Defective peripheral B cell selection in common variable immune deficiency patients with autoimmune manifestations

Highlights

- Central B cell tolerance is intact in common variable immunodeficiency (CVID)
- Peripheral B cell selection is defective in CVID patients with autoimmune disease
- The process of somatic hypermutation works suboptimally in CVID
- Activated naive B cells only partially induce the mismatch repair genes in CVID

Authors

Vanda Friman, Isabella Quinti, Alexey N. Davydov, ..., Dmitriy M. Chudakov, Rita Carsetti, Ola Grimsholm

Correspondence

ola.grimsholm@meduniwien.ac.at

In brief

Friman et al. show that the central tolerance in patients with common variable immunodeficiency (CVID) is intact whereas the peripheral selection of CD27 bright memory B cells is defective, particularly in patients with autoimmune manifestations. This could potentially be one of the underlying reasons of autoimmunity in this patient group.

Fruman et al., 2023, Cell Reports 42, 112446
May 30, 2023 © 2023 The Author(s).
https://doi.org/10.1016/j.celrep.2023.112446
Defective peripheral B cell selection in common variable immune deficiency patients with autoimmune manifestations

Vanda Friman,1,22 Isabella Quinti,2,22 Alexey N. Davydov,3,22 Mikhail Shugay,4,5,6,22 Chiara Farroni,7,6,22 Erik Engström,9 Shirin Pour Akaber,10 Sabina Barresi,11 Ahmed Mohamed,9,12 Federica Pulvirenti,13 Cinzia Milito,14 Guido Granata,14 Ezio Giorda,15 Sara Ahlström,1 Johanna Karlsson,1 Emiliano Marasco,16 Valentina Marcellini,16 Chiara Bocci,6 Simona Cascioi,18 Marco Scarsella,18 Ganesh Phad,17 Andreas Tilevik,18 Marco Tartaglia,11 Mats Bemark,19,20 Dmitry M. Chudakov,4,5,6,3 Rita Carsetti,8,21 and Ola Grimsholm10,9,8,23,*

1Department of Infectious Diseases, Institute of Biomedicine, The Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
2Department of Molecular Medicine, Sapienza University of Rome, Rome, Italy
3Central European Institute of Technology, Brno, Czech Republic
4Department of Life Sciences, Skolkovo Institute of Science and Technology, Moscow, Russia
5Institute of Translational Medicine, Pirogov Russian National Research Medical University, Moscow, Russia
6Translational Research Unit, National Institute for Infectious Diseases Lazzaro Spallanzani (IRCCS), Rome, Italy
7B Cell Pathophysiology Unit, Immunology Research Area, Ospedale Pediatrico Bambino Gesù IRCCS, Rome, Italy
8Institute of Rheumatology and Inflammation Research, The Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
9Institute of Pathophysiology and Allergy Research, Centre for Pathophysiology, Infectiology, and Immunology, Medical University of Vienna, Vienna, Austria
10Genetics and Rare Diseases Research Division, Ospedale Pediatrico Bambino Gesù IRCCS, Rome, Italy
11Centre for Primary Immune Deficiency, AUO Policlinico Umberto I, Rome, Italy
12Clinical and Research Department for Infectious Diseases, National Institute for Infectious Diseases L. Spallanzani (IRCCS), 00149 Rome, Italy
13Research Laboratories, Ospedale Pediatrico Bambino Gesù, IRCCS, Rome, Italy
14Division of Rheumatology, Ospedale Pediatrico Bambino Gesù, IRCCS, Rome, Italy
15Institute for Research in Biomedicine (IRB), Università della Svizzera Italiana (USI), Bellinzona, Switzerland
16School of Bioscience, University of Skövde, Skövde, Sweden
17Department of Microbiology and Immunology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
18Department of Clinical Immunology and Transfusion Medicine, Region Västra Götaland, Sahlgrenska University Hospital, Gothenburg, Sweden
20Unit of Diagnostic Immunology, Department of Laboratories, Ospedale Pediatrico Bambino Gesù, IRCCS, Rome, Italy
21These authors contributed equally
22Lead contact
*Correspondence: ola.grimsholm@meduniwien.ac.at
https://doi.org/10.1016/j.celrep.2023.112446

SUMMARY

Common variable immune deficiency (CVID) is a heterogeneous disorder characterized by recurrent infections, low levels of serum immunoglobulins, and impaired vaccine responses. Autoimmune manifestations are common, but B cell central and peripheral selection mechanisms in CVID are incompletely understood. Here, we find that receptor editing, a measure of central tolerance, is increased in transitional B cells from CVID patients and that these cells have a higher immunoglobulin k:λ ratio in CVID patients with autoimmune manifestations than in those with infection only. Contrariwise, the selection pressure in the germinal center on CD27bright memory B cells is decreased in CVID patients with autoimmune manifestations. Finally, functionally, T cell-dependent activation showed that naive B cells in CVID patients are badly equipped for activation and induction of mismatch repair genes. We conclude that central tolerance is functional whereas peripheral selection is defective in CVID patients with autoimmune manifestations, which could underpin the development of autoimmunity.
INTRODUCTION

Antibodies (i.e., immunoglobulins [Igs]), are crucial components of the humoral immune system, contributing to neutralizing pathogens during infection and hindering re-infection. They are composed of two identical heavy chains (HCs) and two light chains (LCs) and are expressed in soluble and membrane-bound forms (B cell receptor [BCR]). The vast repertoire of BCRs in naive B cells is created through random recombination of Ig HC and LC loci. The variability of the BCR repertoire and the functionality of antibodies is increased during immune responses by two antigen-driven mechanisms: somatic hypermutation (SHM) and class switch recombination (CSR), respectively. The selection of a non-autoreactive and efficient BCR repertoire is ensured by tolerance mechanisms, including receptor editing in the bone marrow and antigen-driven selection in the germinal center (GC). Together they ensure that autoreactivity is properly censored in B cells. The Ig HC is considered to be the main determinant of antibody specificity; however, the contribution of the Ig LC to antigen binding should not be overlooked, and receptor editing often changes the specificity of the antibody.

The genes that encode the Ig LC, k, which comprises, in total, 37–42 functional V genes and 5 J genes, and \( \lambda \), which comprises 37–43 functional V genes and 4–5 J genes, are found on two different chromosomes. \( \kappa \) is usually rearranged before \( \lambda \), and in peripheral blood total B cells, the \( \kappa:\lambda \) ratio is normally in the range of 1.5–2.0 in humans. However, in certain diseases, such as human immunodeficiency virus (HIV) infection and systemic lupus erythematosus (SLE), this ratio is altered. It has also been shown that the physicochemical properties of the \( \kappa \) and \( \lambda \) complementarity-determining region 3 (CDR3), which is the most variable part of the LC (and HC) and in closest contact to the antigen, are significantly different from each other.

We have demonstrated that the Ig HC CDR3 is specifically selected in CD27null and post-GC CD27bright memory B cells (MBCs). In these cells, we detected sequentially, from naive via CD27null to CD27bright, shorter CDR3 lengths as well as alterations in the physicochemical properties, indicating that antigen selection had taken place. We also found that CD27bright MBCs have more somatic mutations in their HC CDRs compared with CD27null MBCs, which is in line with the fact that IgG-expressing MBCs are more mutated than those of the IgM isotype.

Common variable immune deficiency (CVID) is a heterogeneous disorder characterized by low levels of serum Igs and impaired production of specific antibodies following vaccination in the absence of any other known cause of hypogammaglobulinemia. The most common clinical presentation is recurrent or chronic sinopulmonary infection. However, the phenotype is complex and heterogeneous and may include, besides increased susceptibility to infection, additional features, such as lymphoproliferation and autoimmunity. Although most patients with CVID have normal or only slightly reduced numbers of peripheral B cells, the numbers of switched MBCs are often severely reduced. Our recent findings showed that an analysis of Ig sequences from CD27bright MBCs has potential for evaluating GC function. Hence, it may also be used for evaluating peripheral selection in patients with primary immunodeficiencies, such as CVID. Knowledge regarding central and peripheral B cell tolerance and selection in this patient group is still very limited.

Earlier studies have reported inconsistent results regarding secondary recombination at the \( \kappa \) locus, giving conflicting views on whether receptor editing is functional in patients with CVID. The method to determine receptor editing developed by Panizrahi et al. intron recombining sequence (IRS)-RS recombination frequency, not only determines central tolerance but also acts as a surrogate marker for peripheral selection during development from the transitional to naive B cell stages. Another study indicated that the process of SHM on the Ig LC might be defective in CVID, but in this case, only two patients were studied with respect to sorted CD27+ MBCs.

Here we examined Ig LC expression during peripheral B cell differentiation, using receptor editing as a measure for the level of central tolerance and Ig LC sequencing of naive and CD27bright MBCs as a measure of peripheral selection of B cells in patients with CVID (divided into two groups: those with only infection and those with autoimmune manifestations) and compared this with healthy donors (HDs). Finally, we performed RNA sequencing on activated naive B cells. We found that the Ig \( \kappa:\lambda \) ratio is increased in transitional B cells from CVID patients with autoimmune manifestations compared with those with infection only. Receptor editing, as measured by IRS-RS recombination frequency, is increased in \( \lambda \) transitional B cells from CVID patients. Furthermore, sequence analysis of Ig LCs from naive and post-GC CD27bright MBCs demonstrates that differentiation of naive into CD27bright MBC pool is aberrant in CVID patients with autoimmune manifestations. The levels of somatic mutations in CD27bright MBCs are dramatically reduced and replacement mutations are heavily affected in CVID patients with autoimmune manifestations compared with HDs. Furthermore, RNA sequencing (RNA-seq) analysis indicates that activation of naive B cells is impaired during full BCR engagement with T cell-dependent stimuli and that mismatch repair genes are not induced properly in CVID patients’ naive B cells. This, together with the lack of selection into the CD27bright MBC pool, could be one of the reasons why patients with CVID are susceptible to development of autoimmunity.

RESULTS

Ig LC expression changes throughout B cell development and differentiation in healthy subjects

In earlier studies conducted in HDs and different diseases (e.g., SLE), the Ig \( \kappa:\lambda \) ratio (1.5–2.0) has been investigated within the total CD19+ B cell compartment. This approach makes it difficult to evaluate at which development stage the selection takes place. In our HDs, the Ig \( \kappa:\lambda \) ratios for the total CD19+ B cell compartment were in line with data reported previously in the literature (Figures 1A and 1B). To investigate the Ig \( \kappa:\lambda \) ratio in greater detail, we next determined it in different HD peripheral blood B cell subsets (transitional, naive, and CD27bright MBCs) (Figure S1). No differences in the Ig \( \kappa:\lambda \) ratio were observed when comparing transitional and naive B cells in HDs (Figure 1C). However, CD27bright MBCs expressed a higher Ig \( \kappa:\lambda \) ratio than naive B cells, which was evident in some
individuals (Figure 1C). Thus, Igκ is, at least in some individuals, more frequently used in CD27bright MBCs than in naive B cells in HDs.

The Igκ:λ ratio is higher in transitional B cells of patients with CVID who suffer from an autoimmune disease

Next, we determined the Igκ:λ ratio on the same B cell subsets from patients with CVID (Table S1). Naive B cells in patients with CVID expressed a lower Igκ:λ ratio compared with transitional B cells in the same samples, but unlike in HDs, no difference was observed between the naive B cell and CD27bright MBC population (Figure 2A). When the patient group was divided into those who had an autoimmune disease (n = 16) and those who did not (n = 23), a comparison between the groups revealed a difference in the Igκ:λ ratio in transitional B cells, with higher values in those with an autoimmune disease (Figure 2B). The Igκ:λ ratio decreased from the transitional to the naive B cell compartment in both groups (Figure 2B). We conclude that λ-expressing B cells are more prevalent in the naive than the transitional compartment in CVID patients, indicating that more edited B cells are allowed to progress along the maturation path. However, in CVID patients who suffer from an autoimmune manifestation, the early part of the process is altered, with fewer λ-expressing B cells present in the transitional compartment, but there is still counter-selection of λ-expressing cells for entry into the naive compartment.

The levels of iRS-RS rearrangement in transitional B cells are increased in patients with CVID compared with HDs

To further investigate Ig LC selection, we explored a mechanism that might explain the increased proportion of λ-expressing naive B cells in subjects with CVID: receptor editing. We decided to use an assay that was developed by Panigrahi et al. that measures deletion of the κ locus, indicating subsequent secondary rearrangement of the second κ allele and/or λ expression. We compared the overall levels of RS rearrangement from transitional and naive B cells (κ+ and λ+ sorted separately) between patients with CVID and HDs. We found that the frequency of iRS-RS rearrangement, in line with previous data, was higher in λ+ than in κ+ in transitional B cells from HDs and patients with CVID (for the sorting strategy, see Figure S1A; Figure 3A). In detail, we observed that the levels of iRS-RS rearrangement in λ+ transitional B cells were increased and a trend toward increased levels in κ+ transitional B cells (p = 0.08) in patients with CVID compared with those in HDs. Next, we investigated whether there were differences between patients with CVID who had an autoimmune manifestation and those with only infection. We found that the
levels of iRS-RS rearrangement were higher in $\lambda^+$ transitional B cells among patients with CVID and autoimmune manifestation compared with their $\lambda^+$ naive B cell counterparts (Figure 3B). Overall, we conclude that the frequency of iRS-RS rearrangement in $\lambda^+$ transitional B cells is increased in patients with CVID, indicating increased receptor editing in the bone marrow in these patients, presumably as a means to censor autoreactivity. However, we cannot exclude that it could also be due to more non-functional rearrangements. Furthermore, the decrease in the levels of IRS-RS rearrangement from $\lambda^+$ transitional to naive B cells in patients with CVID indicate that edited cells are counter-selected between these two stages of B cell differentiation.

Peripheral selection on Ig L-CDR3 length and properties is dysfunctional in CVID patients with autoimmune disease

Next, to examine the peripheral selection of B cells, we sorted naive and CD27$^{bright}$ MBCs of CVID patients with only infection or with autoimmune manifestations (Table S2) and HDs and sequenced the Ig LC genes. The number of rearrangements, unique sequences, and clonotypes derived from 50,000 naive B cells and 30,000 CD27$^{bright}$ MBCs, respectively, were overall comparable between the different groups (Table S3). To evaluate differences in the variable-to-joining (VJ) rearrangement structures of the immunoglobulin kappa (IGK) and immunoglobulin lambda (IGL) light chains between naive and CD27$^{bright}$ MBCs of HDs and patients with CVID, we first computed the length distribution of CDR3, which has been associated previously with B cell selection.18,19 We observed shortening of the LC-CDR3s in CD27$^{bright}$ MBCs compared with the naive ones in HDs, which is in line with the effect observed previously on immunoglobulin heavy (IGH) chains (Figure 4A). For CVID patients, however, LC-CDR3 shortening was only observed in those with only infection, although the length profiles of CD27$^{bright}$ MBCs and naive B cells in CVID patients with autoimmune disease were similar.

We next assessed four basic physicochemical properties (hydrophathy, charge, polarity, and volume) of the amino acid compositions of the IGK and IGL CDR3s in all samples. We observed significant differences in the charge (in IGK sequences; $p = 0.0033637$) and volume (in IGL sequences; $p = 0.0282570$) of the CDR3 amino acids across B cell subset:donor status combinations (Figure 4B). Interestingly, in almost all cases, the CD27$^{bright}$ MBCs in HDs were selected compared with their naive counterpart, whereas there was little difference between naive and CD27$^{bright}$ MBCs in CVID patients with autoimmune manifestations (Figure 4B). There is, however, a lot of heterogeneity in CVID patients. The selection pattern in CVID patients with only infection resembled that of HDs and that of patients with autoimmune manifestations; thus, it was somewhat in between the two. These results can be attributed to selection of CD27$^{bright}$ MBCs in the GC based on specific features of their IGK and IGL CDR3s but not in patients CVID with autoimmune manifestations. This means that counter-selection of bulky
and sticky (charged) CDR3 variants, with increased risk of being autoreactive, is defective in CD27bright MBCs from CVID patients with autoimmune manifestations.

We also analyzed the V gene compositions of the IGK and IGL chains by performing hierarchical clustering of samples based on their gene frequency profiles. We found no clear preferential co-clustering of V gene profiles in samples having the same status (HD/CVID without autoimmune disease [AD]/CVID with AD) or derived from the same B cell subset. Furthermore, we did not observe any differences in Jκ usage or Jλ usage in naive or CD27bright MBCs between HDs and the groups of patients with CVID (Figure S2). Usage of Jκ5 can be used as a measure of induction of secondary recombination of the κ locus and, thus, functional receptor editing, and we observed that this usage was similar in naive B cells when comparing HDs with patients with CVID.

Thus, peripheral selection on the LC-CDR3 properties and length is defective in CD27bright MBCs in CVID patients with autoimmune manifestations, preserving clones with more charged, bulky (volume), and, thus, potentially more sticky CDR3 variants often seen in autoreactive clones.

Ig LC repertoire diversification and selection is altered in CVID patients with autoimmune manifestations

To further examine diversification and selection of the Ig LC repertoire, we computed a rank-size distribution for the CDR3 amino acid sequence variants of the IGK and IGL chains of naive B and CD27bright MBCs from HDs and the two groups of patients with CVID. We observed differences in the frequency distribution of top (ranking from 1–100) CDR3 variants for HDs but not for any of the two groups of CVID patients (Figure 5A).

Ig LC repertoire diversification and selection is altered in CVID patients with autoimmune manifestations

To further examine diversification and selection of the Ig LC repertoire, we computed a rank-size distribution for the CDR3 amino acid sequence variants of the IGK and IGL chains of naive B and CD27bright MBCs from HDs and the two groups of patients with CVID. We observed differences in the frequency distribution of top (ranking from 1–100) CDR3 variants for HDs but not for any of the two groups of CVID patients (Figure 5A).
Next, we observed a substantial diversification of the repertoire in naïve-to-CD27 bright transition for HDs, which may seem counter-intuitive at first glance and is the opposite of what is observed for T cells. A lower degree of diversity and the presence of high-frequency CDR3 variants in the naive B-cell subset can be attributed to convergent rearrangement effects that become evident due to an overall low number of possible LC-CDR3 rearrangements compared with HC-CDR3. Diversification of the CD27 bright MBC repertoire can be attributed to specific selection and, most importantly, to acquisition of SHMs, as demonstrated below. Note that no IGK and IGL repertoire diversification was observed for CVID patients with autoimmune manifestations and that the diversity indices of both subsets in these patients were similar to the values observed for naïve B cells of HDs (Figure 5B). Patients with only infection showed a pattern similar to HDs. Thus, this would indicate that selection into the CD27 bright MBC population is defective in CVID patients with autoimmune manifestations.

We investigated CDR3 variant sharing between CD27 bright and naïve B cell subsets, observing a significantly lower level of sharing in HD subjects compared with CVID subjects with autoimmune manifestations (Figure 5C).

We also computed CDR3 variant sharing within B cell subsets across HDs and patients with CVID. We observed substantially higher degrees of IGK and IGL CDR3 variant sharing in the naïve subset compared with the CD27 bright subset, indicating that CD27 bright MBCs represent a more specialized subset that is specific for each donor compared with highly promiscuous naïve
B cells in terms of IGK and IGL repertoires (Figure 5D). Sharing in the naive B cell subset across donors was lower in the CVID patients with autoimmune manifestations on IGK and IGL CDR3 variants. Furthermore, the difference in IGK and IGL clonotype sharing comparing naive and CD27bright MBCs was less in CVID patients with autoimmune manifestations, where we even see an overlap between the two populations, whereas in that of HDs and CVID patients with only infections, a clear separation was always seen from naive to CD27nerg bright MBCs (Figure 5D). However, because we did still observe a change in clonotype sharing between naive B and CD27bright MBCs in CVID patients with autoimmune manifestations, this would indicate that, although the diversification mechanism is impaired in these, a unique selection mechanism operates to create a specialized subset of CD27bright MBCs that is distinctive for each patient.

Deficient process of SHM of the Ig LC in CD27bright MBCs from CVID patients with autoimmune manifestations

We reasoned that the decreased selection of CDR3 properties and less diversification of the CD27bright MBC repertoire could be due to a dysfunctional GC reaction in CVID. Therefore, to evaluate the function of the GC, we analyzed the frequency of SHM and mutation pattern (replacement-to-silent mutation ratio) in CD27bright MBCs from HDs and the two groups of patients with CVID as above. We have shown recently that IgM- and IgG-expressing CD27bright MBCs are frequently mutated in HD CDRs, as demonstrated for switched CD27*, IgM* IgD*, and IgM-only MBCs by numerous groups. Here, we analyzed the SHM rates in LC sequences of CD27bright MBCs, finding a significant decrease in SHM in CD27bright MBCs of patients with CVID in those with only infection and those with autoimmune manifestations compared with those of HDs (Figure 6A). The number of SHMs for patients with CVID was similar to the background level observed for naive B cells of HDs, although it was still higher than that in naive B cells of patients (the fact that naive B cells have a non-zero number of SHMs can be attributed to contamination with CD27* IgG-expressing MBCs).

Next, to compare the mutation pattern between CD27bright MBCs from HDs and CVID patients, we deciphered the types of nucleotide substitutions that have occurred; i.e., whether a replacement of amino acid occurred. Our sequencing depth revealed clear differences in the mutation spectrum (Figure 6B). Notably, we observed that the replacement-to-silent mutation ratio was significantly reduced, especially in CD27bright MBCs from CVID patients with autoimmune manifestations. These results would indicate that the process of SHM to replace amino acids to increase affinity in CD27bright MBCs is severely impaired in CVID patients with autoimmune manifestations.

Smaller lineage trees are more frequent in CVID patients with autoimmune manifestations

Next, to visualize the mutation process in CD27bright MBCs in HDs and disease groups, we performed a lineage tree analysis. For this purpose, we used a recently released online tool developed by Hoehn et al., called the Dowser tool. The trees from HDs show an expected clonal expansion with many layers in the bigger clones. We observed that CD27bright MBCs in HDs contained more nodes than both groups of CVID patients, although it was only significant between HDs and CVID patients with autoimmune manifestations (Figure 6C, left panel). However, we did not observe any difference between the two patient groups. We also observed that the CD27bright MBCs in HDs contained more trees with singletons; the difference was only significant when compared with the patients with autoimmune manifestations (Figure 6C, center panel). Finally, we calculated the normalized tree length, which showed that both groups of CVID patients had a significantly shorter normalized tree branch length (see STAR Methods for definition) (Figure 6C, right panel). We also visualized the lineage tree analysis, and it showed that the mutation process is inefficient in most patients with CVID (Figure 6D), but in combination with the fact that replacement mutations are less frequent in patients with autoimmune manifestations, we conclude that the GC reaction in the latter patient group is the most defective.

In summary, although the process of SHM can be induced in patients with CVID who have CD27bright MBCs, the process can most likely not be sustained, resulting in a severely reduced frequency of SHM and changed mutation pattern with reduced replacement mutations in CVID patients with autoimmune manifestations. Together, this would indicate that the GC reaction is defective in patients with CVID.

Differentially expressed genes in the B cell activation pathway and suboptimal induction of mismatch repair genes in patients with CVID

In an attempt to shed light on the mechanisms behind the defective process of SHM and possibly non-functional GCs in CVID (separating into two groups as previously), we sorted naive B cells and activated these with a stimulation cocktail that would induce optimal B cell activation through BCR engagement and stimuli involved in T cell-dependent interactions. Hence, the cells were incubated with a cocktail with anti-Ig, CD40L, and interleukin-21 (IL-21) and harvested the cultures 48 h later for transcriptomic analysis using RNA-seq. A recent study reported transcriptomic differences observed mainly in activated MBCs, with only minor changes in the naive B cell subset, in the peripheral blood mononuclear cell (PBMC) fraction when comparing HDs and CVID patients. We started out by first identifying transcriptomic differences between activated HD-derived naive B cells and those of CVID patients, and an unsupervised principal-component analysis (PCA) plot using the 120 differentially expressed genes (DEGs) (Tables S4; S5; adjusted p < 0.05, log2 fold change > 2) indicated that activated naive B cells from HDs cluster closely to each other, whereas those from the two groups of CVID patients did, in almost all cases, cluster separately from the HDs, but there was also a clear heterogeneity within the two CVID patient groups (Figure 7A). Next, we visualized the 50 most up- or down-regulated genes in a heatmap (Figure 7B). We noticed that many of the genes are induced/ repressed during B cell activation; e.g., CXCR4 and CD79B but also MSH6, which is involved in mismatch repair during the process of SHM (Figure 7C). MSH2 is also shown, but the difference was not significant. Furthermore, we found several genes involved in proliferation that were downregulated in CVID patients; e.g., PCNA and CINP (Figure 7D). Next, we focused our attention on the comparison between CVID patients with only
infection with those with autoimmune manifestations, but only minor differences were observed between them (Figure 7C). To examine whether the DEGs we found meant something more functionally, we performed a pathway analysis using Gene Ontology (GO) terms (similar results were obtained with the Kyoto Encyclopedia of Genes and Genomes [KEGG] database) on the DEGs found between the HDs and CVID patients, and this confirmed that many genes belonged to pathways involving genes important for B cell activation, lymphocyte proliferation, and mismatch repair machinery (Figure 7E). In summary, the RNA-seq analysis supports our data from the Ig sequencing because B cell activation-related genes are affected in patients with CVID, and failure to upregulate MSH6 and the mismatch repair pathway might influence the process of SHM in vivo.

DISCUSSION

In the present study, we show that the peripheral B cell selection mechanisms, which normally ensure that tolerance to self is induced, are incomplete in patients with CVID. This was exemplified by the absent shortening of CDR3, reduced repertoire diversity, reduced SHM, and, importantly, the reduced number of
A

Act. naïve B

- HD
- CVID w/o. AD
- CVID w. AD

PC2: 75% variance

PC1: 74% variance

B

HD3
CVID49 w/o. AD
HD1
HD2
HD4
CVID50 w/o. AD
CVID52 w/o. AD
CVID54 w. AD
CVID56 w. AD
CVID61 w. AD
CVID51 w/o. AD
CVID54 w. AD
CVID48 w/o. AD
CVID47 w. AD
CVID40 w. AD

C

IRF4

CD79B

IL2RG

MSH6

MSH2

D

PCNA

CINP

E

Pathway analysis (GO terms)

- Adaptive immune response 35/744
- Regulation of lymphocyte activation 25/573
- B-cell mediated immunity 17/218
- B cell receptor signaling pathway 9/130
- Regulation of lymphocyte proliferation 9/239
- Mismatch repair 3/38

(legend on next page)
replacement mutations in CD27bright MBCs from CVID patients with autoimmune manifestations (Figure S3). This would indicate that defective selection in the GC might be one of the reasons why these patients have an increased risk of developing autoimmunity.

B cell selection starts already in the bone marrow (BM), and here we assessed this in CVID patients by detecting secondary recombination on the Ig LC in a process called receptor editing.3 Our approach of detecting deletion of the κ locus provides strong evidence that the secondary rearrangement is intact or increased in patients with CVID. In this case, we could not detect any major differences between patients with autoimmune symptoms and those with only infection, apart from the fact that λ+ naive B cells were less edited than their transitional B cell counterparts in patients with autoimmune symptoms. This is probably attributable to the counter-selection of edited λ+ cells in the naïve compartment because of autoreactivity. These results are strengthened by the fact that usage of Jκ5, which is the most downstream Jκ segment, is similar in naïve B cells from patients with CVID and HDs, indicating functional secondary recombination (Figure S2). Earlier data have provided conflicting results regarding central selection mechanisms such as receptor editing in patients with CVID,28,29 and it is still possible that certain patients might have a reduced capacity to induce secondary recombination of the Ig LC because our cohort was limited in sample size.

Peripheral B cell selection acts at several stages during B cell differentiation,10,11 but here we focused on what would be imposed on antigen-experienced B cells (i.e., MBCs), most likely within the GC reaction. Our data on LC-CDR3 length clearly demonstrated that there is a selection of CD27bright MBCs expressing shorter CDR3 in HDs, as we showed previously for the HC selection,19 whereas this selection is defective in CVID patients with autoimmune manifestations. The physicochemical properties of the CDR3 are only clearly selected in CD27bright MBCs from HDs, indicating that V segment usage is not the main driving force (as was expected) but that, instead, the CDR3 amino acids dictate this selection.36 Furthermore, the selection of shorter CDR3s as well as amino acid properties that are less bulky and hydrophobic give rise to a selected CD27bright MBC repertoire with less sticky (less charged) and cross-reactive Ig variants,19 as shown for T cell receptors (TCRs).37,38 It has been shown earlier that the immunoglobulin heavy chain V gene (VH) repertoire of naïve B cells in patients with CVID has a reduced richness.39 Our analysis of Vκ and Vλ usage did not reveal any major differences in the naïve B cell repertoire, but we did observe differences in the diversity of the repertoire between naïve and CD27bright MBCs in patients compared with HDs. In the latter, the changes in clonotype sharing between naïve B and CD27bright MBCs was always clearly different where the sharing was much lower in the MBC repertoire between donors. On the other hand, in patients, the changes from naïve to CD27bright MBCs were smaller and also with larger variation. Sharing between patients in the CD27bright MBC repertoire was higher, indicating that the diversification mechanisms work less efficiently in CVID with autoimmune manifestations; this also points toward a defective GC reaction because the MBC repertoire normally shows little overlap between individuals, mirroring the different immune responses we go through during our lives.40,41 Additionally, in HDs, we observed an increased diversity in the CD27bright MBC repertoire, most probably because of the fact that SHMs are introduced and that the LC repertoire is less diverse than the HC repertoire in general, thereby creating new Ig variants because of the SHMs. However, these Ig variants in the CD27bright MBC repertoire are not shared between donors in HDs or in patients with CVID. Thus, although SHM frequency is severely reduced in patients with CVID, they do express Ig variants in the CD27bright MBC repertoire that are different between donors. As expected, the naïve B cell repertoire is more shared between donors, especially in HDs but also in patients with CVID, although to a significantly lesser extent.

Affinity maturation through the process of SHM is crucial for immune homeostasis of an individual because it ensures that re-infection of previously encountered pathogens can be neutralized quickly. We have demonstrated previously that IgM and switched CD27bright MBCs are highly selected with an increased replacement-to-silent mutation ratio.19 Our current data clearly show that, on the LC, the SHM frequency is reduced in CD27bright MBCs in all CVID patients, regardless of the frequency of post-GC CD27bright MBCs in their peripheral blood or disease status, a fact that has also been suggested in a previous brief report on Ig LC mutation status in CVID,31 although few patients were included in that study. Previous data on the Ig HC in CD27+ MBCs from patients with CVID have shown that the SHM frequency is reduced in certain patients, especially those lacking switched MBCs, as well as in CVID patients with autoimmune cytopenia.35,36 In addition, in a study using PBMCs, where plasmablasts probably contribute substantially to what was observed, a reduction of SHM was observed in most CVID patients.43 We have demonstrated previously that the GC reaction is indispensable for formation of CD27bright MBCs, whereas SHM is not, as shown previously in CD27+...
MBCs. In light of this, our present results point in the direction that the GC reaction can be induced in patients with CVID who have circulating CD27\textsuperscript{bright} MBCs, but the process of SHM works inefficiently, especially in patients with autoimmune manifestations because of fewer replacement mutations in these. Fewer replacement mutations have also been observed in MBCs from several ADs; e.g., Sjögren’s syndrome. \cite{55} In line with this, a recent report on three CVID patients where MBCs from peripheral blood and lymph nodes were examined indicated that the number of replacement mutations was reduced in some of these. \cite{46} The development of autoimmunity in CVID patients is most likely caused by several factors, and it has been suggested previously that regulatory T cells are decreased in CVID patients.\cite{42,48} Another explanation for GC dysfunction in patients with CVID, although no apparent difference could be observed between the two patient groups in this analysis.\cite{51,52} Furthermore, looking at the topology of the lineage trees, it was also clear that CD27\textsuperscript{bright} MBCs from HDs formed larger trees than both of the patient groups, indicating reduced efficiency of the process of SHM in CVID, although no apparent difference could be observed between the two patient groups in this analysis.

Our results on SHM frequencies and mutation patterns are in accordance with recent reports indicating that the GC reaction does not function properly (i.e., hyperplastic appearance) in patients with CVID.\cite{55,56} Another explanation for GC dysfunction in patients with CVID could be that dark zone/light zone polarization is also disturbed in our patient cohort, as described earlier described by Unger et al.\cite{57} In our RNA-seq analysis on activated naive B cells, we observed that the mismatch repair pathway was altered (Figure S3), where, e.g., MSH6 was suboptimally induced, and this could potentially be one reason for the defective SHM machinery in CVID.\cite{58} Furthermore, MSH6 is also involved in the repair during CSR,\cite{59} and therefore our results could be one contributing reason why CSR is impaired in many patients with CVID.\cite{60,61}

In summary, we provide results that would argue that peripheral selection is defective in CVID patients with autoimmune manifestations. Furthermore, in vitro functional results supported our Ig sequencing results with reduced induction of the mismatch repair machinery. Our results can at least partly explain why patients with CVID are more prone to developing autoimmune manifestations.

**Limitations of the study**

A limitation of our Ig sequencing data in this study is, however, that we cannot exclude false clonal expansions of MBCs because it is possible that more than one RNA molecule from the same B cell can be included in the analysis. We want to point out, though, that similar amounts of sorted naive (50,000 cells) and CD27\textsuperscript{bright} MBCs (30,000 cells) were used from both HDs and CVID patients to create the Ig seq libraries.\cite{55} We also do realize that our study cohort on both the Ig sequencing and RNA-seq is limited and that a larger cohort might be necessary to definitively prove defective peripheral selection in CVID patients with autoimmune manifestations. In addition, a parallel investigation of regulatory T cells (Tregs) and T follicular helper (TFH) cell function might be of value in such an additional study.\cite{42,46}

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

*KEY RESOURCES TABLE*

*RESOURCE AVAILABILITY*

- Lead contact
- Materials availability
- Data and code availability

*EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS*

- Ethical statement
- Human samples

*METHOD DETAILS*

- Cell isolation and cell sorting
- Stimulation and reagents
- Flow cytometry and antibodies
- Real-time PCR
- Immunoglobulin light chain sequencing
- Bioinformatics analysis of Ig light chain sequencing data
- RNA sequencing
- RNA-seq bioinformatic analysis

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112446.

**ACKNOWLEDGMENTS**

This work has received funding from the European Union’s Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant agreement 754412 (to O.G.). O.G. was also supported by a personal fellowship from...
REFERENCES


AUTHOR CONTRIBUTIONS


DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: October 28, 2020
Revised: March 15, 2023
Accepted: April 13, 2023

REFERENCES

12 Cell Reports 42, 112446, May 30, 2023


### Key Resources Table

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD19 (Clone HIB19) PECy5</td>
<td>BD Biosciences</td>
<td>Cat#555414; RRID: AB_395814</td>
</tr>
<tr>
<td>CD19 (Clone SJ25C1) APCH7</td>
<td>BD Biosciences</td>
<td>Cat#643078; RRID: AB_1645470</td>
</tr>
<tr>
<td>CD19 (Clone SJ25C1) BV510</td>
<td>BD Biosciences</td>
<td>Cat#562947; RRID: AB_2737912</td>
</tr>
<tr>
<td>CD20 (Clone 2H7) FITC</td>
<td>EXBIO Praha</td>
<td>Cat#1F-638-T100; RRID: AB_10736214</td>
</tr>
<tr>
<td>CD24 (Clone ML5) BV711</td>
<td>BD Biosciences</td>
<td>Cat#563450; RRID: AB_2738164</td>
</tr>
<tr>
<td>CD27 (Clone M-T271) PE</td>
<td>BD Biosciences</td>
<td>Cat#557330; RRID: AB_395834</td>
</tr>
<tr>
<td>CD27 (Clone M-T271) APC</td>
<td>BD Biosciences</td>
<td>Cat#561786; RRID: AB_395834</td>
</tr>
<tr>
<td>CD27 (Clone L128) BV605</td>
<td>BD Biosciences</td>
<td>Cat#562655; RRID: AB_2744351</td>
</tr>
<tr>
<td>CD38 (Clone HB7) PE-Cy7</td>
<td>BD Biosciences</td>
<td>Cat#335790; RRID: AB_399969</td>
</tr>
<tr>
<td>CD38 (Clone Hit2) BV421</td>
<td>BD Biosciences</td>
<td>Cat#562445; RRID: AB_11151894</td>
</tr>
<tr>
<td>IgM/IgG/IgA (Polyclonal)</td>
<td>Jackson Immunoresearch</td>
<td>Cat#109-606-064; RRID: AB_2337548</td>
</tr>
<tr>
<td>IgKappa FITC (polyclonal)</td>
<td>Dako</td>
<td>Cat#F0198; RRID: AB_2335707</td>
</tr>
<tr>
<td>IgLambda PE (polyclonal)</td>
<td>Dako</td>
<td>Cat#R0437; RRID: AB_579536</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Biological samples</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Human peripheral blood</td>
<td>Bambino Gesù Children Hospital, Rome, Italy. Sapienza University of Rome, Rome, Italy. Sahlgrenska University Hospital, Gothenburg Sweden</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chemicals, peptides, and recombinant proteins</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SMARTScribe Reverse Transcriptase</td>
<td>Clontech</td>
<td>Cat#639538</td>
</tr>
<tr>
<td>rhIL-2</td>
<td>Peprotech</td>
<td>Cat#200-21</td>
</tr>
<tr>
<td>rhCD40L</td>
<td>Enzo Life Science</td>
<td>Cat#ALX-522-015-C010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Critical commercial assays</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RNeasy Plus Micro Kit</td>
<td>Qiagen</td>
<td>Cat#74034</td>
</tr>
<tr>
<td>AllPrep DNA/RNA Micro Kit</td>
<td>Qiagen</td>
<td>Cat#80284</td>
</tr>
<tr>
<td>Qubit RNA HS Assay Kit</td>
<td>ThermoFisher Scientific</td>
<td>Cat#Q32855</td>
</tr>
<tr>
<td>Agencourt AMPure XP beads</td>
<td>Beckman Coulter</td>
<td>Cat#A63880</td>
</tr>
<tr>
<td>Agilent DNA 1000 chip</td>
<td>Agilent Technologies</td>
<td>Cat#5067-1504</td>
</tr>
<tr>
<td>Qubit dsDNA HS Assay Kit</td>
<td>ThermoFisher Scientific</td>
<td>Cat#Q32854</td>
</tr>
<tr>
<td>MiSeq reagent Kit v3 (600 cycles)</td>
<td>Illumina</td>
<td>Cat#MS-102-3003</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Deposited data</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ig LC repertoire in HD and CVID</td>
<td>This paper</td>
<td><a href="https://figshare.com/s/d7937088b4b9cbf3ec827">https://figshare.com/s/d7937088b4b9cbf3ec827</a></td>
</tr>
<tr>
<td>RNA sequencing in HD and CVID</td>
<td>This paper</td>
<td>NCBI-SRA (GSE227377)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Oligonucleotides</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTB</td>
<td>IDT</td>
<td>HS.PT.39a.22214847</td>
</tr>
<tr>
<td>iRS-KDE Primer 1: ATTAGATGCTGCCGTAGCC</td>
<td>IDT</td>
<td>N/A</td>
</tr>
<tr>
<td>Primer 2: AGGGCTTCTTAGGGAGGTCA</td>
<td>IDT</td>
<td>N/A</td>
</tr>
<tr>
<td>Probe: TCTGCGAGCTGATTTTTGCCA</td>
<td>IDT</td>
<td>N/A</td>
</tr>
<tr>
<td>SmartNNNext: AGATGUGTAUAAGAGACAGNNNUNNNNU</td>
<td>IDT</td>
<td>N/A</td>
</tr>
<tr>
<td>NNNNUCTT(rG)r4</td>
<td>IDT</td>
<td>N/A</td>
</tr>
<tr>
<td>hIgK_r1: GCGTTATCCACCTTCC</td>
<td>IDT</td>
<td>N/A</td>
</tr>
<tr>
<td>hIgL_r1: GCTCCCCGGTGAAGT</td>
<td>IDT</td>
<td>N/A</td>
</tr>
</tbody>
</table>

(Continued on next page)
RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact: Ola Grimsholm, ola.grimsholm@meduniwien.ac.at.

Materials availability
The information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact. This study did not generate new unique reagents.

Data and code availability
- The Ig sequencing data that support the findings of this study is openly available in figshare at https://figshare.com/s/d7937088b4b9cbf3c827. RNA-seq data are available in NCBI-SRA, and the reference number of the paper is GSE227377.
- The associated scripts are available on https://github.com/Grimsholmlab/Friman-et-al.-Cell-Reports-2023. The documentation for each of the softwares is publicly available at the respective website (see softwares in the key resources table).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Ethical statement
Informed consent was obtained for all experiments and the study was performed following the guidelines of the Declaration of Helsinki. This study was approved by the Institutional Review Boards of Gothenburg (Gothenburg, Sweden) and of Sapienza University of Rome (Rome, Italy).
Peripheral blood samples from in total 54 patients with CVID were used (the patients' demographics are listed in Tables S1 and S2 (for Human samples with CVID that expressed a clear CD27 bright MBC population to be able to assess the germinal center function (see Table S2). The CVID who had (n = 10) or did not have (n = 10) an autoimmune disease. For the Ig LC sequencing we selected eleven of the patients anonymous buffy coats (30–60 years of age). We also compared the Ig LC expression and receptor editing profiles of the patients with the national guidelines. CVID participants could also be treated with additional drugs following consolidated clinical practice and were on intravenous or subcutaneous immunoglobulin substitution therapy with trough IgG serum levels above 500 mg/dL according to those included in the Ig LC sequencing and/or RNA sequencing). All patients in this study were thoroughly clinically examined and were on intravenous or subcutaneous immunoglobulin substitution therapy with trough IgG serum levels above 500 mg/dL according to the national guidelines. CVID participants could also be treated with additional drugs following consolidated clinical practice and guidelines. Patients with CVID were compared with adult healthy control subjects (n = 28). All material from healthy controls were from anonymous buffy coats (30–60 years of age). We also compared the Ig LC expression and receptor editing profiles of the patients with CVID who had (n = 10) or did not have (n = 10) an autoimmune disease. For the Ig LC sequencing we selected eleven of the patients with CVID that expressed a clear CD27 bright MBC population to be able to assess the germinal center function (see Table S2). The experimental protocols conformed to the ethical guidelines of the 1975 Declaration of Helsinki.

**METHOD DETAILS**

**Cell isolation and cell sorting**
Heparinized peripheral blood mononuclear cells (PBMCs) were isolated using Ficoll Paque Plus 206 (Amersham PharmaciaBiotech) density-gradient centrifugation. Buffy coats were incubated with the RosetteSep human B-cell enrichment antibody cocktail (STEMCELL Technologies), and thereafter the B cells were isolated by density-gradient centrifugation. Alternatively, B cells were isolated after Ficoll density-gradient centrifugation by negative enrichment with a B-cell cocktail (STEMCELL Technologies). The B cell subsets were then stained with conjugated antibodies directed against CD19 (or CD20), CD24, CD27 and CD38 (EXBIO Praha and BD Biosciences). The B-cell subsets were gated in the FACS machine as follows: transitional B cells as CD24++CD27 - mature-naive B cells as CD24+CD27 - CD38; and CD27 bright MBCs as CD24+CD27+ (for details of the gating strategy, see Figure S1). For the receptor editing qPCR experiments, we sorted Ig κ+ and Ig λ+ transitional and naive B cells. Sorting was performed using the FACSAria III cell sorter (BD Biosciences) or the Sony SH800 instrument (Sony Technologies). Sort purities were >99%.

**Stimulation and reagents**
Sorted naive B cells were stimulated 1 μg/mL rhCD40L (Enzo Life Sciences) plus 10 μg/mL anti-IgM/IgG/IgA (Jackson Immunoresearch) and 20 ng/mL rhIL-21 (Peprotech) in complete medium at a concentration of 2.5x10^6 cells/ml for 48 h.

**Flow cytometry and antibodies**
Cells were stained with the appropriate combination of fluorochrome-conjugated antibodies to identify B-cell subsets using standard techniques. The antibodies used in this study are listed in key resources table. Cells were acquired in the BD LSRFortessa X-20 instrument (BD Biosciences). The FACS data were analyzed with the FlowJo ver. 10 software (Treestar). Dead cells were excluded from analysis by side/forward scatter gating.

**Real-time PCR**
Genomic DNA was isolated from sorted transitional and naive (κ+ and λ+, separately) B cells using the AllPrep DNA/RNA Micro Kit (Qiagen). Quantitative PCR (40° C for 10 min, 95° C for 10 min, followed by 60 cycles of 95° C for 10 s, 60° C for 30 s, and 72° C for 1 s) was performed on 15–50 ng of template DNA. All samples were run in duplicate in a 20-μL reaction volume that contained 2x TaqMan Universal PCR Master Mix (ThermoFisher Scientific), 20× primers (Integrated DNA Technologies), 15–50 ng of DNA, and water. The primers for iRS-RS rearrangement and beta-actin control are shown in key resources table. The qPCR was run in the QuantStudio System. (ThermoFisher Scientific). Expression levels were calculated using the 2ΔΔCt method.

**Immunoglobulin light chain sequencing**
An RNA-based 5’-RACE with unique molecular identifiers (UMIs) sequencing approach was used. RNA samples from sorted naive B cells and CD27 bright MBCs of four HDs and eleven patients with CVID (see Table S2) were prepared as described by Turchaninova et al. We described this method in detail for Ig heavy chain sequencing recently. Total RNA was extracted with RNAeasy Micro kit (QIAGEN) according to manufacturer’s instructions. In brief, up to 700 ng of total RNA were used for cDNA synthesis. The cDNA synthesis primer mix (final concentration of 1 μM each; see key resources table) was incubated with RNA for 4 min at 70° C, and then for 2 min at 42° C. To this was added 1× first-strand Buffer (Clontech, final concentration 1×), DTT (SMARTScribe, 2 mM; Takara Bio), dNTPs (1 mM each), 10 × 5’-Template switch adapter (1 μM), and SMARTScribe reverse transcriptase (10 U/ml; Clontech), and the resulting mix was incubated for 60 min at 42° C. After adding 1 μL of uracil DNA glycosylase (5 U/μl; BioLabs) and incubating for 40 min at 37° C, the cDNA and subsequent PCR products were purified using Agencourt AMPure XP beads (Beckman Coulter).

From the cDNA reaction we used a quantity that corresponded to 30 000 CD27 bright MBCs and 50 000 naive B cells, respectively. We then performed a first PCR amplification (95° C for 90 s (1 x), followed by 95° C for 10 s, 54° C for 20 s, and 72° C for 40 s (18 x), with a final step of 4 min at 72° C). In this first PCR step, we used the reverse primer mix (0.4 μM) and common primer (0.4 μM).

A second PCR amplification was performed (95° C for 90 s (1 x), followed by 95° C for 10 s, 54° C for 20 s, and 72° C for 40 s (10 x), with a final step of 4 min at 72° C). Thereafter, 2 μL of purified product from the first PCR amplification were amplified using the forward common primer (0.8 μM) and reverse common primer (0.8 μM). In the final PCR step, purified DNA from the previous PCR of each
Bioinformatics analysis of Ig light chain sequencing data
The bioinformatics workflow has recently been described in detail. Briefly, de-multiplexing, UMI extraction, and UMI-based consensus assembling were performed using the MGEC software. The UMI-based analysis was previously shown to achieve nearly error-free immune repertoires. Multiplex primers for the IGLC segments were trimmed to avoid confusion with respect to isotype (kappa vs. lambda) identification. Further mapping of reads and clonotype assembly were performed using MiXCR_v2.1.10, as described previously, with some modifications to the MiXCR analysis pipeline (the KAligner alignment algorithm that detects indels longer than 2 nt).

Comparative analyses of IGL repertoire features, such as clonotype overlaps and CDR3 physicochemical properties, were performed using the VDJtools_v1.2.1 software. Of all the CDR3 amino acid properties available in the VDJtools software, those listed in the IMGT amino acid classification resource were used. We assigned each property to a given CDR3 by computing the average across all the amino acid residues. Mean values for every property were then computed by averaging across all the CDR3 amino acid variants present in each donor.

The normalized Shannon-Wiener diversity index implemented in the VDJtools software was computed as the entropy $H = -\sum_i f_i \log f_i$, where $f_i$ is the distribution of CDR3 amino acid sequence variant frequencies divided by log $5^N$, where $N$ is the number of unique variants. The normal overlap between CDR3 amino acid sequence variants was quantified using the F2 metric implemented in VDJtools. The metric was computed as: $F2 = \sum_i \sqrt{f_{i1}f_{i2}}$, where $f_{i1}$ and $f_{i2}$ are the frequencies of $i$ overlapping CDR3 variants in the first and second samples, respectively.

Allelic variants were identified as those supported by at least five clonotypes that were present in at least 40% of both molecules and clonotypes with a given Variable segment. They were filtered from the overall pool of detected mutations prior to the analysis of SHM frequencies. SHM information was obtained and parsed from the MiXCR output. We excluded allelic variants that had a frequency of 33% for a given Variable gene and that were supported by at least three unique CDR3 sequences.

Additional statistical analyses of the SHM frequencies and a comparison of the repertoire statistics inferred with VDJtools were performed using in-house R and Python scripts.

For lineage tree analysis we used the Dowser tool and performed the analysis as previously described in Aranburu et al., 2023. The top 1000 sequences in each sample was used, and the output from IMGT/HighV-QUEST was processed using Change-O version 1.2.0. reads parsed with MakeDb.py and germlines reconstructed with CreateGermlines.py. Reads with shared IGLV, IGLJ genes and L-CDR3 junction region length were grouped into clones using DefineClones.py. Nucleotide hamming distance of the resulting consensus sequences were incorporated at the adaptor ligation step for multiplex sample loading on the flow cells. The resulting constructs were purified by two consecutive AMPure XP beads cleanup steps and enriched by 15 cycles of PCR. The quality and quantity of the libraries were assessed using an Agilent DNA 1000 chip with Agilent Technologies and Qubit dsDNA HS Assay Kit, respectively. Fourteen libraries were pooled in equimolar amounts and loaded on flow cell at 1.2 pM (High Output Kit v 2.5 150 cycles, Illumina). The sequencing run was performed in paired-end mode (2 X 75-bp reads) using the Illumina NextSeq 500 platform. Base-calling was performed by the instrument computer using Illumina Real Time Analysis (RTA) software that is integrated with NextSeq Control Software (NCS) and provides a summary of quality statistics per illumina’s acceptance criteria for sequencing. CASAVA 1.8.2 was used for de-multiplexing and conversion of base calls to paired-end FASTQ files.

RNA sequencing
Total RNA was isolated from activated, sorted naive B cells using RNeasy Plus Micro kit (Qiagen). The concentration of total RNA extracted was quantified using the Qubit RNA HS Assay Kit (ThermoFischer Scientific). The Illumina Stranded mRNA Prep (Illumina) was used for isolation of polyadenylated mRNA with oligo-dT beads, second strand cDNA synthesis and NGS library preparation. During the second strand synthesis, dUTP was incorporated in place of dTTP, thus preventing amplification of this strand during the subsequent PCR step and retaining strand information. Unique indexes, included in the standard TruSeq kit set A (Illumina) were incorporated at the adaptor ligation step for multiplex sample loading on the flow cells. The resulting constructs were purified by two consecutive AMPure XP beads cleanup steps and enriched by 15 cycles of PCR. The quality and quantity of the libraries were assessed using an Agilent DNA 1000 chip with Agilent Technologies and Qubit dsDNA HS Assay Kit, respectively. Fourteen libraries were pooled in equimolar amounts and loaded on flow cell at 1.2 pM (High Output Kit v 2.5 150 cycles, Illumina). The sequencing run was performed in paired-end mode (2 X 75-bp reads) using the Illumina NextSeq 500 platform. Base-calling was performed by the instrument computer using Illumina Real Time Analysis (RTA) software that is integrated with NextSeq Control Software (NCS) and provides a summary of quality statistics per illumina’s acceptance criteria for sequencing. CASAVA 1.8.2 was used for de-multiplexing and conversion of base calls to paired-end FASTQ files.

RNA-seq bioinformatic analysis
The quality of FASTQ files were checked using FastQC High Throughput Sequence QC Report version 0.11.9 (Babraham bioinformatics, n.d.). The sequences were filtered using Trimmomatic tool version 0.39 and low-quality reads with phred cutoff of 20 and SLIDINGWINDOW:4:20 and reads with length less than 30 bases were excluded. The trimmed reads were aligned to the human...
reference genome (GRCh38/hg38) with Gencode 29 annotations (Human genome annotations, 2018) using STAR aligner version 2.6.1a in 2-pass mode. Raw reads per gene were counted by STAR. Differential gene expression analysis was conducted by using the DESeq2 package version 1.38.3 in R using vst function for normalization. Gene expression quantification was performed using featureCounts (included in the Subread package). Visual illustrations were generated using packages ggplot2, pheatmap (https://cran.r-project.org/web/packages/pheatmap/index.html). For DEG analysis adjusted p value of <0.05 was considered statistically significant. Packages biomaRt, edgeR, org.Hs.eg.db, and GO.db in R were used for Pathway analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS

All sample sizes can be found in the figure legends. Depending on if the datasets were normally distributed either Mann-Whitney nonparametric test or unpaired Student’s t test were used. A level of p < 0.05 was considered statistically significant. Basic characteristics of IG repertoires and somatic mutation frequencies were compared using the Tukey’s honest significant difference test. The physicochemical properties of IGK and IGL were analyzed using F-statistic and p values with 1-way ANOVA test. 1-Way ANOVA followed by t-tests were performed on the expression counts, comparing healthy donors with the two groups of CVID patients (only infection and autoimmune manifestations).