Accepted Manuscript

Germline Mutations in *FAN1* Cause Hereditary Colorectal Cancer by Impairing DNA Repair

Nuria Seguí, Leonardo B. Mina, Conxi Lázaro, Rebeca Sanz-Pamplona, Tirso Pons, Matilde Navarro, Fernando Bellido, Adriana López-Doriga, Rafael Valdés-Mas, Marta Pineda, Elisabet Guinó, August Vidal, José Luís Soto, Trinidad Caldés, Mercedes Durán, Miguel Urioste, Daniel Rueda, Joan Brunet, Milagros Balbín, Pilar Blay, Silvia Iglesias, Pilar Garré, Enrique Lastra, Ana Beatriz Sánchez-Heras, Alfonso Valencia, Victor Moreno, Miguel Ángel Pujana, Alberto Villanueva, Ignacio Blanco, Gabriel Capellá, Jordi Surrallés, Xose S. Puente, Laura Valle

PII: S0016-5085(15)00783-0
DOI: 10.1053/j.gastro.2015.05.056
Reference: YGAST 59827

To appear in: *Gastroenterology*
Accepted Date: 28 May 2015

Please cite this article as: Seguí N, Mina LB, Lázaro C, Sanz-Pamplona R, Pons T, Navarro M, Bellido F, López-Doriga A, Valdés-Mas R, Pineda M, Guinó E, Vidal A, Soto JL, Caldés T, Durán M, Urioste M, Rueda D, Brunet J, Balbín M, Blay P, Iglesias S, Garré P, Lastra E, Sánchez-Heras AB, Valencia A, Moreno V, Pujana MÁ, Villanueva A, Blanco I, Capellá G, Surrallés J, Puente XS, Valle L, Germline Mutations in *FAN1* Cause Hereditary Colorectal Cancer by Impairing DNA Repair, *Gastroenterology* (2015), doi: 10.1053/j.gastro.2015.05.056.

Germline Mutations in *FAN1* Cause Hereditary Colorectal Cancer by Impairing DNA Repair

*[1]Nuria Seguí, *[2]Leonardo B Mina, [1]Conxi Lázaro, [3]Rebeca Sanz-Pamplona, [4]Tirso Pons, [1]Matilde Navarro, [1]Fernando Bellido, [3]Adriana López-Doriga, [5]Rafael Valdés-Mas, [1]Marta Pineda, [3]Elisabet Guinó, [6]August Vidal, [7]José Luís Soto, [8]Trinidad Caldés, [9]Mercedes Durán, [10]Miguel Urioste, [11]Daniel Rueda, [12]Joan Brunet, [13]Milagros Balbín, [14]Pilar Blay, [1]Silvia Iglesias, [8]Pilar Garré, [15]Enrique Lastra, [16]Ana Beatriz Sánchez-Heras, [4]Alfonso Valencia, [3,17]Victor Moreno, [18]Miguel Ángel Pujana, [18]Alberto Villanueva, [1]Ignacio Blanco, [1]Gabriel Capellá, [2]Jordi Surrallés, [5]Xose S. Puente, [1]Laura Valle.

*These authors equally contributed to the study.

1. Hereditary Cancer Program, Catalan Institute of Oncology, IDIBELL, Hospitalet de Llobregat, Spain
2. Genome Instability and DNA Repair Group, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, and Center for Biomedical Network Research on Rare Diseases (CIBERER), Barcelona, Spain
3. Unit of Biomarkers and Susceptibility, Catalan Institute of Oncology, IDIBELL and CIBERESP, Hospitalet de Llobregat, Spain
4. Structural Biology and Biocomputing Program, Spanish National Cancer Research Center (CNIO), Madrid, Spain
5. Departamento de Bioquímica y Biología Molecular, Instituto Universitario de Oncología del Principado de Asturias, Universidad de Oviedo, Oviedo, Spain
6. Department of Pathology, Bellvitge University Hospital, IDIBELL, Hospitalet de Llobregat, Spain

7. Molecular Genetics Laboratory, Elche University Hospital, Elche, Spain
8. Laboratorio de Oncología Molecular, Servicio de Oncología Médica, Hospital Clínico San Carlos, Madrid, Spain
9. Instituto de Biología y Genética Molecular, IBGM-UVA-CSIC, Valladolid, Spain
10. Familial Cancer Clinical Unit, Human Cancer Genetics Program, Spanish National Cancer Centre and Center for Biomedical Network Research on Rare Diseases, Madrid, Spain
11. Molecular Biology Laboratory, 12 de Octubre University Hospital, Madrid, Spain
12. Hereditary Cancer Program, Catalan Institute of Oncology, IDIBGi, Girona, Spain
13. Familial Cancer Unit, Department of Medical Oncology, Instituto Universitario de Oncología del Principado de Asturias, Hospital Universitario Central de Asturias, Oviedo, Spain
14. Laboratorio de Oncología Molecular, Instituto Universitario de Oncología del Principado de Asturias, Hospital Universitario Central de Asturias, Oviedo, Spain
15. Department of Oncology, Hospital General Yagüe, Burgos, Spain
16. Unit of Genetic Counseling in Cancer, Elche University Hospital, Elche, Spain
17. Department of Clincal Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain
18. Translational Research Laboratory, Catalan Institute of Oncology, IDIBELL, Hospitalet de Llobregat, Spain

Short title: FAN1 mutations in hereditary colorectal cancer

Grant support

This work has been funded by the Spanish Ministry of Economy and Competitiveness [SAF2012-38885 (LV), SAF2012-31881 (JS), SAF2013-45836R (XSP) and SAF2012-33636 (GC) and Ramón y Cajal Contract to LV]; the Carlos III Health Institute [PI11-01439 (VM), PI13/00285 (CL) and fellowship to NS]; CIBERESP [CB07/02/2005 (VM)]; RTICC [RD12/0036/0031, RD12/0036/0006, RD12/0036/0008 and RD12/0036/0067]; the EU FP7 project ASSET (grant agreement 259348 to AV); L'Oréal-UNESCO "For Women in Science"; the Scientific Foundation Asociación Española Contra el Cáncer; and the Government of Catalonia [2014SGR338 (GC) and 2009SGR0489 (JS)]. JS is also funded by ICREA Academia and the European Regional Development FEDER Funds. CIBERER is an initiative of the Carlos III Health Institute.

Abbreviations

CRC, colorectal cancer

ESP, NHLBI GO Exome Sequencing Project

FA, Fanconi Anemia

ICL, DNA interstrand cross-link

KMIN, karyomegalic interstitial nephritis

MAF, minor allele frequency

MMC, mitomycin C

MMR, DNA mismatch repair

TCGA, The Cancer Genome Atlas

Correspondence

Laura Valle, Ph.D.

Hereditary Cancer Program

Catalan Institute of Oncology, IDIBELL

Av. Gran Via 199-203

08908 Hospitalet de Llobregat, Spain

Tel: (+34) 93 260 7145

Fax: (+34) 93 260 7466

E-mail: lvalle@iconcologia.net

Disclosures

The authors declare no conflict of interest

Author Contributions

LV and CL conceived the study. LV supervised the project, analyzed and interpreted data and wrote the manuscript. XSP, JS and GC provided conceptual and experimental advice and helped write the manuscript. NS, LBM, FB, AVidal and AVillanueva performed experiments and analyzed data. XSP and RV-M analyzed the NGS data (exomes and mutation identification by pool sequencing). RS-P, AL-D and VM analyzed the TCGA exomes (Sequencing data analysis II). TP and AValencia performed and interpreted the 3D structure predictions. EG carried out the statistical analyses. MAP and CL gave conceptual advice. MN, MP, AVidal, JLS, TC, MD, MU, DR, JB, MB, PB, SI, PG, EL, ABS-H, IB, CL and GC provided samples and clinical data. All authors reviewed and provided final approval of the version to be published.

Abstract: Identification of genes associated with hereditary cancers facilitates management of patients with family histories of cancer. We performed exome sequencing of DNA from 3 individuals from a family with colorectal cancer who met the Amsterdam criteria for risk of hereditary nonpolyposis colorectal cancer. These individuals had mismatch repair-proficient tumors and each carried nonsense variant in the FANCD2/FANCI-associated nuclease 1 gene (*FAN1*), which encodes a nuclease involved in DNA inter-strand cross-link repair. We sequenced *FAN1* in 176 additional families with histories of colorectal cancer and performed in vitro functional analyses of the mutant forms of FAN1 identified. We detected *FAN1* mutations in ~3% families with who met the Amsterdam criteria and had mismatch repair-proficient cancers with no previously associated mutations. These findings link colorectal cancer predisposition to the Fanconi anemia DNA repair pathway, supporting the connection between genome integrity and cancer risk.

KEYWORDS: Lynch syndrome; genetic risk factor; susceptibility; DNA mismatch repair

BRIEF REPORT

Familial aggregation of colorectal cancer (CRC) is one of the strongest risk factors for CRC. Germline mutations in the DNA mismatch repair (MMR) genes, *EPCAM*, *APC*, *MUTYH*, *POLE*, *POLD1*, *GREM1*, *SMAD4*, *BMPR1A*, *STK11* and *PTEN* cause hereditary forms of CRC.[1-3] However, part of the observed heritability and familial aggregation of the disease is yet to be explained.

With the aim of identifying new hereditary CRC genes, we sequenced the exomes of three cancer-affected members of a high-risk, Amsterdam I MMR-proficient, CRC family (Figure 1a, Family 1). Out of 32 unreported or rare (MAF<1%) non-synonymous variants shared by all affected relatives (Supplementary Table 1), a nonsense mutation in *FAN1*, c.141C>A (p.C47*), deserved our attention as the coded protein, FANCD2/FANCI-associated nuclease 1 (MIM# 613534), is involved in interstrand cross-link (ICL) repair (Fanconi Anemia; FA) and interacts with MMR components such as MLH1, PMS2 and PMS1, thus playing a role in maintaining genome integrity.[4-8] The identified *FAN1* mutation had not been previously reported (ESP-6500, 1000 Genomes Project) or found in 1648 alleles of Spanish origin, including 286 sporadic CRC patients. *In vitro*, the *FAN1*-deficient phenotype shows lower sensitivity to mitomycin C (MMC) than other FA genes.[9] Even so, heterozygous c.141C>A (p.C47*) cells showed higher sensitivity to relatively high doses of MMC (10-70nM) than wildtype cells (Supplementary Figure 1).

Four additional unreported or rare genetic variants in *FAN1* were identified in 176 MMR-proficient Amsterdam-positive families: a truncating mutation, c.2854C>T (p.R952*), and three missense variants, c.418G>T (p.D140Y), c.1018C>T (p.P340S)

and c.1771C>T (p.R591W). Mutation carrier status could be assessed in 15 members of the *FAN1*-mutated families: all cancer-affected (10 CRC and 1 breast cancer) and three unaffected 21, 43 and 47 year-old individuals were carriers, and one unaffected 53 year-old was non-carrier (Figure 1a). No exonic or splice-site variants were identified in 71 MMR-proficient Bethesda CRC families, in the normal colonic mucosae of 42 Spanish sporadic CRC patients and of 100 TCGA CRC patients,[10] and in 250 Spanish individuals without CRC[11]. However, among the 6503 ESP individuals, a total of 10 nonsense, frameshift or splice-site *FAN1* variants with MAF<1%, were identified in 16 subjects (0.24%). Unfortunately, no information about personal or family history of cancer is available. The limited number of mutation carriers identified (n=14), together with the ascertainment bias due to the study of mostly cancer-affected family members, hampers at this point the estimation of risks and penetrance.

The identification of two truncating mutations in *FAN1* prompted us to investigate whether the other three variants might also affect the protein function. *In silico* algorithms predicted damaging functional effects for p.R591W and p.D140Y (Table 1). p.R591W, located in an evolutionary conserved residue, is also predicted to destabilize the protein structure, being localized in an exposed loop that connects two α-helices in the vicinity of the DNA-binding (SAP) domain (Supplementary Figure 2). c.418G>T (p.D140Y) is located in the first traduced exon, which codes for the UBZ domain, essential for FAN1 localization to sites of damage.[7] Heterozygous c.418G>T (p.D140Y) cells showed similar sensitivity to MMC than heterozygous c.141C>A (p.C47*) (Supplementary Figure 3), suggesting functional implications for c.418G>T. To confirm this, we generated a *FAN1* knock-out HEK293T cell line which recapitulated the MMC-sensitive phenotype observed in $FAN1^{-/-}$ cells (Supplementary Figure 4), and stably

transfected it with wildtype *FAN1*, c.418G>T-mutated *FAN1*and the empty vector. The c.418G>T-transfected cell line showed the same level of sensitivity to MMC as the empty vector (Figure 1b) without affecting FAN1 protein expression (Supplementary Figure 5), strongly suggesting that the missense mutation causes an ICL repair defect.

Five colorectal tumors developed by *FAN1* mutation carriers (three c.141C>A and two c.418G>T) were available for somatic testing. Whole-exome sequencing of the Family 1 proband's tumor identified a total of 236 somatic mutations in transcribed sequences (Supplementary Table 2**)**, with an average mutation rate of 5/Mb, or 1.3/Mb for non-synonymous changes. This mutation burden corresponds to that of non-hypermutant CRCs.[10] However, the mutation spectrum is characterized by an excess of T:A>G:C (10.5%) and C:G>G:C transversions (12.5%), both exceeding the 95th percentiles observed in non-hypermutant TCGA CRCs (Supplementary Figure 6). On the other hand, no clear evidence of somatic *FAN1* second hits was obtained: no LOH (0/5) or somatic mutation (0/3) (Supplementary Figure 7). Furthermore, neither loss of RNA expression of the wildtype allele nor reduction of expression of the FAN1 protein were observed in the tumor developed by a c.141C>A (p.C47*) carrier. However, FAN1 protein levels of normal colon mucosa from the *FAN1* c.141C>A carrier were lower than those of a wildtype individual (Supplementary Figure 8). These observations together with the deficient ICL repair observed in lymphoblastoid cells from heterozygous mutation carriers (Supplementary Figures 1 and 3), suggest that *FAN1* haploinsufficiency might cause a bias towards a specific type of errors due to defective DNA maintenance.

FAN1 interacts with MMR proteins and the nuclease function is required for fully functional MMR.[7,8,12] However, being MMR proficiency an inclusion criteria in our study, tumors developed by *FAN1* mutation carriers showed microsatellite stability and/or normal expression of MMR proteins (Supplementary Figure 9). The fact that a plethora of nucleases, including FAN1, EXO1 and MRE11, can carry out the required nuclease activity for the MMR function,[12] might explain the absence of MMR deficiency in *FAN1* mutation carriers' tumors.

FAN1 deficiency causes distinct milder phenotypes than other components of the FA pathway. Biallelic loss of *FAN1* does not cause FA, but karyomegalic interstitial nephritis (KMIN; MIM# 614817), a very rare recessive disease (~20 families reported so far) characterized by slow progressive renal failure that leads to end-stage renal disease before age 50.[13] Despite the lack of information on cancer history of monoallelic carriers, development of cancer at early ages has been described in two families: an autopsy performed in a 30 year-old KMIN individual revealed a rectal adenocarcinoma and another affected individual died of hepatocellular carcinoma at age 22.[14,15] Interestingly, while KMIN-associated biallelic mutations in *FAN1* localize towards the C-terminus of the protein, after the SAP domain, monoallelic mutations associated with hereditary CRC do not show preferential gene location (Supplementary Figure 10).[7]

Our findings implicate *FAN1* mutations in the inherited susceptibility to CRC. The analysis of larger familial CRC series will provide information about the prevalence of *FAN1* mutations (2.8% of Amsterdam-positive MMR-proficient families in our series) and allow the estimation of lifetime cancer risks for mutation carriers. Likewise, a thorough analysis of genetic and genomic alterations found in *FAN1*-associated tumors

will clarify the underlying repair defects accumulated and therefore the mechanism of action of *FAN1* in colorectal carcinogenesis. Our findings further support the relationship between defective DNA repair and cancer predisposition, being the first unequivocal evidence linking the FA pathway and CRC through *FAN1*, bridge between FA and MMR DNA repair pathways.

ACKNOWLEDGMENTS

We thank Dr. John Rouse (University of Dundee, Dundee, UK) for kindly providing the FAN1 antibody and Dr. Fernando Setién (IDIBELL, Barcelona, Spain) for his technical support.

REFERENCES

1. Kastrinos F, et al. Cancer J 2011;17:405-15.
2. **Palles C**, **Cazier JB**, et al. Nat Genet 2013;45:136-44.
3. **Jaeger E**, **Leedham S**, et al. Nat Genet 2012;44:699-703.
4. **Kratz K**, **Schöpf B**, **Kaden S**, et al. Cell 2010;142:77-88.
5. **Liu T**, **Ghosal G**, et al. Science 2010;329:693-6.
6. MacKay C, et al. Cell 2010;142:65-76.
7. Smogorzewska A, et al. Mol Cell 2010;39:36-47.
8. Cannavo E, et al. J Biol Chem 2007;282:2976-86.
9. **Trujillo JP**, **Mina LB**, et al. Blood 2012;120:86-9.
10. Cancer Genome Atlas Network. Nature 2012;487:330-7.
11. Quesada V, et al. Nat Genet 2011;44:47-52.
12. .Desai A, Gerson S. DNA Repair (Amst) 2014;21:55-64.
13. **Zhou W**, **Otto EA**, et al. Nat Genet 2012;44:910-5.

14. Moch H, et al. Pathologe 1994;15:44-8.

15. Burry AF. J Pathol 1974;113:147-50.

Author names in bold designate shared co-first authorship

FIGURE LEGENDS

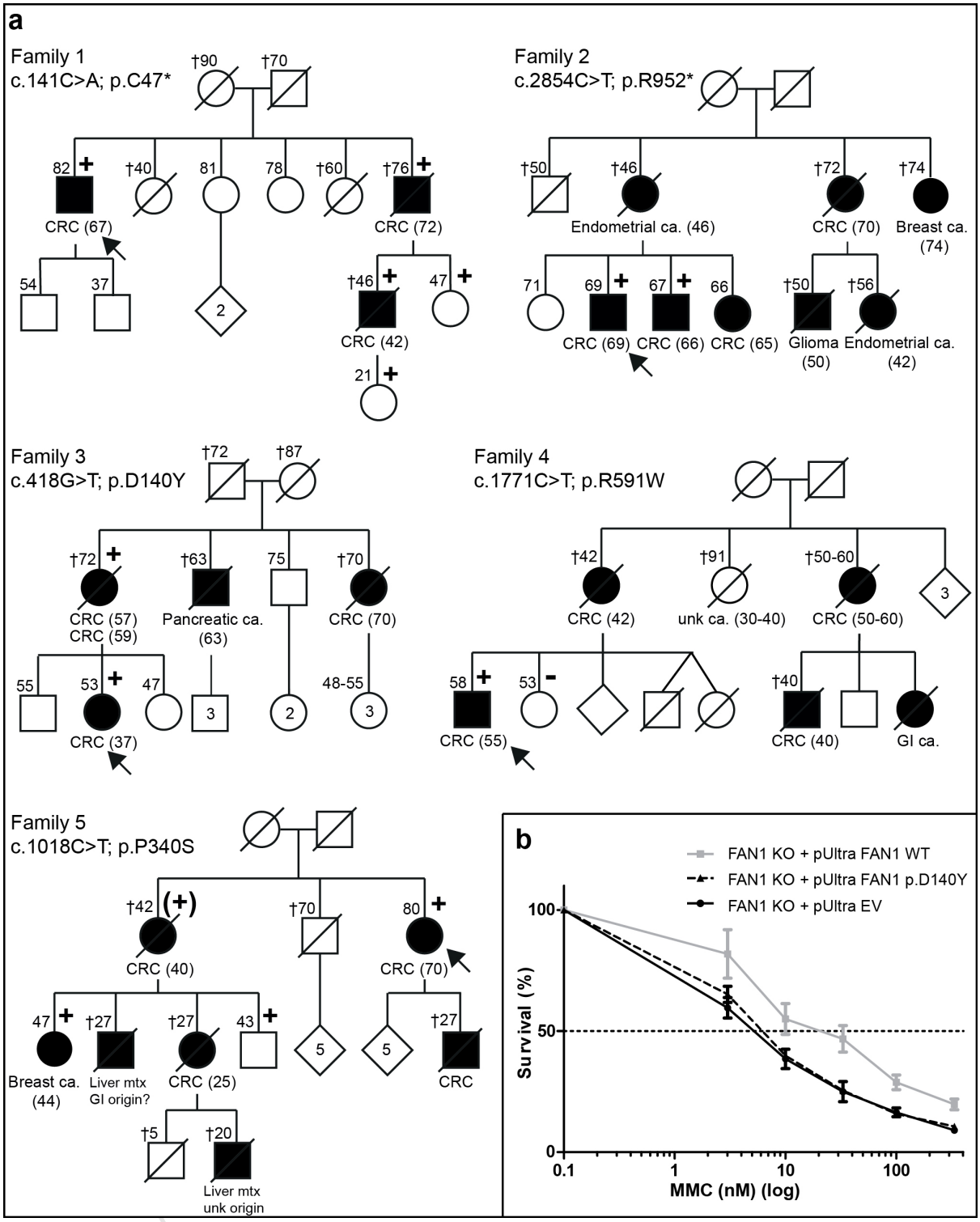
Figure 1. a) Pedigrees of the families with germline *FAN1* mutations. Filled symbol, cancer; +, mutation carrier; (+), obliged mutation carrier; -, wildtype; arrow, index case. Ages at information gathering or at death, when available, are indicated on the top-left corner of each individual's symbol. Ca., cancer; GI, gastrointestinal; mtx, metastasis; unk, unknown location. b) MMC sensitivity assay with the *FAN1* KO HEK293T cell line stably transfected with a pUltra empty vector (EV), the vector with wildtype *FAN1*, and the vector with c.418G>T (p.D140Y)-mutated *FAN1*.

Table 1. Germline *FAN1* mutations identified in 176 MMR-proficient Amsterdam-positive CRC families. In bold, evidence that supports the damaging nature of the variants.

Family	*FAN1* genetic variant	Protein prediction (score)			Structure prediction	Splicing prediction (HSF)	ICL repair status	Population MAF (%) (dbSNP/ESP)
		PolyPhen-2 (HumDiv / HumVar)	SIFT	Condel				
1	c.141C>A (p.C47*)	-	-	-	**Protein truncation**	**-**	**Deficient**[2]	**0/0**
2	c.2854C>T (p.R952*)	-	-	-	**Protein truncation**	**-**	N.P.	**0.05/0**
3	c.418G>T (p.D140Y)	Benign (0.03/0.019)	**Damaging (0.04)**	**Deleterious (0.708)**	N.I.[1]	**New ESS, broken ESE**	**Deficient**[3]	**0/0**
4	c.1771C>T (p.R591W)	**Probably damaging (1/0.998)**	**Damaging (0)**	**Deleterious (1)**	**Protein destabilization**	**New ESS**	N.P.	**0/0.0154**
5	c.1018C>T (p.P340S)	Benign (0.221/0.024)	Tolerated (0.09)	Neutral (0.056)	N.I.[1]	No change	N.P.	**0/0**

1. D140 and P340 are located in a region predicted to have a disordered structure.
2. Supplementary Figure 1.
3. Figure 1b and Supplementary Figure 3.

Abbreviations: ESP, NHLBI GO Exome Sequencing Project; HSF, Human Splicing Finder v.3.0; ICL, DNA interstrand cross-link; MAF, minor allele frequency;; N.I., not informative; N.P, not performed.


a
Family 1
c.141C>A; p.C47*
†90
†70
82
†40
81
78
†60
†76
CRC (67)
CRC (72)
54
37
2
†46
47
CRC (42)
21
Family 2
c.2854C>T; p.R952*
†50
†46
†72
†74
Endometrial ca. (46)
CRC (70)
Breast ca. (74)
71
69
67
66
†50
†56
CRC (69)
CRC (66)
CRC (65)
Glioma (50)
Endometrial ca. (42)
Family 3
c.418G>T; p.D140Y
†72
†87
†72
†63
75
†70
CRC (57)
CRC (59)
Pancreatic ca. (63)
CRC (70)
55
53
47
48-55
3
2
3
CRC (37)
Family 4
c.1771C>T; p.R591W
†42
†91
†50-60
3
CRC (42)
unk ca. (30-40)
CRC (50-60)
58
53
†40
CRC (55)
CRC (40)
GI ca.
Family 5
c.1018C>T; p.P340S
†42
(+)
†70
80
CRC (40)
CRC (70)
47
†27
†27
43
5
5
†27
Breast ca. (44)
Liver mtx GI origin?
CRC (25)
CRC
†5
†20
Liver mtx unk origin
b
FAN1 KO + pUltra FAN1 WT
FAN1 KO + pUltra FAN1 p.D140Y
FAN1 KO + pUltra EV
Survival (%)
100
50
0
0.1
1
10
100
MMC (nM) (log)