Accepted versión of: Bogliolo, M., et al. “Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia “ in American journal of human genetics *(Cell Press)*, Vol. 92 (May 2013) , p. 800-806. The final version is available at DOI: 10.1016/j.ajhg.2013.04.002

Mutations of the DNA repair endonuclease ERCC4/XPF cause Fanconi anemia

Massimo Bogliolo[1,2,10], Beatrice Schuster[3,10], Chantal Stoepker[4], Burak Derkunt[5], Yan Su[5], Anja Raams[6], Juan P. Trujillo[1], Jordi Minguillón[1], María J. Ramírez[1,2], Roser Pujol[1,2], José A. Casado[2,7], Rocío Baños[2,7], Paula Rio[2,7], Kerstin Knies[3], Sheila Zúñiga[8], Javier Benítez[2,9], Juan A. Bueren[2,7], Nicolaas G.J. Jaspers[6], Orlando D. Schärer[5], Johan P. de Winter[4], Detlev Schindler[3,11] and Jordi Surrallés[1,2,11]

[1]Genome Instability and DNA Repair Group, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Bellaterra, 08193, Barcelona, Spain.

[2]Centre for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, Bellaterra, 08193, Barcelona, Spain.

[3]Department of Human Genetics, University of Wurzburg, D-97074, Wurzburg, Germany.

[4]Department of Clinical Genetics, VU University Medical Center, 1081 BT, Amsterdam, The Netherlands.

[5]Department of Pharmacological Sciences and Chemistry, Stony Brook University, 11794-3400, Stony Brook, New York, USA.

[6]Department of Genetics, Erasmus University Medical Center, 3015 GE, Rotterdam, The Netherlands

[7]Hematopoiesis and Gene Therapy Division, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), 28040, Madrid, Spain.

[8]Department of Bioinformatics. Sistemas Genómicos SL, 46980, Valencia, Spain.

[9]Human Genetics Group, Spanish National Cancer Center, CNIO, E-28029, Madrid, Spain.

[10]Contributed equally to this work

[11]Correspondence should be addressed to JS (jordi.surralles@uab.es) or DS (schindler@biozentrum.uni-wuerzburg.de)

Fanconi anemia (FA) is a rare genomic instability disorder characterized by progressive bone marrow failure and predisposition to cancer. FA gene products are involved in the repair of DNA interstrand cross-links. Fifteen FA genes have been identified, but the genetic basis in some individuals still remains unresolved. Here we used whole exome and Sanger sequencing on DNA of unclassified FA individuals and discovered biallelic germline mutations in *ERCC4/XPF*, a structure-specific nuclease previously connected to xeroderma pigmentosum and segmental progeria type XFE. Genetic reversion and wild-type *ERCC4/XPF* cDNA complemented the phenotype of the FA cell lines, providing genetic evidence that mutations in *ERCC4/XPF* cause this FA subtype. Further biochemical and functional analysis demonstrate that the identified FA-causing *ERCC4/XPF* mutations strongly disrupt the function of XPF in DNA interstrand cross-link repair without severely compromising nucleotide excision repair. Our data show that depending on the type of *ERCC4/XPF* mutation and the resulting balance between both DNA repair activities, individuals present with one of the three clinically distinct disorders, highlighting the multifunctional nature of the XPF endonuclease in genome stability and human disease.

Fanconi anemia (FA) is characterized by bone marrow failure (BMF), congenital malformations, hypersensitivity to DNA interstrand cross-linking (ICL) agents, chromosome fragility and a high susceptibility to cancer. Since the discovery of the first FA gene twenty years ago, fifteen FA genes have been identified, *FANCA, FANCB, FANCC, BRCA2/D1, FANCD2, FANCE, FANCF, ERCC9/FANCG, FANCI,*

BRIP1/FANCJ, PHF9/FANCL, FANCM, PALB2/N, RAD51C/FANCO and *SLX4/FANCP*[1; 2]([MIM 227645; 227646; 227650; 300514; 600901; 603467; 605724; 609053; 609054; 610832; 613390; 613951; 614082; 614083; 614087, respectively]). Studies to unravel the genetic basis of this rare disorder uncovered a genome maintenance pathway that protects dividing cells against replication-blocking DNA lesions. To identify additional FA genes, the SOLiD™ 4 platform was used for whole exome sequencing on peripheral blood DNA from a Spanish FA individual (FA104) who was previously excluded from all known FA complementation groups. FA104 was born from unrelated parents and diagnosed neonatally due to a malformative syndrome suggestive of FA, including bilateral absent thumbs, microsomy, esophageal atresia, ventrally translocated anus and dysplastic and low-set ears. She did not show any dermatological abnormality such as skin hyperpigmentation, photosensitivity, sunlight-induced scarring or atrophy. FA104 developed BMF at the age of two and died due to a hemorrhagic shock after bone marrow transplantation at age 4. A positive chromosome breakage test unambiguously confirmed the FA diagnosis, with 92% of the cells showing on average 4.4 diepoxybutane (DEB)-induced breaks. Lymphoblasts from this individual were hypersensitive to mitomycin-C (MMC) and melphalan, but insensitive to the topoisomerase I inhibitor camptothecin and the PARP inhibitor KU58948 (data not shown), and showed normal FANCD2 monoubiquitination and RAD51 focus formation[3]. This suggests a defect downstream in the FA/BRCA pathway, which does not involve homologous recombination. Based on a recessive mode of inheritance, exome sequencing identified 17 candidate disease genes for FA104 (Supplemental Table 1), of which *ERCC4/XPF* (xeroderma pigmentosum group F [MIM 133520]) immediately caught our

attention given the involvement of the XPF-ERCC1 structure-specific nuclease in ICL repair[4]. Both *ERCC4/XPF* mutations were predicted to be pathogenic: a 5 base pair (bp) deletion in exon 8 (c.1484_1488delCTCAA) leading to a frameshift and a premature stop codon (p.(Thr495Asnfs*6)), and a missense mutation in exon 11 (c.2065C>A; p.(Arg689Ser); RefSeq: NG_011442.1 NM_005236.2 and NP_005227.1) that changes a highly conserved arginine within the nuclease active site of XPF. Sanger sequencing on blood DNA confirmed these mutations (Fig. 1A) and their correct segregation (data not shown). In a MMC-resistant FA104 lymphoblasts (FA104R) obtained after long-term exposure to low dose of MMC we detected a mutation that restored the *ERCC4/XPF* reading frame (Supplemental Fig. 1A), supporting the notion that MMC sensitivity is due to *ERCC4/XPF* mutations. Consistently, XPF protein levels were reduced in FA104 lymphoblasts but normalized in the reverted FA104R lymphoblast (Supplemental Fig. 1B). Western blotting did not detect a truncated XPF protein, indicating that only the XPF-R689S mutant protein is present in the FA104 cell line.

Sanger sequencing on 18 unclassified FA individuals from Germany revealed biallelic *ERCC4/XPF* mutations in another individual (1333). Individual 1333 was born in 2002 and was unambiguously diagnosed with FA at the age of 5 years due to multiple FA-related features such as perinatal growth retardation, short stature, pronounced microcephaly, café-au-lait spots, ostium primum defect, biliary atresia with fibrosis of the liver, BMF and a positive chromosome fragility test (0.2, 6.7 and 9.4 breaks per cell at 0, 50 and 100 ng ml^{-1} MMC, respectively). Individual 1333 was redhead with pale skin color but no spontaneous or UV light-induced skin lesions were reported at the age of 10.

Similar to FA104, lymphoblasts from individual 1333 were normal in FANCD2 monoubiquitination and RAD51 focus formation, sensitive to MMC and melphalan but insensitive to the topoisomerase I inhibitor camptothecin and to the PARP inhibitor KU58948 (data not shown). Individual 1333 carried a 28 bp duplication in exon 11 of the maternal allele (c.2371_2398dup28; p.Ile800Thrfs*24; Fig. 1B), which is predicted to result in a truncated XPF protein that lacks the double helix-hairpin-helix (HhH_2) domain involved in heterodimerization with ERCC1 and DNA binding[5]. The paternal allele contained a missense mutation that changed a highly conserved amino acid residue within the helicase-like domain (c.689T>C; p. (Leu230Pro); Fig. 1B)). Western blot analysis showed that a missense mutant and a truncated 90-95 kDa XPF protein are present at very low levels (Fig. 1C). As expected, the truncated XPF protein was undetectable with an antibody against the C-terminal HhH_2 domain of XPF (amino acids 866-916, data not shown). Interestingly, the truncated XPF protein was absent in a MMC-resistant lymphoblastoid cell line (1333R) generated by long-term exposure to MMC, and near normal XPF levels were detected in this reverted cell line (Supplemental Fig. 1C). PCR amplification and sequence analysis revealed that the 28 bp duplication had disappeared in 1333R (Supplemental Fig. 1D) restoring the wild type sequence. Both, the inherited duplication and the somatic reversion may have been triggered by an inverted 5 bp repeat flanking the region.

Genetic complementation of MMC sensitivity in lymphoblasts from both FA individuals was achieved by lentiviral transduction of wild type *ERCC4/XPF* cDNA (Fig 1D,E). In addition, we expressed wild type and mutant human *ERCC4/XPF* cDNAs in embryonic

fibroblasts (MEFs) from *Ercc4/Xpf* null mice and found that *ERCC4/XPF* mutants L230P and R689S did not complement MMC sensitivity of these MEFs (Fig. 1F), providing additional evidence that the *ERCC4/XPF* missense mutations found in both FA individuals inactivate XPF. The genetic and functional data show that mutations in *ERCC4/XPF* cause FA in two unrelated non-consanguineous individuals. Since mutations in *ERCC4/XPF* cause an additional FA subtype (FA-Q), we propose *FANCQ* as an alias for *ERCC4*/*XPF*.

ERCC4/*XPF* mutations have been linked to the skin photosensitive and nucleotide excision repair (NER) deficient disorders xeroderma pigmentosum (XP [MIM 278700; 610651; 278720; 278730; 278740; 278760; 278780; 278750])[6] or progeria (XFE [MIM 610965])[7] and therefore we tried to understand why the identified *ERCC4/XPF* variants specifically lead to FA. We hypothesized that these mutants cause an FA phenotype due to a strong deficiency in ICL repair, but have sufficient NER activity to prevent clinically relevant skin photosensitivity and other NER-related features. FA104 lymphoblasts, were indeed not sensitive to UVC when compared to an XPC lymphoblast (Fig. 2A). Since UV survival experiments are challenging in lymphoblastoid cell lines, we studied skin fibroblasts from individual 1333 (FA104 fibroblasts were not available) and found that the UV sensitivity in FA individual 1333 was milder than in XPF individual XP2YO (Fig 2B). In addition, the FA-specific XPF mutant proteins XPF-L230P and XPF-R689S rescue 100% of the UVC sensitivity of XP2YO fibroblasts (Supplemental Fig. 2A) and approximately 80% of the UVC sensitivity of *Ercc4/Xpf* null MEFs (Supplemental Fig. 2B), while both mutants were unable to complement MMC sensitivity (Fig. 1F).

Furthermore, XFE and 1333 fibroblasts responded typically like FA cells on MMC-induced survival (Fig. 2C), DEB-induced chromosome breakage (Fig. 2D) and MMC-induced G2 phase arrest (Fig. 2E), while XPF cells showed milder MMC sensitivity and lacked DEB-induced chromosome fragility or MMC-induced cell cycle arrest (Fig. 2C-E). Previous experiments in CHO cells also demonstrated that the XFE-specific XPF mutant protein R153P does not rescue MMC or UV sensitivity[8]. Therefore we conclude that XP, XFE and FA cells with *ERCC4/XPF* mutations clearly have a distinct response on UV and MMC (Supplemental Table 2).

To further investigate the extent of NER deficiency in the FA individuals, we measured UV-induced unscheduled DNA synthesis (UDS) in primary skin fibroblasts from individual 1333 and from an XPF individual with mild clinical UV sensitivity (XP42RO) and found 24±4% and 21±3% residual UDS activity, respectively (Fig. 3A). We also determined UDS in *Ercc4/Xpf* null MEFs expressing the FA-specific XPF missense mutant proteins, XPF-L230P or XPF-R689S. The levels of UDS activity were 39,7% and 48,4% for XPF-L230P or XPF-R689S, respectively (Fig. 3B), enough to complement 80% of UVC-sensitivity of these MEFs (Supplemental Fig. 2B). In XPF-deficient human XP2YO fibroblasts, XPF-L230P and XPF-R689S rather efficiently corrected the defective removal of 6-4 photoproducts at sites of local UV damage (Fig. 3C). In contrast, XP2YO cells expressing the *ERCC4/XPF* mutant with the 28 bp duplication was completely deficient in NER activity as predicted from the disruption of the ERCC1 and DNA-binding domain of this truncated protein. The studies presented in Fig. 2 and 3 demonstrate that FA cells with *ERCC4/XPF* mutations are fully deficient in ICL repair

but retain significant levels of NER activity.

Cell lines from XPF individuals show a characteristic failure of the mutant XPF protein to properly translocate to the nucleus, through aggregation of the protein in the cytoplasm[8]. This feature is evident for XP-causing mutations and accentuated in cells from the individual with the XFE syndrome. However, FA-causing XPF missense mutant proteins can actually translocate to the nucleus where they are recruited to sites of active NER (Supplemental Fig.3A,B) and can interact with SLX4 and ERCC1 (Supplemental Fig.3C, D). This result may be functionally important, as a recent article reports that SLX4 interaction with XPF is crucial for interstrand crosslink repair and that SLX4 KO mice phenocopy FA[9]. Using Xenopus extracts, J.C. Walter's group reported that the FA upstream genes are required to regulate a nuclease that makes DNA incisions near the interstrand crosslink[10]. Given that FA-specific mutant XPF proteins can reach the site of damage, we then investigated their ability to cleave DNA. For this aim, XPF-R689S protein was purified as a heterodimer with ERCC1 as previously described[11]. Subsequently, NER reactions were performed with the purified mutant protein, extracts from XPF-deficient XP2YO cells and a plasmid containing an NER substrate (1,3-cisplatin intrastrand cross-link)[12]. Consistent with the functional data above, purified ERCC1-XPF-R689S is proficient in the excision step of NER similar to wild type XPF, as it restored the ability to cleave and remove the site-specific intrastrand cross-link from the plasmid in XP2YO cell extracts (Fig. 4A). Nevertheless, the excision reaction is not perfect as the excised fragments are, on average, 1 nucleotide longer than expected from a normal reaction with wild-type XPF-ERCC1 dimer (Fig.4A, lane 4). We also

performed *in vitro* nuclease activity assays with purified ERCC1-XPF-R689S and ERCC1-XPF-R689A on a stem-loop model DNA substrate. Unlike wild type XPF and XPF mutants causing XP (R799W) or XFE progeria (R153P)[8], XPF-R689S is unable to cleave such substrate (Fig. 4B), indicating that the nuclease-type activity of XPF-R689S is grossly abnormal. Unfortunately, we could not perform these biochemical experiments with the XPF-L230P mutant, as we were unable to express and purify ERCC1-XPF-L230P due to its low stability and tendency to aggregate. We finally checked whether the FA-specific mutant XPF proteins ectopically expressed in *Ercc4/Xpf* null MEFs can perform the incision step of ICL repair. Both XPF-L230P and XPF-R689S completely restored the incision defect of *Ercc4/Xpf* null MEFs as measured by the COMET assay (data not shown), while the cells remained hypersensitive to ICL (Fig 1F). Although additional biochemical experiments are required, our results suggest that ICL-sensitivity of FA104 and 1333 is not directly linked to absence of XPF nuclease activity. It seems unlikely that the defect is a downstream step of homologous recombination since FA104 and 1333 cells are not sensitive to PARP inhibitors and are normal in Rad51 focus formation. Given that the nuclease activity of the FA-specific XPF-R689S is grossly abnormal, it is tempting to speculate that the ICL-unhooking step in these FA cells leaves an intermediate aberrant substrate which is irreparable by subsequent ICL repair factors.

Our genetic, biochemical and functional studies, along with the characterization of previous XPF and XFE alleles, provide a model for the mechanistic understanding of how mutations in *ERCC4/XPF* lead to three distinct diseases (Supplemental Table 2). Most of the presently known XPF individuals suffer from a relatively mild form of XP[13].

Cells from these individuals have a reduced level of XPF protein in the nucleus as the mutant XPF protein has the tendency to aggregate in the cytoplasm[8]. This reduced level of nuclear XPF is insufficient to mediate complete NER, while it still has enough ICL repair-specific functions to prevent chromosome fragility, cell cycle arrest and subsequent FA clinical manifestations. A second set of *ERCC4/XPF* mutations, characterized in this study, allow localization of the protein to the nucleus, where they exert a certain level of NER activity but are fully deficient in ICL repair. XPF-R689S is a stable and NER-proficient protein with an active site structure that prevents it from properly processing ICL repair intermediates. XPF-L230P is more similar to previously described *ERCC4/XPF* mutations in that it is less stable and might have a tendency to aggregate in the cytoplasm. However, sufficient amounts of the protein are properly folded and reach chromatin where it appears to have some activity in the removal of 6-4PPs. Residual NER activity in the skin tissue of individual 1333 *in vivo* may explain why this individual has no clinically relevant skin photosensitivity although we cannot exclude that dermatological problems will arise later in life. A final category of *ERCC4*/*XPF* mutations is associated with the progeroid syndrome XFE, which is characterized by very low levels of nuclear XPF, apparently insufficient to support either NER or ICL repair. Importantly, the only XFE individual described suffered from both skin photosensitivity and anemia[7; 13], and shared some cellular features with XP (NER defect, UV-sensitivity) and FA (extreme ICL-sensitivity, DEB-induced chromosome fragility and MMC-induced cell cycle arrest) suggesting that XFE is characterized by a combination of XP and FA manifestations (Supplemental Table 2). Exhaustion of hematopoietic stem cells is also an attribute of ERCC1-XPF hypomorphic mice that

mimic XFE (Laura Niedernhofer, personal communication). Microsomy, microcephaly and liver fibrosis are likewise observed in FA individual 1333, in *Ercc1*- and *Ercc4/Xpf*-deficient mice and in the unique *ERCC1*-deficient individual, which all lack ICL repair functions[14-18]

In a broader sense, this study demonstrates that depending on the type of *ERCC4/XPF* mutation and the balance between NER and ICL repair activities, affected individuals present with one of three clinically distinct disorders. This resembles the *XPD* gene, which is involved in XP subtype D, trichothiodystrophy (MIM 601675) or Cockayne syndrome (MIM 216400) depending on the type of mutation[19], and highlights the value of characterizing rare genetic disorders to gain insight into the mechanisms of genome maintenance and human disease. XPF/FANCQ has a central role in preventing genome instability, cancer, BMF, developmental abnormalities and premature ageing. Like other breast and ovarian cancer susceptibility genes mutated in FA[20; 21], the product of *ERCC4/XPF*/*FANCQ* also acts downstream FANCD2 monoubiquitination. Therefore, it is important to study *FANCQ* as a candidate gene in hereditary breast and ovarian cancer.

Supplemental data

Supplemental file includes three figures and two tables.

Acknowledgements

The use of *FANCQ* as an alias for the *ERCC4/XPF* gene was approved by the HUGO Gene Nomenclature Committee. This study was approved by the Institutional Committee on Ethical Research in Human Samples and proper informed consent was obtained. We would like to thank Fanconi anemia families and their clinicians for providing samples and clinical data and Dr María A. Blasco (Centro Nacional de Investigaciones Oncológicas, Madrid) for providing *Ercc4/Xpf* deficient MEFs. JAB laboratory is funded by grants from the European Program "7FWP, Health" (PERSIST; Agreement no: 222878), the Spanish Ministry of Science and Innovation (Refs110-90.1 and SAF 2009-07164), Programa RETICS-RD06/0010/0015 ISCIII and Fundación Botín. ODS acknowledges funding from the National Institutes of Health (GM080454 and CA092584). CS is funded by CCA/V-ICI Amsterdam. DS and BS received grants from the Deutsche Fanconi-Anaemie-Hilfe, Aktionskreis Fanconi-Anaemie and the Schroeder-Kurth Fund. JS's laboratory is funded by the Generalitat de Catalunya (SGR0489-2009), the ICREA-Academia award, the Spanish Ministry of Science and Innovation (CIBERER CB06/07/0023, SAF2009-11936 and SAF2012-31881), and the European Regional Development FEDER Funds. CIBERER is an initiative of the Instituto de Salud Carlos III, Spain.

Web resources section

URL for Online Mendelian Inheritance in Man: http://www.omim.org.

References

1. Stoepker, C., Hain, K., Schuster, B., Hilhorst-Hofstee, Y., Rooimans, M.A., Steltenpool, J., Oostra, A.B., Eirich, K., Korthof, E.T., Nieuwint, A.W., et al. (2011). SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. Nature genetics 43, 138-141.
2. Kim, Y., Lach, F.P., Desetty, R., Hanenberg, H., Auerbach, A.D., and Smogorzewska, A. (2011). Mutations of the SLX4 gene in Fanconi anemia. Nature genetics 43, 142-146.
3. Antonio Casado, J., Callen, E., Jacome, A., Rio, P., Castella, M., Lobitz, S., Ferro, T., Munoz, A., Sevilla, J., Cantalejo, A., et al. (2007). A comprehensive strategy for the subtyping of patients with Fanconi anaemia: conclusions from the Spanish Fanconi Anemia Research Network. Journal of medical genetics 44, 241-249.
4. Deans, A.J., and West, S.C. (2009). FANCM connects the genome instability disorders Bloom's Syndrome and Fanconi Anemia. Molecular cell 36, 943-953.
5. de Laat, W.L., Sijbers, A.M., Odijk, H., Jaspers, N.G., and Hoeijmakers, J.H. (1998). Mapping of interaction domains between human repair proteins ERCC1 and XPF. Nucleic acids research 26, 4146-4152.
6. Sijbers, A.M., de Laat, W.L., Ariza, R.R., Biggerstaff, M., Wei, Y.F., Moggs, J.G., Carter, K.C., Shell, B.K., Evans, E., de Jong, M.C., et al. (1996). Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. Cell 86, 811-822.
7. Niedernhofer, L.J., Garinis, G.A., Raams, A., Lalai, A.S., Robinson, A.R., Appeldoorn, E., Odijk, H., Oostendorp, R., Ahmad, A., van Leeuwen, W., et al. (2006). A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. Nature 444, 1038-1043.
8. Ahmad, A., Enzlin, J.H., Bhagwat, N.R., Wijgers, N., Raams, A., Appledoorn, E., Theil, A.F., JH, J.H., Vermeulen, W., NG, J.J., et al. (2010). Mislocalization of XPF-ERCC1 nuclease contributes to reduced DNA repair in XP-F patients. PLoS genetics 6, e1000871.
9. Crossan, G.P., van der Weyden, L., Rosado, I.V., Langevin, F., Gaillard, P.H., McIntyre, R.E., Gallagher, F., Kettunen, M.I., Lewis, D.Y., Brindle, K., et al. (2011). Disruption of mouse Slx4, a regulator of structure-specific nucleases, phenocopies Fanconi anemia. Nature genetics 43, 147-152.
10. Knipscheer, P., Raschle, M., Smogorzewska, A., Enoiu, M., Ho, T.V., Scharer, O.D., Elledge, S.J., and Walter, J.C. (2009). The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science 326, 1698-1701.
11. Enzlin, J.H., and Scharer, O.D. (2002). The active site of the DNA repair endonuclease XPF-ERCC1 forms a highly conserved nuclease motif. The EMBO journal 21, 2045-2053.
12. Moggs, J.G., Yarema, K.J., Essigmann, J.M., and Wood, R.D. (1996). Analysis of incision sites produced by human cell extracts and purified

proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct. The Journal of biological chemistry 271, 7177-7186.
13. Gregg, S.Q., Robinson, A.R., and Niedernhofer, L.J. (2011). Physiological consequences of defects in ERCC1-XPF DNA repair endonuclease. DNA repair 10, 781-791.
14. Weeda, G., Donker, I., de Wit, J., Morreau, H., Janssens, R., Vissers, C.J., Nigg, A., van Steeg, H., Bootsma, D., and Hoeijmakers, J.H. (1997). Disruption of mouse ERCC1 results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. Current biology : CB 7, 427-439.
15. McWhir, J., Selfridge, J., Harrison, D.J., Squires, S., and Melton, D.W. (1993). Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. Nature genetics 5, 217-224.
16. Tian, M., Shinkura, R., Shinkura, N., and Alt, F.W. (2004). Growth retardation, early death, and DNA repair defects in mice deficient for the nucleotide excision repair enzyme XPF. Molecular and cellular biology 24, 1200-1205.
17. Jaspers, N.G., Raams, A., Silengo, M.C., Wijgers, N., Niedernhofer, L.J., Robinson, A.R., Giglia-Mari, G., Hoogstraten, D., Kleijer, W.J., Hoeijmakers, J.H., et al. (2007). First reported patient with human ERCC1 deficiency has cerebro-oculo-facio-skeletal syndrome with a mild defect in nucleotide excision repair and severe developmental failure. American journal of human genetics 80, 457-466.
18. Gregg, S.Q., Gutierrez, V., Robinson, A.R., Woodell, T., Nakao, A., Ross, M.A., Michalopoulos, G.K., Rigatti, L., Rothermel, C.E., Kamileri, I., et al. (2012). A mouse model of accelerated liver aging caused by a defect in DNA repair. Hepatology 55, 609-621.
19. Cleaver, J.E., Lam, E.T., and Revet, I. (2009). Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. Nature reviews Genetics 10, 756-768.
20. Meindl, A., Hellebrand, H., Wiek, C., Erven, V., Wappenschmidt, B., Niederacher, D., Freund, M., Lichtner, P., Hartmann, L., Schaal, H., et al. (2010). Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. Nature genetics 42, 410-414.
21. Levy-Lahad, E. (2010). Fanconi anemia and breast cancer susceptibility meet again. Nature genetics 42, 368-369.
22. Almarza, E., Rio, P., Meza, N.W., Aldea, M., Agirre, X., Guenechea, G., Segovia, J.C., and Bueren, J.A. (2007). Characteristics of lentiviral vectors harboring the proximal promoter of the vav proto-oncogene: a weak and efficient promoter for gene therapy. Molecular therapy : the journal of the American Society of Gene Therapy 15, 1487-1494.
23. Trujillo, J.P., Mina, L.B., Pujol, R., Bogliolo, M., Andrieux, J., Holder, M., Schuster, B., Schindler, D., and Surralles, J. (2012). On the role of FAN1 in Fanconi anemia. Blood 120, 86-89.

24. Limsirichaikul, S., Niimi, A., Fawcett, H., Lehmann, A., Yamashita, S., and Ogi, T. (2009). A rapid non-radioactive technique for measurement of repair synthesis in primary human fibroblasts by incorporation of ethynyl deoxyuridine (EdU). Nucleic acids research 37, e31.

Legends to figures

Figure 1: *ERCC4/XPF* mutations and XPF-deficiency in Fanconi anemia individuals. (**A**) Sequence analysis of blood DNA from FA104 revealed a missense mutation in exon 11 (c.2065C>A; p. (Arg689Ser)) (upper panel) and a 5 bp deletion in exon 8 leading to a frameshift and premature termination of translation (c.1484_1488delCTCAA; p. (Thr495Asnfs*6)) (lower panel). (**B**) Sequence analysis of blood DNA from 1333 revealed a missense mutation in exon 4 (c.689T>C; p.L230P; upper panel) and a 28 bp duplication in exon 11 (lower panel) leading to a frameshift and a premature stop codon (c.2371_2398dup28; p.Ile800Thrfs*24). (**C**) Western blot analysis showing XPF expression in lymphoblasts from 1333 and FA104. Lymphoblasts from a healthy individual (Con), the parents of 1333 (1333-F, 1333-M) and an unrelated *ERCC4/XPF* mutation carrier (Het) were used as controls. XPF levels are expressed as ratio relative to the loading control (Rad50). (**D**) Genetic complementation of MMC sensitivity in FA104 lymphoblasts by wild type XPF but not by XPF mutant R689S. Site-directed mutagenesis was used to introduce point mutations in pFastBac1-XPF HA tagged plasmid using the QuickChange method (Stratagene) as described[11]. Lentiviral supernatants production and transduction was done as previously described[22] and cells were grown for 10 days in presence of MMC. Data represent a typical result of at least three independent experiments. (**E**) Genetic complementation of MMC sensitivity of 1333 lymphoblasts by wild type XPF (experiments were performed as in D). (**F**) MMC-induced growth inhibition of *Ercc4/Xpf* KO MEFs transduced with lentiviral particles coding for GFP (negative control vector), wild type XPF, XPF-R689S and XPF-L230P. Data represent means and SD of at least three independent experiments.

Figure 2: UV and ICL sensitivities of *ERCC4/XPF* mutants leading to Fanconi anemia. (**A**) UVC-induced apoptosis in FA104 lymphoblasts. Cells were analyzed for UVC-induced apoptosis 24h post irradiation using the Annexin-V-FLUOS Staining Kit (Roche). Data represent means and SD of at least three independent experiments. (**B**) UVC-induced growth inhibition of human *ERCC4/XPF*-deficient immortal fibroblast cell lines from XP and FA individuals (XP2YO and 1333, respectively) transduced with lentiviral particles carrying a cDNA coding for XPF-WT. The results are expressed as percentage of viable UVC treated cells relative to that of untreated controls. Data represent means and SD of two independent experiments. (**C**) MMC sensitivity of human *ERCC4/XPF*-deficient primary fibroblasts from XP, FA and XFE individuals (XP42RO, 1333 and XP51RO, respectively). Data represent means and SD of two independent experiments. (**D**) DEB-induced chromosome fragility test in human *ERCC4/XPF*-deficient primary fibroblasts from XP, FA and XFE individuals (XP42RO, 1333 and

XP51RO, respectively). (**E**) MMC-induced G2/M cell cycle arrest in the same cells as in (D). Experiments presented in D, E and F were performed as reported earlier[23].

Figure 3: NER analysis of *ERCC4/XPF* mutants *in vivo*. (**A**) Unscheduled DNA synthesis (UDS) in primary fibroblasts representing global NER activity was measured using 5-ethynyl-deoxyuridine, grossly as previously described[24]. Cells (arrows) were compared to mixed-in normal fibroblasts preloaded with polystyrene microbeads (no arrows), used as an internal control. UDS signal was quantified from 20-40 random G1/G2 nuclei and expressed as a percentage of control wt cells (**B**) UDS signals in *Ercc4/Xpf-/-* MEFs measured as in (A), expressed as a percentage of control wt MEFs. *Ercc4/Xpf-/-* cells were stably expressing an empty vector or one of various *ERCC4/XPF* cDNAs (wild type, L230P and R689S). (**C**) Repair kinetics of UV-induced DNA damage by FA-specific *ERCC4/XPF* mutants in *ERCC4/XPF*- and NER-deficient human cells (XP2YO). Cells expressing wild type XPF, XPF-R689S, XPF-L230P or XPF-28 bp duplication were locally UV-irradiated, cultured for the indicated times, fixed and stained for (6-4)PPs and HA-tag of XPF using specific antibodies. Data represent the percentage of cells with (6-4)PP spots at various time points; means and SD of at least two independent experiments are shown. For each experiment 100 cells were counted.

Figure 4: Nuclease activity of XPF mutants. (**A**) NER activity of wild type and mutant ERCC1-XPF dimer. A plasmid containing a site-specific 1,3-intrastrand cis-Pt DNA cross-link was incubated with whole cell extracts from HeLa cells or XPF-deficient cells (XP2YO) complemented with recombinant ERCC1-XPF purified from Sf9 insect cells as reported[9]. The excised DNA fragments of 24-32 nucleotides are shown. The position of a 25mer is indicated. (**B**) Incision of a stem-loop substrate with wild type and mutant XPF. The 3' Cy5-labeled substrate was incubated with recombinant ERCC1-XPF in the presence of 2 mM $MgCl_2$ or 0.4 mM $MnCl_2$, and the products analyzed by denaturing PAGE. The incision reaction was performed essentially as described earlier[8; 11]

Figure 1

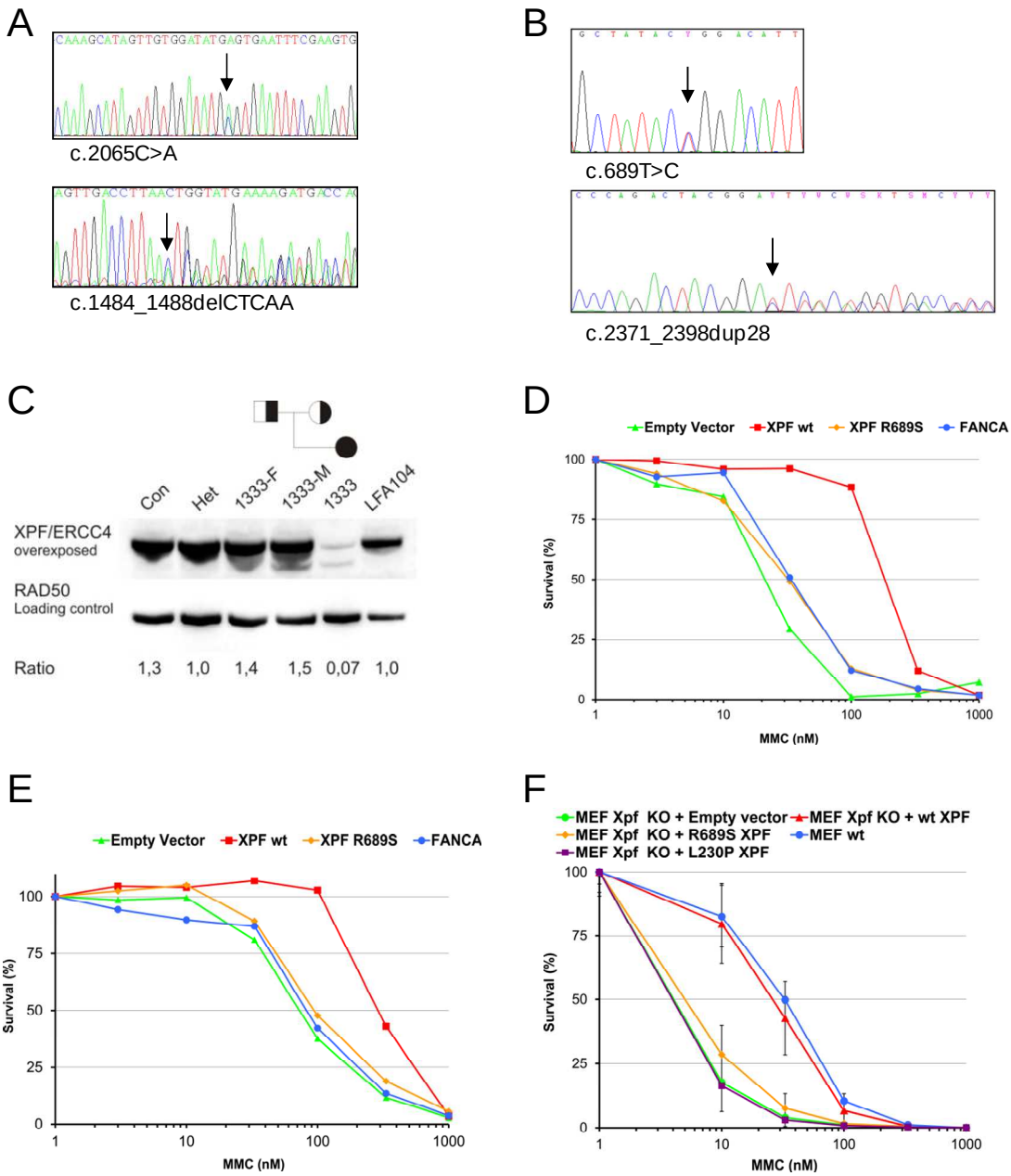

A
c.2065C>A
c.1484_1488delCTCAA
B
c.689T>C
c.2371_2398dup28
C
Con
Het
1333-F
1333-M
1333
LFA104
XPF/ERCC4
overexposed
RAD50
Loading control
Ratio
1,3
1,0
1,4
1,5
0,07
1,0
D
Empty Vector
XPF wt
XPF R689S
FANCA
Survival (%)
MMC (nM)
E
Empty Vector
XPF wt
XPF R689S
FANCA
Survival (%)
MMC (nM)
F
MEF Xpf KO + Empty vector
MEF Xpf KO + wt XPF
MEF Xpf KO + R689S XPF
MEF wt
MEF Xpf KO + L230P XPF
Survival (%)
MMC (nM)

Figure 2

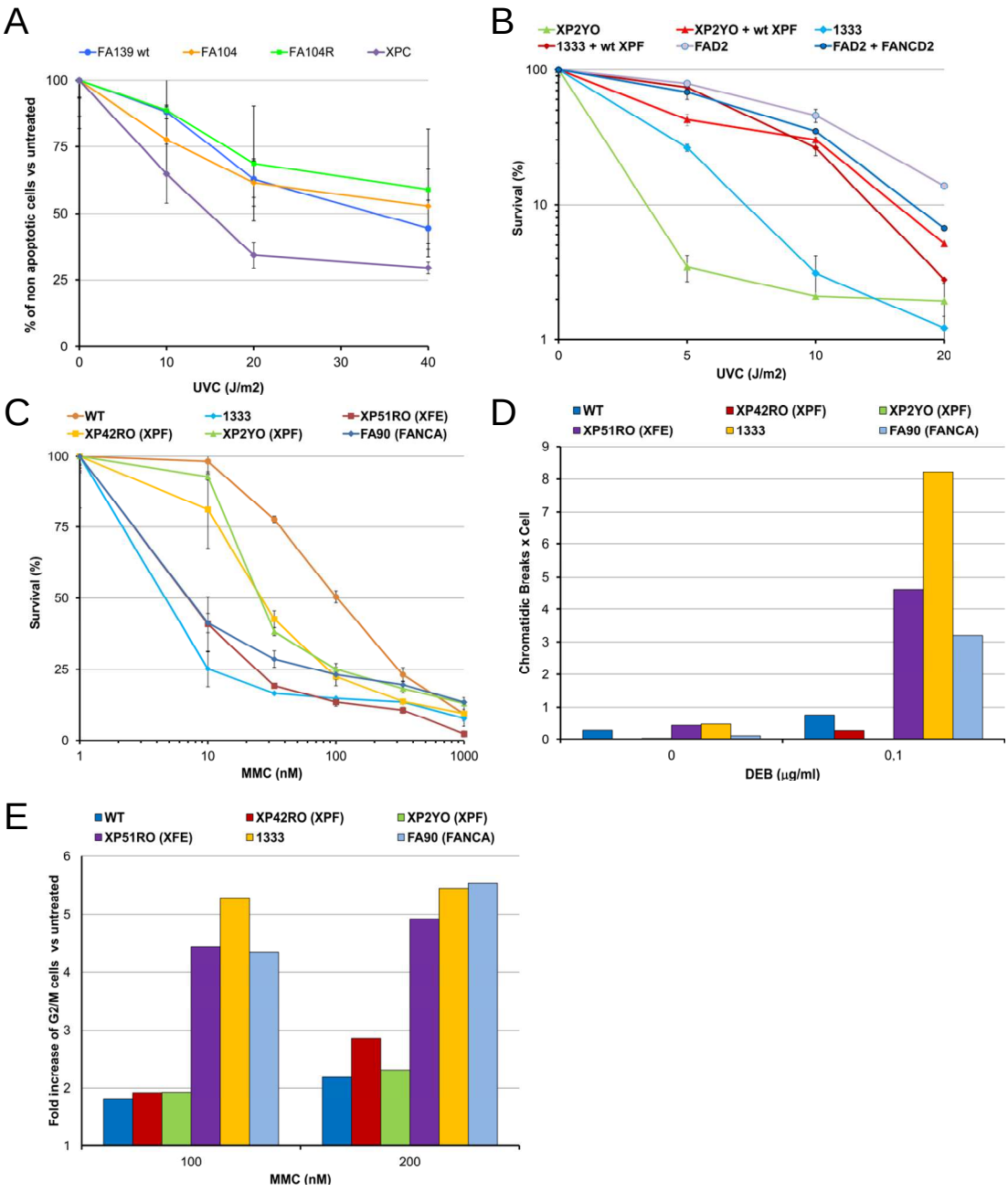

A
FA139 wt
FA104
FA104R
XPC
% of non apoptotic cells vs untreated
UVC (J/m2)
B
XP2YO
XP2YO + wt XPF
1333
1333 + wt XPF
FAD2
FAD2 + FANCD2
Survival (%)
UVC (J/m2)
C
WT
1333
XP51RO (XFE)
XP42RO (XPF)
XP2YO (XPF)
FA90 (FANCA)
Survival (%)
MMC (nM)
D
WT
XP42RO (XPF)
XP2YO (XPF)
XP51RO (XFE)
1333
FA90 (FANCA)
Chromatidic Breaks x Cell
DEB (µg/ml)
E
WT
XP42RO (XPF)
XP2YO (XPF)
XP51RO (XFE)
1333
FA90 (FANCA)
Fold increase of G2/M cells vs untreated
MMC (nM)

Figure 3

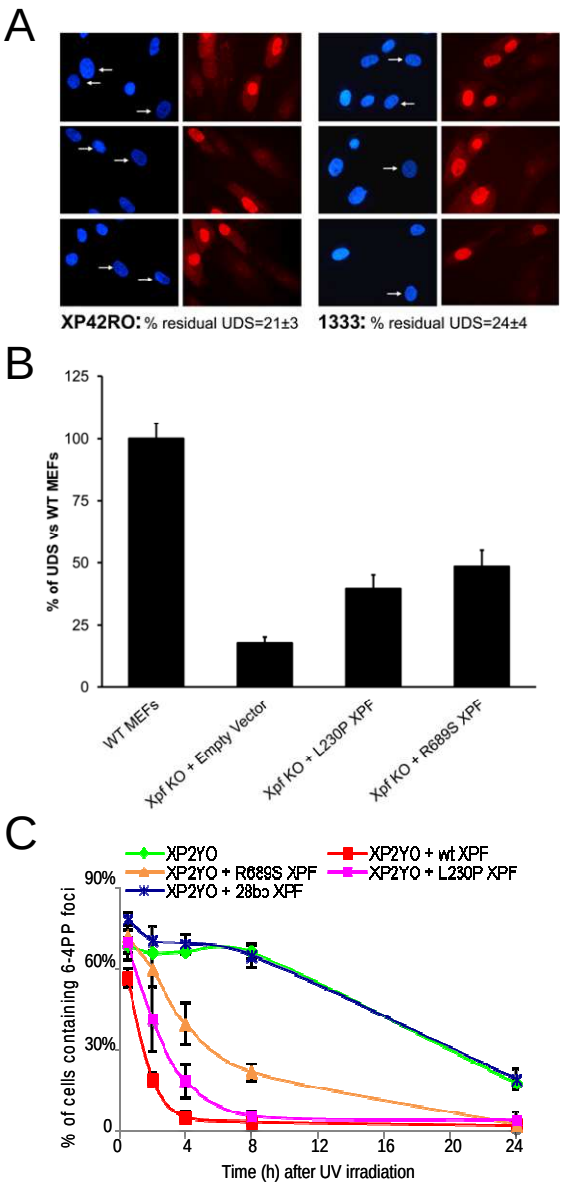
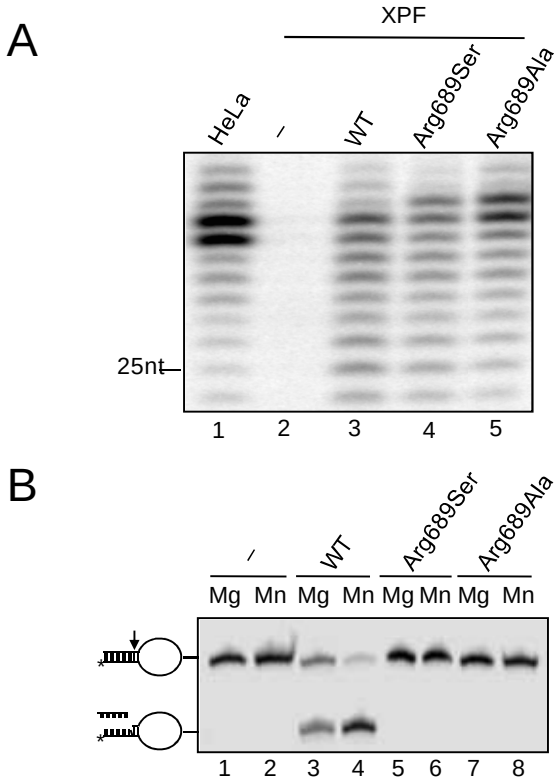

A
XP42RO: % residual UDS=21±3
1333: % residual UDS=24±4
B
% of UDS vs WT MEFs
125
100
75
50
25
0
WT MEFs
Xpf KO + Empty Vector
Xpf KO + L230P XPF
Xpf KO + R689S XPF
C
XP2YO
XP2YO + R689S XPF
XP2YO + 28bp XPF
XP2YO + wt XPF
XP2YO + L230P XPF
% of cells containing 6-4PP foci
90%
60%
30%
0
0
4
8
12
16
20
24
Time (h) after UV irradiation

Figure 4

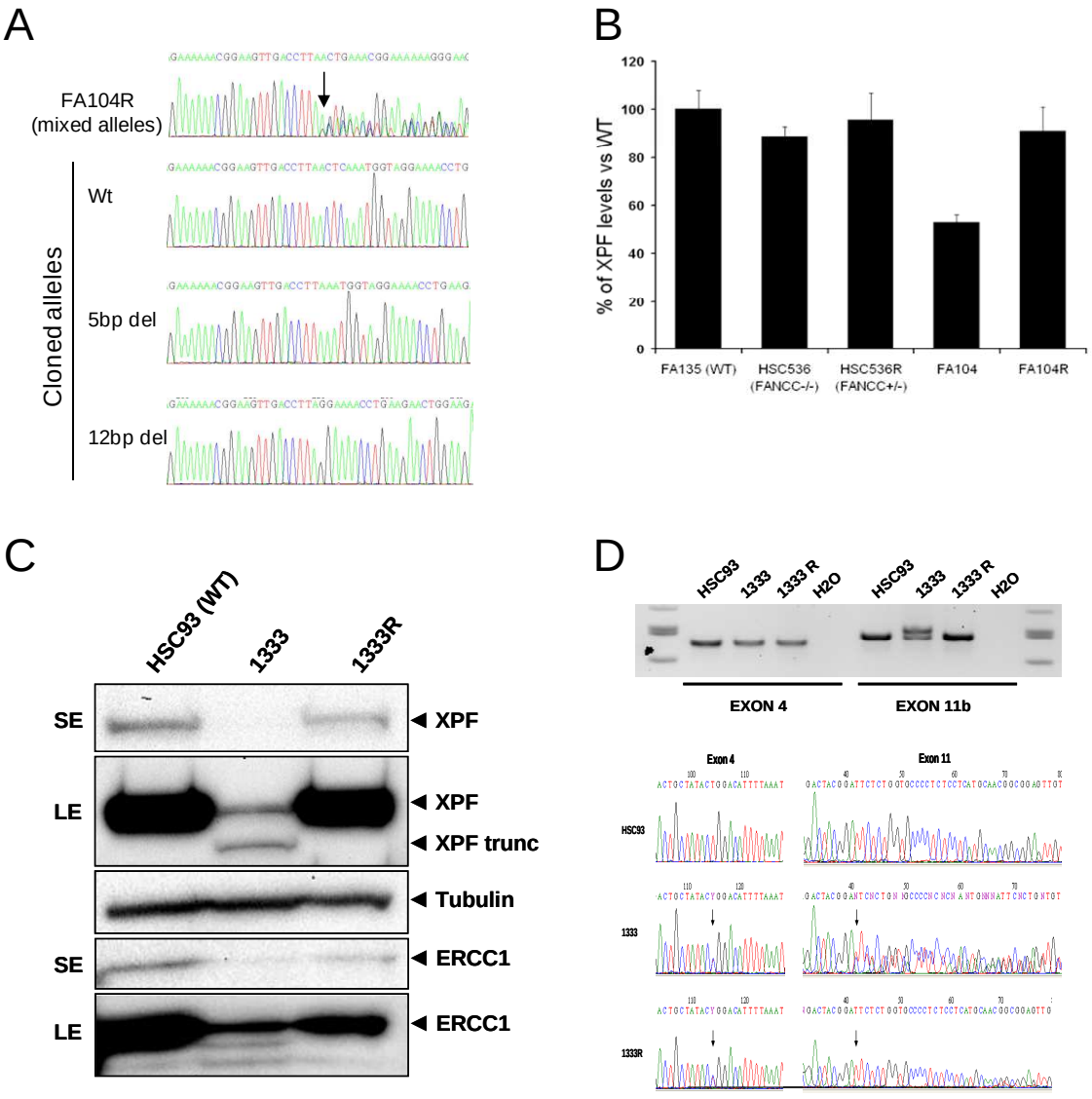

A
XPF
HeLa
–
WT
Arg689Ser
Arg689Ala
25nt
1 2 3 4 5
B
–
WT
Arg689Ser
Arg689Ala
Mg Mn Mg Mn Mg Mn Mg Mn
1 2 3 4 5 6 7 8

Supplemental data

Mutations of the DNA repair endonuclease ERCC4/XPF cause Fanconi anemia

Massimo Bogliolo, Beatrice Schuster, Chantal Stoepker, Burak Derkunt, Yan Su, Anja Raams, Juan P. Trujillo, Jordi Minguillón, María J. Ramírez, Roser Pujol, José A. Casado, Rocío Baños, Paula Rio, Kerstin Knies, Sheila Zúñiga, Javier Benítez, Juan A. Bueren, Nicolaas G.J. Jaspers, Orlando D. Schärer, Johan P. de Winter, Detlev Schindler and Jordi Surrallés

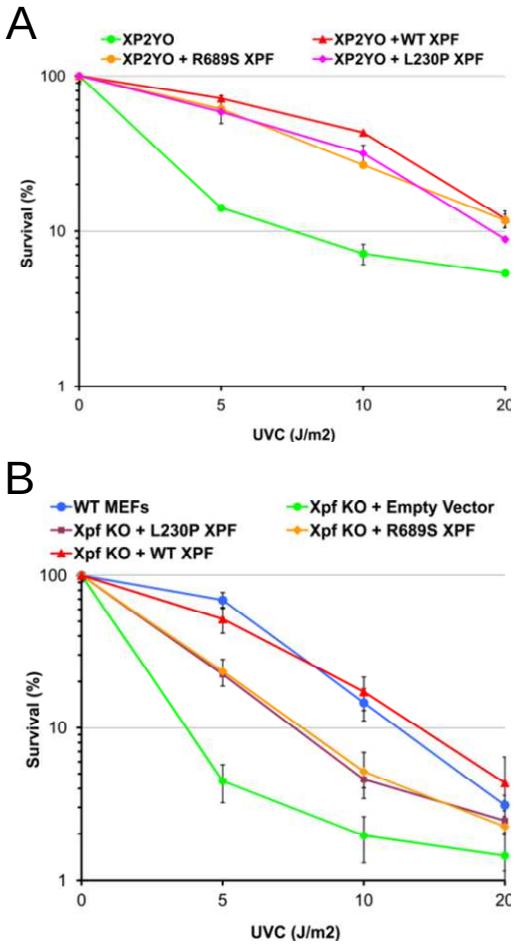
Supplemental figure 1



Legend: Genetic analysis of back mutations in *ERCC4/XPF* in reverted MMC-resistant FA cell lines. (**A**) Sequence analysis of individual exon 8 alleles cloned from the FA104R cell line. Exons 8 was amplified from FA104R DNA and the PCR products were cloned with the Topo TA Cloning kit (Invitrogen) and transfected into Library Efficiency DH5alpha Competent Cells (Invitrogen). The plasmids from single bacterial colonies were prepared with the NucleoSpin® Plasmid QuickPure Kit (Macherey-Nagel). Sequencing of single bacterial clones revealed the presence of a 12 bp deletion in exon 8 encompassing the pathogenic 5 bp deletion and restoring the reading frame of the *ERCC4/XPF* gene. (**B**) Quantification of XPF expression by Western blot in lymphoblasts from FA104, FA104R, HSC536 (FA-C), HSC536R (HSC536 reverted to

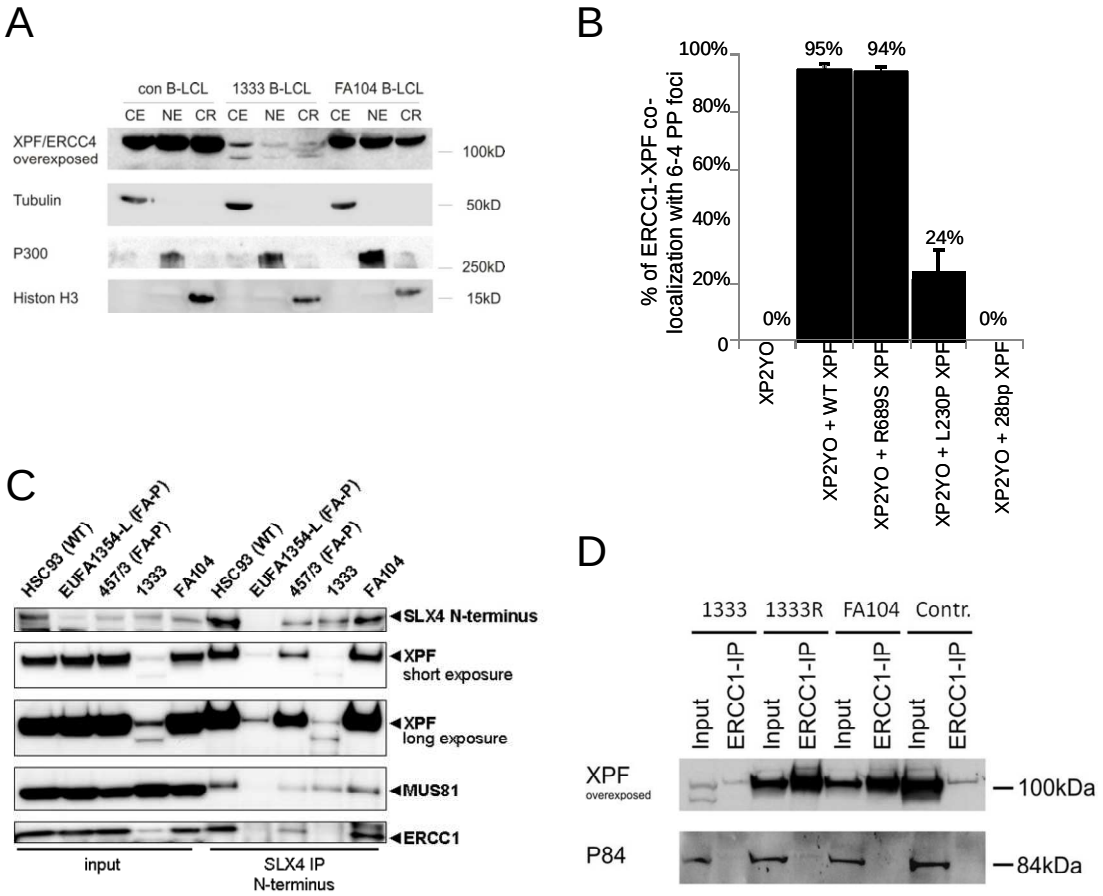
wt) and FA139 (wt). XPF levels are expressed as a ratio of the loading control (vinculin). Further details on antibodies used can be obtained upon request. The histogram represents XPF levels in the different cell lines normalized to the levels of the loading control. Means and SEM of at least three independent experiments are shown. (**C**) Western blot analysis showing low levels of two XPF proteins in 1333 and a normal size XPF protein in the reverted cell line 1333R. (**D**) Absence of the 28 bp duplication in *ERCC4/XPF* exon 11 in 1333R eliminating the longer *ERCC4/XPF* mutant allele with the 28 bp duplication (upper panel) and restoring the wt sequence in exon 11 (lower panel).

Supplemental figure 2



Legend: XPF mutants leading to FA restore UVC resistance of NER deficient human and mouse fibroblast. (**A**) UVC-induced growth inhibition of human XPF-deficient immortal cell lines from XP and FA individuals (XP2YO and 1333, respectively) transduced with lentiviral particles carrying cDNAs coding for XPF-WT, XPF-R689S and XPF-L230P. Data represent means and SD of two independent experiments. (**B**) UV-induced growth inhibition of *Ercc4/Xpf* KO MEFs transduced with lentiviral particles expressing GFP (negative control vector), wild type XPF, XPF-R689S and XPF-L230P. Data represent means and SD of at least four independent experiments. Lentivirus mediated cDNA transduction and survival analysis were performed as shown in Fig. 1 and 2 (main text).

Supplemental figure 3



Legend: XPF relocation to DNA damage and protein-protein interactions. (**A**) Western blot analysis of XPF protein in cytoplasmic (CE), nucleoplasmic (NE) and chromatin (CR) fractions from wild type, 1333 and FA104 lymphoblast cell lines. FA104 and FA104 show an abundance and a distribution of XPF between the cytoplasmic, nuclear and chromatin compartments comparable to a normal control, whereas 1333 reveals reduced abundance and two species of that protein with sizes predicted by its mutations but, of note, XPF is still detected in the nucleus and on chromatin with grossly unaffected ratios to the cytoplasmic fraction. Specificity of the separation of extracts from lymphoblasts is confirmed by the compartment-specific marker proteins tubulin, p300 and histone H3. Subcellular Protein Fractionation Kit from Pierce (Thermo Scientific) following the manufacturer's instructions. Further details on antibodies used and subfractioning conditions can be optained upon request (**B**) Graphical representation of the percent co-localization of XPF with (6-4)PP in XP2YO cells expressing various forms of XPF. XP2YO cells were transduced with HA-tagged wild type XPF, XPF-R689S, XPF-L230P, or XPF-28bp dup, irradiated with UVC (120 J m^{-2}) through a polycarbonate filter with 5 μm pores, incubated for 0.5 h, fixed and stained with antibodies to (6-4)PP and antibodies to HA. Data represent the average of at least 3 independent experiments ± the SD. For each experiment 100 cells were counted. (**C**)

Normal SLX4 interactions in XPF-deficient FA individuals. SLX4 was immunoprecipitated with a polyclonal antibody raised against the first 300 amino acids of SLX4 (SLX4 N-terminus; gift from J. Rouse, Dundee). Precipitated proteins were visualized by Western blotting with antibodies to SLX4 N-terminus, XPF, ERCC1 and MUS81. Reduced XPF and ERCC1 protein expression was found in lymphoblasts of individual 1333. In these cells, full-length and truncated XPF and MUS81 were coprecipitated with SLX4, whereas ERCC1 is barely detectable. In lymphoblasts of individual FA104, the interaction between SLX4 and its binding partners XPF-ERCC1 and MUS81 is normal. Wild type lymphoblasts (HSC93) and lymphoblasts of FA-P individuals (EUFA1354-L and 457/3) were used as controls. (**D**) Normal ERCC1-XPF interactions in FA104 and 1333 lymphoblast cell lines. ERCC1 was immunoprecipitated with a polyclonal antibody against ERCC1 and the precipitated proteins were visualized by Western bloting with antibodies against XPF and P83 as internal control. Further technical details on co-immunoprecipitation conditions and antibodies used can be obtained upon request.

Supplemental Table 1. List of candidate genes with biallelic mutations after whole exome sequencing

Chrom	Pos	Ref	Alt	Ensembl pred	AA change	GN	NRR	SNV Q	GT Q
1	169489751	A	W	SS	-	F5	42	171	171
1	169525877	T	Y	SS	-	F5	52	36	36
2	73675227	-	CTC	NFC	S/SP	ALMS1	16	N/A	N/A
2	73678183	G	R	NSC	G1509D	ALMS1	156	120	120
3	49094490	G	S	NSC	N381K	QRICH1	122	228	228
3	49095011	C	S	NSC	G208R	QRICH1	109	43	43
4	126238305	C	M	NSC	P247T	FAT4	52	178	178
4	126355484	C	M	NSC	A2368E	FAT4	56	190	190
5	156479444	TTG	-	NFC	TS/S	HAVCR1	61	N/A	N/A
5	156479568	-	GTT	NFC	T/TT	HAVCR1	106	N/A	N/A
6	31238942	G	W	NSC	A176V	HLA-C	23	61	39
6	31239577	A	C	NSC	S48A	HLA-C	21	90	90
6	32709309	A	R	SS	-	HLA-DQA2	29	84	84
6	32713044	C	Y	NSC	T64M	HLA-DQA2	192	228	228
6	32713188	C	Y	SS	-	HLA-DQA2	126	228	228
6	38840915	A	R	NSC	I2479V	DNAH8	72	216	216
6	38879340	A	T	NSC	E3267D	DNAH8	12	34	34
7	100686777	C	Y	NSC	T4027M	MUC17	323	228	228
7	100687107	G	R	SS	-	MUC17	66	79	79
8	30700598	T	Y	NSC	N1979S	TEX15	33	97	97
8	30701995	A	M	NSC	D1513E	TEX15	141	228	228
10	69682773	T	Y	NSC	D920G	HERC4	64	69	69
10	69785435	-	A	SS	-	HERC4	9	N/A	N/A
16	14029271	AACTC	-	FC	-	ERCC4	22	N/A	N/A
16	14041518	C	M	NSC	R689S	ERCC4	121	228	228
16	72137553	C	S	NSC	Q564E	DHX38	56	85	85
16	72142141	A	R	NSC	S994G	DHX38	52	106	106

17	74272839	C	Y	NSC	V1593M	QRICH2	54	33	33
17	74277009	T	Y	NSC	Q1264R	QRICH2	23	81	81
18	14105016	C	M	NSC	R508I	ZNF519	136	228	228
18	14105853	C	M	NSC	R229I	ZNF519	23	51	51
19	51918360	A	R	NSC	S445P	SIGLEC12	43	39	39
19	52004795	G	CT	FC	-	SIGLEC12	19	N/A	N/A
X	53561632	A	W	NSC	F4226I	HUWE1	42	53	53
X	53642759	C	M	NSC	E665D	HUWE1	16	33	33

Chrom: chromosome number; Pos: genomic position (GRCh37/hg19);Ref: reference allele; Alt: sample allele; Ensembl pred: consequence prediction of variants on transcript according to Ensembl v59. This column contains one of the following values: SS=splice site, NSC=non-synonymous coding, FC=frameshift coding, NFC=non-frameshift coding; AA change: amino acid change in the affected protein; GN: Gene name; NRR: Number of non-redundant reads; SNV Q: the phred-scaled likelihood that the genotype is identical to the reference; GT Q: Phred-scaled likelihood that the genotype is wrong.

Supplemental Table 2: Comparative summary of clinical and cellular/molecular features of XP, XFE and FA individuals with mutations in *ERCC4/XPF*.

Clinical/cellular features	XPF	XFE	FA
Skin phosensitivity	mild	severe	no
Atrophic epidermis	variable	yes	no
Neurologic features	rare	yes	no
Hematology	normal	anemia[a]	anemia, BMF
Growth retardation[b]	no	yes	yes
Premature death	no	16yo	4yo (FA104). 1333 alive at age 10
UV sensitivity	mild	severe[c]	none (FA104)[d], mild (1333)
UDS defect	mild	severe[c]	mild (1333), ND in (FA104)[d]
MMC sensitivity	mild	severe	severe
DEB-test	negative	positive	positive
MMC induced G2/M arrest	negative	positive	positive
Nuclease activity on stem loop susbtrates	yes[e]	yes[e]	no

[a]It is not known whether anemia evolved to BMF in the XFE individual (Laura Niedernhofer, personal communication). [b]Include microsomy in 1333, FA104 and XFE and microcephaly in XFE and 1333. [c]Reported in Niedernhofer et al., 2005. [d]UDS assay was not done in FA104 due to the lack of skin fibroblasts but FA104 lymphoblasts were resistant to UV. [e]Reported in Ahmad et al., 2010 for XP mutation R799W and XFE mutation R153P. Typically XP and FA features are marked in yellow and green, respectively.