Decoding the genetic landscape in granulomatous diseases: pathways to improved diagnosis and therapy

- Type of fellowship: Short Term Research Fellowship (3 months)
- Home institution: University of Padova (Padova, Italy)
- Home institution supervisor: Prof. Francesco Cinetto
- Host institution: Erasmus University Medical Center (Rotterdam, The Netherlands)
- Host institution supervisor: Prof. Virgil A. S. H. Dalm
- Project co-supervisor: Prof. Peter J. van der Spek
- Period covered: 1st April – 1st July 2025

Introduction: sarcoidosis and GLILD as challenging rare granulomatous diseases

Sarcoidosis and Granulomatous Lymphocytic Interstitial Lung Diseases (GLILD) represent two rare and clinically variable disorders primarily marked by the development of Interstitial Lung Disease (ILD), subjected to systemic immune dysregulation. Of note, these two conditions present overlapping aspects of their immunological background and pathological features, most notably the presence of granulomas as a defining common characteristic [1], [2].

They both represent challenging cases in internal medicine due to the lack of good-quality and evidence-based treatment guidelines, in addition to standardized management protocols. Given their similar traits, sarcoidosis and GLILD are often studied in parallel in an effort to uncover new insights. However, many questions are still open on their etiology and pathogenesis, thus representing a critical clinical unmet need and, most importantly, a highly emerging scientific quest [2].

On one hand, GLILD can be regarded as one of the most severe non-infectious complications in 15-20% of patients suffering from Common Variable Immunodeficiency (CVID), the most common symptomatic primary immunodeficiency in adults [1], [3], [4]. Normally, granulomas, as cellular aggregates which can potentially arise in every district of the body, form as a

response to the presence of hard-to-clear antigens; however, in this particular case there is still no evidence of the main responsible trigger [1], [2].

At the histological level, GLILD is characterized by the presence of peculiar granulomas, similar to those who can be observed in sarcoidosis and/or polyclonal lymphocytic infiltrates in the lungs, often showing extrapulmonary involvement (e.g., lymph nodes, liver, spleen), as in sarcoidosis [2], [3]. As no standard first-line therapeutic regimens are available, treatment strategies are mainly based on case series [5].

Sarcoidosis, in contrast, is classified as a granulomatous disease of unknown etiology which occurs in immunocompetent individuals; it often presents with hilar adenopathy and interstitial lung disease, but it can potentially involve any organ or tissue [6]. The most up-to-date therapeutic guidelines are based on low-quality evidence. In addition, while sarcoidosis has been studied for decades, GLILD, in contrast, represents a relatively recent rare entity with a worse impact on patient overall survival and quality of life when compared to sarcoidosis [5].

State of the art

The immune component's impact has been investigated from multiple perspectives: it is well known that Th17, Th17.1 and regulatory T (Treg) lymphocytes play a fundamental role in the recruitment and activation of macrophages in sarcoidosis; however, it is still unclear whether their action might have a protective or enhancing effect on the pathological process [7]. Additionally, CVID patients with ILD complications have shown elevated serum levels of sCD25, sTIM-3, IFN-γ and TNF, indicating a strong activation and subsequent exhaustion of T lymphocytes [8]. B lymphocytes have only recently been implicated in the pathogenesis of sarcoidosis, with a still yet uncertain role, as the effectiveness of rituximab, an anti-CD20 monoclonal antibody, has been reported in some cases of refractory sarcoidosis. A significant role of B cells has also been highlighted in GLILD: patients with X-linked agammaglobulinemia, lacking mature B cells, do not develop GLILD, while granuloma formation has been reported in monogenic forms of CVID associated to a deregulation of B cell regulatory signaling pathways, with a consistent increase in the $CD21^{low}$ B cell subpopulation. In this regard, a significant deregulation of the PI3K-mTOR signaling pathway, involving the activation of FOXO1 and HIF-1α, has been further observed in naïve B cells of CVID patients [9].

Recent insights have drawn attention to the subset of $CD21^{low}$ B cells, which are increasingly recognized as key players in the pathogenesis of immune dysregulation in CVID and, more specifically, in GLILD [10], [11]. These cells exhibit a distinctive anergic phenotype characterized by low expression of CD21, increased autoreactivity, and impaired antigen-induced proliferation. In CVID patients with GLILD, this subset is markedly expanded, and it correlates with disease severity, whereas its presence is also becoming significant in sarcoid lesions [12], [13], [14], [15], [16].

In sarcoidosis, moreover, the JAK/STAT and mTORC1 signaling pathways have been reported as of particular therapeutic interest; they role has recently been investigated by using *in vivo* and *in vitro* models [8], [9]. RNA sequencing analyses have highlighted significantly elevated expression of downstream effectors, such as STAT1 and STAT3 in cutaneous sarcoidosis biopsies, compared to healthy controls [17], [18]. Additionally, STAT1 and STAT3 are related to Th1 and Th17 differentiation, while the treatment with tofacitinib, a JAK inhibitor, has promoted remission of cutaneous lesions sarcoidosis patients [17]. Recently, thanks to major histological and spatial proteomics analyses, van Stigt and colleagues highlighted that, although these two diseases share specific immune mechanisms and cellular players, CVID granulomas appear to be less organized and characterized by a higher influx of CD20+ B cells. In this regard, CD163+ macrophages, neutrophils and B cells have shown to be essential for granuloma maintenance, thus unveiling a unique biological and metabolic signature. Moreover, tissue analyses alone have shown to be particularly helpful in distinguishing CVID from sarcoidosis based on histological sections [19].

To further validate these observations, we aim to investigate the interacting network and transcriptional profile of $CD21^{low}$ B cells from publicly available databases in the final perspective of observing their presence, distribution and interactions with other cells in lung biopsies coming from patients with GLILD and sarcoidosis. By integrating preliminary GWAS and bioinformatics data with spatial and single-cell transcriptomic approaches, we hope to delineate the specific immune landscape of these conditions and better understand the role of aberrant B cell responses in granulomatous inflammation, unveiling potential targets for novel mechanism-based therapies.

Aims of this project

As previously reported, several genetic alterations may be involved in immune dysregulation and in the pathogenesis of sarcoidosis and GLILD, as in the development of granulomatous lesions and disease progression. In order to gain novel insights into what lies behind granuloma formation in rare granulomatous diseases, the following goals were addressed:

1. Performing Genome-Wide Association Studies (GWAS) on sarcoidosis and GLILD databases to identify genetic susceptibility and discriminate overlapping/divergent signatures.
2. Conducting preliminary bioinformatic analyses for further spatial transcriptomics experiments on granuloma tissue samples coming from each mentioned condition to further define the gene expression landscape.
3. Investigate the potential immunopathogenic role of $CD21^{low}$ B cells in sarcoidosis and GLILD.

Preliminary remarks

In the initial version of this project, it was also planned to investigate the role of T cells in these pathological contexts. However, due to time constraints, training period and the exploratory nature of this first phase, it was decided to only focus on the B cell compartment, particularly the $CD21^{low}$ subset, to develop a more targeted and feasible approach.

Bioinformatics tools training and set up

During the first month of the research stay, a significant portion of the time was dedicated to gaining familiarity with a range of essential bioinformatics tools and databases, which served as a foundation for all subsequent analyses:

- GeneCards [a]: a comprehensive database which provides information about all human genes, including gene function, expression patterns, diseases and related pathways.
- Human Protein Atlas [b]: an open-access resource for protein expression data in several human tissues and cell types, which can be used to investigate tissue-specific protein localization.
- STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) [c]: an open-access database of known and predicted protein-protein interactions, which has been employed to explore functional relationships between proteins of interest.
- NCBI GEO (Gene Expression Omnibus) [d]: a public repository for gene expression and other functional genomics datasets. This tool has been essential for retrieving all transcriptomic data used in hereby reported analyses.
- HGMD (Human Gene Mutation Database, Qiagen®) [e]: a curated collection of published gene mutations associated with inherited human diseases employed to investigate known variants linked to immune dysregulation.
- GWAS Catalog (Genome-Wide Association Studies Catalog) [f]: a publicly available, manually curated resource which reports published genome-wide association studies (GWAS) reporting associations between genetic variants and human traits or diseases. It is supported and maintained by the NHGRI and EMBL-EBI, and includes standardized information related to SNP-associations, mapped traits by using ontologies such as MONDO, and study metadata.
- WebCSEA [g]: an online tool which performs cell type-specific enrichment analysis by linking gene lists to single-cell transcriptomic datasets, helping to identify which cell populations predominantly express genes of interest.

In addition to database querying, specific training was also provided for the use of:

- Omniviz™: a software tool, available for use at the department of Bioinformatics at Erasmus MC, for multidimensional data visualization and pattern recognition, which supported the exploratory analysis of gene expression data.
- Ingenuity Pathway Analysis (IPA, Qiagen®): a commercial bioinformatics platform designed to analyze omics data within the context of biological pathways, upstream regulators and causal networks. This tool was particularly useful in identifying common signaling pathways and potential regulatory molecules shared between sarcoidosis and CVID-related granulomatous disease.

Section 1. Comparative GWAS analysis between sarcoidosis and CVID patients integrated with HGMD investigation of mutations

Introduction

As previously reported, GLILD represents a rare and poorly understood pulmonary complication of CVID. It is characterized by a heterogeneous histological pattern involving granulomas, lymphoid hyperplasia and interstitial inflammation. Due to its rarity and diagnostic complexity, no GWAS datasets are currently available specifically for this condition.

A GWAS data analysis was conducted as an initial exploratory step to examine potential contributors and shared susceptibility loci, with a focus on sarcoidosis and CVID.

Aims

By exploring publicly available GWAS datasets for these two diseases, the main aim was to:

1. Explore genetic variants potentially associated with granuloma formation in both immunocompetent (Sarcoidosis) and immunocompromised (CVID+GLILD) individuals.
2. Explore shared genes and molecular pathways between the two conditions using a multi-step bioinformatic approach.
3. Lay the foundation for future analysis when GLILD-specific genetic data will become available.

Materials and methods

Data sources

The following disease traits were selected from the GWAS Catalog for variant extraction:

- Sarcoidosis (GWAS Catalog trait, MONDO:0019338; **Figure 1**)
- Common Variable Immunodeficiency (GWAS Catalog trait, MONDO:0015517; **Figure 2**)

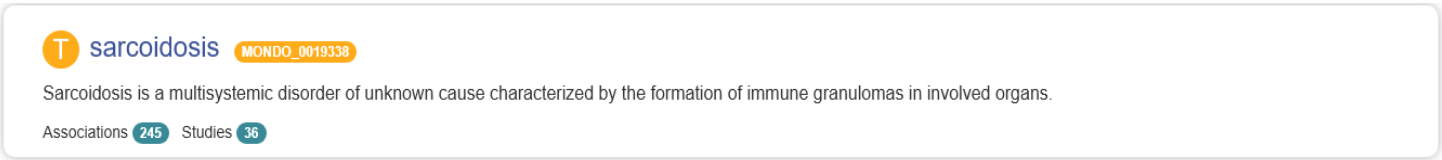


Figure 1. Sarcoidosis GWAS catalog entry.

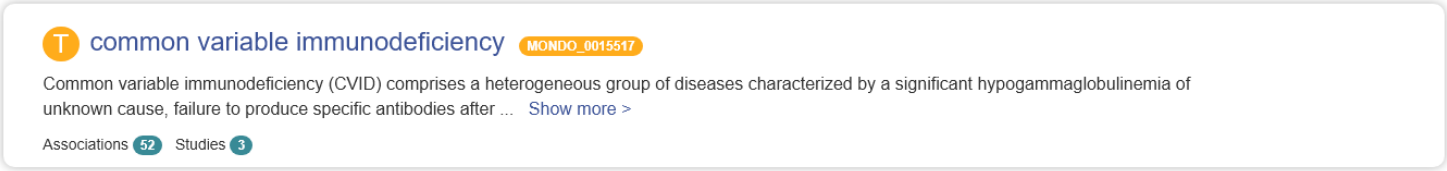


Figure 2. Common variable immunodeficiency (CVID) GWAS catalog entry.

The following gene expression datasets were selected from Genome Expression Omnibus (GEO):

- **GSE157671** (lung biopsies: sarcoidosis vs. healthy controls; **Figure 3**)

Series GSE157671	Query DataSets for GSE157671
Status	Public on Dec 09, 2020
Title	Differential transcriptomics in sarcoidosis lung and lymph node granulomas with comparisons to pathogen-specific granulomas
Organism	Homo sapiens
Experiment type	Expression profiling by high throughput sequencing
Summary	Rationale: Despite the availability of multi-"omics" strategies, insights into the etiology and pathogenesis of sarcoidosis have been elusive. This is partly due to the lack of reliable preclinical models and a paucity of validated biomarkers. As granulomas are a key feature of sarcoidosis, we speculate that direct genomic interrogation of sarcoid tissues, may lead to identification of dysregulated gene pathways or biomarker signatures. Objective: To facilitate the development sarcoidosis genomic biomarkers by gene expression profiling of sarcoidosis granulomas in lung and lymph node tissues (most commonly affected organs) and comparison to infectious granulomas (coccidiodomycosis and tuberculosis).
Overall design	Transcriptomic profiles of immune-related gene from micro-dissected lungs and mediastinal lymph nodes sarcoidosis granulomas was compared to infectious granulomas. Differentially-expressed genes (DEGs) were profiled, compared among the three granulomatous diseases and analyzed for functional enrichment pathways. Grant ID: CDA22 Grant title: University of Arizona Career Development Funding source: The University of Arizona Grantee: Nancy Casanova Grant ID: NA Grant title: Sarcoidosis Mini-grant Funding source: Foundation for Sarcoidosis Research Grantee: Nancy Casanova
Contributor(s)	Casanova NG, Gonzalez-Garay ML, Sun B, Bime C, Knox KS, Crouser E, Sun X, Sammani N, Gregory T, Natt B, Chaudhary S, Lussier Y, Garcia J
Citation(s)	Casanova NG, Gonzalez-Garay ML, Sun B, Bime C et al. Differential transcriptomics in sarcoidosis lung and lymph node granulomas with comparisons to pathogen-specific granulomas. *Respir Res* 2020 Dec 4;21(1):321. PMID: 33276795 Casanova NG, Reyes-Hernon V, Gregory T, Sun B et al. Biochemical and genomic identification of novel biomarkers in progressive sarcoidosis: HBEGF, eNAMPT, and ANG-2. *Front Med (Lausanne)* 2022;9:1012827. PMID: 36388923

Figure 3. GSE157671 dataset description on GEO.

- **GSE16538** (lung biopsies: active pulmonary sarcoidosis vs. healthy controls; **Figure 4**)

Series GSE16538	Query DataSets for GSE16538
Status	Public on Jun 11, 2009
Title	Genome-wide gene expression profile analysis in pulmonary sarcoidosis
Organism	Homo sapiens
Experiment type	Expression profiling by array
Summary	We hypothesized that tissue genome-wide gene expression analysis, coupled with gene network analyses of differentially expressed genes, would provide novel insights into the pathogenesis of pulmonary sarcoidosis. Keywords: Disease state analysis
Overall design	Genome-wide gene expression profiles were compared in tissues derived from subjects with active pulmonary sarcoidosis (n=6) and those with normal lung anatomy (n=6). Differentially expressed genes were analyzed by gene network analysis
Contributor(s)	Crouser ED, Culver DA, Knox KS, Julian MW, Shao G, Abraham S, Liyanarachchi S, Macre JE, Wewers MD, Gavrilin MA, Ross P, Abbas A, Eng C
Citation(s)	Crouser ED, Culver DA, Knox KS, Julian MW et al. Gene expression profiling identifies MMP-12 and ADAMDEC1 as potential pathogenic mediators of pulmonary sarcoidosis. *Am J Respir Crit Care Med* 2009 May 15;179(10):929-38. PMID: 19218196

Figure 4. GSE16538 dataset description on GEO.

- **GSE51406**: this publicly available dataset (**Figure 5**) was reprocessed with the support of Sigrid Swagemakers (Bioinformatics Department, Erasmus MC) to facilitate stratified analysis of CVID patient blood samples. Samples were categorized based on the presence or absence of clinical complications, specifically focusing on those exhibiting granulomatous manifestations. The subset of granulomatous CVID samples was isolated, and the dataset was reformatted and uploaded into Ingenuity Pathway Analysis (IPA) for further comparative analyses between granulomatous and non-granulomatous CVID phenotypes.

Series GSE51406	Query DataSets for GSE51406
Status	Public on Jan 09, 2014
Title	Interferon Signature in the Blood in Inflammatory Common Variable Immune Deficiency
Organism	Homo sapiens
Experiment type	Expression profiling by array
Summary	This SuperSeries is composed of the SubSeries listed below.
Overall design	Refer to individual Series
Citation(s)	Park J, Munagala I, Xu H, Blankenship D et al. Interferon signature in the blood in inflammatory common variable immune deficiency. *PLoS One* 2013;8(9):e74893. PMID: 24069364

Figure 5. GSE51406 dataset description on GEO.

The GSE51406 dataset was reformatted as follows:

1. Matrix data were first downloaded from GEO.
2. Matrix data were analyzed using BeadStudio (v. 1.5.1.3) by applying the average chip normalization method.
3. The normalized expression matrix was then imported into Omniviz™ for exploratory data analysis, including Significance Analysis of Microarrays (SAM) to identify differentially expressed genes. Selection criteria were set to a minimum fold change of 1.5 and a false discovery rate (FDR) threshold of 5%. The analysis was conducted in order to compare CVID samples with granulomatous complications vs. CVID samples without granulomatous complications.
4. The resulting data were further analyzed using Ingenuity Pathway Analysis (IPA) for functional analyses.

Bioinformatics tools

- **HGMD** (QIAGEN® Human Gene Mutation Database): for identifying genetic mutations linked to sarcoidosis and CVID.
- **IPA** (QIAGEN® Ingenuity Pathway Analysis): for upstream regulator and causal network identification.
- **STRING**: for protein-protein interaction analysis.
- **Microsoft Excel**: for manual comparison of gene lists.

Analytical workflow

1. GWAS Catalog Comparison: initially, we retrieved all genome-wide significant variants reported in the NHGRI-EBI GWAS Catalog for sarcoidosis and for CVID. A direct comparison of these lead variants showed that none was shared between the two conditions.
2. Shared gene analysis: we repeated the analysis at the gene level. Each lead variant was linked to the gene(s) annotated in the GWAS catalog, and the resulting gene lists were

intersected. This second step revealed a small set of genes which are common to both diseases, summarized in the Excel table reported in **Figure 6**.

	A	B	C	D
1	shared genes	*CVID*	Sarcoidosis	
24	C1orf141	C1orf141	C1orf141	
80	HLA-DQB1	HLA-DQB1	HLA-DQB1	
96	IL12B	IL12B	IL12B	
99	IL23R	IL23R	IL23R	
127	LINC01845	LINC01845	LINC01845	

Figure 6. Excel sheet table reporting disease-associated genes shared between CVID and sarcoidosis.

3. STRING analysis: the gene list reported in **Figure 6** was uploaded into STRING database to highlight interesting protein-protein interaction networks. Results are reported in **Figure 9A**.
4. Upstream regulator analysis (IPA): Upstream Regulator Analysis was conducted using IPA software to identify potential transcriptional regulators and, thus, interesting, related pathways. The input data consisted of the list of gene found to be shared between sarcoidosis and CVID reported in **Figure 6**. Each gene from the list was queried into the investigated databases uploaded in IPA, where the software exploits its curated knowledge base to predict upstream molecules, such as transcription factors, cytokines and other regulators, that could explain such genetic patterns. The analysis incorporated both activation and inhibition predictions based on the expression directionality and network relationships, providing insights into key regulatory mechanisms. In addition, we applied Causal Network Analysis within IPA to extend the upstream regulator analysis beyond direct interactions. This algorithm identifies higher order regulatory networks by integrating known cause-effect relationships from literature, allowing the identification of master regulators that act indirectly through intermediary molecules. The activation state of each regulator was predicted based on the direction of gene expression changes, and is reported using IPA's z-score algorithm, which infers activation or inhibition status.

Integration with rare variant data (HGMD)

In parallel to GWAS-based analyses, we explored rare pathogenic variants associated with sarcoidosis and CVID by querying the Human Gene Mutation Database (HGMD). This complementary approach aimed to capture additional insights into genetic contributions (specifically rare, high impact variants) which are typically not detected by GWAS due to their low allele frequency in the general population.

For CVID, we searched the HGMD using the terms “Common variable immune deficiency” and “Common variable immunodeficiency”, which yielded the following set of associated genes (**Figure 7**):

A

results for search term *'Common variable immune deficiency'*

1 to 1 of 1 results

Common variable immune deficiency	TNFRSF13B NFKB2 LRBA	3 mutations

Designed by P.D.Stenson HGMD®

Copyright © Cardiff University 2024

B

results for search term *'Common variable immunodeficiency'*

1 to 10 of 10 results

Common variable immunodeficiency	NFKB1 LRBA CD3G GTF3A STAT3 TNFRSF13B IFNK NFKB2 IRF2BP2	14 mutations
Common variable immunodeficiency 2	TNFRSF13B	1 mutation
Common variable immunodeficiency 8 with autoimmunity	LRBA	1 mutation
Common variable immunodeficiency with autoimmune cytopaenia	CTNNBL1	1 mutation
Common variable immunodeficiency with impaired T-cell function, association with	CCDC28B	1 mutation
Common variable immunodeficiency with loss of B cells	IKZF1	6 mutations
Common variable immunodeficiency, autoimmune haemolytic anaemia & pancytopenia	IKZF1	1 mutation
Common variable immunodeficiency, cytopaenias with splenomegaly & nodular regenerative hyperplasia of the liver	NFKB1	1 mutation
Common variable immunodeficiency/IgG subclass deficiency, association with	TNFRSF13B	1 mutation
Gastrointestinal issues in common variable immunodeficiency	NOD2	1 mutation

Figure 7. Human Gene Mutation Database (HGMD) rare pathogenic variants query for (A) “Common variable immune deficiency” and (B) “Common variable immunodeficiency”.

- *TNFRSF13B*
- *NFKB1*
- *NFKB2*
- *LRBA*
- *CD3G*
- *GTF3A*
- *STAT3*
- *IFNK*
- *IRF2BP2*
- *CTNNBL1*
- *CCDC28B*

- *IKZF1*
- *NOD2*

By searching for the term “sarcoidosis” in the HGMD database, the following genes were retrieved (**Figure 8**):

results for search term *'sarcoidosis'*

1 to 17 of 17 results

Cardiac sarcoidosis	FLNC	1 mutation
Disease progression in sarcoidosis, association with	NOD2	1 mutation
Erythema nodosum in sarcoidosis	TNF	1 mutation
Erythema nodosum in sarcoidosis females	LTA	1 mutation
Neurosarcoidosis	RNASEH2A	1 mutation
Ocular sarcoidosis, association with	CFH	1 mutation
Pulmonary sarcoidosis	ZC3H12A	1 mutation
Sarcoidosis	NOD2 CLEC7A EPHA2 NCF2 KALRN MADD ANXA11 IL17RA FAT1	10 mutations
Sarcoidosis-related uveitis, association with	RAB23	1 mutation
Sarcoidosis, association with	BTN3A2 BTNL2 MIF MUC22 FCGR1A NFKBIA TLR7 CCHCR1 IL7R IL23R SLC11A1 NOTCH4	16 mutations
Sarcoidosis, early onset	NOD2	10 mutations
Sarcoidosis, early-onset	NOD2	2 mutations
Sarcoidosis, increased risk, association with	IFNA10 IFNA17	2 mutations
Sarcoidosis, modifer of	FCGR3A	1 mutation
Sarcoidosis, protection, association	ANXA11	1 mutation
Sarcoidosis, susceptibility to	CFTR	2 mutations
Sarcoidosis, susceptibility, association with	HLA-DRB1	1 mutation

Figure 8. Human Gene Mutation Database (HGMD) rare pathogenic variants query for “sarcoidosis”.

- *FLNC*
- *NOD2*
- *TNF*
- *LTA*
- *RNASEH2A*
- *CFH*
- *ZC3H12A*
- *CLEC7A*
- *EPHA2*
- *NCF2*
- *KALRN*
- *MADD*
- *ANXA11*
- *IL17RA*
- *FAT1*
- *RAB23*

- *BTN3A2*
- *BTNL2*
- *MIF*
- *MUC22*
- *FCGR1A*
- *NFKBIA*
- *TLR7*
- *CCHCR1*
- *IL7R*
- *IL23R*
- *SLC11A1*
- *NOTCH4*
- *IFNA10*
- *IFNA17*
- *FCGR3A*
- *ANXA11*
- *CFTR*
- *HLA-DRB1*

Results

GWAS analysis between sarcoidosis and CVID patients

No genome-wide significant SNPs were found to be shared between the available GWAS datasets for sarcoidosis and CVID. While this finding points to potentially disease-specific risk patterns, we cannot rule out shared genetic influences that fall below current significance thresholds, differ across populations, or converge at the pathway, rather than single-variant level.

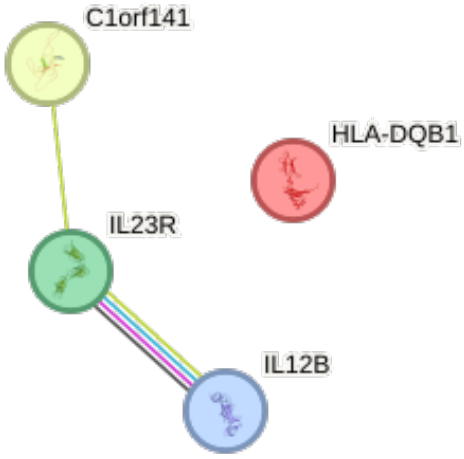
To expand the initial GWAS-based comparison, we focused on a gene-level approach, in order to identify genes which carry or are located near disease-associated SNPs in each dataset. This analysis revealed a small subset of shared genes:

- *C1orf141*
- *HLA-DBQ1*
- *IL12B*
- *IL23R*
- *LINC01845*

While these genes are not linked to the same SNPs or allelic variants, their recurrent association across two distinct disease contexts suggests potential converging immunopathogenic mechanisms that might be related to granuloma formation.

The STRING network analysis (**Figure 9A**) showed that IL23R and IL12B are strongly functionally linked. This association is supported by multiple evidence sources (four edges) in the STRING database, suggesting a well-established and biologically relevant interaction, likely reflecting their shared involvement in IL-12/IL-23-mediated immune signaling. On the other hand, the connection between C1orf and IL23R is only supported by a single source of evidence (single edge), suggesting a weaker or less characterized association which may require further investigation, or that could be related to a limited functional relevance. Moreover, HLA-DQB1 appeared as an unconnected node indicating no direct associations within the network. This may reflect a more general role for HLA-DQB1 in antigen presentation, rather than a direct involvement in the same signaling axis.

A



B

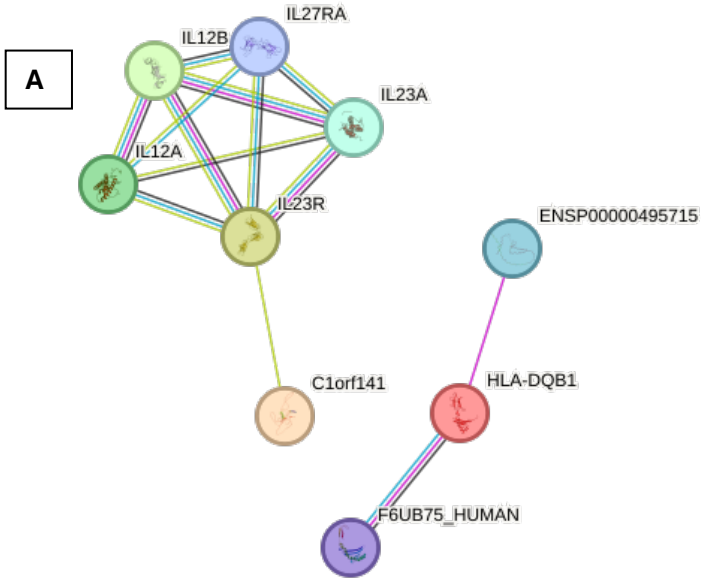
Biological Process (Gene Ontology)					
GO-term	*description*	*count in network*	*strength*	*signal*	*false discovery rate*
GO:2000330	Positive regulation of T-helper 17 cell lineage commitment	2 of 4	3.39	1.67	0.0073
GO:0051142	Positive regulation of NK T cell proliferation	2 of 10	2.99	1.6	0.0087
GO:0043382	Positive regulation of memory T cell differentiation	2 of 9	3.04	1.6	0.0087
GO:0010536	Positive regulation of activation of Janus kinase activity	2 of 7	3.15	1.6	0.0087
GO:0032819	Positive regulation of natural killer cell proliferation	2 of 13	2.88	1.59	0.0087

KEGG Pathways					
pathway	*description*	*count in network*	*strength*	*signal*	*false discovery rate*
hsa05321	Inflammatory bowel disease	2 of 59	2.22	1.28	0.0189

Figure 9. (A) STRING network analysis of shared disease associated genes based on the comparison between Sarcoidosis and CVID. (B) Functional enrichments in the resulting network.

The first shell expansion of this network revealed a broader enriched network with additional proteins and related pathways of interest (**Figure 10A**).

A



B

Biological Process (Gene Ontology)

GO-term	description	count in network	strength	signal	false discovery rate
GO:0051142	Positive regulation of NK T cell proliferation	4 of 10	2.94	4.21	5.24e-08
GO:0032819	Positive regulation of natural killer cell proliferation	4 of 13	2.83	4.13	6.22e-08
GO:0032660	Regulation of interleukin-17 production	5 of 41	2.43	4.02	3.62e-08
GO:0002827	Positive regulation of T-helper 1 type immune response	4 of 19	2.66	3.84	1.63e-07
GO:2000330	Positive regulation of T-helper 17 cell lineage commitment	3 of 4	3.22	3.51	1.09e-06

KEGG Pathways

pathway	description	count in network	strength	signal	false discovery rate
hsa05321	Inflammatory bowel disease	4 of 59	2.17	2.98	1.98e-06
hsa04630	JAK-STAT signaling pathway	5 of 158	1.84	2.65	1.50e-06
hsa04060	Cytokine-cytokine receptor interaction	5 of 282	1.59	2.04	8.51e-06
hsa05133	Pertussis	3 of 73	1.95	1.79	0.00038
hsa04659	Th17 cell differentiation	3 of 99	1.82	1.59	0.00074

Figure 10. (A) First shell expansion of the STRING network investigated in Figure 9A. (B) Functional enrichments in the resulting network.

HGMD-based analysis of shared mutations

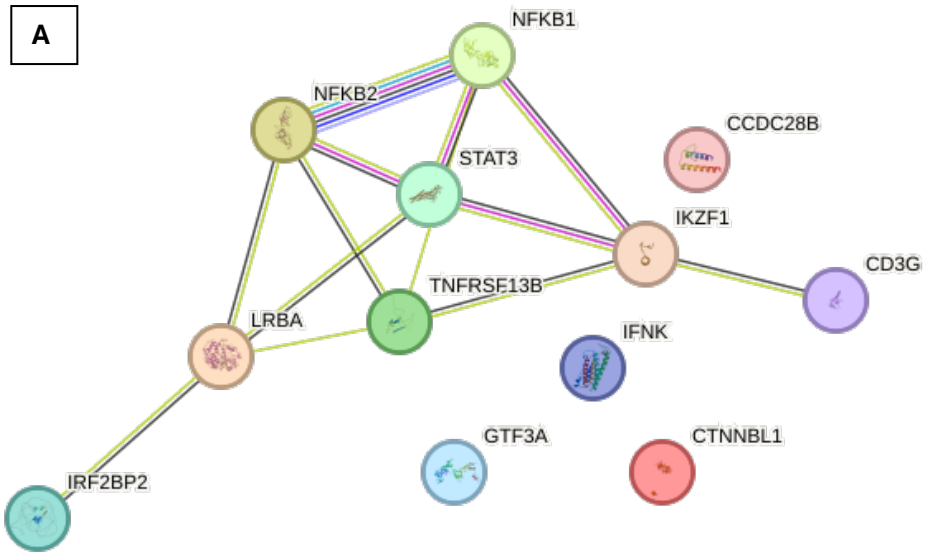
The interrogation of the HGMD database retrieved multiple genes associated with monogenic forms of Common Variable Immune Deficiency (CVID), including *TNFRSF13B*, *NFKB1*, *NFKB2*, *STAT3*, *IKZF1* and *LRBA*. These genes are primarily involved in B cell development, immune tolerance and inflammatory regulation. For sarcoidosis, a broader array of immune-related genes was identified, such as *TNF, IL23R, IL17RA, NFKBIA* and *HLA-DRB1*, many of which are known to promote granulomatous inflammation, cytokine signaling and antigen presentation.

Among all investigated genes, *NOD2* was the only gene found to carry mutations in both pathological contexts. In particular, this gene has been linked to the development of granulomas in other chronic inflammatory conditions such as Crohn's disease and Blau syndrome. The identification of *NOD2* mutations in both sarcoidosis and CVID supports its potential role as a shared pathogenic node in granulomatous inflammation.

Although HGMD mutated genes were disease-specific, convergence at the pathway level was observed. Going into deeper details, both diseases involve alterations in the NF-kB signaling axis, with mutations in *NFKB1* and *NFKB2* in CVID, and *NFKBIA, TNF and LTA* in sarcoidosis. Similarly, dysregulation of the IL-12/IL-23-Th17 pathway was suggested by mutations in *STAT3* in CVID and *IL23R* and *IL17RA* in sarcoidosis. In addition, variants affecting lymphocyte development and peripheral tolerance, like *TNFRSF13B, IKZF1* for CVID and *IL7R, BTNL2* and *BTN3A2* for sarcoidosis, were observed in both diseases.

At the same time, TNF/TNFRSF signaling emerged in both datasets, with *TNF* mutated in sarcoidosis and *TNFRSF13B* (TACI) in CVID, further supporting overlapping inflammatory networks. Additionally, sarcoidosis-associated mutations in *IL23R* and *IL17RA*, together with CVID-associated mutations in *STAT3*, indicate the involvement of the IL-12/IL-23-Th17 axis, laying the basis for upstream regulator predictions.

HGMD variant data have then been plotted in STRING to explore functional connections (**Figure 11**).



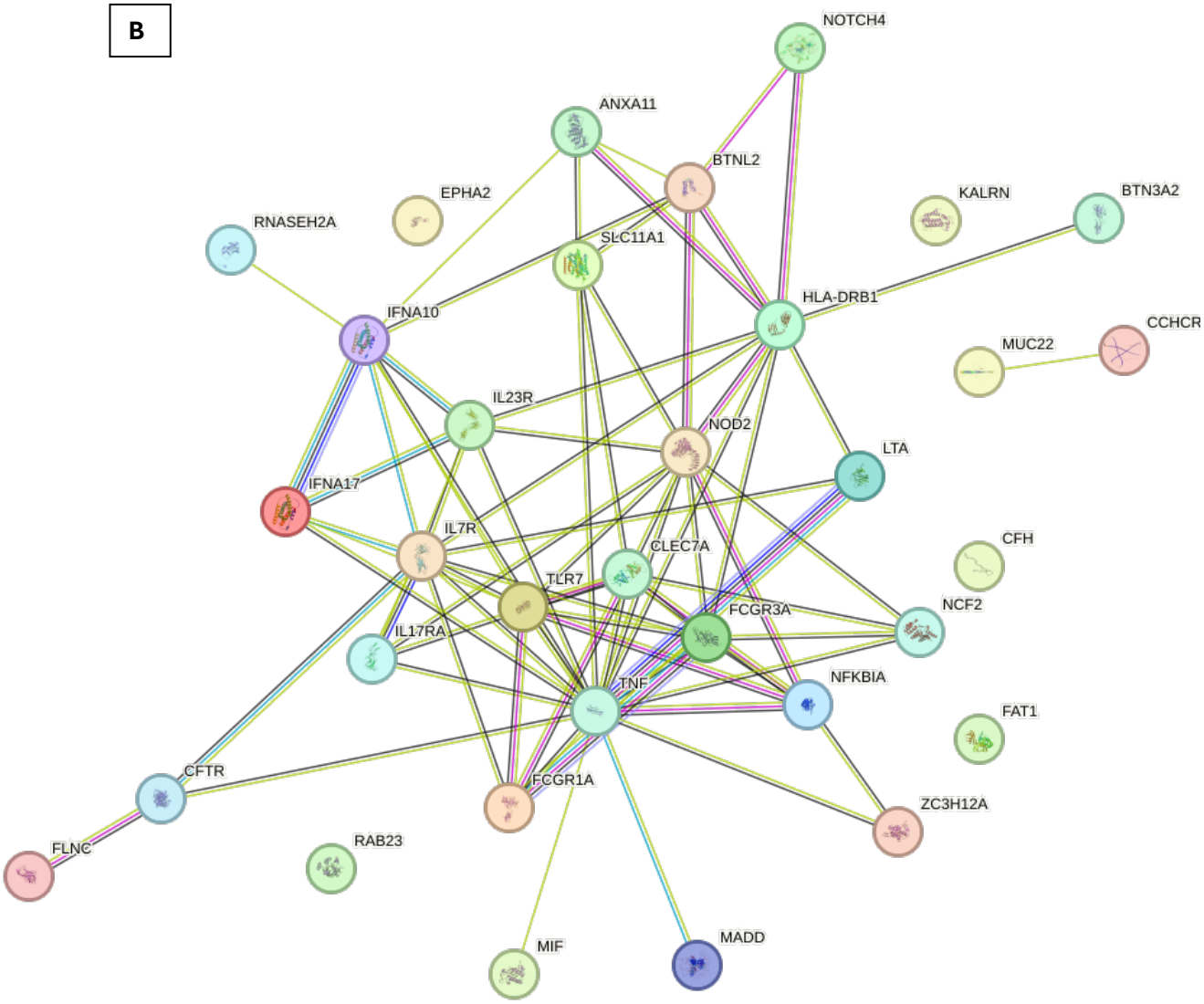


Figure 11. STRING network analysis of (A) CVID and (B) sarcoidosis HGMD variants.

The CVID-associated gene network (**Figure 11A**) is displayed as dense and centralized, with a strong interconnection between NFKB1, NFKB2, STAT3, IRF2BP2, TNFRSF13B, LRBA, IKZF1 e CD3G. These pathways clustered around pathways involved in B cell activation, NF-kB signaling and immune tolerance, while GTF3A, IFNK, CTNNBL1 and CCDC28B remained functionally unconnected within the network.

In contrast, the sarcoidosis-associated network (**Figure 11B**) was markedly more interconnected and widespread, with multiple interconnected hubs such as TNF, IL23R and CLEC7A. This reflects a more heterogeneous genetic architecture, possibly involving genetic factors affecting various immune modules such as Th17 signaling, innate immunity and MHC II-mediated antigen presentation.

IPA analysis

The results of the upstream regulator analysis are presented below as causal network diagrams generated with Ingenuity Pathway Analysis (IPA). Each diagram highlights the predicted activation or inhibition state of the upstream regulator, along with its downstream targets. The networks are based on curated cause-effect relationships from the IPA knowledge base. The resulting diagrams are reported below, each related to the original reference dataset. Each causal network highlights the central upstream regulator (orange: activated; blue: inhibited), its target genes (red: upregulated in dataset; blue: downregulated), and the causal relationships (solid edges: direct; dashed edges: indirect), based on IPA's literature-curated knowledge base.

- **GSE157671**: in lung sarcoidosis biopsies, compared to healthy controls, **IL12B** and **IL23R** emerged as significant activated regulators, as reported, respectively, in causal network analysis diagrams (**Figure 12** and **Figure 13**).

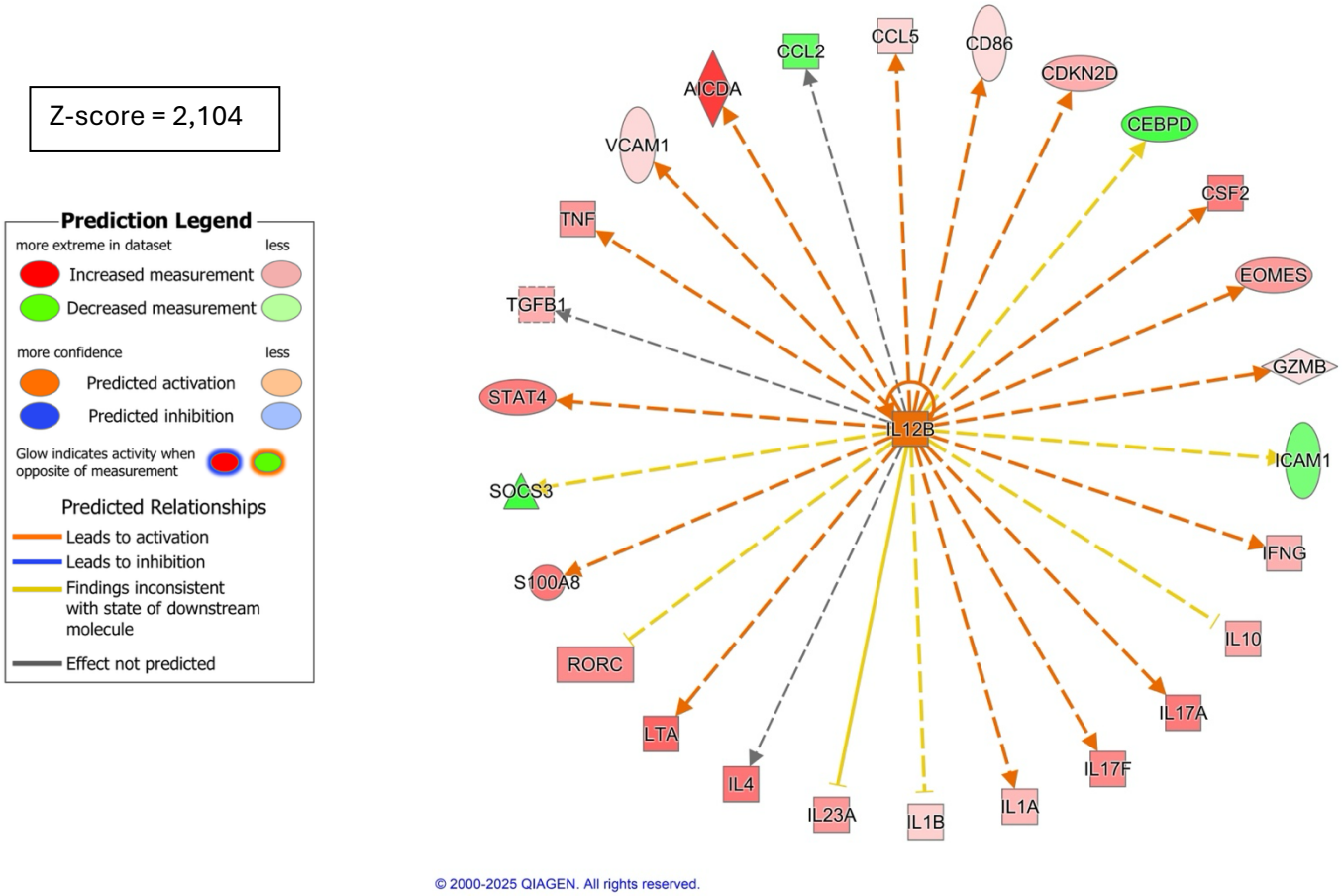


Figure 12. Causal network analysis diagram for IL12B in GSE157671 dataset.

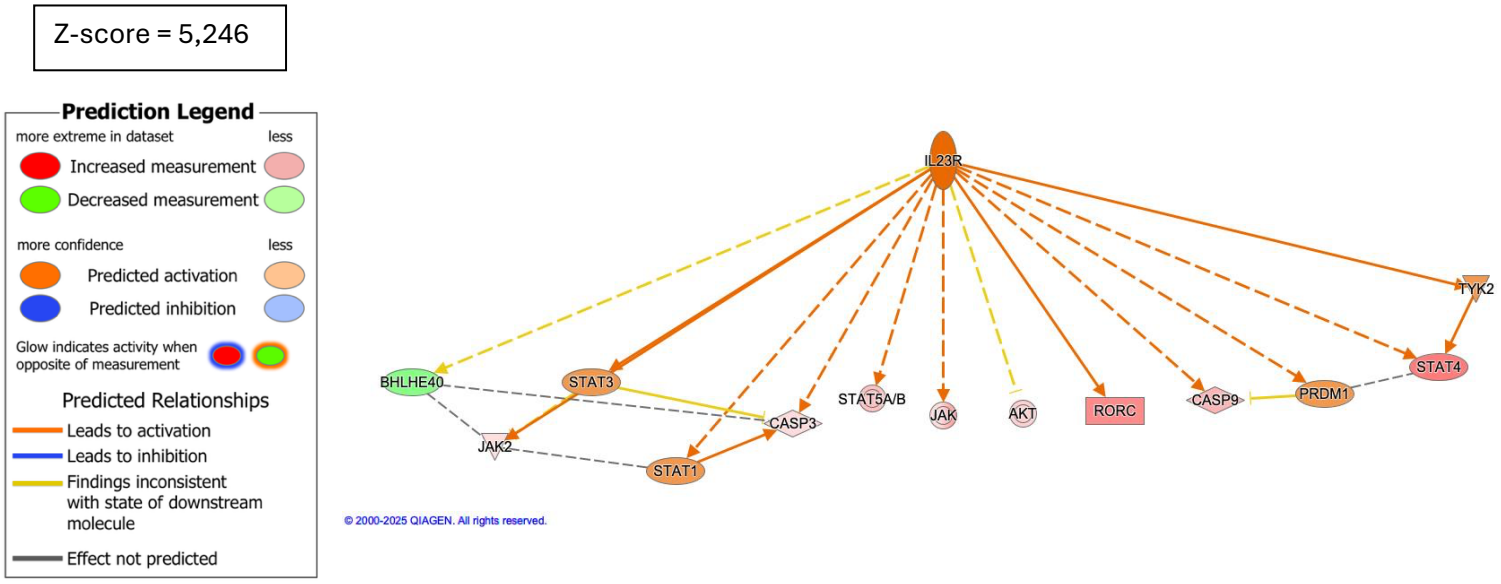


Figure 13. Causal network analysis diagram for IL23R in GSE157671 dataset.

- **GSE16538**: this analysis confirmed the presence of **IL12B** as a potential activated upstream regulator in sarcoidosis vs. healthy lung tissue (**Figure 14**).

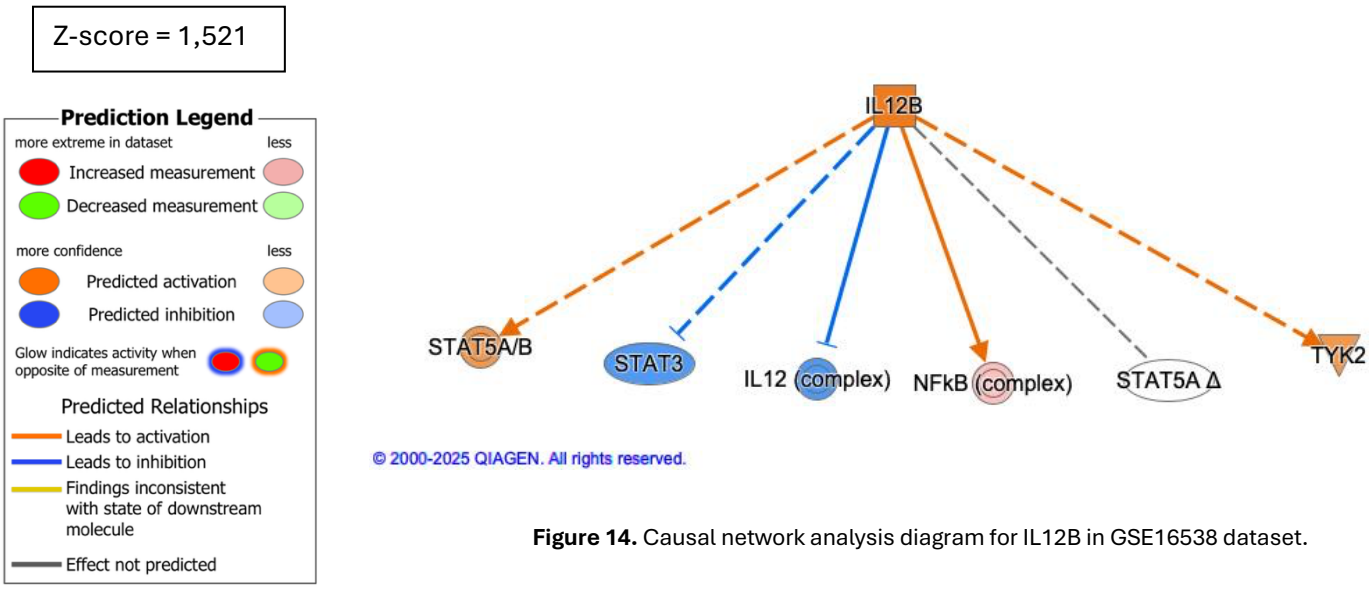


Figure 14. Causal network analysis diagram for IL12B in GSE16538 dataset.

- **GSE51406**: only IL12B was consistently identified as an activated upstream regulator in the resulting causal network diagram (**Figure 15**).

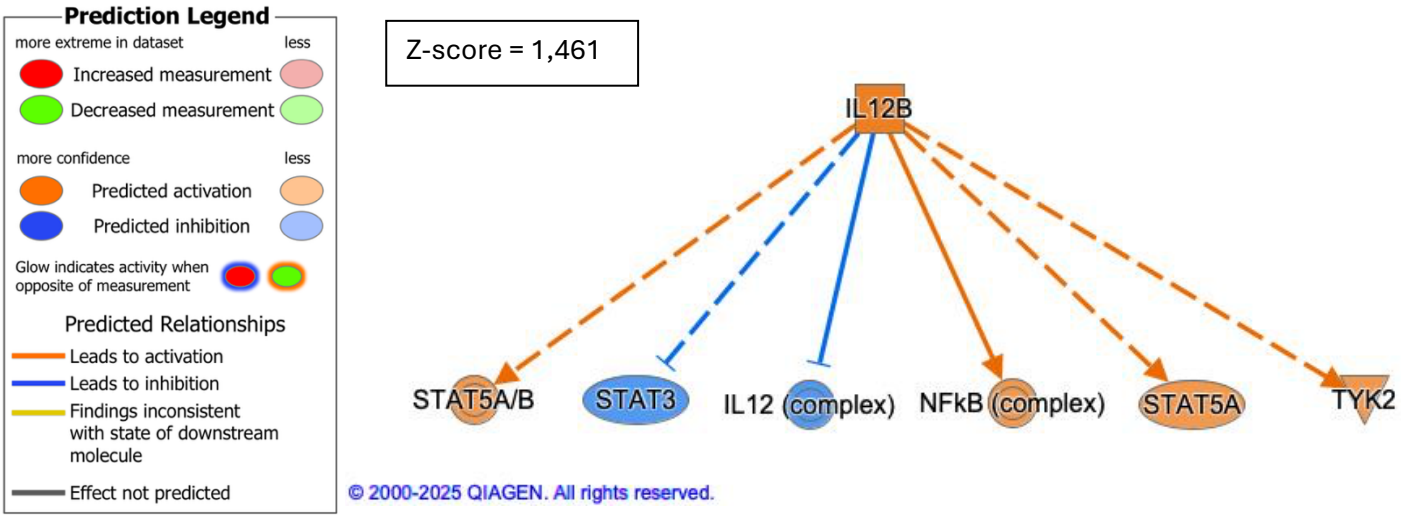


Figure 15. Causal network analysis diagram of IL12B in GSE51406.

IL12B was predicted as an upstream regulator in three independent datasets which cover different biological sources and clinical settings:

- **GSE157671**: lung tissue from sarcoidosis patients vs. healthy controls (orange node, predicted activation)
- **GSE16538**: lung tissue biopsies from active pulmonary sarcoidosis patients vs. healthy controls (orange node, predicted activation)
- **GSE51406**: whole-blood samples from CVID patients with or without granulomatous complications (orange node, predicted activation).

The consistent detection of *IL12B* across tissue and blood-derived datasets reinforces its central involvement in granuloma-associated immune regulation. Despite the different tissue sources, *IL12B* emerged as a shared upstream regulator, with predicted activation in both disease contexts. Together, these results suggest *IL12B* as a key, yet dynamically regulated mediator of Th1/Th17-driven inflammation in both sarcoidosis and granulomatous CVID, and that *IL12B*-mediated signaling pathways may represent a shared mechanistic axis in granulomatous inflammation.

Notably, in GSE157671 the *IL23R* gene, which encodes the receptor subunit for IL23, was also identified as a predicted upstream regulator (orange node, activation). This result, observed only in the direct comparison between sarcoidosis and healthy controls further supports the involvement of the IL12/IL23-Th17 axis specifically in sarcoid granulomatous inflammation.

Discussion

Results summary:

1. No overlap was observed at the SNPs level between sarcoidosis and CVID GWAS.
2. Some disease-associated genes were found to be shared across the two diseases.
3. *IL12B* was found to consistently appear as an activated upstream regulator in all transcriptomic datasets, suggesting a potential role in granuloma formation irrespective of the immune status of selected patients.
4. *IL23R* was identified as an activated upstream regulator in the sarcoidosis-specific dataset only.

Taken together, these results suggest a converging inflammatory pathway centered around *IL12B*, potentially linking granuloma formation in both immunocompetent and immunocompromised settings. The consistent identification of *IL12B* as an upstream regulator across multiple transcriptomic datasets in our analysis suggests a potential shared immunological mechanism between sarcoidosis and GLILD, despite their differing immunological backgrounds.

Taken together, these analyses show that, even if genetic heterogeneity is present between sarcoidosis and CVID, common inflammatory pathways may be involved. Future studies with larger and updated datasets will provide further solid evidence about its role in supporting the granulomatous microenvironment and, hopefully, highlight novel therapeutic targets.

Highlights: functional role of IL12B in granuloma formation

The *IL12B* gene encodes the p40 subunit, which is shared by two key cytokines involved in immune responses: IL-12 and IL-23. These cytokines are essential for orchestrating Th1 and Th17 immune responses, respectively, both of which play significant roles in granulomatous inflammation [20].

IL-12 promotes the differentiation of naïve T cells into Th1 cells, which in turn produce interferon-gamma (IFN-γ). IFN-γ represents a critical activator of macrophages, promoting their aggregation and differentiation into epithelioid and multinucleated giant cells, the hallmark of granuloma formation [21]. Indeed, deficiencies in IL-12 signaling or *IL12B* mutations have been linked to impaired granuloma formation and increased susceptibility to intracellular pathogens in both human and murine studies. Moreover, the impairment of IL12 and IL12B has been associated to dysfunctional granuloma formation and increased susceptibility to infections [21], [22].

Meanwhile, IL-23, which shares the p40 subunit with IL-12, supports the maintenance and expansion of Th17 cells. These cells secrete IL-17A, a pro-inflammatory cytokine which

contributes to neutrophil recruitment and granuloma persistence [23], [24]. The role of Th17 cells and IL-23 has been increasingly recognized in sarcoidosis, with increased expression of IL-23 receptor observed in affected lung tissues [25]. Although less well characterized in GLILD, Th17-related pathways may similarly contribute to chronic inflammation in immunodeficient contexts.

Study limitations

While our exploratory analysis provides initial insight into potential shared pathways, several limitations must be acknowledged:

- First, GWAS data specific for GLILD are currently unavailable. As a result, our comparison relied on GWAS data for CVID in general, which may not fully capture the genetic contributions to granulomatous complications in these patients.
- Second, we compared gene-level associations rather than specific risk alleles, due to the lack of overlap in SNP-level data. Although useful for exploratory purposes, this approach may overlook subtle but significant genetic differences between the two conditions.
- Third, the datasets used for gene expression analysis originated from different tissue sources: lung biopsies in sarcoidosis and peripheral blood in CVID. This heterogeneity in sample type could limit the direct comparability of gene expression signatures.
- Fourth, the reliance on curated databases such as IPA and HGMD, while powerful, introduces potential biases due to differences in annotation depth, update frequency, and coverage.

Finally, our findings, including the central role of *IL12B*, are based on bioinformatic inference and require experimental validation

Future directions

Future experimental work should aim to:

- Analyze larger and more specific cohorts, particularly including GLILD patients with granulomatous complications.
- Investigate the functional consequences of *IL12B* variants *in vitro* and *in vivo*.
- Integrate multi-omics approaches (transcriptomics, proteomics, cytokine profiling) to refine our understanding of granuloma-related pathways.
- Explore the therapeutic targeting of IL-12/IL-23 signaling as a potential avenue for both sarcoidosis and GLILD.

Section 2. Bioinformatics analysis for spatial transcriptomics

Introduction

To further elucidate the cellular architecture and spatial organization of granulomatous inflammation in GLILD and sarcoidosis, we focused on the implementation of a spatial transcriptomics approach focusing on B cells and, in particular, on the $CD21^{low}$ B subpopulation, within lung biopsies.

These cells are characterized by an activated phenotype, reduced signaling threshold, and enhanced pro-inflammatory potential. Despite their well-documented presence in peripheral blood, their tissue distribution and *in situ* interactions within granulomatous lesions still remains poorly characterized.

Aims

The preparatory phase of the spatial transcriptomics project focused on the following specific aims:

1. Characterizing the transcriptional profile of $CD21^{low}$ B cells in the context of CVID-associated GLILD, through the analysis of publicly available datasets and literature, in order to define relevant activation markers and signaling molecules.
2. Identifying candidate protein markers involved in signaling and potential cell-cell interactions by performing protein-protein interaction analysis using the STRING database, with the goal of selecting key targets for downstream spatial transcriptomics detection and annotation.
3. Highlight the possible interactions between $CD21^{low}$ B cells and surrounding immune cell populations, in order to build hypotheses regarding their functional positioning within the granulomatous microenvironment and to guide probe/antibody selection for future experiments.

Materials and methods

Bioinformatics tools

- **Omniviz**: for dataset normalization, analysis and visualization.
- **STRING**: for protein-protein interaction analysis.

Identification of CD21^{low} B cell-associated transcriptional signatures

As a preparatory step, we downloaded and analyzed the GSE13917 dataset from the NCBI GEO. This dataset contains transcriptional profiles of B cell subsets isolated from CVID and Rheumatoid Arthritis patients, including CD21^{low} conventional B cells and CD21^{low} anergic B cells (**Figure 16**).

Series GSE13917	Query DataSets for GSE13917
Status	Public on Dec 01, 2009
Title	Complement receptor 2/CD21– human naive B cells contain mostly autoreactive unresponsive clones
Organism	Homo sapiens
Experiment type	Expression profiling by array
Summary	Complement receptor 2–negative (CR2/CD21–) B cells have been found enriched in patients with autoimmune diseases and in common variable immunodeficiency (CVID) patients who are prone to autoimmunity. However, the physiology of CD21–/lo B cells remains poorly characterized. We found that some rheumatoid arthritis (RA) patients also display an increased frequency of CD21–/lo B cells in their blood. A majority of CD21–/lo B cells from RA and CVID patients expressed germline autoreactive antibodies, which recognized nuclear and cytoplasmic structures. In addition, these B cells were unable to induce calcium flux, become activated, or proliferate in response to B-cell receptor and/or CD40 triggering, suggesting that these autoreactive B cells may be anergic. Moreover, gene array analyses of CD21–/lo B cells revealed molecules specifically expressed in these B cells and that are likely to induce their unresponsive stage. Thus, CD21–/lo B cells contain mostly autoreactive unresponsive clones, which express a specific set of molecules that may represent new biomarkers to identify anergic B cells in humans.
Overall design	RNA was extracted from batch sorted CD19+CD21+CD10-CD27- and CD19+CD21-CD10-CD27- naïve B cells isolated from donors using the Absolutely RNA microprep kit (Stratagene). 100-200 ng of RNA was obtained per sample, and the quality of the purified RNA was assessed by the Bioanalyzer from Agilent. Using the Ovation biotin system kit from Nugen, 30-50ng of RNA was amplified and labeled to produce cDNA. Labeled cDNA was hybridized on chips containing the whole human genome (Human Genome U133 2.0 from Affymetrix). Data from CD21+ and CD21- B cell populations were compared in order to determine the gene signature of the newly described CD21- B cells.
Contributor(s)	Isnardi I, Meffre E
Citation(s)	Isnardi I, Ng YS, Menard L, Meyers G et al. Complement receptor 2/CD21- human naive B cells contain mostly autoreactive unresponsive clones. *Blood* 2010 Jun 17;115(24):5026-36. PMID: 20231422

Figure 16. GSE13917 dataset description from GEO.

The dataset was normalized prior to analysis to ensure comparability across samples. The expression matrix was subsequently uploaded into the Omniviz software, which was then used for interactive visual exploration, clustering, and comparison of gene expression levels between the two B cell subsets. The analysis aimed to identify differentially expressed genes characterizing functional differences between conventional and anergic CD21^{low} B cells. The analysis went on by investigating the top 25 genes co-expressed with *CR2* (CD21). The resulting heatmaps are reported in **Figure 17**.

This list of the top 25 genes co-expressed with *CR2* was then used as input for protein-protein interaction analysis using the STRING database (version 12.0). This allowed us to explore

potential functional interactions, shared signaling pathways, and the broader molecular network associated with CD21^{low} B cell biology, with the aim of identifying candidate markers and interaction partners relevant for future spatial transcriptomics profiling (**Figure 18**).

Protein-protein interaction analysis (STRING)

The 25 genes most strongly co-expressed with *CR2* in the GSE13917 dataset were used as an input for protein-protein interaction (PPI) analysis via the STRING database (version 12.0). The analysis was performed using the default parameters (interaction medium confidence score ≥ 0.4), interacting evidence from experimental data, co-expression, curated databases, and text mining. The resulting protein-protein interaction network is shown in **Figure 18**.

Cell-type specific expression analysis via WebCSEA

To further investigate the cellular context of the genes co-expressed with CR2, we exploited WebCSEA, a tool which maps gene lists onto reference single-cell transcriptomic datasets to determine enrichment patterns across immune and non-immune cell types. The list of the top 25 CR2 co-expressed genes was used as an input. The output is shown in **Figure 19**.

Results

CD21^{low} B cell transcriptomic profile

During the transcriptional profile analysis, particular attention was given to the expression of *CR2* (CD21), the canonical surface marker defining this B cell population. As expected, *CR2* was found to be downregulated in the "anergic" CD21^{low} subset compared to the conventional counterpart (**Figure 17A**). This observation is consistent with previous findings and supports the hypothesis that the anergic CD21^{low} B cell population may play a role in the pathogenesis of GLILD, due its altered activation state and dysregulated signaling potential.

To further investigate the molecular context of the *CR2* downregulation, we identified the top 25 genes co-expressed with CR2 across the dataset (**Figure 17B**). These genes were analyzed to explore possible functional interactions and network dynamics. This analysis provided additional insights into the regulatory landscape surrounding CD21^{low} B cells and allowed the prioritization of candidate molecules potentially involved in granulomatous inflammation or B cell dysfunction in both sarcoidosis and GLILD.

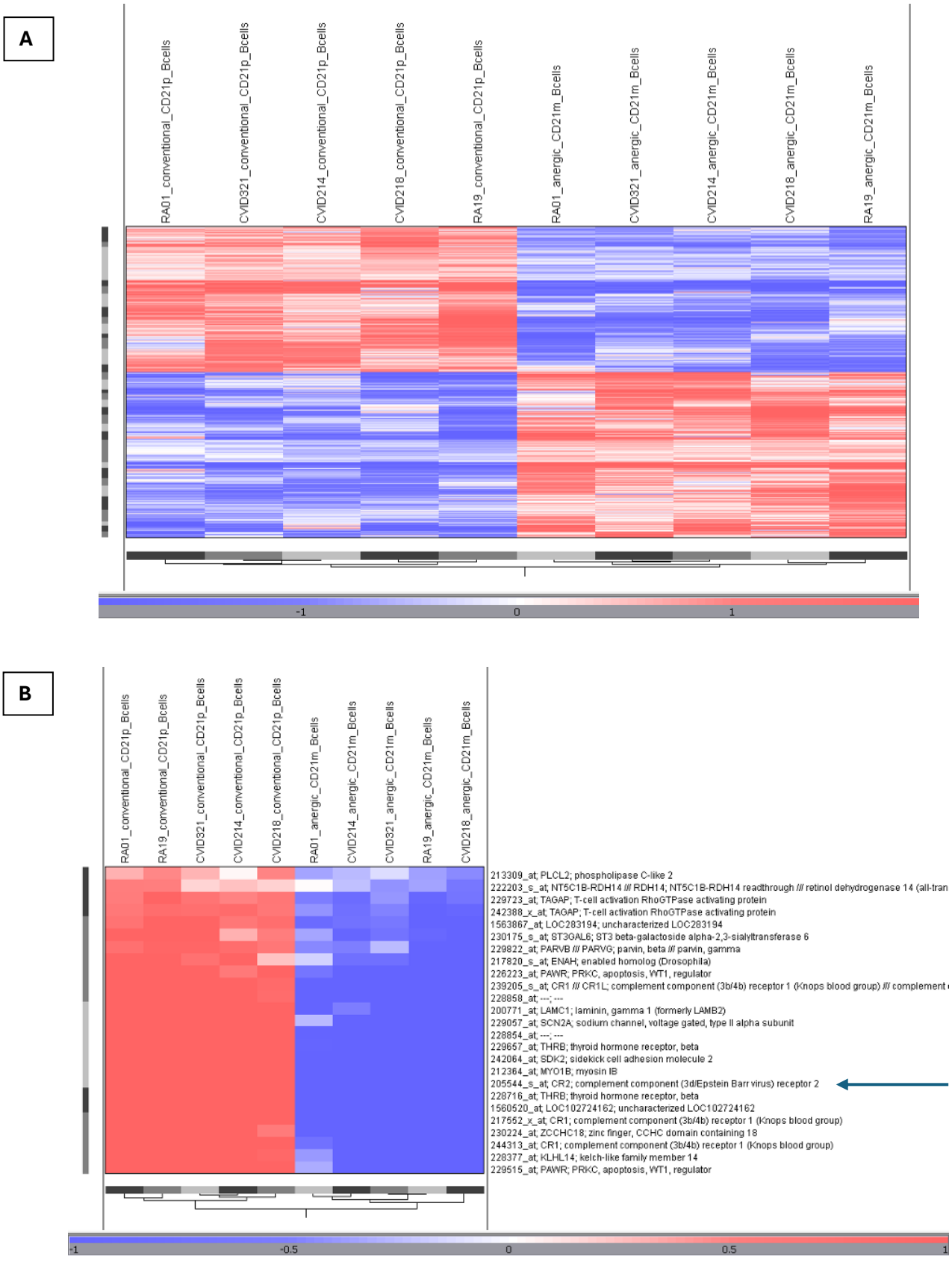
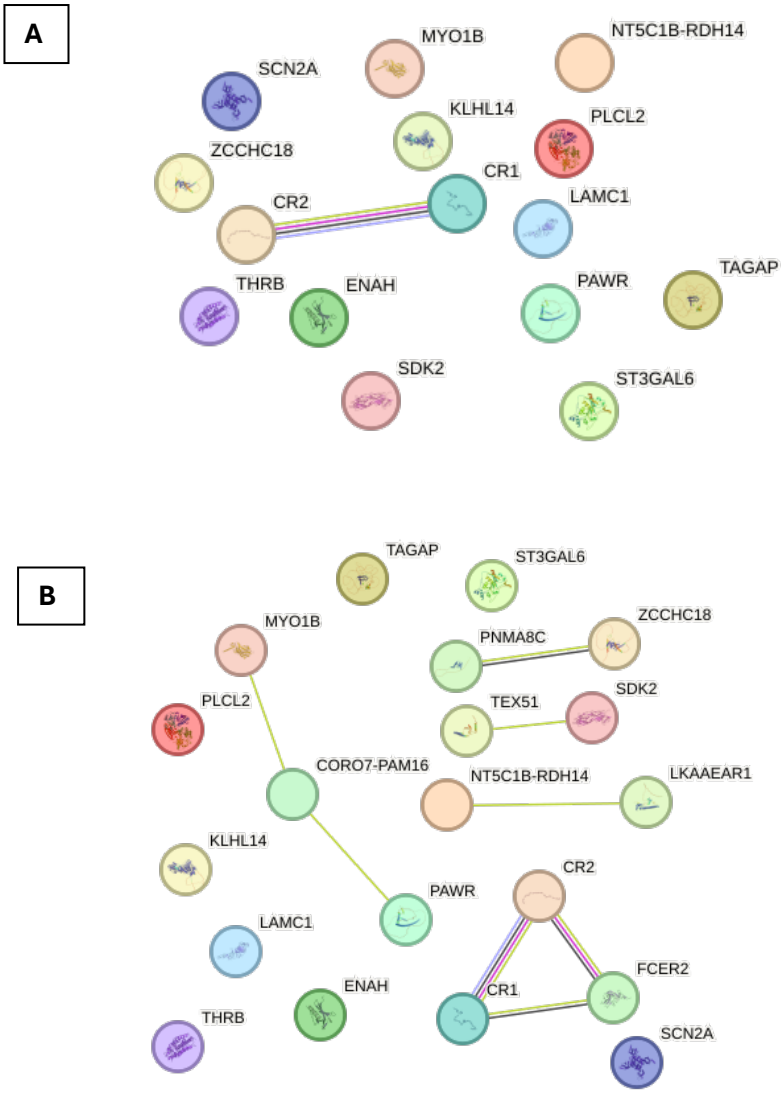


Figure 17. (A) Heatmap resulting from GSE13917 SAM analysis. (B) Heatmap resulting from the analysis of the top 25 CD21 co-expressed genes.

CD21^{low} B cell network analysis

In addition to transcriptomic profiling, the STRING network analysis (**Figure 18A**) did not highlight particularly connected nodes or functional interactions but confirmed the strong interconnection between CR1 and CR2. This analysis could not provide a detailed description of possible molecular interactions that could be helpful in unveiling the CD21^{low} B cell network, so further analysis was supported by a first shell expansion on STRING (**Figure 18B**). This second investigation highlighted another important key mediator in the CR2 network; indeed, CR2 was found to be functionally linked to both CR1 and FCER2, which reflect their coordinated role in mediating B cell activation and complement receptor signaling. The functional enrichment table shown in **Figure 18C** displays high enrichment in the "Regulation of humoral immune response mediated by circulating immunoglobulin".



C

Biological Process (Gene Ontology)					
GO-term	description	count in network	strength	signal	false discovery rate
GO:0002923	Regulation of humoral immune response mediated by circulating im...	3 of 19	2.19	0.76	0.0213

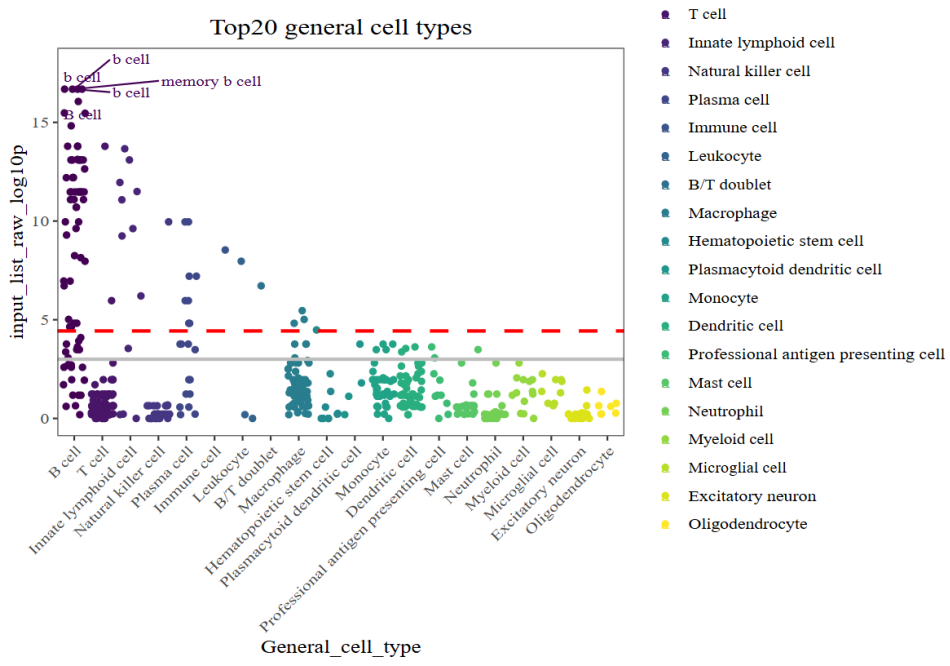
WikiPathways					
pathway	description	count in network	strength	signal	false discovery rate
WP2806	Complement system	4 of 96	1.61	1.07	0.0022

Disease-gene Associations (DISEASES)					
disease	description	count in network	strength	signal	false discovery rate
DOID:7848	Interdigitating dendritic cell sarcoma	2 of 4	2.69	0.62	0.0448
DOID:6262	Follicular dendritic cell sarcoma	2 of 3	2.82	0.62	0.0448

Tissue Expression (TISSUES)					
tissue	description	count in network	strength	signal	false discovery rate
BTO:0004267	Follicular dendritic cell	2 of 5	2.6	0.6	0.0496

Figure 18. (A) Protein-protein interaction network of the top 25 genes co-expressed with CR2, generated using STRING (v12.0). (B) first shell expansion of STRING network showed un 13A. (C) Functional enrichments displayed in the network.

Finally, the resulting WebCSEA analyses highlighted an enrichment in signaling in B cells, particularly within activated and memory-like B cell subsets and mature B cells, consistent with the expected expression of CD21^{low} anergic B cells. The WebCSEA output, shown in **Figure 19**, provided further support for the immune relevance of selected genes and reinforced their potential as markers for spatial localization in granulomatous tissues.



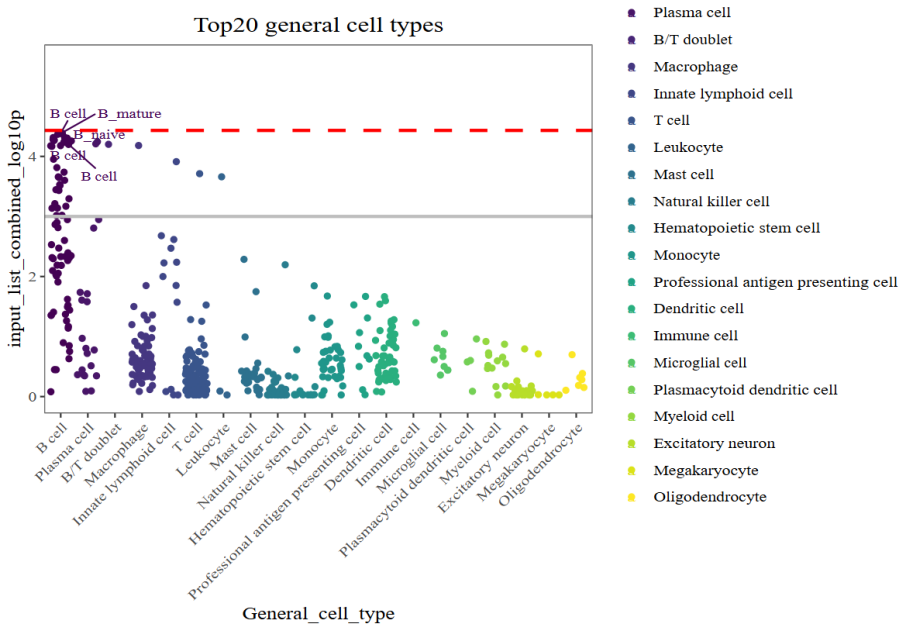


Figure 19. WebCSEA enrichment plot showing predicted cell-type specificity of the 25 CR2 co-expressed genes, based on the integration with single-cell expression atlases. High enrichment scores were observed in memory and mature B cells.

The combined results from protein-protein interaction analysis (STRING) and cell-type specific expression mapping provided a first descriptive framework for prioritizing candidate genes to be included in the future spatial transcriptomics experiment. The integration and implementation of co-expression patterns, molecular network connectivity, and cell-specific expression profiles will allow us to rationally select a panel of markers likely to be biologically relevant and spatially informative within CD21^{low} B cell-enriched immune niches in sarcoidosis and GLILD-affected lung tissues.

Additional notes

Following the experimental design phase, and in preparation for potential downstream implementation, I attended a specialized training course offered by 10X Genomics at Erasmus MC on the Visium Spatial Transcriptomics platform. The course covered key aspects of tissue sectioning, RNA quality control, library preparation, and data processing via the Space Ranger and Loupe Browser pipelines. This training allowed me to gain a clearer understanding of the experimental workflow, and the challenges associated with spatial transcriptomics analysis. It also equipped me with the necessary background to engage with both wet-lab and bioinformatics teams during future implementation phases, whatever the commercial chosen platform will be.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to Prof. Francesco Cinetto, my supervisor in Italy, for introducing me to this field and for inspiring my passion for immunology. His belief in my potential and his unwavering support have been instrumental in giving me the opportunity to never give up, pursue this project and apply for this fellowship. Without his support, none of this would have been possible.

I am also immensely grateful to Prof. Virgil Dalm, my host supervisor in Rotterdam, for welcoming me in his group and for providing me with all the support I needed before and after my arrival. From the very beginning, he showed genuine interest in my project and in my scientific development, making sure I had all the resources and contacts to grow at Erasmus MC. His confidence in my potential gave me the space to take ownership of my work and explore new approaches without fear. His trust and encouragement have been essential to the success of this experience.

A huge, heartfelt thank you goes to Prof. Peter van der Spek, who guided me through every step of the bioinformatic analysis included in this report. His passion for his work, daily support, and ability to strengthen my scientific confidence have deeply impacted me, both as a researcher and as a person. Beyond the science, I will always be grateful to him for showing me that doing research can (and should) also mean having fun and enjoying the journey.

In addition, I would like to sincerely thank Sigrid Swagemakers for her invaluable scientific and technical support throughout the project. Her guidance and patience have been crucial in structuring my analyses and navigating the complexities of dataset interpretation and reorganization.

I also wish to warmly thank Matthijs van Wijngaarden, the kind of colleague anyone would wish for, especially when you are experiencing something completely new abroad. From the very first day, he made me feel at home. His presence, support, and warmth have had a fundamental impact on my experience in Rotterdam, even in the hardest days. He made space for me to feel safe, valued and encouraged me to go beyond my limits. Thanks to him, I learned that science can be discussed with openness and humanity, and that this environment is not only possible, but also essential for deep, lasting growth.

These months have flown by, yet at the same time they felt incredibly rich and expansive: so much I have learned, both as a person and as a researcher. What made this experience truly transformative was all I have received from each person I have met at Erasmus MC. Saying goodbye wasn't easy, but I'll carry with me all the growth and memories we shared. This is not a farewell, indeed, it's only a *see you soon*.

In the end, I must express my heartfelt gratitude to EAACI for awarding me this fellowship. It has had a tremendous impact on my professional and personal growth, far beyond what I could ever imagine, as I will always remember this as a life changing experience for real. I would

strongly encourage any junior member to apply: this opportunity can truly shape you career and open new doors to extraordinary experiences.

References:

[1] H. Buso *et al.*, "Sarcoidosis versus Granulomatous and Lymphocytic Interstitial Lung Disease in Common Variable Immunodeficiency: A Comparative Review," *Biomedicines*, vol. 12, no. 7, Art. no. 7, Jul. 2024, doi: 10.3390/biomedicines12071503.

[2] A. C. Van Stigt, G. Gualtiero, F. Cinetto, V. A. S. H. Dalm, H. IJspeert, and F. Muscianisi, "The biological basis for current treatment strategies for granulomatous disease in common variable immunodeficiency," *Curr. Opin. Allergy Clin. Immunol.*, Oct. 2024, doi: 10.1097/ACI.0000000000001032.

[3] J. R. Hurst *et al.*, "British Lung Foundation/United Kingdom Primary Immunodeficiency Network Consensus Statement on the Definition, Diagnosis, and Management of Granulomatous-Lymphocytic Interstitial Lung Disease in Common Variable Immunodeficiency Disorders," *J. Allergy Clin. Immunol. Pract.*, vol. 5, no. 4, pp. 938–945, 2017, doi: 10.1016/j.jaip.2017.01.021.

[4] F. A. Bonilla *et al.*, "International Consensus Document (ICON): Common Variable Immunodeficiency Disorders," *J. Allergy Clin. Immunol. Pract.*, vol. 4, no. 1, pp. 38–59, 2016, doi: 10.1016/j.jaip.2015.07.025.

[5] J. W. Verbsky and J. M. Routes, "Sarcoidosis and common variable immunodeficiency: similarities and differences," *Semin. Respir. Crit. Care Med.*, vol. 35, no. 3, pp. 330–335, Jun. 2014, doi: 10.1055/s-0034-1376862.

[6] P. Spagnolo, G. Rossi, R. Trisolini, N. Sverzellati, R. P. Baughman, and A. U. Wells, "Pulmonary sarcoidosis," *Lancet Respir. Med.*, vol. 6, no. 5, pp. 389–402, May 2018, doi: 10.1016/S2213-2600(18)30064-X.

[7] H. Zhang, U. Costabel, and H. Dai, "The Role of Diverse Immune Cells in Sarcoidosis," *Front. Immunol.*, vol. 12, Nov. 2021, doi: 10.3389/fimmu.2021.788502.

[8] M. S. A. Fraz *et al.*, "Raised Serum Markers of T Cell Activation and Exhaustion in Granulomatous-Lymphocytic Interstitial Lung Disease in Common Variable Immunodeficiency," *J. Clin. Immunol.*, vol. 42, no. 7, pp. 1553–1563, 2022, doi: 10.1007/s10875-022-01318-1.

[9] I. Harder *et al.*, "Dysregulated PI3K Signaling in B Cells of CVID Patients," *Cells*, vol. 11, no. 3, p. 464, Jan. 2022, doi: 10.3390/cells11030464.

[10] J. Ng, K. Wright, M. Alvarez, G. M. Hunninghake, and D. R. Wesemann, "Rituximab Monotherapy for Common Variable Immune Deficiency-Associated Granulomatous-Lymphocytic Interstitial Lung Disease," *Chest*, vol. 155, no. 5, pp. e117–e121, May 2019, doi: 10.1016/j.chest.2019.01.034.

[11] G. Tessarin *et al.*, "Rituximab Monotherapy Is Effective as First-Line Treatment for Granulomatous Lymphocytic Interstitial Lung Disease (GLILD) in CVID Patients," *J. Clin. Immunol.*, vol. 43, no. 8, pp. 2091–2103, Nov. 2023, doi: 10.1007/s10875-023-01587-4.

[12] D. Friedmann *et al.*, "Bronchoalveolar Lavage Fluid Reflects a TH1-CD21low B-Cell Interaction in CVID-Related Interstitial Lung Disease," *Front. Immunol.*, vol. 11, Feb. 2021, doi: 10.3389/fimmu.2020.616832.

[13] S. Phalke *et al.*, "Age-associated B Cells Appear in Patients with Granulomatous Lung Diseases," *Am. J. Respir. Crit. Care Med.*, vol. 202, no. 7, pp. 1013–1023, Oct. 2020, doi: 10.1164/rccm.201911-2151OC.

[14] A. Raniszewska *et al.*, "Imbalance of B-Cell Subpopulations in the Microenvironment of Sarcoidosis or Lung Cancer," *Cells*, vol. 13, no. 15, p. 1274, Jul. 2024, doi: 10.3390/cells13151274.

[15] N.-S. Lee, L. Barber, S. M. Akula, G. Sigounas, Y. P. Kataria, and S. Arce, "Disturbed Homeostasis and Multiple Signaling Defects in the Peripheral Blood B-Cell Compartment of Patients with Severe Chronic Sarcoidosis," *Clin. Vaccine Immunol.*, vol. 18, no. 8, pp. 1306–1316, Aug. 2011, doi: 10.1128/CVI.05118-11.

[16] M. E. Reincke *et al.*, "The Antigen Presenting Potential of CD21low B Cells," *Front. Immunol.*, vol. 11, Oct. 2020, doi: 10.3389/fimmu.2020.535784.

[17] W. Damsky, D. Thakral, N. Emeagwali, A. Galan, and B. King, "Tofacitinib Treatment and Molecular Analysis of Cutaneous Sarcoidosis," *N. Engl. J. Med.*, vol. 379, no. 26, pp. 2540–2546,

Dec. 2018, doi: 10.1056/NEJMoa1805958.

[18] T. Krausgruber *et al.*, "Single-cell and spatial transcriptomics reveal aberrant lymphoid developmental programs driving granuloma formation," *Immunity*, vol. 56, no. 2, pp. 289-306.e7, Feb. 2023, doi: 10.1016/j.immuni.2023.01.014.

[19] A. C. van Stigt *et al.*, "Granulomas in Common Variable Immunodeficiency Display Different Histopathological Features Compared to Other Granulomatous Diseases," *J. Clin. Immunol.*, vol. 45, no. 1, p. 22, Oct. 2024, doi: 10.1007/s10875-024-01817-3.

[20] P. J. Lupardus and K. C. Garcia, "The structure of interleukin-23 reveals the molecular basis of p40 subunit sharing with interleukin-12," *J. Mol. Biol.*, vol. 382, no. 4, pp. 931–941, Oct. 2008, doi: 10.1016/j.jmb.2008.07.051.

[21] A. M. Cooper, J. Magram, J. Ferrante, and I. M. Orme, "Interleukin 12 (IL-12) is crucial to the development of protective immunity in mice intravenously infected with mycobacterium tuberculosis," *J. Exp. Med.*, vol. 186, no. 1, pp. 39–45, Jul. 1997, doi: 10.1084/jem.186.1.39.

[22] C. Picard *et al.*, "Inherited Interleukin-12 Deficiency: IL12B Genotype and Clinical Phenotype of 13 Patients from Six Kindreds," *Am. J. Hum. Genet.*, vol. 70, no. 2, pp. 336–348, Feb. 2002, doi: 10.1086/338625.

[23] T. Zelante *et al.*, "IL-23 and the Th17 pathway promote inflammation and impair antifungal immune resistance," *Eur. J. Immunol.*, vol. 37, no. 10, pp. 2695–2706, 2007, doi: 10.1002/eji.200737409.

[24] S. L. Gaffen, R. Jain, A. V. Garg, and D. J. Cua, "IL-23-IL-17 immune axis: Discovery, Mechanistic Understanding, and Clinical Testing," *Nat. Rev. Immunol.*, vol. 14, no. 9, pp. 585–600, Sep. 2014, doi: 10.1038/nri3707.

[25] S. Ringkowski, P. S. Thomas, and C. Herbert, "Interleukin-12 family cytokines and sarcoidosis," *Front. Pharmacol.*, vol. 5, Oct. 2014, doi: 10.3389/fphar.2014.00233.

Online resources and consulted databases:

[a] GeneCards – https://www.genecards.org

[b] Human Protein Atlas – https://www.proteinatlas.org

[c] STRING – https://string-db.org

[d] NCBI GEO – https://www.ncbi.nlm.nih.gov/geo/

[e] HGMD - https://www.qiagen.com/us/products/discovery-and-translational-research/clinical-and-epigenetics/hgmd

[f] GWAS Catalog - https://www.ebi.ac.uk/gwas/

[g] WebCSEA - https://bioinfo.uth.edu/webcsea/index.php