Somatic uniparental disomy with a rare *EFL1* variant causes Shwachman-Diamond syndrome through dysregulating ribosomal protein synthesis

Sangmoon Lee[1,12], Chang Hoon Shin[2,12], Che Ry Hong[3], Jun-Dae Kim[4], Ah-Ra Kim[5], Jawon Lee[1], Oleksandr Kokhan[6], Taekyeong Yoo[1], Young Bae Sohn[7], Ok-Hwa Kim[8], Jung Min Ko[3], Tae-Joon Cho[9], Nathan T. Wright[6], Suk-Won Jin[4,5], Hyoung Jin Kang[3,10,13], Hyeon Ho Kim[2,11,13], Murim Choi[1,3,13]

[1]Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Republic of Korea.
[2]Department of Health Sciences and Technology, Samsung Advanced Institute for Health Sciences and Technology, Sungkyunkwan University, Seoul, Republic of Korea.
[3]Department of Pediatrics, Seoul National University College of Medicine, Seoul, Republic of Korea.
[4]Yale Cardiovascular Research Center, Section of Cardiovascular Medicine, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT, USA.
[5]School of Life Sciences, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea.
[6]Department of Chemistry and Biochemistry, James Madison University, Harrisonburg, VA, USA.
[7]Department of Medical Genetics, Ajou University Hospital, Ajou University School of Medicine, Suwon, Republic of Korea.
[8]Department of Radiology, Woorisoa Children's Hospital, Seoul, Republic of Korea.
[9]Division of Pediatric Orthopaedics, Seoul National University Children's Hospital, Seoul, Republic of Korea.
[10]Seoul National University Cancer Research Institute, Seoul, Republic of Korea.

[11]Institute for Future Medicine, Samsung Medical Center, Seoul, Republic of Korea
[12]These authors contributed equally to this work.
[13]These authors jointly directed this work.
Correspondence should be addressed to M.C. (murimchoi@snu.ac.kr), H.H.K. (hyeonhkim@skku.edu) or H.J.K. (kanghj@snu.ac.kr).

We present three unrelated Korean Shwachman-Diamond syndrome (SDS) patients that carry an incomplete but identical homozygous *EFL1* p.Thr1069Ala variant due to a bone marrow-specific mosaic uniparental disomy (UPD) in chromosome 15. This rare variant is found in 0.017% of East Asians and is asymptomatic in a heterozygous status, but harbors a hypomorphic effect, leading to reduced 80S ribosome assembly of ribosomal protein (RP) transcripts. We propose a novel somatically-induced pathogenesis mechanism and EFL1 dysfunction that eventually leads to aberrant translational control and ribosomopathy.

SDS (OMIM: #260400) presents with hematologic manifestations, exocrine pancreatic dysfunction with fatty infiltration, and skeletal dysplasia resulting in short stature[1]. SDS is mostly caused by variants in *SBDS* (Shwachman-Bodian-Diamond syndrome gene), encoding a protein that plays an important role in ribosome assembly[1-5]. However, 10-20% of the cases clinically diagnosed with SDS do not harbor pathogenic variants in the gene, suggesting the existence of additional genetic mechanisms that lead to the disorder[1,3]. Recent reports demonstrated that variants in *DNAJC21* and *SRP54* are implicated with bone marrow failure syndrome and SDS[6-8], meanwhile, another report suggested that homozygous mutations in elongation factor-like GTPase 1 (*EFL1*) causes a SDS-like syndrome[9]. Given the known role of EFL1 in a direct interaction with SBDS that releases eukaryotic translation initiation factor 6 (eIF6) from the 60S ribosomal subunit for 80S ribosomal assembly[4,10], it is plausible that functional abnormalities in

EFL1 cause diseases similar to SDS. However, detailed molecular mechanisms and *in vivo* evidence that the impaired EFL1 causes SDS-like features are still limited.

We exome-sequenced three unrelated and non-consanguineous Korean SDS patients without plausible mutations in *SBDS* (Supplementary Tables 1 and 2, Fig. 1a and Supplementary Fig. 1). The heterozygous p.Thr1069Ala variant of *EFL1* (chr15:82,422,872 T>C, hg19, NM_024580.5:c.3205A>G) was identified in all three patients (Supplementary Table 3, Fig. 1b). It is a low frequency variant, carried by three individuals among 17,972 alleles of the East Asian population in gnomAD (allele frequency (AF) = 1.67×10^{-4})[11]. It was noted that the variant allele of I-1 and II-1 was dominantly covered compared to the reference allele (86.5 and 77.6%, respectively), making it function like an incomplete homozygous variant, while the ratio was comparable in III-1 (41.1%, Fig. 1c). Surprisingly, all patients carried a partial loss-of-heterozygosity (LOH) in chromosome 15, where *EFL1* locates (Fig 1d, Supplementary Figs. 2 and 3). This LOH was copy-neutral and not seen in healthy parents (Supplementary Fig. 4), suggesting that it was caused by somatic UPD. Furthermore, homozygosity of the *EFL1* locus is not frequently found in healthy Korean individuals (1/3,667, Supplementary Fig. 5). The sizes of the LOH internals were variable among the patients (100, 100 and 27.8% of the entire chromosome span, respectively; Fig. 1d). To test if the UPD event occurred in a mosaic pattern and whether LOH-carrying and non-LOH-carrying cells co-exist, single-cell SNP microarray was performed using bone marrow (I-1) or buccal swabbed epithelial cells (III-1). As expected, complete LOHs in chromosome 15 were observed in a subset of the cells, confirming mosaic UPD events (Supplementary Fig. 6). Therefore, those cells having UPD of chromosome 15 are expected to be *EFL1*$^{p.Thr1069Ala/p.Thr1069Ala}$. II-1 could not be tested due to sample unavailability. Next, we checked the spatial extent of the mosaic UPD by subjecting all available tissue samples of I-1 into a high-depth amplicon sequencing analysis (>40,000X coverage depths). Variant AFs of

p.Thr1069Ala in I-1 were ~0.85 in peripheral blood and bone marrow, while ~0.5 in most tissues, suggesting that the mosaic UPD is restricted at least to the bone marrow (Fig. 1e). These results were concordant with Sanger sequencing results (Supplementary Figs. 7 and 8). The degree of mosaicism of bone marrow tissue changed dynamically over the time course on I-1, without displaying a strong correlation with the clinical status of the patient (Supplementary Fig. 9). The variant residue lies in the C-terminus of the protein and is highly conserved throughout evolution (Fig. 1f). EFL1 protein structure analysis suggests that the p.Thr1069 residue is engaged in hydrogen bonding with p.Gln884, which is presumably involved in stabilizing the protein structure (Fig. 1g and Supplementary Fig. 10). Since EFL1 is known to mediate GTP hydrolysis-coupled eukaryotic translation initiation factor 6 (eIF6) release together with SBDS in maturation of the 60S ribosome subunit[4,10], we observed the ribosomal assembly status of wild-type, siRNA- or CRISPR/Cas9-mediated ablation of *EFL1* (*EFL1*^{KD} or *EFL1*$^{-/-}$) in HeLa and K562 cell lines to further elucidate the molecular mechanism of the mutant protein function. Polysome profiling of *EFL1*^{KD} and *EFL1*$^{-/-}$ cells showed a significantly reduced 80S peak (Fig. 1h and Supplementary Fig. 11). This abnormal polysome profile was completely rescued after introduction of FLAG-tagged wild-type *EFL1* (EFL1$^{-/-}$-WT), but not by the mutant *EFL1* (EFL1$^{-/-}$-T1069A; Fig. 1i and Supplementary Fig. 12). These results indicate that EFL1 plays a crucial role in ribosome assembly and that the EFL1$^{p.Thr1069Ala}$ is hypomorphic. To validate these observations *in vivo*, a morpholino-targeting zebrafish model of *efl1* was subjected to rescue experiments (Supplementary Figs. 13a-b). *efl1* morphants have smaller head and eyes as wells as a slightly bent tail (Fig. 1i), and display an increased number of apoptotic cells during development (Supplementary Fig. 13c). Both primitive erythrocytes and granulocytes were significantly reduced in *efl1* morphants, indicating impaired primitive hematopoiesis in these embryos (Fig. 1j, k, and Supplementary Fig. 13d). All the phenotypes were rescued by the introduction of wild type mRNA, but less so by the mutated mRNAs (Figs. 1i-

k), confirming that the abnormal degree of *efl1* activity causes SDS-like features *in vivo*.

We then investigated the molecular mechanism of the variant in ribosome assembly. The variant function was not mediated via phosphorylation of p.Thr1069, aberrant subcellular localization, or changes in binding affinity with SBDS (Supplementary Figs. 14 and 15). Next, measuring eIF6 level from each ribosome subunit-bound fraction revealed that it was more enriched in 40S and 60S ribosome fractions of EFL1$^{-/-}$ or EFL1$^{-/-}$-T1069A (Fig. 2a and Supplementary Fig. 15) compared to wild-type or EFL1$^{-/-}$-WT. This result, along with an observation that eIF6 and SBDS changes did not stem from a transcription level (Supplementary Fig. 16), implies that blocked exclusion of eIF6 and SBDS from a 60S ribosomal subunit is one of the mechanisms of our mutant protein action, and results in impaired 80S ribosome assembly. To further investigate features of downstream genes that are heavily affected by the reduced function of EFL1, we performed RNA sequencing on fractions from total RNA, 40S-, 80S- and polysome-bound RNAs from wild-type and *EFL1* knockout K562 cells (Fig. 2b and Supplementary Fig. 17). The 248 genes that were decreased in the 80S-bound compared to the 40S-bound fraction in the mutant cells were extracted and a gene ontology analysis strongly suggests that ribosomal protein (RP) genes (GO:0003735) are the major constituents of the genes that were influenced by the absence of *EFL1* (P_{adj} = 5.2 x 10^{-77}; Fig. 2c). There was no significantly enriched gene group in the increased 80S/40S genes. The fractions of transcripts bound by 40S-, 80S- or polysome differ substantially for the RP genes by the absence of *EFL1* (Wilcoxon signed-rank test, $P < 1.0 \times 10^{-13}$ for differences in both 40S- and 80S-bound transcripts), whereas *TP53* target genes (n = 65) or all other genes did not show such change (Fig. 2d, $P > 0.05$). To elucidate the mechanism of RP-specific regulation, we sought to compare whether transcript sizes (Supplementary Fig. 18), expression levels, or consensus sequence elements in 5' UTRs may function as determining factors. Notably, highly-expressed RP transcripts with a terminal oligo-

pyrimidine (TOP) element (5'-CUUYCUUUUNS-3') were specifically altered (Fig. 2f and Supplementary Fig. 19; $P = 6.9 \times 10^{-7}$). Subsequent western blotting of selected RPs confirmed reduced protein synthesis in *EFL1*$^{-/-}$ cells (Fig. 2g and Supplementary Fig. 20).

Assessing the functionality of variants of unknown significance poses a major hurdle in genetic studies. Here we demonstrate a process where an apparently benign, but hypomorphic variant gains pathogenicity through somatic UPD. *EFL1*-mutated patients arise in a recessive manner[9], and heterozygous LoF variants do not cause any defect (gnomAD pLI = 0.00), which suggests that a UPD-derived homozygous p.Thr1069Ala variant drives the protein function below a clinical threshold. According to a recent survey of haematopoietic chromosomal mosaicism events, 117 individuals carried copy-neutral or deletion of the *EFL1* locus among 151,202 apparently normal individuals. Multiplying this to the AF of the p.Thr1069Ala produces 1.30×10^{-7}, which is not too divergent from our odds of observing three patients from the population of Korea, with a centralized medical referral system[12]. The degrees of mosaicism in the three patients varied substantially (Fig. 1d), and does not seem to determine clinical severity (Supplementary Fig. 10). Nonetheless, an altered function of EFL1 specifically influenced translation of RP genes with a TOP element, which leads to the mechanistic mimicry of Diamond-Blackfan anemia (DBA), another ribosomopathy caused by insufficient RP dose[13] (Fig 2f and Supplementary Fig. 17). Activation of TP53 has been considered as a targetable downstream pathway that leads to SDS or a DBA phenotype[14-16]. But our data suggest that loss of *EFL1* does not induce TP53 activation, which is consistent with the previous zebrafish *sbds* model (Fig. 2e)[17]. Here we introduce a mechanism where a somatically-induced mosaic mutation in *EFL1* converted an apparently benign rare variant into a pathogenic one, and phenocopies classical SDS in three unrelated patients. We demonstrated that the defective EFL1 causes impaired 80S ribosomal assembly and the zebrafish models displayed similar features as in human through alteration of 80S ribosome assembly of RP

transcripts. An extensive search of such SDS patients may shed more insights into the development of somatic mosaicism and subsequent molecular cascades that may lead to new avenues of treatment for ribosomopathy.

Acknowledgements

We appreciate the patients and families for participating in this study. We thank Jung-Ah Kim and Hyoung-Jin Kim at Seoul National University Hospital for collecting and interpreting hematological data and for interpreting images. This study was partly supported by grants from the National Research Foundation of Korea (2014M3C9A2064686, 2014M3A6A4075061 to M.C. and 2017R1A2A2A05069691 to H.H.K.) and National Science Foundation (REU CHE-1062629 and RUI MCB-1607024) awards to N.T.W. O.K. gratefully acknowledges the computing resources provided on Bebop, a high-performance computing cluster operated by the Laboratory Computing Resource Center at Argonne National Laboratory.

Author Contributions

M.C., H.J.K. and H.H.K. conceived the project. S.L. and M.C. performed genetic analysis and statistical evaluations. C.H.S., S.L., H.H.K., J.L. and M.C. performed molecular biology and biochemistry experiments and assessed the results. J.-D.K., A.-R.K. and S.-W.J. performed zebrafish experiments. C.R.H., Y.B.S., O.-H.,K., J.M.K., T.-J.C. and H.J.K. provided patient care and generated clinical data. N.T.W., O.K. and T.Y. analyzed protein structure. S.L., C.H.S., C.R.H., M.C., H.J.K. and H.H.K. wrote the manuscript. All authors approved the final version of the manuscript.

References

1. Boocock, G. R. B. *et al.* Mutations in SBDS are associated with Shwachman-Diamond syndrome. *Nat. Genet.* **33,** 97–101 (2003).
2. García-Márquez, A., Gijsbers, A., la Mora, de, E. & Sánchez-Puig, N. Defective Guanine Nucleotide Exchange in the Elongation Factor-like 1 (EFL1) GTPase by Mutations in the Shwachman-Diamond Syndrome Protein. *J Biol Chem* **290,** 17669–17678 (2015).
3. Hashmi, S. K. *et al.* Comparative analysis of Shwachman-Diamond syndrome to other inherited bone marrow failure syndromes and genotype-phenotype correlation. *Clinical Genetics* **79,** 448–458 (2011).
4. Weis, F. *et al.* Mechanism of eIF6 release from the nascent 60S ribosomal subunit. *Nat. Struct. Mol. Biol.* **22,** 914–919 (2015).
5. Menne, T. F. *et al.* The Shwachman-Bodian-Diamond syndrome protein mediates translational activation of ribosomes in yeast. *Nat. Genet.* **39,** 486–495 (2007).
6. Tummala, H. *et al.* DNAJC21 Mutations Link a Cancer-Prone Bone Marrow Failure Syndrome to Corruption in 60S Ribosome Subunit Maturation. *The American Journal of Human Genetics* **99,** 115–124 (2016).
7. Dhanraj, S. *et al.* Biallelic mutations in DNAJC21 cause Shwachman-Diamond syndrome. *Blood* **129,** 1557–1562 (2017).
8. Carapito, R. *et al.* Mutations in signal recognition particle SRP54 cause syndromic neutropenia with Shwachman-Diamond-like features. *J. Clin. Invest.* **127,** 4090–4103 (2017).
9. Stepensky, P. *et al.* Mutations in EFL1, an SBDS partner, are associated with infantile pancytopenia, exocrine pancreatic insufficiency and skeletal anomalies in aShwachman-Diamond like syndrome. *Journal of Medical Genetics* **54,** 558–566 (2017).
10. Finch, A. J. *et al.* Uncoupling of GTP hydrolysis from eIF6 release on the ribosome causes Shwachman-Diamond syndrome. *Genes & Development* **25,** 917–929 (2011).
11. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536,** 285–291 (2016).
12. Loh, P.-R. *et al.* Insights into clonal haematopoiesis from 8,342 mosaic chromosomal alterations. *Nature* **559,** 350–355 (2018).
13. Draptchinskaia, N. *et al.* The gene encoding ribosomal protein S19 is mutated in Diamond-Blackfan anaemia. *Nat. Genet.* **21,** 169–175 (1999).
14. Danilova, N., Sakamoto, K. M. & Lin, S. Ribosomal protein S19 deficiency in zebrafish leads to developmental abnormalities and defective erythropoiesis through activation of p53 protein family. *Blood* **112,** 5228–5237 (2008).
15. McGowan, K. A. *et al.* Ribosomal mutations cause p53-mediated dark skin and pleiotropic effects. *Nat. Genet.* **40,** 963–970 (2008).

16. Elghetany, M. T. & Alter, B. P. p53 protein overexpression in bone marrow biopsies of patients with Shwachman-Diamond syndrome has a prevalence similar to that of patients with refractory anemia. *Arch. Pathol. Lab. Med.* **126,** 452–455 (2002).
17. Provost, E. *et al.* Ribosomal biogenesis genes play an essential and p53-independent role in zebrafish pancreas development. *Development* **139,** 3232–3241 (2012).

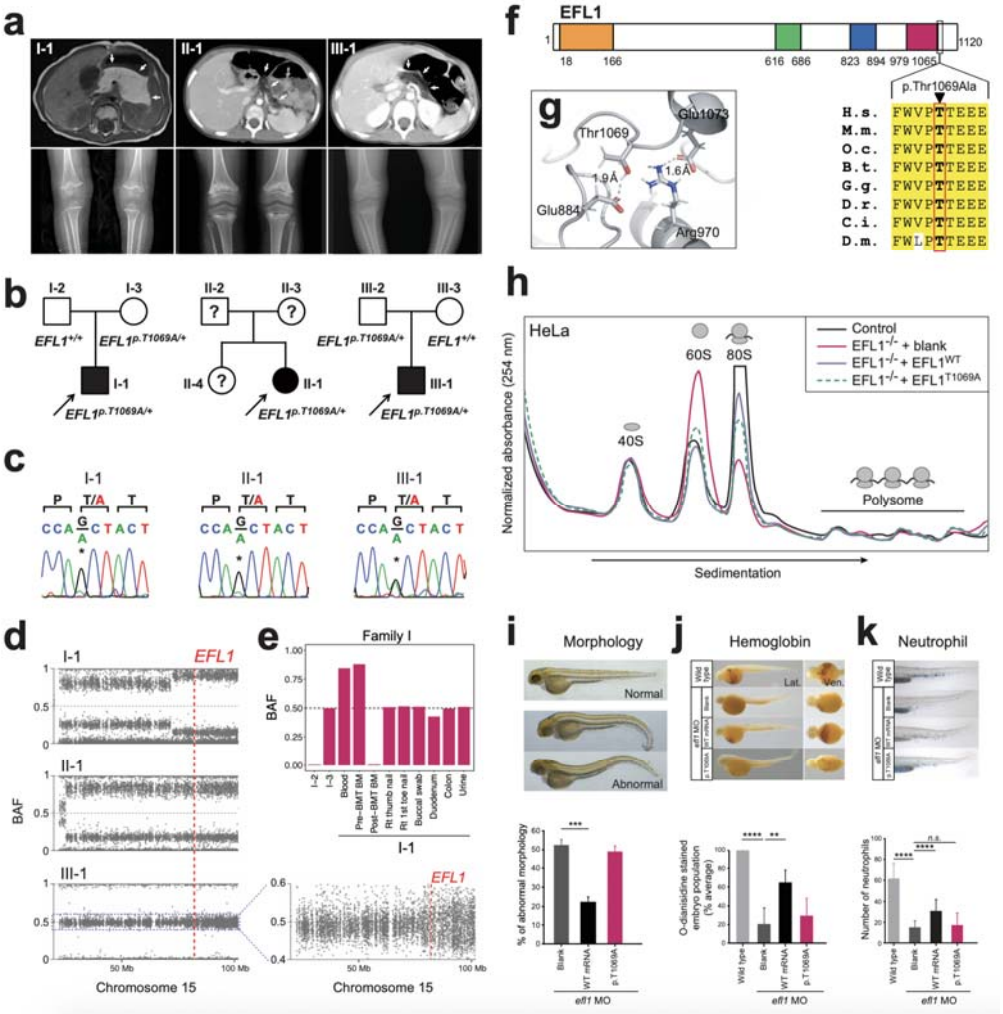

Figure 1. SDS caused by mosaic *EFL1* p.Thr1069Ala. (**a**) Abdominal MRI or CT images showing diffuse enlargement with lipomatosis of the pancreas (arrow) (upper) and X-ray images of both knees showing bilateral metaphyseal widening of femurs (I-1) or irregularity and genu varum (III-1) of the three patients (lower). (**b**) Pedigrees of three families. (**c**) Sanger sequencing traces showing dominant variant allele in the affected individuals. (**d**) All three patients have partial LOH in chromosome 15, where *EFL1* locates. (**e**) Variant AF of *EFL1* variant (chr15:82,422,872 T>C) in multiple samples and tissues from I-1. (**f**) The mutated residue of EFL1 is located in a

C-terminus domain and is evolutionary conserved. (**g**) Structural analysis of EFL1 p.Thr1069 residue based on PDB 5ANC[4]. (**h**) Polysome profiling results of *EFL1*$^{+/+}$ and *EFL1*$^{-/-}$ cells followed by *EFL1*^{WT} or *EFL1*$^{p.Thr1069Ala}$ rescue. (**i**) Phenotypic changes of zebrafish by *efl1* knockdown and rescue experiments by morpholino. (**j**) Hemoglobin stained by O-dianisidine of 48 hpf embryos to measure primitive hematopoiesis. (**k**) Neutrophils stained by sudan black of 48 hpf embryos. MO, morpholino; hpf, hours post fertilization; B.t., *Bos taurus*; C.i., *Ciona intestinalis*; D.m., *Drosophila melanogaster*; D.r., *Danio rerio*; G.g., *Gallus gallus*; H.s., *Homo sapiens*; M.m., *Mus musculus*; O.c., *Oryctolagus cuniculus*.

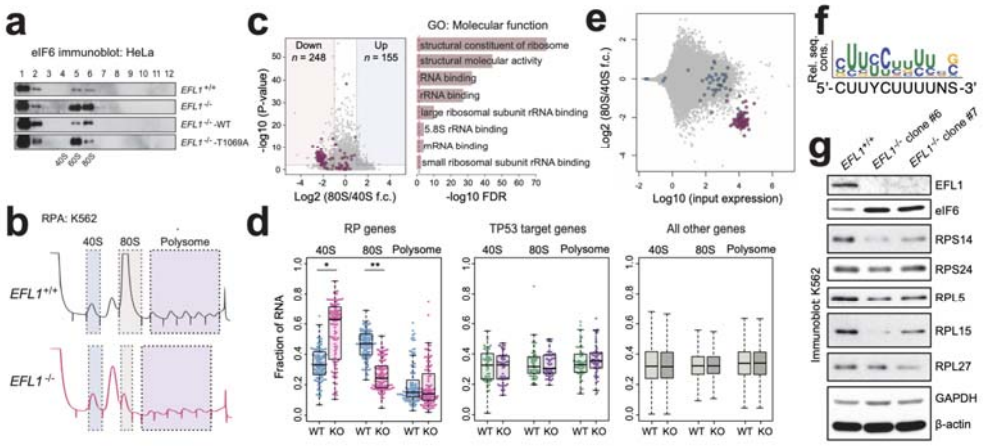

Figure 2. EFL1 is required for eIF6 release from a 60S ribosome and ribosomal protein synthesis. (**a**) Immunoblots of eIF6 from each ribosome fraction. (**b**) Schematic diagram of RNA expression analysis from K562 ribosomal fractions. (**c**) Volcano plot of genes that are differentially enriched in the 80S fraction compared to the 40S fraction (left). GO analysis of the 248 downregulated genes (right). Red dots depict RP genes (GO:0022625 and GO:0022627, *n* = 118). (**d**) Fractions of 40S-, 80S- or polysome-bound RNAs for RP genes (*n* = 108, left), *TP53* target genes (*n* = 51, middle) and all other genes (*n* = 11,812, right) are displayed. Wilcoxon signed-rank test, *. $P = 4.7 \times 10^{-14}$, **. $P = 1.5 \times 10^{-26}$. All other test *P*-values > 0.05. (**e**) Fold change of 80S enriched genes normalized by 40S, depicted against transcriptome. Red; significantly downregulated RP genes in Fig. 2c, blue; other RP genes. (**f**) Consensus sequence profile of downregulated RP gene 5'UTRs, deducing a TOP sequence. Rel. seq. cons. denotes relative sequence conservation from MEME run. (**g**) Immunoblots of selected RPs from K562 wild type or *EFL1*$^{-/-}$ cell lines.