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

Lucy M. McGowan [1,*], George Davey Smith[2], Tom R. Gaunt[2], Tom G. Richardson[2,*]

Abstract

Immune mediated diseases (IMDs) arise from a lack of immune tolerance, causing chronic inflammation. Despite their growing prevalence, targeted therapies to treat IMDs are lacking. Cytokines and their receptors, which mediate inflammation, have been associated with IMD susceptibility. However, the complex signalling networks and multiple cell-types required to orchestrate inflammatory responses have made it difficult to pinpoint specific cytokines and immune cell-types which drive IMDs.

In this study, we developed an analytical framework which integrates Mendelian randomisation (MR) and multiple-trait colocalization (moloc) analyses to determine putative cell-specific drivers of IMDs. We used MR to determine the likelihood of causal associations between the levels of 10 circulating cytokines/cytokine receptors and 9 IMDs within human populations of European descent. Conservative (single SNP) and liberal (multiple SNP) MR analysis supported a causal role for IL-18 in inflammatory bowel disease ($P = 1.17 \times 10^{-4}$) and eczema/dermatitis ($P = 2.81 \times 10^{-3}$), as well as roles for IL-2rα and IL-6R in several IMDs.

Where associations between cytokines/cytokine receptors and IMDs were discovered using MR, we undertook moloc analyses. This was to assess the likelihood that cytokine/cytokine receptor protein quantitative trait loci (pQTL) and IMD-associated loci share a causal genetic variant along with expression QTL (eQTL) using data from 3 immune cell-types: monocytes, neutrophils and T cells. We found a monocyte and neutrophil-driven role for IL-18 in IBD pathogenesis, amongst evidence supporting several other cell-specific inflammatory drivers of IMDs. Our study helps to elucidate causal

[1] Department of Physiology, Pharmacology and Neuroscience, Faculty of Life Sciences, University of Bristol, BS8 1TD, UK.
[2] MRC Integrative Epidemiology Unit, Population Health Sciences Institute, University of Bristol, BS8 2BN, UK.
* Correspondence: lucy.mcgowan@bristol.ac.uk, tom.g.richardson@bristol.ac.uk

pathways for the pathogeneses of IMDs which, together with other studies, highlights possible therapeutic targets for their treatment.

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 pathogeneses and effective drug discovery[2]. Cytokines and growth factors (herein referred to as cytokines) are signalling factors which orchestrate the balance between immune homeostasis and inflammation[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 (GWAS) have been instrumental in identifying large numbers of genetic loci which influence disease risk. This includes associations between loci which harbour 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 (SLE)[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 characterise 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 used a conservative (single SNP) and liberal (multiple SNPs) two-sample MR analysis to investigate associations between 10 circulating inflammatory biomarkers (cytokines or cytokine receptors[18-23]) (Table S1) and 9 IMDs (Table S2). We subsequently used the recently-developed multiple-trait colocalization (moloc) method[24], integrating either immune cell or tissue expression quantitative expression loci (eQTL), cytokine protein QTL (pQTL) and IMD-associated loci data, to identify immune cell and tissue-specific drivers of inflammatory disease.

Methods and Materials

Data Sources

For our two-sample MR analysis, we harnessed genetic instrument data for 10 circulating inflammatory cytokines or soluble cytokine receptors (Table 1) from summary statistics of previously published GWAS[18-20; 25]. 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 1MB distance of the gene which encoded the cytokine or cytokine receptor of interest). Data concerning IMD outcomes (Table 2) were derived from large-scale GWAS using the MR-Base platform[26]. For the moloc analysis, we harnessed human monocyte ($CD14^+$ $CD16^-$), neutrophil ($CD16^+$ $CD66b^+$) and T cell ($CD4^+$ $CD45RA^+$) eQTL data from the BLUEPRINT epigenome project[27]. All data used were derived from populations of European descent.

Mendelian Randomization

Mendelian randomization follows 3 assumptions; that the selected instruments used are robustly associated with the exposure (1), that the selected instruments are unconfounded (2), and that the

selected instruments can only influence the outcome via the exposure (3). 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[26], using two different analysis methods (Figure S1A) as described below, depending upon available data:

Conservative two-sample MR was used to analyse 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 1MB 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:

$$\beta_{Wald\ ratio} = \frac{\hat{\beta}_{Y|Z}}{\hat{\beta}_{X|Z}}$$

$$se(\beta_{Wald\ ratio}) = \sqrt{\frac{se(\hat{\beta}_{Y|Z})^2}{\hat{\beta}_{X|Z}^2} + \frac{\hat{\beta}_{Y|Z}^2 se(\hat{\beta}_{X|Z})^2}{\hat{\beta}_{X|Z}^4} - \frac{2\hat{\beta}_{Y|Z} cov(\hat{\beta}_{X|Z}, \hat{\beta}_{Y|Z})}{\hat{\beta}_{X|Z}^3}}$$

where $\hat{\beta}\ Y|Z$ is the coefficient of the genetic variant in the regression of the exposure (e.g. circulating cytokine level) and $\hat{\beta}\ X|Z$ is the coefficient of the genetic variant in the regression of the outcome.

Where possible, liberal MR was also used to analyse the causal effect of circulating inflammatory biomarkers on immune-mediated diseases. In contrast to the conservative MR analysis, liberal MR used multiple SNPs as genetic instruments which were either acting in *cis* or *trans* (i.e. over 1MB 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 (IVW) method to obtain MR effect estimates:

$$\beta_{IVW} = \frac{\sum_k X_k\ Y_k\ \sigma_{Yk}^{-2}}{\sum_k X_k^2\ \sigma_{Yk}^{-2}}$$

$$se(\hat{\beta}_{IVW}) = \sqrt{\frac{1}{\sum_k X_k^2 \, \sigma_{Yk}^{-2}}}$$

Where X is the exposure, Y is the outcome and the genetic variants are k (where k = 1,…,n).

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:

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

We subsequently undertook reverse MR analyses for all associations which 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[28].

Multiple Trait Colocalization Analysis

We applied the moloc approach[24] to immune cell-specific expression quantitative trait loci (eQTL) data, inflammatory biomarker protein QTL (pQTL) data, and IMD-associated loci data, to identify putative immune cell-specific drivers of IMDs (Figure S1B). The moloc method uses a Bayesian statistical framework to calculate PPA (posterior probability of association) scores to measure the degree of colocalization between gene loci using 3 or more datasets. PPA scores of ≥ 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 ($CD16^+$ $CD66b^+$), monocytes ($CD14^+$ $CD16^-$), and T cells ($CD4^+$ $CD45RA^+$)[27]. We ran independent analyses for each of the 3 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 which 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 moloc analyses to investigate gene expression within different tissue types using data from the GTEx consortium v6p[29]. Our 3 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. As in the immune-cell eQTL moloc analysis, evidence of colocalization was based on a PPA score of ≥ 80%.

All statistical and bioinformatics analyses were undertaken using R statistical software version 3.31 [30]. Plots illustrating multiple trait colocalization were generated using base R graphics, whereas our volcano plot was generated using ggplot[31].

Results

Mendelian Randomization 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 a single cis-acting SNP instrument for the inflammatory biomarker gene of interest. *IP-10* and *TRAIL* were removed from the analysis as the SNP identified as instruments[18] had minor allele frequencies too rare (< 0.05) to undertake formal two-sample MR. Results from the conservative MR analysis are displayed in Figure 1 and 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/dermatits ($P = 1.35 \times 10^{-8}$), rheumatoid arthritis (RA) ($P = 5.67 \times 10^{-8}$), Crohn's disease (CD) ($P = 2.81 \times 10^{-5}$) and Asthma ($P = 1.12 \times 10^{-4}$), as well as an association between IL-2Rα levels and multiple sclerosis (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, Table S1) as instruments. *IL-6R* (SNPs affecting levels of soluble IL-6R only), *MIF*, and *IL-2Rα* were excluded from the liberal analysis as there was only one remaining SNP available for MR after LD clumping to remove SNPs in linkage disequilibrium with one another. We adjusted our Bonferroni corrected P value threshold to account for the change in the number of tests and applied the new threshold of $P \leq 6.14 \times 10^{-4}$ to the liberal MR results. From this analysis we found evidence to suggest there may be causal relationships between circulating levels of IL-18 and inflammatory bowel disease (IBD) ($P = 1.17 \times 10^{-4}$) and eczema/dermatitis ($P = 2.81 \times 10^{-3}$). These associations were not detected in the conservative analysis using the *IL-18* SNP (rs71478720) alone ($P = 1.06 \times 10^{-2}$, E/D: $P = 2.81 \times 10^{-3}$), although using multiple instruments in the liberal MR provided much 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 (Figure 2) and on eczema/dermatitis (Figure S2) survived leave-one-out analyses (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 (Table S6) and the Steiger directionality test[28] (Table S7) showed that reverse causation was unlikely for any of the associations identified in either the conservative or liberal MR analysis.

Multiple-trait Colocalization Uncovers Immune Cell-Specific Drivers of IMD

We next investigated the cell-type specificity for the detected putative drivers of IMD using eQTL data from the BLUEPRINT consortium[27]. For circulating inflammatory biomarkers and IMDs where evidence of a causal association was detected using MR, we applied the moloc method using eQTL data for the gene encoding the target inflammatory biomarker, pQTL data for the circulating inflammatory biomarker itself and GWAS summary statistics for the associated complex trait (Figure S1B). We tested whether immune-cell eQTL ($_a$) colocalised with pQTL ($_b$) and IMD-associated loci ($_c$) The moloc analysis was performed 3 times for each association, using eQTL from either human $CD14^+$ $CD16^-$ monocytes ($_{a1}$), $CD16^+$ $CD66b^+$ neutrophils ($_{a2}$) or $CD4^+$ $CD45RA^+$ T cells ($_{a3}$).

We found evidence of multiple-trait colocalization between 10 combinations of immune-cell eQTL, IMD-associated loci, and inflammatory biomarker pQTL (Table S8). PPA scores measuring colocalization of genetic signals between all 3 traits (PPA_{abc}) indicated that monocytes (PPA_{a1bc} = 0.8812) and neutrophils (PPA_{a2bc} = 0.9657), but not T cells (PPA_{a3bc} = 0.0159), may drive IBD via IL-18 (Figure 3A). Additionally, we found evidence of a T cell-specific role in the disease pathways of eczema/dermatitis (PPA_{a3bc} = 0.9115) driven by soluble IL-6R (Figure 3B). IL-2Rα pQTL, MS-associated loci and T cell eQTL showed evidence of colocalization (PPA_{a3bc} = 0.8826), but there were no *IL-2Rα* eQTL data within monocytes or neutrophils in the BLUEPRINT study.

In the tissue-specific analysis, the association between IL18 and IBD was relatively ubiquitous, as evidence of colocalization was observed within 7 diverse tissue types (Table S9), including thyroid tissue (Figure 4a). Evidence of colocalization for the association between soluble IL6R and eczema/dermatitis was observed in 3 tissue types, most strongly in whole blood (PPA = 99.5%), which may help shed light on the pleiotropic effects observed at this locus (Figure 4b). Lastly, the association between IL2ra and multiple sclerosis colocalized in only two tissues; subcutaneous adipose and spleen (Figure 4c).

Discussion

The prevalence of autoimmune and atopic diseases has increased drastically over recent decades, particularly within European populations[2; 32]. Despite the rising demand, a poor understanding of the molecular pathways driving IMDs means that drugs for their treatment are lacking[33; 34]. Cytokines and their receptors mediate the balance between tolerance and inflammation, by signalling between immune cells, making them attractive therapeutic targets[35]. However, the ways in which the hundreds of functionally diverse cytokines interact and orchestrate inflammatory responses in the body are unclear[36]. Consequently, identifying individual cytokines for drug-targeting which drive IMDs has been difficult. The use of two-sample MR has aided the elucidation of many disease pathways, by measuring the likelihood of causal relationships between exposures and traits in human populations[37]. In this study, we developed a framework that integrates MR with multiple-trait colocalization 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-2Rα and MS, amongst others (Tables S3-4, S6-8). Additionally, we provided evidence to suggest that these associations are likely to be driven in an immune-cell and tissue-specific manner. 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 which are likely to drive disease.

A T cell-mediated role for IL-2/IL-2Ra in MS has already been well established through epidemiological and lab-based studies[10; 38]. The IL-2R-targetting drug daclizumab was given FDA approval for the treatment of MS, but was recently withdrawn due to serious side-effects[39-41]. Additionally, a causal role for IL-18R in atopic dermatitis has recently been described, by integrating MR and pQTL-data[42]. 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 (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 INF-γ production[43]. 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[44]. 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[45]. 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[47]. The T cell-specific eQTL data currently available for moloc analyses concerned naïve $CD4^+$ T cells[27]. 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. However, this data may be difficult to acquire on a large-scale. Although eQTL data for many tissue types are now readily available[29], moloc analysis using tissue-specific eQTL data concerning tissues which are known to become inflamed during IMDs would help to indicate whether cytokine or cytokine receptor expression in certain tissue types drives IMDs. For example, eQTL data from the different layers within the intestine would help to further unravel which roles inflammatory biomarkers are likely play in driving inflammation during IBD, in a tissue-context-dependent manner.

IL-6R is the receptor of the pro-inflammatory cytokine IL-6, which can exist in a membrane-bound state (classical) to the surface of leukocytes and hepatocytes, or in soluble form (trans)[48]. Both, classical and trans IL-6R signalling culminates in the expression of signal transducer and activator of transcription (STAT)3, which promotes inflammation via the expression of genes encoding antiapoptotic proteins and cytokines[49]. However, classical IL-6 signalling can act on few cell types compared to trans IL-6 signalling, which can act on any cell which has the cell-bound signal transducer, gp130[50]. One GWAS previously reported the an association between elevated soluble IL-6R levels resulting from a SNP in *IL-6R* and atopic dermatitis[51]. We provide evidence from the conservative MR and moloc analyses that supports a causal role between soluble IL-6R and eczema/dermatitis. Levels of membrane-bound IL-6R (leading to classical IL-6 signalling) and soluble IL-6R (trans IL-6 signalling) fluctuate as a result of receptor shedding by cells expressing membrane-bound IL-6R, as wells changes to the amount of receptor being synthesised by cells[50]. Moreover, the ratio of classical to trans IL-6 signalling also depends upon levels of IL-6 and cell-bound gp130. Therefore, analyses combining instruments which affect the classical or trans signalling pathways would provide a clearer insight into the role of IL-6 signalling in disease. Interestingly, activated $CD4^+$ T cells have been shown increase levels of soluble IL-6R, by shedding their membrane-bound IL-6R; this mechanism is thought to have a role in the development of autoimmune diseases, which are often mediated by autoreactive T cells [52; 53]. Through the moloc analysis, we showed that eczema/dermatitis was likely to be driven by IL-6R expression in T cells (Figure 3B, Table S8), a finding which 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[51].

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[18-20]; 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.

In conclusion, in this study 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. Indeed, small molecule inhibitors which target and repress IL-18-mediated signalling events are currently under development, although not for the treatment of IBD[54]. Moreover, a recently-published phase-II trial of an IL-18 binding protein (IL-18bp) drug to treat adult-onset Still's disease demonstrated a favourable efficacy safety profile. If further trails are deemed successful, IL-18bp drugs may also be used to target other IMDs such as IBD.

Appendices

No appendices.

Supplemental Data Declaration

Document S1. Supplementary figures S1 & S2 (PDF).

Data S1. Supplementary tables S1-S8 (.xls).

Acknowledgements

This study and L.M.M were financially supported by a 4-year studentship fund from the Wellcome Trust Dynamic Molecular Cell Biology Ph.D. Programme at the University of Bristol: (108907/Z/15/Z). T.G.R is a UKRI Innovation Research Fellow (MR/S003886/1). The research was carried out in the UK Medical Research Council Integrative Epidemiology Unit (MC_UU_00011/4 and MC_UU_00011/1).

This study makes use of data generated by the BLUEPRINT Consortium[27] and data generated by the GTEx Consortium[29].

Declaration of Interests

T.R.G. receives research funding from GlaxoSmithKline, Biogen and Sanofi.

Web Resources

MR-Base Platform: http://www.mrbase.org/

Blueprint: http://dcc.blueprint-epigenome.eu/#/home

GTEx: https://www.gtexportal.org/home/

Figures

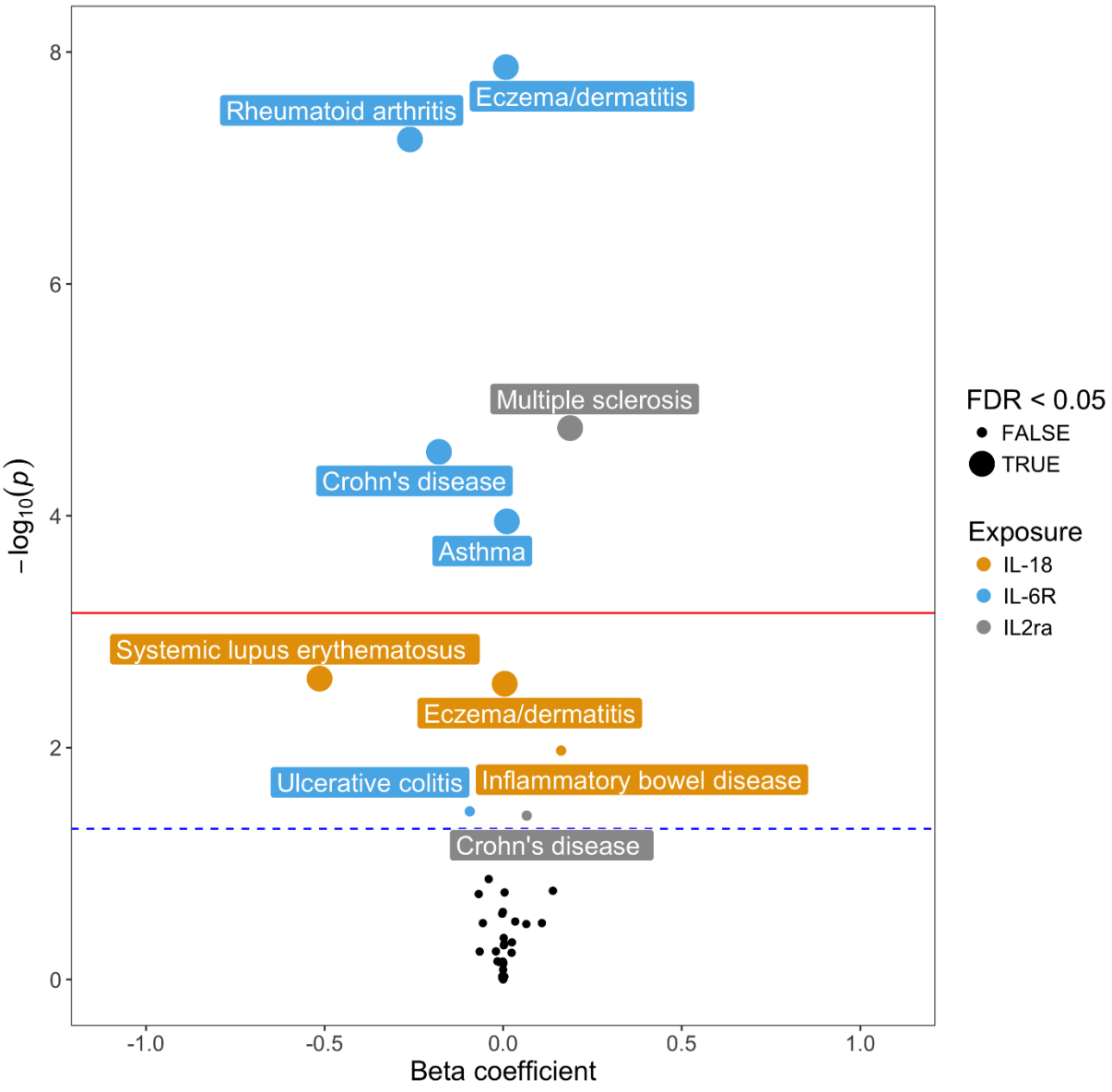


Figure 1. Conservative Mendelian randomization (MR) analysis detects associations between circulating inflammatory biomarkers and IMDs. Volcano plot of conservative MR analysis illustrating associations between inflammatory cytokines and complex traits. Red (upper) line represents Bonferroni corrected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

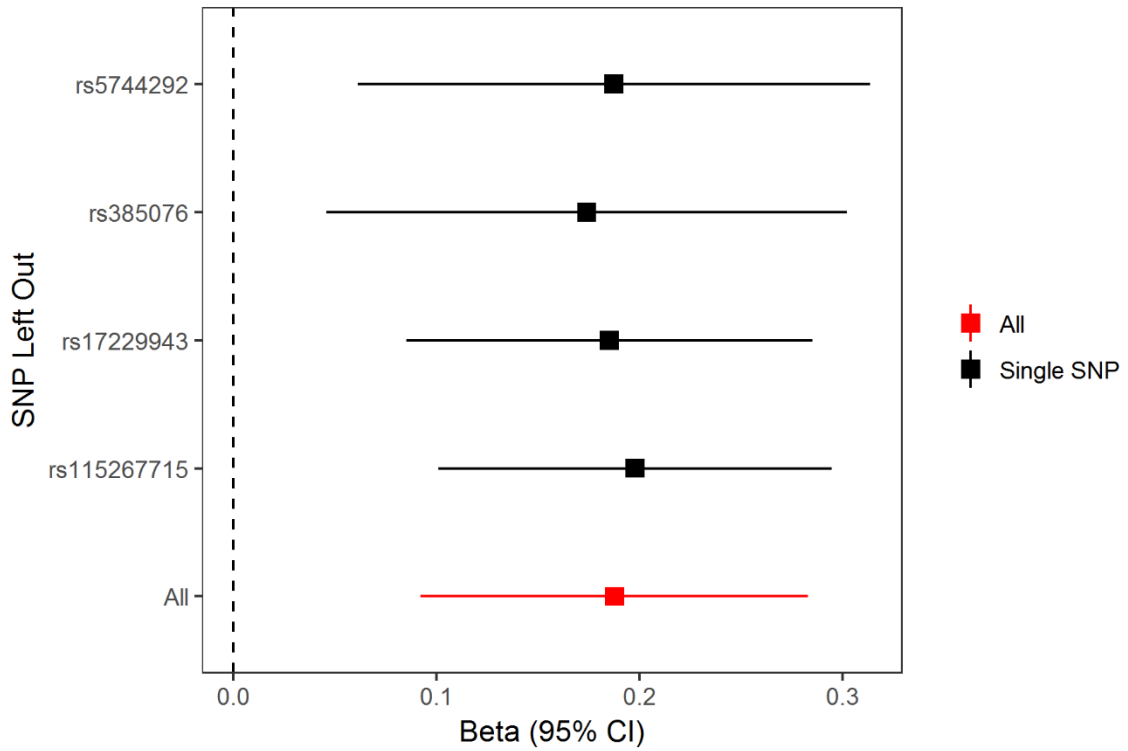


Figure 2. Liberal Mendelian randomization (MR) identifies a putative causal relationship between circulating levels of IL-18 and IBD which survives leave-one out sensitivity analysis. Leave-one out MR analysis for SNPs used as instruments for liberal MR analysis (black). Results show that this effect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

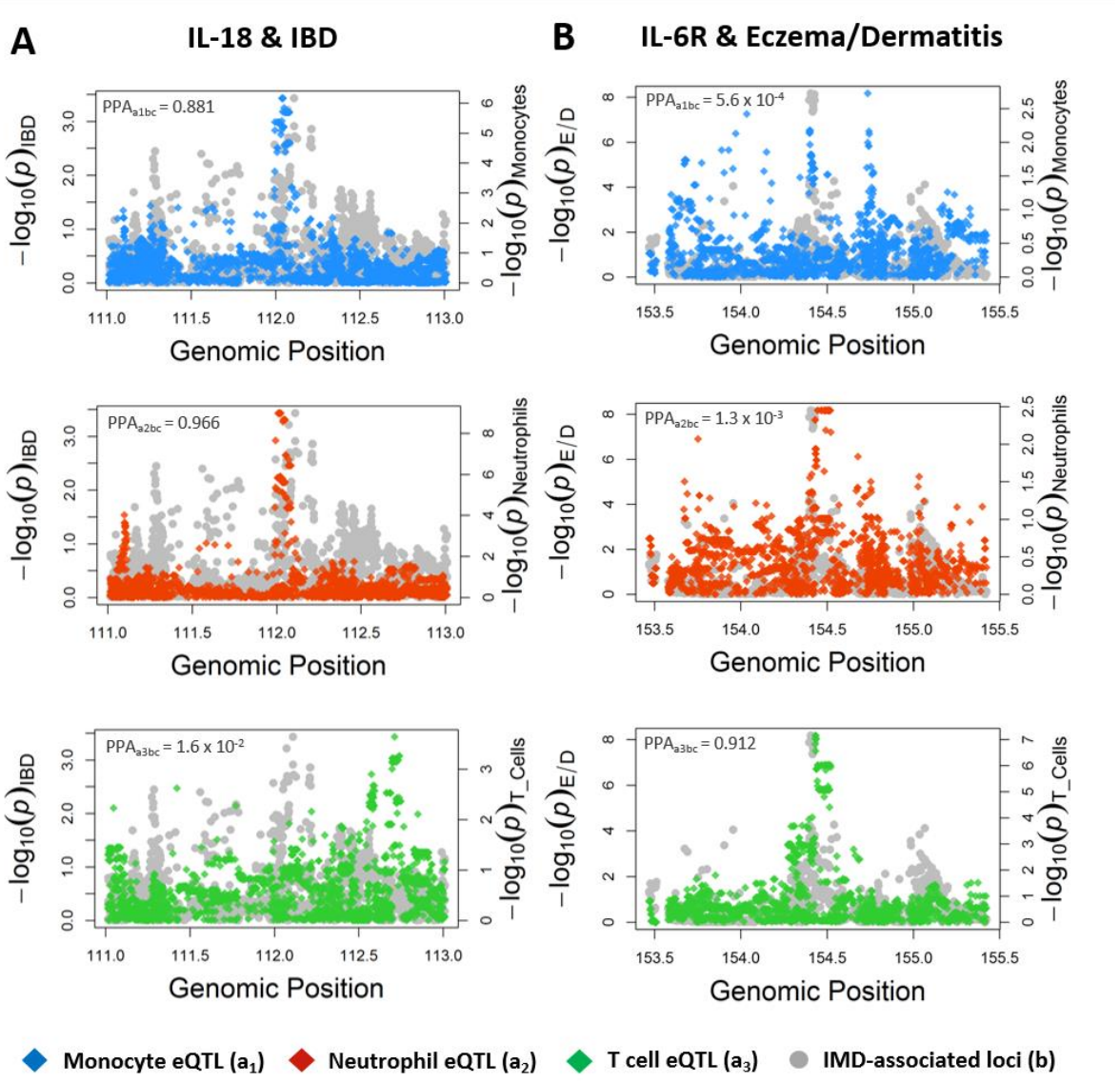


Figure 3. Multiple-trait colocalization reveals immune cell-specific divers of IMDs using immune cell expression quantitative trait loci (eQTL) data, inflammatory cytokine or cytokine receptor protein QTL (pQTL) data and immune-mediated disease GWAS data. These plots illustrate observed effects of genetic variants at the *IL-18* (A) and *IL-6R* (B) loci on inflammatory bowel disease (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, $_{a1}$), neutrophils (middle, $_{a2}$) and T cells (bottom, $_{a3}$). For simplicity, circulating cytokine effects are not displayed within the plots but were used to calculate PPA$_{abx}$ scores. PPA$_{abc}$ 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$_{abc}$

342 ≥ 0.8 indicates evidence of colocalization (i.e. a shared genetic variant between all 3 signals) and
343 suggests that the cytokine (or its receptor) is a putative driver of the IMD when it is expressed within
344 the cell-type of interest.

345

346

347

348

349

350

351

352

353

354

355

356

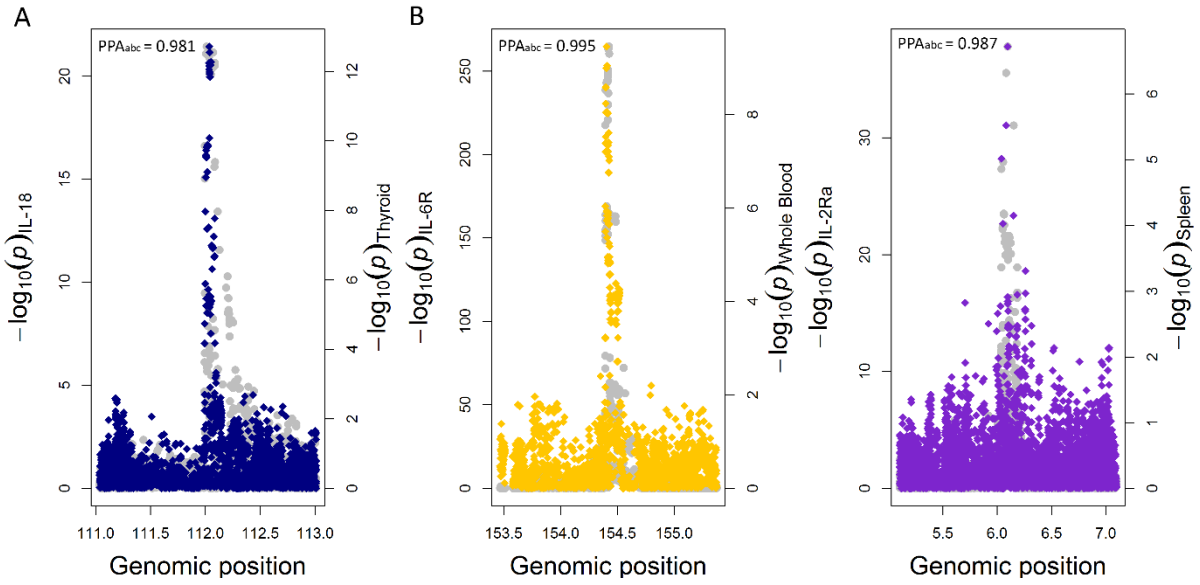


Figure 4. Multiple-trait colocalization indicates tissue specific expression of genes which are associated with circulating cytokine or cytokine receptor levels and immune-mediated disease (IMD) susceptibility. These plots illustrate observed effects of genetic variants at the *IL-18* (A), and *IL-6R* (B) and *IL-2Rα* (C) loci with their corresponding protein product. Effect estimates on the expression of IL-18, IL-6R, IL-2Rα are overlaid in each plot using expression quantitative trait loci data derived from thyroid (A), whole blood (B) and spleen (C). For simplicity, effects on complex traits are not displayed within the plots but were used to calculate PPA_{abx} scores. PPA_{abc} values reflect the likelihood that a causal variant influences the target cytokine ($_b$), associated complex trait ($_c$) and the expression of the corresponding gene ($_a$). $PPA_{abc} \geq 0.8$ indicates evidence of colocalization (i.e. a shared genetic variant between all 3 signals) and suggests that the cytokine (or its receptor) is a putative driver of the IMD when it is expressed within the tissue-type of interest.

References

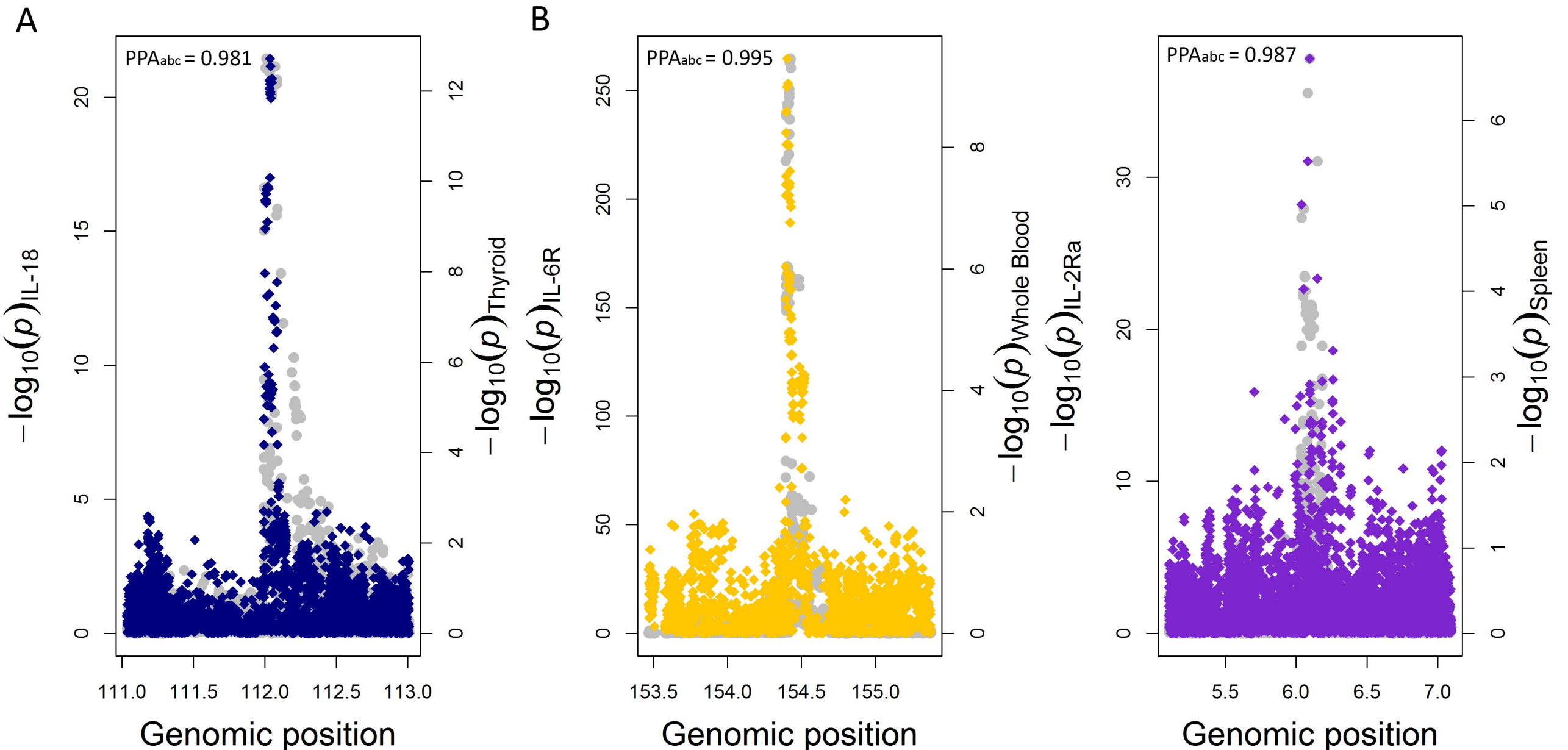
1. Rabin, R.L., and Levinson, A.I. (2008). The nexus between atopic disease and autoimmunity: a review of the epidemiological and mechanistic literature. Clinical and Experimental Immunology 153, 19-30.
2. Bach, J.F. (2018). The hygiene hypothesis in autoimmunity: the role of pathogens and commensals. Nature Reviews Immunology 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. Current Opinion in Immunology 43, 89-97.
4. Smith, G.D., and Ebrahim, S. (2003). 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? International Journal of Epidemiology 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., Sventoraityte, 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. Nature Genetics 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 Genetics 7.
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. Nature Genetics 47, 979-+.
9. Beecham, A.H., Patsopoulos, N.A., Xifara, D.K., Davis, M.F., Kemppinen, 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. Nature Genetics 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 Int Multiple Sclerosis, G. (2011). Genome-wide association study of severity in multiple sclerosis. Genes and Immunity 12, 615-625.
12. Stahl, E.A., Raychaudhuri, S., Remmers, E.F., Xie, G., Eyre, S., Thomson, B.P., Li, Y.H., Kurreeman, F.A.S., Zhernakova, A., Hinks, A., et al. (2010). Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nature Genetics 42, 508-U556.
13. Bentham, J., Morris, D.L., Graham, D.S.C., Pinder, C.L., Tombleson, 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. Nature Genetics 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. Nature Genetics 47, 1449-+.
15. Demenais, F., Margaritte-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. Nature Genetics 50, 42-+.

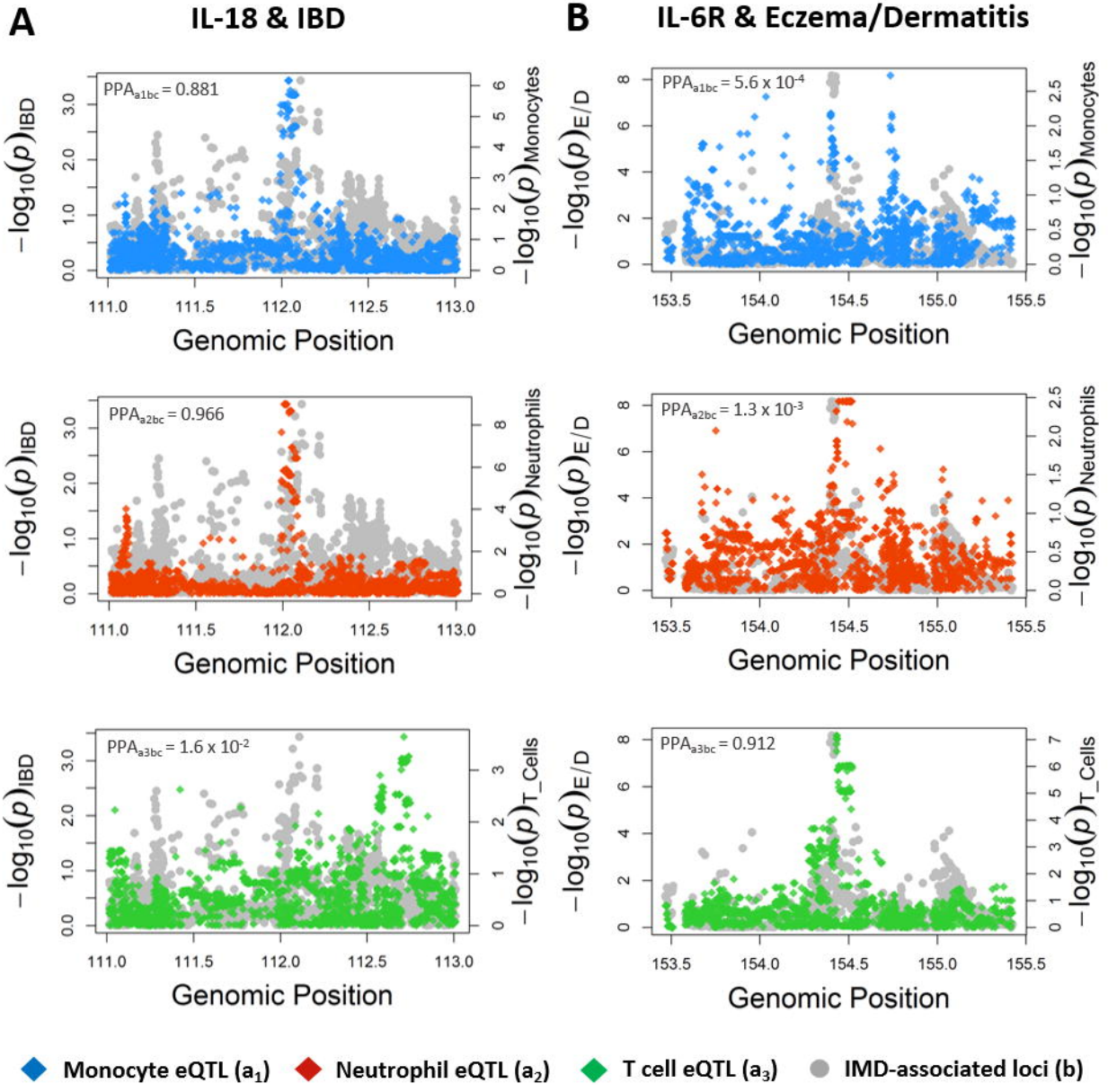
16. Nelson, M.R., Tipney, H., Painter, J.L., Shen, J.D., Nicoletti, P., Shen, Y.F., Floratos, A., Sham, P.C., Li, M.J., Wang, J.W., et al. (2015). The support of human genetic evidence for approved drug indications. Nature Genetics 47, 856-+.
17. Davey Smith, G., and Hemani, G. (2014). Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Human Molecular Genetics 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. American Journal of Human Genetics 100, 40-50.
19. Hartwig, F.P., Borges, M.C., Horta, B.L., Bowden, J., and Smith, G.D. (2017). Inflammatory Biomarkers and Risk of Schizophrenia A 2-Sample Mendelian Randomization Study. Jama Psychiatry 74, 1226-1233.
20. 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.
21. 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.
22. 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. British Medical Journal 342.
23. 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.
24. Giambartolomei, C., Zhenli Liu, J., Zhang, W., Hauberg, M., Shi, H., Boocock, J., Pickrell, J., Jaffe, A.E., The CommonMind, C., Pasaniuc, B., et al. (2018). A Bayesian framework for multiple trait colocalization from summary association statistics. Bioinformatics 34, 2538-2545.
25. 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 Genetics 13.
26. Hemani, G., Zhengn, 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.
27. 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-+.
28. Hemani, G., Tilling, K., and Smith, G.D. (2017). Orienting the causal relationship between imprecisely measured traits using GWAS summary data. Plos Genetics 13.
29. 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. Nature Genetics 45, 580-585.
30. R Core Team. (2017). A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. In. (https://www.R-project.org/.
31. Ginestet, C. (2011). ggplot2: Elegant Graphics for Data Analysis. Journal of the Royal Statistical Society Series a-Statistics in Society 174, 245-245.
32. Okada, H., Kuhn, C., Feillet, H., and Bach, J.F. (2010). The 'hygiene hypothesis' for autoimmune and allergic diseases: an update. Clinical and Experimental Immunology 160, 1-9.
33. Chatenoud, L. (2015). Biotherapies targeting T and B cells: from immune suppression to immune tolerance. Current Opinion in Pharmacology 23, 92-97.

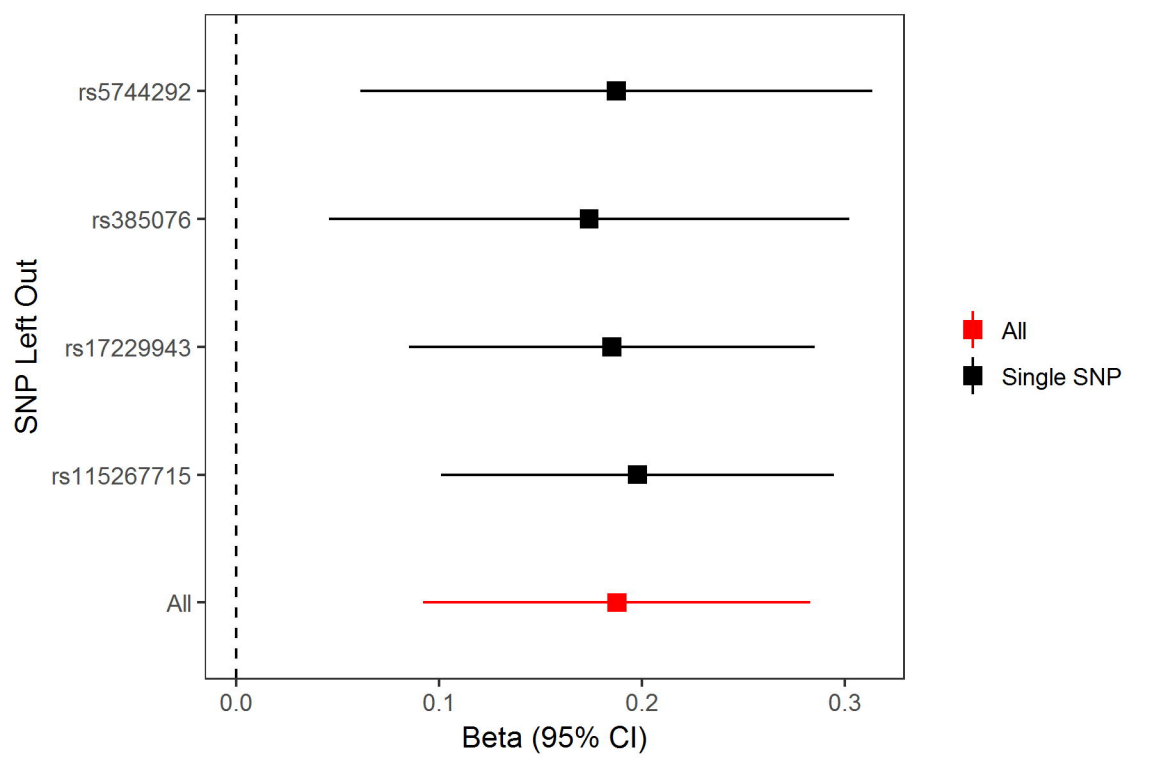
34. Li, P., Zheng, Y., and Chen, X. (2017). Drugs for Autoimmune Inflammatory Diseases: From Small Molecule Compounds to Anti-TNF Biologics. Frontiers in Pharmacology 8.
35. Theofilopoulos, A.N., Kono, D.H., and Baccala, R. (2017). The multiple pathways to autoimmunity. Nature Immunology 18, 716-724.
36. 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 in Immunology 34, 120-128.
37. Zheng, J., Baird, D.,, Borges, C.M., Bowden, J., Hemani, G., Haycock, P., Evans, D.M., and Davey Smith, G. (2017). Recent Developments in Mendelian Randomization Studies. In. (Current Epidemiology Reports), pp 330-345.
38. 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. Nature Communications 5.
39. Chataway, J., Martin, K., Barrell, K., Sharrack, B., Stolt, P., Wraith, D.C., and Grp, A.-M.S. (2018). Effects of ATX-MS-1467 immunotherapy over 16 weeks in relapsing multiple sclerosis. Neurology 90, E955-504.
40. MS Society. (2018). Daclizumab. In. (https://www.mssociety.org.uk/about-ms/treatments-and-therapies/disease-modifying-therapies/daclizumab.
41. Baecher-Allan, C., Kaskow, B.J., and Weiner, H.L. (2018). Multiple Sclerosis: Mechanisms and Immunotherapy. Neuron 97, 742-768.
42. 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-+.
43. Kaplanski, G. (2018). Interleukin-18: Biological properties and role in disease pathogenesis. Immunological Reviews 281, 138-153.
44. 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. Journal of Immunology 162, 6829-6835.
45. 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.
46. Magyari, L., Kovesdi, E., Sarlos, P., Javorhazy, A., Sumegi, K., and Melegh, B. (2014). Interleukin and interleukin receptor gene polymorphisms in inflammatory bowel diseases susceptibility. World Journal of Gastroenterology 20, 3208-3222.
47. Kanai, T., Kamada, N., and Hisamatsu, T. (2013). Clinical Strategies for the Blockade of IL-18 in Inflammatory Bowel Diseases. Current Drug Targets 14, 1392-1399.
48. Rincon, M. (2012). Interleukin-6: from an inflammatory marker to a target for inflammatory diseases. Trends in Immunology 33, 571-577.
49. Hodge, D.R., Hurt, E.M., and Farrar, W.L. (2005). The role of IL-6 and STAT3 in inflammation and cancer. European Journal of Cancer 41, 2502-2512.
50. Rose-John, S. (2012). IL-6 Trans-Signaling via the Soluble IL-6 Receptor: Importance for the Pro-Inflammatory Activities of IL-6. International Journal of Biological Sciences 8, 1237-1247.
51. 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. Journal of Allergy and Clinical Immunology 132, 371-377.
52. 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. Journal of Immunology 180, 7102-7106.
53. 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

524 during Inflammation Underlines a Role for IL-6 Trans Signaling in the Local Maintenance of
525 Th17 Cells. Journal of Immunology 184, 2130-2139.
526 54. Krumm, B., Meng, X.Z., Xiang, Y., and Deng, J.P. (2017). Identification of small molecule inhibitors
527 of Interleukin-18. Scientific Reports 7.

528


A
PPAabc = 0.981
−log10(p)IL-18
−log10(p)Thyroid
Genomic position
B
PPAabc = 0.995
−log10(p)IL-6R
−log10(p)Whole Blood
Genomic position
PPAabc = 0.987
−log10(p)IL-2Ra
−log10(p)Spleen
Genomic position


A
IL-18 & IBD
B
IL-6R & Eczema/Dermatitis
PPA$_{a1bc}$ = 0.881
PPA$_{a1bc}$ = 5.6 x 10^{-4}
PPA$_{a2bc}$ = 0.966
PPA$_{a2bc}$ = 1.3 x 10^{-3}
PPA$_{a3bc}$ = 1.6 x 10^{-2}
PPA$_{a3bc}$ = 0.912
$-\log_{10}(p)_{IBD}$
$-\log_{10}(p)_{E/D}$
$-\log_{10}(p)_{Monocytes}$
$-\log_{10}(p)_{Neutrophils}$
$-\log_{10}(p)_{T\_Cells}$
Genomic Position
111.0
111.5
112.0
112.5
113.0
153.5
154.0
154.5
155.0
155.5
Monocyte eQTL (a_1)
Neutrophil eQTL (a_2)
T cell eQTL (a_3)
IMD-associated loci (b)


rs5744292
rs385076
rs17229943
rs115267715
All
SNP Left Out
0.0
0.1
0.2
0.3
Beta (95% CI)
All
Single SNP