

**GERMLINE TET2 LOSS-OF-FUNCTION CAUSES
CHILDHOOD IMMUNODEFICIENCY AND LYMPHOMA**

Jarmila Stremenova Spegarova^{1*}, Dylan Lawless^{2*}, Siti Mardhiana Binti Mohamad³, Karin R. Engelhardt¹, Gina Doody², Jennifer Shrimpton², Anne Rensing-Ehl⁴, Stephan Ehl⁴, Frederic Rieux-Lauca⁵, Catherine Cargo⁶, Helen Griffin¹, Aneta Mikulasova¹³, Meghan Acres¹, Neil V. Morgan⁷, James A. Poulter², Eamonn G. Sheridan², Philip Chetcuti⁸, Sean O'Riordan⁸, Rashida Anwar², Clive Carter⁹, Stefan Przyborski¹⁰, Kevin Windebank¹¹, Andrew J. Cant^{1,12}, Majlinda Lako¹³, Chris M. Bacon^{11,14}, Sinisa Savic^{9,15**}, Sophie Hambleton^{1,12**}

- 1) Primary Immunodeficiency Group, Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, UK
- 2) Leeds Institute of Medical Research, St James' University Hospital, University of Leeds, Leeds, UK
- 3) Regenerative Medicine Cluster, Advanced Medical and Dental Institute, Universiti Sains Malaysia, Bertam, Penang, Malaysia
- 4) Institute for Immunodeficiency, Center for Chronic Immunodeficiency, Freiburg, Germany
- 5) Institut Imagine, Paris, France
- 6) Hematological Malignancy Diagnostic Service, St James's University Hospital, Leeds, UK
- 7) Institute of Cardiovascular Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK
- 8) Department of Paediatrics, Leeds General Infirmary, Leeds, UK
- 9) Department of Clinical Immunology and Allergy, St James's University Hospital, Beckett Street, Leeds, UK.
- 10) Department of Biosciences, Durham University, Durham, UK
- 11) Wolfson Childhood Cancer Research Centre, Northern Institute for Cancer Research, Newcastle University, Newcastle upon Tyne, UK
- 12) Great North Children's Hospital, Newcastle upon Tyne, UK
- 13) Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK
- 14) Department of Cellular Pathology, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, United Kingdom
- 15) NIHR, Leeds Biomedical Research Centre and Leeds Institute of Rheumatic and Musculoskeletal Medicine, Wellcome Trust Brenner Building, St James's University Hospital, Leeds, UK

* These authors contributed equally to this work

** These authors contributed equally to this work

Contact authors: S.Savic@leeds.ac.uk, sophie.hambleton@newcastle.ac.uk

SUPPLEMENTAL DATA

Contents

Supplemental materials and methods

Supplemental Table S1. Immunoglobulin levels of patients before transplantation.

Supplemental Table S2. Details of HSCT of patients.

Supplemental Table S3. Rare homozygous variants present in siblings P1 and P2, and in unrelated P3.

Supplemental Table S4. *TET2* variants present in our patients previously reported as somatic variants in CMM and AML patient cohort.

Supplemental Table S5. Somatic missense mutations in RAS signalling pathway-related genes *KRAS* (P1) and *ERBIN* (P2), identified by pairwise comparison of lymphoma tissue and germline WES.

Supplemental Table S6. Sequence of primers used in the study.

Supplemental Figure S1. Peripheral blood cell counts in patients before and after hematopoietic stem cell transplantation.

Supplemental Figure S2. Histopathology of lymphoid tumors and other significant pathology in patients.

Supplemental Figure S3. Clinical timelines.

Supplemental Figure S4. The effect of loss of TET2 function on DNA methylation status.

Supplemental Figure S5. B-cell differentiation assay *in vitro*.

Supplemental Figure S6. Characterisation of derived iPSC from patient P1 and P2 and healthy individuals.

Supplemental Figure S7. Hematopoietic differentiation of patients and control iPSC *in vitro*.

SUPPLEMENTAL MATERIALS AND METHODS

Patient cohort

The parents of all patients and healthy donor volunteers provided written informed consent to participate in research protocols approved by the local Research Ethics Committee.

DNA extraction

Genomic DNA was extracted from cells using QIAamp DNA Micro Kit, DNA Blood Mini Kit, and from tissue using Qiagen AllPrep DNA/RNA/miRNA Kit (Qiagen, 51104, 56304, and 80224, respectively) following the instructions of the manufacturer. The purity and concentration of DNA was determined by NanoDrop 2000 spectrophotometer and Qubit dsDNA HS and dsDNA BR Assay Kits (Q32854, Q32850) with the Qubit Fluorometer (Thermo Fisher Scientific).

Whole Exome Sequencing

Patient 1 and 2 genomic DNA, extracted from the early passage primary dermal fibroblasts and lymphoma tissue, was submitted for whole exome sequencing using Nextera Rapid Capture Exomes kit (Illumina) coupled with massively parallel sequencing by the Illumina NovaSeq Sequencing system. The DNA sequences were mapped to the hg19 human genome by NovoAlign (<http://novocraft.com/main>). In parallel, homozygosity mapping was performed using the Affymetrix Genome-Wide Human SNP 5.0 microarray. Homozygous regions were identified using Homozygosity Mapper (<http://homozygositymapper.org>) and further analysed using microsatellite markers.

DNA was purified from whole blood of P3 and prepared using SureSelectXT with All Exon v6 capture library and sequenced on Illumina HiSeq 3000 for 2 × 150-bp paired-end sequencing. Reads were aligned with BWA-MEM to GRCh37/hg19 and variant calling was performed according to GATK-best practices.

Sanger Sequencing

Amplification of genomic DNA for Sanger sequencing was performed by the standard PCR method. PCR clean-up was performed with ExoSAP-IT (Affymetrix). Sanger sequencing was then performed using the same primers, primer sequences are listed in **Supplementary Table S6**. Sanger sequencing using BigDye Terminator Cycle Sequencing Kit, version 3.1 (Applied Biosystems) and analysis on an ABI 3130XL DNA analyzer (Applied Biosystems).

Western Blot for TET2 protein expression

Cell pellets were lysed in RIPA buffer (50mM Tris HCL pH 7.5, 150mM Sodium Chloride, 0.5% Sodium Deoxycholate, 0.1% Sodium Dodecyl Cholate (SDS), 1% NP-40, 1x complete protease inhibitor). Protein concentration was determined using Pierce Bicinchoninic Acid Assay (Thermo Fisher Scientific) according manufacturer's instructions. Cell lysates were denatured at 90°C for 10 minutes, equal amount of lysates were loaded and subjected to NuPAGE 3-8% Tris-Acetate Polyacrylamide protein gel electrophoresis in 1x NuPAGE Tris-Acetate SDS Running Buffer. Proteins were transferred to a Immobilon-P polyvinylidene fluoride membrane (PVDF, 0.45um, Millipore Merck) in 1x NuPAGE Novex Transfer Buffer (Thermo Fisher Scientific) and subsequently blocked with 5% bovine serum albumin (BSA, Tocris) in Tris Buffered Saline/0.1% Tween 20 (TBS/T) for 1hour at room temperature (RT), followed by incubation with anti-human primary antibodies: mouse TET2 (Active Motif, 61389, clone 21F11, 1:1000) and rabbit GAPDH (Cell Signaling Technology, 5174, clone D16H11, 1:2000) overnight at 4°C. The blots were then washed three times with TBS/T and incubated with appropriate Horseradish Peroxidase (HRP)-conjugated secondary antibodies: anti-mouse (Cell Signaling Technology, 7076, 1:5000) and anti-rabbit (Cell Signaling Technology, 7074S, 1:2000) in 5% BSA in TBS/T for 1 hour at RT. The blots were developed with Immobilon ECL Ultra Western HRP substrate (Millipore Merck) according to the manufacturer's instruction. The Chemiluminescent images were captured on a G:BOX Chemi using GeneSnap and Image Studio Lite Software (Syngene, LI-COR Biosciences).

TET2 loss-of-function *in vitro* assay

Wild type TET2 plasmid (a kind gift of Skirmantas Kriaucionis) was mutated by QuikChange II XL Site-Directed Mutagenesis kit and transformation into XL10-Gold ultracompetent cells. Purification

of plasmid DNA was performed using the PureYieldTM Plasmid Midiprep system (Promega) according to the manufacturer's protocol. HEK293T-cells grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% Glutamax (all purchased from Thermo Fisher Scientific), without antibiotics were transfected with wild type TET2 or mutant TET2 p.H1382R plasmids by Fugene HD transfection reagent (Promega). The immunofluorescence staining with 5hmC and Flag M2 antibodies was performed as described below in section Immunofluorescence staining and microscopy.

DNA Methylation

Purified DNA was quantified with the Qubit dsDNA BR Assay Kit and normalized. In duplicate, three aliquots of DNA were prepared for each sample according to the EpiMark analysis kit (E3317S, NEB) manufacturer instructions. One aliquot per sample was treated with T4 β-glucosyltransferase (T4-BGT) (10 units per sample) (M0357S) to convert 5hmC to glucosylated 5hmC (5ghmC) using uridine diphosphate glucose (UPD) (1.24 µl). T4-BGT-dependent glucosylation of 5hmC to form 5ghmC was facilitated by heating at 37°C for 6 hours. After glucosylation, enzyme restriction was performed on all samples. Mspl (R0106S) and Hpall (R0171S) recognize the same DNA sequence ('5 CCGG 3') but are differentially sensitive to methylation status. Mspl cleaves both 5mC and 5hmC. However, Mspl cleavage is blocked by 5ghmC. Hpall cannot cleave modified sites; any modification with 5mC, 5hmC, or 5ghmC at either cytosine will prevent cleavage. All samples were heated at 37°C for 6-12 hours to allow for complete digestion. After digestion the separation and analysis of DNA samples up to greater than 60,000 base pairs was performed using the Agilent Genomic DNA ScreenTape assay (Agilent, 5067- 5365/6) with the Agilent 2200 TapeStation system according to manufacturer specifications, with quantification sensitive down to 5 pg/µl.

High-resolution assay reports for each Genomic DNA ScreenTape were analysed with the use of ImageJ software to plot lane densities¹. DNA runs from high to low molecular weight on the gel from top to bottom; the density plots present this data from left to right. The area under the curve (AUC) at each decile interval was recorded and used for calculation of differences between 5mC and 5hmC concentration. The minimum level of 5hmC genome wide is found when the difference

before and after T4-BGT treatment approaches zero for Mspl-restricted DNA. Similarly, the difference between Mspl and Hpall restriction indicates non-specific methylation.

FasL-induced apoptosis assay by Annexin V/ PI staining

Peripheral blood mononuclear cells (PBMC) were isolated from peripheral blood of affected patients and unaffected controls using Lymphoprep density gradient medium (Stem Cell Technologies) according manufacturer's protocol. PBMC were stimulated for 3 days with 2 μ g/ml Phytohemagglutinin-L (PHA, Sigma-Aldrich Merck, 11249738001) and 20ng/ml IL-2 (Peprotech, 200-02), and subsequently with 20ng/ml IL-2 for another 2 days. Cells were stimulated with a soluble Fas Ligand set (Enzo Lifesciences, ALX-850-014-KI02). Cells were stained with Annexin V-FITC and PI according to manufacturer's protocol (BD Biosciences, 556547) and detected on Canto II flow cytometer and the output data was analysed by FlowJo V10 (Treestar Inc.).

Negative magnetic B-cell isolation

PBMCs were isolated from peripheral blood by density dependent centrifugation with Lymphoprep (Axis Shield). Total B-cells were isolated by negative selection with a memory B-cell isolation kit according manufacturer's instructions (Miltenyi Biotech). Briefly, B-cell Biotin-Antibody Cocktail was added to PBMC pellet in cold MACS buffer and incubated for 20 minutes at 4°C. Subsequently, anti-Biotin Microbeads were added, and incubated for 20-30 minutes at 4°C. Cells were washed by MACS buffer, centrifuged for 10 minutes at 1500rpm at 4°C. The cell pellet was resuspended in cold MACS buffer and applied to the LD column placed in the magnetic field. The unlabelled cells that passed through were collected in a fresh tube. The column was washed twice with cold MACS buffer and the effluent collected in the same tube. The cells were counted and resuspended in Iscove's Modified Dulbecco's Medium (IMDM, Thermo Fisher Scientific) + 10% FBS at 5 x 10⁵ cells/ml.

Preparation of CD40L-L cells

Murine fibroblasts transfected with human CD40L (CD40L-L cells) were irradiated in advance at 50Gy for 50 minutes in a gamma irradiator (NDT) and stored at -80°C. CD40L-L cells were

resuspended in 12ml fresh IMDM + 10% FBS (1×10^6 /plate) and incubated overnight at 37°C + 5% CO₂.

B-cell differentiation assay *in vitro*

B-cells were cultured with IMDM + 10% FBS supplemented with Lipid Mixture 1, chemically defined (Sigma-Aldrich Merck, L0288) and MEM amino acids solution (Sigma-Aldrich Merck, M7145). *Day 0*: B-cells were cultured in 24-well plates at 2.5×10^5 cells/ml in IMDM + 10% FBS + supplements with the addition of 20 U/ml IL-2 (Roche, 11147528001), 50 ng/ml IL-21 (Peprotech, 500-P191). B-cells were stimulated by gamma-irradiated CD40L-L cells and 2 µg/ml F(ab')2 goat anti-human IgM and IgG (Jackson ImmunoResearch, 109-006-127). *Day 3*: B-cells were removed from the CD40L-L cells by gentle pipette mixing and reseeded at 1×10^5 /ml in IMDM + 10% FBS + supplements with the addition of 20 U/ml IL-2 and 50 ng/ml IL-21. *Day 6*: For long-term culture, cells were seeded at 5×10^5 /ml in IMDM + 10% FBS + supplements with the addition of 10 ng/ml IL-6 (Peprotech, 200-06), 50 ng/ml IL-21, 100 U/ml IFN-α (Sigma-Aldrich Merck, SRP4596). *Day 13 onwards*: IL-21 was withdrawn from the media + supplements + IL-6 + IFN-α. Cells were re-fed weekly by replacing half of the volume of media with fresh media + supplements + day 13 cytokines.

Enzyme-linked immunosorbent assay (ELISA) for immunoglobulin IgM and IgG detection

ELISAs were performed with Human IgM ELISA Quantitation Set (E80-100) or Human IgM ELISA Quantitation Set (E80-104) (Bethyl Laboratories Inc) according to the manufacturer's instructions. Briefly, 96-well plates (Corning) were coated with 100µl of affinity purified antibody at 1:100 in coating buffer (0.05M bicarbonate) for 1 hour at RT. Plates were washed four times with ELISA wash buffer (0.5M Tris, 1.4M NaCl, 0.5% Tween) before blocking buffer was added (0.5M Tris, 1.4M NaCl, 1% BSA) and the plates blocked overnight at 4°C. Supernatant collected from the B-cell differentiations was serially diluted 1:20, 1:200, 1:400 in IMDM + 10% FBS. Wells were incubated with 100µl 1:75,000 HRP anti-IgG or anti-IgM detection antibody in wash buffer + 1% BSA for 1 hour at RT. Plates were washed four times followed by development with 100µl TMB

and quenched with 0.18M H₂SO₄. ELISA absorbance values were analysed at 450 nm and Ig concentrations calculated from standard curves.

iPSC generation and culture

Primary fibroblasts established from Patient 1 and Patient 2, two adult healthy volunteers and one neonate were cultured in DMEM supplemented with 10% FBS, 100000 U/l penicillin and 100 mg/l streptomycin (1% Penicillin/Streptomycin) and 1% Glutamax (Thermo Fisher Scientific). Fibroblasts were transduced with CytoTune™ iPSC 2.0 Sendai Reprogramming kit including polycistronic hKlf4–hOct3/4–hSox2, hcMyc and hKlf4 according manufacturer's instructions, growing on inactivated mouse embryonic fibroblasts for 2-3 weeks (Thermo Fisher Scientific). Derived clones were adapted to feeder-free conditions, i.e. were cultured on Matrigel-coated plates (Corning) in mTESR1 culture medium (Stem Cell Technologies) supplemented with 1% Penicillin/Streptomycin in 5% CO₂ atmosphere at 37°C with daily change of medium. The cells were passaged with Versene® (EDTA) 0.02% (Lonza) every 4 to 5 days.

Gene expression by PCR and quantitative RT-PCR

The clearance of Sendai virus vectors and endogenous expression of pluripotent markers were validated by vector- and marker-specific primers, respectively. RNA was extracted by ReliaPrep™ RNA Miniprep systems (Promega), followed by reverse transcription using GoScript™ Reverse Transcription System (Promega) according manufacturer's instructions. qPCR was provided by GoTaq® Green Master Mix and product was detected by standard 2% agarose gel electrophoresis. qRT-PCR reaction was run on the Applied Biosystems® QuantStudio™ 7 Flex Real-Time PCR System (Thermo Fisher Scientific) with the GoTaq qPCR Master Mix (Promega). The data were analysed using the QuantStudio™ software (Thermo Fisher Scientific) and relative gene expression was determined using the 2^{-ΔΔCt} method using *SDHA* and *UBE4A* as a housekeeping genes. Primers sequences are listed in **Supplementary Table S6**.

Immunofluorescence staining and microscopy

Cells were fixed with 4% paraformaldehyde for 15min at RT, washed with PBS, and permeabilized with 0.2% Triton X-100/PBS for 10min at RT, blocked in 10% FBS and 1% BSA in PBS for min 1hour at RT, followed by overnight staining at 4°C with primary anti-human antibodies: mouse TRA1-60 (BD Biosciences, 560071, clone TRA-1-60, 1:100), mouse Oct3/4 (BD Biosciences, 611202, clone 40/Oct-3, 1:100) and mouse SSEA4-AF555 (BD Biosciences, 560218, clone MC813-70, 1:60). Cells were washed and stained with secondary antibody rabbit anti-mouse TRITC (Sigma-Aldrich Merck, T2402, 1:200) for 1hour at RT in the dark.

For 5mC and 5hmC staining of TET2 transfected cells, after the cell permeabilisation step, an acid denaturalization step with 2N HCl for 30min at RT was added, followed by neutralisation with 100mM Tris-HCl buffer (pH 8.0) for 10 min at RT to unmask the epitope. Cells were stained with primary mouse 5mC (Eurogentec, BI-MECY-0100, clone 33D3, 1:1000), rabbit 5hmC (Active Motif, 39769, 1:2000) and mouse anti-Flag (Sigma-Aldrich Merck, F3165, clone M2, 1:500) overnight at 4°C. Secondary antibodies goat anti-mouse FITC (Sigma-Aldrich Merck, F2012, 1:1000), goat anti-rabbit Cy3 (Jackson, 111-165-144, 1:1000), goat anti-mouse AF647 (Invitrogen, A21237, 1:500) and goat anti-rabbit AF488 (Invitrogen, A11070, clone, 1:500) were used for 1hour at RT.

Nuclei were counterstained with DAPI (CyStain® DNA, Sysmex Partec) and cells were mounted with Vectashield mounting medium (Vector Laboratories, USA). Images were taken with Zeiss Axiovert 200M Inverted Widefield microscope with Zeiss AxioCam HRm camera and analysed with AxioVision software.

In vitro differentiation of iPSC into three germ layers

To access the pluripotent potential of derived iPSC to form three germ layers, the cells were allowed to spontaneously differentiate in DMEM/F12, 20% FBS supplemented with 1% Penicillin/Streptomycin solution, 1% non-essential amino acid solution (all from Thermo Fisher Scientific) for 2 weeks in 5% CO₂ atmosphere at 37°C. Differentiated cells were stained with 3-Germ Layer Immunocytochemistry kit containing beta-III tubulin (ectoderm), smooth muscle actin (mesoderm) and alpha-fetoprotein (endoderm) according manufacturer's instructions (Thermo Fisher Scientific, A25538).

Teratoma formation *in vivo*

For *in vivo* teratoma assay 500,000 iPSC cells in 100ul PBS mixed with 100ul Matrigel (Corning) were injected into right flank of male adult NOD/SCID mice. After 8-14 weeks, mice were sacrificed and teratoma tissues were dissected. Samples were fixed with 4% paraformaldehyde at 4°C for 14hrs and subsequently dehydrated with 70%, 80%, 90% and 95% Ethanol for 2 hours, respectively. Tissues were processed, sectioned according to standard procedures and counterstained with either Hematoxylin and Eosin or Weigert's stain. Sections were examined using bright field microscopy. All procedures were carried out in accordance with institutional animal guidelines and permission.

Karyotype

Karyotypes were determined by standard G-banding procedure. At least 20 metaphases were analysed for each sample.

Hematopoietic differentiation of iPSC

For hematopoietic differentiation iPSC were adapted to culture on Vitronectin-coated plates in StemPro-hESC medium supplemented with 1% Penicillin/Streptomycin and passaged mechanically by StemPro EZPassage tool (Thermo Fisher Scientific). Hematopoietic differentiation was achieved using a protocol slightly modified from Olivier et al 2016². Briefly, cells were manually passaged using passaging tool and plated on low-attachment plates to form embryoid bodies in Stemline HSC expansion medium (Sigma-Aldrich, Merck) supplemented with cytokines, such as BMP4, VEGF, Wnt-3a (R&D Systems), Activin A (Peprotech) and GSK-3β Inhibitor VIII (Calbiochem, Merck). On day 2 of differentiation FGF-alpha (Perotech), SCF (Thermo Fisher Scientific) and beta-Estradiol (Sigma-Aldrich, Merck) were added. On day 3, the embryoid bodies were dissociated using TrypLE Express (Thermo Fisher Scientific) to single cells for 10min at 37°C and plated into a standard cell culture treated 12-well plate dish at a density of 2×10^5 cells per well in Stemline medium supplemented with BMP4, VEGF, FGF-alpha, SCF, beta-Estradiol, IGF-II and TPO (both Peprotech), Heparin and IBMX (both Sigma-Aldrich, Merck) and cultured in 5%

CO₂ atmosphere at 37°C. Fresh cytokines were added every 2-3 days and complete change of medium was performed every 5-6 days.

Flow Cytometry analysis

For pluripotency marker expression iPSC were collected using Accutase (Thermo Fisher Scientific) for 5min at 37°C, fixed with 4% paraformaldehyde for 10min at 37°C and permeabilized with cold Methanol for 30min at -20°C. Cells were stained with mouse anti-human antibodies: TRA1-60-AF555 (BD Biosciences, 560121, clone TRA-1-60, 1:20), Oct3/4-AF488 (BD Biosciences, 560217, clone 40/Oct-3, 1:20), SSEA-4-PerCpCy5.5 (BD Biosciences, 561565, clone MC813-70, 1:20) and rabbit Nanog-AF647 (Cell Signaling Technology, 5448S, clone D73G4, 1:100) for at least 40min in 2% FBS in PBS at RT in the dark.

For hematopoietic differentiation markers, cells were collected using TrypLE Express (Thermo Fisher Scientific) and stained with directly conjugated mouse anti-human antibodies: KDR-PE (BD Biosciences, 560494, clone 89106, 1:5), CD34-APC (BD Biosciences, 555824, clone 581, 1:5), CD43-FITC (Life Technologies, MHCD4301, clone L10, 1:20), CD41a-APC-H7 (BD Biosciences, 561422, clone HIP8, 1:20), CD235a-BV421 (BD Biosciences, 562938, clone GA-R2, 1:20), CD45-PerCP-Cy5.5 (BD Biosciences, 564105, clone HI30, 1:20) and DAPI for nuclear staining (Thermo Fisher Scientific).

For B-cell differentiation assay, dead cells were excluded by staining with 7AAD (Biolegend, 420404, 1:20). The cells were stained with: CD19-PE (Miltenyi Biotec, 130-091-247, clone HIB19, 1:50), CD20-e450 (Thermo Fisher Scientific, 48-0209-42, clone 2H7, 1:50), CD27-FITC (BD Biosciences, 555440, clone M-T271, 1:40), CD38-PE-Cy7 (BD Biosciences, 335825, clone HB7, 1:200), CD138-APC (Miltenyi Biotec, 130-091-250, clone 44F9, 1:50). Cell counts were determined with CountBright Absolute Counting Beads (Thermo Fisher Scientific, C36950).

For immunophenotyping of PBMC two panels of directly-conjugated mouse anti-human antibodies were used. For T-cells and NK-cells panel: TCR gamma/delta-FITC (Biolegend, 331208, clone B1, 1:20), CD127-PerCP-Cy5.5 (Biolegend, 351322, clone A019D5, 1:20), CCR6-PE (BD Biosciences, 559562, clone 11A9, 1:50), CD56-PE-CF594 (BD Biosciences, 564849, clone NCAM16.2, 1:100), CCR7-PE-Cy7 (Biolegend, 353226, clone G043H7, 1:33), CXCR5-AF647 (Biolegend, 356906,

clone J252D4, 1:10), PD1-AF700 (Biolegend, 329952, clone EH12.2H7, 1:33), CD45RO-APC-H7 (BD Biosciences, 561137, clone UCHL1, 1:50), CXCR3-BV421 (BD Biosciences, 562558, clone LS177-1C6, 1:33), CD57-BV605 (Biolegend, 393304, clone QA17A04, 1:100), CD16-BV650 (BD Biosciences, 563692, clone 3G8, 1:100), CD95-BV711 (BD Biosciences, 563132, clone DX2, 1:33), CD4-BV785 (Biolegend, 344642, clone SK3, 1:100), CD3-BUV396 (BD Biosciences, 563546, clone UCYT1, 1:33), CD8-BUV496 (BD Biosciences, 564804, clone RPA-T8, 1:100), CD25-BUV737 (BD Biosciences, 564385, clone 2A3, 1:100). For B-cells and Monocytes panel: CD27-FITC (BD Biosciences, 555440, clone M-T271, 1:50), CD38-PerCP-eF710 (eBiosciences, 46-0388-42, clone HB7, 1:33), IgA-PE (MACS Miltenyi Biotec, 130-099-108, clone IS11-8E10, 1:50), CD11c-PE-CF594 (Biolegend, 337228, clone Bu15, 1:100), CD21-PE-Cy7 (Biolegend, 354912, clone Bu32, 1:100), CD19-APC (BD Biosciences, 55415, clone HIB19, 1:10), IgD-AF700 (Biolegend, 348230, clone IA6-2, 1:10), CD5-APC-Cy7 (Biolegend, 302340, clone 2H7, 1:100), CD138-BV421 (Biolegend, 356516, clone MI15, 1:20), CD20-BV510 (Biolegend, 302340, clone 2H7, 1:100), CD16-BV650 (BD Biosciences, 563692, clone 3G8, 1:100), HLA-DR-BV711 (Biolegend, 307644, clone L243, 1:100), IgG-BV786 (BD Biosciences, 564230, clone G18-145, 1:100), CD123-BUV395 (BD Biosciences, 564195, clone 7G3, 1:20), CD14-BUV737 (BD Biosciences, 612764, clone M5E2, 1:33). Dead cells were excluded by 7AAD.

Cells were analysed by BD Fortessa X20 or BD Symphony flow cytometer, data were evaluated with FACSDiva (BD Biosciences), FlowJo V10 software (Tree Star Inc) or FCS Express 7 (DeNovoSoftware).

Hematopoietic Colony Forming Unit Assay

Twenty thousand unsorted cells from hematopoietic differentiation were plated into semi-solid methylcellulose medium MethoCult H4434 (Stem Cell Technologies) in 10mm gridded dishes in duplicate. After 12-14 days of incubation in 5% CO₂ at 37°C, the colonies were counted and scored according to their morphology under a light microscope.

Statistical analysis

The data are shown as mean \pm SD of at least two independent experiments. Nonparametric One-Way ANOVA Kruskal-Wallis with Dunn's multiple comparisons test, two-tailed t-test, and unpaired t-test were used to calculate statistical significance with $p<0.05$ using GraphPad Prism 7.02 software.

Supplemental Table S1. Immunoglobulin levels of patients before transplantation.

[g/l]	Patient 1 age 3.5 years	Patient 2 age 1 month 3 months*			Patient 3	
		age 2.5 years	age 10.4 years*			
IgG	22.40 (4.9 - 16.1)	<0.50	13.2* (2.4 - 8.8)	13.50 (4.9 - 16.1)	9.9* (5.4 - 16.1)	
IgA	1.24 (0.4 - 2.0)	<0.06	<0.25 (0.1 - 0.5)	<0.06 (0.4 - 2.0)	<0.06 (0.8 - 2.8)	
IgM	1.16 (0.5 - 2.0)	0.70	0.28 (0.2 - 1.0)	0.24 (0.5 - 2.0)	0.22 (0.5 - 1.9)	
IgG1	20.78 (3.5 - 9.4)			13.40 (3.2 - 9.0)		
IgG2	2.47 (0.6 - 3.0)			0.17 (0.5 - 2.8)		
IgG3	2.76 (0.1 - 1.3)			2.39 (0.1 - 1.2)		

Age matched reference values are displayed in brackets. Bold values: abnormal laboratory values, asterisk: for patient P2 and P3 the values were taken on immunoglobulin supplementation.

Supplemental Table S2. Details of HSCT of patients.

	P1			P2			P3		
Conditioning	Alemtuzumab 1mg/kg D-14 to D-10	Melphalan 140mg/m ² D-2	Fludarabine 150mg/m ² D-7 to D-3	Alemtuzumab 1mg/kg D-8 to D-4	Treosulfan 42mg/m ² D-6 to -4	Fludarabine 150mg/m ² D-6 to D-2	Alemtuzumab 1mg/kg D-7 to D-3	Treosulfan 42mg/m ² D-6 to -4	Fludarabine 150mg/m ² D-5 to D-1
HSCT	Donor, stem cell source HLA matching Cell doses CD34+ (/Kg) Cell doses CD3+ (/Kg)	sibling red-cell depleted BM single antigen mismatch (11/12) 10.2×10^6 3×10^7	maternal CD3/CD19-depleted PBSC haploidentical (6/12) 2×10^6 1.2×10^5		unrelated donor PBSC single antigen mismatch (11/12) 9×10^6 2.3×10^8				
GVHD prophylaxis	cyclosporin		cyclosporin	MMF		cyclosporin	MMF		
Day of neutrophil engraftment	D+13		D+14			D+13			
Day of recorded T-cell count >200/uL	D+26		D+8			D+38			
First measured T-cell chimerism									
Day/ % Donor T-cells	D+28/ 22%		D+46/ 0%			D+52/ 9%			
Additional serotherapy									
Drug	ATG	Alemtuzumab		ATG					
Dose	4mg/kg	0.9mg/kg							
Day	D-2 to D-1	D+28 to D+30		D+56, 3x5mg/kg					
Top-up graft	sibling red cell-depleted BM 12×10^6 4×10^7								
Cell doses CD34+ (/Kg)									
Cell doses CD3+ (/Kg)									

ATG: anti-thymocyte globulin, HSCT: hematopoietic stem cell transplantation, D: day, BM: bone marrow, PBSC: peripheral blood stem cells, MMF: Mycophenolate mofetil.

Supplemental Table S3. Rare homozygous variants present in siblings P1 and P2, and in unrelated P3.

Patient	Gene	Chromosome	Position	Reference/ Alteration	AminoAcid Change	ExAC (No. homo.)	CADD	Polyphen
P1/P2	TET2	4	106190867	A / G	p.His1382Arg	.	25	1
	MT-CYB	MT	15884	G / A	p.Ala380Thr	MitoMap 0.8%	23	0.494
	FAM186A	12	50745863	C/CCTGCTGAGGGGTGAGAG GGATCCCCAGGGCCTGG	p.Glu1586AlafsTer92	0.0007601 (1)	20	NA
	MT-ND4	MT	10845	C / T	p.Thr29Ile	MitoMap 0.1%	17	0.997
	MT-CYB	MT	14766	C / T	p.Thr7Ile	MitoMap 77%	15	0
	MT-CYB	MT	15326	A / G	p.Thr194Ala	MitoMap 98.7%	3	0
P3	TET2	4	106196561	C / T	p.Gln1632Ter	.	41	NA
	ATP6VOA1	17	40646370	C / T	p.Thr405Met	0.00001647 (0)	34	1
	RSAD2	2	7027078	A / G	p.Asp174Gly	0.0001977 (1)	28	0.838
	OBSCN	1	228524756	T / C	p.Val6487Ala	0.002009 (5)	26	0.547
	SLC25A47	14	100795833	C / T	p.His260Tyr	0.00007413 (1)	25	1
	ANKRD50	4	125592798	C / T	p.Gly545Glu	0.0004201 (0)	24	0.796
	MAPT	17	44067289	C / T	p.Leu410Phe	0.0001318 (0)	24	0.984
	DMD	X	32503077	A / G	p.Val921Ala	0.00002471 (1)	23	0.825
	ZBTB39	12	57397627	T / G	p.Ile359Leu	0.00168 (6)	22	0.079
	WDR20	14	102661345	A / G	p.Asn107Ser	0.00005765 (0)	22	0.164
	SYNPO2	4	119951925	G / T	p.Glu665Asp	0.000008236 (0)	19	0.8
	STAT6	12	57490685	C / T	p.Val768Met	0.0001812 (2)	18	0.079
	KRTAP9-4	17	39406076	C / T	p.Thr35Ile	0.00002397 (0)	18	NA
	RAB32	6	146875636	G / T	p.Lys191Asn	0.0005271 (0)	14	0.008
	INTU	4	128626758	T / A	p.Cys527Ser	0.0003048 (2)	12	0.053
	NOTUM	17	79918775	C / T	p.Gly4Glu	0.001314 (1)	12	0
	OBSCN	1	228471294	C / T	p.Ala3372Val	0.0007762 (2)	8	0.375
	OPRM1	6	154360656	G / A	p.Gly86Ser	0.00002486 (0)	8	0
	GPR137C	14	53020002	A / G	p.Gln46Arg	0.000008604 (0)	8	NA
	DCD	12	55039020	C / CT	p.Ala76SerfsTer21	0.001499 (1)	7	NA
	SNAP47	1	227947138	G / T	p.Ala359Ser	0.0004118 (2)	0	0.004
	TCP10	6	167794659	C / T	p.Gly77Arg	0.000006736 (0)	0	0.001

NA – unavailable

Supplemental Table S4. *TET2* variants present in our patients previously reported as somatic variants in CMML and AML patient cohort.

Patient	age	variant	cNomen	pNomen	VAF	depth	SNP array	diagnosis
1	48	TET2_4_106190867_A/G	c.4145A>G	p.H1382R	0.426	2947		AML
2	80	TET2_4_106190867_A/G	c.4145A>G	p.H1382R	0.951	2817		AML
3	59	TET2_4_106190867_A/G	c.4145A>G	p.H1382R	0.946	2605		AML
4	83	TET2_4_106190867_A/G	c.4145A>G	p.H1382R	0.201	3742		CMMI
5	65	TET2_4_106196561_C/T	c.4894C>T	p.Q1632*	0.4	1352		CMMI
6	71	TET2_4_106196561_C/T	c.4894C>T	p.Q1632*	0.436	2090		AML
7	81	TET2_4_106196561_C/T	c.4894C>T	p.Q1632*	0.943	1550	LOH 4q	CMMI
8	70	TET2_4_106196561_C/T	c.4894C>T	p.Q1632*	0.879	2062		AML

Instances of *TET2* p.H1382R and p.Q1632* identified in a cohort of 4324 patients with suspected or confirmed myeloid malignancy, investigated as described (Cargo et al, 2019). The cohort included 1221 cases of acute myeloid leukemia (AML) and 285 cases of chronic myelomonocytic leukemia (CMML). VAF: variant allele frequency, LOH: loss of heterozygosity.

Supplemental Table S5. Somatic missense mutations in RAS signalling pathway-related genes *KRAS* (P1) and *ERBIN* (P2), identified by pairwise comparison of lymphoma tissue and germline WES.

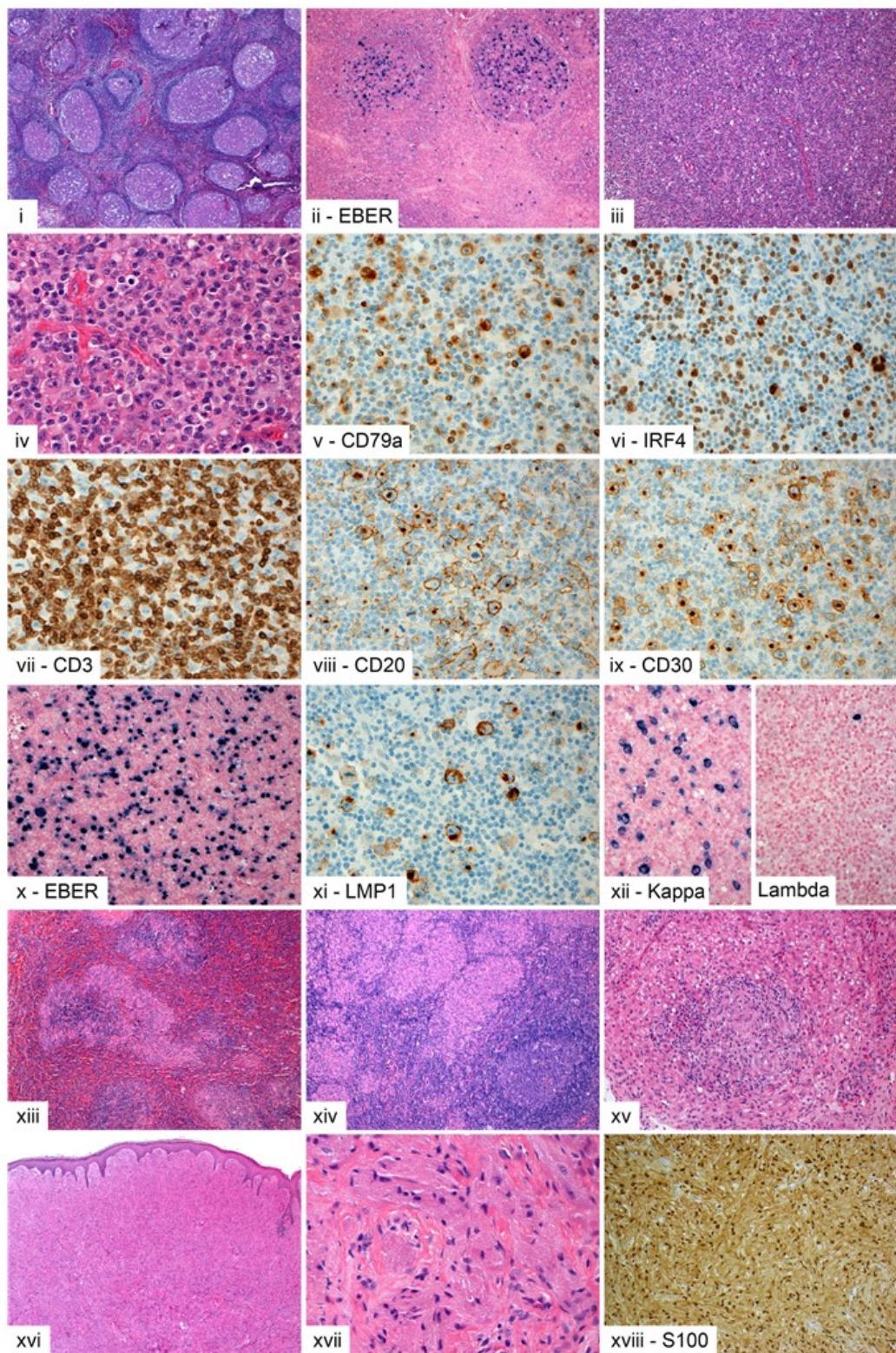
KRAS_p.K117N, c.351A>T rs770248150				ERBIN_p.R1194H, c.3581G>A rs760950077		
Prediction Programme	Score	Cut-off	Prediction	Score	Cut-off	Prediction
CADD	21.2	23	Low impact	15.92	N/A	N/A
PolyPhen	0.998	0.079	Probably damaging	0	N/A	Benign
SIFT	0.011	0.073	Damaging	0.251	0.05	Tolerated
PROVEAN	-4.56	-2.5	Deleterious	1.8	-2.5	Neutral
Mutation Taster	N/A	N/A	Disease-causing	N/A	N/A	Polymorphism
PON-P2	0.728	N/A	Unknown	0.054	N/A	Neutral
	Alleles	Frequency	Homozygous	Alleles	Frequency	Homozygous
ExAc	0	0	0	2	N/A	0
100G	0	0	0	0	0	0
Gnomad	1	3.98E-06	0	3	1.20E-05	0
Associated phenotypes	Colorectal neoplasms Hepatocellular carcinoma Malignant melanoma Multiple myeloma Carcinoma of oesophagus Adenocarcinoma of stomach			No associated phenotypes		
Variants affecting the same codon:	COSM6854421 - large intestine tumour COSM4696721 - large intestine tumour, genital tract tumour					

Supplemental Table S6. Sequence of primers used in the study.

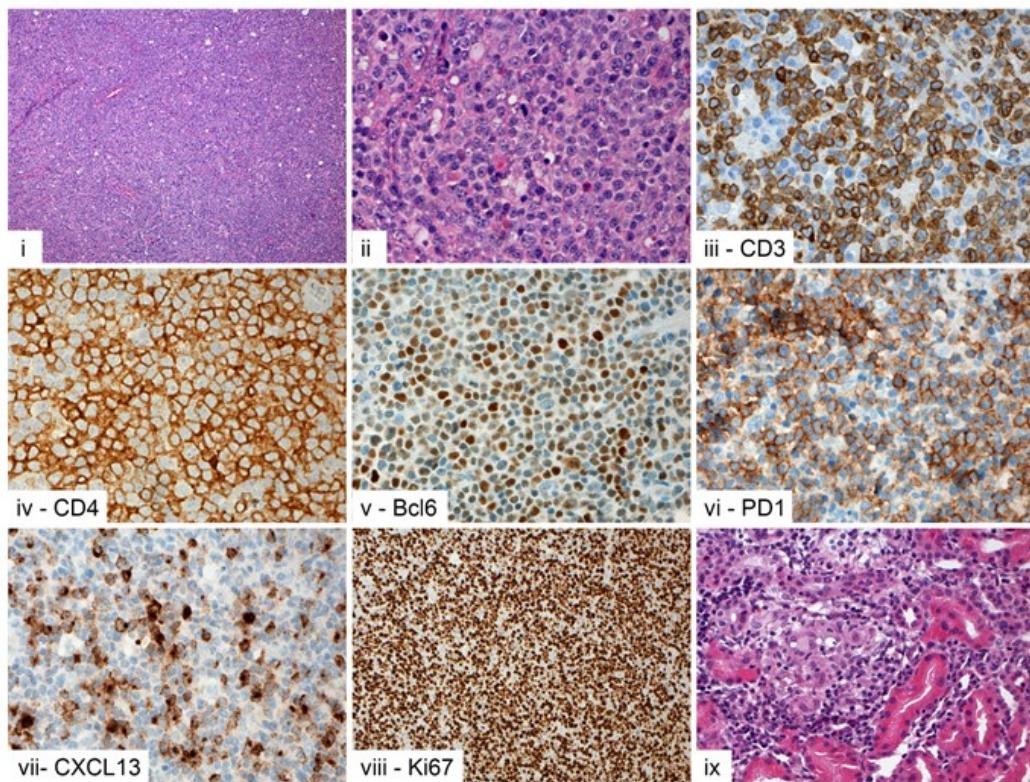
qPCR Primer	Sequence
SeV_forward	GGATCACTAGGTGATATCGAGC
SeV_reverse	ACCAGACAAGAGTTAACAGAGATATGTATC
KOS_forward	ATGCACCGCTACGACGTGAGCGC
KOS_reverse	ACCTTGACAATCCTGATGTGG
Klf4_forward	TTCCTGCATGCCAGAGGGAGCCC
Klf4_reverse	AATGTATCGAAGGTGCTCAA
cMyc_forward	TAACTGACTAGCAGGCTTGTG
cMyc_reverse	TCCACATACAGTCCTGGATGATGATG
Nanog_forward	GAATCTCACCTATGCCTGTG
Nanog_reverse	ATCAGGGCTGTCCTGAATAA
Oct3/4_forward	AGAAACCCCTCGTGCAGGCC
Oct3/4_reverse	CATAGTCGCTGCTTGTGATCGC
qRT-PCR Primer	
SDHA_forward	TGGGAACAAGAGGGCATCTG
SDHA_reverse	CCACCACTGCATCAAATTCTATG
UBE4A_forward	GGATGGACGTTCCATTCCCC
UBE4A_reverse	AGGTCTGCAAGAGACTTGATT
TET1_forward	TCTGTTGTTGTCCTCTGGA
TET1_reverse	GCCTTAAAACTTGGCTTC
TET2_forward	AAAGATGAAGGTCCTTTTATACCC
TET2_reverse	ATAGCTTACCCCTCTGTCCAAC
TET3_forward	CACTCCGGAGAAGATCAAGC
TET3_reverse	GGACAATCCACCCCTCAGAG
Sanger Sequencing	
TET2_forward	CTTCGCATTACACACACTT
TET2_reverse	GAGTTCCCTGCACATGTTC
KRAS_forward	TTGGGTGAGTGGAAACTAGGAA
KRAS_reverse	ACTGGATTAAGAAGCAATGCCCT
ERBIN_forward	ACCCCTTCTGCACGAACAT
ERBIN_reverse	CCCAGTACCAAGAGAAAAACACC

Supplemental Figure S1. Histopathology of lymphoid tumors and other significant pathology in patients. **(A) P1:** i-ii) Lymph node biopsy showing EBV-associated follicular hyperplasia: i, H&E (x40); ii, EBV EBER (x100). iii-xii) Lymph node biopsy showing EBV-positive polymorphic B-cell lymphoproliferative syndrome: iii, H&E (x100); iv, H&E (x600); v, CD79a (x400); vi, IRF4 (x400); vii, CD3 (x400); viii, CD20 (x400); ix, CD30 (x400); x, EBV EBER (x200); xi, EBV LMP1 (x400); xii, Kappa/Lambda immunoglobulin light chains (x400). xiii-xv) Spleen, lymph node and liver showing granulomatous inflammation: xiii, H&E spleen (x100); xiv, H&E lymph node biopsy (x100); H&E liver biopsy (x200). xvi-xvii) Scrotal skin showing granular cell tumor: xvi, H&E (x40); xvii, H&E (x400); xviii, S100 (x200). **(B) P2:** i-viii) Lymph node biopsy showing nodal peripheral T-cell lymphoma with T follicular helper phenotype: i, H&E (x100); ii, H&E (x600); iii, CD3 (x600); iv, CD4 (x600); v, Bcl6 (x600); vi, PD1 (x600); vii, CXCL13 (x600), viii, Ki67 (x200). ix) Renal biopsy showing granulomatous inflammation: H&E (x400). **(C) P3:** i-ii) Lymph node biopsy showing EBV-associated follicular hyperplasia: i, H&E (x40); ii, EBV EBER (x100). iii) Skin showing a cellular dermatofibroma: H&E (x40). iv) Skin showing a pilomatrixoma: H&E (x20). v-xii) Mediastinal mass biopsy showing primary mediastinal large B-cell lymphoma: v, H&E (x200); vi, H&E (x600); vii, CD79a (x400); viii, Bcl6 (x400); ix, CD10 (x400); x, IRF4 (x400); xi, CD30 (X400); xii, Ki67 (x200).

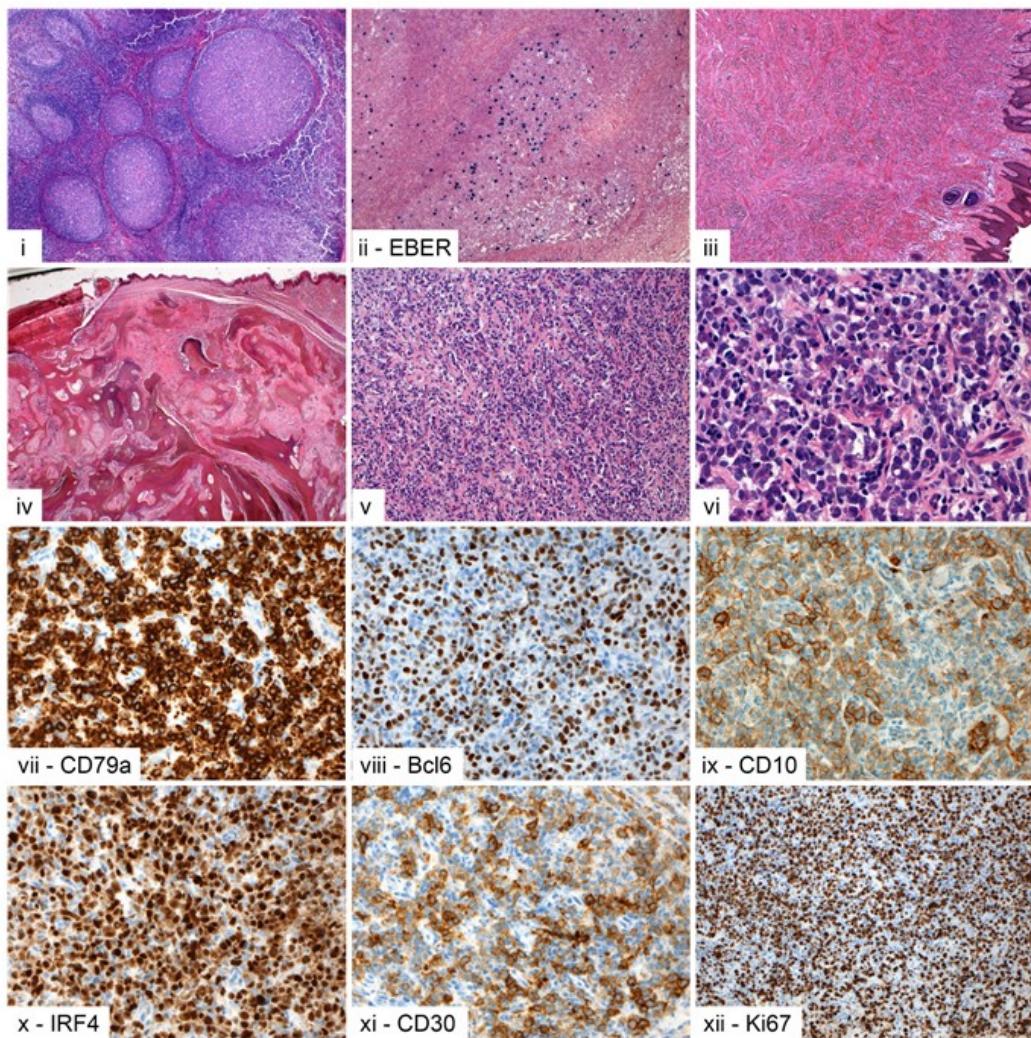
A



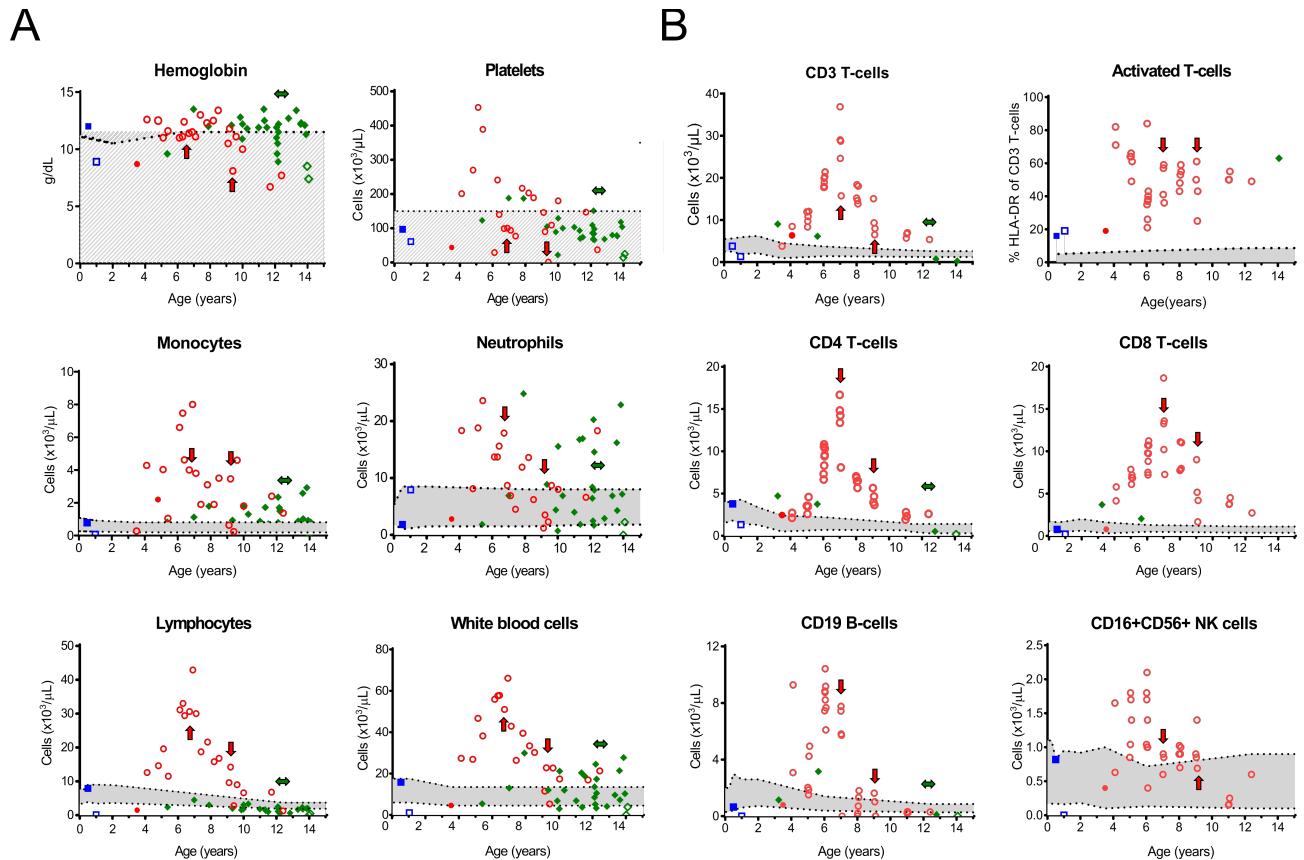
B



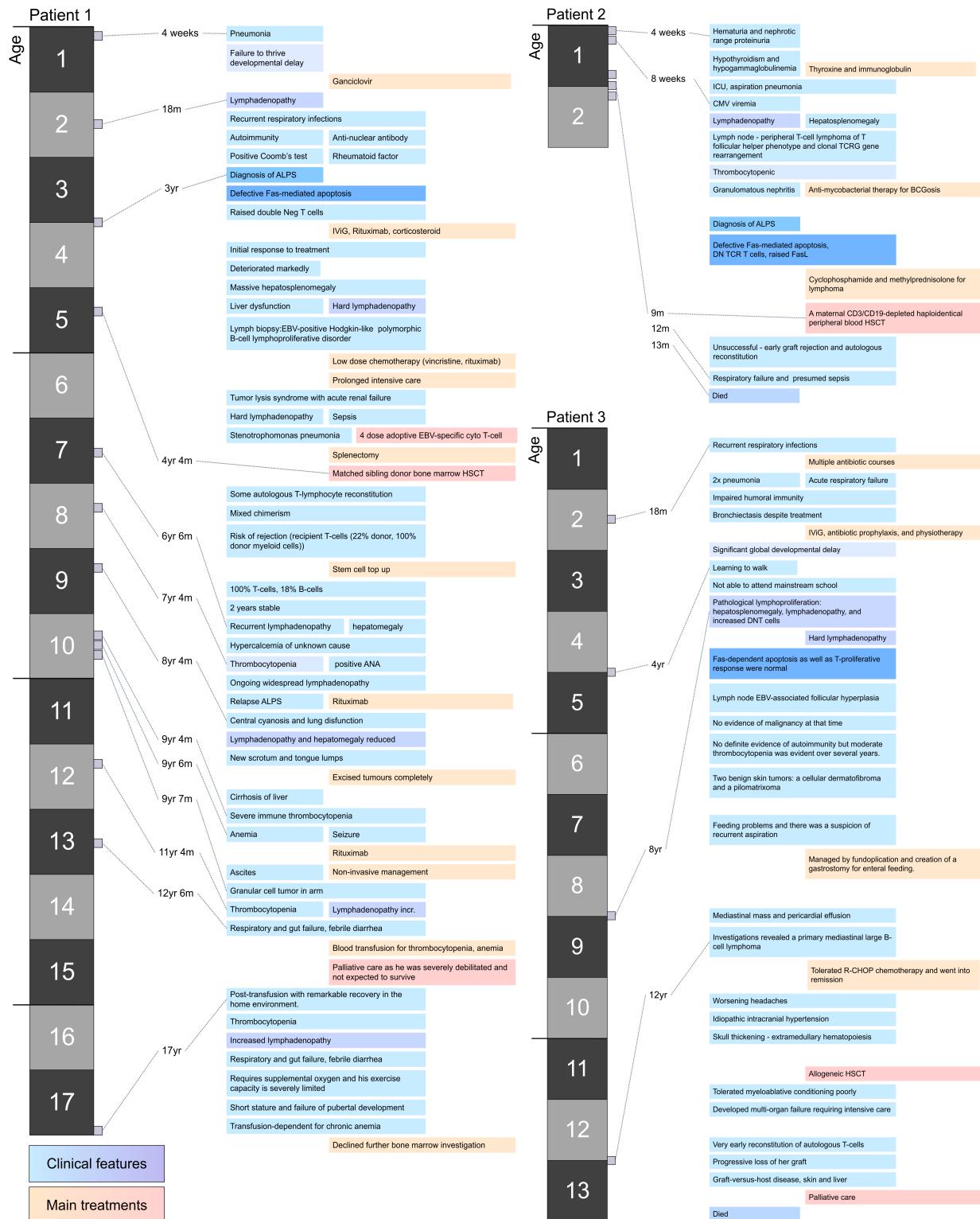
C



Supplemental Figure S2. Peripheral blood cell counts in patients before and after hematopoietic stem cell transplantation. (A) Hemoglobin and total blood cell counts, and (B) absolute numbers of lymphocyte subsets, B-cells, NK-cells and percentages of activated T-cells in peripheral blood of all patients. Hatched area: sub-normal range; grey area: normal range. Red arrow: treatment with Rituximab (R) in patient P1 (age 9 years), green double-ended arrow: R-CHOP in patient P3 (at age 12-12.5 years), red circles: P1; blue squares: P2; green diamonds: P3. Filled symbols: pre-transplantation; open symbols: post-transplantation.

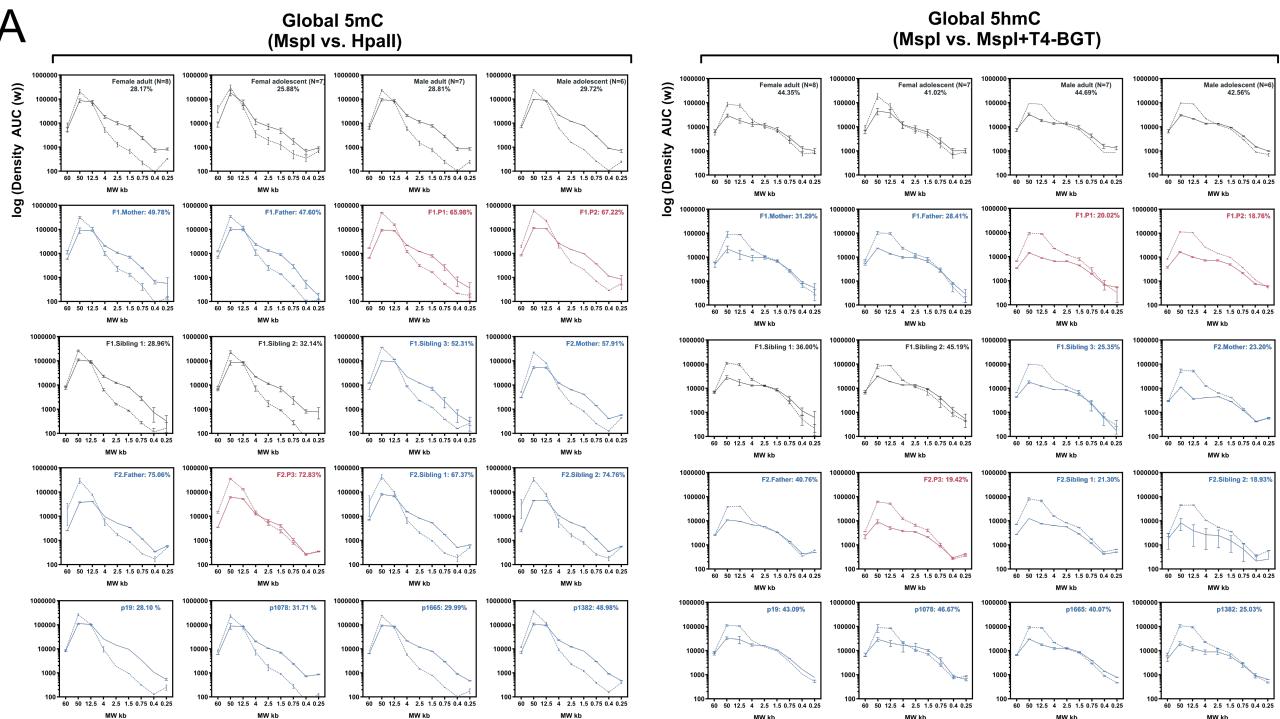
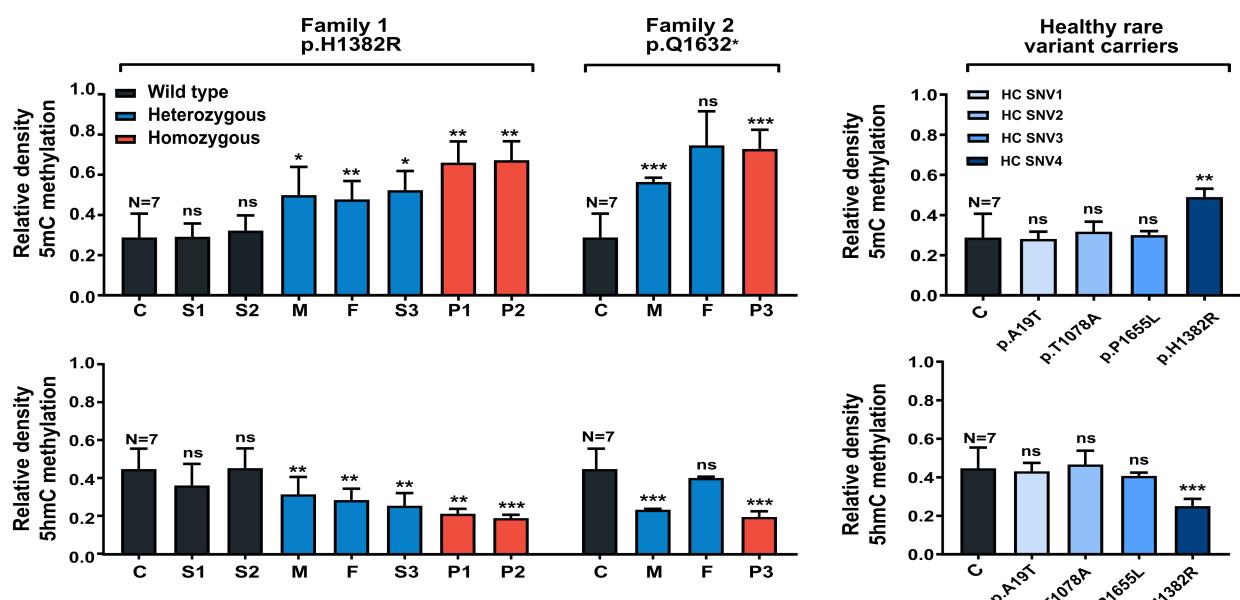
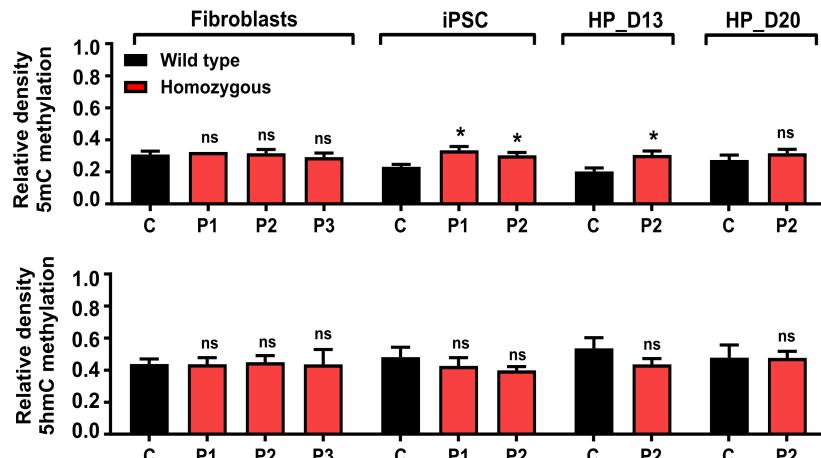


Supplemental Figure S3: Clinical timelines. Blue tone colours are indicative of clinical features, while yellow-red tone boxes indicate treatments. Key events are colour matched between patients 1-3 (on the opposite page). The numbered age boxes represent each year of life. Small purple time-point boxes indicate specifically recorded events or assays.

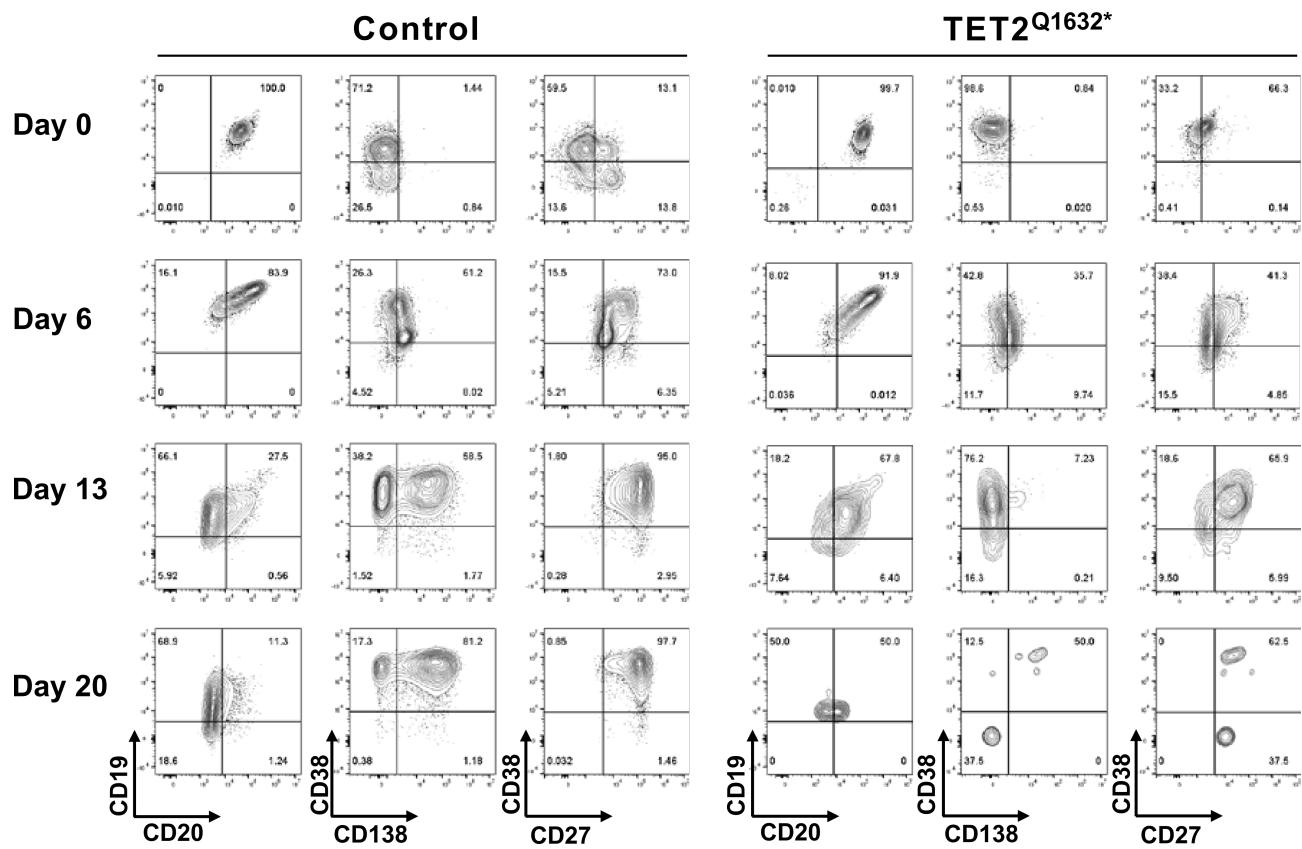


Supplemental Figure S4. The effect of loss of TET2 function on DNA methylation status. (A)

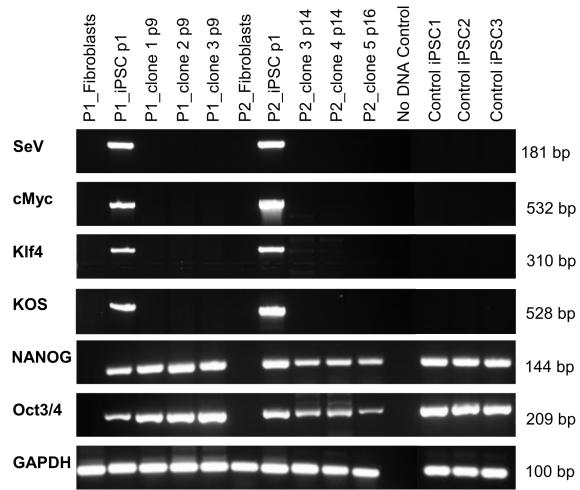
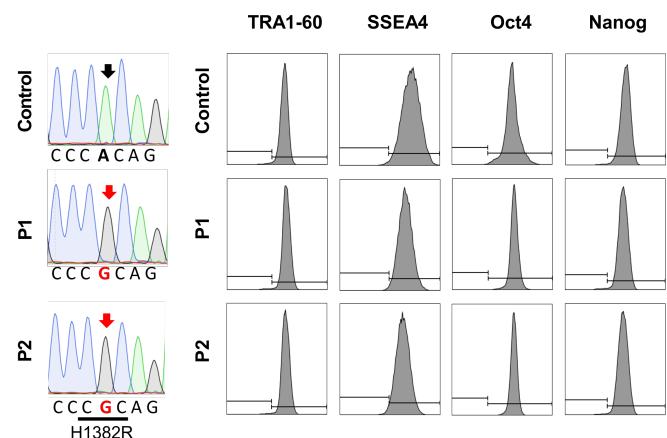
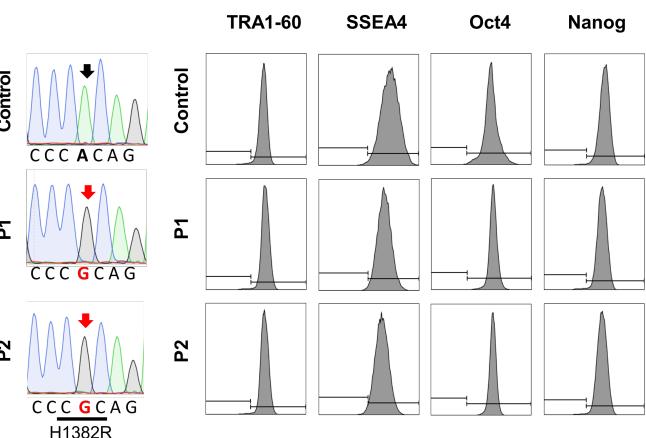
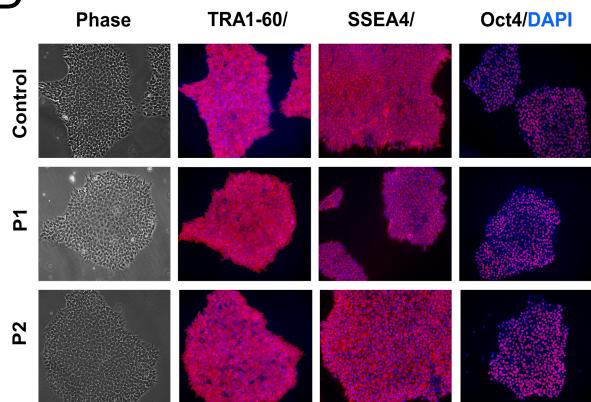
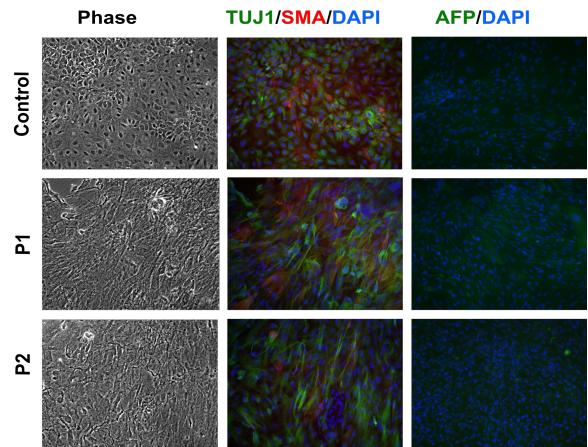
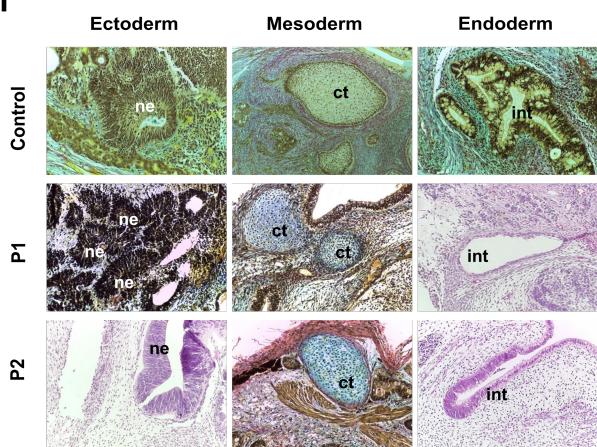
Representative curves of global 5mC and 5hmC methylation in total blood of healthy controls, patients and their relatives. Density AUC(w) after Mspl and Hpall digestion and T4-BGT pre-treatment for production of 5ghmC. **(B)** Bar plots of global 5mC and 5hmC methylation in total blood of controls, patients, their relatives and healthy individuals carrying polymorphisms showed increased 5mC levels and decreased 5hmC in patients, with intermediate values in their heterozygous relatives. Data shown are mean \pm SD from 2 independent experiments and seven healthy controls. **(C)** Bar plots of global 5mC and 5hmC methylation in control and patients fibroblasts, iPSC and haematopoietic progenitors (HP) during haematopoietic differentiation *in vitro* at day 13 (D13) and day 20 (D20). Data are shown as mean \pm SD, n = 2. P-values were determined by unpaired t-tests compared to healthy controls: *, P < 0.05; **, P < 0.01; ***, P < 0.001.

A**B****C**

Supplemental Figure S5. B-cell differentiation assay *in vitro*. A representative example of immunophenotype of TET2-deficient patient P3 and healthy control cells between day 0 and day 20 of an *in vitro* B-cell differentiation assay determined by flow cytometry, showing impaired B-cell maturation and cell survival in patient.



Supplemental Figure S6. Characterisation of derived iPSC from patient P1 and P2 and healthy individuals. **(A)** Clearance of Sendai virus vectors and expression of pluripotency markers Nanog and Oct3/4 detected by PCR in TET2-deficient patient-derived iPSC P1 and P2 and three healthy control iPSC lines in various passage number p1, p9, p14 or p16. **(B)** Sanger sequencing of *TET2* gene in patients and non-affected iPSC confirming *TET2*^{H1382R} mutation. **(C)** Representative pictures of expression of pluripotency markers TRA1-60, SSEA4, Oct4 and Nanog detected by flow cytometry and **(D)** by immunofluorescence in iPSC lines. Blue: DAPI, Red: pluripotency markers. **(E)** Differentiation of iPSC into 3-germ layers *in vitro* and **(F)** *in vivo*. Ectoderm - beta-III tubulin (TUJ1, green), Mesoderm - smooth muscle actin (SMA, red) and Endoderm - alpha-fetoprotein (AFP, red). Ne - neuroepithelium, ct - cartilage, int - intestinal epithelium. Blue: DAPI. **(G)** Karyotyping of fibroblasts and iPSC lines, number of positive metaphases in brackets. Isodicentric chromosome 20 in 44% of P1 clone1 cells, and trisomy of chromosome 14 in 30% of P1 clone 3 cells, highlighted by arrow.

A**B****C****D****E****F****G**

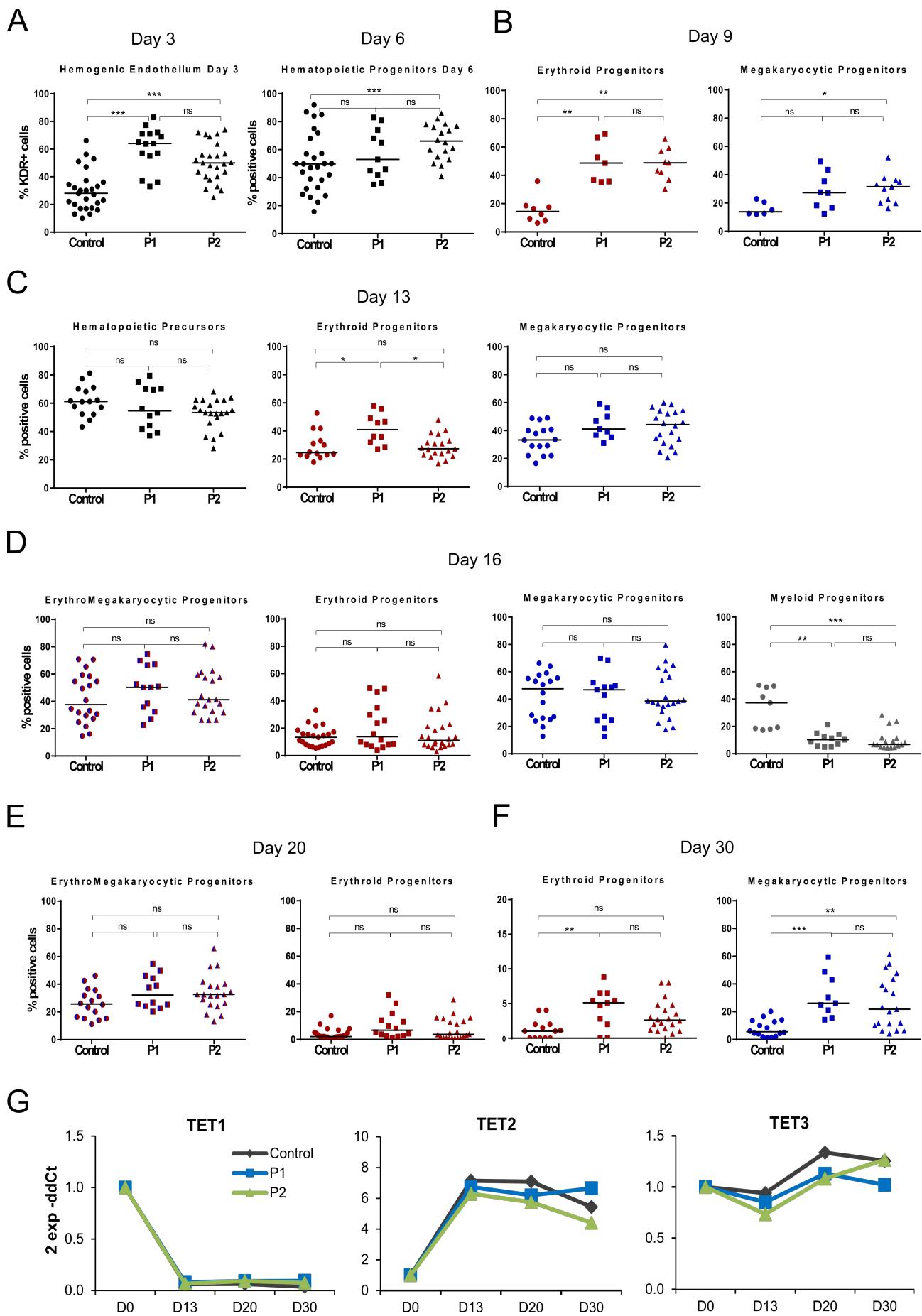
Karyotype	
Control	
Control iPSC 1	46,XY
Control iPSC 2	46,XY
Control iPSC 3	46,XY
P1	
Fibroblasts	46,XY
cl1 iPSC	46,XY, idic 20[44%]/46,XY
cl2 iPSC	46,XY
cl3 iPSC	47,XY, +14[30%]/46,XY
P2	
Fibroblasts	46,XY
cl3 iPSC	46,XY
cl4 iPSC	46,XY
cl5 iPSC	46,XY

P1 clone 1

P1 clone 3

Supplemental Figure S7. Hematopoietic differentiation of patients and control iPSC *in vitro*.

Scatter box plots demonstrating **(A)** the presence of hemogenic endothelium. i.e. percentage of KDR positive cells on Day 3 and hematopoietic progenitors (CD34+/-CD43+) on Day 6 of hematopoietic differentiation. Percentage of positive cells of individual hematopoietic progenitor subpopulations at **(B)** Day 9, **(C)** Day 13, **(D)** Day 16, **(E)** Day 20 and **(F)** Day 30 of differentiation *in vitro*. Erythro-megakaryocytic progenitors (CD43+CD235a+CD41a+), erythroid progenitors (CD43+CD235a+CD41a-), megakaryocytic progenitors (CD43+CD235a-CD41a+) and myeloid progenitors (CD43+CD235a-CD41a-CD45+). Bar represents median value from minimum 6 independent experiments. Statistical significance was calculated using non-parametric Kruskal-Wallis test. **(G)** qRT-PCR of TET1, TET2 and TET3 mRNA expression in hematopoietic precursors during individual time points of hematopoietic differentiation normalized to average house-keeping genes SDHA and UBE4A and to respective expression at Day 0.



1. Schindelin J, Rueden CT, Hiner MC, Eliceiri KW. The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol Reprod Dev* 2015;82:518-29.
2. Olivier EN, Marenah L, McCahill A, Condie A, Cowan S, Mountford JC. High-Efficiency Serum-Free Feeder-Free Erythroid Differentiation of Human Pluripotent Stem Cells Using Small Molecules. *Stem Cell Transl Med* 2016;5:1394-405.