A multimorphic mutation in IRF4 causes human autosomal dominant combined immunodeficiency

Oriol Fornes, Alicia Jia, Hye Sun Kuehn, Qing Min, Ulrich Pannicke, Nikolai Schleussner, Romane Thouenon, Zhijia Yu, María de los Angeles Astbury, Catherine M Biggs, et al.

This is the accepted version of the following article: "IRF4 International Consortium; A multimorphic mutation in IRF4 causes human autosomal dominant combined immunodeficiency. Sci Immunol. 2023 Jan 20. ", which has been published in final format at doi: 10.1126/sciimmunol.ade7953.

Title: A multimorphic mutation in IRF4 causes human autosomal dominant combined immunodeficiency

Authors: IRF4 International Consortium[1*]

Affiliations:

[1]Consortium members and their affiliations are detailed at the end of the manuscript.

*Co-Corresponding authors email: anselm.enders@anu.edu.au, sven.kracker@inserm.fr, ruben.m.barricarte@vumc.org, stephan.mathas@charite.de, srosenzweig@cc.nih.gov, klaus.schwarz@uni-ulm.de, sturvey@bcchr.ca, wang@fudan.edu.cn

Abstract: Interferon regulatory factor 4 (IRF4) is a transcription factor (TF) and key regulator of immune-cell development and function. We report a recurrent heterozygous mutation in IRF4, p.T95R, causing an autosomal dominant combined immunodeficiency (CID) in 7 patients from 6 unrelated families. The patients exhibited profound susceptibility to opportunistic infections, notably *Pneumocystis jirovecii*, and presented with agammaglobulinemia. Patients' B cells showed impaired maturation, decreased immunoglobulin isotype switching, and defective plasma cell differentiation, whereas their T cells contained reduced Th17 and Tfh cells populations and exhibited decreased cytokine production. A knock-in mouse model of heterozygous T95R showed a severe defect in antibody production both at the steady state and after immunization with different types of antigens, consistent with the CID observed in these patients. The variant maps to the TF's DNA-binding domain, alters its canonical DNA-binding specificities, and results in a simultaneous multimorphic combination of dominant loss-, gain- and new-functions for IRF4. This previously undescribed multimorphic IRF4 pathophysiology disrupts normal lymphocyte biology causing human disease.

One-Sentence Summary: An IRF4 mutation with multimorphic impacts on DNA-binding specificity and activity reveals a previously undescribed distinct disease-causing mechanism.

Main Text:

Introduction

Inborn errors of immunity (IEI) are a heterogeneous group of gene defects characterized by increased susceptibility to infections, autoimmunity, autoinflammation, allergy, and in some cases cancer. During recent years, and following the advent of high-throughput sequencing, the rate of discovery of new genetic lesions causing IEI has skyrocketed (1). Yet, a large percentage of patients with an IEI lack a genetic diagnosis. Furthermore, due to the complexity of designing gene-tailored experiments to assess the pathogenicity of genetic variants and the scarcity of patients with similar genetic diseases, the molecular consequences of many IEI-causing mutations are poorly understood (2).

The human interferon regulatory factor (IRF) family is a group of nine transcription factors (TFs) that regulate critical innate and adaptive immune processes (3). The IRF family is typically characterized by their ability to recognize promoters containing the IRF consensus sequence (GAAA) (4). IRF4 is a predominantly hematopoietic TF that is pivotal for the development and function of B and T cells, as well as dendritic cells and macrophages (5-7). IRF4 binds to the canonical IFN-Stimulated Response Elements (ISRE) as a homodimer and it engages Erythroblast transformation-specific (Ets)-IRF Composite Elements (EICE) and activating protein 1 (AP-1)-IRF Composite Elements (AICE1 or 2) as a heterodimer requiring either PU.1 or SPIB, or AP-1 family members, respectively, for high-affinity interactions (5, 8, 9). *Irf4*$^{-/-}$ mice have severely impaired B and T cell activation and differentiation, along with a profound reduction in serum immunoglobulins of all isotypes (3).

Here, we describe a human immune disorder identified in 7 patients from 6 unrelated families across 4 continents. This newly described form of CID is caused by the same heterozygous germline mutation in the DNA-binding domain (DBD) of IRF4 (p.T95R), and mouse knock-in models accurately recapitulated the human phenotype. Particularly notable was the mechanism of this human disease with a simultaneous multimorphic combination of dominant loss-, gain- and new-functions for IRF4.

Results

Heterozygous T95R IRF4 variant identified in seven patients with profound combined immunodeficiency

We investigated 7 patients with profound CID from six kindreds of diverse ethnic origins (Fig. 1A). All affected individuals suffered with early onset (< 1 year of age) recurrent sinopulmonary infections with the opportunistic pathogen *Pneumocystis jirovecii* causing pneumonia in most individuals. Other clinical features indicative of their significant immune compromise were severe viral infections (notably cytomegalovirus (CMV) and Epstein-Barr virus (EBV)), localized infection with weakly pathogenic (i.e., vaccine-strain bacille Calmette-Guerin (BCG)) or pathogenic mycobacteria, (i.e., *M. bovis*), and chronic diarrhea (Fig. 1B and Table S1, patient summaries in Supplementary Materials and Methods). Next-generation sequencing was performed on all patients, with Sanger sequencing used for confirmation and familial segregation (Figs. S1A-S1B). All seven individuals carried the same heterozygous variant in the gene *IRF4* (*IRF4* NM_001195286.2 c.284C>G, p.T95R; henceforth designated IRF4^{T95R}) (Fig. 1C). This variant was *de novo* in all patients, except in the case of kindred six (P6 and P7) where the mother was found to be mosaic with 4/124 reads mapping to the variant in peripheral blood genetic testing

(Fig. S1B). The IRF4^{T95R} variant has not been reported in population databases of controls (Fig. S1C), and is predicted to be pathogenic by a variety of *in silico* pathogenicity models (Table S2). Interestingly, IRF4^{T95R} has been reported in the Catalogue of Somatic Mutations in Cancer (COSMIC) associated with adult T cell leukemia (10).

IRF4 mRNA and protein is expressed normally in IRF4^{T95R} patient lymphocytes but broad immune architecture is disrupted

As IRF4 is highly expressed in lymphocytes, we investigated how the IRF4^{T95R} variant affects total IRF4 mRNA and protein expression in these cells. We found that total *IRF4* mRNA expression was comparable to controls in activated CD4$^+$ T cells (Fig. 1D, left). Sanger sequencing of cDNAs extracted from both activated CD4$^+$ T cells (Fig. 1D, right) and EBV-B cells (Fig. 1E) suggested that the WT and mutant alleles are likely expressed at similar levels. We also found that patient IRF4 protein levels were comparable to controls in unstimulated CD3$^+$ T cells (Fig. 1F); CD19$^+$ B cells (Fig. 1G); expanded T cell blasts (Fig. S1D, left); activated CD4$^+$ T cells (Fig. S1D, right); and EBV-B cells (Fig. S1E). In combination, these results show that the T95R mutation does not affect total IRF4 mRNA and protein levels.

The infection pattern experienced by the individuals carrying IRF4^{T95R} was not consistent with a defect in just one immune compartment, suggesting a broad immunological phenotype. To define the overall immune architecture caused by IRF4^{T95R} in an unbiased manner, we applied cytometry by time of flight (CyTOF). Combining unsupervised analysis with manual clustering, we found that the distribution of B cells, and CD4$^+$ and CD8$^+$ T cells was different between healthy controls and patients (Figs. S2A-S2C). These data suggested that IRF4^{T95R} causes a broad immunological phenotype most notably impacting B and T cells which we set out to systematically decipher.

IRF4^{T95R} patients have multiple defects in B cell development and function

IRF4 is critical for B cell development and differentiation (7, 11). All patients were virtually agammaglobulinemic with severely reduced serum concentrations of IgM, IgG and IgA (Fig. 2A; Tables S3-S9). B cell enumeration revealed that all patients had reduced but not absent total CD19^{+} B cells with a developmental arrest characterized by increased naïve B cells, reduced class-switched memory B cells, and decreased plasmablasts (Fig. 2A; Tables S3-S9). The B cell differentiation defect was further refined in high dimensional space using CyTOF (Figs. 2B-2D; Figs. S2A-S2C), flow-cytometric analysis (Figs. S2D-S2E), and single-cell RNA-sequencing (scRNA-seq) on sorted CD19^{+} B cells (Figs. 2E-2F). Pseudotime inference indicated that IRF4^{T95R} transitional B (TrB) cells matured toward entirely different fates than those from a healthy control (Fig. 2E) and further confirmed an accumulation of TrB and naïve B cells with few terminally differentiated B cell subsets (i.e., memory B and plasma cells) in IRF4^{T95R} B cells (Fig. 2F). Moving beyond peripheral blood analysis, we confirmed that IRF4^{T95R} patient bone marrow (Fig. S2F) and lymph node (Fig. 2G) also had scant CD38^{+} or CD138^{+} plasmablasts.

Having demonstrated a marked impairment in B cell differentiation in patient peripheral blood using multiple complementary technologies (Figs. 2A-2D and S2A-S2E), we designed experiments to determine whether this defect was intrinsic or extrinsic to the IRF4^{T95R} B cells. To test this, we isolated naïve B cells from patients and controls and exposed them to classic B cell stimuli. When compared to healthy controls, purified naïve IRF4^{T95R} B cells or total PBMC generated less class-switched (Figs. 2H and S2G) and memory B cells (Fig. 2I), and the induction of plasmablasts (Figs. 2J-2K, S2H and S2J) as well as immunoglobulin secretion was lost or greatly reduced (Figs. 2L, S2I and S2K). Importantly, reduced B cell differentiation was not due

to impaired cell proliferation (Fig. S2L). Collectively, these data demonstrate that IRF4^{T95R} B cells have an intrinsic defect in differentiation into memory B cells, plasmablasts and plasma cells, as well as immunoglobulin secretion.

IRF4^{T95R} patient T cell differentiation is largely intact but effector function is abnormal

In contrast to the striking impact on B cell differentiation and biology, IRF4^{T95R} patients had relatively normal numbers of total T cells in peripheral blood (Fig. 3A). Appreciating that IRF4 has been linked to both CD4$^+$ and CD8$^+$ T cell effector function and that both *Pneumocystis jirovecii* infection and mycobacterial disease are not commonly found in patients with isolated B cell deficiencies (12), we proceeded with higher resolution phenotyping and functional interrogation. The proportions of naïve, central memory (CM), effector memory (EM) and EM re-expressing CD45RA (EMRA or TEMRA) T cells among the CD4$^+$ population were relatively normal (Fig. 3A, upper panels), whereas CD8$^+$ CM, EM and TEMRA cells were reduced in some patients (Fig. 3A, lower panels). In multidimensional space, both CyTOF and scRNA-seq revealed an abnormal distribution of CD4$^+$ and CD8$^+$ T cells (Figs. 3B-3D and S3A-S3B; Figs. S4-S5). Further functional studies confirmed that in IRF4^{T95R} T cells TCRVβ usage was not affected (Fig. S3C), and T cells were activated and proliferated normally upon TCR (i.e., anti-CD3/CD28) stimulation (Fig. S3D and 3E). Cytokine production is another critical effector function of T cells. Stimulated IRF4^{T95R} T cells produced less IL2 and IFNγ, and had less polyfunctional cells that produced both TNFα and IFNγ (Figs. 3F-3G). T helper (Th) cell commitment, evaluated by flow cytometry and CyTOF, showed that Th1, Th2, Treg and Tfr regulatory cells were similar to healthy controls (Fig. 3H; Table S10). In contrast, Th17 and Tfh cells were reduced in most patients (Fig. 3H; Table S10), and *in vitro* differentiation to Tfh cells was reduced (Figs. 3I-3J). When considered in the context of the patients' infectious history, these data indicate that IRF4^{T95R} T cells have a

defect in subset differentiation and function, although this defect is more subtle than that observed in B cells.

$IRF4^{T95R}$ patient monocyte and NK-cell evaluation showed no gross abnormalities

Monocyte and NK cells were evaluated as part of a comprehensive description of a unique IEI. Frequencies of classical, intermediate, and non-classical monocytes were within normal ranges in the patients analyzed (Fig. S3E; Table S11). CD107a degranulation studies of NK (Fig. S3F) and $CD8^+$ T cells (Fig. S3G) were also within normal ranges in the patients tested. Further and deeper studies will be required to rule out more nuanced defects in these cells.

Heterozygous p.T95R IRF4 knock-in mice recapitulate patient phenotype and provide additional insights into T95R pathophysiology

To obtain additional insights into the pathophysiological mechanisms of the disease caused by $IRF4^{T95R}$, we generated knock-in mice carrying the same p.T95R variant in the heterozygous state ($Irf4^{T95R/+}$) and compared them to $Irf4^{+/+}$, $Irf4^{+/-}$, and $Irf4^{-/-}$ mice. Replicating the patient phenotype, $Irf4^{T95R/+}$ splenocytes and B cells had IRF4 protein expression comparable to WT mice, whereas $Irf4^{+/-}$ had reduced protein expression and $Irf4^{-/-}$ were IRF4 protein deficient (Fig. 4A). Total numbers of B and T cell in the spleen (Fig. S6A) and of $B220^+$ B cells in the bone marrow and lymph nodes (Figs. S6B-S6C) were comparable among $Irf4^{+/+}$, $Irf4^{T95R/+}$, $Irf4^{+/-}$ and $Irf4^{-/-}$ mice, although there was a moderate increase in $B220^+$ cells in peripheral lymph nodes of $Irf4^{T95R/+}$ relative to $Irf4^{+/-}$ mice (Fig. S6C). $Irf4^{-/-}$ mice had barely detectable levels of serum immunoglobulins, while $Irf4^{+/-}$ mice had normal levels of serum IgA and total IgG, but reduced IgM, IgG_{2c} and IgG_3 levels (Fig. 4B). Notably, serum Ig levels were further reduced in $Irf4^{T95R/+}$ mice compared with $Irf4^{+/-}$ mice, and were very close to those found in $Irf4^{-/-}$ mice (Fig. 4B).

Similar to *Irf4*$^{-/-}$ mice, T95R homozygous mice (*Irf4*$^{T95R/T95R}$) had barely detectable levels of serum immunoglobulins (Fig. 4B). Total germinal center (GC) B cells were expanded in the spleen, mesenteric lymph nodes and Peyer's patches of *Irf4*$^{T95R/+}$ mice at the steady state (Fig. S6D) and in the spleen after immunization with sheep red blood cells (SRBC) (Fig. S6E) or *Plasmodium* sporozoites (Fig. S6F). Such expansions of GC B cells were not observed in *Irf4*$^{-/-}$mice consistent with previous studies (13-15) nor in *Irf4*$^{T95R/T95R}$ mice (Fig. S6D). In the *Plasmodium* sporozoite model, GC B cells specific for the immunodominant circumsporozoite protein (CSP) were decreased in *Irf4*$^{T95R/+}$ mice compared with *Irf4*$^{+/+}$ mice (Figs. 4C-4D), suggesting that there is an expansion of nonspecific GC B cells at the expense of antigen-specific GC B cells in *Irf4*$^{T95R/+}$ mice. Following *Plasmodium* immunization, *Irf4*$^{T95R/+}$ mice generated few malaria CSP-specific CD138^{+} plasmablasts (Figs. 4E-4F) and had lower levels of CSP-specific IgM and IgG antibodies (Fig. 4G). Further defining the antibody production defect, IgG_{2c} production after immunization with formalin-fixed *B. pertussis* (Th1-dependent) was reduced in *Irf4*$^{T95R/+}$ mice compared with *Irf4*$^{+/+}$ and *Irf4*$^{+/-}$ mice (Fig. 4H, left), and IgG_1 production against chicken γ-globulin (CGG) immunization (Th2-dependent) was also reduced in *Irf4*$^{T95R/+}$ mice, although the difference between *Irf4*$^{T95R/+}$ and *Irf4*$^{+/-}$ mice was not significant (Fig. 4H, right). Following immunization of 4-hydroxy-3-nitrophenyl-acetyl (NP)-conjugated to CGG (NP-CGG), the production of NP-specific IgM and both low- and high-affinity NP-specific IgG_1 antibodies were barely detectable in *Irf4*$^{T95R/+}$ mice compared with *Irf4*$^{+/-}$ mice (Fig. 4I). In line with earlier findings in *Irf4*$^{-/-}$ mice (16-18), *Irf4*$^{T95R/+}$ mice also showed altered distribution of various T cell subsets (Figs. S7A-S7C).

To clarify whether the impaired production of antibodies in *Irf4*$^{T95R/+}$ mice was due to B cell-intrinsic defects in plasma cell differentiation, we cultured purified splenic B cells in medium alone or in the presence of CD40L+IL4+IL21 (CI21) or LPS for 72h. The percentage of B220^{int}CD138^{+}

cells in cultured *Irf4*$^{T95R/+}$ B cells was lower than those in WT and *Irf4*$^{+/-}$ B cells, and were indistinguishable from those in *Irf4*$^{-/-}$ B cells (Fig. 4J). In addition, the frequency of antibody-forming cells (AFC) was reduced in *Irf4*$^{T95R/+}$ B cells compared with WT or *Irf4*$^{+/-}$ B cells (Figs. 4K-4L). Further defining the B cell defect, we found that when splenic B cells from WT or *Irf4*$^{T95R/+}$ Blimp1-GFP reporter mice were stimulated with LPS, *Irf4*$^{T95R/+}$ B cells were able to activate *Blimp1* transcription but failed to upregulate CD138 expression (Figs. S7D-S7E), suggesting that *Irf4*$^{T95R/+}$ B cells initiate, but fail to complete, plasma cell differentiation.

IRF4^{T95R} *B cells fail to undergo appropriate differentiation due to failure to activate IRF4 target genes critical for plasma cell differentiation*

To understand why IRF4^{T95R} B cells fail to appropriately differentiate into antibody-secreting plasma cells, we leveraged a variety of complementary model systems to assess different aspects of B cell development. First, mouse C57BL/6 splenic B cells were stimulated and transduced with IRF4^{WT}-, IRF4^{T95R}-, or control (MIG-ctrl)-retroviruses (Fig. 5A and S8A). Ectopic expression of IRF4^{WT} robustly enhanced formation of CD138^{high} B220^{low} plasmablasts compared to the MIG-control (Fig. 5A). In contrast, IRF4^{T95R} failed to enhance plasmablast formation (Fig. 5A). We next isolated IRF4^{WT}-, IRF4^{T95R}-, or MIG-control-transduced B cells for RNA-seq analyses. Overall, IRF4^{T95R}-transduced cells clustered separately from control or IRF4^{WT} cells, indicative of a distinct gene expression program (Fig. 5B). Specifically, IRF4^{T95R} regulated less (Fig. 5C, upper panel) and unique (Fig. 5C, lower panel) genes, with only a small overlap with genes differentially regulated by IRF4^{WT} (Figs. 5C and S8B; Data file S2). Through comparison to other gene expression profiles, we found that IRF4^{T95R} virtually lost the ability to regulate IRF4-driven and plasma cell-specific gene expression (Fig. 5D).

To further define the B cell defect, immunoglobulin class-switch recombination (CSR) was studied in CH12 cells transduced with different IRF4 variants (WT and T95R) or GFP alone without further activation, and GFP^+ cells were analyzed for the IgA^+ population. $IRF4^{T95R}$ and GFP alone failed to induce IgA^+ cells and failed to induce *Aicda* transcription (Figs. 5E-5G). To quantify plasma cell differentiation, we transduced the IRF4 variants into a human Burkitt's lymphoma cell line, Raji, which endogenously expresses extremely low levels of IRF4. $IRF4^{T95R}$ failed to upregulate genes encoding CD38, BLIMP1 and XBP1 (Figs. S8C-S8E). To assess antibody secretion, we utilized the fact that Ramos cells (human Burkitt's lymphoma cell line) spontaneously secret low levels of IgM. After disrupting the endogenous *IRF4* gene using CRISPR-Cas9-mediated genome editing (Fig. 5H), IRF4-deficient Ramos cells were transduced with the IRF4 variants, and $IRF4^{T95R}$ was unable to increase the frequency of antibody-secreting cells nor to enhance IgM secretion (Figs. 5I-5J and S8F-S8G). Co-expression of T95R with WT IRF4 in IRF4-deficient Ramos cells suppressed WT IRF4-induced plasma cell differentiation (Figs. S8H-J). In addition, ectopic expression of T95R, but not empty vector, suppressed the generation of $CD138^+$ cells in human tonsil B cells activated with CD40L+IL21 (Fig. S8K). These results revealed the antimorphic properties of T95R. To define the difference between $IRF4^{WT}$ and $IRF4^{T95R}$ at the transcriptome level, RNA-seq was performed on Raji cells expressing either version of IRF4. $IRF4^{WT}$ upregulated 132 genes, including key regulators of B cell development such as *PRDM1, XBP1*, *ERN1/IRE1*, *TNFRSF17*, *SLAMF7* and *CD38* (Figs. 5K-5L). $IRF4^{T95R}$ upregulated the expression of 38 genes, 19 of which are shared with $IRF4^{WT}$. $IRF4^{T95R}$ did not hyper-activate the expression of anti-plasma cell transcription factors such as *BCL6*, *BACH2* and *ID3* (Fig. S8L). Consistent with the predominant role of IRF4 as a transcriptional activator in B cells, few genes were downregulated by either variant. When considered in aggregate, this series of experiments demonstrates that $IRF4^{T95R}$ is unable to induce CSR and plasma cell differentiation

due to failure to upregulate the expression of key target genes, including *AICDA*, *PRDM1* and *XBP1*. However, IRF4^{T95R} also demonstrated neomorphic functions by upregulating a unique set of genes.

IRF4^{T95R} has increased affinity for DNA, enhanced nuclear localization and altered DNA-binding specificity

We next defined the intrinsic molecular mechanism of action of the IRF4^{T95R} allele. Due to the additional positive charge introduced by the arginine, we hypothesized that IRF4^{T95R} would have increased affinity for DNA (which is negatively charged) compared to IRF4^{WT}. Moreover, protein affinity for DNA is associated with increased nuclear localization (19). Using both imaging and immunoblotting, we established that the subcellular localization of IRF4^{T95R} was altered, with an increased nuclear:cytoplasmic ratio compared with IRF4^{WT} (Figs. 6A-6B). Next, using surface plasmon resonance to quantify the interaction between the IRF4 and specific DNA sequences, we demonstrated that IRF4^{T95R} has 2-3-fold higher binding affinity (i.e., lower K_D) than IRF4^{WT} for the ISRE, AICE, and EICE motifs (Fig. 6C). This higher binding affinity of IRF4^{T95R} was independently assessed by single-molecule fluorescence microscopy showing that IRF4^{T95R} makes both more and longer (i.e., specific) interactions with DNA than IRF4^{WT} (Fig. 6D). Furthermore, electrophoretic mobility shift assays (EMSA) biochemically confirmed the higher affinity of IRF4^{T95R} for an ISRE DNA probe (Fig. 6E).

Alignment of the nine human IRF protein sequences revealed that the arginine at position 95 of IRF4^{T95R} is analogous to the arginine at position 78 of IRF3 (Fig. S9A). This arginine allows IRF3 to bind to nonconsensus DNA sites (20). Hence, we hypothesized that IRF4^{T95R} would also be able to bind to nonconsensus DNA sites. High-throughput systematic evolution of ligands by

exponential enrichment (HT-SELEX) analysis revealed that, in addition to binding to the consensus IRF GAAA sequence, IRF4^{T95R} had gained the ability to bind to GATA-containing sequences (Fig. 6F). This neomorphic ability of IRF4^{T95R} to bind to GATA sequences was verified by EMSA, which also demonstrated that IRF4^{WT} is unable to bind to this sequence (Fig. S9B). Together, IRF4^{T95R} showed an increased nuclear localization paired with increased DNA-binding affinity, as well as the ability to target additional DNA motifs which were not recognized by IRF4^{WT}, further supporting the hypermorphic and neomorphic functions associated with IRF4^{T95R}.

While IRF4^{T95R} could bind to both GAAA- and GATA-containing ISRE sites, it remained unclear whether it could actually activate target gene transcription. To test this possibility, we performed luciferase-reporter assays to assess the ability of IRF4^{WT} and IRF4^{T95R} to activate transcription from reporter constructs containing a canonical (GAAACCGAAA) or a noncanonical ISRE (GATACCGATA). While IRF4^{WT} activated the canonical ISRE-construct in a dose-dependent fashion, IRF4^{T95R} exhibited hypomorphic activity (Fig. 6G, upper panel). However, confirming IRF4^{T95R} neomorphic activity, IRF4^{T95R} activated the GATA-containing construct in a dose-dependent fashion, while IRF4^{WT} could not activate this construct at all (Fig. 6G, lower panel).

T95R changes both the genome-wide binding landscape of IRF4 and the transcriptome

To evaluate the genome-wide landscape of IRF4-binding in patient cells, we performed IRF4 chromatin immunoprecipitation followed by sequencing (ChIP-seq) of immortalized EBV-B cells from an IRF4^{T95R} patient and an IRF4^{WT} healthy control (Data file S3). Overall, IRF4^{T95R} B cells had more IRF4 ChIP-seq peaks than IRF4^{WT} B cells (Fig. 6H). Emphasizing the neomorphic activity of IRF4^{T95R}, over 35% of the peaks (vs. <10% in IRF4^{WT}) corresponded to "non-ChIPable" IRF4 regions (*i.e.,* they are not reported in the ReMap database (21), which aggregates IRF4 ChIP-

seq data from B cells, T cells, plasmablasts, and various cell lines), and approximately 33% do not overlap any of the >1 million candidate *cis*-regulatory elements from ENCODE (22) (vs. ~7% in $IRF4^{WT}$). Applying a new deep learning tool, ExplaiNN (explainable neural networks) (23), we separately identified motifs *de novo* in four different datasets, including the patient and healthy control ChIP-seq datasets and two custom datasets describing the binding of IRF4 to either AICE or EICE sites in GM12878 cells. Next, we used these motifs to initialize a "surrogate" ExplaiNN model in a process known as transfer learning with which to evaluate their importance towards the $IRF4^{T95R}$-specific, $IRF4^{WT}$-specific or common component of the ChIP-seq data (Fig. S9C, full data in Data file S4). In the patient dataset, ExplaiNN identified a noncanonical, GATA-containing ISRE motif, matching the results of the HT-SELEX analysis, and various noncanonical AICE motifs that were important for $IRF4^{T95R}$-specific peaks but detrimental for $IRF4^{WT}$-specific peaks (Fig. 6H). In contrast, identified EICE motifs in the different datasets were regarded as detrimental for $IRF4^{T95R}$-specific peaks compared to $IRF4^{WT}$-specific and common peaks (Fig. 6H). PU.1 co-factor-independent binding of $IRF4^{T95R}$ was observed at EICE, however increasing amounts of $IRF4^{T95R}$ outcompeted formation of PU.1-$IRF4^{T95R}$ composite complexes at this site (Fig. S9D), providing a possible explanation for the irrelevance of EICE motifs in the patient-derived ChIP-seq peaks. Furthermore, increased co-factor-independent $IRF4^{T95R}$ binding to EICE depended on the presence of a C in position +1 in relation to the GAAA motif (Fig. S9E), suggesting that nucleotides adjacent to the core-binding motif influence $IRF4^{T95R}$ binding capacity. Together, these data confirmed the neomorphic DNA-binding capacity of $IRF4^{T95R}$ to noncanonical ISRE as well as noncanonical AICE sites in primary patient cells.

Next, through matching of RNA-seq and ChIP-seq data on the same cells, we found a clear association between IRF4 binding and gene-expression levels (Fig. 6I). Specifically, genes that

were highly expressed in IRF4^{T95R} patient cells compared to IRF4^{WT} cells had more IRF4^{T95R} ChIP-seq peaks around them (longest transcript ±50 kb) (Fig. 6I; Data file S3). To validate the functional impact of this neomorphic binding activity of IRF4^{T95R}, we focused on CXCL13. CXCL13 is a key chemokine involved in germinal-center activity and lymph node architecture (24) and *CXCL13* was the most differentially expressed gene in patient IRF4^{T95R} B cells when compared to healthy controls (log2 expression fold change = 11.98; Data file S3). The ±50 kb region around the longest *CXCL13* transcript (hg38:chr4:77,511,752-77,611,834) harbored 15 IRF4^{T95R} peaks, of which two overlapped with the only two IRF4^{WT} peaks in the region (Fig. S9F). Two of the 15 IRF4^{T95R} peaks encompassed noncanonical AICE sites (designated as CXCL13 sites A, B and C), none of which contained the canonical GAAA half-ISRE motif. EMSA analyses confirmed strongly increased (CXCL13 site A) or exclusive (CXCL13 site C) neomorph binding of IRF4^{T95R} compared to IRF4^{WT} in the presence of the AP-1 factors JUNB and BATF (Fig. 6J and S9G). However, at CXCL13 site B, which is located within the same peak and only 81 bp away from site A, both IRF4^{T95R} and IRF4^{WT} showed similar binding patterns. To investigate the functional impact of this binding pattern, we analyzed the activity of a luciferase-reporter construct containing the two AICE sites, CXCL13 site A and CXCL13 site B (Fig. 6K). Whereas IRF4^{WT} only slightly increased luciferase activity above the AP-1-induced activity, IRF4^{T95R} induced robust luciferase activity in the presence of AP-1. Finally, to analyze whether this IRF4^{T95R}-mediated transcriptional increase is mirrored by biochemical changes *in vivo*, we measured CXCL13 serum or plasma levels of five IRF4^{T95R} patients and found that they had consistently 3-4.5-fold increased CXCL13 levels compared to healthy controls (Fig. 6L). In combination, these data confirm the functional impact of the neomorphic IRF4^{T95R} binding activity.

Discussion

Foundational to our understanding of genetics are 'Muller's morphs', where Muller suggested that a mutant allele can be classified into one of five types: amorph, hypomorph, hypermorph, neomorph, and antimorph (25). Here we have defined a previously undescribed mechanism of human disease caused by the 'multimorphic' activity of IRF4^{T95R}. The T95R amino acid change results in an unanticipated combination of hypermorphic, hypomorphic, and neomorphic functions in the TF causing the clinical and immunological phenotype we document in the 7 patients with CID. IRF4^{T95R} behaves as a gain-of-function hypermorph by binding to DNA with higher affinity than IRF4^{WT}. Despite this increased affinity for DNA, the transcriptional activity on IRF4 canonical genes is reduced, showcasing a hypomorphic activity of IRF4^{T95R}. Simultaneously, IRF4^{T95R} functions as a neomorph by binding to noncanonical DNA sites to alter the gene expression profile, including the transcription of genes exclusively induced by IRF4^{T95R} but not by IRF4^{WT}. In combination, this multimorphic activity of the IRF4^{T95R} variant leads to a transcriptional space which is detrimental to activation and differentiation of immune cells.

Similar neomorphic functions have been reported in other TFs. In most cases, these neomorphs and the ensuing altered expression of genes predisposes to malignancies (26-28). Furthermore, neomorphic germline mutations in *SP7* cause osteogenesis imperfecta (29), in *EKLF* anemia (30), and in *NOTCH3* arteriopathy (31), to mention a few. However, we are not aware of human diseases caused by the multimorphic combination of hypermorphic, hypomorphic, and neomorphic activity as displayed by IRF4^{T95R}. Overall, our findings highlight a distinct molecular mechanism for human disease and expand the spectrum of diseases originated from mutations in IRF4.

The heterozygous IRF4^{T95R} variant found in multiple unrelated families caused a fully-penetrant, severe very early-onset immunodeficiency characterized by greatly enhanced susceptibility to opportunistic pathogens such as *Pneumocystis jirovecii* and weakly pathogenic mycobacteria. In contrast, IRF4 haploinsufficiency caused by the heterozygous loss-of-function variant IRF4^{R98W} found in a single family was associated with age-dependent, incomplete penetrance of Whipple's disease, an inflammatory intestinal disease caused by *Tropheryma whipplei* (32). Early and severe disease onset was also reported in a single IRF4-deficient patient carrying a homozygous splicing mutation in *IRF4* due to uniparental isodisomy (33). Given the very complex genetics of this isolated case of IRF4-deficiency, the possibility of other contributing factors cannot be excluded. Nevertheless, some immunological features are shared between the heterozygous IRF4^{T95R} patients we describe here, the IRF4-deficient patient (33), and *Irf4* knockout mice (3, 34), including low percentages of memory B lymphocytes in peripheral blood, very low plasma cells in the bone marrow or lymph nodes, agammaglobulinemia, polyclonal T cell repertoire, and low percentages of Th17 and Tfh cells.

Our analysis of heterozygous p.T95R Irf4 knock-in mice recapitulates the human phenotype and further establishes the causal role of the IRF4^{T95R} mutation. The reduced production of antibodies following immunizations and infections, and absence of plasma cells after *in vitro* differentiation points to a failure of IRF4^{T95R} to support terminal B cell differentiation. However, our finding that *Irf4*$^{T95R/+}$ mice have reduced formation of antigen-specific GC B cells, despite an overall increase in the number of GC B cells observed in naïve *Irf4*$^{T95R/+}$ mice, suggests a broader dysregulation of the GC reaction and a very different mechanism for the lack of high-affinity antibodies than a simple loss of IRF4. Together, our data suggest that altered subcellular localization, higher DNA-binding affinity and noncanonical binding of IRF4^{T95R} perturbs the dynamic spatial and temporal

expression of IRF4, resulting in dysregulated GC reactions, impaired immunoglobulin CSR, disturbed T cell differentiation, and decreased plasma cell differentiation.

In conclusion, we describe a human immune disorder caused by the heterozygous germline IRF4^{T95R} mutation. The simultaneous multimorphic combination of dominant loss-, gain- and new-functions for IRF4^{T95R} represents a new mechanism for human disease. We suggest this new human disease be named 'multimorphic IRF4 combined immunodeficiency' (or MICI). Since IRF4 is primarily expressed by cells of the immune system, IRF4^{T95R} manifests disease as a CID. However, we anticipate that variants with multimorphic activity may be more widespread in health and disease, particularly variants in transcription factors that alter the genome-wide binding and transcriptional landscapes.

Materials and Methods

Study design

The objective of this study was to elucidate the pathophysiology of a combined immunodeficiency caused by a recurrent, de novo missense heterozygous *IRF4* variant. We performed extensive phenotyping of the patients' peripheral blood cells by CytoF, scRNA-seq, and conventional flow cytometry to reveal the immune cell abnormalities associated with the disease. Functional defects of B cells, T cells, and myeloid cells were analyzed by inducing their activation and differentiation under in vitro culture conditions. We further generated knock-in mice with the identical T95R substitution and confirmed that the IRF4^{T95R} heterozygous mice recapitulated the immunodeficient phenotypes observed in these patients. We uncovered the altered DNA binding specificity and transcriptional activity of IRF4^{T95R} by multiple approaches, including surface plasmon resonance, single-molecule fluorescence microscopy, EMSA, HT-SELEX, luciferase assay, ChIP-seq, and a deep learning tool (ExplaiNN).

Patients and healthy controls

All study participants and/or their parents/guardians provided written informed consent. Research study protocols were approved by local institutions, specifically: The University of British Columbia Clinical Research Ethics Board (H15-00641, H18-02853), Ethics Committee of the Children's Hospital of Fudan University (registration no. NCT03383380), *Comité de Protection des Personnes Ile de France II,* Paris, France (reference: 2015-01-05; 2015-01-05 MS2), Vanderbilt University Medical Center Human Research Protection Program (IRB#. 200412) and the French Advisory Committee on Data Processing in Medical Research (*Comité Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé*, Paris, France;

reference: 15.297bis), University of Ulm Ethics Board, Germany (application number 156/11), National Institute of Allergy and Infectious Diseases, NIH.

Mice (Shanghai)

Irf4^{T95R} mutant mice and conventional *Irf4*$^{-/-}$ mice were generated by CRISPR/Cas9-mediated genome engineering. For the *Irf4*^{T95R} mutant allele, a gRNA that targets exon 3 of the *Irf4* gene and single-stranded donor oligonucleotides were synthesized at Integrated DNA Technologies. The gRNA, Cas9, and donor oligo were co-injected into C57BL/6 fertilized eggs (Cyagen, Suzhou, China). The sequences of the gRNA and donor oligo are as follows: gRNA, 5'-CGTAATCTTGTCTTCCAAGT-3'; donor oligo, 5'-GGCTTGGGCATTGTTTAAAGGCAAGTTCCGAGAAGGGATCGACAAGCCAGATCCTC CTACGTGGAAGAGAAGATTACGATGTGCTCTGAACAAGAGCAATGACTTTGAGGAA TTGGTCGAGAGGAGCCAG-3'. For conventional *Irf4* knockout, double gRNAs were designed to delete exons 3-6 of the *Irf4* gene. The sequences of gRNAs are as follows: gRNA1, 5'-AGGGACTCACACGGGGTCGG-3'; gRNA2, 5'-TGTCTGCTTCCACGGAGTGT-3'. The gRNAs and Cas9 were co-injected into C57BL/6 fertilized eggs (Cyagen, Suzhou, China).

Mice at 8-12 weeks of age (both males and females as indicated in the figure legend) were used for the functional and biochemical analysis. All mouse strains were bred and maintained in the specific pathogen-free animal facility of Fudan University. All animal studies were conducted in accordance with the Animal Experiment Committee of Fudan University and used the Committee's approved protocols.

Mice (ANU)

Irf4^{T95R} and *Irf4*$^{-/-}$ mice were generated by CRISPR/Cas9-mediated genome engineering in an in-house facility. For generation of the T95R point mutation, gRNA1 gRNA1 5'-CGTAATCTTGTCTTCCAAGT-3' and a single-stranded donor Oligo 5'-TCTGGAACAATCCTGTACACCTTGTATGGGTCAGAGATATCCAGCTGGCTCCTCTCG ACCAATTCCTCAAAGTCATTGCTCTTGTTCAGAGCACATCGTAATCTTCTCTTCCAAG TAGGAGGATCTGGCTTGTCGATCCCTTCTCGGAACTTGCCTTTAAACAATGCCCAAG CCTAAGAACGATGAGA-3' from Integrated DNA Technologies were co-injected into fertilized C57BL6/N zygotes (Cas9 50 ng/µl, gRNA 10 ng/µl and ss oligo 100 ng/µl) in a form of a ribonucleoprotein complex. For generation of *Irf4*$^{-/-}$ mice gRNA1 5'- CGTAATCTTGTCTTCCAAGT-3' and gRNA2 5'-GAACAAGAGCAATGACTTTG-3' were co-injected into fertilized C57BL/6N zygotes (Cas9 50 ng/µl and sgRNA 10 ng/µl). *Irf4*$^{-/-}$ mice contain a 47 bp deletion leading to a frame shift from AA92 of the IRF4 protein and a premature stop codon. Lack of protein expression was confirmed by Western blot.

C57BL/6 mice, *Irf4*$^{T95R/+}$ and *Irf4*$^{+/-}$ were maintained on a C57BL/6N background and bred in pathogen-free conditions at Australia National University (ANU, ACT, Australia) Phenomics Facility. Unless stated otherwise, all mice used were aged between 8-12 weeks. All animal experiments were approved by the ANU Animal Experimentation Ethics Committee on protocols A2014/058, A2017/54 and A2020/45 according to the National Health and Medical Research Council (NHMRC) Australian code of practice.

Cells

HEK293T, Ampho, Raji and Ramos cells were obtained from ATCC. CH12 cells were kindly provided by Prof. Tasuku Honjo (Kyoto University).

Genetic analysis

Genomic DNA was extracted from the whole blood or peripheral blood mononuclear cells (PBMC) and fibroblasts of the patients and their parents. Next generation sequencing of the genomic DNA was performed using an Illumina sequencing platform using either whole exome or a targeted panel approach. Bioinformatics analysis for detection of rare sequence variants were performed as described previously (35, 36).

Mutations in *IRF4* were confirmed by using Sanger sequencing. Primers used for sequencing were listed in Table S13.

Functional predictions of the IRF4 variant

Predictions of the functional relevance of the detected IRF4 variant were performed using the dbNSFPv4 database (https://sites.google.com/site/jpopgen/dbNSFP) (37).

Measurement of immunoglobulin plasma levels

Plasma IgG, IgA and IgM levels were determined by an automated clinical chemistry analyzer (Erba, Model: XL-200) as previously described (36) or by Immunoturbidimetric assays (Roche Diagnostics, Switzerland). Serum IgE was measured by UniCAP (Pharmacia).

Cell isolation and culture

Human $CD4^+$ and $CD8^+$ naive T cells (defined as $CD45RA^+$ $CCR7^+$ or $CD45RA^+$) were isolated (>98% purity) with a FACS AriaII cell sorter (BD Biosciences) or using the MojoSortTM Human CD4 Naïve T Cell Isolation Kit (BioLegend, USA), according to the user manual. Human naïve B cells were isolated from PBMC using Human naïve B cell enrichment kit (Stemcell technology) or MojoSort Human Naïve B Cell Isolation Kit (Biolegend) together with biotinylated anti-human

IgG beads to remove IgG^+ cells or human pan B cell isolation kit (Biolegend) (Memory B cells were removed by a second isolation step using biotinylated anti-human CD27 and F(ab')2 anti-human IgG and IgA antibodies binding to streptavidin-coupled paramagnetic beads and a MoJo magnet as described (38-41)). Mouse BM cells were flushed from femurs and tibias with cold FACS buffer (PBS supplemented with 2% FBS), spleen and Peyer's patch cells were obtained by gently teasing these tissues onto a 70-µm cell strainer. Spleen B cells were purified using a negative sorting kit (BD Biosciences) or a MACS B cell isolation kit (Miltenyi Biotec). Erythrocytes were removed by using ACK buffer.

Purified human naïve B cells, human naïve T cells, human PBMC, mouse purified spleen B cells, and Raji, Ramos, CH12 cells were cultured in RPMI 1640 (GIBCO) containing 5×10^{-5} M 2-mercaptoethanol (Sigma-Aldrich), 100 U/ml penicillin and 100 µg/ml streptomycin (GIBCO), supplemented with 10% heat-inactivated fetal bovine serum (GIBCO). Stimuli were freshly added as described in the figure and figure legends. 293T and Ampho cells were cultured in DMEM medium (GIBCO) supplemented with 10% FBS. All cells were cultured at 37℃ in humidified air and 5% CO_2.

For *in vitro* plasma cells differentiation, enriched human naïve B cells were activated with 2 µg/ml $F(ab')_2$ anti-human plus 0.6 µM CpGODN 2006 and 10 ng/ml IL2 (CMIL2) or 200 ng/ml CD40L and 50 ng/ml IL21. Mouse B cell were activated with 0.1V CD40L plus 20 ng/ml IL4 and 30 ng/ml IL21 or 20 µg/mL LPS.

For the induction of the differentiation into Tfh and Tfr cells, purified naïve human $CD4^+$ T cells were incubated in AIM V Medium (Gibco, USA) supplemented with 5 ng/ml IL12 (Peprotech, Germany), 100 ng/ml Activin A (Peprotech, Germany) and one bead per cell Dynabeads Human T-Activator CD3/CD28 (Gibco, USA) for the specified times (42).

Establishment of EBV-B cell lines

EBV-B immortalization was performed as previously reported (43).

CyTOF

The Maxpar® Direct™ Immune Profiling Assay™ (Fluidigm, #201325) in combination with Maxpar® Direct™ T cell Expansion Panel 2 (Fluidigm, #201406) was used for the high-dimensional immune profiling of PBMC. For each sample, cells were thawed, washed once, and checked for viability (>80%). After a 10 min incubation with FC Blocker (TruStain FCX, Biolegend) in MaxPar staining buffer, cells were directly transferred into the antibody-containing tube. Cells were then incubated for 10 min in 1.6% formaldehyde solution, washed once, transferred into Intercalator ID solution, and incubated overnight at 4°C. Immediately before acquisition, the cells were washed, resuspended in Maxpar Cell Acquisition Solution (1 million cells per mL) and mixed with 10% v/v EQ Beads. An average of 500,000 events were acquired per sample on a Helios mass cytometer. The acquisition data were analyzed with CyTOF software (version 6.7.1014, Fluidigm), enabling Maxpar®Pathsetter™ automated single-cell analysis at the Cytometry Facility in Pitié-Salpetriere Hospital (Paris, France).

For multidimensional analysis, FCS files were pre-gated in FlowJo 10.7.2 (Becton, Dickinson & Company, Ashland, OR) following standard procedures (44). The pre-gated data was then imported into RStudio and analyzed using the package CATALYST (45). Unsupervised clustering and cluster characterization were performed with FlowSOM (46) and marker enrichment modeling (MEM) (47). Additional gating was performed in FlowJo. We used ggplot2 and Prism 9 (GraphPad Software, San Diego, CA) for data representation.

Single-cell RNA-sequencing

To compare gene expression in single B cells, we sorted $CD19^+$ cells from PBMC using MoFlo XDP (Beckman). Either $CD19^+$ B cells or PBMC were processed using the 10x Genomic Chromium Single Cell Platform at a concentration of 700-1,000 cells per microliter, as described in the manufacturer's protocol. Approximately 8,000-10,000 cells were loaded onto 30 library chips following the Single Cell 30 Reagent Kits (v2) User Guide. The cells were then partitioned into GelBeads in Emulsion in the GemCode instrument, where cell lysis and barcoded reverse transcription of RNA occurred, followed by amplification, enzymatic fragmentation, 50 adaptor, and sample index attachment. The libraries were constructed using Chromium Single Cell 30 Reagent Kits (v2): Single Cell 30 Library & Gel Bead Kit v2 (PN-120237), Single Cell 30 Chip Kit v2 (PN-120236), and i7 Multiplex Kit (PN-120262) (10x Genomics). The generated scRNA-seq libraries were sequenced using an Illumina HiSeq 4000 platform as 150-bp paired-end reads, at one full lane per sample.

Shanghai NovelBio Co., Ltd performed scRNA-seq data analysis using the NovelBrain Cloud Analysis Platform. Single-cell RNA-sequencing reads were aligned to the Genome Reference Consortium Human Build 38 (GRCh38), normalized for batch effects, and filtered for cell events using the Cell Ranger software (v3.0.0). Cells expressing fewer than 200 genes and cells with a mitochondrial UMI rate higher than 20% were excluded. Mitochondrial genes were removed from the expression matrix but used for cell expression regression to avoid the cell status effect for clustering analysis and marker analysis of each cluster. The Seurat package (version: 2.3.4) was used for normalization based on the expression matrix according to the UMI counts of each sample and percent of mitochondria rate to obtain the scaled data (Butler et al., 2018). Utilizing the graph-based cluster method, we acquired the unsupervised cell cluster results based on the PCA top 20 principal. To identify differentially expressed genes (DEGs) between the control and disease

groups of the same cell type, the Seurat package FindMarkers function using the Wilcox rank-sum test algorithm under the following criteria: 1) logFC > 0.25; 2) p value < 0.05; 3) min.pct > 0.1.

Immunohistochemistry (IHC) of lymph node tissue

Immunostaining of formalin-fixed and paraffin-embedded tissue sections (2 μm) were carried out by the avidin-biotin-complex method (48). The antibodies used were listed in Table S12.

Mice Immunizations

For humoral immune responses, mice were immunized i.p. with 25 μg of NP-CGG (Biosearch Technologies) in 4.5% alum in 200 μL of PBS. NP-specific IgM and IgG_1 Abs were measured by ELISA as described (49). Antibody responses to chicken γ-globulin (CGG) and formalin fixed *Bordetella pertussis* in mice were determined 14 days post immunization as described previously (50).

Germinal Centre analysis in mice

For induction of GC reactions, each mouse was intravenously injected with 200 μL of the SRBC suspension containing $2x10^8$ SRBCs and spleen were analyzed at day 10. Alternatively, *P. berghei* parasites engineered to express *P. falciparum* CSP in place of the endogenous *P. berghei* CSP molecule (Pb-PfSPZ) (51) were used. Parasites were maintained by serial passage through *Anopheles stephensi* mosquitoes. Mice were immunized IV with $5 x10^4$ irradiated (15kRad) Pb-PfSPZ dissected by hand from the salivary glands of *Anopheles stephensi* mosquitoes as described previously.

For detection of PfCSP specific cells (NANP)9 peptide was sourced from Biomatik (Ontario, Canada) and was biotinylated with the Sulfo-NHS-LC-Biotinylation Kit (ThermoFisher) at a ratio

of 1:1 according to the manufacturer's instructions. Biotinylated Ags were incubated with premium-grade SA-PE and SA-APC (Molecular Probes) at a molar ratio of 4:1, added four times with 15 min incubation at room temperature.

Quantitative PCR

Cells were lysed in RNAiso Plus (TaKaRa) and stored at -80℃. Total RNA was extracted following the manufacturer's instructions. cDNA was synthesized using a 1st Strand cDNA Synthesis Kit (YEASEN). Quantitative PCR was performed using a Light Cycler® 480 Instrument II (Roche) with TB Green premix Ex Taq (TaKaRa). Primers used for Q-PCR were listed in Table S13.

Establishment of IRF4-deficient Ramos cells.

Two guide RNAs were designed to target genomic DNA within the second and third exons of the *IRF4* gene by online software (http://www.crispr-cas.org/), which predicted high-specificity and protospacer adjacent motif target sites in the human exome. Construction of lentiCRISPR vector and collection of lentivirus have been described (52). Ramos cells were cultured in the presence of viral supernatant and selected for the correctly targeted clones as described (52).

ELISPOT assay

This assay was performed as described previously (49). Briefly, multiscreen high-throughput screening plates (Millipore) were coated with 50 μg/mL of goat anti-human Ig (SouthernBiotech) or rabbit anti-mouse Ig (SouthernBiotech). Serially diluted cells were added to individual wells in triplicate and then incubated at 37 °C for 2 hours in a CO_2 incubator. The plates were further incubated with biotin-anti-human IgM or IgG (Southern Biotech), followed by alkaline

phosphatase-conjugated streptavidin (Southern Biotech) or directly with AP-conjugated anti-mouse IgM or IgG_1 (Southern Biotech). Spots were revealed by 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium reagent (MOSS Inc.), and colonies were counted by using an ELISPOT reader (AID).

RNAseq of splenic B cells

RNAseq data from mouse splenic B cells was processed using PiGx-RNA-seq pipeline (53). In short, the data were mapped onto the GRCm38/mm10 version of the mouse transcriptome (downloaded from the ENSEMBL database (54)) using SALMON (55). The quantified data was processed using tximport (56), and the differential expression analysis was done using DESeq2 (57). Genes with less than 5 reads in all biological replicates of one condition were filtered out before the analysis. Two groups of differentially expressed genes were defined - a relaxed set containing genes with an absolute log2 fold change of 0.5, and a stringent set containing games with an absolute log2 fold change of 1. The fold change was deemed significant if the adjusted p-value was less than 0.05 (Benjamini-Hochberg corrected).

Data integration and visualization. Splenic B cell per-sample heatmap was constructed by calculating the pairwise Pearson correlation coefficient between samples. The expression values were normalized using DESeq2. The heatmap was visualized using the ComplexHeatmap package (58). Human and mouse genes were mapped through the orthologous assignment using the ENSEMBL database. Monocyte, B cell, and plasma cell expression profiles were extracted from the ARCHS4 database (59).

RNA-seq of Raji cells

Total RNA was extracted using RNAiso Plus (TaKaRa) from Raji cells transduced with retrovirus expressing WT or mutant IRF4. RNA-seq was performed by BGI (Beijing Genomic Institute, ShenZhen, China) using the BGISEQ-500 platform, pair end 100 bases read lengths. The sequencing data were filtered with SOAPnuke (v1.5.2). The clean reads were mapped to the reference genome using HISAT2 (v2.0.4). After that, Ericscript (v0.5.5) and rMATS (V3.2.5) were used to identify genes and differentially spliced genes (DSGs), respectively. Bowtie2 (v2.2.5) was applied to align the clean reads to the gene set, a database for this organism built by BGI with known and novel coding transcripts included, then the expression level of gene was calculated by RSEM (v1.2.12). Heatmaps were drawn by pheatmap (v1.0.8) according to the gene expression levels in different samples. Essentially, differential expression analysis was performed using the DESeq2 (v1.4.5) under the following criteria: 1) logFC > 1; 2) q value < 0.001; 3) read counts > 10 and FPKM > 0.5.

Immunofluorescence

Transiently transfected HEK293T cells were fixed by incubation with 4% paraformaldehyde for 15 min on coverslips and permeabilized in a 10 min incubation with 1X PBS 5% BSA/0.1% Triton™ X-100. Cells were washed with filtered PBS, blocked by a 40 min incubation with PBS 5% BSA and stained for IRF4 (Cell Signaling, #4948) for 1 h, washed, and incubated with a secondary antibody (anti-rabbit FITC Jackson, #50784,) DAPI (Chemometech, #910-3018) and phalloidin (Cell Signaling, #8940) were added for 40 min. After several washes fluorescence was detected with a confocal Leica SP8 microscope. Images were analyzed with Fiji software.

HT-SELEX

HT-SELEX with the randomized 40 bp-based synthetic ligand was performed as described in (60). To compute 8-mer (i.e., DNA sequences of length 8) enrichments, for each dataset, the number of occurrences of each 8-mer in the first and last SELEX cycles was obtained using Jellyfish (version 2.2.10) (61). Then, the enrichment of each 8-mer was computed as the logarithm to the base 2 resulting from the division between the number of occurrences of that 8-mer in the last and first SELEX cycles. Motifs were obtained using ExplaiNN (23) (see "Deep learning models" section).

ChIP-seq

For ChIP-seq analysis $10x10^6$ Epstein-Barr virus (EBV)-immortalized lymphoblastoid B cells (B-EBV cells) from P3 and a healthy control C1 were fixed with freshly prepared 1% formaldehyde (#F-8775, Sigma) for 15 min, quenched with 0.125 M glycine (Sigma #G-7403) for 5 min and washed twice with 1x PBS containing 0.5% Igepal CA-630 (Sigma #I-8896) and (in the second wash only) 1 mM phenylmethanesulfonyl fluoride. Cell pellets were snap frozen and subjected to ChIP-Seq. Chromatin extraction, immunoprecipitation with an anti-IRF4 antibody (Cell Signaling, #4948), library preparation, next-generation sequencing, and a model-based analysis of the ChIP-Seq data (47) were performed by Active Motif.

Matched ChIP-/RNA-seq analysis

Genes were grouped into 21 different bins, ranging from -10 to +10 according to their log2 expression fold change in patient (i.e., $IRF4^{T95R}$) vs. healthy control cells (i.e., $IRF4^{WT}$; Data file S3). For example, genes whose log2 expression fold change was smaller than -9.5 were assigned to the "-10" bin, between -9.5 and -8.5 to the "-9" bin, between -8.5 and - 7.5 to "-8" bin, etc. For each gene, the number of $IRF4^{T95R}$-specific, $IRF4^{WT}$-specific, and common ChIP-seq peaks were obtained using BEDTools intersect on the genomic window encompassing the longest transcript

of that gene, according to RefSeq definitions (62), ±50 kb. Peak counts were further normalized by applying a normalization factor, which, for each gene, was obtained by dividing the median genomic window size across all genes (i.e., 127,269 bp) by the genomic window size of that gene.

Electrophoretic mobility shift assay (EMSA)

Preparation of whole cell and nuclear extracts as well as western blotting were performed as previously described (63, 64). Custom single-stranded IRDye 700-labelled oligonucleotides were ordered from Integrated DNA Technologies, and double stranded oligonucleotides were generated by annealing in TE buffer (pH 7.9) supplemented with 33.3 mM NaCl and 0.67 mM MgCl2. The annealing conditions were 95°C for 3 min, followed by cooling overnight at room temperature. Supershift assays were performed with 15 µg of nuclear protein lysate incubated on ice for 30 min with either IRF4 (#4964S, Cell Signaling) or IgG (611-145-002, Rockland) or HA-tag (C29F4, Cell Signaling) antibodies, then incubated at room temperature for 20 min with probes shown in Table S13. Protein-oligonucleotide-antibody mixtures were then subjected to electrophoresis in 5% acrylamide/bis-acrylamide 19:1 gels in 1% TBE migration buffer for 60 min at 70V. A LI-COR Odyssey infrared imager (LI-COR Bioscience) was used for imaging.

Luciferase reporter assay

For canonical and noncanonical ISRE reporter assay, HEK293T cells in a 24 well plate were transiently transfected with 500 ng of a 1x*ISRE* reporter plasmid (which contains one *ISRE* sequence) or pGL3 basic backbone plasmid, TK-cypridina vector (an internal control, 50 ng/well) and increasing amounts of pFLAG-CMV-5a vector expressing WT or mutant IRF4 using the Hieff Trans™ Liposomal Transfection Reagent (YEASEN). Cells were harvested 24 h after the transfection and the luciferase activity was measured with the Dual Luciferase Reporter Gene

Assay Kit (Beyotime) according to the manufacturer's protocol. Each transfection was performed in duplicate and reporter activity is expressed as fold induction relative to cells transfected with the empty vector. At least three independent experiments were performed for each reporter plasmid. For the CXCL13 reporter assay, a gBlock DNA fragment was cloned into the pGL3-Basic plasmid (Promega, USA) using the SmaI restriction site. The sequences of the DNA fragment were listed in Table S13. The plasmid was checked by Sanger sequencing using the primer pGL3-Basic/PromF. The HEK293-cell line was cultured as previously described (63). For analysis of luciferase activity, HEK293 cells were transfected by electroporation in OPTI-MEM I using Gene-Pulser II (Bio-Rad) with 960 µF and 0.18 kV with 5 µg of pGL3-based reporter constructs, together with 150 ng pRL-TKLuc as an internal control. Where indicated, cells were additionally transfected with 5 µg pcDNA3-FLAG-JUNB, 5 µg pcDNA-FLAG-BATF, or 40 µg of the respective pHEBO-IRF4 variants. 48 hours after transfection, the ratio of the two luciferases was determined (Dual luciferase kit; Promega).

Statistical analyses

For Figs. 1-6, SciPy (version 1.7.1) (65) was used for mul. Two group data were compared using the Welch's t test (one-tailed), and multiple groups data using the Tukey's honestly significant difference test. For Figs. S1-S9, Graphpad Prism (version 8) was used for statistical analysis. Multiple group data were compared using one-way ANOVA and post-hoc test, and two group data were assessed by two-tailed unpaired Student's t test. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$. P values smaller than 0.05 are considered statistically significant. Error bars show the mean ± standard deviation.

Supplementary Materials

Supplemental Materials and Methods

Clinical summaries of the patients

Fig. S1. Identification of a unique heterozygous mutation in the *IRF4* gene in seven CID patients.

Fig. S2. Overt immunological phenotype, impaired Ig gene CSR and defective PC differentiation in IRF4^{T95R} patients.

Fig. S3. Altered T cell differentiation and function in IRF4^{T95R} patients.

Fig. S4. Heat map of Top 20 differentially expressed (DE) genes in the four CD4 T cell clusters, colored as in Fig. S3A.

Fig. S5. Heat map of Top 20 DE genes in the four CD8 T cell clusters, colored as in Fig. S3B.

Fig. S6. Spontaneous GC B-cell expansion but defective antibody production in *Irf4*$^{T95R/+}$ mice.

Fig. S7. T-cell subset distribution in the spleen of *Irf4*$^{+/+}$, *Irf4*$^{T95R/+}$, *Irf4*$^{+/-}$ and *Irf4*$^{-/-}$ mice.

Fig. S8. IRF4^{T95R} was unable to induce PC differentiation due to the inability to activate the expression of genes involved in the PC differentiation program.

Fig. S9. IRF4^{T95R} showed increased affinity for DNA, altered specificity, and a different IRF4-binding landscape.

Table S1. Clinical features of the patients.

Table S2. Features of the IRF4 T95R mutation.

Table S3. Immunological features of patient 1.

Table S4. Immunological features of patient 2.

Table S5. Immunological features of patient 3.

Table S6. Immunological features of patient 4.

Table S7. Immunological features of patient 5.

Table S8. Immunological features of patient 6.

Table S9. Immunological features of patient 7.

Table S10. Th cell subset distributions in T95R patients and HC.

Table S11. Monocyte subset distributions in T95R patients and HC.

Table S12. List of antibodies and stimuli used in the present study.

Table S13. List of oligonucleotides used in the current study.

Data file S1. Raw data.

Data file S2. Differential gene expression of transduced splenic B cells.

Data file S3. RNA-Seq and CHIP-Seq of EBV B cells.

Data file S4. ExplaiNN results.

References

1. S. G. Tangye, W. Al-Herz, A. Bousfiha, C. Cunningham-Rundles, J. L. Franco, S. M. Holland, C. Klein, T. Morio, E. Oksenhendler, C. Picard, A. Puel, J. Puck, M. R. J. Seppänen, R. Somech, H. C. Su, K. E. Sullivan, T. R. Torgerson, I. Meyts, Human Inborn Errors of Immunity: 2022 Update on the Classification from the International Union of Immunological Societies Expert Committee. *J. Clin. Immunol.*, 1-35 (2022).
2. I. Meyts, A. Bousfiha, C. Duff, S. Singh, Y. L. Lau, A. Condino-Neto, L. Bezrodnik, A. Ali, M. Adeli, J. Drabwell, Primary Immunodeficiencies: A Decade of Progress and a Promising Future. *Front. Immunol.* **11**, 625753 (2020).
3. H. W. Mittrücker, T. Matsuyama, A. Grossman, T. M. Kündig, J. Potter, A. Shahinian, A. Wakeham, B. Patterson, P. S. Ohashi, T. W. Mak, Requirement for the transcription factor LSIRF/IRF4 for mature B and T lymphocyte function. *Science* **275**, 540-543 (1997).
4. T. Tamura, H. Yanai, D. Savitsky, T. Taniguchi, The IRF family transcription factors in immunity and oncogenesis. *Annu. Rev. Immunol.* **26**, 535-584 (2008).
5. A. Agnarelli, T. Chevassut, E. J. Mancini, IRF4 in multiple myeloma-Biology, disease and therapeutic target. *Leuk. Res.* **72**, 52-58 (2018).
6. J. Acquaviva, X. Chen, R. Ren, IRF-4 functions as a tumor suppressor in early B-cell development. *Blood* **112**, 3798-3806 (2008).
7. U. Klein, S. Casola, G. Cattoretti, Q. Shen, M. Lia, T. Mo, T. Ludwig, K. Rajewsky, R. Dalla-Favera, Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nat. Immunol.* **7**, 773-782 (2006).
8. E. Glasmacher, S. Agrawal, A. B. Chang, T. L. Murphy, W. Zeng, B. Vander Lugt, A. A. Khan, M. Ciofani, C. J. Spooner, S. Rutz, J. Hackney, R. Nurieva, C. R. Escalante, W.

Ouyang, D. R. Littman, K. M. Murphy, H. Singh, A genomic regulatory element that directs assembly and function of immune-specific AP-1-IRF complexes. *Science* **338**, 975-980 (2012).

9. S. Sundararaj, S. Seneviratne, S. J. Williams, A. Enders, M. G. Casarotto, Structural determinants of the IRF4/DNA homodimeric complex. *Nucleic Acids Res.* **49**, 2255-2265 (2021).

10. K. Kataoka, Y. Nagata, A. Kitanaka, Y. Shiraishi, T. Shimamura, J. Yasunaga, Y. Totoki, K. Chiba, A. Sato-Otsubo, G. Nagae, R. Ishii, S. Muto, S. Kotani, Y. Watatani, J. Takeda, M. Sanada, H. Tanaka, H. Suzuki, Y. Sato, Y. Shiozawa, T. Yoshizato, K. Yoshida, H. Makishima, M. Iwanaga, G. Ma, K. Nosaka, M. Hishizawa, H. Itonaga, Y. Imaizumi, W. Munakata, H. Ogasawara, T. Sato, K. Sasai, K. Muramoto, M. Penova, T. Kawaguchi, H. Nakamura, N. Hama, K. Shide, Y. Kubuki, T. Hidaka, T. Kameda, T. Nakamaki, K. Ishiyama, S. Miyawaki, S. S. Yoon, K. Tobinai, Y. Miyazaki, A. Takaori-Kondo, F. Matsuda, K. Takeuchi, O. Nureki, H. Aburatani, T. Watanabe, T. Shibata, M. Matsuoka, S. Miyano, K. Shimoda, S. Ogawa, Integrated molecular analysis of adult T cell leukemia/lymphoma. *Nat. Genet.* **47**, 1304-1315 (2015).

11. K. Ochiai, M. Maienschein-Cline, G. Simonetti, J. Chen, R. Rosenthal, R. Brink, A. S. Chong, U. Klein, A. R. Dinner, H. Singh, R. Sciammas, Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. *Immunity* **38**, 918-929 (2013).

12. T. Smith, C. Cunningham-Rundles, Primary B-cell immunodeficiencies. *Hum. Immunol.* **80**, 351-362 (2019).

13. S. N. Willis, K. L. Good-Jacobson, J. Curtis, A. Light, J. Tellier, W. Shi, G. K. Smyth, D. M. Tarlinton, G. T. Belz, L. M. Corcoran, A. Kallies, S. L. Nutt, Transcription factor IRF4

regulates germinal center cell formation through a B cell-intrinsic mechanism. *J. Immunol.* **192**, 3200-3206 (2014).

14. N. Bollig, A. Brüstle, K. Kellner, W. Ackermann, E. Abass, H. Raifer, B. Camara, C. Brendel, G. Giel, E. Bothur, M. Huber, C. Paul, A. Elli, R. A. Kroczek, R. Nurieva, C. Dong, R. Jacob, T. W. Mak, M. Lohoff, Transcription factor IRF4 determines germinal center formation through follicular T-helper cell differentiation. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 8664-8669 (2012).

15. R. Sciammas, A. L. Shaffer, J. H. Schatz, H. Zhao, L. M. Staudt, H. Singh, Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation. *Immunity* **25**, 225-236 (2006).

16. A. Brüstle, S. Heink, M. Huber, C. Rosenplänter, C. Stadelmann, P. Yu, E. Arpaia, T. W. Mak, T. Kamradt, M. Lohoff, The development of inflammatory T(H)-17 cells requires interferon-regulatory factor 4. *Nat. Immunol.* **8**, 958-966 (2007).

17. F. Raczkowski, J. Ritter, K. Heesch, V. Schumacher, A. Guralnik, L. Höcker, H. Raifer, M. Klein, T. Bopp, H. Harb, D. A. Kesper, P. I. Pfefferle, M. Grusdat, P. A. Lang, H. W. Mittrücker, M. Huber, The transcription factor Interferon Regulatory Factor 4 is required for the generation of protective effector CD8+ T cells. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15019-15024 (2013).

18. J. Wu, H. Zhang, X. Shi, X. Xiao, Y. Fan, L. J. Minze, J. Wang, R. M. Ghobrial, J. Xia, R. Sciammas, X. C. Li, W. Chen, Ablation of Transcription Factor IRF4 Promotes Transplant Acceptance by Driving Allogenic CD4(+) T Cell Dysfunction. *Immunity* **47**, 1114-1128.e1116 (2017).

19. M. J. Lyst, R. Ekiert, J. Guy, J. Selfridge, M. V. Koerner, C. Merusi, D. De Sousa, A. Bird, Affinity for DNA Contributes to NLS Independent Nuclear Localization of MeCP2. *Cell Rep.* **24**, 2213-2220 (2018).

20. C. R. Escalante, E. Nistal-Villán, L. Shen, A. García-Sastre, A. K. Aggarwal, Structure of IRF-3 bound to the PRDIII-I regulatory element of the human interferon-beta enhancer. *Mol. Cell* **26**, 703-716 (2007).

21. F. Hammal, P. de Langen, A. Bergon, F. Lopez, B. Ballester, ReMap 2022: a database of Human, Mouse, Drosophila and Arabidopsis regulatory regions from an integrative analysis of DNA-binding sequencing experiments. *Nucleic Acids Res.* **50**, D316-d325 (2022).

22. J. E. Moore, M. J. Purcaro, H. E. Pratt, C. B. Epstein, N. Shoresh, J. Adrian, T. Kawli, C. A. Davis, A. Dobin, R. Kaul, J. Halow, E. L. Van Nostrand, P. Freese, D. U. Gorkin, Y. Shen, Y. He, M. Mackiewicz, F. Pauli-Behn, B. A. Williams, A. Mortazavi, C. A. Keller, X. O. Zhang, S. I. Elhajjajy, J. Huey, D. E. Dickel, V. Snetkova, X. Wei, X. Wang, J. C. Rivera-Mulia, J. Rozowsky, J. Zhang, S. B. Chhetri, J. Zhang, A. Victorsen, K. P. White, A. Visel, G. W. Yeo, C. B. Burge, E. Lécuyer, D. M. Gilbert, J. Dekker, J. Rinn, E. M. Mendenhall, J. R. Ecker, M. Kellis, R. J. Klein, W. S. Noble, A. Kundaje, R. Guigó, P. J. Farnham, J. M. Cherry, R. M. Myers, B. Ren, B. R. Graveley, M. B. Gerstein, L. A. Pennacchio, M. P. Snyder, B. E. Bernstein, B. Wold, R. C. Hardison, T. R. Gingeras, J. A. Stamatoyannopoulos, Z. Weng, Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699-710 (2020).

23. G. Novakovsky, O. Fornes, M. Saraswat, S. Mostafavi, W. W. Wasserman, ExplaiNN: interpretable and transparent neural networks for genomics. 2022.2005.2020.492818 (2022).

24. K. M. Ansel, V. N. Ngo, P. L. Hyman, S. A. Luther, R. Förster, J. D. Sedgwick, J. L. Browning, M. Lipp, J. G. Cyster, A chemokine-driven positive feedback loop organizes lymphoid follicles. *Nature* **406**, 309-314 (2000).

25. H. J. Muller, Further studies on the nature and causes of gene mutations. *Proc. Sixth Int. Congr. Genet.* **1**, 213–255 (1932).

26. A. Arruabarrena-Aristorena, J. L. V. Maag, S. Kittane, Y. Cai, W. R. Karthaus, E. Ladewig, J. Park, S. Kannan, L. Ferrando, E. Cocco, S. Y. Ho, D. S. Tan, M. Sallaku, F. Wu, B. Acevedo, P. Selenica, D. S. Ross, M. Witkin, C. L. Sawyers, J. S. Reis-Filho, C. S. Verma, R. Jauch, R. Koche, J. Baselga, P. Razavi, E. Toska, M. Scaltriti, FOXA1 Mutations Reveal Distinct Chromatin Profiles and Influence Therapeutic Response in Breast Cancer. *Cancer Cell* **38**, 534-550.e539 (2020).

27. Y. K. Xia, Y. R. Zeng, M. L. Zhang, P. Liu, F. Liu, H. Zhang, C. X. He, Y. P. Sun, J. Y. Zhang, C. Zhang, L. Song, C. Ding, Y. J. Tang, Z. Yang, C. Yang, P. Wang, K. L. Guan, Y. Xiong, D. Ye, Tumor-derived neomorphic mutations in ASXL1 impairs the BAP1-ASXL1-FOXK1/K2 transcription network. *Protein Cell* **12**, 557-577 (2021).

28. M. E. Figueroa, O. Abdel-Wahab, C. Lu, P. S. Ward, J. Patel, A. Shih, Y. Li, N. Bhagwat, A. Vasanthakumar, H. F. Fernandez, M. S. Tallman, Z. Sun, K. Wolniak, J. K. Peeters, W. Liu, S. E. Choe, V. R. Fantin, E. Paietta, B. Löwenberg, J. D. Licht, L. A. Godley, R. Delwel, P. J. Valk, C. B. Thompson, R. L. Levine, A. Melnick, Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* **18**, 553-567 (2010).

29. J. C. Lui, A. Raimann, H. Hojo, L. Dong, P. Roschger, B. Kikani, U. Wintergerst, N. Fratzl-Zelman, Y. H. Jee, G. Haeusler, J. Baron, A neomorphic variant in SP7 alters sequence specificity and causes a high-turnover bone disorder. *Nat Commun* **13**, 700 (2022).

30. A. Planutis, L. Xue, C. D. Trainor, M. Dangeti, K. Gillinder, M. Siatecka, D. Nebor, L. L. Peters, A. C. Perkins, J. J. Bieker, Neomorphic effects of the neonatal anemia (Nan-Eklf) mutation contribute to deficits throughout development. *Development* **144**, 430-440 (2017).

31. T. Pippucci, A. Maresca, P. Magini, G. Cenacchi, V. Donadio, F. Palombo, V. Papa, A. Incensi, G. Gasparre, M. L. Valentino, C. Preziuso, A. Pisano, M. Ragno, R. Liguori, C. Giordano, C. Tonon, R. Lodi, A. Parmeggiani, V. Carelli, M. Seri, Homozygous NOTCH3 null mutation and impaired NOTCH3 signaling in recessive early-onset arteriopathy and cavitating leukoencephalopathy. *EMBO Mol. Med.* **7**, 848-858 (2015).

32. A. Guérin, G. Kerner, N. Marr, J. G. Markle, F. Fenollar, N. Wong, S. Boughorbel, D. T. Avery, C. S. Ma, S. Bougarn, M. Bouaziz, V. Béziat, E. Della Mina, C. Oleaga-Quintas, T. Lazarov, L. Worley, T. Nguyen, E. Patin, C. Deswarte, R. Martinez-Barricarte, S. Boucherit, X. Ayral, S. Edouard, S. Boisson-Dupuis, V. Rattina, B. Bigio, G. Vogt, F. Geissmann, L. Quintana-Murci, D. Chaussabel, S. G. Tangye, D. Raoult, L. Abel, J. Bustamante, J. L. Casanova, IRF4 haploinsufficiency in a family with Whipple's disease. *Elife* **7**, (2018).

33. M. Bravo García-Morato, F. J. Aracil Santos, A. C. Briones, A. Blázquez Moreno, Á. Del Pozo Maté, Á. Domínguez-Soto, M. J. Beato Merino, L. Del Pino Molina, J. Torres Canizales, A. V. Marin, E. Vallespín García, M. Feito Rodríguez, D. Plaza López Sabando, A. Jiménez-Reinoso, Y. Mozo Del Castillo, F. J. Sanz Santaeufemia, R. de Lucas-Laguna, P. P. Cárdenas, L. Casamayor Polo, M. Coronel Díaz, M. Valés-Gómez, E. Roldán Santiago, A. Ferreira Cerdán, J. Nevado Blanco, L. Corbí Á, H. T. Reyburn, J. R. Regueiro, E. López-Granados, R. Rodríguez Pena, New human combined immunodeficiency caused

by interferon regulatory factor 4 (IRF4) deficiency inherited by uniparental isodisomy. *J. Allergy Clin. Immunol.* **141**, 1924-1927.e1918 (2018).

34. M. Huber, M. Lohoff, IRF4 at the crossroads of effector T-cell fate decision. *Eur. J. Immunol.* **44**, 1886-1895 (2014).

35. M. A. Field, V. Cho, T. D. Andrews, C. C. Goodnow, Reliably Detecting Clinically Important Variants Requires Both Combined Variant Calls and Optimized Filtering Strategies. *PLoS One* **10**, e0143199 (2015).

36. N. Lai, L. Liu, L. Lin, C. Cui, Y. Wang, Q. Min, E. Xiong, W. Wang, W. Ying, Q. Zhou, J. Hou, J. Sun, J. Y. Wang, X. Wang, Effective and safe treatment of a novel IL2RA deficiency with rapamycin. *J Allergy Clin Immunol Pract* **8**, 1132-1135.e1134 (2020).

37. X. Liu, C. Li, C. Mou, Y. Dong, Y. Tu, dbNSFP v4: a comprehensive database of transcript-specific functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Genome Med.* **12**, 103 (2020).

38. H. Sic, H. Kraus, J. Madl, K. A. Flittner, A. L. von Münchow, K. Pieper, M. Rizzi, A. K. Kienzler, K. Ayata, S. Rauer, B. Kleuser, U. Salzer, M. Burger, K. Zirlik, V. Lougaris, A. Plebani, W. Römer, C. Loeffler, S. Scaramuzza, A. Villa, E. Noguchi, B. Grimbacher, H. Eibel, Sphingosine-1-phosphate receptors control B-cell migration through signaling components associated with primary immunodeficiencies, chronic lymphocytic leukemia, and multiple sclerosis. *J. Allergy Clin. Immunol.* **134**, 420-428 (2014).

39. C. R. Smulski, P. Kury, L. M. Seidel, H. S. Staiger, A. K. Edinger, L. Willen, M. Seidl, H. Hess, U. Salzer, A. G. Rolink, M. Rizzi, P. Schneider, H. Eibel, BAFF- and TACI-Dependent Processing of BAFFR by ADAM Proteases Regulates the Survival of B Cells. *Cell Rep.* **18**, 2189-2202 (2017).

40. E. Sevdali, V. Block, M. Lataretu, H. Li, C. R. Smulski, J. S. Briem, Y. Heitz, B. Fischer, N. J. Ramirez, B. Grimbacher, H. M. Jäck, R. E. Voll, M. Hölzer, P. Schneider, H. Eibel, BAFFR activates PI3K/AKT signaling in human naive but not in switched memory B cells through direct interactions with B cell antigen receptors. *Cell Rep.* **39**, 111019 (2022).

41. C. R. Smulski, L. Zhang, M. Burek, A. Teixidó Rubio, J. S. Briem, M. P. Sica, E. Sevdali, M. Vigolo, L. Willen, P. Odermatt, D. Istanbullu, S. Herr, M. Cavallari, H. Hess, M. Rizzi, H. Eibel, P. Schneider, Ligand-independent oligomerization of TACI is controlled by the transmembrane domain and regulates proliferation of activated B cells. *Cell Rep.* **38**, 110583 (2022).

42. M. Locci, J. E. Wu, F. Arumemi, Z. Mikulski, C. Dahlberg, A. T. Miller, S. Crotty, Activin A programs the differentiation of human TFH cells. *Nat. Immunol.* **17**, 976-984 (2016).

43. G. Tosato, J. I. Cohen, Generation of Epstein-Barr Virus (EBV)-immortalized B cell lines. *Curr. Protoc. Immunol.* **Chapter 7**, Unit 7.22 (2007).

44. C. E. Roe, M. J. Hayes, S. M. Barone, J. M. Irish, Training Novices in Generation and Analysis of High-Dimensional Human Cell Phospho-Flow Cytometry Data. *Curr Protoc Cytom* **93**, e71 (2020).

45. M. Nowicka, C. Krieg, H. L. Crowell, L. M. Weber, F. J. Hartmann, S. Guglietta, B. Becher, M. P. Levesque, M. D. Robinson, CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets. *F1000Res* **6**, 748 (2017).

46. S. Van Gassen, B. Callebaut, M. J. Van Helden, B. N. Lambrecht, P. Demeester, T. Dhaene, Y. Saeys, FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A* **87**, 636-645 (2015).

47. K. E. Diggins, A. R. Greenplate, N. Leelatian, C. E. Wogsland, J. M. Irish, Characterizing cell subsets using marker enrichment modeling. *Nat Methods* **14**, 275-278 (2017).

48. U. Pannicke, M. Hönig, I. Hess, C. Friesen, K. Holzmann, E. M. Rump, T. F. Barth, M. T. Rojewski, A. Schulz, T. Boehm, W. Friedrich, K. Schwarz, Reticular dysgenesis (aleukocytosis) is caused by mutations in the gene encoding mitochondrial adenylate kinase 2. *Nat. Genet.* **41**, 101-105 (2009).

49. E. Xiong, Y. Li, Q. Min, C. Cui, J. Liu, R. Hong, N. Lai, Y. Wang, J. Sun, R. Matsumoto, D. Takahashi, K. Hase, R. Shinkura, T. Tsubata, J. Y. Wang, MZB1 promotes the secretion of J-chain-containing dimeric IgA and is critical for the suppression of gut inflammation. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 13480-13489 (2019).

50. K. L. Randall, T. Lambe, A. L. Johnson, B. Treanor, E. Kucharska, H. Domaschenz, B. Whittle, L. E. Tze, A. Enders, T. L. Crockford, T. Bouriez-Jones, D. Alston, J. G. Cyster, M. J. Lenardo, F. Mackay, E. K. Deenick, S. G. Tangye, T. D. Chan, T. Camidge, R. Brink, C. G. Vinuesa, F. D. Batista, R. J. Cornall, C. C. Goodnow, Dock8 mutations cripple B cell immunological synapses, germinal centers and long-lived antibody production. *Nat. Immunol.* **10**, 1283-1291 (2009).

51. D. A. Espinosa, D. Christensen, C. Muñoz, S. Singh, E. Locke, P. Andersen, F. Zavala, Robust antibody and CD8(+) T-cell responses induced by P. falciparum CSP adsorbed to cationic liposomal adjuvant CAF09 confer sterilizing immunity against experimental rodent malaria infection. *NPJ Vaccines* **2**, (2017).

52. J. Liu, E. Xiong, H. Zhu, H. Mori, S. Yasuda, K. Kinoshita, T. Tsubata, J. Y. Wang, Efficient Induction of Ig Gene Hypermutation in Ex Vivo-Activated Primary B Cells. *J. Immunol.* **199**, 3023-3030 (2017).

53. R. Wurmus, B. Uyar, B. Osberg, V. Franke, A. Gosdschan, K. Wreczycka, J. Ronen, A. Akalin, PiGx: reproducible genomics analysis pipelines with GNU Guix. *Gigascience* **7**, (2018).

54. D. R. Zerbino, P. Achuthan, W. Akanni, M. R. Amode, D. Barrell, J. Bhai, K. Billis, C. Cummins, A. Gall, C. G. Girón, L. Gil, L. Gordon, L. Haggerty, E. Haskell, T. Hourlier, O. G. Izuogu, S. H. Janacek, T. Juettemann, J. K. To, M. R. Laird, I. Lavidas, Z. Liu, J. E. Loveland, T. Maurel, W. McLaren, B. Moore, J. Mudge, D. N. Murphy, V. Newman, M. Nuhn, D. Ogeh, C. K. Ong, A. Parker, M. Patricio, H. S. Riat, H. Schuilenburg, D. Sheppard, H. Sparrow, K. Taylor, A. Thormann, A. Vullo, B. Walts, A. Zadissa, A. Frankish, S. E. Hunt, M. Kostadima, N. Langridge, F. J. Martin, M. Muffato, E. Perry, M. Ruffier, D. M. Staines, S. J. Trevanion, B. L. Aken, F. Cunningham, A. Yates, P. Flicek, Ensembl 2018. *Nucleic Acids Res.* **46**, D754-d761 (2018).

55. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417-419 (2017).

56. C. Soneson, M. I. Love, M. D. Robinson, Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res* **4**, 1521 (2015).

57. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

58. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847-2849 (2016).

59. A. Lachmann, D. Torre, A. B. Keenan, K. M. Jagodnik, H. J. Lee, L. Wang, M. C. Silverstein, A. Ma'ayan, Massive mining of publicly available RNA-seq data from human and mouse. *Nat Commun* **9**, 1366 (2018).

60. L. Zheng, J. Liu, L. Niu, M. Kamran, A. W. H. Yang, A. Jolma, Q. Dai, T. R. Hughes, D. J. Patel, L. Zhang, S. G. Prasanth, Y. Yu, A. Ren, E. C. Lai, Distinct structural bases for sequence-specific DNA binding by mammalian BEN domain proteins. *Genes Dev.* **36**, 225-240 (2022).

61. G. Marçais, C. Kingsford, A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**, 764-770 (2011).

62. N. A. O'Leary, M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B. Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y. Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M. Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K. Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O'Neill, S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz, H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully, A. R. Vatsan, C. Wallin, D. Webb, W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T. D. Murphy, K. D. Pruitt, Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733-745 (2016).

63. S. Kreher, M. A. Bouhlel, P. Cauchy, B. Lamprecht, S. Li, M. Grau, F. Hummel, K. Köchert, I. Anagnostopoulos, K. Jöhrens, M. Hummel, J. Hiscott, S. S. Wenzel, P. Lenz, M. Schneider, R. Küppers, C. Scheidereit, M. Giefing, R. Siebert, K. Rajewsky, G. Lenz, P. N. Cockerill, M. Janz, B. Dörken, C. Bonifer, S. Mathas, Mapping of transcription factor motifs in active chromatin identifies IRF5 as key regulator in classical Hodgkin lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E4513-4522 (2014).

64. S. Mathas, M. Janz, F. Hummel, M. Hummel, B. Wollert-Wulf, S. Lusatis, I. Anagnostopoulos, A. Lietz, M. Sigvardsson, F. Jundt, K. Jöhrens, K. Bommert, H. Stein, B. Dörken, Intrinsic inhibition of transcription factor E2A by HLH proteins ABF-1 and Id2 mediates reprogramming of neoplastic B cells in Hodgkin lymphoma. *Nat. Immunol.* **7**, 207-215 (2006).

65. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* **17**, 261-272 (2020).

66. Q. Min, X. Meng, Q. Zhou, Y. Wang, Y. Li, N. Lai, E. Xiong, W. Wang, S. Yasuda, M. Yu, H. Zhang, J. Sun, X. Wang, J. Y. Wang, RAG1 splicing mutation causes enhanced B cell differentiation and autoantibody production. *JCI Insight* **6**, (2021).

67. S. Morita, T. Kojima, T. Kitamura, Plat-E: an efficient and stable system for transient packaging of retroviruses. *Gene Ther.* **7**, 1063-1066 (2000).

68. L. Hipp, J. Beer, O. Kuchler, M. Reisser, D. Sinske, J. Michaelis, J. C. M. Gebhardt, B. Knöll, Single-molecule imaging of the transcription factor SRF reveals prolonged chromatin-binding kinetics upon cell stimulation. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 880-889 (2019).

69. M. Reisser, A. Palmer, A. P. Popp, C. Jahn, G. Weidinger, J. C. M. Gebhardt, Single-molecule imaging correlates decreasing nuclear volume with increasing TF-chromatin associations during zebrafish development. *Nat Commun* **9**, 5218 (2018).

70. A. P. Popp, J. Hettich, J. C. M. Gebhardt, Altering transcription factor binding reveals comprehensive transcriptional kinetics of a basic gene. *Nucleic Acids Res.* **49**, 6249-6266 (2021).

71. T. Kuhn, J. Hettich, R. Davtyan, J. C. M. Gebhardt, Single molecule tracking and analysis framework including theory-predicted parameter settings. *Sci. Rep.* **11**, 9465 (2021).

72. P. K. Koo, M. Ploenzke, Improving representations of genomic sequence motifs in convolutional networks with exponential activations. *Nat Mach Intell* **3**, 258-266 (2021).

73. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).

74. A. Khan, R. Riudavets Puig, P. Boddie, A. Mathelier, BiasAway: command-line and web server to generate nucleotide composition-matched DNA background sequences. *Bioinformatics* **37**, 1607-1609 (2021).

75. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, G. Louppe, Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, (2012).

76. M. Asif, Y. Orenstein, DeepSELEX: inferring DNA-binding preferences from HT-SELEX data using multi-class CNNs. *Bioinformatics* **36**, i634-i642 (2020).

77. D. P. Kingma, J. J. C. Ba, Adam: A Method for Stochastic Optimization. **abs/1412.6980**, (2015).

IRF4 International Consortium Authors and Contributions

(Contributions described using the CRediT taxonomy)

Co-First Authors (these authors contributed equally; listed alphabetically by last name)

Oriol Fornes[1], Alicia Jia[2], Hye Sun Kuehn[3], Qing Min[4], Ulrich Pannicke[5], Nikolai Schleussner[6,7,8], Romane Thouenon[9], Zhijia Yu[10] (Contributions=Conceptualization, Methodology, Formal Analysis, Investigation, Visualization, Writing-original draft preparation, Writing-reviewing and editing)

Clinician Authors (listed alphabetically by last name)

María de los Angeles Astbury[11], Catherine M. Biggs[2], Miguel Galicchio[12], Jorge Alberto Garcia-Campos[13], Silvina Gismondi[14], Guadalupe Gonzalez Villarreal[15], Kyla J. Hildebrand[2], Manfred Hönig[16], Jia Hou[4], Despina Moshous[17,18], Stefania Pittaluga[19], Xiaowen Qian[20], Jacob Rozmus[2], Ansgar S. Schulz[16], Aidé Tamara Staines-Boone[13], Bijun Sun[4], Jinqiao Sun[4], Schauer Uwe[21], Edna Venegas-Montoya[13], Wenjie Wang[4], Xiaochuan Wang[4], Wenjing Ying[4], Xiaowen Zhai[20], Qinhua Zhou[4] (Contributions= Investigation, Resources, Writing-reviewing and editing)

Scientist Authors (listed alphabetically by last name)

Altuna Akalin[22], Isabelle André[9], Thomas F.E. Barth[23], Bernd Baumann[24], Anne Brüstle[25], Gaetan Burgio[25], Jacinta C. Bustamante[26], Jean-Laurent Casanova[27,28,29,30], Marco G. Casarotto[31], Marina Cavazzana[32], Loïc Chentout[9], Ian A. Cockburn[25], Mariantonia Costanza[6,7,8], Chaoqun Cui[33], Oliver Daumke[34], Kate L. Del Bel[2], Hermann Eibel[35], Xiaoqian Feng[33], Vedran Franke[22], J Christof M. Gebhardt[36], Andrea Götz[5], Stephan Grunwald[34], Bénédicte Hoareau[37], Timothy R. Hughes[38], Eva-Maria Jacobsen[16], Martin Janz[6,7,8], Arttu Jolma[38], Chantal Lagresle-Peyrou[32], Nannan Lai[39], Yaxuan Li[33], Susan Lin[2], Henry Y. Lu[2], Saul O. Lugo-Reyes[40], Xin Meng[33], Peter Möller[23], Nidia Moreno-Corona[9], Julie E. Niemela[3], Gherman Novakovsky[1], Jareb J. Perez-Caraballo[41,42],

Capucine Picard[43,44,45,46,47], Lucie Poggi[9], Maria-Emilia Puig-Lombardi[48], Katrina L. Randall[49,25], Anja Reisser[36], Yohann Schmitt[50], Sandali Seneviratne[25], Mehul Sharma[2], Jennifer Stoddard[3], Srinivasan Sundararaj[31], Harry Sutton[25], Linh Q. Tran[41,42], Ying Wang[4], Wyeth W. Wasserman[1], Zichao Wen[33], Wiebke Winkler[6,7,8], Ermeng Xiong[33], Ally W.H. Yang[38], Meiping Yu[4], Lumin Zhang[39], Hai Zhang[4], Qian Zhao[51], Xin Zhen[41,42] (Contributions= Investigation, Resources, Writing-reviewing and editing)

Co-Senior Co-Corresponding Authors (these authors contributed equally; listed alphabetically by last name)

Anselm Enders[10], Sven Kracker[9], Ruben Martinez-Barricarte[41,42], Stephan Mathas[6,7,8], Sergio D. Rosenzweig[3], Klaus Schwarz[5,52], Stuart E. Turvey[2], Ji-Yang Wang[4,33,53,54,55] (Contributions=Conceptualization, Methodology, Investigation, Validation, Visualization, Resources, Funding acquisition, Project administration, Supervision, Writing-original draft preparation, Writing-reviewing and editing)

Affiliations

[1]Centre for Molecular Medicine and Therapeutics, Department of Medical Genetics, BC Children's Hospital Research Institute, University of British Columbia, Vancouver, BC, Canada.[2]Department of Pediatrics, BC Children's Hospital, University of British Columbia, Vancouver, BC, Canada.[3]Immunology Service, Department of Laboratory Medicine, Clinical Center, National Institutes of Health (NIH), Bethesda, MD, USA.[4]Department of Clinical Immunology, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai, China.[5]Institute for Transfusion Medicine, University of Ulm, Ulm, Germany.[6]Charité - Universitätsmedizin Berlin, Hematology, Oncology and Tumor Immunology, Berlin, Germany.[7]Max-Delbrück-Center for Molecular Medicine, Group Biology of Malignant

Lymphomas, Berlin, Germany.[8]Experimental and Clinical Research Center (ECRC), a joint cooperation between Charité and MDC, Berlin, Germany.[9]Université Paris Cité, Laboratory of Human Lymphohematopoiesis, Imagine Institute, INSERM UMR 1163, Paris, France.[10]Centre for Personalised Immunology and Division of Immunology and Infectious Disease, John Curtin School of Medical Research, The Australian National University, Canberra, ACT, Australia.[11]Infectología pediátrica, Hospital Español de Rosario, Rosario, Argentina.[12]Allergy and Immunology Service, Hospital de Niños VJ Vilela, Rosario, Argentina.[13]Immunology Service, Hospital de Especialidades Unidad Médica de Alta Especialidad (UMAE) 25 del Instituto Mexicano del Seguro Social (IMSS), Monterrey, México.[14]Servicio de Pediatría, Hospital Español de Rosario, Rosario, Argentina.[15]Hematology Service, Hospital de Especialidades Unidad Médica de Alta Especialidad (UMAE) 25 del Instituto Mexicano del Seguro Social (IMSS), Monterrey, México.[16]Department of Pediatrics and Adolescent Medicine, University Medical Center Ulm, Ulm, Germany.[17]Pediatric Immunology, Hematology and Rheumatology Department, Necker-Enfants Malades University Hospital, Assistance Publique Hôpitaux de Paris, Paris, France.[18]Laboratory "Genome Dynamics in the Immune System", INSERM UMR 1163, Imagine Institute, Université Paris Cité, Paris, France.[19]Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD, USA.[20]Department of Hematology/Oncology, Children's Hospital of Fudan University, Shanghai, China.[21]University Children's Hospital, Ruhr University Bochum, Bochum, Germany.[22]Bioinformatics and Omics Data Science Platform, Berlin Institute for Medical Systems Biology, Max Delbrück Center, Berlin, Germany.[23]Department of Pathology, University Medical Center Ulm, Ulm, Germany.[24]Institute of Physiological Chemistry, University of Ulm, Ulm, Germany.[25]Division of Immunology and Infectious Disease, John Curtin School of Medical Research, The Australian National University, Canberra, ACT, Australia.[26]Université de Paris Cité, Necker Hospital, Study

Center for Primary Immunodeficiencies, Imagine Institute, Assistance publique des hôpitaux de Paris (APHP), Paris, France.[27]St. Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, Rockefeller University, New York, NY, USA.[28]Université Paris Cité, Laboratory of Human Genetics of Infectious Diseases, Necker Branch INSERM U1163, Necker Hospital for Sick Children, Imagine Institute, INSERM UMR 1163, Paris, France.[29]Pediatric Immunology-Hematology Unit, Necker Hospital for Sick Children, 75015, Paris, France.[30]Howard Hughes Medical Institute (HHMI), Rockefeller University, New York, NY, USA.[31]Research School of Biology, The Australian National University, Canberra, ACT, Australia.[32]Université de Paris Cité, Necker Hospital, Biotherapy and Clinical Investigation Centre-APHP, Imagine Institute, Paris, France.[33]Department of Immunology, School of Basic Medical Sciences, Fudan University, Shanghai, China.[34]Group Structural Biology, Max-Delbrück-Center for Molecular Medicine, Berlin, Germany.[35]Dept., of Rheumatology and Clinical Immunology, Medical Center and Faculty of Medicine, University of Freiburg and Center for Chronic Immunodeficiency, Medial Center and Faculty of Medicine, University of Freiburg, Freiburg, Germany.[36]Department of Physics, Institute of Biophysics, University of Ulm, Ulm, Germany.[37]Plateforme de Cytométrie CyPS, UMS-37 PASS, Faculté de Médecine de Sorbonne Université, Paris, France.[38]Donnelly Centre, University of Toronto, Toronto, ON, Canada.[39]Key laboratory of whole-period monitoring and precise intervention of digestive cancer, Shanghai Municipal Health Commission (SMHC), Minhang Hospital, Fudan University, Shanghai, China.[40]Immune Deficiencies Laboratory, National Institute of Pediatrics Secretariat of Health, Mexico City, México.[41]Division of Genetic Medicine, Department of Medicine, Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA.[42]Division of Molecular Pathogenesis, Department of Pathology, Microbiology, and Immunology, Vanderbilt Center for Immunobiology, Vanderbilt Institute for Infection, Immunology, and Inflammation, Vanderbilt University Medical Center, Nashville, TN,

USA.[43]Université Paris Cité, Paris, France.[44]Study Center for Primary Immunodeficiencies, Necker-Enfants Malades Hospital, Assistance Publique Hôpitaux de Paris (APHP), Paris, France.[45]Laboratory of Lymphocyte Activation and Susceptibility to EBV infection, Inserm UMR 1163, Institut Imagine, Paris, France. [46]Department of Pediatric Immunology, Hematology and Rheumatology, Necker-Enfants Malades Hospital, APHP, Paris, France.[47]Centre de références des déficits immunitaires Héréditaires (CEREDIH), Necker-Enfants Malades Hospital, APHP, Paris, France.[48]Paris-Descartes Bioinformatics Platform, Imagine Institute, Paris, France.[49]Medical School, The Australian National University, Canberra, ACT, Australia.[50]Genomics Core Facility, Institut Imagine-Structure Fédérative de Recherche Necker, INSERM U1163 et INSERM US24/CNRS UMS3633, Paris Descartes Sorbonne Paris Cite University, Paris, France.[51]Department of Neurology, Zhongshan Hospital, Fudan University, Shanghai, China.[52]Institute for Clinical Transfusion Medicine and Immunogenetics Ulm, German Red Cross Blood Transfusion Service Baden-Württemberg-Hessen and University Hospital Ulm, Ulm, Germany.[53]Department of Microbiology and Immunology, College of Basic Medical Sciences, Zhengzhou University, Zhengzhou, China.[54]Division of Biochemistry, Faculty of Pharmacy and Graduate School of Pharmaceutical Science, Keio University, Tokyo, Japan.[55]Shanghai Huashen Institute of Microbes and Infections, Shanghai, China.

Acknowledgements:

We thank the patients and their families for supporting the study. We acknowledge the extended clinical care team for supporting these patients, including the Rare Disease Discovery Hub at BC Children's Hospital and for their support of this patient-centered research project. We thank the Bioinformatics Team of Fudan University Children's Hospital, Jing Qian and Yiwei Zhong at

Shanghai Medical College Fudan University for their excellent technical support, and the Animal Facility of Fudan University for maintaining the mice. We thank Marita Führer, Christina Kellerer, Juliane Nell, Evi Rump, Sarah Wölfle, Benjamin Winter, and Dr Sofia Omari for excellent technical assistance. We thank the Necker Imagine Centre de Resources Biologiques for generating EBV-transformed lymphoblastoid cell lines; the clinical research team at the Imagine Institute for their support; the Etablissement Français du Sang for blood supply from healthy donors. We thank Nay-Chi Khin, Nikki Ross and Dr Jenna Lowe for assistance in generating the *Irf4*^{T95R} and *Irf4* KO mouse strains at ANU. SK is a Centre National de la Recherche Scientifique staff researcher. SET holds a Tier 1 Canada Research Chair in Pediatric Precision Health and the Aubrey J. Tingle Professor of Pediatric Immunology. HSK and SDR state that the content of this article does not necessarily reflect the views or policies of the US Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government.

Funding:

This work was supported by the following grants: National Natural Science Foundation of China (grant #91942302, #31870898 and #82011540008, to JYW), Ministry of Science and Technology of China (grant #2019YFE0100600 to JYW), Canadian Institutes for Health Research (grant #PJT-178054 to SET and grant # FDN-148403 to TRH), BC Children's Hospital Foundation (SET), Canada Research Chairs Program (STE), National Institute Of Allergy And Infectious Diseases of the National Institutes of Health (grant #R21AI171466 to RMB), National Institutes of Health, intramural research program, NIH Clinical Center and NIAID to SR, Agence Nationale de la Recherche (grant #ANR-19-CE17-0012-01 to SK and ANR-19-CE17-0012-02; ANR-19-CE17-

0012-04), the French state (via the Agence Nationale de la Recherche's "Investissments d'avenir" program (ANR-10-IAHU-01 to Institute Imagine), the Ligue Contre le Cancer–Comité de Paris, and INSERM to SK, the National Health and Medical Research Council of Australia (GNT2012498) to AE, and by the Phenomics Translation Initiative, an Medical Research Future Funds funded program (#EPCD000035) and the National Collaborative Research Infrastructure, Strategy (NCRIS) via Phenomics Australia. CG was supported by the German Research Foundation (grant no. 316249678–SFB 1279, subproject B05). OF, GN, and WWW were supported by grants from the Canadian Institutes of Health Research (PJT-162120), Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN-2017-06824), and the BC Children's Hospital Foundation and Research Institute.

Competing interests:

Authors declare that they have no competing interests.

Data and materials availability:

The ChIP-Seq and RNA-seq data of EBV immortalized B cells can be accessed in the Gene Expression Omnibus (GEO) repository, under the accession numbers GSE199684 and GSE199685 (GSE199686 SuperSeries). The scRNA-seq data of human B cells and PBMC can be accessed in the GEO repository, under the accession numbers GSE215936 and GSE215938. The RNA-Seq data of Raji cells can be accessed by DOI: 10.5281/zenodo.7198935. The RNA-seq data of mouse splenic B cells can be accessed in the ArrayExpress under the ID: E-MTAB-12352. All other data needed to support the conclusions of the paper are present in the paper or the

Supplementary Materials. The IRF4^{T95R} mice are available from Ji-Yang Wang or Anselm Enders upon request.

Figure 1

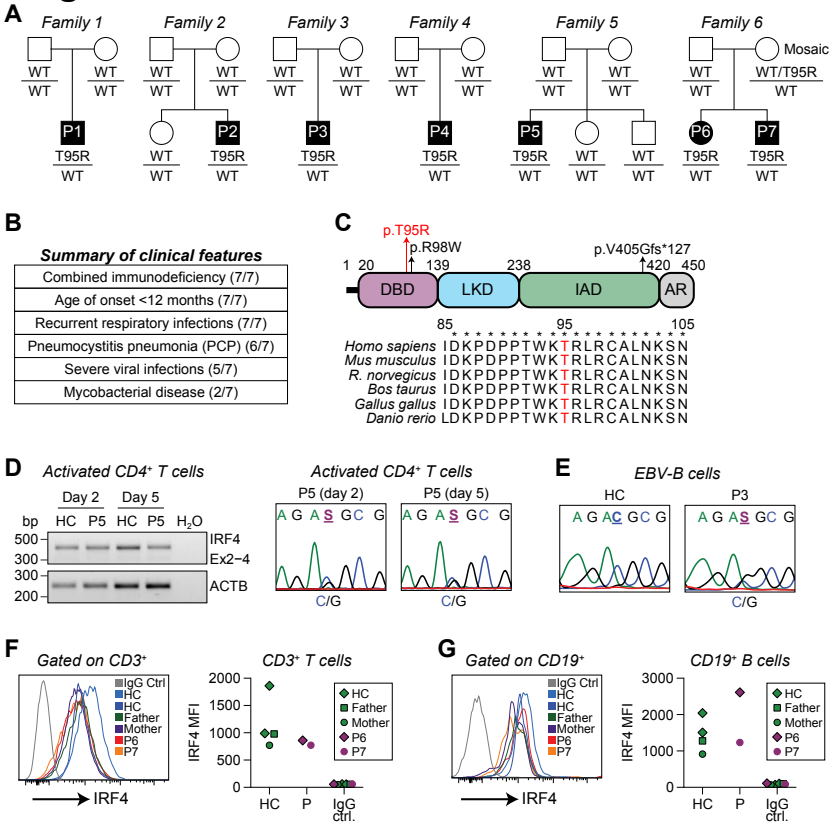

A
Family 1
Family 2
Family 3
Family 4
Family 5
Family 6
Mosaic
WT/T95R
P1
P2
P3
P4
P5
P6
P7
T95R
WT
B
Summary of clinical features
Combined immunodeficiency (7/7)
Age of onset <12 months (7/7)
Recurrent respiratory infections (7/7)
Pneumocystitis pneumonia (PCP) (6/7)
Severe viral infections (5/7)
Mycobacterial disease (2/7)
C
p.T95R
p.R98W
p.V405Gfs*127
DBD
LKD
IAD
AR
Homo sapiens IDKPDPPTWKTRLRCALNKSN
Mus musculus IDKPDPPTWKTRLRCALNKSN
R. norvegicus IDKPDPPTWKTRLRCALNKSN
Bos taurus IDKPDPPTWKTRLRCALNKSN
Gallus gallus IDKPDPPTWKTRLRCALNKSN
Danio rerio LDKPDPPTWKTRLRCALNKSN
D
Activated CD4+ T cells
Day 2
Day 5
HC
P5
H2O
IRF4 Ex2−4
ACTB
P5 (day 2)
P5 (day 5)
C/G
E
EBV-B cells
P3
F
Gated on CD3+
IgG Ctrl
Father
Mother
IRF4
CD3+ T cells
IRF4 MFI
IgG ctrl.
G
Gated on CD19+
CD19+ B cells

Fig. 1. Identification of a unique heterozygous mutation in the *IRF4* gene in seven combined immunodeficiency patients. (**A**) Pedigrees of seven patients from six unrelated families with an identical *IRF4* mutation (c.284C>G, p.T95R). Affected individuals are indicated by the filled symbols. (**B**) Summary of the major clinical features in these patients. (**C**) Schematic representation of the IRF4 protein (Isoform 2; NCBI accession#: NP_001182215.1). DBD, DNA binding domain; LKD, linker domain; IAD, IRF-association domain; AR, auto-regulatory region. The T95R substitution is indicated by a red arrow. A recently identified heterozygous LOF mutation (c.292C>T, R98W) associated with Whipple's disease and a homozygous splicing mutation (c.1213-2A>G, p.V405Gfs*Ter127) causing human CID are also indicated. Below, alignment of the amino acids in the DBD of IRF4 from different species. T95 is shown in red. (**D**) Left, $CD4^+$ naive T cells from P5 were stimulated with anti-CD3/CD28 beads, IL12 and Activin A for 2 and 5 days. Total RNA was isolated from 500 cells and analyzed by RT-PCR. Right, Sanger sequencing of RT-PCR products. HC, healthy controls. (**E**) Sanger sequencing of RT-PCR products of EBV-immortalized B cells from P3. (**F** and **G**) Flow cytometric analysis (intracellular staining) for IRF4 protein expression in gated $CD3^+$ T cells (**F**) or $CD19^+$ B cells (**G**). Left, representative profiles; right, mean fluorescence intensity (MFI) of IRF4 protein from P6 and P7, their parents and two HC.

Figure 2

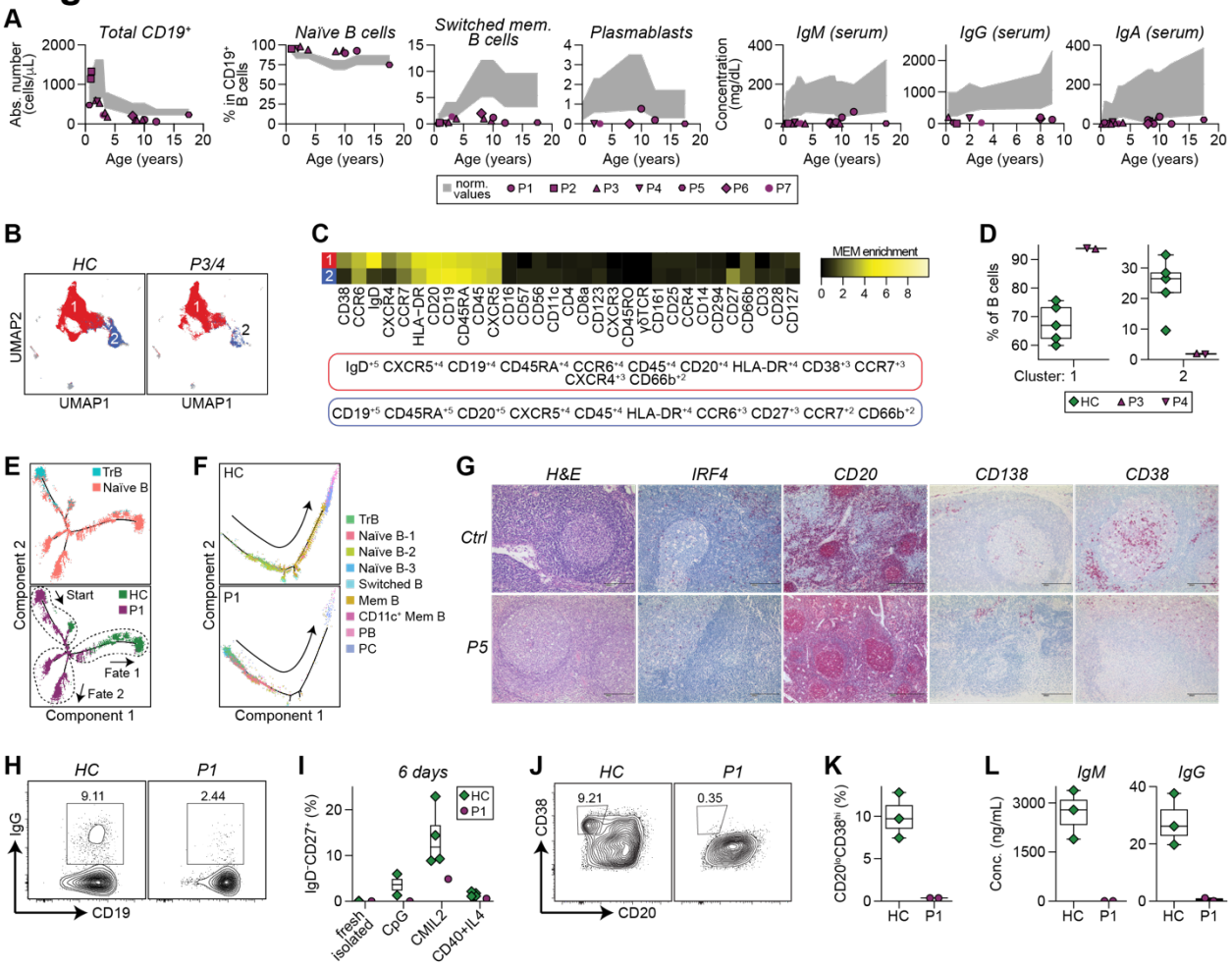


Fig. 2. Impaired Ig gene class switch recombination (CSR) and defective memory B and plasma cell differentiation in IRF4^{T95R} patients. (**A**) Total number of B cells, proportions of naive B, switched memory B and plasmablasts among CD19^{+} B cells, and serum Ig levels in the seven patients. Age-matched reference ranges are shown in grey. Detailed data are shown in Supplemental tables 3-9. (**B-D**) CyTOF analysis of the B cell cluster from Fig. S2A-C performed in P3, P4 and 5 HC. (**B**) Dimensional reduction by UMAP of the 2 major populations obtained by unsupervised clustering using FlowSOM. **(C)** Marker Enrichment Modeling (MEM) heatmap and tags showing the markers that characterize each population. (**D**) Proportions of clusters 1 and 2 among total B cells from Fig. S2A-C. **(E** and **F)** Pseudotime analysis based on Single-cell RNA-seq of purified peripheral blood B cells performed in a HC and P1. (**E**) Pseudotime analysis of transitional and naïve B cell subpopulations. (**F**) Pseudotime analysis of B cell differentiation into memory B and plasma cells. (**G**) Immunohistochemical staining of lymph nodes of P5 and an unaffected control (Ctrl). (**H**) Purified naïve B cells of an age-matched HC and P1 were cultured with CD40L plus IL4 for 6 days and analyzed for the percentages of IgG^{+} cells by flow cytometry. (**I**) Purified naïve B cells of 4 HC and P1 were stimulated with CpG ODN 2006 alone, CpG + F(ab')$_2$ α-IgM and IL2 (CMIL2) or CD40L + IL4 for 6 days and analyzed for the generation of CD19^{+}IgD^{-}CD27^{+} memory B cells. (**J-L**) Purified naïve B cells of HC and P1 were stimulated with CMIL2 for 6 days and analyzed for the induction of CD20^{-}CD38^{+} plasmablasts and Ig

secretion. (**J**) Representative FACS profiles. (**K**) Results of three HC and two experiments of P1. (**L**) IgM and IgG levels in the culture supernatants.

Figure 3

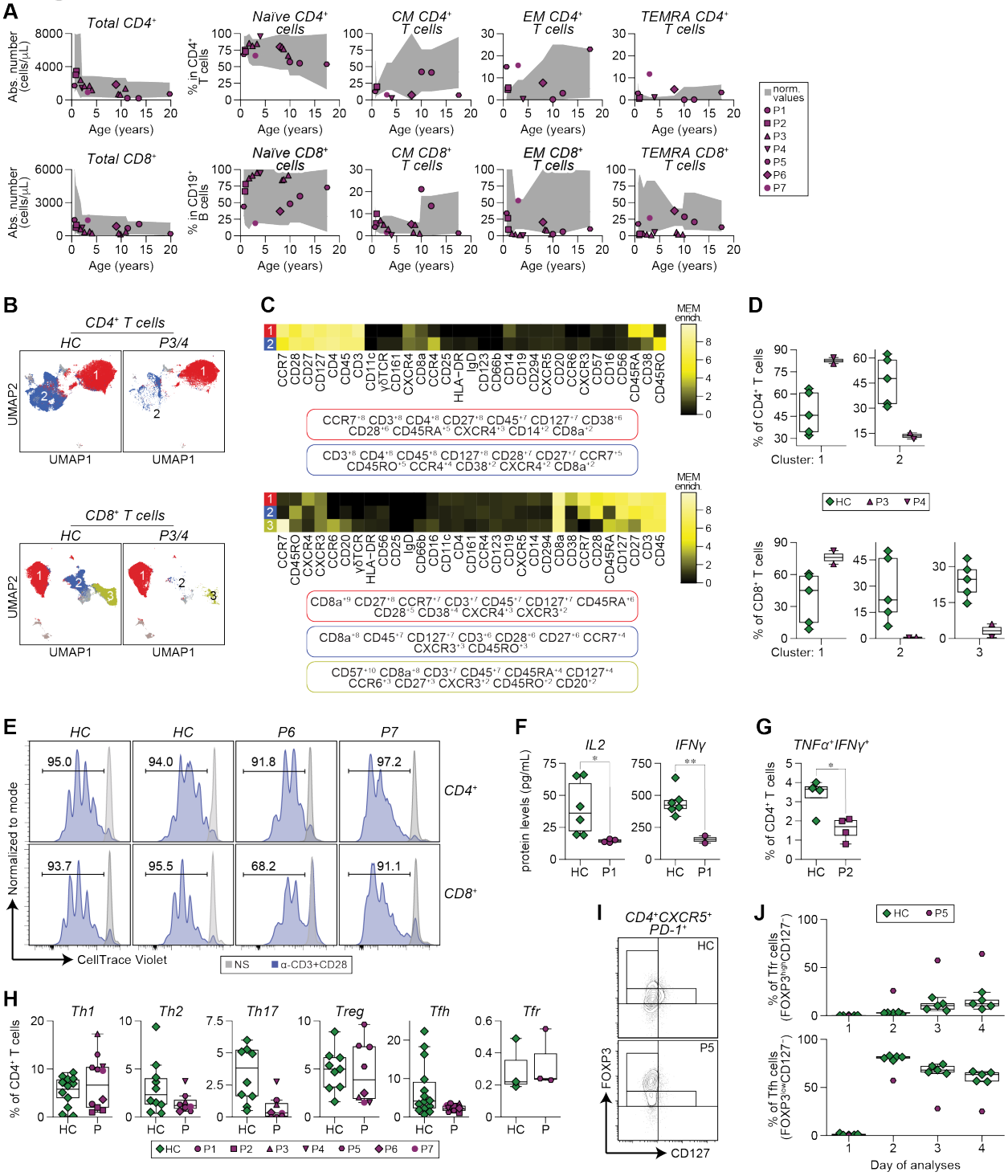


Fig. 3. Altered T cell differentiation and reduced T cell activation and cytokine production in IRF4^{T95R} patients. (**A**) Total number of CD4$^+$ (upper left) and CD8$^+$ (lower left) T cells, and percentages of naïve, central memory (CM), effector memory (EM) and EM re-expressing CD45RA (TEMRA) cells among CD4$^+$ (upper panels) and CD8$^+$ (lower panels) T cells. (**B-D**)

CyTOF analysis of the $CD4^+$ and $CD8^+$ cell clusters in P3, P4 and 5 HC from Fig. S2A-C. (**B**) Dimensional reduction by UMAP showing the 2 major clusters in the $CD4^+$ population and the 3 major clusters of the $CD8^+$ population obtained by unsupervised clustering using FlowSOM. **(C)** Marker Enrichment Modeling (MEM) heatmap and tags showing the markers that characterize each population. (**D**) Proportions of the different clusters among $CD4^+$ (Top) and $CD8^+$ (Below) cells from Fig. S2A-C. (**E**) PBMC from 2 HC and P6 and P7 were labeled with celltrace violet and stimulated with anti-CD3/anti-CD28 dynabeads for 4 days. Upper panels, gated $CD4^+$ T cells; Lower panels, gated $CD8^+$ T cells. (**F**) PBMC from HC and P1 were stimulated with anti-CD3 and anti-CD28 for 24 h and analyzed for the amount of IL2 and IFNγ in the culture supernatants. Results of six HC and 2-4 independent experiments of P1 are shown. (**G**) Intracellular staining of TNFα and IFNγ in $CD4^+$ T cells from a HC and P2 before and after PMA + ionomycin stimulation. Mean ± S.D. of four experiments was shown. (**H**) Th-cell subset distributions in T95R patients and HC. The proportions of Th1, Th2, Th17, Treg, Tfh and Tfr among peripheral $CD4^+$ T cells were determined by flow cytometric analysis for P1, P2, and P5-P7 and by CyTOF for P3 and P4 (supplemental Table 10). (**I**) Purified naive T cells ($CD3^+CD4^+CD45RA^+$) of P5 and 5 HC were subjected to a Tfh/Tfr differentiation assay and were analyzed by flow cytometry on days 0, 4, 5 and 6. Left, representative FACS profiles on day 6; Right, Proportions of Tfr-like ($FOXP3^{high}CD127^-$) and Tfh-like ($FOXP3^{low}CD127^{+/-}$) cells among the $CD4^+PD\text{-}1^+CXCR5^+$ T cells were quantified. Red dots show means of 3 biological independent P5 replicates. Bars represent mean percentages of HC. *$p < 0.05$; **$p < 0.01$.

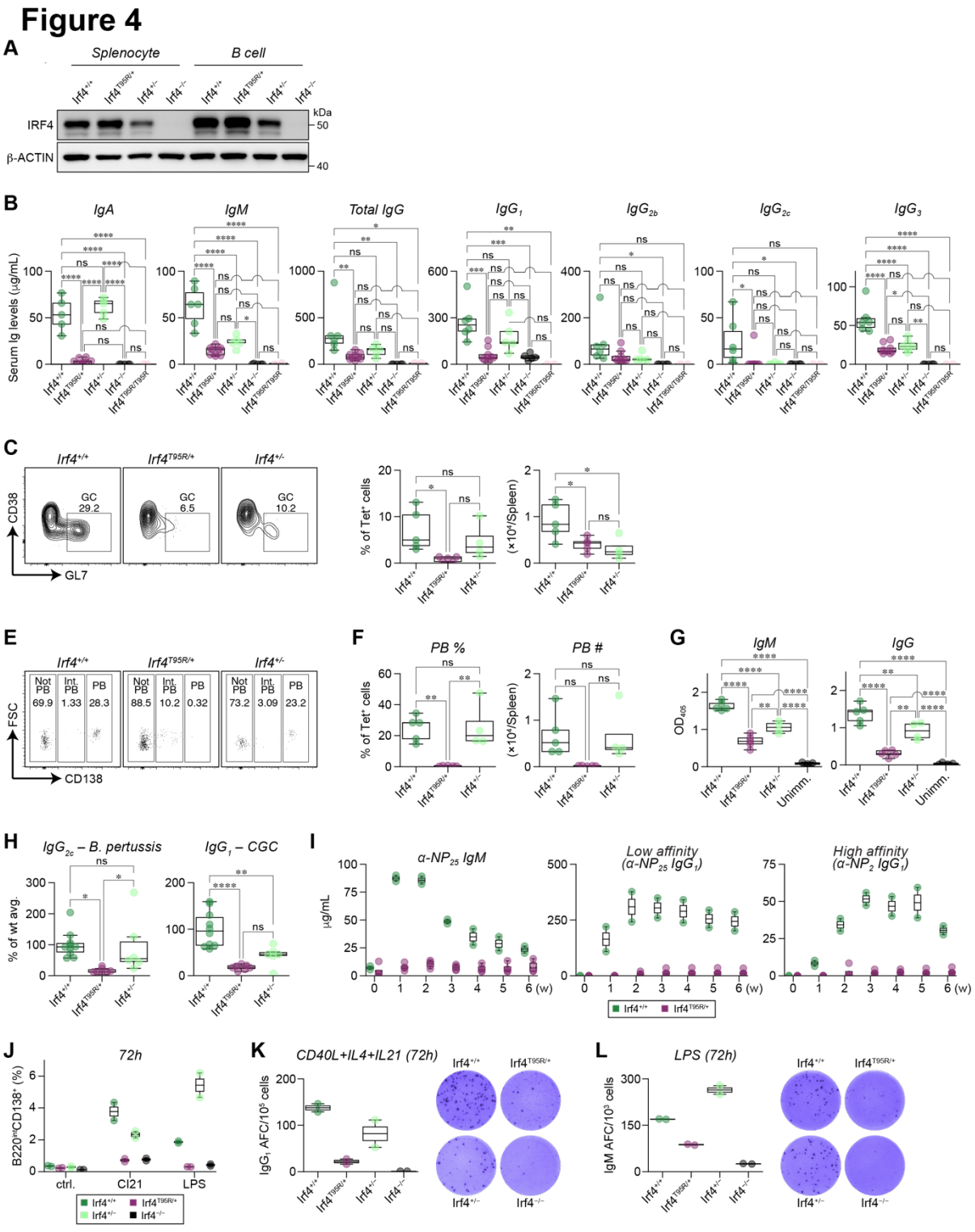


Fig. 4. Heterozygous p.T95R IRF4 knock-in mice recapitulate the immunodeficiency phenotypes observed in IRF4^{T95R} patients. (**A**) WT and mutant IRF4 protein expression in total splenocytes and purified B cells from WT (*Irf4*$^{+/+}$), *Irf4*$^{T95R/+}$, *Irf4*$^{+/-}$ and *Irf4*$^{-/-}$ male mice. (**B**) Serum Ig levels in male and female mice were measured by ELISA. (**C**) *Irf4*$^{+/+}$, *Irf4*$^{T95R/+}$ and

Irf4$^{+/-}$ mice were immunized with *Plasmodium* sporozoites and analyzed for the proportions of GC B cells within the Tet^{+} circumsporozoite protein (CSP)-specific B cells. (**D**) Percentages of Tet^{+} (Ag-specific) cells within all GC B cells (left) and number (right) of Tet^{+} (Ag-specific) GC B cells (left). (**E**) *Irf4*$^{T95R/+}$ mice were unable to generate CD138^{high} PB. (**F**) Percentages (left) and number (right) of Tet^{+} PB in the spleen. (**G**) The production of CSP-specific IgM and IgG antibodies. Absorbance at 405 nm was measured and the area under the curve was calculated in Prism 8 from the log (dilution) on the x-axis and the A405 on the y axis, fitting a sigmoidal curve. (**H**) Left, IgG_{2c} production following immunization with formalin-fixed *B. pertussis*; Right, IgG_1 production in response to CGG immunization. (**I**) WT and *Irf4*$^{T95R/+}$ female mice were immunized *i.p.* with 25 µg of NP-CGG in alum. Serum levels of NP-specific IgM and low and high-affinity IgG_1 were determined each week by ELISA. (**J-L**) Naïve B cells purified from WT and mutant male mice were cultured for 72 h in the presence of CD40L + IL4 + IL21 (CI21) or LPS. The cells were then analyzed for the generation of CD138^{+} plasma cell by flow cytometry (**J**) and IgG_1 and IgM secretion by ELISPOT (**K** and **L**). Each dot represents data from an individual mouse. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

Figure 5

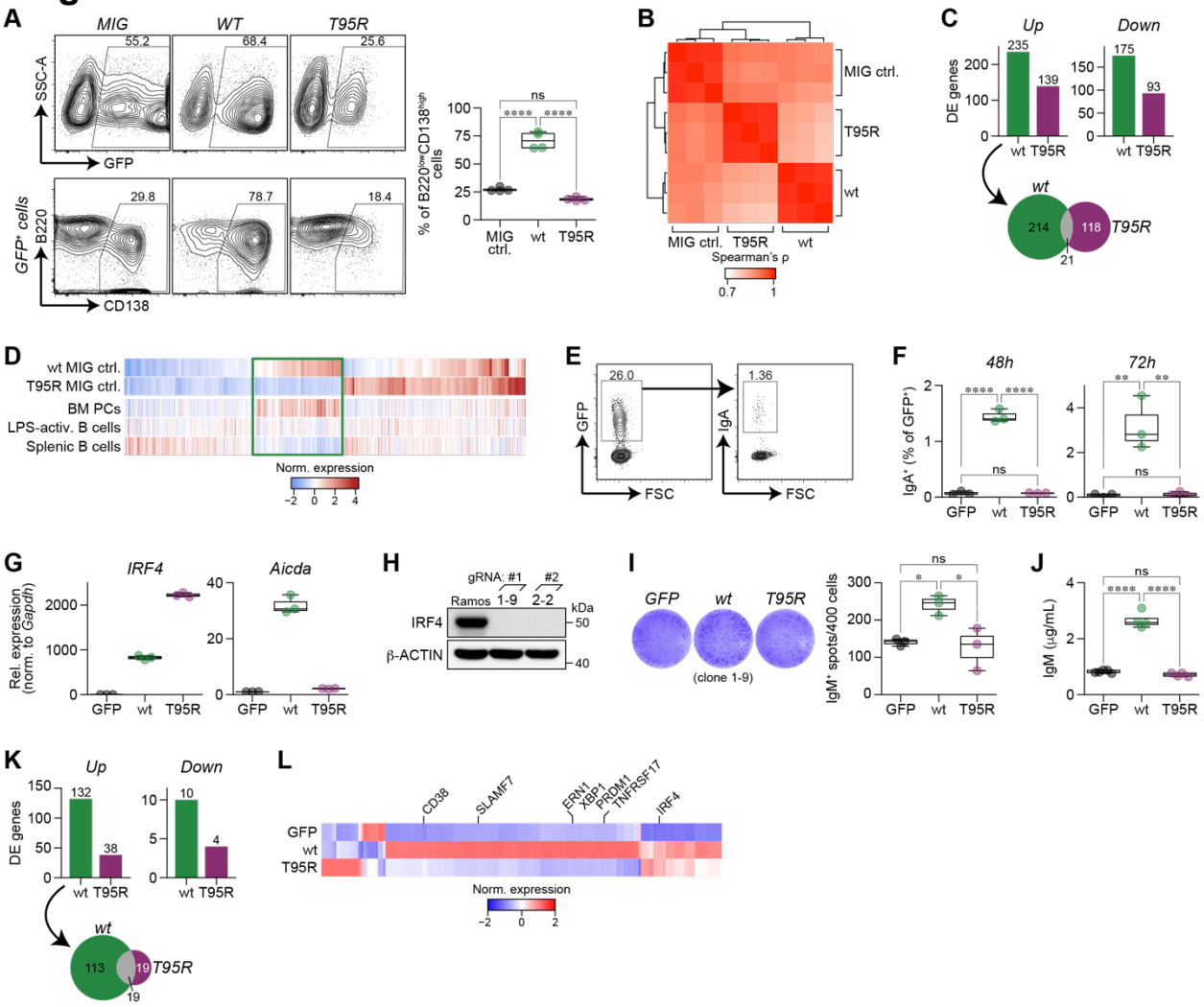


Fig. 5. IRF4^{T95R} failed to induce plasma cell **differentiation due to the inability to activate IRF4 target genes involved in** plasma cell **differentiation. (A)** Mouse C75BL/6 splenic B cells were cultured with LPS + IL4, transduced with control retrovirus (MIG-ctrl), IRF4^{WT} or IRF4^{T95R}, and analyzed for CD138 and B220 expression in transduced GFP^{+} cells. Left upper panels, representative FACS profiles. The percentage of GFP^{+} cells is indicated. Left lower panels, CD138 and B220 expression in gated GFP^{+} cells; the percentage of CD138^{+}B220^{low} cells is indicated. Right panel, summary of 4 independent experiments. Mean ± S.D. is shown. (**B**) Spearman correlation coefficient between RNA-seq derived expression values of isolated mouse splenic B cells transduced with IRF4^{WT}, IRF4^{T95R}, or MIG-control (MIG-ctrl). (**C**) Upper panels, number of differentially expressed genes in IRF4^{WT}, and IRF4^{T95R} when compared to MIG-ctrl samples. Lower panel, genes differentially upregulated by IRF4^{WT}, and IRF4^{T95R} show limited overlap. (**D**) Comparison of genes differentially regulated by IRF4^{T95R} with gene expression of selected lymphoid cell types. Genes associated with plasma cell differentiation are marked by a rectangle. (**E-G**) CH12 B cells were transduced with retrovirus expressing GFP alone, IRF4^{WT} or IRF4^{T95R} and analyzed for the proportion of IgA^{+} cells in gated GFP^{+} cells 48 h and 72 h later. (**E**)

Representative FACS profiles showing the virus-transduced GFP^+ fraction (left) and IgA expression among the GFP^+ cells (right). (**F**) Class switch to IgA at 48 h and 72 h after transduction of retrovirus expressing GFP alone, $IRF4^{WT}$ or $IRF4^{T95R}$. Mean ± S.D. of 3 independent experiments is shown. (**G**) Real-time PCR analysis of *IRF4* and *Aicda* expression in sorted GFP^+ CH12 cells after retrovirus transduction. The level of *IRF4* and *Aicda* in CH12 cells expressing GFP alone was set to 1. (**H**) Generation of IRF4-deficient Ramos cells. Immunoblot for IRF4 protein expression in WT and IRF4-deficient Ramos clones 1-9 and 2-2 derived from $1^{\#}$ and $2^{\#}$ gRNA. (**I** and **J**) IRF4-deficient Ramos cells (clone 1-9) were transduced with retrovirus expressing GPF alone, $IRF4^{WT}$ or $IRF4^{T95R}$. The GFP^+ cells were sorted 3 days later and analyzed for IgM-secreting cells by ELISPOT. (**I**) Left, representative images of ELISPOT; Right, the number of IgM-secreting spots. Mean ± S.D. of triplicate wells is shown. (**J**) The sorted cells were further cultured for 5 days and analyzed by ELISA for IgM secreted into the culture supernatant. (**K** and **L**) Raji cells were transduced with retrovirus expressing GPF alone, $IRF4^{WT}$ or $IRF4^{T95R}$ and the GFP^+ cells were sorted for RNA sequencing. (**K**) The number of differentially expressed genes in Raji cells expressing $IRF4^{WT}$ or $IRF4^{T95R}$, as compared with Raji cells expressing GFP alone. (**L**) Expression heat map depicting the differentially expressed genes shown in K. Average TPM values of 3 independent samples standardized by z-score are shown. $*p < 0.05$; $**p < 0.01$; $****p < 0.0001$.

Figure 6

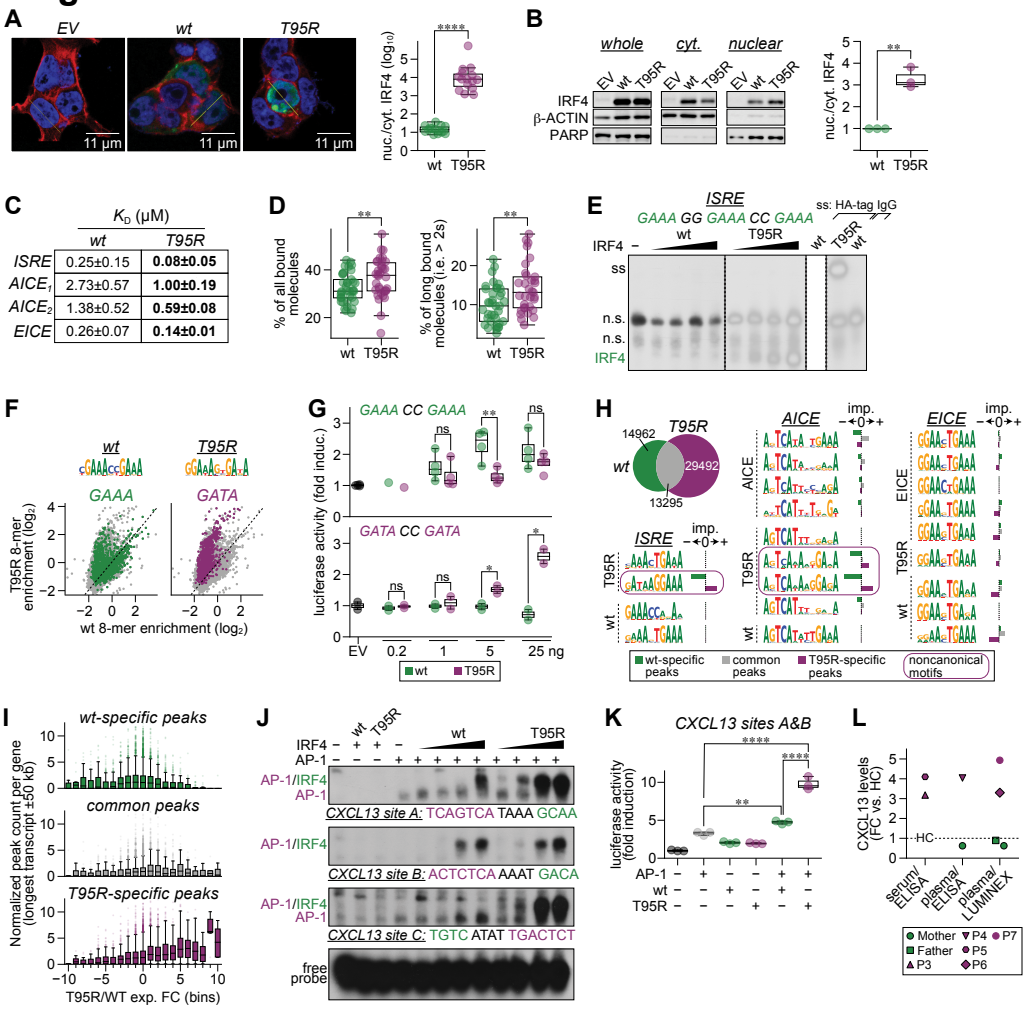

A
EV
wt
T95R
11 µm
nuc./cyt. IRF4 (log10)
wt T95R
B
whole
cyt.
nuclear
IRF4
β-ACTIN
PARP
nuc./cyt. IRF4
C
KD (µM)
wt
T95R
ISRE 0.25±0.15 0.08±0.05
AICE1 2.73±0.57 1.00±0.19
AICE2 1.38±0.52 0.59±0.08
EICE 0.26±0.07 0.14±0.01
D
% of all bound molecules
% of long bound molecules (i.e. > 2s)
E
ISRE
GAAA GG GAAA CC GAAA
ss: HA-tag IgG
IRF4
ss
n.s.
n.s.
IRF4
F
wt
T95R
GAAA
GATA
T95R 8-mer enrichment (log2)
wt 8-mer enrichment (log2)
G
GAAA CC GAAA
GATA CC GATA
luciferase activity (fold induc.)
EV 0.2 1 5 25 ng
wt
T95R
H
14962
29492
13295
T95R
wt
AICE
EICE
ISRE
imp.
wt-specific peaks
common peaks
T95R-specific peaks
noncanonical motifs
I
wt-specific peaks
common peaks
T95R-specific peaks
Normalized peak count per gene (longest transcript ±50 kb)
T95R/WT exp. FC (bins)
J
IRF4
AP-1
wt
T95R
AP-1/IRF4
AP-1
CXCL13 site A: TCAGTCA TAAA GCAA
CXCL13 site B: ACTCTCA AAAT GACA
CXCL13 site C: TGTC ATAT TGACTCT
free probe
K
CXCL13 sites A&B
luciferase activity (fold induction)
AP-1
wt
T95R
L
CXCL13 levels (FC vs. HC)
HC
serum/ ELISA
plasma/ ELISA
plasma/ LUMINEX
Mother
Father
P3
P4
P5
P6
P7

Fig. 6. IRF4^{T95R} showed increased nuclear localization and affinity for DNA, altered specificity, and a different IRF4-binding landscape. (**A**) 293T cells were transfected with an empty vector (EV) or a vector expressing IRF4^{WT} or IRF4^{T95R}. Nuclei were stained with DAPI (blue), cytoplasm with phalloidin (red), and IRF4 with an anti-IRF4 antibody (green). Left, representative images; right, a summary of randomly chosen cells. (**B**) Ratio of nuclear to cytoplasmic IRF4 in Raji cells transduced with retrovirus expressing IRF4^{WT} or IRF4^{T95R}. Left, representative immunoblot; right, mean ± S.D. of three independent experiments. (**C**) IRF4^{T95R} showed increased affinity for an ISRE, two AICEs and an EICE site. (**D**) Fractions of all bound SiR-HaloTag- IRF4^{WT} and IRF4^{T95R} molecules (left) and molecules long bound for > 2s (right) as determined by single-molecule fluorescence microscopy with interlaced time-lapse illumination. (**E**) HEK293 cells were transfected with control plasmid (-), IRF4^{WT} or IRF4^{T95R}. Nuclear extracts were analyzed by EMSA using 3xGAAA ISRE. Supershifts (ss) of WT and T95R extracts using HAtag antibody or IgG control are shown at the far right. Note, that IRF4^{T95R} binds more strongly to ISRE compared to IRF4^{WT}. Dashed lines indicate cuts of the scan for presentation. (**F**) Top, IRF4^{WT} (left) and IRF4^{T95R} (right) motifs discovered in the HT-SELEX data. Bottom, 8-mers containing GAAA (left) or GATA (right) enriched in IRF4^{WT} (x-axis) or IRF4^{T95R} (y-axis). (**G**) 293T cells were transfected with a TK-Cypridina luciferase vector (an internal control) and either a canonical (ISRE)1-driven luciferase vector (upper panel) or a noncanonical (ISRE)1-driven luciferase vector (lower panel), together with a pFLAG-CMV empty vector (400 ng) or increasing amounts of plasmids encoding IRF4^{WT} or IRF4^{T95R}. The luciferase activity was compared to that induced by the empty vector, which was set to 1. Mean ± S.D. of 2-4 independent experiments is shown. (**H**) ChIP-seq analysis of immortalized B cells from P3 compared to a HC. Top left, overlay of IRF4 ChIP-seq peaks in EBV-B cells of P3 and HC. From left to right, ISRE, AICE and EICE motifs discovered in IRF4^{T95R}, IRF4^{WT}, AICE or EICE ChIP-seq data (indicated at the left of the motifs). The importance of each motif towards the IRF4^{T95R}-specific (purple), IRF4^{WT}-specific (green) or common (grey) component of the ChIP-seq data is shown to the right of each motif. Noncanonical motifs are surrounded by a purple line. (**I**) Normalized IRF4^{WT}-specific (green, top), common (grey, middle) and IRF4^{T95R}-specific (purple, bottom) ChIP-seq peak counts (y-axis) for different groups of differentially expressed genes. (**J**) HEK293 cells were transfected with AP-1 (JUNB and BATF) with or without IRF4^{WT} or IRF4^{T95R}, as indicated. Nuclear extracts were analyzed for binding to various CXCL13 sites, as indicated. Note, that IRF4^{T95R} shows strongly increased (CXCL13-A) or exclusive (CXCL13-B) binding compared to IRF4^{WT}. (**K**) HEK293 cells were transfected with CXCL13 reporter construct encompassing CXCL13 sites A and B together with AP-1 (JUNB and BATF) and IRF4 variants, as indicated. Luciferase activity is shown as fold activation compared to control transfected cells (far left), which is set as 1. Mean ± S.D. of 3 independent experiments is shown. (**L**) Fold change of CXCL13 levels in serum or plasma from P3-P7 compared to HC.

Supplementary Materials

Supplemental Materials and Methods

RNA isolation and RT-PCR

Total RNA of 500 activated naïve $CD4^+$ T cells was isolated and reverse transcribed by the CellLyse Microlysis and cDNA Synthesis Kit (TATAA Biocenter AB, Sweden). Primers used for the PCR amplification of IRF4 and ACTB cDNAs were listed in Table S13.

Flow cytometry

Flow cytometric phenotyping was performed either by using 100 µl fresh, heparinized full blood or cryopreserved PBMC. For surface staining, single cell suspensions were first incubated with either human BD Fc block or anti-mouse CD16/32 antibody on ice to block human or mouse FcγR and then stained with fluorophore-conjugated antibodies. 7-AAD Viability Staining Solution (eBioscience, Thermo Fisher Scientific) was used for live versus dead cell discrimination. For determination of regulatory T cells (Treg) by intracellular staining of FOXP3, we used the Human Regulatory T Cell Staining Kit from eBioscience (Thermo Fisher Scientific, USA) according to the manufacturers' instructions. For IRF4 intracellular staining, PBMCs were fixed and permeabilized after T and B surface marker staining using FoxP3 staining kit (eBioscience) according to the manufacturer's instructions. Cells were incubated with an Alexa Fluor® 647 anti-IRF4 Antibody (Biolegend) on ice for 30 min. Cells were washed with the staining buffer.

Th1, Th2 and Th17 cells of P5-P7 were detected by intracellular staining of cytokines: PBMC ($1x10^6$/ml) were either left unstimulated or stimulated with 10 ng/ml or 100 ng/ml PMA (Phorbol-12-Myristat13-Acetate) and 1 µg/ml Calcium-Ionophore (Merck, Germany) under addition of 1 µg/ml Brefeldin A (BD Biosciences, USA) for 6 or 12 h at 37°C. Cells then were harvested and washed twice with PBS/1% FCS. For surface staining, cells were incubated with anti-CD45RO-PE, anti-CD3-APC, anti-CD4-APCAlexaFluor750 and anti-CD45-KromeOrange or anti-CD3-PerCP and anti-CD4-PEcy7 for 30 min at 4°C. After being washed twiced with PBS/1% FCS, cells were fixed and permeabilized with 100 µl Cytofix/ CytopermTM (BD Biosciences, USA). For intracellular staining, cells were incubated with anti-IFNγ, anti-IL4 and anti-IL17A and anti-CD4 for 45 min at 4°C. Cells were washed twice with Perm/Wash Buffer (BD Biosciences, USA) and diluted in 500 µl PBS/1% FCS.

Samples were analyzed with a FACSVerse flow cytometer (BD Biosciences), a NaviosTM Flow Cytometer (Beckman Coulter, USA), a FACS Canto II flow or a Fortessa LSR cytometer (BD Biosciences, USA). Data analysis was performed using FlowJo software (BD Biosciences, USA).

Immunoblot

Whole-cell lysates extracted with medium radioimmunoprecipitation assay (RIPA) lysis buffer (Cwbiotech) and cytoplasmic and nuclear lysates extracted with Nuclear and Cytoplasmic Protein Extraction Kit (Sangon Biotech) were resolved by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to an immobilon-P membrane (MILLIPORE). Immunoblots were performed with specific primary antibodies overnight at 4℃ followed by horseradish peroxidase (HRP)-conjugated secondary antibodies for 1 h at room temperature and developed with an enhanced chemiluminescence light (ECL) reagent (MILLIPORE). For immunoblotting in T cell blasts, total PBMCs were stimulated with anti-CD3 and anti-CD28 antibodies in the presence of IL2 (10 ng/ml) for 10 days. Fresh IL2 was added every 2-3 days. T cell blasts were washed with PBS once and lysed in the lysis buffer (50mM Tris pH 7.4, 150mM

NaCl, 2mM EDTA, 0.5% Triton X-100 and 0.5% NP40 and protease and phosphatase inhibitor cocktail [Sigma, PPC1010]). Cell lysates were subjected to the immunoblotting.

Evaluation of cell proliferation

Total PBMCs were incubated with CellTrace violet Cell Proliferation Kit (1µM; Invitrogen) according to manufacturer's instructions. Cells were stimulated with anti-CD3 antibody and anti-CD28 antibodies (1 µg/mL each, eBioscience) for four days. For B cell proliferation, cells were stimulated with indicated agonists; anti-IgM (10 µg/ml, Jackson Lab), CD40L (100ng/ml, Enzo), IL4 (50 ng/ml, Peprotech), CpG (500 nM, Enzo) for four days. Cells were stained with fluorochrome-conjugated CD4, CD8, or CD19 antibodies (BD Biosciences, Clone RPA-T4, RPA-T8, HIB19 respectively) for 30 minutes at 4°C (dark). Cells were washed with PBS twice, and cells were acquired and analyzed by flow cytometry (Becton Dickinson FACSCanto II) and FlowJo software (FlowJo 10.5.2, TreeStar).

Enzyme linked immunosorbent assay (ELISA)

Human IgM and IgG levels in culture supernatants were measured as described (*66*) or measured by ELISA kits (Abcam; Cat# ab195215, ab214568, ab196263). IL2 and IFNγ cytokine production was analyzed by ELISA (Thermo Fisher Scientific) according to the manufacturer's instructions. Levels of mouse IgA, IgM, IgG, IgG_1, IgG_{2b}, IgG_{2c}, and IgG_3 and IgM and IgG1 in serum or culture supernatants were measured by ELISA performed as described previously (*49*). Levels of CXCL13 in patient and healthy controls sera were measured using the Quantikine ELISA Human CXCL13/BLC/BCA-1 Kit (R&D Systems; USA).

Flow cytometry for TCR Vbeta repertoire analysis

The Vβ T-cell repertoire was evaluated according to the manufacturer's instructions with the Beta Mark TCR Vbeta Repertoire Kit (Beckman-Coulter, USA), which allows detection of 24 different Vβ T-cell receptors.

NK-cell degranulation (CD107a assay)

Cryopreserved PBMCs ($2x10^6$/ml) were thawed, washed and incubated overnight at 37°C in either cell-culture medium (RPMI/ 10% FCS, PAN Biotech, 1% Penicillin/Streptomycin, 14 mM Hepes, and 2 mM L-Glutamin (Gibco/Thermo Fisher Scientific, USA), alone or in medium supplemented with 600 U/ml IL2 (PeproTech/tebu-bio, Germany). After harvesting and washing, 100 µl ($2x10^5$) of cells were cocultivated with 100 µl ($2x10^5$) K562 cells (DSMZ, Germany) or 100 µl of medium alone after adding of 5 µl/well anti-CD107a FITC in a 96 well U-bottom plate for 2.5 h at 37°C. Cells were stained with anti-CD107a FITC, anti-CD56 ECD, anti-CD3 APC and CD45 Krome Orange. Samples were acquired in 500 µl IF-Medium with a NaviosTM Flow Cytometer and analyzed with the NaviosTM 1.2. Software (Beckman Coulter, USA).

CD8 T-cell degranulation (CD107a assay)

Cryopreserved PBMCs ($1x10^6$/ml) were thawed, washed and cultured in cell-culture medium with 100 U/ml IL2 and with or without adding of 2% PHA (Gibco/Thermo Fisher) for 4 days at 37°C. After harvesting and washing, $2x10^5$ cells/200 µl were either restimulated with 15 µl CD3/CD28 beads (Gibco/ Thermo Fisher Scientific, USA) or incubated in culture-medium alone after adding of 5 µl/well anti-CD107a FITC in a 96 well U-bottom plate for 3 h at 37°C. Cells were then stained with anti-CD107a FITC, CD3 APC, CD8 Pacific Blue, CD45 Krome Orange. Samples were acquired in 500 µl IF-Medium with a NaviosTM Flow Cytometer (Beckman Coulter, USA) and analyzed with the NaviosTM 1.2. Software (Beckman Coulter, USA).

Construction of IRF4-expression vectors and plasmid transfection, and retrovirus and lentiviral transduction

Amplification of WT and T95R IRF4 cDNA was performed with KOD-plus polymerase (TOYOBO) using the first-strand cDNA generated from total RNA of PBMC from the patient and HC. Primers were listed in Table S13. The PCR products were inserted into pFlag-CMV-5a (Sigma-Aldrich) expression vector, pMX-IRES-GFP and pMX-IRES-hCD8a retroviral vectors. To generate the pcDNA3-FLAG-JUNB and pcDNA3-FLAG-BATF expression constructs, full length human JUNB and BATF were amplified from cDNA of the human cell line L428, and cloned via BamHI and XhoI into pcDNA3-FLAG (Invitrogen). The pHEBO-IRF4-HAtag expression construct and its empty control pHEBO-CMV-HAtag were kindly provided by L. Pasqualucci (New York, USA). The c.284C>G mutation was introduced into the pHEBO-IRF4-HAtag construct by use of the QuikChange Multi Site-Directed Mutagenesis Kit (Stratagene). For the retroviral transduction experiments of the C57BL/6 splenic B cells, the coding sequences for human IRF4 (WT or T95R) were amplified from the pHEBO-constructs using the IRF4_XhoI_forw 5′- ACC TCG AGG CCA CCA TGA ACC TGG AGG GCG GCG GCC GA - 3′and the IRF4_EcoRI_rev 5′- ACG AAT TCT TAA GGC CCT GGA CCC AAA GAA GCG TAA TC - 3′ primers and cloned in front of the IRES sequence of the MSCV-IRES-GFP (MIG) plasmid (a kind gift of F. Rosenbauer, Münster, Germany) via XhoI and EcoRI. All constructs were verified by sequencing.

Each retroviral construct was transfected either into Ampho or PlatE packaging cells (*67*) and the virus supernatants collected after 48 h and 72 h were filtered and concentrated with PEG8000. Virus transduction of Raji, Ramos, IRF4-deficient Ramos (clone 1-9 and 2-2), and CH12 cells was carried out in the presence of 8 µg/ml polybrene reagent (Sigma). For mouse splenic B cell transduction, purified B cells (density: $1x10^6$ cells/mL; $4x10^6$ cells per well) were cultured in the presence of recombinant mouse IL4 (25 ng/ml) and LPS (20 µg/ml) over night. 24 h after isolation, B cells were collected (300g, 5min, 4°C) and resuspended in B cell medium supplemented with 8 µg/mL polybrene (EMD Millipore, #TR-1003) at a density of $2x10^6$ cells/ml. To introduce WT and T95R IRF4, $4x10^6$ B cells per well were plated in 2 ml on 6-well plates that had been coated with RetroNectin (Takara, #T100B; 25 µg/mL; 4°C, overnight), blocked with 2% BSA (in PBS, 1h, RT) and pre-loaded with the respective retroviral particles (1 h, 37°C). Retroviral transduction was performed by the addition of 2 ml of the respective retroviral supernatant and subsequent centrifugation (800g, 90min, 32°C). 24 h after transduction, B cells were collected (300g, 5min, 4°C), resupended in new B cell medium and cultured (density: $1x10^6$ cells/ml; $4x10^6$ cells per well) for another 72 h (FACS for RNA-seq, flow cytometric analysis of plasma cell differentiation) in the presence of recombinant mouse IL4 (25 ng/ml) and LPS (20 µg/ml).

For transient transfection of 293T cells, Opti-MEM (GIBCO) and Hieff Trans™ Liposomal Transfection Reagent (YEASEN) were used following the company's protocol. For analysis of luciferase activity, HEK293 cells were transfected by electroporation in OPTI-MEM I using Gene-Pulser II (Bio-Rad) with 960 µF and 0.18 kV with 5 µg of pGL3-based reporter constructs, together with 150 ng pRL-TKLuc as an internal control. Where indicated, cells were additionally transfected with 5 µg pcDNA3-FLAG-JUNB, 5 µg pcDNA-FLAG-BATF, or 40 µg of the respective pHEBO-IRF4 variants. 48 hours after transfection, the ratio of the two luciferases was determined (Dual luciferase kit; Promega).

Surface plasmon resonance

- Protein Expression and purification

Codon-optimised gene constructs of $IRF4^{WT}$ and $IRF4^{T95R}$ DBD domains were cloned into pJ411KanR (ATUM) plasmids and transformed into *E. coli* BL21 (DE3) (Novagen) for expression as an N-terminal His_6 tag-fusion protein. Bacterial cultures were grown up in 2X YT media with over expression of constructs induced at 18 °C with 0.6mM IPTG. Following overnight incubation, cell pellets were suspended in 50 mM sodium phosphate buffer (pH 7.0), 500 mM NaCl and 30 mM imidazole (Buffer A) supplemented with protease inhibitor cocktail (Sigma), 3 mM β-mercaptoethanol, 0.5% Triton X-100, 4 mM MgCl2 and lysed via French press at 1500 psi. The DBD proteins were purified via Nickle affinity purification using a 5 ml HisTrap column (Cytiva) in Buffer A, against a linearly increasing imidazole gradient of up to a concentration of 500 mM.

Following isolation, the N-terminal His_6 tag was cleaved from DBD using overnight HRV3c digestion at 4 °C. Through a second round of Nickle affinity purification (as previously described), $IRF4^{WT}$ and $IRF4^{T95R}$ DBD proteins were isolated from the cleaved His tags. Further purification of $IRF4^{WT}$ and $IRF4^{T95R}$ DBD proteins were achieved via size exclusion chromatography (SEC) using a superdex 200 16/600 gel filtration column (Cytiva) in a 20 mM Tris buffer (pH 7.4), 150 mM NaCl, 1 mM TCEP buffer.

- Affinity determination

Affinity values of both IRF4 DBD constructs were acquired on a Biacore 8K (Cytiva) at 20°C with HBS buffer (10 mM HEPES-HCl, pH 7.4, and 150 mM NaCl) with additional 3 mM ETDA and 0.05 % P20 used as the running buffer. The biotinylated DNA motifs of interest (ISRE, EICE1, AICE1, and AICE2) (Integrated DNA technologies) were immobilised (up to ~ 2,000 RU) on a Series S streptavidin (SA) chip (Cytiva) as per manufacturer's instructions. Affinity measurements were gained by passing serially increasing concentrations of $IRF4^{WT}$ and $IRF4^{T95R}$ DBD proteins (up to 5 µM) over each of the coupled DNA motifs at a flow rate of 30 µl/min. The final response unit was calculated by subtracting the response unit of the reference flow cell. The steady-state multi cycle affinity data were fitted using the Biacore 8K BIAevaluation software.

Single-molecule fluorescence microscopy

- Generation of stable cell lines

HeLa cells, which stably express the IRF4-HaloTag fusion protein were produced by lentiviral transduction (*68*). Briefly, HEK293T cells were transiently transfected with the plasmids psPAX2 (Addgene, USA), pMD2.G (Addgene, USA) and pLV-tetO IRF4-HaloTag using JetPrime (PolyPlus, France) according to the manufacturers protocol. The virus in the supernatant of these cells was harvested through a 0.45 µm filter after 48 h. With this, HeLa cells were infected for 72 h at 37°C and 5% CO2.

- Preparation of cells for imaging

Cells were seeded on a heatable glass bottom Dela T dish (Bioptechs, USA) one day before imaging. Prior to imaging, cells were incubated in 2 pM silicon rhodamine (SiR) HaloTag ligand (kindly provided by Kai Johnson, MPI, Heidelberg, Germany) for 15 min following the HaloTag staining protocol (Promega, Germany). Subsequently, the cells were washed with PBS and recovered for 30 min in DMEM at 37 °C and 5% CO2. For imaging, the cells were washed three times with PBS and imaged in 2 ml OptiMEM.

- Microscope setup

We used a custom-built fluorescence microscope for single-molecular fluorescence imaging (*69*). It was built around a TiE Nikon microscope body and equipped with a 638 nm laser (IBEAM-

SMART-640-S, 150 mW, Toptica), AOTF (AOTFnC-400.650-TN, AA Opto-Elctronic, France) and a high-NA objective (100x, NA 1.45, Nikon). The fluorescence signal passed a multiband emission filter (F72-866, AHF, Germany) and was detected by an EMCCD camera (iXon Ultra DU 897, UK).

- Interlaced time-lapse illumination and data analysis

We illuminated the cells with a highly inclined light beam (*70*) using an interlaced time-lapse illumination scheme (ITM) (*69*). In ITM, we repeated a pattern of two consecutive images with 50 ms camera integration time followed by a dark-time of 2 s. To detect and localize fluorescent molecules within an image and to track molecules across consecutive images we used the analysis software TrackIt v1.0.1 (*71*). Parameters for detection and tracking were: 'threshhold factor' 3, 'tracking radius' 2, 'min. track length' 2, 'gap frames' 0, 'min. track length before gap frame' 0. We classified molecules that were only detected within a single image as unbound, molecules that were detected in two consecutive images within an area of 0.35 μm^2 as short-bound and molecules tracked over at least one dark-time period as long-bound molecules. The ratio of all bound molecules (including short- and long-bound molecules) to all molecules (including long-, short- and unbound molecules) and of long-bound molecules to all molecules was calculated for each imaged cell. Significance between IRF4-WT and IRF4-T95R was tested with an unpaired, non-parametric t-test (Mann-Withney-test) using GraphPad Prism 9.0.1.

Deep learning models

Six different ExplaiNN models, each with 100 units, were trained on either $IRF4^{T95R}$, $IRF4^{WT}$, AICE or EICE ChIP-seq data, or the two *in vitro* datasets (i.e., HT-SELEX). The architecture of each unit was as follows:

- 1st convolutional layer with 1 filter (26x4; 19x4 for training on the HT-SELEX data), batch normalization, exponential activation to improve the representation of the learnt sequence motifs (*72*) and max pooling (7x7);
- 1st fully connected layer with 100 nodes, batch normalization, ReLU activation and 30% dropout; and
- 2nd fully connected layer with 1 node, batch normalization and ReLU activation.

The four models trained on $IRF4^{T95R}$, $IRF4^{WT}$, AICE and EICE ChIP-seq data were further used in a transfer learning strategy to initialize the filter weights of an additional "surrogate" ExplaiNN model with 400 units (100 units x 4 models), which was fine-tuned on $IRF4^{T95R}$ and $IRF4^{WT}$ ChIP-seq data (Fig. S9C). During fine-tuning, filter weights were frozen to prevent them from being refined (i.e., the surrogate model was only allowed to learn the weights of the fully connected and final linear layers).

For training the $IRF4^{T95R}$ and $IRF4^{WT}$ models, ChIP-seq peaks were resized to 201 bp by extending their summits 100 bp in each direction using BEDTools slop (version 2.30.0) (*73*). Negative sequences were obtained by dinucleotide shuffling each dataset using BiasAway (version 3.3.0) (*74*). Sequences were randomly split into training (80%), validation (10%) and test (10%) sets using the "train_test_split" function from scikit-learn (version 0.24.2) (*75*) (datasets were always randomly split in this way).

For training the AICE and EICE models, BATF (ENCFF728KFD), IRF4 (ENCFF113VGD), JUNB (ENCFF912OPT) and PU.1 ChIP-seq data (ENCFF492ZRZ) from GM12878 cells were retrieved from ENCODE (*22*). ChIP-seq peaks were resized to 201 bp, similarly to the $IRF4^{T95R}$ and $IRF4^{WT}$ datasets. AICE peaks were obtained by intersecting the BATF, IRF4 and JUNB peaks using BEDTools intersect. Similarly, EICE peaks were obtained by intersecting IRF4 and PU.1

peaks. In both cases, non-overlapping IRF4 peaks were used as negative sequences, after which sequences were randomly split into training (80%), validation (10%) and test (10%) sets.

For fine-tuning the transfer learning model, $IRF4^{T95R}$ and $IRF4^{WT}$ 201 bp-long ChIP-seq peaks were intersected using BEDTools. Since the number of $IRF4^{T95R}$-specific, $IRF4^{WT}$-specific and intersected (i.e., common) peaks was different, we subsampled them to an equal number while accounting for their %GC content distributions. The resulting sequences were randomly split into training (80%), validation (10%) and test (10%) sets.

For training the HT-SELEX models, each cycle was treated as an independent class as in Asif and Orenstein (*76*), thereby removing the need for negative sequences. Reads were randomly split into training (80%), validation (10%) and test (10%) sets while preserving the proportions between reads from each cycle.

Models were trained as described in ExplaiNN. Briefly, using the Adam optimizer (*77*) and binary cross entropy as loss function, applying one-hot encoding, setting the learning rate to 0.003 (10 times lower for fine-tuning) and batch size to 100, and using an early stopping criteria to prevent overfitting. Models were also interpreted following the specifications from ExplaiNN. The filter of each unit was converted into a motif by aligning al sub-sequences activating that filter's unit by ≥50% of its maximum activation value in correctly predicted sequences. The importance of each motif was calculated as the product of the activation of its unit for each correctly predicted sequence activating that unit by ≥50% of its maximum activation value times the weight of the final layer of that unit for the prediction of $IRF4^{T95R}$-specific, $IRF4^{WT}$-specific, or common ChIP-seq peaks.

Clinical summaries of the patients

Family 1, Patient 1

Patient 1 (P1) was born to healthy, nonconsanguineous parents. He received all pediatric vaccines, including BCG at birth, that were all well tolerated. He was healthy until age of 6 months when he developed his first pneumonia. From the age of 6 months to 3 years, he experienced recurrent upper respiratory tract infections that responded appropriately to antibiotics. From age 3 to 8 years, he was admitted several times for recurrent pneumonias. Pathogens identified in the sputum included *Streptococcus pneumoniae, Haemophilus influenzae, Mycoplasma pneumoniae* and adenovirus. At age 7, he was diagnosed with hypoglobulinemia (IgG 1.96 g/L, IgM 0.08 g/L, IgA 0.02 g/L) and had reduced, but not absent, B cells in peripheral blood (CD19 104/uL). By the age of 8, bronchiectases were revealed by chest CT and IgG replacement therapy was initiated. He also received long-term antimicrobial prophylaxis including amoxicillin-clavulanate, clarithromycin, and fluconazole. Since then, the frequency of respiratory infections was reduced. Gastrointestinal manifestations were also prevalent in P1. He suffered from chronic diarrhea since the age of 3 years, and progressively worsening abdominal pain since age 11 years. Endoscopic examination showed proctocolitis, as well as mucosal eosinophilia (eosinophil count per high power field: 50 in the intestine, 100 in the descending colon, and 110-120 in the rectum). Abdominal CT indicated multiple enlarged lymph nodes in the abdominopelvic cavity and enhanced signal in part of the intestinal mucosa. Treatments with cyproheptadine, omeprazole, acyclovir, and sirolimus seemed to be ineffective, although abdominal pain was partially relieved by prednisone. Since the age of 10 years, P1 developed a non-itchy maculopapular rash on his face, limbs, the back of the hands, and feet. Peripheral blood EBV viral load at that time was positive at 1550 copies/mL. A skin biopsy showed scattered lymphocyte infiltrates around small vessels in the dermis while EBV immunohistochemistry was negative. At age 13 years, P1 received a hematopoietic stem cell transplant (HSCT) from a matched unrelated donor. At the time of this report, 7 months post-HSCT, the patient is fully engrafted with full chimerism and with complete resolution of his enteropathy.

Family 2, Patient 2

Patient 2 (P2) was born to non-consanguineous parents and no family history of immune disorders apart from a maternal history of rheumatoid arthritis. This previously healthy male tolerated all routine immunizations including rotavirus. At 11 months of age, he presented with respiratory failure requiring mechanical ventilation. Rhinovirus/enterovirus, *Pneumocystis jirovecii* pneumonia (PCP), cytomegalovirus (CMV) viremia and *Escherichia coli* ventilator-associated pneumonia ensued. The patient's infections were treated and resolved with standard antimicrobial therapies. He was maintained on a regimen of monthly intravenous IgG replacement therapy together with prophylactic cotrimoxazole and valganciclovir.

Initial immunological assessment was noteworthy for agammaglobulinemia (IgG <0.3 g/L, IgA < 0.04 g/L and IgM < 0.03 g/L), undetectable titers to tetanus and diphtheria toxoids and reduced lymphocyte proliferation stimulation index (SI) to phytohemagglutinin with a normal response to pokeweed mitogen. The presence of lymphocytes including peripheral T and B lymphocytes was confirmed (x 10e9 cells/L): CD3 cells 3.74, CD4 cells 2.82, CD19 cells 1.14, NK cells 0.04 (all normal for age). Normal T cell memory subsets. B cell memory subsets demonstrated a low proportion of class switched memory B cells and borderline increase in naïve B cells.

At the age of 2 years the patient received his first HSCT from a matched unrelated donor. HSCT was complicated by loss of donor graft three times due to ongoing multi-drug resistant CMV infection, CMV end-organ disease and *Aspergillus* lung infection. P2 is currently fully engrafted

with a CMV viral load of zero following a fourth HSCT from a haploidentical donor although follow-up time is only 6 months at the time of writing.

Family 3, Patient 3

Patient 3 (P3) was born to healthy, non-consanguineous parents. The patient presented in the first month of life with bronchiolitis and recurrent respiratory tract infections. He had diarrhea with documented rotavirus infection at 1.5 month of age and developed failure to thrive requiring supplemental enteral feeding. At the age of 6 months, he suffered from PCP, norovirus, enterovirus, and rhinovirus infection. Immunological investigations revealed agammaglobulinemia while B lymphocytes were present in the blood. Since then, he has been on immunoglobulin replacement therapy and cotrimoxazole prophylaxis. The percentage of memory B cells ($CD19^+CD27^+$) was always below reference values. IgA levels were undetectable (<0.03 g/L) until the age of 4 years when they were found at 0.14 g/L, and levels remained low (0.15 g/L at 8 years, 0.40 g/L at 9 years 10 months, respectively). However, IgM serum level remained undetectable (<0.05 g/L). T cell $CD8^+$ naïve cells were increased in frequencies in contrast to CD8+ effector memory and $CD8^+$ terminal effector memory RA+ cells, which were below age matched values. The patient height and weight caught up after his initial failure to thrive and had normal development during childhood. He was well and attending regular school without significant health problems until the age of 9.5 years at which he suffered from a second episode of PCP infection due to poor adherence to cotrimoxazole prophylaxis.

Family 4, Patient 4

Patient 4 (P4) is a five-year-old boy born to non-consanguineous parents. He was born at term and remained healthy until he turned 6 months old, when he was admitted at the pediatric intensive care unit with a suspected PCP. Since then, he has had recurrent respiratory infections of unknown etiology, some of which required hospitalization. At the age of 15 months, he presented with prolonged diarrhea (non-responsive to antibiotics) that was accompanied by a 2.5kg weight loss. When he was 2y, he developed a localized regional BCGitis which responded to isoniazid, rifampicin, pyrazinamide, and ethambutol treatment. Immunologically, he showed reduced levels of circulating immunoglobulins (IgG 1.69 g/L, IgA 0.06 g/L, IgM 0.09 g/L) consistent with a diagnosis of hypogammaglobulinemia. Since the age of 2 years, he has been on prophylactic treatment with fluconazole, cotrimoxazole, and IgG replacement therapy. At the age of 5 years, he received a HSCT from his haploidentical mother (1-month post-HSCT at the time of manuscript submission).

Family 5, Patient 5

Patient 5 (P5) is a male patient who was born to non-consanguineous parents. He presented at the age of 4 months with recurrent respiratory infections; at the age of 7 months, he was admitted with a positive *Pneumocystis jirovecii* and CMV respiratory infection. With absent serum levels for IgG, IgM and IgA, absent specific antibodies for tetanus toxoid and *Haemophilus influenzae Type B* (after three vaccinations), a combined immunodeficiency was suspected. He recovered well on an anti-infectious regimen containing cotrimoxazole and ganciclovir. Immunophenotyping revealed normal counts for T, B and NK cells with a normal proportion of naïve CD4 positive cells. T-cell function after non-specific (PHA, anti-CD3/-CD28) and specific (tetanus toxoid, mixed lymphocyte culture) stimulation was found positive, but resulted negative against CMV. The patient was placed on IgG replacement therapy and cotrimoxazole prophylaxis. At the age of 20 months (7 months after cotrimoxazole discontinuation) he experienced a second *Pneumocystis jirovecii* pneumonia which required mechanical ventilation. He recovered with therapeutic

cotrimoxazole and remained on pneumocystis prophylaxis since then. Extensive onychomycosis involving his finger- and toenails has been persistent since the age of 3 years. In the next 20 years of follow up, no other serious infectious complications occurred, whilst the patient continued prophylaxis and he did not develop any relevant complications or organ damage.

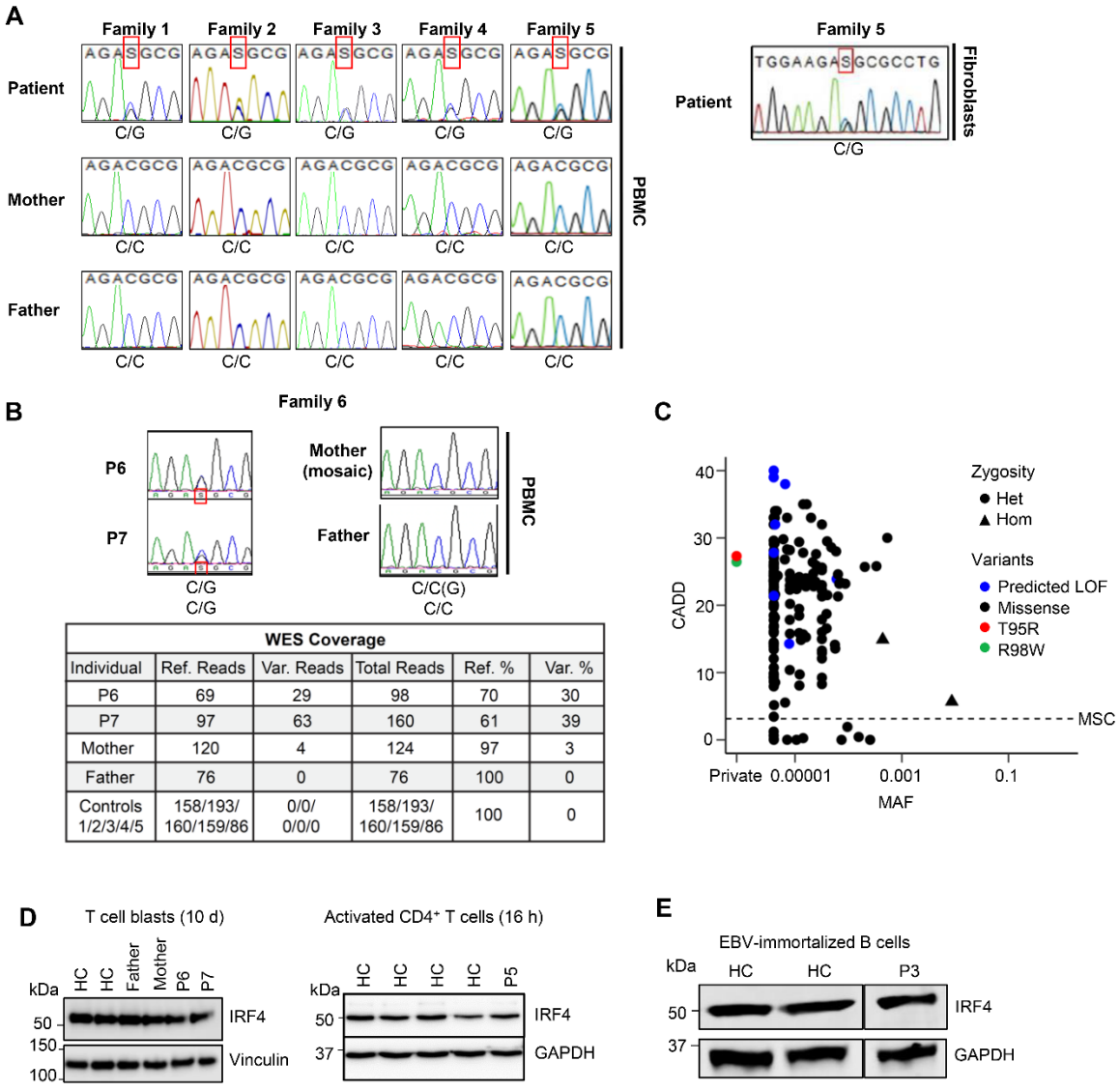
Family 6, Patient 6 and Patient 7

Patient 6 (P6) and Patient 7 (P7) were born to healthy, non-consanguineous Argentine parents. Both patients received all pediatric vaccines, including BCG at birth, that were well tolerated.

P6, the elder sister, was healthy until the age of 5 months when she started experiencing recurrent episodes of hypoxia, several of them requiring hospitalization. At age 6 month she was admitted with pneumonitis, suspected to be due to *Pneumocystis jirovecii* (no bronchoalveolar lavage or specific stains were performed, rapid response to cotrimoxazole was achieved). At age 1 year she who diagnosed with agammaglobulinemia (IgG: 0.10 g/L; IgA and IgM: not detectable) with B cells present in peripheral blood (CD19 3293/uL; 28% of total lymphocytes). Since early childhood she also presented with repeated upper sinopulmonary infections including suppurative otitis media, adenoiditis, sinusitis and oto-mastoiditis; the latter complicated with lateral sinus thrombosis for which she needed prolonged anticoagulation. Due to severe adenoid hypertrophy, at age 8 years an adenoidectomy was performed, and bilateral tympanostomy tubes were placed. She is currently on IgG replacement therapy and cotrimoxazole prophylaxis.

P7, 5 years younger than his sister P6, was also healthy until the age of 5 months when he started experiencing recurrent episodes of bronchospasms. At the age of 7 months, he was admitted with *Pneumocystis jirovecii* pneumonia, which responded adequately to cotrimoxazole therapy. During this admission he was diagnosed with agammaglobulinemia (IgG: 0.24 g/L; IgA: 0.02 g/L; IgM: 0.03 g/L) with reduced but present B cells in peripheral blood (CD19 463/uL; 7.7% of total lymphocytes). Due to severe tonsil hypertrophy, the patient had a tonsillectomy at age 4 years; soon after, a fast growing left cervical lymph node was also excised. The left tonsil and the left lymph node showed multiple acid-fast bacilli and granuloma formation by histopathologic analysis. The organism cultured from the lymph node was characterized as *Mycobacterium bovis*. Besides his IgG replacement therapy and cotrimoxazole, the mycobacterial infection is currently being treated with isoniazid, rifampin, and clarithromycin.

Supplemental Figure 1 (related to Figure 1)



WES Coverage					
Individual	Ref. Reads	Var. Reads	Total Reads	Ref. %	Var. %
P6	69	29	98	70	30
P7	97	63	160	61	39
Mother	120	4	124	97	3
Father	76	0	76	100	0
Controls 1/2/3/4/5	158/193/ 160/159/86	0/0/ 0/0/0	158/193/ 160/159/86	100	0

Fig. S1. Identification of a unique heterozygous mutation in the *IRF4* gene in seven CID patients. (**A**) Sanger sequencing of *IRF4* genomic DNA from PBMC of families 1-5 (left) and from fibroblasts of P5 (right). (**B**) Upper, Sanger sequencing of PBMC from P6, P7 and their parents. The mother appeared to carry a mosaic mutation. Below, Whole-exome sequencing (WES) from DNA obtained from peripheral blood revealed that 4 out of 124 sequences from the mother were mutated. (**C**) Minor allele frequency (MAF) and combined annotation-dependent depletion (CADD) score of all coding and essential splicing variants of IRF4. (**D**) Left, PBMC from P6 and P7, their parents and two HC were stimulated with anti-CD3/CD28 beads in the presence of IL2 for 10 days, and analyzed for IRF4 protein expression by immunoblot. Vinculin was used as a loading control; Right, $CD4^+$ T cells purified from P5 and 4 HC were stimulated with anti-CD3/CD28 beads for 16 h and analyzed for IRF4 expression by immunoblot. GAPDH was used as a loading control. (**E**) IRF4 protein levels in EBV-immortalized B cell lines from P3 and two HC. GAPDH was used as a loading control.

Supplemental Figure 2 (related to Figure 2)

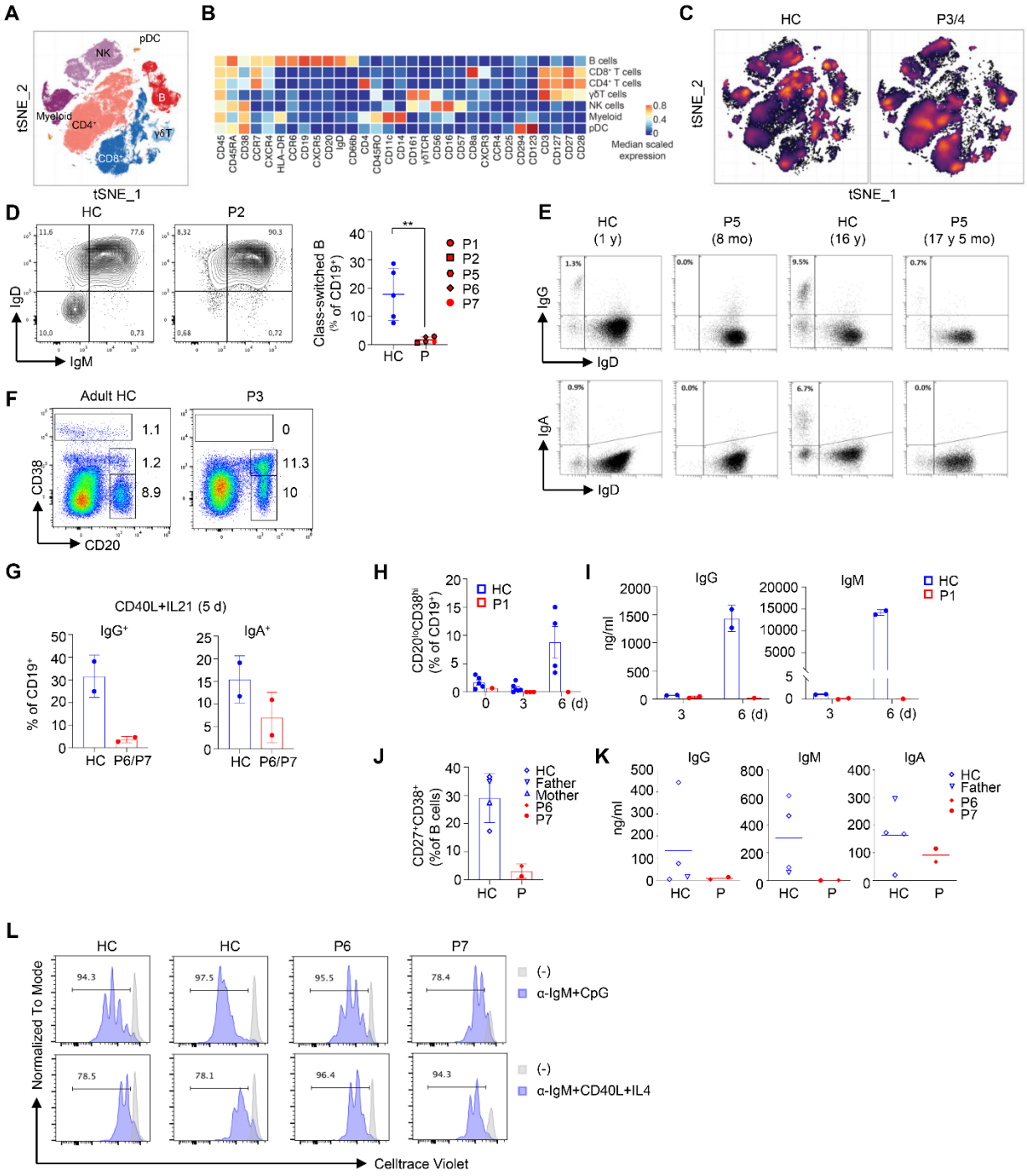


Fig. S2. Overt immunological phenotype, impaired Ig gene CSR and defective PC differentiation in IRF4^{T95R} patients. (A-C) CyTOF analysis of PBMC from P3, P4 and 5 healthy controls. **(A)** Dimensional reduction by t-SNE of the 31 markers used for immunophenotyping by mass cytometry. Each color represents a cell population obtained by manual clustering according to their surface marker expression. (**B**) Heatmap showing marker expression of the populations in Fig. S2A. (**C**) Density tSNE showing comparing cell distribution among and within the populations shown in Fig. S2A. (**D**) Left, representative flow cytometric profiles showing the IgM$^-$ IgD$^-$ class-switched B cells in gated CD19$^+$ cells; Right, summary of the data from P1, P2 and P5-P7 and their age-matched HC. (**E**) Frequencies of class-switched (IgG$^+$IgD$^-$ or IgA$^+$IgD$^-$) cells among CD19$^+$ B cells in PBMC from P5 (at ages of 8 m and 17 y 5 m) and two HC (1 y and 16 y). (**F**) CD38^{hi} plasmablasts were barely detectable in the BM of P3 compared with an adult HC. (**G**) PBMC of P6 and P7 and two HC were stimulated with CD40L + IL21 for 5 days. Percentages of IgG$^+$ and IgA$^+$ cells among CD19$^+$ B cells are shown. Data represent mean ± S.D. (**H**) PBMC of six HC and P1 were cultured with CMIL2 for three and six days and analyzed for PC differentiation. (**I**) IgG and IgM levels in the culture supernatants of Fig. S2H. (**J**) PBMC of two HC, P6 and P7 and their parents were cultured with CD40L + IL21 for four days and analyzed for PC differentiation. Data represent mean ± S.D. (**K**) Naïve B cells were stimulated with CD40L and IL21 for 6 days. IgG, IgM and IgA levels in the culture supernatants were measured. The horizontal lines indicate mean values. (**L**) Total PBMCs from P6 and P7 and two HC were cultured with the indicated stimuli for 4 days and B-cell proliferation was assessed by celltrace violet dilution. $^{**}p < 0.01$.

Supplemental Figure 3 (related to Figure 3)

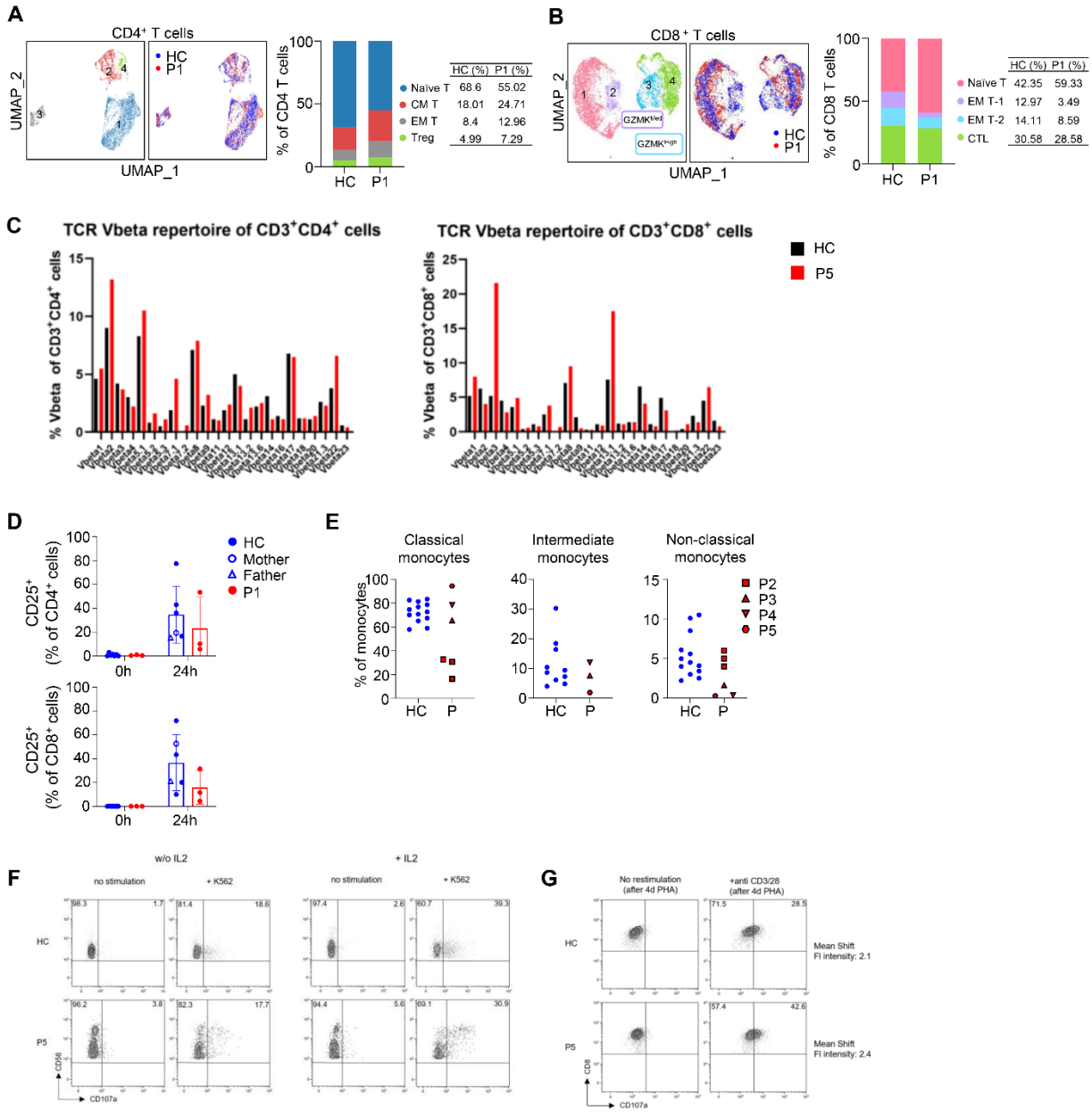


Fig. S3. Altered T cell differentiation and function in IRF4^{T95R} patients. (**A** and **B**) scRNA-seq of peripheral T cells from P1. (**A**) Left, UMAP profiles of CD4$^+$ T cells; Right, frequencies of naïve, CM, EM and Treg subsets among CD4$^+$ cells. (**B**) Left, UMAP profiles of CD8$^+$ T cells; Right, frequencies of naïve, EM T-1, EM T-2 and CTL subsets among CD8$^+$ cells. (**C**) TCR Vβ repertoire in CD3$^+$CD4$^+$ (left) and CD3$^+$CD8$^+$ (right) T cells of P5 and HC. (**D**) CD25 expression in CD4$^+$ (upper) and CD8$^+$ (lower) T cells of HC and P1 before (0 h) and after (24 h) stimulation with anti-CD3 + anti-CD28. (**E**) Proportions of classical (left), intermediate (middle) and non-classical monocytes among total monocytes in PBMC of P2- P5 (Supplemental Table 11). (**F**) NK-cell degranulation assay using NK cells from P5 PBMC as effector cells and K562 cells as target cells. Assays were performed without IL2 and with IL2. FACS profiles showing CD107A release by NK cells. (**G**) CD8$^+$ T cells from HC and P5 were stimulated with PHA for 4 days and then re-stimulated with anti-CD3/anti-CD28, and analyzed for CD107a expression (degranulation).

Supplemental Figure 4 CD4 T (related to Figure 3)

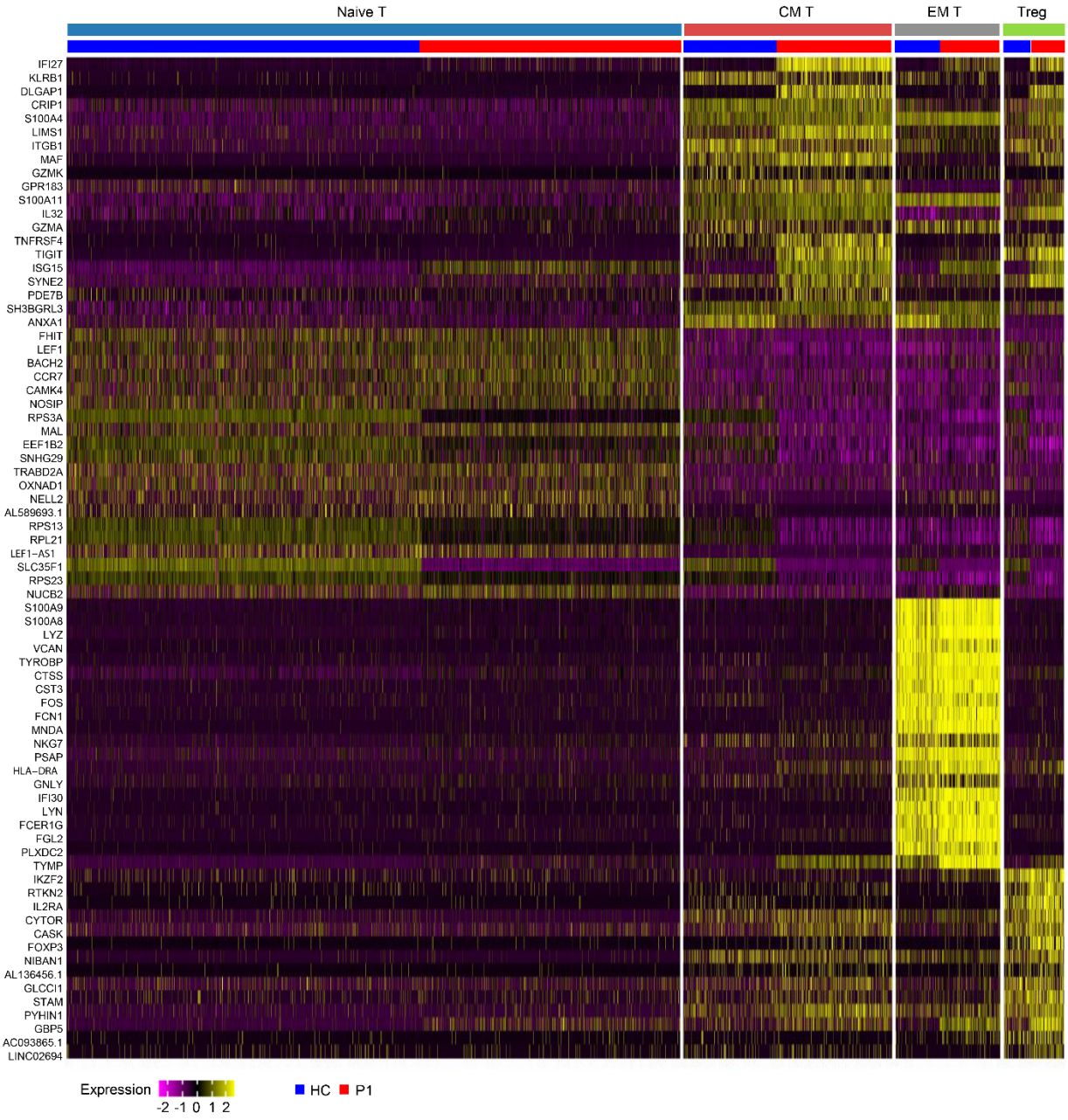


Fig. S4. Heat map of Top 20 differentially expressed (DE) genes in the four CD4 T cell clusters, colored as in Fig. S3A.

Supplemental Figure 5 CD8 T (related to Figure 3)

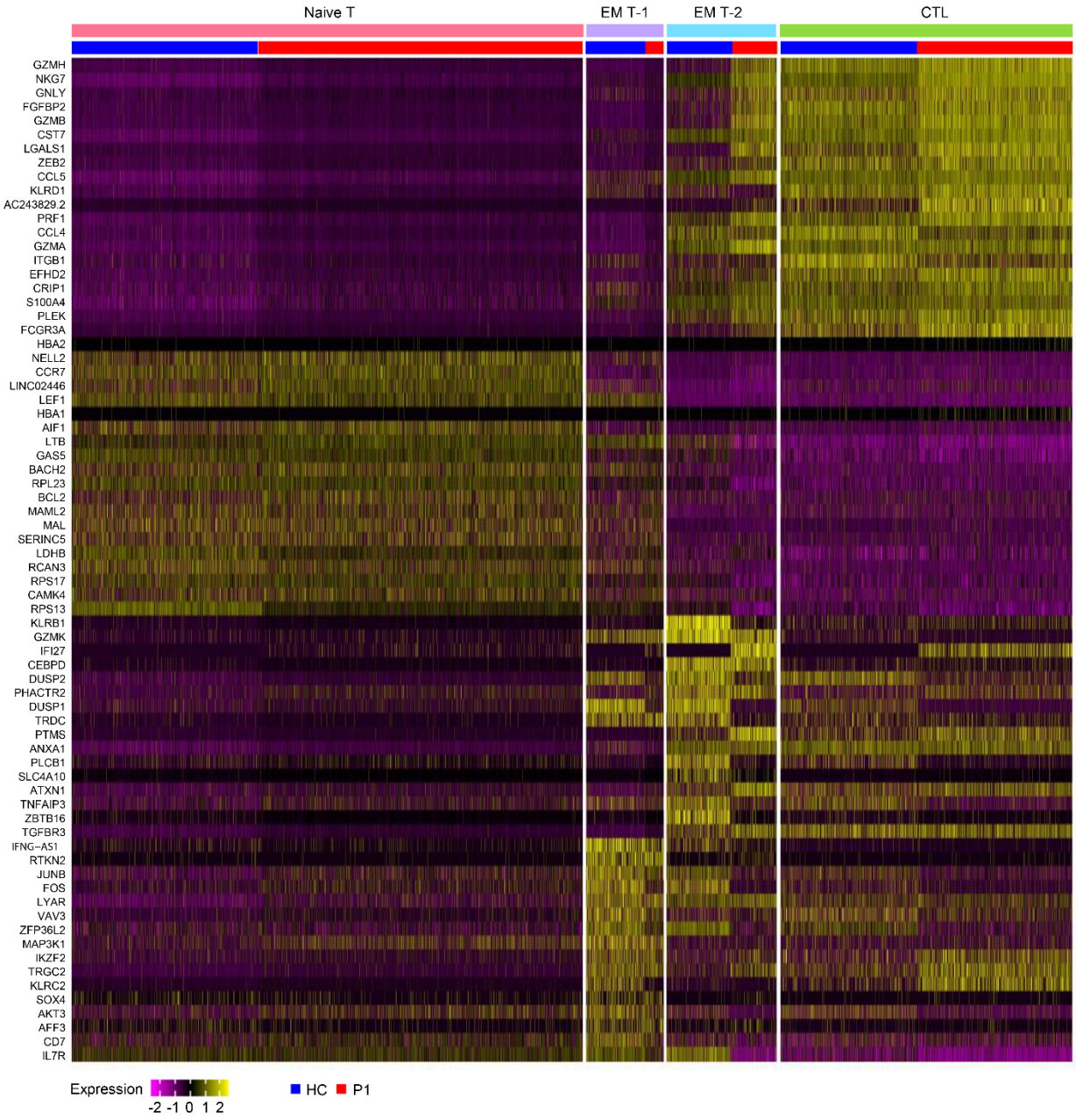


Fig. S5. Heat map of Top 20 DE genes in the four CD8 T cell clusters, colored as in Fig. S3B.

Supplemental Figure 6 (related to Figure 4)

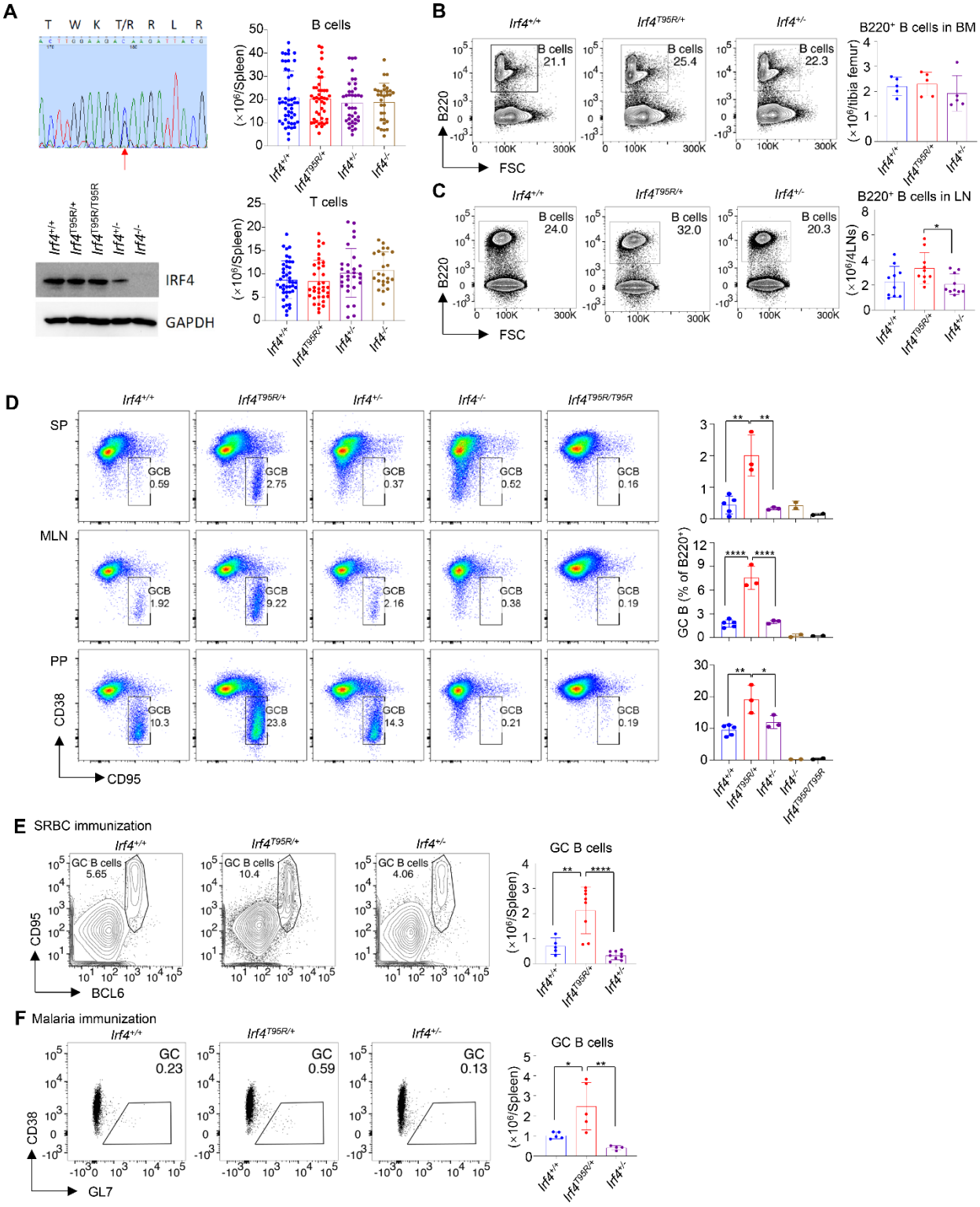


Fig. S6. Spontaneous GC B-cell expansion but defective antibody production in *Irf4*$^{T95R/+}$ mice. (**A**) Upper left, sequencing profile of the founder mouse generated in ANU. Lower left, an IRF4 immunoblot using spleen B cells purified from different genotypes. Number of B (upper right) and T (lower right) cells in the spleen of WT and mutant mice. (**B**) Representative proportions and summary of numbers of $B220^+$ cells in the BM. (**C**) Representative proportions and summary of numbers of $B220^+$ cells in the LN. (**D**) Percentages of the $CD38^-CD95^+$ GC B cells among $B220^+$ cells in the spleen (SP), mesenteric lymph nodes (MLN) and Peyer's patches (PP) of *Irf4*$^{+/+}$, *Irf4*$^{T95R/+}$, *Irf4*$^{+/-}$ *Irf4*$^{-/-}$ and *Irf4*$^{T95R/T95R}$ mice. Left, representative flow cytometric profiles; Right, summary data. (**E**) Representative proportions and summary of numbers of $BCL6^+CD95^+$ GC B cells among $B220^+$ cells in the spleen after SRBC immunization. (**F**) Representative proportions and summary of numbers of $GL7^+CD38^-$ GC B cells among $B220^+$ cells in the spleen after immunization with irradiated *Plasmodium* sporozoites. *$p < 0.05$; **$p < 0.01$; ****$p < 0.0001$.

Supplemental Figure 7 (related to Figure 4)

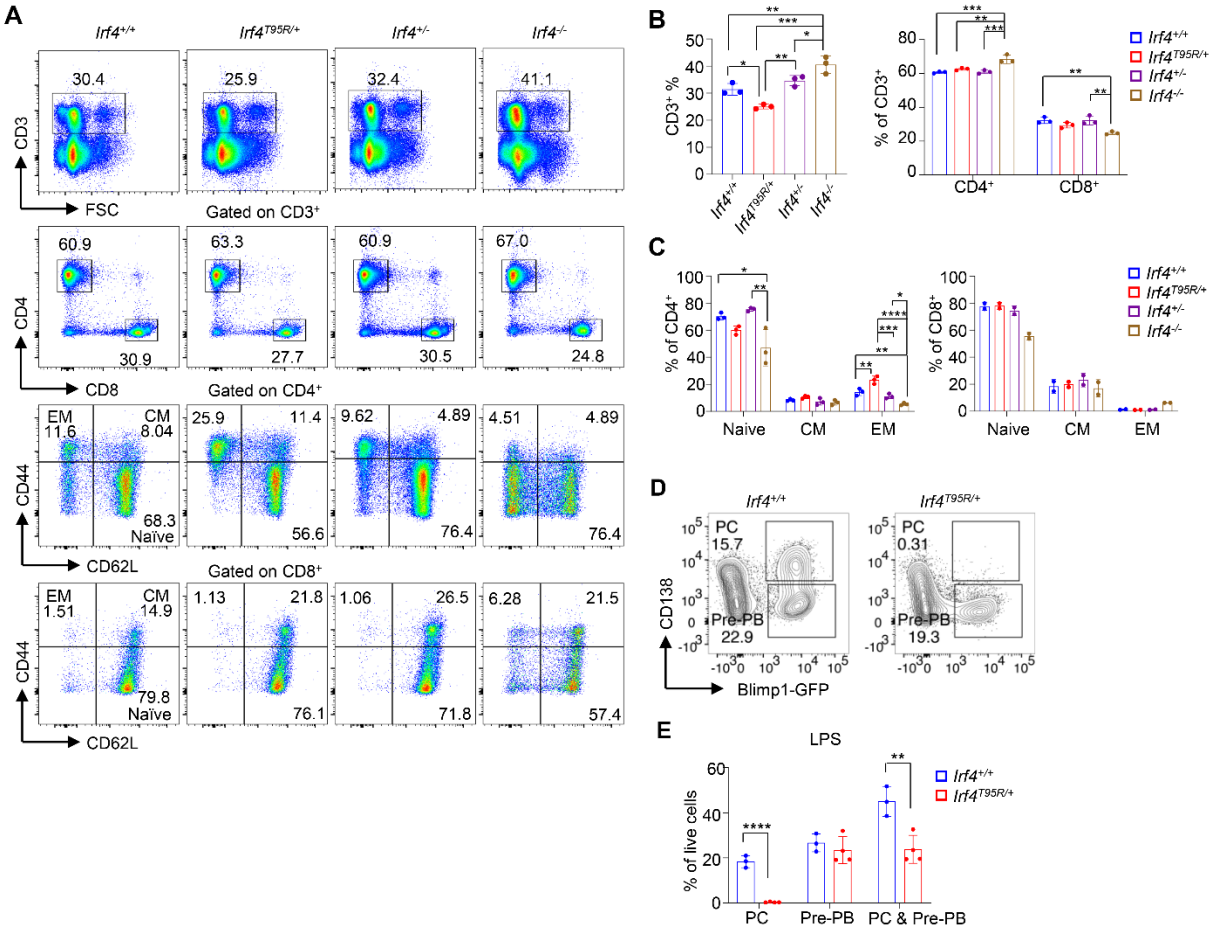


Fig. S7. T-cell subset distribution in the spleen of *Irf4*$^{+/+}$, *Irf4*$^{T95R/+}$, *Irf4*$^{+/-}$ and *Irf4*$^{-/-}$ mice. (**A**) From top to bottom, representative flow cytometric profiles of CD3 vs. FSC, CD4 vs. CD8 in gated CD3$^+$ cells, and CD44 vs. CD62L in gated CD4$^+$ or CD8$^+$ T cells. (**B**) Percentages of CD3$^+$ T cells (left) among live cells and CD4$^+$ and CD8$^+$ T cells among CD3$^+$ cells (right). (**C**) Proportions of naïve (CD62L$^+$CD44$^-$), CM (CD62L$^+$CD44$^+$) and EM (CD62L$^-$CD44$^+$) T cells among total CD4$^+$ (left) and CD8$^+$ (right) cells. (**D**) Spleen B cells purified from WT or *Irf4*$^{T95R/+}$ Blimp1-GFP reporter mice were cultured for 3 days in the presence of LPS and analyzed for the induction of GFP and CD138 expression. *Irf4*$^{T95R/+}$ B cells were able to activate Blimp1 transcription but failed to upregulate CD138 expression. (**E**) Summary of 3 independent experiments shown in Fig. S7D. Data are presented as mean ± S.D. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

Supplemental Fig 8 (related to Figure 5)

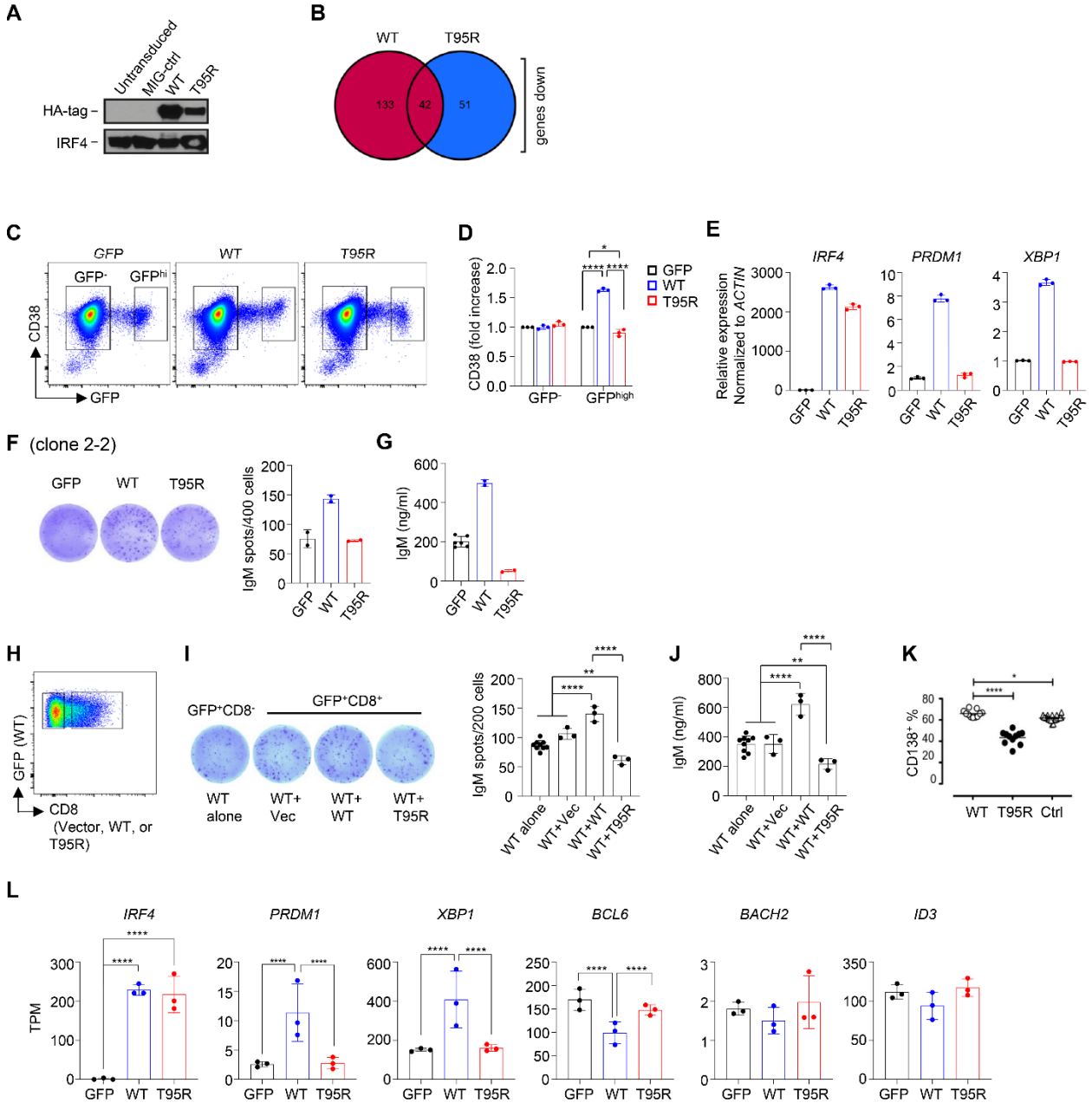


Fig. S8. IRF4^{T95R} was unable to induce PC differentiation due to the inability to activate the expression of genes involved in the PC differentiation program. (A) Mouse C75BL/6 splenic B cells were cultured with LPS+IL4, transduced with control retrovirus (MIG-ctrl), IRF4^{WT} or IRF4^{T95R}, or not transduced, and analyzed for protein levels of IRF4 by western blot using an antibody against HAtag for detection of ectopically expressed IRF4, or an antibody to IRF4 for detection of IRF4 overall protein levels. **(B)** Genes differentially downregulated by IRF4^{WT}, and IRF4^{T95R} show limited overlap. (**C** and **D**) Raji cells were transduced with retrovirus expressing GFP alone, IRF4^{WT} or IRF4^{T95R} and analyzed for CD38 expression in gated GFP^{-} or GFP^{high} cells 3 days later. (**C**) Representative FACS profiles showing the gate for GFP^{-} and GFP^{high} cells. (**D**) CD38 level in Raji cells expressing IRF4^{WT} or IRF4^{T95R}, relative to that in Raji cells expressing GFP alone, which was set to 1. Mean ± S.D. of 3 independent experiments is shown. (**E**) Raji cells were transduced with retrovirus as in Fig. S8C and analyzed for *IRF4*, *PRDM1* and *XBP1* expression in sorted GFP^{+} cells by real-time PCR. The value in Raji cells expressing GFP alone was set to 1. Representative results of 3 independent experiments are shown. (**F** and **G**) IRF4-deficient Ramos cells (clone 2-2) were transduced with retrovirus expressing GFP alone, IRF4^{WT} or IRF4^{T95R}. The GFP^{+} cells were sorted 3 days later and analyzed for IgM-secreting cells by ELISPOT. (**F**) Left, representative images; right, mean ± S.D. of duplicate wells. (**G**) The sorted cells were further cultured for 5 days and analyzed by for IgM levels in the culture supernatant by ELISA. Mean ± S.D. is shown. (**H-J**) IRF4-deficient Ramos cells were transduced with retrovirus expressing IRF4^{WT}-IRES-GFP together with retrovirus expressing IRF4^{WT}-IRES-CD8, IRF4^{T95R}-IRES-CD8 or an empty vector (**H**) and 3 days later GFP^{+}CD8^{-} and GFP^{+}CD8^{+} cells were sorted for the analyses of antibody-secreting cells by ELISPOT (**I**) and antibody secretion after 5 days' culture (**J**). (**K**) B cells were purified from tonsils of a HC by magnetic cell sorting, activated in vitro with CD40L and IL21 and transduced by lentivirus expressing WT IRF4 or T95R. After 6 days, cells were analyzed for CD138 expression. (**L**) Raji cells were transduced with retrovirus expressing GPF alone, IRF4^{WT} or IRF4^{T95R} and analyzed for the expression (Transcripts Per Kilobase Million, TPM) of *IRF4*, *PRDM1*, *XBP1*, *BCL6*, *BACH2* and *ID3* by RNA-Seq. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

Supplemental Figure 9 (related to Figure 6)

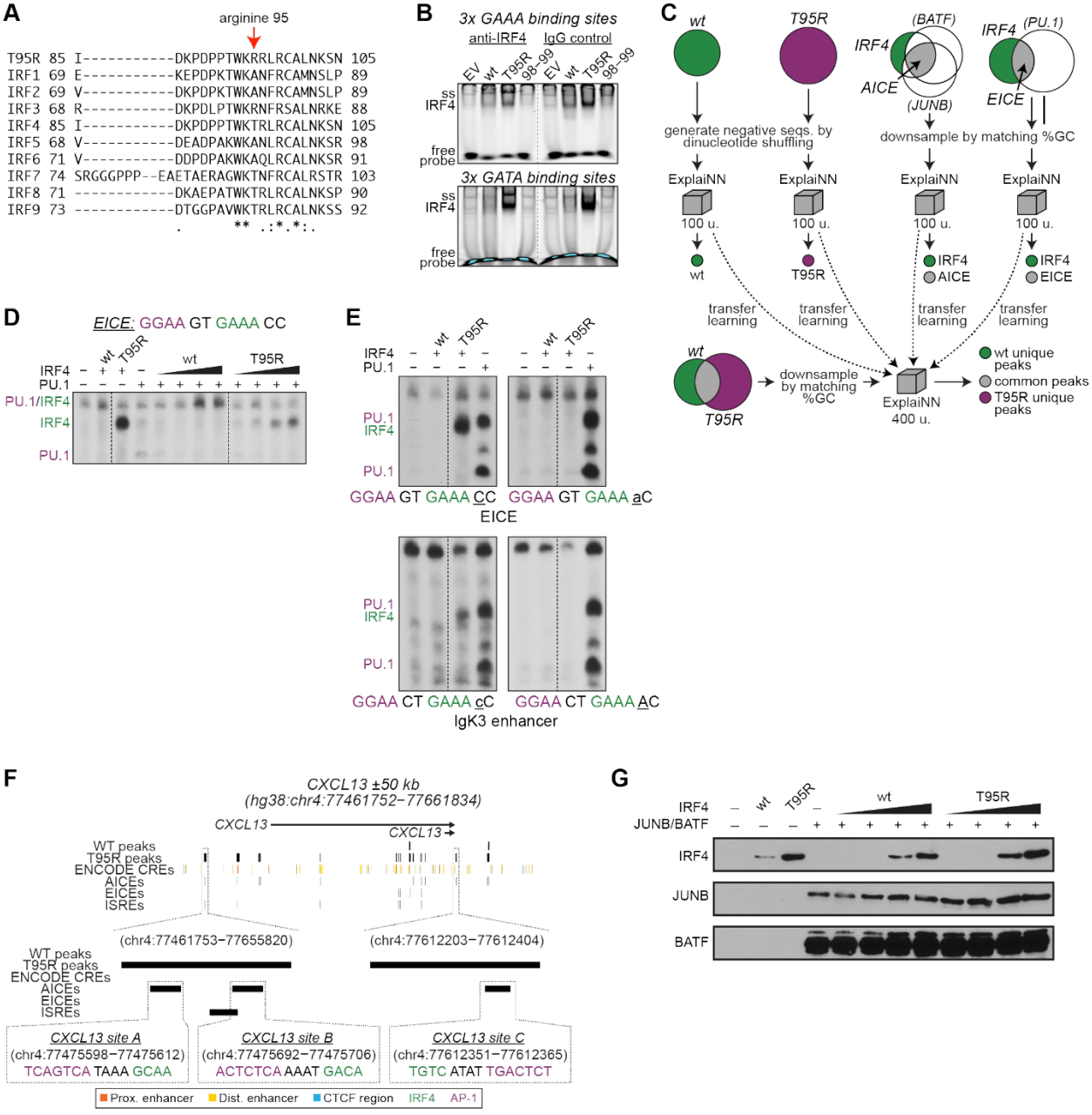


Fig. S9. IRF4^{T95R} showed increased affinity for DNA, altered specificity, and a different IRF4-binding landscape. (**A**) Partial multiple sequence alignment of the nine human IRF proteins and IRF4^{T95R}. (**B**) EMSA showing that IRF4^{T95R} bound to ISRE containing 3x GAAA sites more strongly than did IRF4^{WT} (upper) but also bound to a mutant ISRE with 3x GATA sites (lower). (**C**) Transfer learning strategy used in this work: pre-training four individual ExplaiNN models on IRF4^{WT}, IRF4^{T95R}, AICE or EICE ChIP-seq data, and fine-tuning a multi-task ExplaiNN model on IRF4^{WT}-specific, IRF4^{T95R}-specific, and common ChIP-seq data, previously initialized with the weights of the pre-trained models. (**D**) IRF4^{T95R} alone was able to bind to an EICE but failed to cooperate with PU.1. HEK293 cells were transfected with PU.1 with or without IRF4^{WT} or IRF4^{T95R}, as indicated. Nuclear extracts were analyzed for binding to an EICE site, as indicated. Dashed lines indicate cuts of the scan for presentation. (**E**) HEK293 cells were transfected with PU.1, IRF4^{WT} or IRF4^{T95R}, as indicated. Nuclear extracts were analyzed for binding to two different EICE sites (named EICE and IgK3 enhancer). Note, that increased binding affinity of IRF4^{T95R} depends on a C present in position +1 relative to the GAAA binding motif (present in wildtype configuration of probe EICE, indicated by capital, whereas the IgK3 enhancer probe contains an A in wildtype configuration at +1). Dashed lines indicate cuts of the scan for presentation. (**F**) Schematic overview of the genomic region encompassing the longest *CXCL13* transcript ±50 kb. The presence of IRF4^{WT} and IRF4^{T95R} peaks, as well as ENCODE *cis*-regulatory elements (CREs) and predicted AICE, EICE and ISRE sites in the region is shown. The two IRF4^{T95R} IRF4 ChIP-seq peaks encompassing the CXCL13 sites A, B and C are highlighted. (**G**). Detection of IRF4, JUNB and BATF protein by western blotting in nuclear extracts used for EMSA analyses in panel J of Figure 6.

Table S1. Clinical features of the patients

	P1	P2	P3	P4
Sex	Male	Male	Male	Male
Age	12y	12m	9y10m	4y8m
Age of onset	6m	11m	1m	6m
Clinical symptoms	Recurrent respiratory infections Chronic diarrhea	Recurrent respiratory infections	Recurrent respiratory infections Diarrhea	Recurrent respiratory infections Chronic diarrhea
Respiratory tract/lung	Sinusitis Bronchiectasis	Respiratory failure	Bronchiolitis	Bronchospasms
Skin	Limb rash			
Lymphadenopathy	+			+
Hepatosplenomegaly	+			
Intestinal tract	Mild colitis			
Others	Liver dysfunction			
Infections	*Streptococcus pneumoniae* *Haemophilus influenzae* Salmonella Enterovirus EBV Candida albicans	*Pneumocystis jirovecii* Enterovirus Rhinovirus CMV	*Pneumocystis jirovecii* (2x) Enterovirus Rhinovirus Norovirus Rotavirus	*Pneumocystis jirovecii* suspected BCGitis
Treatment	Antibiotics, IgG replacement HSCT	Antibiotics IgG replacement HSCT (4x)	Antibiotics IgG replacement	Antibiotics IgG replacement HSCT

	P5	P6	P7
Sex	Male	Female	Male
Age	22y	8y	3y
Age of onset	4m	6m	6m
Clinical symptoms	Recurrent respiratory infections	Recurrent infections	Recurrent bronchospasms
Respiratory tract/lung			Bronchospasms
Skin	Onychomycosis		
Lymphadenopathy			
Hepatosplenomegaly			
Intestinal tract			
Others			
Infections	*Pneumocystis jirovecii* (2x) CMV	*Pneumocystis jirovecii* suspected	*Pneumocystis jirovecii* *Mycobacterium bovis*
Treatment	Antibiotics Antimycotics Pneumocystis prophylaxis IgG replacement	Pneumocystis prophylaxis IgG replacement	Antibiotics Pneumocystis prophylaxis, IgG replacement

EBV, Epstein-Barr virus. CMV, cytomegalovirus. HSCT, hematopoietic stem cell transplantation

Table S2. Features of the IRF4 T95R mutation

Variant Annotation	
Chromosome	6
Genomic Position (GRCh38)	394888
cDNA Position (NM_001195286.2)	284
Nucleotide Reference	C
Nucleotide variant	G
Protein Variant (NP_001182215.1)	Thr95Arg; T95R
Zygosity	Heterozygous
Inheritance	P1, P2, P3, P4, P5: de novo P6, P7: maternal mosaic
dbSNP153	No entry
gnomAD (v3.1.1)	No entry
COSMIC (v95)	Somatic report, 2 entries (Mutation ID COSV66704481)
***In silico* Pathogenicity Prediction Models**	
CADD (v1.6)	27.3
SIFT	Deleterious (0)
Polyphen-2	Probably damaging (1)
LRT	Deleterious (0)
MutationTaster	Disease causing (1)
PROVEAN	Deleterious (-5.64)
MetaSVM	Deleterious (1.04)
M-Cap	Possibly pathogenic (0.40)
fathmm_MKL-coding	Deleterious (0.96)

Table S3. Immunological features of patient 1

	8y	Reference value	9y	Reference value	10y	Reference value	12y	Reference value
Serum Abs								
IgG (g/L)	**1.96**	6.09-12.85	**1.24**	6.09-12.85	11.8*	6.09-12.85	13.1*	6.98-14.26
IgA (g/L)	**0.02**	0.52-2.16	**UD**	0.52-2.16	**0.36**	0.52-2.16	**UD**	0.92-2.50
IgM (g/L)	**0.08**	0.67-2.01	**0.08**	0.67-2.01	**0.31**	0.67-2.01	0.599	0.56-2.16
IgE (kU/L)	UD	<100		<100	18.14	<100	UD	<100
B lymphocytes								
$CD19^+$ (/µl)					**104**	200-600	**55**	200-600
$CD19^+$ (%)			**5.72**	9.19-19.48	**6.02**	9.19-19.48	**2.78**	9.19-19.48
Naïve B (% of $CD19^+$) (IgD^+CD27^-)					**89.6**	51.84-77.61	**92.66**	51.84-77.61
Memory B (% of $CD19^+$) (IgD^-CD27^+)					**1.19**	8.96-24.09	**0.19**	8.96-24.09
MZB-Like (% of $CD19^+$) (IgD^+CD27^+)					2.12		1	
Transitional B (% of $CD19^+$) ($CD24^{hi}CD38^{hi}$)					**60.5**	2.50-9.07	**37.95**	2.50-9.07
T lymphocytes								
$CD3^+$ (/µl)					1177	1100-2200	1521	1100-2200
$CD3^+$ (%)			**83.52**	57.10-73.43	63.9	57.10-73.43	**76.97**	57.10-73.43
$CD4^+$ (/µl)					**359**	600-1600	**344**	600-1600
$CD4^+$ (% of $CD3^+$)			18.82	24.00-38.72	31.6	24.00-38.72	**17.43**	24.00-38.72
CD4 naïve T (% of $CD4^+$) ($CD45RA^+CD27^+$)					57.4	39.72-69.59	55.45	39.72-69.59
CD4 central memory T (% of $CD4^+$) ($CD45RA^-CD27^+$)					42.4	24.24-52.73	41.48	24.24-52.73
CD4 effector memory T (% of $CD4^+$) ($CD45RA^-CD27^-$)					**0.15**	3.40-11.17	**3**	3.40-11.17
CD4 TEMRA (% of $CD4^+$) ($CD45RA^+CD27^-$)					**0**	0.10-0.29	**0.07**	0.10-0.29
$CD8^+$ (/µl)					767	500-1200	1100	500-1200

$CD8^+$ (% of $CD3^+$)	61.07	21.01-33.94	**63.9**	21.01-33.94	**55.65**	21.01-33.94
CD8 naïve T (% of $CD8^+$) ($CD45RA^+CD27^+$)			48.10	41.41-73.04	59.74	41.41-73.04
CD8 central memory T (% of $CD8^+$) ($CD45RA^-CD27^+$)			21.10	13.21-37.89	13.50	13.21-37.89
CD8 effector memory T (% of $CD8^+$) ($CD45RA^-CD27^-$)			2.23	1.53-15.39	5.89	1.53-15.39
CD8 TEMRA (% of $CD8^+$) ($CD45RA^+CD27^-$)			**28.6**	2.01-21.65	20.87	2.01-21.65
NK cells						
$CD16^+CD56^+$ (/µl)			420	300-600	399	300-600
$CD16^+CD56^+$ (%)	**9.75**	10.01-26.98	24.26	10.01-26.98	20.23	10.01-26.98

UD, undetectable. TEMRA, EM re-expressing CD45RA. Boldface, out of range value. * Ig replacement therapy.

Table S4. Immunological features of patient 2

	11m	12m	Reference value
Serum Abs			
IgG (g/L)	**UD**		4.0-8.3
IgA (g/L)	**UD**		0.08-0.8
IgM (g/L)	**UD**		0.06-1.45
IgE (µg/L)	**UD**		<55
Corynebacterium diphtheriae toxin antibody	**UD** (<0.01 IU/mL)		Detectable
Clostridium tetani toxin antibody	**UD** (<0.01 IU/mL)		Detectable
B lymphocytes			
$CD19^+$ (/µl)	1140	1330	600-2700
$CD19^+$ (%)	0.23	0.23	0.15-0.39
Naïve B (% of $CD19^+$) (IgD^+CD27^-)	**95.2**		71-94
Switched memory B (% of $CD19^+$) ($IgM^-IgD^-CD27^+CD38^{low}$)	**0.2**		1-11
Un-switched memory B (% of $CD19^+$)	3.4		2-10
$CD21^{low}$ (% of $CD19^+$) ($CD38^{lo}CD21^{lo}$)	**0.2**		1-7
Transitional B (% of $CD19^+$) ($CD24^{hi}CD38^{hi}$)	**1.3**		8-27
T lymphocytes			
$CD3^+$ (/µl)	3740	4190	1600-6700
$CD3^+$ (%)	76	73	54-76
$CD4^+$ (/µl)	2820	3230	1000-4600
$CD4^+$ (% of $CD3^+$)	**59**	**56**	31-54
CD4 naïve T (% of $CD4^+$) ($CD45RA^+CCR7^+$)	75.7	73.3	54-80
CD4 central memory T (% of $CD4^+$) ($CD45RA^-CCR7^+$)	19.3	20	10-26

CD4 effector memory T (% of $CD4^+$) ($CD45RA^-CCR7^-$)	4.6	5.5	3-16
CD4 TEMRA (% of $CD4^+$) ($CD45RA^+CCR7^+$)	**0.5**	**1.2**	3-12
$CD8^+$ (/µl)	810	1030	400-2100
$CD8^+$ (% of $CD3^+$)	17	18	12-28
CD8 naïve T (% of $CD8^+$) ($CD45RA^+CCR7^+$)	66.90	**77.90**	34-73
CD8 central memory T (% of $CD8^+$) ($CD45RA^-CCR7^+$)	5.60	9.90	3-15
CD8 effector memory T (% of $CD8^+$) $CD45RA^-CCR7^-$)	26.4	9.3	9-47
CD8 TEMRA (% of $CD8^+$) ($CD45RA^+CCR7^+$)	**1.2**	**2.9**	7-25
NK cells			
$CD16^+CD56^+$ (/µl)	40	200	
$CD16^+CD56^+$ (%)	0.01	0.04	

UD, undetectable. TEMRA, EM re-expressing CD45RA. Boldface, out of range value.

Table S5. Immunological features of patient 3

	2.5m	1y9m	2y5m	3y4m	3y9m	Reference value	8y5m	8y9m	9y8m	Reference value
Serum Abs										
IgG (g/L)	**2**	4.63*	5.83*		10.66*	2.7-11.8	**12.54***	**13.6***	9.1*	5.82-11.54
IgA (g/L)	**<0.03**	**<0.05**	**0.05**		**0.06**	0.1-1.3	**0.17**	**0.15**	**0.36**	0.46-1.57
IgM (g/L)	**<0.02**	**<0.05**	**0.05**		**<0.05**	0.36-1.1	**<0.05**	**<0.05**	**<0.05**	0.54-1.55
IgE (kUI/L)	<2								15.7	(<148)
B lymphocytes										
$CD19^+$ (/µl)		598	527	**286**	**176**	390-1400	**90**	**91**	**109**	219-509
$CD19^+$ (%)		**13**	17	**10**	**8**	14-33	6	7	5	4.8-24.3
Naïve B (% of $CD19^+$) (IgD^+CD27^-)			**98**		**94**	78.5-84.6	**92**		**93**	58.5-84.6
Memory B (% of $CD19^+$) ($CD27^+$)		**1**	**1**		**3**	11.9-18.2	**8**		**7**	9-35
Memory B (% of $CD19^+$) (IgD^-CD27^+)			**0.3**		**1**	4.4-7.9	**1**		**0.6**	4.4-20.5
Switched memory B (% of $CD19^+$) ($IgM^-IgD^-CD27^+$)			**0.2**							
MZB-Like (% of $CD19^+$) (IgD^+CD27^+)		**1**	**1**		**1**	5.9-10.8	7		6	3-21.1
$CD21^{low}$ (% of $CD19^+$) ($CD38^{lo}CD21^{lo}$)		0.3	2		2					
Transitional B (% of $CD19^+$) ($CD24^{hi}CD21^{hi}$)		82	77							
T lymphocytes										
$CD3^+$ (/µl)		3358	2356	2485	1892	1400-3700	1332	**1146**	2005	1200-2600
$CD3^+$ (%)		73	**76**	**87**	**86**	56-75	**89**	**88**	**91**	60-76
$CD4^+$ (/µl)		**2300**	1643	1685	1232	700-2200	793	729	1383	650-1500
$CD4^+$ (% of $CD3^+$)		**50**	**53**	**59**	**56**	28-47	**53**	**56**	**63**	31-47
CD4 naïve T (% of $CD4^+$) ($CD45RA^+$)		86	82	85		73-86	**80.5**	**77**	67	58-70

CD4 naïve RTE (% of CD4$^+$) (CD31$^+$CD45RA$^+$)	**43**	**46**	**47**		57-65	52	47	44	43-55
CD8$^+$ (/µl)	690	**465**	514	**374**	490-1300	389	**299**	490	370-1100
CD8$^+$ (% of CD3$^+$)	**15**	**15**	18	17	16-30	26	23	22	18-35
CD8 naïve T (% of CD8$^+$) (CD45RA$^+$CCR7$^+$)	**87**	**91**	**95**		52-68	**84**	**85**	**94**	52-68
CD8 central memory T (% of CD8$^+$) (CD45RA$^-$ CCR7$^+$)	**7**	**5**	3		3-4	**1.5**	**5**	**1**	3-4
CD8 effector memory T (% of CD8$^+$) (CD45RA$^-$ CCR7$^-$)	**3**	**2**	**1**		11-20	**9**	**8**	**2**	11-20
CD8 TEMRA (% of CD8$^+$) (CD45RA$^+$CCR7$^-$)	**3**	**2**	**1**		16-28	**5.5**	**2**	**2**	16-28
NK cells									
CD16$^+$CD56$^+$ (/µl)	598	186	**86**	**66**	130-720	**90**	**52**	**88**	100-480
CD16$^+$CD56$^+$ (%)	13	6	**3**	**3**	4-17	6	4	4	4-17

RTE, Recent thymus emigrant. TEMRA, EM re-expressing CD45RA. Boldface, out of range value. * Ig replacement therapy.

Table S6. Immunological features of patient 4

	1y5m	4y	Reference value
Serum Abs			
IgG (mg/dl)	**169**		685 (424-1051)
IgA (mg/dl)	**9**		47 (14-123)
IgM (mg/dl)	**6**		95 (48-168)
IgE (mg/dl)			
B lymphocytes			
$CD19^+$ (%)		**7.83**	11.12 (8.17-15.3)
Naïve B (% of $CD19^+$) (IgD^+CD27^-)		**92.3**	61.3 (44.9-77.5)
Switched memory B (% of $CD19^+$) (IgD^-CD27^+)		**1.22**	21.24 (10.4-30.4)
Un-switched memory B (% of $CD19^+$) (IgD^+CD27^+)		**4.27**	10.61 (3.62-22.2)
Transitional B (% of $CD19^+$) ($CD20^+CD38^{hi}$)		**9.46**	3.01 (1.73-4.08)
T lymphocytes			
$CD3^+$ (%)		53.6	60.26 (41.4-71.7)
$CD4^+$ (% of $CD3^+$)		**44.6**	59.6 (45.4-72.8)
CD4 naïve T (% of $CD4^+$) ($CD45RA^+CCR7^+$)		**96.3**	81.16 (72.5-91.5)
CD4 central memory T (% of $CD4^+$) ($CD45RA^-CCR7^+$)		**2.29**	14.8 (4.91-22.5)
CD4 effector memory T (% of $CD4^+$) ($CD45RA^-CCR7^-$)		**0.3**	1.226 (0.6-1.94)
CD4 TEMRA (% of $CD4^+$) ($CD45RA^+CCR7^+$)		**1.07**	2.824 (1.52-4.07)
$CD8^+$ (% of $CD3^+$)		**44.4**	31.68 (19.1-40.1)
CD8 naïve T (% of $CD8^+$) ($CD45RA^+CCR7^+$)		**94.4**	52.58 (36.3-69.9)

CD8 central memory T (% of $CD8^+$) ($CD45RA^-CCR7^+$)	**0.14**	4.38 (0.4-8.08)
CD8 effector memory T (% of $CD8^+$) $CD45RA^-CCR7^-$)	**0.29**	5.452 (1.87-7.48)
CD8 TEMRA (% of $CD8^+$) ($CD45RA^+CCR7^+$)	**5.14**	37.56 (27.8-48.9)
NK cells		
$CD16^+CD56^+$ (%)	19.2	10.552 (3.45-22.7)

TEMRA, EM re-expressing CD45RA. Boldface, out of range value.

Table S7. Immunological features of patient 5

	7.5m	Reference value	8m	Reference value	17y6m	Reference value
Serum Abs						
IgG (g/L)	**<0.35**	2.42-11.08			*	4.79-14.33
IgA (g/L)	0.07	0.02-1.26			**0.21**	0.6-2.3.49
IgM (g/L)	**<0.05**	0.21-2.15			**<0.19**	0.26-2.32
IgE (IU/ml)	7	2-34			ND	
B lymphocytes						
$CD19^+$ (/µl)			1590	130-6300	231	64-820
$(IgM^{++}CD38^{++})$ (/µl)			**7**	15-1700	4	1-100
(IgM^+IgD^+) (/µl)			1557	110-5300	177	28-550
(IgD^+CD27^-) (/µl)			1573	461-1939	173	126-1939
Switched memory B (/µl) $(CD20^+IgM^-CD27^+)$			**0**	1.5-82	**0.6**	4.5-130
T lymphocytes						
$CD3^+$ (/µl)			2981	400-11500	1201	780-3000
$CD3^+DR^+$ (/µl)			**497**	135-232	89	73-162
$CD3^+TCR\alpha/\beta^+$ (/µl)			2881	1200-11500	1140	600-3300
$CD3^+TCR\gamma/\delta^+$ (/µl)			99	38-890	37	25-200
$CD3^+CD4^+$ (/µl)			1689	1000-7200	801	500-2000
CD4 naïve T (/µl) $(CD4^+CD45RA^+CCR7^+)$			1165	800-7600	432	199-2300
CD4 central memory T (/µl) $(CD4^+CD45RO^+CCR7^+)$			219	83-1300	**57**	180-1100
CD4 effector memory T (/µl) $(CD4^+CD45RO^+CCR7^-)$			**253**	1-72	184	13-220
CD4 TEMRA (/µl) $(CD4^+CD45RA^+CCR7^-)$			51	1-400	27	1-68
$CD8^+$ (/µl)			1143	200-5400	339	200-1200
CD8 naïve T (/µl) $(CD8^+CD45RA^+CCR7^+)$			503	150-3200	247	16-1000

CD8 central memory T (/µl) ($CD8^+CD45RO^+CCR7^+$)	45	2-150	**4**	5-120
CD8 effector memory T (/µl) ($CD8^+CD45RO^+CCR7^-$)	388	8-1400	**34**	40-640
CD8 TEMRA (/µl) ($CD8^+CD45RA^+CCR7^-$)	205	17-280	44	25-280
NK cells (/µl) ($CD3^-CD16^+CD56^+$)	248	68-3900	**31**	100-1200
NKT cells (/µl) ($CD3^+CD16^+CD56^+$)	0	4.4-510	**12**	23-410

TEMRA, EM re-expressing CD45RA. Boldface, out of range value. * Ig replacement therapy.

Table S8. Immunological features of patient 6

	8y	Reference value
Serum Abs		
IgG (g/L)	**1.54**	5.14-16.72
IgA (g/L)	**<0.05**	0.52-2.26
IgM (g/L)	**<0.05**	0.16-1.88
IgE (IU/ml)	8	2-403
B lymphocytes		
$CD19^+$ (/µl)	201	100-800
Switched memory B (/µl) ($CD20^+IgM^-CD27^+$)	**4**	7-51
T lymphocytes		
$CD3^+$ (/µl)	2936	770-4000
$CD3^+CD4^+$ (/µl)	1805	400-2500
CD4 naïve T (/µl) ($CD4^+CD45RA^+CD62L^+$)	1451	200-2500
CD4 central memory T (/µl) ($CD4^+CD45RA^-CD62L^+$)	131	4-510
CD4 effector memory T (/µl) ($CD4^+CD45RA^-CD62L^-$)	138	3-170
CD4 TEMRA (/µl) ($CD4^+CD45RA^+CD62L^-$)	**86**	1-25
$CD8^+$ (/µl)	921	200-1700
CD8 naïve T (/µl) ($CD8^+CD45RA^+CD62L^+$)	339	42-1300
CD8 central memory T (/µl) ($CD8^+CD45RA^-CD62L^+$)	**48**	6-43
CD8 effector memory T (/µl) ($CD8^+CD45RA^-CD62L^-$)	187	45-410
CD8 TEMRA (/µl) ($CD8^+CD45RA^+CD62L^-$)	**347**	57-340
NK cells (/µl) ($CD3^-CD16^+$ and/or $CD56^+$)	**593**	70-590

NKT cells (/μl) ($CD3^+CD16^+$ and/or $CD56^+$)	336	12-340

UD, undetectable. TEMRA, EM re-expressing CD45RA. Boldface, out of range value.

Table S9. Immunological features of patient 7

	3y	Reference value
Serum Abs		
IgG (g/L)	**0.24**	4.85-11.60
IgA (g/L)	**<0.05**	0.14-2.12
IgM (g/L)	**<0.05**	0.26-1.55
IgE (IU/ml)	2	2-307
B lymphocytes		
$CD19^+$ (/µl)	222	180-1300
Switched memory B (/µl) ($CD20^+IgM^-CD27^+$)	3	2.2-25
T lymphocytes		
$CD3^+$ (/µl)	2538	850-4300
$CD3^+CD4^+$ (/µl)	991	500-2700
CD4 naïve T (/µl) ($CD4^+CD45RA^+CD62L^+$)	664	300-2300
CD4 central memory T (/µl) ($CD4^+CD45RA^-CD62L^+$)	**76**	160-660
CD4 effector memory T (/µl) ($CD4^+CD45RA^-CD62L^-$)	**155**	4-89
CD4 TEMRA (/µl) ($CD4^+CD45RA^+CD62L^-$)	**116**	1-16
$CD8^+$ (/µl)	1404	200-1800
CD8 naïve T (/µl) ($CD8^+CD45RA^+CD62L^+$)	261	53-1100
CD8 central memory T (/µl) ($CD8^+CD45RA^-CD62L^+$)	21	4-64
CD8 effector memory T (/µl) ($CD8^+CD45RA^-CD62L^-$)	**745**	24-590
CD8 TEMRA (/µl) ($CD8^+CD45RA^+CD62L^-$)	377	25-530
NK cells (/µl) ($CD3^-CD16^+$ and/or $CD56^+$)	289	61-510

NKT cells (/µl) ($CD3^+CD16^+$ and/or $CD56^+$)	**322**	15-250

TEMRA, EM re-expressing CD45RA. Boldface, out of range value.

Table S10. Th cell subset distributions in T95R patients and HC

	HC	HC	P1	HC	HC	HC	HC	P2	P2	P2	P2
Age	12y	12y	12y	3y	3y	3y	3y	2y	2y	2y	2y
Tfh (% of CD4+)	2.82	3.38	3.61	4.04	2.2	6.86	1.81	1.8	1.9	2.77	1.1
Th1 (% of CD4+)	0.52	0.22	2.7	2	4	4	4	1	2	1.4	2
Th2 (% of CD4+)	0.51	0.39	0.99								
Th17 (% of CD4+)	1.75	1.62	1.32								
Treg (% of CD4+)		4.99	7.29								
Tfr (% of CD4+)											

	HC	HC	HC	HC	HC	P3	P4	HC	HC	HC	P5	P5	P5	P6	P7
Age	NA	39y	31y	18y8m	2y	9y	4y	4y11m	6y10m	9y8m	9m	1y2m	1y9m	8y	3y
Tfh (% of CD4+)	22.3	16.3	18.3	1.39	9.8	3.6	2.46	5.39	4.27	0.66	1.70	1.71	1.28	2.27	3.03
Th1 (% of CD4+)	8.44	6.97	9.25	6.48	7.89	17.3	10.5	8.0	5.1	7.2	10.3	9.9	13.0	3.66	9.65
Th2 (% of CD4+)	3.64	5.4	4.12	9.39	2.96	1.6	3.73	1.5	1.3	1.7	0.6	2.2	1.3	0.6	1.08
Th17 (% of CD4+)	5.28	6	5.27	4.94	5.1	0.97	2.76	0.8	0.5	2.7	0	0	0.1	0.34	0.3
Treg (% of CD4+)	4.69	3.12	5.36	1.6	3.02	2.06	1.51	6.36	6.18	8.87	5.25	7.42	9.64	2.49	1.51
Tfr (% of CD4+)								0.19	0.49	0.22	0.23	0.55	0.24		

Th-cell subsets of P1-P4, P6-P7 and their corresponding HC were determined by surface chemokine receptor staining, P5 and corresponding HC were determined by intracellular cytokine staining. P1, P2, and P5-P7 were analyzed by flow cytometry and P3 and P4 by CyTOF.

NA, not available.

Table S11. Monocyte subset distributions in T95R patients and HC

	HC	HC	HC	P2	P2	P2	HC	HC	HC	HC	HC	P3	P4
Age	3y	3y	3y	2y	2y	2y	NA	39y	31y	18y8m	2y	9y	4y
Classical monocytes (% of monocytes)	67.4	70.7	65	32.7	16.4	30.6	76.8	57.9	70.1	74.6	58.9	65.5	78.4
Intermediate monocytes (% of monocytes)							7.27	30.2	18.4	8.65	16.4	7.6	12
Non-classical monocytes (% of monocytes)	5.6	3	4	4	5	6	2.52	2.22	4.98	3.43	4.52	1.64	0.38

	HC	HC	HC	HC	HC	P5
Age	12y	18y	18y	18y	20y	17y6m
Classical monocytes (% of monocytes)	78.7	72.6	82.5	81.3	83.4	94.3
Intermediate monocytes (% of monocytes)	10.4	9.4	4.0	4.8	6.1	1.9
Non-classical monocytes (% of monocytes)	6.1	10.5	8.5	10.1	4.1	0.3

P2 and P5 were analyzed by flow cytometry and P3 and P4 by CyTOF. NA, not available.

Table S12. List of antibodies and stimuli used in the present study

REAGENT or RESOURCE	COMPANY	IDENTIFIER
Antibodies		
Percp/Cy5.5 Anti-human CD38 Clone HIT2	BD Biosciences	561106
PE/Cy7 Anti-human CD38 Clone HIT2	Biolegend	303515
APC Anti-human CD19 Clone HIB19	Biolegend	302212
PE/Cy7 Anti-human IgG Clone G18-145	BD Biosciences	561298
PE Anti-human IgD Clone IA6-2	BD Biosciences	555779
APC/Cy7 Anti-human CD27 Clone O323	Biolegend	356424
FITC Anti-human CD20 Clone 2H7	Biolegend	302304
PE Anti-human HLADR Clone G46-6	BD Biosciences	556644
PE/Cy7 Anti-human CD69 Clone FN50	BD Biosciences	557745
FITC Anti-human CD80 Clone L307.4	BD Biosciences	557226
APC Anti-human CD86 Clone IT2.2	Biolegend	305412
PE Anti-human CD25 Clone 24212	R&D	FAB1020P
PE Anti-human CD24 Clone ML5	BD Biosciences	555428
APC-H7 Anti-human CD3 Clone SK7	BD Biosciences	560176
Percp/Cy5.5 Anti-human CD3 Clone SK7	BD Biosciences	340949
FITC Anti-human CD4 Clone RPA-T4	BD Biosciences	555346
BV510 Anti-human CD8 Clone RPA-T8	BD Biosciences	563256
APC Anti-human CD8a Clone RPA-T8	Biolegend	301049
PE/Cy7 Anti-human CD45RA Clone HI30	BD Biosciences	557748
APC Anti-human CD27 Clone M-T271	BD Biosciences	558664
PE-Cy7 Anti-human CD138	Biolegend	356513
Percp/Cy5.5 Anti-human CD38	Biolegend	356614
PE Anti-human CD27	Biolegend	356406

Alexa Fluor 647 Anti-human BCMA	Biolegend	357506
APC/Cy7 Anti-human CD19	Biolegend	302218
Alexa Fluor 647 Goat F(ab′)2 anti-human IgA	Jackson ImmunoResearch	109-496-011
PerCP Anti-human CD3	BD Biosciences	345766
PE Anti-human CD4	Biolegend	300508
APC Anti-human CD45RA	BD Biosciences	550855
PerCP Anti-human CXCR5	Thermo Fisher Scientific	46-9185-42
APC Anti-human PD-1	Thermo Fisher Scientific	17-9969-42
FITC Anti-human CD127	Biolegend	351312
Pacific Blue Anti-human CCR7	Biolegend	353210
APC Anti-human CD3	Beckman Coulter	IM2467
APC Alexa Fluor 750 Anti-human CD4	Beckman Coulter	A94682
APC-Alexa Fluor 700 Anti-human CD8	Beckman Coulter	B49181
Pacific Blue Anti-human CD8	Beckman Coulter	B49182
FITC Anti-human CD14	Dako/ Agilent	F0844
ECD Anti-human CD16	Beckman Coulter	B49216
APC-Alexa Fluor 700 Anti-human CD19	Beckman Coulter	B49212
PerCP Anti-human CD20	Biolegend	302324
Horizon V450 Anti-human CD27	Becton Dickinson	560448
PE-Cy7 Anti-human CD38	Becton Dickinson	335825
Krome-Orange Anti-human CD45	Beckman Coulter	B36294
PE-Cy7 Anti-human CD45RA	Beckman Coulter	B10821
FITC Anti-human CD45RO	Dako/ Agilent	F0800
ECD Anti-human CD56	Beckman Coulter	B49214
FITC Anti-human CD107a	Becton Dickinson	555800
PE Anti-human HLA-DR	Beckman Coulter	IM1639

FITC Anti-human IgD	Dako/ Agilent	F0189
APC Anti-human IgM	Biolegend	314510
PE-Cy7 Anti-human IL-4	Biolegend	500824
eFlour 450 Anti-human IL-17A	eBioscience (Thermo Fisher Scientific)	48-7179-42
FITC Anti-human IFN-g	Biolegend	502506
Anti-Human FOXP3 Staining Set PE	Thermo Fisher Scientific	72-5774-40
Human regulatory T-cell staining Kit #3	eBioscience (Thermo Fisher Scientific)	88-8995-41
Beta Mark TCR Vbeta Repertoire Kit	Beckman Coulter	IM3497
TruStain FcX	Biolegend	101320
PerCP/cy5.5 Anti-mouse B220	Biolegend	103235
PE Anti-mouse B220	Biolegend	102504
APC Anti-Mouse CD138 Clone 281-2	BD Biosciences	558626
PE-Cy7 Anti-Mouse CD45R/B220 Clone RA3-6B2	BD Biosciences	553093
APC Anti-Mouse CD38 Clone 90	Biolegend	102712
FITC Anti-Mouse FAS Clone Jo2	BD Biosciences	554257
APC-Cy7 Anti-Mouse CD3 Clone 17A2	Biolegend	100221
FITC Anti-Mouse CD4 Clone H129.19	BD Biosciences	553650
PE-Cy7 Anti-Mouse CD8 Clone RPA-T8	Biolegend	301012
PE Anti-Mouse CD44 Clone 2M7	BD Biosciences	553134
APC Anti-Mouse CD62L Clone MEL-14	BD Biosciences	553152
PE Anti-mouse IgA Clone mA-6E1	eBioscience	12-4204-82
BUV737 Anti-Human CD19	BD Biosciences	741829
BUV615 Anti-Human IgD	BD Biosciences	751595
BV605 Anti-Human CD27	BD Biosciences	751673
APC Anti-Human CD38	BD Biosciences	560980

BV421 Anti-Human HLA-DR	BioLegend	307635
BV750 Anti-Human CD14	BD Biosciences	746920
BUV496 Anti-Human CD16	BD Biosciences	612945
BB660 Anti-Human CD11c	BD Biosciences	Custom
BUV661 Anti-Human CD56	BD Biosciences	750478
Fixable Viability Dye eFluor 780	BD Biosciences	565388
BV510 Anti-Human CD3	BD Biosciences	563109
BUV395 Anti-Human CD4	BD Biosciences	563550
BV570 Anti-Human CD8	BioLegend	301037
BV510 Anti-Human IgM	BD Biosciences	563113
BUV395 Anti-Human TNFα	BD Biosciences	563996
BV711 Anti-Human IFNγ	BD Biosciences	564793
PerCp-Cy5.5 Anti-Human CD45RA	BD Biosciences	563429
BV421 Anti-Human CXCR5	BD Biosciences	562747
BV786 Anti-Human ICOS	BD Biosciences	741017
PE Anti-Human IRF4	BD Biosciences	566649
BV650 Anti-Human IL2	BD Biosciences	564166
Alexa Fluor® 647 anti-IRF4 Antibody	BioLegend	646408
V450 Mouse Anti-Human IL-4	BD	561595
PE anti-human IFN-γ Antibody	Biolegend	506507
Alexa Fluor® 647 Mouse anti-Human IL-17A	BD	560491
APC Mouse Anti-Human CD4; RPA-T4	BD	555349
FITC Mouse Anti-Human CD8; RPA-T8	BD	555366
APC/Cyanine 7 anti-human CD19; HIP19	Biolegend	302218
7-amino-actinomycin D	Invitrogen	00-6993-50
Anti-IRF4 antibody (E8H3S)	Cell Signaling Technology	62834

Anti-IRF4 antibody (P173)	Cell Signaling Technology	4948
Anti-IRF4 antibody	Cell Signaling Technology	4964S
Anti-β-actin antibody	Cell Signaling Technology	3700S
Anti-GAPDH antibody	Cell Signaling Technology	2118S
Anti-Lamin B antibody	Santa Cruz Biotechnology	Sc-6216
IRDye 680RD Goat anti-Mouse	LI-COR Bioscience	925-68070
IRDye 800CW Goat anti-Rabbit	LI-COR Bioscience	926-32211
Anti-PARP antibody (46D11)	Cell Signaling Technology	9532
Anti-DYKDDDDK Tag antibody	Cell Signaling Technology	2368
Anti-Flag antibody (M2)	Sigma	F1804
Anti-HA antibody (C29F4)	Cell Signaling Technology	3724
Normal rabbit IgG antibody	Cell Signaling Technology	2729
Anti-AID antibody (L7E7)	Cell Signaling Technology	4975
Anti-β-actin antibody	Sigma	A1978
Anti-Vinculin	Santa Cruz	sc-73614
Rabbit monoclonal to human GAPDH	Abcam	ab181602
Rabbit monoclonal to MUM1 (IRF4)	Abcam	ab124691
Goat anti-rabbit IgG (H+L)-HRP conjugate	BioRad	170-6515
Mouse monoclonal to human CD20 (L26)	Dako	M0755
Mouse monoclonal to human CD38 (SPC32)	Leica Biosystems	NCL-CD38-290
Mouse monoclonal to human CD138 (MI15)	Dako	7228
Mouse monoclonal to human IRF4/ MUM1 (MUM1p)	Dako	7259
Goat Anti-Human Ig-UNLB	SouthernBiotech	2010-01
Goat Anti-Human IgM-BIOT	SouthernBiotech	2020-08
Goat Anti-Human IgG-BIOT	SouthernBiotech	2040-08

Human IgM Lambda-UNLB	SouthernBiotech	0158L-01
Human IgG-UNLB	SouthernBiotech	0150-01
HRP Avidin	Biolegend	405103
Neutralite Avidin-AP	SouthernBiotech	7200-04
Purified NA/LE Mouse Anti-Human CD3	BD Biosciences	555336
Purified NA/LE Mouse Anti-Human CD28	BD Biosciences	555725
F(ab')2 anti-human IgM	SouthernBiotech	2022-14
Mouse IgA-UNLB	SouthernBiotech	0106-01
Mouse IgM-UNLB	SouthernBiotech	0101-01
Mouse IgG-UNLB	SouthernBiotech	0107-01
Mouse IgG_1-UNLB	SouthernBiotech	0102-01
Mouse IgG_{2b}-UNLB	SouthernBiotech	0104-01
Mouse IgG_{2c}-UNLB	SouthernBiotech	0122-01
Mouse IgG_3-UNLB	SouthernBiotech	0105-01
Anti-mouse Ig-UNLB	SouthernBiotech	1010-01
Goat Anti-Mouse IgA-HRP	SouthernBiotech	1040-05
Goat Anti-Mouse IgM-HRP	SouthernBiotech	1020-05
Goat Anti-Mouse IgG-HRP	SouthernBiotech	1030-05
Goat Anti-Mouse IgG_1, Human ads-HRP	SouthernBiotech	1070-05
Goat Anti-Mouse IgG_{2b}, Human ads-HRP	SouthernBiotech	1090-05
Goat Anti-Mouse IgG_{2c}-HRP	SouthernBiotech	1078-05
Goat Anti-Mouse IgG_3, Human ads-HRP	SouthernBiotech	1100-05
Goat Anti-Mouse IgG_1-AP	SouthernBiotech	1071-04
Peroxidase-AffiniPure Goat Anti-Rabbit IgG (H+L)	Jackson Immunoresearch	111-035-003
Chemicals, Peptides, and Recombinant Proteins		

Recombinant Human IL2 GMP Protein, Carrier free	R&D	202-GMP-010
Recombinant Human IL4	BD Biosciences	554605
CpG ODN 2006	Sangon Biotech	NA
CpG ODN 2006	ENZO	ALX 746-006-C100
Recombinant Mouse IL4	R&D	404-ML-010
Recombinant Mouse IL21	Biolegend	574502
MEGACD40L® Protein	ENZO	ALX-522-110-C010
IL21	Peprotech	200-21
IL2	Peprotech	200-02
IL4	Peprotech	200-04
Anti-Human IgM	Jackson Immunoresearch	109-006-129
Anti-CD3	eBioscience	16-0037-85
Anti-CD28	eBioscience	16-0289-85
Phorbol 12-myristate 13-acetate (PMA)	Sigma	P8139
Brefeldin A	Sigma	B7651
Ionomycin	Thermo Fisher Scientific	I24222
LPS (E. coli O111:B4)	Sigma-Aldrich	L4391/L2880
CD40L		NA

Table S13. List of oligonucleotides used in the current study

Primers for mutation confirmation		
IRF4-for P1	Forward	AGGACCTATGCGCCATTCTT
	Reverse	GGCAGGCAGGCAATACAAAA
humIRF4_gDNA_Ex1	Forward	AAGTCCCTCTCCCCAGTC
	Reverse	ACCTCTGGTTCGCGCTC
humIRF4_gDNA_Ex2	Forward	CCTCGTGGTCACTGGCGC
	Reverse	CCTCCTCCTGCGGCTCCG
humIRF4_gDNA_Ex3	Forward	TGGGCAGCAGAGCAGGAC
	Reverse	TAAGGTGCCTCAAGGATCTG
humIRF4_gDNA_Ex4	Forward	AACACCGTGTTATGCATTCT
	Reverse	CTGGGCTGCTGTGGTTCTC
humIRF4_gDNA_Ex5	Forward	GATGATAAAATGCTTCGGCTG
	Reverse	GGAAAGAGCTTTGGTGCTG
humIRF4_gDNA_Ex6	Forward	TCCCAGGCTTCACACACAC
	Reverse	CTAAAGTCCCATCGAATCTGC
humIRF4_gDNA_Ex7	Forward	AGGTGCTTGGCTCTGTGGAG
	Reverse	CAGGAGGAAGACCTCAGCC
humIRF4_gDNA_Ex8	Forward	TTCACATCAAGAGCCCCAC
	Reverse	TTCTAAATGAAACTCTGGCC
humIRF4_gDNA_Ex9	Forward	TTCTAGGATGTAACTTTGGGC
	Reverse	CCTGGGAGACAGAGCAAGAC
Primers for RT-PCR		
ACTBcDNA	Forward	CCAACACAGTGCTGTCTG
	Reverse	CAACTAAGTCATAGTCCACC
IRF4cDNA	Forward	GGGAAGCTCCGCCAGTGG
	Reverse	GGGTAAGGCGTTGTCATGGTG
gRNA for IRF4 CRISPR/CAS9		

#1 gRNA	Oligo1	CACCGCAAGCAGGACTACAACCGCG
	Oligo2	AAACCGCGGTTGTAGTCCTGCTTGC
#2 gRNA	Oligo1	CACCGCCACCTGGAAGACGCGCCTG
	Oligo2	AAACCAGGCGCGTCTTCCAGGTGGC
Primers for vector generation		
IRF4	Forward	CGGGGATCCCCACCATGAACCTGGAGGGCGGC
	Reverse	CGGGGATCCTCATTCTTGAATAGAGGAATGGC
	dAR-Reverse	CCGGAATTCTCATTGTTGAGCAAAATAATATAG TTGT
IRF4-FLAG	Forward	CCGGAATTCCCACCATGAACCTGGAGGGCGGC
	Reverse	CGGGGATCCTCA CTTATCGTCGTCATCCTTGTAATC TTCTTGAATAGAGGAATGGC
IRF4-HA	Forward	CCGGAATTCCCACCATGAACCTGGAGGGCGGC
	Reverse	CGGGGATCCTCA CTTATCGTCGTCATCCTTGTAATC TTCTTGAATAGAGGAATGGC
pGL3-Basic	PromF	TAGTACTAACATACGCTCTCCATC
Primer for QPCR		
hIRF4	Forward	TCCGAGAAGGCATCGACAAG
	Reverse	AGGCGTTGTCATGGTGTAGG
hPRDM1	Forward	GCAACTGGATGCGCTATGTG
	Reverse	TCTCAGTGCTCGGTTGCTTT
hXBP1	Forward	TTCCGGAGCTGGGTATCTCA
	Reverse	GAAAGGGAACCCCCGTATCC
hACTIN	Forward	AGCGAGCATCCCCCAAAGTT
	Reverse	GGGCACGAAGGCTCATCATT
mAID	Forward	TGGACAGCCTTCTGATGA
	Reverse	GTCTGAGATGTAGCGTAGG

mGAPDH	Forward	TGTGAACGGATTTGGCCGTA
	Reverse	ACTGTGCCGTTGAATTTGCC
gBlock DNA Fragment for luciferase assay		
CXCL13NewA wt / B wt		AGAATATACGTTCTTATCTGCAATCTTCTCATCTAAAATTGACCACACGCTCAGTCATAAAGCAAGTCTTAAAATCAAAAATATCAATATTAAGCATCTTCTCACACCACAGTGAAATAAAAATAGAAATTAATATCAAAAGGAACTCTCAAAATGACACAAATATATGGAAACAAAACAACTTCTGAATTACTTTTGGGT
Probes for EMSA		
ISRE 3xGAAA	Forward	5'IRDye700-gatcGGGAAAGGGAAACCGAAACTGAA
	Reverse	TTCAGTTTCGGTTTCCCTTTCCCgatc
ISRE 3xGATA	Forward	5'IRDye700-gatcGGGATAGGGATACCGATACTGAA
	Reverse	TTCAGTATCGGTATCCCTATCCCgatc
EMSA_ISRE_3xGAAA	Forward	AGC TGG GAA AGG GAA ACC GAA ACT G
	Reverse	AGC TCA GTT TCG GTT TCC CTT TCC C
EMSA_EICE	Forward	AGCTATAAAAGGAAGTGAAACCAAG
	Reverse	AGCTCTTGGTTTCACTTCCTTTTAT
EMSA_EICE_mut+1A	Forward	AGCTATAAAAGGAAGTGAAAACAAG
	Reverse	AGCTCTTGTTTTCACTTCCTTTTAT
EMSA_EICE_IgK3_enhancer	Forward	ACGTAAGACCCTTTGAGGAACTGAAAACAGAACC
	Reverse	AGCTGGTTCTGTTTTCAGTTCCTCAAAGGGTCTT
EMSA_EICE_Igk3_enhancer_mut+1C	Forward	ACGTAAGACCCTTTGAGGAACTGAAACCAGAACC
	Reverse	AGCTGGTTCTGGTTTCAGTTCCTCAAAGGGTCTT
EMSA_CXCL13_A	Forward	AGCTACGCTCAGTCATAAAGCAAGTCT
	Reverse	AGCTAGACTTGCTTTATGACTGAGCGT

EMSA_CXCL13_B	Forward	AGCTAGGAACTCTCAAAATGACACAAA
	Reverse	AGCTTTTGTGTCATTTTGAGAGTTCCT
EMSA_CXCL13_C	Forward	AGCTACTGTGTCATATTGACTCTTAAA
	Reverse	AGCTTTTAAGAGTCAATATGACACAGT