Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Patterns of Inheritance of Inborn Errors of Immunity in a Cohort of B-Cell Lymphoproliferative Disorders: an in Silico Approach

María Palacios-Ortega
Hospital Clínico San Carlos
Blanca García-Solís
Hospital Clínico San Carlos
Teresa Guerra-Galán
Hospital Clínico San Carlos
Marc Pérez-Guzmán
Hospital Clínico San Carlos
María Dolores Mansilla-Ruiz
Hospital Clínico San Carlos
Ángela Villegas-Mendiola
Interdepartmental Unit of Immunodeficiencies
Ascensión Peña-Cortijo
Hospital Clínico San Carlos
Eduardo Anguita-Mandly
Hospital Clínico San Carlos
Marta Polo-Zarzuela
Hospital Clínico San Carlos
Estefanía Bolaños-Calderón
Hospital Clínico San Carlos
Cristina Pérez-López
Hospital Clínico San Carlos
Fiorella Medina
Hospital Clínico San Carlos
Eduardo de la Fuente-Muñoz
Hospital Clínico San Carlos
María Ruiz-del-Río
Hospital Clínico San Carlos
Reynaldo Homen
Hospital Clínico San Carlos
Miguel Fernández-Arquero

Hospital Clínico San Carlos

Juliana Ochoa-Grullón

Interdepartmental Unit of Immunodeficiencies

María Guzmán-Fulgencio

Hospital Clínico San Carlos

Celina Benavente-Cuesta

Hospital Clínico San Carlos

Rebeca Pérez-de-Diego

IdiPAZ Institute for Health Research, La Paz University Hospital

Silvia Sánchez-Ramón

ssramon@salud.madrid.org

Hospital Clínico San Carlos

Research Article

Keywords:

Posted Date: September 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10767852/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: Competing interest reported. Silvia Sánchez-Ramón is an Associate Editor of this journal

---

Abstract

The immunogenetic profile of patients with B-cell lymphoproliferative disorders (B-CLPD) is still unknown. Even though a substantial proportion of patients manifest clinical pictures that resemble those of primary immunodeficient patients (PID) or inborn errors of immunity (IEI), little is known about their genetic origin and their possible relationship with PIDs.

We performed an observational study and a statistical and *in silico* analysis of 78 B-CLPD patients with secondary immunodeficiency (SID), through Whole Exome Sequencing (WES), and compared the results with the published data from the general population, and the results of a cohort with PID diagnosis.

The genetic variants found in our SID patients associated with B-CLPD resembled those of PID patients, with a similar mutational frequency, above the 50%, which is significantly higher than in the general population. However, the IUIS categories of affected genes and the patterns of inheritance differed from SID to PID cohorts: SIDs showed increased ratios of mutations in genes related to combined immunodeficiencies and immune dysregulation, and more likely pathogenic/pathogenic variants in autosomal recessive genes, more variants of unknown significance in autosomal dominant genes, and also, more co-occurrence of mutations in those genes.

This is the first study to address the genetic similarities and differences between B-CLPD and PID patients. The results are compelling and may give novel hints to lymphomagenesis, IEI genes, and PIDs.

INTRODUCTION

Diagnosis and management of B-cell lymphoproliferative disorders (B-CLPDs) (i.e., chronic lymphocytic leukemia, Hodgkin lymphoma and non-Hodgkin lymphoma, multiple myeloma) is well established has dramatically changed in the last decade towards better immune evaluation and immunotherapeutic strategies. However, clinically relevant questions remain unanswered, particularly regarding the intrinsic role of the immune system in B-CLPD oncogenesis and the potential impact of genetic variants that may predispose individuals to a higher risk of developing these disorders.

Many genetic defects, including inborn errors of immunity (IEI)[1], are known to hinder the normal function of the immune system, and, notably, patients with IEI carry a higher risk of developing malignancies, mostly hematological[2]. While it is described that monogenic defects appear in about 30–45% of all patients with clinically diagnosed primary immunodeficiencies (PID) [3,4,5], still, a currently unknown proportion of B-CLPDs are associated with IEI variants. In a previous paper we covered the real difficulties to distinguish between actual secondary immunodeficiencies (SID) associated with B-CLPD and “suspected PID” behind this diagnosis. We found that 66% of the 59 B-CLPD patients studied carried, at least, one genetic variant related to an IEI[6], a finding which could have a real clinical relevance and impact on these patients and their families, as early identification of IEIs could help anticipate severe infectious and non-infectious complications with a more personalized-based management and treatment strategy, potentially prolonging free-from-disease years and a better quality of life.

Moreover, the fact that IEI can underlie B-CLPD patients arise other questions. What are the inheritance patterns of IEIs in these patients? Do they follow a specific distribution? Does it differ from the general population and from PID patients? Are there any specific biological pathway defects overrepresented?

Therefore, we performed a novel *in silico* study to analyze the clinical exome of a total of 78 patients with B-CLPDs with suspected PID, aiming to describe the genetic landscape of the IEI gene lesions in these population and their possible relationships in terms of biology and gene ontogeny with their hematological malignancy.

MATERIALS AND METHODS

Patients and study design

We performed a single-center observational and exploratory study in a cohort of 78 SID patients associated with B-CLPD evaluated at our institution between 2015 and 2025, as previously described. Their B-CLPD diagnosis was established by the Haematology Department according to the WHO Classification of Tumors of Haematopoietic and Lymphoid Tissue[7]. SID was diagnosed based on the last clinical guidelines[8]: recurrent or severe infections, a reduction in serum IgG, IgA, and/or IgM by at least two standard deviations below the normal mean, and a failure to mount an antibody response to polysaccharide and/or protein antigens, in the presence of diagnosed malignancy. We followed a non-biased approach, where B-CLPD patients were included attending solely to one criterion: their hematological condition (non-Hodgkin lymphoma (NHL)). No analysis or study of their immunodeficiency was performed, and no diagnostic algorithm was followed to differ whether they could fit into the 'Suspected-PID' group or not. Also, to increase the strength of the study, we used, as an internal control, our cohort of 56 patients with a PID diagnosis, to compare the results of their clinical exome to those of the SID patients. We also performed a database search to find the prevalence of the mutations in the general population (GP), using *GenIE* (Genetic Prevalence and Incidence Estimator) software[9], which takes prevalence data from *ClinVar* and *gnomAD* public databases.

All procedures followed were in accordance with the ethical standards of the institutional research committee and with the Helsinki Declaration of 1975, as revised in 2000. The study received approval from the hospital's Institutional Research Ethics Committee (20/072-E and 19/219-E).

Genetic analysis

For the present study, peripheral blood samples were taken from 78 patients with B-CLPD and 56 with PID, having previously signed the proper informed consent form approved by the hospital's Institutional Research Ethics Committee (20/072-E and 19/219-E). A single EDTA-tube was obtained from each patient and was sent to *Health in Code* laboratories, where the received sample went through whole exome sequencing (WES) by NGS method following the guidelines of HIC-PT-BIO-06 and HIC-PA-BIO-01. The genes included in the analysis were selected based on clinical criteria, considering their association

with a specific phenotype, and classified according to the level of evidence supporting this association into priority genes, secondary genes, and candidate genes. The regions of interest comprise the coding regions and at least 10 nucleotides of the flanking intronic regions of the selected genes. For certain genes with sufficient evidence, the region under study was extended to deep intronic regions or UTRs with clinical evidence. The list of analyzed genes along with their sequencing metrics is available upon request, if required.

The sensitivity and specificity of this assay exceed 99% for single nucleotide variants (SNVs) and 94% for small insertions and deletions. Copy number variant (CNV) detection yields sensitivity values exceeding 99%. Prioritization and interpretation of genetic variants is conducted by considering clinical, genetic, and population-based criteria, cross-referencing these data against multiple population, disease, gene, and variant databases, and employing computational prediction tools. In accordance with the current guidelines of the American College of Medical Genetics and Genomics (ACMG)[10], variant classification is performed using five categories: pathogenic (P), likely pathogenic (LP), variant of uncertain significance (VUS), likely benign (LB), and benign (B). The current guidelines, however, also allow the use of additional subclasses, like the subdivision of VUS (e.g. VUS-high and VUS-low), due to its wide range of probability for causality (from 10% to 90)%[11, 12]. For the purpose of this study, only P, LP and VUS-high variants were considered. For simplicity of the analysis, P and LP variants were considered together.

Statistical analysis

Data were collected and analyzed using Excel spreadsheet (Microsoft, Inc., Redmond, WA, USA). Genetic variants, as categorical variables, were described as counts and percentages of subjects and grouped according to the latest IUIS Classification for IEIs[1]. All categorical variables were compared using the *Chi-squared* test. *Pearson Chi-Square* value was reported for most categorical variables, while Fischer's Exact Test was used when the number of events in a subgroup were less than 5. T- Tests were performed to compare continues variables. Statistical significance was considered when 2-sided p-value was below 0.05.

GraphPad Prism Version 9.5.1 was used for the statistical analysis.

Clustering

To deepen the biochemical and functional understanding of the relationship of the proteins coded by the genes found in the study, we used two informatic platforms designed to carry out clustering and enrichment analysis. These types of analysis allow the determinations of statistically enriched biological pathways within a set of genes or proteins[13]. First, we used 'STRING: functional protein-protein interaction networks' software version 12.0 for gene-phenotyping clustering, and employed *K-Means* clustering, which is a well-established algorithm for studying genetic associations[14]. We then used

'*Cytoscape* software' version 3.10.4 to perform enrichment analysis in relation to their ontology and disease phenotypes, to better understand the most over-represented relationships between the different genes. P-values were obtained for statistical significance of the functional enrichment and correlate to the hypergeometric p-value after correction for multiple testing. Only p-values < 0.05 were considered significant.

RESULTS

1. IEI-associated genetic variants in patients with B-CLPD

We identified 43 out of the 78 patients carrying one or more genetic variants related to IEIs, representing 55.13% of the SID patient's cohort.

There was a total of 62 genetic variants affecting 35 different genes. Of all the genes involved, 97.14% (34/35) were autosomal: 62.86% (22/35) were described as autosomal recessive (AR) in the scientific literature, and 17.14% (6/35) as autosomal dominant (AD). The remaining 17.14% (6/35) were genes where both types of inheritance (AR and AD) have been notified, depending on the type of defect present. Only one of the genes affected (NRAS) (2.86%) was considered somatic, as germinal NRAS mutations rarely to occur. No genes with X-linked inheritance were found in our cohort.

Regarding the type of genetic variants, missense mutations were the most represented among the cohort, found in a total of 22 genes (62.86%). Frameshift mutations were described in 4 genes (11.43%). The rest of the genes carried other types of mutations: nonsense, splicing, and somatic mutations. Each of those were detected in 2 genes, respectively. Only 1 gene (PIK3CD) was affected by intronic variants, and 2 genes had more than one type of genetic mutation (PRKDC and SBDS) (Supplementary table 1).

In terms of the predicted pathogenicity of the variants, 32.26% (20/62) were described as LP/P, and 67.74% (42/60) as high-VUS[10,12]. Patients with low-VUS, LB or B variants were considered IEI negative.

All the variants appeared in heterozygosity, however, we observed that more than a third of the IEI-positive patients (30.23%; 13/43) had mutational co-occurrence, which is defined as having more than one gene affected by one or more variants: a 69.23% (9/13) carried 2 different mutations and 3 (23.08%) carried 3 mutations. Only 1 patient (6.25%) carried more than 3 gene variants. Compound heterozygosity was found in 2 patients (6.98%) (Fig. 1).

The co-occurrence of 2 or more LP/P variants was 2.33% (1 patient). However, the co-occurrence of LP/P mutations with VUS raised to almost a fifth of the patients with IEI (16.18%), slightly more than the co-occurrence of 2 or more VUS (13.96%). Interestingly, when analyzing the type of inheritance and the predicted pathogenicity altogether, we found three major patterns among the patients with IEI: 1) Isolated VUS in AR genes (25.58%), 2) isolated VUS in AD genes (16.28%), 3) LP/P variants in AR genes (16.28%).

Furthermore, almost 40% (39.53%) of the 43 patients with IEI-associated genes had, at least, one LP/P variant, which corresponds to a 21.79% of the whole cohort. Noteworthy, as well, is that 18 patients (41.86%) carried, at least, one mutation in an AD gene (Supplementary Fig. 2).

There were no sex differences, the same proportion of variants were found in females and males (50% in each sex group).

2. The majority of the genes affected fell under three main groups of IEIs from the 10th IEI IUIS categories.

In our cohort, most of the genes detected fell under three major groups: 1. Group 1 + 2 (Immunodeficiencies affecting humoral and cellular immunity + CIDs with associated or syndromic features, which we have considered together for the purpose of this study) (SCID/CID) with a 45.90% (n = 27); 2. Group 4 (Diseases of immune dysregulation) (PIRD) representing a 21.31% (n = 13); and group 3 (Predominantly antibody deficiencies) (PAD/CVID) with a 13.11% (n = 9). Five additional genes (8.20%) belonged into group 5 (Congenital defects of phagocyte number or function) (Phagocyte defects), and the remaining belonged to groups 7 (Autoinflammatory disorders), 9 (Bone marrow failure), 6 (Defects in intrinsic and innate immunity) (Innate defects) and 10 (Phenocopies of IEIs associated with autoantibodies or somatic variants) (IEI phenocopies), with 4, 2, 1 and 1 genes found in each of the groups, respectively (Supplementary table 1).

When looking at the 20 LP/P variants, we found that the majority belonged to groups 1 + 2, and 4, with 35.00% (7/20) and 30.00% (6/20), respectively. This finding might suggest that deep immune defects such as those of combined immunodeficiencies, and increased dysregulation and affected immunosurveillance could increase the risk of developing lymphoproliferative disorders. Only 2 genes of group 3 had LP/P variants. Of note as well, is that the only two genes of group 9 were LP/P.

3. More than 60% of the genetic variants found in the 3 most represented groups of IEI either appeared in co-occurrence with another variant, or in genes with described AD.

When deepening on the analysis of the relationship between the inheritance pattern of the genes, the predicted pathogenicity of the mutations and the presence of co-occurrence in the most represented groups of IEIs of our cohort, a 66.13% of the total of the genetic variants happened in AD genes or presented in co-occurrence with another variants (46.77% appeared in co-occurrence, and 19.35% affected AD genes). Looking at each of the groups, we found that Groups 1 + 2 and 4, had a higher representation of co-occurrence (27.42% and 12.90%, respectively) (Fig. 1), with a lower presence of mutations on AD genes. On the contrary, in group 3, most of the mutations were described as VUS, but at the same time, most of the genes affected had AD inheritance (66.66% of group 3), representing a 9.68% of all the mutations. But also, a no lesser number of variants of group 3 presented in co-occurrence, a 6.45% of the 62 mutations

4. IEI variants are more represented in the SID cohort than in the general population (GP).

Even though we observed that IEI variants were highly represented in the SID cohort, we sought whether the frequency of IEI carriers in the GP was similar, because, in that case, it would be less likely that the mutations observed were related to the disease of the patients. We looked up the prevalence of the heterozygous genetic variants in *GenIE* (Genetic Prevalence and Incidence Estimator) software[9], which takes prevalence data from *ClinVar* and *gnomAD* public databases. Intriguingly, the frequency of carriers in the GP was significantly lower than that in the SID cohort (Supplementary Fig. 1), with a $p < 0.0001$, which probably indicates that IEI variants are overrepresented in the SID population.

5. Comparison of patterns of IEI mutations observed in B-CLPD and PID patients.

We then compared the analysis with that of our PID cohort (N = 56).

First of all, we observed an almost identical proportion of the 56 PID patients with a detected IEI variant (53.57%) as those with B-CLPD and SID (55.13%). We then studied and compared the distribution of the genetic variants according to the IUIS classification[1] and observed a similar pattern between both groups. As described above for the SID cohort, most of the variants in the PID cohort also fell under three major groups: group 1 + 2 (SCID/CID): 11 variants (23.40%); group 3: 13 variants (27.66%); and, alike the SID cohort, group 5 was slightly more represented (6 variants, 12.77%) than group 4 (5 variants, 10.64%), but with no statistically significant difference. Another visible difference, as seen on Fig. 2-A, although also non-significant, is that the representation of variants related to group 3 is bigger in the PID group than in the SID one (13.11% vs 27.66%). Conversely, groups 1 + 2 (45.90% vs 23.40%) and 4 (21.31% vs 10.64%) were more frequently represented in SID patients compared to PIDs.

To deepen on those differences between the most represented IEI groups in both cohorts, we compared the rates of VUS and LP/P variants in each of them (Fig. 2. B) with some interesting findings. We found no statistically significant differences between the rates of VUS and LP/P variants between SIDs and PIDs for groups 1 + 2 and 5; however, there was a significant difference for groups 3 and 4 ($p < 0.05$ and $p < 0.001$, respectively).

LP/P variants in group 3 were more prevalent in PID patients, but more frequent of group 4 in the SID cohort.

6. SID associated with B-CLPD patients tend to have either more VUS or more AR variants, but also a greater rate of co-occurrence of LP/P variants or variants in AD genes, compared to PID patients.

The factors conveying a particular variant expression is a difficult one to fully understand, including penetrance and expressivity of a given gene. And therefore, it is an intricate task to try to understand genotype-phenotype associations. For that reason, we analyzed the type of inheritance, described pathogenicity and the presence of co-occurrence with other genetic variants combined in the patients, to compare the different patterns among our two cohorts.

First, we studied each of the three variables separately in our patients of both groups, finding no significant differences (Fig. 3). However, in the patients, these variables act altogether in different combinations, leading to different inheritance and expression outcomes. For that reason, we then analyzed and compared the different possible patterns in both groups and found out that they significantly differ between SIDs and PIDs.

If we studied the variants of each patient in terms of pathogenicity and type of inheritance, we were able to detect that mutations happening in AD genes were more prevalent in PIDs, especially, those described as LP/P. Regarding co-occurrence and types of inheritance, we found that single AD variants were increased in PIDs, while single AR were in SIDs.

In alignment with that finding, when studying the combination of co-occurrence and pathogenicity, we observed that single VUS variants were more described in SIDs, and single LP/P variants in PIDs. However, the prevalence of co-occurrence of LP/P variants with other types of variants was higher in SIDs than PIDs (Fig. 3).

Finally, we studied the patterns resulting from the combination of the three variables in the patients to try to make a more reliable approach. Interestingly, we noticed clear differences: there was a higher prevalence of PID patients with single variants in genes that have an AD inheritance, whether VUS or LP/P. Secondly, there were more SID patients with single AR variants, also both VUS or LP/P, but also, more frequent co-occurrence of VUS variants in AD genes, and with co-occurrence of LP/P variants in AR genes (Fig. 3).

7. In Silico functional cluster analysis revealed four major cellular roles of IEIs in B-CLPD patients

Using a clustering software database (*String*), we were able to classify the resulting genes of the SID patients in 4 groups or clusters. Clustering was obtained through a functional enrichment analysis, through which different gene groups are created based on the biological and functional processes in which the genes are involved. Nodes represent the different genes or proteins involved in the search, and the edges represent their biological associations, whether functional and/or physical, obtained through deep computational research. These interactions are classified in ‘known interactions’ from curated databases (blue line) or experimentally determined (pink line), in ‘predicted interactions’ such as gene neighborhood (green line), and ‘other interactions’ obtained by ‘text-mining’ (yellow line).

The network resulting from the 35 genes studied had 83 edges, which means that the genes and their derived proteins have more interactions among themselves than what would be expected for a random set of proteins of the same size and degree distribution drawn from the genome ($p = < 1.0e-16$). Thus, indicating that the genes studied are at least partially biologically connected, as a group. Another finding that may suggest that the mutations affecting IEI genes in patients who develop B-CLPD may not be just a result of pure chance, as they share similar clinical phenotypes along with mutations in biologically connected genes.

The 4 clusters of genes obtained had clear implications for the immune system's function, as can be observed in Fig. 4. The 'Red group', constituted by 14 genes related to adaptative immune responses, regulation of leukocyte and T cell apoptotic processes, as well as 'Primary immunodeficiency disease', was the biggest cluster obtained. The next most represented group was the 'Yellow group' (N = 8), related to somatic diversification of immune receptors and DNA repair complex. The 'Green' and 'Blue' groups were made up of 4 and 3 genes, respectively. The first one is associated to leukocyte-mediated cytotoxicity, cytolytic granules and vesicle trafficking; while the other, is related with DNA methylation, transcriptional regulation and "abnormal lymphocyte counts". Other genes, represented in white nodes (N = 5: ERCC6L2, NCF1, TRNT1, ADA2, CLPB), were not classified in any of the clusters mentioned above. However, when studying them separately, even though they do not seem to share any direct relationship, enrichment analysis suggested that they could possibly belong to sets of genes involved in functions that can lead to 'abnormal cellular immune system morphology" (Fig. 4).

These results highlight that main biological processes that are essential for the maintenance of the immune system's normal function against all kinds of potential aggressors, could be impaired in B-CLPD patients.

In alignment with the previous results, this analysis shows that the most affected biological processes are those which correlate with the IEI groups that are mostly represented.

Specifically, defects in T-and B-lymphocyte activation and differentiation can lead towards reduced responses against outer and inner challenges, resulting in combined immunodeficiency phenotypes; altered DNA repair processes and somatic hypermutation are known to increase mutagenesis risk, as well as impaired antibody responses. Impaired vesicular trafficking, degranulation and cytotoxicity are associated with decreased immunosurveillance by cytotoxic T-cells (CTLs) and NK-cells, which, in turn, can increase the risk of malignancy via reduced viral clearance and reduced cytotoxicity. Lastly, altered development of immune cells can also lead to increased risk of transformation and malignancy, mostly, when apoptotic pathways are also challenged.

8. Human disease phenotypes 'Haematological malignancy' and 'Recurrent infections' share similar genetic background.

Further enrichment analysis, with *Cytoscape* software, was performed to look at the relationships regarding gene ontology (GO), and the related human disease phenotype of the genes studied.

We identified 4 over-represented GO terms ('leukocyte activation', 'cell activation', 'immune system development' and 'immune system process') (Fig. 5), which correspond to the most statistically significant associations, and which also align with the results obtained in the previous enrichment analysis of biological processes.

When looking at human phenotype enrichment, we found an over-representation of the following phenotypes: 'abnormal leukocyte count', 'abnormal leukocyte morphology', 'abnormal cellular immune

system morphology' and 'abnormal immune system morphology'. All of them represent the importance of the genes identified in both the normal function and normal cellular morphology of the immune system's cells, which could also be referred to as a higher risk of immune dysregulation and malignant cell development.

Lastly, we overlapped this last enrichment analysis with two clinical outcomes: "Recurrent infections", and "Haematological malignancy" (Fig. 6). We observed that both phenotypes are largely related to almost all the genes studied, especially in the case of the "Recurrent infections" phenotype, which suggests a possible crossroads and connection between them, in terms of ontology, as well as biological origin.

DISCUSSION

Following our previous study[6], we broaden the genetic study of patients with SID associated with B-CLPD to a bigger cohort of 78 patients. But, in contrast to that study, the only inclusion criterion was having NHL, making it a novel and non-biased-approach study. Confirming our previous findings, we found that more than half of these patients carry, at least, 1 IEI-related variant.

In the last decade, different studies with varying approaches have supported the concept that patients with SID and B-CLPD, especially lymphomas, are not that different from PID patients with infections or other inflammatory complications. Xiaofei et al.[15] studied, in 2022, both somatic tumor DNA and germline DNA of a cohort of 23 patients with suspected immunodeficiency and lymphomas, finding LP germline mutations related to common variable immunodeficiency (CVID) in more than half of them. Guevara-Hoyer. K et al.[16] found a relevant 25% rate of CVID-related somatic variants (half of them with up to ten concurrent variants) in series of 1,309 NHL database, publicly available. Others, like Allain.V. et al[17], also found, in 2023, similarities between patients who developed a lymphoma before the onset of their hypogammaglobulinemia and those who did in the reverse order.

Our study encompasses a different and novel approach where WES was performed in a wide cohort of SID patients with B-CLPD, without including any other variables, for the purpose of studying their genetic background without bias, comparing the results to the GP and to a cohort of clinically diagnosed PIDs. Noteworthy, upholding our previous results, we identified 62 genetic variants in 35 different genes of IEI, in a total of 43 out of the 78 B-CLPD patients studied, revealing that 55.13% carried, at least, one germinal IEI-related-variant. Of those affected, 39.53% carried variants with confirmed or predicted pathogenicity (LP/P), which, overall, implies that more than 20% of the 78 SID patients had LP/P mutations. Very interestingly, we also observed that an almost identical frequency of a cohort of adult PID patients (53.57%) carried germinal mutations in known IEI-related genes, also with a similar yield of LP/P variants (26.80%). This genetic yield is consistent with other studies that have performed NGS in PID patients [18, 19, 20]. Beyond that, we found that all IEI variants were heterozygous in our SID cohort and significantly higher incidence than in the GP. These findings show that B-CLPD patients with SID are

genetically and phenotypically closer to PID patients than to the GP, strengthening the hypothesis that mutations in IEI genes could play a key role in their disease.

Even though SIDs with B-CLPD and PID patients seem to have similar rates of IEI mutations, we observed significant differences in the patterns of inheritance that could explain their different clinical phenotypes. Gene expression depends on a high variety of variables and mechanisms such as splicing, epigenomic modifications, the effect of other regulatory genes, chromosome skewing, mosaicism, or the appearance of secondary somatic mutations, just to mention some. However, to make a exploratory approach into the genetic inheritance of IEI variants in our cohort of patients through 3 classical variables of gene expression: 1) the predicted pathogenicity of the mutations (LP/P, VUS); 2) the type of inheritance of the genes affected (AD vs AR); and 3) the presence of co-occurrence. We found out that PID patients carry more single LP/P mutations in AD genes, while SIDs more single variants in AR genes. However, SIDs also had an unexpected higher rate of co-occurrence of LP/P mutations in AR genes, and of VUS mutations AD genes

In terms of gene ontogeny and the biological pathways affected by the mutations, we studied how the genetic variants were distributed among the IEI groups between both cohorts. Most of the genes detected in the B-CLPD patients fell under 4 IEI categories. The distribution of variants broadly mirrored that observed in the PID cohort, with SCID/CID, PAD and phagocyte defects remaining the predominant categories. Interestingly, SID was enriched for PIRD-associated variants, whereas autoinflammatory disorders were underrepresented. Moreover, B-CLPD patients showed a significantly higher prevalence of LP/P variants affecting group 4 ($p < 0.001$), and a significantly lower one for group 3 ($p < 0.05$) compared to the PID cohort.

Summing up, SID to B-CLPD carry more variants related to immune dysregulation, while in PIDs, genes associated to antibody production roles were much more affected. Also, the patterns of inheritance were neatly different. Although preliminary, our data suggest that these differences reinforce, rather than undermine, the hypothesis of a biological overlap between PID and SID associated with B-CLPD. B-CLPD patients with SID belong to the PID spectrum, where differences could explain the different clinical behavior, with later disease onsets and a higher burden of malignancy. What appears to be clear is that the increased presence of mutational co-occurrence and the different combinations between predicted pathogenicity and the types of inheritance point out to a much more intricate genetic landscape than that of monogenic IEI/PIDs, converging to a more pleiotropic phenotype of PID.

We may also speculate that germline IEI variants in heterozygosis could fit the hypothesis of ‘oncogenic competence’ [21], where their presence increases the odds of developing lymphoid malignancy as well as immunodeficiency, but without severely affecting the fitness of the cells, as Schmitz et al.[22] describes for some somatic variants. In other words, they could represent a first ‘hit’, and other factors, such as: the types of genes affected (IEI groups), their genetic landscape (type of inheritance, pathogenicity and co-occurrence), different environmental exposures leading to changes in the epigenome and somatic mutations, and the increased mutational rate cell burden of lymphoid cells[2]; all could contribute to a

second ‘hit’ with definite consequences on the cells’ ‘fitness’, shaping the penetrance and expression of the germline mutations and leading to specific phenotypes.

In our cohort of B-CLPD patients most variants belonged to IEI groups 1 + 2 and 4. The *in silico* functional enrichment analysis also reproduced similar results, showing two major gene clusters that could be correlated with those two IEI groups of genes: ‘combined immunodeficiencies-like phenotypes and altered DNA repair mechanisms’, and ‘diseases of immune dysregulation’(Fig. 4).

These mutations, when present in homozygosis, lead to known life-threatening immunodeficiencies, typically in the first years of life. However, in heterozygosis, and in co-occurrence with other variants, they may act as predisposing factors for cellular dysfunction and increased inflammation. Chronic mild T- and/or B-cells dysfunction, as well as defects in pathways related to inflammation, cytotoxicity and stress responses could increase the odds of lymphomagenesis, which could, therefore, worsen the underlying immunodeficiency, as well, by adding more dysfunction to the whole immune system.

Mutations related to group 3 (antibody deficiencies) have also been linked to a higher risk of lymphomagenesis[15,23, 24] mostly, in relation to chronic viral and EBV exposure[2] which can lead to uncontrolled B-cell lymphoproliferation. However, in our study, we observed that these variants were clearly more represented in the PID cohort. Could this mean that the mutations in these genes have a lower mutagenesis burden, or that they specifically need the aid of a viral agent to increase that risk? We do not know. However, we hypothesize that, maybe, in the case of these types of genes, the theory of ‘oncogenic competence’ could be switched to one of ‘infectious competence’, where the second ‘hit’ appears as a consequence of an already settled immunodeficiency that increases the infectious vulnerability and the mutagenic risk of the host.

More studies are warranted to confirm our results, which cannot set direct correlations between the genes affected and the clinical picture of the patients studied. Overall, our findings demonstrate that a substantial proportion of patients with B-CLPD harbor variants in genes classically associated with severe IEI and PIRD, at a frequency exceeding that observed in the general population. These results support the existence of a shared genetic landscape between B-CLPD and PID and raise fundamental questions for future research, including whether B-CLPD may represent part of the IEI spectrum and whether genetic profiling should complement, or even precede, clinical presentation in the diagnosis of immunodeficiency.

Abbreviations

ACMG
American College of Medical Genetics and Genomics
AI
artificial intelligence
AD

autosomal dominant
AR
autosomal recessive
B-CLPD
B-cell lymphoproliferative disorder
CID
combined immunodeficiency
CNV
copy number variant
CTLs
cytotoxic T-cells
GO
gene ontology
GP
general population
IEI
inborn errors of immunity
LP
likely pathogenic
LP/P
likely pathogenic / pathogenic
NHL
non-Hodgkin lymphoma
NGS
next generation sequencing
P
pathogenic
PAD
predominantly antibody deficiencies
PID
primary immunodeficiency
PIRD
primary immune regulatory disorders
SCID
severe combined immunodeficiency
sFLC
serum Free Light Chains
SID
secondary immunodeficiency
SNV

single nucleotide variant
VUS
variant of unknown significance

Declarations

Competing Interests

Silvia Sánchez-Ramón is an Associate Editor of this journal

Funding

Pharming ®

Author Contribution

MPO: performed the conceptualization and methodology, the writing of the main manuscript text, prepared all figures and took part in the review and editing process. BGS: took part in the conceptualization and investigation, and the review and editing process. TGG and MPG intervened in the formal analysis and investigation process. MDMR, AVM, EFM, MRR, RH, JOG: performed investigation, helped with resources. APC, EAM, MPZ, EBC, CPL, FM: provided resources, and validated the study. MFA, MGF, CBC, RPD: did conceptualization, supervision and validated the study. SSR: performed the conceptualization, methodology, investigation, writing of the original draft and the review and editing of the final manuscript, as well as, the supervision and project administration. All authors reviewed the manuscript.

Acknowledgement

We are immensely grateful for the patients and their families, because they have made possible this study. The authors would also like to acknowledge the nursing staff of the Hospital Day Care of the Hospital Clínico San Carlos, for their excellent care of our patients, and their thoughtful help; and Dr. José María García-Aznar from Health In Code company, for his kind predisposition and his help in the genetic analysis.

Data Availability

All the main data supporting the findings of this study are available within the paper and its Supplementary Information. Any other data that might have not been included is available under request.

Genetic data from the general population are publicly available at GenIE (Genetic Prevalence and Incidence Estimator) software (URL: https://genie.broadinstitute.org/)

References

1. Tangye SG et al. Human Inborn Errors of Immunity: 2022 Update on the Classification from the International Union of Immunological Societies Expert Committee. *Journal of Clinical Immunology* vol. 42Springer US, (2022).
2. Riaz I, Bin, Faridi W, Patnaik MM, Abraham RS. A systematic review on predisposition to lymphoid (B and T cell) neoplasias in patients with primary immunodeficiencies and immune dysregulatory disorders (inborn errors of immunity). Front Immunol 10, (2019).
3. Sakura F, et al. A complementary approach for genetic diagnosis of inborn errors of immunity using proteogenomic analysis. PNAS Nexus. 2023;2:1–12.
4. Abolhassani H, et al. Global systematic review of primary immunodeficiency registries. Expert Rev Clin Immunol. 2020;16:717–32.
5. Kienzler AK, Hargreaves CE, Patel SY. The role of genomics in common variable immunodeficiency disorders. Clin Exp Immunol. 2017;188:326–32.
6. Palacios-Ortega M et al. Dissecting Secondary Immunodeficiency: Identification of Primary Immunodeficiency within B-Cell Lymphoproliferative Disorders. J Clin Immunol 45, (2025).
7. Alaggio R, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. Leukemia. 2022;36:1720–48.
8. Jolles S, Chapel H, Litzman J. When to initiate immunoglobulin replacement therapy (IGRT) in antibody deficiency: a practical approach. Clin Exp Immunol. 2017;188:333–41.
9. Baxter SM et al. The Power of Partnership: Democratizing Genetic Prevalence to Empower Patient Advocacy. (2026) 10.64898/2026.03.30.26349539
10. Richards S, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17:405–24.
11. Mighton C, et al. Points to consider for the reporting of variants of uncertain significance in germline genetic and genomic testing: A statement of the American College of Medical Genetics and Genomics (ACMG). Genet Med. 2026;102583. 10.1016/j.gim.2026.102583.
12. Tavtigian SV, et al. Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines. Hum Mutat. 2021;41:1734–7.
13. Jüri Reimand R, Isserlin V, Voisin M, Kucera C, Tannus-Lopes A, Rostamianfar L, Wadi M, Meyer J, Wong C, Xu DM. D. B. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nat Protoc. 2019;14:482–517.
14. James M. Some methods for classification and analysis of multivariate observations. *Proc. fifth Berkeley Symp. Math. Stat. Probab.* 1, 281–297 (1967).

15. Ye X, et al. Genomic characterization of lymphomas in patients with inborn errors of immunity. Blood Adv. 2022;6:5403–14.
16. Guevara-Hoyer K, et al. Genomic crossroads between non-Hodgkin's lymphoma and common variable immunodeficiency. Front Immunol. 2022;13:1–15.
17. Allain V, et al. Lymphoma as an Exclusion Criteria for CVID Diagnosis Revisited. J Clin Immunol. 2023;43:181–91.
18. Chinn IK, Orange JS. A 2020 update on the use of genetic testing for patients with primary immunodeficiency. Expert Rev Clin Immunol. 2020;16:897–909.
19. Similuk MN, et al. Clinical exome sequencing of 1000 families with complex immune phenotypes: Toward comprehensive genomic evaluations. J Allergy Clin Immunol. 2022;150:947–54.
20. Vorsteveld EE et al. Clinical exome sequencing data from patients with inborn errors of immunity: Cohort level diagnostic yield and the benefit of systematic reanalysis. Clin Immunol 268, (2024).
21. Pavinato L, Baggiolini A. Oncogenic competence: balancing mutations, cellular state, and microenvironment. Trends Cancer. 2025;11:276–85.
22. Schmitz EG, Griffith M, Griffith OL, Cooper MA. Identifying genetic errors of immunity due to mosaicism. J Exp Med. 2025;222:1–15.
23. Hauck F, Voss R, Urban C, Seidel MG. Intrinsic and extrinsic causes of malignancies in patients with primary immunodeficiency disorders. J Allergy Clin Immunol. 2018;141:59–e684.
24. Mayor PC, et al. Cancer in primary immunodeficiency diseases: Cancer incidence in the United States Immune Deficiency Network Registry. J Allergy Clin Immunol. 2018;141:1028–35.

Figures

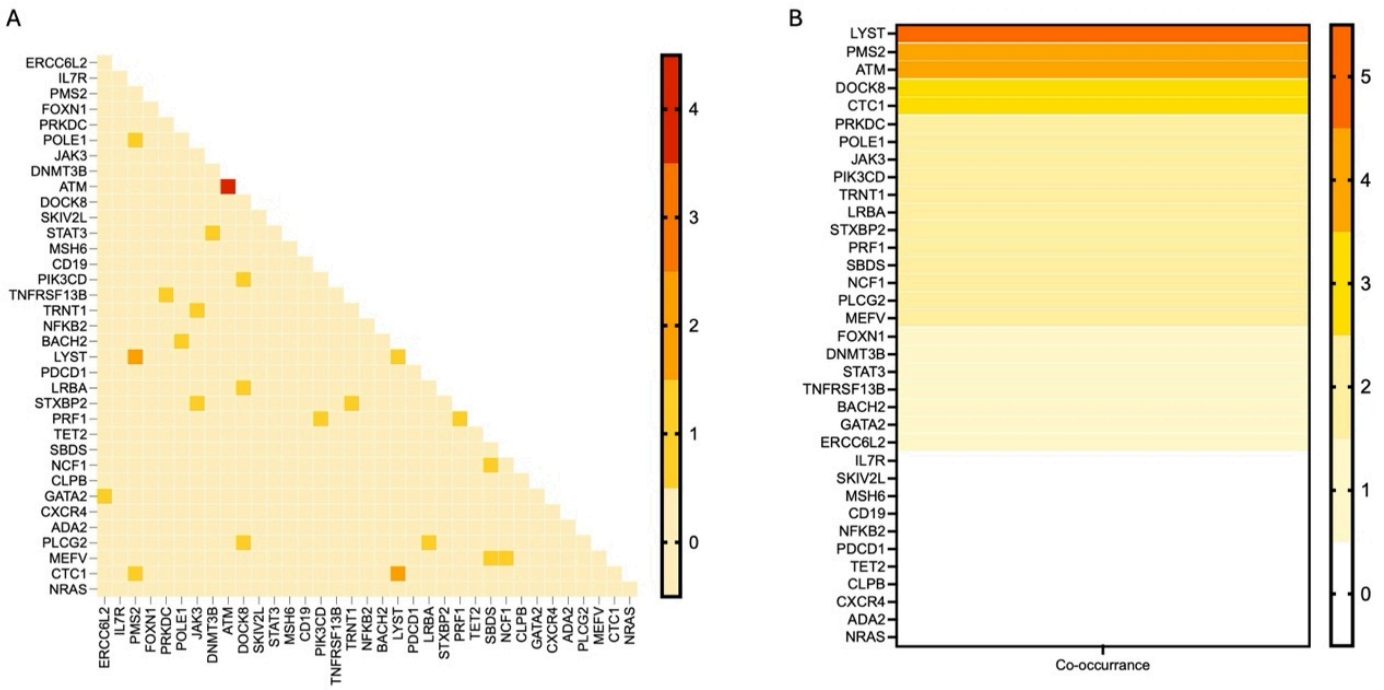


Figure 1. Co-occurrence of IEI gene variants in the SID cohort. (A) Matrix showing the number of times each gene variant appears in co-occurrence with other/s mutations in other genes. (B) Heat-map representing the rates of mutational co-occurrence of every gene variant in the SID patient cohort. The color gradient represents the number of events.

Figure 1

See image above for figure legend

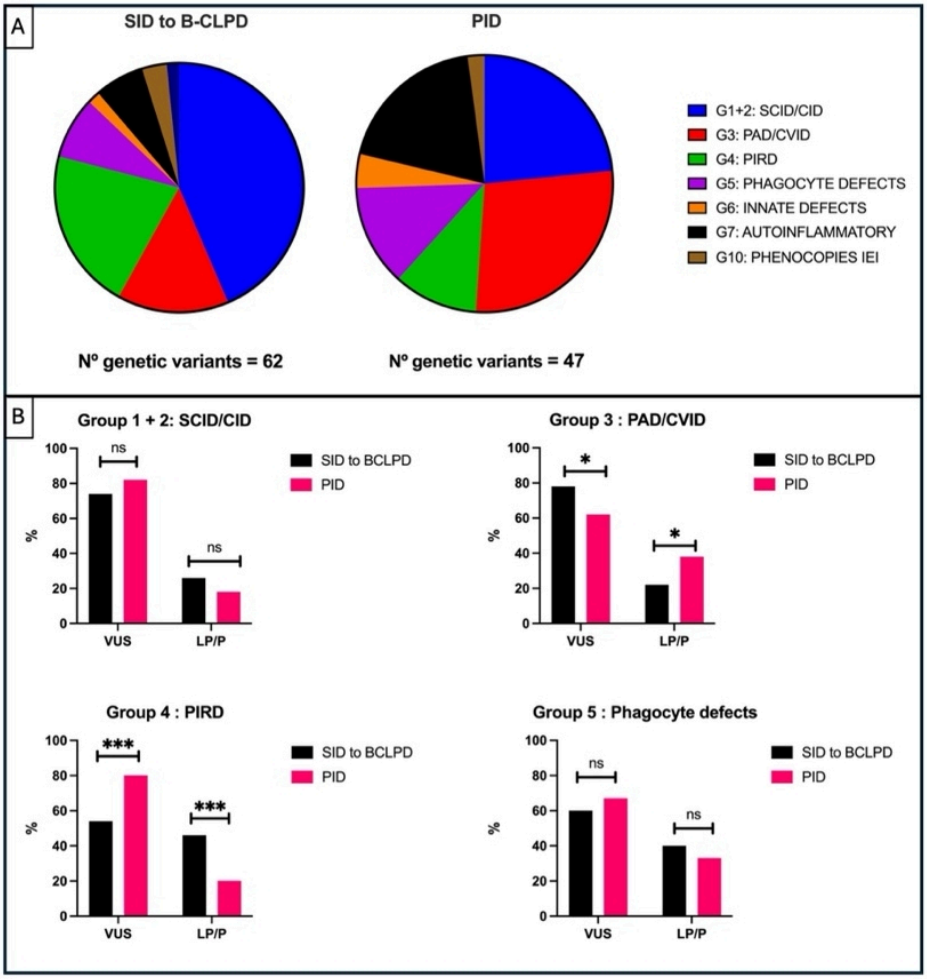


Figure 2. Comparison of the IEI variants between the cohorts of B-CLPD with SID and PID. (A) Representation and comparison of the numeric distribution of IEI variants between the SID to BCLPD and PID patient cohorts. We observed a similar pattern in the distribution of the variants between both groups. However, the cohort of SID patients had a greater representation of variants related to combined immunodeficiencies (group 1+2) and immune dysregulation (group 4), while in PID there were more genetic mutations in relation to antibody deficiencies (group 3) and autoinflammatory disorders; (B) Comparison of VUS and LP/P variants of IEI groups 1+2, 3, 4 and 5, between the cohorts of SID to B-CLPD and PID. The percentage of genetic variants found in IEI groups 1 and 2, and group 5 are similar between our patients with SID to B-CLPD and PID. Groups 3 and 4, however, were found to have a significant difference between both groups. There was a greater presence of LP/P group 3 variants in patients with PID, but higher of group 4 in SID to B-CLPD. $p < 0.05$ is represented as *; $p < 0.001$ is represented as ***; no significant difference is shown as 'ns'.

Figure 2

See image above for figure legend

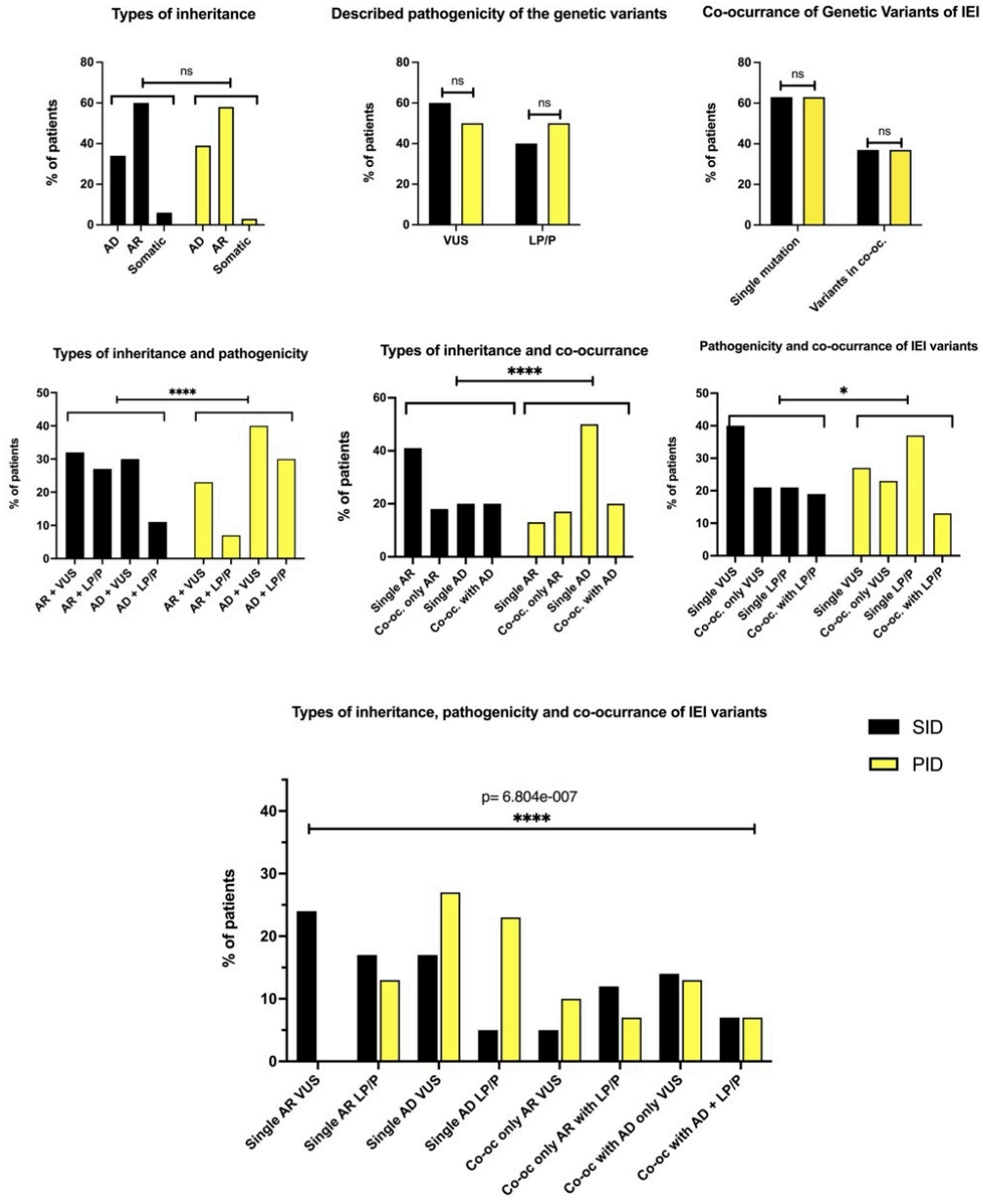


Figure 3. Comparison of the types of mutations found between SID to B-CLPD and PID patients. There was no significant difference between the groups when studying, separately, the type of inheritance, the presented pathogenicity and the presence of mutation co-occurrence in the patients (represented by the three graphs in the top of the figure). However, when studying the different patterns and combinations the three mutation variables present in the patients, we could observe significant differences between both groups (represented by the rest of the graphs). Percentages of SID patients are shown in the black bars; and percentages of PID patients are shown un yellow bars. $p < 0.05$ is represented as *; $p < 0.01$ is represented as **; $p < 0.001$ is represented as ***; $p < 0.0001$ is represented as ****; no significant difference is shown as 'ns'.

Figure 3

See image above for figure legend

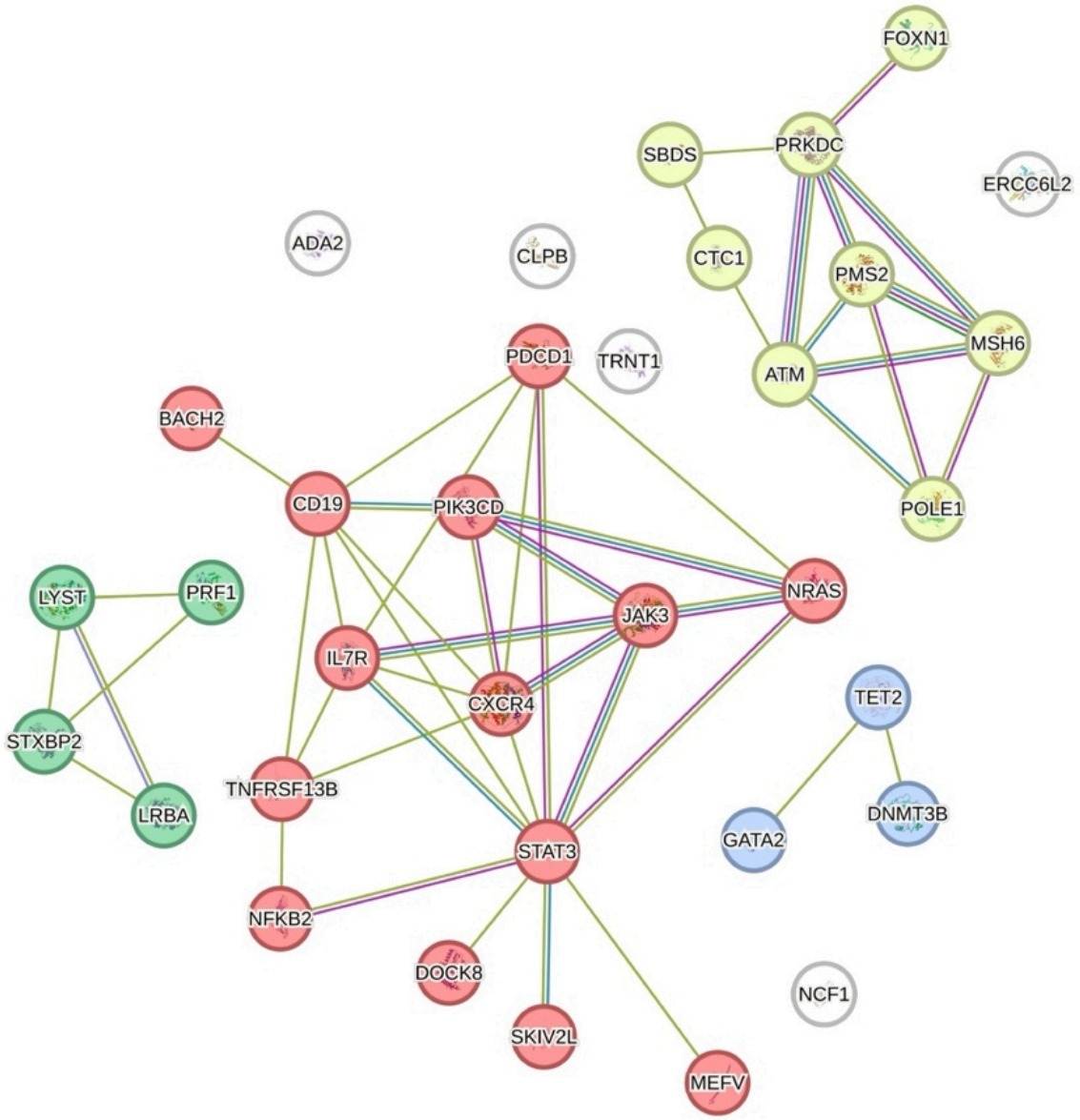


Figure 4. Clustering of the 35 genes by functional enrichment analysis in relation to biological processes. Obtained with String software. Color legend: Blue (DNMT3B, TET2, GATA2): DNA methylation or demethylation / transcriptional regulation of pluripotent cells; Yellow (FOXN1, PRKDC, SBDS, CTC1, POLE, PMS2, MSH6, ATM): somatic diversification of immune receptors, DNA repair complex; Green (PRF1, STXBP2, LYST, LRBA): leukocyte mediated cytotoxicity, degranulation, vesicle trafficking; Red (IL7R, JAK3, NRAS, PIK3CD, CXCR4, DOCK8, STAT3, NFKB2, TNFRSF13B, CD19, PDCD1, BACH2, SKIV2L, MEFV): primary immunodeficiency, lymphocyte activation and differentiation, B-cell receptor signaling pathway, Negative regulation of lymphocyte apoptotic process.

Figure 4

See image above for figure legend

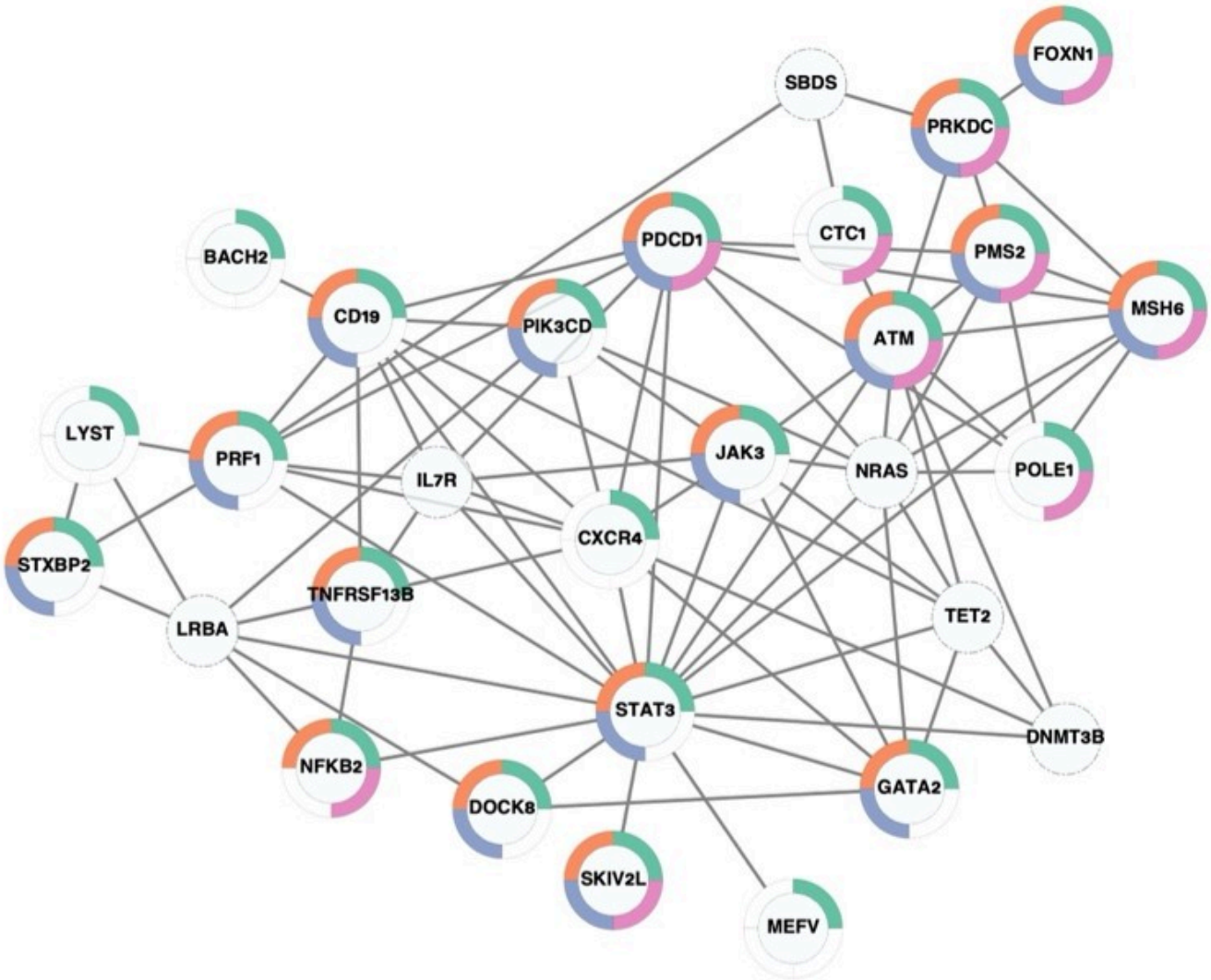


Figure 5: **Gene ontology enrichment analysis**. Obtained with *Cytoscape* software. Color legend: blue: 'leukocyte activation' (p-value = 1.36 E-11), orange: 'cell activation' (p-value = 5.03 E-12); green: 'immune system process' (p-value = 1.01 E-11); and pink: "immune system development" (p-value = 1.10 E-10). p-value correlates with the hypergeometric p-value after correction for multiple testing and describes how significant the enrichment is.

Figure 5

See image above for figure legend

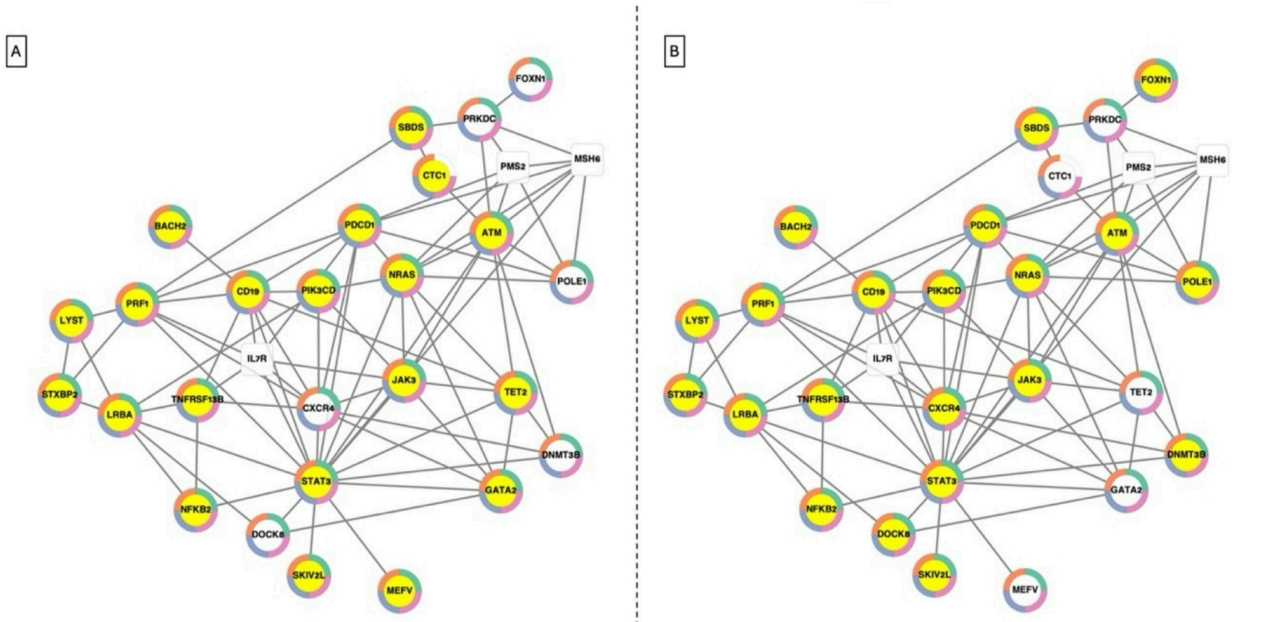


Figure 6. Phenotype enrichment analysis. (A) Association with 'Recurrent infections' phenotype. Yellow filling of nodes: 'Recurrent infections' (p-value: 2.44 E-13). (B) Association with 'Haematological malignancy' phenotype. Yellow filling of nodes: 'Lymphoma' and 'Haematological malignancy' (p-values = 2.74 e-13 and p = 1.17 E-11). / Color legend A and B: green: 'abnormal leukocyte count' (p-value = 2.51 E-19); orange: 'abnormal leukocyte morphology' (p-value = 1.22 E-18); blue: abnormal cellular immune system morphology' (p-value = 1.22 E-18), and pink: 'abnormal immune system morphology' (p-value: 3.61 E-18). P-value correlates with the hypergeometric P-value after correction for multiple testing and describes how significant the enrichment is. Obtained with *Cytoscape software*.

Figure 6

See image above for figure legend

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- graphicalabstractpatternsofinheritanceinIEI.jpg
- SUPLEMENTARYFIGURES.docx