Table S4. Analysis of the Splice Site Predictions of the Two Mutations

		ESEfinder3.0 (score)	NetGene2 (confidence)	HSF 2.4.1[a] (value)	SpliceView (score)	BDGP[a] (score)
c.378+1G>A	reference	11.9514	0.67	96.91	92	0.99
	mutation	<6.67 (under threshold)	under threshold	70.07	under threshold	<0.40 (under cutoff)
	assessment	***abolished***	***abolished***	***site broken***	***abolished***	***abolished***
c.664+1G>A	reference	9.8861	0.75	87.83	81	1.00
	mutation	<6.67 (under threshold)	under threshold	61	under threshold	<0.40 (under cutoff)
	assessment	***abolished***	***abolished***	***site broken***	***abolished***	***abolished***

[a]Human Splicing Finder Version 2.4.1

[b]Berkeley Drosophila Genome Project

nature genetics

Mutations in *UVSSA* cause UV-sensitive syndrome and impair RNA polymerase IIo processing in transcription-coupled nucleotide-excision repair

Yuka Nakazawa[1–3,16], Kensaku Sasaki[4,16], Norisato Mitsutake[1,3,16], Michiko Matsuse[1,3], Mayuko Shimada[1–3], Tiziana Nardo[5], Yoshito Takahashi[6], Kaname Ohyama[1,7], Kosei Ito[1,8], Hiroyuki Mishima[4], Masayo Nomura[4], Akira Kinoshita[1,4], Shinji Ono[4], Katsuya Takenaka[9], Ritsuko Masuyama[8], Takashi Kudo[10], Hanoch Slor[11], Atsushi Utani[12], Satoshi Tateishi[13], Shunichi Yamashita[3,14], Miria Stefanini[5], Alan R Lehmann[15], Koh-ichiro Yoshiura[4] & Tomoo Ogi[1–3]

UV-sensitive syndrome (UV^{S}S) is a genodermatosis characterized by cutaneous photosensitivity without skin carcinoma[1–4]. Despite mild clinical features, cells from individuals with UV^{S}S, like Cockayne syndrome cells, are very UV sensitive and are deficient in transcription-coupled nucleotide-excision repair (TC-NER)[2,4,5], which removes DNA damage in actively transcribed genes[6]. Three of the seven known UV^{S}S cases carry mutations in the Cockayne syndrome genes *ERCC8* or *ERCC6* (also known as *CSA* and *CSB*, respectively)[7,8]. The remaining four individuals with UV^{S}S, one of whom is described for the first time here, formed a separate UV^{S}S-A complementation group[1,9,10]; however, the responsible gene was unknown. Using exome sequencing[11], we determine that mutations in the *UVSSA* gene (formerly known as *KIAA1530*) cause UV^{S}S-A. The UVSSA protein interacts with TC-NER machinery and stabilizes the ERCC6 complex; it also facilitates ubiquitination of RNA polymerase IIo stalled at DNA damage sites. Our findings provide mechanistic insights into the processing of stalled RNA polymerase and explain the different clinical features across these TC-NER–deficient disorders.

We performed exome sequencing on two cell lines, Kps3 and XP24KO, derived from two individuals with UV^{S}S-A (cell lines described in **Supplementary Table 1**; exome described in **Table 1**, Online Methods, **Supplementary Table 2a–c** and **Supplementary Note**). Using a recessive model of inheritance, we directly identified overlapping mutations in *KIAA1530* (NCBI Gene 57654), a predicted gene at 4p16.3 encoding a 709 amino acid protein of unknown function (**Table 1** and **Supplementary Table 2c**). *KIAA1530* was subsequently renamed *UVSSA* (encoding UV-stimulated scaffold protein A) because of this finding, with support from the Human Gene Nomenclature Committee (HGNC). Affected individuals were homozygous for a c.367A>T mutation, which led to the introduction of a premature stop codon, p.Lys123*, in the UVSSA protein (**Fig. 1a,b**). We identified the same homozygous mutation in subject Kps2 (a sibling of Kps3) and a homozygous c.87delG mutation causing a p.Ile31Phefs*9 frameshift alteration in Israeli subject UV^{S}S24TA (**Fig. 1b,c**, **Supplementary Fig. 1** and **Supplementary Note**). The identified mutations are summarized (**Fig. 1d**). We did not detect the 80-kDa UVSSA protein in any of the individuals with UV^{S}S-A (**Fig. 1e**). We also examined several mild xeroderma pigmentosum cases; in one case, XP70TO[12] (**Supplementary Table 1**), we identified a homozygous missense mutation (encoding p.Cys32Arg) in *UVSSA* (**Fig. 1c,d**), which implies that XP70TO is also in the UV^{S}S-A complementation group. The mutant protein was stably expressed in XP70TO cells, although band intensity was faint (**Fig. 1f** and **Supplementary Fig. 2a–d**).

The mutated allele encoding p.Lys123* was observed in the heterozygous state in 1 of 576 control individuals (allele frequency of 0.09%) evaluated by direct sequencing or high-resolution melting

[1]Nagasaki University Research Centre for Genomic Instability and Carcinogenesis (NRGIC), Nagasaki, Japan. [2]Department of Molecular Medicine, Atomic Bomb Disease Institute, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan. [3]Department of Radiation Medical Sciences, Atomic Bomb Disease Institute, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan. [4]Department of Human Genetics, Atomic Bomb Disease Institute, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan. [5]Istituto di Genetica Molecolare, Consiglio Nazionale delle Ricerche, Pavia, Italy. [6]Innovative Beauty Science Laboratory, Kanebo Cosmetics Inc., Odawara, Japan. [7]Department of Environmental and Pharmaceutical Sciences, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan. [8]Department of Cell Biology, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan. [9]Department of Molecular Genetics, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan. [10]Department of Radioisotope Medicine, Atomic Bomb Disease Institute, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan. [11]Human Molecular Genetics and Biochemistry, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel. [12]Department of Dermatology, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan. [13]Institute of Molecular Embryology and Genetics, Kumamoto University, Kumamoto, Japan. [14]Fukushima Medical University, Fukushima, Japan. [15]Genome Damage and Stability Centre, University of Sussex, Brighton, UK. [16]These authors contributed equally to this work. Correspondence should be addressed to T.O. (togi@nagasaki-u.ac.jp).

Received 26 September 2011; accepted 29 February 2012; published online 1 April 2012; doi:10.1038/ng.2229

Table 1 Direct identification of the UV^{S}S-A causal gene by exome sequencing of subjects Kps3 and XP24KO

Filter	Kps3	XP24KO
Single-nucleotide variants (SNVs)		
Total SNVs identified	75,368	94,628
Nonsynonymous missense, stop-gain, stop-loss, or splice-site[a] variants	7,161	6,904
Not reported in dbSNP131	383	340
Not reported in dbSNP131 or 1000 GP[b]	231	210
Not reported in dbSNP131, 1000 GP[b] or 7 control exomes[c] (new functionally significant SNVs)	217	202
Insertion and/or deletions (indels)		
Total indels identified	15,966	19,665
Frameshift indels, in-frame stop-gain, in-frame stop-loss or splice-site[a] variants	327	307
Not reported in dbSNP131	198	172
Not reported in dbSNP131 or 1000 GP[b]	85	76
Not reported in dbSNP131, 1000 GP[b] or 7 control exomes[c] (new functionally significant indels)	46	41
New functionally significant variants		
New SNVs or indels	263	243
Candidate genes under a recessive model		
Homozygous	9	17
Compound heterozygous	9	6
Potential causative genes shared among the subjects[d]	1 (*UVSSA*)	

The number of variants or genes that meet the criteria is indicated for each subject.
[a]Splice-site acceptor or donor variants within 2 bp of exon-intron boundaries. [b]1000 Genomes Project full Phase 1 data, November 2010. [c]Seven Japanese in-house exome sequences. [d]Under a recessive model, only one candidate gene, *UVSSA*, was identified.

analysis (HRMA) (**Supplementary Fig. 3a**); the mutations encoding p.Ile31Phefs*9 and p.Cys32Arg were not detected. Haploinsufficiency for *UVSSA* was negligible, as the parents of Kps2 and Kps3 showed no symptoms[4]. In parallel with exome sequencing, we performed whole-genome SNP genotyping to identify runs of homozygosity (ROHs) shared among the affected individuals. We identified three overlapping ROHs (of >1 Mb) on autosomes, one of which encompassed the *UVSSA* locus (**Fig. 1g**, **Supplementary Fig. 3b,c** and

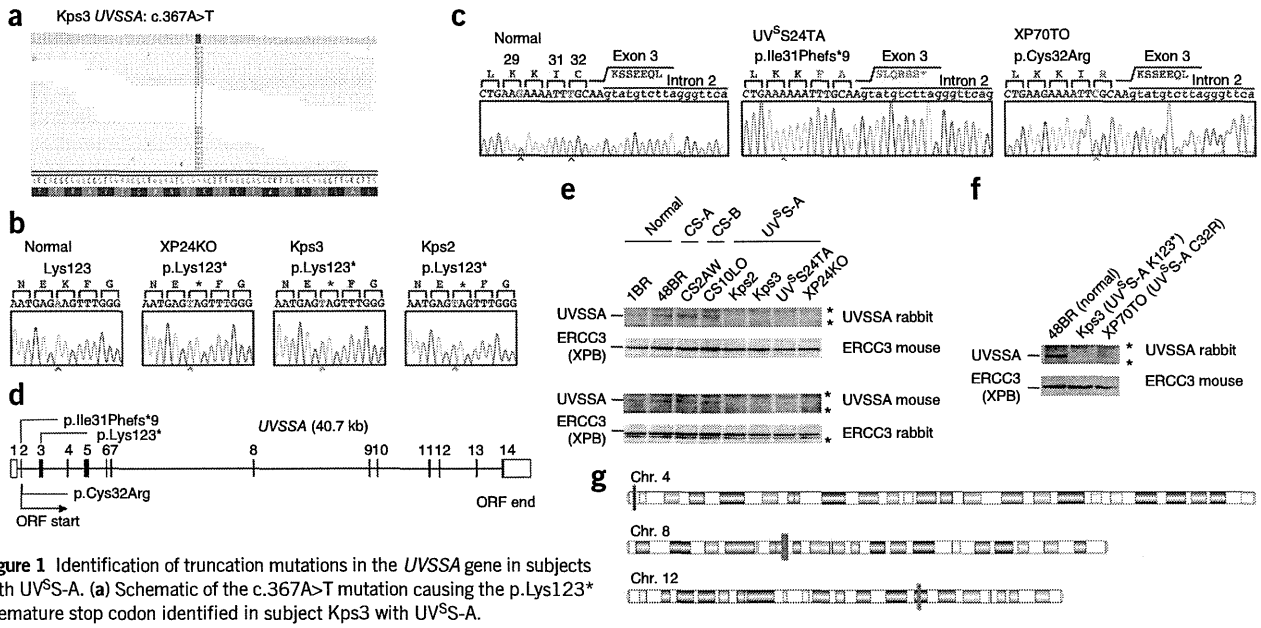

Figure 1 Identification of truncation mutations in the *UVSSA* gene in subjects with UV^{S}S-A. (**a**) Schematic of the c.367A>T mutation causing the p.Lys123* premature stop codon identified in subject Kps3 with UV^{S}S-A. Exome-sequencing data were aligned, and individual reads are shown as beige lines; the position of the A-to-T substitution is in red. (**b**) Capillary Sanger sequencing showed that subjects XP24KO and Kps3 with UV^{S}S-A are homozygous for the c.367A>T SNV in *UVSSA* exon 3; the same mutation was also found in Kps2, the sibling of Kps3. The altered amino acid, Lys 123, is in red. (**c**) The c.87delG frameshift mutation and the p.Cys32Arg missense mutation were identified in subjects UV^{S}S24TA and XP70TO with UV^{S}S-A, respectively. Electropherograms of *UVSSA* exon 2 (uppercase) and intron 2 (lowercase) are shown. The deletion in the third position of the codon for Lys29 causes the p.Ile31Phefs*9 premature stop codon; the transition in the first position of the codon for Cys32 causes the p.Cys32Arg change (in red). Arrowheads indicate the position of the mutations. (**d**) The genomic structure of the *UVSSA* gene and the positions of the alterations in the subjects with UV^{S}S-A. (**e**,**f**) Immunoblots showing the 80-kDa *UVSSA* gene product in normal and Cockayne syndrome cells, which is missing in the UV^{S}S-A cells analyzed (**e**). UVSSA protein expression is reduced in XP70TO cells (**f**). ERCC3 is a loading control. 48BR and 1BR, normal cells; CS2AW, CS-A cells; CS1OLO, CS-B cells; Kps2, Kps3, UV^{S}S24TA, XP24KO and XP70TO, UV^{S}S-A cells. Asterisks indicate nonspecific bands. (**g**) Shared ROH segments between subjects XP24KO, Kps3, Kps2 and UV^{S}S24TA with UV^{S}S-A. ROHs overlapping in all four subjects are shown as blue bars, and the overlapping region that contains the *UVSSA* gene is shown as a red bar (see **Supplementary Fig. 3b** for ROHs identified in the individual subjects and **Supplementary Fig. 3c** for a magnified view of the overlapped region found in chromosome 4).

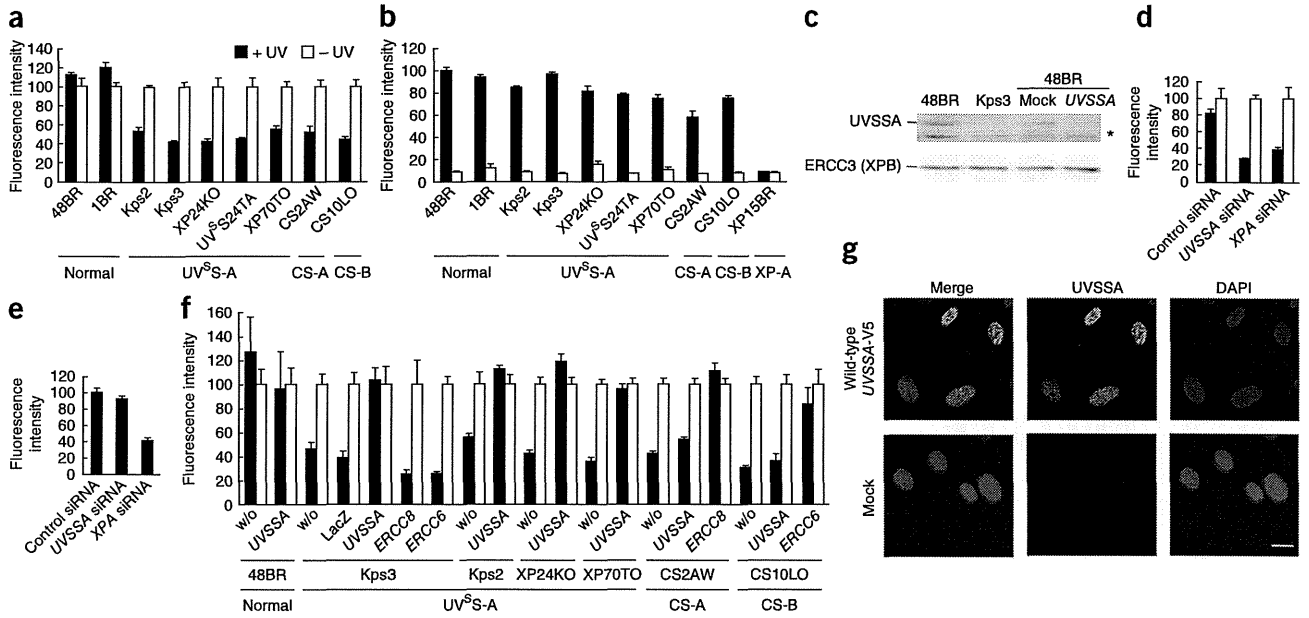

Figure 2 RNA synthesis recovery after UV irradiation requires *UVSSA* gene expression. (**a**) RRS activity after UV irradiation was reduced in the UV^{S}S-A cells (filled bars, 10 J/m^2 of 254 nm UV-C; open bars, no UV). (**b**) Close-to-normal UDS activity in the UV^{S}S-A cells (filled bars, 20 J/m^2 of UV-C; open bars, no UV). (**c**) *UVSSA* gene expression was diminished by siRNA. Normal 48BR cells were mock transfected or transfected with siRNA targeting *UVSSA*. Knockdown efficiency was detected by immunoblotting of the UVSSA protein, with ERCC3 as a loading control. The asterisk indicates a nonspecific band. (**d**) RRS activity after UV irradiation was completely abolished by abrogation of *UVSSA* gene expression. Normal 48BR cells were mock transfected or transfected with siRNA targeting *UVSSA* or *XPA* (filled bars, 10 J/m^2 of UV-C; open bars, no UV). (**e**) UDS activity was unaffected by siRNA knockdown of *UVSSA* (filled bars, 20 J/m^2 of UV-C). (**f**,**g**) Ectopic expression of wild-type *UVSSA* cDNA in UV^{S}S-A cells by recombinant lentivirus infection completely rescued the RRS deficiency (filled bars, 10 J/m^2 of UV-C; open bars, no UV) (**f**). Viral infection efficiency (>85%) was confirmed by immunofluorescent staining of V5-tagged UVSSA protein (**g**). Scale bar, 20 μm. In **a**,**d**,**f**, RRS was normalized to activity in non-irradiated cells. UDS activity were normalized to that of normal 48BR cells in **b** or to that in cells mock transfected with siRNA in **e**. Error bars, s.d. of medians of nuclear fluorescence measurements in quintuplicate samples in **a**,**b**,**d**–**f**. w/o, without.

Supplementary Table 3a,**b**). No chromosome copy-number variation was detected (**Supplementary Fig. 3d**).

The findings strongly suggest that the mutations in *UVSSA* in the subjects with UV^{S}S-A are causal for the disease; therefore, we next examined NER activities in the UV^{S}S-A cells (**Fig. 2**). Unscheduled DNA synthesis (UDS[13]; defective in xeroderma pigmentosum) was nearly normal; however, RNA synthesis recovery (RRS[14]; defective in UV^{S}S and in Cockayne syndrome) was diminished in all cell lines with mutated *UVSSA* (**Fig. 2a**,**b**; UDS and RRS were measured using a recently developed rapid, nonradioactive system[15,16]). Similarly, small interfering RNA (siRNA)-based depletion of *UVSSA* transcripts (**Fig. 2c**) caused a marked reduction in RRS (**Fig. 2d** and **Supplementary Fig. 4**), whereas UDS was unaffected (**Fig. 2e**). Ectopic expression of wild-type *UVSSA* cDNA in UV^{S}S-A cells restored normal RRS (**Fig. 2f**; V5-tagged UVSSA immunofluorescent staining shown in **Fig. 2g**), whereas it did not affect RRS in normal, CS-A or CS-B cells; neither *ERCC8* and *ERCC6* cDNA expression in UV^{S}S-A cells restored RRS. We conclude that *UVSSA* is the causal gene for UV^{S}S-A.

Mutations in the *ERCC8* and *ERCC6* genes underlie both Cockayne syndrome and UV^{S}S[7,8]. To evaluate whether *UVSSA* mutations may also result in Cockayne syndrome phenotypes, we sequenced the *UVSSA* gene in 61 individuals with Cockayne syndrome whose genetic cause had not yet been determined (**Supplementary Table 4**). We found no obvious mutations, except for four new heterozygous changes. These changes, as well as the SNPs that were also found in control individuals and in persons with UV^{S}S-A, did not affect RRS activity (**Supplementary Fig. 5**). These data suggest that *UVSSA* function is distinct from that of genes involved in Cockayne syndrome (**Supplementary Note**).

Amino-acid sequences encoded by human *UVSSA* and its orthologs have no obvious similarity to other protein families. A domain of unknown function, DUF2043 (EMBL-EBI IPR018610) is located near the C terminus (**Fig. 3a**). We performed three-dimensional structure prediction using the PHYRE web server[17] and identified a motif of 143–163 amino acids near the N terminus, which had substantial homology with the Vps-27, Hrs and STAM (VHS) domain[18] (**Fig. 3a** and **Supplementary Fig. 6a**,**b**). VHS domain proteins have recently been implicated in ubiquitin binding[19]. A crystallographic study determined that Trp26, located in the α2 helix of the VHS domain of STAM1, interfaces with Ile44 of ubiquitin[20] (**Supplementary Fig. 6c**). Cys32 in UVSSA is located in the corresponding α2 helix (**Supplementary Fig. 6b**), and the p.Cys32Arg mutation can be superimposed onto the structure of the STAM1-ubiquitin complex, indicating that the Arg32 residue in the mutant UVSSA protein is oriented in the same direction as Trp26 in STAM1 (**Supplementary Fig. 6d**). These findings suggest that the p.Cys32Arg change in UVSSA that occurs in subject XP70TO may obstruct interactions between the VHS domain and ubiquitinated proteins.

To investigate the importance of the DUF2043 and VHS domains in TC-NER, we transduced *UVSSA* truncation mutants into UV^{S}S-A cells and assayed complementation of the RRS defect (**Fig. 3b**). None of the truncation mutants lacking either the VHS or DUF2043 domain was able to restore RRS activity in these cells (**Fig. 3c**, **Supplementary Fig. 7a** (note that mutants T1 and T3 were unstable) and **Supplementary Fig. 7b**).

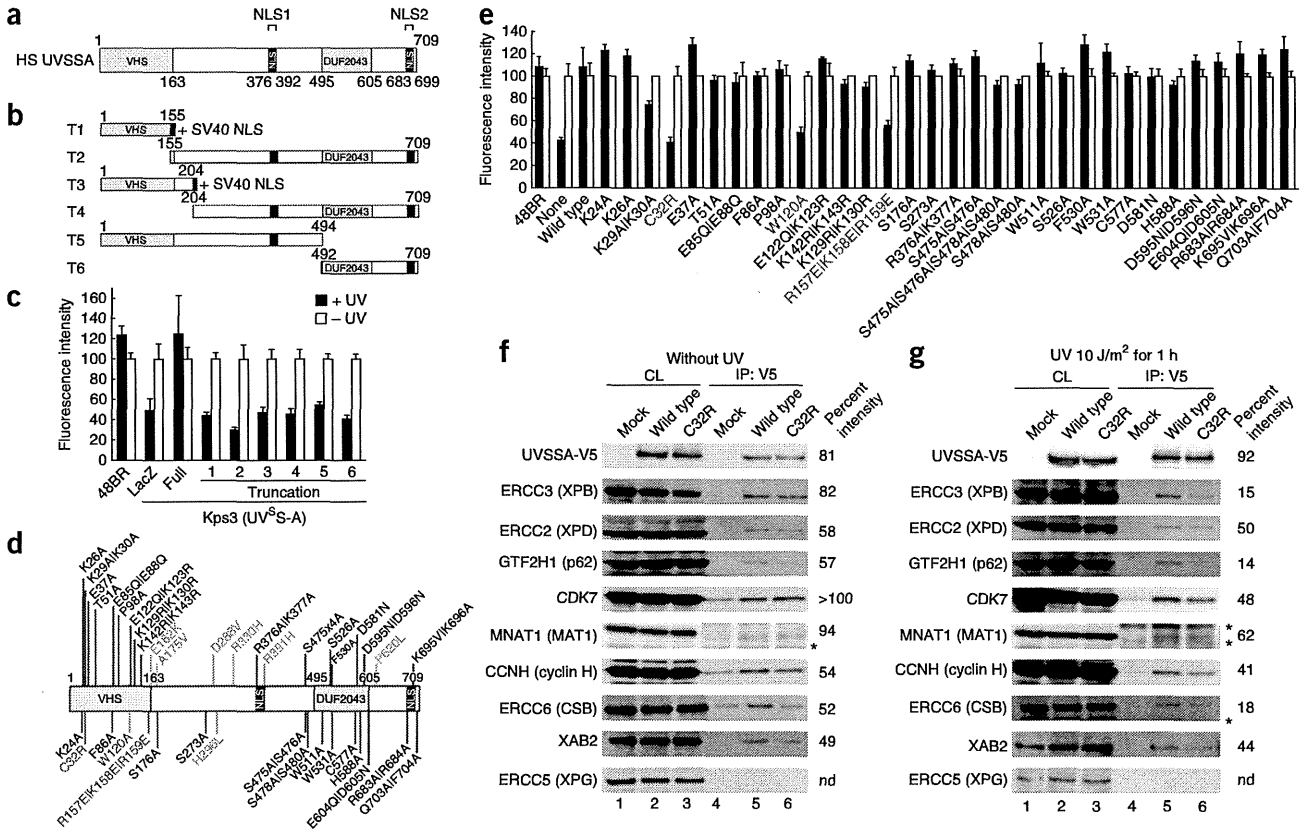

Figure 3 The N-terminal VHS domain of the UVSSA protein is essential for RRS activity and TFIIH interaction. (**a**,**b**) Schematic of human full-length UVSSA (**a**) and UVSSA truncation mutants (**b**) used for the RRS assays. NLS1 and NLS2, putative nuclear-localization signals; T1–T6, truncation mutants. In T1 and T3, the SV40 NLS was added to the C terminus of the mutants. (**c**) RRS was measured in Kps3 cells expressing an intact (full) or truncated (T1–T6) form of V5-tagged UVSSA in **b** (filled bars, 10 J/m^2 of UV-C; open bars, no UV). (**d**) Schematic of the UVSSA mutants that were used for the RRS assays. Heterozygous alterations identified in individuals with Cockayne syndrome are shown in blue (see details in **Supplementary Table 4**); nonsynonymous SNPs found in normal, control individuals are shown in orange; changes shown in black and red are the amino-acid substitutions at residues conserved from Nematoda to humans as well as the XP70TO disease-causing mutation p.Cys32Arg. The three mutants shown in red did not complement the RRS deficiency of Kps3 cells. (**e**) RRS was measured in Kps3 cells expressing UVSSA mutants in **d** (filled bars, 10 J/m^2 of UV-C; open bars, no UV) (see also **Supplementary Fig. 5**). Expression of wild-type or mutant UVSSA in Kps3 cells was confirmed by immunofluorescent staining of V5-tagged protein (transfection efficiency was >85%; **Supplementary Figs. 7a,b** and **9**). In **c**,**e**, RRS activity was normalized to that in non-irradiated cells. Error bars, s.d. of medians of nuclear fluorescence measurements in quintuplicate samples. (**f**,**g**) Wild-type and UVSSA C32R protein interactions with NER factors were assayed by immunoprecipitation without irradiation (**f**). The same interactions were assayed 1 h after 10 J/m^2 of UV irradiation (**g**). Interactions were detected by immunoblotting with antibodies to the V5 tag (UVSSA), core TFIIH factors (ERCC3, ERCC2, GTF2H1), CAK components (CDK7, CCNH, MNAT1) and ERCC6, XAB2 and ERCC5. CL, crude lysate (33% load); IP, immunoprecipitate. The intensities of the bands corresponding to factors binding to the mutant proteins are expressed as the percentages of that of the wild-type protein. Asterisks indicate nonspecific bands.

Next, we mutated residues conserved from Nematoda to humans (**Fig. 3d** and **Supplementary Fig. 8**). We found that, out of 32 mutants, only 3 carrying amino-acid changes to the 5 conserved residues located in the VHS domain (p.Cys32Arg, p.Trp120Ala and p.[Arg157Glu; Lys158Glu; Arg159Glu]) failed to restore normal RRS (**Fig. 3e** and **Supplementary Fig. 9**). Although we could not determine the critical amino-acid residues responsible for DUF2043 function, overall these findings indicate that the ubiquitin-binding VHS domain and the DUF2043 domain of UVSSA are crucial for TC-NER activity.

We next examined the association of UVSSA with TC-NER factors by immunoprecipitation. UVSSA interacted with several subunits of the core general transcription factor IIH (TFIIH) subcomplex (ERCC3 (also called XPB or p89), ERCC2 (also called XPD) and GTF2H1 (also called p62)) and the cyclin-dependent kinase (CDK)-activating kinase (CAK) subcomplex (CDK7, CCNH (also called cyclin H) and MNAT1 (also called MAT1)), as well as with ERCC6 and the XPA-binding protein XAB2 (**Fig. 3f**). No robust interaction was observed between UVSSA and ERCC5 (also known as XPG). Similar interactions were also observed after UV irradiation (**Fig. 3g**). UVSSA is thus a new factor associated with the TC-NER complexes. The absence of ERCC5 protein in the TFIIH-UVSSA complex seems to conflict with a report showing that ERCC5 is important for stabilizing TFIIH[21]; this discrepancy could partly be explained by the possibility that ERCC5 and UVSSA may share the same binding interface of TFIIH.

In non-irradiated cells, UVSSA binding capabilities were unaffected by mutations in the VHS domain (**Fig. 3f**, lane 6, and **Supplementary Fig. 10a**, lane 6). However, following UV irradiation, its interactions with some TFIIH subunits, namely GTF2H1 and ERCC3, and with ERCC6 were substantially weakened by amino-acid substitutions in the VHS domain (**Fig. 3g**, lane 6, and **Supplementary Fig. 10b**, lane 6). These interactions were confirmed by reverse immunoprecipitation (**Supplementary Fig. 10c**). These data indicate that UVSSA interacts transiently with the NER machinery,

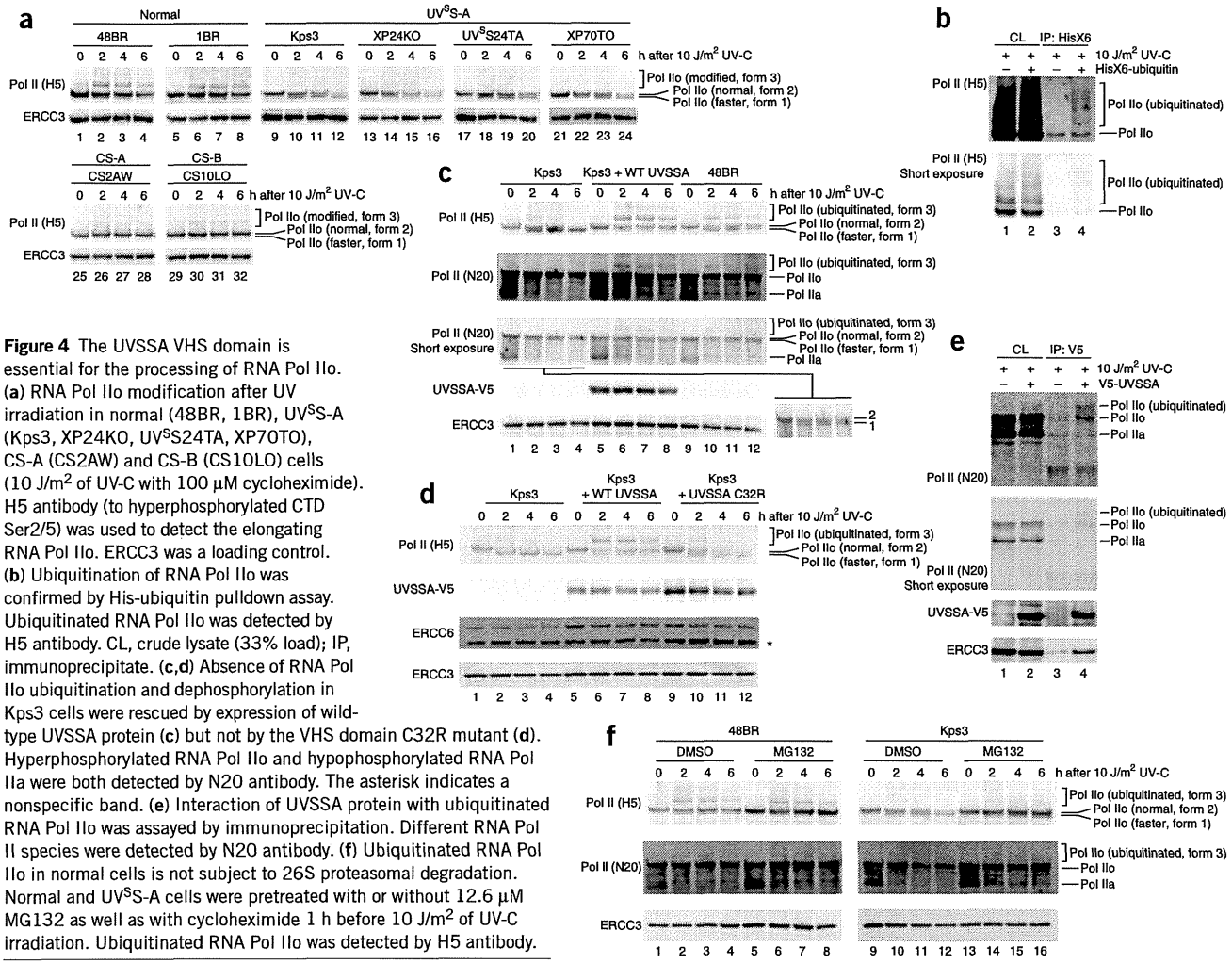

Figure 4 The UVSSA VHS domain is essential for the processing of RNA Pol IIo. (**a**) RNA Pol IIo modification after UV irradiation in normal (48BR, 1BR), UV^{S}S-A (Kps3, XP24KO, UV^{S}S24TA, XP70TO), CS-A (CS2AW) and CS-B (CS10LO) cells (10 J/m^2 of UV-C with 100 μM cycloheximide). H5 antibody (to hyperphosphorylated CTD Ser2/5) was used to detect the elongating RNA Pol IIo. ERCC3 was a loading control. (**b**) Ubiquitination of RNA Pol IIo was confirmed by His-ubiquitin pulldown assay. Ubiquitinated RNA Pol IIo was detected by H5 antibody. CL, crude lysate (33% load); IP, immunoprecipitate. (**c**,**d**) Absence of RNA Pol IIo ubiquitination and dephosphorylation in Kps3 cells were rescued by expression of wild-type UVSSA protein (**c**) but not by the VHS domain C32R mutant (**d**). Hyperphosphorylated RNA Pol IIo and hypophosphorylated RNA Pol IIa were both detected by N20 antibody. The asterisk indicates a nonspecific band. (**e**) Interaction of UVSSA protein with ubiquitinated RNA Pol IIo was assayed by immunoprecipitation. Different RNA Pol II species were detected by N20 antibody. (**f**) Ubiquitinated RNA Pol IIo in normal cells is not subject to 26S proteasomal degradation. Normal and UV^{S}S-A cells were pretreated with or without 12.6 μM MG132 as well as with cycloheximide 1 h before 10 J/m^2 of UV-C irradiation. Ubiquitinated RNA Pol IIo was detected by H5 antibody.

and the VHS domain specifically supports direct contact with the TFIIH core complex and ERCC6 after UV irradiation.

In TC-NER, stalled RNA polymerase II (RNA Pol II) first has to be displaced by backtracking or degradation to allow access of the NER machinery[6]. ERCC8 and ERCC6 complexes (also called CS protein complexes; defective in Cockayne syndrome) are essential to this process[22]. During this step of TC-NER, some of the elongating form of RNA Pol II (RNA Pol IIo), which is phosphorylated at Ser2 and Ser5 in the C-terminal domain (CTD), is ubiquitinated[23], although the precise function of this ubiquitination is not known. We examined RNA Pol IIo modifications after UV irradiation in the presence of cycloheximide in normal, Cockayne syndrome and UV^{S}S-A cells (**Fig. 4a**). (TC-NER activity was unaffected by cycloheximide in the experimental conditions used (**Supplementary Fig. 11**)). In normal cells (**Fig. 4a**, lanes 1–8), a slow-migrating fraction of RNA Pol IIo was observed after UV irradiation, and this modified band (form 3) decreased in intensity over time. Form 3 was not detected in normal cells after induction of oxidative DNA damage by treatment with hydrogen peroxide (**Supplementary Fig. 12**), suggesting that the modification is a UV damage–specific event. We confirmed that the RNA Pol IIo modification was ubiquitination as reported previously[23] (**Fig. 4b**). In contrast, in the UV^{S}S cells, the ubiquitinated band was barely detectable, but a fast-migrating form of RNA Pol IIo (form 1) was observed, and the normally migrating form 2 decreased in intensity over the 6-h period after irradiation (**Fig. 4a**, lanes 9–24). In the CS2AW (CS-A) or CS10LO (CS-B) cells (**Fig. 4a**, lanes 25–32), neither ubiquitination (see also ref. 23) nor form 1 was detected. Transduction of wild-type *UVSSA* cDNA into UV^{S}S-A Kps3 cells restored the ubiquitination of RNA Pol IIo after UV irradiation (**Fig. 4c**, H5 in lanes 5–8). *UVSSA* overexpression was also associated with a substantial reduction in the levels of form 1 after UV irradiation (**Fig. 4c**, compare H5 in lanes 5–8 and 1–4, see also **Fig. 4c**, N20 (short exposure) in lanes 1–4 and its stretched image on the right (form 1–form 2 doublets are visible, regardless of the antibodies used)).

During transcription, RNA Pol IIo is eventually dephosphorylated to RNA Pol IIa to recycle RNA Pol II for another round of transcription initiation. In TC-NER, displaced RNA Pol IIo is also dephosphorylated and recycled (**Fig. 4c**, N20 in lanes 9–12). Dephosphorylation was substantially inhibited in UV^{S}S-A cells 4–6 h after UV irradiation (**Fig. 4c**, N20 in lanes 1–4), as previously reported for Cockayne syndrome cells[24], and was restored following expression of *UVSSA* cDNA (**Fig. 4c**, N20 in lanes 5–8). RNA Pol IIo ubiquitination was not restored in Kps3 cells by expression of mutants with amino-acid substitutions in the VHS domain (**Fig. 4d**, H5 in lanes 9–12, and **Supplementary Fig. 13**, H5 in lanes 9–12), indicating that the VHS domain is crucial for RNA Pol IIo processing.

We noticed that ERCC6 protein was degraded in UV^{S}S-A cells after UV irradiation, indicating that UVSSA contributes to the stabilization

of the ERCC6 complex in TC-NER (**Fig. 4d**, ERCC6 in lanes 1–4). The reduction of ERCC6 protein levels in Kps3 cells was restored following expression of either wild-type or VHS domain–mutated (p.Cys32Arg) UVSSA (**Fig. 4d**, ERCC6 in lanes 5–12, and **Supplementary Fig. 14**). These data indicate that the absence of RNA Pol IIo ubiquitination in the UV^{S}S-A cells is not a side effect of Cockayne syndrome protein depletion (**Fig. 2f**; ectopic expression of *ERCC6* did not restore RRS in UV^{S}S-A cells). We further showed that UVSSA specifically binds to RNA Pol IIo and ubiquitinated RNA Pol IIo but not to RNA Pol IIa and that ubiquitinated RNA Pol IIo is enriched in the UVSSA-binding fraction (**Fig. 4e**).

These findings suggest that UVSSA recruits an E3 ubiquitin ligase and facilitates ubiquitination of RNA Pol IIo. K48-linked polyubiquitination triggers 26S proteasomal degradation of the targeted protein, whereas monoubiquitination and K63-linked polyubiquitination contribute to functional modifications of various DNA repair factors. The BRCA1-BARD1 complex and the NEDD4 E3 ubiquitin ligase, as well as the ERCC8 complex, have been reported to be involved in damage-dependent polyubiquitination and degradation of RNA Pol IIo[25–27]. To determine whether RNA Pol IIo is degraded as a consequence of UVSSA-mediated ubiquitination, we analyzed RNA Pol IIo modifications after UV irradiation and subsequent treatment with the 26S proteasome inhibitor MG132 in normal and UV^{S}S-A cells (**Fig. 4f**). The band corresponding to ubiquitinated form 3 of RNA Pol IIo did not increase substantially in intensity in normal cells treated with MG132 (**Fig. 4f**, compare H5 in lanes 1–4 and 5–8); conversely, increases in intensity for bands corresponding to the normal form 2 of RNA Pol IIo as well as to RNA Pol IIa were observed, both in normal and UV^{S}S-A cells, in the presence of inhibitor (**Fig. 4f**, lanes 5–8 and 13–16). Taken together, these data indicate that UVSSA-dependent RNA Pol IIo ubiquitination is not subject to 26S proteasomal degradation and may therefore be K63 linked, although a substantial amount of stalled RNA Pol IIo is degraded by UVSSA-independent ubiquitination pathways. We speculate that the TFIIH core factor GTF2H2 (p44) is a candidate for an alternative E3 ubiquitin ligase[28], which may form stable K63-linked polyubiquitin chains on stalled RNA Pol IIo during TC-NER. Further studies will address this hypothesis. We conclude that UVSSA has an important role in the processing of RNA Pol II molecules stalled at sites of UV damage.

In conclusion, we have newly identified *UVSSA* as the gene causing UV-sensitive syndrome, and the UVSSA protein is involved in TC-NER of UV damage. UVSSA interacts with TFIIH, ERCC6 and RNA Pol IIo, and the VHS domain is indispensable for TC-NER activity and for ubiquitination and dephosphorylation in the processing of stalled RNA Pol IIo. We hypothesize that UVSSA directly recruits TFIIH at sites where RNA Pol IIo is stalled at UV damage (but not at oxidative damage sites; **Supplementary Fig. 12**)[5,7,29] and facilitates RNA Pol IIo ubiquitination and promotes its backtracking to allow access to the NER machinery. We also found that UVSSA contributes to stabilization of the ERCC6 complex; in accompanying reports by Zhang *et al.*[30] and Schwertman *et al.*[31], the UVSSA-USP7 complex is shown to be involved in the regulation of ERCC6 ubiquitination. These various UVSSA functions coordinate the removal of stalled RNA Pol IIo and formation of the TC-NER pre-incision complex to promote the subsequent steps of NER and transcription resumption at UV-damaged sites.

Cockayne syndrome but not UV^{S}S cells are sensitive to oxidative DNA damage[5,7,29,32]; this differential sensitivity could explain the difference between symptoms in the two syndromes. Our findings, however, suggest that aberrant RNA Pol IIo processing may also contribute to clinical outcome (**Supplementary Fig. 15**). RNA Pol IIo stalled at DNA damage is normally stably ubiquitinated and backtracked in a process dependent on ERCC8 and ERCC6 complexes and UVSSA, facilitating removal of DNA damage and prompt transcription resumption (**Supplementary Fig. 15a–d**). In UV^{S}S-A cells, stalled RNA Pol IIo can still be ubiquitinated in an alternative pathway that is dependent on ERCC8 and ERCC6 complexes and independent of UVSSA that leads to 26S proteasomal degradation of RNA Pol IIo, so that transcription resumption does not occur (**Supplementary Fig. 15e,f**). In individuals with Cockayne syndrome, neither of these pathways is operative (**Supplementary Fig. 15h**), suggesting that prolonged arrest of RNA Pol IIo at DNA lesions (leading to a signal for apoptosis)[6], as in Cockayne syndrome cells, is more deleterious than degradation that occurs in UV^{S}S-A cells, and this might contribute to the more severe clinical features in Cockayne syndrome relative to those in UV^{S}S. UVSSA is thus a key factor that controls the fate of RNA Pol IIo stalled at DNA damage.

URLs. NRGIC, http://www.nrgic.prj.nagasaki-u.ac.jp/; NCBI, http://www.ncbi.nlm.nih.gov/; EMBL-EBI nucleotide sequence database, http://www.ebi.ac.uk/embl/; Phyre2, http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index; Picard version 1.38, http://picard.sourceforge.net/index.shtml; UCSC Genome Bioinformatics, http://genome.ucsc.edu/; dbSNP Build 131, http://www.ncbi.nlm.nih.gov/projects/SNP/; 1000 Genomes Project full Phase 1 data, http://www.openbioinformatics.org/annovar/download/hg19_ALL.sites.2010_11.txt.gz; ANNOVAR, http://www.openbioinformatics.org/annovar/; SIFT dataset, http://www.openbioinformatics.org/annovar/download/hg19_avsift.txt.gz; NIPPON EGT (siRNA oligonucleotides), http://www.n-egt.com/products/sirna/specification.html; TaKaRa (qPCR primers), http://www.takara-bio.co.jp/prt/intro.htm.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We are grateful to P. Hanawalt and G. Spivak for their helpful comments on the manuscript. We are grateful to C. Hayashida, M. Kawamichi and H. Fawcett for technical assistance. This work was supported by Special Coordination Funds for Promoting Science and Technology from the Japan Science and Technology Agency (JST) (to Y.N., K.O., K.I., R.M. and T.O.), a grant-in-aid for Scientific Research KAKENHI (22710056) from the Japanese Society for the Promotion of Science, a science research grant from the Inamori Foundation, a cancer research grant from The Sagawa Foundation for Promotion of Cancer Research, a medical research grant from Mochida Memorial Funds for Medical and Pharmaceutical Research, a medical research grant from the Daiichi-Sankyo Foundation of Life Science, a medical research grant from the Takeda Science Foundation, a grant-in-aid for Seeds Innovation (Type-A) from JST (to T.O.), a Global Centers of Excellence (COE) Program from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to Y.N., K.S., N.M., M.M., M. Shimada, S.Y., K.Y. and T.O.), grants from the Ministry of Health, Labour and Welfare (to K.Y.), the Associazione Italiana per la Ricerca sul Cancro (to M. Stefanini), a Medical Research Council (MRC) programme grant and an EC-RTN and integrated project (to A.R.L.).

AUTHOR CONTRIBUTIONS

N.M. and T.O. designed the study and experiments. Y.N., N.M., M.M., M. Shimada, T.N., K.O., K.I., K.T., R.M. and T.O. performed molecular and cell biology experiments. Y.N., K.S., M. Shimada, Y.T., M.N., A.K., S.O., K.Y. and T.O. performed genetic experiments. K.S., M. Shimada, Y.T., H.M., M.N., A.K., S.O., K.Y. and T.O. analyzed the genetic data. Y.T., H.S., A.U., S.T., M. Stefanini and A.R.L. contributed Cockayne syndrome and UV^{S}S subject materials. N.M., Y.T., T.K., A.U., S.Y., M. Stefanini, A.R.L., K.Y. and T.O. coordinated the study. N.M., M. Stefanini, A.R.L. and T.O. wrote the manuscript. All authors commented on the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Spivak, G. UV-sensitive syndrome. *Mutat. Res.* **577**, 162–169 (2005).
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**, 1267–1276 (1995).
3. 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**, 256–263 (1981).
4. 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**, 233–248 (1994).
5. Spivak, G. *et al.* Ultraviolet-sensitive syndrome cells are defective in transcription-coupled repair of cyclobutane pyrimidine dimers. *DNA Repair (Amst.)* **1**, 629–643 (2002).
6. Hanawalt, P.C. & Spivak, G. Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell Biol.* **9**, 958–970 (2008).
7. 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. USA* **106**, 6209–6214 (2009).
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. USA* **101**, 15410–15415 (2004).
9. Cleaver, J.E. & Thomas, G.H. Clinical syndromes associated with DNA repair deficiency and enhanced sun sensitivity. *Arch. Dermatol.* **129**, 348–350 (1993).
10. 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**, 1022–1029 (2000).
11. Ng, S.B. *et al.* Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**, 272–276 (2009).
12. Kawada, A., Satoh, Y. & Fujiwara, Y. Xeroderma pigmentosum complementation group E: a case report. *Photodermatol.* **3**, 233–238 (1986).
13. Stefanini, M. *et al.* Genetic heterogeneity of the excision repair defect associated with trichothiodystrophy. *Carcinogenesis* **14**, 1101–1105 (1993).
14. 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**, 1473–1478 (1982).
15. Limsirichaikul, S. *et al.* A rapid non-radioactive technique for measurement of repair synthesis in primary human fibroblasts by incorporation of ethynyl deoxyuridine (EdU). *Nucleic Acids Res.* **37**, e31 (2009).
16. Nakazawa, Y., Yamashita, S., Lehmann, A.R. & Ogi, T. A semi-automated non-radioactive system for measuring recovery of RNA synthesis and unscheduled DNA synthesis using ethynyluracil derivatives. *DNA Repair (Amst.)* **9**, 506–516 (2010).
17. Kelley, L.A. & Sternberg, M.J. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* **4**, 363–371 (2009).
18. Lohi, O., Poussu, A., Mao, Y., Quiocho, F. & Lehto, V.P. VHS domain—a longshoreman of vesicle lines. *FEBS Lett.* **513**, 19–23 (2002).
19. Dikic, I., Wakatsuki, S. & Walters, K.J. Ubiquitin-binding domains—from structures to functions. *Nat. Rev. Mol. Cell Biol.* **10**, 659–671 (2009).
20. Ren, X. & Hurley, J.H. VHS domains of ESCRT-0 cooperate in high-avidity binding to polyubiquitinated cargo. *EMBO J.* **29**, 1045–1054 (2010).
21. Ito, S. *et al.* XPG stabilizes TFIIH, allowing transactivation of nuclear receptors: implications for Cockayne syndrome in XP-G/CS patients. *Mol. Cell* **26**, 231–243 (2007).
22. Fousteri, M., Vermeulen, W., van Zeeland, A.A. & Mullenders, L.H. Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II *in vivo*. *Mol. Cell* **23**, 471–482 (2006).
23. Bregman, D.B. *et al.* UV-induced ubiquitination of RNA polymerase II: a novel modification deficient in Cockayne syndrome cells. *Proc. Natl. Acad. Sci. USA* **93**, 11586–11590 (1996).
24. Rockx, D.A. *et al.* UV-induced inhibition of transcription involves repression of transcription initiation and phosphorylation of RNA polymerase II. *Proc. Natl. Acad. Sci. USA* **97**, 10503–10508 (2000).
25. Anindya, R., Aygun, O. & Svejstrup, J.Q. Damage-induced ubiquitylation of human RNA polymerase II by the ubiquitin ligase Nedd4, but not Cockayne syndrome proteins or BRCA1. *Mol. Cell* **28**, 386–397 (2007).
26. Starita, L.M. *et al.* BRCA1/BARD1 ubiquitinate phosphorylated RNA polymerase II. *J. Biol. Chem.* **280**, 24498–24505 (2005).
27. Kleiman, F.E. *et al.* BRCA1/BARD1 inhibition of mRNA 3′ processing involves targeted degradation of RNA polymerase II. *Genes Dev.* **19**, 1227–1237 (2005).
28. Takagi, Y. *et al.* Ubiquitin ligase activity of TFIIH and the transcriptional response to DNA damage. *Mol. Cell* **18**, 237–243 (2005).
29. Spivak, G. & Hanawalt, P.C. Host cell reactivation of plasmids containing oxidative DNA lesions is defective in Cockayne syndrome but normal in UV-sensitive syndrome fibroblasts. *DNA Repair (Amst.)* **5**, 13–22 (2006).
30. Zhang, X. *et al.* Mutations in *UVSSA* cause UV-sensitive syndrome and destabilize ERCC6 in transcription-coupled DNA repair. *Nat. Genet.* published online (1 April 2012); doi:10.1038/ng.2228.
31. Schwertman, P. *et al.* UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. *Nat. Genet.* published online (1 April 2012); doi:10.1038/ng.2230.
32. D'Errico, M. *et al.* The role of CSA in the response to oxidative DNA damage in human cells. *Oncogene* **26**, 4336–4343 (2007).

ONLINE METHODS

Human studies. Samples from affected individuals and controls were obtained with local ethical approval (Nagasaki University Ethical, Legal and Social Implications (ELSI) committee). Written informed consent was obtained from the subjects.

Exome sequencing. Genomic DNA of affected individuals (Kps3 and XP24KO) was enriched using the Agilent SureSelect Human All Exon Kit (Agilent, G3362), according to the manufacturer's instructions. The kit covers 1.22% of human genomic regions, which correspond to the CCDS exons. Briefly, genomic DNA was fragmented (~150–200 bp) and ligated to the Illumina sequencing adaptor oligonucleotides. The adaptor-ligated fragments were then amplified by ligation-mediated PCR (LM-PCR) and were then hybridized to the SureSelect Biotinylated RNA library for exon enrichment. The hybridized fragments were captured by streptavidin-coated magnetic beads. The captured genomic fragments were sequenced on a single lane of the Illumina Genome Analyzer IIx (GAIIx) sequencer using 75-bp paired-end reads.

Bioinformatic analysis for the exome data. Low-quality sequences of each read end were filtered out (**Supplementary Table 2a**). In this study, low-quality sequences were defined as reads that contained more than six unknown bases or >40 continuous identical bases or those contaminated by adaptor sequences during the capture and sequencing steps. The filtered sequences were aligned to the human reference genome (GRCh37/hg19) with the Burrows-Wheeler Aligner (BWA) version 0.5.8 (ref. 33) (**Supplementary Table 2a**). Picard version 1.38 was used to convert, sort and index the aligned data files. Picard was also used to identify and remove duplicate reads from the samples. Base-quality scores were recalibrated, and sequence reads were locally realigned with the Genome Analysis Toolkit (GATK) version 1.0.4905 (refs. 34,35). SNVs were identified by the Unified Genotyper program in GATK (**Supplementary Table 2b**). Low-quality variants were then filtered out using the Variant Filtration Walker tools in GATK if they met the following criteria: (i) SNV clustering and proximity to indels, (ii) confidence (QUAL) score of <30, (iii) greater than 10% of aligned reads at a site with mapping quality of 0 (MAPQ0), (iv) strand-bias score of >−0.1, (v) quality-over-depth score of <5 or (vi) largest contiguous homopolymer run of variant allele in either direction (HRun) of >5. Small indels were detected with the detection of indels (Dindel) program, according to the basic procedure for calling indels from a diploid sample in the Dindel user guide version 1.0 (ref. 36). We filtered out any indels matching the following criteria: (i) quality score of <20, (ii) reference homopolymer length of >10 or (iii) indel allele not covered by at least one read on both strands. All variants were annotated with ANNOVAR[37]. Annotation of variants was based on NCBI and UCSC databases. Variants identified in the following SNP databases were removed: dbSNP (Build 131), the 1000 Genomes Project full Phase 1 data that are based on variants from 629 individuals and seven in-house Japanese exome sequencing data sets (**Table 1**). The effect of each mutation was predicted using the Sorting Intolerant from Tolerant (SIFT) score[38] (**Supplementary Table 2c**). We obtained SIFT scores from the ANNOVAR-SIFT database. According to a recessive inheritance model, we selected genes that carried at least one new homozygous or more than two heterozygous changes in one gene locus (**Table 1**) (see **Supplementary Note** for further details).

Run-of-homozygosity (ROH) analysis. Whole-genome SNP genotyping of the subjects with $UV^{S}S$-A was performed on the Genome-Wide Human SNP Array 6.0 (Affymetrix), according to the manufacturer's instructions. The genotype data were generated using the birdseed v2 algorithm in the Genotyping Console 4.0 (Affymetrix). Homozygosity mapping and copy-number analysis were performed by the Partek Genomics Suite v6.5. We determined the runs of homozygosity (ROHs) of at least 1 Mb, allowing up to 1% error. The shared ROHs among the affected individuals were determined. Genomic positions of the identified SNPs were based on the human genome (GRCh37/hg19).

Quantitative RT-PCR analysis. mRNA from affected individuals was purified using the RNeasy mini kit (QIAGEN). cDNA was synthesized from mRNA using a mixture of oligo(dT) and random hexamer primers with the SuperScript II reverse transcription system (Invitrogen). *UVSSA* mRNA transcript levels in 48BR, Kps3 and XP70TO cells were determined by qRT-PCR assays using SYBR Premix Ex-TaqII polymerase (Perfect Real Time system, TaKaRa) and the Thermal Cycler Dice Real Time System II (TaKaRa). Transcripts from the *HPRT1* allele were used as a quantification control. Experiments were performed in triplicate, and qPCR results were analyzed by the $\Delta\Delta C_T$ method. The sequences of qPCR primers used for the assay can be obtained from the company website.

High-resolution melting analysis (HRMA). HRMA of *UVSSA* exons 2 and 3 in 576 unrelated Japanese control individuals (in house) was performed to determine the allele frequencies of the stop-gain SNVs p.Lys123* and p.Ile31Phefs*9 identified in the individuals with $UV^{S}S$-A. The HRMA profile was collected on the LightCycler 480 Real-Time PCR system (Roche). Melting curves were analyzed by LightCycler 480 Gene Scanning software (Roche).

Recovery of RNA synthesis (RRS) assays. Experimental details have been described previously[16]. Cells were plated in plastic 96-well plates. Virus infection was performed 48 h before RRS assays. Cells were UV irradiated (10 J/m^2 of 254 nm UV-C) and incubated for 12 h for RNA synthesis recovery. RRS was measured by the fluorescence-based ethynyluridine incorporation assay[16]. Briefly, recovered cells were incubated for 2 h in medium supplemented with 100 μM 5-ethynyluridine followed by ethynyluridine detection with a copper-catalyzed fluorescent azide conjugation reaction (Click reaction). Cells were fixed and permeabilized in 2% paraformaldehyde and 0.5% Triton X-100 in PBS; after washing with PBS, cells were incubated with 15 μM Alexa Fluor 488 azide (Invitrogen) in 50 mM Tris-HCl, pH 7.3, 4 mM $CuSO_4$, 10 mM sodium ascorbate and 20 ng/ml DAPI for 1 h. Cells were then washed with PBST (PBS with 0.05% Tween20). Nuclear fluorescent image acquisition and data processing were automated using the In-Cell-Analyzer system (GE Healthcare).

Unscheduled DNA synthesis (UDS) assays. Experimental details have been described previously[15,16]. Cells were plated in 96-well plates. siRNA transfection was performed 72 h before UDS assays. UDS was measured by the fluorescence-based ethynyldeoxyuridine incorporation assay[15,16]. Cells were UV irradiated (20 J/m^2 of 254 nm UV-C) and incubated for 4 h in medium supplemented with 5 μM 5-ethynyl-2′-deoxyuridine (EdU). Incorporated EdU was detected by a copper-catalyzed fluorescent azide conjugation reaction (Click reaction). Nuclear fluorescent image acquisition and data processing were automated using the In-Cell-Analyzer system (GE Healthcare).

Immunofluorescence detection of UVSSA protein. Cells were plated in glass-bottomed 96-well plates. Lentivirus particles that expressed either wild-type UVSSA or series of mutant UVSSA proteins (V5 tagged) were used to infect cells 48 h before immunodetection. Cells were fixed and permeabilized in 2% paraformaldehyde and 0.2% Triton X-100 in PBS; after washing with PBS, cells were incubated with mouse antibody to V5 diluted 1:200 in PBST for 1 h and were extensively washed with PBST. Cells stained with primary antibody were then incubated for 1 h with DAPI (5 ng/ml) and a 1:1,000 dilution of rabbit antibody to mouse IgG conjugated with Alexa Fluor 488 (Molecular Probes). Cells were then washed with PBST. Photographs of the cells were captured with an AxioObserver Z1 microscope (Zeiss) equipped with a charge-coupled device (CCD) camera; captured images were analyzed with Axiovision software (Zeiss).

Coimmunoprecipitation. To investigate the interaction of the UVSSA protein with TC-NER factors, HEK293T cells were transfected with expression plasmids encoding V5-tagged UVSSA (wild type and the C32R and W120A

mutants) and incubated for 24 h. In reverse immunoprecipitation experiments, V5-tagged human ERCC3, GTF2H1, CDK7, ERCC6 and ERCC5 were used. Whole-cell lysates were prepared using the CelLytic Nuclear Extraction kit (Sigma). Coimmunoprecipitation was performed using rabbit antibody to V5 conjugated to agarose beads (MBL).

Immunoblotting. Whole-cell lysates and immunoprecipitated samples were resolved by SDS-PAGE (5–20% gradient gels). Resolved protein samples were transferred to PVDF membrane for immunodetection, unless otherwise noted.

Antibodies. Antibodies used for this study were rabbit polyclonal antibody to UVSSA raised against the N-terminal VHS domain of the human UVSSA protein (NBP1-32598, GeneTex), mouse polyclonal antibody to UVSSA raised against full-length human UVSSA protein (H00057654-B01, Abnova), mouse monoclonal antibody to the V5 tag (1H6, MBL), rabbit polyclonal antibody to the V5 tag (PM003, MBL), mouse monoclonal antibody to p89/ERCC3 (AB3, CRUK), rabbit polyclonal antibody to p89/ERCC3 (S-19, Santa Cruz Biotechnology), mouse monoclonal antibody to XPD/ERCC2 (2F6, CRUK), mouse monoclonal antibody to p62/GTF2H1 (G10, Santa Cruz Biotechnology), mouse monoclonal antibody to CDK7 (MO1, MBL), rabbit polyclonal antibody to MNAT1 (MAT1) (FL-309, Santa Cruz Biotechnology), mouse monoclonal antibody to CCNH (cyclin H) (1B8, Abnova), mouse monoclonal antibody to CSB/ERCC6 (553C5a, BMR), mouse monoclonal antibody to XAB2 (5-17, Santa Cruz Biotechnology), mouse monoclonal antibody to XPG/ERCC5 (8H7, CRUK); mouse monoclonal antibody to RNA polymerase II large subunit C-terminal domain Ser2 phosphorylated (H5, Covance), rabbit polyclonal antibody to RNA polymerase II (N20, Santa Cruz Biotechnology) and rabbit polyclonal antibody to p53 phosphorylated at Ser15 (Cell Signaling Technology).

Detection of the elongating RNA Pol IIo after UV irradiation. Cells were cultured in medium supplemented with 100 μM cycloheximide for 1 h before UV irradiation. Cells were irradiated (10 J/m^2 of UV-C) and incubated for the indicated time periods in medium containing cycloheximide. Whole-cell lysates were resolved on 6% SDS-PAGE gels and transferred to PVDF membrane. RNA polymerase II in elongation mode (RNA Pol IIo) was detected with H5 antibody (to phosphorylated CTD-Ser2/5).

Ubiquitin pulldown assays. Crude lysates were prepared from HEK293T cells that were mock transfected or transfected with a plasmid encoding N-terminally 6× histidine–tagged ubiquitin. Cells were irradiated (10 J/m^2 of UV-C) and incubated for 1 h. Extracts were immunoprecipitated with agarose-conjugated antibody to the His tag. Ubiquitinated RNA Pol IIo was detected by H5 antibody.

siRNA experiments. siRNA oligonucleotides targeting *UVSSA* and *XPA* were purchased from Nippon EGT (sequences can be obtained from the company's website). A mixture of three different siRNA oligonucleotides designed for different regions of each gene was used for all experiments, unless otherwise noted. Individual siRNA oligonucleotides were also used for the experiments presented in **Supplementary Figure 4**. siRNA transfection was performed using X-tremeGENE (Roche) transfection reagent, according to the manufacturer's instructions. In typical experiments, 10 nM of siRNA was transfected in suspension in a reverse transfection, and an additional transfection was carried out 24 h after the first (double transfection). Experiments were performed 72 h after the first siRNA transfection. Knockdown efficiency was confirmed by protein blot.

Lentivirus production. Human *UVSSA* cDNA was cloned in frame with sequences encoding a C-terminal V5 tag into pLenti6/V5-D-TOPO (Invitrogen) to generate plenti6/*UVSSA*-V5. Sequences encoding the truncation and amino-acid substitution mutants were generated from plenti6/*UVSSA*-V5 by site-directed PCR mutagenesis using specific primer sets (primer sequences are available from T.O. upon request), PrimeSTAR HS high-fidelity DNA polymerase (TaKaRa) and the DpnI restriction enzyme (NEB). For lentivirus production, HEK293FT cells were transfected with *UVSSA*-encoding plasmids together with ViraPower Packaging Mix (Invitrogen) containing pLP1, pLP2 and pLP/VSVG using Lipofectamine 2000 (Invitrogen). Viral particles were collected 48 h after transfection and concentrated using PEG-it Virus Precipitation Solution (System Biosciences).

Cell lines and culture. The following cell lines were used for this study: 48BR and 1BR (normal human primary fibroblasts); Kps2, Kps3, XP24KO and UV^{S}S24TA (primary fibroblasts from individuals with UV^{S}S-A); XP70TO (primary fibroblast from a mild case of xeroderma pigmentosum); CS2AW (primary fibroblast from individuals with CS-A); CS10LO (primary fibroblasts from individuals with CS-B); XP15BR (primary fibroblasts from individuals with xeroderma pigmentosum–A); and HEK293T and HEK293FT (human embryonic kidney lines). All cells were maintained in DMEM (WAKO) supplemented with 10% FCS (Hyclone) and antibiotics, unless otherwise noted.

33. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
34. DePristo, M.A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
35. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
36. Albers, C.A. *et al.* Dindel: accurate indel calls from short-read data. *Genome Res.* **21**, 961–973 (2011).
37. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
38. Ng, P.C. & Henikoff, S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**, 3812–3814 (2003).

nature genetics

Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome

Yoshinori Tsurusaki[1], Nobuhiko Okamoto[2], Hirofumi Ohashi[3], Tomoki Kosho[4], Yoko Imai[5], Yumiko Hibi-Ko[5], Tadashi Kaname[6], Kenji Naritomi[6], Hiroshi Kawame[7,8], Keiko Wakui[4], Yoshimitsu Fukushima[4], Tomomi Homma[9], Mitsuhiro Kato[10], Yoko Hiraki[11], Takanori Yamagata[12], Shoji Yano[13], Seiji Mizuno[14], Satoru Sakazume[15], Takuma Ishii[15,16], Toshiro Nagai[15], Masaaki Shiina[17], Kazuhiro Ogata[17], Tohru Ohta[18], Norio Niikawa[18], Satoko Miyatake[1], Ippei Okada[1], Takeshi Mizuguchi[1], Hiroshi Doi[1], Hirotomo Saitsu[1], Noriko Miyake[1] & Naomichi Matsumoto[1]

By exome sequencing, we found *de novo SMARCB1* mutations in two of five individuals with typical Coffin-Siris syndrome (CSS), a rare autosomal dominant anomaly syndrome. As *SMARCB1* encodes a subunit of the SWItch/Sucrose NonFermenting (SWI/SNF) complex, we screened 15 other genes encoding subunits of this complex in 23 individuals with CSS. Twenty affected individuals (87%) each had a germline mutation in one of six SWI/SNF subunit genes, including *SMARCB1*, *SMARCA4*, *SMARCA2*, *SMARCE1*, *ARID1A* and *ARID1B*.

Chromatin remodeling factors regulate the gene accessibility and expression by dynamic alteration of chromatin structure. SWI/SNF complexes have important roles in lineage specification, maintenance of stem cell pluripotency and tumorigenesis[1–5]. These complexes are composed of evolutionarily conserved core subunits and variant subunits. Brahma-associated factor (BAF) and Polybromo BAF (PBAF) complexes constitute two major subclasses[1–5]. It has been suggested that the BAF complex is similar to the yeast SWI/SNF complex and that the PBAF complex is more like the chromatin remodelling complex (RSC) in yeast, which is required for cell cycle progression through mitosis[6]. However, several subunits that are common to both BAF and PBAF complexes are predicted to be related to the regulation of lineage- and tissue-specific gene expression[2].

Coffin-Siris syndrome (MIM 135900) is a rare congenital anomaly syndrome characterized by growth deficiency, intellectual disability, microcephaly, coarse facial features and hypoplastic nail of the fifth finger and/or toe (**Fig. 1** and **Supplementary Table 1**)[7]. The majority of affected individuals represent sporadic cases, which is compatible with an autosomal dominant inheritance mechanism. The genetic cause for this syndrome has not been elucidated.

To identify the genetic basis of CSS, we performed whole-exome sequencing of five typical affected individuals (**Supplementary Methods**). Taking into account our model that assumes that an abnormality in a causal gene would be shared in two or more subjects, 51 variants were identified as candidates (**Supplementary Table 2**). All the variants were also examined by Sanger sequencing of PCR products amplified using genomic DNA from the five affected individuals and their parents. Nine variants were found to be false positives, 40 were inherited from either the father or mother, and 2 *de novo* heterozygous mutations of *SMARCB1* were found in 2 affected individuals (c.1130G>A (p.Arg377His) and c.1091_1093del AGA (p.Lys364del)) (**Table 1**, **Supplementary Fig. 1** and **Supplementary Methods**). Two *de novo* coding-sequence mutations occurring within a specific gene is an extremely unlikely event[8], supporting the idea that *SMARCB1* is a causative gene in CSS. Next, we screened *SMARCB1* in 23 individuals with CSS by high-resolution melting analysis[9] and identified the mutation encoding the p.Lys364del alteration in two additional individuals, including one of Arab descent (subject 22) (**Table 1** and **Supplementary Fig. 1**). As the mutation detection rate was relatively low (4 of 23, only 17.4%), we screened 15 additional genes encoding other SWI/SNF subunits (**Supplementary Table 3**). Unexpectedly, four other subunits, *SMARCA4* (also known as *BRG1*), *SMARCE1*, *ARID1A* and *ARID1B* were also found to be mutated (**Table 1** and **Supplementary Figs. 2–5**). In subject 10, a, c.2144C>T mutation in *ARID1B* (encoding p.Pro715Leu) was found in addition to the c.5632delG mutation in *ARID1B*. RT-PCR products that were amplified from total RNA from this subject's lymphoblastoid cells were cloned into the pCR4-TOPO vector. The two mutations were present on different alleles, according to sequencing of clones containing each allele (data not shown). As the c.5632delG mutation is

[1]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan. [2]Division of Medical Genetics, Osaka Medical Center and Research Institute for Maternal and Child Health, Izumi, Japan. [3]Division of Medical Genetics, Saitama Children's Medical Center, Iwatsuki, Japan. [4]Department of Medical Genetics, Shinshu University School of Medicine, Matsumoto, Japan. [5]Division of Pediatrics, Japanese Red Cross Medical Center, Tokyo, Japan. [6]Department of Medical Genetics, University of the Ryukyus Faculty of Medicine, Okinawa, Japan. [7]Department of Genetic Counseling, Graduate School of Humanities and Sciences, Ochanomizu University, Tokyo, Japan. [8]Division of Medical Genetics, Nagano Children's Hospital, Azumino, Japan. [9]Division of Pediatrics, Yamagata Prefectural and Sakata Municipal Hospital Organization, Nihonkai General Hospital, Sakata, Japan. [10]Department of Pediatrics, Yamagata University Faculty of Medicine, Yamagata, Japan. [11]Hiroshima Municipal Center for Child Health and Development, Hiroshima, Japan. [12]Department of Pediatrics, Jichi Medical University, Tochigi, Japan. [13]Genetics Division, Department of Pediatrics, Los Angeles County and University of Southern California Medical Center, Keck School of Medicine, University of Southern California, Los Angeles, California, USA. [14]Department of Pediatrics, Central Hospital, Aichi Human Service Center, Kasugai, Japan. [15]Department of Pediatrics, Koshigaya Hospital, Dokkyo University School of Medicine, Koshigaya, Japan. [16]Nakagawa-No-Sato, Hospital for the Disabled, Saitama, Japan. [17]Department of Biochemistry, Yokohama City University Graduate School of Medicine, Yokohama, Japan. [18]Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Ishikari-Tobetsu, Japan. Correspondence should be addressed to N. Matsumoto (naomat@yokohama-cu.ac.jp) or N. Miyake (nmiyake@yokohama-cu.ac.jp).

Received 29 September 2011; accepted 10 February 2012; published online 18 March 2012; doi:10.1038/ng.2219

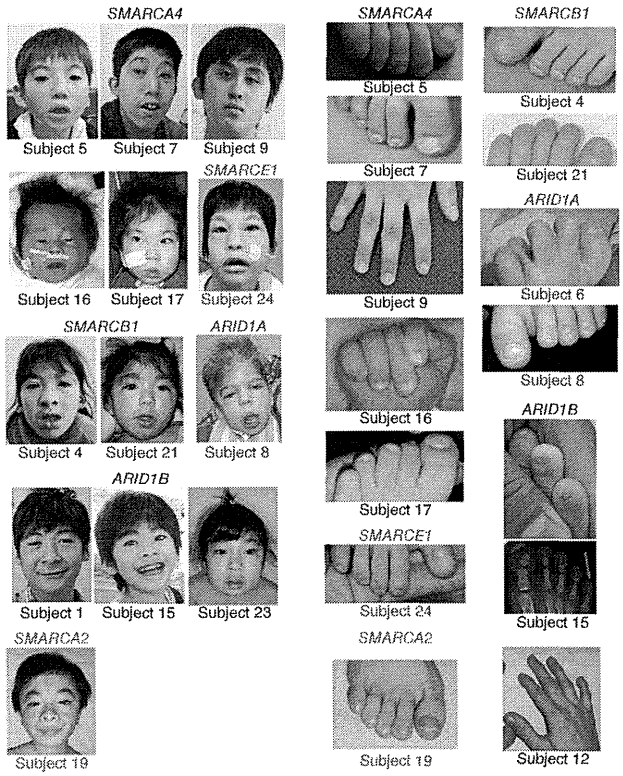

Figure 1 Photographs of individuals with Coffin-Siris syndrome. The faces (left) and hypoplastic-to-absent nail of the fifth finger or toe (right) of affected individuals are shown with the color-coded names of the corresponding mutated genes. The green arrow indicates the absence of the distal phalanx in the fifth toe. No obvious hypoplastic nails were observed in subjects 12 or 19. Consent for all the photographs was obtained from the families of the affected individuals.

very likely to be deleterious (as it results in a truncated protein), the c.2144C>T mutation is likely to be a rare polymorphism. Of note, subject 12, who presented an atypical facial appearance and indistinct hypoplastic nails, had two interstitial deletions at 6q25.3–q27 involving *ARID1B*, as detected by a SNP array (**Supplementary Fig. 6** and **Supplementary Methods**). Furthermore, subject 14 was found to have an interstitial deletion of *SMARCA2* by a SNP array (**Supplementary Fig. 7** and **Supplementary Methods**). No other copy-number changes involving genes encoding SWI/SNF complex components were found in subjects 2, 14 or 18 by array analysis. The overall mutation detection rate was 87%. In total, 20 of the 23 subjects had a mutation affecting one of the six SWI/SNF subunits.

Mutations in CSS were identified in the BAF-specific subunits *ARID1A* and *ARID1B* but not in PBAF-specific subunits (*BRD7*, *ARID2* and *PBRM1*) (**Supplementary Table 3**). In addition, mutations were identified in *SMARCA4* (*BRG1*) as well as in *SMARCA2* (*BRM*) (**Supplementary Table 3**). The BRG1 and BRM proteins are mutually exclusive catalytic ATP subunits in mammalian SWI/SNF complexes. Of note, the majority of heterozygous *Smarca4*-null mice survive with susceptibility to neoplasia, with a minority dying after birth because of exencephaly, whereas homozygous *Smarca2*-null mice are viable and fertile[4]. In *Smarca2*-null mice, Brg1 is upregulated, suggesting that Brg1 can functionally replace Brm in mice[10]. However, in humans, abnormalities in both *SMARCA4* and *SMARCA2* are found in CSS, indicating that the in-frame partial deletion of the gene encoding BRM in subject 19 has a specific mutational effect different from that of simple inactivation in mice. These data support the idea that abnormalities in the BRG1-BAF and BRM-BAF complexes can cause the abnormal neurological development in CSS.

All the mutated genes found in CSS, except for *SMARCE1*, have been reported to be associated with tumorigenesis[1,2]. Among the 23 subjects with CSS, only subject 3 with an *ARID1A* mutation presented with hepatoblastoma. To our knowledge, haploinsufficiency and/or homozygous inactivation of *ARID1A* have been found in several types of cancer but not in hepatoblastoma. Malignancies were not detected in any of the other subjects with CSS examined here. It remains to be seen whether malignancies are robustly associated with CSS.

Given the fact that all the mutations in *ARID1A* and *ARID1B* in CSS were predicted to cause protein truncation, we proposed that haploinsufficiency of these two genes must be able to cause CSS. cDNA analysis of lymphoblastoid cell lines from subjects 1, 6 and 23 indicated that the mutated transcripts were subject to nonsense-mediated mRNA decay (**Supplementary Fig. 8**). In subject 10, the *ARID1B* mutation associated with the creation of a premature stop codon in the last exon did not result in nonsense-mediated mRNA decay as expected (**Supplementary Fig. 8**).

In regard to the other mutated genes, germline heterozygous truncation mutations in *SMARCB1* and *SMARCA4* have been reported

Table 1 Mutations in individuals with Coffin-Siris syndrome

Subject ID	Gene	Mutation	Alteration	Type	Control allele frequency[a]
4	*SMARCB1*	c.1091_1093del AGA	p.Lys364del	*De novo*	0/502
11	*SMARCB1*	c.1130G>A	p.Arg377His	*De novo*	0/500
21	*SMARCB1*	c.1091_1093del AGA	p.Lys364del	NC	0/502
22	*SMARCB1*	c.1091_1093del AGA	p.Lys364del	NC	0/502
9	*SMARCA4*	c.1636_1638del AAG	p.Lys546del	*De novo*	0/350
7	*SMARCA4*	c.2576C>T	p.Thr859Met	*De novo*	0/368
5	*SMARCA4*	c.2653C>T	p.Arg885Cys	*De novo*	0/368
16	*SMARCA4*	c.2761C>T	p.Leu921Phe	*De novo*	0/368
25	*SMARCA4*	c.3032T>C	p.Met1011Thr	NC	0/372
17	*SMARCA4*	c.3469C>G	p.Arg1157Gly	*De novo*	0/368
19	*SMARCA2*	Partial deletion		*De novo*	–
24	*SMARCE1*	c.218A>G	p.Tyr73Cys	*De novo*	0/368
3	*ARID1A*	c.31_56del	p.Ser11Alafs*91	NC	0/330
6	*ARID1A*	c.2758C>T	p.Gln920*	NC	0/376
8	*ARID1A*	c.4003C>T	p.Arg1335*	*De novo*	–
1	*ARID1B*	c.1678_1688del	p.Ile560Glyfs*89	*De novo*	–
15	*ARID1B*	c.1903C>T	p.Gln635*	*De novo*	–
23	*ARID1B*	c.3304C>T	p.Arg1102*	*De novo*	–
10	*ARID1B*	c.2144C>T	p.Pro715Leu	NC	0/368
10	*ARID1B*	c.5632del G	p.Asp1878Metfs*96	NC	0/374
12	*ARID1B*	Microdeletion		NC	–

NC, not confirmed because parental samples were unavailable.
[a]The numbers indicate the observed allele frequency (alleles harboring the change/total tested alleles) in Japanese controls. None of the mutations was found in dbSNP132, the 1000 Genomes database or the National Heart, Lung, and Blood Institute (NHLBI) GO exomesequencing project database. –, not tested.

in individuals with rhabdoid tumor predisposition syndromes 1 (RTPS1; MIM 609322) and 2 (RTPS2; MIM 613325)[11,12], and various types of *SMARCB1* mutations (missense, in-frame deletion, nonsense and splice site) have been found in the germline of individuals with familial and sporadic schwannomatosis (MIM 162091)[13,14]. Furthermore, mice with heterozygous knockout of *Smarca4* or *Smarcb1* were prone to tumor development[2]. All the mutations in *SMARCA4* and *SMARCB1* in individuals with CSS were nontruncating (either missense or in-frame deletions), implying that they exert gain-of-function or dominant-negative effects (excluding haploinsufficiency as a cause). It is noteworthy that comparable germline mutations in *SMARCB1* have such different phenotypic consequences in their association with the phenotypes of CSS and schwannomatosis. The *SMARCB1* mutations in CSS and those in schwannomatosis are indeed different according to the Human Gene Mutation Database. With regard to the *SMARCA2* interstitial deletion in CSS, the change maintained the coding sequence reading frame but removed exons 20–27 that encode the HELICc domain. RT-PCR analysis confirmed the deletion of exons 20–27 at the cDNA level (**Supplementary Fig. 7**). These data suggest the importance of the HELICc domain in the SMARCA2 protein.

The various types of mutations in the genes encoding different SWI/SNF components resulted in similar CSS phenotypes. This suggests that the SWI/SNF complexes coordinately regulate chromatin structure and gene expression. This is the first report, to our knowledge, of germline mutations in SWI/SNF complex genes associated with a multiple congenital anomaly syndrome, highlighting new biological aspects of SWI/SNF complexes in humans. Similarly, genes encoding SNF2-related proteins, which are implicated as chromatin remodeling factors outside of SWI/SNF complexes, are mutated in different syndromes, including in α-thalassaemia/mental retardation syndrome X-linked (ATRX; *ATRX* mutations) and in coloboma, heart defect, atresia choanae, retarded growth and development, genital abnormality and ear abnormality (CHARGE) syndrome (*CHD7* haploinsufficiency)[3]. We expect that more mutations affecting chromatin remodeling factors will be found in different human diseases.

URLs. Human Gene Mutation Database, https://portal.biobase-international.com/cgi-bin/portal/login.cgi.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank all the family members for participating in this study. This work was supported by research grants from the Ministry of Health, Labour and Welfare (to N. Miyake, H.S. and N. Matsumoto), the Japan Science and Technology Agency (to N. Matsumoto), the Strategic Research Program for Brain Sciences (to N. Matsumoto), the Japan Epilepsy Research Foundation (to H.S.) and the Takeda Science Foundation (to N. Matsumoto and N. Miyake). This study was also funded by a Grant-in-Aid for Scientific Research on Innovative Areas (Foundation of Synapse and Neurocircuit Pathology) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to N. Matsumoto), a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (to N. Matsumoto), a Grant-in-Aid for Young Scientists from the Japan Society for the Promotion of Science (to N. Miyake and H.S.) and a Grant for 2011 Strategic Research Promotion of Yokohama City University (to N. Matsumoto). This study was performed at the Advanced Medical Research Center at Yokohama City University. Informed consent was obtained from all the families of affected individuals. The Institutional Review Board of Yokohama City University approved this study.

AUTHOR CONTRIBUTIONS

Y.T., S. Miyatake, I.O., H.D., H.S. and N. Miyake performed exome sequencing and Sanger sequencing. Y.T., M.S., K.O., I.O., T.M., H.D., H.S. and N. Miyake performed data management and analysis. N.O., H.O., T. Kosho, Y.I., Y.H.-K., T. Kaname, K.N., H.K., K.W., Y.F., T.H., M.K., Y.H., T.Y., S.Y., S. Mizuno, S.S., T.I., T.N., T.O. and N.N. provided clinical materials after careful evaluation. Y.T., N. Miyake and N. Matsumoto wrote the manuscript. N. Matsumoto designed and oversaw all aspects of the study.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Reisman, D., Glaros, S. & Thompson, E.A. *Oncogene* **28**, 1653–1668 (2009).
2. Wilson, B.G. & Roberts, C.W. *Nat. Rev. Cancer* **11**, 481–492 (2011).
3. Clapier, C.R. & Cairns, B.R. *Annu. Rev. Biochem.* **78**, 273–304 (2009).
4. Bultman, S. *et al. Mol. Cell* **6**, 1287–1295 (2000).
5. Hargreaves, D.C. & Crabtree, G.R. *Cell Res.* **21**, 396–420 (2011).
6. Xue, Y. *et al. Proc. Natl. Acad. Sci. USA* **97**, 13015–13020 (2000).
7. Coffin, G.S. & Siris, E. *Am. J. Dis. Child.* **119**, 433–439 (1970).
8. Bamshad, M.J. *et al. Nat. Rev. Genet.* **12**, 745–755 (2011).
9. Wittwer, C.T., Reed, G.H., Gundry, C.N., Vandersteen, J.G. & Pryor, R.J. *Clin. Chem.* **49**, 853–860 (2003).
10. Reyes, J.C. *et al. EMBO J.* **17**, 6979–6991 (1998).
11. Schneppenheim, R. *et al. Am. J. Hum. Genet.* **86**, 279–284 (2010).
12. Taylor, M.D. *et al. Am. J. Hum. Genet.* **66**, 1403–1406 (2000).
13. Boyd, C. *et al. Clin. Genet.* **74**, 358–366 (2008).
14. Hadfield, K.D. *et al. J. Med. Genet.* **45**, 332–339 (2008).

Identification of the First ATRIP–Deficient Patient and Novel Mutations in ATR Define a Clinical Spectrum for ATR–ATRIP Seckel Syndrome

Tomoo Ogi[1,2☯*], Sarah Walker[3☯], Tom Stiff[3], Emma Hobson[4], Siripan Limsirichaikul[2¤], Gillian Carpenter[5], Katrina Prescott[4], Mohnish Suri[6], Philip J. Byrd[7], Michiko Matsuse[2], Norisato Mitsutake[1,2], Yuka Nakazawa[1,2], Pradeep Vasudevan[8], Margaret Barrow[8], Grant S. Stewart[7], A. Malcolm R. Taylor[7*], Mark O'Driscoll[5*], Penny A. Jeggo[3*]

1 Nagasaki University Research Centre for Genomic Instability and Carcinogenesis (NRGIC), Nagasaki University, Sakamoto, Nagasaki, Japan, **2** Department of Molecular Medicine, Atomic Bomb Disease Institute, Nagasaki University, Sakamoto, Nagasaki, Japan, **3** Double Strand Break Repair Laboratory, Genome Damage and Stability Centre, University of Sussex, Brighton, United Kingdom, **4** Department of Clinical Genetics, Chapel Allerton Hospital, Leeds, United Kingdom, **5** Human DNA Damage Response Disorders Group, Genome Damage and Stability Centre, University of Sussex, Brighton, United Kingdom, **6** Clinical Genetic Service, City Hospital, Nottingham, United Kingdom, **7** School of Cancer Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom, **8** University Hospitals of Leicester NHS Trust, Leicester Royal Infirmary, Leicester, United Kingdom

Abstract

A homozygous mutational change in the *Ataxia-Telangiectasia and RAD3 related* (*ATR*) gene was previously reported in two related families displaying Seckel Syndrome (SS). Here, we provide the first identification of a Seckel Syndrome patient with mutations in *ATRIP*, the gene encoding ATR-Interacting Protein (ATRIP), the partner protein of ATR required for ATR stability and recruitment to the site of DNA damage. The patient has compound heterozygous mutations in *ATRIP* resulting in reduced ATRIP and ATR expression. A nonsense mutational change in one *ATRIP* allele results in a C-terminal truncated protein, which impairs ATR–ATRIP interaction; the other allele is abnormally spliced. We additionally describe two further unrelated patients native to the UK with the same novel, heterozygous mutations in *ATR*, which cause dramatically reduced ATR expression. All patient-derived cells showed defective DNA damage responses that can be attributed to impaired ATR–ATRIP function. Seckel Syndrome is characterised by microcephaly and growth delay, features also displayed by several related disorders including Majewski (microcephalic) osteodysplastic primordial dwarfism (MOPD) type II and Meier-Gorlin Syndrome (MGS). The identification of an ATRIP-deficient patient provides a novel genetic defect for Seckel Syndrome. Coupled with the identification of further ATR-deficient patients, our findings allow a spectrum of clinical features that can be ascribed to the ATR–ATRIP deficient sub-class of Seckel Syndrome. ATR–ATRIP patients are characterised by extremely severe microcephaly and growth delay, microtia (small ears), micrognathia (small and receding chin), and dental crowding. While aberrant bone development was mild in the original ATR–SS patient, some of the patients described here display skeletal abnormalities including, in one patient, small patellae, a feature characteristically observed in Meier-Gorlin Syndrome. Collectively, our analysis exposes an overlapping clinical manifestation between the disorders but allows an expanded spectrum of clinical features for ATR–ATRIP Seckel Syndrome to be defined.

Citation: Ogi T, Walker S, Stiff T, Hobson E, Limsirichaikul S, et al. (2012) Identification of the First ATRIP–Deficient Patient and Novel Mutations in ATR Define a Clinical Spectrum for ATR–ATRIP Seckel Syndrome. PLoS Genet 8(11): e1002945. doi:10.1371/journal.pgen.1002945

Editor: Andrew O. M. Wilkie, University of Oxford, United Kingdom

Received April 25, 2012; **Accepted** July 26, 2012; **Published** November 8, 2012

Funding: This work has been supported by Special Coordination Funds for Promoting Science and Technology from the Japan Science and Technology Agency (JST) to TO and YN; a Global COE Program from the Ministry of Education, Culture, Sports, Sciences, and Technology of Japan to TO, MM, NM, SL, and YN; and by a Grant in Aid for Scientific Research KAKENHI (24681008) from the Japan Society for the Promotion of Science, a science research grant from Inamori Foundation, a medical research grant from Mochida Memorial Founds for Medical and Pharmaceutical Research, a medical research grant from Daiichi-Sankyo Foundation of Life Science, and a medical research grant from Takeda Science Foundation to TO. The PAJ laboratory is supported by the Medical Research Council, the Association for International Cancer Research, and the Wellcome Research Trust. The MO laboratory is supported by Cancer Research UK, Medical Research Council, and Leukaemia Lymphoma Research. MO and GSS are CR-UK Senior Cancer Research Fellows. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: m.o-driscoll@sussex.ac.uk (MO); A.M.R.Taylor@bham.ac.uk (AMRT); p.a.jeggo@sussex.ac.uk (PAJ); togi@nagasaki-u.ac.jp (TO)

¤ Current address: Faculty of Pharmacy, Silpakorn University, Nakhon Pathom, Thailand

☯ These authors contributed equally to this work.

Introduction

Seckel Syndrome (SS) (OMIM 216000) is an autosomal recessive disorder characterised by marked microcephaly, intrauterine and post-natal growth retardation, developmental delay and characteristic facial features, encompassing micrognathia (small and receding chin), receding forehead and pronounced nose [1]. Majewski (microcephalic) osteodysplastic primordial dwarfism (MOPD) type II and Meier-Gorlin Syndrome (MGS) also display microcephaly and primordial dwarfism [2,3]. How-

Author Summary

Seckel Syndrome (SS) is a rare human disorder characterised by small head circumference and delayed growth. Patients can show additional features including abnormal bone development, receding chins, sloping foreheads, and small ears. In 2003, we identified *ataxia telangiectasia and Rad3 related (ATR)* as a causal genetic defect in two related families displaying SS. However, additional patients with mutations in *ATR* have not hitherto been identified. Here, we describe two further patients with novel mutations in *ATR*. Additionally, we identify a patient with mutations in *ATRIP*, which encodes an interacting partner of ATR, representing a novel genetic defect causing SS. ATR functions to promote the ability of cells to recover from difficulties encountered during replication. We show that patient-derived cells have reduced ATR and ATRIP protein levels and defective ATR/ATRIP function. Our identification of further ATR–ATRIP defective patients and a consideration of their clinical features aids the characterisation and identification of this form of SS and provides insight into the role played by the ATR–ATRIP complex during development.

ever, each of these disorders display an additional spectrum of features conferring clinical distinction. Despite this, on an individual basis, assigning patients to a specific classification is difficult. Additionally, primary microcephaly represents a disorder displaying pronounced microcephaly without marked impact on growth [4]. Five loci conferring SS have been described with four genes identified [5,6]. The first causal genetic defect identified for SS was the *Ataxia-Telangiectasia and RAD3 related (ATR)* gene [7]. A homozygous mutation in *ATR* was identified in two related SS families and cell-based studies provided strong evidence for an impact on ATR function in patient cell lines. This sub-class of SS was designated ATR–SS. More recently, mutations in *CTIP* were identified in a SS patient as well as in a family described as displaying Jawad Syndrome [8]. Additionally, mutations in CENPJ and CEP152, two centrosomal proteins, have been described in SS patients, although mutations in these genes more frequently confer primary microcephaly [9,10]. Mutations in *PCNT*, which encodes a centrosomal protein, and *ORC1L*, a component of the original licensing complex, were reported in patients originally classified as SS although in both cases retrospective analysis revealed that such mutations more frequently cause MOPD type II or MGS, respectively, highlighting the diagnostic challenge faced in the clinic [11–15]. These studies demonstrate that evaluation of multiple patients is required to provide insight into the spectrum of clinical features conferred by specific gene defects, which ultimately aids an understanding of the role of the defective protein during development. To date all ATR–SS patients belong to one of two related families, which harbour the identical homozygous mutation in *ATR*, thereby limiting the characterisation of the clinical phenotype conferred by ATR deficiency. Furthermore, no patients deficient in ATR interacting protein, ATRIP, which is required for ATR stability, have hitherto been described.

ATR, like the related Ataxia-Telangiectasia mutated (ATM) protein, is a phosphoinositol-3 kinase (PI3)-like kinase that functions at the centre of a signal transduction network activated by DNA damage, and most importantly, by replication fork stalling [16]. ATR and ATM share phosphorylation targets but whilst ATM is activated by DNA double strand breaks (DSBs) that arise, for example, following exposure to ionising radiation (IR), ATR is activated by single stranded (ss) regions of DNA that arise following replication fork stalling or exposure to agents that induce bulky DNA adducts [17,18]. Importantly, since replication fork stalling occurs during most cycles of replication, ATR is essential. ATM, in contrast, is non-essential presumably because endogenous DSBs arise infrequently. ATR forms a stable complex with ATR–interacting protein (ATRIP), which is required for ATR stability [19]. Further, ATRIP is required for ATR localisation to ssDNA regions and hence for ATR activation. Consequently, in a range of organisms loss of ATRIP or its homologue, phenocopies ATR deficiency [17,20–22]. Although ATM and ATR share overlapping substrates, ATR specifically phosphorylates Chk1 whilst ATM phosphorylates the related kinase, Chk2. The major functions of ATR are to activate cell cycle checkpoint arrest, stabilise stalled replication forks and promote replication fork restart, which is achieved through its ability to phosphorylate a range of substrates including p53 and H2AX [18,23,24]. Interestingly, in the context of SS, CtIP promotes DNA end resection, which leads to ss DNA formation, the lesion activating ATR. Hence, CtIP functions in a mechanism leading to ATR activation. It is noteworthy that cells derived from *PCNT*-mutated MOPD type II patients are also defective in ATR–dependent G2/M checkpoint arrest although upstream steps in the ATR–signalling pathway are activated normally [11]. These findings suggest that PCNT is required for an important end-point of ATR function. Additionally, the origin licensing complex, components of which are mutated in MGS, is required for the initiation of replication and *ORC1L*-deficient MGS cell lines display slow S phase progression [13]. Similarly, ATR promotes S phase progression by facilitating recovery from replication fork stalling. Together, these findings demonstrate mechanistic overlap between ATR, PCNT and ORC1L, which may underlie some clinical overlap in the disorders conferred by mutations in the genes encoding these proteins.

Here, we provide the first description of a SS patient with mutations in *ATRIP*. Interestingly, the mutational change in one *ATRIP* allele causes impaired ATR–ATRIP interaction and our extensive cellular analysis confirms a deficiency in ATR signalling and damage responses. Additionally, we describe two further, unrelated patients with mutations in *ATR*. The identification and clinical description of an ATRIP patient and two further ATR patients provides a more definitive characterisation of the clinical phenotype conferred by ATR deficiency.

Results

Cells derived from patient CV1720 display a compromised DNA damage response

Patient CV1720 displayed severe microcephaly, growth delay and dysmorphic facial features and was classified as a SS patient (see Table 1 and Figure S1A for details of the clinical features). Cell line CV1720 is a lymphoblastoid cell line (LBL) derived from the patient; fibroblasts were not available. Cells from the previously described ATR–SS (DK0064) patient display impaired DNA damage responses and phosphorylation of ATR substrates [7]. To determine whether CV1720 cells are defective in ATR–dependent G2/M checkpoint arrest, the mitotic index (MI) was monitored at 2 h following UV exposure, a form of DNA damage known to activate ATR–dependent checkpoint arrest. Whilst WT LBLs show a significantly reduced MI following UV exposure compared to undamaged cells, CV1720 cells showed only a modest decrease similar to that observed in DK0064 (ATR–SS) cells (Figure 1A). Cells from the parents of patient CV1720

Table 1. Clinical features of ATR/ATRIP–deficient patients.

	ATRIP–SS		ATR–SS	27-4BI	19-8BI
Ethnicity	**Gujarati-Indian (consanguineous)**		**Pakistani (consanguineous)**	**English**	**English**
Birth.					
OFC (cm)	27.1		24 *(-8SD)*	27	24.2
Wgt (Kg)	2.06		1.1 *(-3SD)*	1.15	0.77
Hgt (cm)	NR		NR	36	NR
Age.	**14 mts**	**3 yrs 3 mts**	**9 yrs**	**20 mts**	**4.5 yrs**
OFC (cm)	-9SD	-10SD	-12SD	-10SD	-10SD
Wgt (Kg)	-5SD	-6SD	-3.3SD	-8SD	-7SD
Hgt (cm)	-5SD	-6.5SD	NR	-8SD	-8SD
Face	Micrognathia, receding forehead, prominent nose.		Micrognathia, receding forehead, prominent nose.	Micrognathia, prominent nose, hypoplastic alae nasi, low set columella, deep set short palpebral fissures.	Micrognathia, blepharophimosis, short palpebral fissures. Prominent nose; high nasal bridge. High anterior hairline.
Teeth	Dental crowding.		Dental crowding and malocclution.	4 teeth at 20 months.	Dental crowding.
Ears	Small lobes.		Posteriorly rotated with absent lobes.	Small, round, low set with poorly formed antihelix tragus & antitragus. Absent lobes.	Small ears with absent lobes
Hands	Bilateral 5th finger clinodactyly.		Multiple ivory epiphysis.	Small, tapering fingers.	Bilateral 5th finger clinodactyly. 5th metacarpels appear short. Blue colouration to both thenar eminence.
Skeletal Survey	Delayed bone age (wrist & hips), symmetric dwarfism.		Microcrania with fuse sutures. Mild thoracic kyphosis. Ribs angulated posteriorly. Narrow iliac blades, cox valga and minor subluxation of the hips. No disslocation of the radial heads	Symmetric dwarfism. Small patellae. No joint hypermobility or kyphoscoliosis.	Symmetric dwarfism. Copper beaten skull. No ossification of the patellae (age 4 yrs). Marked hip & shoulder flexibility. No kyphosis.
Endo-crinology	Normal IGF1, TFT, LH, FSH & cortisol.		NA	NA	NA
MRI	14 mts:generalised cerebral atrophy, normal ventricular systems. Delayed myelination in the anterior limb of the internal capsule. Pituitary is present though of unusual shape with absent fossa.		NA	NA	2 yrs: abnormal gyration in posterior aspect of the cingulated gyrus extending into the parietal occipital region. Hypoplastic corpus collasum.
Other	NR		Developmental delay. Walked at 7 yrs.	No abnormal skin pigmentation. Small feet with metatarsus adductus	No abnormal skin pigmentation. Developmental delay. Sat at 15 mts, walked at 3 yrs 10 mts. High pitched voice, asthma, multiple chest infections, feeding difficulties-reflux (gastrostomy fed). Multiple liver cysts consistent with Caroli's disease found at 17 mts.

NR; not recorded. NA; not assessed.

doi:10.1371/journal.pgen.1002945.t001

(CV1780 and CV1783) displayed normal G2/M checkpoint arrest.

We have previously observed that cells from other SS patients display defects in ATR–dependent G2/M checkpoint arrest but activate upstream steps in the ATR signalling cascade normally [25]. This is exemplified by cell lines from MOPD type II patients with mutations in *PCNT*, which are defective in ATR–dependent G2/M checkpoint arrest but proficient in ATR phosphorylation events [11]. Therefore, next, we examined whether CV1720 LBLs efficiently activate upstream steps in ATR signalling. Since these assays predominantly reflect the response of replicating phase cells, we first verified that CV1720 and control LBLs harbour a similar percentage of S phase cells (Figure S2). Pan-nuclear phosphorylation of H2AX (γH2AX) after replication fork stalling represents an ATR–specific damage response [24]. Strikingly, whilst exposure to 5 mM HU for 2 h resulted in an elevated number of cells staining positively for γH2AX in WT cells, this was not observed in either CV1720 or DK0064 (ATR–SS) cells (Figure 1B). We note that although previous studies have shown that ATM can be activated and phosphorylate γH2AX at DSBs arising following HU treatment in the absence of ATR due to enhanced fork collapse, this was not observed at 2 h post 5 mM HU exposure in these patient cells most likely due to residual ATR activity and/or the early times examined [26,27]. Chk1 represents an important ATR substrate required for G2/M checkpoint arrest. To examine Chk1 activation, we carried out Western Blotting using p-Chk1 antibodies. Following the same UV exposure conditions (2 h post 5 Jm^{-2}) employed to examine G2/M checkpoint arrest, we

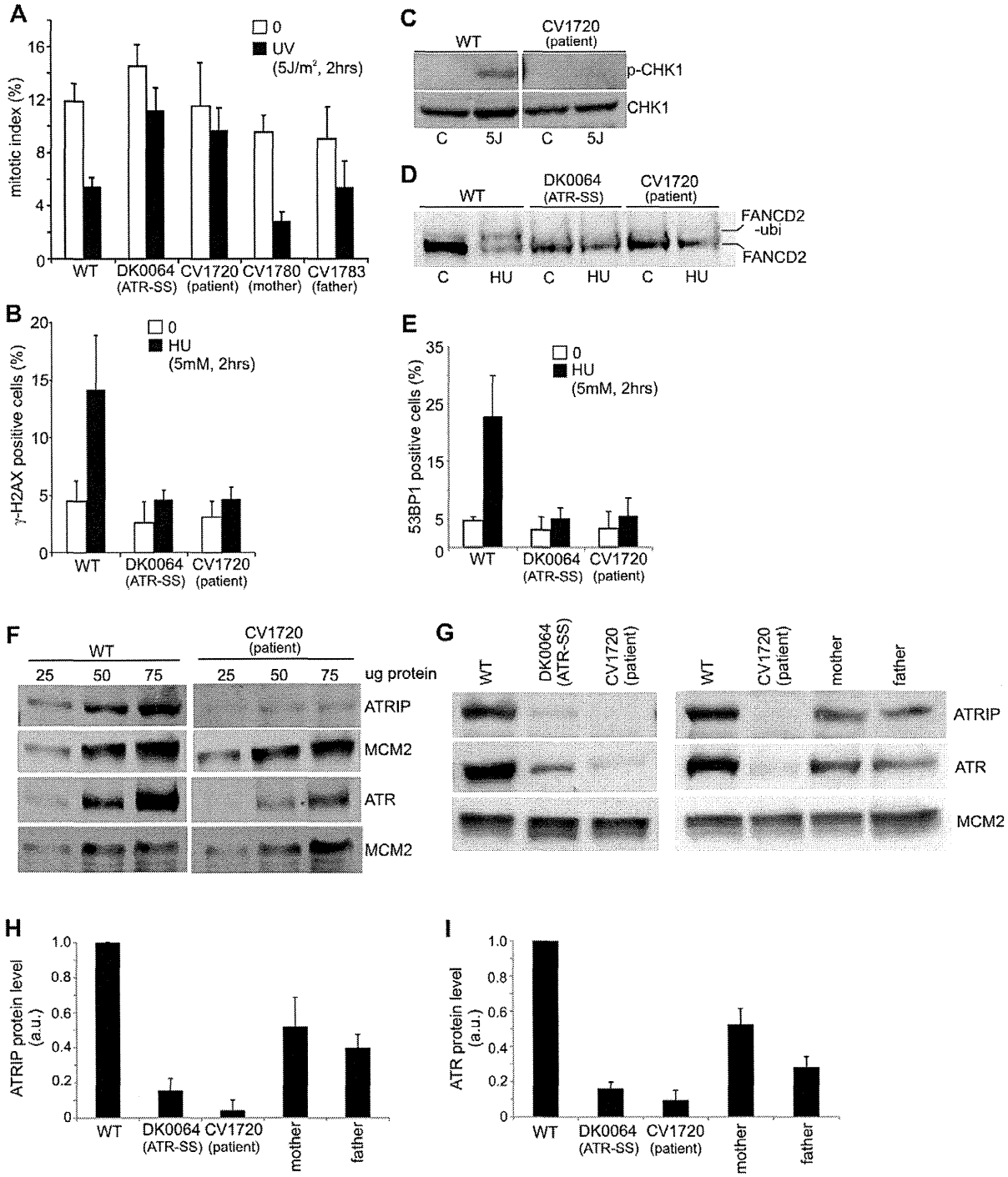

Figure 1. CV1720 cells show impaired ATR-dependent DNA damage responses. A) WT, DK0064 (ATR–SS), CV1720 (patient), CV1780 (patient's mother) and CV1783 (patient's father) cells were exposed to 5 Jm^{-2} UV and the mitotic index (MI) assessed 2 h post exposure. A greater than two fold decrease in mitotic index is observed in WT and both paternal cell lines but not in DK0064 (ATR–SS) or CV1720 (patient) cells. B) Cells were exposed to 5 mM HU for 2 h and the percentage of p-H2AX (γ-H2AX) positive cells assessed by immunofluorescence. Note that HU causes pan nuclear p-H2AX formation rather than defined foci as observed after exposure to ionising radiation. Thus, the percentage of γ-H2AX positive cells was scored. C) Cells were exposed to UV (5 Jm^{-2}) and subjected to Western Blotting (WB) using p-Chk1 (p-Ser317) antibodies at 2 h. Chk1 expression was shown to be similar in WT and patient cells (lower panel). D) Cells were exposed to 3 mM HU for 2 h and whole cell extracts analysed by WB using FANCD2 antibodies. The ubiquitylation of FANCD2, detectable by a product with reduced mobility, is diminished in DK0064 (ATR–SS) and CV1720 cells compared to WT cells. E) Cells were exposed to 5 mM HU and examined for the percentage of cells showing >5 53BP1 foci at 2 h post exposure. 53BP1 foci formation is reduced in DK0064 (ATR–SS) and CV1720 cells compared to WT cells. F–I) The indicated cells were processed by WB using

ATRIP or ATR antibodies. MCM2 was used as a loading control. F shows the analysis of a range of protein levels for accurate comparison. CV1720 (patient) cells show markedly reduced ATR and ATRIP protein levels. G shows that both parental lines have approximately half the level of ATR and ATRIP compared to two WT cell lines. DK0064 (ATR–SS) and CV1720 cells, in contrast, have more dramatically reduced ATR and ATRIP protein levels. 50 ug protein was loaded. WT in all panels was GM2188. Patient, mother and father were as shown in panel A. H and I show the quantification of ATRIP and ATR protein levels from at least three independent WB experiments.
doi:10.1371/journal.pgen.1002945.g001

observed a pronounced p-Chk1 band in WT LBLs but not in CV1720 cells although Chk1 levels were similar in the two lines (Figure 1C). These results provide strong evidence that CV1720 show impaired ATR–dependent substrate phosphorylation.

A further ATR–dependent response is mono-ubiquitylation of FANCD2 following exposure to HU [28]. Mono-ubiquitinated FANCD2 can be detected by the presence of a slowly migrating isoform of FANCD2 generated post exposure to 3 mM HU. Whilst such a product was detected in WT cell extracts, it was absent in CV1720 and DK0064 (ATR–SS) cell extracts (Figure 1D). Finally, ATR also regulates the formation of 53BP1 foci following replication fork stalling via a Chk1-dependent process. We observed a failure to form 53BP1 foci following exposure to 5 mM HU in CV1720 and DK0064 (ATR–SS) LBLs in contrast to WT LBLs (Figure 1E), consistent with the diminished levels of p-Chk1 observed in CV1720 cells.

Collectively, these studies provide strong evidence that CV1720 cells are defective in an upstream step of the ATR–dependent signalling response defining them as distinct to the majority of previously characterised SS cell lines, which, though defective in UV-induced G2/M checkpoint arrest, are proficient in upstream steps of the ATR signalling response [25].

Reduced ATR and ATRIP protein expression in CV1720 cells

Given the overlapping cellular phenotype between CV1720 and DK0064 (ATR–SS) cells, we examined CV1720 cells for expression of ATR and ATRIP protein by Western Blotting. Strikingly, we observed markedly reduced levels of both ATR and ATRIP in CV1720 cells (Figure 1F). Since ATRIP stabilises ATR, this does not distinguish whether the primary defect lies in ATR or ATRIP and indeed a similar reduced level of ATR and ATRIP was observed in DK0064 (ATR–SS) cells (Figure 1G). Significantly, we observed reduced ATRIP and ATR in both parental LBLs (CV1780 and CV1783), which was approximately 50% of the level in WT LBLs (Figure 1G–1I).

Sequencing analysis reveals mutational changes in *ATRIP* in CV1720 cells

To examine whether the causal genetic defect in CV1720 lies in *ATR* or *ATRIP*, we carried out sequencing of the two genes. First, we undertook PCR-based gDNA sequencing of the 47 exons and neighbouring exon-intron boundaries of the human *ATR* gene from CV1720 cells and failed to observe any mutational changes likely to be of functional significance. Next, we undertook gDNA sequencing of *ATRIP* exons and observed a heterozygous mutational change, c.2278C>T, in exon12 which generated a stop codon predicting a truncated protein at position arginine 760 (p.Arg760*) (Figure S3). However, no mutational changes in any other exons were identified although we detected several novel intronic changes that could potentially impact on splicing (Table S1). Significantly, the c.2278C>T mutational change was observed as a heterozygous change in the patient's mother but not in the father (Figure S3).

We also performed RT-PCR sequencing of *ATRIP* cDNA from CV1720 and both parents. These analyses revealed a low level of a smeared PCR product following amplification of the 5′*ATRIP* cDNA region using patient but not control cDNA (data not shown). Following multiple analyses, we found specifically that RT-PCR amplification using primers located in exons 1 and 4, reproducibly yielded a smeared product from CV1720 cDNA with discrete bands at 458 bp (the expected product size) and 325 bp whereas only the 458 bp product was observed using cDNA from WT cells (Figure 2A). Direct sequencing of the gel purified smaller (325 bp) and full-length (458 bp) RT-PCR products showed that the small fragment specifically lacked exon 2. Sequencing analysis of the RT-PCR product of CV1720 cDNA using the same primers revealed the predicted double sequence with the product lacking exon 2 being less than 50% of the product containing exon 2 (Figure 2B). It is notable that there were also some PCR products larger than the full length product although a discrete band was not evident. In sequencing the RT-PCR product, we observed some that harboured intron 2 sequences although these represented a minor product relative to that lacking exon 2. Collectively, these findings strongly suggested that there could be mis-splicing in CV1720 cells with loss of exon 2 being the major product.

To assess this further, qRT-PCR was undertaken using sets of primers that allow selective amplification of the WT and mutant products (c.2278C>T mutant as well as the mis-spliced product). The aim was to determine if the mis-spliced product originated from the paternal allele and if it impacted upon the transcript level. Primer pairs, P1 and P3C, located in exons 12 and 13, respectively, allow selective amplification of the WT (paternal) c.2278C allele whilst primers P2 and P3C selectively amplify the mutated (maternal) c. 2278C>T allele (Figure 2C). As expected, the mutant (c.2278C>T)-allele-specific PCR product (right columns, red bars) was only detected in the patient and mother whereas the WT-specific PCR product (left columns, blue bars) was detected in all samples, demonstrating that the primers distinguished the two alleles (Figure 2C). The results also showed that the c.2278C>T and the WT (c.2278C) alleles were expressed at nearly equal levels (normalised against *HPRT1*) in the mother (compare blue and red bars labelled 'mother' in Figure 2C), suggesting that the c.2278C>T *ATRIP* mRNA is not subject to nonsense mediated RNA decay (NMD) (Figure 2C).

To evaluate the expression level of the mis-spliced *ATRIP* mRNA, we designed primers located at the exon 2/exon 3 boundary (primer P4) and within exon 3 (primer 6C) to allow selective amplification of the correctly spliced mRNA (Figure 2D); primers located at the exon 1/exon 3 boundary (primer P5) and within exon 3 (primer 6C) selectively amplify the mis-spliced mRNA. Whilst the correctly spliced product was amplified to similar (although slightly different) extents from father, mother and patient mRNA (compare the column heights, left panel in Figure 2D), the mis-spliced product was more abundant in the patient and father, suggesting that mis-splicing is a consequence of a mutational change linked to the paternal allele (compare the column heights, right panel in Figure 2D). Since we observed nearly equal expression levels of the wild type (c.2278C) and mutant (c.2278C>T) alleles in the mother (Figure 2C, compare the right and left panels), we considered that the PCR products derived from the mother using primers P4/P6C or P5/P6C would

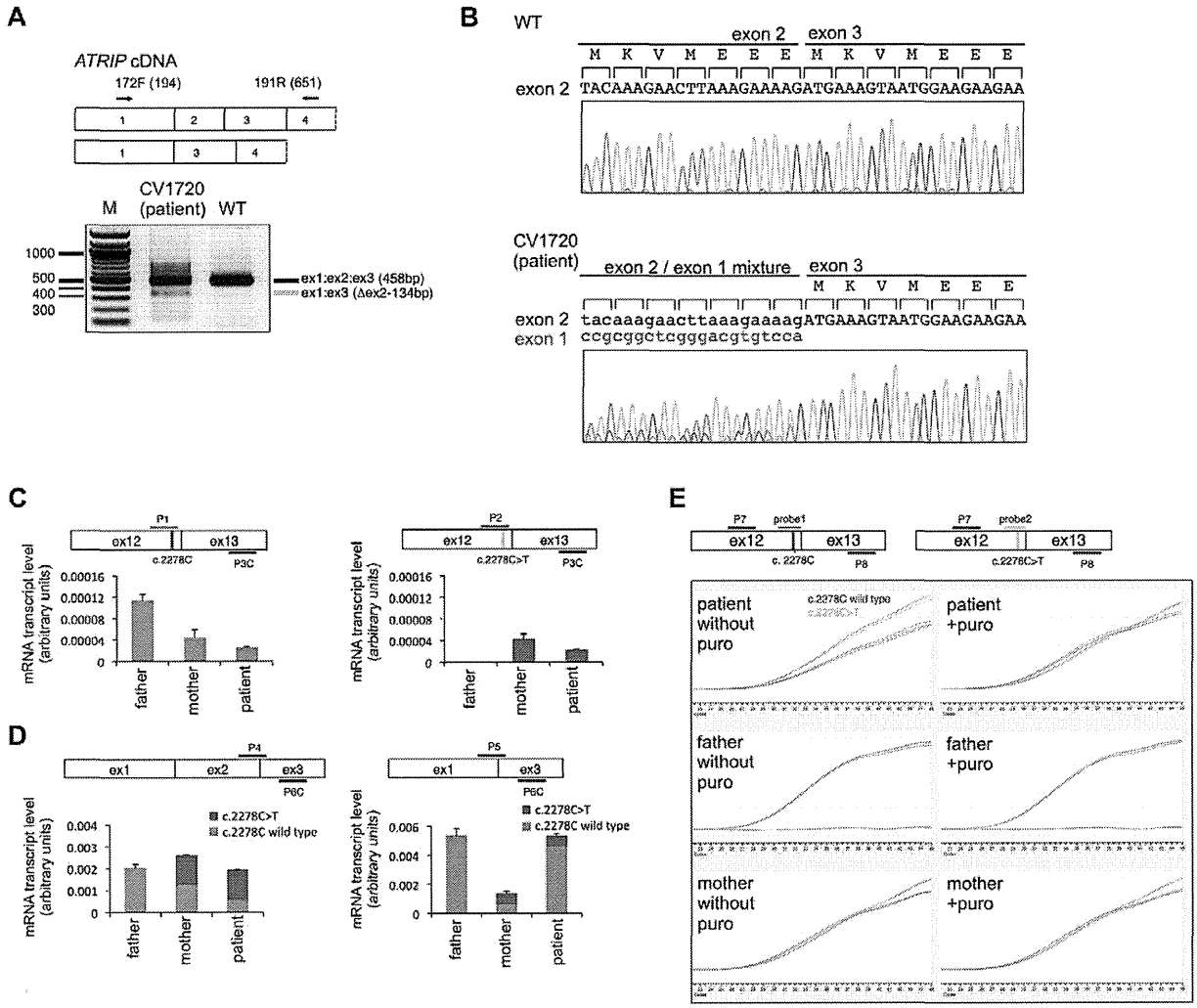

Figure 2. Identification of mutational changes in *ATRIP* in CV1720. A) Upper panel shows primer pairs used to distinguish cDNA products encompassing or lacking exon 2. Lower panel shows RT-PCR products obtained using the primers shown in the upper panel. RT-PCR from patient CV1720 generated a smeared product with a defined band of 458 bp, as observed in WT cells, and a weaker band of 325 bp. The latter band was not detected using cDNA from WT cells (MRC5). A similar single 458 bp band was obtained using the same primers with cDNA derived from a distinct wild type cell line (GM2188; data not shown). (B) Sequencing of the RT-PCR products derived from WT (MRC5) and patient (CV1720) cells. A double sequence pattern at the exon 2–3 boundary is observed using patient CV1720 cDNA. C) Selective quantitative amplification of the WT or 2278C>T *ATRIP* alleles. Primers located in *ATRIP* exon 12 and 13 were designed to selectively amplify the WT (c.2278C) (P1 and P3C) versus the mutated (c.2278C>T) (P2 and P3C) alleles. The WT PCR product is shown in blue and the c.2278C>T PCR product in red. The exon 12 mutated allele is only observed in the patient and mother cDNA whilst the WT allele is observed in the patient, mother and father cDNA although the level is reduced in the patient and mother. D) qRT-PCR analysis of *ATRIP* splicing variants from patient CV1720 and parental cells. qRT-PCR analysis of the level of the normally spliced (encompassing exons 1-2-3) and the aberrantly spliced (Δexon2) *ATRIP* cDNA in the patient and parent cells. PCR primers were designed at the exon2-exon3 or exon1-exon3 boundaries to selectively amplify the splicing variants. Transcripts from *HPRT1* were used as a quantification control. The correctly spliced transcript from the paternal allele of the patient (wild type c.2278C, blue fraction in the cumulative bar labelled, 'patient', at the left panel) was estimated to be ~25% of the normal level. (E) The mis-spliced paternal allele is subject to nonsense mediated mRNA decay (NMD). Cycleave-qPCR confirmed that the *ATRIP* c.2278C>T mutant allele was expressed exclusively in the patient and the mother. The *ATRIP* exon12-13 fragment was amplified with PCR primers P7/P8 as shown in the figure. A set of fluorescent probes were used to distinguish the WT versus c.2278C>T allele (probe1 and probe2, respectively). In the patient, the paternal mRNA transcript level (emerald lines) is low because of NMD (top left). Puromycin treatment eliminated the NMD and the paternal transcript level returned to the normal level. In all panels WT represented MRC5, patient was CV1720 and parents were as shown in Figure 1A.
doi:10.1371/journal.pgen.1002945.g002

be equally derived from the c.2278C and c.2278C>T alleles, which have, therefore, been depicted as equal sized contributions (shown in red or blue in mother columns in Figure 2D). Similarly, the mutant c.2278C>T allele is likely to be expressed at an equal level in the patient as in the mother (shown in red in patient columns in Figure 2D). Based on these assumptions, we estimated that the normally spliced WT mRNA is reduced to around 1/4 of the WT level in the patient and to 3/4 in the father (shown in blue in the left hand panel in Figure 2D). Assuming that the c.2278C>T allele is fully inactivated (see below), the patient

therefore has around 25% of ATRIP activity compared to a normal individual.

The findings above suggested that the mis-spliced mRNA, which generates an out of frame cDNA, is subject to NMD. To examine this and substantiate our findings, qRT-PCR was also carried out using fluorescent cycleave probes with or without exposure to puromycin, an antibiotic which prevents NMD (Figure 2E) [29]. Primers (P7 and P8) and fluorescent probes (probe 1 and 2) were designed to allow amplification of a product encompassing exon 12–13 that distinguished the maternal (probe 2) from the paternal (probe 1) allele. We confirmed detection of the c.2278C>T allele exclusively in the patient and mother as well as the WT allele in all samples (Figure 2E). In the mother, the wild type (c.2278C) and mutant (c.2278C>T) signals were detected at equal levels regardless of whether puromycin was added, indicating that the alleles are equally expressed and are not subjected to NMD. In patient CV1720, the WT product was reduced relative to the mutant product in the absence of puromycin but was at similar levels in the presence of puromycin (Figure 2E). These findings are consistent with the notion that the mRNA expressed from the parental allele is aberrantly spliced and partially subject to NMD. Perhaps surprisingly, we did not detect any obvious difference of the WT product following puromycin treatment in the father; however, in this case, we anticipate a 25% decreased product, which is unlikely to be detected without an internal control. However, despite this, there was evidence for abnormal splicing in the paternal cDNA from analysis of the PCR products spanning exons 1–3 (Figure 2B, 2D).

Finally, to gain insight into the basis underlying mis-splicing, we sequenced introns 1 and 2 from the patient, mother, and father and identified a previously unreported mutational change in intron 2, 13 bp from the intron-exon 2 boundary in the patient and paternal gDNA (Table S1). However, given the modest impact on splicing we did not attempt to examine whether this represented the causal mutational change affecting splicing.

Arg760* ATRIP does not promote ATR-dependent G2/M arrest and reduces ATR–ATRIP interaction

It is likely that *ATRIP* c.2278C>T causes an impacting mutational change since the low levels of ATRIP protein (10–20% WT levels) in CV1720 cells suggest that p.R760* ATRIP is unstable (given that the mRNA level of this allele is normal). To substantiate that p.R760* expression impairs the ATR–dependent response to DNA damage, we examined whether its expression could complement the G2/M checkpoint defect of CV1720 cells. We also examined whether p.R760* might exert a dominant negative impact (since this represented a possible explanation for the low ATRIP protein level in CV1720 cells). The c.2278C>T mutational change was introduced into *ATRIP* cDNA by site directed mutagenesis. cDNA encoding WT *ATRIP* and/or R760* ATRIP was transiently transfected into LBLs and G2/M checkpoint arrest examined at 2 h post exposure to 5 Jm^{-2} UV. Consistent with previous findings, WT but not CV1720 cells showed a G2/M checkpoint arrest (Figure 3A). Whilst transfection with WT *ATRIP* cDNA completely rescued the G2/M checkpoint defect of CV1720 cells, no correction was observed in CV1720 cells following expression of c.2278C>T *ATRIP* cDNA (encoding R760* ATRIP). Surprisingly, expression of WT *ATRIP* cDNA also corrected the G2/M checkpoint defect in DK0064 (ATR–SS) cells, which we propose could result from elevated ATRIP expression causing stabilisation of residual ATR protein, since ATR–SS cells have low ATR and ATRIP expression. Significantly, c.2278C>T *ATRIP* cDNA was unable to rescue ATR–SS cells. Finally expression of c.2278C>T *ATRIP* cDNA in WT cells did not affect G2/M checkpoint arrest demonstrating that p.R760* ATRIP does not exert a dominant negative impact. Collectively, we conclude that p.R760* ATRIP impacts upon ATRIP function.

Next we examined how loss of the ATRIP C-terminus might impact upon ATRIP function. Two studies have previously observed that the C-terminal region of ATRIP is required for interaction with ATR [21,30]. Falck *et al* [30] reported that ATR–ATRIP interaction required the C-terminal 32 amino-acids of ATRIP (769–791) whilst Ball *et al* [21] found that interaction was abolished in a protein that lacked exon 11, which encompasses amino-acids 658–684. Arg760 lies close to these regions. To examine whether p.R760* ATRIP can interact with ATR, HA-tagged *WT* or c.2278C>T (ATRIP R760*) cDNA was co-expressed with untagged WT *ATR* cDNA in HEK293 cells. Following IP with HA-agarose, the level of co-immunoprecipitated ATR was assessed by Western Blotting. Although there was a low level of non-specific ATR binding to the HA beads, the level of ATR present after HA-R760* ATRIP expression (derived from c.2278C>T *ATRIP* cDNA) was substantially lower than after HA-WT ATRIP expression (Figure 3B left panel). Both WT and R760* ATRIP were efficiently expressed, however (Figure 3B right panel). Thus, we conclude that R760* impairs the binding of ATRIP to ATR.

Identification of further patients with mutations in *ATR*

In the course of our functional characterisation of cell lines from SS patients, we examined LBLs derived from two SS patients, 27-4BI and 19-8BI (see Figure 4A, Figure S4, and Table 1 for clinical details). Western Blotting revealed that both cell lines displayed substantially reduced ATR protein whilst showing normal expression of other DNA damage response components, including CtIP, TOPBP1 and RAD17 (Figure 4B). 27-4BI also had reduced ATRIP levels. Additionally, the 27-4BI cell line expressed normal levels of PCNT, excluding MOPD type II as a potential genetic diagnosis, since most of these patients exhibit severely reduced PCNT expression. These findings raised the possibility that the patients could harbour mutations affecting ATR or ATRIP expression. Sequencing of *ATR* cDNA revealed the same c.3477G>T mutational change in both patients (Figure S5A). This change causes an amino acid substitution, p.Met1159Ile, which lies within a conserved UME (NUC010) domain of ATR. UME domains, and particularly the methionine residue within the domain, are highly conserved in ATR species, including yeast although their function is unknown (Figure 4C and 4D).

The second *ATR* mutation identified was c.6897+464C>G;p.Val2300Gly fs75*, which, surprisingly, was also present in both patients. RT-PCR sequencing showed that a 142 bp sequence, which originated from a repeat region present in intron 40, was inserted at the boundary between exon 40 and 41 in both patients (Figure S5C). Genomic sequencing revealed the presence of a single C>G mutation in intron 40, which generates a preferred splice signal causing insertion of the intron sequence to the start of exon 41 (Figure S5D for further details). This insertion causes a frameshift and the generation of a stop codon at c.6978 in exon 41. Sequencing of *ATRIP* cDNA in patient 27-4BI failed to reveal any mutational changes. Thus, our findings provide strong evidence that mutational changes in ATR underlie the reduced ATR/ATRIP expression observed in both patients.

To verify that these mutational changes impact upon ATR function, we examined whether 27-4BI cells could activate UV-induced ATR–dependent G2/M checkpoint arrest. Significantly, we observed an inability to activate UV-induced G2/M checkpoint arrest in 27-4BI cells similar to that observed in DK0064