SUPPLEMENTARY INFORMATION

KIAA1530/UVSSA is responsible for UV-sensitive syndrome that facilitates damage-dependent processing of stalled RNA polymerase IIo in TC-NER

Yuka Nakazawa, Kensaku Sasaki, Norisato Mitsutake, Michiko Matsuse, Mayuko Shimada, Tiziana Nardo, Yoshito Takahashi, Kaname Ohyama, Kosei Ito, Hiroyuki Mishima, Masayo Nomura, Akira Kinoshita, Shinji Ono, Katsuya Takenaka, Ritsuko Masuyama, Takashi Kudo, Hanoch Slor, Atsushi Utani, Satoshi Tateishi, Shunichi Yamashita, Miria Stefanini, Alan R Lehmann, Koh-ichiro Yoshiura & Tomoo Ogi

Supplementary Tables 1, 2a-2c, 3a-3b, and 4

Supplementary Figures 1-15

Supplementary Note (exome details, characterisation of $UV^{S}S24TA$, and CS-patients sequencing)

Correspondence should be addressed to T.O. (mailto: togi@nagasaki-u.ac.jp)

Supplementary Table 1 The reported UV^{S}S cases and cell strains.

Patient	Complementation group	Mutated gene and causal mutation[#]	*UVSSA* SNPs[##]	Reference
Kps2$^+$	UV^{S}S-A	*UVSSA* p.Lys123* (Hom)	R391H (Hom) / P620L (Hom)	1-3
Kps3$^+$	UV^{S}S-A	*UVSSA* p.Lys123* (Hom)	R391H (Hom) / P620L (Hom)	1-3
XP24KO	UV^{S}S-A	*UVSSA* p.Lys123* (Hom)	R391H (Hom) / P620L (Hom)	4,5
UV^{S}S24TA	UV^{S}S-A	*UVSSA* p.Ile31Phefs*9 (Hom)	P620L (Hom)	This study
XP70TO†	inaccurately assigned to XP-E	*UVSSA* p.Cys32Arg (Hom)	R391H (Hom)	6
UVS1KO	UV^{S}S/CS-B	*ERCC6* p.Arg77* (Hom)		2,7,8
CS3AM	UV^{S}S/CS-B	*ERCC6* p.Arg77* (Hom)		9
UV^{S}S1VI	UV^{S}S/CS-A	*ERCC8* p.Trp361Cys (Hom)		10

$^+$siblings. †XP70TO was originally assigned to XP-E based on its failure to complement XP24KO (UV^{S}S-A) by cell fusion assay[6]. This assignment turned out to be erroneous as XP24KO was eventually reassigned from XP-E to UV^{S}S-A[5]. Hom, homozygous for the mutant allele; SNPs, single nucleotide polymorphisms. $^{\#,\#\#}$ The *UVSSA* (*KIAA1530*) causal mutations and SNPs are all identified in this study.

Supplementary Table 2a Overview of the exome sequencing performance in the UV^{S}S-A patients Kps3 and XP24KO.

	Kps3	**XP24KO**
Total reads	69,064,204	69,859,530
Total bases (bp)	5,179,815,300	5,239,464,750
Mapped reads	67,389,162	68,283,271
Mapped reads (%)	97.57	97.74
Uniquely mapped reads	60,935,422	62,123,499
Uniquely mapped reads (%)	90.42	90.98
Uniquely mapped bases (bp)	4,533,582,839	4,612,452,744
Mean read length (bp)	74.61	74.55
Total bases mapped to targets (bp)	2,156,375,625	2,064,988,074
Target bases with at least 2x coverage (%)	96.8	96.5
Target bases with at least 5x coverage (%)	93.9	92.8
Target bases with at least 10x coverage (%)	89.0	86.8
Target bases with at least 20x coverage (%)	78.6	75.5
Target bases with at least 30x coverage (%)	67.7	64.9
Mean target coverage (%)	55.64	53.28

bp, base pairs.

Supplementary Table 2b Overview of the variants identified in the UV^{S}S-A patients Kps3 and XP24KO.

	Kps3	XP24KO
Single nucleotide variants (SNVs)		
Total	75,368	94,628
coding	15,460	14,986
synonymous	8,353	8,131
non-synonymous missense	7,042	6,795
stopgain	54	50
stoploss	11	10
splice site*	54	49
Insertion and deletions (indels)		
Total	15,966	19,665
coding	304	287
frameshift deletion	62	63
frameshift insertion	96	68
in-frame deletion	84	96
in-frame insertion	61	56
in-frame stopgain	1	2
in-frame stoploss	0	2
splice site*	23	20

The number of variants that meet the left criteria is indicated for each patient. *Splice site, splice-site acceptor or donor variants within 2-bp away from exon/intron boundaries.

Supplementary Table 2c Novel homozygous and compound heterozygous variants in the candidate genes identified by the exome sequencing of the UV^{S}S-A patients Kps3 and XP24KO.

Patient	Chr.	Position	Genotype	Gene	Mutation type	RefSeq accession	Nucleotide change	Amino-acid change	SIFT
Kps3	1	908,382	Hom	*PLEKHN1*	ns	001160184	c.C1180G	p.H394D	0.13
XP24KO	1	3,511,910	Hom	*MEGF6*	ns	001409	c.G368A	p.C123Y	0
Kps3	1	22,149,832	Het	*HSPG2*	ns	005529	c.A13153G	p.N4385D	0
Kps3	1	22,178,067	Het	*HSPG2*	ns	005529	c.G7130A	p.R2377H	0.02
XP24KO	1	197,479,954	Het	*DENND1B*	ns	001195215	c.G1964T	p.G655V	0.1
XP24KO	1	197,479,955	Het	*DENND1B*	ns	001195215	c.G1963A	p.G655S	0.42
XP24KO	1	197,480,875	Het	*DENND1B*	ns	001195215	c.A1798G	p.I600V	0.04
Kps3	1	203,053,820	Het	*MYOG*	ns	002479	c.A508T	p.S170C	none
Kps3	1	203,054,659	Het	*MYOG*	ns	002479	c.G431A	p.R144H	0.25
XP24KO	1	237,947,044	Hom	*RYR2*	fs del	001035	c.12032delA	p.E4011fs	none
XP24KO	1	248,813,343	Hom	*OR2T27*	fs del	001001824	c.843delT	p.L281fs	none
Kps3	2	73,675,228	Hom	*ALMS1*	ins	015120	c.1571_1572insTCA	p.S524delinsSH	none
XP24KO	2	207,169,533	Hom	*ZDBF2*	ns	020923	c.A281G	p.E94G	0.21
Kps3 / XP24KO	**4**	**1,343,580**	**Hom**	***UVSSA (KIAA1530)***	**sg**	**020894**	**c.A367T**	**p.K123X**	**none**
Kps3	4	152,567,721	Hom	*FAM160A1*	fs ins	001109977	c.1238_1239insT	p.P413fs	none
Kps3	5	115,298,503	Het	*AQPEP*	ns	173800	c.G189C	p.R63S	0.28
Kps3	5	115,329,552	Het	*AQPEP*	sp				none
Kps3	5	130,764,655	Het	*RAPGEF6*	ns	016340	c.A4720G	p.N1574D	0.02
Kps3	5	130,928,145	Het	*RAPGEF6*	ns	001164387	c.C212A	p.A71D	0
Kps3	6	2,766,471	Hom	*WRNIP1*	fs ins	020135	c.615_616ins7G1C	p.G205fs	none
XP24KO	6	3,287,129	Hom	*SLC22A23*	ns	021945	c.G667A	p.A223T	0.01
Kps3	6	31,593,814	Het	*PRRC2A*	fs ins	004638	c.857_858insG	p.A286fs	none
Kps3	6	31,599,212	Het	*PRRC2A*	ns	004638	c.G2762A	p.R921H	0.01
Kps3	6	32,021,315	Hom	*TNXB*	ns	019105	c.G8635T	p.G2879W	0.01
Kps3	7	39,612,003	Het	*C7orf36*	ns	020192	c.C379T	p.P127S	0.2
Kps3	7	39,612,012	Het	*C7orf36*	ns	020192	c.G388T	p.V130F	0.03
XP24KO	7	98,529,126	Het	*TRRAP*	ns	003496	c.C3690G	p.I1230M	0.02
XP24KO	7	98,565,286	Het	*TRRAP*	ns	003496	c.A7402G	p.M2468V	0.03
Kps3	7	103,202,149	Het	*RELN*	ns	005045	c.C5359T	p.R1787W	0.01
Kps3	7	103,368,614	Het	*RELN*	ns	005045	c.G697A	p.A233T	0.68
XP24KO	7	117,368,252	Het	*CTTNBP2*	ns	033427	c.C3946T	p.R1316C	0.04
XP24KO	7	117,450,990	Het	*CTTNBP2*	ns	033427	c.T243A	p.N81K	0.01
XP24KO	7	141,464,321	Het	*TAS2R3*	ns	016943	c.G363C	p.W121C	0
XP24KO	7	141,464,322	Het	*TAS2R3*	ns	016943	c.C364T	p.L122F	0.01
XP24KO	9	107,560,803	Het	*ABCA1*	ns	005502	c.G5020A	p.V1674I	0.03
XP24KO	9	107,584,945	Het	*ABCA1*	ns	005502	c.G2660T	p.C887F	0.36
Kps3	9	136,320,698	Hom	*ADAMTS13*	ns	139025	c.G3541A	p.G1181R	0.06
XP24KO	11	67,790,188	Hom	*ALDH3B1*	ns	000694	c.G1073A	p.G358E	0
XP24KO	12	32,137,670	Hom	*C12orf35*	ns	018169	c.A3781G	p.K1261E	0.01
XP24KO	12	33,030,861	Hom	*PKP2*	ns	004572	c.A953C	p.H318P	0.16

XP24KO	12	40,012,658	Hom	*ABCD2*	ns	005164	c.A760G	p.I254V	0.4
XP24KO	12	52,699,141	Hom	*KRT86*	ns	002284	c.A853T	p.I285F	0
XP24KO	12	53,291,324	Hom	*KRT8*	ns	002273	c.C1340T	p.A447V	0.5
XP24KO	12	76,881,324	Hom	*OSBPL8*	ns	020841	c.G8T	p.G3V	0.03
XP24KO	12	77,417,815	Hom	*E2F7*	ns	203394	c.C2716T	p.P906S	1
XP24KO	13	100,617,741	Hom	*ZIC5*	ns	033132	c.G1882A	p.A628T	0.11
XP24KO	14	101,348,750	Hom	*RTL1*	ns	001134888	c.C2376A	p.N792K	none
Kps3	16	29,998,996	Het	*TAOK2*	ns	016151	c.C3403T	p.R1135C	none
Kps3	16	30,002,411	Het	*TAOK2*	ns	004783	c.G2672T	p.R891L	0
Kps3	16	58,071,385	Hom	*MMP15*	fs ins	002428	c.172_173insG	p.R58fs	none
XP24KO	17	79,166,385	Het	*AZI1*	ns	001009811	c.C2330T	p.A777V	0.08
XP24KO	17	79,166,386	Het	*AZI1*	ns	001009811	c.G2329T	p.A777S	0.27
Kps3	19	9,577,032	Het	*ZNF560*	sp				none
Kps3	19	9,577,374	Het	*ZNF560*	ns	152476	c.G2249A	p.R750H	0.04
Kps3	19	43,773,571	Hom	*PSG9*	fs ins	002784	c.13_14insT	p.P5fs	none
XP24KO	19	54,803,484	Hom	*LILRA3*	fs del	001172654	c.340delC	p.L114fs	none

Abbreviations used in this table: Chr., chromosome; position, refseq position of the variant; Genotype, Hom / homozygous, Het / heterozygous; Gene, gene symbol in which the variant is located; Mutation type, ns / nonsynonymous missense, fs del / frameshift deletion, fs ins / frameshift insertion, ins / in-frame insertion, sg / nonsynonymous stop gain, sp / splice site variant; RefSeq accession, RefSeq gene accession no.; Exon, exon number of the gene in which the variant is located; Nucleotide change, base change and base number resulting from the observed variant; Amino acid change, amino acid change and codon number resulting from the observed variant; SIFT, SIFT score of the predicted effect of the substitution on protein function.

Supplementary Table 3a Shared run of homozygosity (ROHs) between the four UV^{S}S-A patients Kps2, Kps3, UV^{S}S24TA, and XP24KO.

Chr.	intersection start	intersection end	Length (bps)	Cytoband	# of SNP markers	RefSeq Genes
4	1,002,125	2,035,597	1,033,472	4p16.3	170	*C4orf42, CRIPAK, CTBP1, FAM53A, FGFR3, FGFRL1,* ***UVSSA*** **(*KIAA1530*)**, *LETM1, LOC100130872, LOC100130872, SPON2, MAEA, MIR943, RNF212, SCARNA22, SLBP, SPON2, TACC3, TMEM129, WHSC1, WHSC2*
8	46,940,022	48,945,620	2,005,598	8q11.1-q11.21	204	*BEYLA, CEBPD, KIAA0146, MCM4, PRKDC, UBE2V2*
12	88,295,662	89,331,382	1,035,720	12q21.32-q21.33	235	*C12orf29, C12orf50, CEP290, KITLG, TMTC3*

Chr., chromosome.

Supplementary Table 3b Details of the ROHs on chromosome 4.

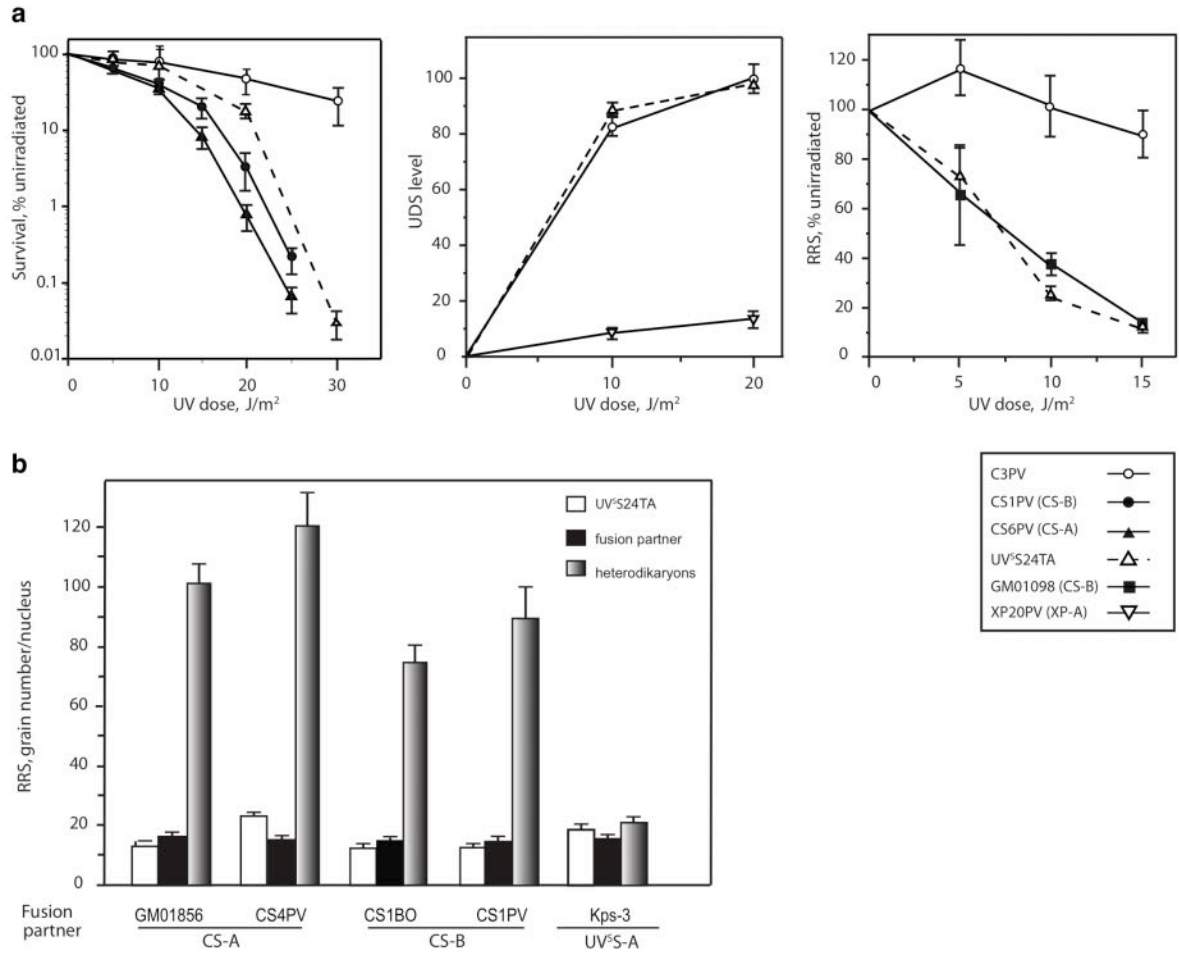
Chr.	intersection start	intersection end	Length (bps)	Cytoband	heterozygous rate	# of SNP markers	UV^{S}S-A patient
4	1,002,125	2,370,341	1,368,217	4p16.3	0.000000	237	Kps2
4	1,002,125	2,370,341	1,368,217	4p16.3	0.000000	237	Kps3
4	956,047	5,139,076	4,183,030	4p16.3-p16.2	0.002717	1,104	UV^{S}S24TA
4	45,410	2,035,597	1,990,188	4p16.3	0.006780	295	XP24KO
Overlap	1,002,125	2,035,597	1,033,472	4p16.3		170	-

Chr., chromosome.

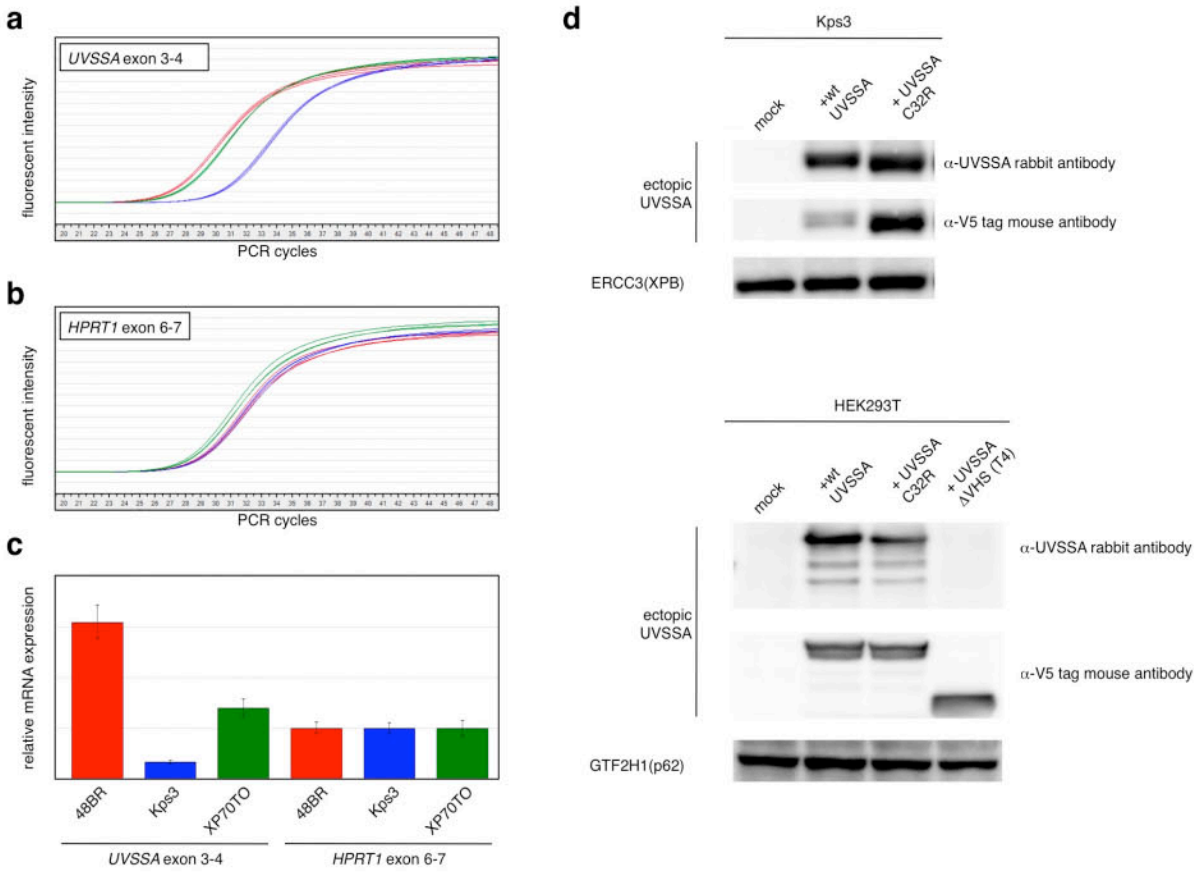
Supplementary Table 4 List of CS-patients of unassigned genotype and novel nonsynonymous variants detected in the *UVSSA* gene.

Cell strain No.	Code	Origin	UVSSA novel SNVs	Cell strain No.	Code	Origin	UVSSA novel SNVs
1	CS2LO	UK		37	CS24BR	UK	p.D288V (het)
2	CS4LO	UK		38	CS2SO	UK	
3	CS1SO	UK		39	CS3LE	UK	
4	CS9LO	UK		40	CS12MA	UK	
5	CS1BEL	UK		42	CS1NE	UK	
6	CS11LO	UK		43	CS1PR	UK	
7	CS2BEL	UK		44	CS32LO	UK	
8	CS3BI	UK		45	CS2LI	UK	
9	CS2MA	UK		46	CS33LO	UK	
10	CS1BL	UK		47	CS13MA	Pakistan	
12	CS14LO	UK		48	CS3SH	UK	
13	CS15LO	UK		49	CS1GLO	UK	
14	CS2LE	UK		50	CS1BR	France	
15	CS6MA	UK		51	CS1GO	Sweden	
16	CS5MA	UK		52	CS2GO	Sweden	
17	CS19LO	UK		53	CS4TAN	Turkey	
18	CS3BL	UK		54	CS1GR	Austria	
19	CS1LI	UK		55	CS8TAN	Turkey	
20	CS20LO	UK		58	CS2GR	Austria	
21	CS1WR	UK		59	CS12RO	Italy	p.H296L (het)
22	CS7MA	UK		60	CS1NY	USA	
23	CS8MA	UK		61	CS1GGO	Austria	
24	CS14BR	UK		63	CS1HAI	Israel	
25	CS24LO	UK		65	CS9TAN	Turkey	
27	CS26LO	UK		66	CS1SYA	Australia	
28	CS27LO	UK		67	CS1USAU	USA	
29	CS17BR	UK		68	CS18BR	Germany	
30	CS19BR	UK	p.E162K (het)	70	CS3ROC	Taiwan	
32	CS10MA	UK	p.A175V (het)	72	CS22BR	Brazil	
33	CS11MA	UK					
34	CS1OX	?					
36	CS23BR	UK					

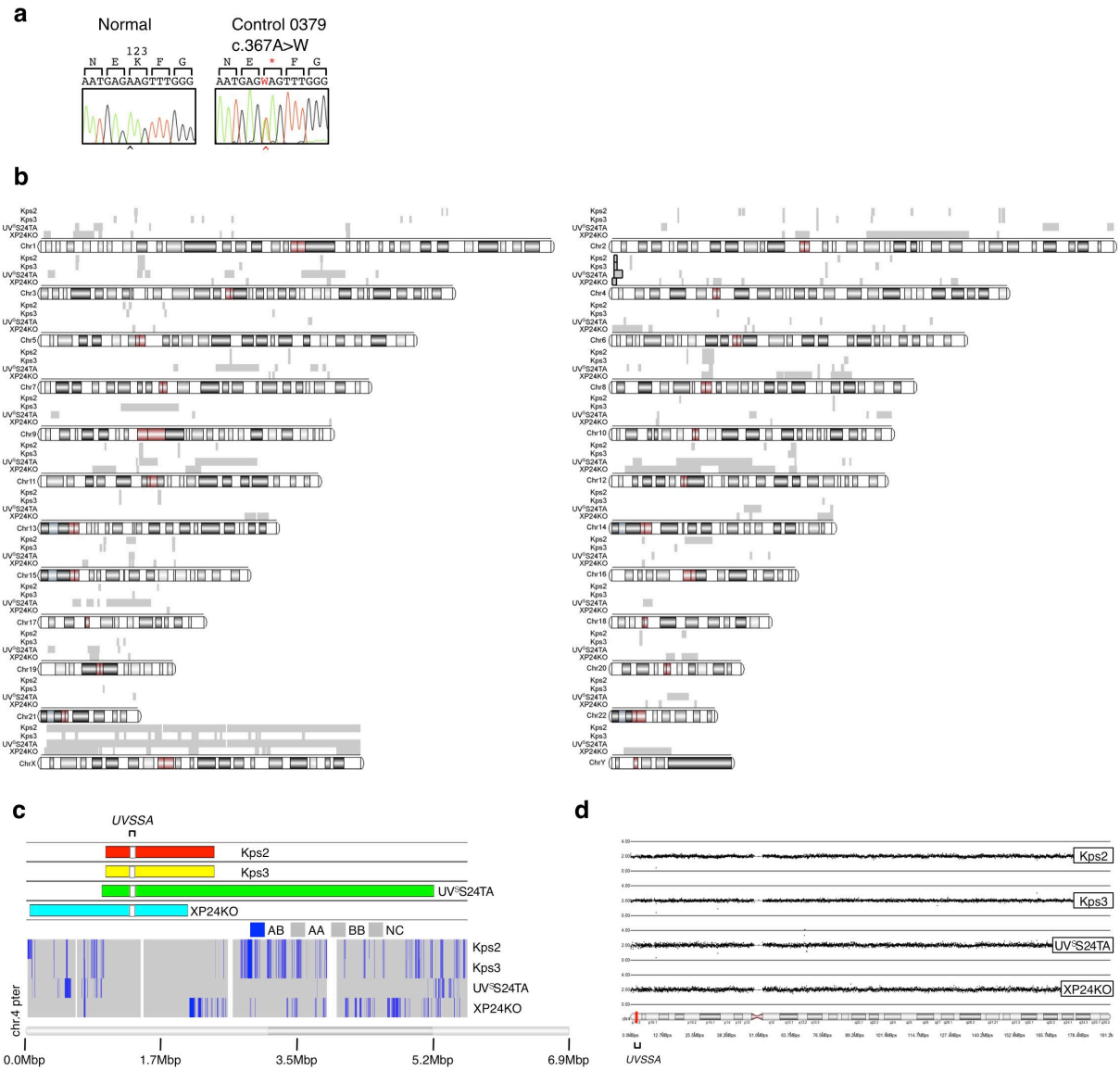
het, heterozygous for the allele



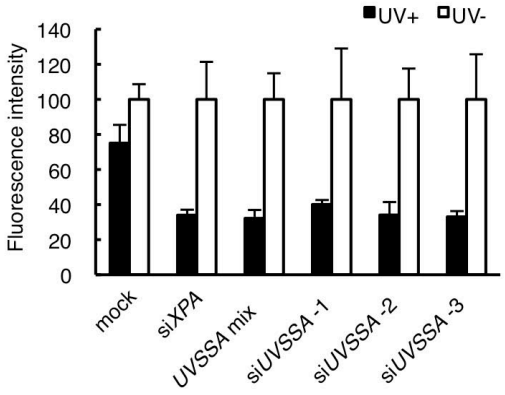
Supplementary Figure 1 The patient UV^{S}S24TA shows the cellular response to UV-irradiation typical of CS and belongs to the UV^{S}S-A complementation group. (**a**) UV-irradiated UV^{S}S24TA fibroblasts show normal level of UV-induced DNA repair synthesis (UDS) and drastically reduced levels of survival and recovery of RNA synthesis (RRS). The analysis was performed using standard procedures (see **Supplementary Note**)[11-13] (**b**) RRS levels in the heterodikaryons obtained after fusion of UV^{S}S24TA cells with the UV^{S}S-A Kps3 cells are similar to those in the homodikaryons, indicating the occurrence of the same genetic defect in the two fusion partners. Conversely, fusions with CS cells representative of groups A and B result in increased RNA synthesis levels in the heterodikaryons compared to those in the homodikaryons, indicating the presence of different genetic defects in the fusion partners. Genetic analysis of the DNA repair defect in UV^{S}S24TA cells was carried out by evaluating the recovery of RNA synthesis after UV-irradiation in classical complementation tests based on somatic cell hybridization[14] (see **Supplementary Note**).



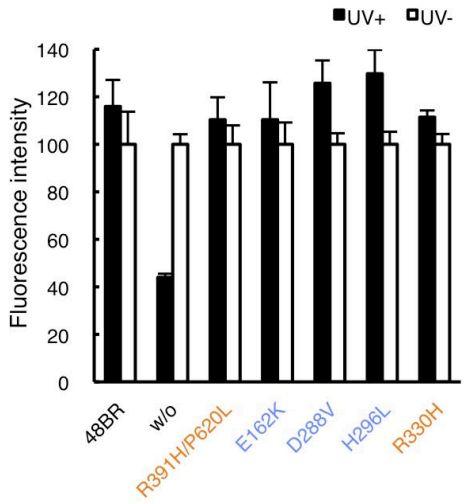
Supplementary Figure 2 *UVSSA* expression is reduced in XP70TO cells. (**a**-**c**) Quantitative-PCR (qPCR) amplification curves of the *UVSSA* (**a**) and the control *HPRT1* (**b**) mRNA transcript levels in 48BR (normal, red), Kps3 (UV^{S}S-A / p.Lys123*, blue), and XP70TO (UV^{S}S-A / p.Cys32Arg, green) cells. The relative transcript levels analysed by ΔΔCT method is shown in (**c**). Note that the *UVSSA* transcript level is reduced in XP70TO cells as well as in Kps3 cells because of nonsense mediated mRNA decay (NMD); in XP70TO cells, the p.Cys32Arg change may affect the splicing efficiency of the *UVSSA* intron 4 as the causal base change locates only 4bp-upstream of the exon-intron boundary (see **Fig. 1c**). (**d**) The immunoblots show weak affinity of the rabbit anti-UVSSA antibody to the Cys32Arg mutant protein. Kps3 cells were mock treated or infected with lentiviruses that express either wild type or Cys32Arg-mutated V5-tagged *UVSSA* cDNA. HEK293T cells were mock treated or transfected with plasmids that express either wild type, Cys32Arg-mutated, or VHS-domain truncated (T4 in **Fig. 3b**) *UVSSA* cDNA (V5-tagged). Ectopically expressed UVSSA proteins (note that the endogenous UVSSA protein is not visible because of its low level expression) were detected by the rabbit anti-UVSSA antibody or a mouse anti-V5-tag antibody. ERCC3 (XPB) and GTF2H1 (p62) are loading controls. In the top panel, V5-tagged UVSSA protein is less expressed in the cells infected with the wild type *UVSSA* compared with the Cys32Arg infectant (41% by image quantification), although the bands intensity are nearly the same (Cys32Arg is 84% of the wild type) when probed with the anti-UVSSA antibody. A similar tendency is seen in the bottom panel.



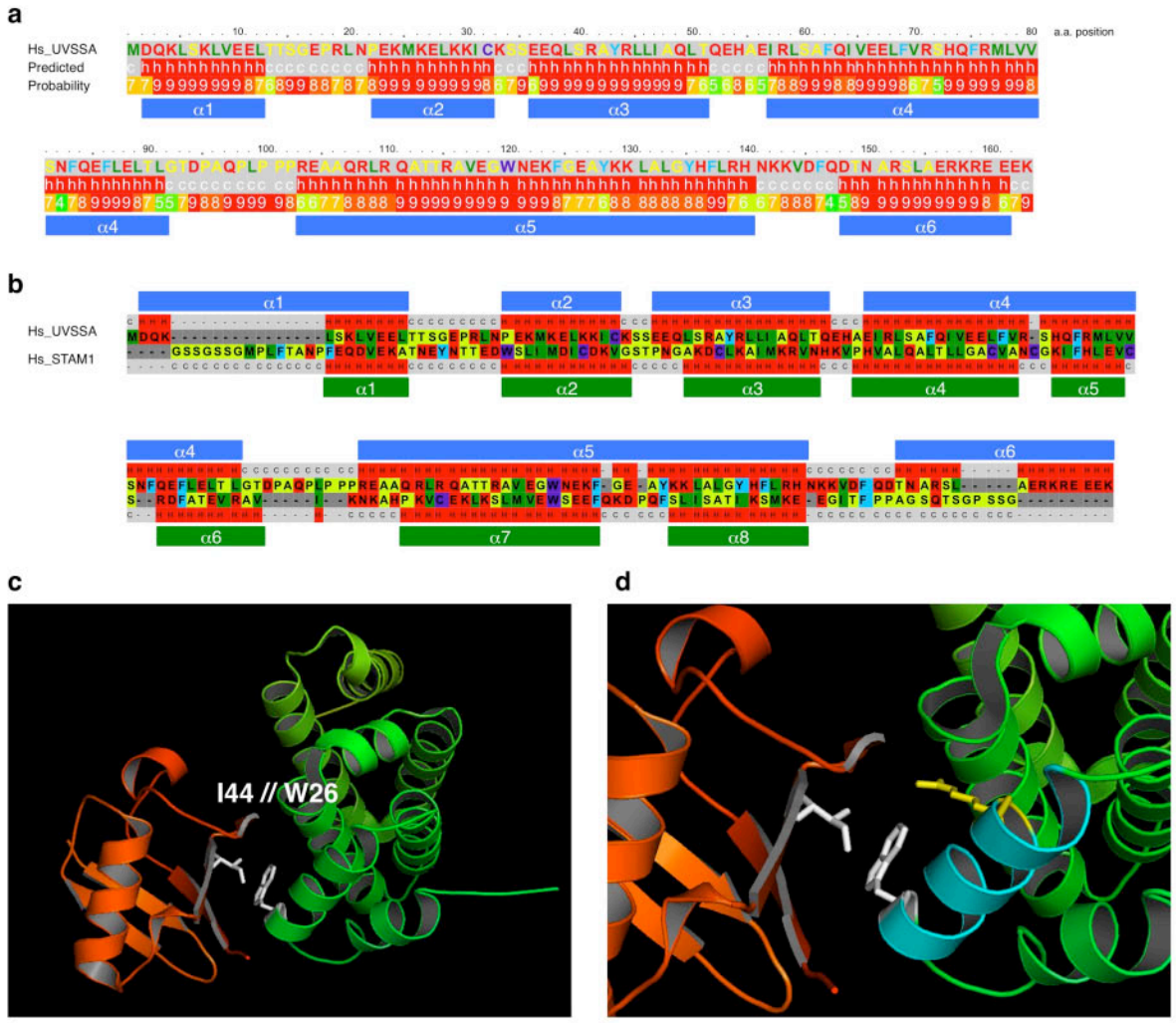
Supplementary Figure 3 Overlapping run of homozygosity (ROH) in the four UV^{S}S-A patients Kps2, Kps3, UV^{S}S24TA, and XP24KO. (**a**) The heterozygous c.367A>T single nucleotide variant was identified in a Japanese control individual (ID:0379). The altered lysine 123 residue is shown in red (*). Normal, 48BR. (**b**) ROHs identified in the individual UV^{S}S-A patients. Gray bars on each chromosome indicate ROHs identified by genome-wide SNP-genotyping. Bars with thick line on chromosome 4 are magnified in (**c**). (**c**) Magnified view of the shared ROHs on chromosome 4, which encompasses the *UVSSA* gene locus. ROHs identified in the individual UV^{S}S-A patients are simplified and shown in different colors. *UVSSA* gene loci are shown as white boxes. SNP-genotyping details are shown in the bottom. AA (homozygous); BB (homozygous); AB (heterozygous); NC (no call). (**d**) SNP 6.0 array results for chromosome 4 in the four UV^{S}S-A patients. Black dots represent chromosome copy number state (0.0 ~ 4.0), and the red line represents *UVSSA* gene locus. No copy number change was detected in the *UVSSA* gene locus in the patients.



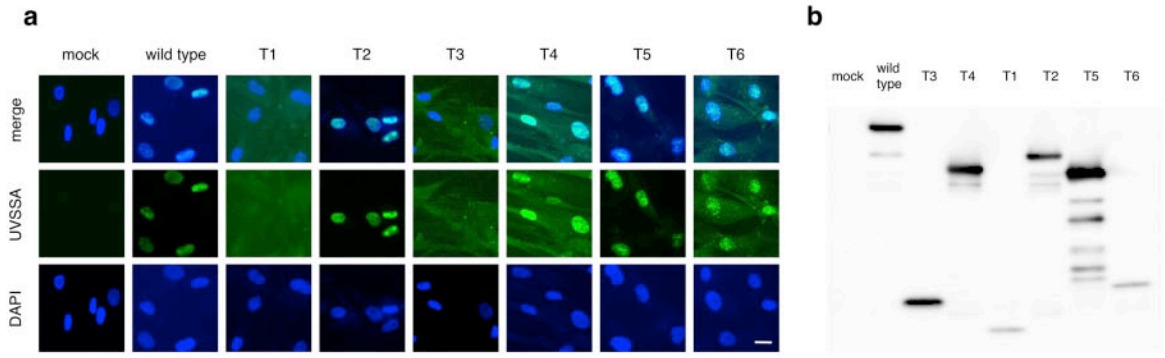
Supplementary Figure 4 Abrogation of *UVSSA* gene expression by individual siRNA oligos diminishes RRS levels. 48BR cells were mock siRNA transfected, or transfected with three individual siRNA oligos targeting the *UVSSA* gene, mixture of the three oligos, or siRNA targeting the *XPA* gene. RRS levels were measured as described in **Online Methods**; closed bars, 10J/m^2 UVC; open bars, no UV.



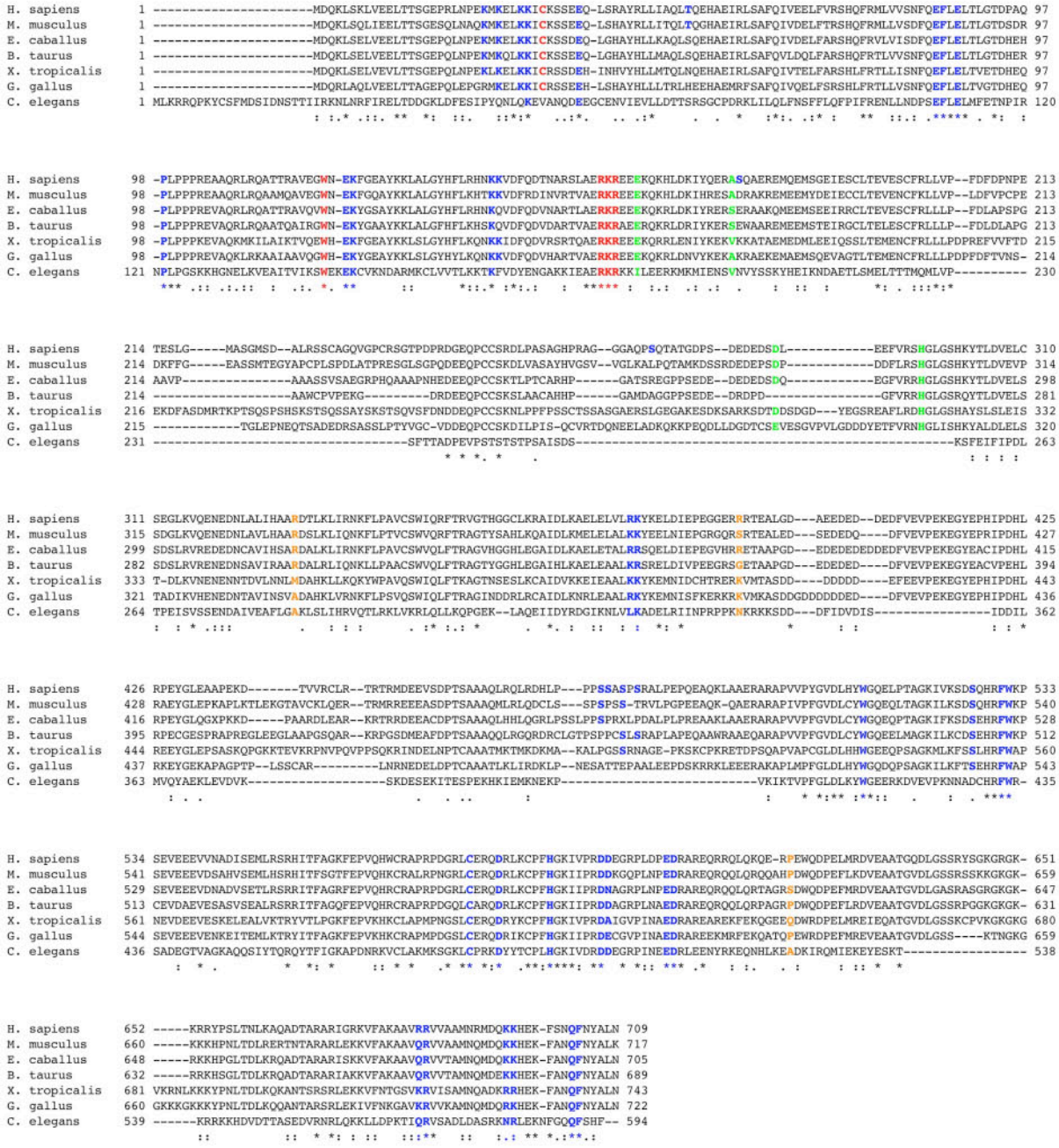
Supplementary Figure 5 UVSSA amino acid substitutions identified in CS-patients, and normal controls do not affect the RRS activity. Lentiviruses expressing amino acid substitution mutants of the V5-tagged *UVSSA* cDNA were infected in Kps3 cells. 48 h after infection, cells were UV irradiated (closed bars, 10J/m^2 UVC; open bars, no UV), followed by 12h incubation for recovery. RRS levels were measured as described in **Online Methods**. Heterozygous mutations identified in CS patients are shown in blue (see details in **Supplementary Note** and **Supplementary Table 4**); nonsynonymous SNPs found in normal control individuals are shown in orange.



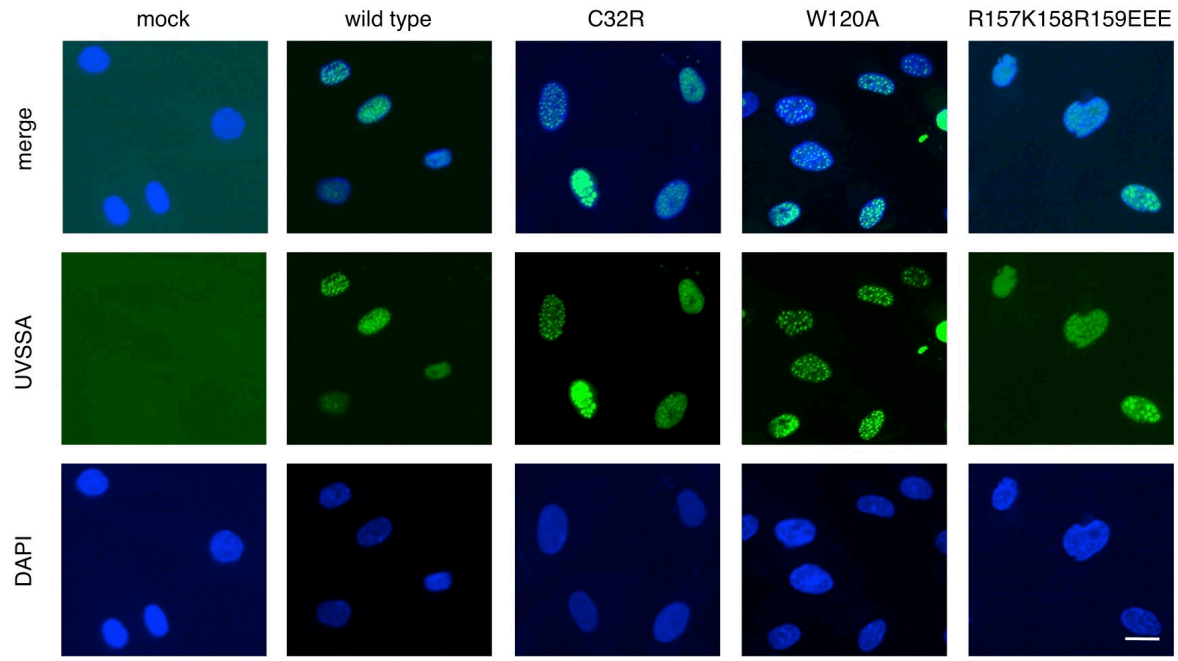
Supplementary Figure 6 3D structure prediction identified a structural similarity between the UVSSA protein N-terminal region and the VHS-domain. (**a**) Secondary structure of the UVSSA protein (NP_065945, 709 amino acids) was predicted by the PHYRE web-server program[15]. α-helix (h, in red, and blue boxes) and coil (c, in gray) structures, and the prediction probabilities (0, low to 9, high) are shown. (**b**) The program predicted a motif of 143-163 amino acids in the N-terminus, which had a 3D-structural similarity with 6 VHS-domain motifs with high probabilities (>95%). 1x5b is the PDB entry for the VHS-domain of human STAM2 protein, which has eight α-helices with 163 amino acids (green boxes). (**c**) 3D structure of the STAM1 VHS-domain (green) / ubiquitin (red) complex (PDB: 3LDZ)[16]. Ubiquitin-Ile44 and STAM1-Trp26 residues, both located at the periphery of the binding interface, are shown as white sticks. (**d**) Magnified view of the binding interface in (**c**). The XP70TO causative Cys32Arg mutation (STAM1 corresponding residue, Cys33, is mutated to arginine) is superimposed onto the structure (shown as a yellow stick). α2-helix is highlighted in blue.



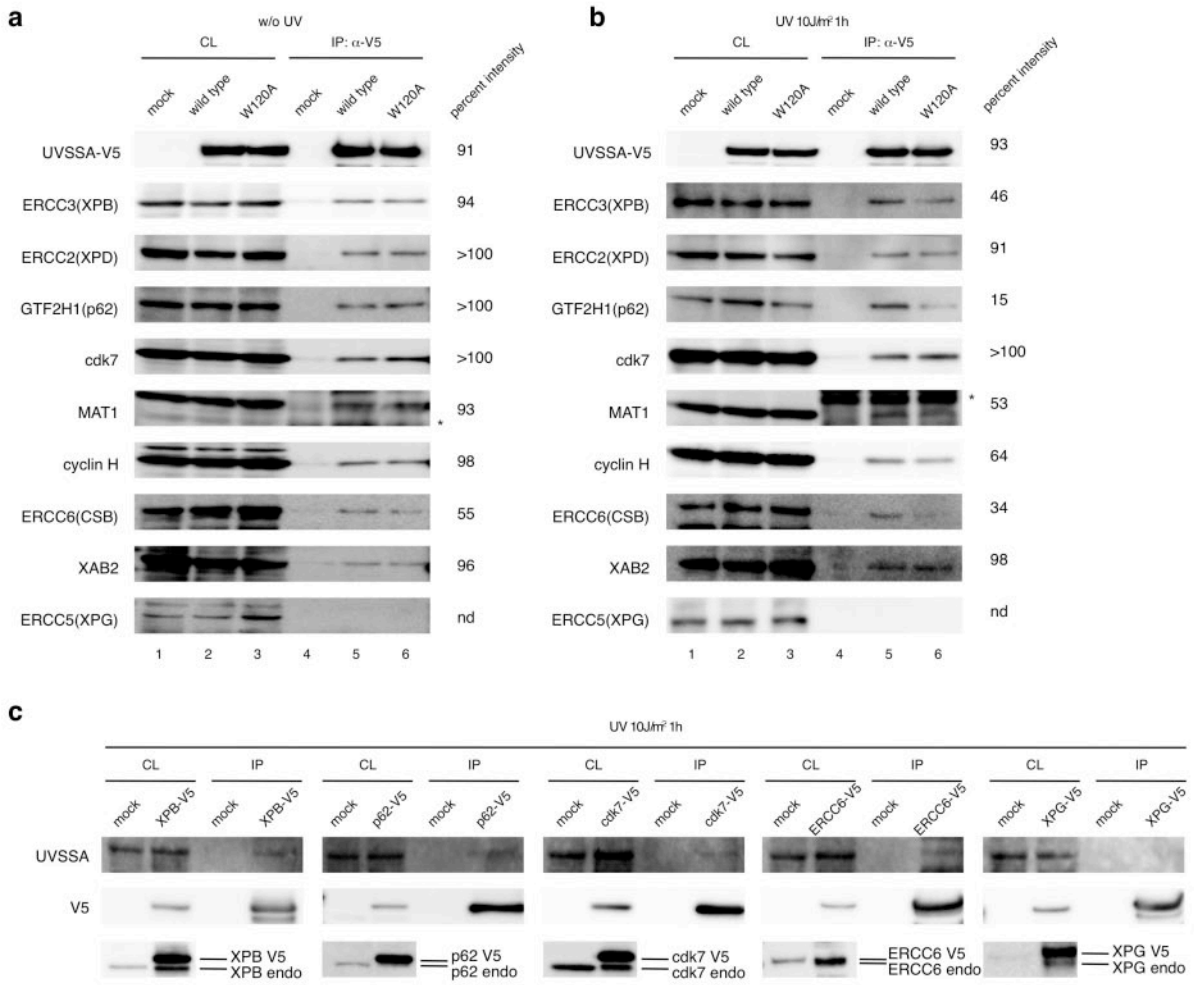
Supplementary Figure 7 Ectopic expression of the wild type and UVSSA truncation mutants in UV^{S}S-A cells. (**a**) Kps3 cells were mock-infected or infected with a lentivirus that expresses V5-tagged- wild type or UVSSA truncation mutants (T1-T6 in **Fig. 3b**). 48 h after infection, the cells were fixed and immunostained with mouse anti-V5 antibody (green). DAPI, DAPI stain (blue); merge, merged picture. Scale bar, 20 μm. (**b**) The expression of T1 and T3 truncation mutants are confirmed by immunoblot. Ectopically-expressed wild type UVSSA and truncation mutant proteins were detected by the mouse anti-V5 antibody. Minor bands are degradation products.



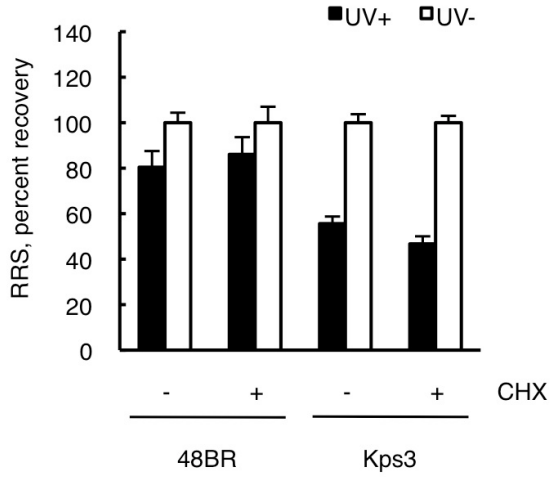
Supplementary Figure 8 Multiple sequence alignment of the UVSSA orthologues. 60 UVSSA protein (human, NP_065945) homologues were identified in a BLASTP search against the NCBI non-redundant protein database (E value <1). The following seven proteins were aligned with CLUSTAL W program. *H. sapiens*, NP_065945; *M. musculus*, NP_001074570; *E. caballus*, XP_001488394; *B. taurus*, XP_587703; *X. tropicalis*, NP_001107306; *G. gallus*, XP_420845; *C. elegans*, NP_505012. Asterisks indicate identical residues, whereas strongly or weakly conserved residues are indicated by colons and periods, respectively. Amino acids shown in bold colors are designed mutants (red, RRS deficient phenotype; blue, no phenotype) or reported SNPs and novel amino acid substitutions identified in CS-patients (green) and normal controls (orange).



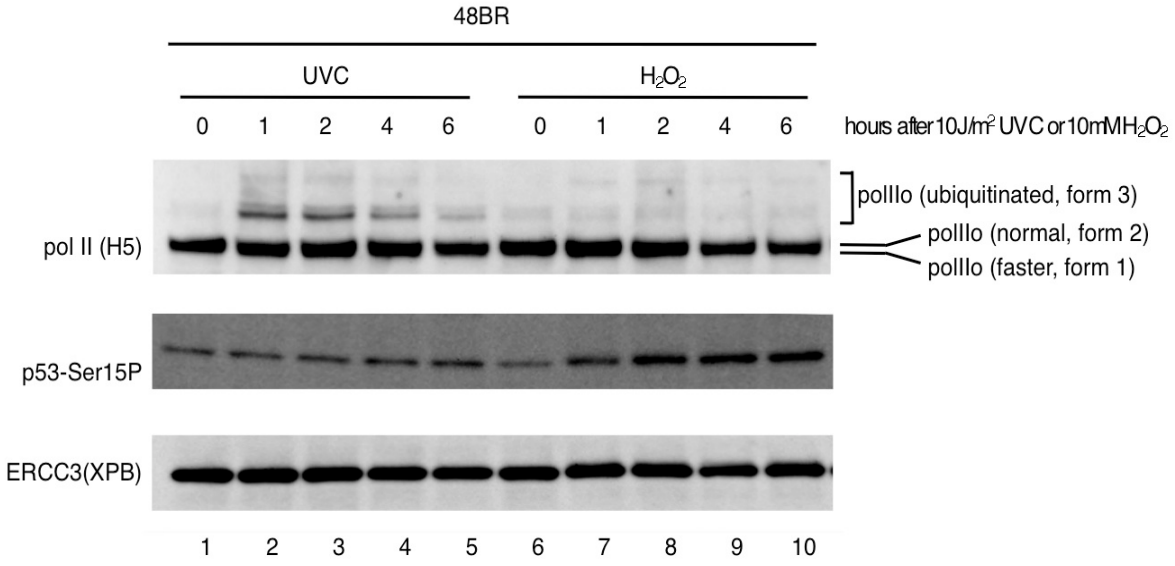
Supplementary Figure 9 Ectopic expression of the wild type and amino acid substitution UVSSA mutant proteins in UV^{S}S-A cells. Kps3 cells were mock-infected or infected with a lentivirus that expresses wild type or indicated amino acid substitution mutants of the V5-tagged *UVSSA* cDNA. 48 h after infection, the cells were fixed and immunostained with mouse anti-V5 antibody (green). DAPI, DAPI stain (blue); merge, merged picture. Scale bar, 20 μm.



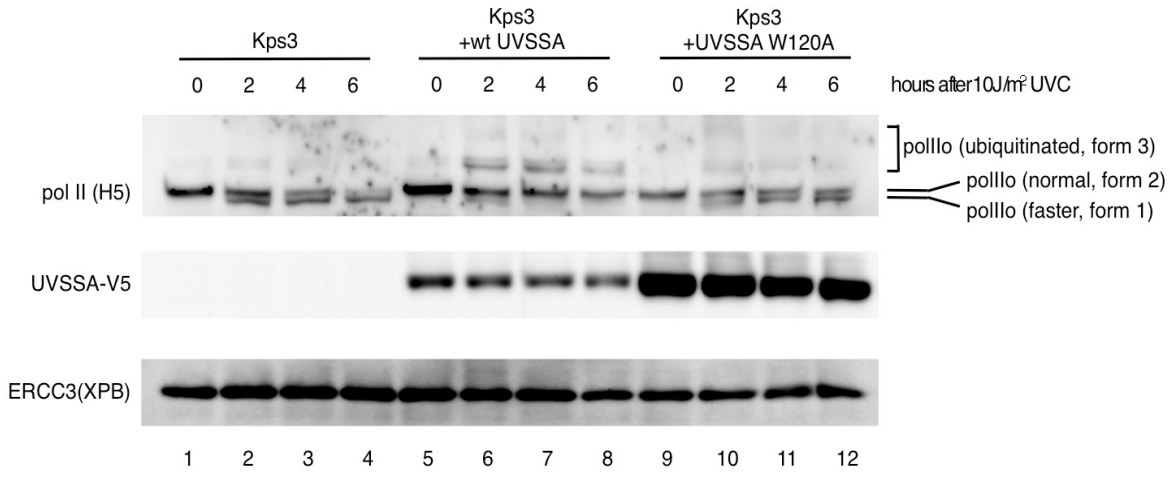
Supplementary Figure 10 The N-terminal VHS-domain is essential for the interaction of UVSSA with TFIIH and ERCC6 (CSB) upon UV-irradiation. (**a-c**) The wild type and the VHS-domain mutated UVSSA protein interactions with the core-TFIIH and CAK sub-complexes were assayed (**a**) without or (**b**, **c**) 1 h after 10J/m^2 UV-irradiation. Crude lysates were prepared from HEK293T cells either mock transfected, or transfected with (**a**, **b**) the wild type or the Trp120Ala VHS-domain mutated *UVSSA* cDNA (V5-tagged), or (**c**) cDNAs encoding the indicated NER factors (V5-tagged). The extracts were immunoprecipitated and detected as described in the main text with antibodies against: (**a**,**b**) V5-tag (mouse), core-TFIIH (ERCC3/XPB, ERCC2/XPD, and GTF2H1/p62), CAK (CDK7, cyclin H, and MAT1) components, and ERCC6 (CSB), XAB2, and ERCC5 (XPG); (**c**) UVSSA (mouse), V5-tag (mouse), and the indicated NER factors (XPB, p62, CDK7, ERCC6, and XPG). CL, crude lysate (33% load); IP, immunoprecipitate. The intensities of the bands corresponding to factors binding to the mutant proteins are expressed as percentages of those of the wild type protein. Asterisks indicate non-specific bands.



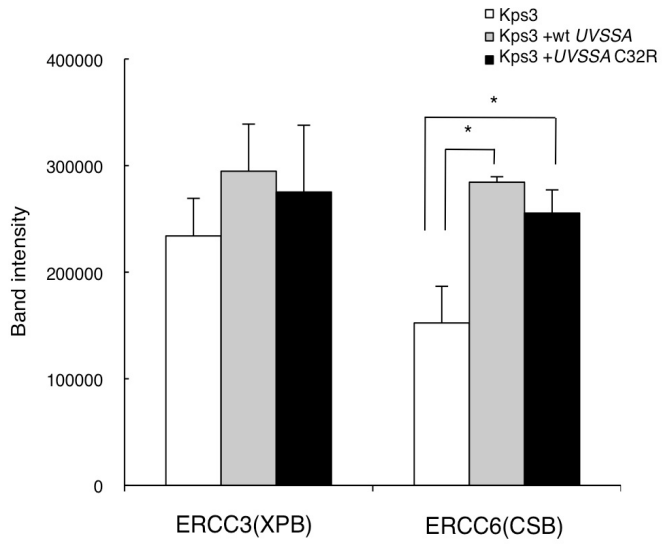
Supplementary Figure 11 Cycloheximide treatment does not alter RNA synthesis recovery after UV-irradiation. Normal (48BR), and UV^{S}S-A (Kps3) cells were pre-incubated in media containing 100μM cycloheximide (CHX) for 1 h, followed by UV-irradiation (closed bars, 10J/m^2 254nm UVC; open bars, no UV). Cells were then incubated for 4 h in media containing 100μM CHX for RNA synthesis recovery. RRS levels were measured as described in **Online Methods** (2h EU labeling was performed in the presence of 100μM CHX). Note that total RNA synthesis levels were 1/3 of that of without CHX.



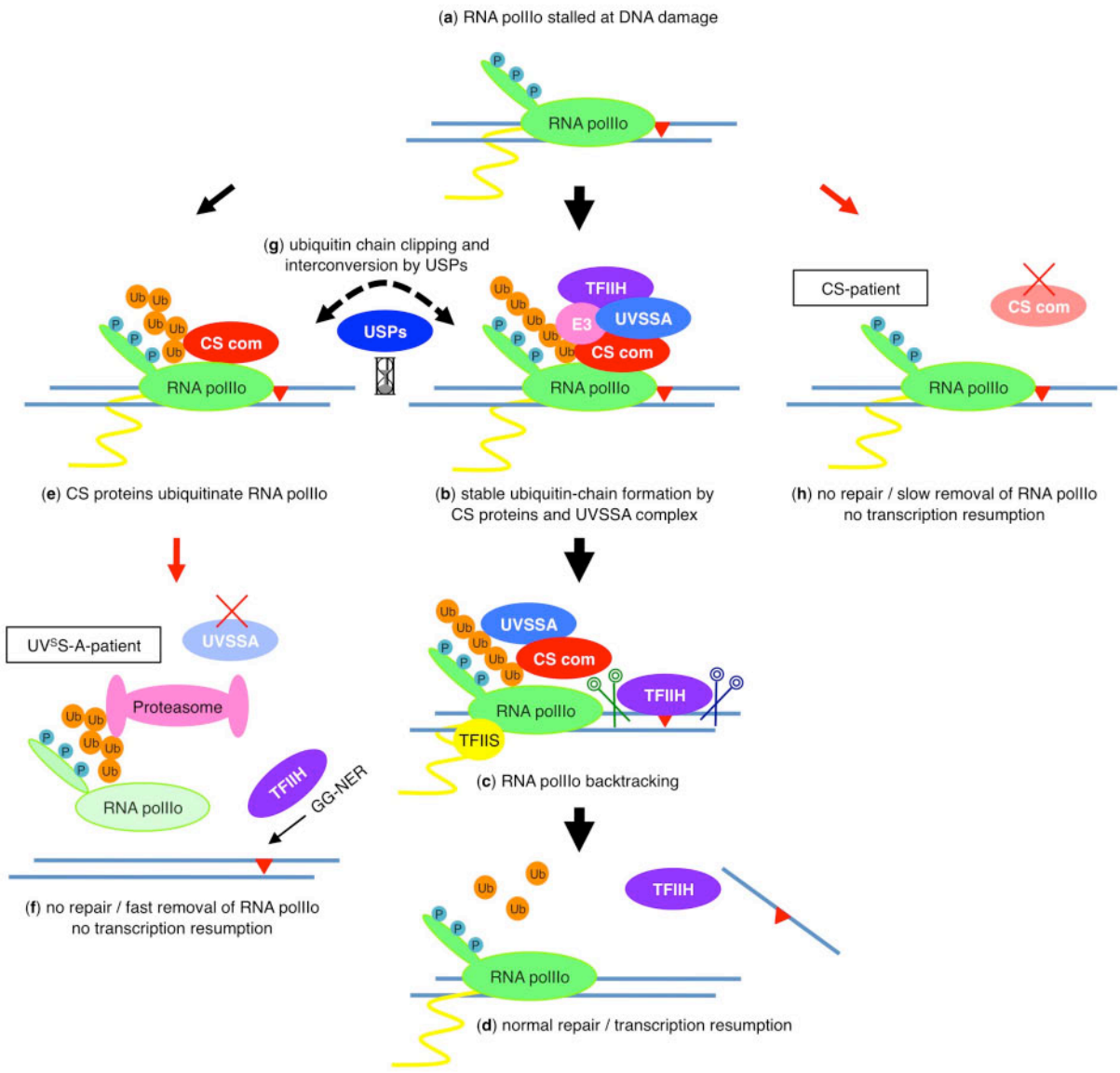
Supplementary Figure 12 Hydrogen peroxide treatment does not induce RNA polIIo modification in normal cells. 48BR (normal) cells were pre-incubated in media containing 100μM cycloheximide (CHX) for 1 h, followed by 10J/m^2 UVC irradiation or 30min treatment with 10mM H_2O_2. Cells were then incubated for indicated time periods in media containing CHX. RNA polIIo modification was assayed as described in **Online Methods**. DNA damage-induced p53 Ser15 phosphorylation was observed after H_2O_2 treatment.



Supplementary Figure 13 The VHS-domain Trp120Ala mutation fails to complement the RNA polIIo processing alterations in UV^{S}S-A cells. Lack of RNA polIIo ubiquitination in Kps3 cells was not rescued by expression of the VHS-domain Trp120Ala mutant. Lentiviruses that express wild type or Trp120Ala-mutated V5-tagged *UVSSA* cDNA were infected in Kps3 cells. 48 h after infection, RNA polIIo ubiquitination was assayed as described in **Online Methods**.



Supplementary Figure 14 The reduced expression of ERCC6 (CSB) protein in the UV^{S}S-A cells was restored by expression of either wild-type- or VHS-domain mutated- (Cys32Arg) UVSSA proteins. The ERCC6 and ERCC3 (XPB) protein expression levels in the indicated Kps3 derivatives (6 h after UV irradiation) were quantified from three sets of immunoblots including **Fig. 4d**, and displayed as a bar chart (n=3; error bars represent standard deviations). The protein expression level differences between the derivatives were examined for statistical significance with one-way ANOVA ($p<0.01$) followed by Turkey's post hoc test. Asterisks with brackets indicate statistically significant differences between the groups (P values < 0.01) in the post test.



Supplementary Figure 15 UVSSA facilitates stable ubiquitination of RNA polIIo and promotes backtracking, a working model. (**a**) RNA polIIo stalled at UV-DNA damage. (**b**) RNA polIIo is then ubiquitinated with a K63-linkage in a process dependent on both the CS-proteins and UVSSA. (**c**) We hypothesise that this allows backtracking of RNA polIIo mediated by TFIIS to take place to allow access to the NER machinery. (**d**) After removal of DNA damage, transcription resumes. (**e**) In a minor pathway RNA polIIo can be ubiquitinated in a K48-linkage, also dependent on the CS-proteins, but not on UVSSA. (**f**) In UV^{S}S-A patients, this becomes the major pathway and leads to rapid proteasomal degradation of RNA polIIo. (**g**) Alternatively, deubiquitinases (USPs) may permit the two ubiquitinated forms of RNA polIIo depicted to be interconverted, as indicated by the dashed arrow, providing a "ubiquitin timer", as suggested by Nouspikel [17]. (**h**) In CS-patients neither of these pathways are operative and the RNA polIIo remains blocked at the lesion. CS com, the ERCC8 (CSA) and ERCC6 (CSB) protein complexes (CS-complexes).

SUPPLEMENTARY NOTE

Identification of the UV^{S}S-A causative gene by exome sequencing: We performed whole exome sequencing of two unrelated Japanese UV^{S}S-A patients, Kps3 and XP24KO, using the Agilent SureSelect Exome Target Enrichment System, followed by paired-end sequencing on the Illumina GAIIx sequencer. We obtained ~69 and 70 million sequencing reads for Kps3 and XP24KO, respectively (about 5.2 Giga base pairs per sample). Of the initial sequencing read-outs, 4.5 Gbp for Kps3 and 4.6 Gbp for XP24KO are uniquely mapped to the human reference sequence (hg19). Of the mapped sequences, 2.2 Gbp (47.56%) for Kps3 and 2.1 Gbp (44.77%) for XP24KO are further mapped to the targeted bases with a mean coverage of 55.64x (Kps3) and 53.28x (XP24KO). In the Kps3 sequencing, 96.8% of the target bases were read more than 2 times coverage, and the mean target coverage was 55.64-fold. In XP24KO, 96.5% of the target bases were read more than 2 times coverage, and the mean target coverage was 53.28-fold (**Supplementary Table 2a**). We identified a total of 75,368 single nucleotide variants (SNVs) and 15,966 insertions and deletions (indels) in Kps3, and 94,628 SNVs and 19,665 indels in XP24KO (**Supplementary Table 2b**). To identify potential pathogenetic changes, we focused on 'functionally significant variants', which are non-synonymous missense variants, nonsense-stopgains and -stoplosses, splice-acceptor and -donor site variants, and short coding indels; we identified 7,488 and 7,211 functionally significant variants in Kps3 and XP24KO, respectively (**Supplementary Table 2b**). The identified variants were then filtered out to extract 'novel functionally significant variants': we compared the variants with dbSNP131, 1000 Genomes Project full phase 1 data (SNVs and indels from 629 individuals), and 7 in-house Japanese exome sequencing data; identical variants found in the datasets were removed (**Table 1**). As a consequence of these filtering strategies, 263 and 243 novel functionally significant variants were determined in Kps3 and XP24KO, respectively. Under the assumption of a recessive inheritance model, we found 18 candidate genes in Kps3 and 23 in XP24KO (**Table 1**, **Supplementary Table 2c**). We considered potential causative genes shared among the patients, and we identified a single candidate gene, *KIAA1530* (NM_020894). We later named this gene *UVSSA*.

Cellular response to UV-irradiation and genetic analysis of the repair defect in UV^{S}S24TA cells: The response to UV-light was analyzed in fibroblasts by measuring UDS, RRS and cell survival. Procedures for the evaluation of these cellular parameters are routinely used in the laboratory of MS and have all been described previously[11-13]. Briefly, UDS was

determined by counting the number of grains on at least 50 non-S-phase cells in autoradiographic preparations of cultures incubated for 3 h after UV-irradiation in medium containing ^{3}H-thymidine (^{3}H-TdR, Amersham, specific activity 25 Ci/mmol)[11]. The recovery of RNA synthesis rate was analysed by measuring the radioactivity incorporated in cells labelled, 24 h after UV-irradiation, for 1 h with ^{3}H-uridine[12]. Cell survival was analyzed by measuring the viability in stationary phase fibroblasts[13]. Complementation analysis was performed by measuring the recovery of RNA synthesis in hybrids obtained by fusing patient's cells with CS reference strains, as previously described[14]. Briefly, fibroblast strains used as partners in the fusion were grown for three days in medium containing latex beads of different sizes (0.8 μm and 1.7 μm) that were incorporated into the cytoplasm as a marker. The cells were fused using polyethylene glycol (PEG-4000, Merck), incubated for 48 h at 37°C, UV-irradiated (20 J/m^2), incubated again for a further 24 h at 37°C, labelled for 1 h with ^{3}H-Urd, and then processed for autoradiography. One sample was treated in parallel but without UV-irradiation. In irradiated preparations, the grains over nuclei in 25 homodikaryons (identified as binuclear cells containing beads of one size) and in 25 heterodikaryons (identified as binuclear cells containing beads of different sizes) were counted. Two cell strains were classified in the same complementation group if the heterodikaryons failed to recover normal RNA synthesis levels after UV-irradiation.

CS-patients sequencing: Recent reports identified mutations in the *ERCC8* and *ERCC6* (also known as *CSA* and *CSB*, respectively) genes as responsible for UV^{S}S[8,10]. Two UV^{S}S cases (UV^{S}1KO, CS3AM) carried the same homozygous termination mutation in the *ERCC6* gene (c.229C>T), which causes a severe truncation near the N-terminus of the 140 kDa ERCC6 (CSB) protein (p.Arg77X), indicating that individuals with null *ERCC6* mutations can have the features of UV^{S}S, even though other mutations in the *ERCC6* gene can cause the much more severe features of CS[8,18,19]. Another case (UV^{S}S1VI) carried a homozygous mutation in the *ERCC8* gene (c.1083G>T), causing a missense change close to the C-terminus of the 40kDa ERCC8 (CSA) protein (p.Trp361Cys)[10]. From these findings, we considered the possibility that deleterious amino acid substitutions, or truncation mutations occurring in the middle or the C-terminal part of UVSSA protein, may result in diseases with more severe clinical manifestations than UV^{S}S. To evaluate this possibility, we sequenced the coding exons of the *UVSSA* gene of 61 genetically unassigned CS-patients with defective RRS (**Supplementary Table 4**). We observed neither homozygous nor compound heterozygous

novel amino acid changes in the *UVSSA* gene in any of these patients (**Supplementary Table 4**). We did however detect four novel heterozygous amino acid changes in UVSSA in 4 of the 61 CS-patients. These changes as well as the SNPs, R391H and P620L, and a novel change, R330H, found in control and UV^{S}S-A individuals, were analyzed for their effect on RRS after UV irradiation; all these amino acid changes were able to restore normal RRS levels following expression in UV^{S}S-A cells (**Supplementary Fig. 5**). Taken together, these data suggest that it is unlikely that mutations in the *UVSSA* gene can result in the clinical features of CS, implying that its function is distinct from but overlaps with that of *ERCC8/ERCC6*. In our screen of cells from 61 CS-patients, we found 6 patients with mutation(s) in the *ERCC8* gene, and 44 patients with mutation(s) in the *ERCC6* gene. Importantly, we found thirteen patients who do not have any causal mutations in the *ERCC8*, *ERCC6*, or *UVSSA* genes. This implies that additional genes involved in TC-NER remain to be discovered. Further exome analysis is ongoing and will determine the disease-gene(s) for these unassigned CS-patients.

SUPPLEMENTARY REFERENCES

1. Itoh, T., Ono, T. & Yamaizumi, M. A new UV-sensitive syndrome not belonging to any complementation groups of xeroderma pigmentosum or Cockayne syndrome: siblings showing biochemical characteristics of Cockayne syndrome without typical clinical manifestations. *Mutat Res* **314**, 233-48 (1994).
2. Itoh, T., Fujiwara, Y., Ono, T. & Yamaizumi, M. UVs syndrome, a new general category of photosensitive disorder with defective DNA repair, is distinct from xeroderma pigmentosum variant and rodent complementation group I. *Am J Hum Genet* **56**, 1267-76 (1995).
3. Itoh, T. *et al.* Clinical characteristics of three patients with UVs syndrome, a photosensitive disorder with defective DNA repair. *Br J Dermatol* **134**, 1147-50 (1996).
4. Fujiwara, Y., Uehara, Y., Ichihashi, M., Yamamoto, Y. & Nishioka, K. Assignment of 2 patients with xeroderma pigmentosum to complementation group E. *Mutat Res* **145**, 55-61 (1985).
5. Itoh, T., Linn, S., Ono, T. & Yamaizumi, M. Reinvestigation of the classification of five cell strains of xeroderma pigmentosum group E with reclassification of three of them. *J Invest Dermatol* **114**, 1022-9 (2000).
6. Kawada, A., Satoh, Y. & Fujiwara, Y. Xeroderma pigmentosum complementation group E: a case report. *Photodermatol* **3**, 233-8 (1986).
7. Fujiwara, Y., Ichihashi, M., Kano, Y., Goto, K. & Shimizu, K. A new human photosensitive subject with a defect in the recovery of DNA synthesis after ultraviolet-light irradiation. *J Invest Dermatol* **77**, 256-63 (1981).

8. Horibata, K. *et al.* Complete absence of Cockayne syndrome group B gene product gives rise to UV-sensitive syndrome but not Cockayne syndrome. *Proc Natl Acad Sci U S A* **101**, 15410-5 (2004).
9. Miyauchi-Hashimoto, H., Akaeda, T., Maihara, T., Ikenaga, M. & Horio, T. Cockayne syndrome without typical clinical manifestations including neurologic abnormalities. *J Am Acad Dermatol* **39**, 565-70 (1998).
10. Nardo, T. *et al.* A UV-sensitive syndrome patient with a specific CSA mutation reveals separable roles for CSA in response to UV and oxidative DNA damage. *Proc Natl Acad Sci U S A* **106**, 6209-14 (2009).
11. Stefanini, M. *et al.* DNA repair investigations in nine Italian patients affected by trichothiodystrophy. *Mutat Res* **273**, 119-25 (1992).
12. Mayne, L.V. & Lehmann, A.R. 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**, 1473-8 (1982).
13. Stefanini, M. *et al.* Genetic heterogeneity of the excision repair defect associated with trichothiodystrophy. *Carcinogenesis* **14**, 1101-5 (1993).
14. Stefanini, M., Fawcett, H., Botta, E., Nardo, T. & Lehmann, A.R. Genetic analysis of twenty-two patients with Cockayne syndrome. *Hum Genet* **97**, 418-23 (1996).
15. Kelley, L.A. & Sternberg, M.J. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**, 363-71 (2009).
16. Ren, X. & Hurley, J.H. VHS domains of ESCRT-0 cooperate in high-avidity binding to polyubiquitinated cargo. *EMBO J* **29**, 1045-54 (2010).
17. Nouspikel, T. Multiple roles of ubiquitination in the control of nucleotide excision repair. *Mech Ageing Dev* **132**, 355-65 (2011).
18. Spivak, G. UV-sensitive syndrome. *Mutat Res* **577**, 162-9 (2005).
19. Spivak, G. The many faces of Cockayne syndrome. *Proc Natl Acad Sci U S A* **101**, 15273-4 (2004).