OPEN ACCESS

Original research

First genome-wide association study reveals immune-mediated aetiopathology in idiopathic achalasia

Sandeep Grover ,[1] Ines Gockel,[2] Anna Latiano,[3] Anna Mokrowiecka,[4] Pouria Dasmeh,[5] Mira M Wouters,[6] Zuzana Vackova,[7] Stephan L Haas,[8] Tania Triantafyllou,[9] Nicole Kreuser,[10] Jessica Trautmann,[11] Stefan Niebisch,[10] Timo Hess,[12] Rene Thieme,[10] Jessica Bigge,[12] Hubert Louis,[13] Eric Quertinmont,[14] Aline Meirhaeghe,[15] Manon Muntaner,[15] Philippe Amouyel,[15] Guillaume Gourcerol ,[16] Stanislas Bruley des Varannes,[17] Francois Mion,[18] Michael Vieth,[19] Nikolaos Scarmeas,[20,21] Orazio Palmieri,[22] Francesca Tavano,[23] Roberto De Giorgio ,[24] Daniela Galimberti,[25,26] Andrea Arighi,[26] Beatrice Arosio,[26,27] Marco Bruno,[28] Justyna Wasielica-Berger,[29] Magdalena Gawron-Kiszka,[30] Maria Janiak,[31] Magdalena Siepsiak,[31] Krystian Adrych,[31] Tomasz Marek,[30] Andrzej Dabrowski,[29] Marek Majewski,[32] Piotr Gietka,[33] Maciej Gonciarz,[33] Julio Pérez de la Serna,[34] Laisy Zacarías Martínez,[35] Vilmantas Giedraitis,[36] Lena Kilander,[36] Laura Fratiglioni,[37] Luis Miguel Real,[38,39] Julius Spicak,[7] Jan Tack ,[40] Stefanie Heilmann-Heimbach,[41] Markus Nöthen,[42,43] Martin Ingelsson,[44,45] Caroline Graff,[46,47] Agustín Ruiz,[48,49] Jean-Charles Lambert,[15] Alfredo Ramirez,[50,51] Alexander J Eckardt,[52] Michaela Müller,[53] Michael Knapp,[54] Thaddäus T Wissinowski,[55] Jutta Keller ,[56] Christiane Josephine Bruns,[57] Christian Gerges,[58] Horst Neuhaus,[59] Thomas Rösch ,[60] Britta Siegmund ,[61] Brigitte Schumacher,[62] Marino Venerito ,[63] Antonio Ruiz de León,[64] Riccardo Rosati,[65] Vito Annese,[66] Uberto Fumagalli,[67] Luigi Laghi ,[68] Elena Urcelay,[69] Fabienne Vavasseur,[17] Sabine Roman ,[70] Pinghong Zhou ,[71] Quanlin Li ,[72] Zuqiang Liu,[71] Burkhard H A von Rahden,[73] Dimitris Theodorou,[9] Ewa Malecka-Panas,[74] Carlo Maj,[75] Ana G Vigo,[76] Jan Martinek ,[77] Guy Boeckxstaens ,[78] Johannes Schumacher[12]

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/gutjnl-2024-334498).

For numbered affiliations see end of article.

Correspondence to
Dr Johannes Schumacher; johannes.schumacher@uni-marburg.de

SG, IG, AL and AM contributed equally.
AGV, JM, GB and JS contributed equally.

Received 3 December 2024
Accepted 20 September 2025

To cite: Grover S, Gockel I, Latiano A, *et al*. *Gut* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/gutjnl-2024-334498

ABSTRACT

Background Idiopathic achalasia (IA) is characterised by the degeneration of neurons in the myenteric plexus leading to an irreversible impaired oesophageal function. Although immune-mediated mechanisms have been proposed, the underlying aetiopathology of IA remains poorly understood.

Objective This study aimed to uncover the genetic risk architecture of IA.

Design We carried out the first genome-wide association study (GWAS) on 4602 European patients with IA and 10 766 ethnically-matched controls.

Results A single nucleotide polymorphism (SNP) in *HLA-DQB1* leading to an 8-amino acid insertion on the protein level conferred strongest IA risk (PQGPPPAG: $p=3.27\times10^{-68}$, OR=2.45). Conditional analyses within the HLA locus revealed a complex genetic risk architecture. Three additional amino acid positions showed independent IA association (Omnibus $p<5\times10^{-8}$). These refer to positions 41 and 130 in HLA-DQα1, position 45 in HLA-DQβ1 and position 86 in HLA-DRβ1. Together, these findings highlight the pivotal role of class II HLA genetic variation in IA pathogenesis. Outside HLA, three independent variants showed IA association ($p<5\times10^{-8}$). One leads to an amino acid substitution with functional effect in PTPN22. Another risk variant leads to a downregulated expression of *TNFSF8*, *TNFSF15* and *TNC* in immune cells. The third risk SNP is located near *ZNF365*, but the exact underlying cellular mechanism remains unknown. Beyond the single marker level, polygenic risk scores revealed that patients with IA

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Only one genetic association study on a sufficiently powered sample with idiopathic achalasia (IA) has been carried out so far. This led to the identification of IA-associated risk variants in the genes *HLA-DQB1* and *HLA-DQA1*.

can be stratified based on their genetic risk. In addition, IA shows a shared aetiopathology with Crohn's disease (r_g=0.335). Integrating GWAS and single-cell RNA-sequencing data from the myenteric plexus showed that the memory T-cell type $FOS^+Tc4^+CD8^+$ plays a central role in IA development ($p=2.50\times10^{-19}$).

Conclusion This GWAS led to the identification of SNPs, cellular mechanisms and cell types that are involved in IA aetiopathology.

INTRODUCTION

Achalasia is a gastrointestinal motility disorder characterised by the degeneration of enteric neurons in the myenteric plexus leading to the absence of peristalsis and impaired relaxation of the lower oesophageal sphincter (LOS).[1] Symptoms include dysphagia, regurgitation of undigested food and weight loss. If left untreated, achalasia may progress to end-stage disease with oesophageal dilatation, known as megaoesophagus.[2] Two distinct disease types exist. Primary achalasia, also known as idiopathic achalasia (IA) due to its unknown origin, and secondary achalasia, which emerges as a consequence of conditions like organic obstructions due to cancer.[1] Although IA is uncommon with a lifetime prevalence of 1:10 000,[2] it constitutes the majority of cases among Europeans.

Studies have pointed to a multifactorial aetiopathology underlying IA.[3] Of interest, detailed examination of IA resection specimens shows infiltration of cytotoxic T lymphocytes within myenteric ganglia,[4] suggesting an aberrant immune reaction against enteric neurons. A previous viral infection with herpes simplex virus-1 (HSV-1) has been put forward to act as an auto-immune trigger in IA,[1 5] based on the observation that T cells from patients with achalasia proliferate in response to HSV-1.[6] Using Illumina Immunochip, which targets a small subset of loci in the human genome that are relevant for immune-mediated diseases, we recently discovered a highly significant IA-associated risk variant that leads to an 8-amino acid insertion in the cytoplasmic tail of HLA-DQβ1.[7] Although the functional impact of this insertion on antigen presentation still needs to be investigated, our observation could explain why patients with IA develop an aberrant immune response to a viral infection, including HSV-1.

In this study, we extended our analysis by carrying out the first genome-wide association study (GWAS) using a sufficiently powered case–control sample to characterise the genetic disease architecture of IA on the genome-wide level.

WHAT THIS STUDY ADDS

⇒ This first genome-wide association study (GWAS) confirms the IA association of variants in *HLA-DQB1* and *HLA-DQA1*, but also points to a more complex genetic risk architecture at this locus that involves an IA risk variant in HLA-DRB1. Moreover, the GWAS resulted in the identification of three novel disease variants outside HLA. One leads to an amino acid substitution with functional effect in PTPN22. One further novel IA risk variant leads to a downregulated expression of *TNFSF8*, *TNFSF15* and *TNC* in immune-relevant cells. The remaining disease variant is located near *ZNF365*, but the cellular pathogenic mechanism remains unknown.

⇒ On the polygenic level, this study provides the first IA heritability estimate and shows that immune-mediated mechanisms that are shared with Crohn's disease (CD) contribute to IA aetiopathology. The study further implies that the application of polygenic risk scores (PRS) might be useful to stratify patients with IA according to their genetic burden.

⇒ The integration of GWAS and transcriptome data from the myenteric plexus implies that the memory T-cell type $FOS^+Tc4^+CD8^+$ promotes disease development.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The genes with non-synonymous or expression-regulatory risk variants and the identified risk-conferring T-cell type ($FOS^+Tc4^+CD8^+$) represent promising candidates for future follow-up studies to determine the disease mechanisms of IA in detail.

⇒ The shared disease aetiology of IA and CD should be studied in the future to identify the underlying pathophysiological mechanisms.

⇒ PRS that determine individual genetic disease risk should be studied in future to test whether PRS-stratified patients with IA show higher susceptibility to potential environmental risk factors (eg, infectious diseases) or exhibit differences in age of onset and disease progression.

METHODS

Study participants

A total of 4602 patients with IA (2231 women, 2371 men) and 10 766 ethnically-matched controls (6270 women, 4496 men) were enrolled in this study. All individuals were of European ancestry, and patients were recruited from Sweden (N=166), Belgium (N=419), the Netherlands (N=306), Poland (N=260), Germany (N=1885), the Czech Republic (N=409), France (N=237), Spain (N=335), Greece (N=156) and Italy (N=429). In all patients, IA diagnosis was made according to standard procedures, including barium swallow oesophagography and/or high-resolution manometry. Online supplemental table 1 gives an overview of all included patients with IA and controls, including information on their origin and involved recruitment centres.

Genotyping and imputation

Genotyping of patients was conducted using the Global Screening Array (Illumina, USA) followed by standardised quality control (QC) procedures.[8] For controls, genotypes were obtained from previous studies (online supplemental table 1).[9] Ancestral outliers (patients and controls) were identified through principal components (PCs) analysis (online supplemental figure 1). Imputation of all post-QC genotypes was carried out using the Michigan Imputation Server (https://imputationserver.sph.umich.edu/index.html). Details regarding QC procedures and imputation methodologies are outlined in the online supplemental methods.

Association analyses

The GWAS consisted of eight case–control samples that were of Swedish (N=1249), Belgian/Dutch (N=1247), Polish (N=646), German/Czech (N=6648), French (N=991), Spanish (N=1773), Greek (N=1120) and Italian (N=1694) origin (online supplemental table 2). Participants from Belgium and the Netherlands, as well as from Germany and the Czech Republic, were analysed together based on their ethnic similarity (online supplemental figure 2). In each cohort, association analysis of imputed dosages with IA as outcome was performed using an additive logistic regression model in PLINK V.2.0 (https://www.cog-genomics.org/plink/2.0/). All analyses were adjusted for sex and PCs 1–10. We also employed the Firth bias-corrected logistic regression model using REGENIE (V.4.1) for sensitivity analysis,[10] because standard logistic regression methods can

produce inflated results when analysing unbalanced case–control datasets. Following single-nucleotide polymorphism (SNP) associations at the individual cohort level, a meta-analysis was conducted using METAL (V.2011-03-25) with a fixed-effects model. Genetic variants exhibiting significant heterogeneity (heterogeneity (HET) $p<0.05$) were excluded from the analysis. Furthermore, only variants present in more than four cohorts were retained. We applied a standard clumping algorithm via Functional Mapping and Annotation of GWAS (FUMA) (V.1.5.2) (https://fuma.ctglab.nl/) to identify independent loci. To refine association signals and identify the most likely causal variants, we performed fine-mapping analysis using the Sum of Single Effects (SuSiE) framework.[11] Lastly, we validated our association findings by conducting exploratory analysis in the UK Biobank (UKB) using individual-level data accessed under application ID 135122 (April 2019 release).[12] Additionally, we used publicly available GWAS summary statistics from the population-based cohort FinnGen.[13] See online supplemental methods for details on association analyses.

Pleiotropic effects

The most significant non-HLA disease SNP per locus and closely linked variants ($r^2 \geq 0.8$) were compared with data from the National Human Genome Research Institute GWAS catalogue (updated December 2023) to identify associations ($p<5\times10^{-08}$) with other diseases and traits. The direction of genetic effects was manually confirmed using the original research publications. When multiple GWAS were available for the same trait, the study with the largest sample size was prioritised.

HLA-imputation and association analyses

The genomic region encompassing HLA genes on chromosome 6p21 (28–34 Mb) was extracted, after which the post-QC genotypic data from all cohorts were merged to create an input file for HLA-imputation. We used the Michigan Imputation Server to impute alleles for three HLA class-I genes (*HLA-A*, *HLA-C*, *HLA-B*) and five class-II genes (*HLA-DRB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DPA1* and *HLA-DPB1*). Logistic regression analysis of imputed dosages for HLA alleles and amino acids was then performed with IA as outcome assuming additive and non-additive models. The analysis was adjusted for sex and PCs 1–10. See online supplemental methods for details on the imputation reference panel and downstream HLA analyses.

Functional gene mapping

We identified genes of potential functional impact for IA by applying gene mapping strategies using FUMA (V.1.5.2). SNPs located within non-HLA risk loci that were either genome-wide significant or in linkage disequilibrium (LD) ($r^2>0.6$) with one of the independently associated SNPs were assigned to genes in FUMA using three complementary approaches: positional mapping, eQTL mapping and chromatin interaction mapping (see online supplemental methods for details).

Gene set analyses

We performed enrichment analyses using the GENE2FUNC module of FUMA on two gene sets: (1) genes identified through functional gene mapping (as described in the previous section) and (2) 327 genes identified through MAGMA-based gene-level analysis when IA associations with $p<0.01$ were used as threshold. The gene list prioritised through MAGMA is provided in online supplemental table 3. Both gene sets were interrogated for pathway and disease enrichment analyses (see online supplemental methods for details).

Genetic heritability and correlations

We used LD score regression (LDSC) (V.1.0.1) to estimate the heritability of IA,[4] which was computed on the observed and liability scales. For the estimation on the liability scale, we assumed a lifetime disease prevalence of 1 in 10 000.[3]

Using LDSC, we further conducted cross-trait analyses to examine the genetic correlations between IA and related phenotypes. In a targeted analysis, we included the largest available GWAS datasets for eight immune-mediated and five neurodegenerative diseases (online supplemental table 4), as biological processes related to these disease categories have been proposed to influence the neuronal loss in the myenteric plexus.[2] To account for multiple comparisons, we applied Bonferroni corrections.

We also carried out a hypothesis-free genetic correlation analysis using GWAS summary statistics for 203 International Classification of Diseases, 10th Revision (ICD-10) diagnosed conditions in the UKB[12] (see online supplemental methods for details).

Polygenic risk score analysis

To evaluate the cumulative genetic risk of IA, we constructed polygenic risk scores (PRS) based on the genome-wide significant IA-associated variants identified in the GWAS meta-analysis. The UKB served as an independent target dataset, as described in the association analysis subsection. Effect size estimates were used to calculate PRS for each individual in the UKB target sample as a weighted sum of risk alleles. The analysis was implemented in PRSice (V.2.3.5).[14]

We performed logistic regression to test the association between individual PRS and IA status, adjusting for sex and the first 10 PCs. For clinical interpretability, we stratified individuals into PRS deciles and compared the top 5% of the PRS distribution against the remaining 95% to assess relative enrichment of disease risk. ORs and 95% CIs were calculated to quantify the increase in IA risk in the high-PRS group.

GWAS enrichment analysis in cell types from the plexus myentericus

We investigated the preferential expression of IA-associated genes across different cell types using the single-cell RNA sequencing (scRNA-seq) atlas by Liu *et al*.[15] The atlas includes 30 469 immune cells (21 828 lymphoid and 8641 myeloid cells) from LOS tissue surrounding the myenteric plexus of four non-IA individuals. For the enrichment analysis, we used the 327 IA-associated genes that were identified through MAGMA (see gene set analysis) and calculated their preferential expression within immune cell types from LOS/myenteric plexus samples using single-cell disease relevance score (scDRS) methodology[16] (see online supplemental methods for details).

RESULTS

Association analysis

The GWAS meta-analysis resulted in a dataset of 6.4 million SNPs. The quantile–quantile (Q–Q) plot showed no inflation in the meta-analysed data set ($\lambda=1.014$) (online supplemental figure 3). Also, on the individual cohort level, Q–Q plots showed no inflation (λ between 0.976 in the Belgian/Dutch cohort and 1.075 in the Greek cohort) (online supplemental figure 3), which is in line with plots from PC analyses

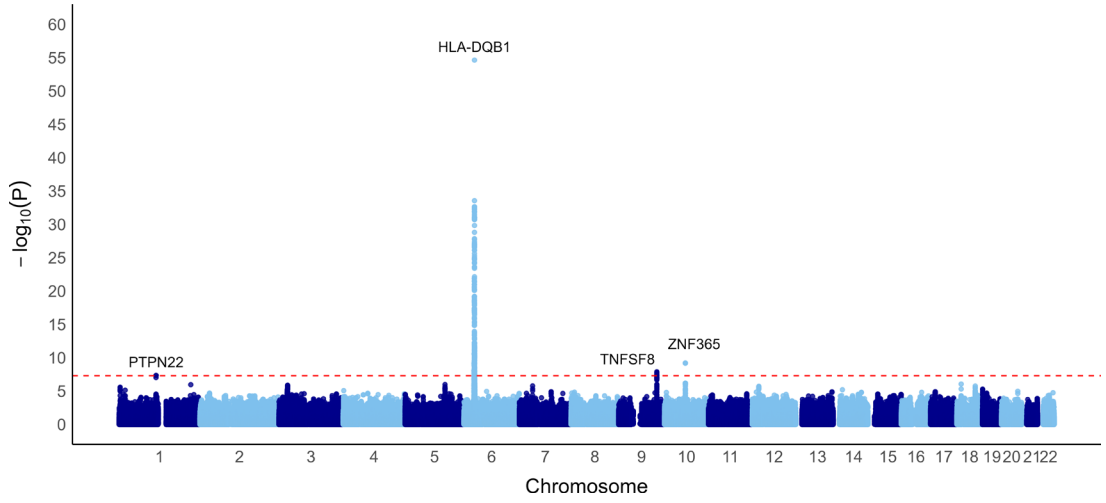

Figure 1 Manhattan plot of the GWAS meta-analysis on IA. Association p values from the fixed-effect meta-analysis of logistic regression analyses of 6.4 million SNPs from individual GWAS of eight different European cohorts (4602 patients with IA and 10 766 ethnically-matched controls). The p values are shown as $-\log_{10}(P)$ for each SNP along the vertical axis, with genomic position for each SNP on the horizontal axis. Each point corresponds to an SNP. The red dashed line indicates the threshold for genome-wide significance ($p=5\times10^{-08}$). GWAS, genome-wide association studies; IA, Idiopathic achalasia; SNP, single nucleotide polymorphism.

that point to no differences between each case–control sample (online supplemental figure 1).

In total, we identified genome-wide significant IA associations at four independent loci (figure 1). The strongest associated variant ($p=2.33\times10^{-55}$, OR=2.54, 95% CI=2.36 to 2.86) is located in the HLA region on chromosome 6p21. This refers to rs28688207, which is located in the splice acceptor site of the gene *HLA-DQB1* and leads to an 8-amino acid insertion in the cytoplasmic tail of HLA-DQβ1. While this SNP has been identified previously,[7] all associations at the remaining three loci are novel. They are located at the genes *PTPN22* (rs2476601) on chromosome 1p13, *TNFSF8* (rs3181374) on chromosome 9q33 and *ZNF365* (rs12774545) on chromosome 10q21 (table 1). One of these variants represents a non-synonymous variant (rs2476601, c.1858G>A) leading to an arginine to tryptophan substitution (R620W) in PTPN22.

Online supplemental figure 4 provides association plots for all identified IA risk loci, including LD patterns within the identified regions and information on whether the tested variants were directly genotyped or imputed.

Although all QC procedures provided no evidence for false positive association results, we carried out different tests to assess the robustness of our findings: (1) At all implicated risk loci, forest plots showed that the IA associations are present in almost all included individual cohorts (online supplemental figure 5). (2) In addition, we applied REGENIE,[10] another method to calculate GWAS meta-analyses considering that the ratio of cases and controls differs at the individual cohort level. Here, all implicated risk loci also showed IA associations ($p<9.66\times10^{-5}$) (online supplemental table 5). (3) Finally, we tested all IA-contributing variants for disease association using two population-based cohorts, FinnGen (646 patients with achalasia, 423 785 controls) and UKB (353 patients with achalasia, 484 086 controls). Although the number of patients with achalasia was relatively small in both cohorts and the proportion of secondary achalasia cases remains unknown, the risk alleles at all four identified IA-associated variants were more frequent among patients in FinnGen. In UKB, three of the four variants showed higher allele frequencies in patients, with the exception of rs2476601. Notably, our top implicated IA variant, rs28688207, showed a strong and statistically significant association in UKB ($p=1.24\times10^{-5}$, OR=2.03, 95% CI=1.48 to 2.80) (online supplemental table 6).

Only two non-HLA loci—*PTPN22* and *TNFSF8*—were prioritised for fine-mapping based on functional evidence from expression studies (see Functional gene mapping). At *PTPN22*, the top credible set consisted of a single variant with a posterior inclusion probability of 1.0, matching the lead GWAS SNP (data not shown). At *TNFSF8*, however, the lead SNP rs3181374 was not included in the top credible set. Instead, SuSiE prioritised rs2974, a nearby SNP in perfect LD ($r^2=1$, $D'=1$) (data not shown). This underscores the limitation of fine-mapping in regions of high LD where distinguishing the true causal variant is challenging.

Table 1 Lead associations of genome-wide significant associated IA loci

SNP	Chromosome	Position (bp, hg19)	Gene	Function	Effect allele (EA)	Other allele	EA frequency	P value	OR (95% CI)
rs2476601	1p13	114 377 568	*PTPN22*	Non-synonymous SNP	A	G	0.108	4.40×10^{-08}	1.29 (1.18 to 1.41)
rs28688207	6p21	32 628 660	*HLA-DQB1*	Splicing effect	C	T	0.061	2.33×10^{-55}	2.54 (2.36 to 2.86)
rs3181374	9q33	117 665 187	*TNFSF8*	UTR3	A	G	0.552	1.28×10^{-08}	1.18 (1.11 to 1.25)
rs12774545	10q21	64 382 662	*ZNF365*	intronic	A	G	0.185	6.08×10^{-10}	1.25 (1.17 to 1.35)

The associations are shown for the EA in the entire IA sample. P values, ORs and the corresponding 95% CIs are shown. The frequencies of the EA are derived from the overall study population (N=15 368 patients and controls). The conditional analysis of all non-HLA SNPs failed to reveal any independent associations (data not shown).
IA, idiopathic achalasia; SNPs, single-nucleotide polymorphism.

Table 2 Geospatial north-south gradients for the 8-amino acid insertion in HLA-DQβ1 (rs28688207) and R620W in PTPN22 (rs2476601) in the European population

				rs28688207 (HLA-DQB1): effect allele (EA) C				rs2476601 (PTPN22): EA A			
Sample	Region	Cases (n)	Controls (n)	Cases EAF	Controls EAF	P value	OR (95% CI)	Cases EAF	Controls EAF	P value	OR (95% CI)
SW	North	166	1083	0.081	0.024	2.41×10^{-06}	3.64 (2.13 to 6.22)	0.145	0.109	0.026	1.50 (1.05 to 2.14)
BE/NL	Central	725	522	0.092	0.028	2.13×10^{-06}	2.87 (1.86 to 4.44)	0.106	0.100	0.170	1.22 (0.92 to 1.62)
PL	Central	260	386	0.094	0.019	5.11×10^{-06}	4.41 (2.33 to 8.35)	0.142	0.128	0.842	0.96 (0.67 to 1.39)
GE/CR	Central	2294	4354	0.080	0.032	3.91×10^{-22}	2.36 (1.98 to 2.81)	0.136	0.109	6.07×10^{-07}	1.34 (1.20 to 1.51)
FR	Central	237	754	0.080	0.038	0.722	1.13 (0.58 to 2.2)	0.095	0.100	0.341	0.78 (0.48 to 1.29)
ES	South	335	1438	0.096	0.040	7.42×10^{-09}	2.84 (1.99 to 4.05)	0.093	0.072	0.013	1.51 (1.09 to 2.09)
GR	South	156	964	0.176	0.081	4.16×10^{-05}	2.14 (1.49 to 3.09)	0.048	0.034	0.244	1.43 (0.78 to 2.61)
IT	South	429	1265	0.179	0.060	6.63×10^{-14}	2.83 (2.16 to 3.72)	0.047	0.042	0.739	1.07 (0.72 to 1.58)

The association results of the rs28688207 in *HLA-DQB1* and rs2476601 in *PTPN22* are presented for all eight analysed European samples including their location within Europe (region), the number of included cases and controls and the effect allele frequencies (EAF).
BE/NL, Belgium/The Netherlands; ES, Spain; FR, France; GE/CR, Germany/The Czech Republic; GR, Greece; IT, Italy; PL, Poland; SW, Sweden.

Geospatial association effects

Forest plots showed that the IA associations at all loci are present in almost all included cohorts (online supplemental figure 5). However, we identified geospatial north-south gradients for the 8-amino acid insertion in HLA-DQβ1 (rs28688207) and R620W in PTPN22 (rs2476601) (table 2). The frequency of the 8-amino acid insertion in HLA-DQβ1 is more than two times as common in southern compared with northern Europeans (17.9% in Italian patients, 8.1% in Swedish patients). In contrast, R620W in PTPN22 is more than two times as common in northern than in southern Europeans (14.5% in Swedish patients, 4.7% in Italian patients).

Pleiotropic effects

The cross-annotation of our IA-associated SNPs along with those in LD ($r^2\geq0.80$) (online supplemental table 7) against studies in the GWAS catalogue identified associations with multiple diseases and traits. R620W in PTPN22 exhibits the highest pleiotropy (online supplemental table 8). This variant, along with those in LD, is associated with 11 other immune-mediated diseases and traits. Here, concordant associations are observed for type 1 diabetes (T1D), rheumatoid arthritis (RA), juvenile idiopathic arthritis (JIA), systemic lupus erythematosus, Hashimoto's thyroiditis, Addison's disease, Graves' disease, myasthenia gravis, vitiligo and polymyositis, while Crohn's disease (CD) shows a discordant association. Pleiotropic effects were also observed for endocrine and oncological diseases, with hypothyroidism and basal cell carcinoma showing concordant associations. Additionally, pleiotropy was observed for white blood cell counts, including granulocytes (such as neutrophils and basophils) and agranulocytes (such as lymphocytes and monocytes), which showed discordant associations. Moreover, rs3181374 and its linked variants ($r^2\geq0.80$) in *TNFSF8* showed pleiotropic effects. These variants are concordantly associated with TNFSF8 levels in blood plasma and with cell counts of granulocytes, including eosinophils and neutrophils as well as HLA-DR^+ HLA-$CD4^+$ and HLA-$CD8^+$ T cells (online supplemental table 8).

HLA association analyses

To identify the causal IA variants in the HLA region, we carried out an HLA imputation followed by conditional association tests under an additive model.[17] After imputation, we tested 120 classical HLA alleles at two-digit resolution, 222 classical HLA alleles at four-digit resolution and 2320 amino acids corresponding to 401 polymorphic sites. The conditional tests using the amino acid positions revealed a complex genetic risk architecture at this locus that involved four independent genome-wide significant IA associations (Omnibus $p<5\times10^{-8}$), namely HLA-DQβ1 position 227, HLA-DQα1 positions 41 and 130, HLA-DQβ1 position 45 and HLA-DRβ1 position 86 (online supplemental table 9, online supplemental figure 6). The first three independent IA associations involving HLA-DQβ1 and HLA-DQα1 have been reported previously using a smaller sized sample.[7] Table 3 further shows the associations of individual residues at the amino acid positions identified in the conditional analyses. These include the 8-amino acid insertion (PQGPPPAG) at position 227 in HLA-DQβ1 ($p=3.27\times10^{-68}$, OR=2.45, 95% CI=2.22 to 2.71), lysine (K) and alanine (A) at positions 41 and 130 in HLA-DQα1 that share both perfect LD ($p=1.26\times10^{-24}$, OR=1.56, 95% CI=1.43 to 1.69) as well as glutamic acid (E) at position 45 in HLA-DQβ1 ($p=2.06\times10^{-16}$, OR=1.28, 95% CI=1.20 to 1.37). The fourth

Table 3 Logistic regression and conditional analyses of IA-associated amino acids in the HLA region

Protein	Amino acid (AA) position*	Risk AA/other AA	Cases EAF	Controls EAF	P value	OR (95% CI)	Haplotype
HLA-DQβ1	227–234	PQGPPPAG/-	0.097	0.040	3.3×10^{-68}	2.45 (2.22 to 2.71)	**05:03*, **06:01*
HLA-DQα1	41 130	K/R A/S	0.121	0.075	1.3×10^{-24}	1.56 (1.43 to 1.69)	**01:03*, **01:10*
HLA-DQβ1	45	E/G	0.246	0.218	2.1×10^{-16}	1.28 (1.20 to 1.37)	**03:01*, **03:04*, **03:69*
HLA-DRβ1	86	G/V†	0.538	0.512	1.5×10^{-18}	1.28 (1.21 to 1.35)	**01:01*, **04:01*, **11:01*†

The AA position within the respective HLA, the effect alleles (EA) and their frequencies (EAF) in cases and controls are shown. In addition, the HLA haplotypes are shown on which the associated disease variants are located.
*According to IPD-IMGT/HLA database (release 3.57).
†Most frequently observed in the present study population.
IA, idiopathic achalasia.

independent IA association has not been reported previously and involves position 86 in HLA-DRβ1. Here, two common amino acid residues exist. While valine (V) confers protection against IA, glycine (G) increases IA risk ($p=1.54\times10^{-18}$, OR=1.28, 95% CI=1.21 to 1.35) (table 3). Of note, at each conditional step, the implicated amino acid residues showed stronger association than the most significant associated classical HLA alleles (online supplemental figure 6).

We further evaluated whether HLA associations under alternative models provided stronger evidence than those observed under the additive model. While the dominant model showed similar results compared with the additive model, no genome-wide significant IA associations were detected under the non-additive model (online supplemental tables 10–12). Finally, we excluded the possibility that the observed IA associations on *HLA-DRB1*, *HLA-DQA1* and *HLA-DQB1* and haplotypes are influenced by the complex LD pattern in this region. An analysis of all haplotype combinations involving class II alleles demonstrated that the associated amino acid residues are distributed across different haplotypes (online supplemental table 13).

Functional gene mapping

FUMA analysis mapped a total of 30 genes to the three non-HLA risk loci identified in this study. 5 genes were identified through positional mapping, 12 genes through eQTL mapping and 29 genes through chromatin interaction mapping (online supplemental tables 14–16). On chromosome 1p13, the leading GWAS variant rs2476601 (R620W) in *PTPN22* was predicted as being deleterious with a Combined Annotation Dependent Depletion (CADD) score of 17.25 (online supplemental table 7). In addition, rs2476601 (R620W) represents the strongest eQTL at this locus leading to a reduced expression of *PTPN22* in specific immune cells of the blood of IA risk allele carriers, namely B cells ($p=1.14\times10^{-7}$, False Discovery Rate, $FDR=5.72\times10^{-4}$) and monocytes ($p=3.25\times10^{-7}$, $FDR=1.63\times10^{-3}$) (online supplemental table 15). SNP rs2476601 (R620W) is also an eQTL in immune cells of the blood for the expression of three additional genes at this locus, namely *DCLRE1B*, *BCL2L15* and *AP4B1* (online supplemental tables 15). On chromosome 9q33, rs145815687 was predicted as being deleterious with a CADD score of 15.97 (online supplemental table 7), because its location in a cis-regulatory element 196 bp upstream of *TNFSF8*. This variant is in strong LD to the lead IA-associated SNP at this locus ($r^2=0.91$), but was not analysed in our GWAS as it represents a 4 bp insertion-deletion (indel) polymorphism. Although the indel is the best functional risk variant at this locus, it was probably not analysed in most eQTL analyses either. Here, SNPs in strong LD to the lead GWAS variant ($r^2>0.8$) showed eQTL effects for the expression of three different genes that are almost exclusively present in immune cells of the blood. In IA risk allele carriers, we observed a downregulated expression of *TNFSF8* in monocytes (best rs10817683: $p=5.10\times10^{-16}$, $FDR=2.55\times10^{-12}$), $CD4^+$ T cells (best rs1006026: $p=2.09\times10^{-11}$, $FDR=1.05\times10^{-7}$), natural killer (NK) cells (best rs2418325: $p=6.97\times10^{-8}$, FDR=0.049) and $CD8^+$ T cells (best rs2418325: 1.09×10^{-7}, $FDR=5.44\times10^{-4}$) (online supplemental table 15). A downregulated expression in IA risk allele carriers was also observed for *TNC* in $CD4^+$ T cells (best rs3181372: $p=8.43\times10^{-11}$, FDR=0.049) and NK cells (best 7032773: $p=1.24\times10^{-6}$, FDR=0.049) as well as for *TNFSF15* in monocytes (best rs12347977: $p=1.39\times10^{-9}$, $FDR=6.96\times10^{-6}$) (online supplemental table 15). In contrast to the above-mentioned loci, the functional gene mapping analyses with FUMA showed no convincing results at the IA GWAS locus on chromosome 10q21 (*ZNF365*) (online supplemental table 15).

Gene set analyses

The 30 genes that were mapped by FUMA provided no significant findings in the pathway enrichment analysis (data not shown). However, in the disease enrichment analysis, this gene set showed significant overlap with GWAS genes for two immune-mediated diseases, namely alopecia areata (AA) ($p=3.33\times10^{-20}$, $FDR=1.47\times10^{-16}$) and CD ($p=6.98\times10^{-11}$, $FDR=1.54\times10^{-7}$) (online supplemental table 17).

The MAGMA gene-based analysis identified 327 significant genes associated with IA at a suggestive p value of 0.01. However, after Bonferroni correction, only *TNFSF8* on chromosome 9q33 remained significantly IA-associated ($p=3.82\times10^{-8}$) (online supplemental table 3). The pathway enrichment analysis using the MAGMA gene set confirmed the involvement of immune-mediated processes in IA aetiology. Genes involved in development of rheumatoid arthritis ($p=5.05\times10^{-6}$, $FDR=9.87\times10^{-3}$), 16p11.2 distal deletion syndrome ($p=6.39\times10^{-6}$, $FDR=9.87\times10^{-3}$), cytokine–cytokine receptor interaction ($p=6.87\times10^{-5}$, $FDR=1.28\times10^{-2}$) and GATA3 pathway ($p=3.91\times10^{-5}$, $FDR=4.02\times10^{-2}$) were enriched among the IA GWAS genes (online supplemental table 18). Also in the disease enrichment analysis, the MAGMA gene set showed significant overlap with GWAS genes for immune-mediated diseases. Besides CD ($p=8.07\times10^{-11}$, $FDR=3.57\times10^{-7}$) and AA ($p=3.38\times10^{-8}$, $FDR=3.86\times10^{-5}$), which already showed significant overlap with the FUMA gene set, this included IBD ($p=9.19\times10^{-9}$, $FDR=2.03\times10^{-5}$), asthma ($p=2.21\times10^{-7}$, $FDR=1.63\times10^{-4}$), ankylosing spondylitis ($p=5.09\times10^{-5}$, $FDR=1.73\times10^{-2}$) and childhood onset asthma ($p=1.03\times10^{-4}$, $FDR=3.22\times10^{-2}$) (online supplemental table 17).

SNP-based heritability and genetic correlations

Using our GWAS summary-level data we observed an SNP-h^2 heritability estimate of 10.04% (±3.36% SD) for IA, corresponding to a liability-scale estimate of 3.64% (±1.13% SD) in the general population (assuming a lifetime IA prevalence of 1 in 10 000). Here, the exclusion of the HLA region had no significant impact on these estimates.

Using a hypothesis-driven approach with GWAS summary statistics for eight immune-mediated and five neurodegenerative diseases, we observed only one positive correlation, namely with CD ($r_g=0.335$, p=0.015) (online supplemental figure 7, online supplemental table 4). We further observed a negative correlation with JIA ($r_g=-0.528$, p=0.032) (online supplemental figure 7, online supplemental table 4). Excluding the HLA region also had no impact on these correlation estimates (data not shown). Given that IA showed a positive correlation with CD and a negative correlation with JIA, we also performed a genetic correlation analysis between CD and JIA. Consistent with our IA findings, CD and JIA showed a negative genetic correlation on the polygenic level ($r_g=-0.340$, $p=3.99\times10^{-5}$) (data not shown).

Using a hypothesis-free approach, we analysed GWAS summary statistics for 211 ICD-10-diagnosed conditions in the UKB. Here, other non-infective gastro-enteritis and colitis (ICD code K52) classified under diseases of the digestive system showed the most significant positive genetic correlation with IA ($r_g=0.524$, $P_{uncorrected}=0.003$) (online supplemental table 19). Also, CD (ICD code K50) showed a positive correlation in the UKB ($r_g=0.190$), but was not significant (p=0.298) (online

supplemental table 19). The reason why this was not significant is probably due to the small number of patients with CD in the UKB (N=968) compared with the GWAS sample of the targeted approach (N=12 194).

Polygenic risk score analysis

PRS derived from the top IA-associated variants showed a significant association with IA status in the UKB target dataset ($p=8.7\times10^{-5}$). When stratified by PRS percentiles, individuals in the top 5% of the PRS distribution had markedly higher odds of IA compared with those in the remaining 95% ($p=6.0\times10^{-5}$, OR=2.06, 95% CI=1.44 to 2.94).

Relevant immune cell types conferring disease risk

To investigate how IA-associated genes contribute to disease, we prioritised immune cell types that show a preferential expression of IA-associated genes. We used an scRNA-seq atlas[15] consisting of 30 469 immune cells (21 828 lymphoid and 8641 myeloid cells) from LOS tissue surrounding the myenteric plexus and partitioned into 26 subclusters of lymphocytes (online supplemental figure 8) and 11 subclusters of myeloid cells. We then performed scDRS to prioritise disease-relevant cell types based on their cell type-specific expression of disease-associated genes as described in Methods. Within the lymphocyte cells, we identified the memory T-cell type $FOS^{+}Tc4^{+}CD8^{+}$ that showed significant IA enrichment ($p=2.50\times10^{-19}$) (online supplemental figure 8, online supplemental table 20). In contrast, no significant IA enrichment was observed among the myeloid cells (online supplemental table 20).

DISCUSSION

This study represents the first IA GWAS and highlights that immune-mediated mechanisms influenced by genetic risk are of major relevance for disease development.

Some genetic risk factors share characteristics with other immune-mediated diseases and traits. At the single marker level, this includes R620W in PTPN22, which encodes the protein tyrosine phosphatase non-receptor type 22 and is also a risk factor for 10 other immune-mediated diseases. Our study further demonstrates that R620W in PTPN22 is particularly relevant for northern European IA patients due to its higher frequency in these populations. Because of its prominent role in immune-mediated diseases, R620W in PTPN22 has been the subject of many functional studies. Most of these studies have focused on its biological effect on T-cell receptor (TCR) signalling.[18] A recent study suggests that R620W contributes to immune-mediated risk by enhancing TCR-signalling and activation in lower avidity self-reactive T cells.[19] At the polygenic level, IA shares characteristics with CD ($r_g=0.335$). This points to pathobiological pathways that are common to both immune-mediated GI diseases, such as impaired mucosa-barrier functions. In contrast, both IA ($r_g=-0.528$) and CD ($r_g=-0.340$) showed a negative genetic correlation with JIA. The findings suggest that genetic variants that confer IA and CD risk are associated with reduced JIA risk on the polygenic level. Similar negative correlations have been found for other autoimmune diseases, including RA and T1D[20] or ulcerative colitis and primary biliary cirrhosis (PBC).[21] At the single locus level, the negative genetic correlation between IA and JIA is not present for *PTPN22*, as R620W represents a risk factor for both diseases.

IA also shows genetic characteristics that are—according to current knowledge—not shared with other immune-mediated diseases. Here, the IA associations with genetic variants in the HLA-DQ receptor are particularly relevant. The disease associations of the 8-amino acid insertion and glutamic acid at position 45 in HLA-DQβ1 as well as lysine and alanine at position 41 and 130 in HLA-DQα1 have been previously reported.[7] The 8-amino acid is the strongest IA-associated variant (OR=2.45) and occurs more frequently among southern Europeans. The insertion is located within the cytoplasmic tail of HLA-DQβ1 and may lead to altered intracellular trafficking, as the C-terminal tails of HLA-molecules are thought to facilitate these processes.[22] The other IA risk variant in HLA-DQβ1, glutamic acid at position 45, is located near the peptide-binding site and might therefore perturb antigen presentation.[7] Although lysine and alanine at position 41 and 130 in HLA-DQα1 have no apparent structural consequences, evidence suggests that position 41 represents the binding site necessary for the interaction between HLA-DQ and the peptide exchange chaperone HLA-DM.[23] However, it would be premature to conclude that position 41 in HLA-DQα1 is more relevant than position 130 without additional functional studies. The fourth IA association at the HLA locus has not been described so far and concerns position 86 in HLA-DRB1, which is located in the peptide-binding site and influences antigen presentation.[24 25] At this position, glycine (G) confers increased IA risk, whereas valine (V) is protective. Notably, the opposite effect has been reported in other immune-mediated diseases, where V constitutes the risk residue for multiple sclerosis,[26 27] pemphigus vulgaris[28] and PBC.[29]

The genes at chromosome 9q33 are also involved in immune response, but have not been connected to immune-mediated diseases through GWAS so far. The leading risk SNP, rs3181374, is located in the 3' untranslated region of *TNFSF8*, which encodes the tumour necrosis factor superfamily member 8. According to FUMA, this SNP and variants in strong LD lead to a downregulation of *TNFSF8* expression in monocytes, naïve $CD4^{+}$ T cells, naïve $CD8^{+}$ T cells and NK cells. TNFSF8 is primarily involved in T cell-dependent immune response.[30] However, the risk SNPs at this locus also represent eQTLs for the expression of *TNC* and *TNFSF15* in immune cells of the blood. Both genes also play an important role in immune response and tissue remodelling.[31 32]

The implicated IA gene at the third new locus is functionally not well characterised. *ZNF365* encodes the zinc finger protein 365 that seems to be involved in homologous recombination repair during DNA replication.[33] Genetic variability at this locus also contributes to Vogt-Koyanagi-Harada (VKH) syndrome,[34] which is characterised by a T cell-mediated autoimmune response against melanocytes. However, the VKH-relevant variant[34] is not linked to the IA risk variant in the present study ($r^2<0.06$).

Our analysis further implies that the application of PRS might be useful to stratify patients with IA according to their genetic burden. Accordingly, IA risk among individuals in the top 5% of the PRS distribution has a 2.06-fold higher risk compared with remaining 95%. Thus, PRS should be tested in future whether patients with IA with high or low genetic burden show other susceptibilities to potential environmental risk factors (eg, infectious diseases) or exhibit differences in age of onset and disease progression.

The integration of our GWAS data and scRNA-seq data from LOS/myenteric plexus revealed that one memory T-cell type is particularly relevant for IA development. This refers to $FOS^{+}Tc4^{+}CD8^{+}$ T cells that showed significant IA enrichment ($p=2.50\times10^{-19}$) and were significantly expanded and localised surrounding the myenteric plexus tissue of patients with achalasia in a previous study.[15] $FOS^{+}Tc4^{+}CD8^{+}$ cells represent tissue-resident memory T cells, which are known for their ability to remain localised in tissues and provide rapid immune

responses on reactivation. Their disease relevance and expansion in patients with IA suggests that these tissue-resident memory T cells might actively contribute to chronic inflammation and immune dysregulation in IA.

Despite all new insights into the genetic architecture of IA, this study has limitations. We have used patient and control samples that were collected by different centres among different European populations. Although we observed no evidence for population stratification effects, we cannot fully exclude that our study design has influenced the present IA findings. Moreover, as this study focused exclusively on European patients with IA, it is not possible to generalise our conclusions to non-European populations. A further limitation of this study is the lack of age information across all cohorts, preventing its inclusion as a covariate. While the impact of age on IA risk is unknown, differences in age distribution between cases and controls might have introduced residual confounding, warranting validation in age-adjusted analyses.

In summary, the first GWAS for IA identified new risk variants and disease genes. Among these, protein-coding variants in class II HLA genes conferred the strongest IA risk. Outside the HLA region, R620W in PTPN22 (chromosome 1p13) represents an IA risk variant and shows biological effects at the cellular level. In addition, the IA risk variants on chromosome 9q33 lead to a downregulated expression of *TNFSF8*, *TNFSF15* and *TNC* in immune-relevant cells of the blood. In contrast, the biological mechanism through which genetic variants at *ZNF365* (chromosome 10q21) contribute to IA risk needs to be elucidated in future studies. At the polygenic level, this study provides the first IA heritability estimate and PRS for future patient stratification. In addition, the data show that immune-mediated mechanisms that are shared with CD confer risk to IA development. Finally, integration of GWAS and scRNA-seq data revealed that $FOS^{+}Tc4^{+}CD8^{+}$ T cells in the LOS/myenteric plexus are most relevant for IA development.

Author affiliations
[1]Center for Human Genetics, University of Marburg, Marburg, Germany
[2]Department of General and Visceral Surgery, Hospital of Magdeburg, Magdeburg, Germany
[3]Unit of Gastroenterology, IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy
[4]Department of Digestive Tract Diseases, Medical University of Lodz Faculty of Medicine, Lodz, Poland
[5]University Hospital of Marburg, Marburg, Germany
[6]Translational Research Center for Gastrointestinal Disorders (TARGID), Leuven University, Leuven, Belgium
[7]Hepatogastroenterology, Institute for Clinical and Experimental Medicine, Prague, Czech Republic
[8]Department of Upper GI Diseases, Karolinska University Hospital and Unit of Gastroenterology and Rheumatology, Karolinska Institute, Stockholm, Sweden
[9]Foregut Surgery Department, 1st Propaedeutic Surgical Clinic, Hippokration General Hospital of Athens, Athens, Greece
[10]Department of Visceral, Transplant, Thoracic and Vascular Surgery, University Hospital of Leipzig, Leipzig, Germany
[11]Institute of Human Genetics, University Hospital Bonn Institute of Human Genetics, Bonn, Germany
[12]University of Marburg, Marburg, Germany
[13]Department of Gastroenterology, Université Libre de Bruxelles, Bruxelles, Belgium
[14]Université Libre de Bruxelles, Bruxelles, Belgium
[15]Institut Pasteur de Lille U1167, Lille, France
[16]Physiology Department - UMR INSERM 1073, University Hospital Centre Rouen, Rouen, France
[17]CHU Nantes, Institut des Maladies de l'Appareil Digestif, Nantes, France
[18]Hopital Edouard Herriot, Lyon, France
[19]Institute for Pathology, Friedrich-Alexander University Erlangen-Nürnberg, Klinikum Bayreuth, Bayreuth, Germany
[20]1st Department of Neurology, National and Kapodistrian University of Athens, Athens, Greece
[21]Department of Neurology, Columbia University, New York, New York, USA
[22]Division of Gastroenterology, IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy
[23]Gastroenterology Unit, Fondazione IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy
[24]University of Ferrara Faculty of Medicine and Surgery, Ferrara, Italy
[25]Department of Biomedical, Surgical and Dental Sciences, University of Milan, Milano, Italy
[26]Ospedale Maggiore Policlinico, Milan, Italy
[27]Department of Clinical Sciences and Community Health, University of Milan, Milano, Italy
[28]Gastroenterology and Hepatology, Erasmus Medical Centre, Rotterdam, The Netherlands
[29]Department of Gastroenterology and Internal Medicine, Medical University of Bialystok, Poland, Poland
[30]Department of Gastroenterology and Hepatology, Medical University of Silesia, Katowice, Poland
[31]Department of Gastroenterology and Hepatology, Medical University of Gdansk, Gdansk, Poland
[32]Department of Plastic and Reconstructive Surgery, Medical University of Lublin, Lublin, Poland
[33]Department of Gastroenterology and Internal Medicine, Military Institute of Medicine, Warsaw, Poland
[34]Immunology and Gastroenterology Departments, Instituto de Investigacion Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain
[35]Hospital Universitari Germans Trias i Pujol, Badalona, Spain
[36]Department of Public Health and Caring Sciences, Clinical Geriatrics, Uppsala University, Uppsala, Sweden
[37]Department NVS, Center for Alzheimer Research, Solna, Sweden
[38]Unidad Clinica de Enfermedades Infecciosas y Microbiología, Hospital Universitario de Valme, Seville, Spain
[39]Instituto de Biomedicina de Sevilla (IBIS), Seville, Spain
[40]Pathophysiology, University Of Leuven, Leuven, Belgium
[41]Institute of Human Genetics, Rheinische Friedrich-Wilhelms-Universitat Bonn, Bonn, Germany
[42]Department of Genomics, Life and Brain Center, University of Bonn, Bonn, Germany
[43]Institute of Human Genetics, University of Bonn, Bonn, Germany
[44]Department of Public Health and Caring Sciences, Molecular Geriatrics, Uppsala University, Uppsala, Sweden
[45]University Health Network, Toronto, Ontario, Canada
[46]Department NVS, Bioclinicum, Solna, Sweden
[47]Karolinska University Hospital, Stockholm, Sweden
[48]University of Texas Health Science Center, San Antonio, Texas, USA
[49]Ace Alzheimer Center Barcelona, Barcelona, Spain
[50]Department of Psychiatry and Psychotherapy, University of Cologne, Koln, Germany
[51]Department of Neurodegenerative Disease and Geriatric Psychiatry, University of Bonn Medical Center, Bonn, Germany
[52]Gastro-Praxis Wiesbaden, Wiesbaden, Albania
[53]Department of Gastroenterology, Endocrinology, Metabolism and Infectiology, University Hospital of Giessen and Marburg Campus Marburg, Marburg, Germany
[54]Institute for Medical Biometry, Informatics and Epidemiology (IMBIE), University of Bonn, Bonn, Germany
[55]Center for Inner Medicine II, Klinikum Chemnitz gGmbH, Chemnitz, Germany
[56]Internal Medicine, Israelitic Hospital, Hamburg, Germany
[57]Department of General, Visceral and Cancer Surgery, University Hospital of Cologne, Cologne, Germany
[58]Department of Interventional Gastroenterological Endoscopy, University Hospital Essen, Essen, Germany
[59]Department of Gastroenterology, EVK Evangelisches Krankenhaus Duesseldorf, Duesseldorf, Germany
[60]Interdisciplinary Endoscopy, University Hospital Hamburg-Eppendorf, Hamburg, Germany
[61]Medical Department I, Charite, Campus Benjamin Franklin, Berlin, Germany
[62]Gastroenterology, Elisabeth-Krankenhaus-Essen GmbH, Essen, Germany
[63]Department of Gastroenterology, Hepatology and Infectious Diseases, Otto-von-Guericke-Universitat Magdeburg, Magdeburg, Germany
[64]Department of Gastroenterology, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain
[65]Department of Gastrointestinal Surgery, , San Raffaele Hospital, Vita-Salute San Raffaele University, Milan, Italy
[66]Department of Gastroenterology, Careggi Hospital, University of Florence, Florence, Italy
[67]Department of Gastroenterology, Humanitas Clinical and Research Center, Istituto Clinico Humanitas IRCCS, Rozzano - Milan, Italy
[68]Gastroenterology Department and Laboratory of Molecular Gastroenterology, Humanitas Clinical and Research Center, Rozzano, Italy
[69]Immunology, Hospital Clínico San Carlos-Madrid, Madrid, Spain
[70]Digestive Physiology, Hopital E Herriot, Hospices Civils de Lyon, Université de Lyon, Lyon, France

[71]Endoscopy Center and Endoscopy Research Institute, Zhongshan Hospital Fudan University, Shanghai, China
[72]Endoscopy Center and Endoscopy Research Institute, Zhongshan Hospital, Fudan University, Shanghai, Shanghai, China
[73]Department of Surgery, Paracelsus Medical University, Salzburg, Austria
[74]Department of Digestive Tract Diseases, Medical University of Lodz, Lodz, Poland
[75]Center of Human Genetics, University Hospital of Giessen and Marburg Campus Marburg, Marburg, Hessen, Germany
[76](6) Immunology and Gastroenterology Departments, Instituto de Investigacion Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain
[77]Hepatogastroenterology, Institute of Clinical and Experimental medicine, Praha, Czech Republic
[78]KU Leuven Department of Chronic Diseases Metabolism and Ageing, Laboratory for Intestinal Neuroimmune Interactions, Translational Research Center for Gastrointestinal Disorders, Leuven, Belgium

Contributors Manuscript (writing first draft): SGr, PDa, ZLi, CMa, GBo, JSc. Samples/phenotypic data acquisition: IGo, Ala, AMo, MWo, ZVa, SHa, TTr, NKr, SNi, RTh, HLo, EQu, AMe, MMu, PAm, GGo, SVa, ELe, FMi, MVi, NSc, OPa, FTa, RGi, DGa, AAr, BAr, MBr, JWa, MGa, MJa, MSi, KAd, TMa, ADą, MMa, PGi, MGo, JSe, LMa, VGi, LKi, LFr, LRe, JSp, JTa, MIn, CGr, ARu, JLa, ARa, AEc, MMü, Twi, JKe, CBr, CGe, HNe, Trö, BSi, BSc, MVe, ALe, RRo, VAn, UFu, LLa, EUr, FVa, SRo, PZh, QLi, ZLi, Bra, DTh, EMa, AVi, JMa, GBo. Analysis: SGr, JTr, THe, JBi, SHe, MNö, MKn, ZLi, PDa, CMa, GBo, JSc. Guarantor: JSc.

Funding We thank all patients and controls for participating in this study. This study was supported by the German Research Foundation (DFG) (grant number SCHU 1596/9-1 and KN 378/4-1). This study was supported by the following (awarded to NSc): IIRG-09-133014 from the Alzheimer's Association and 189 10,276/8/9/2011 from the National Strategic Reference Framework - EU programme; Excellence Grant which is co-funded by the European Social Fund and Greek; National resources and ΔΥ2β/οικ.51657/14.4.2009 from the Ministry for Health and Social Solidarity (Greece). We acknowledge the French Society of Gastroenterology (SNFGE, Société Nationale Française de GastroEntérologie) for supporting clinical data collection through a research grant (COMAD 2010). This grant was awarded to SVa, GGo, SRo and FMi who contributed to the clinical data collection. Biological samples from Lyon, France, were provided by the NeuroBioTec – Hospices Civils de Lyon biobank in accordance with a material transfer agreement. We thank the many study participants, researchers and staff for collecting and contributing to the data, the high-performance computing service at the University of Lille and the staff at CEA-CNRGH for their help with sample preparation and genotyping and excellent technical assistance. This work was funded by a grant (EADB) from the EU Joint Programme – Neurodegenerative Disease Research. INSERM UMR1167 is also funded by the INSERM, Institut Pasteur de Lille, Lille Métropole Communauté Urbaine and French government's LABEX DISTALZ programme (development of innovative strategies for a transdisciplinary approach to AD).

Competing interests MBr has declared the following conflicts of interest: he has received consultancy fees and support for both industry and investigator-initiated studies from Boston Scientific and Cook Medical. He has also served as a consultant and received support for investigator-initiated studies from Pentax Medical and AMBU. Additionally, he has received support for investigator-initiated studies from Mylan and ChiRoStim. BSi consulted for AbbVie, Abivax, Boehringer Ingelheim, Bristol Myers Squibb, Dr. Falk Pharma, Eli Lilly, Endpoint Health, Falk, Galapagos, Gilead, Janssen, Landos, Lilly, Materia Prima, PredictImmune, Pfizer and Takeda. BSi also received speaker fees from AbbVie, AlfaSigma, BMS, CED Service GmbH, Dr. Falk Pharma, Eli Lilly, MSD, Ferring, Galapagos, Janssen, Pfizer and Takeda. Additionally, BSi was supported in the past by grants from Pfizer. GBo and TRö hold positions as Editor or an Editorial Board Member of the journal GUT. The remaining authors declare no competing interests.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the Faculties of Medicine at the Universities of Bonn (Germany) (061/17) and Marburg (Germany) (34/19). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iDs
Sandeep Grover https://orcid.org/0000-0003-2615-4916
Guillaume Gourcerol https://orcid.org/0000-0001-8220-9155
Roberto De Giorgio https://orcid.org/0000-0003-0867-5873
Jan Tack https://orcid.org/0000-0002-3206-6704
Jutta Keller https://orcid.org/0000-0002-5884-1115
Thomas Rösch https://orcid.org/0000-0003-2270-2495
Britta Siegmund https://orcid.org/0000-0002-0055-958X
Marino Venerito https://orcid.org/0000-0001-8581-0974
Luigi Laghi https://orcid.org/0000-0003-4187-1059
Sabine Roman https://orcid.org/0000-0002-7798-7638
Pinghong Zhou https://orcid.org/0000-0002-5434-0540
Quanlin Li https://orcid.org/0000-0002-9108-8786
Jan Martinek https://orcid.org/0000-0002-1415-4719
Guy Boeckxstaens https://orcid.org/0000-0001-8267-5797

REFERENCES

1 Boeckxstaens GE, Zaninotto G, Richter JE. Achalasia. *Lancet* 2014;383:83–93.
2 Gockel I, Müller M, Schumacher J. Achalasia--a disease of unknown cause that is often diagnosed too late. *Dtsch Arztebl Int* 2012;109:209–14.
3 Gockel HR, Schumacher J, Gockel I, *et al*. Achalasia: will genetic studies provide insights? *Hum Genet* 2010;128:353–64.
4 Clark SB, Rice TW, Tubbs RR, *et al*. The nature of the myenteric infiltrate in achalasia: an immunohistochemical analysis. *Am J Surg Pathol* 2000;24:1153–8.
5 Boeckxstaens GE. Achalasia: virus-induced euthanasia of neurons? *Am J Gastroenterol* 2008;103:1610–2.
6 Facco M, Brun P, Baesso I, *et al*. T cells in the myenteric plexus of achalasia patients show a skewed TCR repertoire and react to HSV-1 antigens. *Am J Gastroenterol* 2008;103:1598–609.
7 Gockel I, Becker J, Wouters MM, *et al*. Common variants in the HLA-DQ region confer susceptibility to idiopathic achalasia. *Nat Genet* 2014;46:901–4.
8 Marees AT, de Kluiver H, Stringer S, *et al*. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *Int J Methods Psychiatr Res* 2018;27:e1608.
9 Bellenguez C, Küçükali F, Jansen IE, *et al*. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet* 2022;54:412–36.
10 Mbatchou J, Barnard L, Backman J, *et al*. Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet* 2021;53:1097–103.
11 Wang G, Sarkar A, Carbonetto P, *et al*. A simple new approach to variable selection in regression, with application to genetic fine mapping. *J R Stat Soc Series B Stat Methodol* 2020;82:1273–300.
12 Bycroft C, Freeman C, Petkova D, *et al*. The UK Biobank resource with deep phenotyping and genomic data. *Nature New Biol* 2018;562:203–9.
13 Kurki MI, Karjalainen J, Palta P, *et al*. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature New Biol* 2023;613:508–18.
14 Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc* 2020;15:2759–72.
15 Liu Z-Q, Dai H, Yao L, *et al*. A single-cell transcriptional landscape of immune cells shows disease-specific changes of T cell and macrophage populations in human achalasia. *Nat Commun* 2023;14:4685.
16 Zhang MJ, Hou K, Dey KK, *et al*. Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data. *Nat Genet* 2022;54:1572–80.
17 Hu X, Deutsch AJ, Lenz TL, *et al*. Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes risk. *Nat Genet* 2015;47:898–905.
18 Bottini N, Peterson EJ. Tyrosine phosphatase PTPN22: multifunctional regulator of immune signaling, development, and disease. *Annu Rev Immunol* 2014;32:83–119.
19 Anderson W, Barahmand-Pour-Whitman F, Linsley PS, *et al*. *PTPN22* R620W gene editing in T cells enhances low-avidity TCR responses. *Elife* 2023;12:e81577.
20 Wen YP, Yu ZG. Identifying shared genetic loci and common risk genes of rheumatoid arthritis associated with three autoimmune diseases based on large-scale cross-trait genome-wide association studies. *Front Immunol* 2023;14:1160397.
21 Wang Y, Yang Y, Jia X, *et al*. Identification of the shared genetic architecture underlying seven autoimmune diseases with GWAS summary statistics. *Front Immunol* 2023;14:1303675.

22 Khalil H, Brunet A, Saba I, *et al*. The MHC class II beta chain cytoplasmic tail overcomes the invariant chain p35-encoded endoplasmic reticulum retention signal. *Int Immunol* 2003;15:1249–63.
23 Pos W, Sethi DK, Call MJ, *et al*. Crystal structure of the HLA-DM-HLA-DR1 complex defines mechanisms for rapid peptide selection. *Cell* 2012;151:1557–68.
24 Lang B, Navarrete C, LoGalbo PR, *et al*. Further DNA sequence microheterogeneity of the HLA-DR4/Dw13 haplotype group: importance of amino acid position 86 of the DR beta 1 chain for T-cell recognition. *Hum Immunol* 1990;27:378–89.
25 Newton-Nash DK, Eckels DD. Differential effect of polymorphism at HLA-DR1 beta-chain positions 85 and 86 on binding and recognition of DR1-restricted antigenic peptides. *J Immunol* 1993;150:1813–21.
26 Teutsch SM, Bennetts BH, Buhler MM, *et al*. The DRB1 Val86/Val86 genotype associates with multiple sclerosis in Australian patients. *Hum Immunol* 1999;60:715–22.
27 Cocco E, Sardu C, Pieroni E, *et al*. HLA-DRB1-DQB1 haplotypes confer susceptibility and resistance to multiple sclerosis in Sardinia. *PLoS One* 2012;7:e33972.
28 Drenovska K, Ivanova M, Vassileva S, *et al*. Association of specific HLA alleles and haplotypes with pemphigus vulgaris in the Bulgarian population. *Front Immunol* 2022;13:901386.
29 Hov JR, Kosmoliaptsis V, Traherne JA, *et al*. Electrostatic modifications of the human leukocyte antigen-DR P9 peptide-binding pocket and susceptibility to primary sclerosing cholangitis. *Hepatology* 2011;53:1967–76.
30 Marín ND, García LF. The role of CD30 and CD153 (CD30L) in the anti-mycobacterial immune response. *Tuberculosis (Edinb)* 2017;102:8–15.
31 Roll L, Faissner A. Tenascins in CNS lesions. *Semin Cell Dev Biol* 2019;89:118–24.
32 Xu WD, Li R, Huang AF. Role of TL1A in Inflammatory Autoimmune Diseases: A Comprehensive Review. *Front Immunol* 2022;13:891328.
33 Zhang Y, Park E, Kim CS, *et al*. ZNF365 promotes stalled replication forks recovery to maintain genome stability. *Cell Cycle* 2013;12:2817–28.
34 Hou S, Du L, Lei B, *et al*. Genome-wide association analysis of Vogt-Koyanagi-Harada syndrome identifies two new susceptibility loci at 1p31.2 and 10q21.3. *Nat Genet* 2014;46:1007–11.