TITLE PAGE

TITLE: Evaluation of rare variants in the new Fanconi Anemia gene *ERCC4 (FANCQ)* as familial breast/ ovarian cancer susceptibility alleles.

AUTHORS:

Ana Osorio[1,2*], Massimo Bogliolo[2,3*], Victoria Fernández[1], Alicia Barroso[1], Miguel de la Hoya[4], Trinidad Caldés[4,] Adriana Lasa[5], Teresa Ramón y Cajal[6], Marta Santamariña[2,7] Ana Vega [2,8], Francisco Quiles[9], Conxi Lázaro[9], Orland Díez[10], Daniel Fernández[11], Rogelio González-Sarmiento[11], Mercedes Durán[12], José Fernández Piqueras[2,13], Maria Marín[3], Roser Pujol[2,3], Jordi Surrallés[2,3+], Javier Benítez[1,2+].

[1]Human Genetics Group, Human Cancer Genetics Programme, Spanish National Cancer Research Centre, CNIO, Madrid, Spain. [2]Centre for Biomedical Network Research on Rare Diseases (CIBERER). [3]Genome Instability and DNA Repair Group, Department of Genetics and Microbiology, Universitat Autonoma de Barcelona (UAB), Barcelona, Spain. [4]Molecular Oncology Laboratory. Hospital Clínico San Carlos. Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC). [5]Genetics Service, [6]Oncology Service, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain..[7] Grupo de Medicina Xenómica -USC, University of Santiago de Compostela, CIBERER, IDIS, Santiago de Compostela, Spain. [8] Fundación Pública Galega de Medicina Xenómica-SERGAS Santiago de Compostela, Spain.[9] Hereditary Cancer Program, Catalan Institute of Oncology (ICO), Hospital Duran i Reynals, Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet, Barcelona, Catalonia, Spain. [10] Oncogenetics Laboratory, Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Research Institute (VHIR), Universitat Autonoma de Barcelona, and University Hospital of Vall

d'Hebron, Barcelona, Spain. [11] Instituto de Biología Molecular y Celular del Cancer (IBMCC), laboratory 14. Universidad de Salamanca-CSIC, Spain. [12] Institute of Biology and Molecular Genetics. Universidad de Valladolid (IBGM-UVA), Valladolid, Spain. [13] Centro de Biología Molecular Severo Ochoa (CBM), Consejo Superior de Investigaciones Científicas-Universidad Autónoma de Madrid, Spain.

[*]Contributed equally to this work.

[+]Correspondence should be addressed to JB (jbenitez@cnio.es) or JS (jordi.surralles@uab.es)

Word count: 1042

Abstract:

It has been recently reported that bi-allelic mutations in the *ERCC4* (*FANCQ*) gene cause Fanconi Anemia (FA) subtype FA-Q. To investigate the possible role of *ERCC4* in breast and ovarian cancer susceptibility, as occurs with other FA genes, we screened the 11 coding exons and exon-intron boundaries of *ERCC4* in 1573 index cases from high risk Spanish familial breast and ovarian cancer pedigrees that had been tested negative for *BRCA1* and *BRCA2* mutations and 854 controls. The frequency of *ERCC4* mutation carriers does not differ between cases and controls, suggesting that *ERCC4* is not a cancer susceptibility gene. Interestingly, the prevalence of *ERCC4* mutation carriers (1 in 288) is similar to that reported for *FANCA* while there are approximately 100 fold more FA-A than FA-Q patients, indicating that most bi-allelic combinations of *ERCC4* mutations are embryo lethal. Finally, we identified additional bone fide FA *ERCC4* mutations specifically disrupting interstrand crosslink repair.

The *ERCC4* (*FANCQ*) gene (MIM 133520) encodes for a DNA repair endonuclease (XPF) that plays essential roles in nucleotide excision repair (NER) and interstrand crosslink repair (ICLR) (Gregg, et al., 2011). Bi-allelic mutations in *ERCC4* have been linked to Xeroderma Pigmetosum (XP [MIM 278700; 610651; 278720; 278730; 278740; 278760; 278780; 278750]) (Sijbers, et al., 1996) and progeria (XFE [MIM 610965]) (Niedernhofer, et al., 2006) diseases. Very recently, it has been published that bi-allelic mutations in the gene are also causative of disease in previously unclassified Fanconi Anemia (FA) patients subtype FA-Q and, therefore, the use of *FANCQ* as an alias for *ERCC4* was approved by the HUGO Gene Nomenclature Committee(Bogliolo, et al., 2013). Individuals harbouring mutations in *ERCC4* show high variability on their clinical manifestations, ranging from mild XP symptoms (sun sensitivity, freckling of the skin and basal or squamous cell carcinomas) to the dramatic accelerated aging symptoms of a progeroid syndrome, and finally, as recently shown (Bogliolo, et al., 2013) in FA patients, to malformations and progressive bone marrow failure. It has been also demonstrated that a balance between NER and ICLR activities due to different effects caused by distinct mutations in *ERCC4* determined the final clinical outcome (Bogliolo, et al., 2013). On the other hand, given that FA-Q patients are normal in FANCD2 monoubiquitination, it has been suggested that mono-allelic mutations in *ERCC4* could be linked to the breast and ovarian hereditary cancer syndrome as it occurs with other FA genes such as *BRCA2/FAND1* (Howlett, et al., 2002), *BRIP1/FANCJ* (Seal, et al., 2006), *PALB2/FANCN* (Rahman, et al., 2007) and *RAD51C/FANCO* (Meindl, et al., 2010; Osorio, et al., 2012) (MIM 300514, 610832, 614082, and 614083, respectively) (Levy-Lahad, 2010).

To investigate the possible role of *ERCC4* in breast and ovarian cancer susceptibility, we screened by DHPLC (Denaturing High Performance Liquid Chromatography) and direct sequencing the 11 coding exons and exon-intron boundaries of the *ERCC4* gene in 1573 index cases from high risk Spanish familial breast and ovarian cancer pedigrees that had been tested negative for mutations in *BRCA1* and *BRCA2* and 854 controls without personal or familial antecedents of cancer. Criteria for inclusion of cases and controls, and methods of screening for mutations in *BRCA1/2* have been previously published (Romero, et al., 2011; Fernandez-Rodriguez, et al., 2012; Osorio, et al., 2012; Bonache, et al., 2013) . We identified five and four unique variants among cases and controls among which three and four respectively were considered as putatively deleterious. Missense mutations were evaluated using the program CONDEL that predicts the pathogenicity of non synonymous variants using a consensus deleteriousness score that combines various tools such as SIFT, Polyphen2 and MutationAssessor (http://bg.upf.edu/condel/home). Those missense variants predicted to be deleterious by CONDEL were considered as predicted to affect protein function and this was studied later by functional analysis; those predicted to be neutral were considered as such and not further evaluated (Table 1). Regarding cases, one of the mutations found, c.584+1G>A in intron 3, was confirmed to cause skipping of exon 3 of the gene and a premature stop codon (PTC) (data not shown). The other two mutations, p.Arg150Cys in exon 3 and p.Ser786Phe in exon 11 were later functionally investigated. Regarding controls, we identified two different frameshift mutations c.540_541delAG in exon 3 (found in two independent controls) and c.2291delG in exon 11, both predicted to cause PTCs. Even though this later PTC is very C-terminal and could potentially result in a shorter but partially functional protein, this mutation results in a truncated XPF protein that lacks the double helix-hairpin-helix

(HhH2) domain involved in heterodimerization with ERCC1 and DNA binding (de Laat, et al., 1998), very similar to a pathogenic *ERCC4* mutation (c.2371_2398dup28; p.Ile800Thrfs*24) that functionally disrupts NER and ICLR activities (Bogliolo, et al., 2013). The last deleterious variant found in controls was the missense p.Arg689Ser in exon 11 previously found in a FA patient, and demonstrated to cause abnormal nuclease activity and to specifically disrupt ICLR (Bogliolo, et al., 2013). No differences were found regarding localization of mutations in the gene among cases and controls (Table 1). All variants reported have been submitted to the Leiden Open Variation Database (LOVD).

To evaluate the functional impact of the missense variants, we cloned a- HA-tagged WT *ERCC4* cDNA in a pBABE-puro retroviral vector (Addgene plasmid 14430, kindly shared by Dr. LM Martins) in IRES with the *GFP* cDNA, and the c.448C>T, c.2065C>A and c.2357C>T variants were introduced by site directed mutagenesis (Bogliolo, et al., 2013). The resulting constructs were transduced in NER and ICLR deficient *Ercc4* KO mouse embryonic fibroblasts (MEFs) and, after puromicin selection, the green cells were sorted to achieve a purity of over 98% by FACS (Bogliolo, et al., 2013). Due to the bicistronic nature of the IRES construct, we were able to assess the stability of the mutant XPF proteins using GFP as a reference, since both proteins are encoded by the same mRNA (Pelletier and Sonenberg, 1988). The p.Arg689Ser variant reduced by 40% the stability of XPF, while the p.Arg150Cys and p.Ser786Phe variants had no impact on protein stability (Fig. 1A and 1B). UVC sensitivity of Xpf KO MEFs was complemented with the expression of both p.Arg150Cys- and p.Ser786Phe-XPF (Fig. 1C), and only the Xpf KO MEFs expressing p.Ser786Phe-XPF or p.Arg689Ser showed a FA phenotype in terms of mitomycin C (MMC) sensitivity (Fig. 1D), MMC-induced cell cycle arrest at the G2/M phase (Fig. 1E) and DEB-induced chromosome fragility (Fig. 1F). These data

confirms that, resembling p.Arg689Ser, p.Ser786Phe specifically disrupts ICLR and, therefore, is a bone-fide FA mutation. Interestingly, both mutations are located in the nuclease domain of XPF. Despite a mild MMC sensitivity (Fig. 1D and 1E), *Ercc4* KO MEFs expressing p.Arg150Cys-XPF did not show DEB-induced chromosome fragility (Fig 1F). These data, together with the protein stability and UVC sensitivity data, indicate a null impact of the c.448C>T variant on XPF NER functions and a mild effect on ICLR activity.

In conclusion, the frequency of Spanish individuals heterozygous for pathogenic mutations in the *ERCC4* gene is approximately 0,3% and it does not differ between familial breast/ovarian cancer patients and healthy controls (p=0.251), suggesting that mono-allelic mutations in *ERCC4* are not linked to cancer susceptibility in the general population. Similar results were found with *SLX4* that, like *ERCC4*, acts downstream FANCD2 monoubiquitination but upstream the homologous recombination step of ICL (Fernandez-Rodriguez, et al., 2012). The prevalence of *ERCC4* mutation carriers (1 in 288) is similar to that reported for *FANCA*. However, there are approximately 100 fold more FA-A than FA-Q individuals, suggesting that over 90% of bi-allelic combinations of *ERCC4* mutations are embryo lethal in humans. All reported XP patients subtype XPF world-wide have at least one missense mutation disrupting NER while all missense mutations found in 2.427 Spanish individuals have substantial NER activity explaining why there are no reported XPF families in Spain.

References:

Bogliolo M, Schuster B, Stoepker C, Derkunt B, Su Y, Raams A, Trujillo JP, Minguillon J, Ramirez MJ, Pujol R, Casado JA, Banos R, et al. 2013. Mutations in ERCC4, Encoding the DNA-Repair Endonuclease XPF, Cause Fanconi Anemia. Am J Hum Genet 92:800-806.

Bonache S, de la Hoya M, Gutierrez-Enriquez S, Tenes A, Masas M, Balmana J, Diez O. 2013. Mutation analysis of the SHFM1 gene in breast/ovarian cancer families. Journal of cancer research and clinical oncology 139:529-532.

de Laat WL, Sijbers AM, Odijk H, Jaspers NG, Hoeijmakers JH. 1998. Mapping of interaction domains between human repair proteins ERCC1 and XPF. Nucleic Acids Res 26:4146-4152.

Fernandez-Rodriguez J, Quiles F, Blanco I, Teule A, Feliubadalo L, Valle JD, Salinas M, Izquierdo A, Darder E, Schindler D, Capella G, Brunet J, et al. 2012. Analysis of SLX4/FANCP in non-BRCA1/2-mutated breast cancer families. BMC Cancer 12:84.

Gregg SQ, Robinson AR, Niedernhofer LJ. 2011. Physiological consequences of defects in ERCC1-XPF DNA repair endonuclease. DNA Repair (Amst) 10:781-791.

Howlett NG, Taniguchi T, Olson S, Cox B, Waisfisz Q, De Die-Smulders C, Persky N, Grompe M, Joenje H, Pals G, Ikeda H, Fox EA, et al. 2002. Biallelic inactivation of BRCA2 in Fanconi anemia. Science 297:606-609.

Levy-Lahad E. 2010. Fanconi anemia and breast cancer susceptibility meet again. Nat Genet 42:368-369.

Meindl A, Hellebrand H, Wiek C, Erven V, Wappenschmidt B, Niederacher D, Freund M, Lichtner P, Hartmann L, Schaal H, Ramser J, Honisch E, et al. 2010. Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. Nat Genet 42:410-414.

Niedernhofer LJ, Garinis GA, Raams A, Lalai AS, Robinson AR, Appeldoorn E, Odijk H, Oostendorp R, Ahmad A, van Leeuwen W, Theil AF, Vermeulen W, et al. 2006. A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. Nature 444:1038-1043.

Osorio A, Endt D, Fernandez F, Eirich K, de la Hoya M, Schmutzler R, Caldes T, Meindl A, Schindler D, Benitez J. 2012. Predominance of pathogenic missense variants in the RAD51C gene occurring in breast and ovarian cancer families. Hum Mol Genet 21:2889-2898.

Pelletier J, Sonenberg N. 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature 334:320-325.

Rahman N, Seal S, Thompson D, Kelly P, Renwick A, Elliott A, Reid S, Spanova K, Barfoot R, Chagtai T, Jayatilake H, McGuffog L, et al. 2007. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. Nature genetics 39:165-167.

Romero A, Perez-Segura P, Tosar A, Garcia-Saenz JA, Diaz-Rubio E, Caldes T, de la Hoya M. 2011. A HRM-based screening method detects RAD51C germ-line deleterious mutations in Spanish breast and ovarian cancer families. Breast cancer research and treatment 129:939-946.

Seal S, Thompson D, Renwick A, Elliott A, Kelly P, Barfoot R, Chagtai T, Jayatilake H, Ahmed M, Spanova K, North B, McGuffog L, et al. 2006. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. Nature genetics 38:1239-1241.

Sijbers AM, de Laat WL, Ariza RR, Biggerstaff M, Wei YF, Moggs JG, Carter KC, Shell BK, Evans E, de Jong MC, Rademakers S, de Rooij J, et al. 1996. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. Cell 86:811-822.

Trujillo JP, Mina LB, Pujol R, Bogliolo M, Andrieux J, Holder M, Schuster B, Schindler D, Surralles J. 2012. On the role of FAN1 in Fanconi anemia. Blood 120:86-89.

Acknowledgements

This study was approved by the Institutional Committee on Ethical Research in Human Samples and proper informed consent was obtained. CNIO study was partially supported by Mutua Madrileña Foundation (FMMA), Intrasalud PI12/00070, Acción Cooperativa y Colaborativa Intramural-CIBER2012 and SAF2010-20493. UAB laboratory is funded by the Generalitat de Catalunya (SGR0489-2009), the ICREA-Academia award, the Spanish Ministry of Economy and Competitiveness (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. MH contribution was supported by Fondo de Investigación Sanitaria (FIS) research grant PI 12/00539, an initiative of the Insituto de Salud Carlos III (Spain), partially supported by European Regional Development FEDER Funds. ICO study is funded by Asociación Española Contra el Cáncer, Spanish Health Research Fund; Carlos III Health Institute; Catalan Health Institute and Autonomous Government of Catalonia. Contract grant numbers: ISCIIIRETIC RD06/0020/1051, RD12/0036/008, PI10/01422, PI10/00748 and 2009SGR290. We wish to thank all members of the ICO Hereditary Cancer Program team led by Dr. Gabriel Capella. CBM laboratory was supported by Acción Cooperativa y Colaborativa Intramural-CIBERER 2012, the OncoCycle Programme (S2011/BMD-2470) from the Comunidad de Madrid, and SAF2012-36556. RGS' study was supported by Fondo de Investigación Sanitaria. (PI10/00219).

All authors of this manuscript made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data, collaborated in drafting the article or revising it critically for important intellectual content and gave final approval of the version to be published.

The authors declare no competing interests of any kind.

Web resources section

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

URL for Leiden Open Variation Database: http://databases.lovd.nl/shared/login

Legend to figure

Figure 1: Funcional studies of *ERCC4* missense variants. **(A)** Western blot analysis showing HA-XPF and GFP expressions in the transduced MEFs as a measure of XPF protein stability. XPF levels were quantified with ImageJ software and are expressed as a ratio relative to the GFP levels. **(B)** XPF levels were quantified as in (A) and expressed as a percentage of the WT. Mean and SEM of 6 experiments are shown (* p=0.005 t-student test). **(C)** UVC sensitivity of Ercc4 KO MEFs expressing the different alleles of *ERCC4*. Cells were analysed 72 h post irradiation. Data represent means and SD of three independent experiments. **(D)** MMC sensitivity in the same cells as in C. Data represent means and SD of two independent experiments. **(E)** MMC-induced G2/M cell cycle arrest in the same cells as in (C). **(F)** DEB-induced chromosome fragility test arrest in the same cells as in (C). Chromatid breaks were scored in 25 metaphase per point. Methods for western blot, site-directed mutagenesis, cDNA transduction, antibodies used and the

experiments presented in D, E and F were performed as reported earlier (Trujillo, et al., 2012; Bogliolo, et al., 2013).

Table 1. Unique genetic variants identified in the *ERCC4* gene in 1573 cases and 854 controls.

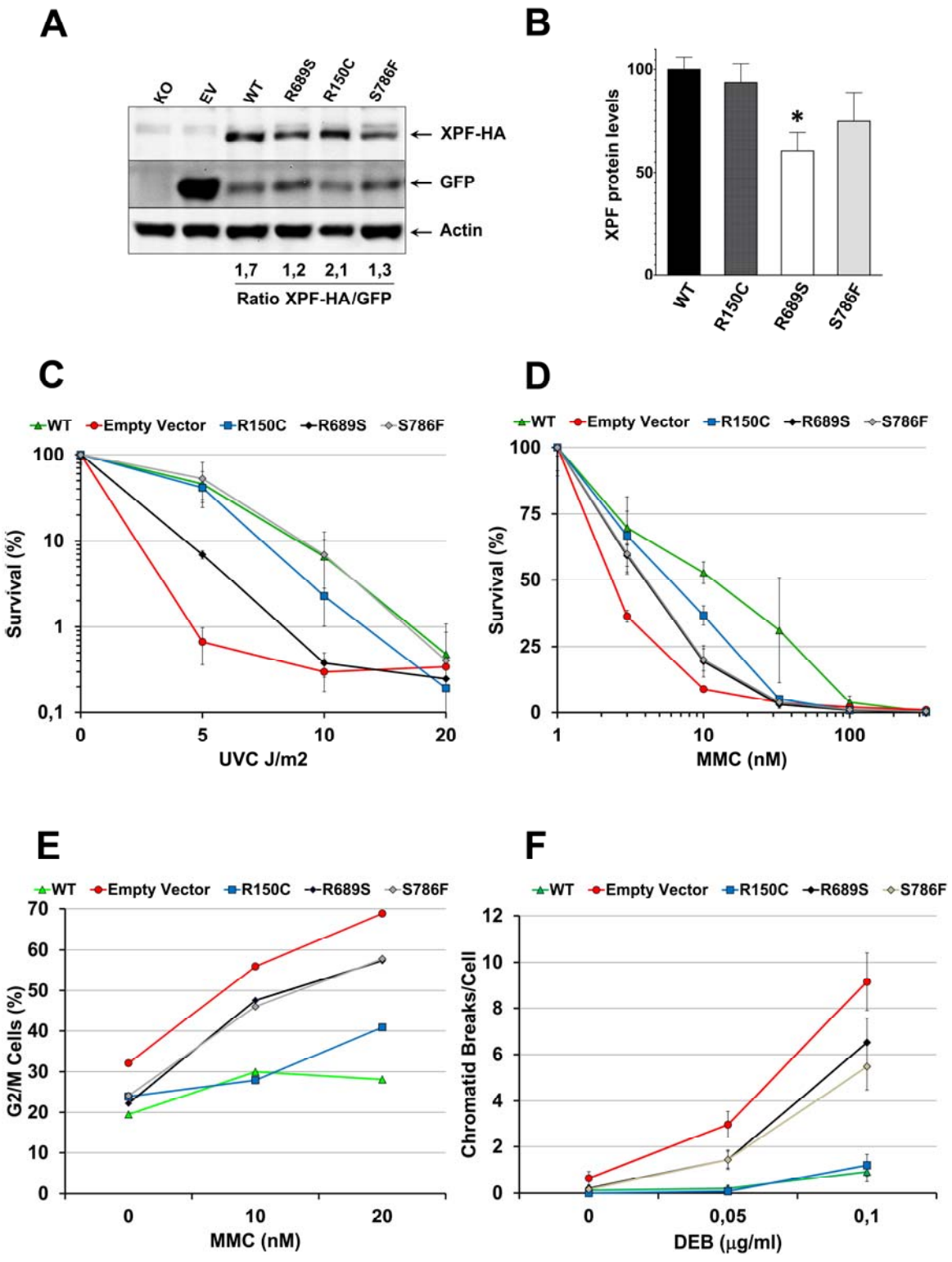
Exon	Nucleotide change[a]	Protein change	Previously described	MAF reported[c]	Phenotype	Predicted to affect protein function[e]	Functional assay[f]
3	c.448C>T[b]	p.Arg150Cys	rs145402255	0.0	Case	Yes	mild effect on ICLR
3	c. 540_541delAG	p.Arg180SerfsX30	_	_	Control[d]	Yes	_
IVS3	c.584+1G>A[b]	p.Ile131SerfsX2	_	_	Case	Yes	_
5	c.800G>A	p.Arg267His	rs143479220	0.0001	Case	No	_
9	c.1861A>G	p.Thr621Ala	_	_	Case	No	_
11	c.2065C>A	p.Arg689Ser	rs149364215	0.0001	Control	Yes	disrupts ICLR
11	c.2291delG	p.Ser764ThrfsX53	_	_	Control	Yes	_
11	c.2357C>T[b]	p.Ser786Phe	rs143081574	_	Case	Yes	disrupts ICLR

[a] Mutation nomenclature listed uses GenBank reference sequences NM_005236.2 with numbering starting at the A of the first ATG, following the HGVS guidelines (www.hgvs.org/mutnomen).

[b] Apart from the 854 controls in which *XPF/ERCC4* was fully sequenced, these variants were specifically analyzed in 300 additional controls. The mutation gave rise to skipping of exon 3 (r.389_584del) that was confirmed at the cDNA level (data not shown).

[c] MAF (Minor Allele Frequency) in European American Population as described in the NHLBI Exome Sequencing Project (ESP) (http://evs.gs.washington.edu/EVS/)

[d] Mutation found in two independent controls.


A
KO
EV
WT
R689S
R150C
S786F
XPF-HA
GFP
Actin
1,7
1,2
2,1
1,3
Ratio XPF-HA/GFP
B
XPF protein levels
100
50
0
*
WT
R150C
R689S
S786F
C
WT
Empty Vector
R150C
R689S
S786F
Survival (%)
100
10
1
0,1
0
5
10
20
UVC J/m2
D
WT
Empty Vector
R150C
R689S
S786F
Survival (%)
100
75
50
25
0
1
10
100
MMC (nM)
E
WT
Empty Vector
R150C
R689S
S786F
G2/M Cells (%)
70
60
50
40
30
20
10
0
0
10
20
MMC (nM)
F
WT
Empty Vector
R150C
R689S
S786F
Chromatid Breaks/Cell
12
10
8
6
4
2
0
0
0,05
0,1
DEB (µg/ml)