nature genetics

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

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 patients with inflammatory bowel disease (IBD) who had developed pancreatitis within 3 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 2,035 controls with IBD. We identified strong evidence of association within the class II HLA region, with the most significant association identified at rs2647087 (odds ratio 2.59, 95% confidence interval 2.07–3.26, $P = 2 \times 10^{-16}$). We replicated these findings in an independent set of 78 cases and 472 controls with IBD 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, whereas homozygotes have a 17% risk.

Almost one million prescriptions for the thiopurines azathioprine and mercaptopurine were dispensed in England alone in 2011 (ref. 1). Mercaptopurine and its prodrug azathioprine are the most commonly prescribed immunosuppressive agents used to maintain corticosteroid-free remission and prevent postoperative recurrence in patients with IBD[2–4]. Outside of gastroenterology, they are widely used as antirejection agents after solid organ transplantation and as steroid-sparing agents for conditions such as rheumatoid arthritis. Despite their widespread use, it has been estimated that 17% of patients taking these medications develop side effects that necessitate drug withdrawal[5]. Acute pancreatitis after thiopurine therapy is a well-recognized, idiosyncratic, dose-independent adverse drug reaction with an incidence of ~4–7% 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 characterized cases by genome-wide association study (GWAS) 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 (OR) 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 United States and Australia, HLA-B*57:01 testing is now mandatory before prescribing abacavir[8]. We aimed to characterize 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 so we removed them, leaving 433 patients in the combined cohort. An expert panel of gastroenterologists reviewed each case, and we assessed a causal link between azathioprine or mercaptopurine and pancreatitis development using a modified version of the Liverpool Causality Assessment Tool[9]. We required that individuals classified as definite cases developed recurrent pancreatitis after thiopurine rechallenge. Cases classified as probable must have demonstrated a clear temporal relationship with thiopurine administration, with no other identifiable risk factors for pancreatitis, including the concomitant use of other drugs recognized as causing pancreatitis. We took only definite and probable cases forward for analyses. Details of case recruitment, the adjudication process and results are available in the Online Methods and **Supplementary Note**. 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 GWAS with 217 of the 248 patients with thiopurine-induced pancreatitis and 2,035 previously genotyped control cases with Crohn's disease or ulcerative colitis[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 (Online Methods), 172 cases remained. To account for the different genotyping platforms used within the control and case cohorts, we first performed an imputation analysis and used genotype dosages from all individuals for the analysis. We imputed the case cohort and the Crohn's disease and ulcerative colitis control cohorts separately to the 1000 Genomes European phase 1 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.

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

Received 17 March; accepted 22 August; published online 14 September 2014; doi:10.1038/ng.3093

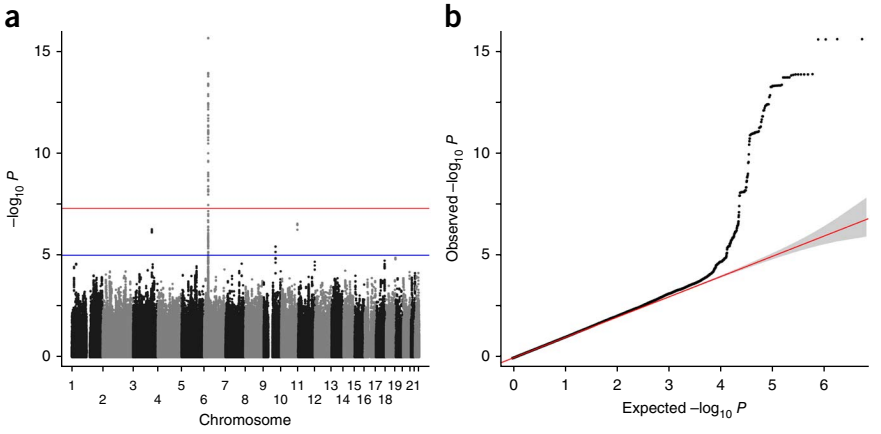

Figure 1 Genome-wide association results for all SNPs after 1000 Genomes and HLA imputation. (**a**) Manhattan plot. Red line, $P = 5 \times 10^{-8}$; blue line, $P = 1 \times 10^{-5}$. (**b**) Quantile-quantile plot.

We identified an association with pancreatitis development at rs2647047 (OR = 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 method enabled us to refine the identified association with thiopurine-induced pancreatitis to the SNP rs2647087 within the class II HLA region (OR = 2.59, 95% CI 2.07–3.26, $P = 2 \times 10^{-16}$) (**Fig. 1**). This association was robust to principal component correction (OR = 2.62, $P = 2 \times 10^{-15}$ after adjustment for 20 principal components). 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 ($r^2 = 0.49$). The significant HLA allele associations are displayed in **Table 1**, and the complete results are displayed in **Supplementary Table 4**. To validate the HLA imputation, we performed classical HLA sequence-specific primer (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 Fig. 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 ($n = 37$) in the analysis increased the OR marginally (OR = 3.18), but because of 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 whether smoking, a recognized risk factor for pancreatitis, was associated with the rs2647087 genotype, we subdivided the cases by smoking status at the time of their pancreatitis. rs2647087 had an OR of 2.99 (95% CI 2.18–4.11) in patients who were smoking at the time of their pancreatitis and an OR of 2.19 (95% CI 1.58–2.96) in patients who were not smoking ($P = 0.17$), 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 (OR = 2.50, $P = 3.4 \times 10^{-13}$). Thiopurine methyltransferase (*TPMT*) genotype (TPMT*3A, TPMT*3C, TPMT*2, TPMT*4 or TPMT*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 controls with IBD 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, rs6935723 had the highest genotype success rate (98.2%) and demonstrated robust replication of the association (OR = 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 been reported to be associated with ileal Crohn's disease[15], with a relative risk of 1.6. To ensure that the association was not due to a disparity in the anatomical location of Crohn's disease between cases and controls (i.e., ileal compared to colonic disease), we performed an association analysis with disease location as a covariate in the regression analysis (rs2647087, OR = 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, $n = 38$, OR = 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 ~4–7%. Our results suggest that individuals heterozygous at rs2647087 will be ~2.5 times more likely to get pancreatitis and individuals homozygous at this risk allele will be ~5 times more likely to get pancreatitis than individuals homozygous for common alleles. In a clinical setting, this means that for every 1,000 patients tested for rs2647087, 77 risk allele homozygotes will be identified, and these individuals will have an approximate risk of 17% for the development of pancreatitis (based on a relative risk of 4.31; Online Methods). Patients heterozygous at rs2647087 would have a 9% risk of developing pancreatitis. If azathioprine and mercaptopurine were

Table 1 Top imputed HLA association signals ($P < 5 \times 10^{-8}$)

HLA allele	Position (hg19)	GWAS control risk allele frequency	GWAS case risk allele frequency	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}

Risk is described for presence of the listed allele.

Table 2 Top SNP association signals from the GWAS analysis ($P < 5 \times 10^{-8}$)

Stage	SNP	Position (hg19)	Risk Allele	Other Allele	Control risk allele frequency	GWAS case risk allele frequency	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}

All SNPs listed are located on chromosome 6.

subsequently avoided in all individuals homozygous at this risk allele, it would equate to an overall need to test of 76 patients to prevent 1 case of pancreatitis.

Pancreatitis was first described in renal transplant recipients treated with azathioprine in the early 1960s[16]. It was not until 1972, however, that azathioprine was implicated as the causal agent when a patient with Crohn's disease was rechallenged with azathioprine and developed pancreatitis for a second time[17]. Candidate gene case-control studies conducted before 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, $P = 0.177$).

So far, six pharmacogenetic GWAS 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 the development of pancreatitis of 23.8 days (95% CI 21.2–26.4 days), would be consistent with a delayed immunological or T cell–mediated reaction. The extended HLA DRB1*07-DQA1*02 haplotype identified here has been described as being associated with drug-induced liver injury resulting from 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 new 'self' antigen presentation or new 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 constructed a computational docking model (**Supplementary Fig. 2**) predicting that mercaptopurine is able to bind the HLA DRB1*07:01 molecule; however, the effects of this binding on T cell activation are unknown at this stage.

We have undertaken large-scale clinical and genetic analyses of thiopurine-induced pancreatitis and identified an association with a common variant in the class II HLA region. To our knowledge, 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 generalizable to many other patient populations in which azathioprine and mercaptopurine 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 on treatment options.

URLs. qqman R Script, http://cran.r-project.org/web/packages/qqman/index.html; International Serious Adverse Event Consortium Data Access Site, https://dataportal.saeconsortium.org/; European Genome-Phenome Archive, https://www.ebi.ac.uk/ega/home.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

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 thank Crohn's and Colitis UK for funding support and publicizing 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, Cambridge, Massachusetts, USA. We thank all the clinicians who assisted with sample collection as part of the IBD Pharmacogenetics Study Group (listed in the **Supplementary Note**) and the International IBD Genetics Consortium, as well as S. Marriott for her assistance during the trial initiation. We acknowledge The International Serious Adverse Events Scientific Management Committee members (listed in the **Supplementary Note**), T. Frayling, S. Lin and K. Hunt for kindly providing comments on the draft manuscript, as well as C. Heard and M. Parkinson for their ongoing administrative support to the study. We also thank all the patients for their time and participation.

AUTHOR CONTRIBUTIONS

G.A.H., K.S., T.K.D., D.A.v.H., P.C.D., A.L.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.N.W. and T.A. analyzed the data and participated in writing the paper. K.P. performed DNA preparation. The International Serious Adverse Events Consortium advised on the GWAS, performed genotyping and participated in writing the paper. A.C., T.C., P.C.D., E.G., P.M.I., J.O.L., J. Mawdsley, 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., T.K.D., R.N.F., T.H.F., D.R.G., E.G., J.H., A.H., P.M.I., G.J., A.K., I.C.L., J.C.L., C.L., R.L.-T., J.O.L., J. Mansfield, J. Mawdsley, 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. and S.Z. contributed a substantial number of samples in line with International Committee of Medical Journal Editors criteria. All authors assisted in the writing, reviewing and approval of the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Health and Social Care Information Centre. *Quarterly Prescription Cost Analysis, England* (UK Government Statistical Service, London, 2013).
2. Prefontaine, E., Sutherland, L.R., Macdonald, J.K. & Cepoiu, M. Azathioprine or 6-mercaptopurine for maintenance of remission in Crohn's disease. *Cochrane Database Syst. Rev.* **1**, CD000067 (2009).
3. Timmer, A., MacDonald, J.W., Tsoulis, D.J. & Macdonald, J.K. Azathioprine and 6-mercaptopurine for maintenance of remission in ulcerative colitis. *Cochrane Database Syst. Rev.* **9**, CD000478 (2012).
4. Cosnes, J. *et al.* Early administration of azathioprine vs conventional management of Crohn's disease: a randomized controlled trial. *Gastroenterology* **145**, 758–765 (2013).
5. Chaparro, M. *et al.* Safety of thiopurine therapy in inflammatory bowel disease: long-term follow-up study of 3931 patients. *Inflamm. Bowel Dis.* **19**, 1404–1410 (2013).
6. Gearry, R.B., Barclay, M.L., Burt, M.J., Collett, J.A. & Chapman, B.A. Thiopurine drug adverse effects in a population of New Zealand patients with inflammatory bowel disease. *Pharmacoepidemiol. Drug Saf.* **13**, 563–567 (2004).
7. Daly, A.K. *et al. HLA-B*5701* genotype is a major determinant of drug-induced liver injury due to flucloxacillin. *Nat. Genet.* **41**, 816–819 (2009).
8. Mallal, S. *et al. HLA-B*5701* screening for hypersensitivity to abacavir. *N. Engl. J. Med.* **358**, 568–579 (2008).
9. Gallagher, R.M. *et al.* Development and inter-rater reliability of the Liverpool adverse drug reaction causality assessment tool. *PLoS ONE* **6**, e28096 (2011).
10. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
11. Barrett, J.C. *et al.* Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the *HNF4A* region. *Nat. Genet.* **41**, 1330–1334 (2009).
12. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
13. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G.R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* **44**, 955–959 (2012).
14. Jia, X. *et al.* Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS ONE* **8**, e64683 (2013).
15. Ahmad, T. *et al.* The molecular classification of the clinical manifestations of Crohn's disease. *Gastroenterology* **122**, 854–866 (2002).
16. Tilney, N.L., Collins, J.J. & Wilson, R.E. Hemorrhagic pancreatitis. A fatal complication of renal transplantation. *N. Engl. J. Med.* **274**, 1051–1057 (1966).
17. Nogueira, J.R. & Freedman, M.A. Acute pancreatitis as a complication of Imuran therapy in regional enteritis. *Gastroenterology* **62**, 1040–1041 (1972).
18. Marinaki, A.M. *et al.* Adverse drug reactions to azathioprine therapy are associated with polymorphism in the gene encoding inosine triphosphate pyrophosphatase (ITPase). *Pharmacogenetics* **14**, 181–187 (2004).
19. Illing, P.T., Vivian, J.P., Purcell, A.W., Rossjohn, J. & McCluskey, J. Human leukocyte antigen–associated drug hypersensitivity. *Curr. Opin. Immunol.* **25**, 81–89 (2013).
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.* **8**, 186–195 (2008).
21. Spraggs, C.F. *et al. HLA-DQA1*02:01* is a major risk factor for lapatinib-induced hepatotoxicity in women with advanced breast cancer. *J. Clin. Oncol.* **29**, 667–673 (2011).
22. Illing, P.T. *et al.* Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. *Nature* **486**, 554–558 (2012).

Graham A Heap[1,2,47], Michael N Weedon[2,47], Claire M Bewshea[1,2], Abhey Singh[1], Mian Chen[3], Jack B Satchwell[3], Julian P Vivian[4], Kenji So[1], Patrick C Dubois[5], Jane M Andrews[6], Vito Annese[7], Peter Bampton[8], Martin Barnardo[3], Sally Bell[9], Andy Cole[10], Susan J Connor[11], Tom Creed[12], Fraser R Cummings[13], Mauro D'Amato[14], Tawfique K Daneshmend[1], Richard N Fedorak[15], Timothy H Florin[16], Daniel R Gaya[17], Emma Greig[18], Jonas Halfvarson[19], Alisa Hart[20], Peter M Irving[21], Gareth Jones[22], Amir Karban[23], Ian C Lawrance[24], James C Lee[25], Charlie Lees[22], Raffi Lev-Tzion[26], James O Lindsay[27], John Mansfield[28], Joel Mawdsley[29], Zia Mazhar[30], Miles Parkes[25], Kirstie Parnell[2], Timothy R Orchard[31], Graham Radford-Smith[32,33], Richard K Russell[34], David Reffitt[35], Jack Satsangi[22], Mark S Silverberg[36], Giacomo C Sturniolo[37], Mark Tremelling[38], Epameinondas V Tsianos[39], David A van Heel[40], Alissa Walsh[41], Gill Watermeyer[42], Rinse K Weersma[43], Sebastian Zeissig[44], Jamie Rossjohn[4], Arthur L Holden[45], International Serious Adverse Events Consortium[46], IBD Pharmacogenetics Study Group[46] & Tariq Ahmad[1,2]

[1]IBD Pharmacogenetics, Royal Devon and Exeter Hospital, Exeter, UK. [2]Precision Medicine Exeter, University of Exeter, Exeter, UK. [3]Oxford Transplant Centre, Oxford University Hospital National Health Service (NHS) Trust, Oxford, UK. [4]Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria, Australia. [5]Department of Gastroenterology, King's College Hospital, London, UK. [6]IBD Service, Department of Gastroenterology and University of Adelaide at Royal Adelaide Hospital, Adelaide, South Australia, Australia. [7]Division of Gastroenterology, Azienda Ospedaliero Universitaria Careggi, Florence, Italy. [8]Flinders Medical Centre, Flinders University of South Australia, Adelaide, South Australia, Australia. [9]Department of Gastroenterology, St. Vincent's Hospital, Fitzroy, Victoria, Australia. [10]Gastroenterology and Hepatology, Royal Derby Hospital, Derby, UK. [11]Department of Gastroenterology and Hepatology, Liverpool Hospital, Sydney, New South Wales, Australia. [12]Joint Clinical Research Unit, University Hospitals Bristol NHS Foundation Trust, Bristol, UK. [13]Department of Gastroenterology, University Hospital Southampton NHS Foundation Trust, Southampton, UK. [14]Department of Biosciences and Nutrition, Karolinska Institute, Stockholm, Sweden. [15]Division of Gastroenterology, University of Alberta, Edmonton, Alberta, Canada. [16]The University of Queensland School of Medicine, South Brisbane, Queensland, Australia. [17]Gastroenterology Unit, Glasgow Royal Infirmary, Glasgow, UK. [18]Department of Gastroenterology, Taunton and Somerset NHS Foundation Trust, Taunton, UK. [19]Division of Gastroenterology, Örebro University Hospital and School of Health and Medical Sciences, Örebro University, Örebro, Sweden. [20]Department of Medicine, St. Mark's Hospital and Academic Institute, North West London Hospitals NHS Trust, London, UK. [21]Department of Gastroenterology, Guy's and St. Thomas' NHS Foundation Trust, London, UK. [22]Department of Gastroenterology, Western General Hospital, Edinburgh, UK. [23]Department of Gastroenterology, Rambam Medical Center, Haifa, Israel. [24]Centre for Inflammatory Bowel Diseases, University of Western Australia, Fremantle Hospital, Fremantle, Western Australia, Australia. [25]Department of Gastroenterology, Cambridge University Hospitals NHS Trust, Cambridge, UK. [26]Paediatric Gastroenterology and Nutrition Unit, Shaare Zedek Medical Centre, Jerusalem, Israel. [27]Department of Gastroenterology, Barts and The London NHS Trust, London, UK. [28]Department of Gastroenterology, Newcastle University Hospitals NHS Trust, Newcastle, UK. [29]Department of Gastroenterology, West Middlesex University Hospital NHS Trust, Isleworth, UK. [30]Department of Gastroenterology, Basildon and Thurrock Hospital NHS Trust, Basildon, UK. [31]Department of Medicine, Imperial College Healthcare NHS, London, UK. [32]Department of Gastroenterology, Royal Brisbane and Women's Hospital, Brisbane, Queensland, Australia. [33]IBD Group, Queensland Institute of Medical Research and University of Queensland School of Medicine, Herston Campus, Brisbane, Queensland, Australia. [34]Department of Paediatric Gastroenterology, Yorkhill Hospital, Glasgow, UK. [35]Department of Gastroenterology, Lewisham and Greenwich NHS Trust, London, UK. [36]Inflammatory Bowel Disease Group, Zane Cohen Centre for Digestive Diseases, Mount Sinai Hospital, Toronto, Ontario, Canada. [37]Gastroenterologia, Univerita di Padova, Padova, Italy. [38]Department of Gastroenterology, Norfolk and Norwich Hospital NHS Trust, Norwich, UK. [39]1st Division of Internal Medicine and Division of Gastroenterology, Faculty of Medicine, University of Ioannina, Ioannina, Greece. [40]Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [41]Department of Gastroenterology, St. Vincent's Hospital, Sydney, New South Wales, Australia. [42]Gastrointestinal Clinic, Groote Schuur Hospital, Cape Town, South Africa. [43]Department of Gastroenterology and Hepatology, University Medical Center Groningen and the University of Groningen, Groningen, the Netherlands. [44]Department of Internal Medicine, University Medical Center Schleswig-Holstein, Kiel, Germany. [45]The International Serious Adverse Events Consortium, Chicago, Illinois, USA. [46]A complete list of all consortia members is provided in the **Supplementary Note**. [47]These authors contributed equally to this work. Correspondence should be addressed to G.A.H. (grahamheap@nhs.net) or T.A. (tariq.ahmad1@nhs.net).

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 study 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: (i) onset of acute severe abdominal pain within 3 months of starting mercaptopurine or azathioprine treatment for ulcerative colitis or Crohn's disease; (ii) greater than or equal to a twofold rise in amylase or lipase as defined by the research site's local laboratory; and (iii) 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 6-ml 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[9] (**Supplementary Fig. 3**). All cases were reviewed independently by four medically trained experts and assigned a causality category on the basis of the assessment tool. Confounding medications or conditions that classified the patient as a possible cause of thiopurine-induced pancreatitis included the following: (i) gallstones; (ii) alcohol; (iii) hyperlipidemia (in particular, hypertriglyceridemia); (iv) concomitant administration of other medications from Badalov *et al.*[23]; (v) infection (for example, viruses such as mumps, coxsackie, hepatitis B, cytomegalovirus, varicella zoster or herpes simplex virus); (vi) post–endoscopic retrograde cholangiopancreatography; (vii) ischemia; and (viii) 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 Note** and help illustrate the robust phenotype assessment undertaken for all cases. Coadministration 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 3 months before 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-stabilized blood using the Qiagen Autopure LS with Puregene chemistry. A total of 248 samples were genotyped on the Illumina Infinium HumanCoreExome beadchip (Illumina), which contains 264,909 tagging SNP markers and 244,593 exome-focused markers, by the Broad Institute (Cambridge, Massachusetts, USA).

A total of 472 patients with IBD 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. Analysis was limited to the 78 samples that self identified their ethnicity 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 on the basis of 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 characterized as a greater than or equal to a twofold increase above the upper limit of normal for the research site's local laboratory. Relative risks for genotypes were calculated from ORs based on the formula[24] relative risk = OR/(1 − baseline risk + (baseline risk × OR)).

All statistical analyses were undertaken in R (Version 3.0.2) and Stata (Version 13). Manhattan and quantile-quantile 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 Hardy-Weinberg equilibrium (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 >4 s.d. (five samples removed). We used zCall to improve the calling of low frequency variants[26]. After running zCall, SNPs were excluded if they had HWE $P < 0.0001$, minor allele frequency (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 1,748 Crohn's disease control samples genotyped on the Affymetrix 500K SNP chip and 2,361 ulcerative colitis samples genotyped on the Affymetrix 6 SNP chip available for this analysis. Preliminary quality control had already been performed on the 1,748 Crohn's disease and 2,361 ulcerative colitis samples[27,28]. From these two control cohorts, we excluded SNPs with a genotyping success rate <0.99, MAF < 0.01 and 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 Fig. 4**). We excluded 40 cases and 66 control samples for being >4 s.d. 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. Nine Crohn's disease control samples were excluded because they were diagnosed with pancreatitis and were 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 1,669 Crohn's disease patients that passed quality control 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 phase 1 version 3 (20101123) SNPs and indels reference panel. 76% of the 9,412,474 variants with MAF > 1% were imputed at $R^2 > 0.6$ in the cases: 75% in the Crohn's disease controls and 82% in the ulcerative colitis 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 $R^2 > 0.95$ 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. In total, 8,321 of the 8,961 variants in the T1DGC panel were captured with an INFO score >0.8. We used mach2dat[32] to perform the association analyses for the genome-wide analysis, and we used PLINK[33] to perform the association analyses for the HLA-imputed analysis. We tested for deviation from additivity by including a dominance deviation (0/1/0) term in the regression model.

***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, HLA-DRB1*07:01 was homology modeled from the crystal structure of HLA-DRB1*04:01 (Protein Data Bank (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 used 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 are freely available on 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 are freely available to researchers on request. For further data access details, please visit the International Serious Adverse Event Consortium Data Access Site (see URLs). Genotype data for the WTCCC ulcerative colitis and Crohn's disease cases are available from the European Genome-Phenome Archive (see URLs).

23. Badalov, N. *et al.* Drug-induced acute pancreatitis: an evidence-based review. *Clin. Gastroenterol. Hepatol.* **5**, 648–661 (2007).
24. Grant, R.L. Converting an odds ratio to a range of plausible relative risks for better communication of research findings. *Br. Med. J.* **348**, f7450 (2014).
25. Oliphant, A., Barker, D.L., Stuelpnagel, J.R. & Chee, M.S. BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques* (suppl.), 56–58 (2002).
26. Goldstein, J.I. *et al.* zCall: a rare variant caller for array-based genotyping: genetics and population analysis. *Bioinformatics* **28**, 2543–2545 (2012).
27. Barrett, J.C. *et al.* Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.* **40**, 955–962 (2008).
28. Anderson, C.A. *et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat. Genet.* **43**, 246–252 (2011).
29. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
30. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
31. Sebastiani, P. *et al.* Retraction. *Science* **333**, 404 (2011).
32. Li, Y., Willer, C.J., Ding, J., Scheet, P. & Abecasis, G.R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.* **34**, 816–834 (2010).
33. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
34. Scally, S.W. *et al.* A molecular basis for the association of the *HLA-DRB1* locus, citrullination, and rheumatoid arthritis. *J. Exp. Med.* **210**, 2569–2582 (2013).
35. Morris, G.M. *et al.* AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).
36. Trott, O. & Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).

doi:10.1038/ng.3093