Supplemental Information - IDENTIFICATION OF GERMLINE MONOALLELIC MUTATIONS IN *IKZF2* IN PATIENTS WITH IMMUNE DYSREGULATION

Supplemental clinical reports

**Family A (R291X)**
The index patient, female, developed at 10.9 years non-specific skin-lesions, photosensitivity, polyarthritis involving small joints and mild immune thrombocytopenia. Protein electrophoresis showed polyclonal hypergammaglobulinemia at 16 g/L. Anti-GPIb, anti-GPIIb IIIa, antinuclear and anti-DNA antibodies were present. The child responded well to hydroxychloroquine, non-steroidal anti-inflammatory drugs and sun protection. At the latest follow-up, at the age of 16, the child was in clinical and biologic remission on low-dose hydroxychloroquine.

**Family B (Y359C)**
Patient B1 is a 24-year-old female who developed at 15 years of age a chronic ITP associated with lupus biology including antinuclear and anti-DNA autoantibodies, and low complement. She never presented organ involvement of systemic lupus. Flares of ITP were treated by corticosteroids and she is currently well controlled under hydroxychloroquine. She never received monoclonal immunosuppressants other than steroids. Her mother (Patient B2) developed at adult age a systemic lupus with ITP and glomerulonephritis. She was successfully treated with steroids and hydroxychloroquine without other immunosuppressive drugs.

**Family C (V347M)**
Patient C1 is a 27-year-old male from a mixed ethnic background (mother is Afro-Caribbean, father is North European). He experienced isolated seizures at the age of five and had recurrent otitis media during adolescence. At 25 years of age, he developed nephrotic syndrome (focal segmental glomerulonephritis specifically) and was given prednisone and cyclosporine. A few months later, he presented with pericardial effusion that was drained, and hypogammaglobinemia that was followed by a dose of IVIG. At the age of 26 years, he presented with persistent fevers leading to the detection of chronic active EBV (CAEBV) and full-blown hemophagocytic lymphohistiocytosis (HLH) fulfilling 7 out of 8 HLH criteria:
fever, splenomegaly, hyperferritinemia, hypofibrinogenemia/hypertriglyceridemia, hemophagocytosis in bone marrow, elevated sIL2R and cytopenias (supplemental Table S1). NK-cell cytotoxicity was tested by chromium release at the time but showed normal function. Lymphopenia, particularly that of B cells and CD4\(^+\) T cells was apparent in both bone marrow and peripheral blood (supplemental Table 1). Bone marrow analysis excluded any indications of leukemia or lymphoma and showed a hypocellular marrow with scattered polytypic plasma cells (<5%) and many CD68\(^+\)CD163\(^+\) histiocytes containing lucencies and nucleated cells, but minimal fibrosis. Immunophenotyping of peripheral lymphocytes show very low naïve CD4\(^+\) and CD8\(^+\) T cells with particular expansion of effector memory CD8\(^+\) T cells (supplemental Table 1). Additionally, EBV infection of T cells and NK cells was particularly noted (real-time (RT) PCR results of EBV in peripheral B cells: detected but below 5700 IU/mL; T cells: 299145 IU/mL; NK cells: 363060 IU/mL). The presence of a range of autoantibodies were also tested and shown to be negative. Moreover, the patient had recurrent stomatitis due to HSV infection, serositis, joint inflammation (with limited range of motion of the bilateral elbows), hepatomegaly, edema, diarrhea and seizures. Several immunomodulatory drugs were given at the time (dexamethasone, etoposide, rituximab, anakinra, emapalumab and ruxolitinib) together with subcutaneous immunoglobulin. As the patient did not significantly improve within 8 months of therapy, hematopoietic stem cell transplantation (HSCT) was performed at the age of 27 years. He responded well and showed 100 % donor engraftment with no EBV detected in the blood by RT-PCR, starting from 3.5 months post-HSCT. His kidneys, though still poor, showed improved function (from an estimated glomerular filtration rate by cystatin C of 10 ml/min before HSCT, to 46-63 ml/min post-HSCT). The clinical course and treatment of this patient were recently published\(^1\).

**Family D (R106W)**

Patient D1 is a 29-year-old male born to healthy non-consanguineous parents of Caucasian origin. His sister received a diagnosis of urothelial bladder cancer at 26 years of age and his brother had a past history of atopy. Both parents were healthy. He presented at the age of 23 years with acute onset of fever and myalgia with enlarged lymph with confirmed infectious mononucleosis and HLH fulfilling 5/8 HLH criteria (fever, splenomegaly, hyperferritinemia, increased TG, cytopenia, sCD25 and NK activity not available).
The patient was affected by recurrent maxillary sinusitis starting at 20 years of age requiring repeated antibiotic prescriptions. In the follow up of severe infectious mononucleosis, the patient displayed persistent low grade fever, migratory papular rash, painful inflammatory joint disease and an increased leukocyte counts (WBC 14 400, PMN 8500, Leukocytes 4770, Eosinophils 0, monocytes 720) leading to the diagnosis of Adult onset Still’s disease. After a prolonged 6-month treatment with corticosteroids (prednisone, max. dose 60 mg/d) the patient is currently asymptomatic and does not require pharmacotherapy.

**Family E (N220S)**
Adopted child. At 2 years of age, patient had classical immune thrombocytopenia and was given corticosteroids and IVIG. At 7 years of age the patient developed autoimmune hemolytic anemia (AIHA) and the diagnosis of Evan’s syndrome was made.
Supplemental Material and Methods

Study approval

Samples from the patients and 13 age-marched healthy donors were obtained following an informed written consent. Samples from healthy donors were shipped together with the samples. Approval was given by the Institutional Review Boards of the Medical University of Vienna, Tehran University of Medical Sciences, National Institutes of Health, Institute Imagine and Children’s Hospital of Philadelphia.

Data and material availability

Raw and processed sequencing data for scRNA-seq have been deposited at the European Genome-phenome Archive (EGA), which is hosted by the EBI and the CRG, under accession number no. EGAS00001005675 (for controls 1-4; Control: 1: EGAN00003417282; Control 2: EGAN00003417168; Control 3: EGAN00003417281; Control 4: EGAN00003417169), and accession number no. EGAS00001005874 (for patient R291X). R codes used for the analysis of scRNA-seq are available on GitHub: https://github.com/cancerbits/shahin2021_ikzf2_het. The dataset from the BioID LC-MS analysis is available in the MassIVE repository under the identifier MSV000088165.

Genetics methods

For patients Helios\textsubscript{N220S}, Helios\textsuperscript{R291X} and Helios\textsuperscript{Y359C}-daughter whole exome sequencing (WES) was performed. Samples of DNA were prepared from the patients’ whole peripheral blood, using standard extraction methods. After quality control, genomic DNA (3 μg) intended for WES was captured using an in-solution enrichment method (Human All Exon v5—50 Mb, Agilent Technologies, CA, USA). Exome-enriched libraries (~20,000 targeted genes) were prepared on an automated system, using the NGSx robot (Perkin Elmer Inc, MA, USA) and the Bravo robot (Agilent Technologies, CA, USA) respectively, according to the manufacturers’ instructions (SureSelect, Agilent Technologies). After normalization and quality control, exome-enriched libraries were sequenced on a HiSEQ 2000 system (Illumina Inc., CA, USA) as paired-end 100b reads. The mean sequencing coverage was at least 60 to 70x for each sample. Image analysis and base calling were performed using the Illumina Real
Time Analysis Pipeline. Sequence quality parameters were assessed daily throughout the 12-day sequencing run. Sequences were aligned with the hg19 reference human genome, using the Burrows-Wheeler Aligner. Downstream processing was carried out with the Genome Analysis Toolkit (GATK), SAMtools, and Picard, in line with documented best practices (http://www.broadinstitute.org/gatk/guide/topic?name=best-practices). Variants were called with the GATK Unified Genotyper. All calls with a read coverage of ≤2X or a Phredscaled SNP quality score of ≤20 were removed from the analysis. All variants were annotated using a software system developed by the Paris Descartes University Bioinformatics platform. All the annotation procedures were based on the latest release of the Ensembl database. Sanger sequencing of the family members of these patients was performed to confirm the next-generation sequencing results and to analyze each mutation’s familial segregation. Purified PCR products were directly sequenced using BigDye Terminators (version 1.1) and a 3500xL Genetic Analyzer (Applied Biosystems).

For patient Helios$^{R106W}$ genomic DNA was extracted from whole blood according to the standard methods, and whole exome sequencing and data analysis were performed as previously described. Sanger sequencing was then applied to the patient’s family members to confirm NGS findings and analyze segregation of the mutation within the family. Genomic DNA regions flanking the IKZF2 R106W mutation were amplified using Q5 High-Fidelity DNA Polymerase (NEB) according to the manufacturer’s recommendations, gel-purified with a High Pure PCR Product Purification Kit (Roche), sequenced with a BigDye terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, University Park, Ill), and analyzed on a 3500xL Genetic Analyzer (Applied Biosystems). All collected sequences were analyzed by using a DNADynamo (BlueTractorSoftware).

For patient Helios$^{V347M}$ WES was performed as previously described and the mutation was confirmed and segregated within the family using Sanger sequencing as previously described.

**Plasmids**

Human $IKZF2$ (NM_016260; HELIOS; Origene (RC214813)), $IKZF1$ (NM_006060; IKAROS) and $IKZF3$ (NM_012481; AIOLOS; GenScript (Clone ID: Ohu21008D)) cDNA
were synthesized and subcloned into a CMV6 driven C-terminal Myc-Flag-tagged mammalian expression vector, or into pTO-Strep-HA plasmid by gateway recombination to generate an N-terminal Streptavidin-Hemagglutinin tag. Patient variants were generated using the Q5® Site-Directed Mutagenesis Kit (NEB) or AccuPrime Pfx DNA Polymerase followed by DpnI treatment.

**Light Shift chemiluminescent EMSA**

To assess the ability of HELIOS variants to bind to DNA, HEK293T cells were transfected with pCMV6-C-Myc-Flag-HELIOS WT or mutant and electrophoretic mobility shift assays were performed with LightShift Chemiluminescent EMSA kit (Thermo Fisher Scientific) as previously described. IKAROS consensus-binding sequence 1 (IK-bs1) was used for EMSA assay (forward: 5′-BIOTIN-TCAGCTTTTGGGAATACCCTGTCA; reverse: 5′-BIOTIN-TGACAGGGTATTCCCCAAAGCTGA).

**Immunophenotyping and HELIOS staining**

Cryopreserved peripheral blood mononuclear cells (PBMCs) from patients and healthy controls were thawed and stained with fluorescent antibodies (Supplemental Table 4) for the determination of frequencies of subpopulations of T, B and NK cells. HELIOS and FOXP3 staining was performed using the eBioscience™ Foxp3/transcription factor staining buffer set (#00552300, ThermoFisher). Stained cells were acquired with an LSR-Fortessa (BD Biosciences).

**Co-immunoprecipitation (Co-IP)**

The ability of HELIOS variants to homo/heterodimerize and bind to protein interaction partners was assessed by Co-IP studies. HEK293T cells were co-transfected with Strep-HA-HELIOS WT or mutants, together with either pCMV6-C-Myc-Flag-HELIOS WT or mutant, pCMV6-C-Myc-Flag-IKAROS or AIOLOS, or pCMV-HA-HDAC1. After 24 hours, cells were washed with PBS and lysed with IP lysis buffer (50mM Tris (pH 7.5), 150nM NaCl, 2mM EDTA, 0.5% Triton X-100, 1mM PMSF and protease inhibitor cocktail) for 30 minutes on ice. Lysates were centrifuged for 10 minutes at 13,000 rpm and some taken for input. Rest of lysates were incubated with Strep-Tactin® Sepharose® 50% suspension by rotating at 4 °C for 2 hours. Beads were then washed three times with lysis buffer and resuspended with 1.5X SDS loading buffer. Beads and input were analyzed by running on 8% SDS-PAGE and Western blotting with the
following primary antibodies: HELIOS (D8W4X, CST 42427S), HDAC1 (ab7028, Abcam), HA-HRP (HA-7, Sigma Aldrich), FLAG (F7425, Sigma Aldrich) and GAPDH (G-9, sc-365062).

**Immunofluorescence microscopy**

To assess the ability of HELIOS variants to form foci around pericentromeric heterochromatin, NIH3T3 cells adherent on glass slides in 6-well plates were transfected with pCMV6-C-Myc-Flag-HELIOS WT or mutant using the X-tremeGENE™ 9 DNA Transfection Reagent (Sigma Aldrich), following manufacturer’s protocol. After 20 hours of transfection, cells were fixed with 4% formaldehyde (#28908, Thermo Scientific) for 15 minutes, permeabilized (0.1% Triton X-100 in PBS) for 10 minutes and blocked (0.1% Triton X-100 and 10% FBS in PBS) for 30 minutes at room temperature. Cells were then incubated with HELIOS (D8W4X, CST 42427S) antibody for 2 hours, stained with Rabbit IgG-Alexa Fluor 488 secondary antibody (#A11034, Invitrogen) for 1 hour (both in blocking buffer at 4 °C and in the dark) and then with DAPI (#63351, Roth). Coverslips were mounted in Fluorescent Mounting Medium (Dako). Axio Imager M2-1 (Zeiss) microscope was used to acquire the images with a 40x objective. The formation of foci and number of foci per cell for each mutant were quantified using Cell Profiler (https://cellprofiler.org). We used the Speckle Counting CellProfiler Pipeline provided on the website and modified it further to add a filtering step where only transfected cells were analyzed.

**ScRNAseq**

Frozen PBMCs from the index patient (R291X) and four controls were thawed, washed in RPMI media and resuspended in sterile PBS with 0.04% bovine serum albumin (BSA). A single-cell suspension was obtained by passing one million cells into a 5 mL FACS tube through a cell strainer and sorting for the live lymphocytes and monocytes based on the forward and side-scatter using the FACSaria™ Fusion (BD). Single-cell RNA-seq was then performed on the live samples using the 10x Genomics Chromium Controller with the Chromium Single Cell 3’ Reagent Kit (v3 chemistry) following the manufacturer’s instructions. 23,946 Libraries were sequenced on the Illumina HiSeq 4000 platform in 2x75bp paired-end mode. Supplemental Table 5 includes an overview of sequenced samples. We used the CellRanger v3.0.2 software (10x Genomics) for demultiplexing and alignment to GRCh38 human reference transcriptome. The R statistics software v 4.0.2 was used to analyze the
processed data. For optimal reproducibility, we used a Docker container containing R and all dependent libraries pre-installed (cancerbits/dockR:mo-ikzf2-v1 available from https://hub.docker.com/r/cancerbits/dockr). Briefly, CellRanger outputs were loaded into R to perform quality control (removing cells with less than 500 genes, mitochondrial content more than 10%, or log_{10} doublet score >3 [calculated using the function doubletCells from scran (Lun et al., 2016) v1.14.6] and downstream analysis using Seurat v 3.9.9.9038. Each dataset was normalized to generate corrected, log-transformed counts, and Pearson residuals by regressing out mitochondrial content. We then used these normalized data to integrate all datasets using Harmony v1.0. Integrated data was used for low-dimensional projection using UMAP based on the top 30 principal components and for clustering cells (resolution = 0.1). Differential gene expression (DGE) analysis for this rare disease dataset suffers from lack of replication. In an attempt to reduce the impact of inter-individual differences (which would confound comparisons with variable number of cells from different individuals on both sides of a comparison), we used a stratified comparison scheme in which we performed comparisons per pair of individuals and focused on consistent differences across all pairwise comparisons. Specifically, statistics from sample-wise DGE tests were combined using the minimum log fold change (“worst fold change”) and Fisher’s method for p-values. Cluster-sample strata with less than 40 cells, as well as ribosomal and mitochondrial genes, were excluded from DGE analysis. We used MAST as a statistical test with FDR-adjusted p-value threshold of q < 0.05 and log fold change ≥ 0.5. Genes that satisfied these thresholds and identified with the same direction of change in at least three of the four were recorded as differentially expressed genes (supplemental Table 2).

**Interaction proteomics**

Proximity-dependent Biotin Identification (BioID) technique, coupled with liquid chromatography mass spectrometry (LC-MS) was performed according to Liu et al. Flip-In T-Rex HEK293 cell lines were co-transfected with C-terminally MAC-tagged Helios^{WT} or Helios^{Y359C}, Helios^{V347M}, Helios^{R106W}, Helios^{N220S}, Helios^{R291X} or GFP (with nuclear localization signal) together with the pOG44 vector for stable expression. Positive clones were selected and amplified. Cells from three biological replicates were induced with Tetracyclin and biotin induction for 24 hours, harvested after induction,
and affinity purified using a harsh lysis buffer. Lysates were cleared and loaded onto Bio-Rad spin columns with Strep-Tactin beads (IBA), followed by stringent washes and elution. The eluate was then reduced, alkylated, trypsin digested, desalted and dried as described, followed by reconstitution in 0.1% TFA and 1% Acetonitrile. Analysis of the sample was performed on a Q-Exactive mass spectrometer with an EASY-nLC 1000 system via an electrospray ionization sprayer (Thermo Fisher Scientific), using Xcalibur version 3.0.63. Database search was performed with Proteome Discoverer 1.4 (Thermo Scientific) using the SEQUEST search engine on the Reviewed human proteome in UniProtKB/SwissProt databases (http://www.uniprot.org, downloaded May 2020). All reported data were based on high-confidence peptides assigned in Proteome Discoverer (FDR < 0.05).

**Statistical analysis**

Significance Analysis of INTeractome (SAINT)-express version 3.6.3 and Contaminant Repository for Affinity Purification (CRAPome, http://www.crapome.org) were used to discover statistically significant interactions from the LC-MS data. Final results represent proteins with a BFDR score lower than 0.05, and present in all three replicates. MaxQuant version 1.6.10.43 with the Andromeda search engine was used to compare differences in LFQ intensity between WT and Y359C, V347M, R106W, N220S, or R291X samples, two-tailed two-sample equal variance Student’s t-test was performed and p-values lower than 0.05 were considered significant.

Mann-Whitney U test was used to address statistically significant differences in the subpopulation of immune cells analyzed by flow cytometry (FACS) in this study.
Supplemental Table 1

| SAMPLE NAME | GENOTYPE | AGE | SEX | SAMPLE DESCRIPTION | ORGANISM | TISSUE | CELL TYPE |
|-------------|----------|-----|-----|--------------------|----------|--------|-----------|
| PATIENT R291X | R291X-Het | 14 | Female | Sorted viable PBMCs from frozen sample | Human | Blood | PBMCs |
| ND-1819 | WT | 13 | Male | Sorted viable PBMCs from frozen sample | Human | Blood | PBMCs |
| ND-1999 | WT | 11 | Female | Sorted viable PBMCs from frozen sample | Human | Blood | PBMCs |
| ND-2000 | WT | 31 | Male | Sorted viable PBMCs from frozen sample | Human | Blood | PBMCs |
| ND-2001 | WT | 26 | Female | Sorted viable PBMCs from frozen sample | Human | Blood | PBMCs |

Supplemental Table 1. Overview of samples sequenced via single cell RNA-seq.
Supplemental Table 2. Clinical and genetic data of patients with variants in *IKZF2*; AA, amino acid; AIHA, autoimmune hemolytic anemia; EBV, Epstein-Barr virus; FSGN, focal segmental glomerulonephritis; HLH, hemophagocytic lymphohistiocytosis; ITP, immune-thrombocytopenia; IVIG, intravenous immunoglobulin substitution; NSAIDs, non-steroidal anti-inflammatory drugs; NT, nucleotide; SLE, systemic lupus erythematosus; *Patient suffered low grade fever, migratory popular rash, inflammatory joint disease; #Patient was affected by hepatomegaly, edema, diarrhea and seizures.

|               | P1 (R291X) | P2 – DAUGHTER (Y359C) | P3 – MOTHER (Y359C) | P4 (N220S) | P5 (V347M) | P6 (R106W) |
|---------------|------------|-----------------------|---------------------|------------|------------|------------|
| **GENETIC DATA** |            |                       |                     |            |            |            |
| POSITION      | 213872794  | 213872589             | 213872589           | 213866770  | 213872626  | 213921647  |
| NT REF        | G          | T                     | T                   | T          | C          | G          |
| NT ALT        | A          | C                     | C                   | C          | T          | A          |
| AA CHANGE     | R291X      | Y359C                 | Y359C               | N220S      | V347M      | R106W      |
| CADD          | 38         | 25.2                  | 25.2                | 22.80      | 24.2       | 32         |
| POLYPHEN      | NA         | probably damaging     | probably damaging   | probably damaging | probably damaging | probably damaging |
| Gnomad AF     | 0.000004125| 0.00007809            | 0.00007809          | 0.00004385 | NA         | 0.00003186 |
| **CLINICAL DATA** |            |                       |                     |            |            |            |
| CURRENT AGE   | 16         | 24                    | 56                  | unknown    | 27         | 29         |
| AGE AT ONSET  | 10         | 15                    | unknown             | 2          | 25         | 23         |
| AUTOIMMUNE/AUTOINFLAM. FEATURES | SLE (polyarthritis, ITP, skin rash) | ITP | SLE (ITP, glomerulonephritis) | Evan's Syndrome (ITP, AIHA) | EBV HLH, FSGN, serositis, arthritis | EBV associated Adult onset Still's Disease*. |
| CONCOMITANT FEATURES | -         | -                     | -                   | -          | +*        | -          |
| POSITIVE AUTO-ANTIBODIES | Anti-nuclear, anti-DNA, anti-GPllb | Anti-nuclear, anti-DNA, anti-platelet | Anti-nuclear, anti-DNA | -         | -         | -          |
| REC. RESPIRATORY TRACT INFECTIONS | -         | -                     | -                   | +         | +         | +          |
| HYPOGAMMLOBULINEMIA | -         | -                     | -                   | +         | +         |            |
| TREATMENT     | Hydroxychloroquine, NSAIDs, full sun protection | Hydroxychloroquine, corticosteroids | Hydroxychloroquine, corticosteroids | Corticosteroids, IVIG | Refractory to corticosteroids, etoposide, IVIG, rituximab, anakinra, emapalumab, ruxolitinib | In remission since HSCT at 27yrs |
### Supplemental Table 3

| clust | gene     | min_lfc | adj_pval | clust | gene     | min_lfc | adj_pval |
|-------|----------|---------|----------|-------|----------|---------|----------|
| 1_a   | H3F3B    | 0.502727645 | 0 | 3     | HSP90AB1 | 0.936581295 | 0 |
| 1_a   | CTNNB1   | 0.540186584  | 0 | 3     | YBX1     | 0.55496621  | 4.64E-281 |
| 1_a   | MAP3K8   | 0.5463886869 | 0 | 3     | EIF5A    | 0.583918872 | 1.00E-214 |
| 1_a   | S100A10  | 0.55338442  | 0 | 4     | HLA-DRB5 | -2.742503778 | 3.54E-266 |
| 1_a   | MALAT1   | 0.570824855 | 0 | 4     | H3F3B    | 0.537018237 | 2.66E-123 |
| 1_a   | CTNNB1   | 0.540186584 | 0 | 3     | YBX1     | 0.55496621  | 4.64E-281 |
| 1_a   | KLF3     | 0.619402801 | 0 | 4     | EIF5A    | 0.583918872 | 1.00E-214 |
| 1_a   | MALAT1   | 0.570824855 | 0 | 4     | H3F3B    | 0.537018237 | 2.66E-123 |
| 1_a   | CEBP8    | 0.717656578 | 0 | 4     | FOS      | 0.605584724 | 1.61E-62  |
| 1_a   | CEBP8    | 0.717656578 | 0 | 4     | FOS      | 0.605584724 | 1.61E-62  |
| 1_a   | CEBP8    | 0.717656578 | 0 | 4     | FOS      | 0.605584724 | 1.61E-62  |
| 1_a   | PHACTR1  | 0.764809363 | 0 | 4     | CEBP8    | 0.52378203  | 7.29E-47  |
| 1_a   | C5AR1    | 0.776935716 | 0 |        |          |          |          |
| 1_a   | FOS      | 1.013524343 | 0 |        |          |          |          |
| 1_b   | HLA-DRB5 | -1.502776124 | 1.57E-150 | 1_b   | LITAF    | 0.990912749 | 1.18E-80  |
| 1_b   | LITAF    | 0.990912749 | 1.18E-80 | 1_b   | LITAF    | 0.990912749 | 1.18E-80  |
| 1_b   | LYZ      | 1.228789642 | 7.85E-58 | 1_b   | NINJ1    | 0.687561848 | 8.82E-57  |
| 1_b   | NINJ1    | 0.687561848 | 8.82E-57 | 1_b   | CD63     | 0.559003274 | 2.18E-44  |
| 1_b   | CD63     | 0.559003274 | 2.18E-44 | 1_b   | BCL2A1   | 0.567942429 | 2.17E-34  |
| 1_b   | BCL2A1   | 0.567942429 | 2.17E-34 | 1_b   | CDV3     | 0.63282745  | 3.15E-34  |
| 1_b   | CDV3     | 0.63282745  | 3.15E-34 | 1_b   | KLF3     | 0.53137921  | 6.57E-34  |
| 2_a   | KLF3     | 0.53137921  | 6.57E-34 | 2_b   | BTG1     | 0.768166619 | 5.10E-299 |
| 2_b   | BTG1     | 0.768166619 | 5.10E-299| 2_b   | ZFP36    | 0.855491874 | 1.30E-243 |
| 2_b   | ZFP36    | 0.855491874 | 1.30E-243| 2_b   | GNLY     | -1.412357846 | 3.11E-230 |
| 2_b   | GNLY     | -1.412357846 | 3.11E-230| 2_b   | HSPA5    | 0.744855672 | 1.48E-228 |
| 2_b   | HSPA5    | 0.744855672 | 1.48E-228| 2_b   | DUSP2    | 0.667837697 | 4.41E-214 |
| 2_b   | DUSP2    | 0.667837697 | 4.41E-214| 2_b   | SBDS     | 0.700021801 | 1.37E-201 |
| 2_b   | SBDS     | 0.700021801 | 1.37E-201| 2_b   | HSP90AB1 | 0.756435329 | 2.56E-180 |
| 2_b   | HSP90AB1 | 0.756435329 | 2.56E-180| 2_b   | LDHA     | 0.762446136 | 1.80E-170 |
| 2_b   | LDHA     | 0.762446136 | 1.80E-170| 2_b   | PABPC1   | 0.564332115 | 1.49E-130 |
| 2_b   | PABPC1   | 0.564332115 | 1.49E-130| 2_b   | SRSF5    | 0.500741978 | 1.86E-90  |
| 2_b   | SRSF5    | 0.500741978 | 1.86E-90 | 2_b   | CCL5     | -0.608539489 | 6.63E-90  |

Supplemental Table 3. Single-cell RNA sequencing of PBCMs from P1 (R291X).
## Supplemental Table 4

### Lymphocyte Subsets

| Age at evaluation | R291X (11 yrs) | Y359C-Daughter (24 yrs) | Y359C-Mother (55 yrs) | V347M (26 yrs)* | R106W (29 yrs) | N220S (13 yrs) |
|-------------------|----------------|-------------------------|----------------------|----------------|----------------|----------------|
| Absolute Lymphocyte Count, cells/μl | 1607 (1710-3060) | (1140-3380) | (1140 – 3380) | 600 (1140 – 3380) | 4589 (1140 – 3380) | 1527 (1710-3060) |
| CD3+ T cells, cells/μl | 1093 (1070-2270) | (780-2240) | (780-2240) | 352 (780-2240) | 3849 (780-2240) | 1145 (1070-2270) |
| CD4+ T cells, cells/μl | 562 (640-1290) | (490-1640) | (490-1640) | 54 (490-1640) | 1646 (490-1640) | 809 (640-1290) |
| CD45RACCR7+, % | 58.1 (25-63) | 57.2 (17-67) | 53.1 (17-67) | 0.54 (17-67) | 66.2 (17-67) | 24 (25-63) |
| CD45RACCR7+, % | 20.9 (11-25) | 12.9 (24-68) | 30.3 (24-68) | 31.4 (24-68) | 24 (24-68) | 24 (11-25) |
| CD45RACCR7+, % | 20.3 (12-30) | 24.7 (3-25) | 15.8 (3-25) | 66.5 (3-25) | 8.6 (3-25) | 24 (12-30) |
| CD45RACCR7+, % | 0.7 (4-24) | 5.3 (0-15) | 0.7 (0-15) | 0.1 (0-15) | 1.2 (0-15) | 4.5 (4-24) |
| CD8+ T cells, cells/μl | 466 (380-880) | (170-880) | (170-880) | 259 (170-880) | 2115 (170-880) | 275 (380-880) |
| CD45RACCR7+, % | 48 (45 (22-58) | 78.2 (19-79) | 64.2 (19-79) | 3.54 (19-79) | 61.1 (19-79) | 48 (22-58) |
| CD45RACCR7+, % | 7 (2-15) | 6.9 (2-27) | 6.9 (2-27) | 5.01 (2-27) | 8.1 (2-27) | 5 (2-15) |
| CD45RACCR7+, % | 45 (24-58) | 11.6 (2-22) | 23.4 (2-22) | 86.6 (2-22) | 15 (2-22) | 29 (2-22) |
| CD45RACCR7+, % | 7 (7-26) | 7.7 (9-61) | 4.5 (9-61) | 4.9 (9-61) | 15.8 (9-61) | 18 (7-26) |
| CD4/CD8 | 1.2 (0.9-3.4) | 1.1 (0.9-5.0) | 0.6 (0.9-5.0) | 0.21 (0.9-5.0) | 0.8 (0.9-5.0) | 2.9 (0.9-3.4) |
| CD56*CD16+ NK cells/μl | 143 (170-530) | (80-690) | (80-690) | (80-690) | (80-690) | (170-530) |
| CD19+ B cells, cells/μl | 402 (170-630) | (80-490) | (80-490) | a-CD20 start | a-CD20 start | a-CD20 start |
| CD27IgD+, % | 78 (64-84) | 1.1 (32-84) | 0.6 (32-84) | n/a (32-84) | 61.5 (32-84) | n/a (64-84) |
| CD27IgD+, % | 8 (4-14) | 17.6 (2-18) | 2.9 (2-18) | n/a (2-18) | 20.9 (2-18) | 7 (4-14) |
| CD27IgD+, % | 6 (6-16) | 17.6 (7-41) | 2.9 (7-41) | n/a (7-41) | 15.8 (7-41) | 1.4 (7-41) |

### Age at evaluation

- 26 yrs (prior to a-CD20)
- 29 yrs

### Immunoglobulin

| | | | | | | |
|---|---|---|---|---|---|
| IgG (g/l) | 0.172 (8-18) | | 9.12 (8-18) | | |
| IgM (g/l) | 0.20 (0.6-2.5) | | 1.22 (0.6-2.5) | | |
| IgA (g/l) | <0.40 (0.9-4.5) | | 0.72 (0.9-4.5) | | |

**Supplemental Table 4.** Laboratory test performed on six patients with rare variants in IKZF2. Patients were compared to age matched published reference datasets. Blue and red font indicate levels that are above or below reference values, respectively. n/a, not applicable; a-CD20, anti-CD20 monoclonal antibody therapy (rituximab), * laboratory features of patient Helios(V347M) potentially influenced by treatment with cyclosporine and steroids a time of analysis.
### Supplemental Table 5

| Antibody-Fluorophore | Clone | Producer          |
|-----------------------|-------|-------------------|
| CD3-APC-H7            | SK7   | BD Biosciences    |
| CD4-BV605             | RPA-T4| BD Pharmigen      |
| CD4-PerCPCy5.5        | RPA-T4| eBioscience       |
| CD4-BV421             | RPA-T4| BD Pharmigen      |
| CD4-PE                | 13B8.2| Beckman Coulter   |
| CD8-V450              | RPA-T8| BD Horizon        |
| CD8-V500              | RPA-T8| BD Pharmigen      |
| CD8-PECy7             | SFCi21Thy2D3 | Beckman Coulter |
| CD8-FITC              | HIT8a | BD Pharmigen      |
| CD16-FITC             | 3G8   | Beckman Coulter   |
| CD19-PerCPCy5.5       | HIB19  | eBioscience       |
| CD19-PECy7            | J3-119| Beckman Coulter   |
| CD21-PE               | B-ly4  | BD Pharmigen      |
| CD25-PE               | MA-A251| BD Pharmigen     |
| CD25-BV605            | 2A3   | BD Horizon        |
| CD27-V450             | M-T271| BD Pharmigen      |
| CD28-APC              | CD28.2| eBioscience       |
| CD31-APC              | WM59  | eBioscience       |
| CD38-PECy7            | HIT2  | BD Pharmigen      |
| CD45RA-AF700          | HI100  | BD Pharmigen      |
| CD56-V450             | B159  | BD Pharmigen      |
| CD57-PE               | TB01  | eBioscience       |
| CD95-PECy7            | DX2   | BD Pharmigen      |
| CD127-PE-CD594        | HIL-7R-M21 | BD Horizon   |
| PD1-PE                | EH12.1| BD Pharmigen      |
| TCRαβ-FITC            | WT31  | BD Biosciences    |
| TCRγδ-PE              | 11F2  | BD Biosciences    |
| TCR γδ24-PECy7        | C15   | Beckman Coulter   |
| TCRβ111-FITC          | C21   | Beckman Coulter   |
| IgD-FITC              | IA6-2 | BD Pharmigen      |
| Antibody                  | Code     | Vendor            |
|--------------------------|----------|-------------------|
| IgM-APC                  | G20-127  | BD Pharmigen      |
| CCR4-PECy7               | 1G1      | BD Pharmigen      |
| CCR6-BV605               | G034E3   | Biolegend         |
| CCR7-PE-CF594            | 150503   | BD Pharmigen      |
| CXCR3-BV711              | 1C6      | BD Horizon        |
| CXCR5-APC                | 51505    | R&D Systems       |
| HELIOS-PE                | 22F6     | BD Pharmigen      |
| FOXP3-APC                | PCH101   | eBioscience       |
| IL-17A-eF450             | eBio64DEC17 | eBioscience    |
| IFNγ-FITC                | 4S.B3    | eBiosciences      |

**Supplemental Table 5.** List of fluorescent antibodies used for flow cytometry.
Supplemental Figures

Supplemental Figure 1. Identification of 4 additional heterozygous variants in *IKZF2* and population genetics data of *IKZF2* (A) Haploinsufficiency susceptibility prediction of *IKZF2*. The pLI score for *IKZF2* is extremely high (0.99), compared to other known haploinsufficient genes, such *CDH7* or *IKZF1*. (B) Family pedigrees of the 4 additional patients. In family B, a missense variant was detected and carried by the mother (B1, 54 years) and daughter (B2, 24 years) leading to a tyrosine to cytosine substitution (c.1076A>G, p.Tyr359Cys, rs751327038). A unique missense variant (c.1038 G>A, Val347Met) was detected in a 27-year-old male patient (C1). Patient D1 carries a missense variant in exon 3, before the DNA binding zinc-finger domains of HELIOS (c.316 C>T, p.Arg106Trp, rs750531870). Patient E1 is an adopted child carrying a heterozygous missense variant in exon 5, immediately after the DNA binding domain (c.659A>G, p.Asn220Ser, rs566824591). Half-filled square represents asymptomatic family member carrying the *IKZF2* variant. (C) Schematic representation of Helios protein domains showing the location of all the variants described (red arrows).
Supplemental Figure 2. The ability of forming heterodimers with other members of the IKAROS family is intact in the additional *IKZF2* heterozygous variants (A-B) Co-immunoprecipitation (Co-IP) experiments following co-transfection of HEK293T cells with Strep-HA-HeliosWT, HeliosR291X, HELIOSY359C, HELIOSN220S, HELIOSV347M, HELIOSR106W or an Empty Vector (EV) together with either FLAG-Ikaros (A) or FLAG-Aiolos (B). IP was performed with Strep-beads and Western blot analysis was done by running both the IP and whole cell lysate (input) on a gel and blotting with HA, FLAG and GAPDH antibodies. Results are representative of 3-5 independent experiments. Note that panels shown in Fig.1 for the R291X variant are also shown here for better comparison with the other variants.
Supplemental Figure 3. Single cell RNA sequencing methodology. (A) Cell-type inference at two levels of annotation resolution using reference mapping to the CITE-
seq dataset of 211,000 human PBMCs from Hao et al., 2021⁹. (B) Heatmaps showing differentially expressed genes in UMAP cell clusters (see Figure 1). (C) The percentage of cells expressing Helios in each cell subtype (using the cell type labels from panel B).

Supplemental Figure 4. Laboratory features of patients carrying HELIOS Y359C, HELIOS N220S, HELIOS V347M, HELIOS R106W or HELIOS R291X variants. (A-F) Graphs representing the frequency of lymphocytes or lymphocyte subsets in patients with variants in IKZF2. PBMC from healthy controls or patients were labelled with mAbs against CD3, CD4, CD8, CD16, CD19, CD25, CD27, CD56 and CD127. (A) Graph representing the percentage of lymphocytes in the patients with variant in IKZF2 versus healthy donors as monitored by the expression of CD19. (B) Proportion of CD27⁺IgD⁺ class-switched B-cells within CD3⁻CD19⁺B-cells, with cells of Helios Y347M patient excluded due to absence of CD3⁻CD19⁺B-cells. (C) Proportion of follicular T-cells within CD3⁺CD4⁺ cells. (D) Proportion of NK cells
(CD16^+CD56^{dim}) within lymphocytes. (E) Frequency of CD4^+ and CD8^+ cells within total (CD3^+) T-cells. (F) Proportion of regulatory T cells (CD127^-CD25^-FOXP3^+) within CD3^-CD4^+. (G) HELIOS mean fluorescence intensity (MFI) within CD127^-CD25^-FOXP3^+ regulatory T cells. (H) Graph representing FOXP3^+ Treg cells within CD3^+CD4^+CD25^-Helios^+. (I) Graph representing mean fluorescence intensity (MFI) of CD3^+CD4^+CD25^-Helios^+FOXP3^+ Treg cells. Note that for all the patients, one time point is represented, except for patient Helios^{V347M} for whom two time points are represented in C, D, and E. These two time points were taken 6 months apart. At the time of both time points, the patient was being treated with cyclosporine for nephrotic syndrome and steroids for EBV HLH. The second blood sample was shortly before conditioning for HSCT, when the patient was very ill with refractory HLH due to EBV, chronic diarrhea, TPN-dependent malnutrition and nephrotic syndrome. Different symbols represent each of the patients with IKZF2 variants. Mann-Whitney U test was used for statistical analysis, taking the mean value from the two time points of patient Helios^{V347M}. 
Supplemental Figure 5. DNA and dimerization properties of the *IKZF2* additional heterozygous variants. (A) Electromobility shift assay (EMSA) was performed using nuclear extracts from HEK293T cells to show the ability of HELIOS^WT, HELIOS^R291X, HELIOS^Y359C, HELIOS^N220S, HELIOS^V347M and HELIOS^R106W to bind to the IK-BS1 probe as dimers and multimers. The red asterisk indicates HELIOS^R291X binding to the IK-BS1 probe as a monomer. Representative image of 3-5 independent experiments; EV: empty vector. Subfigure also shown in Figure 1C. (B) Representative confocal images of immunofluorescence stainings of NIH3T3 cells transfected with HELIOS^WT and HELIOS^Y359C, HELIOS^N220S, HELIOS^V347M and
HELIOS\textsuperscript{R106W} using anti-HELIOS antibody, showing the formation of HELIOS-positive loci at pericentromeric heterochromatin regions. Subfigure also shown in Figure 1D. (C) Quantification of the number of loci per cell represented in B. Between 554-1183 cells were quantified per condition. Results represent 3 independent experiments. Unpaired t-test was performed. (D) Graph representing the percentage of cells showing 1 or more loci in cells expressing either HELIOS\textsuperscript{WT} and HELIOS\textsuperscript{Y359C}, HELIOS\textsuperscript{N220S}, HELIOS\textsuperscript{V347M} or HELIOS\textsuperscript{R106W}. Note that panels shown in Fig.1 for the R291X variant are also shown here for better comparison with the other variants.
Supplemental Figure 6. The ability of forming homodimers and heterodimers with HeliosWT is intact in the additional IKZF2 heterozygous variants. (A-B) Co-immunoprecipitation (Co-IP) experiments following co-transfection of HEK293T cells with Strep-HA-HeliosWT, HeliosR291X, HELIOSY359C, HELIOSN220S, HELIOSV347M, HELIOSR106W or an Empty Vector (EV) together with FLAG-HeliosWT, HeliosR291X, HELIOSY359C, HELIOSN220S, HELIOSV347M or HELIOSR106W. IP was performed with Strep-beads and Western blot analysis was done by running both the IP and whole cell lysate (input) on a gel and blotting with HA, FLAG and GAPDH antibodies. Results are representative of 3-5 independent experiments. Note that panels shown in Fig.1 for the R291X variant are also shown here for better comparison with the other variants.
Supplemental Figure 7. HELIOS^{N220S} and HELIOS^{R106W} show similar aberrant interactome as HELIOS^{R291X} Bio-ID proximity labelling assay. Heatmap showing the hierarchical
clustering of HELIOS mutants based on their high-confidence interactions. Heatmap colour displays the detected interactor abundance (AvgSpec).
Supplemental Figure 8. Helios variants show statistically significant changes in the interactions with some components of the NuRD, SWI/SNF or Cohesin complexes compared to the Helios$^{WT}$. Heatmap showing significantly altered Average PSM values. Heat values is the fold change between the Prey AvgSpec and WT AvgSpec. Asteriks represent statistically significant differences.

References
1. Triebwasser MP, Barrett DM, Bassiri H, et al. Combined use of emapalumab and ruxolitinib in a patient with refractory hemophagocytic lymphohistiocytosis was safe and effective. *Pediatr Blood Cancer*. 2021;68(7):e29026.
2. Izawa K, Martin E, Soudais C, et al. Inherited CD70 deficiency in humans reveals a critical role for the CD70-CD27 pathway in immunity to Epstein-Barr virus infection. *J Exp Med*. 2017;214(1):73-89.
3. Le Coz C, Nguyen DN, Su C, et al. Constrained chromatin accessibility in PU.1-mutated agammaglobulinemia patients. *J Exp Med*. 2021;218(7).
4. Boutboul D, Kuehn HS, Van de Wyngaert Z, et al. Dominant-negative IKZF1 mutations cause a T, B, and myeloid cell combined immunodeficiency. *J Clin Investig*. 2018;128(7):3071-3087.
5. Liu Z, Yang L, Cui Y, et al. IL-21 enhances NK cell activation and cytolytic activity and induces Th17 cell differentiation in inflammatory bowel disease. *Inflammatory bowel diseases*. 2009;15(8):1133-1144.
6. Garcia-Prat M, Álvarez-Sierra D, Aguiló-Cucurull A, et al. Extended immunophenotyping reference values in a healthy pediatric population. *Cytometry B Clin Cytom*. 2019;96(3):223-233.
7. Jentsch-Ullrich K, Koenigsmann M, Mohren M, Franke A. Lymphocyte subsets' reference ranges in an age- and gender-balanced population of 100 healthy adults--a monocentric German study. *Clin Immunol*. 2005;116(2):192-197.
8. Yi JS, Rosa-Bray M, Staats J, et al. Establishment of normative ranges of the healthy human immune system with comprehensive polychromatic flow cytometry profiling. *PLoS One*. 2019;14(12):e0225512.
9. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573-3587.e3529.