## **Supplementary data**

### **Supplementary Figure 1**



**Figure S1. (a)** Direct Sanger sequencing of the *EFL1* gene in control, individual P1, his parents and his healthy sister. Sequencing of genomic DNA (left) and cDNA from B-LCL (right). **(b)** Direct Sanger sequencing of the *EFL1* gene in control, individual P2, her parents and her healthy siblings. **(c)** Direct Sanger sequencing of the *EFL1* gene in control, individual P3, her parents and her healthy brother.



**Figure S2.** Expression of P3's maternal *EFL1* allele is defective in P3 and her mother. (**a**) Sequencing analysis of the single nucleotide variant rs4725 detected in genomic DNA (gDNA) at a heterozygous state in B-LCL from a control, P3's mother and P3 is almost undetectable in cDNA from P3's mother and P3. (**b**) Western blot analysis showed a reduction of EFL1 amount in B-LCL lysate from the P3's mother confirming the defective *EFL1* expression of one allele detected in (**a**). **(c)** Haplotype analysis using different genetic markers (listed on the right) indicates that P3's brother, unlike P3, possesses the maternal wild type *EFL1* allele. (**d**) Polysome profiles from B-LCL extracts from the individual P3 (n=9) and her mother (n=3) compared with wild type control (n=9). **(e)** Quantification of the 60S:80S ribosomal subunit ratio is indicated as a bar chart in B-LCL cells lines from individual P3 ( $n=3$ ) and her mother (n=2) relative to wild type control cells (n=3). (**f**) Relative rate of protein synthesis assessed by OP-Puro incorporation in B-LCL cells lines from individual P3  $(n=3)$  and her mother  $(n=3)$ , relative to wild type control cells (n=3) quantified 1 hr after OP-Puro administration.

a Metaphyseal chondrodysplasia of the knees<br>and the hips in individual P1 at 27 years of age



**Figure S3. (a)** X-rays of individual P1 and P3 at 27 years of age and 6 years of age, respectively, showing poly-epiphyseal dysplasia in P1 and metaphyseal chondrodysplasia in P3. **(b)** Variation of absolute neutrophil count (x  $10^9$ /L), hemoglobin (g/dL), and platelets (x  $10^9$ /L) in P1, P2 and P3, according to age. Red line indicates lower normal range in healthy individuals. **(c)** Hyposegmented neutrophils in bone marrow aspirates from individuals P1, P2 and P3 but not in healthy control.



**Figure S4. (a)** Flow cytometry strategy to quantify indicated bone marrow subpopulations. **(b)** Number of hematopoietic stem and progenitor cells in the bone marrow (n=8 per genotype). For all tests, p values: \*< 0.05; \*\* < 0.01; \*\*\* < 0.001.



**Figure S5. (a)** Localization of the *EFL1* gene (green arrow) and *EFL1P1* pseudogene (red arrow) on chromosome 15. Capture image obtained from UCSC website (https://genome.ucsc.edu/). **(b)** Schematic representation of a gene conversion event between *EFL1P1* and *EFL1.* **(c)** Sequence alignment of *EFL1*  and *EFL1P1* indicating gene-specific (green) and pseudogene-specific (red) sequences and the position of the codon encoding amino-acid F505 mutated in P1.

**Supplementary Table 1.** Characterization of *EFL1* mutations identified in individuals P1, P2 and P3.



Abbreviations: N, no; Y, yes; NA, not applicable.

**Supplementary Table2.** Characteristics of the segmental duplication and pseudogenes of the *EFL1*

gene.





**Segmental duplication** 





## **Supplementary Table 3. Antibodies.**

#### **Supplementary Material and Methods**

**Cell culture.** Fibroblast cell lines and B-lymphoblastoid cell lines (B-LCL) were cultured at 37 °C with 5 % CO2 in DMEM and RPMI Medium 1640 (Gibco™ GlutaMAX™), respectively, supplemented with 10 % fetal bovine serum and 1 % Penicillin-Streptomycin (Sigma).

**Plasmids.** 12URA-U expression vector; gift from Scott Gradia, Addgene # 48306. pCW57-GFP-P2A, Addgene #71783. p.lenti7.3/V5 TOPO, Invitrogen.

**Whole exome sequencing.** Exome capture was performed using the SureSelect Human All Exon kit (Agilent Technologies®, Santa Clara, CA). Agilent SureSelect Human All Exon (54 Mb, Clinical research Exome) libraries were prepared from 3 µg of genomic DNA sheared with an Ultrasonicator (Covaris®, Woburn, MA) as recommended by the manufacturer. Barcoded exome libraries were pooled and sequenced using a HiSeq2500 (Illumina®, San Diego, CA) generating 130 x 130 paired-end reads. After demultiplexing, sequences were mapped on the human genome reference (NCBI build37/hg19 version) with BWA. The mean depth of coverage obtained for the exome library was 138x with > 99 % of the targeted exonic bases covered at least by 15 independent reads and  $> 97\%$  by at least 30 independent sequencing reads ( $> 99\%$  at 15 x and  $> 97\%$  at 30 x). Variant calling was carried out with the Genome Analysis Toolkit (GATK), SAMtools and Picard Tools. Single nucleotide variants were called with GATK Unified Genotyper, whereas indel calls were made with the GATK IndelGenotyper y2. All variants with a read coverage  $\leq 2$  x and a Phred-scaled quality of  $\leq 20$  were filtered out. All the variants were annotated and filtered using PolyWeb, an in-house developed annotation software.

# **Targeted resequencing by NGS (capture by hybridization approach) to detect copy number variations and intronic mutations.**

Briefly, 1 to 3 µg of each patient's genomic DNA was mechanically fragmented to a median size of 200 bp using an ultrasonicator (Covaris). 100 ng of double strand fragmented DNA was end-repaired and adaptors containing a specific 8 base barcode were ligated to the repaired ends (one specific barcode per patient). DNA fragments were PCR amplified to get the final precapture barcoded libraries that were pooled at equimolar concentrations (a pool of 15 libraries was prepared). The capture process was performed using the SureSelect reagents (Agilent), 750 ng of this pool of precapture libraries and homemade biotinylated probes. The biotinylated single stranded DNA probes were designed and prepared to cover a 194 kb chromosomal region including the complete *EFL1* gene on chromosome 15. The limits of the targeted chromosomal region are Chr15:82,085,926-82,280,215 according to the GRCh38.p2 assembly of the human reference genome. During the capture process, barcoded library molecules complementary to the biotinylated beads were retained by streptavidin coated magnetic beads on a magnet and PCR amplified to generate a final pool of postcapture libraries covering the targeted chromosomal region on chromosome 15. In total a pool of 6 libraries (DNA from P3, P3's parents, and 3 controls) covering a 194 kb territory which includes the entire *EFL1* gene was sequenced on an Illumina HiSeq2500 (Paired-End sequencing 130x130 bases, High Throughput Mode, 7 samples on half of a FlowCell lane). After demultiplexing, sequences were aligned to the reference human genome hg19 using the Burrows-Wheeler Aligner (Li and Durbin, 2010). The mean depth of coverage obtained per sample was  $\geq 600X$  to enable more accurate copy number variant **analysis**. Downstream processing was carried out with the Genome Analysis Toolkit (GATK), SAMtools and Picard, following documented best practices (http://www.broadinstitute.org/gatk/guide/topic?name=best-practices). Variant calls were made with the GATK Unified Genotyper. The annotation process was based on the latest release of the Ensembl database. Variants were annotated, analyzed and prioritized using the Polyweb/PolyDiag software interface designed by the Bioinformatics platform of University Paris Descartes.

**Immunoblotting.** Cell pellets were lyzed in RIPA buffer (20 mM HEPES at pH 7.4, 20 mM βglycerophosphate, 10 mM NaF, 0.5 mM EDTA, 0.5 mM EGTA, 0.2 M NaCl, 1 % [v/v] Nonidet P-40, 0.5 % [w/v] sodium deoxycholate, 0.1 % [w/v] SDS) with complete EDTA-free protease inhibitors (Roche), and incubation on ice for 10 min with occasional vortexing. Cell lysates were cleared in a microfuge and normalized for protein concentration using a BCA protein assay kit (Pierce). Protein samples were separated on SDS-PAGE gels and transferred to PVDF membranes. Blots were subjected to immunoblotting analysis using specific antibodies listed in **Supplementary Table 3**. Protein bands were visualized by chemiluminescence (SuperSignal West Pico Chemiluminescent Substrate, Thermo Fisher). Specific protein signal was measured using Fiji (Image  $J^1$ .

Measurement of protein synthesis. Protein synthesis was measured as described<sup>2</sup>. OP-Puro (Invitrogen; final concentration 50 µM) was added to the culture medium (Dulbecco's Modified Eagle Medium (DMEM, Gibco™ GlutaMAX™), 10 % fetal bovine serum (Sigma) and 1 % Penicillin-Streptomycin (Pen-Strep, Sigma)) for 60 min. AHA (Invitrogen; 0.2 mM final concentration) was added to the culture medium (methionine-free Dulbecco's Modified Eagle's Medium (Sigma) supplemented with 200  $\mu$ M L-cysteine (Sigma), 50  $\mu$ M 2-mercaptoethanol (Sigma), 1 mM L-glutamine (Gibco) and 0.1% bovine serum albumin (BSA; Sigma)) for 3 h. Cells were removed from wells and washed twice in ice-cold Ca<sup>2+</sup>- and Mg<sup>2+</sup>-free phosphate buffered saline (PBS) (Invitrogen) with 100  $\mu$ g/ml cycloheximide. Cells were fixed and permeabilized using the Cytofix/Cytoperm Fixation Permeabilization Kit (BD Biosciences). Azide-alkyne cycloaddition was performed using the Click-iT Cell Reaction Buffer Kit (Invitrogen) with azide conjugated to Alexa Fluor 488 or Alexa Fluor 594 (Invitrogen) at 5 µM final concentration. Following the 30 min reaction, cells were washed twice in PBS supplemented with 2% fetal bovine serum, resuspended in PBS and analyzed by flow cytometry (Becton Dickinson LSR Fortessa analyzer). To inhibit OP-Puro or AHA incorporation, cycloheximide (Sigma) was added 15 min before OP-Puro or AHA at a final concentration of 100  $\mu$ g ml<sup>-1</sup>. 3 x 10<sup>5</sup> cells were used per Click-iT assay. All cultures were incubated at  $37^{\circ}$ C in  $5\%$  CO<sub>2</sub> with constant humidity. Flow cytometry data analysis was performed using FlowJo v10.1 (FlowJo, Ashland, OR). 'Relative rates of protein synthesis' were calculated by normalizing OP-Puro or AHA signals to control cells after subtracting background fluorescence (cells without OP-Puro or AHA incorporation).

**Sucrose density gradients.** Ribosomal subunits were separated by sucrose density gradients as described<sup>3,4</sup>. Briefly, fibroblasts (60 % - 80 % confluence) and B-LCL cells (0.6 x 10<sup>6</sup>/mL) were treated with cycloheximide at final concentration of 100 µg/mL for 15 min at 37 °C before harvesting. Cells were then lyzed in lysis buffer (5 mM Tris-HCl at pH 7.5, 1.5 mM KCl, 2.5 mM  $MgCl_2$ , 0.5 % [v/v] NP-40, 0.5 % [w/v] deoxycholate, 1 % [v/v] Tween 20, 100  $\mu$ g/mL cycloheximide) with complete EDTA-free protease inhibitors (Roche) and 500 U/mL RNase inhibitor (RNaseOUTTM Invitrogen) and incubated for 10 min on ice. Lysates were cleared in a microcentrifuge at 4 °C. Equal amounts (typically 1.5  $A_{254}$  U) were applied to a 5-45 % (w/v) sucrose gradient in 14 mL of gradient buffer (20) mM HEPES pH 7.5, 100 mM KCl, 5 mM  $MgCl<sub>2</sub>$ , 100  $\mu g/mL$  cycloheximide) with complete EDTA-free protease inhibitors (Roche) and centrifuged (Beckmann SW40 rotor at 40,000 rpm for 2 hour at 4 °C). The sucrose gradient was made using a Biocomp Gradient Master. Centrifugation samples were unloaded using a Brandel gradient fractionator, polysome profiles detected using an ÄKTAprime plus system (GE Healthcare), and 0.5 mL fractions collected. Proteins were precipitated from sucrose gradient fractions with 20 % (w/v) trichloroacetic acid, separated on SDS-PAGE gels and transferred to PVDF membranes for immunoblotting.

Cell fractionation. Subcellular fractions were separated from 6 x 10<sup>6</sup> cells using ProteoExtract® Subcellular Proteome Extraction Kit (Merk) or PER™ Nuclear and Cytoplasmic Extraction Reagents (Thermo Fisher Scientific). Nuclear and cytoplasmic fractions were used for immunoblotting.

**eIF6 release assay.** 60S ribosomal subunit fractions were pooled from several polysome profiles in HEPES 20 mM, KCl 100 mM,  $Mg(OAc)<sub>2</sub> 2.5$  mM pH 7.4. 60S subunits were centrifuged at 444000 g for 35 min in an Optima MAX XP Ultracentrifuge (Beckmann Coulter) using a MLA-80 rotor. The supernatant was discarded and the pellet resuspended in 60S buffer (HEPES 20 mM, KCl 150 mM,  $Mg(OAc)<sub>2</sub>2.5$  mM pH 7.4) to a concentration of 450 nM. Reaction mixtures for eIF6 release included 60S subunits, EFL1 variants, SBDS, and GTP at final concentrations of 30 nM, 1.5 µM, 5 µM, and 1 mM respectively in a final volume of 50  $\mu$  60S buffer. EFL1 was omitted in the negative control. 50 mM EDTA was added as the positive control to disassemble 60S subunits. Reactions were incubated at 25 ºC for 1 h, added to 50 µl of a 15 % sucrose cushion in 60S buffer and centrifuged in an Optima MAX XP Ultracentrifuge using a TLA-100 rotor at 476000g for 30 min. The top 80 µl ("free fraction") was removed from the cushion and the pellet resuspended in the remaining 20 µ of solution ("bound" fraction). Samples in LDS buffer were electrophoresed in 10 % Bis-Tris gels, blotted onto nitrocellulose membranes saturated in 5 % fat-free milk and incubated with anti-eIF6 (GeneTex gtx 117971, dilution 1:1000) and anti-RPL23 (uL14) (Abcam ab112587, dilution 1:500) antibodies overnight at 4 ºC. Membranes were washed in 0.1 % Tween-20 containing PBS and incubated with anti-rabbit IgG antibodies (Abcam ab6721, dilution 1:5000), washed again and chemiluminescent signal obtained by incubation with ECL substrate (Thermo SuperSignal West Pico PLUS). Images were recorded on a BioRad ChemiDoc MP system.

**EFL1 cloning, expression, and purification.** Human EFL1 mutants were generated by PCR using the wild-type EFL1 cDNA (long isoform ENST00000268206.11) I.M.A.G.E. plasmid IRAUp969G11102D as a template. Following the method of Liu and Naismith<sup>5</sup>, mutations were inserted in the primers and the plasmid was PCR amplified, DpnI digested and transformed in DH10β cells. A 6x His tag was added to all sequence variants. After further PCR amplification using 5'- TACTTCCAATCCAATCGATGGTGCTCAACAGTTTGGA-3' and 5'-TTATCCACTTCCAATGTTATTATTAATGGTGATGGTGATGGTGCTTATTTTTGCTGAGTGTC CTCTG-3' primers, EFL1 mutants were inserted into 12URA-U (Addgene plasmid # 48306) by ligation independent cloning. Mutations were verified by sequencing. Plasmids were transformed into *S. cerevisiae* INVSc1 cells (Thermo Fisher Scientific) by heat shock and plated on YEPD agar. Individual colonies were grown in liquid YEPD medium at 30 ºC and diluted in YEP supplemented with 2% raffinose and 0.5% glucose medium. During exponential growth, 2% galactose was added and cells incubated at 25 ºC for 20 hours. Cells were harvested by centrifugation and lysed in Buffer A (HEPES 50 mM, NaCl 200 mM, Imidazole 10 mM, pH 7.5) using a Retsch Mill MM400 following the manufacturer's instructions. The lysate was loaded onto a 1 ml HisTrap HP column (GE Healthcare) using an ÄKTA pure system (GE Healthcare). After extensive washing of the column with buffer A, proteins were eluted in 2 ml of Buffer B (HEPES 50 mM, NaCl 200 mM, Imidazole 150 mM, pH 7.5) and loaded on a HiLoad 26/600 Superdex 200 pg column (GE Healthcare) equilibrated in Buffer C

(HEPES 50 mM, NaCl 200 mM, pH 7.5). All EFL1 variants were retrieved under a single elution peak yielding  $\sim$  2 mg of pure protein per litre of culture. EFL1 proteins ( $>95\%$  pure by SDS-PAGE) were snap-frozen in liquid  $N_2$  at a concentration of 15  $\mu$ M.

**EFL1 cloning, expression, and purification.** Human EFL1 mutants were generated by PCR using the wild-type EFL1 cDNA (long isoform ENST00000268206.11). I.M.A.G.E. plasmid IRAUp969G11102D as a template. Following the method of Liu and Naismith<sup>5</sup>, mutations were inserted in the primers and the plasmid was PCR amplified, DpnI digested and transformed in DH10β cells. A 6x His tag was added to all sequence variants. After further PCR amplification using 5'- TACTTCCAATCCAATCGATGGTGCTCAACAGTTTGGA-3' and 5'-

### TTATCCACTTCCAATGTTATTATTAATGGTGATGGTGATGGTGCTTATTTTTGCTGAGTGTC

CTCTG-3' primers, EFL1 mutants were inserted into 12URA-U (Addgene plasmid # 48306) by ligation independent cloning. Mutations were verified by sequencing. Plasmids were transformed into *S. cerevisiae* INVSc1 cells (Thermo Fisher Scientific) by heat shock and plated on YEPD agar. Individual colonies were grown in liquid YEPD medium at 30 ºC and diluted in YEP supplemented with 2% raffinose and 0.5% glucose medium. During exponential growth, 2% galactose was added and cells incubated at 25 ºC for 20 hours. Cells were harvested by centrifugation and lysed in Buffer A (HEPES 50 mM, NaCl 200 mM, Imidazole 10 mM, pH 7.5) using a Retsch Mill MM400 following the manufacturer's instructions. The lysate was loaded onto a 1 ml HisTrap HP column (GE Healthcare) using an ÄKTA pure system (GE Healthcare). After extensive washing of the column with buffer A, proteins were eluted in 2 ml of Buffer B (HEPES 50 mM, NaCl 200 mM, Imidazole 150 mM, pH 7.5) and loaded on a HiLoad 26/600 Superdex 200 pg column (GE Healthcare) equilibrated in Buffer C (HEPES 50 mM, NaCl 200 mM, pH 7.5). All EFL1 variants were retrieved under a single elution peak yielding  $\sim$  2 mg of pure protein per litre of culture. EFL1 proteins ( $>95\%$  pure by SDS-PAGE) were snap-frozen in liquid  $N_2$  at a concentration of 15  $\mu$ M.

**Mice.** The  $EFLI^{K983R}$  mutant strain was already identified but lacked phenotypic information reported<sup>6</sup>.

All work was performed under appropriate UK Home Office animal license. All mice were produced by intercrossing heterozygous mutant carriers (*Efl1K938R/+*) after backcrossing onto the parental C3H strain for at least 3 generations, to minimize the possibility of other ENU-induced mutations interfering with the phenotype. Whole genome sequencing excluded additional coding mutations within the critical region of mouse chromosome 7. For all experiments, littermates for all genotypes produced by intercrossing *Efl1K938R/+* mice were used. All experiments were performed blind to genotype following ARRIVE guidelines<sup>7</sup>, except that data are presented only for females due to the aggressiveness of *Efl1K938R/K938R* males. Briefly, phenotypic analysis was done as follows: Weights were recorded at selected time points. Whole body lean and fat mass was assessed at selected time points using Echo MRI as per the manufacturer's instructions (Echo Medical Systems). At least 10 female mice per genotype were used. Dual-energy X-ray Absorptiometry (DEXA) was used to quantify bone mineral content and density following manufacturer instructions at selected time points. At least 7 mice per genotype and time point were used. Monitoring of in-cage wheel running activity was performed in female mice at 3 months of age using the PhenoMaster (TSE systems) system. 12 mice per genotype were used. Recordings were calculated from the first hour of the night when animals are most active using PhenoMaster automated data software, as previously<sup>8</sup>. Gait was assessed qualitatively by using a numeric scoring system at selected time points:  $1 = normal$  (normal tail height, leg position and undercarriage height);  $2 =$  slightly flat (slightly low undercarriage height, normal or wide leg position, normal or high tail position);  $3 = \text{flat}$  (low undercarriage position, normal or wide leg position, normal, high or low tail position);  $4 = \text{very flat (undercarriage dragging on ground, wide led position, low tail)}$ position). At least 10 female mice per genotype were used per time point. The Y maze alternation test was performed at 6 months of age on female mice to assess working memory. At least 6 mice per genotype were used. Briefly, first, one of the arms was blocked and mice were placed in the starting arm and left to explore the starting arm and a familiar arm for 5 minutes. Mice were then removed from the Y maze apparatus and placed back into the home cage for one minute. After that, mice were placed back into the start arm, and the arm that was previously blocked off (the novel arm) was opened for exploration. The mouse was allowed to explore for 5 minutes, and the time spent in each arm was

recorded using Any-MAZE video-tracking software.

Primary mouse embryonic fibroblasts (MEFs) from Efl1<sup>K983R/K983R</sup> mutant mice were produced from E14.5 embryos derived from Efl1<sup>K938R/+</sup> intercrosses, established following standard protocols and maintained on DMEM media supplemented with 10% fetal bovine serum.

Bone marrow cells were isolated from 4-8 week old mice by crushing hip bones, femurs and tibiae in PBS containing 2 % FCS (Thermo Fisher Scientific) and counted using the scil Vet ABC ™Hematology Analyzer. FACS analysis of the myelo-erythroid compartment in the bone marrow was performed as previously described<sup>9</sup>. The applied gating strategy and antibodies used are shown in the

supplementary figure 4 and Supplementary Table S3.

**cDNA analysis.** For RT-PCR, total RNA from patient and control fibroblasts or B-LCL was extracted using Trizol reagent (Invitrogen, Grand Island, NY) according to the manufacturer instructions. Reverse transcription was performed using a SuperScript First-Strand Synthesis Kit (Invitrogen), and cDNAs were used as templates to amplify by PCR the *EFL1* gene (long isoform ENST00000268206.11). Nucleotide numbering reflects cDNA numbering with  $+1$  corresponding to the A of the ATG translation initiation codon in the reference sequence. The initiation codon is codon 1.

#### **References**

1. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biologicalimage analysis. *Nat Methods*. 2012;9(7):676-682.

2. Signer RA, Magee JA, Salic A, Morrison SJ. Haematopoietic stem cells require a highly regulated protein synthesis rate. *Nature*. 2014;509(7498):49-54.

3. Finch AJ, Hilcenko C, Basse N, et al. Uncoupling of GTP hydrolysis from eIF6 release on the ribosome causes Shwachman-Diamond syndrome. *Genes Dev*. 2011;25(9):917-929.

4. Wong CC, Traynor D, Basse N, Kay RR, Warren AJ. Defective ribosome assembly in Shwachman-Diamond syndrome. *Blood*. 2011;118(16):4305-4312.

5. Liu H, Naismith JH. An efficient one-step site-directed deletion, insertion, single and multiplesite plasmid mutagenesis protocol. *BMC Biotechnol*. 2008;8:91.

6. Potter PK, Bowl MR, Jeyarajan P, et al. Novel gene function revealed by mouse mutagenesis screens for models of age-related disease. *Nat Commun*. 2016;7:12444.

7. Kilkenny C, Browne W, Cuthill IC, Emerson M, Altman DG, Group NCRRGW. Animal research: reporting in vivo experiments: the ARRIVE guidelines. *Br J Pharmacol*. 2010;160(7):1577- 1579.

8. Joyce PI, McGoldrick P, Saccon RA, et al. A novel SOD1-ALS mutation separates central and peripheral effects of mutant SOD1 toxicity. *Hum Mol Genet*. 2015;24(7):1883-1897.

9. Pronk CJ, Rossi DJ, Mansson R, et al. Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell Stem Cell*. 2007;1(4):428-442.