Published in final edited form as:

Nat Cell Biol. 2021 June ; 23(6): 595–607. doi:10.1038/s41556-021-00688-9.

ELOF1 is a transcription-coupled DNA repair factor that directs RNA polymerase II ubiquitylation

Yana van der Weegen[1,*], **Klaas de Lint**[2,*], **Diana van den Heuvel**[1], **Yuka Nakazawa**[3,4], **Tycho E.T. Mevissen**[5], **Janne J.M. van Schie**[2], **Marta San Martin Alonso**[1,9], **Daphne E.C. Boer**[1], **Román González-Prieto**[6], **Ishwarya V. Narayanan**[7], **Noud H.M. Klaassen**[1], **Annelotte P. Wondergem**[1], **Khashayar Roohollahi**[2], **Josephine C. Dorsman**[2], **Yuichiro Hara**[3,4], **Alfred C.O. Vertegaal**[6], **Job de Lange**[2], **Johannes C. Walter**[5], **Sylvie M. Noordermeer**[1,9], **Mats Ljungman**[7,8], **Tomoo Ogi**[3,4], **Rob M.F. Wolthuis**[2,†], **Martijn S. Luijsterburg**[1,†]

[1] Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

[2] Department of Clinical Genetics, Section Oncogenetics, Cancer Center Amsterdam, Amsterdam University Medical Center, Amsterdam, the Netherlands

[3] Department of Genetics, Research Institute of Environmental Medicine (RIeM), Nagoya University, Nagoya, Japan

[4] Department of Human Genetics and Molecular Biology, Nagoya University Graduate School of Medicine, Nagoya, Japan

[5] Howard Hughes Medical Institute and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA.

[6] Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands

[7] Department of Radiation Oncology, University of Michigan, Ann Arbor, MI, USA

[†]Correspondence to: m.luijsterburg@lumc.nl; r.wolthuis@amsterdamumc.nl.
[*]These authors contributed equally to this work.

Author contributions:

YvdW generated plasmids, U2OS and RPE1-iCas9 single KO and double KO cells, U2OS Flp-In cell-lines, performed clonogenic survivals, co-IP experiments, RRS and DRB-RRS experiments, western blot analyses, in situ proximity-ligation assays, generated samples for pan-RNAPII ChIP-seq, ser2-RNAPII ChIP-seq, ATAC-seq, BruDRB-seq, and Bru-seq, generated the figures and wrote the paper. KdL optimized, performed and analyzed all CRISPR screens and RNA-seq experiments, performed and analyzed drug-sensitivity assays, generated gene-gene interaction network and Venn diagrams, performed γH2AX foci and DNA fiber experiments and helped writing the paper. DvdH performed Co-IP experiments, developed tools and analyzed pan-RNAPII ChIP-seq, ser2-RNAPII ChIP-seq, ATAC-seq, BruDRB-seq, Bru-seq and helped writing the paper. YN performed Co-IP experiments and generated TCR-seq libraries. TETM generated recombinant xlELOF1, xlCSB, and xlCRL4^{CSA}, purified xlRNAPII and performed pull-down and in vitro ubiquitylation assays. JJMvS created RPE1-iCas9 cell line, performed CRISPR screens, performed γH2AX foci and DNA fiber analyses. MSMA performed and analyzed DRIP-qPCR. IVN analyzed BruDRB-seq and Bru-seq. DECB generated stable cell-lines and performed clonogenic survival assays. RG-P analyzed the mass spectrometry samples with support from ACOV. NK generated U2OS single KO clones, Flp-In cell-lines, performed clonogenic survivals and Co-IP experiments. APW performed UDS experiments and generated plasmids. KR processed and analyzed RNA-seq data, performed gene-gene interaction network analysis, and performed and analyzed CRISPR screens. YH analyzed ser2-RNAPII ChIP-seq. JdL supervised JJMvS. JCD supervised KR. JCW supervised TETM. SMN supervised MSMA. ML supervised IVN, and analyzed BruDRB-seq and Bru-seq. TO supervised YN and YH, and analyzed Ser2-RNAPII ChIP-seq. RMFW supervised KdL, co-supervised KR and JJMvS, conceived, coordinated and supervised the project and helped writing the paper. MSL supervised YvdW, DvdH, DECB, NK, APW, conceived, coordinated, and supervised the project, generated all cryo-EM pictures and wrote the paper.

Competing interests: Authors declare no competing interests.

[8] Department of Environmental Health Sciences, Rogel Cancer Center and Center for RNA Biomedicine, University of Michigan, Ann Arbor, MI, USA

[9] Oncode Institute, Utrecht, The Netherlands

Abstract

Cells employ transcription-coupled repair (TCR) to eliminate transcription-blocking DNA lesions. DNA damage-induced binding of the TCR-specific repair factor CSB to RNA polymerase II (RNAPII) triggers RNAPII ubiquitylation of lysine K1268 by the CRL4^{CSA} ubiquitin ligase. How CRL4^{CSA} is specifically directed toward K1268 is unknown. Here, we identify ELOF1 as the missing link that facilitates RNAPII ubiquitylation, a key signal for the assembly of downstream repair factors. This function requires its constitutive interaction with RNAPII close to K1268, revealing ELOF1 as a specificity factor that binds and positions CRL4^{CSA} for optimal RNAPII ubiquitylation. Drug-genetic interaction screening also reveals a CSB-independent pathway in which ELOF1 prevents R-loops in active genes and protects cells against DNA replication stress. Our study offers key insights into the molecular mechanisms of TCR and provides a genetic framework of the interplay between transcriptional stress responses and DNA replication.

Transcription of protein-coding and non-coding genes requires RNA polymerase II (RNAPII), which synthesizes RNA transcripts complementary to the DNA template strand. The presence of DNA lesions in the template strand causes stalling of elongating RNAPII (RNAPIIo) leading to a genome-wide transcriptional arrest[1–3]. It is essential that cells overcome this arrest and restore transcription. The transcription-coupled repair (TCR) pathway efficiently removes transcription-blocking DNA lesions through the CSB, CSA, and UVSSA proteins[4–6]. Inactivating mutations in *CSB* and *CSA* cause Cockayne syndrome (CS), which is characterized by severe neurological dysfunction related to persistent RNAPII arrest at DNA lesions[2, 7].

The sequential and cooperative actions of CSB, CSA, and UVSSA recruit TFIIH to DNA damage-stalled RNAPII to initiate DNA repair[6]. In addition to protein-protein contacts, efficient transfer of TFIIH onto RNAPII requires ubiquitylation of a single lysine on the largest subunit of RNAPII (RPB1-K1268), which is essential for efficient TCR[2]. This DNA damage-induced modification of RNAPII is dependent on cullin-ring type E3-ligases (CRLs) and strongly decreased in CSA-deficient cells[2], indicating that the CRL4^{CSA} E3 ligase complex drives RNAPII ubiquitylation.

CSB binds to DNA upstream of RNAPII[8] (Extended Data Fig. 1a) and recruits the CRL4^{CSA} complex through an evolutionary conserved motif in its C-terminus[6]. However, how the CRL4^{CSA} ubiquitin ligase activity is specifically directed toward the K1268 site remains to be elucidated.

Results

A CRISPR screen identifies *ELOF1* as a putative TCR gene

To identify unknown TCR genes, we performed a genome-wide CRISPR screen in the presence of the compound Illudin S, which induces transcription-blocking DNA lesions that

are eliminated by TCR[9]. RPE1-iCas9 cells were transduced with the pLCKO-TKOv3 library containing 70,948 ORF-targeting sgRNAs[10] and cultured for 12 population doublings after which sgRNA contents were analyzed (Extended Data Fig. 1b).

Using an FDR cutoff of 0.01, we found 104 sensitizer hits and 18 hits conferring resistance to Illudin S (see source data). The strongest resistance was conferred by gRNAs targeting *PTGR1,* in line with its known role in bio-activating Illudin[11] (Fig. 1a, Extended Data Fig. 1c). Nine known core TCR genes, including *CSB*, *CSA* and *UVSSA*, but also genes connected to transcription recovery after UV (*HIRA*[12], *DOT1L*[13], and *STK19*[14]; Fig. 1a, b) were required for Illudin S tolerance. Consistent with known effects of Illudin S on replication[9], we found the 9–1-1 complex, translesion synthesis and sister-chromatid cohesion components (Fig. 1b). Our screen also identified the uncharacterized *ELOF1* gene as a top hit, which we decided to study in detail (Fig. 1a, b).

We generated single *PTGR1*, *ELOF1* and *CSB* knockouts in RPE1-iCas9 cells (Extended Data Fig. 1d) and exposed these to compounds that generate transcription stress. Drug-sensitivity assays confirmed that single *ELOF1*-KO and *CSB*-KO clones were highly sensitive to Illudin S, Irofulven, and cisplatin, while *PTGR1*-KO cells were resistant to Illudin S and Irofulven (Fig. 1c, Extended Data Fig. 1e). Clonogenic survival experiments confirmed that *ELOF1*-KO cells are nearly as sensitive to Illudin S as *CSB*-KO cells (Extended Data Fig. 1f). Re-expression of GFP-tagged ELOF1 in *ELOF1*-KO U2OS cells completely restored their Illudin S tolerance (Fig. 1d, Extended Data Fig. 2a, b).

ELOF1 interacts with RNAPII through conserved residues

ELOF1 is a small zinc-finger protein that is evolutionary conserved across eukaryotes, except for its C-terminal acidic tail, which is absent in metazoan orthologues (Extended Data Fig. 2c)[15]. ELF1, the yeast orthologue of ELOF1, is a putative transcription elongation factor that interacts with RNAPII, through a central RNAPII-binding helix (Fig. 1e), and promotes passage through nucleosomes[16, 17]. Mutations in residues that disrupt this interaction (ELF1$^{S71K/D72K}$) compromise its function[16, 17].

To study human ELOF1, we immunoprecipitated GFP-tagged ELOF1 variants followed by mass spectrometry (MS). Pulldown of ELOF1^{WT} revealed an interaction with nine RNAPII subunits and three transcription elongation factors (Fig. 1f, Extended Data Fig. 2d). The ELOF1$^{S72K/D73K}$ (S71K/D72K in yeast) mutant showed a severely reduced interaction with RNAPIIo (Fig. 1g, Extended Data Fig. 2d), which was confirmed by reciprocal pull-down of RNAPIIo (Fig. 1h). The RNAPIIo-ELOF1 interaction was constitutive and not affected by UV irradiation (Fig. 1h, Extended Data Fig. 2e). Re-expression of ELOF1^{WT}, but not ELOF1$^{S72K/D73K}$, fully rescued the Illudin S-sensitive phenotype of *ELOF1*-KO cells (Fig. 1i, Extended Data Fig. 2f). These findings reveal similar RNAPII-binding modes between yeast ELF1 and human ELOF1 and show that the ELOF1-RNAPII interaction protects against transcription-blocking DNA damage.

Loss of ELOF1 decreases RNAPII elongation rates

To identify the framework of genetic interactions with ELOF1, we performed a genome-wide CRISPR screen in *ELOF1*-KO RPE1-iCas9 cells. This approach identified synthetic

lethality between *ELOF1* and several transcription elongation factors, including *SUPT4H1*, *TCEA1, SUPT6H* and the PAF1 complex subunits *CTR9* and *LEO1* (Fig. 2a, b). This is consistent with genetic studies showing that deletion of yeast ELF1 causes synthetic lethality with mutations in elongation factors[18], likely caused by functional redundancy[16, 17]. Several of these elongation factors also reside in a protein complex with ELOF1 (Fig. 1f).

To investigate a role of ELOF1 in transcription elongation, we performed genome-wide BruDRB-seq[19]. Cells were treated with the transcription inhibitor 5,6-dichlorobenzimidazole1-β-D-ribofuranoside (DRB) for 3,5 h followed by its removal to trigger synchronized transcription elongation into gene bodies. Nascent transcripts were labelled with bromo-uridine (BrU), isolated and deep-sequenced[19]. Averaging nascent RNA profiles of 400 genes revealed that RNAPII progressed ~60 kb within 30 min in both WT and *ELOF1*-KO cells (Fig. 2c, d, e). Within 60 min, RNAPII progressed ~120 kb in WT cells, but only ~90 kb in *ELOF1*-KO cells (Fig. 2c, d, e). Transcription wave-front measurements revealed that RNAPII elongates with ~2 kb/min in WT cells and ~1.5 kb/min in *ELOF1*-KO cells (Extended Data Fig. 3a).

Basal transcription after 5-EU incorporation was reduced to ~40% in *ELOF1*-KO cells compared to WT cells (Fig. 2f, g). Analysis of 400 genes of at least 100 kb by Bru-seq revealed no difference in initiation while transcription decreased progressively toward the end of longer genes in *ELOF1*-KO cells (Fig. 2h), which was milder for shorter genes (Extended Data Fig. 3b). Consistently, *ELOF1*-KO cells did not display differences in chromatin accessibility at transcription start sites (TSSs) of ~3,000 genes measured by ATAC-seq (Extended Data Fig. 3c). These findings show that human ELOF1 acts as a transcription elongation factor.

ELOF1 is required for transcription restart following UV

TCR-deficient cells fail to resume transcription following genotoxic stress[20]. To assess whether ELOF1 is involved in transcription restart, we measured recovery of RNA synthesis (RRS) by 5-EU labelling after UV irradiation. Transcriptional arrest was visible 3 h after UV irradiation in WT, *ELOF1*-KO and *CSB*-KO cells. Although WT cells completely recovered RNA synthesis within 24 h after UV, both *CSB*-KO and *ELOF1*-KO cells failed to restart transcription (Fig. 3a, b, Extended Data Fig. 4a). Importantly, this defect is not due to decreased transcription elongation since *ELOF1*-KO cells were able to restart transcription after treatment with transcription inhibitor DRB (Extended Data Fig. 4b). Re-expression of ELOF1^{WT} fully rescued transcription restart after UV irradiation in *ELOF1*-KO cells, while expression of ELOF1$^{S72K/D73K}$ did not (Fig. 3a, b, Extended Data Fig. 4c).

To explore the recovery of RNA synthesis in a genome-wide manner, we labelled nascent transcripts with BrU after UV followed by RNA isolation and sequencing. At 3 h after UV, nascent transcription was strongly reduced at TSSs and progressively decreased further into gene bodies with a loss beyond 100 kb in both WT and *ELOF1*-KO cells (Fig. 3c, d). Although milder, a similar loss of nascent transcription was also observed toward the end of shorter genes (25–50 kb and 50–100 kb; Extended Data Fig. 4d), which fits with the distribution of photolesions[21] and the probability that RNAPII encounters them. Nascent transcription was partially resumed at 8 h and fully restored at 24 h after UV irradiation in

WT cells, while *ELOF1*-KO cells failed to recover transcription at both time-points (Fig. 3c, d, e).

Loss of ELOF1 causes stalling of RNAPII after UV

Genotoxic stress induces expression of the transcriptional repressor ATF3, which downregulates ~5,000 genes[22]. Both CSB and CSA promote ATF3 degradation, thereby relieving its inhibitory impact on transcription initiation[22]. Western blot analysis revealed an increase in steady-state ATF3 levels 7 h after UV in WT, *ELOF1*-KO and *CSB*-KO cells (Fig. 3f). ATF3 protein levels decreased in WT cells within 48 h after UV, while ATF3 levels strongly increased in both *ELOF1*-KO and *CSB*-KO cells at 24 h and 48 h after UV (Fig. 3f). These findings may explain the reduced RNA synthesis at TSSs following UV irradiation in *ELOF1*-KO cells (Fig. 3d, Extended Data 4d). RNA-sequencing (RNA-seq) showed that ATF3 transcript levels were strongly upregulated 24 h after UV in *ELOF1*-KO and *CSB*-KO cells compared to WT cells (Extended Data Fig. 5a–c), indicating that UV-mediated ATF3 induction is at least in part controlled at the mRNA synthesis level. While *ELOF1*-KO and *CSB*-KO cells showed a strong upregulation of short pro-survival genes, such as *CDKN1A*[23], a set of longer genes was downregulated at 24 h after UV (Extended Data Fig. 5a–d), consistent with the RNA synthesis recovery defect in these cells (Fig. 3a, b).

To monitor genome-wide occupancy of RNAPII after UV irradiation, we performed ChIP-seq using antibodies against unmodified RNAPII (pan-RNAPII). Heatmaps around the TSS showed a strong signal for ~3,000 genes around the promoter in non-irradiated cells, which was comparable between wild-type and *ELOF1*-KO cells (Extended Data Fig. 6a). At 1 h and 8 h after UV, we detected a striking reduction in RNAPII binding at TSSs in *ELOF1*-KO cells, but not in WT cells, which was accompanied by an increase in RNAPII reads in the gene body (Fig. 3g, Extended Data Fig. 6a, b). A similar redistribution of RNAPII after UV irradiation was detected in *CSB*-KO cells, suggesting that this corresponds to increased stalling at DNA lesions in gene bodies (Fig. 3g). This is consistent with our Bru-seq data showing that RNA synthesis is already largely restored around TSSs at 8 h after UV in WT cells, but not in *ELOF1*-KO cells (Fig. 3c, d).

We also detected enrichment of pan-RNAPII after transcription termination sites (TTSs) in non-irradiated cells, reflecting post-transcriptional pausing prior to dissociation (Extended Data Fig. 6b, c). *ELOF1*-KO cells showed poor TTS enrichment at 8 h after UV irradiation compared to WT cells, suggesting that fewer transcripts reach the end of genes (Fig. 2h, Extended Data Fig. 6b, c). ChIP-seq experiments using antibodies against elongating RNAPII (Ser2-RNAPII) revealed that reads throughout gene-bodies and at TTSs indeed recovered within 16 h after UV in WT cells, but not in *ELOF1*-KO and *CSB*-KO cells (Fig. 3h).

TCR-seq reveals that ELOF1 is an essential TCR factor

To determine whether ELOF1 is directly involved in repair or regulates subsequent restart of transcription, we measured genome-wide TCR kinetics using our recently developed strand-specific ChIP-seq technology (TCR-seq)[2]. The DNA fragments that are co-purified with

RNAPIIo are highly enriched for UV-induced photolesions specifically in the transcribed (template) strand, but not in the coding strand[1]. TCR-seq allows strand-specific PCR amplification of fragments without DNA damage, enabling genome-wide quantification of TCR kinetics (Fig. 4a).

Metaplots revealed a strong strand-bias in WT, *ELOF1*-KO, and *CSB*-KO cells 1 h after UV irradiation, which was fully resolved in WT cells over the course of 16 h, while both *ELOF1*-KO and *CSB*-KO cells displayed a strong strand-bias at all time-points analyzed (Fig. 4b, c, Extended Data Fig. 7a). Resolving the strand-bias in WT cells was accompanied by the reappearance of Ser2-RNAPII reads throughout gene bodies and after the TTS, which did not occur in the absence of *ELOF1* or *CSB* (Fig. 4b, c, Extended Data Fig. 7a). Analysis of the genome-wide removal of DNA lesions from transcribed strands (Extended Data Fig. 8a), showed that WT cells cleared DNA lesions within 16 h, while *ELOF1*-KO and *CSB*-KO cells displayed no significant repair within this timeframe (Fig. 4d). Strand-specific analysis after ChIP-seq with another (pan-RNAPII) antibody confirmed these results (Extended Data Fig. 9a, b). Importantly, *ELOF1*-KO cells were fully capable of removing UV-induced DNA lesions by global genome repair (GGR; Extended Data Fig. 10a, b), demonstrating that ELOF1 is specifically involved in TCR-mediated removal of DNA lesions.

ELOF1 is required for RNAPII ubiquitylation

To monitor whether ELOF1 is involved in TCR complex assembly, we immunoprecipitated endogenous RNAPIIo to isolate intact TCR complexes[6]. Both CSB and CSA associated with RNAPII in a UV-specific manner in WT and *ELOF1-KO* cells, while the association of CSA with RNAPIIo was completely abolished in cells lacking *CSB* (Fig. 5a)[6]. The constitutive interaction between ELOF1 and RNAPII was not affected by depletion of CSB (Extended Data Fig. 10c). The TFIIH subunits (p89 and p62) showed a UV-induced interaction with RNAPIIo in WT cells, while this interaction was severely reduced in cells lacking *ELOF1* (Fig. 5a, b). Importantly, re-expressing ELOF1-GFP in *ELOF1-KO* cells restored TFIIH recruitment (Fig. 5b).

RNAPII ubiquitylation on lysine K1268 (RPB1-K1268) in response to UV irradiation is required for the recruitment of the TFIIH complex[2]. Considering that TFIIH recruitment is reduced in *ELOF1*-KO cells, we asked whether ELOF1 is involved in the UV-induced ubiquitylation of RNAPII. Immunoprecipitation showed that RNAPII is robustly ubiquitylated in response to UV in WT cells, which was largely absent in cells lacking either *ELOF1* or *CSB* (Fig. 5c, d). The association of endogenous UVSSA with DNA damage-stalled RNAPIIo was strongly decreased in *ELOF1-KO* cells compared to WT cells, and virtually absent in *CSB-KO* cells (Fig. 5c). In agreement, the accompanying paper by Geijer *et al.* shows similarly defective UVSSA recruitment by live-cell imaging after UV irradiation. Interestingly, recruitment of mono-ubiquitylated UVSSA was particularly affected in *ELOF1*-KO cells[2] (Fig. 5c). This is in line with previous work demonstrating that RNAPII ubiquitylation supports the ubiquitylation of UVSSA and that, in turn, UVSSA ubiquitylation is required for TFIIH recruitment[2]. The loss of UV-induced RNAPII ubiquitylation and UVSSA mono-ubiquitylation, and the reduced association of TFIIH with

DNA damage-stalled RNAPIIo in *ELOF1*-KO cells were fully restored by stable expression of ELOF1^{WT}, while expression of ELOF1$^{S72K/D73K}$ failed to do so (Fig. 5e).

ELOF1 interacts with the CRL4^{CSA} complex

The UVSSA-binding partner USP7 was previously found to protect CSB from UV-induced degradation[5, 24, 25]. Considering that UVSSA recruitment is decreased in *ELOF1*-KO cells, we hypothesized that ELOF1 may indirectly affect RNAPII ubiquitylation by affecting CSB levels. Western blot analysis revealed that CSB is not degraded in response to UV irradiation in any of the cell lines (WT, *CSA*-KO, *ELOF1*-KO, or *UVSSA*-KO), but we noticed that CSA is degraded in response to UV irradiation in *ELOF1-KO* cells and *UVSSA*-KO cells, but not in WT or *CSB-KO* cells (Fig. 6a, Extended Data Fig. 10d). It is feasible that CSA is no longer de-ubiquitylated by the UVSSA-binding partner USP7 in *ELOF1*-KO cells. Indeed, WT cells treated with the USP7 inhibitor FT671 were unable to recover transcription after UV irradiation (Extended Data Fig. 10e)[24] and showed increased degradation of CSA after UV irradiation, but not to the same extent as in *ELOF1*-KO cells (Extended Data Fig. 10f), suggesting both USP7-dependent and -independent mechanisms. Treatment of *ELOF1*-KO cells with the proteasome inhibitor MG132 fully restored CSA protein levels after UV irradiation, but failed to rescue RNAPII ubiquitylation in *ELOF1*-KO cells (Fig. 6b, c), demonstrating that this is not an indirect effect caused by CSA degradation.

The yeast ortholog ELF1 is bound to RNAPII in close proximity to the K1268 ubiquitylation site (K1264 in *K. pastoris*; Fig. 6d). We therefore hypothesized that ELOF1 brings the ubiquitin ligase activity of the CRL4^{CSA} complex in close proximity of the K1268 site on RNAPII. To test this, we employed a recombinant approach using purified *Xenopus laevis* (xl) proteins. Immobilized GST-tagged xlELOF1 interacted robustly with purified xlRNAPII *in vitro* (Fig. 6e), while it failed to associate with the GGR factor xlRAD23B (Extended Data Fig. 10g). A robust interaction was detected between immobilized xlELOF1 and the xlCRL4^{CSA} complex, consisting of CSA, DDB1, CUL4A and RBX1 (Fig. 6f). Despite this direct interaction, the CRL4^{CSA} complex did not ubiquitylate xlELOF1 *in vitro*, whereas xlCSB was efficiently ubiquitylated (Extended Data Fig. 10h). To verify the interaction between ELOF1 and CRL4^{CSA}, we applied *in situ* proximity-ligation assays (PLA) between TY1 epitope-tagged ELOF1 and the RBX1 subunit of CRL4^{CSA}. Nuclear TY1 staining was detected in *ELOF1*-KO cells rescued with stable expression of ELOF1-TY1 (Extended Data Fig. 10i), which fully restored their Illudin S resistance (Extended Data Fig. 2f). Endogenous RBX1, but not ELOF1-TY1, was recruited to sites of local UV-induced DNA damage (Extended Data Fig. 10j). In 50% of the cells an increased PLA signal between RBX1 and ELOF1 was detected at these DNA damage sites (Fig. 6g, h), which was ~2-fold higher than at other nuclear areas (Fig. 6i). Our findings suggest that ELOF1 interacts directly with CRL4^{CSA} and stimulates the CRL4^{CSA}-dependent ubiquitylation of RNAPII K1268 (Fig. 6j)

ELOF1 is involved in a CSB-independent repair pathway

To further compare the roles of CSB and ELOF1 in response to DNA damage, we performed CRISPR screens in *CSB*-KO and *ELOF1*-KO cells in the presence of Illudin S. Loss of *CSA* and *UVSSA* in *CSB*-KO cells was epistatic, while loss of *ELOF1* caused an additive sensitivity to Illudin S (FDR<0.001, Fig. 7a). The screen in *ELOF1*-KO cells confirmed

that loss of core TCR genes *CSA*, *CSB* and *UVSSA* caused an additive sensitivity to Illudin S (FDR<0.01). Loss of *STK19*, previously linked to transcription recovery after UV[14] and sensitizing RPE1 cells to Illudin S (Fig. 1a, b)[26], did not cause additional Illudin S sensitivity in either *CSB*-KO or *ELOF1*-KO cells (Fig. 7a, b). Both *CSB*-KO and *ELOF1*-KO cells revealed a strong dependency on *REV7* (*MAD2L2*) and the 9–1-1 complex (Fig. 7a, b), suggesting that these KO cells require translesion synthesis to deal with unresolved DNA damage during replication. Loss of the *PTEN* tumor suppressor reduced the cytotoxic effects of transcriptional stress in both *ELOF1*-KO and *CSB*-KO cells (Fig. 7a, b). Furthermore, loss of the deubiquitylase *OTUD5*, which inhibits RNAPII elongation in response to DNA double-strand breaks[27], alleviates the sensitivity of *ELOF1*-KO cells to Illudin S, while its loss has the opposite impact on *CSB*-KO cells (Fig. 7a, b).

To validate these genetic interactions, we generated *CSB/ELOF1* and *CSB/CSA* double knockout (dKO) cells. Drug-sensitivity assays confirmed that *CSB/ELOF1*-dKO cells were more sensitive to Illudin S and Irofulven than either single knockout, while this was not the case for *CSB/CSA*-dKO cells (Fig. 7c, d). We propose that *ELOF1*-KO cells are not fully TCR deficient, as they still benefit from remaining CSB expression. Conversely, *CSB*-KO cells rely on ELOF1 for proliferation in the presence of Illudin S-induced DNA damage. Our network analyses show that *CSB*-KO and *ELOF1*-KO cells have overlapping, but also unique genetic dependencies (Fig. 7e). Based on these observations, we hypothesize that in addition to its role in canonical TCR, ELOF1 also acts in a second DNA repair pathway.

ELOF1 regulates a replication stress pathway

Recently published genome-wide CRISPR screens in the presence of 27 genotoxic agents suggested that loss of *ELOF1*, but not of *CSB*, *CSA*, or *UVSSA* causes sensitivity to compounds that interfere with DNA replication (Extended Data Fig. 10k)[26]. In line with this, we found that *ELOF1*-KO cells showed increased sensitivity to a specific inhibitor of DNA polymerase α (CD437) in clonogenic survival experiments, while WT and *CSB*-KO cells were similarly sensitive to CD437 (Fig. 8a). To investigate S-phase-specific DNA repair functions of ELOF1, we monitored γH2AX foci in EdU-positive cells. Both *CSB*-KO and *ELOF1*-KO cells showed a marked increase in γH2AX foci in S-phase cells after Illudin S treatment. This induction increased even further in *CSB/ELOF1*-dKO cells (Fig. 8b, c). Formation of γH2AX foci in non-replicating cells was much lower and poorly induced in *CSB*-KO, *ELOF1*-KO or *CSB*/*ELOF1*-dKO cells (Fig. 8b, c). DNA fiber analyses demonstrated that DNA replication forks progressed at normal rates (~1 kb/min) in mock-treated cells. However, fork speed declined upon 25 nM Illudin S treatment in *CSB*-KO cells and *ELOF1*-KO cells, but not in WT cells. Importantly, knocking out *CSB* and *ELOF1* together significantly reduced fork speed even at low Illudin S concentrations (5 nM; Fig. 8d, e).

To test whether a role of ELOF1 in DNA replication is linked to transcription, we monitored R-loop accumulation after UV irradiation by DRIP-qPCR in the actively transcribed *RPL13A* and *EGR1* genes. Both genes showed increased R-loops in *ELOF1*-KO cells, while *CSB*-KO cells showed levels comparable to WT cells (Fig. 8f). The R-loop signal in *ELOF1*-KO cells was completely lost after RNaseH treatment. These findings reveal a

second TCR-independent role of ELOF1 in protecting cells against DNA damage during replication, which becomes particularly important when canonical TCR fails (Fig. 8g).

Discussion

In this study, we identify and characterize human ELOF1 as an RNAPII-associated transcription elongation factor that is also a core component of the transcription-coupled DNA repair machinery with an additional role in preventing DNA damage during DNA replication.

The UV-induced ubiquitylation of RNAPII (RPB1-K1268) by the $CRL4^{CSA}$ complex is important for TCR[2, 3]. Yeast ELF1 binds to a central cleft in front of RNAPII close to this ubiquitylation site (Fig. 6d)[17]. Our findings suggest that loss of ELOF1 does not affect recruitment of either CSB or $CRL4^{CSA}$ to lesion-stalled RNAPII (Fig. 5a, b), but rather that ELOF1 interacts with $CRL4^{CSA}$ through direct protein-protein contacts and directs the catalytic site of the $CRL4^{CSA}$ complex into close proximity of the K1268 site (Fig. 6e–i). By serving as a specificity factor for RNAPII ubiquitylation, ELOF1 subsequently promotes the recruitment and ubiquitylation of UVSSA, which, in turn, transfers the TFIIH complex from UVSSA onto DNA damage-stalled RNAPIIo to initiate repair (Fig. 6j).

The importance of ELOF1 in TCR is evident from *(1)* the inability of *ELOF1*-KO cells to remove UV-induced lesions from transcribed strands (Fig. 4, Extended Data Fig. 7 – Fig. 9), *(2)* the persistent stalling of RNAPII throughout the genome upon UV irradiation (Fig. 3g–h, Extended Data Fig. 6b), and *(3)* the inability of *ELOF1*-KO cells to recover RNA synthesis following UV irradiation, similar to TCR-deficient *CSB*-KO cells (Fig. 3a–d, Extended Data Fig. 4a, d)[20]. We therefore propose that ELOF1 is a core TCR factor, and hypothesize that hypomorphic *ELOF1* mutations may cause a Cockayne syndrome-like phenotype.

Our genetic-interaction network indicates that ELOF1 is not only involved in TCR, but also in a CSB-independent repair pathway (Fig. 7a–b). Consistently, we find that *ELOF1*-deficient cells but not *CSB*-deficient cells are sensitive to replication stress triggered by DNA polymerase α inhibition (Fig. 8a), and accumulate R-loops in transcribed regions (Fig. 8f). Although the precise nature of this pathway remains to be elucidated, these findings suggest that ELOF1 helps to prevent DNA damage during replication and might prevent collisions between transcription and replication machineries (Fig. 8g). Importantly, our data reveal that cells become more dependent on this secondary ELOF1 pathway when canonical TCR fails, suggesting that the additive impact of the second ELOF1 pathway may be partially masked in TCR-proficient cells. (Fig. 8c, e). Elucidating the components and the mechanism involved in the compensatory ELOF1-dependent pathway are important goals for future research.

Together, our results identify ELOF1 as a transcription-coupled DNA repair factor that promotes CSB-dependent TCR by directing RNAPII ubiquitylation and, at the same time, acts in a compensatory second pathway as a more general sensor of transcription stress during DNA replication (Fig. 8g).

Methods

Cell lines.

All cell lines (listed in Supplementary Table 1) were cultured at 37°C in an atmosphere of 5% CO2 in DMEM (Thermo Fisher Scientific) supplemented with penicillin/streptomycin (Sigma) and 10% fetal bovine serum (FBS; Bodinco BV or Thermo Fischer Scientific (Gibco)). U2OS Flp-In/T-REx cells (from here on out called U2OS (FRT)), which were generated using the Flp-InTM/T-RExTM system (Thermo Fisher Scientific), were a gift from Daniel Durocher[29].

Construction of RPE1-TetOn-Cas9-PuroS-*TP53*-KO.

Human RPE1-hTERT cells were acquired from ATCC (CRL-4000™). Lentiviral particles of pLVX-Tre3G-Cas9 plasmid (Clontech, Cas9 from vector Lenti-Cas9–2A-Blast (Addgene) were produced in HEK293T cells using the Lenti-X HT packaging system (Clontech). Transduced RPE1-hTERT cells were selected using 10 μg/mL puromycin and 400 μg/mL G418 to generate RPE1-TetOn-Cas9[30]. To create knockouts, Cas9 expression was induced by 100 ng/mL doxycycline (Sigma-Aldrich) followed by transfection by RNAiMAX (Invitrogen) with 10 nM tracrRNA (Integrated DNA Technologies, IA) and 10 nM synthetic crRNA targeting *TP53* and the *Streptomyces* puromycin-N-acetyltransferase *PAC1* gene (sgRNAs listed in Supplementary Table 2). The population was selected for *TP53*-KO cells by selection on 10 μM Nutlin-3 (Selleck Chemicals, TX) for one week. Single cell clones were selected based on sensitivity to 1 μg/mL puromycin, and knockout of *TP53* was confirmed by Sanger sequencing and western blot. Cas9 activity of the RPE1-TetOn-Cas9-PuroS-*TP53*-KO cell line (RPE1-iCas9) was assayed by transducing cells with lentivirus produced using vector pXPR-011 (Addgene)[31] containing eGFP, sgGFP and puromycin N-acetyltransferase (Plasmids are listed in Supplementary Table 3). After transduction, transduced cells were selected by puromycin (3 μg/mL, Sigma) and eGFP-expression was assayed every 4 h for up to a week in the absence and presence of doxycycline (100 ng/mL, Sigma) in an IncuCyte Zoom automated live-cell imaging platform (Sartorius, Göttingen, Germany).

CRISPR/Cas9 screens.

For every screen, three populations of RPE1-iCas9 were transduced at an M.O.I. of ~0.2 with a 1:1000 dilution of TKOv3-in-pLCKO lentiviral library in medium containing 8 μg/mL hexadimethrine bromide (Sigma-Aldrich). The library was a gift from Katherine Chan, Amy Tong and Jason Moffat (Donnelly Centre, University of Toronto, Toronto, ON). 24 h after transduction, puromycin (Sigma-Aldrich) was added to 5 μg/mL to select for transduced cells. After all cells in non-transduced control populations had died and dishes with transduced populations had reached 90% confluence, a t=0 sample was taken for each of the three populations. From the remaining cells of each population 30×10^6 (corresponding to a library representation of >400) were grown as a control population, and additionally 30×10^6 were grown in the presence of Illudin S for the drug screens, at a concentration of 25 nM for the screen in RPE1-iCas9, 2 nM for the screen in RPE1-iCas9 *CSB*-KO, and 5 nM for the screen in RPE1-iCas9 *ELOF1*-KO. Doxycycline was added to the medium of all replicates from t=0 onwards to induce expression of Cas9, at a

concentration of 200 ng/mL. After three doublings 30×10^6 cells of each population were passed. After 12 doublings, all populations were harvested.

Sequencing and analysis of CRISPR screens.

Genomic DNA was isolated from each population using the Blood and Cell Culture DNA Maxi Kit (Qiagen). 3 μg gDNA of each population was amplified using the KAPA HiFi ReadyMix PCR Kit (Roche) with the TKO outer Fw and Rv primers (Primers are listed in Supplementary Table 4) followed by a second PCR reaction using reverse primers with different Illumina i7 index sequences for each sample to identify the sample after pooled sequencing as described by Hart *et al.*[32]. The second PCR products of each pool were purified by QIAquick PCR Purification Kit (Qiagen). RNAseq samples were added to 10% to create balanced reads for the first 21 nucleotides sequenced. Up to twelve samples were sequenced in a single HiSeq4000 lane in an SR50 run using standard reagents and conditions and reads were mapped to the TKOv3 library sequences, not allowing any mismatches. The screen data was analyzed using DrugZ version 1.1.0.2[33], Log_2_Fold-Change was calculated by first normalizing all samples to an equal number of total reads (pseudocount +1), then calculating the median fold change per guide for the three replicate screen samples, and finally calculating the Log_2 of the median fold change of all guides targeting the gene. For network analysis, the physical interactions of differentially depleted genes were extracted from the gene interaction database, GeneMANIA. Functional pathway analysis was performed using Reactome (v72). The interaction network was built and visualized in Cytoscape v.3.7. Node colors were adjusted based on FDR. Node border colors were customized based on enriched pathways.

Generation of knock-out cells.

Cells were either transfected with Cas9–2A-GFP (pX458; Addgene #48138) containing a guide RNA from the TKOv3 library (Addgene #125517) or co-transfected with Cas9–2A-GFP (pX458; Addgene #48138) together with pLV-U6g-PPB encoding a guide RNA from the LUMC/Sigma-Aldrich sgRNA library using lipofectamine 2000 (Invitrogen) (sgRNAs are listed in Supplementary Table 2 and plasmids in Supplementary Table 3). Cells were FACS sorted on BFP/GFP and plated at low density after which individual clones were isolated. Alternatively, cells were selected with puromycin (1 μg/mL) for 3 days and seeded at low density after which individual clones were isolated. Isolated knockout clones were verified by western blot analysis and/or Sanger sequencing previously described (Primers are listed in as Supplementary Table 4) [6].

Plasmids.

The Neomycin resistance gene in pcDNA5/FRT/TO-Neo (Addgene #41000) was replaced with a puromycin resistance gene. Fragments spanning GFP-N1 (Clontech) including the multiple cloning site were inserted into pcDNA5/FRT/TO-puro. ELOF1^{WT} was amplified by PCR and inserted into pcDNA5/FRT/TO-puro-GFP-N1. ELOF1 mutants were generated by site-directed mutagenesis PCR (Primers are listed in Supplementary Table 4). A region spanning the PGK promoter was amplified by PCR and used to replace the CMV promoter in pEGFP-C1-IRES-PURO[34]. The ELOF1^{WT}-GFP and ELOF1$^{S72K/D73K}$-GFP

genes were inserted into pPGK-EGFP-C1-IRES-PURO to replace EGFP-C1. All sequences were verified by Sanger sequencing.

Generation of stable cell lines.

U2OS (FRT) ELOF1-KO clone 3–12 was selected and subsequently used to stably express inducible GFP-tagged proteins by co-transfecting pCDNA5/FRT/TO-Puro plasmid encoding GFP-tagged fusion proteins (5 μg), together with pOG44 plasmid encoding the Flp recombinase (0.5 μg) using lipofectamine 2000 (Invitrogen) (Plasmids are listed in Supplementary Table 3). Cells expressing inducible GFP-tagged proteins were selected by incubation with 1 μg/mL puromycin and 4 μg/mL blasticidin S. Expression of these GFP-tagged proteins was induced by the addition of 2 μg/mL doxycycline for 24 h.

Compounds.

In the indicated experiments 10 μM MG132 (Cayman Chemical, #13697), 20 mM N-ethylmaleimide (NEM) (Pierce, #23030), and 10 μM FT671 (Medchem express, #HY-107985) was used.

Immunoprecipitation for Co-IP.

Cells were mock treated or irradiated with UV-C light (20 J/m^2) and harvested 1 h after UV. Chromatin-enriched fractions were prepared by incubating the cells for 20 min on ice in IP buffer (IP-130 for endogenous RNAPII IP and IP-150 for GFP-IP), followed by centrifugation, and removal of the supernatant. For endogenous RNA pol II IPs, the chromatin-enriched cell pellets were lysed in IP-130 buffer (30 mM Tris [pH 7.5], 130 mM NaCl, 2 mM $MgCl_2$, 0.5% Triton X-100, protease inhibitor cocktail (Roche), 250 U/mL Benzonase® Nuclease (Novagen), and 2 μg RNAPII-S2 (ab5095, Abcam) for 2–3 h at 4 °C. For GFP IPs, the chromatin-enriched cell pellets were lysed in IP-150 buffer (50 mM Tris [pH 7.5], 150 mM NaCl, 0.5% NP-40, 2 mM $MgCl_2$, protease inhibitor cocktail (Roche), and 500 U/mL Benzonase® Nuclease (Novagen)) for 1 h at 4 °C. Protein complexes were pulled down by 1.5 h incubation with Protein A agarose beads (Millipore) or GFP-Trap® A beads (Chromotek). For subsequent analysis by western blotting, the beads were washed 6 times with IP-130 buffer for endogenous RNAPII IPs and EBC-2 buffer (50 mM Tris [pH 7.5], 150 mM NaCl, 1 mM EDTA, 0.5% NP-40, and protease inhibitor cocktail (Roche)) for GFP-IPs. The samples were prepared by boiling in Laemmli-SDS sample buffer.

Mass spectrometry.

After pulldown, the beads were washed 4 times with EBC-2 buffer without NP-40 and 2 times with 50 mM ammonium bicarbonate followed by overnight digestion using 2.5 μg trypsin at 37°C under constant shaking. Peptides were desalted using a Sep-Pak tC18 cartridge by washing with 0.1% acetic acid. Finally, peptides were eluted with 0.1% formic acid/60% acetonitrile and lyophilized as described[35]. Samples were analyzed on a Orbitrap Exploris 480 mass spectrometer (Thermo Fisher, Germany) coupled to an Ultimate 3000 UHPLC system (Thermo Fisher, Germany). Digested peptides were separated using a 50 cm long fused silica emitter (FS360–75-15-N-5-C50, New Objective, Massachusetts, US) in-house packed with 1.9 μM C18-AQ beads (Reprospher-DE, Pur, Dr. Maisch, Ammerburch-

Entringen, Germany) and heated to 50°C in a Column Oven for ESI/Nano Spray (Sonation, Germany). Peptides were separated by liquid chromatography using a gradient from 2% to 32% acetonitrile with 0.1% formic acid for 40 min followed by column re-conditioning for 20 min. Scan range for MS was from 300 to 1,600 m/z. Data-Dependent Acquisition (DDA) mode was used with a Top Speed method, resolution of 60,000 and tandem mass spectra (MS/MS) cycles of 3 sec. Internal calibration was enabled based on lock mass corresponding to polysiloxane (445.12003 m/z). Higher-Collisional Dissociation (HCD) normalized collision energy was set to 28% and only precursor ions with charges between 2 and 6 were considered to trigger MS/MS events. Precursors selected for MS/MS analysis were dynamically excluded for 30 sec.

Mass spectrometry data analysis.

Raw mass spectrometry data were further analyzed in MaxQuant v 1.6.14 as previously described[36] using standard settings with the following modifications. Maximum missed cleavages by trypsin was set to 4. Searches were performed against an *in silico* digested database from the human proteome including isoforms and canonical proteins (Uniprot, 8th June 2020) plus GFP protein. Oxidation (M), Acetyl (Protein N-term), GlyGly (K) and Phospho (STY) were set as variable modifications with a maximum of 3. Carbamidomethyl (C) was disabled as fixed modification. Label-free quantification was activated, including calculation of iBAQ and not enabling Fast LFQ. The match between runs feature was activated with default parameters. MaxQuant output data were further processed in the Perseus Computational Platform v1.6.14 as previously described[37]. LFQ intensity values were Log_2 transformed and potential contaminants and proteins identified by site only or reverse peptide were removed. Samples were grouped in experimental categories and proteins not identified in 4 out of 4 replicates in at least one group were also removed. Missing values were imputed using normally distributed values with a 1.8 downshift (Log_2) and a randomized 0.3 width (Log_2) considering whole matrix values. Two-sided t-tests were performed to compare groups. Data were exported from Perseus and further processed in Microsoft Excel 2016 for comprehensive visualization.

Western blotting.

Proteins were separated on 4–12% Criterion XT Bis-Tris gels (Bio-Rad, #3450124) in NuPAGE MOPS running buffer (NP0001–02 Thermo Fisher Scientific), and blotted onto PVDF membranes (IPFL00010, EMD Millipore). Membranes were blocked with blocking buffer (Rockland, MB-070–003) for 1 h at RT. Membranes were then probed with antibodies as indicated (Antibodies are listed in Supplementary Table 5). Proteins stained with HRP-conjugated secondary antibodies were detected using Western Lightning Plus-ECL (PerkinElmer, NEL103001EA).

Clonogenic survival assays.

Cells were seeded at low density and mock-treated or exposed to a dilution series of Illudin S (Santa Cruz; sc-391575) or CD437 (Stem cell; #72722) for 72 h. On day 10, the cells were washed with 0.9% NaCl and stained with methylene blue. Colonies of more than 20 cells were scored.

Dose-response curves.

Cells were seeded in 96-well plates (1000 cells per well for RPE1-iCas9) in 150 µl of medium. 24 h after seeding, drugs were added manually dissolved in 50 µl of medium (Cisplatin) or printed directly into the plate by Tecan D300e digital dispenser (Illudin S, Irofulven). Phenylarsine oxide (PAO; Sigma-Aldrich) was used at 10 µM as a control for zero cell viability. After 72 h of drug exposure, cell viability was assayed by adding 30 µL of CellTiter-Blue (Promega) in medium to each well. Plates were incubated for 4 h at 37°C, and plate fluorescence (560Ex/590Em) was measured on a BioTek plate reader. Relative viability was calculated for each well from the fluorescence measurements as: (value-(value(PAO))/((value(untreated)-(value(PAO)). A four-parameter dose–response curve was fit to the viability values using Graphpad Prism 8 (v8.4.2).

RNA recovery synthesis (RRS).

For the RRS following DRB treatment, cells were treated with 100 µM DRB (Sigma, D1916) for 2 h and either during the DRB treatment or following DRB washout the cells were pulse-labelled with 400 µM 5-ethynyl-uridine (EU; Jena Bioscience) for 1 h. For the RRS following UV irradiation, cells were irradiated with UV-C light (9 J/m^2 or 12 J/m^2), allowed to recover for the indicated time periods, and pulse-labelled with 400 µM 5-ethynyl-uridine (EU; Jena Bioscience) for 1 h followed by a 15 min medium-chase with DMEM without supplements. Cells were fixed with 3.7% formaldehyde in PBS for 15 min, permeabilized with 0.5% Triton X-100 in PBS for 10 min at RT and blocked in 1.5% bovine serum albumin (BSA, Thermo Fisher) in PBS. Nascent RNA was visualized by click-it chemistry, labelling the cells for 1 h with a mix of 60 µM Atto azide-Alexa594 (Atto Tec), 4 mM copper sulfate (Sigma), 10 mM ascorbic acid (Sigma) and 0.1 µg/mL DAPI in a 50 mM Tris-buffer. Cells were washed extensively with PBS and mounted in Polymount (Brunschwig).

Unscheduled DNA synthesis (UDS).

Cells were locally irradiated with UV-C light (5 µm filter; 30 J/m^2) and subsequently pulse-labelled with 20 µM 5-Ethynyl-2´-deoxyuridine (EdU; VWR) and 1 µM FuDR (Sigma Aldrich) for either 1 h or 4 h followed by a 30 min medium-chase with DMEM without supplements. Cells were fixed with 3.7% formaldehyde in PBS for 15 min, permeabilized with 0.5% Triton X-100 in PBS for 20 min at RT and blocked in 3% BSA (Thermo Fisher) in PBS. The incorporated EdU was coupled to Atto azide Alexa Fluor 647 using Click-iT chemistry according to the manufacturer's instructions (Invitrogen). After coupling, the cells were post-fixed with 2% formaldehyde for 10 min and subsequently blocked with 100 mM Glycine. DNA was denatured with 0.5% NaOH for 5 min, followed by blocking with 10% BSA (Thermo Fisher) for 15 min. Next, the cells were incubated with an antibody against CPDs for 2 h, followed by secondary antibodies 1 h, and DAPI for 5 min (Antibodies are listed in Supplementary Table 5). Cells were mounted in Polymount (Brunschwig).

Proximity ligation assay (PLA).

Cells were locally irradiated with UV-V light (5 µm filter; 100 J/m^2; 1 h recovery), washed with CSK buffer (100 mM NaCl, 300 mM Sucrose, 3 mM $MgCl_2$ 10 mM PIPES [pH

6.8]) followed by a pre-extraction with CSK buffer containing 0.25% Triton X-100 for 5 min. Cells were fixed with 3.7% formaldehyde in PBS for 15 min and permeabilized with 0.5% Triton X-100 for 10 min at RT. The PLA was performed according to manufacturer's specifications using Duolink in Situ (Sigma) with the indicated antibodies (Antibodies are listed in Supplementary Table 5) and signal intensities were quantified using Image J.

DNA fiber assay.

Cells were treated with 5 nM or 25 nM Illudin S for 15 min begore labelling. Cells were sequentially incubated for 20 min with 25 μM chlorodeoxyuridine (CldU) followed by 20 min with 250 μM iododeoxyuridine (IdU). Next, cells were lysed in spreading buffer (200 mM Tris-HCl [pH 7.4], 50 mM EDTA and 0.5% SDS) on Superfrost microscope slides. Fibers were spread by slightly tilting slides followed by fixation in methanol:acetic acid (3:1). After denaturation of DNA by 2.5 M HCl for 75 min and blocking in PBS with 1% BSA and 0.1% Tween-20, slides were incubated with antibodies against BrdU (Antibodies are listed in Supplementary Table 5). After fixing using 4% paraformaldehyde, slides were incubated with secondary antibodies and mounted with Prolong Gold Antifade Mountant (Invitrogen). Fiber lengths were measured using ImageJ and values were converted into kilobases using the conversion factor 1 μm = 2.59 kb[38].

γH2AX Foci.

Cells were either mock treated or treated with 5 nM Illudin S for 15 h. Cells were incubated with 10 μM EdU (Jena Bioscience) for 20 min, fixed with 2% paraformaldehyde and permeabilized with 70% ice cold EtOH. Cells were blocked with 3% BSA containing 0.3% Triton X-100. Next, the cells were incubated with an antibody against γH2AX for 1.5 h, followed by secondary antibodies 1 h (Antibodies are listed in Supplementary Table 5). Next, EdU positive cells were visualized by incubating for 30 min with EdU Click-iT reaction cocktail (50 mM Tris-HCl [pH 7.6], 150 mM NaCl, 1 μM Picolyl Azide 5/6-FAM (#CLK-1180, Jena Bioscience), 4 mM $CuSO_4$ and 2 mg/mL Sodium-L-Ascorbate). Cells were mounted using ProLong™ Gold Antifade Mountant with DAPI (Invitrogen) and γH2AX foci per cell were quantified using Image J.

Microscopic analysis of fixed cells.

Images of fixed samples were acquired on a Zeiss AxioImager M2 or D2 widefield fluorescence microscope equipped with 63x PLAN APO (1.4 NA) oil-immersion objectives (Zeiss) and an HXP 120 metal-halide lamp used for excitation. Images were recorded using ZEN 2012 (Blue edition, Version 1.1.0.0) software and analyzed in Image J (1.48v).

DRIP-qPCR.

Cells were irradiated with UV-C light (9 J/m^2) and incubated in conditioned media for 3 h. DNA was extracted by incubating the cells overnight at 37°C in TE buffer containing 0.625% SDS and 0.1 mg/mL proteinase K (Thermo Fisher Scientific; EO0491) followed by phenol-chloroform extraction and ethanol precipitation. The DNA was enzymatically digested with the following cocktail of enzymes and compounds at 37°C overnight: *HindIII*, *EcoRI*, *XbaI*, *SspI*, and *BsrGI*, BSA, NEB buffer 2.1 (New England Biolabs) and

spermidine (0.1 M; Sigma-Aldrich, #05292–1ML-F) as described previously[39]. The DNA was purified using phenol-chloroform and precipitated with ethanol. For the control of the immunoprecipitated material, 10 μg of each sample was separated in an independent tube and treated with RNase H (NEB) overnight at 37°C. The DRIP procedure was performed as previously described with some modifications[40]. 1/10th of the purified DNA was used as input. For the immunoprecipitation 8 μg of DNA was incubated with 3 μg of S9.6 antibody (Kerafast, #ENH001) in 1x binding buffer (10 mM sodium phosphate [pH 7.0], 140 mM NaCl, and 0.05% Triton X-100) overnight at 4°C, followed by pulldown of DNA-RNA hybrids with Protein A Dynabeads (Thermo Fisher; 10002D) for 2 h at 4°C. The samples were washed extensively, after which the samples were eluted with elution buffer (50 mM Tris [pH 8.0], 10 mM EDTA [pH 8.0], and 0.5% SDS). Both the input and the IP samples were treated with proteinase K and incubated for 45 min at 55°C. The samples were purified using phenol-chloroform and precipitated with ethanol. Quantitative qPCR was performed using SYBR-Green Mastermix (Bio-Rad) with the indicated primers (Primers are listed in Supplementary Table 4). The immunoprecipitation efficiency as percentage of input at each locus was calculated as previously described[40].

Protein expression and purification.

Cloning, protein expression and purification of *Xenopus laevis* CSB and CSA-DDB1-CUL4A-RBX1 (CRL4^{CSA}) was done as described previously[6]. Coding sequences of *Xenopus laevis* ELOF1 and RAD23B for bacterial expression were ordered as codon-optimized gene blocks from Integrated DNA Technologies and cloned into pOPINK vectors (Addgene #41143) containing the indicated affinity tags. Proteins were expressed in *E. coli* OverExpress™ C41(DE3) Chemically Competent Cells (Sigma) grown at 37°C in LB medium supplemented with appropriate antibiotics. At an optical density (OD_{600}) of 0.5–0.6, protein expression was induced with 0.5 mM IPTG for 18–20 h at 18°C. For the expression of ELOF1, $ZnSO_4$ was added to a final concentration of 20 μM at the time of induction. Protein purifications were performed at 4°C. Cells were lysed by sonication in Lysis Buffer (25 mM HEPES [pH 8.0], 300 mM NaCl, 10% glycerol, 2 mM DTT, 20 μM $ZnSO_4$ (ELOF1 only)) containing one EDTA-free complete protease inhibitor tablet (Roche) and 1 mg/mL lysozyme (Sigma), and cleared by centrifugation for 1 h at 35,000 rpm. The cleared lysate was incubated with 1–2 mL equilibrated Glutathione Sepharose 4B resin (GE Healthcare) for 1 h at 4°C on a rotating wheel. The resin was washed extensively with GST High-Salt Wash Buffer (25 mM HEPES [pH 8.0], 500 mM NaCl, 5% glycerol, 2 mM DTT, 10 μM $ZnSO_4$ (ELOF1 only)). For the purification of GST and GST-tagged ELOF1, proteins were eluted with GST High-Salt Wash Buffer supplemented with 20 mM reduced L-Glutathione (Sigma). For the purification of untagged ELOF1 and RAD23B, the resin was washed with Gel-Filtration Buffer (25 mM HEPES [pH 8.0], 200 mM NaCl, 5% glycerol, 2 mM DTT, 10 μM $ZnSO_4$ (ELOF1 only)) and the GST tag was cleaved overnight on the resin with GST-tagged PreScission protease. Both GST-tagged and untagged proteins were further purified by gel filtration (Superdex 75) in Gel-Filtration Buffer. Peak fractions were concentrated with 5 mL 3 MWCO spin concentrators (Millipore), frozen in liquid nitrogen, and stored at −80°C. Endogenous *Xenopus laevis* RNA Polymerase II (RNAPII) complex was purified from egg extract (HSS; high-speed supernatant) using the monoclonal 8WG16 antibody immobilized on Protein G Sepharose 4 Fast Flow resin (GE Healthcare).

After overnight incubation at 4°C, the resin was washed with Wash Buffer (25 mM HEPES [pH 7.5], 12.5 mM $MgCl_2$, 0.1 mM EDTA, 10% glycerol, 0.1 mM DTT, 0.01% NP-40) containing 400 mM KCl followed by Wash Buffer containing 200 mM KCl. The protein was eluted with Wash Buffer containing 200 mM KCl and 1.5 mg/mL RNAPII CTD peptide (H_2N-PTSPSYSPTSPSYSPTSPSYS-OH; New England Peptide) and further purified by gel filtration (Superose 6 Increase) in Superose 6 Gel-Filtration Buffer (25 mM HEPES [pH 7.5], 150 mM NaCl, 10% glycerol, 2 mM DTT). Peak fractions were pooled, concentrated with 5 mL 50 MWCO spin concentrators (Millipore), frozen in liquid nitrogen, and stored at −80°C.

Pull-down using immobilized GST and GST-ELOF1.

Purified GST and GST-tagged *Xenopus laevis* ELOF1 were immobilized on equilibrated Glutathione Sepharose 4B resin (GE Healthcare) for 1 h at 4°C. The beads were washed with GST Pull-down Buffer (10 mM HEPES [pH 7.7], 50 mM KCl, 2.5 mM $MgCl_2$, 250 mM sucrose, 0.2 mg/mL BSA, 0.02% Tween) and incubated with the indicated purified proteins for 1 h at 4°C. Subsequently, the resin was washed three times with GST Pull-down Buffer. After the final wash, the beads were resuspended in Laemmli-SDS sample buffer prior to SDS-PAGE analysis and Western blotting.

In vitro ubiquitylation assay.

In vitro ubiquitylation assays were performed as previously described[6]. In short, in vitro neddylated recombinant xlCRL4^{CSA} was incubated at RT with purified ELOF1 or CSB, E1, UBE2D2, ubiquitin, and ATP prior to SDS-PAGE and western blot analysis.

ChIP-sequencing.

Cells were mock treated or irradiated with UV-C light (9 J/m^2) and incubated in conditioned media for different periods of time (1, 4, 8, 16 h). Cells were crosslinked with 0.5 mg/mL disuccinimidyl glutarate (DSG; Thermo Fisher) in PBS for 45 min at RT. Cells were washed with PBS and crosslinked with 1% PFA for 20 min at RT. Fixation was stopped by adding 1.25 M Glycin in PBS to a final concentration of 0.1 M for 3 minutes at RT. Cells were washed with cold PBS and lysed and collected in a buffer containing 0.25% Triton X-100, 10 mM EDTA [pH 8.0], 0.5 mM EGTA [pH 8.0] and 20 mM Hepes [pH 7.6]. Chromatin was pelleted in 5 min at 400 g and incubated in a buffer containing 150 mM NaCl, 1 mM EDTA [pH 8.0], 0.5 mM EGTA [pH 8.0] and 50 mM Hepes [pH 7.6] for 10 minutes at 4°C. Chromatin was again pelleted for 5 min at 400 g and resuspended in ChIP-buffer (0.15% SDS, 1% Triton X-100, 150 mM NaCl, 1 mM EDTA [pH 8.0], 0.5 mM EGTA [pH 8.0] and 20 mM Hepes [pH 7.6]) to a final concentration of 15×10^6 cells/mL. Chromatin was sonicated to approximately 1 nucleosome using a Bioruptor waterbath sonicator (Diagenode). Chromatin of ~5×10^6 cells was incubated with 3 μg antibody (Antibodies are listed in Supplementary Table 5) overnight at 4°C, followed by a 1.5 h protein-chromatin pull-down with a 1:1 mix of protein A and protein G Dynabeads (Thermo Fisher; 10001D and 10003D). ChIP samples were washed extensively, followed by decrosslinking for 4 h at 65°C in the presence of proteinase K. DNA was purified using a Qiagen MinElute kit. For ser2-RNAPII ChIP-seq sample libraries were prepared from 1 ng ChIPed DNA

using NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, E7645S) and index primer sets (NEB, E7335S). Strand-biased library amplification was performed using high-fidelity KAPA HiFi Enzyme (Roche, KK2102). The prepared libraries were sequenced on an Illumina HiSeq 2500 system (Illumina), resulting in 150 bp paired-end reads. For the remaining ChIP-seq samples, libraries were prepared using KAPA HyperPrep kit (Roche) and A-T mediated ligation of full Y-shaped IDT adapters. Samples were sequenced in a 150bp paired-end run on an Illumina HiSeq X system (Macrogen).

Bru-sequencing.

For Bru-sequencing after DRB treatment, cells were pre-treated with DRB (100 μM) for 3.5 h. Then cells were washed with PBS and incubated with 2 mM bromouridine (BrU) at 37°C for 30 or 60 min, as indicated per experiment. For Bru-sequencing after UV irradiation, cells were irradiated with UV-C light (9 J/m^2) and incubated in conditioned media for different periods of time (3, 8, 24 h) before being incubated with 2 mM BrU at 37°C for 30 min. All Bru-sequencing samples were subsequently lysed in TRIzol reagent (Invitrogen) and BrU-containing RNA was isolated as previously described[41]. cDNA libraries were made from the BrU-labeled RNA using the Illumina TruSeq library kit and paired-end 151 bp sequenced using the Illumina NovaSeq platform at the University of Michigan Advanced Genomics Core. Single-end or paired-end sequencing data was used for downstream analyses.

ATAC-seq.

For each experimental condition 50.000 cells were harvested and washed with cold PBS containing protease inhibitors (Roche). The cells were resuspended in 50 μl Transposase mixture, containing 1x Tagment DNA buffer Illumina Kit, (#20034197), 0.5 μl Tagment DNA enzyme (Illumina Kit, #20034197) and 200 ng/μl Digitonin (Promega, #G9441), and incubated at 37°C for 30 min. Fragmented DNA was purified using a Qiagen MinElute kit (#28204). Libraries were prepared by PCR using NEBNext High-Fidelity 2x PCR Master Mix (New England Labs, #M0541) and Nextera XT index Kit v2 (Illumina). Samples were sequenced in a 150bp paired-end run on an Illumina HiSeq X system (Macrogen).

RNA-seq and data analysis.

Cells were irradiated with UV-C light (9 J/m^2) or mock-treated and harvested 24 h later. Total RNA was isolated using the RNeasy mini kit (Qiagen) according to the manufacturer's protocol. Up to 5×10^6 cells per sample were lysed in RLT buffer. For quantification purposes, 24 ng of the eight ArrayControl™ RNA Spikes (Thermo Fisher Scientific) were added to each sample at this stage (3 ng/spike). Samples were purified using the KAPA mRNA HyperPrep kit (Roche) and prepared for sequencing using the TruSeq RNA Library Prep Kit v2 (Illumina) according to the manufacturer's instructions, and sequenced in an SR50 run in a single lane on an Illumina HiSeq4000. For RNA-seq data analysis, obtained sequencing reads were cleaned by 5'-end quality trimming and Illumina-adapter clipping by Trimmomatic (v0.32)[42]. Pre-alignment quality control of the cleaned sequencing reads was done with FastQC (v0.11.8). The alignment to reference genome hg19 of trimmed sequencing reads was done with Hisat2, guided by gene annotation in the refGene UCSC table[43]. The generated SAM files were converted to the binary counterpart BAM, followed by BAM sorting and indexing with SAMtools (v1.10)[44]. Counting reads to the UCSC's

genomic features hg19, performed by Subread and featureCounts (*8*). Differential expression analysis was performed using edgeR (v3.14.0)[45]. Only genes with at least 2 counts per million in at least 33% of samples were included in the analysis. Data were normalized for sample specific effects by the trimmed mean of M-values. This was followed by estimating the dispersion and determining the differentially expressed genes, using general linear model (GLM). FDR-adjusted p-values < 0.05 were considered significant.

ChIP-seq, Bru-seq and ATAC-seq data analyses.

For ser2-RNAPII ChIP-seq, low-quality sequence reads and adapters were filtered out by Trimmomatic (v3.36)[42]. The trimmed reads were aligned to the human reference genome (GRC h37/hg19) with the Burrows-Wheeler Aligner (BWA-v0.7.12-r1039)[46]. Biobambam2 (v2.0.72)[47] was used to remove duplicate reads from the aligned reads. Sequence reads were locally realigned and base-quality scores were recalibrated with the IndelRealigner and BaseRecalibrator programs in Genome Analysis Toolkit (GATK-v3.5)[48].

For pan-RNAPII ChIP-seq, Bru-seq and ATAC-seq, a sequencing quality profile was generated using FastQC (Version 0.11.2). If needed, sequences were trimmed using TrimGalore (Version 0.6.5). Reads were aligned to the Human Genome 38 using bwa-mem tools (BWA (Version 0.7.16a)[46]. Only high-quality reads (> q30) were included in the analyses and duplicates were removed using Samtools (Version 1.6) with fixmate -m and markdup -r settings (Supplementary Table 6). Bam files were converted into stranded TagDirectories and UCSC genome tracks using HOMER tools (Version 4.8.2)[49]. Example genome tracks were generated in IGV (Version 2.4.3). A list of 49,948 genes was obtained from the UCSC genome database (https://genome.ucsc.edu/cgi-bin/hgTables) selecting the knownCanonical table containing the canonical transcription start sites per gene. To prevent contamination of binding profiles, genes should be non-overlapping with at least 2 kb between genes. Subsequent selection of sets of genes is described per analysis.

For ChIP-seq, binding profiles within selected areas of individual genes (e.g. around TSS or TTS), were defined using the AnnotatePeaks.pl tool of HOMER using the default normalization to 10 million reads. Metagene profiles were defined using the makeMetaGeneProfile.pl tool of HOMER, using default settings. Read densities in input samples were subtracted from individual ChIP-seq datasets to background-correct our data in which negative values were converted to 0, to prevent the use of impossible negative read densities in further calculations. Individual datasets were subsequently processed into heatmaps or binding profiles using R (Version 3.5.3) and Rstudio (Version 1.1.423)[50]. ChIP-seq binding and metagene profiles were averaged per set of genes and profiles were normalized to area under the curve (unless described otherwise) to allow proper comparison of the profiles without effects of overall differences in read density.

ATAC-seq and Bru-seq heatmaps were generated as described for ChIP-seq analyses, except for the input subtraction. Bru-seq aggregated profiles of Fig. 3d and Extended Data Fig. 4d, were defined using the AnnotatePeaks.pl tool of HOMER using the default normalization to 10 million reads, as for ChIP-sequencing except for input subtraction. Genes were selected on size and strongest binding of RNAPII at the TSS in undamaged condition, as for ChIP-sequencing. Read counts at 200 bp was put to 0 and averaged binding profiles

were subsequently normalized to nascent transcript levels, as quantified by 5-EU labelling, relative to the control in their specific cell type. For example, WT cells 3 h after UV irradiation showed 35% RNA relative to WT control cells, so we multiplied the expression of the bins by 0.35.

For ChIP-seq, ATAC-seq and Bru-seq analyses, genes were selected based on their lengths for the different analyses (either 3–100 kb, 25–50 kb, 50–100 kb, or greater than 100 kb), and were further selected based on the strongest binding of RNAPII in mock-treated WT cells, as indicated per analysis.

Strand-bias analyses and recovery index.

Pull down of RNAPII from damaged genes results in a relative overrepresentation of damaged transcribed strands in the ChIP-seq samples. The inability of high-fidelity DNA polymerases to amplify damaged DNA strands therefore results in unequal PCR amplification (strand-bias) of the transcribed and non-transcribed strand during ChIP-seq library sample prep. ChIP-seq reads were quantified in transcribed and non-transcribed strands of genes of 3–100 kb (removing small genes that are unlikely to be damaged in our approach and long genes that are unlikely to be repaired within the timeframe of our analyses) with minimal gap of 2 kb between genes using the AnnotatePeaks.pl tool of HOMER with the default normalization to 10 million reads. We focused on the top 3000 genes that showed strongest binding of RNAPII at the TSS in WT undamaged condition to select for actively transcribed genes. A strand specificity index (SSI) was subsequently calculated per gene with the following equation (as described before[2])

$$\text{Strand specificity index} = \frac{\text{reads in “transcribed” strand} - \text{reads in “non-transcribed” strand}}{\text{reads in “transcribed” strand} + \text{reads in “non-transcribed” strand}}$$

Depending on whether genes are present on the + or − strand of DNA, a strand bias results in either positive or negative SSI values, varying around 0 (representing no strand bias). To quantify the amount of damage and correlated level of recovery, we defined a recovery index (RI). For this, we generated frequency distributions of per-gene SSIs. While SSIs frequencies follow a single Gaussian distribution in non-biased ChIP-sequencing samples, the SSI frequencies in biased ChIP-seq samples are expected to follow a mixed Gaussian distribution composed of 3 normal distributions: one distribution representing genes without strand bias (with $\mu_0=0$), and two Gaussian distributions representing strand-biased genes on either the + or − strands of DNA, with μ_1 and μ_2 symmetrically positioned around μ_0 ($-\mu_1 = \mu_2$). The mean distance that the strand-biased Gaussian distributions deviate from 0 (which is represented by $-\mu_1$ or μ_2) was defined as the recovery index (RI). The three Gaussian distributions were fitted using normalmixEM of the mixtools package in R[51].

RNAPII wave-front analyses.

DRB treatment inhibits release of protomer proximal RNAPII into the gene, resulting in a run-off of all elongating RNAPII molecules from transcribed genes. After DRB release, RNAPII molecules restart transcription from the TSS with an average speed of ~2 kb/min in WT cells[52]. To define the speed of RNAPII molecules, an RNAPII wave-front was

defined in BruDRB-seq samples as the average distance from the TSS that newly released RNAPII molecules reach within the indicated time-period (30–60 min after DRB release). Per gene, a positive transcription level was defined as the average BrU readcount/kb within the first 20 kb of a gene. Per gene, the front of the RNAPII pool was subsequently defined as the shortest distance from the TSS where in at least 5 consecutive kb the BrU readcount/kb dropped below 20% of this positive transcription level. Wave-fronts per gene were subsequently recalculated to speed of RNAPII in kb/min and averaged per BruDRB-seq repeat.

Statistics and Reproducibility.

Most experiments were confirmed in multiple cell lines and using complementary approaches. The genome wide CRISPR/Cas9 screens were performed in three independent replicate populations. The ChIP-seq and ATAC-seq experiments were repeated independently twice, except for the pan-RNAPII ChIP-seq U2OS CSB-KO 20 J/m^2 at 8 h, which was performed once. The Bru-seq and BruDRB-seq experiments were repeated independently twice, expect for the BruDRB-seq 30 min, which was repeated three times. RNA-seq experiments were independently repeated three times. All experiments yielding micrographs, pull-down experiments and drug-sensitivity assays were performed independently at least twice and often three times as indicated in the legends. The mass-spectrometry experiments were performed in quadruplicate. Statistical analysis was carried out using two-sided t-tests to compare groups ($P<0.05$).

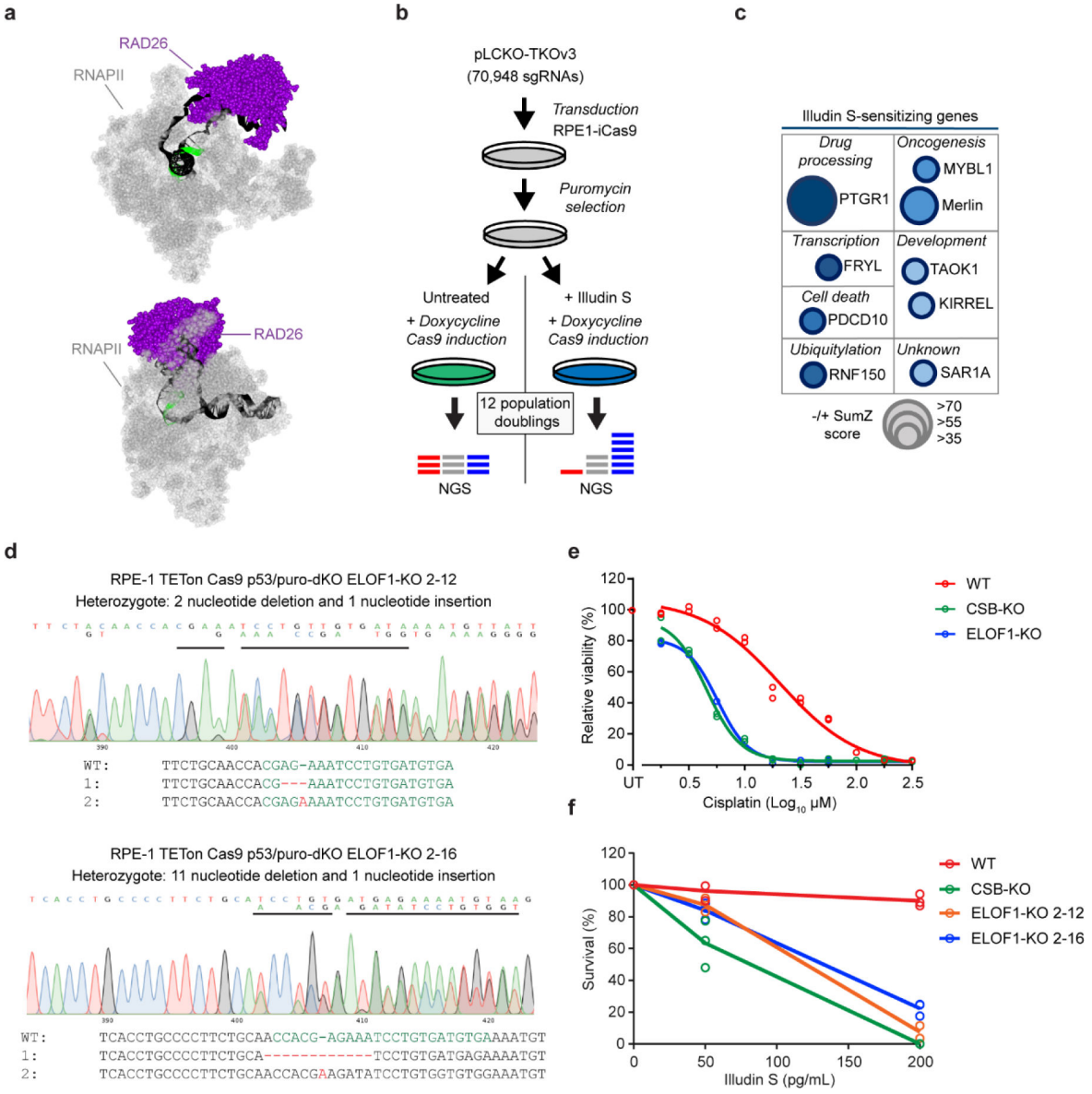
Code availability.

Custom code used for the analysis of NGS data was written in R and is available from GitHub (https://git.lumc.nl/dvandenheuvel/van-der-weegen-etal_elof_ncb2021.git).

Data availability.

Both raw and processed ChIP-seq, Bru-seq, ATAC-seq, and RNA-seq data shown in main Figure 2 – 4 and Extended Data Figure 3 – 9 are deposited in the Gene Expression Omnibus (GEO) under GSE149760. The mass spectrometry proteomics data shown in main figure 1 and Extended Data Figure 2 are deposited to the ProteomeXchange Consortium via the PRIDE partner repository (https://www.ebi.ac.uk/pride/) with the dataset identifier PXD024051. Additionally, publicly available reference datasets of the Hg38 genome and hg19 genome and the known Canonical gene table from the UCSC genome database (https://genome.ucsc.edu/cgi-bin/hgTables; hg38 genome, lifted-over to hg19 when needed) and gene interactions from the GeneMANIA database (https://genemania.org/) have been obtained and used in this manuscript. Published structural information has been obtained for *S. cerevisiae* RAD26 bound to RNAPII (https://www.rcsb.org/ PBD: 5VVS) and *K. pastoris* ELF1 bound to RNAPII (PDB: 5XOG). Source data are provided with this study. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

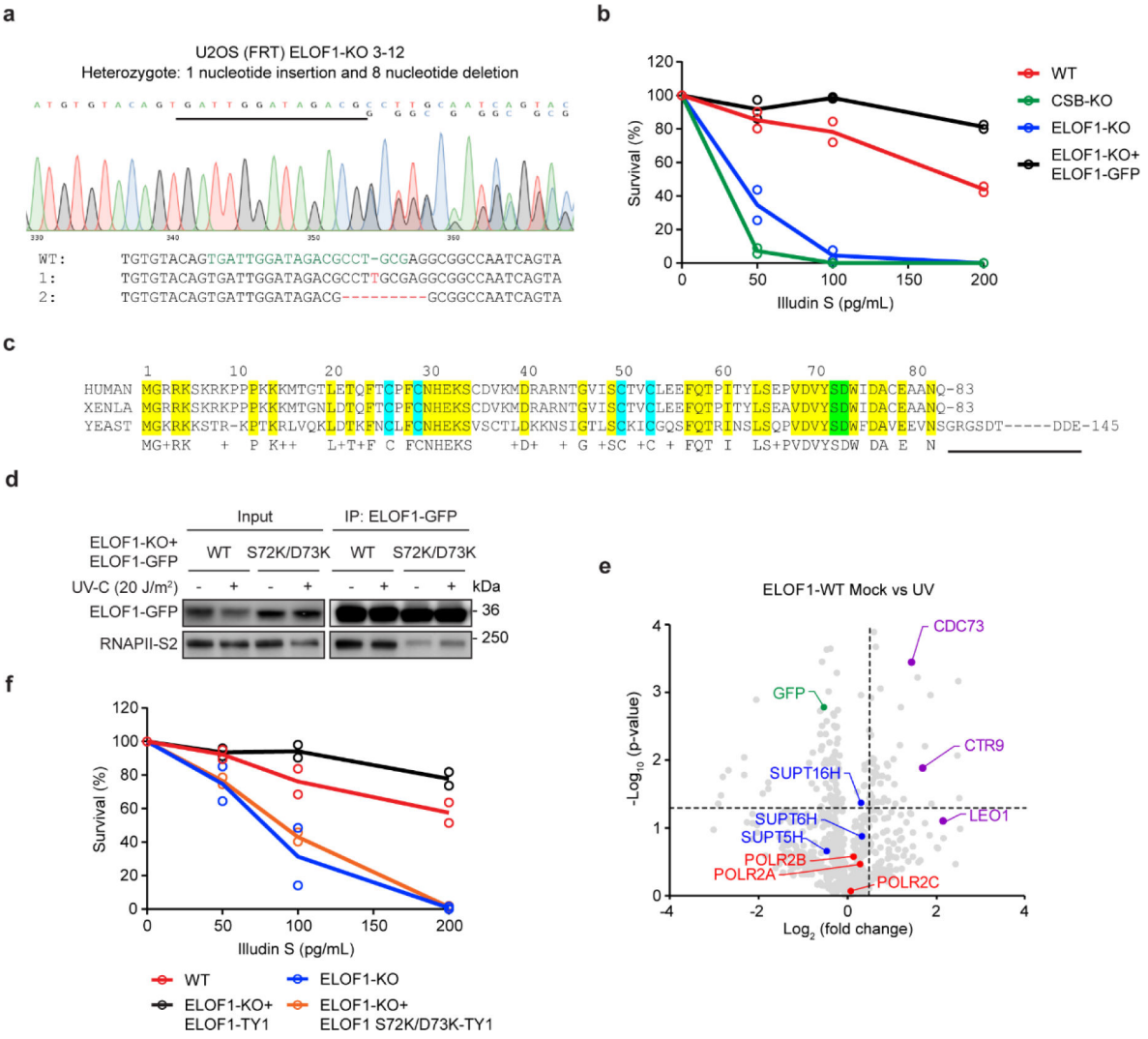
Extended Data



Extended Data Fig. 1. Human ELOF1 protects cells against transcription stress.

(**a**) Front-view and side-view of the yeast orthologue of CSB, *S. cerevisiae* Rad26 (purple), bound to RNAPII (grey) (PDB: 5VVS). (**b**) Schematic representation of the CRISPR/Cas9 screen in RPE1-iCas9 cells in the presence of Illudin S (IC60; 25 nM). (**c**) Network analysis of highly significant hits representing genes that promote Illudin S toxicity. Grey lines reflect known protein-protein interactions (Cytoscape, BioGRID). (**d**) Sanger sequencing of the indicated RPE1-iCas9 single *ELOF1*-KO clones. (**e**) 72 h drug sensitivity assay of indicated RPE1-iCas9 KO clones. Each symbol represents the median of an independent experiment (*n=2*), each containing 6 technical replicates. (**f**) Clonogenic Illudin S survival of the indicated RPE1-iCas9 cells. Each symbol represents the mean of an independent

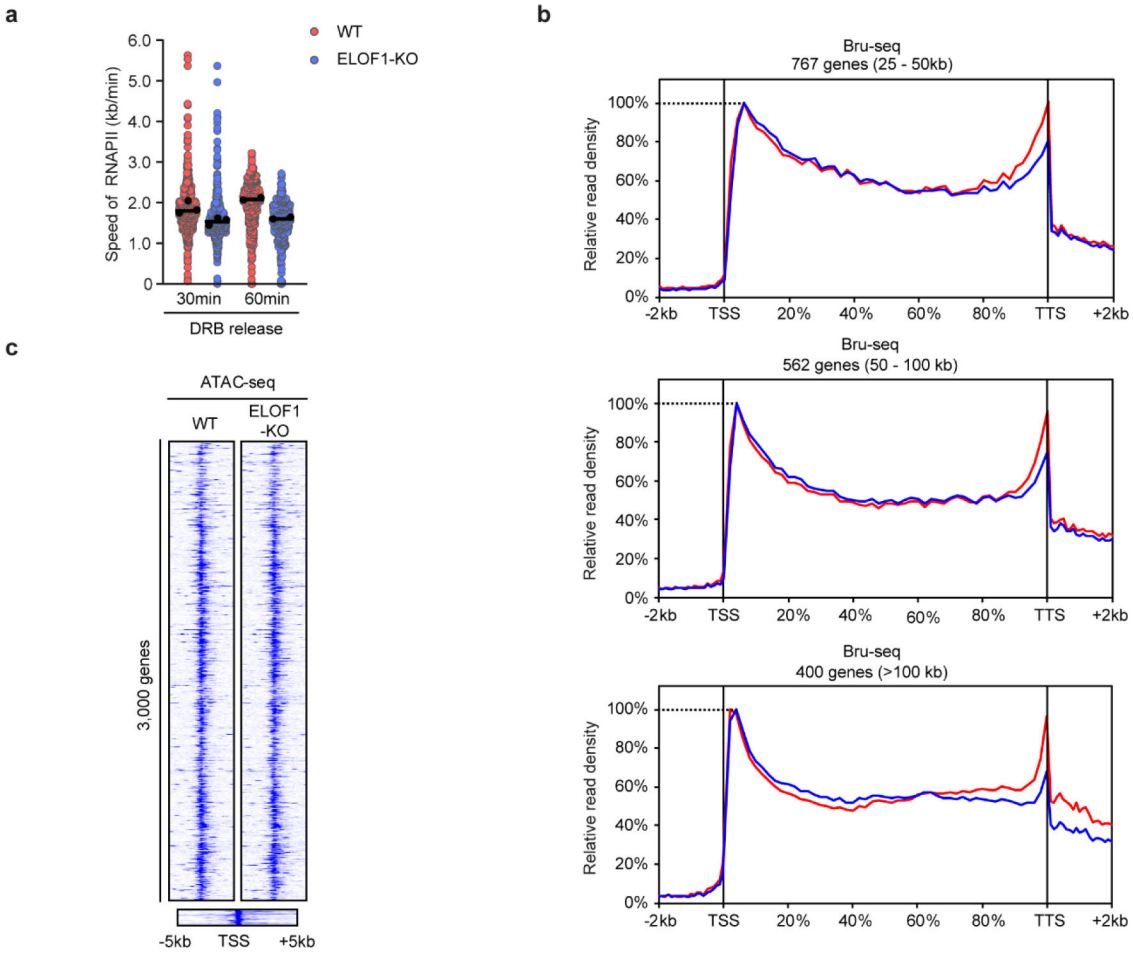
experiment (*n=3* for all except for *ELOF1*-KO 2–12 which is *n=2*), each containing 2 technical replicates.



Extended Data Fig. 2. Human ELOF1 and yeast ELF1 show similar RNAPII-binding modes.

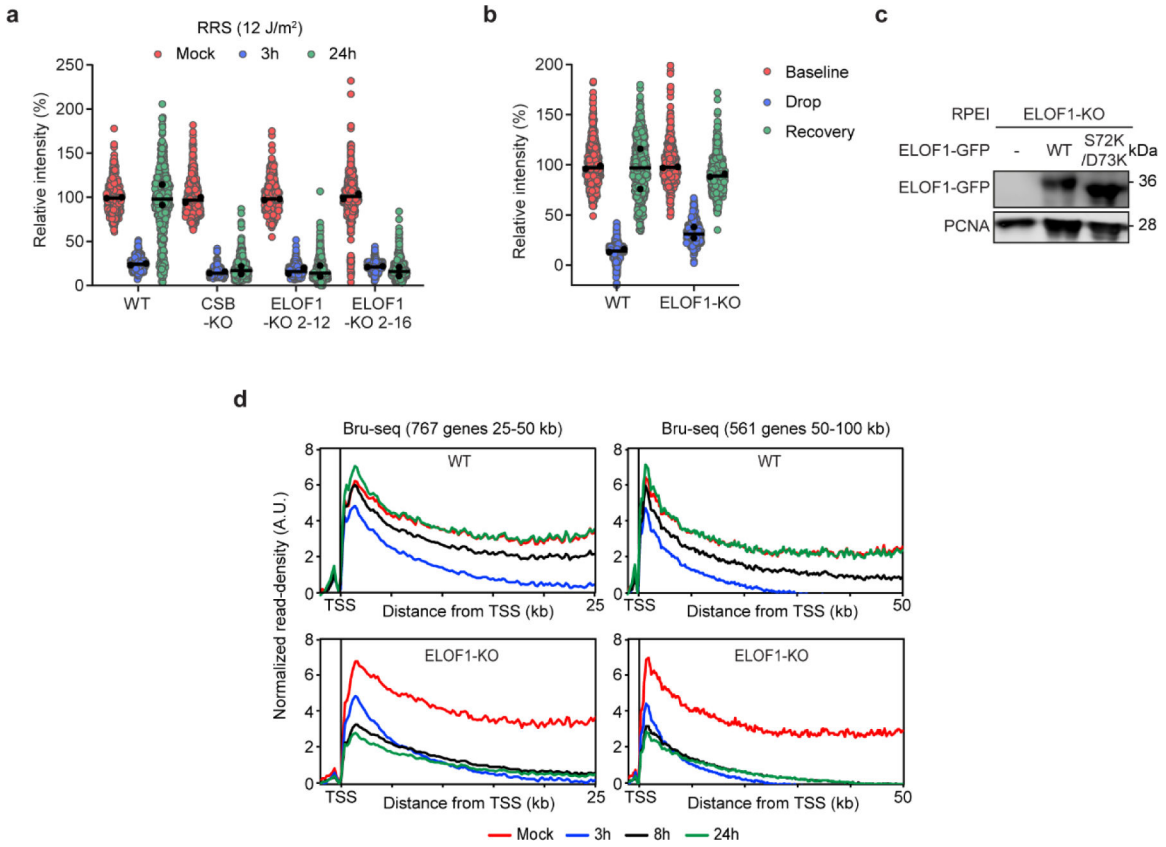
(**a**) Sanger sequencing of the indicated U2OS (FRT) *ELOF1*-KO clone. (**b**) Clonogenic Illudin S survival of U2OS (FRT) WT, *CSB*-KO, *ELOF1*-KO, and *ELOF1*-KO complemented with ELOF1^{WT}-GFP. Each symbol represents the mean of an independent experiment (*n=2*), each containing 2 technical replicates. (**c**) Alignment of human ELOF1, *Xenopus leavis* ELOF1, and *S. cerevisiae* ELF1. Conserved residues are indicated in yellow, zinc-finger cysteines in magenta, residues involved in the RPB1 or RPB2 interaction in green. Note that the C-terminus of *S. cerevisiae* ELF1 (83–145) is absent in human ELOF1 and *Xenopus leavis* ELOF1. (**d**) Co-immunoprecipitation (IP) of ELOF1^{WT}-GFP and ELOF1$^{S72K/D73K}$-GFP on the combined soluble and chromatin fraction (*n=4*). (**e**) Volcano plot depicting the statistical differences between 4 replicates of the MS analysis on ELOF1^{WT}-GFP pull-down in mock treated and UV-irradiated samples. The fold change (Log_2) is plotted on the x-axis and the significance (t-test -Log_{10} p-value) is plotted on the

y-axis. RNAPII subunits are indicated in red, elongation factors are indicated in blue, GFP is indicated in green, and PAF1 subunits are indicated in purple. (**f**) Clonogenic Illudin S survival of U2OS (FRT) WT, *ELOF1*-KO, and TY1-tagged ELOF1 rescue cell lines. Each symbol represents the mean of an independent experiment (*n=2*), each containing 2 technical replicates.



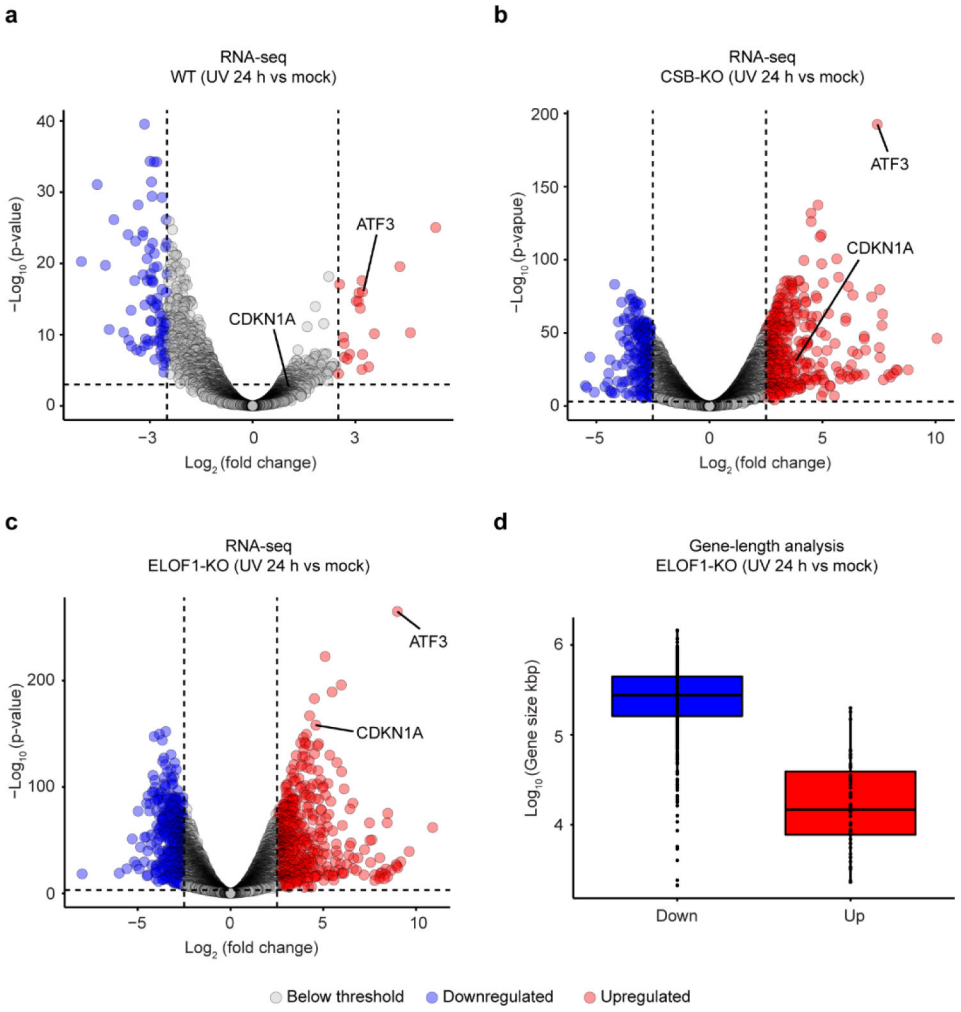
Extended Data Fig. 3. ELOF1 promotes transcription elongation.

(**a**) Calculated speed of RNAPII, quantified by wave-front analyses, in WT (red) or *ELOF1*-KO (blue) cells 30 min (*n=3*) or 60 min (*n=2*) after DRB release. Data represents the median per condition (black line) and median within individual replicates (black circles). (b) Metaplots of BrU signal (nascent transcription) from 2 kb before the TSS to 2 kb after the TTS in 767 genes between 25–50 kb (upper), 562 genes between 50–100 kb (middle), or 400 genes of >100 kb in WT (red) or *ELOF1*-KO (blue) cells. Profiles are normalized to 100% at promoter-proximal BrU peaks instead of area under the curve for better comparison of transcription profiles. Profiles are averages of 2 independent replicates. (c) Heatmaps of ATAC-seq data around the TSS (−5 kb until +5 kb) of 3,000 genes of 3–100 kb in unirradiated RPE1-iCas9 cells (WT or *ELOF1*-KO).



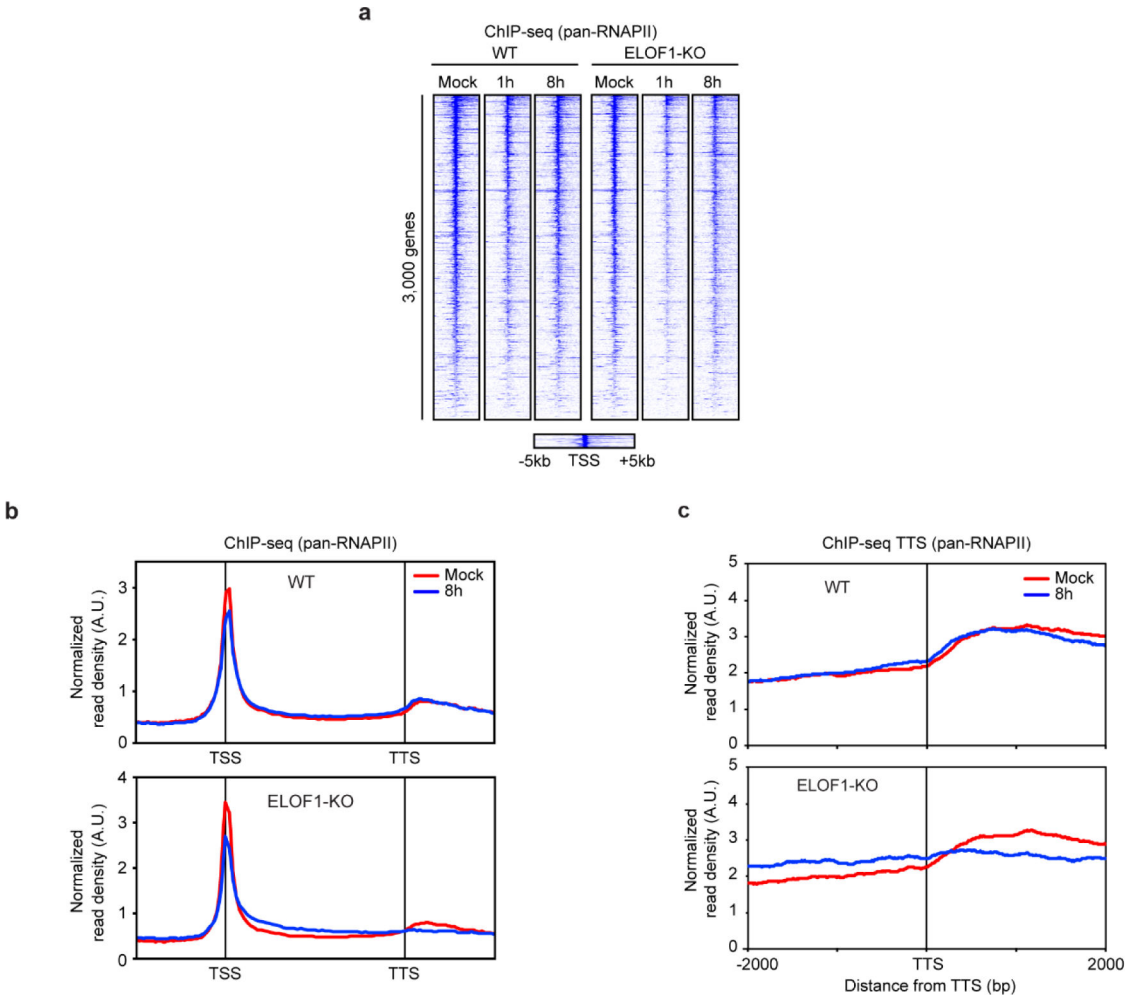
Extended Data Fig. 4. ELOF1 promotes genome-wide transcription recovery.

(**a**) Quantification of 5-EU levels of the indicated RPE1-iCas9 cells normalized to mock-treated levels for each cell line. Cells were either mock-treated or UV-irradiated (3 h or 24 h; 12 J/m^2). Each black circle represents the median of an independent experiment (*n=2*), each containing 2 technical replicates, >80 cells collected per technical replicate. The black line represents the median of all the cells collected. (**b**) Quantification of 5-EU levels normalized to the baseline level before DRB treatment for each condition. Each black circle represents the median of an independent experiment (*n=2*), each containing 2 technical replicates, >80 cells collected per technical replicate. The black line represents the median of all the cells collected. (**c**) Western blot analysis of RPE1-iCas9 *ELOF1*-KO cells complemented with GFP-tagged versions of ELOF1 (*n=3*). (**d**) Metaplots of BrU signal (nascent transcription) in 767 genes between 25–50 kb, or in 561 genes between 50–100 kb in WT (upper) or *ELOF1*-KO (lower) cells after mock treatment (red), or 3 h (blue), 8 h (black), and 24 h (green) after UV irradiation (9 J/m^2). Profiles are averages of 2 independent replicates.



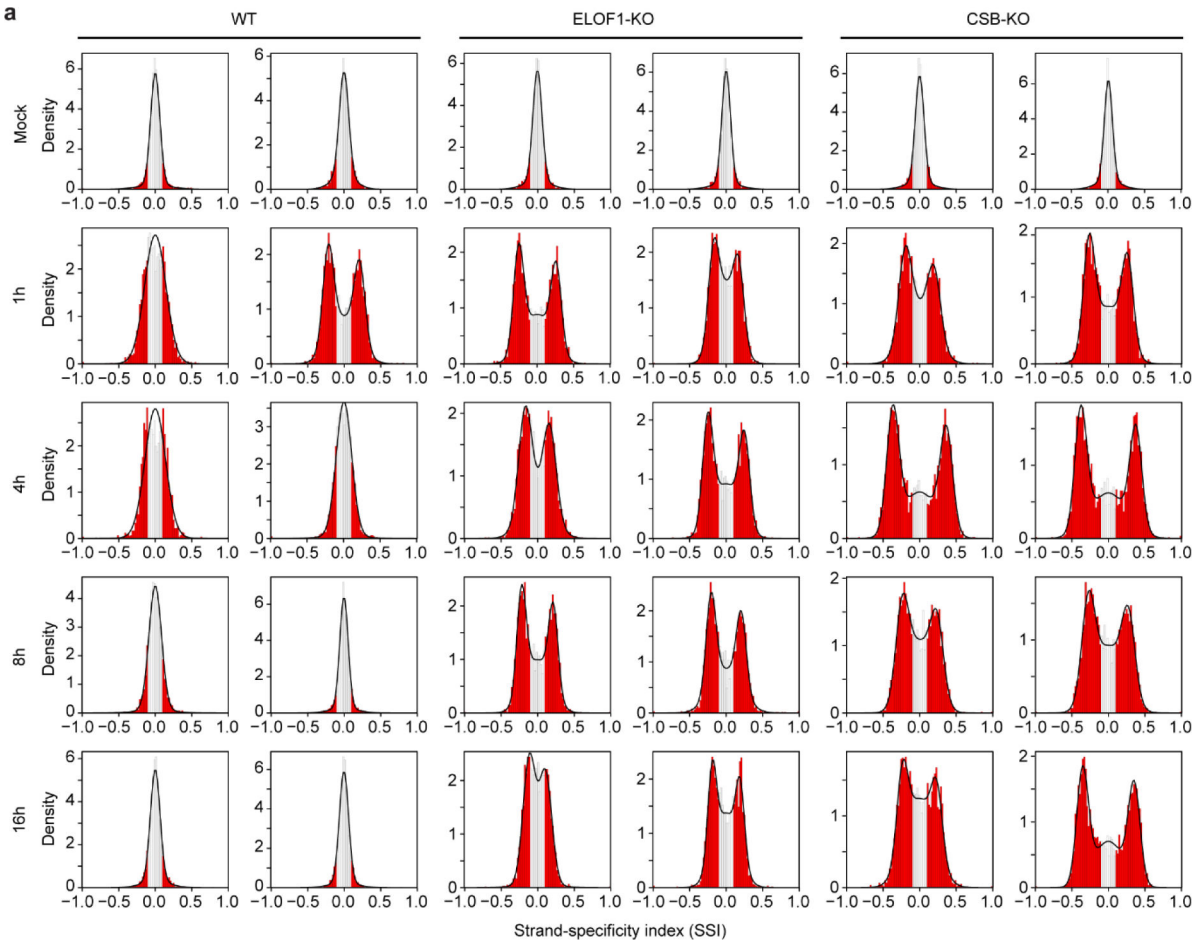
Extended Data Fig. 5. Global gene-expression changes in response to UV irradiation.

Volcano plots of RNA-seq in the indicated RPE1-iCas9 cell lines depicting the downregulation (blue) or upregulation (red) of gene expression in response to UV irradiation (24 h; 9 J/m^2). The fold change (Log_2) is plotted on the x-axis and the significance ($-Log_{10}$ p-value) is plotted on the y-axis. (**a**) WT, (**b**) *CSB*-KO, (**c**) *ELOF1*-KO. Only genes indicating at least 2 counts per million (CPM) in at least 33% of samples were included in the analysis. FDR-adjusted p-values < 0.05 were considered significant. Two short UV-response genes (*ATF3*, *CDKN1A*) are highlighted. (**d**) Box plot depicting the gene length of the 650 most significantly downregulated genes (in blue) or the 650 most significantly upregulated genes (in red). The horizontal line represents the median (center); Upper Bound: gene length scores larger than 75% of all data points; Lower Bound: gene lengths scores shorter than 75% of all data points. Points above and below the box represent the outliers.



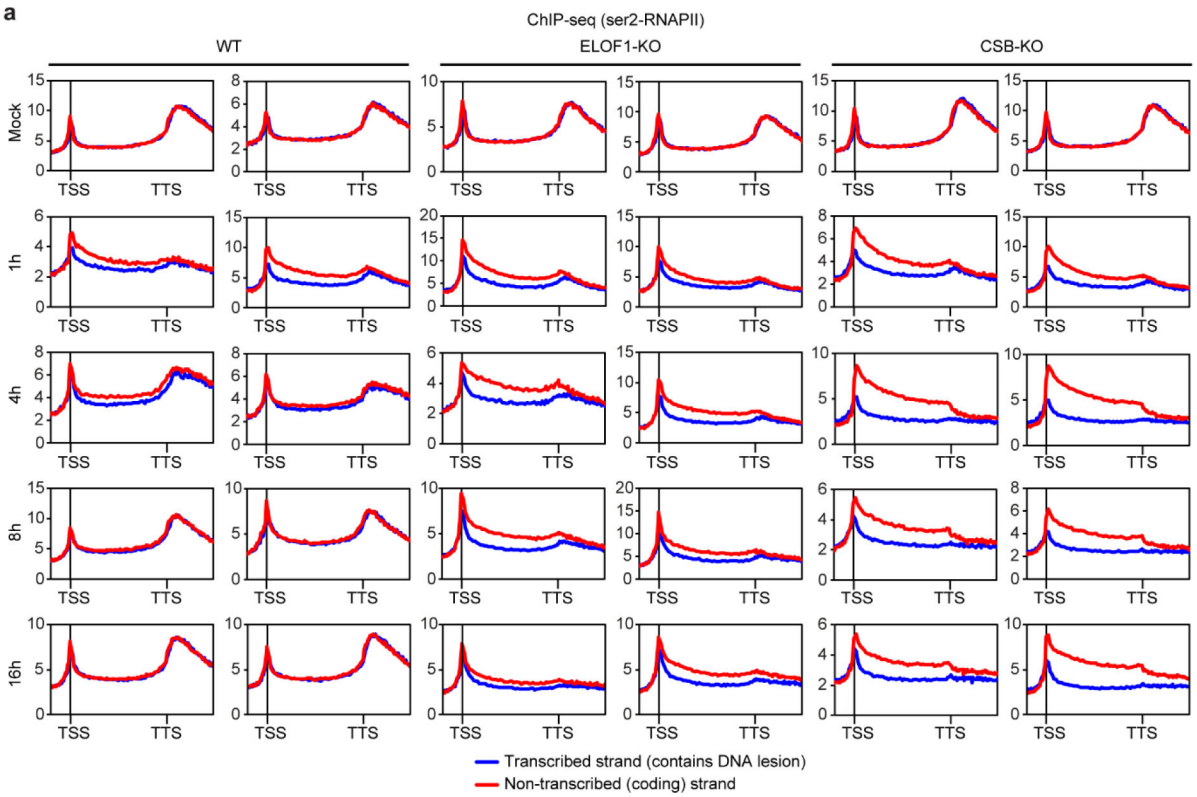
Extended Data Fig. 6. Genome-wide redistribution of RNAPII in response to UV irradiation

(**a**) Heatmaps of pan-RNAPII ChIP-seq data around the TSS of 3,000 genes of 3–100 kb, ranked according to RNAPII signal in mock-treated WT cells. Heatmaps of the same genes are shown after UV irradiation (9 J/m^2) in WT or *ELOF1*-KO cells. (**b**) Averaged metaplots of pan-RNAPII ChIP-seq of 3,000 genes of 3–100 kb from the TSS until the TTS in the indicated RPE1-iCas9 cells after mock-treatment (red) or at 8 h (blue) after UV irradiation (9 J/m^2). (**c**) As in **b** showing the area around the TTS (−2 kb until +2 kb).



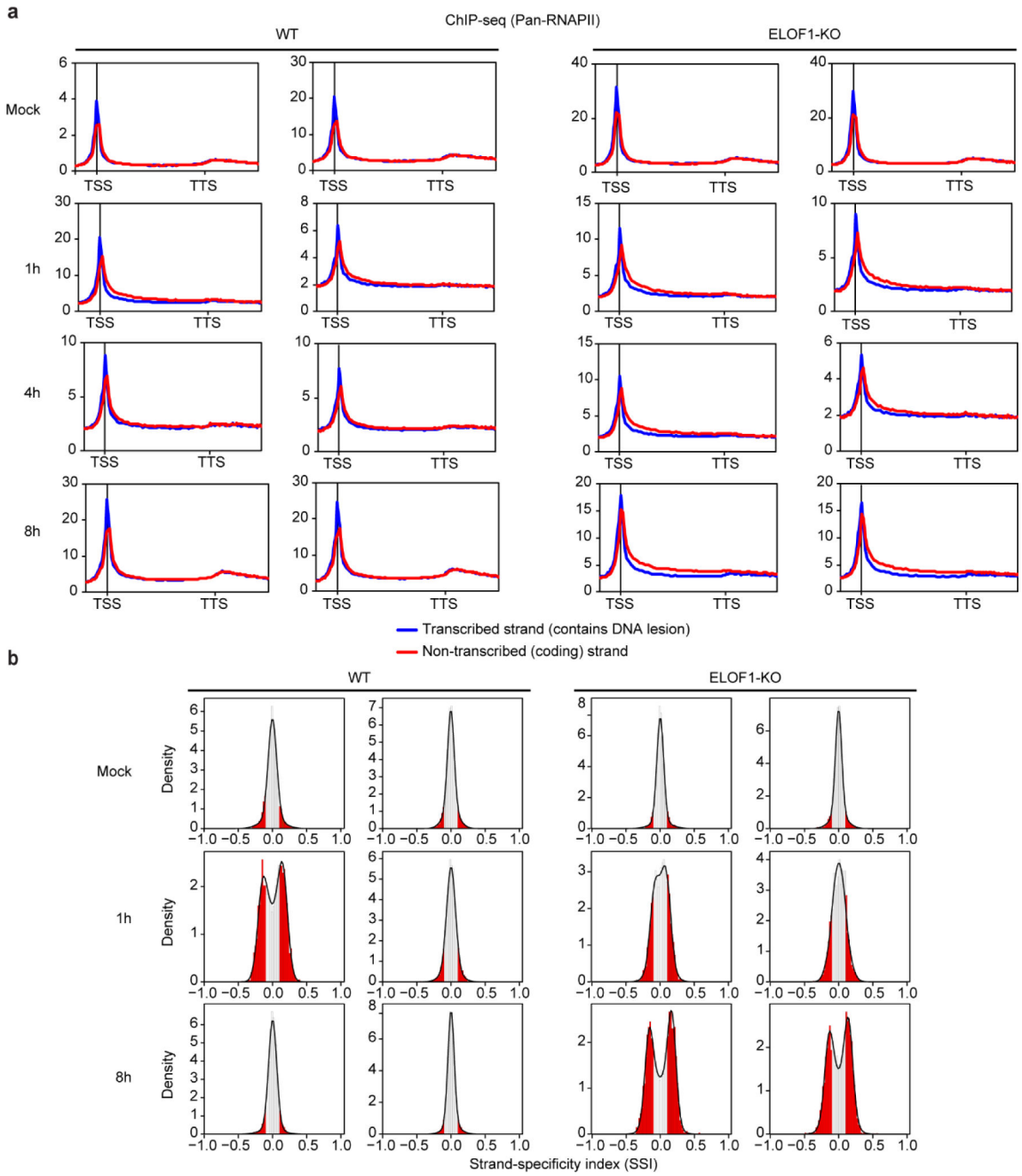
Extended Data Fig. 7. Metaplots of TCR-seq using a Ser2-RNAPII antibody

(**a**) Individual metaplots (two replicates for each condition) of ser2-RNAPII TCR-seq of 3,000 genes for 3–100 kb from the TSS until the TTS (−5 kb and +5 kb, respectively) in the indicated RPE1-iCas9 cells after mock-treatment or at 1 h, 4 h, or 8 h after UV irradiation (9 J/m^2). The coding (non-transcribed) strand in shown in red, while the template (transcribed) strand is shown in blue.



Extended Data Fig. 8. Histogram plots of TCR-seq using a Ser2-RNAPII antibody.

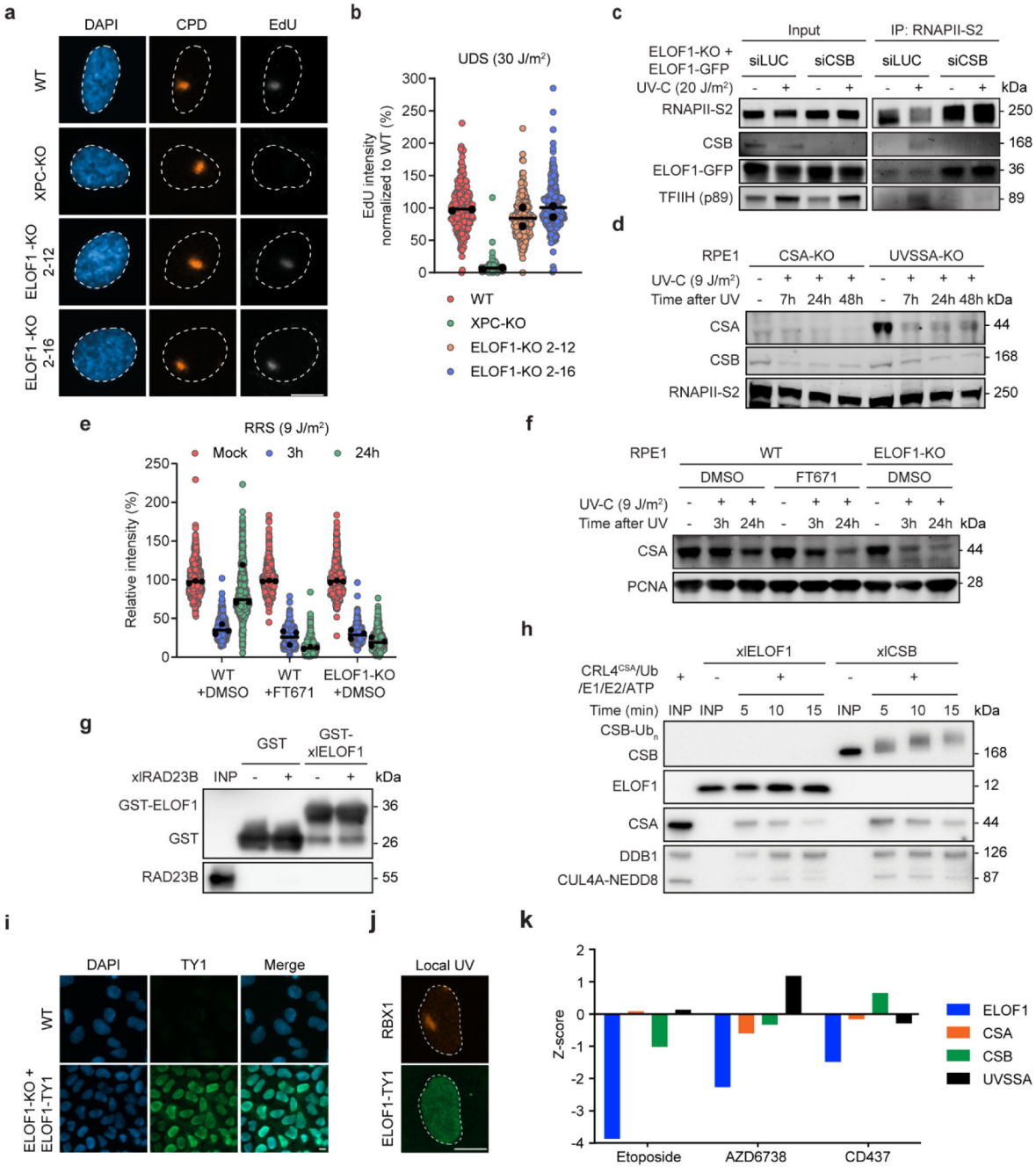
(**a**) Frequency distribution plots of the gene-by-gene ser2-RNAPII strand-specificity index (SSI). SSIs below −0.1 or above 0.1 are presented in red. A unimodal distribution indicates no strand-bias (and thus no DNA damage in the template strand), while a trimodal distribution reflects a strand-bias caused by DNA damage in the template strand.



Extended Data Fig. 9. Validation of TCR-seq with a pan-RNAPII antibody.

(**a**) Individual metaplots (two replicates for each condition) of pan-RNAPII TCR-seq of 3,000 genes of 3–100 kb from the TSS until the TTS (−5 kb and +5 kb, respectively) in the indicated RPE1-iCas9 cells after mock-treatment or at 1 h, 4 h, or 8 h after UV irradiation (9 J/m^2). The coding (non-transcribed) strand in shown in red, while the template (transcribed) strand is shown in blue. (**b**) Frequency distribution plots of the gene-by-gene pan-RNAPII strand-specificity index (SSI). SSIs below −0.1 or above 0.1 are presented in red. A unimodal distribution indicates no strand-bias (and thus no DNA damage in the

template strand), while a trimodal distribution reflects a strand-bias caused by DNA damage in the template strand.



Extended Data Fig. 10. ELOF1 is not involved in global genome repair.

(**a-b**) Unscheduled DNA synthesis (UDS) in the indicated RPE1-iCas9 cells following local UV irradiation (30 J/m^2; 1 h). DNA damage was identified by CPD staining. (a) Representative images (scale bar = 10 μm) and (b) quantification of EdU levels normalized to WT cells. Each black circle represents the median of an independent experiment (*n=2*), each containing 2 technical replicates, >80 cells collected per technical replicate. The black

line represents the median of all the cells collected. (c) Endogenous RNAPIIo Co-IP on U2OS (FRT) *ELOF1*-KO cells complemented with ELOF1^{WT}-GFP after knockdown of CSB (siCSB) or as a control luciferase (siLUC) (*n=2*). (d) Western blot analysis of CSA protein levels in the indicated RPE1-iCas9 cells after mock treatment, or 7 h, 24 h, and 48 h after UV irradiation (9 J/m^2; *n=2*). (e) Quantification of 5-EU levels of the indicated RPE1-iCas9 cells normalized to mock-treated levels for each cell line. Cells were either mock-treated or UV-irradiated (3 h or 24 h; 9 J/m^2). 10 μM FT671 (USP7 inhibitor) was added to the indicated cells 24 h prior to UV irradiation. Each black circle represents the median of an independent experiment (*n=3*), each containing 2 technical replicates, >50 cells collected per technical replicate. The black line represents the median of all the cells collected. (f) Western blot analysis of CSA protein levels from e. (g) GST pull-down of immobilized recombinant *Xenopus laevis* (xl) ELOF1 incubated with recombinant xlRAD23B (*n=2*). (h) In vitro ubiquitylation of recombinant xlELOF1 and xlCSB with recombinant xlCRL4^{CSA}, E1, E2, ubiquitin, and ATP. In vitro ubiquitylation reactions were stopped at the indicated times (*n=3*). (i) Representative images of staining with TY1 antibodies in U2OS (FRT) WT cells and *ELOF1*-KO cells complemented with TY1-tagged ELOF1 (*n=2*). (j) Representative image of staining with TY1 and RBX1 antibodies at 1 h after local UV irradiation (50 J/m^2) in U2OS (FRT) *ELOF1*-KO cells complemented with TY1-tagged ELOF1 (*n=2*). (k) Results of mining a recent CRISPR screen repository[27]. Shown are the Z-scores for the indicated sgRNAs (targeting ELOF1 (blue), CSA (orange), CSB (green), UVSSA (black) after exposure to the indicated genotoxic agents.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge Angela Kragten, Jesper Balk, Jos Poell, Kana Kato, Mayuko Shimada, Susan Kloet, Michelle Paulsen, Maja Vukic, Daniel Warmerdam, Anisha Ramadhin and Lucia Daxinger for help during this project. We also thank Jason Moffat, Katie Chan and Amy Tong for sharing the pLCKO-TKOv3 library prior to publication. We thank the Amsterdam UMC NGS sequencing facilities for support. We thank Peter van Veelen and Arnoud de Ru for mass spectrometry equipment maintenance. This work was funded by an LUMC Research Fellowship, an NWO-ENW-M grant (OCENW.KLEIN.090) and an NWO-VIDI grant (ALW.016.161.320) to MSL, a Leiden University Fund (LUF) grant to DvdH (W18355-2-EM), a KWF/Alpe Young lnvestigator 10701 grant (JdL), a CCA proof-of-concept grant (KdL, RW), an Amsterdam UMC Innovation Grant (CRISPR Expertise Center, 2019; RW), UM1 HG009382 and R01 CA213214 NCI grants to ML, a KWF Young Investigator grant 11367 to RG-P, and an ERC starting grant 310913 to ACOV. JCW was supported by NIH grant HL098316 and is a Howard Hughes Medical Institute (HHMI) Investigator and an American Cancer Society Research Professor. TETM was supported by an EMBO Long-term fellowship (ALTF 1316-2016) and an HHMI fellowship of The Jane Coffin Childs Memorial Fund for Medical Research.

References

1. Brueckner F, Hennecke U, Carell T & Cramer P CPD damage recognition by transcribing RNA polymerase II. Science 315, 859–862 (2007). [PubMed: 17290000]
2. Nakazawa Y et al. Ubiquitination of DNA Damage-Stalled RNAPII Promotes Transcription-Coupled Repair. Cell 180, 1228–1244.e1224 (2020). [PubMed: 32142649]
3. Tufegdžić Vidaković A et al. Regulation of the RNAPII Pool Is Integral to the DNA Damage Response. Cell 180, 1245–1261.e1221 (2020). [PubMed: 32142654]

4. Nakazawa Y et al. Mutations in UVSSA cause UV-sensitive syndrome and impair RNA polymerase IIo processing in transcription-coupled nucleotide-excision repair. Nat Genet 44, 586–592 (2012). [PubMed: 22466610]
5. Schwertman P et al. UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. Nat Genet 44, 598–602 (2012). [PubMed: 22466611]
6. van der Weegen Y et al. The cooperative action of CSB, CSA, and UVSSA target TFIIH to DNA damage-stalled RNA polymerase II. Nat Commun 11, 2104 (2020). [PubMed: 32355176]
7. Laugel V et al. Mutation update for the CSB/ERCC6 and CSA/ERCC8 genes involved in Cockayne syndrome. Human mutation 31, 113–126 (2010). [PubMed: 19894250]
8. Xu J et al. Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. Nature 551, 653–657 (2017). [PubMed: 29168508]
9. Jaspers NG et al. Anti-tumour compounds illudin S and Irofulven induce DNA lesions ignored by global repair and exclusively processed by transcription- and replication-coupled repair pathways. DNA Repair (Amst) 1, 1027–1038 (2002). [PubMed: 12531012]
10. Mair B et al. Essential Gene Profiles for Human Pluripotent Stem Cells Identify Uncharacterized Genes and Substrate Dependencies. Cell reports 27, 599–615 e512 (2019). [PubMed: 30970261]
11. Yu X et al. Up-regulation of human prostaglandin reductase 1 improves the efficacy of hydroxymethylacylfulvene, an antitumor chemotherapeutic agent. J Pharmacol Exp Ther 343, 426–433 (2012). [PubMed: 22895897]
12. Adam S, Polo SE & Almouzni G Transcription recovery after DNA damage requires chromatin priming by the H3.3 histone chaperone HIRA. Cell 155, 94–106 (2013). [PubMed: 24074863]
13. Oksenych V et al. Histone methyltransferase DOT1L drives recovery of gene expression after a genotoxic attack. PLoS Genet 9, e1003611 (2013). [PubMed: 23861670]
14. Boeing S et al. Multiomic Analysis of the UV-Induced DNA Damage Response. Cell reports 15, 1597–1610 (2016). [PubMed: 27184836]
15. Daniels JP, Kelly S, Wickstead B & Gull K Identification of a crenarchaeal orthologue of Elf1: implications for chromatin and transcription in Archaea. Biology direct 4, 24 (2009). [PubMed: 19640276]
16. Ehara H et al. Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. Science 363, 744–747 (2019). [PubMed: 30733384]
17. Ehara H et al. Structure of the complete elongation complex of RNA polymerase II with basal factors. Science 357, 921–924 (2017). [PubMed: 28775211]
18. Prather D, Krogan NJ, Emili A, Greenblatt JF & Winston F Identification and characterization of Elf1, a conserved transcription elongation factor in Saccharomyces cerevisiae. Mol Cell Biol 25, 10122–10135 (2005). [PubMed: 16260625]
19. Veloso A et al. Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. Genome research 24, 896–905 (2014). [PubMed: 24714810]
20. Mayne LV & Lehmann AR Failure of RNA synthesis to recover after UV irradiation: an early defect in cells from individuals with Cockayne's syndrome and xeroderma pigmentosum. Cancer Res 42, 1473–1478 (1982). [PubMed: 6174225]
21. Perdiz D et al. Distribution and repair of bipyrimidine photoproducts in solar UV-irradiated mammalian cells. Possible role of Dewar photoproducts in solar mutagenesis. J Biol Chem 275, 26732–26742 (2000). [PubMed: 10827179]
22. Epanchintsev A et al. Cockayne's Syndrome A and B Proteins Regulate Transcription Arrest after Genotoxic Stress by Promoting ATF3 Degradation. Mol Cell 68, 1054–1066 e1056 (2017). [PubMed: 29225035]
23. Bugai A et al. P-TEFb Activation by RBM7 Shapes a Pro-survival Transcriptional Response to Genotoxic Stress. Mol Cell 74, 254–267 e210 (2019). [PubMed: 30824372]
24. Zhang X et al. Mutations in UVSSA cause UV-sensitive syndrome and destabilize ERCC6 in transcription-coupled DNA repair. Nat Genet 44, 593–597 (2012). [PubMed: 22466612]
25. Fei J & Chen J KIAA1530 protein is recruited by Cockayne syndrome complementation group protein A (CSA) to participate in transcription-coupled repair (TCR). J Biol Chem 287, 35118–35126 (2012). [PubMed: 22902626]

26. Olivieri M et al. A Genetic Map of the Response to DNA Damage in Human Cells. Cell 182, 481–496 e421 (2020). [PubMed: 32649862]
27. de Vivo A et al. The OTUD5-UBR5 complex regulates FACT-mediated transcription at damaged chromatin. Nucleic Acids Res 47, 729–746 (2019). [PubMed: 30508113]
29. Panier S et al. Tandem protein interaction modules organize the ubiquitin-dependent response to DNA double-strand breaks. Mol Cell 47, 383–395 (2012). [PubMed: 22742833]
30. Benedict B et al. WAPL-Dependent Repair of Damaged DNA Replication Forks Underlies Oncogene-Induced Loss of Sister Chromatid Cohesion. Dev Cell 52, 683–698 e687 (2020). [PubMed: 32084359]
31. Doench JG et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 32, 1262–1267 (2014). [PubMed: 25184501]
32. Hart T et al. High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. Cell 163, 1515–1526 (2015). [PubMed: 26627737]
33. Colic M et al. Identifying chemogenetic interactions from CRISPR screens with drugZ. Genome medicine 11, 52 (2019). [PubMed: 31439014]
34. Typas D et al. The de-ubiquitylating enzymes USP26 and USP37 regulate homologous recombination by counteracting RAP80. Nucleic Acids Res (2015).
35. Rappsilber J, Mann M & Ishihama Y Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nature protocols 2, 1896–1906 (2007). [PubMed: 17703201]
36. Tyanova S, Temu T & Cox J The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nature protocols 11, 2301–2319 (2016). [PubMed: 27809316]
37. Tyanova S et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature methods 13, 731–740 (2016). [PubMed: 27348712]
38. Parra I & Windle B High resolution visual mapping of stretched DNA by fluorescent hybridization. Nat Genet 5, 17–21 (1993). [PubMed: 8106079]
39. Garcia-Rubio M, Barroso SI & Aguilera A Detection of DNA-RNA Hybrids In Vivo. Methods Mol Biol 1672, 347–361 (2018). [PubMed: 29043635]
40. Sanz LA & Chedin F High-resolution, strand-specific R-loop mapping via S9.6-based DNA-RNA immunoprecipitation and high-throughput sequencing. Nature protocols 14, 1734–1755 (2019). [PubMed: 31053798]
41. Andrade-Lima LC, Veloso A, Paulsen MT, Menck CF & Ljungman M DNA repair and recovery of RNA synthesis following exposure to ultraviolet light are delayed in long genes. Nucleic Acids Res 43, 2744–2756 (2015). [PubMed: 25722371]
42. Bolger AM, Lohse M & Usadel B Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120 (2014). [PubMed: 24695404]
43. Kim D, Langmead B & Salzberg SL HISAT: a fast spliced aligner with low memory requirements. Nature methods 12, 357–360 (2015). [PubMed: 25751142]
44. Li H et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079 (2009). [PubMed: 19505943]
45. Robinson MD, McCarthy DJ & Smyth GK edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140 (2010). [PubMed: 19910308]
46. Li H Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2 [q-bio.GN] (2013).
47. Tischler G & Leonard S biobambam: tools for read pair collation based algorithms on BAM files. Source Code for Biology and Medicine 9, 13 (2014).
48. McKenna A et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research 20, 1297–1303 (2010). [PubMed: 20644199]
49. Heinz S et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576–589 (2010). [PubMed: 20513432]
50. Team RCR: A language and environment for statistical computing. (2019).

51. Benaglia T CD, Hunter DR, Young D mixtools: An R Package for Analyzing Finite Mixture Models. Journal of Statistical Software 32, 1–29 (2009).

52. Danko CG et al. Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. Mol Cell 50, 212–222 (2013). [PubMed: 23523369]

53. Perez-Riverol Y et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res 47, D442–D450 (2019). [PubMed: 30395289]

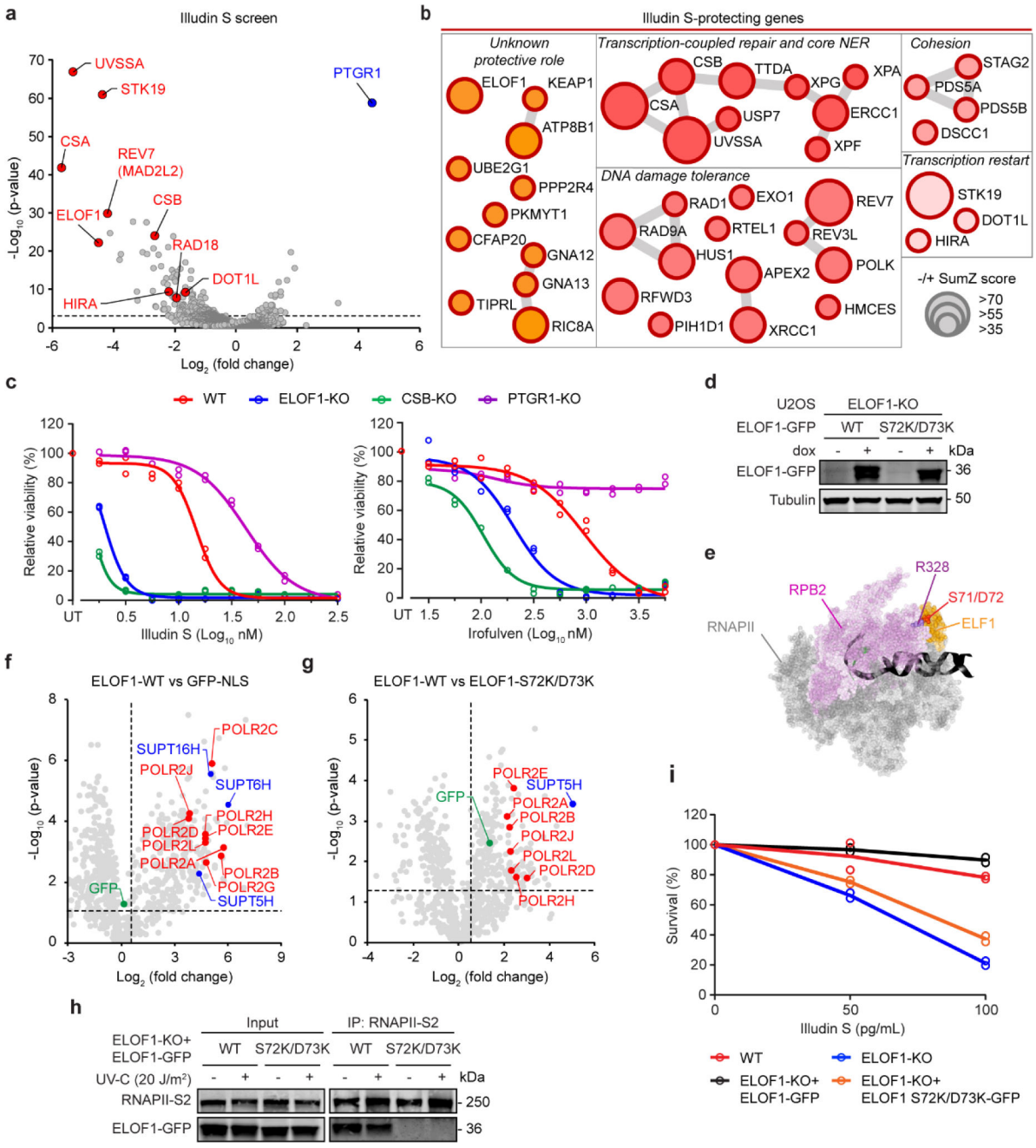


Fig. 1. RNAPII-associated ELOF1 is a putative TCR gene.

(**a**) Volcano plot depicting gene-knockouts sensitizing (red) or conferring resistance (blue) to Illudin S. Fold changes (Log_2) are plotted on the x-axis and significance ($-Log_{10}$ p-value) is plotted on the y-axis (full analysis results in the Source data). (**b**) Network analysis of highly significant hits representing genes essential for Illudin S resistance. Grey lines reflect known protein-protein interactions (Cytoscape, BioGRID). (**c**) 72 h drug sensitivity assays of indicated RPE1-iCas9 cells. The experiment has been performed twice and each symbol represents the median of 6 technical replicates of an independent experiment. (**d**) Western blot analysis of U2OS (FRT) *ELOF1*-KO cells complemented with inducible GFP-

tagged versions of ELOF1. Data shown represent *3* independent experiments. (**e**) Side-view of the structure (PDB: 5XOG) of *K. pastoris* ELF1 (orange) bound to RNAPII (grey) with RPB2 in purple. Residues important for the ELF1-RNAPII interaction are indicated. (**f-g**) Volcano plots depicting the statistical differences between 4 replicates of the MS analysis after GFP immunoprecipitation of mock treated cells comparing (**f**) ELOF1^{WT}-GFP with GFP-NLS (**g**) ELOF1^{WT}-GFP with ELOF1$^{S72K/D73K}$-GFP. The fold change (Log_2) is plotted on the x-axis and the significance (t-test $-Log_{10}$ p-value) is plotted on the y-axis. RNAPII subunits are indicated in red, elongation factors are indicated in blue, and GFP is indicated in green. (**h**) Endogenous RNAPIIo Co-IP on U2OS (FRT) *ELOF1*-KO cells complemented with ELOF1^{WT}-GFP and ELOF1$^{S72K/D73K}$-GFP. Data shown represent *4* independent experiments. (**i**) Clonogenic Illudin S survival of U2OS (FRT) WT, *ELOF1*-KO, and GFP-tagged ELOF1 rescue cell lines. The experiment has been performed twice and each symbol represents the mean of 2 technical replicates of an independent experiment. Uncropped blots and numerical data are provided in Source data fig. 1.

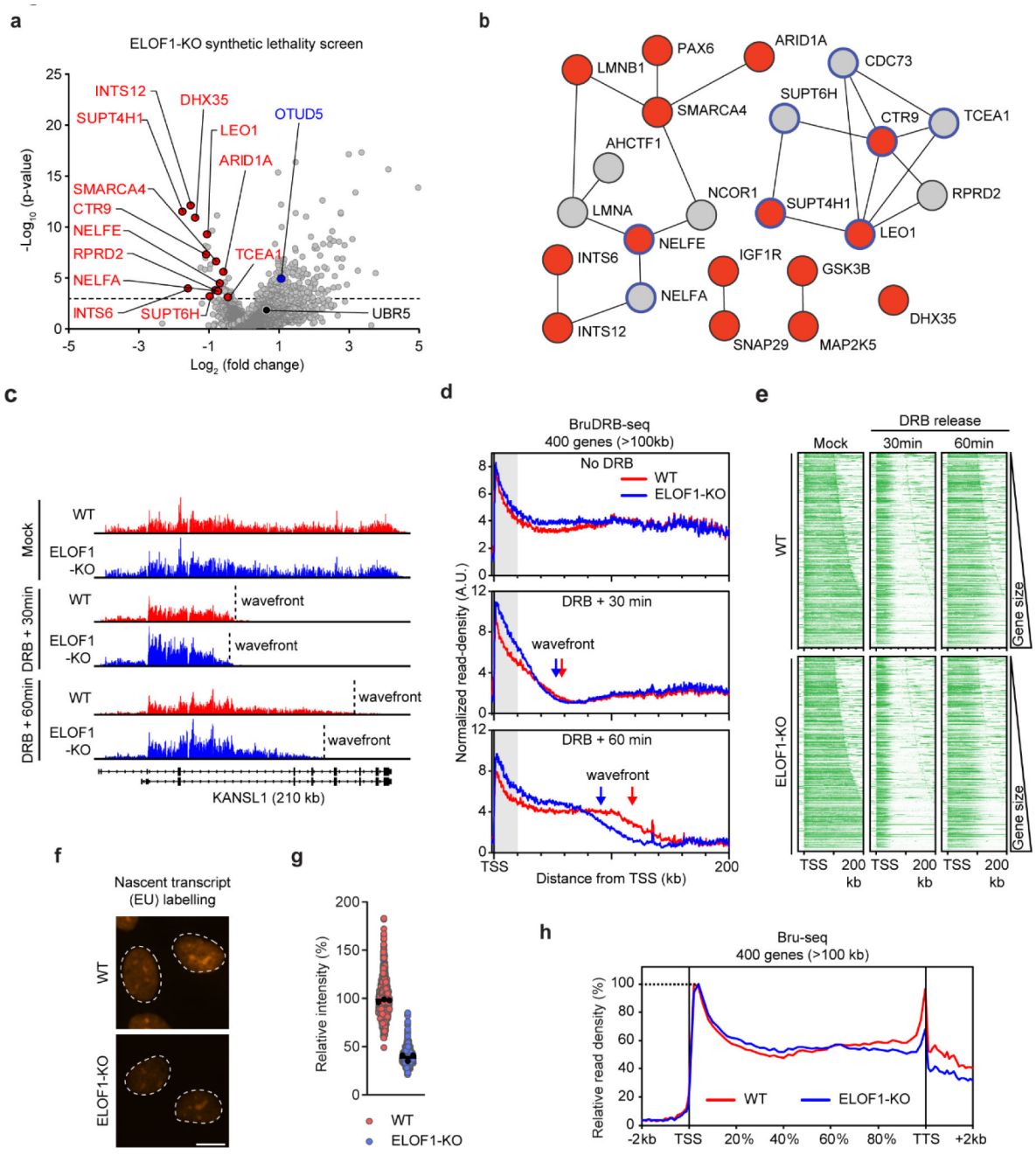


Fig. 2. ELOF1 is required for efficient transcription elongation.

(**a**) Volcano plot depicting gene-knockouts depleted (red) or enriched (blue) in proliferating *ELOF1*-KO cells as compared to WT. The fold change (Log_2) is plotted on the x-axis and significance ($-Log_{10}$ p-value) is plotted on the y-axis (full analysis results in the Source data). (**b**) Network of interacting hit genes (FDR<0.1, red) depleted in *ELOF1*-KO cells compared to WT cells, and depleted interactors (normalized Z-score<−3, grey). Blue edges reflect RNAPII interactors and elongation factors, grey lines indicate other protein-protein interactions. (**c**) UCSC genome browser track showing read density of BrU signal across the *KANSL1* gene after mock treatment, or after DRB wash-out (30 min or 60 min) showing

transcription wavefronts in WT (red) and *ELOF1*-KO cells (blue). (**d**) Metaplots of nascent transcription in 400 genes of >100 kb in WT (red) or *ELOF1*-KO (blue) cells after mock treatment, or after DRB wash-out (30 min or 60 min). Wavefronts are defined as the distance from the TSS where the BrU signal drops below 20% of the positive signal. (**e**) Heatmaps of Bru-seq data from the TSS into the first 200 kb of 400 genes of >100 kb with the highest Bru-seq signal. Genes are ranked according to gene length (left panel). Heatmaps of the same genes after DRB wash-out (30 min or 60 min) in WT or *ELOF1*-KO cells. (**f-g**) Representative images (scale bar = 10 μm) (**f**) and quantification (**g**) of 5-EU levels of RPE1-iCas9 *ELOF1*-KO cells normalized to RPE1-iCas9 WT cells. The experiment has been performed three times and each black circle represents the median of 2 technical replicates of an independent experiment, >80 cells collected per technical replicate. The black line represents the median of all the cells collected. (**h**) Metaplots of BrU signal (nascent transcription) from 2 kb before the TSS to 2 kb after the TTS in 400 genes of >100 kb in WT (red) or *ELOF1*-KO (blue) cells. Profiles are normalized to 100% at promoter-proximal BrU peaks instead of area under the curve for better comparison of transcription profiles. Profiles are averages of 2 independent replicates. Numerical data are provided in Source data fig. 2.

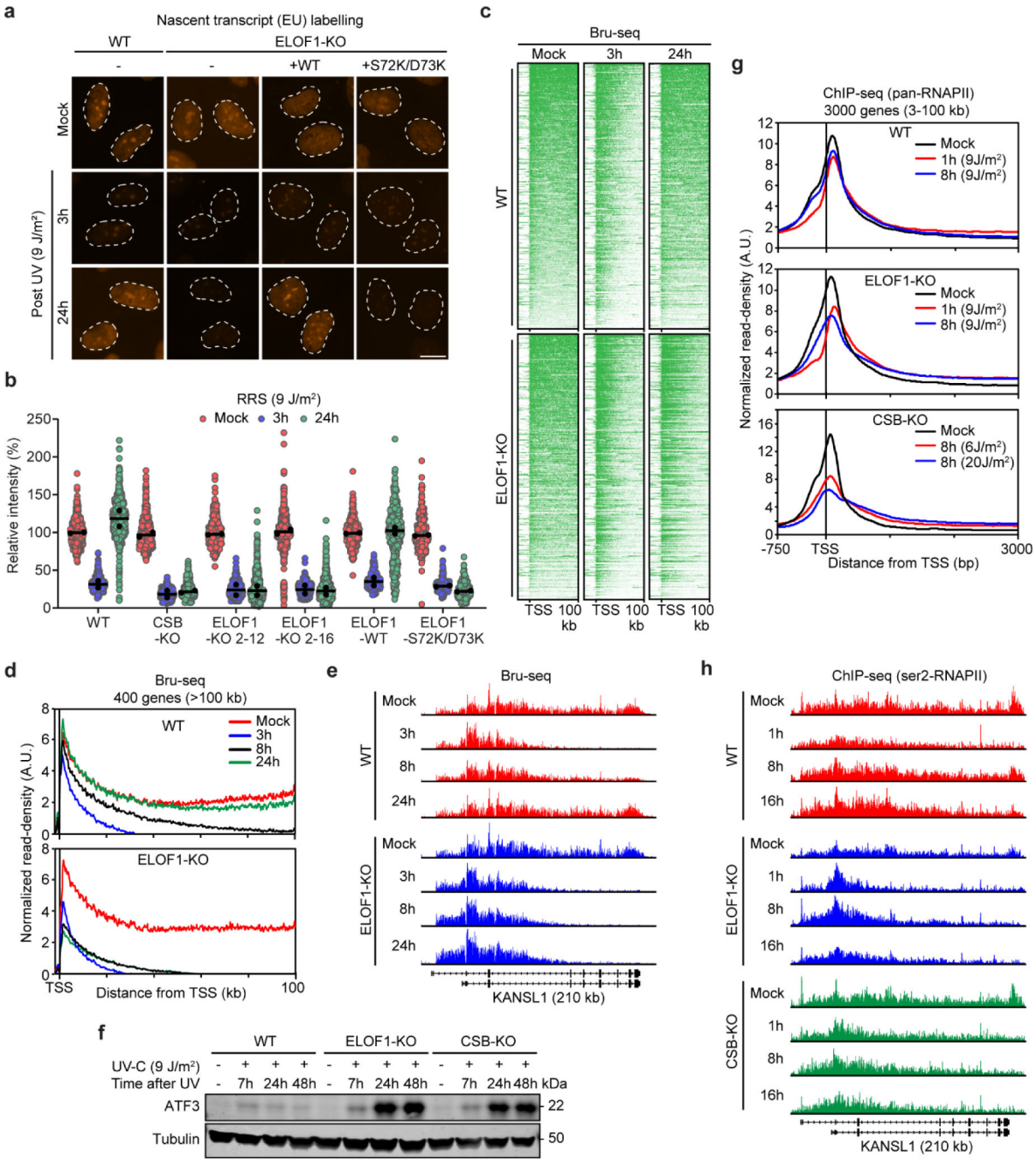


Fig. 3. ELOF1 is essential for transcription recovery after UV.

(**a-b**) Recovery of RNA synthesis (RRS) in the indicated RPE1-iCas9 cells following UV irradiation (3 h or 24 h; 9 J/m^2). (**a**) Representative images (scale bar = 10 μm) and (**b**) quantification of 5-EU levels normalized to mock for each cell line. The experiment has been performed twice and each black circle represents the median of 2 technical replicates of an independent experiment, >80 cells collected per technical replicate. The black line represents the median of all cells. (**c**) Heatmaps of Bru-seq data from the TSS into the first 100 kb of 400 genes of >100 kb ranked according to BrU signal in mock-treated cells (left panel). Heatmaps of the same genes after UV irradiation (9 J/m^2). (**d**) Metaplots

of BrU signal in 400 genes of >100 kb in WT (upper) or *ELOF1*-KO (lower) cells after mock treatment or UV irradiation (9 J/m^2). (**e**) UCSC genome browser track showing BrU read density across the *KANSL1* gene after mock treatment or UV irradiation (9 J/m^2) in WT (red) or *ELOF1*-KO cells (blue). (**f**) Western blot analysis of ATF3 protein levels in the indicated RPE1-iCas9 cell-lines after mock treatment, or UV irradiation (9 J/m^2). Data shown represent *4* independent experiments for WT and *ELOF1*-KO, and 3 independent experiments for *CSB*-KO. (**g**) Averaged metaplots of pan-RNAPII ChIP-seq of 3,000 genes of 3–100 kb around the TSS in RPE1-iCas9 WT (upper) or RPE1 *ELOF1*-KO (middle) after mock-treatment (black) or at 1 h (red) or 8 h (blue) after UV irradiation (9 J/m^2). The lower panel shows metaplots in U2OS *CSB*-KO cells after mock-treatment (black) or 8 h after UV irradiation with either 6 J/m^2 (red) or 20 J/m^2 (blue). (**h**) UCSC genome browser track showing the read density of the ser2-RNAPII signal across the *KANSL1* gene after mock treatment or UV irradiation (9 J/m^2) in WT (red), *ELOF1*-KO (blue) or *CSB*-KO cells (green). Uncropped blots and numerical data are provided in Source data fig. 3.

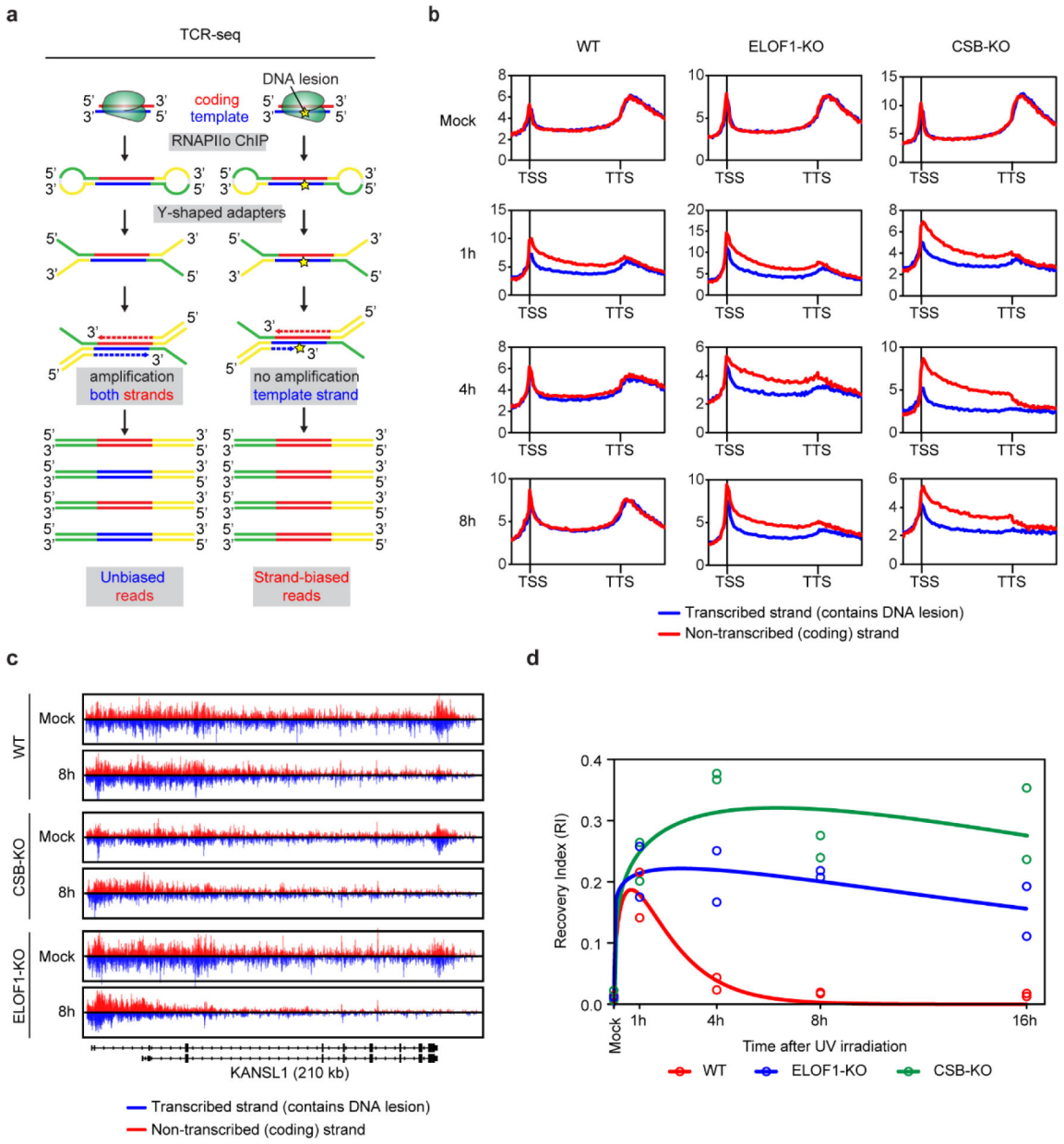


Fig. 4. ELOF1 is essential for repair of transcription-blocking DNA damage.

(**a**) Schematic representation of TCR-seq. The use of asymmetrical adapters allows the preservation of strand-specific information. Without UV the reads in the transcribed strand are similar to the reads in the non-transcribed strand (unbiased reads). After UV irradiation the DNA that is co-purified with RNAPIIo after ChIP is highly enriched for DNA lesions in the transcribed strand, but not in the non-transcribed strand. These DNA lesions prevent PCR amplification during library preparation, resulting in less reads in the transcribed strand compared to the non-transcribed strand (biased reads). (**b**) Averaged metaplots of Ser2-RNAPII TCR-seq of 3,000 genes from the TSS until the TTS (−5 kb, +5 kb respectively) in the indicated RPE1-iCas9 cells after mock-treatment or at 1 h, 4 h, or 8 h after UV irradiation (9 J/m^2). The coding (non-transcribed) strand is shown in red, while

the template (transcribed) strand is shown in blue. The experiment has been performed twice. See Extended Data Fig. 6a for individual replicates and additional timepoints. (**c**) UCSC genome browser track showing read densities of TCR-seq data based on strand-specific Ser2-RNAPII signal across the *KANSL1* gene after mock treatment, or at 8 h after UV irradiation (9 J/m^2) in WT, *CSB*-KO or *ELOF1*-KO cells. Reads from the coding (non-transcribed) strand are shown in red, while reads from the template (transcribed) strand are shown in blue. (**d**) Time-course of the indicated RPE1-iCas9 cells showing the recovery index (RI), representing genome-wide TCR repair kinetics. Per sample, a frequency distribution plot was generated of the per-gene strand specificity index (SSI; defined as the relative difference in read density between the transcribed and non-transcribed strands) of 3,000 genes of 3–100 kb. The RI is subsequently obtained by fitting a mixture of 3 Gaussian distributions, corresponding to undamaged gene fractions (SSI=0) and two unrepaired gene fractions (i.e. |SSI|<>0). The RI was calculated from duplicate time-course experiments shown in Extended Data Fig. 8a.

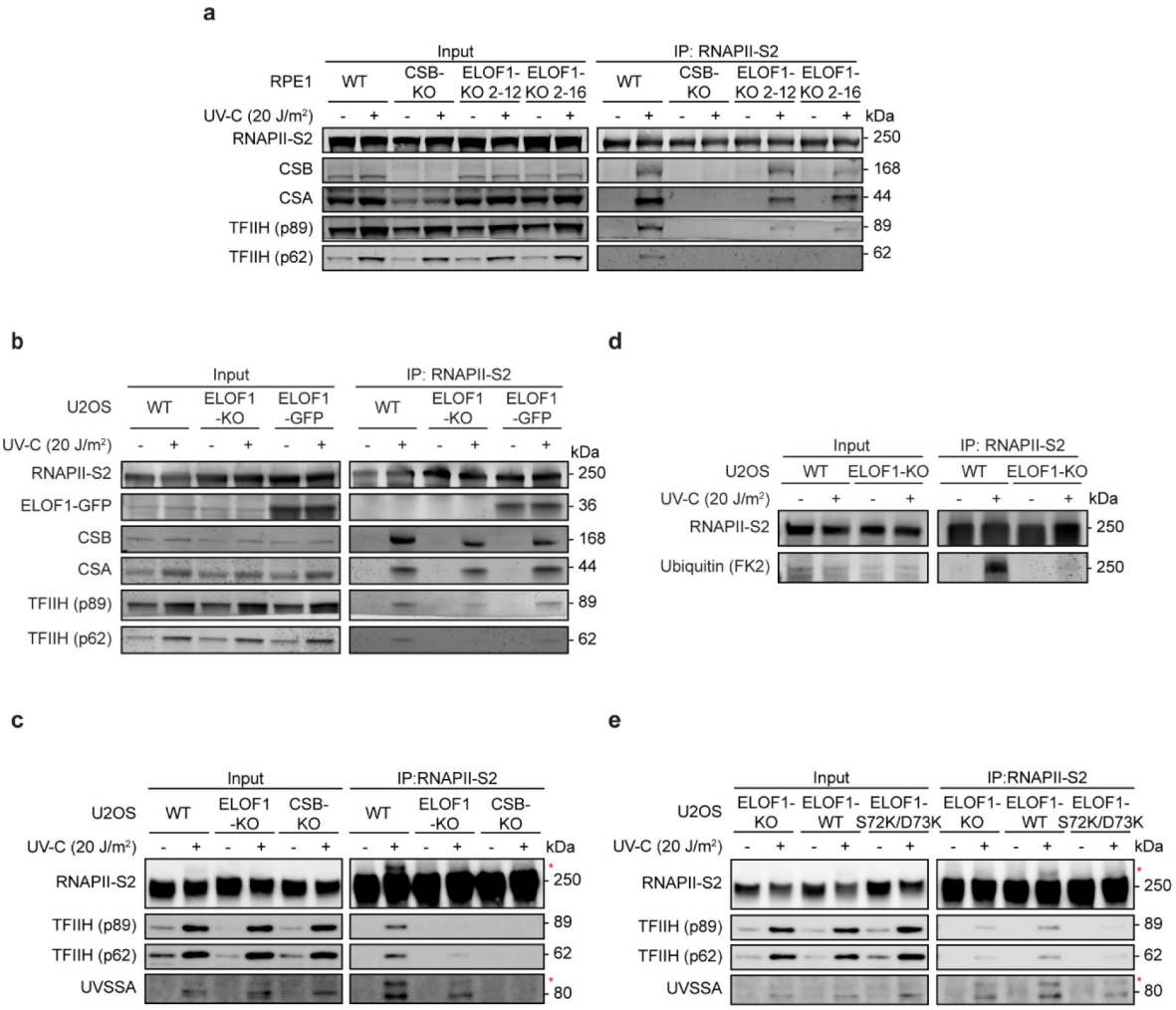


Fig. 5. ELOF1-RNAPII interaction is required for RNAPII ubiquitylation

(**a**) Endogenous RNAPIIo Co-IP on RPE1-iCas9 WT, *CSB*-KO, and *ELOF1*-KO clones. (**b**) Endogenous RNAPIIo Co-IP on U2OS (FRT) WT, *ELOF1*-KO and *ELOF1*-KO complemented with ELOF1^{WT}-GFP. (**c**) Endogenous RNAPIIo Co-IP on U2OS (FRT) WT, *ELOF1*-KO, and *CSB*-KO clones. The ubiquitylated form of RNAPII and UVSSA, detected as higher migrating bands, are indicated with an asterisk. (**d**) Endogenous RNAPIIo co-IP on U2OS (FRT) cells (WT and *ELOF1*-KO) in the presence of deubiquitylase inhibitor N-ethylmaleimide (NEM). The ubiquitylated form of RNAPII is detected using FK2 antibody. (**e**) Endogenous RNAPIIo Co-IP on U2OS (FRT) *ELOF1*-KO and *ELOF1*-KO complemented with ELOF1^{WT}-GFP or ELOF1$^{S72K/D73K}$-GFP. Data shown represent 3 independent experiments for a and d and 2 experiments for b, c and e. Uncropped blots are provided in Source data fig. 5.

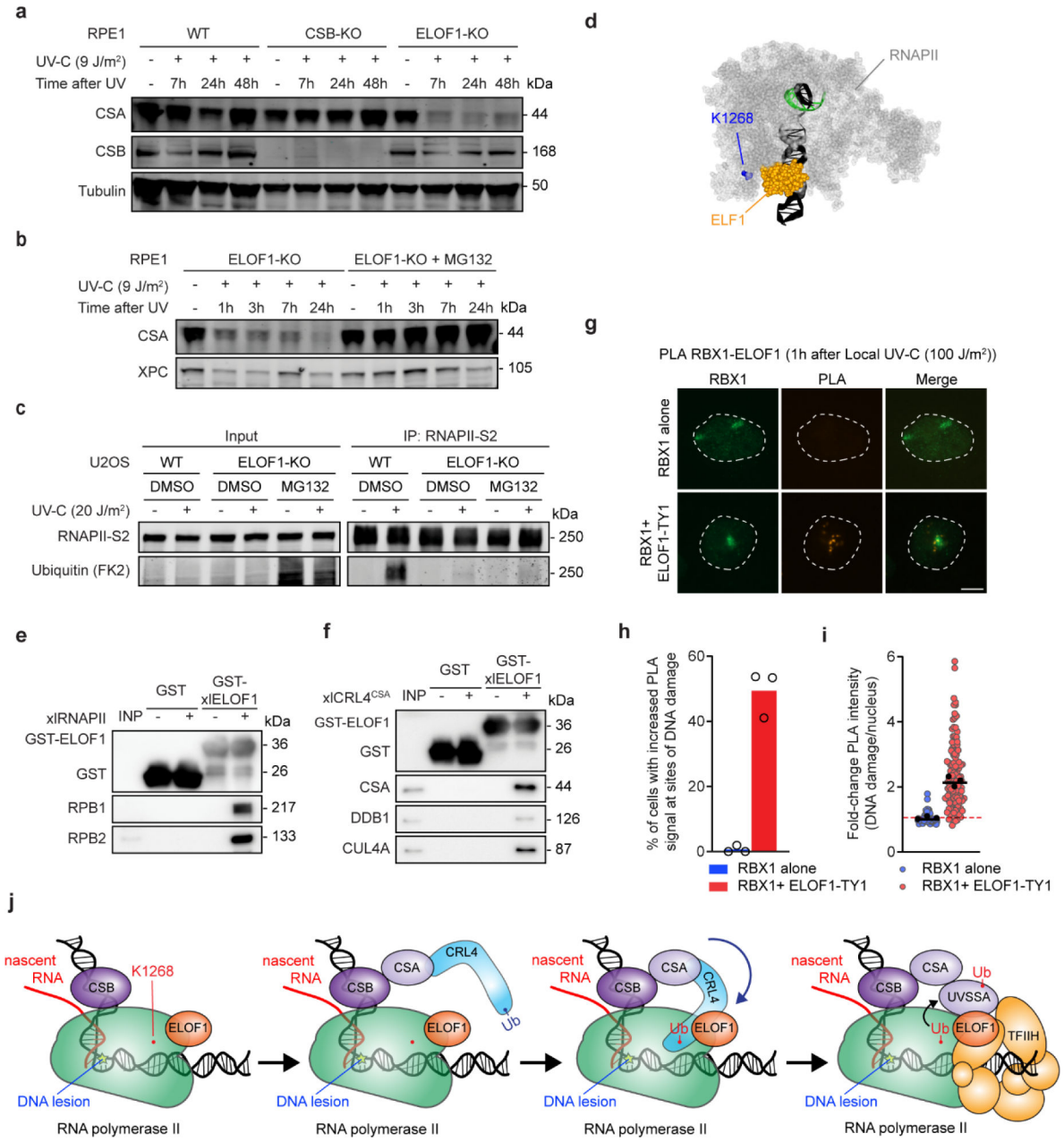


Fig. 6. ELOF1 interacts with the CRL4^{CSA} complex.

(**a-b**) Western blot analysis of CSA protein levels of the indicated RPE1-iCas9 cell-lines after mock treatment or UV irradiation (9 J/m^2). 20 μM MG-132 was added to the indicated cells 1 h prior to UV irradiation. (**c**) Endogenous RNAPIIo Co-IP on U2OS (FRT) cells (WT and *ELOF1*-KO) in the presence of NEM. 20 μM MG-132 was added to the indicated cells 1 h prior to UV irradiation. (**d**) Top-view of the structure (PDB: 5XOG) of *K. pastoris* ELF1 (orange) bound to RNAPII (grey). The RPB1-K1268 ubiquitylation site (K1264 in *K. pastoris*) is indicated in blue. (**e-f**) GST pull-down of immobilized recombinant *Xenopus laevis* (xl) ELOF1 incubated with (**e**) purified xlRNAPII or (**f**) recombinant xlCRL4^{CSA} complex. Data in a-c, e and f represent 2 independent experiments. (**g-i**) Proximity ligation

assay (PLA) between RBX1 and ELOF1-TY1 on sites of local DNA damage (1 h; 100 J/m^2). DNA damage was identified by RBX1 staining and RBX1 alone is a single antibody control. (**g**) Representative images of PLA (scale bar = 5 μm). (**h**) Quantification of the percentage of cells with increased PLA signals at sites of local DNA damage. The experiment has been performed three times and each symbol represents the mean of an independent experiment (*>30 cells collected per experiment*). (**i**) Quantification of the fold change of PLA signal intensities at sites of local DNA damage compared to other nuclear areas. The experiment has been performed three times, each symbol presents the median of an independent experiment and the median of all cells collected is shown as a black line (*>30 cells collected per experiment*). (**j**) Model of how ELOF1 serves as a specificity factor for RNAPII ubiquitylation during TCR. Uncropped blots and numerical data are provided in Source data fig. 6.

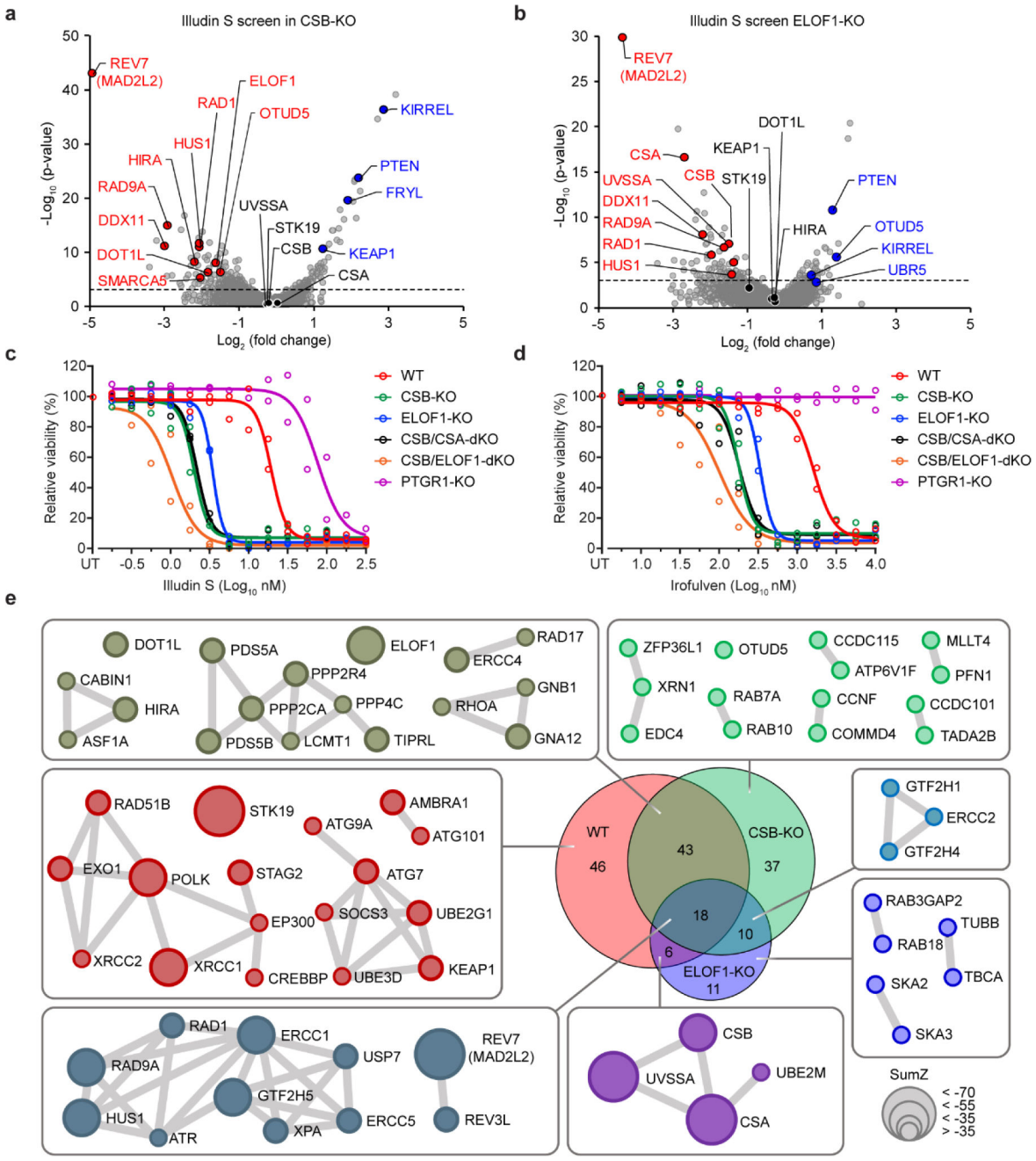


Fig. 7. CRISPR screens identify determinants of Illudin S sensitivity in the absence of ELOF1 or CSB

(**a, b**) Volcano plot depicting gene-knockouts sensitizing (red) or conferring resistance (blue) to Illudin S. (**a**) *CSB*-KO (IC40; 2 nM Illudin S) or (**b**) *ELOF1*-KO (IC40; 5 nM Illudin S) cells. The fold change (Log_2) is plotted on the x-axis and the significance ($-Log_{10}$ p-value) is plotted on the y-axis (full analysis results in the Source data). (**c-d**) 72 h drug-sensitivity assays in the indicated RPE1-iCas9 single and double KO clones exposed to (**c**) Illudin S or (**d**) Irofulven. The experiment has been performed twice and each symbol represents the median of 4 technical replicates of an independent experiment. (**e**) Venn diagram and

network analysis of overlapping and unique hits in the Illudin S screens of RPE1-iCas9 WT, *CSB*-KO and *ELOF1*-KO cell lines. Grey lines reflect known protein-protein interactions (Cytoscape, BioGRID). Numerical data are provided in Source data fig. 7.

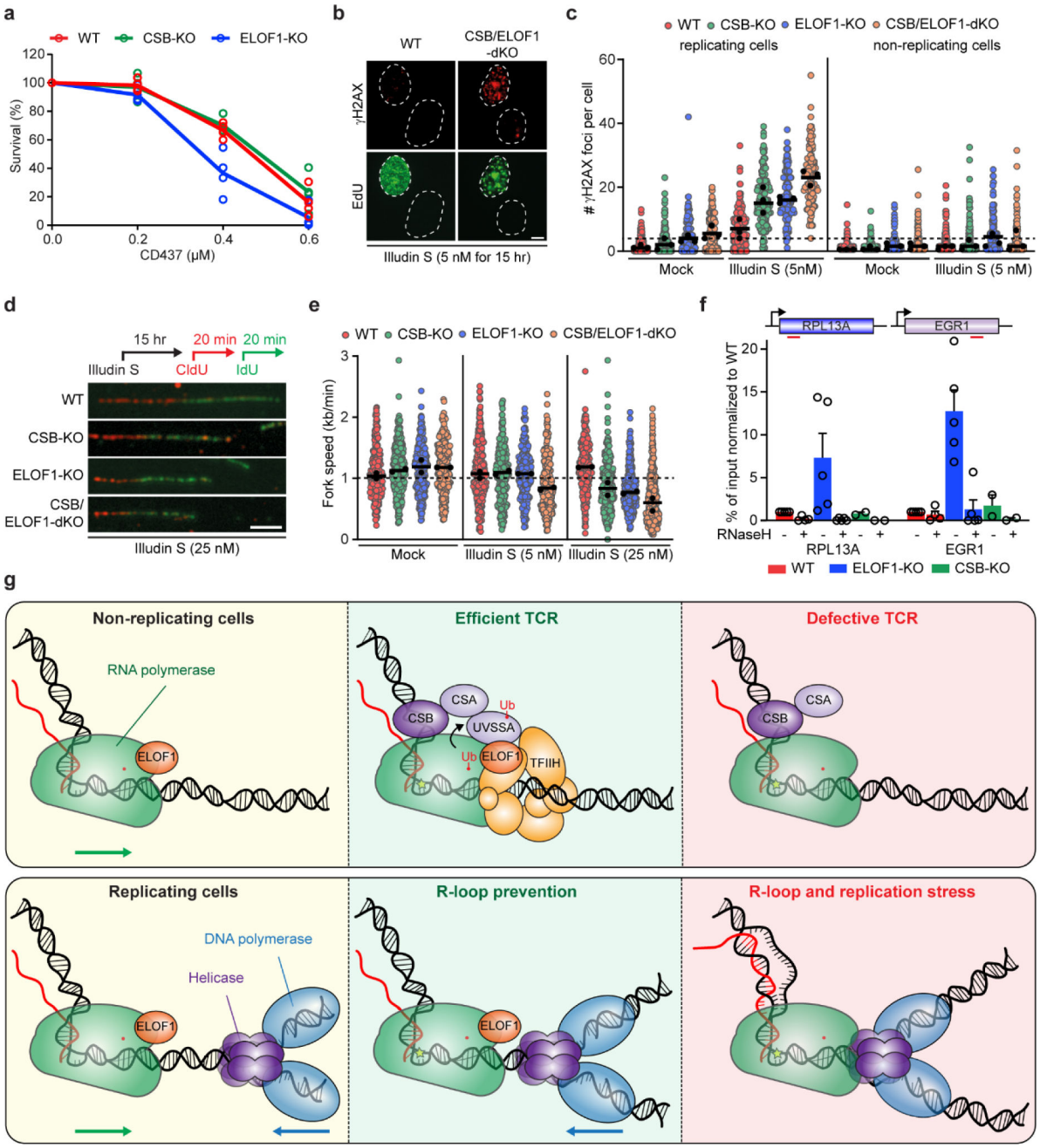


Fig. 8. ELOF1 protects cells against DNA damage during replication.

(a) Clonogenic survival of the indicated RPE1-iCas9 cells after treatment with CD437. The experiment has been performed 4 times and each symbol represents the mean of 2 technical replicates of an independent experiment. (**b-c**) Representative images (scale bar = 5 μm) (**b**) and quantification **(c)** of γH2AX foci in the indicated RPE1-iCas9 cells after treatment with Illudin S (5 nM). Replicating cells were identified by EdU labelling. The experiment has been performed three times and black circles represent the median of an independent experiment (*>26 cells collected per experiment*). The black line represents the median of all the cells collected. (**d**) Top: Schematic representation of the DNA fiber assay. Cells were

mock treated or treated with Illudin S (5 nM or 25 nM) for 15 h, followed by sequential incubation with CldU (red) and IdU (green), each for 20 min. Bottom: Representative images of DNA fibers in the indicated RPE1-iCas9 knockout clones after treatment with Illudin S (25 nM; scale bar = 5 μm). (**e**) Quantification of DNA replication fork speed in the indicated RPE1-iCas9 knockout clones after mock treatment and Illudin S treatment (5 nM or 25 nM). The experiment has been performed twice and black circles represent the median of an independent experiment (*>100 forks scored per experiment*). The black line represents the median of all the scored forks. (**f**) Top: A schematic representation of the RPL13A and EGR1 locus depicting relative position of primer pairs used for DRIP-qPCR (red). Bottom: DRIP-qPCR analysis of the RPL13A and EGR1 genes in the indicated RPE1-iCas9 cells with and without RNase H treatment. Each symbol represents the relative level of DNA-RNA hybrids normalized to input and WT without RNaseH per independent experiment. The bars indicate the mean +SEM of all the experiments (*n=5* for WT and *ELOF1*-KO, *n=2* for *CSB*-KO). (**g**) Model of the role of ELOF1 in non-replicating and replicating cells showing how ELOF1 depletion leads to defective TCR, R-loop accumulation, and replication stress upon encountering transcription-blocking DNA damage. Numerical data are provided in Source data fig. 8.