁷Gastroenterology Unit, Glasgow Royal Infirmary, Glasgow, United Kingdom ¹⁸Department of Gastroenterology, Taunton and Somerset NHS Foundation Trust, Taunton, United Kingdom ¹⁹Division of Gastroenterology, Örebro University Hospital & School of Health and Medical Sciences, Örebro University, Örebro, Sweden ²⁰Department of Medicine, St Mark's Hospital and Academic Institute, North West London Hospitals NHS Trust, United Kingdom ²¹Department of Gastroenterology, Guy's & St Thomas' NHS Foundation Trust, London, United Kingdom ²²Department of Gastroenterology, Western General Hospital, Edinburgh, United Kingdom ²³Department of Gastroenterology, Rambam Medical Center, Haifa, Israel ²⁴Centre for Inflammatory Bowel Diseases, University of Western Australia, Fremantle Hospital, Fremantle, WA, Australia ²⁵Department of Gastroenterology, Cambridge University Hospitals NHS Trust, Cambridge, United Kingdom ²⁶Paediatric Gastroenterology and Nutrition Unit, Shaare Zedek Medical Centre, Jerusalem, Israel ²⁷Department of Gastroenterology, Barts and The London NHS Trust, London, United Kingdom ²⁸Department of Gastroenterology, Newcastle University Hospitals NHS Trust, Newcastle, United Kingdom ²⁹Department of Gastroenterology, West Middlesex University Hospital NHS Trust, Isleworth, United Kingdom ³⁰Department of

Gastroenterology, Basildon and Thurrock Hospital NHS Trust, Basildon, United Kingdom ³¹Department of Medicine, Imperial College Healthcare NHS, London, United Kingdom ³²Department of Gastroenterology, Royal Brisbane and Women's Hospital, Brisbane, Queensland, Australia ³³IBD Group, Queensland Institute of Medical Research and University of Queensland School of Medicine, Herston Campus, Brisbane, Queensland, Australia ³⁴Department of Paediatric Gastroenterology, Yorkhill Hospital, Glasgow, UK ³⁵Department of Gastroenterology, Lewisham and Greenwich NHS Trust, London, United Kingdom ³⁶Inflammatory Bowel Disease Group, Zane Cohen Centre for Digestive Diseases, Mount Sinai Hospital, Toronto, Ontario, Canada ³⁷Gastroenterologia, Univerita di Padova, Padova, Italy ³⁸Department of Gastroenterology, Norfolk and Norwich Hospital NHS Trust, Norwich, United Kingdom ³⁹1st Division of Internal Medicine and Division of Gastroenterology, Faculty of Medicine, University of Ioannina, Ioannina, Greece ⁴⁰Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK ⁴¹Department of Gastroenterology, St Vincent's Hospital, Sydney, NSW, Australia ⁴²GIT Clinic, Groote Schuur Hospital, Cape Town, South Africa ⁴³Department of Gastroenterology and Hepatology, University Medical Center Groningen and the University of Groningen, Groningen, The Netherlands ⁴⁴Department of Internal Medicine, University Medical Center Schleswig-Holstein, Kiel, Germany ⁴⁵The International Serious Adverse Events Consortium, Chicago, IL, United States

Acknowledgements

The International Serious Adverse Events Consortium (iSAEC) funded the sample collection and genotyping. The UK National Institute for Health Research (NIHR) provided research nurse support to facilitate recruitment at all UK research sites. We would like to thank Crohn's and Colitis UK for funding support and publicising this study to its members. A Wellcome Trust Institutional Strategic Support Award (WT097835MF) generously supported the work in this study. Genotyping was undertaken at the Broad Institute, USA. We would like to thank all the clinicians who assisted with sample collection as part of the IBD Pharmacogenetics Study Group (listed in the Supplementary Information) and Suzie Marriott for her assistance during the trial initiation. We would like to acknowledge The International Serious Adverse Events Scientific Management Committee members (listed in the Supplementary Information), Dr Tim Frayling, Dr Siying Lin & Dr Karen Hunt for kindly providing comments on the draft manuscript as well as Clare Heard and Marian Parkinson for their on-going administrative support to the study. We would also like to thank all the patients for their time and participation.

References

1. Health and Social Care Information Centre. Quarterly Prescription Cost Analysis, England. 2013.
2. Prefontaine E, Lr S, Jk M, Cepoiu M. Azathioprine or 6-mercaptopurine for maintenance of remission in Crohn's disease (Review). Cochrane database Syst Rev. 2010
3. Timmer A, et al. Azathioprine and 6-mercaptopurine for maintenance of remission in ulcerative colitis (Review). Cochrane database Syst Rev. 2012
4. Cosnes J, et al. Early administration of azathioprine vs conventional management of Crohn's Disease: a randomized controlled trial. Gastroenterology. 2013; 145:758–65.e2. [PubMed: 23644079]
5. Chaparro M, et al. Safety of Thiopurine Therapy in Inflammatory Bowel Disease: Long-term Follow-up Study of 3931 Patients. Inflamm Bowel Dis. 2013; 19:1–7. [PubMed: 22532319]

6. Gearry RB, Barclay ML, Burt MJ, Collett JA, Chapman BA. Thiopurine drug adverse effects in a population of New Zealand patients with inflammatory bowel disease. *Pharmacoepidemiol Drug Saf.* 2004; 13:563–7. [PubMed: 15317038]
7. Daly AK, et al. HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. *Nat Genet.* 2009; 41:816–819. [PubMed: 19483685]
8. Mallal S, et al. HLA-B*5701 screening for hypersensitivity to abacavir. *N Engl J Med.* 2008; 358:568–579. [PubMed: 18256392]
9. Gallagher RM, et al. Development and inter-rater reliability of the Liverpool adverse drug reaction causality assessment tool. *PLoS One.* 2011; 6:e28096. [PubMed: 22194808]
10. WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007; 447:661–678. [PubMed: 17554300]
11. Barrett JC, et al. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nat Genet.* 2009; 41:1330–1334. [PubMed: 19915572]
12. Abecasis GR, et al. A map of human genome variation from population-scale sequencing. *Nature.* 2010; 467:1061–73. [PubMed: 20981092]
13. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through prephasing. *Nat Genet.* 2012; 44:955–959. [PubMed: 22820512]
14. Jia X, et al. Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One.* 2013; 8:e64683. [PubMed: 23762245]
15. Ahmad T, et al. The molecular classification of the clinical manifestations of Crohn's disease. *Gastroenterology.* 2002; 122:854–866. [PubMed: 11910336]
16. Tilney NL, Collins JJ, Wilson RE. Hemorrhagic pancreatitis. A fatal complication of renal transplantation. *N Engl J Med.* 1966; 274:1051–7. [PubMed: 5326140]
17. Nogueira JR, Freedman MA. Acute pancreatitis as a complication of Imuran therapy in regional enteritis. *Gastroenterology.* 1972; 62:1040–1. [PubMed: 5029069]
18. Marinaki AM, et al. Adverse drug reactions to azathioprine therapy are associated with polymorphism in the gene encoding inosine triphosphate pyrophosphatase (ITPase). *Pharmacogenetics.* 2004; 14:181–187. [PubMed: 15167706]
19. Illing PT, Vivian JP, Purcell AW, Rossjohn J, McCluskey J. Human leukocyte antigen-associated drug hypersensitivity. *Curr Opin Immunol.* 2013; 25:81–9. [PubMed: 23141566]
20. Kindmark, a; , et al. Genome-wide pharmacogenetic investigation of a hepatic adverse event without clinical signs of immunopathology suggests an underlying immune pathogenesis. *Pharmacogenomics J.* 2008; 8:186–95. [PubMed: 17505501]
21. Spraggs CF, et al. HLA-DQA1*02:01 is a major risk factor for lapatinib-induced hepatotoxicity in women with advanced breast cancer. *J Clin Oncol.* 2011; 29:667–73. [PubMed: 21245432]
22. Illing PT, et al. Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. *Nature.* 2012; 486:554–8. [PubMed: 22722860]
23. Badalov N, et al. Drug-induced acute pancreatitis: an evidence-based review. *Clin Gastroenterol Hepatol.* 2007; 5:648–661. [PubMed: 17395548]
24. Grant RL. Converting an odds ratio to a range of plausible relative risks for better communication of research findings. *BMJ.* 2014; 348:f7450–f7450. [PubMed: 24464277]
25. Oliphant A, Barker DL, Stuelpnagel JR, Chee MS. BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques.* 2002; (Suppl):56–8.
26. Goldstein JI, et al. zCall: a rare variant caller for array-based genotyping: genetics and population analysis. *Bioinformatics.* 2012; 28:2543–5. [PubMed: 22843986]
27. Barrett JC, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet.* 2008; 40:955–962. [PubMed: 18587394]
28. Anderson CA, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet.* 2011; 43:246–252. [PubMed: 21297633]
29. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet.* 2011; 88:76–82. [PubMed: 21167468]

30. Manichaikul A, et al. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010; 26:2867–73. [PubMed: 20926424]
31. Sebastiani P, et al. Retraction. *Science*. 2011; 333:404.
32. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol*. 2010; 34:816–34. [PubMed: 21058334]
33. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81:559–575. [PubMed: 17701901]
34. Scally SW, et al. A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. *J Exp Med*. 2013; 210:2569–82. [PubMed: 24190431]
35. Morris GM, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*. 2009; 30:2785–2791. [PubMed: 19399780]
36. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2010; 31:455–461. [PubMed: 19499576]

Figure 1. Genome-wide association results for all SNPs post 1000 Genomes and HLA imputation.

A Manhattan plot

B Q-Q Plot

Table 1

Top imputed HLA association signals ($P < 5 \times 10^{-8}$).

Risk described for presence of the listed allele

HLA Allele	Position (hg19)	GWAS control risk allele Freq.	GWAS case risk allele Freq.	GWAS OR (95% CI)	GWAS P
HLA-DRB1*07:01	32660042	0.16	0.33	2.55 (2.01, 3.23)	1×10^{-14}
HLA-DQA1*02:01	32716284	0.16	0.33	2.54 (2.00, 3.22)	2×10^{-14}

Table 2

Top SNP association signals from the GWAS analysis ($P < 5 \times 10^{-8}$).

All SNPs located on Chromosome 6.

Stage	SNP	Position (hg19)	Risk Allele	Other Allele	Control risk allele Freq.	GWAS case risk allele Freq.	OR (95% CI)	GWAS P
GWAS	rs7745656	32680970	T	G	0.27	0.49	2.59 (2.07,3.26)	2×10^{-16}
Replication					0.26	0.46	2.22 (1.57, 3.14)	6×10^{-6}
Combined							2.47 (2.05, 2.99)	2×10^{-21}
GWAS	rs2647087	32681049	C	A	0.27	0.49	2.59 (2.06,3.26)	2×10^{-16}
Replication					0.26	0.48	2.37 (1.68, 3.34)	1×10^{-6}
Combined							2.52 (2.09, 3.05)	1×10^{-20}
GWAS	rs6935723	32681669	C	T	0.27	0.49	2.59 (2.07,3.26)	2×10^{-16}
Replication					0.26	0.46	2.21 (1.58, 3.10)	4×10^{-6}
Combined							2.46 (2.04, 2.98)	1×10^{-21}
GWAS	rs2647089	32681568	C	T	0.27	0.49	2.59 (2.07,3.25)	2×10^{-16}
Replication					0.26	0.45	2.17 (1.54, 3.06)	9×10^{-6}
Combined							2.45 (2.03, 2.97)	6×10^{-20}