

## **The Genomic Architecture of Circulating Cytokine Levels Points to Drug Targets for Immune-Related Diseases**

Marek J. Konieczny<sup>1</sup>, Murad Omarov<sup>1</sup>, Rainer Malik<sup>1</sup>, Tom G. Richardson<sup>2,3</sup>, Sebastian-Edgar Baumeister<sup>4</sup>, Jürgen Bernhagen<sup>1,5,6b</sup>, Martin Dichgans<sup>1,5,6,6b</sup>, Marios K. Georgakis<sup>1,5,7\*</sup>

<sup>1</sup> Institute for Stroke and Dementia Research (ISD), LMU University Hospital, LMU Munich, Munich, Germany

<sup>2</sup> Medical Research Council (MRC) Integrative Epidemiology Unit (IEU), University of Bristol, Bristol, United Kingdom

<sup>3</sup> Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom

<sup>4</sup> Institute of Health Services Research in Dentistry, University of Münster, Münster, Germany

<sup>5</sup> Munich Cluster for Systems Neurology (SyNergy), Munich, Germany

<sup>6</sup> German Center for Neurodegenerative Diseases (DZNE), Munich, Germany

<sup>6b</sup> German Centre for Cardiovascular Research (DZHK, Munich), Munich, Germany

<sup>7</sup> Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA

\* Corresponding author

### **Competing interests**

The authors declare the following competing interests: MJK is an employee of Alexion Pharmaceuticals – AstraZeneca Rare Disease unrelated to this work. TGR is an employee of GlaxoSmithKline unrelated to this work.

### **Materials & Correspondence**

Marios K. Georgakis, MD, PhD  
Institute for Stroke and Dementia Research  
Ludwig-Maximilians-University (LMU) Hospital  
Feodor-Lynen-Str. 17, 81377 Munich, Germany  
Tel: +49-89-4400-46154, Fax: +49-89-4400-46040  
E-mail: [marios.georgakis@med.uni-muenchen.de](mailto:marios.georgakis@med.uni-muenchen.de)

### **Acknowledgements**

This work is supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) by an Emmy Noether grant (GZ: GE 3461/2-1, ID 512461526 to MG), by a clinician-scientist grant to MG and project grants to JB and MD from the Munich Cluster for Systems Neurology (EXC 2145 SyNergy, ID 390857198), and by CRC1123 project A3 to JB. Additional support comes from a research grant from the Fritz-Thyssen Foundation (Ref. 10.22.2.024MN to MG) and a research fellowship by the Hertie Foundation (Hertie Network of Excellence in Clinical Neuroscience, ID P1230035 to MG).

## Abstract

Circulating cytokines orchestrate immune reactions and are promising drug targets for immune-mediated and inflammatory diseases. Exploring the genetic architecture of circulating cytokine levels could yield key insights into causal mediators of human disease. Here, we performed genome-wide association studies (GWAS) for 40 circulating cytokines in meta-analyses of 74,783 individuals. We detected 359 significant associations between cytokine levels and variants in 169 independent loci, including 150 *trans*- and 19 *cis*-acting loci. Integration with transcriptomic data point to key regulatory mechanisms, such as the buffering function of *ACKR1* acting as scavenger for multiple chemokines and the role of *TRAFD1* in modulating the cytokine storm triggered by TNF signaling. Applying Mendelian randomization (MR), we detected a network of complex cytokine interconnections with TNF- $\beta$ , VEGF, and *IL-1ra* exhibiting pleiotropic downstream effects on multiple cytokines. Drug target *cis*-MR paired with colocalization revealed *G-CSF/CSF-3* and *CXCL9/MIG* as potential causal mediators of asthma and Crohn's disease, respectively, but also a potentially protective role of TNF- $\beta$  in multiple sclerosis. Our results provide an overview of the genetic architecture of circulating cytokines and could guide the development of targeted immunotherapies.

## Introduction

Chronic inflammation contributes to multiple human diseases, including allergic and autoimmune diseases, cardiometabolic diseases, and cancer. Inflammatory proteins like cytokines, chemokines and growth factors (hereafter collectively referred to as “cytokines”) orchestrate the immune response underlying inflammation.<sup>1,2</sup> Circulating cytokines are readily accessible and therefore attractive targets for therapeutic modulation, as they represent soluble ligands that execute downstream mechanisms through binding to membrane receptors or other circulating agents.<sup>3</sup> While several immunotherapies targeting circulating cytokines have been successfully introduced into the clinic, the lack of efficacy in other indications and the usually associated susceptibility to infection underscore the need for targeted approaches.<sup>4,5</sup> Prioritizing specific downstream mediators is critical to minimize safety signals and ensure adherence to a life-long pharmacotherapy.<sup>5,6</sup>

Recent advances in human genetics have enabled an *in silico* prioritization of drug targets<sup>7,8</sup>, with approval rates more than two times higher than targets without genetic support.<sup>9</sup> Mendelian randomization (MR) uses data from genome-wide association studies (GWAS) and offers a statistical framework for exploring associations between variants in genes encoding drug targets and disease traits.<sup>10</sup> Previous MR analyses have illustrated the potential of integrating GWAS data for circulating proteins, including cytokines, with disease outcomes to discover novel drug targets.<sup>11–16</sup> However, existing efforts have been largely restricted by the small sample sizes of GWAS studies for circulating cytokines. For example, the largest-to-date targeted GWAS which focused specifically on, circulating cytokines included up to 8,293 individuals and allowed the detection of 27 significant genomic loci for 41 cytokines.<sup>17</sup>

Novel proteomic platforms, such as the aptamer-based SOMAScan<sup>®</sup> and the proximity extension assay Olink<sup>®</sup>, gain popularity in quantifying at scale large numbers of proteins including cytokines. Here, we performed cross-assay comparisons in the genetic architecture of 40 cytokines quantified with 3 approaches (multiplex bead-based immunoassay, aptamer-based assay, proximity extension assay) and pooled data in GWAS meta-analyses including up to 74,783 individuals. This effort allowed the detection of 359 significant associations between 169 independent genomic loci and one or more of the 40 cytokines offering novel insights into mechanisms regulating circulating cytokine levels. Applying MR, we establish a causal cytokine network including upstream mega-regulator cytokines that exert influence on a range of other cytokines. Finally, integrating these data with GWAS data for relevant disease endpoints, we provide genetic support for putative anti-inflammatory drug targets.

## Results

### Study cohorts and cross-assay reproducibility rate of significant genomic loci

We leveraged summary-level GWAS data for 40 circulating cytokines from 3 published datasets summing up to 74,783 individuals: the Cardiovascular Risk in Young Finns Study (YFS) and FINRISK studies that measured cytokines in serum using Luminex bead-based multiplex immunoassays (N=8,293); the Systematic and Combined AnaLysis of Olink Proteins (SCALLOP) study that measured cytokines in plasma using the proximity extension assay-based Olink<sup>®</sup> platform (N=30,931); and the dataset provided by deCODE that measured cytokines in plasma using the aptamer-based SOMAScan<sup>®</sup> assay (N=35,559, **Figure 1**).

**Figure 1. Flowchart of the study design.** Illustration of the analytical pipeline steps applied in this study to decipher the genetic architecture of circulating cytokines and their relation to allergic and autoimmune, cardiometabolic and cancer outcomes. SCALLOP, Systematic and Combined AnaLysis of Olink Proteins; SNP,

Single-nucleotide polymorphism; TWAS-MR, Transcriptome-wide Mendelian randomization analysis; YFS & FINRISK, Cardiovascular Risk in Young Finns Study.

Given the known differences across the assaying methods, we first tested the replication rate of significant variants detected in each dataset in the other two datasets.<sup>18</sup> Although the GWAS in SCALLOP led to a considerably lower number of genome-wide significant loci for available cytokines, these variants showed the highest rate of reproducibility ( $p < 5 \times 10^{-5}$  and directionally consistent) in the other two datasets (median 67% in YFS & FINRISK and 63% in deCODE, **Figure 2, Supplementary Table S1**). A lower reproducibility rate was found for significant variants detected in YFS & FINRISK (median 4% in deCODE and 11% in SCALLOP) and deCODE (median 21% in YFS & FINRISK and 19% in SCALLOP). The cytokines that showed the highest relative proportion of reproducible SNPs across all 3 datasets independently of the measuring assay were CC chemokine ligand 2 (CCL2/MCP-1) and vascular endothelial growth factor (VEGF).

**Figure 2. Comparisons of significant genomic loci for 40 circulating cytokines across 3 proteomics assays.** (a) Number of reproducible and non-reproducible SNPs per cytokine (depicted as saturated and light-colored bars, respectively) for deCODE, SCALLOP and YFS & FINRISK cohorts. (b) Median proportion of replicated SNPs across the three platforms (error bars represent the 25<sup>th</sup> and 75<sup>th</sup> percentiles). Colored bars represent deCODE consortium in red, SCALLOP consortium in blue and YFS & FINRISK cohorts in yellow. SNP, Single-nucleotide polymorphism; SCALLOP, Systematic and Combined AnaLysis of Olink Proteins; YFS & FINRISK, Cardiovascular Risk in Young Finns Study.

### GWAS meta-analysis reveals novel *trans*- and *cis*-acting variants

Next, we performed GWAS meta-analyses across the 3 datasets. We identified a total of 359 significant associations between variants at 169 independent genomic loci and the circulating levels of one or more of the 40 cytokines ( $p < 5 \times 10^{-8}$  in fixed-effects meta-analysis, **Figure 3, Supplementary Table S2**). Variants that showed significant heterogeneity between the three cohorts (HetPval < 0.1) are reported in **Supplementary Table S3** (48% of the significant loci; range 0% to 100% across the 40 cytokines). The lambda values ranged between 0.96 for interleukin (IL)-16 and 1.04 for basic fibroblast growth factor (FGF-b), indicating absence of overall inflation in the test statistics (**Supplementary Table S3**). According to GWAS catalog (<https://www.ebi.ac.uk/gwas/>)<sup>19</sup>, 156 of the loci have not been associated with circulating levels of the 40 cytokines in previous GWASs (**Supplementary Table S4**). The proportion of explained variance by the significant variants ranged from 0.0008 for IL-17 to 0.033 for stem cell growth factor beta (SCGF-b) (**Supplementary Table S2**).

**Figure 3. Genetic architecture of the circulating levels of the 40 cytokines.** Circular Manhattan plot of genomic loci significantly associated with circulating levels of 40 cytokines in a meta-analysis of the three datasets. The 359 genome-wide significantly associated variants at  $p < 5 \times 10^{-8}$  are depicted as black dots for GWAS meta-analyses in YFS & FINRISK, SCALLOP, and deCODE cohorts. The horizontal and vertical location of dots in each single rectangle depict genomic positioning (increasing from left to right) and p-value (decreasing from bottom to top), respectively. SCALLOP, Systematic and Combined AnaLysis of Olink Proteins; YFS & FINRISK, Cardiovascular Risk in Young Finns Study.

As expected due to larger effects sizes for rare genetic variants, we found a strong inverse correlation between minimum allele frequency and effect size (Spearman's  $\rho = -0.827$ ,  $p = 5 \times 10^{-30}$ , **Figure 4a**). The majority of the significant loci (150 out of 169) represented *trans*- (distant-) acting variants. When excluding the human leukocyte antigen (HLA) region on chromosome 6, we found 33 pleiotropic variants showing associations with >1 cytokine

among the significant *trans*- variants (**Figure 4b**). A locus hotspot associated with multiple cytokines was found at the region of the gene encoding *complement factor H* (*CFH*). This soluble mediator plays an essential role by interacting with the C3 convertase for regulation of inflammatory responses exerted by the complement system, which could possibly explain the associations with multiple cytokines.<sup>20</sup> While at least 1 significant *trans*-variant was present for all studied cytokines (median number of variants per cytokine = 5, range 1 to 22), we found significant *cis*- (local-) acting variants in the vicinity of their encoding gene for 19 cytokines (**Supplementary Table S2**). The lead *cis*-acting variants showed stronger associations with cytokine levels (mean absolute beta: 0.18, range: 0.05-0.94) than *trans*-acting variants (mean absolute beta: 0.08, range: 0.03-0.55, *p*-for-comparison=0.03, **Figure 4c**).

**Figure 4. *Trans*- and *cis*-acting genetic variants underlying circulating cytokines.** (a) Inverse correlation between minimum allele frequency and effect size illustrated for *trans*- and *cis*-acting loci. (b) Number of pleiotropic loci associated with circulating cytokines (excluding the HLA region on chromosome 6). (c) *Cis*-acting variants showed stronger associations with cytokine levels when compared with *trans*-acting variants. Bars and lines represent median and 95% confidence intervals, respectively. HLA, human leukocyte antigen; SD, standard deviation.

To map causal variants responsible for associations between cytokine serum levels and genes within each of the 169 independent genomic loci we used SuSiE fine mapping. Employing a Bayesian framework, fine mapping identifies credible sets of variants with posterior probability (PP) 95%. The number of variants within credible sets ranged from 2 to 50. The highest numbers of variants within a credible set were found at 15q21.3 for stem cell factor (SCF) (n=50), at 19q13.33 for SCGF-b (n=49), and at 6p21.1 for VEGF (n=44). SuSiE mapped the association test lead variant to the credible sets for 49 genomic loci, identifying the GWAS lead as most likely causal mutations (**supplementary Table S5**).

### Functional follow-up analyses highlight immune response regulatory mechanisms

To understand the biological significance and downstream functional impact of the identified variants, we performed follow-up analyses. A MAGMA gene-based analysis showed 829 significant associations with the levels of circulating cytokines at a Bonferroni-defined significance level (**Supplementary Table S6**). In total, 626 uniquely mapped genes were associated with at least 1 cytokine. The number of genes mapped to cytokines ranged from 1 for beta nerve growth factor (bNGF), cutaneous T-cell attracting (CCL27/CTACK), IL-10 and tumor necrosis factor-alpha (TNF- $\alpha$ ) to up to 95 genes mapped for SCGF-b, 92 genes for macrophage inflammatory protein-1 $\beta$  (CCL4/MIP—1b) and 51 genes for CCL11/eotaxin-1. In line with our GWAS results, the gene that was mapped for most cytokines (n=16) was *CFH*. The genes with the lowest p-value were *H4C14* for monocyte specific chemokine 3 (CCL7/MCP-3) ( $p=1\times 10^{-50}$ ), *DBA4* for IL-16 ( $p=1.3\times 10^{-45}$ ), and *ABC1* for SCF ( $p=1.9\times 10^{-29}$ ). A gene-property analysis revealed that the cytokine-related genes were primarily enriched for expression in the liver ( $p=4.9\times 10^{-10}$ ), in line with its well-established role as a main source of production of many cytokines. Other enriched tissues included the spleen ( $p=4.9\times 10^{-4}$ ) and lung ( $p=5.9\times 10^{-4}$ , **Supplementary Table S7**). A MAGMA gene-set analysis prioritized 41 pathways for 12 cytokines that reached a Bonferroni-adjusted significance level ( $p<1.2\times 10^{-7}$ , **Supplementary Table S8**). The identified pathways were primarily related to immune response with a small cluster involved in metabolic and developmental processes.

Positional mapping ascribed 75% of significant variants to intronic (54%) and intergenic (21%) regions, suggesting that the identified variants primarily determine gene transcription or gene expression profiles (**Supplementary Table S2**).<sup>21</sup> Thus, we integrated our GWAS data with transcriptomic data and performed a transcriptome-wide association study (TWAS) using Mendelian randomization (MR) for a deeper elaboration on the transcriptional effects underlying our GWAS results. Summary statistics for expression quantitative trait loci (eQTLs) in whole blood were obtained from the eQTLGen consortium including transcriptomic profiles for 31,684 individuals of primarily European ancestry.<sup>22</sup> Using *cis*-eQTLs as genetic instruments, we identified 245 significant associations between genetically proxied gene expression in whole blood and cytokines levels (**Figure 5, Supplementary Table S9**). The number of significant genes per cytokine ranged from 1 to 18. While most significant genes (78%) influenced the levels of a single cytokine, the genetically proxied expression of 54 genes showed an effect on circulating levels of up to 9 cytokines (n[*SKIV2L*]=9, n[*HLA-DRB5*]=9, n[*NELFE*]=7, n[*ACKR1*]=5, n[*FCERIA*]=4, n[*TRAFD1*]=4, n[*LCMT2*]=4). Interestingly, we found significant *cis*-effects of the encoding gene expressions on the circulating levels of only 3 of the 40 respective cytokines. This is in line with previous eQTL-pQTL comparisons and aligns with the fact that the circulating proteome is not the direct product of the whole-blood transcriptome.<sup>23,24</sup>

**Figure 5. Genetically predicted gene expression in peripheral blood partly explains the genetic architecture of 40 circulating cytokine levels.** The dots represent genes, the blood expression of which was significantly associated with circulating cytokine levels in a Mendelian randomization-based transcriptome-wide association study.

Excluding genes within the very dense HLA region (i.e. *SKIV2L*, *HLA-DRB5* and *NELFE*) we explored deeper the biological relevance of the pleiotropic *ACKR1*, *TRAFD1* and *LCMT2* genes. The genetically proxied mRNA levels of *ACKR1* were associated with circulating CCL2/MCP-1, CCL7/MCP-3, CCL11/eotaxin-1, growth regulated oncogene- $\alpha$  (CXCL1/GRO $\alpha$ ), CXCL8/IL-8. *ACKR1* codes for a cell-surface receptor that binds, internalizes and transports multiple CC and CXC chemokines and promotes leukocyte transcytosis into the circulation.<sup>25,26</sup> By acting as scavenger receptor, *ACKR1* modulates the bioavailability of cytokines and thereby affects inflammatory responses.<sup>27,28</sup> The identified associations were driven by rs12075, a well-characterized missense variant in *ACKR1*, resulting in less efficient chemokine binding to *ACKR1* due to loss of a necessary amino-acid sulfation (**Figure 6a**).<sup>29</sup> The impaired receptor binding leads to elevated circulatory levels of chemokines and might in turn result in a compensatory increase in *ACKR1* expression which could explain the positive association between genetically proxied *ACKR1* and its ligands.<sup>30</sup> We replicated previously reported associations between *ACKR1* and levels of CCL2/MCP-1, CCL7/MCP-3, CCL11/eotaxin-1, and CXCL1/GRO $\alpha$  and additionally showed an association with CXCL8/IL-8 levels.<sup>17,23</sup> The genetically proxied expression of *tumor necrosis factor receptor-associated factor 1* (*TRAFD1*) was also associated with multiple circulatory cytokine levels, including CCL7/MCP-3, monokine induced by interferon-gamma (CXCL9/MIG), interferon gamma-induced protein 10 (CXCL10/IP-10), and tumor necrosis factor-beta (TNF-b) (**Figure 6b**). *TRAFD1* functions as an adaptor protein that binds to the intracellular domain of TNF receptors expressed on both innate and adaptive immune cells. It regulates downstream signaling also involving the NF- $\kappa$ B pathway and thereby modulates the production of several pro-inflammatory cytokines and inflammatory responses.<sup>31-33</sup> *TRAFD1* is a master regulator of genes involved in interferon- $\gamma$  (IFN $\gamma$ ) signaling and T-cell receptor activation.<sup>34</sup> Genetically proxied expression of the gene encoding for *LCMT2* showed associations with beta nerve growth factor (bNGF), CXCL8/IL-8, CXCL10/IP-10 and platelet-derived growth factor-bb (PDGFbb). *LCMT2* is involved in amino-acid metabolism

presumably regulating hypothalamic gene expression but there is only limited knowledge on its biological function.<sup>35,36</sup>

**Figure 6. Gene expression of *ACKR1* and *TRAFDI* exert pleiotropic effects on multiple cytokine levels.** (a) Schematic illustrating the impact of *cis*-eQTLs for *ACKR1* on receptor function. Left hand side shows *ACKR1* gene encoding the atypical chemokine receptor 1 functioning as sink for multiple chemokines which are buffered intracellularly in venular endothelial cells. Depicted on the right, is the missense variant rs12075 coding for a dysfunctional receptor with less efficient chemokine binding efficacy. This leads to higher levels of circulating CCL2/MCP-1, CCL7/MCP-3, CCL11/eotaxin-1, CXCL1/GROa, and CXCL8/IL-8 and possibly to a compensatory increase in *ACKR1* expression and receptor density. (b) Schematic illustrating how genetically proxied *TRAFDI* expression regulates multiple cytokine levels (CCL7/MCP-3, CXCL9/MIG, CXCL10/IP-10, TNF-b), supporting its regulatory role in TNF-mediated NF- $\kappa$ B signaling. ACKR1, atypical chemokine receptor 1; TNF-R, TNF-receptor.

### Genetic associations point to network interactions between circulating cytokines

As a next step, we explored cross-trait genetic correlations between the circulating levels of the 40 studied cytokines (**Figure 7a, Supplementary Table S10**). One third of the between-cytokine correlations were significant at  $p < 0.05$ ; the vast majority of the significant associations (96%) were positive. Furthermore, to understand causal interconnections between circulating cytokines levels, we performed MR analysis using *cis*-variants from our GWAS meta-analysis. We found significant (FDR-corrected  $p < 0.05$ ) associations between 65 cytokine pairs (53 positive associations and 12 negative associations, **Figure 7b, Supplementary Table S11**). Genetically proxied levels of CCL7/MCP-3, stromal cell-derived factor-1alpha (CXCL12/SDF-1a), granulocyte colony-stimulating factor (G-CSF/CSF-3), IL-9, TNF-b, and VEGF were positively associated with the levels of  $>2$  other cytokines, whereas genetically proxied CXCL1/GROa and IL-1 receptor antagonist (IL-1ra) were negatively associated with lower level of  $>2$  other cytokines. Most significant associations were detected for TNF-b ( $n=13$ ), VEGF ( $n=9$ ), IL-1ra ( $n=7$ ), IL-9 ( $n=7$ ), and G-CSF/CSF-3 ( $n=7$ ). The negative associations between IL-1ra and several proinflammatory cytokines (CCL7/MCP-3, IL-9, TNF-a, TNF-b), chemokines (macrophage inflammatory protein-1 $\alpha$ , CCL3/MIP-1a), and growth factors (hepatocyte growth factor, HGF; VEGF) align well with the immunoregulatory role of the IL-1 pathway and the inhibitory effect of IL-1ra on downstream IL-1 signaling.<sup>37,38</sup> TNF-b emerged as a significant player in our network analysis, demonstrating characteristics of a master regulator by showing significant associations with higher circulating levels of 13 mostly pro-inflammatory cytokines. Furthermore, TNF-b exhibited significant positive LDSC genetic correlations with 7 of the 13 cytokines, suggesting a shared genetic architecture within the TNF- $\beta$  network (**Figure 7a**). While certain interactions with TNF-b, such as those involving IL-1ra, TNF-a, TNF-related apoptosis inducing ligand (TRAIL), and VEGF, are well-documented, the majority of interactions have not been reported previously and merit additional investigation.<sup>39-41</sup>

**Figure 7. Cross-cytokine genetic associations.** (a) Genetic correlations with LD-score regression across cytokine serum levels are depicted as correlation heatmap. Stars highlight significance level \*, 0.05; \*\*, 0.0001; \*\*\*, 0.00001. LD-score correlation coefficients are illustrated according to the legend below spanning from -1 in blue to +1 in red, missing correlation coefficients are depicted in grey. (b) *Cis*-Mendelian randomization between genetically proxied circulating cytokine levels. Arrow heads show the direction of causal influence, color gradient indicates the effect estimate and line width the logarithm-adjusted Benjamin-Hochberg corrected significance level. LD, linkage disequilibrium.

## ***Cis*-Mendelian randomization and colocalization highlight potential drug targets for immune-related diseases**

For insights into the clinical consequences of genetically proxied levels of the circulating cytokines, we analyzed associations with allergic and autoimmune, cardiometabolic, and cancer outcomes in two-sample MR followed by colocalization analyses (**Figure 8** and **Supplementary Table S12**). We used *cis*-acting genetic variants as instruments due to their lower likelihood of influencing cytokine levels through pleiotropic mechanisms. We further complemented these analyses with Bayesian colocalization to prioritize associations less likely to be influenced by pleiotropy due to linkage disequilibrium of studied variants with neighboring genes.<sup>42</sup> Following correction for multiple comparisons, we found 24 significant MR associations between genetically proxied cytokine levels and disease outcomes (14 positive and 10 negative associations). Our MR findings partially confirmed established pathogenetic associations with diseases and therapeutic drug targets that are already in clinical application. For example, there is solid evidence linking IL-2 receptor subunit alpha (IL-2ra) increasing variants to elevated risk for multiple sclerosis (MS) and Crohn's disease (CD).<sup>17</sup> Aldesleukin, a recombinant form of IL-2 approved for cancer indications, is currently under investigation in a phase-2 clinical trial for CD (ClinicalTrials.gov ID: NCT04263831).<sup>43</sup> Also, compounds targeting IL-1 signaling, anakinra or canakinumab, represent established treatment algorithms for inflammatory joint diseases like rheumatoid arthritis (RA) or juvenile arthritis.<sup>44,45</sup>

Of the 24 signals, 4 also showed evidence of significant colocalization, that is a PP of association >80% for shared causal variants between cytokine levels and disease outcomes (**Figure 8, Supplementary Table S13**), thus providing even stronger evidence for causality. These included associations of higher genetically proxied G-CSF/CSF-3 levels with asthma, lower genetically proxied G-CSF/CSF-3 and higher genetically proxied CXCL9/MIG levels with CD, as well as lower genetically proxied TNF-b levels with MS. Furthermore, the association between genetically proxied IL-1 receptor antagonist (IL-1ra) levels and lower risk of RA reached a PPA of 68% for a shared causal variant in colocalization analysis. These results are consistent with data from preclinical studies,<sup>46-48</sup> observational studies in humans,<sup>49-53</sup> and clinical trials,<sup>54,55</sup> thus providing support for potentially promising targeted immunotherapies for these indications.

**Figure 8. *Cis*-Mendelian randomization associations and colocalization analyses between genetically proxied cytokine levels and disease risk.** Significant associations between circulating cytokine levels and disease outcomes are shown for allergic and autoimmune, cardiometabolic, and cancer outcomes. Effect sizes and log-transformed, Benjamin-Hochberg corrected p-values are illustrated by color gradient and circle size, respectively. Only cytokines and disease endpoints with at least 1 significant association are depicted. Numbers at the top indicates average number of *cis*-acting genetic variants used as instruments in MR analyses. Stars highlight significant genetic colocalizations (posterior probability of association >80%) for shared causal variants between circulating cytokine levels and disease risk.

## **Integration of cytokine-disease MR and TWAS-MR results implicates additional mediators of disease mechanisms that could represent promising drug targets**

As a last step, we aimed to integrate the cytokine-disease MR results with the TWAS MR results with the goal of also detecting upstream regulators of the potentially causal cytokines. We performed MR analyses between genetically proxied expression of genes significantly associated with G-CSF/CSF-3, CXCL9/MIG, and TNF-b in our TWAS-MR analyses and the associated disease outcomes. We found that higher genetically proxied expression of



*PPP1R37* is associated with lower levels of G-CSF/CSF-3, as well as with a lower risk of asthma (**Figure 9a**). We also found higher genetically proxied expression of *TRAFD1* to be associated with higher CXCL9/MIG levels and higher risk of CD (**Figure 9b**).

**Figure 9. Causal associations between genetic regulators for cytokines, circulating cytokine levels and disease risk.** (a) Genetically proxied mRNA for *PPP1R37*, *PVR*, *RTN2* and *IGFBP2* affect serum G-CSF/CSF-3 levels leading to increased risk for asthma. In turn, *PPP1R37* directly lowers disease risk for asthma. (b) Genetically proxied mRNA for 11 genes underlying CXCL9/MIG serum levels differentially affect circulating cytokine concentrations which influence the risk for Crohn's disease. Independently, *TRAFD1*, *ATF6B* and *C4A* also modulate disease risk for Crohn's disease.

## Discussion

Pooling data from up to 74,783 patients from 3 independent GWAS cohorts, we identified 169 independent genomic loci influencing the circulating concentration of 40 cytokines, 156 of which have not been associated with circulating cytokine levels in previous GWASs. Integrating our results with transcriptomic data, a TWAS-MR analysis revealed 245 potentially causal associations between gene expression of mostly immunoregulatory genes in peripheral blood and circulating cytokine levels. Analyzing regulatory interactions between cytokines, we found TNF-b, VEGF, and IL-1ra as master controllers of the circulating levels of multiple cytokines. Finally, we provide genetic evidence that the circulating levels of 3 cytokines (G-CSF/CSF-3, CXCL9/MIG, TNF-b) might be causally involved in the pathogenesis of asthma, CD, and MS, thus offering insights for the development of more specific immunotherapies.

Our MR and colocalization analyses provided evidence for potential causality for four cytokine-indication pairs, thus offering genetic support for potentially promising targeted immunotherapies for asthma, CD, and MS. These results are highly consistent with preclinical, epidemiological, and occasionally clinical data. For example, G-CSF/CSF-3 is a pro-inflammatory cytokine involved in neutrophil differentiation and systemic mobilization and has been implicated in the pathogenesis of neutrophilic atopic asthma.<sup>46,56</sup> Several preclinical studies in asthma models showed that blockage of upstream inductors or the receptor of G-CSF/CSF-3 reduced circulating cytokine levels, alleviated the airway inflammatory response, and improved disease outcome.<sup>46-48,57</sup> Furthermore, G-CSF/CSF-3 levels in the sputum of asthma patients have been suggested as a marker of airway neutrophilic inflammation.<sup>58</sup> Our results provide genetic support for the concept of targeting G-CSF/CSF-3 in asthma, potentially focusing on patients with neutrophilic asthma.

Using integrated results from our GWAS and TWAS findings, we identified an upstream mechanism of yet unknown relevance: higher *PPP1R37* gene expression was associated with reduced G-CSF/CSF-3 serum levels and lower risk of asthma. The gene encodes a regulatory subunit that acts as phosphatase inhibitor and has, so far, not been associated with airway diseases.<sup>59</sup> Previous studies investigating related regulatory subunits have unveiled potential biological mechanisms through which these subunits may influence the immune response<sup>60,61</sup> and genetic studies have provided support for the significance of other protein phosphatase regulatory subunits as contributing factor to airway diseases.<sup>62-64</sup> For example, genetically proxied expression of *PPP1R3D* was associated with disease characteristics of asthma including mucosal immunity, cell metabolism, and airway remodeling and predicted responsiveness to omalizumab therapy.<sup>62</sup> Although the specific biological mechanism

underlying our finding is unknown, one might speculate that the diverse range of functions associated with PP1, including cell progression, apoptosis, and muscle contraction, might underlie the observed findings.

The MR results supported a connection between genetically proxied circulating levels of 6 cytokines and CD. Among them, associations of higher CXCL9/MIG and lower G-CSF/CSF-3 with CD were also supported by colocalization evidence. CXCL9/MIG, a proinflammatory IFN $\gamma$ -induced CXC-chemokine, is released by various immune cells including macrophages to attract and activate T-cells and forms together with neighboring CXCL10/IP-10, CXCL11/IP-9 and their cognate CXCR3 receptor an axis with specific relevance in inflammatory bowel disease.<sup>65,66</sup> In clinical studies, elevated CXCL9/MIG serum levels have been associated with CD relapses, affirming CXCL9/MIG as a risk factor in CD.<sup>53</sup> Apart from pro-inflammatory actions, exogenous administration of G-CSF/CSF-3 has been associated with immunoregulatory effects, such as modulation of T-cell responses.<sup>67</sup> Two open-label studies have indeed demonstrated that subcutaneous G-CSF/CSF-3 is effective in inducing clinical remission, fostering mucosal healing, and normalizing cell counts and cytokine responses in CD patients.<sup>54,55</sup>

Integrating our results linking CD with circulating CXCL9/MIG levels, we identified genetically proxied expression of *TRAFD1* as a potential upstream causal regulator of CXCL9/MIG levels and subsequently risk of CD.<sup>53,68,69</sup> TRAFD1 binds as homodimer or in interaction with TRAFD2, to TNF receptors, impacting pro-inflammatory cytokine production and modulating inflammatory responses in immune cells.<sup>32,33,70</sup> In celiac disease, *TRAFD1* was recognized as an upstream regulator of IFN $\gamma$  signaling and thereby activating cytotoxic T-cells, an important pathomechanism.<sup>34</sup> The identified function of *TRAFD1* as inductor of IFN $\gamma$  signaling aligns well with the literature and also with our findings showing increased CXCL9/MIG serum levels underlying *TRAFD1* expression.<sup>66</sup> Substantiating the risk-increasing associations in our analysis, elevated expression of *TRAFD1* and *TRAFD2* was noticed in acutely inflamed mucosal biopsies of CD patients.<sup>71</sup> Together these reports confirm our findings for the importance of *TRAFD1* in the pathology of CD and as modulator of inflammatory reactions through effects on cytokine levels.

Our analysis also provided evidence for an inverse association between genetically proxied circulating TNF-b and MS risk. TNF-b, is a pro-inflammatory cytokine within the TNF superfamily with substantial genetic correlation to TNF-a and binding affinity to pro-inflammatory as well as anti-inflammatory TNF receptors.<sup>72</sup> In a randomized phase 2 trial the TNF-inhibitor lenercept was tested for safety and efficacy in MS but had to be terminated prematurely after the interim analysis detected a dose-dependent increase in the frequency and severity of MS exacerbations.<sup>73</sup> In contrast to the TNF blockers infliximab, adalimumab and golimumab, lenercept provides equal inhibitory efficacy for TNF-a and TNF-b.<sup>72,74</sup> While TNF inhibition has demonstrated success in treating autoimmune diseases such as rheumatoid arthritis or psoriasis, patients undergoing anti-TNF therapy for these indications are at risk for developing demyelinating CNS lesions, indicating a disease-specific effect.<sup>40,75</sup> Supporting the clinical findings, GWAS studies in MS identified TNF lowering alleles for both cytokines TNF-a and TNF-b that were associated with higher risk for MS.<sup>23,76</sup> We provide additional genetic evidence in line with observational, clinical, and GWAS findings for a potentially protective role of TNF-b in MS.

In our network analysis, over 80 percent of the significant cytokine-cytokine interactions led to an increase in downstream cytokine concentrations. Given that the majority of the involved cytokines were pro-inflammatory implies a self-perpetuating feedback mechanism leading to

strong inflammatory responses and suggests a global trend where cytokines mutually activate each other to amplify their immune reaction. For the IL-1ra and TNF-b network, specifically, we successfully expanded the list of downstream affected cytokines. TNF-b exhibited notable downstream effects on 13 other cytokines, signifying its role as a master regulator. This observation aligns with existing reports on *TRAFDI*'s downstream effects, influencing various cytokine-encoding genes like CXCL10/IP-10 and IFN $\gamma$ .<sup>34</sup>

Our study has limitations. First, our meta-analysis was based on 3 cohorts that used different affinity-based assaying approaches for quantifying circulating cytokines. The different approaches might yield varying measurements for the same proteins with only moderate correlations across the assays.<sup>18</sup> This might explain the difference in replication rates across the 3 cohorts. Interestingly, we found a higher replication rate for signals detected with the Olink assay. In a previous cross-assay comparison between Olink and SomaScan, the proportion of assays with detected pQTLs was also higher with the Olink-based assay.<sup>18</sup> The differences across the panels should be further explored at a larger scale to explore the extent to which it would be possible to scale genetic explorations across cohorts utilizing different proteomic platforms. Second, due to differences in reporting of effect sizes for genetic variants across the GWAS source data, we could perform only p-value-based meta-analyses and only indirectly estimate the pooled effect sizes based on the derived p-values and the variant allele frequencies. Inaccuracies in this estimation could influence downstream analyses heavily relying on effect sizes. Third, due to data availability, our analyses were based on 40 selected cytokines. Future endeavors utilizing solely high-throughput proteomic data could scale up to analyses including more inflammatory proteins. Fourth, our analyses are based on individuals of European, Finnish, and Icelandic ancestry and as such might not be generalizable to individuals of a different ancestry background. Fifth, due to the large number of cytokines we adjusted the significance level for multiple testing which might have neglected important findings due to non-significance. Using a hypothesis-driven approach, future studies should follow-up on our results to identify additional targets we might have missed for a comprehensive view around our findings. Sixth our GWAS meta-analysis was based on population-based cohorts without predominant inflammatory diseases. Genetic variants might influence cytokine levels in specific contexts such as a response to infection or other pro-inflammatory stimuli; our approach could not detect such signals.

## Conclusion

In conclusion, our study, leveraging data from 74,783 individuals across 3 cohorts, identified 169, mostly novel, genomic loci influencing circulating cytokine levels. Follow-up analyses of the detected signals reveal interesting underlying pathways, which enhance our understanding of the biology of the immune response. Integrating our data with genetic data for human disease risk, our analyses suggest potential targets like G-CSF/CSF-3, CXCL9/MIG, and TNF-b for immune-related diseases including asthma, Crohn's disease, and multiple sclerosis, warranting further exploration in clinical trials. The summary statistics from our study offer a valuable resource for future omics analyses, aiding data integration for the identification of potential drug targets for human diseases.

## References

1. Propper DJ, Balkwill FR. Harnessing cytokines and chemokines for cancer therapy. *Nat Rev Clin Oncol*. 2022;19(4):237-253. doi:10.1038/s41571-021-00588-9
2. Deckers J, Anbergen T, Hokke AM, et al. Engineering cytokine therapeutics. *Nat Rev Bioeng*. 2023;1(4):286-303. doi:10.1038/s44222-023-00030-y
3. Attwood MM, Jonsson J, Rask-Andersen M, Schiöth HB. Soluble ligands as drug targets. *Nat Rev Drug Discov*. 2020;19(10):695-710. doi:10.1038/s41573-020-0078-4
4. Lutgens E, Atzler D, Döring Y, Duchene J, Steffens S, Weber C. Immunotherapy for cardiovascular disease. *Eur Heart J*. 2019;40(48):3937-3946. doi:10.1093/eurheartj/ehz283
5. Soehnlein O, Libby P. Targeting inflammation in atherosclerosis - from experimental insights to the clinic. *Nat Rev Drug Discov*. 2021;20(8):589-610. doi:10.1038/s41573-021-00198-1
6. Tímár J, Uhlyarik A. On-Target Side Effects of Targeted Therapeutics of Cancer. *Pathol Oncol Res POR*. 2022;28:1610694. doi:10.3389/pore.2022.1610694
7. Holmes MV, Richardson TG, Ference BA, Davies NM, Davey Smith G. Integrating genomics with biomarkers and therapeutic targets to invigorate cardiovascular drug development. *Nat Rev Cardiol*. 2021;18(6):435-453. doi:10.1038/s41569-020-00493-1
8. Minikel EV, Karczewski KJ, Martin HC, et al. Evaluating drug targets through human loss-of-function genetic variation. *Nature*. 2020;581(7809):459-464. doi:10.1038/s41586-020-2267-z
9. King EA, Davis JW, Degner JF. Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. Marchini J, ed. *PLOS Genet*. 2019;15(12):e1008489. doi:10.1371/journal.pgen.1008489
10. Gill D, Georgakis MK, Walker VM, et al. Mendelian randomization for studying the effects of perturbing drug targets. *Wellcome Open Res*. 2021;6:16. doi:10.12688/wellcomeopenres.16544.2
11. Kappelmann N, Arloth J, Georgakis MK, et al. Dissecting the Association Between Inflammation, Metabolic Dysregulation, and Specific Depressive Symptoms: A Genetic Correlation and 2-Sample Mendelian Randomization Study. *JAMA Psychiatry*. 2021;78(2):161-170. doi:10.1001/jamapsychiatry.2020.3436
12. Mokry LE, Zhou S, Guo C, et al. Interleukin-18 as a drug repositioning opportunity for inflammatory bowel disease: A Mendelian randomization study. *Sci Rep*. 2019;9(1):9386. doi:10.1038/s41598-019-45747-2
13. Chong M, Sjaarda J, Pigeyre M, et al. Novel Drug Targets for Ischemic Stroke Identified Through Mendelian Randomization Analysis of the Blood Proteome. *Circulation*. 2019;140(10):819-830. doi:10.1161/CIRCULATIONAHA.119.040180
14. Bouras E, Karhunen V, Gill D, et al. Circulating inflammatory cytokines and risk of five cancers: a Mendelian randomization analysis. *BMC Med*. 2022;20(1):3. doi:10.1186/s12916-021-02193-0
15. Georgakis MK, Malik R, Gill D, et al. Interleukin-6 Signaling Effects on Ischemic Stroke and Other Cardiovascular Outcomes: A Mendelian Randomization Study. *Circ Genomic Precis Med*. 2020;13(3):e002872. doi:10.1161/CIRCGEN.119.002872
16. Georgakis MK, Gill D, Rannikmäe K, et al. Genetically Determined Levels of Circulating Cytokines and Risk of Stroke. *Circulation*. 2019;139(2):256-268. doi:10.1161/CIRCULATIONAHA.118.035905
17. Ahola-Olli AV, Würtz P, Havulinna AS, et al. Genome-wide Association Study Identifies 27 Loci Influencing Concentrations of Circulating Cytokines and Growth Factors. *Am J Hum Genet*. 2017;100(1):40-50. doi:10.1016/j.ajhg.2016.11.007
18. Eldjarn GH, Ferkingstad E, Lund SH, et al. Large-scale plasma proteomics comparisons through genetics and disease associations. *Nature*. 2023;622(7982):348-358. doi:10.1038/s41586-023-06563-x

19. Sollis E, Mosaku A, Abid A, et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res.* 2023;51(D1):D977-D985. doi:10.1093/nar/gkac1010
20. Parente R, Clark SJ, Inforzato A, Day AJ. Complement factor H in host defense and immune evasion. *Cell Mol Life Sci CMLS.* 2017;74(9):1605-1624. doi:10.1007/s00018-016-2418-4
21. Rose AB. Introns as Gene Regulators: A Brick on the Accelerator. *Front Genet.* 2019;9. Accessed April 7, 2023. <https://www.frontiersin.org/articles/10.3389/fgene.2018.00672>
22. Vösa U, Claringbould A, Westra HJ, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet.* 2021;53(9):1300-1310. doi:10.1038/s41588-021-00913-z
23. Zhao JH, Stacey D, Eriksson N, et al. Genetics of circulating inflammatory proteins identifies drivers of immune-mediated disease risk and therapeutic targets. *Nat Immunol.* 2023;24(9):1540-1551. doi:10.1038/s41590-023-01588-w
24. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science.* 2020;369(6509):1318-1330. doi:10.1126/science.aaz1776
25. Schnabel RB, Baumert J, Barbalic M, et al. Duffy antigen receptor for chemokines (Darc) polymorphism regulates circulating concentrations of monocyte chemoattractant protein-1 and other inflammatory mediators. *Blood.* 2010;115(26):5289-5299. doi:10.1182/blood-2009-05-221382
26. Korbecki J, Barczak K, Gutowska I, Chlubek D, Baranowska-Bosiacka I. CXCL1: Gene, Promoter, Regulation of Expression, mRNA Stability, Regulation of Activity in the Intercellular Space. *Int J Mol Sci.* 2022;23(2):792. doi:10.3390/ijms23020792
27. Szpakowska M, D'Uonno G, Luís R, et al. New pairings and deorphanization among the atypical chemokine receptor family — physiological and clinical relevance. *Front Immunol.* 2023;14:1133394. doi:10.3389/fimmu.2023.1133394
28. Crawford KS, Volkman BF. Prospects for targeting ACKR1 in cancer and other diseases. *Front Immunol.* 2023;14:1111960. doi:10.3389/fimmu.2023.1111960
29. Jiménez-Sousa, Gómez-Moreno, Pineda-Tenor, et al. Impact of DARC rs12075 Variants on Liver Fibrosis Progression in Patients with Chronic Hepatitis C: A Retrospective Study. *Biomolecules.* 2019;9(4):143. doi:10.3390/biom9040143
30. Chen MH, Raffield LM, Mousas A, et al. Trans-ethnic and Ancestry-Specific Blood-Cell Genetics in 746,667 Individuals from 5 Global Populations. *Cell.* 2020;182(5):1198-1213.e14. doi:10.1016/j.cell.2020.06.045
31. Takechi R, Galay RL, Matsuo T, et al. Role of the tumor necrosis factor receptor-associated factor-type zinc finger domain containing protein 1 (TRAFD1) from the hard tick *Haemaphysalis longicornis* in immunity against bacterial infection. *Ticks Tick-Borne Dis.* 2016;7(1):36-45. doi:10.1016/j.ttbdis.2015.08.002
32. Park HH. Structural feature of TRAFs, their related human diseases and therapeutic intervention. *Arch Pharm Res.* 2021;44(5):475-486. doi:10.1007/s12272-021-01330-w
33. Lalani AI, Zhu S, Gokhale S, Jin J, Xie P. TRAF Molecules in Inflammation and Inflammatory Diseases. *Curr Pharmacol Rep.* 2018;4(1):64-90. doi:10.1007/s40495-017-0117-y
34. Van Der Graaf A, Zorro MM, Claringbould A, et al. Systematic Prioritization of Candidate Genes in Disease Loci Identifies TRAFD1 as a Master Regulator of IFN $\gamma$  Signaling in Celiac Disease. *Front Genet.* 2021;11:562434. doi:10.3389/fgene.2020.562434
35. Fantus RJ, Na R, Wei J, et al. Genetic Susceptibility for Low Testosterone in Men and Its Implications in Biology and Screening: Data from the UK Biobank. *Eur Urol Open Sci.* 2021;29:36-46. doi:10.1016/j.euros.2021.04.010

36. Gao MM, Hu F, Zeng XD, et al. Hypothalamic proteome changes in response to nicotine and its withdrawal are potentially associated with alteration in body weight. *J Proteomics*. 2020;214:103633. doi:10.1016/j.jprot.2020.103633
37. Boersma B, Jiskoot W, Lowe P, Bourquin C. The interleukin-1 cytokine family members: Role in cancer pathogenesis and potential therapeutic applications in cancer immunotherapy. *Cytokine Growth Factor Rev*. 2021;62:1-14. doi:10.1016/j.cytogfr.2021.09.004
38. Basu A, Krady JK, Levison SW. Interleukin-1: A master regulator of neuroinflammation. *J Neurosci Res*. 2004;78(2):151-156. doi:10.1002/jnr.20266
39. Okada T, Kawada T. [Function of implanted prosthetic heart valves]. *Kokyu To Junkan*. 1987;35(10):1053-1058.
40. Ding Si, Xie J, Wan Q. Association Between Cytokines and Their Receptor Antagonist Gene Polymorphisms and Clinical Risk Factors and Acute Rejection Following Renal Transplantation. *Med Sci Monit Int Med J Exp Clin Res*. 2016;22:4736-4741. doi:10.12659/msm.898193
41. Daniel D, Wilson NS. Tumor necrosis factor: renaissance as a cancer therapeutic? *Curr Cancer Drug Targets*. 2008;8(2):124-131. doi:10.2174/156800908783769346
42. Zuber V, Grinberg NF, Gill D, et al. Combining evidence from Mendelian randomization and colocalization: Review and comparison of approaches. *Am J Hum Genet*. 2022;109(5):767-782. doi:10.1016/j.ajhg.2022.04.001
43. Allegretti JR, Mitsialis V, Canavan JB, Low-Dose IL2 UC Study Group, Snapper SB. Low-Dose Interleukin 2 for the Treatment of Moderate to Severe Ulcerative Colitis. *Gastroenterology*. 2023;165(2):492-495.e2. doi:10.1053/j.gastro.2023.03.230
44. Bedaiwi MK, Almaghlouth I, Omair MA. Effectiveness and adverse effects of anakinra in treatment of rheumatoid arthritis: a systematic review. *Eur Rev Med Pharmacol Sci*. 2021;25(24):7833-7839. doi:10.26355/eurev\_202112\_27630
45. Shaul E, Conrad MA, Dawany N, et al. Canakinumab for the treatment of autoinflammatory very early onset- inflammatory bowel disease. *Front Immunol*. 2022;13:972114. doi:10.3389/fimmu.2022.972114
46. Ouyang S, Liu C, Xiao J, Chen X, Lui AC, Li X. Targeting IL-17A/glucocorticoid synergy to CSF3 expression in neutrophilic airway diseases. *JCI Insight*. 2020;5(3):e132836, 132836. doi:10.1172/jci.insight.132836
47. Kwak DW, Park D, Kim JH. Leukotriene B4 Receptor 2 Mediates the Production of G-CSF That Plays a Critical Role in Steroid-Resistant Neutrophilic Airway Inflammation. *Biomedicines*. 2022;10(11):2979. doi:10.3390/biomedicines10112979
48. Wang H, FitzPatrick M, Wilson NJ, et al. CSF3R/CD114 mediates infection-dependent transition to severe asthma. *J Allergy Clin Immunol*. 2019;143(2):785-788.e6. doi:10.1016/j.jaci.2018.10.001
49. Åkesson J, Hojjati S, Hellberg S, et al. Proteomics reveal biomarkers for diagnosis, disease activity and long-term disability outcomes in multiple sclerosis. *Nat Commun*. 2023;14(1):6903. doi:10.1038/s41467-023-42682-9
50. Romme Christensen J, Börnsen L, Hesse D, et al. Cellular sources of dysregulated cytokines in relapsing-remitting multiple sclerosis. *J Neuroinflammation*. 2012;9(1):215. doi:10.1186/1742-2094-9-215
51. Hojjati S, Ernerudh J, Vrethem M, Mellergård J, Raffetseder J. Dimethyl fumarate treatment in relapsing remitting MS changes the inflammatory CSF protein profile by a prominent decrease in T-helper 1 immunity. *Mult Scler Relat Disord*. 2023;80:105126. doi:10.1016/j.msard.2023.105126
52. Huang J, Khademi M, Fugger L, et al. Inflammation-related plasma and CSF biomarkers for multiple sclerosis. *Proc Natl Acad Sci*. 2020;117(23):12952-12960. doi:10.1073/pnas.1912839117

53. Walshe M, Nayeri S, Ji J, et al. A Role for CXCR3 Ligands as Biomarkers of Post-Operative Crohn's Disease Recurrence. *J Crohns Colitis*. 2022;16(6):900-910. doi:10.1093/ecco-jcc/jjab186
54. Dejaco C, Lichtenberger C, Miehsler W, et al. An open-label pilot study of granulocyte colony-stimulating factor for the treatment of severe endoscopic postoperative recurrence in Crohn's disease. *Digestion*. 2003;68(2-3):63-70. doi:10.1159/000074517
55. Korzenik JR, Dieckgraefe BK. An open □labelled study of granulocyte colony □stimulating factor in the treatment of active Crohn's disease. *Aliment Pharmacol Ther*. 2005;21(4):391-400. doi:10.1111/j.1365-2036.2005.02287.x
56. Tsioumpekou M, Krijgsman D, Leusen JHW, Olofsen PA. The Role of Cytokines in Neutrophil Development, Tissue Homing, Function and Plasticity in Health and Disease. *Cells*. 2023;12(15):1981. doi:10.3390/cells12151981
57. Wang H, Aloe C, McQualter J, et al. G-CSFR antagonism reduces mucosal injury and airways fibrosis in a virus-dependent model of severe asthma. *Br J Pharmacol*. 2021;178(8):1869-1885. doi:10.1111/bph.15415
58. Kim YM, Kim H, Lee S, et al. Airway G-CSF identifies neutrophilic inflammation and contributes to asthma progression. *Eur Respir J*. 2020;55(2):1900827. doi:10.1183/13993003.00827-2019
59. PPP1R37 protein phosphatase 1 regulatory subunit 37 [Homo sapiens (human)]. Gene - NCBI. Published December 31, 2023. <https://www.ncbi.nlm.nih.gov/gene/284352>.
60. Yadav L, Tamene F, Göös H, et al. Systematic Analysis of Human Protein Phosphatase Interactions and Dynamics. *Cell Syst*. 2017;4(4):430-444.e5. doi:10.1016/j.cels.2017.02.011
61. Haystead TAJ. ZIP kinase, a key regulator of myosin protein phosphatase 1. *Cell Signal*. 2005;17(11):1313-1322. doi:10.1016/j.cellsig.2005.05.008
62. Kidwai S, Barbiero P, Meijerman I, et al. A robust mRNA signature obtained via recursive ensemble feature selection predicts the responsiveness of omalizumab in moderate □to □severe asthma. *Clin Transl Allergy*. 2023;13(11):e12306. doi:10.1002/ctt2.12306
63. Freidin MB, Polonikov AV. Validation of PPP1R12B as a candidate gene for childhood asthma in Russians. *J Genet*. 2013;92(1):93-96. doi:10.1007/s12041-013-0210-x
64. Andiappan AK, Sio YY, Lee B, et al. Functional variants of 17q12-21 are associated with allergic asthma but not allergic rhinitis. *J Allergy Clin Immunol*. 2016;137(3):758-766.e3. doi:10.1016/j.jaci.2015.08.038
65. Tokunaga R, Zhang W, Naseem M, et al. CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation - A target for novel cancer therapy. *Cancer Treat Rev*. 2018;63:40-47. doi:10.1016/j.ctrv.2017.11.007
66. Ding Q, Lu P, Xia Y, et al. CXCL9: evidence and contradictions for its role in tumor progression. *Cancer Med*. 2016;5(11):3246-3259. doi:10.1002/cam4.934
67. Martins A, Han J, Kim SO. The multifaceted effects of granulocyte colony □stimulating factor in immunomodulation and potential roles in intestinal immune homeostasis. *IUBMB Life*. 2010;62(8):611-617. doi:10.1002/iub.361
68. Liu B, Qian Y, Li Y, et al. Circulating levels of cytokines and risk of inflammatory bowel disease: evidence from genetic data. *Front Immunol*. 2023;14:1310086. doi:10.3389/fimmu.2023.1310086
69. Huang J, Su B, Karhunen V, et al. Inflammatory Diseases, Inflammatory Biomarkers, and Alzheimer Disease: An Observational Analysis and Mendelian Randomization. *Neurology*. 2023;100(6). doi:10.1212/WNL.0000000000201489
70. Edilova MI, Abdul-Sater AA, Watts TH. TRAF1 Signaling in Human Health and Disease. *Front Immunol*. 2018;9. Accessed January 13, 2024. <https://www.frontiersin.org/articles/10.3389/fimmu.2018.02969>

71. Qiao YQ, Shen J, Gu Y, et al. Gene expression of tumor necrosis factor receptor associated factor ( TRAF ) 1 and TRAF 2 in inflammatory bowel disease. *J Dig Dis*. 2013;14(5):244-250. doi:10.1111/1751-2980.12044
72. Ruddle NH. Lymphotoxin and TNF: How it all began—A tribute to the travelers. *Cytokine Growth Factor Rev*. 2014;25(2):83-89. doi:10.1016/j.cytogfr.2014.02.001
73. The Lenercept Multiple Sclerosis Study Group and The University of British Columbia MS/MRI Analysis Group. TNF neutralization in MS: Results of a randomized, placebo-controlled multicenter study. *Neurology*. 2011;77(14):1382-1382. doi:10.1212/01.wnl.0000406608.85830.82
74. Mitoma H, Horiuchi T, Tsukamoto H, Ueda N. Molecular mechanisms of action of anti-TNF- $\alpha$  agents – Comparison among therapeutic TNF- $\alpha$  antagonists. *Cytokine*. 2018;101:56-63. doi:10.1016/j.cyto.2016.08.014
75. Li L, Aviña-Zubieta JA, Bernstein CN, et al. Risk of Multiple Sclerosis Among Users of Antitumor Necrosis Factor  $\alpha$  in 4 Canadian Provinces: A Population-Based Study. *Neurology*. 2023;100(6). doi:10.1212/WNL.0000000000201472
76. Gregory AP, Dendrou CA, Attfield KE, et al. TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. *Nature*. 2012;488(7412):508-511. doi:10.1038/nature11307
77. Ferkingstad E, Sulem P, Atlason BA, et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet*. 2021;53(12):1712-1721. doi:10.1038/s41588-021-00978-w
78. Folkersen L, Gustafsson S, Wang Q, et al. Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. *Nat Metab*. 2020;2(10):1135-1148. doi:10.1038/s42255-020-00287-2
79. 1000 Genomes Project Consortium, Abecasis GR, Altshuler D, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010;467(7319):1061-1073. doi:10.1038/nature09534
80. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinforma Oxf Engl*. 2010;26(17):2190-2191. doi:10.1093/bioinformatics/btq340
81. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet*. 2016;48(5):481-487. doi:10.1038/ng.3538
82. Greco M FD, Minelli C, Sheehan NA, Thompson JR. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat Med*. 2015;34(21):2926-2940. doi:10.1002/sim.6522
83. Suzuki K, Hatzikotoulas K, Southam L, et al. Genetic drivers of heterogeneity in type 2 diabetes pathophysiology. *Nature*. Published online February 19, 2024. doi:10.1038/s41586-024-07019-6
84. Zheng J, Erzurumluoglu AM, Elsworth BL, et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinforma Oxf Engl*. 2017;33(2):272-279. doi:10.1093/bioinformatics/btw613
85. Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. 2015;47(3):291-295. doi:10.1038/ng.3211
86. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015;47(11):1236-1241. doi:10.1038/ng.3406
87. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575. doi:10.1086/519795



88. Wellcome Trust Case Control Consortium, Maller JB, McVean G, et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat Genet.* 2012;44(12):1294-1301. doi:10.1038/ng.2435
89. Park JH, Wacholder S, Gail MH, et al. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat Genet.* 2010;42(7):570-575. doi:10.1038/ng.610
90. Kamat MA, Blackshaw JA, Young R, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinforma Oxf Engl.* 2019;35(22):4851-4853. doi:10.1093/bioinformatics/btz469
91. Staley JR, Blackshaw J, Kamat MA, et al. PhenoScanner: a database of human genotype-phenotype associations. *Bioinforma Oxf Engl.* 2016;32(20):3207-3209. doi:10.1093/bioinformatics/btw373
92. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun.* 2017;8(1):1826. doi:10.1038/s41467-017-01261-5
93. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol.* 2015;11(4):e1004219. doi:10.1371/journal.pcbi.1004219
94. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol.* 2013;37(7):658-665. doi:10.1002/gepi.21758
95. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol.* 2015;44(2):512-525. doi:10.1093/ije/dyv080
96. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol.* 2016;40(4):304-314. doi:10.1002/gepi.21965
97. Schmidt AF, Finan C, Gordillo-Marañón M, et al. Genetic drug target validation using Mendelian randomisation. *Nat Commun.* 2020;11(1):3255. doi:10.1038/s41467-020-16969-0
98. Tsuo K, Zhou W, Wang Y, et al. Multi-ancestry meta-analysis of asthma identifies novel associations and highlights the value of increased power and diversity. *Cell Genomics.* 2022;2(12):100212. doi:10.1016/j.xgen.2022.100212
99. Liu JZ, van Sommeren S, Huang H, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet.* 2015;47(9):979-986. doi:10.1038/ng.3359
100. International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science.* 2019;365(6460):eaav7188. doi:10.1126/science.aav7188
101. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature.* 2018;562(7726):203-209. doi:10.1038/s41586-018-0579-z
102. Okada Y, Wu D, Trynka G, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature.* 2014;506(7488):376-381. doi:10.1038/nature12873
103. Klarin D, Lynch J, Aragam K, et al. Genome-wide association study of peripheral artery disease in the Million Veteran Program. *Nat Med.* 2019;25(8):1274-1279. doi:10.1038/s41591-019-0492-5
104. Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet.* 2015;47(10):1121-1130. doi:10.1038/ng.3396

105. Mishra A, Malik R, Hachiya T, et al. Stroke genetics informs drug discovery and risk prediction across ancestries. *Nature*. 2022;611(7934):115-123. doi:10.1038/s41586-022-05165-3
106. Zhang H, Ahearn TU, Lecarpentier J, et al. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat Genet*. 2020;52(6):572-581. doi:10.1038/s41588-020-0609-2
107. Burrows K, Haycock P. Genome-wide Association Study of Cancer Risk in UK Biobank. Published online 2021. doi:10.5523/BRIS.AED0U12W0EDE20OLB0M77P4B9
108. McKay JD, Hung RJ, Han Y, et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat Genet*. 2017;49(7):1126-1132. doi:10.1038/ng.3892
109. Rashkin SR, Graff RE, Kachuri L, et al. Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. *Nat Commun*. 2020;11(1):4423. doi:10.1038/s41467-020-18246-6
110. Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. Williams SM, ed. *PLoS Genet*. 2014;10(5):e1004383. doi:10.1371/journal.pgen.1004383
111. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res*. 2014;42(D1):D1001-D1006. doi:10.1093/nar/gkt1229

## Materials & Methods

### Study populations and design

The study cohorts and a flowchart of the study design are depicted in **Figure 1**. We downloaded publicly available GWAS summary statistics for the circulating levels of up to 40 cytokines from 3 independent cohorts. Details of the study protocols have been published elsewhere.<sup>17,77,78</sup> Human genome assembly GRCh37 (hg19) from Genome Reference Consortium was used for genomic positioning.<sup>79</sup> Before further computations, all 3 databases were harmonized regarding data structure. For the GWAS meta-analyses and downstream computations we included all cytokines that were available in at least 2 cohorts. To ensure that the available cytokines were identical between cohorts we used information provided on the NIH (<https://www.ncbi.nlm.nih.gov/gap/>) and GeneCards (<https://www.genecards.org/>) websites and verified synonyms and aliases in the abbreviated and full names of the cytokines. We excluded 1 cytokine (IL-12) from the analyses, because different subunits of the protein were quantified in the two cohorts (IL-12p70 in YFS & FINRISK and IL-12p40 in deCODE).

### *YFS and FINRISK*

Genomic data for 40 cytokines were drawn from up to 8,293 individuals of Finnish background that were included in the YFS & FINRISK cohorts 1997 and 2002, respectively.<sup>17</sup> The mean age across all studies was 49 years (standard deviation 8 years). The cytokine measurements were carried out in EDTA plasma for the FINRISK 1997 cohort, in heparin plasma for the FINRISK 2002 cohort and in serum for the YFS cohort using cytokine Luminex®-based multiplex immunoassays from Bio-Rad®. Genotyping was completed using the Illumina HT12 platform for the YFS study and the Illumina 670k HumanHap array for both FINRISK studies. Imputation was performed using the 1000 Genomes reference panel across all cohorts.<sup>79</sup> The GWAS meta-analyzing all 3 studies normalized the cytokine distribution using inverse transformation and adjusted the genetic analyses for age, sex and ancestral principal components 1-10. The reported effect sizes were scaled per standard deviation increment in inverse-transformed cytokine levels.

### *SCALLOP Consortium*

Genomic data for 16 cytokines were drawn from up to 30,931 individuals with European background from the SCALLOP consortium, a collaborative framework analyzing gene-protein associations across 13 studies.<sup>78</sup> The cytokine measurements were carried out in plasma samples using the proximity extension assay-based Olink® platform. Genotyping methods across the studies included Cardiometabochip, Immunochip, PsychChip, Illumina HumanCoreExome, Illumina OmniExpress, Metabochip, Illumina OmniExpress 2.5, Affymetrix Axiom UK Biobank array, HumanCytoSNP-12 BeadChip, HapMap300v2, Human Exome, Illumina HumanOmniExpressExome-8 v1, Illumina HumanHap300v1, Omni1, OmniX, Illumina HumanHap300v1 and Infinium PsychArray-24 v1.2. Imputation was performed using the following panels: 1000G phase v5, 1000G phase v3, UK10K reference panel, HRC, HRC r1.1. The GWAS meta-analyzing all 13 studies adjusted the cytokines for age, gender, site, OLINK batch, Olink plate, MDS components, storage time, bleed to processing time (days), smoking status, oral contraceptive usage, blood cell counts, season of venipuncture and ancestral principal components 1-10. The log<sub>2</sub>-based normalized expression values (NPX) for each protein were rank-based inverse normal transformed and standardized to units of standard deviation.

### *deCODE*

Genomic and proteomic data for 39 cytokines were taken from 35,559 Icelandic individuals included in deCODE.<sup>77</sup> The mean age was 55 years (standard deviation 18 years). The cytokine measurements were carried out in plasma samples using the aptamer-based SOMAScan<sup>®</sup> assay. Genotyping was completed using Illumina SNP Chip. Imputation was based on an in-house developed whole genome sequencing reference panel. The genetic analyses were adjusted for age and sex. Cytokine measurements were normalized using rank-inverse normal transformation and standardized to standard deviation increment. To allow alignment with other datasets, we excluded all SNPs that were not covered by the 1000 Genomes reference panel.

### **Cross-assay comparisons in the genetic architecture of circulating cytokines**

To explore differences in the genomic architecture of cytokines levels between the 3 studies that applied different measurement assays, we compared the proportion of overlapping SNPs between datasets confined to significant ( $p$ -value  $< 0.05$ ) and directionally concordant (same direction of effect estimates across all 3 databases) variants. Using the raw data, we analyzed overlapping SNPs by taking one dataset as reference comparing it to the other two.

### **GWAS meta-analyses**

We performed fixed-effects inverse variance-weighted meta-analysis for each cytokine across the available cohorts using METAL software (v.2011-03-25, number of cohorts and sample sizes per cytokine GWAS are provided in **Supplementary Table S14**).<sup>80</sup> Due to differences in scaling of the derived effect estimates across the 3 datasets, we applied a z-score-based meta-analysis (SCHEME SAMPLESIZE). Subsequently, we estimated standardized beta coefficients using p-values, minimum allele frequency, and direction of effects, weighted according to sample sizes, as previously described.<sup>81</sup> For estimation of heterogeneity of effect sizes between data sources we calculated chi-square test statistics for all included markers. To control for genomic inflation we calculated lambda statistics for each cytokine (**Supplementary Table S3**).<sup>82</sup> Significant variants were defined based on the established genome-wide significance level ( $p < 5 \times 10^{-8}$ ). To detect independent variants following correction for linkage disequilibrium, we clumped across the significant ones using `clump_data` (TwoSampleMR R package version 0.5.6) at an  $r^2 < 0.001$  based on the European 1000 Genomes Project reference panel.<sup>79</sup> We defined independent loci as SNPs that were separated by more than 1 Mb from the next SNPs in the 3' and 5' direction, as reported earlier.<sup>83</sup>

### **Linkage disequilibrium score regression (LDSC)**

Using the LD score v1.0.1. tool we applied LDSC regression with reference data from the European 1000 Genomes project for calculation of cross-trait LDSC genetic correlations between all 40 cytokines using the meta-analysis results.<sup>79,84–86</sup>

### **Fine mapping, functional annotation, pathway and gene-set analysis**

To identify causal variants responsible for variations in circulating cytokine concentrations, we investigated significant loci associated with cytokines. We employed PLINK v1.9 to compute LD score correlation matrices and further refined the results using SuSiE (susieR R package version 0.12.16) to derive sets of variants, ensuring the inclusion of at least one causal variant with a cumulative probability  $\geq 95\%$ .<sup>87,88</sup> Subsequently, the causal variants were utilized to estimate the total variance explained by the identified loci for individual cytokines.<sup>89</sup> For functional analyses we used phenoscanner (MendelianRandomization R package version 0.6.0) which ascribes functional consequences (intron, intergenic, exon, upstream, downstream, etc.) of single variants using positional mapping (physical distance).<sup>90,91</sup> Gene-property analyses was conducted for identification of the tissue specificity

of cytokines using the FUMA Gene2Func web database.<sup>92</sup> Lastly, MAGMA gene-based and gene-set analyses were conducted. Gene-based analysis initially calculates p-value association tests for variants mapped to protein coding genes which are then used to calculate gene-set p-values in the gene-set analysis. Using predefined gene-sets, variants with significant associations to genes can then be analyzed to determine their underlying functional or process-related feature, i.e. gene-sets belonging to molecular functions or biological processes.<sup>93</sup>

### **Transcriptome-wide Mendelian randomization analysis (TWAS-MR)**

To further explore whether variant effects on expression of specific genes underlie the genetic underpinnings of circulating cytokine levels, we performed transcriptome-wide inverse variance-weighted 2-sample MR analysis, as has been previously described.<sup>94</sup> Sensitivity analyses were conducted using MR Egger regression and the weighted median estimator to control for horizontal pleiotropy.<sup>95,96</sup> For calculation of effect estimates we used the `mr` command from the TwoSampleMR R Package (TwoSampleMR version 0.5.6) with *cis*-expression quantitative trait loci (eQTL) gene instruments from the eQTLGen Consortium as exposure (clumped at  $r^2 < 0.01$ ) and the GWAS meta-analysis results of our cytokine panel as outcome. The eQTL consortium included 31,684 individuals of primarily European ancestry (detailed methods have been described previously and are available online).<sup>22</sup>

### **Mendelian randomization analyses**

We performed MR analyses exploring (i) the effects of circulating cytokine levels on other cytokines, (ii) the effects of circulating cytokine levels on allergic and autoimmune, cardiometabolic, and cancer disease endpoints, (iii) and, depending on the outcomes of the 2<sup>nd</sup> MR analyses the effects of gene transcripts (eQTL) upstream of promising cytokines on allergic and autoimmune endpoints. We used *cis*-acting variants as genetic instruments for our MR analyses, as they are associated with a lower risk of pleiotropic effects when compared to *trans*-acting variants.<sup>97</sup> We filtered the GWAS meta-analysis results for variants within 300 kb around the gene encoding the respective cytokine. We selected variants associated at  $p < 5 \times 10^{-5}$  and clumped the data at  $r^2 < 0.1$ . We applied fixed-effects inverse variance-weighted MR analysis as our main analytical approach.<sup>94</sup> Again, MR Egger regression and the weighted median estimator were used as sensitivity analyses.<sup>95,96</sup> After harmonization of the effect alleles across cytokines we used `mr` command from the TwoSampleMR R Package (TwoSampleMR version 0.5.6) to extract the respective effect estimates.

### **Disease outcome GWASs**

For the disease endpoints, we downloaded summary level data from the largest publicly available GWAS and performed MR analyses for 3 independent disease groups. For allergic and autoimmune phenotypes we analyzed asthma (121,940 cases, 1,254,131 controls)<sup>98</sup>, Crohn's disease (5,956 cases, 14,927 controls)<sup>99</sup>, ulcerative colitis (6,968 cases, 20,464 controls)<sup>99</sup>, multiple sclerosis (47,429 cases, 68,374 controls)<sup>100</sup>, psoriasis (4,815 cases, 415,646 controls)<sup>101</sup>, and rheumatoid arthritis (14,361 cases, 43,923 controls)<sup>102</sup>. For cardiometabolic phenotypes we analyzed peripheral vascular disease (31,307 cases, 211,753 controls)<sup>103</sup>, coronary artery disease (60,801 cases, 123,504 controls)<sup>104</sup>, large artery stroke (9,219 cases, 1,503,898 controls)<sup>105</sup> and diabetes mellitus type II (242,283 cases, 1,569,730 controls)<sup>83</sup>. For cancer phenotypes we analyzed breast cancer (133,384 cases, 113,789 controls)<sup>106</sup>, colorectal cancer (5,657 cases, 372,016 controls)<sup>107</sup>, lung cancer (29,266 cases, 56,450 controls)<sup>108</sup>, non-Hodgins lymphoma (2,400 cases, 410,350 controls)<sup>109</sup>, and skin cancer (23,694 cases, 372,016 controls)<sup>107</sup>. The data sources are detailed in **Supplementary Table S15**.

### **Colocalization analysis**

To analyze shared causal variants between SNPs for circulating cytokines and disease outcomes showing significant associations in MR analyses, we used the “coloc” v3 R package. COLOC is a variant colocalization method that performs tests on shared causal variants in the locus. Colocalization methods consider the GWAS and disease outcome summary statistics at a locus jointly and probabilistically test if the two signals are likely to be generated by the same causal variant.<sup>110</sup> We used the meta-analyses summary statistics for the significant cytokines restricted to a flanking region  $\pm 300\text{Kb}$  around the genetic location of each cytokine and mapped disease-associated variants by their rsID.

### **Database search**

To assess previously reported associations a database search was conducted using the NHGRI-GWAS catalogue<sup>111</sup> on February 15<sup>th</sup>, 2023. We analyzed our GWAS hits for associations with any of the 40 cytokines reported here (**supplementary Table S4**), restricting the results for European-ancestry associations.

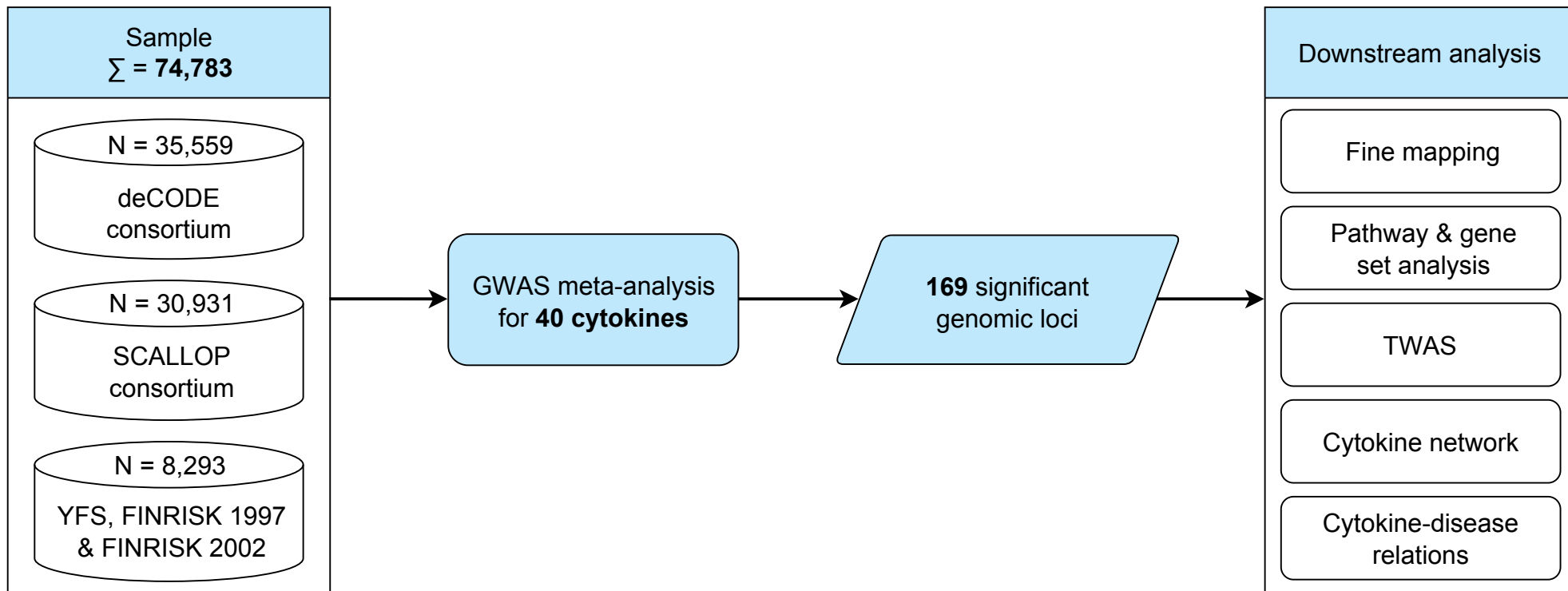
### **Data availability**

The data sources used in the current study are publicly available (download links for the summary statistics are available in **supplementary Table S14**). Ethical approval was not required due to usage of publicly available summary-level data.<sup>17,77,78</sup> GWAS meta-analysis summary statistics will be available at GWAS catalog.

### **Code availability**

Codes used to generate the results and figures are available on request from the corresponding author.

**Figure 1**



**Figure 2**

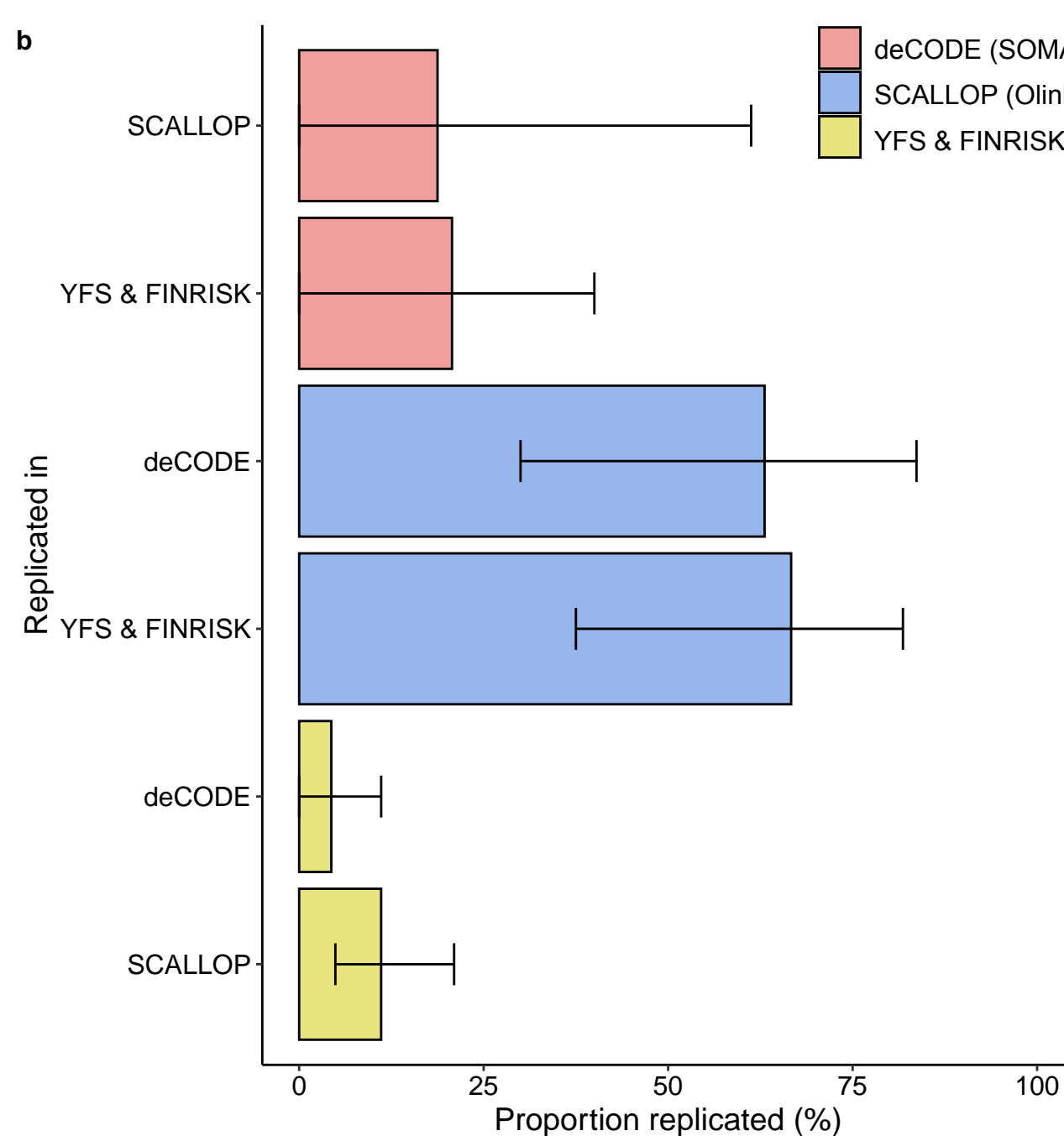
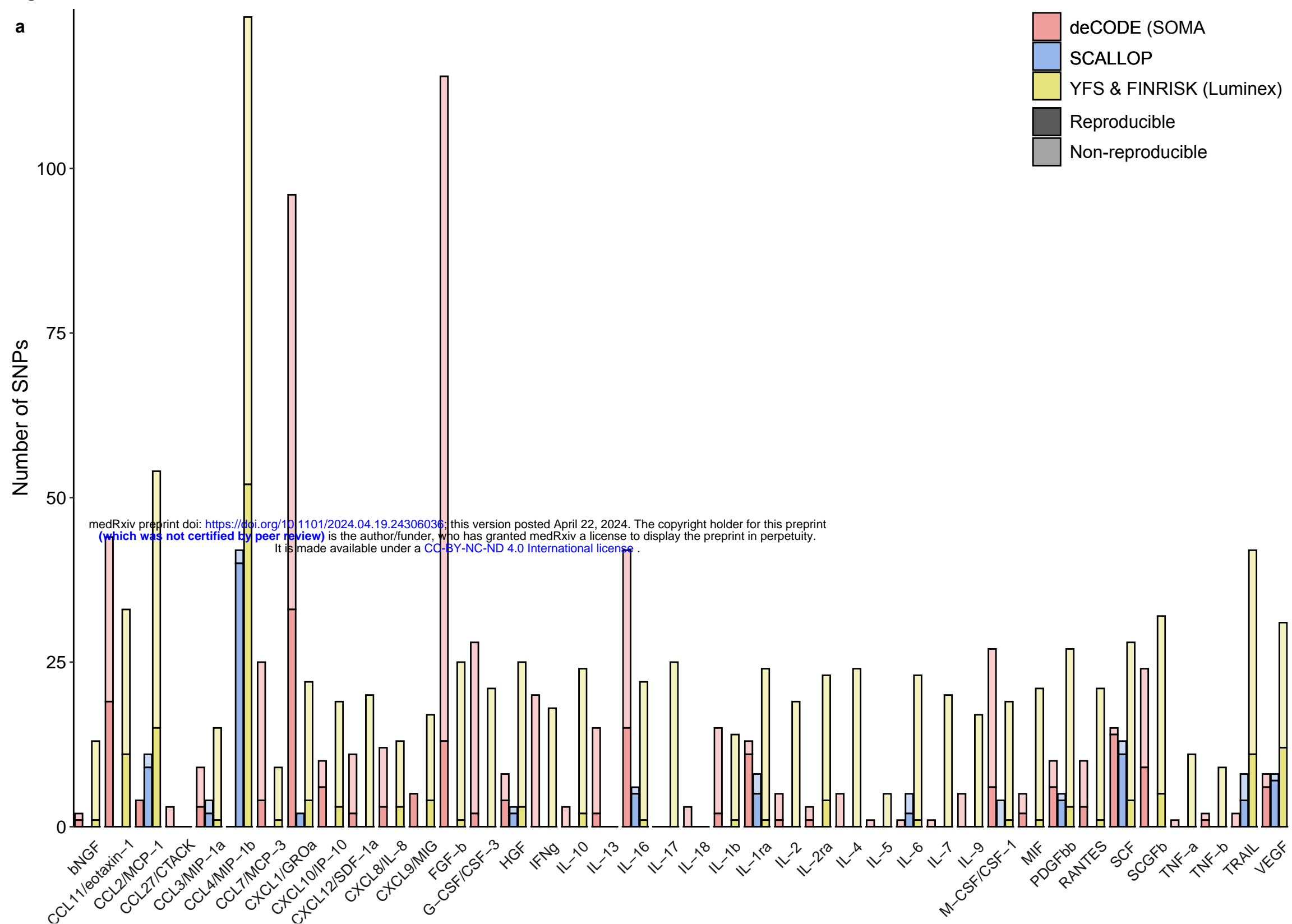
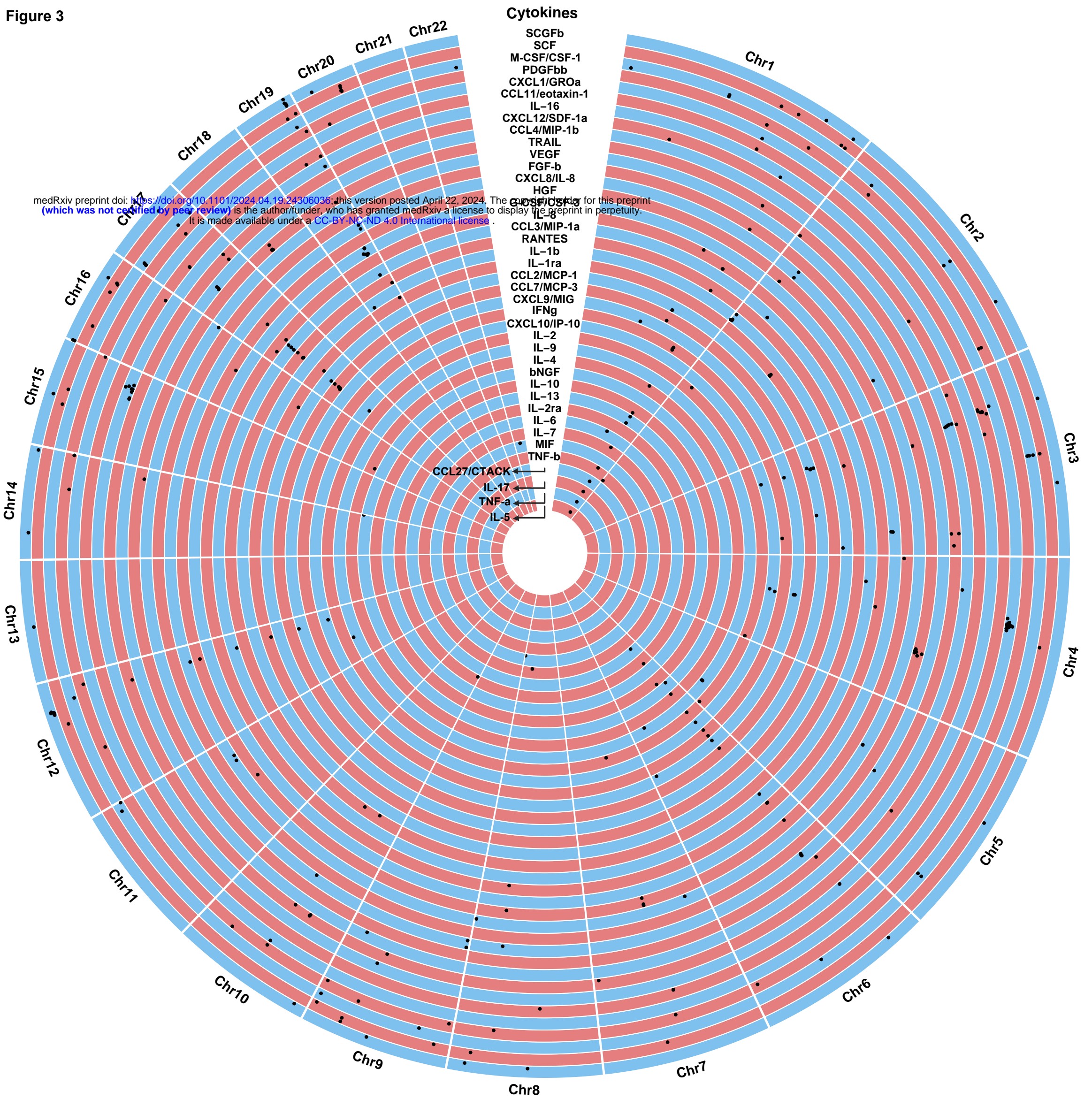




Figure 3



**Figure 4**

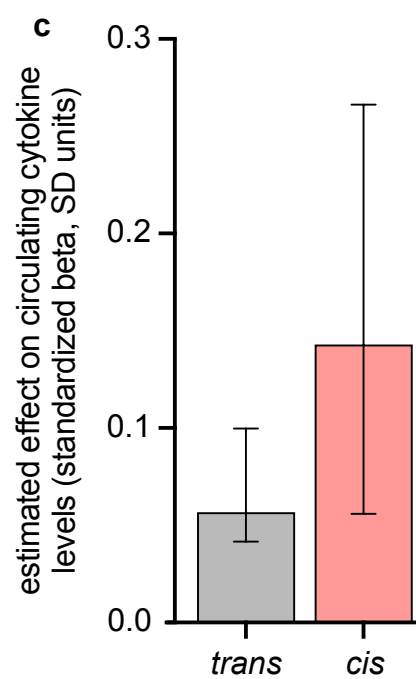
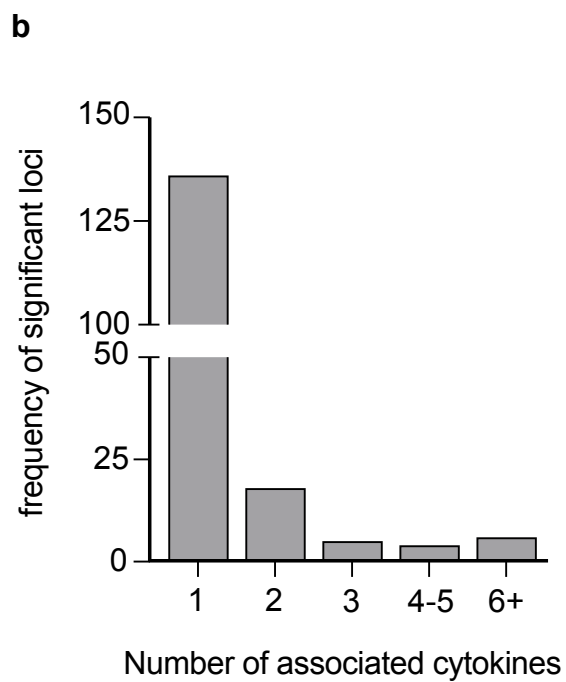
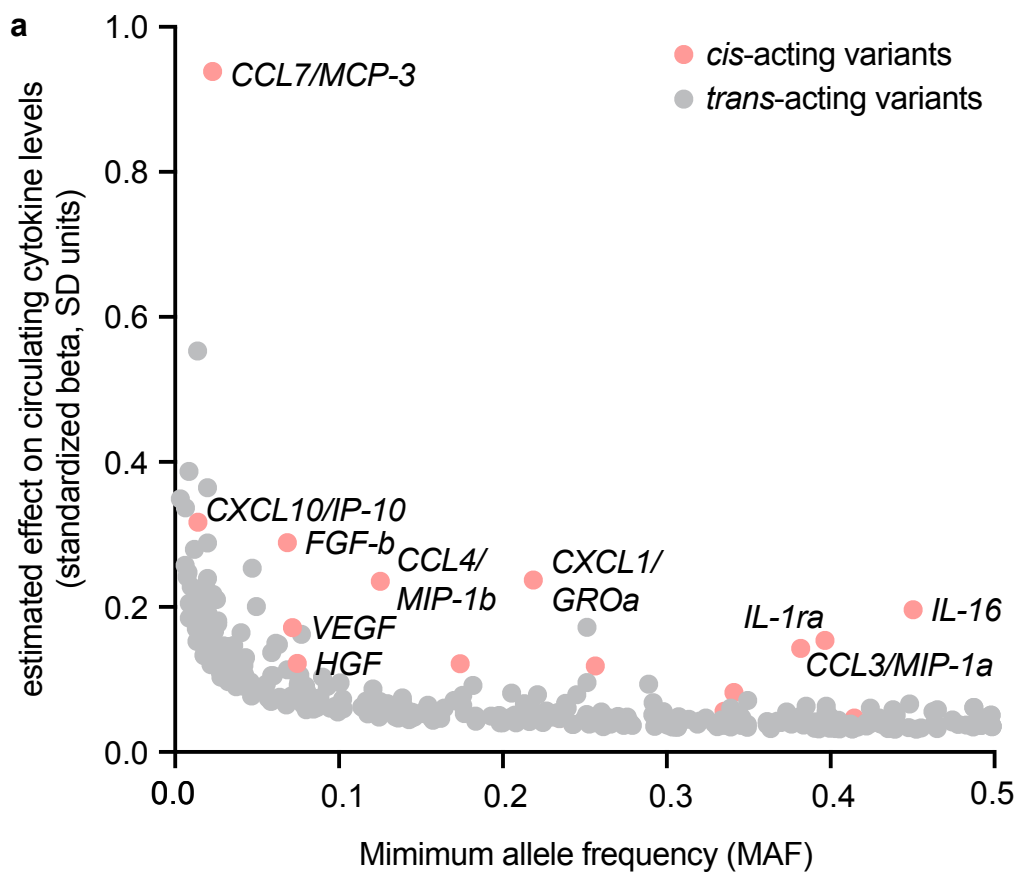


Figure 5

$-\log_{10}(\text{P-value})$

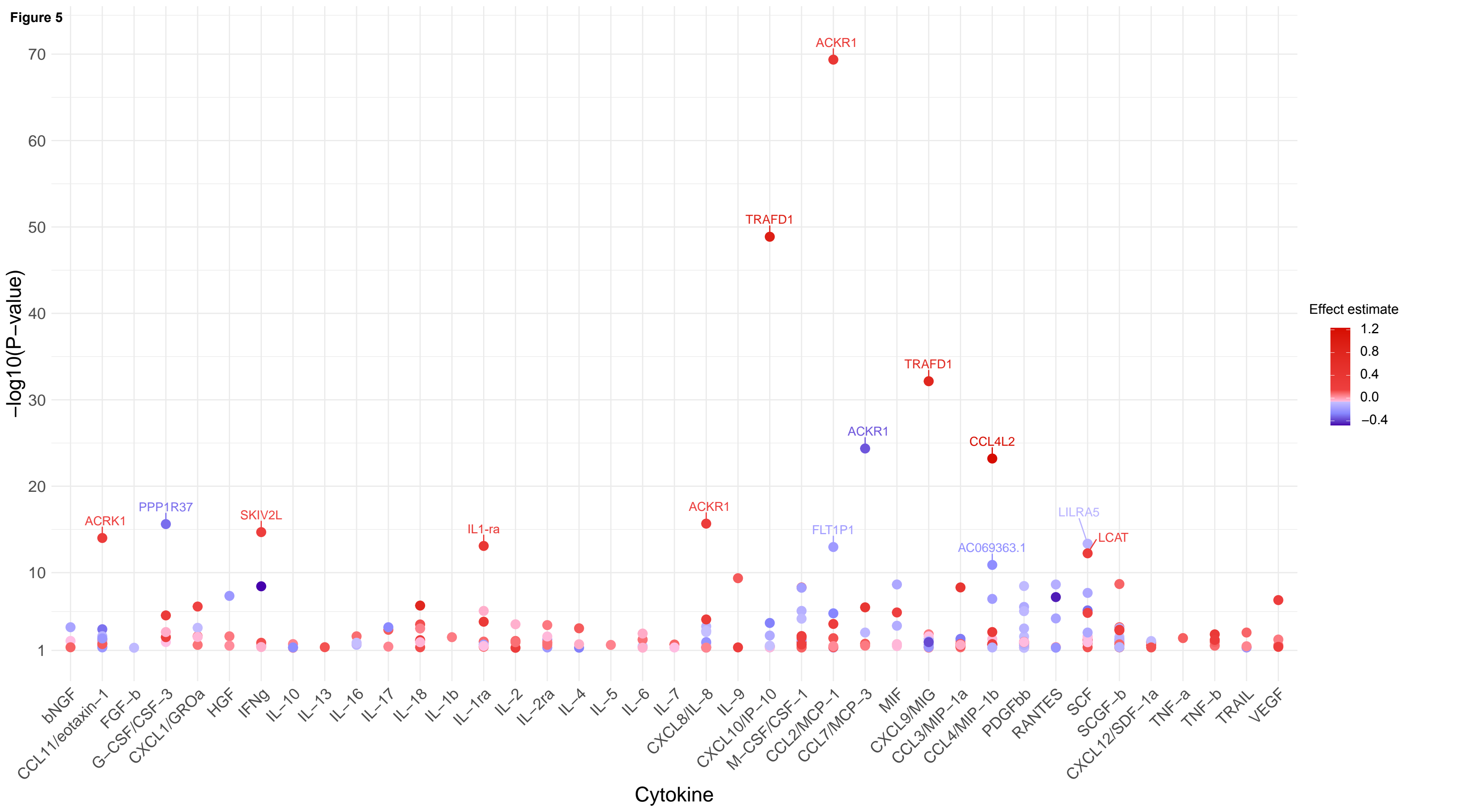
Effect estimate



70  
60  
50  
40  
30  
20  
10  
1

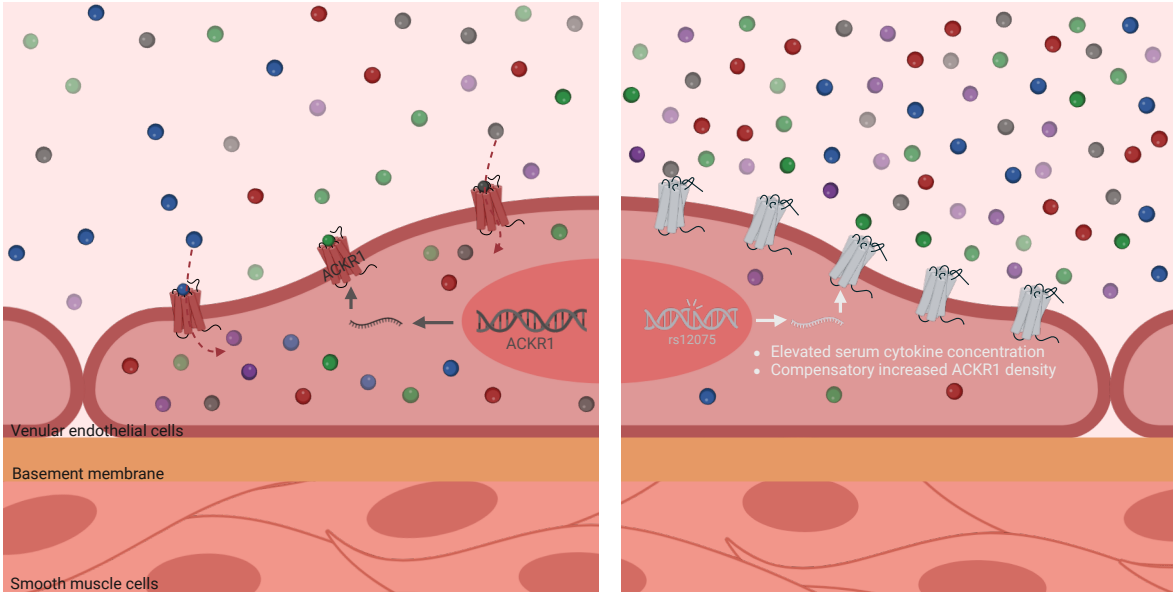
bNGF  
CCL11/eotaxin-1  
FGF-b  
G-CSF/CSF-3  
CXCL1/GROa  
HGF  
IFNg  
IL-10  
IL-13  
IL-16  
IL-17  
IL-18  
IL-1b  
IL-1ra  
IL-2  
IL-2ra  
IL-4  
IL-5  
IL-6  
IL-7  
CXCL8/IL-8  
IL-9  
CXCL10/IP-10  
M-CSF/CSF-1  
CCL2/MCP-1  
CCL7/MCP-3  
MIF  
CXCL9/MIG  
CCL3/MIP-1a  
CCL4/MIP-1b  
PDGFbb  
RANTES  
SCF  
SCGF-b  
CXCL12/SDF-1a  
TNF-a  
TNF-b  
TRAIL  
VEGF

Cytokine



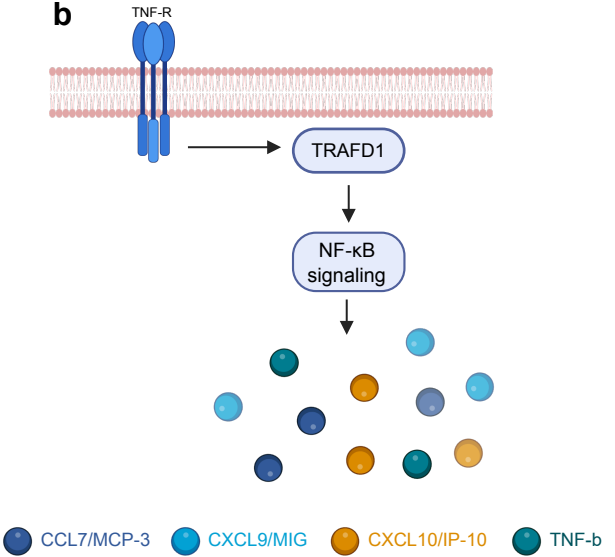
**Figure 6**

**a**



- CCL2/MCP-1
- CCL7/MCP-3
- CCL11/eotaxin-1
- CXCL1/GRO $\alpha$
- CXCL8/IL-8

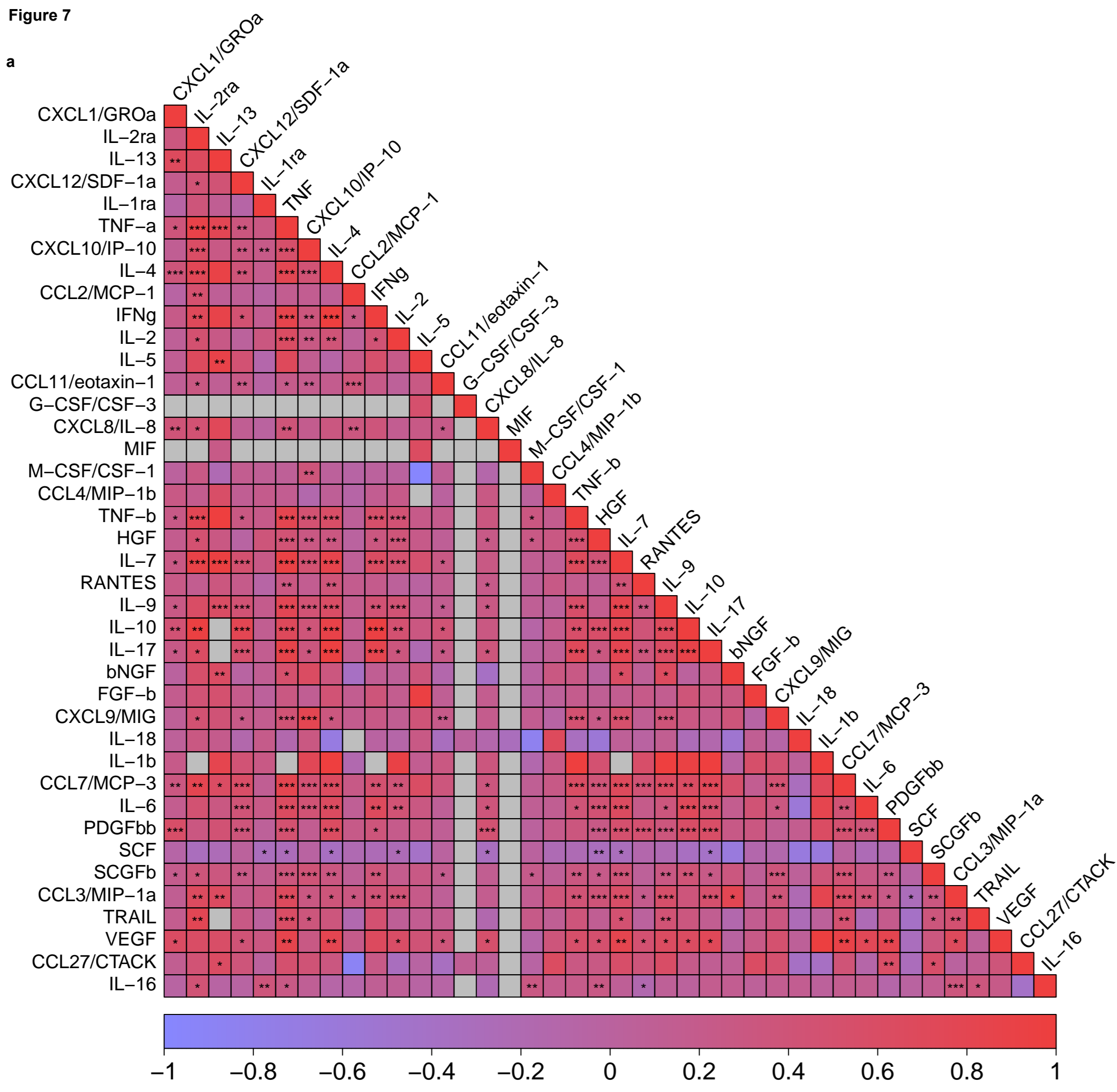
**b**



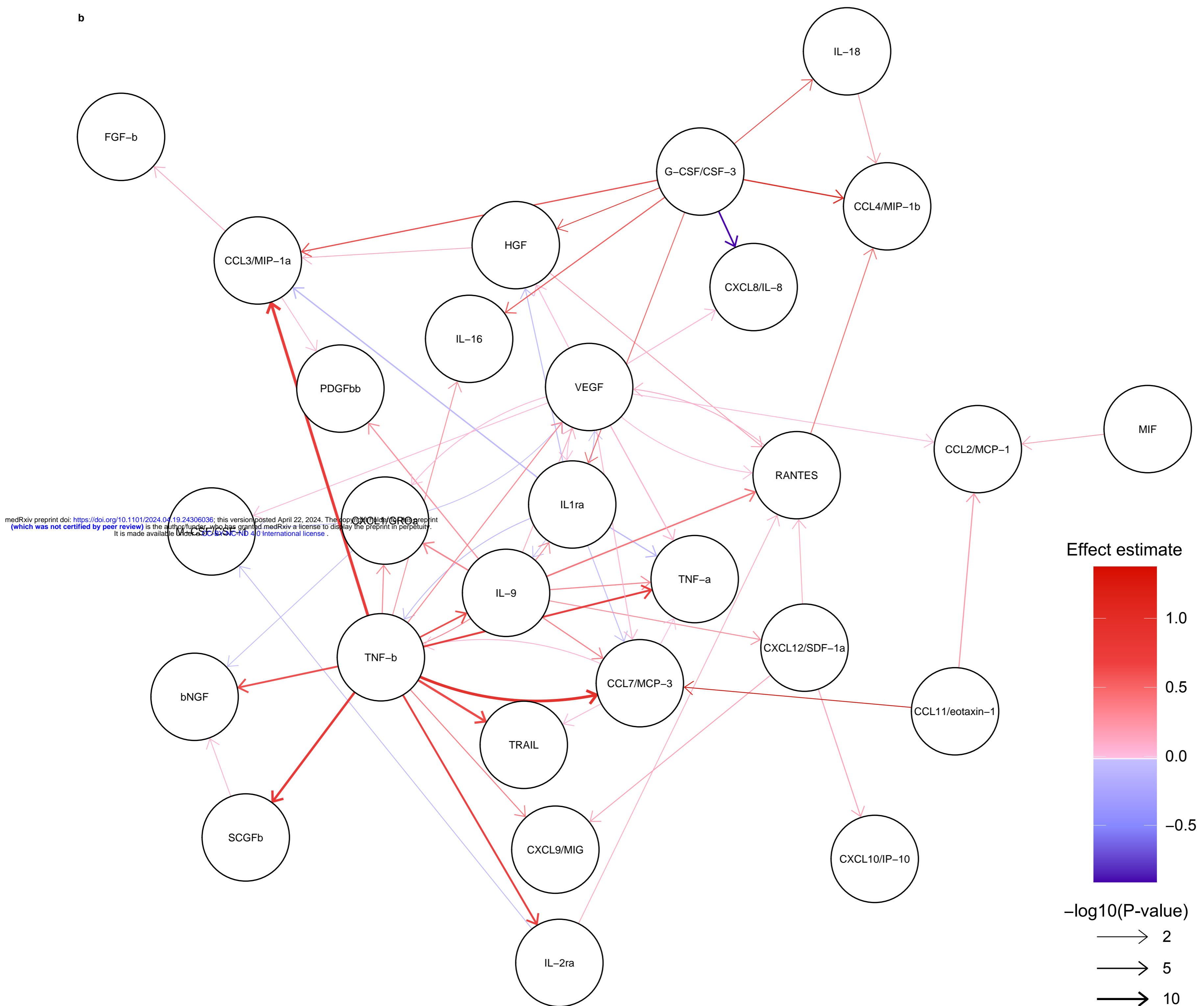
- CCL7/MCP-3
- CXCL9/MIG
- CXCL10/IP-10
- TNF-b

Figure 7

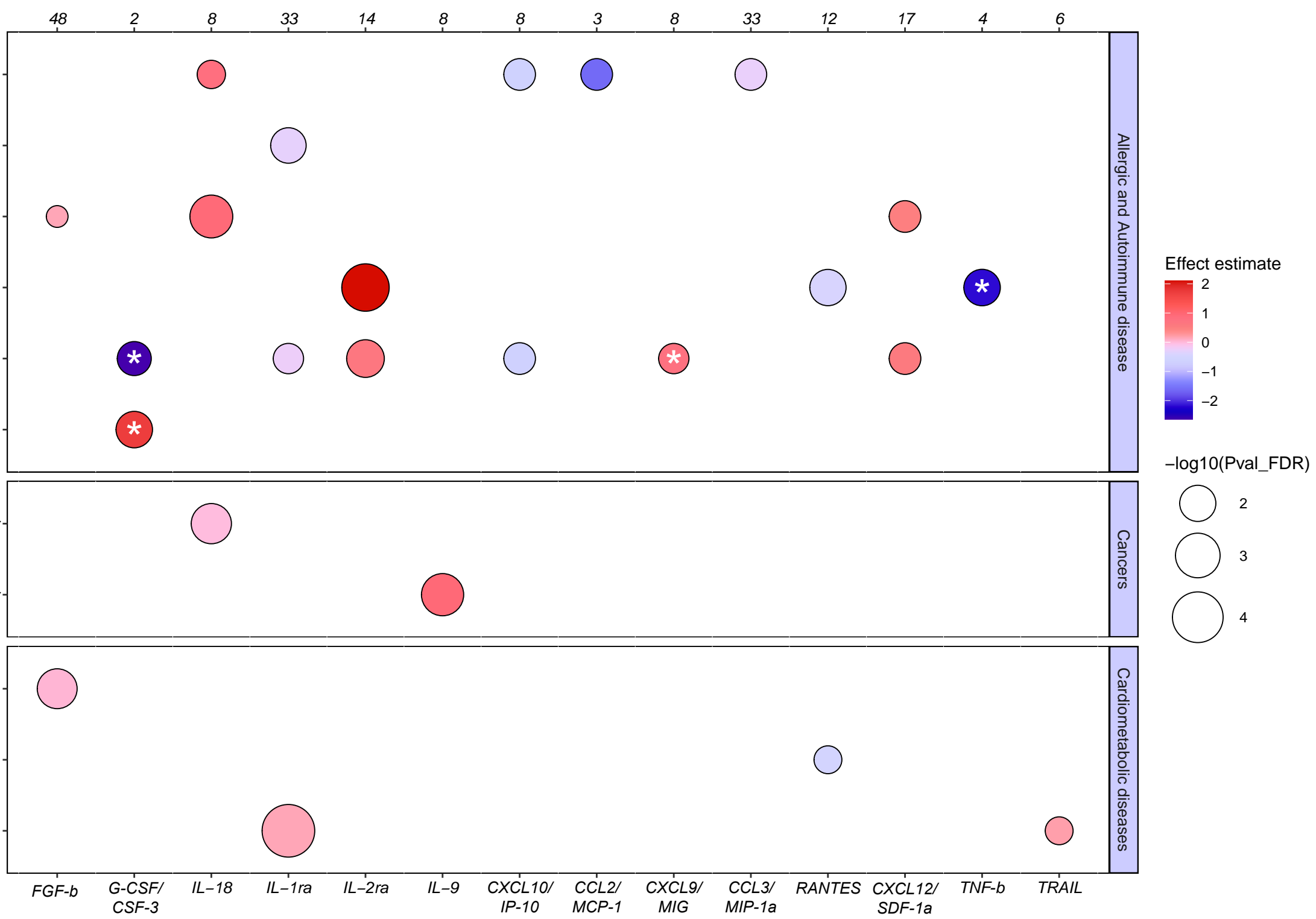
a



b



**Figure 8**



**Figure 9**

