Supplemental Data

Detailed clinical descriptions of the patients

Patient 1

Congenital anaemia

A Japanese female patient followed up for intrauterine growth restriction (IUGR) was born at 29 weeks of gestation with a weight of 726 g (−3.3 standard deviation [SD]) and a height of 32 cm (−2.9 SD). Her Apgar score at 5 min was 6. She did not require surfactant replacement therapy, but was on a ventilator for 9 days. Her haematological characteristics are shown in Supplemental Table 1 and Supplemental Figure 1. At birth, she had severe macrocytic anaemia without reticulocytopenia (red blood cells [RBC] 0.95 × 10^{12}/L, haemoglobin [Hgb] level 43 g/L, mean corpuscular volume [MCV] 123 fL, reticulocyte counts 158 × 10^{9}/L, white blood cells [WBC] 2.3 × 10^{9}/L, and platelets 396 × 10^{9}/L). Although she was treated with oral iron supplementation and erythropoietin administration with intermittent transfusion of packed RBC considering anaemia of prematurity, her anaemia did not improve. Oral iron supplementation and erythropoietin administration were discontinued in early infancy. She required repetitive transfusions of RBC once or twice a month. WBC counts (including differential counts) and platelets were normal (Supplemental Figure 1). Reticulocyte counts were mostly within the reference interval, but without marked reticulocytopenia in contrast to those observed in Diamond-Blackfan anaemia (DBA). She was referred to the National Center for Child Health and Development (Tokyo, Japan) during infancy for the evaluation and treatment of anaemia and bloody diarrhoea (details of gastrointestinal symptoms are described below). The examination of bone marrow aspirate revealed erythroid dysplasia, but without marked erythroid hypoplasia (M/E ratio 1.5) (Supplemental Table 1). These findings did not fit the diagnosis of DBA. Immunostaining for TP53 of bone marrow cells showed enhanced signals (Figure 1). The panel sequencing for known DBA-associated genes, including *RPS10, RPS14, RPS17, RPS19, RPS24, RPS26, RPL5, RPL11, RPL35A,* and *TP53*, did not detect any pathogenic variant. In toddlerhood, the follow-up bone marrow test demonstrated hypercellularity and revealed dysplasia of trilineage cells with excess blasts (7.5%) that fit myelodysplastic syndrome (MDS), classified into refractory anaemia with excess blasts

(RAEB)-1 using the World Health Organization (WHO) classification. G-banding karyotyping of the bone marrow aspirate showed a normal female karyotype as 46,XX.

Inflammatory bowel disease

The patient developed chronic bloody diarrhoea and underwent lower gastrointestinal endoscopy during infancy. Inflammatory bowel disease (IBD) (unclassified) was diagnosed based on the endoscopic and pathological findings. Her IBD was intractable to 5-aminosalicylate and exclusive enteral nutrition. Although it responded to corticosteroids, bloody diarrhoea relapsed soon after the discontinuation of corticosteroids. She was dependent on corticosteroids and required a central intravenous catheter for intravenous hyperalimentation to treat malnutrition caused by IBD.

Other clinical features

She had facial dysmorphism with epicanthic folds. She had no nail dystrophy, skin pigmentation, genitourinary anomalies, or café-au-lait patches. Despite adequate caloric intake, she showed no catch-up growth, and her height SD score was –3.1 SD (height 74 cm) in toddlerhood. At this point, hormone-secreting capacities of the anterior pituitary gland and adrenal cortex were investigated, but no abnormalities were found. Magnetic resonance imaging (MRI) showed microencephaly and skull thickening, and a computed tomography (CT) scan detected calcification at the right brachiocephalic artery. Her development was mildly delayed; she began to roll, sat steady, and crawled in late infancy, and walked in late toddlerhood. In early childhood, she could run, use several words, and imitate actions, but could not use two-word phrases, build a tower of two bricks, use a spoon, or express wants with pointing. Clinical features compared to those of FILS and IMAGE-I syndromes are shown in Supplemental Table 2.

Patient 2

Congenital anaemia

Patient 2, the younger female sibling of Patient 1, was born at 37 weeks of gestation with a weight of 2,124 g (–1.6 SD) and height of 43 cm (–1.9 SD). She had a low Apgar score (4 at 5 min) and required

ventilator support for 2 days. She also had epicanthic folds resembling her sibling's. Her haematological characteristics are shown in Supplemental Table 1 and Supplemental Figure 2. Her Hgb level was 54 g/L at birth, there were no decreases in reticulocyte counts, WBC counts, or platelet counts (RBC 1.71×10^{12}/L, MCV 117 fL, reticulocyte counts 30×10^{9}/L, platelets 306×10^{9}/L, and WBC 12.8×10^{9}/L). Bone marrow examination in infancy showed erythroid dysplasia, but without marked erythroid hypoplasia (M/E ratio 1.5) (Supplemental Table 1). Immunostaining for TP53 of bone marrow cells showed enhanced signals (Figure 1). She required repetitive transfusions of packed RBC once or twice a month. In early toddlerhood, the follow-up examination of the bone marrow aspirate demonstrated hypercellularity and dysplasia of trilineage cells with excess blasts (8.0%) that fit MDS, classified into RAEB-1. G-banding karyotyping showed a normal female karyotype as 46,XX.

Other clinical features

She had no nail dystrophy, skin pigmentation, genitourinary anomalies, or café-au-lait patches. She showed growth failure with height 71 cm (–2.6 SD) in toddlerhood. At this point, hormone-secreting capacities of the anterior pituitary gland and adrenal cortex were investigated, but no abnormalities were found. She showed the following typical developmental milestone achievements: began to roll, sat alone, walked, and used words during infancy. In toddlerhood, considering the abnormal intracranial findings of Patient 1, brain MRI, MR angiography, and CT scans of head, chest, and abdomen and pelvis were performed, showing no abnormal findings. Clinical features compared to those of FILS and IMAGE-I syndromes are shown in Supplemental Table 2.

She had not developed bloody diarrhoea by the follow-up period. She underwent a colonoscopy for her mild diarrhoea during infancy, but endoscopic and pathological findings of IBD were not detected.

Supplemental methods:

Immunostaining of bone marrow samples

Clot sections of the bone marrow obtained from Patient 1, Patient 2, and controls (a patient with

immune thrombocytopenia in early childhood and a patient with autoimmune neutropenia in infancy) were treated with high pH target retrieval solution (Dako, Santa Clara, CA) at 98 °C for 40 min. Immunostaining was performed with mouse anti-TP53 antibody (DO-7, Dako Nichirei Bioscience, Tokyo, Japan) and a polymer reagent Simple Stain MAX-PO (MULTI; Nichirei Bioscience) in an autostainer, i.e., Histostainer (48A, Nichirei Bioscience), for 30 min at room temperature.

Exome sequencing

Genomic screening was conducted under the "Initiative for Rare and Undiagnosed Diseases in Paediatrics" program supported by the Japan Agency for Medical Research and Development after obtaining written informed consent from the patients' parents. Exome sequencing of the patients and their parents were performed with SureSelect Human All ExomeV6 (Agilent Technologies, Santa Clara, CA) and the Illumina HiSeq 2500 system (Illumina, San Diego, CA), as previously described.[1] We filtered the dataset using the Japanese variant database JPN38K (https://jmorp.megabank.tohoku.ac.jp). To obtain a list of candidate pathogenic variants, the following filtering criteria were applied: (1) non-synonymous or frameshift variants in coding exons or splicing sites, (2) allele frequencies below 0.5% in 38KJPN, and (3) absence of the variant in our in-house exome data to filter pipeline-specific systematic errors. Potential pathogenicity was assumed from inheritance patterns — *de novo*, homozygous, compound heterozygous, or hemizygous — and functional effects were predicted using SIFT, PolyPhen2, and CADD.

Direct sequencing was performed to confirm the presence of the identified *POLE* variants. The genomic DNA samples from the patients were PCR-amplified for the region containing the variant using AmpliTaq Gold 360 PCR Master Mix (Thermo Fisher Scientific, Waltham, MA). The PCR products were sequenced with BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific) on an ABI 3130 sequencer (Thermo Fisher Scientific).

Functional characterization of POLE variants

A vector containing human *POLE* cDNA (FHC24944) was purchased from Kazusa DNA Research Institute (Chiba, Japan). We modified the pBQM812A-1 vector (System Biosciences, Palo Alto, CA)

by replacing the cumate-inducible promoter with a doxycycline-inducible promoter and the CymR repressor sequence with the reverse tetracycline transactivator sequence.[2] *POLE* cDNA was cloned into the modified vector by adding the haemagglutinin (HA) sequence to the N-terminus using the Gibson assembly technique (NEBuilder HiFi DNA Assembly Master Mix; New England Biolabs, Ipswich, MA). We introduced c.3392del and c.5672_5674del into the wildtype (WT)-POLE1-expressing vector using a standard PCR-based method. HEK293 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 50 U/mL penicillin, 50 μg/mL streptomycin, and 10% foetal bovine serum (FBS). The cells were transfected with each vector using Lipofectamine 3000 reagent (Thermo Fisher Scientific), and inducible stable cell lines were established according to the protocol for pBQM812A-1.

For Western blotting, nuclear and cytoplasmic extracts were prepared from inducible stable HEK293 cells expressing each HA-POLE1 (WT, Asp1131fs or Thr1891del). The cells cultured in 10-cm dishes were treated with 1 μg/mL doxycycline for 24 h. The cells were harvested using 0.25% trypsin-ethylene diamine tetra acetic acid (EDTA) solution and centrifuged at 2,000 rpm for 3 min. Cell pellets were lysed in five times the pellet volume (PV) of hypotonic lysis buffer (10 mM Tris, pH 8.0, 1.5 mM $MgCl_2$, 10 mM KCl, and 1 mM dithiothreitol with a protease inhibitor cocktail) and allowed to swell on ice for 15 min. Triton X-100 was added to a final concentration of 0.6%, and the samples were vortexed for 10 seconds. The homogenates were pelleted by centrifugation at 10,000 × *g* for 30 seconds, and the supernatants were stored as cytoplasmic lysates. The nuclear pellets were suspended in extraction buffer (20 mM Tris pH 8.0, 1.5 mM $MgCl_2$, 420 mM NaCl, 0.2 mM EDTA, 25% glycerol, and 1 mM dithiothreitol with a protease inhibitor cocktail) at two-thirds volume of the PV and rotated for 15 min at 4 °C. The homogenates were centrifuged at 20,000 × *g* for 5 min, and the supernatants were stored as cytoplasmic lysates. Nuclear and cytoplasmic extracts were separated using 10% sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Immunoblotting was performed with rat anti-HA antibody (clone 3F10; Sigma-Aldrich, St. Louis, MO), rat anti-tubulin antibody (YL1/2; Abcam, Cambridge, UK) and rabbit anti-histone H3 antibody (ab1791; Abcam) as primary antibodies, and horseradish peroxidase (HRP)-conjugated goat anti-rat IgG (Sigma-Aldrich) and HRP-conjugated goat anti-rabbit IgG (Sigma-Aldrich) as secondary

antibodies.

To visualize subcellular localization of each HA-POLE1 protein (WT, Asp1131fs or Thr1891del), inducible stable HEK293 cells were seeded on glass-bottom dishes and grown with 1 μg/mL doxycycline for 24 h. The cells were then fixed in 4% formaldehyde in phosphate-buffered saline (PBS) at room temperature for 10 min. Blocking and plasma membrane permeabilization were performed by incubating the cells with 4% Block Ace (KAC, Kyoto, Japan) and 0.1% Triton X-100 at room temperature for 1 h. Immunofluorescence was performed using rat anti-HA antibody (clone 3F10) and rabbit anti-POLE2 antibody (HPA02755; Thermo Fisher Scientific) as primary antibodies, and Alexa Fluor 488-conjugated donkey anti-rat IgG antibody (A21208, Thermo Fisher Scientific) and Alexa Fluor 647-conjugated goat anti-rabbit IgG antibody (A27040; Thermo Fisher Scientific) as secondary antibodies. Nuclei were stained with Hoechst 33342 (Dojindo Laboratories, Kumamoto, Japan). The cells were observed under an FV3000 confocal microscope (Olympus, Tokyo, Japan).

Generation of patient-derived induced pluripotent stem cells (iPSCs)

Human iPSCs were generated and maintained on iMatrix-511 (Nippi, Inc., Tokyo, Japan) in StemFit AK02N (Reprocell, Inc., Kanagawa, Japan) or StemFlex (Thermo Fisher Scientific) media. For Patient 1, a small intestinal tissue biopsy specimen obtained from the patient was chopped into small pieces and seeded on a dish in DMEM/F12 (Thermo Fisher Scientific) containing 20% FBS. The medium was changed every 2 days until confluence. Intestinal epithelial cells were maintained in DMEM/F12 containing 20% FBS until reprogramming. On the first day of reprogramming, we changed the culture medium to NutriStem Medium (Sartorius Stedim Biotech, Göttingen, Germany) and used the StemRNA-3rd Gen Reprogramming Kit (Reprocell, Inc.) to induce six reprogramming factors, including OCT4, SOX2, KLF4, c-MYC, NANOG, and LIN28, following the manufacturer's protocol. From days 2-4, we replaced the fresh NutriStem medium and transfected a reprogramming cocktail using Lipofectamine RNAiMAX (Thermo Fisher Scientific) every day. From day 5, the old medium was changed every 2 days until colony pickup.

For Patient 2, peripheral blood mononuclear cells (PBMCs) were isolated from the whole blood of the patient by centrifugation on Ficoll-Paque gradient using Leucosep (Greiner Bio-One

International GmbH, Frickenhausen, Germany) following the manufacturer's protocol. The CytoTune-iPS 2.0 Sendai Reprogramming Kit (Thermo Fisher Scientific) was used to induce four reprogramming factors, including OCT4, SOX2, KLF4, and L-MYC. PBMCs were infected with Sendai virus vector at 37 °C in a 5% CO_2 incubator for 2 h. Next, the infected PBMCs were seeded into iMatrix-511-coated 12-well plates (AGC Techno Glass Co., Ltd., Shizuoka, Japan) in KBM 501 medium (Kohjin Bio Co., Ltd., Saitama, Japan) with 10% FBS (Thermo Fisher Scientific). On days 2 and 4, 1 mL of StemFit AK02N medium was added gently to avoid disturbing cell adhesion, after which the entire medium was changed every 2 days until colony pickup.

From days 14-21, individual colonies were picked into iMatrix-511-coated 4-well plates (Thermo Fisher Scientific) in StemFit AK02N medium supplemented with 10 μM Y-27632 (Fujifilm Wako Pure Chemical Corporation, Osaka, Japan) or in StemFlex medium supplemented with 10 μM Y-27632. The medium was changed every other day and cells were passaged approximately once per week using enzymatic (TrypLE Select Enzyme or Accutase Cell Dissociation Reagent; Thermo Fisher Scientific) or mechanical methods.

RNA sequencing (RNA-seq) analysis

Total RNA was extracted from the iPSCs generated from the *POLE* variant carriers and control individuals (three lines for Patient 1, one line for Patient 2, and one line each from the two controls) using TRIzol reagent (Thermo Fisher Scientific) following the manufacturer's instructions. The RNA samples were treated with TruSeq Stranded Total RNA Library Prep Kit (Illumina), and the libraries were sequenced on DNBSEQ-T7 (MGI Tech Co., Ltd., Shenzhen, China). Reads were aligned to the reference human genome hg38, visualized with Integrative Genomics Viewer (https://software.broadinstitute.org/software/igv/), and reads per gene were counted using STAR-2.7.9a software.[3] Within-sample normalization was conducted using the fragments per kilobase million method, and between-sample normalization was performed using the trimmed mean of the M-values method.[4]

Western blotting of TP53 protein

For Western blotting, nuclear proteins were extracted from iPSCs generated from the patients, and normal controls (two iPSC lines each). First, the cells grown in 10-cm dishes were harvested using 0.25% trypsin-EDTA solution and centrifuged at 2,000 rpm for 3 min. Cell pellets were lysed in five times the PV of hypotonic lysis buffer (10 mM Tris, pH 8.0, 1.5 mM $MgCl_2$, 10 mM KCl, and 1 mM dithiothreitol with a protease inhibitor cocktail) and allowed to swell on ice for 15 min. Triton X-100 was added to a final concentration of 0.6%, and the samples were vortexed for 10 s. The homogenates were pelleted by centrifugation at 10,000 × *g* for 30 seconds and the supernatants were discarded. The nuclear pellets were suspended in extraction buffer (20 mM Tris pH 8.0, 1.5 mM $MgCl_2$, 420 mM NaCl, 0.2 mM EDTA, 25% glycerol, and 1 mM dithiothreitol with a protease inhibitor cocktail) at two-thirds volume of the PV and rotated for 15 min at 4 °C. The homogenates were centrifuged at 20,000 × *g* for 5 min, and the supernatants (nuclear fraction) were separated using 10% SDS-PAGE. Immunoblotting was performed with mouse anti-TP53 antibody (DO-7) and rabbit anti-histone H3 antibody (ab1791) as primary antibodies, and HRP-conjugated goat anti-mouse IgG (Sigma-Aldrich) and HRP-conjugated goat anti-rabbit IgG (Sigma-Aldrich) as secondary antibodies.

Generation of Pole-deficient medaka

Supplemental Figure 3 shows the sequence alignment of human and medaka *Pole* orthologues. We designed two crRNAs that recognize the sequences ACG TAG AAT ACA TCA CCA AC (site 1) and GAA TAC ATC ACC AAC AGG TC (site 2) of the medaka DNA polymerase epsilon (*Pole*) gene to establish a *Pole* mutant using CRISPR/Cas9. Partial sequences of medaka *Pole*, two crRNA recognition sites, and PCR primers for the amplification of the target region are shown in Supplemental Figure 4. The formation of duplex guide ribonucleoproteins (dgRNP) is in accordance with that reported by Hoshijima *et al*.[5] The crRNA, tracrRNA, and Cas9 proteins were purchased from Integrated DNA Technologies (IA). The PCR primers (CAT CTT GTG CAC CAA AAA GC and AAG TCT TGG GGT TTT GAA AT) were purchased from Fasmac (Kanagawa, Japan). Microinjection and establishment of *Pole* knockout strains are in accordance with “Medaka: Biology, Management, and Experimental Protocol, Volume 2”.[6] Briefly, microinjected with dgRNPs into embryos at the 1 cell stage, eggs were incubated for 4 days at 25 °C, and the genomic DNA was extracted and subjected to

PCR amplification. The efficiency of genome editing was evaluated using an automated electrophoresis system MCE-202 with a DNA-500 reagent kit (Shimadzu, Kyoto, Japan). Two dgRNPs were designed to efficiently edit the target DNA regions. Following the selection of four pairs of founder fish, we observed the phenotypes of the next generation. Two out of the four pairs of founder fish laid eggs with a specific phenotype. We selected four founder fish, No. 5, No. 6, No. 7, and No. 8, and mated them with WT fish. Supplemental Figure 5 shows the types of insertions or deletions identified in the target region of the F1 fish using amplicon sequencing. As a result, we established two mutant lines, one with a 4 bp deletion mutation using one crRNA (site 2) and another with a 9 bp deletion mutation using another crRNA (site 1) at the *Pole* gene. We named these two mutant strains *Pole^del4* and *Pole^del9*, respectively. Two mutant strains *Pole^del4* and *Pole^del9* showed essentially similar phenotypes, and thus we used the *Pole^del9* strain for further analysis. As shown in Supplemental Figure 4, the *Pole^del9* strain deleted the splice-acceptor site and was expected to be non-functional.

Morphological observation of medaka

Medaka has a more extended period time from fertilization to hatching than zebrafish. For this reason, the length of the body and other developmental phenotypes in pre-hatched embryos cannot be accurately analysed. Therefore, we dissolved the chorion using hatching enzymes[7,8] and observed the embryos. The hatching enzyme was supplied by NBRP Medaka (https://shigen.nig.ac.jp/medaka/). Seven days after fertilization, the eggs were transferred to waterproof sandpaper (#1200) (AS One, Osaka, Japan) placed in the lid of a 10-cm Petri dish. Breeding water was removed, and a moderate amount of Iwamatsu's balanced salt solution (BSS)[9] was added to prevent the drying of the eggs. The eggs were slowly rolled to remove outer surface hairs and lightly scratch the surface of the eggs. The eggs were then transferred to a 24-well plate, and the BSS was removed. The hatching enzyme diluted three times in BSS was added so that the eggs were immersed and incubated at 27 °C. As soon as the embryos emerged from the chorion, embryos were transferred to a Petri dish containing 1 × BSS. The embryos were captured using a microscope camera (Zeiss Axiocam 208 colors; Carl Zeiss Meditec, Jena, Germany) attached to a stereomicroscope (model M165C; Leica, Wetzlar, Germany).

Blood cell counts of medaka

Following anaesthesia with tricaine methane sulfonate (MS222; Sigma-Aldrich), the fish's heart was exposed by dissection. Blood was collected from the bulbus arteriosus using glass capillaries (GD-1; Narishige, Tokyo, Japan) coated with 10,000 U/mL heparin sodium salt (Wako, Osaka, Japan).[10] In our experiments, we used Dulbecco's modified PBS (DPBS), treated to remove Mg^{2+} and Ca^{2+} ions [DPBS (–)] to prevent coagulation during the dilution of whole blood when necessary. Total blood cells were diluted to 1:200, stained with Shaw's diluent, and counted using a haemocytometer.[11] Blood cells (2×10^5 cells) diluted in DPBS (–) containing 2.5% FBS were centrifuged (Cytopro 7620; Wescor Inc., Logan, UT) at 1,000 rpm for 4 min to prepare cell centrifuge specimens and transferred onto glass slides. After air drying, cells were stained with o-dianisidine (Wako) and Giemsa (Wako) as previously described.[12] Blood cells were examined by light microscopy (model BX51; Olympus, Tokyo, Japan).

Statistical analysis

Data were analysed using the GraphPad Prism 9.1.2 software (GraphPad Software, Inc., San Diego, CA). The two-group datasets were analysed using Student's *t*-test. Statistical significance was set at $P < 0.05$.

References

1. Sato T, Samura O, Kato N, et al. Novel TFAP2A mutation in a Japanese family with branchio-oculo-facial syndrome. *Hum Genome Var* 2018;5:5.

2. Shima H, Koehler K, Nomura Y, et al. Two patients with MIRAGE syndrome lacking haematological features: role of somatic second-site reversion SAMD9 mutations. *J Med Genet* 2018;55:81–5.

3. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29:15–21.

4. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 2010;11:R25.

5. Hoshijima K, Jurynec MJ, Klatt Shaw D, et al. Highly efficient CRISPR-Cas9-based methods for generating deletion mutations and F0 embryos that lack gene function in Zebrafish. *Dev Cell* 2019;51:64–557.e644.

6. Murata K, Kinoshita M, Naruse K, et al. Heteroduplex mobility assay – A simple method to detect targeted genome modification. In: Murata K, Kinoshita M, Naruse K, et al. eds. Medaka: Biology, Management, and Experimental Protocols. Vol. 2. Iowa: Wiley-Blackwell 2019:244–60.

7. Kinoshita M, Murata K, Naruse K, et al. Medaka: biology, management, and experimental protocols. New York: John Wiley & Sons 2009:252–56.

8. Porazinski SR, Wang H, Furutani-Seiki M. Dechorionation of medaka embryos and cell transplantation for the generation of chimeras. *J Vis Exp* 2010;46:e2055.

9. Iwamatsu T. A new technique for dechorionation and observations on the development of the naked egg in Oryzias latipes. *J Exp Zool* 1983;228:83–9.

10. Kayo D, Oka Y, Kanda S. Examination of methods for manipulating serum 17β-estradiol (E2) levels by analysis of blood E2 concentration in medaka (Oryzias latipes). *Gen Comp Endocrinol* 2020;285:113272.

11. Hadji-Azimi I, Coosemans V, Canicatti C. Atlas of adult Xenopus laevis laevis hematology. *Dev Comp Immunol* 1987;11:807–74.

12. Okui T, Yamamoto Y, Maekawa S, et al. Quantification and localization of erythropoietin-receptor-expressing cells in the liver of Xenopus laevis. *Cell Tissue Res* 2013;353:153–64.