

## GENERAL ARTICLE

# Integrating Mendelian randomization and multiple-trait colocalization to uncover cell-specific inflammatory drivers of autoimmune and atopic disease

Lucy M. McGowan<sup>1</sup>, George Davey Smith<sup>2</sup>, Tom R. Gaunt<sup>2</sup> and Tom G. Richardson<sup>2,\*</sup>,<sup>†</sup>

<sup>1</sup>School of Physiology, Pharmacology and Neuroscience, Faculty of Life Sciences, University of Bristol, Bristol, BS8 1TD, UK and <sup>2</sup>MRC Integrative Epidemiology Unit, Population Health Sciences Institute, University of Bristol, Bristol, BS8 2BN, UK

\*To whom correspondence should be addressed at: MRC Integrative Epidemiology Unit, Population Health Sciences Institute, University of Bristol, Bristol, BS8 2BN, UK. Tel: +44 1173313370; Fax: +44 1173310123; Email: Tom.G.Richardson@bristol.ac.uk

## Abstract

Immune-mediated diseases (IMDs) arise when tolerance is lost and chronic inflammation is targeted towards healthy tissues. Despite their growing prevalence, therapies to treat IMDs are lacking. Cytokines and their receptors orchestrate inflammatory responses by regulating elaborate signalling networks across multiple cell types making it challenging to pinpoint therapeutically relevant drivers of IMDs. We developed an analytical framework that integrates Mendelian randomization (MR) and multiple-trait colocalization (moloc) analyses to highlight putative cell-specific drivers of IMDs. MR evaluated causal associations between the levels of 10 circulating cytokines and 9 IMDs within human populations. Subsequently, we undertook moloc analyses to assess whether IMD trait, cytokine protein and corresponding gene expression are driven by a shared causal variant. Moreover, we leveraged gene expression data from three separate cell types (monocytes, neutrophils and T cells) to discern whether associations may be attributed to cell type-specific drivers of disease. MR analyses supported a causal role for IL-18 in inflammatory bowel disease (IBD) ( $P = 1.17 \times 10^{-4}$ ) and eczema/dermatitis ( $P = 2.81 \times 10^{-3}$ ), as well as associations between IL-2 $\alpha$  and IL-6R with several other IMDs. Moloc strengthened evidence of a causal association for these results, as well as providing evidence of a monocyte and neutrophil-driven role for IL-18 in IBD pathogenesis. In contrast, IL-2 $\alpha$  and IL-6R associations were found to be T cell specific. Our analytical pipeline can help to elucidate putative molecular pathways in the pathogenesis of IMDs, which could be applied to other disease contexts.

<sup>†</sup>Tom G. Richardson, <http://orcid.org/0000-0002-7918-2040>

Received: March 17, 2019. Revised: June 7, 2019. Accepted: June 26, 2019

© The Author(s) 2018. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

## Introduction

Autoimmune and atopic diseases may arise due to a lack of immune tolerance towards self-antigen or harmless allergens, respectively (1). Loss of immune tolerance results in recurrent or chronic inflammation, causing damage to healthy tissues and extensive morbidity. The incidence of immune-mediated diseases (IMDs) has drastically increased in recent decades, highlighting the need for a clearer understanding of their pathogenesis and effective drug discovery (2). Cytokines and growth factors (herein referred to as cytokines) are signalling factors that orchestrate the balance between immune homeostasis and inflammation via complex signalling pathways (3). However, traditional observational epidemiological studies are prone to confounding and reverse causation, making it challenging to disentangle causal effects of individual cytokines on IMDs (4).

Genome-wide association studies (GWASs) have been instrumental in identifying large numbers of genetic loci that influence disease risk. This includes associations between genes responsible for the synthesis of cytokines and their receptors with autoimmune diseases such as inflammatory bowel disease (IBD) (5–8), multiple sclerosis (MS) (9–11), rheumatoid arthritis (RA) (12) and systemic lupus erythematosus (13), as well as atopic diseases such as eczema (14) and asthma (15). This suggests that particular inflammatory cytokines may have a causal effect on the development of these diseases (4). Previous studies have not yet integrated genome-wide association and cytokine loci data with cell or tissue-specific gene expression loci data to characterize the molecular basis of IMD pathogenesis. Identifying immune cell-specific disease drivers, as well as putative causal relationships between cytokines and IMDs, will help to elucidate complex IMD pathways and identify drug target candidates for therapeutic intervention. Furthermore, targets supported by evidence from statistical analyses of human genetic data are thought to have double the success rate in clinical development (16).

Mendelian randomization (MR) is an increasingly popular statistical method used to strengthen causal inference with respect to exposure–disease associations within human populations, in the absence of confounding variables. MR uses single nucleotide polymorphisms (SNPs), identified through GWAS, as unconfounded proxies for an exposure of interest, analogous to a randomized controlled trial (17). In this study, we have used a conservative (single SNP) and liberal (multiple SNPs) two-sample MR analysis to investigate associations between 11 circulating inflammatory biomarkers (cytokines or cytokine receptors (18–23); Supplementary Material, Table S1) and 9 IMDs (Supplementary Material, Table S2).

We subsequently applied multiple-trait colocalization (moloc) (24) to investigate whether associations detected by MR are due to IMD and circulating cytokine protein sharing a causal variant. Evidence suggesting that they do can strengthen findings from MR that variation in cytokine protein levels has a causal influence on IMD risk. Furthermore, we also integrated data concerning the gene expression of associated cytokines derived from three different cell types (monocytes, neutrophils and T cells) (25). This allowed us to discern whether associations may be due to genetic drivers that influence IMD risk via changes in cell type-specific gene expression. As such, our findings can help develop mechanistic understanding into the regulatory mechanisms underlying associations from GWAS.

## Results

### MR analyses identify putative causal relationships between circulating cytokine/cytokine receptor levels and IMDs

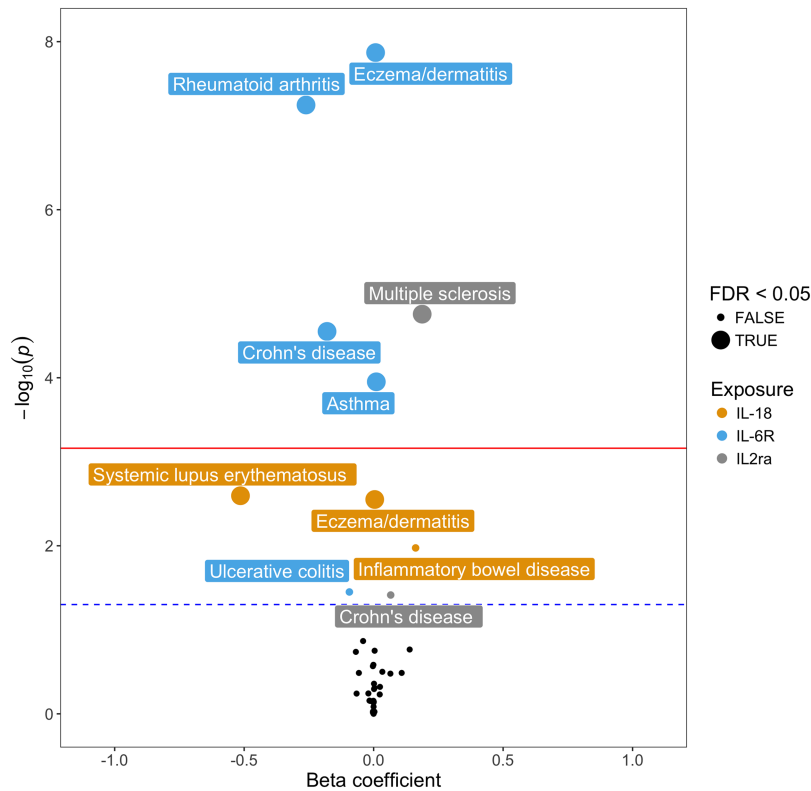
We first used a conservative MR approach to detect associations between inflammatory biomarkers and IMDs, using single cis-acting SNPs (i.e. within 1 MB distance of their associated gene) as instruments based on findings from a previous GWAS that reported genetic associations circulating cytokine levels (18). Genes assessed were *IL-18*, *IL-2R $\alpha$* , *VEGFA*, *MIF*, *IL-16*, *HGF*, *IL-6R*, *CRP* and *IL-1R*. We were unable to investigate *IP-10* and *TRAIL* as their cis-acting SNPs had minor allele frequencies too rare (<0.05) to undertake formal two-sample MR. Results from the conservative MR analysis are displayed in Fig. 1 and Supplementary Material, Table S3. Based on a Bonferroni corrected P-value threshold of  $P \leq 6.17 \times 10^{-4}$ , we identified associations between soluble IL-6R levels and eczema/dermatitis ( $P = 1.35 \times 10^{-8}$ ), RA ( $P = 5.67 \times 10^{-8}$ ), Crohn's disease ( $P = 2.81 \times 10^{-5}$ ) and asthma ( $P = 1.12 \times 10^{-4}$ ), as well as an association between *IL-2R $\alpha$*  levels and MS ( $P = 1.75 \times 10^{-5}$ ).

We next performed a liberal MR analysis using all available instruments (i.e. acting in either cis or trans), associated with our inflammatory biomarkers (Exposure, Supplementary Material, Table S1) as instruments. *IL-6R* (SNPs affecting levels of soluble IL-6R only), *MIF* and *IL-2R $\alpha$*  were excluded from the liberal analysis as they only had a single cis-acting instrument after undertaking linkage disequilibrium (LD) clumping of  $r^2 < 0.001$ . Our Bonferroni corrected P-value threshold for the number of Liberal MR analyses undertaken was  $P \leq 6.14 \times 10^{-4}$ . There was strong evidence of association between circulating levels of *IL-18* and IBD ( $P = 1.17 \times 10^{-4}$ ) and eczema/dermatitis ( $P = 2.81 \times 10^{-3}$ ). Although these associations were only moderately detected in the conservative analysis using the *IL-18* SNP (rs71478720) alone ( $P = 1.06 \times 10^{-2}$ , E/D:  $P = 2.81 \times 10^{-3}$ ), using multiple instruments in the liberal MR provided stronger evidence of association.

One of the advantages of the liberal MR analysis is that sensitivity analyses can be performed to test the robustness and the direction of putative inferred causal relationships. Thus, we next conducted a leave-one-out analysis to ensure that no single SNP from the instruments was responsible for the observed effect. Both the analyses of *IL-18* on IBD (Fig. 2) and on eczema/dermatitis (Supplementary Material, Fig. S2) survived leave-one-out analyses (Supplementary Material, Table S5), as the removal of any individual SNP from the analysis had little effect on observed effect estimates. These results provide evidence to support a causal role for circulating *IL-18* levels in the pathogenesis of IBD and eczema/dermatitis. Moreover, reverse MR (Supplementary Material, Table S6) and the MR directionality test (26) (Supplementary Material, Table S7) showed that reverse causation was unlikely for any of the associations identified in either the conservative or liberal MR analysis.

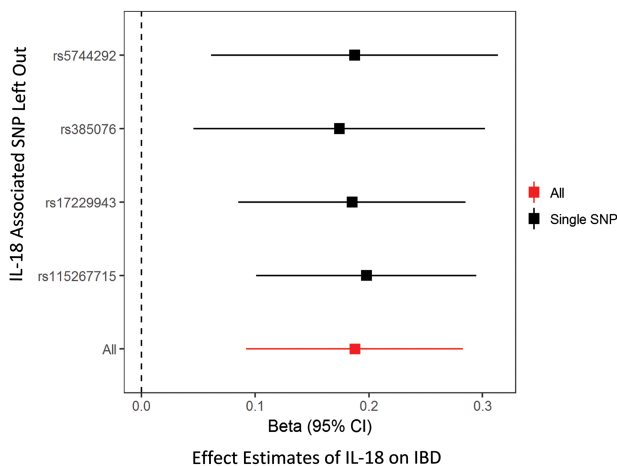
### Moloc uncovers immune cell-specific drivers of IMD

For cytokines and IMDs where evidence of a causal association was detected using MR, we applied the moloc method (24) to test whether gene expression for the gene encoding the cytokine, the cytokine protein itself and associated IMD trait all shared the same causal variant. Supplementary Material, Fig. S1B provides an overview of this analysis. At each loci assessed, expression quantitative trait loci (eQTL), protein quantitative trait loci (pQTL) and GWAS effect estimates were used by moloc to assess whether they shared a common causal variant. Furthermore,



**Figure 1.** Conservative MR analysis detects associations between circulating inflammatory biomarkers and IMDs. Volcano plot of conservative MR analysis illustrating associations between inflammatory cytokines and complex traits. Red (upper) line represents Bonferroni corrected threshold ( $P \leq 6.17 \times 10^{-4}$ ) and black dotted line (bottom) represents unadjusted threshold ( $P \leq 0.05$ ).

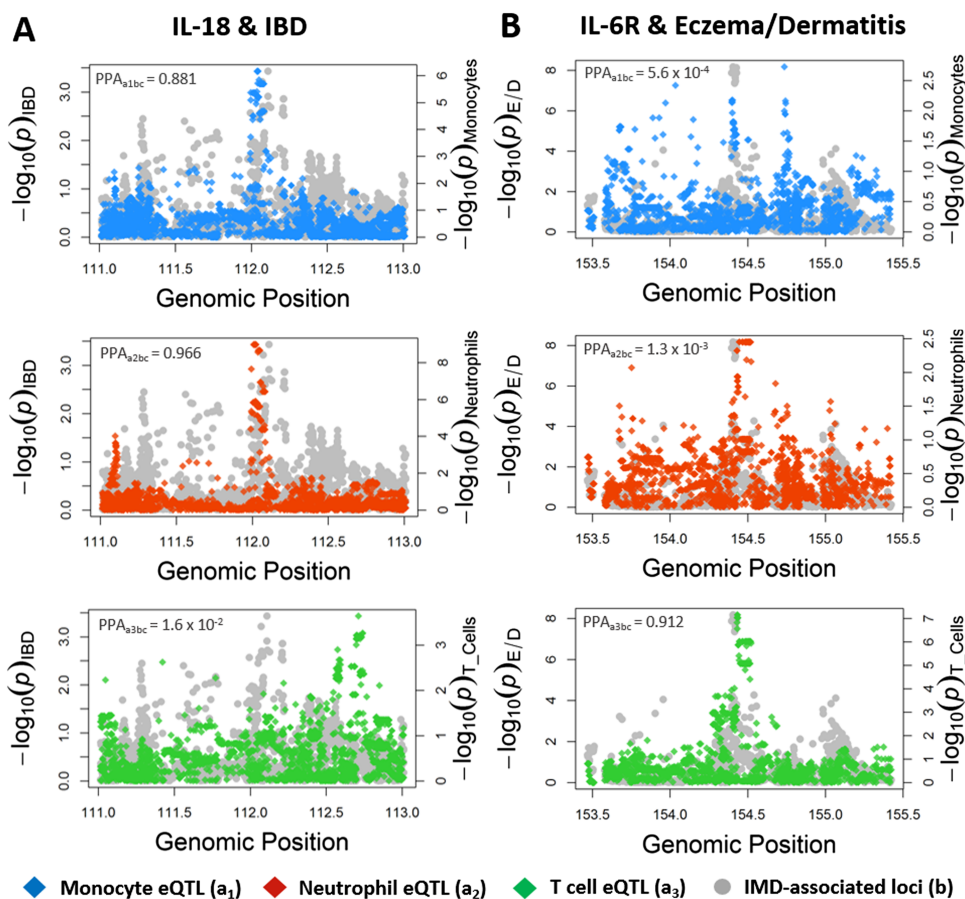
the moloc analysis was performed 3 times for each association, using eQTL from the BLUEPRINT project derived from either human  $CD14^+ CD16^-$  monocytes,  $CD16^+ CD66b^+$  neutrophils or  $CD4^+ CD45RA^+$  T cells (25).



**Figure 2.** Liberal MR identifies a putative causal relationship between circulating levels of IL-18 and IBD that survives leave-one-out sensitivity analysis. Leave-one-out MR analysis for SNPs used as instruments for liberal MR analysis (black). Results show that this effect is not likely to be due to an individual SNP when compared to the observed effect of all SNPs (red). All SNPs in the leave one out analysis are *trans*-acting.

We identified evidence of moloc between 10 combinations of immune cell eQTL, IMD-associated loci and inflammatory biomarker pQTL based on a posterior probability of association (PPA)  $\geq 80\%$  (Supplementary Material, Tables S8 and S9). SNPs from each dataset used in the analysis with the strongest P-values were reported where evidence of colocalization was detected (Supplementary Material, Table S10). Results indicated that IL-18 plays a role in IBD risk due to changes in gene expression within monocytes (PPA<sub>abc</sub> = 88.12%) and neutrophils (PPA<sub>abc</sub> = 96.57%), but not T cells (PPA<sub>abc</sub> = 1.59%) (Fig. 3A). Additionally, we found evidence of a T cell-specific role in the disease pathways of eczema/dermatitis (PPA<sub>abc</sub> = 91.15%) driven by soluble IL-6R (Fig. 3B). Likewise, evidence of moloc was identified between IL-2R $\alpha$  and MS using T cell-derived eQTL data (PPA<sub>abc</sub> = 88.26%), although this gene was not expressed in monocytes or neutrophils based on findings from the BLUEPRINT study.

We also repeated analyses except using tissue-specific eQTL data derived from the GTEx consortium (27). The association between IL-18 and IBD was relatively ubiquitous, as evidence of colocalization was observed within seven diverse tissue types (Supplementary Material, Table S11), including thyroid tissue (Supplementary Material, Fig. S3a). Evidence of colocalization for the association between soluble IL-6R and eczema/dermatitis was observed in three tissue types, most strongly in whole blood (PPA = 99.52%), which may help shed light on the pleiotropic effects observed at this locus (Supplementary Material, Fig. S3b). Lastly, the association between IL-2R $\alpha$  and MS colocalized in two tissue types: subcutaneous adipose and spleen (Supplementary Material, Fig. S3c).



**Figure 3.** Moloc reveals immune cell-specific drivers of IMDs using immune cell eQTL data, inflammatory cytokine or cytokine receptor pQTL data and IMD GWAS data. These plots illustrate observed effects of genetic variants at the IL-18 (A) and IL-6R (B) loci on IBD and eczema/dermatitis (E/D), respectively. Effect estimates on the expression of IL-18 and IL6R are overlaid in each plot using eQTL data derived from monocytes (top, a<sub>1</sub>), neutrophils (middle, a<sub>2</sub>) and T cells (bottom, a<sub>3</sub>). For simplicity, circulating cytokine effects are not displayed within the plots but were used to calculate PPA<sub>abc</sub> scores. PPA<sub>abc</sub> values reflect the likelihood that a causal variant influences the target cytokine (b), associated complex trait (c) and the expression of the corresponding gene (a). PPA<sub>abc</sub> ≥ 0.8 indicates evidence of colocalization (i.e. a shared genetic variant between all three signals) and suggests that the cytokine (or its receptor) is a putative driver of the IMD when it is expressed within the cell type of interest.

However, other tissue-specific relationships may be masked by the lack of statistical power in current tissue eQTL datasets. As increasing sample sizes of tissue-specific eQTL data emerge, this method could be utilized to test the relationships within a larger number of tissue types.

## Discussion

The ways in which hundreds of functionally diverse cytokines interact and orchestrate inflammatory responses in the pathogenesis of IMDs are unclear (28). In this study, we developed a framework that integrates MR with moloc to gain insights into the molecular basis of IMD pathogenesis. Using this framework, we found evidence to support causal relationships between levels of circulating IL-18 and IBD, as well as eczema/dermatitis, circulating soluble IL-6R and eczema/dermatitis and circulating IL-2R $\alpha$  and MS, amongst others (Supplementary Material, Tables S3–4, S6–8). Additionally, we provided evidence to suggest that these associations are likely to be driven in an immune cell-specific manner and explored the potential of this method to pinpoint tissue-specific drivers of IMDs.

A T cell-mediated role for IL-2/IL-2R $\alpha$  in MS has already been well established through epidemiological and lab-based studies (10,29). The IL-2R-targeting drug daclizumab was given FDA (food and drug administration) approval in the United States for the treatment of MS, but was recently withdrawn due to serious side effects (30–32). Additionally, a causal role for IL-18R in atopic dermatitis has recently been described, by integrating MR and pQTL data (33). Our results provide additional evidence to support these existing findings, as well as identifying monocytes and neutrophils as potential drivers of the relationship between IL-18 and eczema/dermatitis (Supplementary Material, Table S8). However, the associations between soluble IL-6R variation and eczema/dermatitis or IL-18 and IBD within human populations are less well understood. Our analyses not only help establish evidence for causal relationships between these inflammatory biomarkers and IMDs but also help characterize their cell type-specific nature.

IL-18, a member of the IL-1 superfamily of cytokines, is a potent inducer of Th1-mediated inflammation and IFN- $\gamma$  production (34). This cytokine was first linked to IBD nearly 20 years ago, where it was shown to be highly expressed in intestinal tissues derived from IBD patients, compared to healthy control patients (35). Using murine IBD models, deletion

of *il-18* or its receptor *il-18r1* has been shown to be protective against inducible colitis, by controlling goblet cell function and maintaining intestinal barrier homeostasis (36). Despite strong evidence to suggest a role for IL-18 in IBD in mice, whether there was a causative role within human populations remained unclear. Our study provides evidence that IL-18 is likely the causal gene responsible for the association with IBD at this locus, as well as out MR analysis supporting a causal role for IL-18 in the disease pathogenesis of IBD within human populations. Furthermore, the moloc analyses suggested that innate immune cells such as monocytes and neutrophils are likely to drive this association, supporting the current dogma that innate production of IL-18 stimulates Th1/Th17-mediated autoimmunity in IBD (37). The T cell-specific eQTL data currently available for moloc analyses concerned naïve CD4<sup>+</sup> T cells (38). More eQTL data concerning activated and differentiated immune cell subsets, such as macrophages, dendritic cells and T helper cell subsets (i.e. Th1, Th2, Th17, Treg), are required for additional immune cell subset-specific moloc analysis, to further elucidate the molecular pathways which drive IMDs. Interestingly, small molecule inhibitors that target and repress IL-18-mediated signalling events are currently under development, although not for the treatment of IBD (39). A recently published phase II trial of an IL-18 binding protein (IL-18 bp) drug to treat adult-onset Still's disease demonstrated a favourable efficacy safety profile (40). If further trials are deemed successful, IL-18 bp drugs may also be used to target other IMDs such as IBD.

IL-6R is the receptor of the pro-inflammatory cytokine IL-6, which can exist in a membrane-bound state on the surface of leukocytes and hepatocytes (classical) or in soluble form (*trans*) (41). Both classical and *trans* IL-6R signalling culminates in the expression of signal transducer and activator of transcription 3, which promotes inflammation via the expression of genes encoding antiapoptotic proteins and cytokines (42). However, classical IL-6 signalling can act on few cell types compared to *trans* IL-6 signalling, which can act on any cell that has the cell-bound signal transducer, gp130 (43). One GWAS previously reported an association between elevated soluble IL-6R levels resulting from an SNP in *IL-6R* and atopic dermatitis (44). We provide evidence from the conservative MR and moloc analyses that supports a causal role between soluble IL-6R and eczema/dermatitis. The ratio of classical to *trans* IL-6R fluctuates as a result of levels cell-bound gp130, receptor shedding by cells expressing membrane-bound IL-6R and changes to the amount of receptor being synthesized by cells (43). Therefore, analyses combining instruments that affect the classical or *trans*-signalling pathways would provide a clearer insight into the role of IL-6 signalling in disease. Interestingly, activated CD4<sup>+</sup> T cells have been shown to increase levels of soluble IL-6R via IL-6R shedding; this mechanism is thought to have a role in the development of autoimmune diseases, which are often mediated by autoreactive T cells (45,46). Through the moloc analysis, we showed that eczema/dermatitis was likely to be driven by IL-6R expression in T cells (Fig. 3B; Supplementary Material, Table S8). This finding is supported by evidence of increased IL-6R shedding leading to an increase in soluble IL-6R in people diagnosed with atopic dermatitis compared to healthy controls (44).

Relatively few SNPs have been associated with changes in the circulating inflammatory biomarkers investigated in this study, as most of the GWAS summary statistics used to identify our instruments were derived from cohorts with fewer than 9000 people (18,19,23); this is likely due to the high cost of quantification of circulating cytokines and cytokine receptors from blood.

Our reverse MR analysis may therefore have been underpowered to evaluate evidence of reverse causation, although supplementing this analysis using the Steiger directionality test also suggested that this was unlikely for the associations we identified. As sample sizes for GWAS of circulating cytokines increase, the analysis pipeline illustrated by this study will have further power to detect novel relationships between markers of inflammation and complex disease. Moreover, our moloc analysis assumes only one causal variant is responsible for associations. While current sample sizes of molecular traits can typically only be instrumented using a single independent *cis*-QTL, future studies uncovering multiple independent QTL on a large scale should benefit from alternative approaches to genetic colocalization. Furthermore, applying MR analyses as proposed in our study in a phenome-wide manner can help elucidate potential adverse side effects of therapeutic intervention (47).

In conclusion, we have found strong evidence supporting new and known causal, immune cell-driven relationships between inflammatory biomarkers and IMDs. Triangulation of results from these analyses, with published results from experimental models of IMDs and GWAS, suggests that targeting IL-18 or its receptor, IL-18R, may be promising for the treatment of IBD. We believe our analysis framework could be applied by other studies with alternative hypotheses, as a way of disentangling complex biochemical cell signalling pathways and identifying molecules and cell types that are likely to drive disease for drug target prioritization.

## Materials and Methods

### Data sources

For our two-sample MR analysis, we harnessed genetic instrument data for 10 circulating inflammatory cytokines or soluble cytokine receptors (Supplementary Material, Table S1) from summary statistics of previously published GWAS (18,19,23,48). The SNP used as an instrument for *IL-6R* affected levels of soluble IL-6R. The SNPs chosen in this analysis have been shown to be robustly associated with a change in circulating levels of a cytokine ( $P < 5 \times 10^{-8}$ ) and are in *cis* with the gene of interest (i.e. the SNP was located within a 1 MB distance of the gene that encoded the cytokine or cytokine receptor of interest). Data concerning IMD outcomes (Supplementary Material, Table S2) were derived from large-scale GWAS using the MR-Base platform (49). For the moloc analysis, we harnessed human monocyte (CD14<sup>+</sup> CD16<sup>-</sup>), neutrophil (CD16<sup>+</sup> CD66b<sup>+</sup>) and T cell (CD4<sup>+</sup> CD45RA<sup>+</sup>) eQTL data from the BLUEPRINT epigenome project (38). All data used were derived from populations of European descent.

### Mendelian randomization

MR follows three assumptions: (1) that the selected instruments used are robustly associated with the exposure, (2) that the selected instruments are unconfounded and (3) that the selected instruments can only influence the outcome via the exposure. Using randomly inherited unmodifiable SNPs associated with circulating inflammatory cytokine levels through GWAS as genetic instruments for MR satisfies assumptions 1 and 2. We performed two-sample MR with the MR-Base platform (49), using two different analysis methods (Supplementary Material, Fig. S1A) as described below, depending upon available data.

Conservative two-sample MR was used to analyze the causal effect of cytokines using SNPs at target genes encoding the inflammatory biomarkers of interest. As such, for this analysis

we used a single SNP acting in cis as a genetic instrument (i.e. located within a 1 MB distance of the target gene encoding the cytokine with  $P < 5 \times 10^{-08}$ ). As only one genetic instrument was used in this analysis, effect estimates were based on the Wald ratio test (50). For single genetic variant  $j$ , this can be calculated by taking the ratio of the gene–outcome association (denoted by  $\hat{\Gamma}_j$ ) and the gene–exposure association (denoted  $\hat{\gamma}_j$ ) estimates

$$\hat{\beta}_j = \frac{\hat{\Gamma}_j}{\hat{\gamma}_j}.$$

Where possible, liberal MR was also used to analyze the causal effect of circulating inflammatory biomarkers on IMDs. In contrast to the conservative MR analysis, liberal MR used multiple SNPs as genetic instruments that were either acting in cis or trans (i.e. over 1 MB distance from the target gene encoding the cytokine with  $P < 5 \times 10^{-08}$ ).

To identify instruments for the liberal MR we undertook genome-wide LD clumping based on  $P < 5 \times 10^{-08}$  and  $r^2 < 0.001$ . A leave-one-out MR analysis was performed in parallel with liberal MR to ensure that causal effects were not observed due to the influence of a single SNP. As two or more genetic instruments were available for liberal MR, we used the inverse variance weighted method to obtain MR effect estimates (50).

$$\hat{\beta}_{IVW} = \frac{\sum_j \hat{\gamma}_j^2 \sigma_{\gamma_j}^{-2} \hat{\beta}_j}{\sum_j \hat{\gamma}_j^2 \sigma_{\gamma_j}^{-2}},$$

where  $\sigma_{\gamma_j}$  is the standard error of the gene–outcome estimate for variant  $j$ .

For both conservative and liberal MR analyses, we used the Bonferroni correction to calculate an adjusted  $P$ -value threshold (conservative:  $P \leq 6.17 \times 10^{-4}$ , liberal:  $P \leq 6.94 \times 10^{-4}$ ) to account for multiple testing, where

$$\text{Threshold} = \frac{0.05}{\text{No.Exposures} \times \text{No.Outcomes}}.$$

We subsequently undertook reverse MR analyses for all associations that survived multiple testing to investigate reverse causation (i.e. the likelihood that genetic predisposition to disease has an influence on inflammatory cytokine levels). As an additional sensitivity analysis, we used the MR Steiger directionality test to assess the directionality of associations between inflammatory cytokines and complex traits (26).

### Multiple-trait colocalization analysis

We applied the moloc approach (24) to immune cell-specific eQTL data, inflammatory biomarker pQTL data and IMD-associated loci data to identify putative immune cell-specific drivers of IMDs (Supplementary Material, Fig. S1B). The moloc method uses a Bayesian statistical framework to calculate PPA scores to measure the degree of colocalization between gene loci using three or more datasets. PPA scores of  $\geq 80\%$  are likely to share a common genetic causal variant based on evaluations undertaken by the authors (24). We obtained eQTL data derived from concerning neutrophils (CD16<sup>+</sup> CD66b<sup>+</sup>), monocytes (CD14<sup>+</sup> CD16<sup>-</sup>) and T cells (CD4<sup>+</sup> CD45RA<sup>+</sup>) (38). We ran independent analyses for each of the three immune cell types, testing the degree of colocalization between immune cell eQTL, with inflammatory biomarker pQTL and IMD-associated loci.

We chose to perform moloc for inflammatory biomarkers and IMDs in this instance based upon a  $P$ -value threshold  $< 0.05$  in either the conservative or liberal MR analysis. Along with investigating cell-type specificity for identified associations, this analysis was used to detect evidence of a coordinated system that is consistent with causality (i.e. gene expression and respective protein products colocalize with the associated complex traits). As such these findings can complement evidence from MR to detect putative causal effects between biomarkers and disease.

Finally, we undertook exploratory moloc analyses to investigate gene expression within different tissue types using data from the GTEx consortium v6p (51). Our three traits in each analysis were the circulating inflammatory cytokine using pQTL data, the complex trait with strongest evidence of association with the cytokine in either the liberal or conservative MR analyses and tissue-specific gene expression for the cytokine using GTEx eQTL data. We only investigated tissue types with at least one eQTL ( $P < 1.0 \times 10^{-04}$ ) for the target cytokine gene due to the small sample sizes concerning tissue-specific eQTL in the GTEx database. As in the immune cell eQTL moloc analysis, evidence of colocalization was based on a PPA score of  $\geq 80\%$ .

All statistical and bioinformatics analyses were undertaken using R statistical software version 3.31 (52). Plots illustrating multiple trait colocalization were generated using base R graphics, whereas our volcano plot was generated using ggplot (53).

### Supplementary Material

Supplementary Material is available at HMG online.

### Acknowledgements

This study makes use of and is grateful for open-access data generated by the BLUEPRINT Consortium (38) and the GTEx Consortium (51). We are also thankful to the genome-wide association studies that have made their summary statistics available for the benefit of this research.

*Conflict of Interest statement.* T.R.G. receives research funding from GlaxoSmithKline, Biogen and Sanofi.

### Funding

Wellcome Trust Dynamic Molecular Cell Biology PhD Programme at the University of Bristol (108907/Z/15/Z); UK Research and Innovation (UKRI) Innovation Research Fellowship (MR/S003886/1 to T.G.R.); UK Medical Research Council Integrative Epidemiology Unit (MC\_UU\_00011/4 and MC\_UU\_00011/1)

### References

1. Rabin, R.L. and Levinson, A.I. (2008) The nexus between atopic disease and autoimmunity: a review of the epidemiological and mechanistic literature. *Clin. Exp. Immunol.*, **153**, 19–30.
2. Bach, J.F. (2018) The hygiene hypothesis in autoimmunity: the role of pathogens and commensals. *Nat. Rev. Immunol.*, **18**, 105+.
3. Hirahara, K., Schwartz, D., Gadina, M., Kanno, Y. and O'Shea, J.J. (2016) Targeting cytokine signaling in autoimmunity: back to the future and beyond. *Curr. Opin. Immunol.*, **43**, 89–97.

4. Davey Smith, G. and Ebrahim, S. (2003) 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.*, **32**, 1–22.
5. Duerr, R.H., Taylor, K.D., Brant, S.R., Rioux, J.D., Silverberg, M.S., Daly, M.J., Steinhart, A.H., Abraham, C., Regueiro, M., Griffiths, A. et al. (2006) A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science*, **314**, 1461–1463.
6. Franke, A., Balschun, T., Karlsen, T.H., Svventoraityte, J., Nikolaus, S., Mayr, G., Domingues, F.S., Albrecht, M., Nothnagel, M., Ellinghaus, D. et al. (2008) Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. *Nat. Genet.*, **40**, 1319–1323.
7. Festen, E.A.M., Goyette, P., Green, T., Boucher, G., Beauchamp, C., Trynka, G., Dubois, P.C., Lagace, C., Stokkers, P.C.F., Hommes, D.W. et al. (2011) A meta-analysis of genome-wide association scans identifies IL18RAP, PTPN2, TAGAP, and PUS10 as shared risk loci for Crohn's disease and celiac disease. *PLoS Genet.*, **7**, e1001283.
8. Liu, J.Z., van Sommeren, S., Huang, H.L., Ng, S.C., Alberts, R., Takahashi, A., Ripke, S., Lee, J.C., Jostins, L., Shah, T. et al. (2015) Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat. Genet.*, **47**, 979+.
9. Beecham, A.H., Patsopoulos, N.A., Xifara, D.K., Davis, M.F., Kempainen, A., Cotsapas, C., Shah, T.S., Spencer, C., Booth, D., Goris, A. et al. (2013) Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat. Genet.*, **45**, 1353+.
10. Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C.C.A., Patsopoulos, N.A., Moutsianas, L., Dilthey, A., Su, Z., Freeman, C., Hunt, S.E. et al. (2011) Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*, **476**, 214–219.
11. Briggs, F.B.S., Shao, X.R., Goldstein, B.A., Oksenberg, J.R., Barcellos, L.F., De Jager, P.L. and International Multiple Sclerosis Genetics Consortium (2011) Genome-wide association study of severity in multiple sclerosis. *Genes Immun.*, **12**, 615–625.
12. Stahl, E.A., Raychaudhuri, S., Remmers, E.F., Xie, G., Eyre, S., Thomson, B.P., Li, Y.H., Kurzeeman, F.A.S., Zhernakova, A., Hinks, A. et al. (2010) Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat. Genet.*, **42**, 508–556.
13. Benthams, J., Morris, D.L., Graham, D.S.C., Pinder, C.L., Tomblason, P., Behrens, T.W., Martin, J., Fairfax, B.P., Knight, J.C., Chen, L.Y. et al. (2015) Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat. Genet.*, **47**, 1457+.
14. Paternoster, L., Standl, M., Waage, J., Baurecht, H., Hotze, M., Strachan, D.P., Curtin, J.A., Bonnelykke, K., Tian, C., Takahashi, A. et al. (2015) Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nat. Genet.*, **47**, 1449+.
15. Demenais, F., Margartite-Jeannin, P., Barnes, K.C., Cookson, W.O.C., Altmuller, J., Ang, W., Barr, R.G., Beaty, T.H., Becker, A.B., Beilby, J. et al. (2018) Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. *Nat. Genet.*, **50**, 42+.
16. Nelson, M.R., Tipney, H., Painter, J.L., Shen, J.D., Nicoletti, P., Shen, Y.F., Floratos, A., Sham, P.C., Li, M.J., Wang, J.W. et al. (2015) The support of human genetic evidence for approved drug indications. *Nat. Genet.*, **47**, 856+.
17. Davey Smith, G. and Hemani, G. (2014) Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.*, **23**, R89–R98.
18. Ahola-Olli, A.V., Wurtz, P., Havulinna, A.S., Aalto, K., Pitkanen, N., Lehtimaki, T., Kahonen, M., Lyytikainen, L.P., Raitoharju, E., Seppala, I. et al. (2017) Genome-wide association study identifies 27 loci influencing concentrations of circulating cytokines and growth factors. *Am. J. Hum. Genet.*, **100**, 40–50.
19. Sarwar, N., Butterworth, A.S., Freitag, D.F., Gregson, J., Willeit, P., Gorman, D.N., Gao, P., Saleheen, D., Rendon, A., Nelson, C.P. et al. (2012) Interleukin-6 receptor pathways in coronary heart disease: a collaborative meta-analysis of 82 studies. *Lancet*, **379**, 1205–1213.
20. Dehghan, A., Dupuis, J., Barbalic, M., Bis, J.C., Eiriksdottir, G., Lu, C., Pellikka, N., Wallaschofski, H., Kettunen, J., Henneman, P. et al. (2011) Meta-analysis of genome-wide association studies in > 80 000 subjects identifies Multiple loci for C-reactive protein levels. *Circulation*, **123**, 731–U151.
21. Wensley, F., Gao, P., Burgess, S., Kaptoge, S., Di Angelantonio, E., Shah, T., Engert, J.C., Clarke, R., Davey-Smith, G., Nordestgaard, B.G. et al. (2011) Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. *Br. Med. J.*, **342**, d548.
22. Matteini, A.M., Li, J., Lange, E.M., Tanaka, T., Lange, L.A., Tracy, R.P., Wang, Y., Biggs, M.L., Arking, D.E., Fallin, M.D. et al. (2014) Novel gene variants predict serum levels of the cytokines IL-18 and IL-1ra in older adults. *Cytokine*, **65**, 10–16.
23. Hartwig, F.P., Borges, M.C., Horta, B.L., Bowden, J. and Davey Smith, G. (2017) Inflammatory biomarkers and risk of schizophrenia a 2-sample Mendelian randomization study. *JAMA Psychiatry*, **74**, 1226–1233.
24. Giambartolomei, C., Zhenli Liu, J., Zhang, W., Hauberg, M., Shi, H., Boocock, J., Pickrell, J., Jaffe, A.E., CommonMind Consortium, Pasaniuc, B. et al. (2018) A Bayesian framework for multiple trait colocalization from summary association statistics. *Bioinformatics*, **34**, 2538–2545.
25. Martens, J.H. and Stunnenberg, H.G. (2013) BLUEPRINT: mapping human blood cell epigenomes. *Haematologica*, **98**, 1487–1489.
26. Hemani, G., Tilling, K. and Davey Smith, G. (2017) Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.*, **13**, e1007149.
27. Carithers, L.J. and Moore, H.M. (2015) The genotype-tissue expression (GTEx) project. *Biopreserv. Biobank*, **13**, 307–308.
28. Liu, X.B., Fang, L., Guo, T.B., Mei, H.K. and Zhang, J.W.Z. (2013) Drug targets in the cytokine universe for autoimmune disease. *Trends Immunol.*, **34**, 120–128.
29. Hartmann, F.J., Khademi, M., Aram, J., Ammann, S., Kockum, I., Constantinescu, C., Gran, B., Piehl, F., Olsson, T., Codarri, L. et al. (2014) Multiple sclerosis-associated IL2RA polymorphism controls GM-CSF production in human T-H cells. *Nat. Commun.*, **5**, Article number: 5056.
30. Chataway, J., Martin, K., Barrell, K., Sharrack, B., Stolt, P., Wraith, D.C. and ATX-MS1467 Study Group (2018) Effects of ATX-MS-1467 immunotherapy over 16 weeks in relapsing multiple sclerosis. *Neurology*, **90**, E955–E504.
31. MS Society. (2018) <https://www.mssociety.org.uk/about-ms/treatments-and-therapies/disease-modifying-therapies/daclizumab>, Vol. 2018. Date accessed: 17 Aug 2018.

32. Baecher-Allan, C., Kaskow, B.J. and Weiner, H.L. (2018) Multiple sclerosis: mechanisms and immunotherapy. *Neuron*, **97**, 742–768.
33. Sun, B.B., Maranville, J.C., Peters, J.E., Stacey, D., Staley, J.R., Blackshaw, J., Burgess, S., Jiang, T., Paige, E., Surendran, P. et al. (2018) Genomic atlas of the human plasma proteome. *Nature*, **558**, 73+.
34. Kaplanski, G. (2018) Interleukin-18: biological properties and role in disease pathogenesis. *Immunol. Rev.*, **281**, 138–153.
35. Pizarro, T.T., Michie, M.H., Bentz, M., Woraratanadharm, J., Smith, M.F., Foley, E., Moskaluk, C.A., Bickston, S.J. and Cominelli, F. (1999) IL-18, a novel immunoregulatory cytokine, is up-regulated in Crohn's disease: expression and localization in intestinal mucosal cells. *J. Immunol.*, **162**, 6829–6835.
36. Nowarski, R., Jackson, R., Gagliani, N., de Zoete, M.R., Palm, N.W., Bailis, W., Low, J.S., Harman, C.C.D., Graham, M., Elinav, E. et al. (2015) Epithelial IL-18 equilibrium controls barrier function in colitis. *Cell*, **163**, 1444–1456.
37. Kanai, T., Kamada, N. and Hisamatsu, T. (2013) Clinical strategies for the blockade of IL-18 in inflammatory bowel diseases. *Curr. Drug Targets*, **14**, 1392–1399.
38. Chen, L., Ge, B., Casale, F.P., Vasquez, L., Kwan, T., Garrido-Martin, D., Watt, S., Yan, Y., Kundu, K., Ecker, S. et al. (2016) Genetic drivers of epigenetic and transcriptional variation in human immune cells. *Cell*, **167**, 1398+.
39. Krumm, B., Meng, X.Z., Xiang, Y. and Deng, J.P. (2017) Identification of small molecule inhibitors of Interleukin-18. *Sci. Rep.*, **7**, 483.
40. Gabay, C., Fautrel, B., Rech, J., Spertini, F., Feist, E., Kotter, I., Hachulla, E., Morel, J., Schaeffer, T., Hamidou, M.A. et al. (2018) Open-label, multicentre, dose-escalating phase II clinical trial on the safety and efficacy of tadekinig alfa (IL-18BP) in adult-onset Still's disease. *Ann. Rheum. Dis.*, **77**, 840–847.
41. Rincon, M. (2012) Interleukin-6: from an inflammatory marker to a target for inflammatory diseases. *Trends Immunol.*, **33**, 571–577.
42. Hodge, D.R., Hurt, E.M. and Farrar, W.L. (2005) The role of IL-6 and STAT3 in inflammation and cancer. *Eur. J. Cancer*, **41**, 2502–2512.
43. Rose-John, S. (2012) IL-6 trans-Signaling via the soluble IL-6 receptor: importance for the pro-inflammatory activities of IL-6. *Int. J. Biol. Sci.*, **8**, 1237–1247.
44. Esparza-Gordillo, J., Schaarschmidt, H., Liang, L.M., Cookson, W., Bauerfeind, A., Lee-Kirsch, M.A., Nemat, K., Henderson, J., Paternoster, L., Harper, J.I. et al. (2013) A functional IL-6 receptor (IL6R) variant is a risk factor for persistent atopic dermatitis. *J. Allergy Clin. Immunol.*, **132**, 371–377.
45. Briso, E.M., Dienz, O. and Rincon, M. (2008) Cutting edge: soluble IL-6R is produced by IL-6R ectodomain shedding in activated CD4 T cells. *J. Immunol.*, **180**, 7102–7106.
46. Jones, G.W., McLoughlin, R.M., Hammond, V.J., Parker, C.R., Williams, J.D., Malhotra, R., Scheller, J., Williams, A.S., Rose-John, S., Topley, N. et al. (2010) Loss of CD4(+) T cell IL-6R expression during inflammation underlines a role for IL-6 trans Signaling in the local maintenance of Th17 cells. *J. Immunol.*, **184**, 2130–2139.
47. Richardson, T.G., Hemani, G., Gaunt, T.R., Relton, C.L. and Davey Smith, G. (2019) A transcriptome-wide Mendelian randomization study to uncover tissue-dependent regulatory mechanisms across the human phenome. *bioRxiv* doi: <https://doi.org/10.1101/563379>.
48. Folkersen, L., Fauman, E., Sabater-Lleal, M., Strawbridge, R.J., Franberg, M., Sennblad, B., Baldassarre, D., Veglia, F., Humphries, S.E., Rauramaa, R. et al. (2017) Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease. *PLoS Genet.*, **13**, e1006706.
49. Hemani, G., Zheng, J., Elsworth, B., Wade, K.H., Haberland, V., Baird, D., Laurin, C., Burgess, S., Bowden, J., Langdon, R. et al. (2018) The MR-Base platform supports systematic causal inference across the human phenome. *Elife*, **7**, e34408.
50. Bowden, J., Davey Smith, G., Haycock, P.C. and Burgess, S. (2016) Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.*, **40**, 304–314.
51. Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N. et al. (2013) The genotype-tissue expression (GTEx) project. *Nat. Genet.*, **45**, 580–585.
52. R Core Team (2017) <https://www.R-project.org/>, Vol. 2018. Accessed: 16 Aug 2018.
53. Ginestet, C. (2011) ggplot2: elegant graphics for data analysis. *J. R. Stat. Soc. A Stat. Soc.*, **174**, 245–245.