



UNIVERSITÉ
DE NAMUR

Institutional Repository - Research Portal Dépôt Institutionnel - Portail de la Recherche

researchportal.unamur.be

RESEARCH OUTPUTS / RÉSULTATS DE RECHERCHE

The identification of blood-derived response eQTLs reveals complex effects of regulatory variants on inflammatory and infectious disease risk

Dubussy, Christophe; Liefferinckx, Claire; Stern, David; Perée, Hélène; Bottieau, Jérémie; Mayer, Alice; Quertinmont, Eric; Tafciu, Viola; Minsart, Charlotte; Petrov, Vyacheslav; Kvasz, Alex ; Coppieters, Wouter; Karim, Latifa; Rahmouni, Souad; Georges, Michel; Franchimont, Denis

Published in:
PLoS Genetics

DOI:
[10.1371/journal.pgen.1011599](https://doi.org/10.1371/journal.pgen.1011599)

Publication date:
2025

Document Version
Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (HARVARD):

Dubussy, C, Liefferinckx, C, Stern, D, Perée, H, Bottieau, J, Mayer, A, Quertinmont, E, Tafciu, V, Minsart, C, Petrov, V, Kvasz, A, Coppieters, W, Karim, L, Rahmouni, S, Georges, M & Franchimont, D 2025, 'The identification of blood-derived response eQTLs reveals complex effects of regulatory variants on inflammatory and infectious disease risk', *PLoS Genetics*, vol. 21, no. 4, e1011599, pp. 1-22.
<https://doi.org/10.1371/journal.pgen.1011599>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

RESEARCH ARTICLE

The identification of blood-derived response eQTLs reveals complex effects of regulatory variants on inflammatory and infectious disease risk

Claire Liefverinckx^{1,2*}, David Stern³, H el ene Per ee⁴, J er emie Bottieau¹, Alice Mayer³, Christophe Dubussy⁴, Eric Quertinmont², Vjola Tafciu², Charlotte Minsart^{1,2}, Vyacheslav Petrov⁴, Alex Kvasz⁵, Wouter Coppieters⁶, Latifa Karim⁶, Souad Rahmouni⁴, Michel Georges^{4,7}, Denis Franchimont^{1,2}

1 Center for the study of IBD, Laboratory of Experimental Gastroenterology, Universit e libre de Bruxelles, Brussels, Belgium, **2** Department of Gastroenterology, Hepatopancreatology, and Digestive Oncology, HUB H opital Erasme, Universit e Libre de Bruxelles, Brussels, Belgium, **3** GIGA Bioinformatics Platform, GIGA Institute, University of Li ge, Li ge, Belgium, **4** Unit of Animal Genomics, GIGA Institute, University of Li ge, Li ge, Belgium, **5** Software development, University of Li ge, Li ge, Belgium, **6** GIGA Genomics Platform, GIGA Institute, University of Li ge, Li ge, Belgium, **7** WEL Research Institute & Faculty of Veterinary Medicine, Li ge, Belgium

* claire.liefferinckx@hubruxelles.be



OPEN ACCESS

Citation: Liefverinckx C, Stern D, Per ee H, Bottieau J, Mayer A, Dubussy C, et al. (2025) The identification of blood-derived response eQTLs reveals complex effects of regulatory variants on inflammatory and infectious disease risk. *PLoS Genet* 21(4): e1011599. <https://doi.org/10.1371/journal.pgen.1011599>

Editor: Lin Chen, The University of Chicago, UNITED STATES OF AMERICA

Received: September 19, 2024

Accepted: January 29, 2025

Published: April 10, 2025

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pgen.1011599>

Copyright:   2025 Liefverinckx et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution,

Abstract

Hundreds of risk loci for immune mediated inflammatory and infectious diseases have been identified by genome-wide association studies (GWAS). Yet, what causal variants and genes in risk loci underpin the observed associations remains poorly understood for most. The identification of colocalized cis-expression Quantitative Trait Loci (cis-eQTLs) is a promising way to identify candidate causative genes. The catalogue of cis-eQTLs of the immune system is likely incomplete as many cis-eQTLs may be context-specific. We built a large cohort of 406 healthy individuals and expanded the immune cis-regulome through their whole blood transcriptome obtained after stimulation with specific toll-like receptor (TLR) agonists and T-cell receptor (TCR) antagonist. We report three mechanisms that may explain why an eQTL could only be revealed after immune stimulation. More than half of the cis-eQTLs detected in this study would have been overlooked without specific immune stimulations. We then mined this new catalogue of response (r) eQTLs, with public GWAS summary statistics of three diseases through a colocalization approach: inflammatory bowel diseases, rheumatoid arthritis and COVID-19 disease. We identified reQTL-specific colocalizations for risk loci for which no matching eQTL were reported before, revealing interesting new candidate causal genes.

Author summary

Although many risk loci have been identified by GWAS for immune and infectious diseases, the causal variants and genes that underpin the observed associations remain unknown for most. The identification of colocalized cis-expression

and reproduction in any medium, provided the original author and source are credited.

Data availability statement: Data can be found at <https://ega-archive.org/datasets/EGAD50000001320> Additionally, we provide an interactive web browser to facilitate user-friendly exploration of our dataset at <https://tools.giga.uliege.be/cedar/public/w>.

Funding: CL received supporting grant from the Belgian Inflammatory Bowel Disease Research and Development group (<https://www.birdgroup.be/en>) CL received salary from Fonds National de la Recherche scientifique (<https://www.frs-fnrs.be/en/>) CL received salary and supporting grant from Fonds Erasme (<https://fondserasme.org/fr/chercheurs>) DF received supporting grants from Fonds National de la Recherche scientifique (<https://www.frs-fnrs.be/en/>) The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Quantitative Trait Loci (cis-eQTLs) is a recognized way to identify candidate causative genes. However, the catalogue of cis-eQTLs of the immune system is likely incomplete as many cis-eQTLs may be context-specific.

We built a large cohort of 406 healthy individuals and expanded the immune cis-regulome by analyzing the whole blood transcriptome after specific immune stimulations.

We observed that more than half of the cis-eQTLs revealed in our study would have been overlooked without specific immune stimulations. After characterization of this expanded cis-eQTL catalogue, we mined it with publicly GWAS summary statistics of three diseases using a colocalization approach. The specific exploration of reQTLs expanded the number of risk loci with matching eQTL and identified new candidate genes for inflammatory bowel diseases, rheumatoid arthritis and COVID-19. This work highlights the importance of using meticulously selected healthy cohorts at population based-level to improve our knowledge on the causal drivers of disease.

Introduction

Genome-wide association studies (GWAS) have uncovered numerous risk loci for all studied common complex diseases [1]. Most of the underlying causative variants appear to be regulatory as opposed to coding, complicating the identification of the cognate causative genes [2,3]. Accordingly, identifying colocalized or matching cis-expression Quantitative Trait Loci (cis-eQTLs) in disease relevant tissues has proven an effective route to identify candidate genes in risk loci [4]. Cis-eQTLs are said to be matching if they exhibit SNP association patterns that are very similar to the SNP association patterns for the disease [5]. Previous studies have shown that matching cis-eQTL can be found for 25% of risk loci with the available eQTL data sets [6,7]. This raises the question as to the molecular drivers of the remaining risk loci. One possible explanation is that the matching cis-eQTL only manifest in specific, as of yet uninterrogated cell types and/or developmental stages. Another is that some of the matching cis-eQTL only manifest after onset of the disease process, such as in inflammatory conditions. Of note, most existing eQTL datasets were generated using heterogenous cohorts of healthy individuals.

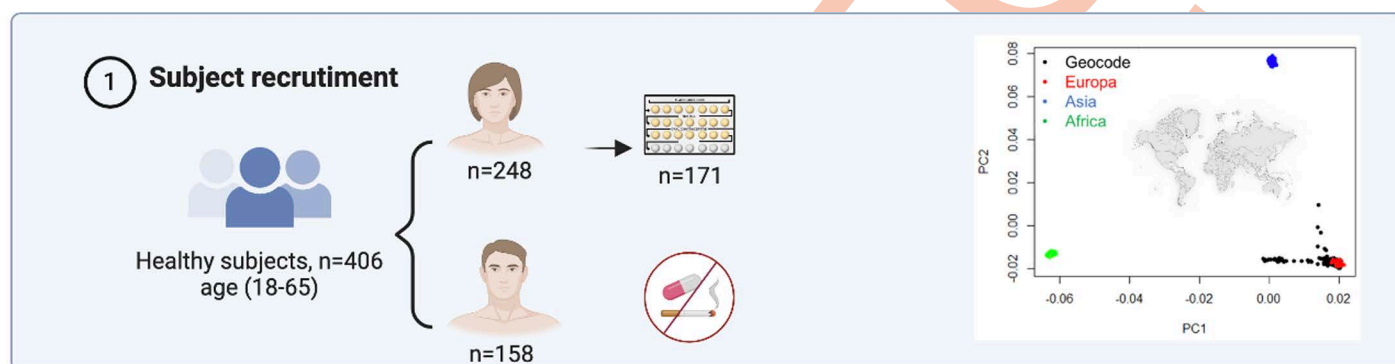
To address the latter limitation, and in order to expand the catalogue of human cis-eQTLs that might underpin immune or infectious diseases, we studied the whole blood transcriptome of a highly selected cohort of 406 healthy individuals of white Europeans ancestry, both in resting state as well as after stimulation with specific toll-like receptor (TLR) agonists (TLR4 activated through lipopolysaccharide (LPS) and TLR7/8 activated through resiquimod (R848)) and T-cell receptor (TCR) antagonist (using anti-CD3 and anti-CD28 antibodies). The underlying assumption was that the corresponding perturbations might mimic inflammatory conditions and elicit novel

eQTLs (also known as “response eQTLs” or reQTLs) of which some may contribute to inter-individual variation in predisposition to immune mediated inflammatory (IMIDs) or infectious diseases [8–18].

Results

Leveraging both genome-wide association study and transcriptomics identifies genetic determinants of TCR stimulation

The GEOCODE cohort consists of 406 individuals, with details and methodologies outlined in Fig 1, including inclusion criteria and sample handling (see Methods). The inter-individual variability in circulating immune cell composition and baseline characteristics has been previously described [19]. We validated the effectiveness of the stimulations by assessing the levels of the corresponding cytokines (Table A in S1 Table). We applied an in-silico cell deconvolution technique [20] to assess the impact of immune stimulation on altering immune cell proportions. Interestingly, we observed that the proportions of immune cells remained unchanged after 24 hours of in vitro stimulation (S1 Data).



Methods

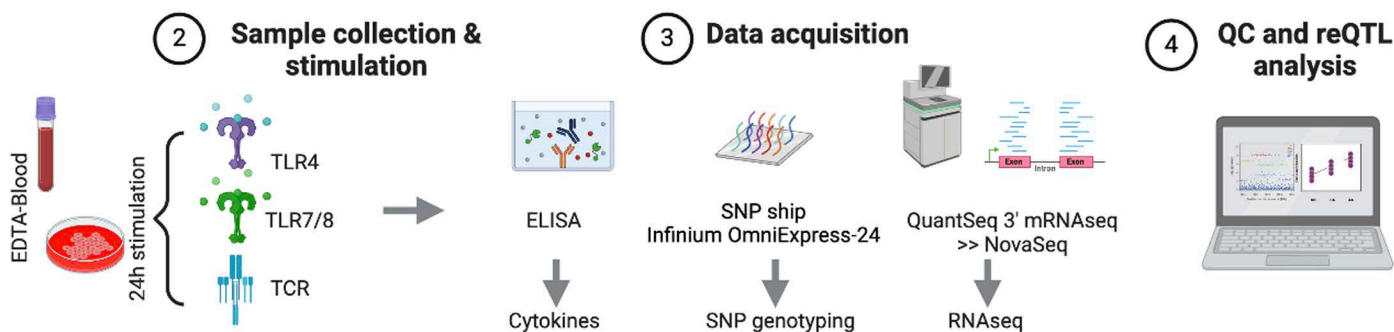


Fig 1. GEOCODE study design. **1. Subject Recruitment:** The GEOCODE cohort consisted of highly selected healthy individuals with no history of disease, medication use, nor smoking habits. Among the 248 female participants recruited, 69% reported contraceptive pill use. The plot displays PCA mapping from the HapMap consortium for ancestry comparison alongside PCA mapping of the GEOCODE cohort. **2. Sample collection and stimulation:** Whole blood samples were collected, and cell cultures and stimulations were performed within three hours of collection. Three stimulation conditions were applied during a 24-hour incubation period: Resiquimod (TLR7/8 agonist), Lipopolysaccharide (LPS) (TLR4 agonist) Anti-CD3/anti-CD28 (TCR stimulation). **3. Data acquisition:** Cytokine production following TLR and TCR stimulation was quantified using a standard ELISA method, with all measurements expressed in pg/ml. Genotyping was performed using Illumina’s Human OmniExpress BeadChips, and RNA sequencing was carried out with the QuantSeq 3’ mRNA-Seq Library Prep Kit FWD at the GIGA-Genomics platform, resulting in a total of 1305 samples. **4. QC and reQTL analysis:** Cis-eQTL analysis was performed using QTLtools with 10,000 permutations to compute adjusted p-values. The resulting cis-eQTLs were processed following the guidelines outlined in reference 6. Images used in the figure have been generated with a Biorender Academic License.

<https://doi.org/10.1371/journal.pgen.1011599.g001>

All participants were genotyped with Illumina's OmniExpress SNP array and genotype data imputed to whole genome using TOPMed. Transcriptome analysis was conducted by RNA-Seq (Fig 1 and Methods). Gene expression profiles segregated samples in multidimensional expression space by conditions (Fig 2A-B). In general, principal component (PC) 1 separated stimulated samples (TLR7/8-TLR4-TCR) from controls, while PC2 separated TCR from TLR stimulated samples. However, we observed a continuum of individuals between TCR stimulation and control groups (Fig 2A). This is also evident in Fig 2B, where a subset of TCR stimulated samples (in purple) appears similar to the control group. Using a threshold correlation of $|0.5|$, PC1 was found to be influenced by 398 genes, with 61 showing positive and 337 showing negative correlation (Table B in S1 Table). Positively expressed genes highlighted Reactome pathways related to "Neutrophil degranulation" (R-HSA-6798695) and "Chemokine receptors bind chemokines" (R-HSA-380108) among the top significant pathways, while negatively expressed genes highlighted "Interleukin-10 signaling" (R-HSA-6783783) and "Interferon alpha/beta signaling" (R-HSA-909733). The full list of pathways associated with these genes is presented in Table C in S1 Table.

The scale represents Euclidean distance, where shorter distances indicate higher correlations in gene expression between subjects. The Ctrl group (red) clearly segregates from the stimulated conditions. Among the stimulation conditions, the TLR-stimulated groups—LPS stimulation (TLR4, green) and R848 stimulation (TLR7/8, blue)—exhibit similar correlation patterns and are closer to each other compared to the TCR-stimulated group (anti-CD3/anti-CD28, purple), which forms a distinct cluster.

To further dissect the genetic architecture underlying this continuum, PCs were then calculated for each paired condition (control-stimulated condition). By considering individual PC values as new phenotypes in GWAS, rs1801274 on chromosome 1 was found to be significantly associated ($p = 4e-45$) with PC1 for TCR stimulation but not with other PCs (S1 Fig). As expected, 313 of the 398 genes driving PC1 obtained with the 4 groups were showing trans-eQTL effects for rs1801274 (Table B in S1 Table).

Rs1801274, is a missense mutation (G>A) (G wild type and A variant) in the FCGR2A gene on chromosome 1, encoding FcγRIIa receptor, a low-affinity receptor for the constant fragment (Fc) of immunoglobulin G (IgG) [21,22]. The A variant, common in Europeans (MAF G 0.49), results in an arginine (R) to histidine (H) substitution at position 131 or

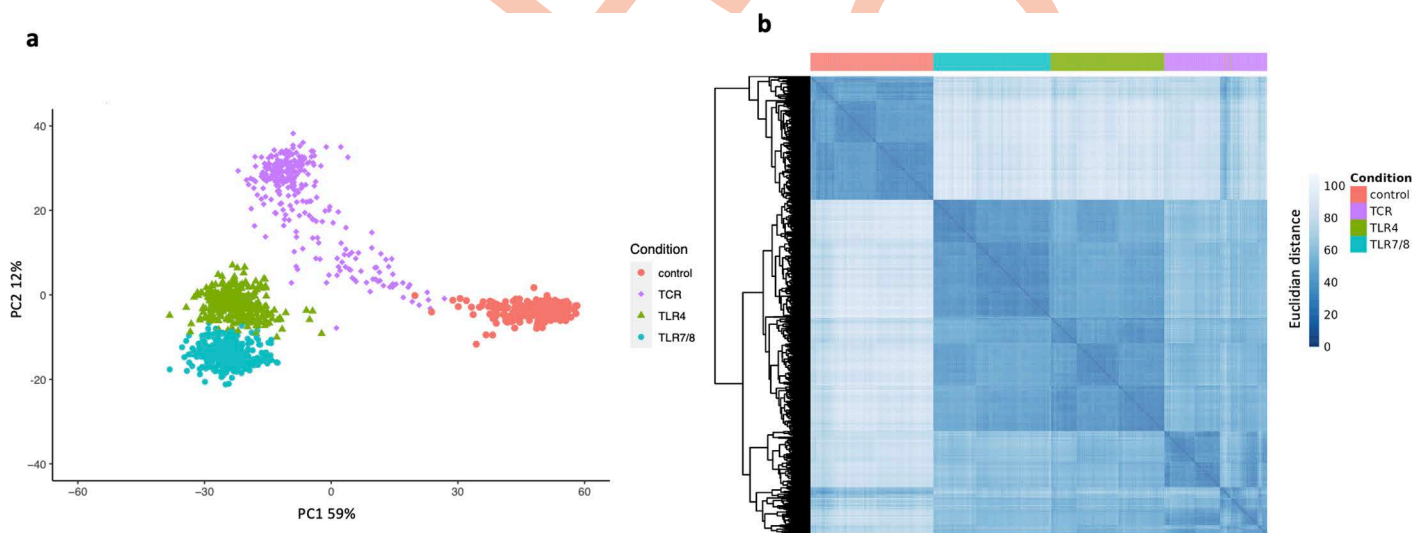


Fig 2. Transcriptomic analysis. A. Principal component analysis (PCA) of the expression of the most variable 500 genes across the four conditions of stimulation (Control (Ctrl), TLR4, TLR7/8 and TCR stimulations). **B. Heatmap depicting the correlation patterns of the 500 most variable genes across subjects after 24 hours of stimulation, as assessed by RNA-seq.**

<https://doi.org/10.1371/journal.pgen.1011599.g002>

166 (depending on the numbering in the literature) in the extracellular domain of the receptor protein, and was reported to impact the binding affinity for the Fc region of different IgG subclasses, affecting the cytokine production, clearance of immune complexes and phagocytosis of opsonized bacteria by granulocytes [23,24]. We confirmed that rs1801274 affect cytokine production and was strongly associated with IFN γ ($p=7.72e-40$) and IL-2 levels ($p=2.63e-15$) following anti-CD3/anti-CD28 whole blood stimulation as reported by others [25,26] (S2 Fig). This suggests that part of the TCR stimulation in whole blood cell cultures might be mediated by the engagement of the Fc γ RIIA, expressed monocytes/macrophages, neutrophils and dendritic cells, by circulating IgG2 -containing immune complexes or released IgG immunoglobulins (including the co-stimulatory signal of the soluble anti-CD28 antibody). This underscores the role of monocytes/macrophages and granulocytes in the elicitation and perpetuation of the TCR stimulation thanks to the ex vivo model/assay of whole blood cell cultures. Interestingly, whole blood cell cultures specifically depleted in neutrophils demonstrated a reduction in IFN γ release, indicating that not only monocytes/macrophages but also neutrophils play a critical role in the cytokine release cascade through Fc:Fc γ R interactions [27].

Expanding the immune cis-eQTL catalogue and clustering cis-eQTLs into cis-regulatory modules

We performed cis-eQTL analyses using QTL-tools, and gene expression levels pre-corrected for defined (sex, age, BMI) and hidden confounders (29–38 expression principal components). We identified a total of 13,679 autosomal cis-eQTL (FDR 0.05) corresponding to 6,496 eQTL genes (eGenes) (Tables D and E in S1 Table).

Cis-eQTL affecting the same gene in more than one condition were merged in a “cis regulatory module” (cRM) when sharing near-identical “association patterns”, following ref. 6 (Figs 3A and S3). This yielded 8,401 cRM operating on average in 1.63 of the four tested conditions. cRM operating either in one specific condition or in all four conditions were the most common, accounting for 79.5% of the modules (Fig 3B). The predominance of modules operating in one condition only was unlikely to be due to limited cis-eQTL detection power, as allowing the merger of 12,950 non-significant cis-eQTL that would nevertheless match at least one significant cis-eQTL with 0.6 did not change the pattern (S4 Fig and Methods). It is noteworthy that 4,695 of cRM (55.8%) are not detected in resting condition, requiring either TLR or TCR stimulation to be detected.

We uncovered three mechanisms that may explain why a cRM may be active in condition “a” but not in condition “b”. The first is that the target gene is expressed at too low levels in condition “b” for the eQTL to be detectable. Several genes exemplified this scenario such as *IL36G* and *IL36RN* associated with psoriasis [28], and *CSF3* associated with severe neutropenia [29] (Figs 3C and S5). We estimate that this first scenario accounts for 2.6% of cases in our dataset (see Methods). The second mechanism is the loss of a genotype effect in condition “b”, although the gene remains expressed at high enough levels for the eQTL to be detected if it exists. We estimated that this second scenario accounts for 72.8% of cases (see Methods). Examples of this scenario include *NCF2* associated with chronic granulomatous diseases [30], and *LILRB2* (important immune mediator) [31] (Figs 3D and S6). Many eQTLs were specifically lost after “TCR” stimulation, and this was often associated with an increase in expression variance of the corresponding gene. One such example is the *CISD1* gene associated with inflammatory bowel diseases (IBD) [32] (Fig 3D and S6). The third mechanism occurs when a gene (eGene) switches cRM between conditions: the regulatory variants governing gene expression in the four conditions, although possibly overlapping, are distinct enough to generate non-matching EAPs (Fig 3E). We estimated that this last scenario accounts for 16.5% of cases. Accordingly, 1,578 eGenes were assigned to two or more cRM (Table E in S1 Table). These comprised a high proportion of genes (55.7%) that were subject to one eQTL in resting condition, and another eQTL after stimulation, with occasionally distinct stimulation-specific EAPs (Figs 3E and S7). When a gene was influenced by the same cRM in more than one condition, the signs of the SNP effects were generally consistent across conditions (either up- or down-regulation). There were only few, yet noteworthy exceptions for which the allelic effects switched sign between control and TLR and/or TCR stimulation, including *ADCY3*, *BMP8A* and *HIP1* (Fig 3F).

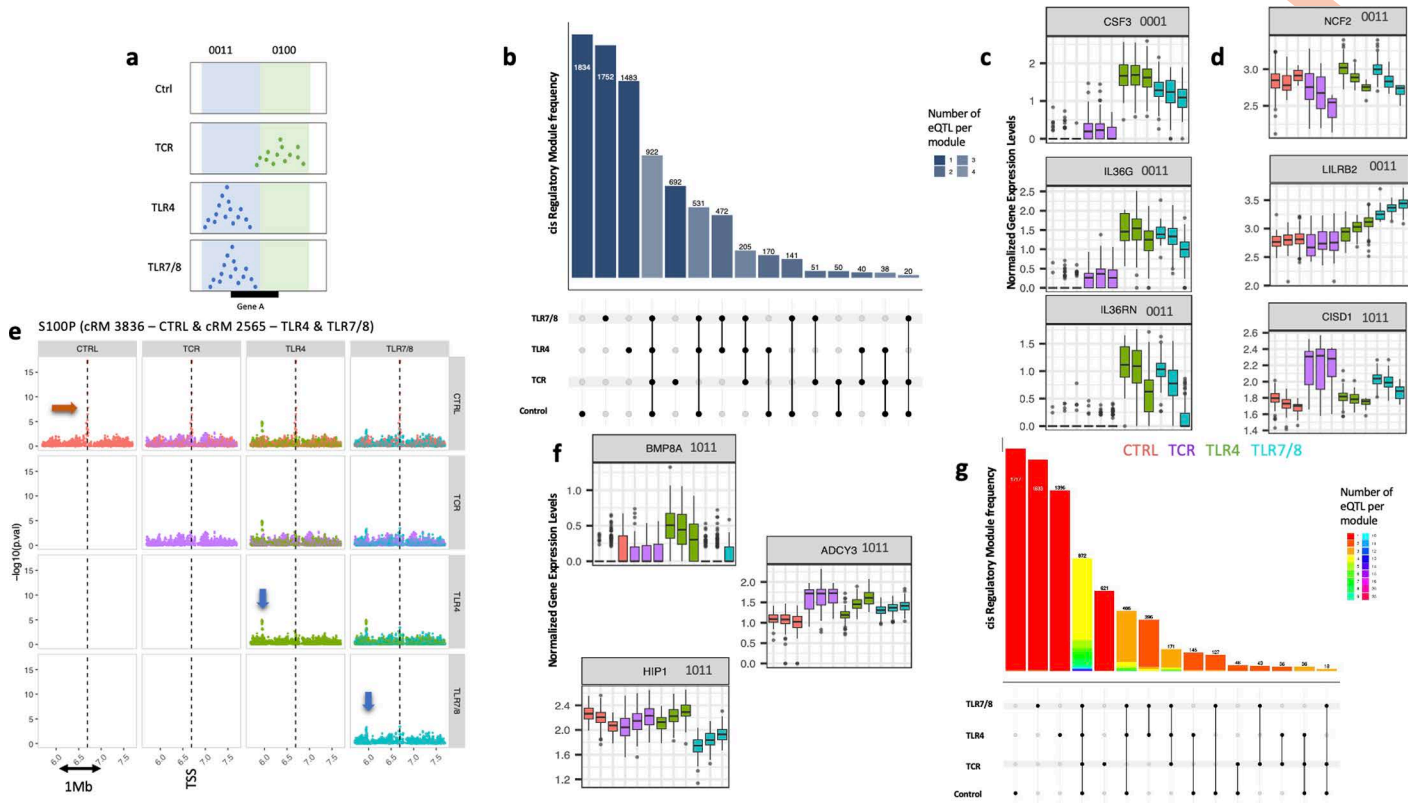


Fig 3. Exploration of Response (R) cis-eQTLs in stimulated whole blood cell cultures from healthy subjects (GEOCODE cohort). A. Schematic illustration of the eQTL colocalization methodology. The reQTL association patterns (EAPs) for TLR4 and TLR7/8 conditions are similar and are merged into the same cis regulatory module (cRM) associated with Gene A. This cRM is designated according to the following label 0011, indicating no eQTL in Control or TCR conditions, but identical reQTLs association patterns in TLR4 and TLR7/8 conditions. Gene A is associated with another eQTL whose EAP does not colocalize with the reQTLs for TLR4 and TLR7/8 conditions. This eQTL exists in its own cRM, labelled as 0100. B. Upset plot showing the distribution of Cis regulatory modules (cRMs) affecting the same gene under resting or TCR – TLR- conditions of stimulation. The plot highlights the predominance of condition-specific cRM with the majority being reQTLs. Since this graph only includes monogenic cRMs, the maximum number of eQTLs per gene is four, corresponding to the number of conditions in our dataset. C. First mechanism that may explain why an eQTL is active in one condition but not in another condition: For each eQTL, the x-axis represents the four conditions with the three alleles per genetic variant. The expression level of the target gene's is too low in Ctrl (resting) condition for the eQTL to be detectable, but it becomes detectable under stimulation conditions. The Boxplots illustrate gene expression levels, with the y-axis showing normalized gene expression in arbitrary units and x-axis showing different stimulation (Control (Ctrl) in red, TCR in purple, TLR4 in green, TLR7/8 in blue). The genes CFS3 (Colony Stimulating Factor 3), IL-36G (Interleukin 36 Gamma) and IL-36RN (Interleukin 36 Receptor Antagonist) are used as examples. D. Second mechanism that may explain why an eQTL is only active in some conditions despite the expression of the gene being measured in all four conditions: For each eQTL, the x-axis represents the four conditions with the three alleles per genetic variant. This scenario occurs when a genotype effect is not consistently detected across all conditions, even though the gene is significantly expressed. For example, the cRM for the genes NCF2 (Neutrophil Cytosolic Factor 2) and LILRB2 (leukocyte immunoglobulin like receptor B2) is labelled 0011, indicating eQTLs are detected only in TLR7/8 and TLR4 conditions, although the gene expression levels are sufficiently high in Ctrl and TCR conditions for potential eQTL detection. Likewise, the cRM for the gene CISD1 (CDGSH Iron Sulfur Domain 1) is labelled 1011, indicating eQTLs are detected in Ctrl, TLR4 and TLR7/8 conditions but not in TCR condition. E. Third mechanism that may explain why an eQTL is active in one condition but not in another condition: The target gene switches between different cRM under various conditions. The plots illustrate the EAPs for the target gene S100P (S100 Calcium Binding Protein P) across different stimulation conditions, color-coded as follow: red for Ctrl, purple for TCR, green for TLR4, and blue for TLR7/8 conditions. The y-axis shows the distribution of $-\log_{10}(p)$ values for variants in the region around the top cis-eQTL, while the x-axis represents the genomic region centred on the Transcriptional Start Site (TSS) of the gene. The eGene S100P is influenced by two different cRMs. The peak indicating the first cRM, 3886, is highlighted by the red arrow, and the peak indicating the second cRM, 2565, is highlighted by the blue arrow. In the latter case, the EAPs for the TLR4 and TLR7/8 conditions are similar and combined within a single cRM. F. Examples of cis-eQTLs where the direction of the allelic effect changes between the control and the conditions of stimulation. For each eQTL, the x-axis represents the four conditions with the three alleles per genetic variant. For the *BMP8A* gene (Bone Morphogenetic Protein 8a), the cRM is labelled 1011, indicating eQTLs detected in Ctrl, TLR4 and TLR7/8 conditions but not in the TCR condition. The eQTL effect is positive in Ctrl and TLR7/8 conditions but negative in the TLR4 condition. For the *ADCY3* (*Adenylate Cyclase 3*) and *HIP1* (*Huntingtin Interacting Protein 1*) genes, the cRM is also labelled

1011. The eQTL effect is negative in Ctrl but positive in TLR7/8 and TLR4 conditions. G. Upset plot showing the distribution of Cis regulatory modules (cRMs) across resting, TCR and TLR conditions of stimulation, considering both mono- and multi- genic cRMs. The plot highlights the predominance of condition-specific cRM with most of them being reQTLs. Since the graph includes both monogenic and multigenic cRMs, the number of eQTLs per cRM can be as high as 25. A cRM can result from the merge of both several eQTLs linked to one gene and/or eQTLs linked to several genes. The red bars, representing single eQTL per cRM (monogenic and mono-condition eQTL), demonstrates that monogenic reQTLs are the most prevalent in our dataset.

<https://doi.org/10.1371/journal.pgen.1011599.g003>

We then merged the previously described “monogenic” cRMs that shared similar association patterns across genes (see Methods) (Fig 3G). This yielded 453 multigenic cRMs, encompassing an average of 2.38 genes per module. While modules that are only active following stimulation account for 56% (4296/7723) of all cRMs, they only account for 20% (92/453) of multigenic cRM ($p < 10^{-6}$). Thus, stimulation-specific reQTL tend to be more gene-specific than eQTL that are also active under base-line conditions. Nevertheless, 36 multigenic cRM were activated by both TCR and TLR stimulation, and could have a large impact on the immune response (Table F in S1 Table). An open-access website has been developed to visualize cRMs within their genomic context (<https://tools.giga.uliege.be/cedar/publicdw>).

Increasing the number of disease risk loci with matching cis-eQTLs

We then mined our catalogue for cis-eQTLs that colocalize with GWAS-identified risk loci for IMIDs and infectious diseases. We focused on inflammatory bowel diseases (IBD) [33] and rheumatoid arthritis (RA) [34], two IMIDs with worldwide growing prevalence, as well as on the infectious COVID-19 disease [35], a recent major healthcare issue, and involving the TLR7/8 signaling pathways [36]. Colocalization analyses were conducted following ref. 6. This analysis assumes that if a risk locus influences disease susceptibility by altering transcript levels of a given gene in cis, the disease association-pattern (DAP) for this locus and the EAP of the gene should be similar, quantified by the θ metric.

We found 187 significant DAP-EAP correlations ($|\theta| > 0.6$, $p < 0.05$) spanning 39 of the 244 IBD risk loci (16.2%) and encompassing 68 genes, as outlined in Table G in S1 Table. Of interest, 19.2% of these correlations (36 out of 187) were with reQTLs, including 11 risk loci that would have been overlooked in the absence of TLR and TCR stimulations. An intriguing example of the added value of immune stimulation is the long non-coding (lnc) gene *AJ009632.2* (with unknown function) which is subject to a cis-eQTL/cRM active in both control conditions and after TLR4 stimulation. Yet, the effect on gene expression has opposite sign: the variants that increase IBD risk increase the expression levels of *AJ009632.2* in resting state, while doing the opposite after TLR4 stimulation (Fig 4A). The function of this lnc gene is for now poorly defined. *CTSS*, encoding cathepsin S, a lysosomal enzyme expressed by immune and epithelial cells in response to inflammation, regulating antigen presentation and taking part in degradation of extracellular matrix [37], is another interesting example. Its expression is modulated by one cRM in resting condition and a distinct cRM after TLR4 stimulation (S8 Fig). This *CTSS* TLR4 reQTL matched the DAP of the rs4845604 IBD risk locus, located on chromosome 1 (150,100,000–151,120,000). Additional examples of relevant DAP-EAP correlations linked to reQTLs that highlight new candidate genes worthy of further exploration in the context of IBD are listed in Table 1. In addition, even in resting conditions, the GEOCODE eQTL catalogue reveals new DAP-EAP correlations that shed light on previously unknown candidate genes (Table 2).

With regards to RA, we identified 56 significant correlations ($|\theta| > 0.6$, $p < 0.01$) involving 14 of the 124 tested risk loci (11.3%) and 15 genes. Of interest, 12.5% of these correlations (7 out of 56) were associated with reQTLs, including 3 risk loci that would have been overlooked in the absence of TLR and TCR stimulations (Table H in S1 Table). Risk locus rs10497813, located on chromosome 2 (197,200,000–198,800,000), is an interesting region which has been associated with the risk to develop RA, where the DAP’s top SNP corresponds to an intronic variant in the *PLCL1* gene [38]. Interestingly, the same region has been associated with the risk to develop Crohn’s disease (rs6738825, Chr2: 197,310,000–198,110,000) [39]. DAP related to this risk locus matched reQTLs for *PLCL1* (Fig 4B), thereby confirming *PLCL1* as an interesting candidate gene to be further evaluated. This gene has recently been suggested to be linked to RA by modulating the inflammatory response in synoviocytes [40]. Strikingly, we observed that the sign of θ is positive for Crohn’s

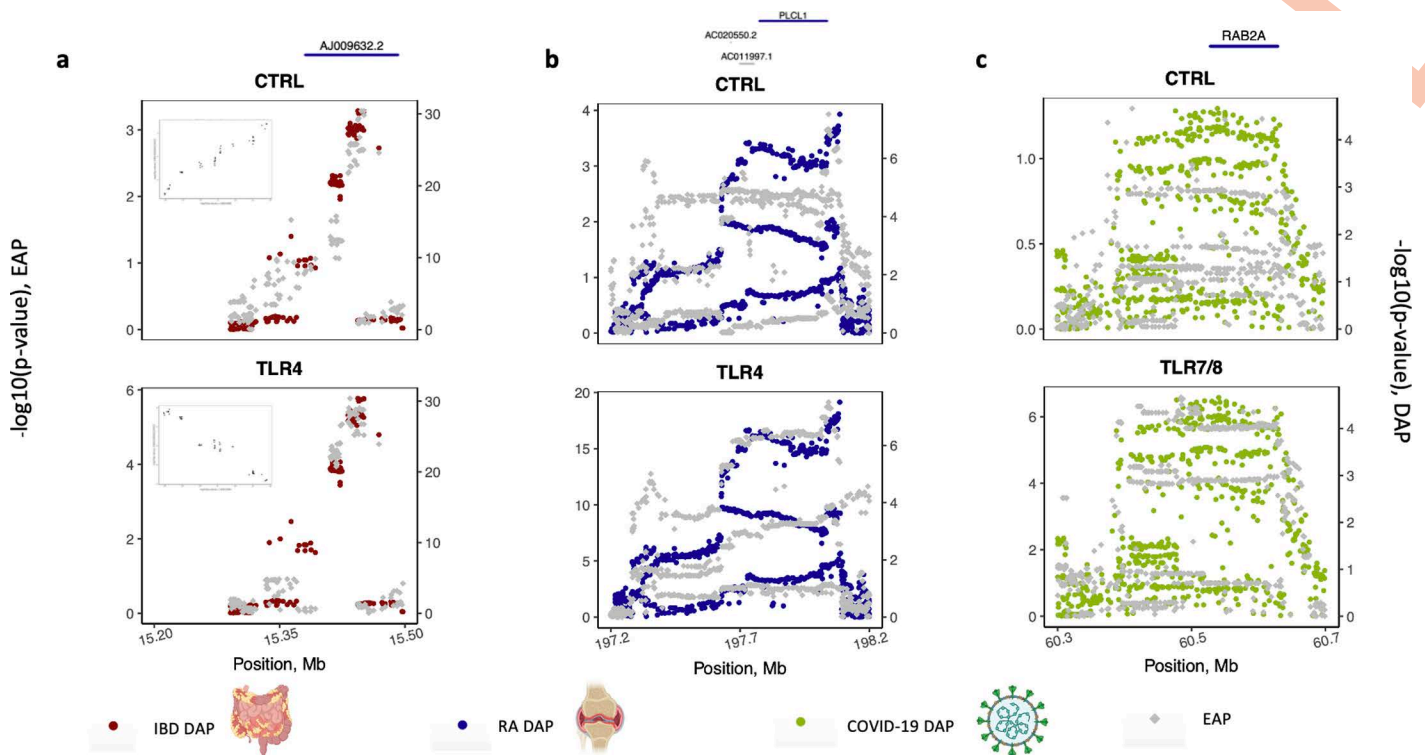


Fig 4. Zoom plots of Disease Association Pattern (DAP) and eQTL association pattern (EAP). **A.** Zoom plot of a locus on Chromosome 21 displaying the correlation between a Disease Association Pattern (DAP) for Crohn's disease (Red) and an eQTL association pattern (EAP) for the *AJ009632.2* gene (grey). A p-value was calculated for the association between each surrounding SNP and the expression level of the *AJ009632.2* gene under different conditions of stimulation within a defined window around the target gene. The EAP on the left Y-axis represents the distribution of association $-\log(p)$ values for all variants within 1Mb of the eGene Transcriptional Start Site (TSS). Similarly, a p-value was extracted from GWAS summary statistics for each association between neighbouring variants of the top SNP and IBD phenotype. The DAP on the right Y axis represents the distribution of association $-\log(p)$ values for all variants within approximately 200 kb of rs2823286, an IBD risk locus from a public dataset. In this example, the DAP significantly correlates with the EAP for the Ctrl condition ($\theta = 0.91$, $p = 0.0004$) and the EAP for the TLR4 condition ($\theta = -0.9$, $p = 0.001$). The plots in the upper left panel show the correlation of each SNP in the EAP and DAP: the correlation is positive for the DAP-EAP in the Ctrl condition and negative for the DAP-EAP in the TLR4 condition. **B.** Zoom plot of a locus on Chromosome 2 showing the correlation between a DAP for Rheumatoid arthritis (blue) and an EAP for *PLCL1* gene (grey). A p-value was calculated for the association between each surrounding SNP and the expression level of the *PLCL1* gene within a defined window around this gene. The EAP on left Y-axis represents the distribution of association $-\log(p)$ values for all variants within 1Mb of the eGene Transcriptional Start Site (TSS), based on our dataset. Similarly, a p-value was extracted from GWAS summary statistics for each association between neighbouring variants of the top SNP and RA phenotype. The DAP on right Y-axis represents the distribution of association $-\log(p)$ values for all variants within approximately 200 kb of rs10497813, a RA risk locus from a public dataset. In this example, the DAP was not significantly correlated with the EAP for the Ctrl condition ($\theta = -0.04$, $p = \text{NS}$) but was significantly correlated with the EAP for the TLR4 condition ($\theta = -0.73$, $p = 0.004$). This example highlights that the eQTL linked to the *PLCL1* gene is a response eQTL. **C.** Zoom plot of a locus on Chromosome 8 illustrating the correlation between a DAP for COVID-19 (green) and an EAP for *RAB2A* gene (grey). A p-value was calculated for each association between each surrounding SNP and the expression level of the *RAB2A* gene within a specifically defined window around this gene. The EAP on the left Y-axis represents the distribution of association $-\log(p)$ values for all variants within 1Mb of the eGene Transcriptional Start Site (TSS), based on our dataset. Similarly, a p-value was extracted from GWAS summary statistics for each association between neighbouring variants of the top SNP and COVID-19 phenotype. The DAP on right Y-axis represents the distribution of association $-\log(p)$ values for all the variants within the approximately 200 kb of rs2875968, a COVID-19 risk locus from a public dataset. In this example, the DAP was not significantly correlated with the EAP for the Ctrl condition but significantly correlated with the EAP for the TLR7/8 condition ($\theta = 0.73$, $p = 0.0001$). This example highlights that the eQTL linked to *RAB2A* gene is a response eQTL. Images used in the figure have been generated with a Biorender Academic License.

<https://doi.org/10.1371/journal.pgen.1011599.g004>

disease, while it is negative for RA (Table 1 and S9 Fig). This suggests that increased expression of *PLCL1* in stimulated immune cells increases the risk to develop IBD, yet decreases risk to develop RA, and hence that the risk variants for IBD are protective for RA and vice versa. Additional examples of relevant DAP-EAP correlations that highlight new candidate genes worthy of further exploration in the context of RA are listed in Table 1.

Table 1. Relevant reQTL-specific colocalisations with risk loci.

Locus				Best Colocalisations θ (p-val)				Response eQTL	Ref.
Chr.	Coordinates	Diseases	Gene	CTRL	TCR	TLR4	TLR7/8		
Chr01_c	Chr1: 150,100,000–151,120,000 (rs4845604-a)	IBD, CD, UC	HORMAD1	-0.02 (0.1)	-0.58 (0.01)	-0.82 (0.0004)	-0.82 (0.0009)	1#	[57]
			CTSS	-0.05 (0.1)	0 (0.4)	-0.72 (0.004)	-0.45 (0.03)	1*	[58]
Chr01_f	Chr1: 161,050,000–162,050,000 (rs1801274)	UC	FCGR2A	-(-)	0.75 (<0.0001)	-(-)	0 (0.9)	1	[59]
Chr02_a	Chr2: 197,310,000–198,110,000 (rs6738825-b)	CD	PLCL1	0 (0.4)	- (-)	0.75 (0.005)	0.64 (0.02)	1*	[60]
			PLCL1	-0.04 (NS)	- 0.73(0.0005)	-0.73 (0.0004)	-0.73 (0.0009)	1*	[61]
Chr03_b	Chr3: 52,920,000–53,150,000 (rs2581828)	CD	RFT1	0.82 (0.006)	- (-)	0.84 (0.005)	0.22 (0.1)	1*	[57–59]
Chr03_c	Chr3: 141,300,000–141,700,000 (rs724016)	CD	ZBTB38	0.39 (0.03)	- (-)	0.73 (0.0007)	0.48 (0.02)	1*	[57]
Chr05_a	Chr5: 10,600,000–10,850,000 (rs2930047)	IBD	DAP	0 (0.1)	- (-)	-0.59 (0.008)	-0.63 (0.005)	1	[58,59,62]
			DAP	0 (NS)	0 (NS)	-0.23 (NS)	-0.63 (0.004)	1	[61]
Chr05_b	Chr5: 390,000–810,000 (rs4957048)	UC	TPPP	0.01 (0.1)	- (-)	0.8 (0.0005)	0.68 (0.006)	1	–
Chr8	Chr. 8: 60300000–60700000 (rs2875968)	COVID (B2,C2)	CA8	- (-)	- (-)	-0.49 (NA)	-0.66 (0.002)	1	–
Chr12_a	Chr12: 111,200,000–112,600,000 (rs3184504)	RA1,RA2	PHETA1	0 (NS)	-0.76 (NS)	0 (NS)	- (-)	1	–
Chr12	Chr. 12: 112800000–113150000 (rs10850094)	COVID (A2,B2,C2)	OAS1	0.34 (1)	0.85 (0.0002)	0.72 (0.004)	0.91 (<0.0001)	1	[63,64]
Specific case of eQTL for which the effect on gene expression is opposite according to the condition of stimulation									
Chr21	Chr21:15,290,000–15,500,000 (rs2823286)	IBD, CD, UC	AJ009632.2	0.91 (0.0004)	0 (0.4)	-0.9 (0.0009)	-0.66 (0.01)	0	–

The EAP-DAP matching listed in the Table 1 have been extracted from Tables F-H in S1 Table based on several criteria: 1. Only the EAP-DAP matchings with eQTLs defined as *Response* eQTL in our eQTL catalogue, 2. Only the EAP-DAP matching for which the DAP from risk loci showed strong p-values, 3. Only the EAP-DAP matchings for which the eQTL linked to the candidate gene was not extensively reported in the literature.

<https://doi.org/10.1371/journal.pgen.1011599.t001>

COVID-19 serves as another example, where we explored its pathogenesis by leveraging reQTLs generated following the activation of TLR7/8, a pathway relevant to the disease. Twenty-seven DAP-EAP correlations ($|\theta| > 0.6$ p 0.01) were observed, involving 5 risk loci (Tables 1 and I in S1 Table). Of note, 12 DAP-EAP comparisons, involving 2 risk loci, were captured through reQTLs alone and would have been overlooked without immune stimulation. The risk locus rs2875968 spans three genes (*RAB2A*, *LINC01301*, and *CA8*), none of which have strong prior evidence implicating them in COVID-19 physiopathology. Through our analysis, we identified eQTL associations between rs2875968 and the genes *RAB2A* and *CA8*, located on chromosome 8 (60300000 – 60700000) (Fig 4C). These eQTLs are active after TCR and TLR stimulation with opposite sign of effect on *RAB2A* and *CA8*. A recent study demonstrated that increased expression of *RAB2A* was linked to more severe COVID-19 outcome due to its role in viral replication [41]. In contrast, no current evidence supports an association between *CA8* and COVID-19 physiopathology. Beyond this example, colocalization analyses further validated the involvement of *IFNAR2* and *OAS1–3* genes in COVID-19 physiopathology (S10 Fig). Notably, the eQTLs related to the *OAS1–3* genes were part of a multigenic cRM active after both TCR and TLR stimulations.

Discussion

We built a large cohort of healthy individuals, stimulating their whole blood with TLR agonists and a TCR antagonist in a standardized approach to enhance our catalog of immune cis-eQTLs. We discovered that 72.9% of eQTLs, referred to as response eQTLs, were revealed through in vitro stimulation. We distinguished three potential mechanisms to explain the context-specific activity of some eQTLs. By exploiting our enhanced eQTL catalog to the analysis of risk loci for immune-mediated and infectious diseases, we demonstrated that response eQTLs improve our ability to identify new colocalizations and pinpoint candidate genes linked to these conditions.

Enhancing our eQTL catalog through functional studies offers a critical resource for mapping genetic effects on specific phenotypes. In such studies, carefully selecting participants is crucial to avoid noise introduced by disease, medications, lifestyle factors such as tobacco and alcohol use, and other independent health issues [42]. To date, only two large healthy cohorts have been established to investigate the genetic determinants related to stimulation-dependent immune response [43,44]. Our research has meticulously selected a highly homogeneous group of healthy participants—non-smokers, non-obese, under 65 years of age, not using any medications except for contraceptive pills, and with no history of chronic diseases or infections. This strategy minimizes external influences on the immune response. Furthermore, to reduce immune responses influenced by ex-vivo manipulations, without interfering with the cell-cell interactions that resides in their internal milieu, we used whole blood as an in-vitro model that best grasps the phenotype of the profile and magnitude of immune response of each individual [45]. This approach captures the overall immune response with minimal cell manipulation, although it may impair the detection of cell-specific eQTLs.

The majority of eQTLs captured in our study were response eQTLs, aligning with findings from other studies [9,11,13]. Notably, 72.9% of eQTLs were not detected in resting conditions, supporting the hypothesis that reQTLs are under significant selective pressure as supported by Kim-Hellmuth et al [13]. Additionally, response eQTLs were predominantly found in monogenic cis-regulatory modules, emphasizing their gene-specific behavior. Our findings also revealed that 1,578 eGenes were influenced by multiple independent eQTLs, with some genes regulated by distinct eQTLs under resting conditions and different eQTLs upon stimulation. This highlights how external stimuli can impact gene regulation through distinct genetic regulators.

As we identified three mechanisms underlying condition-specific cRMs, mechanistic hypotheses may help elucidate how immune stimulation triggers these genetic effects. Previous studies have shown that eQTLs frequently overlap with active cis-regulatory elements, such as enhancers and promoters, which interact with target gene promoters across various immune cell types [46]. Epigenetic modifications, including H3K27ac at enhancers and H3K4me3 at promoters, could drive chromatin accessibility changes, thereby influencing gene expression in a condition-specific manner. For example, LPS exposure has been shown to induce immune tolerance in monocytes through epigenetic modifications detectable as early as 1 hour post-stimulation [47]. Similarly, recent studies on chromatin accessibility in lymphoblastoid cell lines revealed that chromatin accessibility quantitative trait loci (caQTLs) can also account for immune-mediated disease associations. These findings highlight the critical role of chromatin accessibility variations in shaping the regulatory landscape that modulates gene expression [48]. These findings suggest that chromatin accessibility and histone modification dynamics could underlie the condition-specific activation of cRMs, potentially altering the detection or directionality of eQTLs under different conditions.

We employed the previously described θ -based approach [6] for colocalization analysis. This method compares vectors of $\log(1/p)$ values (referred to as “association patterns” or APs) in a specific chromosomal region for pairs of traits under investigation. These pairs could involve two eQTLs (EAP1 vs. EAP2) or an eQTL and a disease (EAP vs. DAP). Importantly, the p -value vectors do not need to originate from the same cohort, as long as both cohorts represent the same population and thus share a similar LD structure. The comparison metric is Pearson’s correlation coefficient, yet, (i) restricting analysis to SNPs with an association p -value < 0.05 for at least one trait, (ii) assigning greater weight to SNPs

with stronger associations (“peaks” in the AP), (iii) incorporating the coherence of allelic effect signs across SNPs within traits, and (iv) providing a signed θ value (positive or negative) to indicate, for example, whether increased gene expression correlates with increased (positive θ) or decreased (negative θ) disease risk. The statistical significance of θ is assessed by comparing the observed θ value with those generated through permutation (randomizing phenotype-genotype pairs) for one of the traits. Permutations are performed for both traits, and the final p-value is the average of the two empirical p-values. This approach accounts for variable LD patterns across the genome.

The θ -based approach is very similar to the SMR approach [49], with which it has been compared, and shown to provide comparable results [6]. In essence, the SMR approach performs a correlation analysis between the effect sizes of the SNPs (β) rather than their $\log(1/p)$ -values. Neither method aims to fine-map association patterns or resolve whether signals arise from one or multiple causative variants (allelic heterogeneity). For both approaches, concluding that two APs match, requires the same set of causative variants to drive both signals. This set of causative variants, of unknown size, is referred to as a regulatory module (RM). Importantly, two APs influenced by distinct but overlapping sets of regulatory variants will not be deemed as “matching.” In this regard, both methods differ from other colocalization methods, including coloc [50–52] and eCAVIAR [53]. For example, coloc can exploit fine-mapping results (obtained f.i. using SuSiE [54]) and test for colocalization between pairs of multiple independent signals found to affect traits 1 and 2, respectively.

Colocalizing cis-eQTLs with genetic variants from the GWAS catalog has effectively pinpointed candidate genes within risk loci. We observed that many risk loci matching our eQTL dataset under resting condition were previously recognized, validating the relevance of our dataset (Tables G-I in S1 Table). Notably, even in resting condition, we discovered several significant eQTLs without prior links between the eGene and diseases as outlined in Table 2. Additionally, a substantial proportion of these new candidate genes (around 59%), corresponded to recently annotated genes, reflecting ongoing advancements in genome annotation and functional characterization. More importantly, by leveraging the power of our response eQTL dataset, we identified matches with risk loci that are typically missed without immune stimulation. This supports the hypothesis that some cis-eQTLs may manifest only under specific conditions.

Our study has several limitations. Firstly, as previously discussed, employing a whole-blood study design presents both advantages and constraints. Using whole blood for eQTL discovery offers us significant advantages, including accessibility, relevance for immune traits, and the ability to capture dynamic systemic responses [45]. However, we agree that whole blood also presents challenges, such as cell-type heterogeneity, potential signal dilution, and confounding factors that may reduce specific regulatory effects. To address these limitations, we performed *in silico* cell deconvolution techniques to refine the interpretation of our blood-based eQTL findings and to evaluate in which impact immune stimulation can alter immune cell proportions. Additionally, we verified that the detected eQTLs were only minimally influenced by blood cell traits. Secondly, our assessment of cis-eQTL effects was based on transcriptomic data collected after 24 hours of TLR and TCR stimulations. This timepoint, although beneficial for capturing a broad range of gene expressions, may overlook some genes activated in the initial phases of the immune response. Nonetheless, recent research reported that more genes were differentially expressed at 24 hours compared to 4 hours post-stimulation. Ultimately, we applied a colocalization method developed by our team [6], acknowledging that other existing colocalization methods might slightly alter the influence of the clustering of our eQTLs into cis regulatory modules as well as eQTL mapping with risk loci.

In conclusion, this work extended our understanding of the immune cis-regulome by conducting at a population-based level *ex-vivo* stimulations. After leveraging eQTL dataset in health, we applied a colocalization approach to public datasets of IMIDs and COVID-19 disease. The specific exploration of reQTLs expanded our ability to colocalize risk loci, which would have been overlooked under resting baseline conditions, thereby identifying new candidate for IBD, RA and COVID-19. The web-based browser that we developed here will facilitate the dissemination of our data with the research community.

Table 2. Relevant eQTL-specific colocalisations with risk loci.

Locus				Best Colocalisations θ (p-val)				Response eQTL	Ref.
Chr	Coordinates	Diseases	Gene	CTRL	TCR	TLR4	TLR7/8		
Chr01_a	Chr1: 2,470,000–2,750,000 (rs10797432_a)	IBD,UC	TNFRSF14-AS1	0.79 (0.02)	0.82 (0.008)	0.85 (0.005)	0.74 (0.02)	0	–
			AL139246.3	-0.67 (0.05)	- (-)	-0.31 (0.07)	-0.76 (0.004)	0	–
Chr01_b	Chr1: 92,950,000–93,450,000 (rs34856868-b)	IBD	DR1	-0.65 (0.0004)	-0.67 (0.0001)	-0.62 (0.001)	-0.71 (0.0004)	0	–
Chr01_c	Chr1: 150,100,000–151,120,000 (rs4845604-a)	IBD, CD, UC	CTSK	0.15 (0.08)	-0.71 (0.006)	0 (0.2)	0 (0.3)	0	–
Chr02	Chr2: 73,900,000–74,300,000 (rs28421442)	RA1	AC073263.1	-0.46 (NS)	0.62 (0.005)	-0.69 (0.005)	-0.49 (NS)	0	–
Chr06	Chr6:111,060,000–111,910,000 (rs3851228)	IBD, UC	Z97989.1	-0.63 (0.004)	-0.46 (0.02)	-0.61 (0.005)	-0.67 (0.003)	0	–
			AL080317.2	0.75 (0.0005)	- (-)	0 (NS)	- (-)	0	–
Chr06	Chr6:166,860,000–167,230,000 (rs1819333)	CD	Z94721.3	-0.5 (0.005)	-0.73 (0.0002)	-0.55 (0.002)	-0.46 (0.007)	0	–
Chr08_b	Chr8: 48,320,000–48,700,000 (rs7011507_b)	IBD,UC	AC026904.1	0.87 (0.0003)	0 (0.6)	0.72 (0.003)	0.63 (0.007)	0	–
Chr11_b	Chr11: 64,050,000–64,500,000 (rs559928)	IBD, CD	AP003774.2	-0.71 (0.0001)	-0.35 (0.02)	-0.66 (0.0003)	-0.71 (<0.0001)	0	–
Chr13	Chr13:40,800,000–41,000000 (rs7329174)	CD	AL590064.1	-0.77 (0.02)	-0.68 (0.04)	-0.86 (0.004)	-0.61 (0.05)	0	–
Chr15	Chr15: 76,800,000–77,600,000 (rs115284761)	RA1,RA2	TSPAN3	-0.78 (0.0007)	-0.66 (0.005)	-0.6 (NS)	-0.82 (0.0003)	0	–
Chr16_c	Chr16: 68,490,000–68,830,000 (rs1728785)	IBD, UC	AC126773.6	0.84 (0.006)	- (-)	0 (0.4)	- (-)	0	–
Chr17	Chr. 17: 39125000–40100000 (rs3744245)	COVID (B2)	ORMDL3	-0.57 (NS)	-0.42 (NS)	-0.6 (<0.0001)	-0.6 (NS)	0	–
Chr21	Chr21: 44,040,000–44,260,000 (rs7282490_b)	IBD,C-D,UC	AP001505.1	0.78 (0.004)	0 (0.3)	0 (0.2)	0 (0.9)	0	–
			GATD3A	-0.79 (0.004)	-0.81 (0.002)	-0.66 (0.02)	-0.68 (0.01)	0	–
			TRAPPC10	-0.83 (0.002)	-0.76 (0.005)	-0.8 (0.003)	-0.75 (0.006)	0	–

The EAP-DAP matching listed in the [Table 2](#) have been extracted from Tables F-H in [S1 Table](#) based on several criteria: 1 Only the EAP-DAP matchings for which the DAP from risk loci showed strong p-values. 2. Only the EAP-DAP matching if there was no reported link between the genes and diseases on the basis of available eQTL information.

(i)The columns linked to “Locus” refer to the name and chromosomal coordinates of the corresponding risk locus (Locus, Chr, Beg, End) (GRCh38/hg19) as well as the disease where the locus has been reported (CD: Crohn disease, UC: Ulcerative colitis, IBD: Inflammatory bowel disease, RA: rheumatoid arthritis, COVID: COVID-19 disease). A risk locus was associated with a letter (ex: rs7282490_b) refers to a risk locus for which the window has been manually adapted before processing to EAP-DAP matching. If there is no letter associated with the Rs, that means that the boundaries of the risk locus have not been adapted.

(ii)The columns linked to “Best colocalizations” refer to the genes and conditions of stimulation involved in matching DAP–EAP ($|\theta| > 0.6$, p-val <0.05). Among the different EAP-DAP, the best θ -values and corresponding p-values are reported in the [Table 2](#). (NS= non-significant p-value, (-) = no computed DAP-EAP matching).

(iii)The column “Response eQTL” refers to the eQTL characterization as response eQTL (1) or not (0). Both the eQTL catalogue (Table C in [S1 Table](#)) and extended catalogue were interrogated to confirm the response eQTL status. Symbol “*” refers to eGenes with switched modules (including different cRMs: one cRM with resting conditions and one cRM with response eQTL), Symbol “#” refers to presence of eQTL in resting condition in the extended catalogue.

(iv)The column “Ref.” results from an in-depth review of colocalization between immune -mediated inflammatory diseases (IMIDs) GWAS and reported immune eQTLs

<https://doi.org/10.1371/journal.pgen.1011599.t002>

Materials and methods

Ethics statement

The study was approved by the Ethics committee of Erasme Hospital, Brussels, Belgium (Reference number: P2015/425, date approval: 03/11/2015). All used methods were in accordance with approved guidelines and were performed in accordance with the Declaration of Helsinki. Each subject signed an informed consent.

Study population

The study cohort (GEOCODE cohort) involved 406 healthy subjects prospectively included between October 2016 and March 2018. Inclusion criteria included age between 18 and 65 years, smoking-free status, drugs free with exception of hormonal contraception or finasteride, benzodiazepine and proton-pump inhibitors, and self-assessment of being in “good health” ([S1 Data](#)). First, second- and third-degree family members were excluded as well as shift workers (chronic jet lag) or any of following situations within the previous two weeks of inclusion: active allergic disease, episode of body temperature > 38 °C, dentist consultation, endoscopy, vaccination or steroid (systemic or topic) treatment. Gender, age, height, weight, medical and familial history were collected for all participating subjects.

Sample collection, whole blood cell culture, stimulations and cytokine measurements

For all participating subjects, 40 ml of peripheral blood (distributed in several tubes - see [S1 Data](#)) were collected between 07h30 and 10h00 a.m. (to standardize the circadian cycle) after overnight fasting. A fresh EDTA tube was processed within the same day of blood collection for immunophenotyping. Details on immunophenotyping are presented in [S1 Data](#). All immune cell counts were expressed by mm³. Plasma and sera were aliquoted and stored at -20 °C for later measurements. Whole blood cell cultures and stimulations were performed within three hours of blood collection. Two different TLR agonists, Resiquimod (R848) - TLR7/8 and Lipopolysaccharide (LPS) - TLR4, and one TCR antagonist (anti-CD3/anti-CD28) were used ([S1 Data](#)). The IC₅₀ dose (or concentration) of each stimulant was chosen from a ranging dose-stimulation pilot study. Briefly, the blood was diluted 1/4 with pre-warmed FBS-RPMI. TLR agonists or TCR antagonists were added prior to incubation for 24h, at 37°C in 5% CO₂ containing atmosphere. Whole blood from each well was next transferred into a pre-labelled 1.5 ml Eppendorf tubes. After centrifugation, the supernatants were collected and stored at -80°C until use. Cytokines production following TCR and TLR stimulation were measured using a standard ELISA method. IL-6 (DuoSet Human IL-6, R&D Systems) and TNFα (DuoSet Human TNF-α, R&D Systems) were measured for all TLR stimulation conditions. IFNγ (DuoSet Human IFNγ, R&D Systems) and IL-2 (DuoSet Human IL-2, R&D Systems) were measured for TCR stimulation conditions. All cytokine measurements were expressed in pg/ml.

DNA extraction

Human genomic DNA was extracted from EDTA-collected peripheral blood by automated genomic DNA isolation Tecan Freedom. The extraction was performed in batches of 32 samples. Subsequently, concentration and quality DNA were measured by nanodrop ND-1000. DNA concentration standardized for all samples to 50 ng/ml.

SNP genotyping and imputation

The 406 individuals were genotyped for >700K SNPs using Illumina’s Human OmniExpress BeadChips, an iScan system and the Genome Studio software (GIGA genomics core facility, Liège, Belgium). We eliminated variants with call rate ≤0.95, with the minor allele frequency (MAF) <0.05 and deviating from Hardy–Weinberg equilibrium (HWE) (<0.001). European ancestry of all individuals was analyzed by PCA using the HapMap population as reference and used the 3 first PC as covariates. We used The Michigan Imputation Server with TOPMed Imputation Reference panel (<https://imputation.biodatacatalyst.nhlbi.nih.gov>) to impute genotypes at autosomal variants in our population. We removed all SNP with

info < 0.4. A new quality control has been applied after imputation by filtering all SNP with MAF < 0.05 and HWE < 0.001. Finally, 4,939,638 variants were obtained after imputation.

GWAS Analysis

Before starting association analyses, several covariates were integrated for analyses, namely sex, age, Body Mass Index (BMI) and the three first principal components related to ancestry. The association analyses were based on linear regression. Wald test was used for phenotypes association. We corrected the P-value threshold for associations due to the multiple testing of phenotypes. Phenotypes and covariates were forced to follow a gaussian distribution, preserving only the original rank orders, by using a quantile normalization. After controlling for residual test statistic inflation via genomic control, Manhattan plots and quantile-quantile (QQ) plots were used to describe GWAS results.

RNA extraction

RNA was extracted using Trizol Reagent (Thermo Fisher Scientific). Four conditions of stimulation at 24h were selected for each healthy individual (Control, R848, LPS, anti-CD3/antiCD28). In total, 1,574 samples were processed for RNA extraction. Briefly, Trizol Reagent obtaining phase was centrifuged at 3,300g for 5 minutes and clarified phase was subsequently used to remove cell debris. Extraction was performed using RNeasy Plus Universal Mini Kit (QIAGEN). Samples were next loaded on QIAcube to finalize extraction. The quantification of RNA was performed using a Nanodrop spectrophotometer. The RNA quantity was expressed ng/μl with a median of 37 ng/μl (IQR 27.7 – 46.6). The quality of RNA was measured using QIAxcel. The median RNA Integrity Score was 7.2 (IQR 6.7 – 7.8). All samples that met criteria of having a RIS value of 6.0 or higher were batched for RNA sequencing.

RNA sequencing

RNA sequencing was performed using QuantSeq 3' mRNA-Seq Library Prep Kit FWD at the GIGA-Genomics platform for a total of 1407 samples. Briefly, 100 ng of total RNA was used from each sample as the starting material. This method uses oligo dT beads to select poly-A mRNA from the total RNA sample. The selected RNA is then heat fragmented and randomly primed before cDNA synthesis from the RNA template and UMI was added before following steps. The resultant cDNA was next processed through Lexogen library preparation kit procedure following the instructions of the manufacturer and using designed indexed adapters for multiplexing of samples. After enrichment, the samples were qPCR quantified and equimolar pooled before proceeding sequencing on the NovaSeq equipment (Illumina), with a sequence coverage goal of 20M 150bp single-end reads (median achieved was ~ 21M total reads). All samples were sequenced in batches with a mix of the four conditions (control and stimulated conditions) across the batches. A globin clear module was used to remove globin mRNA from whole blood conditions. The globin genes expression (HBA1, HBA2, HBB HBG1, HBG2, HBD) after using globin clear module was around 2% of total reads.

RNA sequence QC and read filtering

Raw RNA-seq data were demultiplexed and trimmed with bcl2fastq then aligned with STAR using Homo_sapiens.GRCh38.97.gtf as reference genome. Reads sharing the same UMI and mapped to the same place in the genome were collapsed using UMI-tools to avoid bias related to qPCR amplification. Quantification of reads was evaluated by FeatureCounts. RNA-seq expression samples were scrutinized using several quality control measures before being included in the final analysis set. A variant calling was applied using QTLtools to evaluate correct sample matching. Samples which failed to QC and/or exposed matching discrepancies were removed for the final analyses. At the end of read filtering and QC, 1305 samples from 359 subjects were used for following analyses (354, 327, 334 and 290 in control, LPS, R848 and TCR groups, respectively).

From the 60,617 annotations included in the analysis, we excluded short RNAs, pseudogenes, and mitochondrial RNAs from the analysis, thus leaving 38,286 annotations. We next filtered out genes with zero CPM in 90% of all samples; leaving 22,272 annotations of which 15,495 were protein coding for downstream analysis.

Transcriptomic analysis

The normalization of gene expression counts as well as the identification of differentially expressed genes were conducted using the DESeq2 v1.36.0 package, which accounts for library depth by calculating a normalization factor for each sample. The process involved the following steps: First, the geometric mean expression of each gene across all samples was computed to serve as a pseudo-reference. Second, the expression of each gene was divided by its pseudo-reference value to calculate ratios for each gene. The normalization factor for each sample was then determined as the median of these ratios specific to that sample. Finally, normalized counts were obtained by dividing the raw counts by the normalization factor. Differentially expressed genes were identified normalizing each pair of comparison (Stimulation vs Control condition) together.

Principal Component Analysis (PCA) was applied taking the 500 most variable genes across the samples to confirm stratification in the dataset regarding the group stimulations. Principal components (PCs) were analyzed to identify the genes with the greatest contribution to each PC, based on their loading values, which indicate the influence of each gene on the PC. Greatest contributor genes were selected accordingly to an absolute value of correlation coefficient to PCs higher than 0.5.

Pathway enrichment analysis was performed on the top differentially expressed genes in PC1 using the public Reactome database. Only biological pathways with a false discovery rate (FDR) <0.05 were included. A heatmap was used to visualize the correlation patterns of the 500 most variable genes across subjects after 24 hours of stimulation.

Covariates

Previously to eQTL analysis, to remove confounding variation in the gene expression data, that might mask or skew the effects of local genetic variation, we calculated the expression principal components (PCs) for each of the four conditions. For each subset of eQTL analyses, the dominant expression PCs were chosen for inclusion as covariates in the model such that the number of significant associations was maximized across all conditions. Before inclusion of PC in analysis, we checked by GWAS if each PC was associated to a specific genetic signal. Gender, age, BMI and three top PC (from genotype data) were also considered as additional covariates. To gain insight into the biological meaning of these factors, the relationship between expression PC and phenotypic covariates were analyzed.

eQTLs analysis and cis-regulatory modules (cRM)

The cis window was defined as 1 megabase up- and down-stream of the transcriptional start site (± 1 Mb). Analysis was run by QTLtools using 10,000 permutations to get adjusted p-values. These adjusted p-values were then used to compute the corresponding false discovery rates (FDR or q-value). eQTLs are considered as true eQTLs regarding a FDR threshold ≤ 0.05 . Of note, for the purpose of a specific analysis (see results), we also used an “extended” catalogue of eQTLs. The extended catalogue contains all eQTLs with an FDR lower or equal to 0.05 plus non significant eQTLs regarding the threshold of FDR but with a gene being part of a significant eQTLs in at least one of the four conditions.

For rs1801274, we conducted a trans-eQTL analysis using the same covariates as in the cis-eQTL analysis. The analysis was performed with QTLtools, employing 10,000 permutations to obtain adjusted p-values

The list of significant cis-eQTLs was used to delineate the corresponding eQTL association patterns (EAP). EAP was defined as the distribution of association $-\log(p)$ values for all the variants in the region of 1Mb around the eGene Transcriptional Start Site (TSS). Hence, the denomination of eQTL referred to the top significant variant of the corresponding EAP.

In a first step, we have analyzed all EAPs related to one given gene (monogenic) in a window of 1 Mb. If the expression of a given gene is influenced by eQTLs in two conditions, the corresponding EAPs are expected to be similar. In this case, the eQTLs are merged in a monogenic cis-regulatory modules (cRM). The maximum number of eQTLs merged in a cRM is equal to the number of conditions (that is 4, CTRL, TLR4, TLR7/8 and TCR conditions of stimulation). In a second step, we have analyzed all EAPs in a given window of 1 megabase. If the expression of different genes is influenced by the same eQTL, across different conditions, the corresponding EAPs are expected to be similar. In this case, the corresponding eQTL is a “multi-genes” eQTL. When “multi-genes” EAP are similar, the eQTLs are merged in multigenic cRMs. θ metrics was used to calculate the correlation between EAPs based on the methodology that we have developed and previously published [6], a minimal of 50 SNPs is considered for computing θ metrics. All EAP-EAP correlations with $|\theta \text{ metrics}| \geq 0.6$ are considered for downstream analyses.

Investigating the distribution of the three mechanisms driving cRM activity

Each monogenic cis-Regulatory Module (cRM) was characterized by indicating whether the gene is part of an eQTL ($FDR \leq 0.05$) (1) or not (0) in the four conditions defined in order CTRL-TCR-LPS-R848. As an example, a gene matching as eQTL in the three conditions of stimulation but not in resting condition is labelled 0111 while another gene matching as eQTL in TCR and R848 conditions only is labelled 0101

We quantified the occurrence of changes in status for a gene regarding different conditions under specific scenarios.

First mechanism, a gene might be part of an eQTL in condition “a” but not in condition “b” due to a too low level expression in the second condition. We calculated the number of cases where an eQTL transitioned from 0 to 1, corresponding to the first mechanism, where gene expression levels are too low in a condition for the eQTL to be detectable. Next, we assessed cases where a gene match as an eQTLs in several conditions but is splitted in multiple modules, representing the third mechanism, where a gene (eGene) switches cRMs between conditions, the genes is part of an eQTL in both conditions but the involved variants are distinct. The remaining activity patterns were attributed to the second mechanism, the gene is detected in both conditions at sufficient levels of expression but only match as an eQTL in one condition. This second mechanism was validated using a permutation-based methodology approach fully described, including the computational and statistical procedures, here: <https://doi.org/10.1101/2024.10.14.24315443>.

Correlations EAP and DAP

We used the disease association patterns (DAP) related to IBD, RA and Covid-19 risk loci reported in a recent GWAS meta-analyses [1–3]. If regulatory variants affect disease risk by altering gene expression, the corresponding DAP and EAP are expected to be similar, even if obtained from different cohorts with the same ethnicity. In the same approach than EAP-EAP correlations, all SNPs with nominal p-value < 0.05 for at least one EAP, present in a specific window around the risk loci, were compared to EAP obtained in our dataset. For each DAP, we manually visualized the distribution of association $-\log(p)$ values to define the comparison window.

To assess whether some eQTL signals might be confounded by blood cell traits, we exploited data from recent GWAS meta-analyses of blood cell phenotypes [55]. Out of 7,122 identified loci, we extracted summary statistics for 2,162 loci associated with four key blood cell traits: eosinophils, neutrophils, lymphocytes, and monocytes. These statistics were used to generate plots illustrating the distribution of association $-\log(p)$ values for each locus. Given the large volume of plots, we focused on loci where the top SNP had an association exceeding $-\log(10^{-15})$. Subsequently, we manually reviewed the distribution of association $-\log(p)$ values to define the comparison window, corresponding to a specific DAP. After validating the DAP window, we computed DAP-EAP in the same way than EAP-EAP previously described.

Softwares

All analyses related to quality controls were performed using PLINK version 1.9 while PLINK version 2.0 was used for association analyses. Of note, PLINK version 2.0 can manage the imputation quality metric (described as phased

dosages meaning the probability of a true imputation) but not PLINK version 1.9. BCF tools v1.17 was used for filtering files obtained after imputation with INFO score. R software – package qqman was used for graphics. QTLtools 1.3.1 was used for eQTLs analysis following methodology developed in reference paper [56].

Supporting information

S1 Fig. Association of rs1801274 with genes related to PC1. A. **Principal component analysis (PCA)** of the expression of the most variable 500 genes across the four conditions of stimulation (Control (Ctrl), TLR4, TLR7/8 and TCR stimulations). B. **Manhattan plot showing significant association of rs1801274 with PC1 phenotype.** Principal Components were calculated for expression of genes in each paired condition (control-stimulated condition). By considering PC as new phenotype in GWAS, rs1801274, located on chromosome 1, was significantly associated with PC1-TCR phenotype ($p=4e-45$). C. **Distribution of rs1801274 genotypes in the TCR group.** TCR group observed in panel A has been highlighted by removing others groups and all individuals have been flagged by genotype of variant rs1801274. Individuals with genotype AA segregate from the two others genotypes (AG and GG) and tend to form the continuum with control group as shown panel A.

(TIFF)

S2 Fig. Association of rs1801274 with IL-2 and IFN γ levels following anti-CD3/anti-CD28 stimulation. A. **Manhattan plot showing significant association of rs1801274 with IFN γ level phenotype.** IFN γ level was considered as phenotype and associated with genotypic data through GWAS. rs1801274, located on chromosome 1, was significantly associated with the phenotype ($p=7.72e-40$). Geocode cohort available for this phenotype ($n=380$). B. **Manhattan plot showing significant association of rs1801274 with IL-2 level phenotype.** IL-2 level was considered as phenotype and associated with genotypic data through GWAS. rs1801274, located on chromosome 1, was significantly associated with the phenotype ($p=2.63e-15$). Only a sub cohort was available for this phenotype ($n=187$).

(TIFF)

S3 Fig. Examples of monogenic and multigenic cis-Regulatory modules (cRMs). A. **Examples of monogenic cRMs.** Graphical representation of one cRM modulating *UBE2L3* (Ubiquitin Conjugating Enzyme E2 L3) and *C5* (Complement C5) genes. Every node corresponds to a condition-specific EAP linked to the Gene and edges connect pairs illustrates a significant matching of EAPs with $\theta \geq |0.6|$. B. **Examples of multigenic cRMs.** Graphical representation of two multigenic cRMs. The first module clusters eQTLs linked to *LLGL1* (LLGL Scribble Cell Polarity Complex Component 1) and *TOP3A* (DNA Topoisomerase III Alpha) genes. The second module clusters eQTLs linked to *TRGJP2* (T Cell Receptor Gamma Joining P2) and *TRGJP1* (T Cell Receptor Gamma Joining P1) genes. Every node corresponds to a condition-specific EAP linked to the genes and edges connect pairs illustrates a significant matching of EAPs with $\theta \geq |0.6|$. The blue line is used in case of positive θ while red line is used in case of negative θ .

(TIFF)

S4 Fig. Upset plot of the distribution of cis-eQTLs across resting and conditions of stimulation, based on an extended catalog including 26.629 cis-eQTLs extended. Despite the larger catalog, the plot reveals the continued predominance of condition-specific eQTLs, with the majority being reQTLs. Since this graph only considers monogenic cRMs, the maximum number of eQTLs per gene is four, corresponding to the number of conditions in our dataset. The bars indicate the number of modules.

(TIFF)

S5 Fig. Additional examples of the first mechanism. These manually curated examples highlight eQTLs where the gene is either not expressed or expressed at too low level for the eQTL to be detected in the resting condition, but becomes detectable after stimulation. The boxplots display gene expression levels, with the Y-axis representing

normalized gene expression and the X-axis indicating different conditions (Control (Ctrl) condition in red, TCR stimulation in purple, TLR4 stimulation in green, TLR7/8 stimulation in blue).

(TIFF)

S6 Fig. Additional examples of the second mechanism. These manually curated examples highlight eQTLs where there is a loss of genotype effect in one specific condition (resting or one condition of stimulation), while the gene remains expressed at sufficiently high levels for the eQTL to be detected if it exists. X-axis shows the different conditions of stimulation (red – Ctrl, purple – TCR, green – TLR4, blue- TLR7/8), and Y-axis shows normalized gene expression. The genes are indicated at the top of each panel. The cRM related to the target gene (in uppercase) are labelled 1000, 0100, 0010 and 0001 for the conditions of stimulation: Ctrl-, TCR-, TLR4 - and TLR7/8 -, respectively. The situations are depicted according to these labels.

(TIFF)

S7 Fig. Manually curated examples of eQTLs when an eGene switches cRM between conditions. The plots represent the EAPs for several selected genes under different conditions of stimulation (red – Ctrl, purple – TCR, green – TLR4, blue- TLR7/8). The Y-axis shows the distribution of $-\log(p)$ values for all variants in the region around the top cis-eQTL. The X-axis represents a genomic region centered on the Transcriptional Start Site (TSS) of the gene. The selected eGenes are modulated by different cRMs. The peak of pattern of the first cRM is highlighted by a red arrow, while the pattern of the second cRM is highlighted by a blue arrow and when three different cRM exist, the pattern of the third cRM is highlighted by a green arrow.

(TIFF)

S8 Fig. Specific example of the CTSS gene which is modulated by two distinct cRMs. The dashed line represents the transcriptional start site (TSS) of the CTSS (Cathepsin S) gene. The Y-axis shows the distribution of $-\log(p)$ values for all the variants in the region around the top cis-eQTL. The X- axis represents a genomic region centered on the TSS of the CTSS gene. The EAP of the different conditions of stimulation are shown in colors (red – Ctrl, purple – TCR, green – TLR4, blue- TLR7/8). The EAP in Ctrl condition (in red) is sufficiently different from EAPs in TLR4 and TLR7/8 conditions to be included in a specific cRM (cRM 4515). EAPs related to TLR4 and TLR7/8 conditions were similar and merged in a same cRM (cRM 2090).

(TIFF)

S9 Fig. Zoom plots of Disease Association Pattern (DAP) and eQTL association pattern (EAP). A Zoom plot of a locus on Chromosome 2 illustrating the correlation between DAP for IBD (red) and EAP for PLCL1 gene (grey). A p-value was calculated for each association between each neighbouring variants of SNP and the expression of the *AJ009632.2* gene within a specifically defined window around this target gene under each condition of stimulation. The **EAP**, shown on left Y-axis, represents the distribution of association $-\log(p)$ values for all the variants in the region in the 1Mb region around the eGene Transcriptional Start Site (TSS), derived from our dataset. Similarly, a p-value was extracted from GWAS summary statistics for each association between neighbouring variants of the top SNP and IBD phenotype. The resulting **DAP** on right Y-axis, represents the distribution of association $-\log(p)$ values for all variants in the ~200 kb region around rs6738825, an IBD risk locus from public dataset (ref 27). In this example, the DAP was not significantly correlated with the EAP linked to the Ctrl ($\theta = -0.04$, $p = \text{NS}$) but was significantly correlated with the EAP linked to the TLR4 condition ($\theta = 0.75$, $p = 0.005$). **B Zoom plot of a locus on Chromosome 2 illustrating the correlation between DAP for RA (blue) and EAP for PLCL1 gene (grey).** A p-value was calculated for each association between each neighbouring variants of SNPs and the expression of the PLCL1 gene within a specific window around this gene. The **EAP on the left Y-axis**, represents the distribution of association $-\log(p)$ values for all variants in the 1Mb region around the eGene Transcriptional Start Site (TSS), derived from our dataset. Similarly, a p-value was extracted from

GWAS summary statistics for each association between neighbouring variants of the top SNP and RA phenotype. The resulting **DAP** on the right Y-axis, represents the distribution of association $-\log(p)$ values for all variants in the ~ 200 kb region around rs10497813, a RA risk locus from a public dataset (ref. 25). In this example, the DAP was not significantly correlated with the EAP linked to the Ctrl condition ($\theta = -0.04$, $p = \text{NS}$) but was significantly correlated with the EAP linked to the TLR4 condition ($\theta = -0.73$, $p = 0.004$). The correlation plots show the correlation of each SNP in EAP and DAP. The add-on plots of correlation show that the correlation is positive for matching DAP-EAP in case of IBD but negative for matching DAP-EAP in case of RA. Images used in the figure have been generated with a Biorender Academic License. (TIFF)

S10 Fig. Zoom plot of a locus on Chromosome 12 showing correlation between DAP for COVID-19 (green) and EAP for OAS1 gene (grey). For each neighbouring variant of the SNP, a p-value was computed indicating the association with the expression of the RAB2A gene within a specific genomic window. **The EAP, represented on left Y-axis,** was defined as the distribution of association $-\log(p)$ values for all variants within 1Mb vicinity around the eGene Transcriptional Start Site (TSS), derived from our dataset. Similarly, a p-value was extracted from GWAS summary statistics for each association between neighbouring variants of the top SNP and COVID-19 phenotype. The resulting **DAP** shown on right Y-axis, represents the distribution of association $-\log(p)$ values across all variants within a manually defined region ~ 200 kb around rs10850094, a COVID-19 risk locus obtained from public dataset (Ref. 26). In this particular example, the DAP exhibited no significant correlation with the EAP associated with Ctrl ($\theta = 0.34$, $p = \text{NS}$), whereas it displayed significant correlations with EAPs linked to TLR7/8 ($\theta = 0.91$, $p < 0.0001$), TCR ($\theta = 0.85$, $p = 0.0002$) and TLR4 conditions ($\theta = 0.72$, $p = 0.004$). This example underscores that the eQTLs associated with *OAS1* gene are *responsive* eQTLs. The add-on correlation plots demonstrate positive correlations between matching DAP-EAPs pairs. Images used in the figure have been generated with a Biorender Academic License. (TIFF)

S1 Table. DAP-EAP correlations about COVID-19 (summary statistics from Nature. 2023;621(7977):E7-E26). (XLSX)

S1 Data. Exploring the deconvolution of whole blood tissue and influence on eQTL discovery. (DOCX)

Acknowledgments

We gratefully thank all participants to the GEOCODE cohort.

Author contributions

Conceptualization: Claire Liefferinckx, Denis Franchimont.

Data curation: Claire Liefferinckx, David Stern.

Formal analysis: Claire Liefferinckx, David Stern, Jérémie Bottieau, Christophe Dubussy.

Funding acquisition: Claire Liefferinckx, Denis Franchimont.

Investigation: Claire Liefferinckx.

Methodology: David Stern, H el ene Per ee, Alice Mayer, Vyacheslav Petrov, Souad Rahmouni, Michel Georges.

Project administration: Claire Liefferinckx, Charlotte Minsart.

Resources: Eric Quertinmont, Vjola Tafciu, Charlotte Minsart, Wouter Coppieters, Latifa Karim.

Software: Alex Kvasz.

Supervision: Michel Georges, Denis Franchimont.

Writing – original draft: Claire Liefverinckx.

Writing – review & editing: H el ene Per ee, J er emie Bottieau, Charlotte Minsart, Souad Rahmouni, Denis Franchimont.

References

1. Claussnitzer M, Cho JH, Collins R, Cox NJ, Dermitzakis ET, Hurlles ME, et al. A brief history of human disease genetics. *Nature*. 2020;577(7789):179–89. <https://doi.org/10.1038/s41586-019-1879-7> PMID: 31915397
2. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science*. 2012;337(6099):1190–5. <https://doi.org/10.1126/science.1222794> PMID: 22955828
3. Gusev A, Lee S, Trynka G, Finucane H, Vilhjalms on B, Xu H, et al. Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. *American Journal of Human Genetics*. 2014;95(5):535–52.
4. Consortium GT. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*. 2020;369(6509):1318–30. <https://doi.org/10.1126/science.aaz1776> PMID: 32913098
5. Umans BD, Battle A, Gilad Y. Where are the disease-associated eQTLs?. *Trends in Genetics*. 2021;37(2):109–24.
6. Momozawa Y, Dmitrieva J, Theatre E, Deffontaine V, Rahmouni S, Charlotiaux B. IBD risk loci are enriched in multigenic regulatory modules encompassing putative causative genes. *Nature Communications*. 2018;9(1):2427.
7. Mostafavi H, Spence JP, Naqvi S, Pritchard JK. Systematic differences in discovery of genetic effects on gene expression and complex traits. *Nat Genet*. 2023;55(11):1866–75. <https://doi.org/10.1038/s41588-023-01529-1> PMID: 37857933
8. Lee MN, Ye C, Villani A-C, Raj T, Li W, Eisenhaure TM, et al. Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science*. 2014;343(6175):1246980. <https://doi.org/10.1126/science.1246980> PMID: 24604203
9. Fairfax BP, Humburg P, Makino S, Naranbhai V, Wong D, Lau E, et al. Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science*. 2014;343(6175):1246949. <https://doi.org/10.1126/science.1246949> PMID: 24604202
10. Ye CJ, Feng T, Kwon H-K, Raj T, Wilson MT, Asinovski N, et al. Intersection of population variation and autoimmunity genetics in human T cell activation. *Science*. 2014;345(6202):1254665. <https://doi.org/10.1126/science.1254665> PMID: 25214635
11. Kim S, Becker J, Bechheim M, Kaiser V, Noursadeghi M, Fricker N, et al. Characterizing the genetic basis of innate immune response in TLR4-activated human monocytes. *Nat Commun*. 2014;5:5236. <https://doi.org/10.1038/ncomms6236> PMID: 25327457
12. Quach H, Rotival M, Pothlichet J, Loh Y-HE, Dannemann M, Zidane N, et al. Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. *Cell*. 2016;167(3):643–656.e17. <https://doi.org/10.1016/j.cell.2016.09.024> PMID: 27768888
13. Kim-Hellmuth S, Bechheim M, P utz B, Mohammadi P, N ed elec Y, Giangreco N, et al. Genetic regulatory effects modified by immune activation contribute to autoimmune disease associations. *Nat Commun*. 2017;8(1):266. <https://doi.org/10.1038/s41467-017-00366-1> PMID: 28814792
14. Alasoo K, Rodrigues J, Mukhopadhyay S, Knights AJ, Mann AL, Kundu K, et al. Shared genetic effects on chromatin and gene expression indicate a role for enhancer priming in immune response. *Nat Genet*. 2018;50(3):424–31. <https://doi.org/10.1038/s41588-018-0046-7> PMID: 29379200
15. Schmiedel BJ, Gonzalez-Colin C, Fajardo V, Rocha J, Madrigal A, Ram irez-Su astegui C, et al. Single-cell eQTL analysis of activated T cell subsets reveals activation and cell type-dependent effects of disease-risk variants. *Sci Immunol*. 2022;7(68):eabm2508. <https://doi.org/10.1126/sciimmunol.abm2508> PMID: 35213211
16. Soskic B, Cano-Gamez E, Smyth DJ, Ambridge K, Ke Z, Matte JC, et al. Immune disease risk variants regulate gene expression dynamics during CD4+ T cell activation. *Nat Genet*. 2022;54(6):817–26. <https://doi.org/10.1038/s41588-022-01066-3> PMID: 35618845
17. H ader A, Sch auble S, Gehlen J, Thielemann N, Buerfent BC, Sch uller V, et al. Pathogen-specific innate immune response patterns are distinctly affected by genetic diversity. *Nat Commun*. 2023;14(1):3239. <https://doi.org/10.1038/s41467-023-38994-5> PMID: 37277347
18. Kumasaka N, Rostom R, Huang N, Polanski K, Meyer KB, Patel S, et al. Mapping interindividual dynamics of innate immune response at single-cell resolution. *Nat Genet*. 2023;55(6):1066–75. <https://doi.org/10.1038/s41588-023-01421-y> PMID: 37308670
19. Liefverinckx C, De Gr eve Z, Toubeau J-F, Per ee H, Quertinmont E, Tafciu V, et al. New approach to determine the healthy immune variations by combining clustering methods. *Sci Rep*. 2021;11(1):8917. <https://doi.org/10.1038/s41598-021-88272-x> PMID: 33903641
20. Aguirre-Gamboa R, de Klein N, di Tommaso J, Claringbould A, van der Wijst MG, de Vries D, et al. Deconvolution of bulk blood eQTL effects into immune cell subpopulations. *BMC Bioinformatics*. 2020;21(1):243. <https://doi.org/10.1186/s12859-020-03576-5> PMID: 32532224
21. Nagelkerke S, Schmidt D, de Haas M, Kuijpers T. Genetic variation in low-to-medium-affinity Fc gamma receptors: Functional consequences, disease associations, and opportunities for personalized medicine. *Frontiers in Immunology*. 2019;10:2237.
22. Williams M, Bruhns P, Saeys Y, Hammad H, Lambrecht BN. The function of Fc gamma receptors in dendritic cells and macrophages. *Nat Rev Immunol*. 2014;14(2):94–108. <https://doi.org/10.1038/nri3582> PMID: 24445665
23. Bredius RG, Fijen CA, De Haas M, Kuijper EJ, Weening RS, Van de Winkel JG, et al. Role of neutrophil Fc gamma RIIa (CD32) and Fc gamma RIIIb (CD16) polymorphic forms in phagocytosis of human IgG1- and IgG3-opsonized bacteria and erythrocytes. *Immunology*. 1994;83(4):624–30. PMID: 7875742

24. Bruhns P, Iannascoli B, England P, Mancardi DA, Fernandez N, Jorieux S, et al. Specificity and affinity of human Fcγ receptors and their polymorphic variants for human IgG subclasses. *Blood*. 2009;113(16):3716–25. <https://doi.org/10.1182/blood-2008-09-179754> PMID: [19018092](https://pubmed.ncbi.nlm.nih.gov/19018092/)
25. Duffy D, Rouilly V, Libri V, Hasan M, Beitz B, David M, et al. Functional analysis via standardized whole-blood stimulation systems defines the boundaries of a healthy immune response to complex stimuli. *Immunity*. 2014;40(3):436–50. <https://doi.org/10.1016/j.immuni.2014.03.002> PMID: [24656047](https://pubmed.ncbi.nlm.nih.gov/24656047/)
26. Finco D, Grimaldi C, Fort M, Walker M, Kiessling A, Wolf B, et al. Cytokine release assays: current practices and future directions. *Cytokine*. 2014;66(2):143–55. <https://doi.org/10.1016/j.cyto.2013.12.009> PMID: [24412476](https://pubmed.ncbi.nlm.nih.gov/24412476/)
27. Rowley TF, Peters SJ, Aylott M, Griffin R, Davies NL, Healy LJ, et al. Engineered hexavalent Fc proteins with enhanced Fc-gamma receptor avidity provide insights into immune-complex interactions. *Commun Biol*. 2018;1:146. <https://doi.org/10.1038/s42003-018-0149-9> PMID: [30272022](https://pubmed.ncbi.nlm.nih.gov/30272022/)
28. Man A, Orasan M, Hoteiuc O, Olanescu-Vaida-Voevod M, Mocan T. Inflammation and psoriasis: A comprehensive review. *International Journal of Molecular Sciences*. n.d.;24(22):1–10. <https://doi.org/10.3390/ijms24221234>
29. Khouj E, Marafi D, Aljamal B, Hajiya A, Elshafie RM, Hashem MO, et al. Human “knockouts” of CSF3 display severe congenital neutropenia. *Br J Haematol*. 2023;203(3):477–80. <https://doi.org/10.1111/bjh.19054> PMID: [37612131](https://pubmed.ncbi.nlm.nih.gov/37612131/)
30. Vignesh P, Rawat A, Kumar A, Suri D, Gupta A, Lau YL, et al. Chronic Granulomatous Disease Due to Neutrophil Cytosolic Factor (NCF2) Gene Mutations in Three Unrelated Families. *J Clin Immunol*. 2017;37(2):109–12. <https://doi.org/10.1007/s10875-016-0366-2> PMID: [28035544](https://pubmed.ncbi.nlm.nih.gov/28035544/)
31. Zhang J, Mai S, Chen H-M, Kang K, Li XC, Chen S-H, et al. Leukocyte immunoglobulin-like receptors in human diseases: an overview of their distribution, function, and potential application for immunotherapies. *J Leukoc Biol*. 2017;102(2):351–60. <https://doi.org/10.1189/jlb.5MR1216-534R> PMID: [28351852](https://pubmed.ncbi.nlm.nih.gov/28351852/)
32. Marigorta U, Denson L, Hyams J, Mondal K, Prince J, Walters T, et al. Transcriptional risk scores link GWAS to eQTLs and predict complications in Crohn’s disease. *Nature Genetics*. 2017;49(10):1517–21.
33. de Lange KM, Moutsianas L, Lee JC, Lamb CA, Luo Y, Kennedy NA, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet*. 2017;49(2):256–61. <https://doi.org/10.1038/ng.3760> PMID: [28067908](https://pubmed.ncbi.nlm.nih.gov/28067908/)
34. Ishigaki K, Sakae S, Terao C, Luo Y, Sonehara K, Yamaguchi K, et al. Multi-ancestry genome-wide association analyses identify novel genetic mechanisms in rheumatoid arthritis. *Nature Genetics*. 2022;54(11):1640–51. <https://doi.org/10.1038/s41588-022-01045-5>
35. Initiative C-H. A second update on mapping the human genetic architecture of COVID-19. *Nature*. 2023;621(7977):E7–26.
36. Mantovani S, Daga S, Fallerini C, Baldassarri M, Benetti E, Picchiotti N, et al. Rare variants in Toll-like receptor 7 results in functional impairment and downregulation of cytokine-mediated signaling in COVID-19 patients. *Genes and Immunity*. 2022;23(1):51–6.
37. Smyth P, Sasiwachirangkul J, Williams R, Scott CJ. Cathepsin S (CTSS) activity in health and disease - A treasure trove of untapped clinical potential. *Mol Aspects Med*. 2022;88:101106. <https://doi.org/10.1016/j.mam.2022.101106> PMID: [35868042](https://pubmed.ncbi.nlm.nih.gov/35868042/)
38. Saad MN, Mabrouk MS, Eldeib AM, Shaker OG. Studying the effects of haplotype partitioning methods on the RA-associated genomic results from the North American Rheumatoid Arthritis Consortium (NARAC) dataset. *J Adv Res*. 2019;18:113–26. <https://doi.org/10.1016/j.jare.2019.01.006> PMID: [30891314](https://pubmed.ncbi.nlm.nih.gov/30891314/)
39. Franke A, McGovern D, Barrett J, Wang K, Radford-Smith G, Ahmad T, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn’s disease susceptibility loci. *Nature Genetics*. 2010;42(12):1118–25. <https://doi.org/10.1038/ng.2010.155>
40. Luo S, Li X-F, Yang Y-L, Song B, Wu S, Niu X-N, et al. PLCL1 regulates fibroblast-like synoviocytes inflammation via NLRP3 inflammasomes in rheumatoid arthritis. *Adv Rheumatol*. 2022;62(1):25. <https://doi.org/10.1186/s42358-022-00252-5> PMID: [35820936](https://pubmed.ncbi.nlm.nih.gov/35820936/)
41. Pairo-Castineira E, Rawlik K, Bretherick AD, Qi T, Wu Y, Nassiri I, et al. GWAS and meta-analysis identifies 49 genetic variants underlying critical COVID-19. *Nature*. 2023;617(7962):764–8. <https://doi.org/10.1038/s41586-023-06034-3> PMID: [37198478](https://pubmed.ncbi.nlm.nih.gov/37198478/)
42. Brodin P, Davis MM. Human immune system variation. *Nat Rev Immunol*. 2017;17(1):21–9. <https://doi.org/10.1038/nri.2016.125> PMID: [27916977](https://pubmed.ncbi.nlm.nih.gov/27916977/)
43. Li Y, Oosting M, Deelen P, Ricano-Ponce I, Smeekens S, Jaeger M, et al. Inter-individual variability and genetic influences on cytokine responses to bacteria and fungi. *Nature Medicine*. 2016;22(8):952–60.
44. Thomas S, Rouilly V, Patin E, Alanio C, Dubois A, Delval C, et al. The Milieu Intérieur study - an integrative approach for study of human immunological variance. *Clin Immunol*. 2015;157(2):277–93. <https://doi.org/10.1016/j.clim.2014.12.004> PMID: [25562703](https://pubmed.ncbi.nlm.nih.gov/25562703/)
45. Muller S, Kroger C, Schultze J, Aschenbrenner A. Whole blood stimulation as a tool for studying the human immune system. *Eur J Immunol*. 2024;54(2):e2350519.
46. Chandra V, Bhattacharyya S, Schmiedel BJ, Madrigal A, Gonzalez-Colin C, Fotsing S, et al. Promoter-interacting expression quantitative trait loci are enriched for functional genetic variants. *Nat Genet*. 2021;53(1):110–9. <https://doi.org/10.1038/s41588-020-00745-3> PMID: [33349701](https://pubmed.ncbi.nlm.nih.gov/33349701/)
47. Novakovic B, Habibi E, Wang S-Y, Arts RJW, Davar R, Megchelenbrink W, et al. β-Glucan Reverses the Epigenetic State of LPS-Induced Immunological Tolerance. *Cell*. 2016;167(5):1354–1368.e14. <https://doi.org/10.1016/j.cell.2016.09.034> PMID: [27863248](https://pubmed.ncbi.nlm.nih.gov/27863248/)
48. Jeong R, Bulyk M. Chromatin accessibility variation provides insights into missing regulation underlying immune-mediated diseases. *eLife*. 2024.
49. Wu Y, Zeng J, Zhang F, Zhu Z, Qi T, Zheng Z, et al. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nature Communications*. n.d.;9(1):918. <https://doi.org/10.1038/s41467-018-03210-4>
50. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet*. 2014;10(5):e1004383. <https://doi.org/10.1371/journal.pgen.1004383> PMID: [24830394](https://pubmed.ncbi.nlm.nih.gov/24830394/)

51. Wallace C. Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS Genet.* 2020;16(4):e1008720. <https://doi.org/10.1371/journal.pgen.1008720> PMID: [32310995](https://pubmed.ncbi.nlm.nih.gov/32310995/)
52. Wallace C. A more accurate method for colocalisation analysis allowing for multiple causal variants. *PLoS Genet.* 2021;17(9):e1009440. <https://doi.org/10.1371/journal.pgen.1009440> PMID: [34587156](https://pubmed.ncbi.nlm.nih.gov/34587156/)
53. Hormozdiari F, van de Bunt M, Segre AV, Li X, Joo JWW, Bilow M, et al. Colocalization of GWAS and eQTL Signals Detects Target Genes. *Am J Hum Genet.* 2016;99(6):1245-60.
54. Wang G, Sarkar A, Carbonetto P, Stephens M. A simple new approach to variable selection in regression, with application to genetic fine mapping. *J R Stat Soc Series B Stat Methodol.* 2020;82(5):1273–300. <https://doi.org/10.1111/rssb.12388> PMID: [37220626](https://pubmed.ncbi.nlm.nih.gov/37220626/)
55. Vuckovic D, Bao EL, Akbari P, Lareau CA, Mousas A, Jiang T, et al. The polygenic and monogenic basis of blood traits and diseases. *Cell.* 2020;182(5):1214–31.
56. Delaneau O, Ongen H, Brown AA, Fort A, Panousis NI, Dermizakis ET. A complete tool set for molecular QTL discovery and analysis. *Nat Commun.* 2017;8:15452. <https://doi.org/10.1038/ncomms15452> PMID: [28516912](https://pubmed.ncbi.nlm.nih.gov/28516912/)
57. Mu Z, Wei W, Fair B, Miao J, Zhu P, Li Yi. The impact of cell type and context-dependent regulatory variants on human immune traits. *Genome Biol.* 2021;22(1):122. <https://doi.org/10.1186/s13059-021-02334-x> PMID: [33926512](https://pubmed.ncbi.nlm.nih.gov/33926512/)
58. Hu S, Uniken Venema WT, Westra H-J, Vich Vila A, Barbieri R, Voskuil MD, et al. Inflammation status modulates the effect of host genetic variation on intestinal gene expression in inflammatory bowel disease. *Nat Commun.* 2021;12(1):1122. <https://doi.org/10.1038/s41467-021-21458-z> PMID: [33602935](https://pubmed.ncbi.nlm.nih.gov/33602935/)
59. Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J. Impact of genetic polymorphisms on human immune cell gene expression. *Cell.* n.d.;175(6):1701–15.
60. Momozawa Y, Dmitrieva J, Théâtre E, Deffontaine V, Rahmouni S, Charleatoux B, et al. IBD risk loci are enriched in multigenic regulatory modules encompassing putative causative genes. *Nat Commun.* 2018;9(1):2427. <https://doi.org/10.1038/s41467-018-04365-8> PMID: [29930244](https://pubmed.ncbi.nlm.nih.gov/29930244/)
61. Ni J, Wang P, Yin KJ, Yang XK, Cen H, Sui C, et al. Novel insight into the aetiology of rheumatoid arthritis gained by a cross-tissue transcriptome-wide association study. *RMD Open.* 2022;8(2).
62. Ota M, Nagafuchi Y, Hatano H, Ishigaki K, Terao C, Takeshima Y, et al. Dynamic landscape of immune cell-specific gene regulation in immune-mediated diseases. *Cell.* 2021;184(11):3006-3021.e17. <https://doi.org/10.1016/j.cell.2021.03.056> PMID: [33930287](https://pubmed.ncbi.nlm.nih.gov/33930287/)
63. Schmiedel BJ, Rocha J, Gonzalez-Colin C, Bhattacharyya S, Madrigal A, Ottensmeier CH, et al. COVID-19 genetic risk variants are associated with expression of multiple genes in diverse immune cell types. *Nat Commun.* 2021;12(1):6760. <https://doi.org/10.1038/s41467-021-26888-3> PMID: [34799557](https://pubmed.ncbi.nlm.nih.gov/34799557/)
64. Zhou S, Butler-Laporte G, Nakanishi T, Morrison DR, Afilalo J, Afilalo M, et al. A Neanderthal OAS1 isoform protects individuals of European ancestry against COVID-19 susceptibility and severity. *Nat Med.* 2021;27(4):659–67. <https://doi.org/10.1038/s41591-021-01281-1> PMID: [33633408](https://pubmed.ncbi.nlm.nih.gov/33633408/)