MOLECULAR ONCOLOGY 7 (2013) 85–100

available at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/molonc

Polymorphic variations in the *FANCA* gene in high-risk non-*BRCA1/2* breast cancer individuals from the French Canadian population

Nadhir Litim[1], *Yvan Labrie*[1], *Sylvie Desjardins*, *Geneviève Ouellette*, *Karine Plourde*, *Pascal Belleau*, *INHERIT BRCAs*[2], *Francine Durocher*[*]

Cancer Genomics Laboratory, Division of Endocrinology and Genomics of CHUQ Research Centre and Laval University, Québec G1V 4G2, Canada

ARTICLE INFO

Article history:
Received 13 September 2011
Received in revised form
17 August 2012
Accepted 21 August 2012
Available online 11 September 2012

Keywords:
Breast cancer susceptibility
FANCA variants
Fanconi anemia
Alternative splicing

ABSTRACT

The majority of genes associated with breast cancer susceptibility, including *BRCA1* and *BRCA2* genes, are involved in DNA repair mechanisms. Moreover, among the genes recently associated with an increased susceptibility to breast cancer, four are Fanconi Anemia (FA) genes: *FANCD1/BRCA2*, *FANCJ/BACH1/BRIP1*, *FANCN/PALB2* and *FANCO/RAD51C*. FANCA is implicated in DNA repair and has been shown to interact directly with BRCA1. It has been proposed that the formation of FANCA/G (dependent upon the phosphorylation of FANCA) and FANCB/L sub-complexes altogether with FANCM, represent the initial step for DNA repair activation and subsequent formation of other sub-complexes leading to ubiquitination of FANCD2 and FANCI. As only approximately 25% of inherited breast cancers are attributable to *BRCA1/2* mutations, FANCA therefore becomes an attractive candidate for breast cancer susceptibility. We thus analyzed *FANCA* gene in 97 high-risk French Canadian non-BRCA1/2 breast cancer individuals by direct sequencing as well as in 95 healthy control individuals from the same population. Among a total of 85 sequence variants found in either or both series, 28 are coding variants and 19 of them are missense variations leading to amino acid change. Three of the amino acid changes, namely Thr561Met, Cys625Ser and particularly Ser1088Phe, which has been previously reported to be associated with FA, are predicted to be damaging by the SIFT and PolyPhen softwares. cDNA amplification revealed significant expression of 4 alternative splicing events (insertion of an intronic portion of intron 10, and the skipping of exons 11, 30 and 31). *In silico* analyzes of relevant genomic variants have been performed in order to identify potential variations involved in the expression of these spliced transcripts. Sequence variants in *FANCA* could therefore be potential spoilers of the Fanconi-BRCA pathway and as a result, they could in turn have an impact in non-BRCA1/2 breast cancer families.

[*] *Corresponding author*. Cancer Genomics Laboratory, Oncology and Molecular Endocrinology Research Centre, Centre Hospitalier Universitaire de Québec and Laval University, 2705 Laurier Boulevard, T2-53, Québec G1V 4G2, Canada. Tel.: +1 418 654 2296; fax: +1 418 654 2761.
E-mail address: Francine.durocher@crchul.ulaval.ca (F. Durocher).
[1] These authors are joint first author.
[2] Other members of INHERIT BRCAs involved in this study are listed in the Appendix.
1574-7891/$ – see front matter © 2012 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.molonc.2012.08.002

1. Introduction

Pathogenic germline mutations in *BRCA1*, *BRCA2*, *TP53*, *ATM*, *CHEK2*, *BRIP1*, *PALB2* and *RAD51C* genes have been associated with an increased breast cancer risk and, together, are found in less than 25% of breast cancer families showing a clear pattern of inheritance (Stratton and Rahman, 2008; Wang, 2007; Moldovan and D'Andrea, 2009). Thus, it is clear that other susceptibility alleles remain to be identified to explain the increased susceptibility in the remnant high-risk families. As the number and characteristics of such alleles are undetermined, a focussed candidate gene approach based on genes closely interacting with the known susceptibility genes (in particular *BRCA1* and *BRCA2*) constitutes a study design of choice to identify rare moderate-penetrance susceptibility alleles.

It is noteworthy that of the genes recently associated with an increased susceptibility to breast/ovarian cancer, four are Fanconi Anemia (FA) genes: *FANCD1/BRCA2*, *FANCJ/BACH1/BRIP1*, *FANCN/PALB2*, and *FANCO/RAD51C* (Somyajit et al., 2010; Cantor et al., 2001, 2004; Seal et al., 2006; Rahman et al., 2007; Erkko et al., 2007; Pang et al., 2011; Zheng et al., 2010; Meindl et al., 2010), therefore connecting FA proteins to homologous recombination repair (HR).

FANCA is the most frequently mutated gene in FA, representing 60–70% of the cases. *FANCA* is highly polymorphic, with over 350 unique mutations reported (Fanconi Mutation Database; http://www.rockefeller.edu/fanconi/mutate), including large genomic deletions mediated by the unusually high density of ALU repetitions found in its genomic sequence (Levran et al., 2005). For *FANCA*, an association of sequence alterations or altered expression have been suggested in some instances of ovarian cancer and leukemia (Thompson et al., 2005; Lensch et al., 2003; Tischkowitz et al., 2008). In addition, homozygous mutations of *FANCA* were also associated with esophageal cancer in Iranian population (Akbari et al., 2011).

FA proteins not only interact with each other but also work in a network of processes implicated in the maintenance of genome integrity during DNA replication and following some types of DNA damage. When activated, the FANCI/D2 complex associates with chromatin and colocalizes in DNA damage-induced S-phase foci with DNA repair response proteins such as BRCA1, FANCD1/BRCA2, FANCJ, RAD51, PCNA and NBS1 to form the complex II (Wang, 2007; Raschle et al., 2008; Knipscheer et al., 2009).

It has been proposed that the formation of FANCA/G (dependent upon the phosphorylation of FANCA) and FANCB/L subcomplexes altogether with FANCM, represents the initial step for DNA repair activation and for regulating the nuclear accumulation of FANCL, and therefore the subsequent formation of other subcomplexes and ubiquitination of FANCD2 and FANCI (Medhurst et al., 2006). It is worth mentioning that FANCA protein has been also shown to interact with ERCC4 (Sridharan et al., 2003), FANCC (Reuter et al., 2000), FANCF (de Winter et al., 2000), FANCE (Medhurst et al., 2001), BRCA1 (Folias et al., 2002) and SMARCA4 (Otsuki et al., 2001).

As a direct and constitutive interaction between FANCA and BRCA1 has been shown to occur in the cell (Folias et al., 2002), and the close connection between the FANC and BRCA1/2 proteins in DNA repair, we screened the proximal promoter, the coding sequence and intron–exon boundaries of the *FANCA* gene in a cohort of 97 *BRCA1/2*-negative (BRCAX) breast cancer families from the French Canadian population as well as 95 healthy controls from the same origin for sequence variations or splicing variants that could modulate breast cancer risk.

2. Material and methods

2.1. Ascertainment of families and DNA purification

Recruitment of high-risk French Canadian breast/ovarian families (i.e. families in which multiple cases of breast/ovarian cancer are present in close relatives – 3 cases in 1st or 4 cases in 2nd degree relatives – or with strong evidence of a familial component) was part of the major interdisciplinary research program INHERIT BRCAs further explained elsewhere, in which a major component was to identify and characterize BRCA1/2 mutations (Simard et al., 2007). A subset of 97 high-risk French Canadian breast/ovarian cancer families was drawn from the initial study based on the absence of detectable BRCA1/2 mutation (so-called BRCAX) and constituted the cohort used for another study specifically aiming at the identification of other susceptibility loci/genes to breast cancer. This latter study obtained ethics approvals from all participating institutions, and each participant, knowing their inconclusive BRCA1/2 test results signed a specific informed consent, had to be at least 18 years of age and mentally capable. One individual affected with breast cancer per family was selected for analysis, with a selection preference for the youngest subject available in the family. The ascertainment criteria of recruitment and BRCA1/2 status analysis have been described previously (Antoniou et al., 2006; Durocher et al., 2006). In all instances, diagnosis of breast cancer was confirmed by pathology reports.

A cohort of 95 healthy unrelated individuals from the French Canadian population was also included in the study. They were obtained from Dr Damian Labuda at the Centre de Cancérologie Charles Bruneau, Hôpital Ste-Justine, Montreal, Canada. The individuals who provided these samples were recruited on a non-nominative basis, in the framework of long-term studies aiming the characterization of the genetic variability in human populations, approved by the Institutional Ethic Review Board. Blood samples were taken for each participant. Details of the clinical procedures and *BRCA1/2* testing methodology are described elsewhere (Simard et al., 2007; Desjardins et al., 2008, 2009). Genomic DNA extraction of the 97 BRCAX breast cancer cases as well as 95 healthy individuals has been performed as previously described (Durocher et al., 2006).

Additional control genotyping data coming from the French Canadian population were extracted either from the CARTaGENE database ($n = 140$) or from the asthma familial collection of Saguenay–Lac-Saint-Jean (NorthEastern region of the province of Quebec in Canada) ($n = 254$ unrelated subjects with asthma frequency similar to general population, i.e. 10%). To assess whether the frequencies of the rare significant DNA variants ($n = 4$, namely c.-1C/A (NA), c.894-8A/G (rs1164881), c.1627-32T/C (rs17226337) and c.3348 + 18A/G

(rs1800347)) detected in our study might be caused by population stratification, these variant were genotyped in a second French Canadian cohort comprising 192 subjects selected at random from the CRCHUQ glaucoma DNA bank (Vincent Raymond).

2.2. RNA isolation from immortalized cell lines and cDNA synthesis

Lymphocytes were isolated and immortalized from 7 to 9 ml of blood samples using the Epstein–Barr Virus (EBV) in 15% RPMI media as previously described (Durocher et al., 2006). Total RNA was extracted from EBV-transformed β-lymphoblastoid cell lines and seven breast cancer cell lines obtained from the American Type Culture Collection (ATCC) including, two estrogen receptor (ER)-negative breast cancer cell lines (BT-20 and MDA -231), five ER + breast cancer cell lines (SUM-40, CAMA-1, MCF7, T47D and SKBR3) as well as the MCF-10A human epithelial cell line which was used as control, using TRI REAGENT® (Molecular Research Center, inc., Cincinnati, OH, USA) according to the manufacturer's instructions. The purified RNA was stored at −80 °C until use. Following RNA extraction, reverse transcription of 5 μg of RNA was performed as previously described (Durocher et al., 2006).

2.3. PCR amplification, mutation analysis and variant characterization

The coding and 5′UTR (comprising the proximal promoter region) sequences, including intron-exon boundaries of the *FANCA* gene (NM_000135.2) was amplified and sequenced using primers listed in Supplemental Table 1. Sequencing of all 43 exons was performed on breast cancer cases and control individuals using the Big Dye 3.1 chemistry and loaded on an ABI3730XL automated sequencer according to the manufacturer's instructions (Applied Biosystems, Foster City, USA). Analysis of sequence data was done using the Staden package. Allelic frequency was evaluated in both series by means of a Chi^2 test. *p*-values less than 0.05 were considered as significant. Protein and nucleotide sequence alignments with other species were performed using data extracted from the National Center for Biotechnology Information (NCBI) and UCSC databases.

2.4. In silico analyzes

The SIFT and PolyPhen web-based softwares were used to predict the effect of amino acid substitution on protein structure (Ng and Henikoff, 2003; Sunyaev et al., 2001; Ramensky et al., 2002). Assessment of the putative deleterious effect of *FANCA* missense substitutions was also performed with the ALIGN-GVGD algorithm (Tavtigian et al., 2006; Mathe et al., 2006) using a full-length alignment of validated and predicted *FANCA* sequences (produced by aligning human *FANCA* exons to the genomic sequences) selected from the NCBI and Ensembl databases. The alignment was made using T-Coffee (Notredame et al., 2000) and was followed by minor handmade adjustments.

The prediction of the effect of a given variant on exonic and intronic splicing modulators was assessed using the Human Splicing Finder (HSF) (Desmet et al., 2009) which contains all available matrices from auxiliary sequence prediction softwares such as ESEfinder 3.0 program (Cartegni et al., 2003; Smith et al., 2006) for binding sites of the 9G8, Tra2-β and hnRNP A1 proteins. The putative impact of intronic variants on splice consensus sites was evaluated using Splice Site Prediction programs using Neural Networks (SSPNN) (Reese et al., 1997) with default parameters.

2.5. Alternative splicing analysis

Screening for evaluating alternative splicing events (ASEs) occurring in *FANCA* gene was performed by PCR amplification on cDNA on a subset of 10 breast cancer cases using the primers described below. Following screening of the whole cDNA sequence, targeted ASE amplification of cDNA regions covering exons 6–15 and 26–33 was performed on cDNA from BRCAX individuals and human cell lines using the following primers: Exons 6–15, Forward primer: 5′-cattgtgagcctgcaagagctg-3′ and Reverse primer: 5′-cttcttgctgcagccatggtag-3′; Exons 26–33, Forward primer: 5′-agcctcctgacctgtaggacga-3′ and Reverse primer: 5′-gaagaactgctcgcatctggc-3′.

2.5.1. QRT-PCR assays

QRT-PCR assays were performed on StepOnePlus™ Real-Time PCR Systems. For all the assays, a reaction mixture was prepared in a final volume of 20 μl with Fast SYBR® Green Master Mix (Applied Biosystems) which included 500 nM of each primer for *FANCA* wild-type and alternative transcript variants. cDNA samples were reverse-transcribed from total RNA. To perform fluorescent-based real-time PCR, the amount of cDNA used for quantitation was 20 ng of total RNA coming from the immortalized cell lines of the carriers of the variants, as well as ten wild type individuals. The primer sequences were; *FANCA* wild-type exon 10–11: forward primer: 5′ cacggttgatgtactgcagagaa 3′, reverse primer: 5′ actgaacactccgaaccagca 3′; *FANCA* deletion of the exon 11 splice transcript: forward primer: 5′ cacggttgatgtactgcagagaa 3′, reverse primer: 5′ ctgaacagcatcagatgcagc 3′; *FANCA* insertion of the intronic fragment 10A splice transcript forward primer: 5′ cacggttgatgtactgcagagaa 3′, reverse primer: 5′ tcctcctcacgcacgttatcg 3′; *FANCA* wild-type exon 30–31: forward primer: 5′ tccgagaggtgttgaaagagg 3′, reverse primer: 5′ ggtcataactccttgagctttgg 3′; *FANCA* deletion of the exon 30 splice transcript forward primer: 5′ tccgagaggtgttgaaagagg 3′, reverse primer: 5′ ggtcataactccttgagcgttca 3′; *FANCA* deletion of the exon 31 splice transcript forward primer: 5′ tccgagaggtgttgaaagagg 3′, reverse primer: 5′ ggtcagctaccatctcctttgg 3′. Expression analyzes were then carried out using the ΔΔCt method.

2.5.2. Polysome analysis

$12x10^7$ cells were suspended in 3 volumes of RNAse-free lysis buffer (250 mM sucrose solution, Cycloheximide 10 mg/ml, KCl 25 mM, MgCl2 5 mM, tris HCl 50 mM, DTT 2 mM) containing NP40 (0.5%), Roche protease tablet and phosphatase inhibitor 1×. The cells were homogenized in the solution and incubated 10 min on ice. The nuclei were pelleted by centrifugation at 750 g for 10 min at 4 °C. The supernatant was then centrifuged at 12,000 g for 10 min to pellet the mitochondria. The supernatant was loaded onto linear gradient of 15–45% sucrose (W/W) and centrifuged at 38,000 rpm for 2 h at 4 °C

in Bekman SW-41 Ti rotor. As a control, equivalent supernatants were prepared and centrifuged in sucrose gradient in buffers in which $MgCl_2$ was replaced by 10 mM EDTA. The absorbance of the gradients at 260 nm was determined and the gradients were fractionated manually. **RNA extraction**: SDS was added to a final concentration of 2% and proteinase K to a final concentration of 100 μg/ml to each sucrose gradient fraction for 1 h at 37 °C. An equal volume of phenol:chloroform (1:1) was then added to each fraction. The sucrose gradient fractions were concentrated by two precipitations with ethanol and dissolved in DEPC water.

2.6. *Electronic databases*

ESE Finder: http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi
Human Splicing Finder (HSF): http://www.umd.be/HSF
Splice Site Prediction Program using Neural Networks (SSPNN): (http://www.fruitfly.org/seq_tools/splice.html)
UCSC Genome Bioinformatics: http://genome.ucsc.edu/
SIFT: http://blocks.fhcrc.org/sift/SIFT.html
PolyPhen: http://genetics.bwh.harvard.edu/pph/

3. Results

Sequencing of *FANCA* genomic and complementary DNA of 97 BRCAX breast cancer subjects from high risk French Canadian breast/ovarian cancer families and 95 healthy controls led to the identification of sequence variants and spliced isoforms. *In silico* approaches provided significant knowledge concerning the potential impact of these sequence variants into splicing of exonic regions of transcripts and consequently into the potential variations involved in the expression of these spliced transcripts. The impact of a given variant in non-BRCA1/2 breast cancer families and the potential disruption of the Fanconi-BRCA pathway could thus justify this screening strategy.

3.1. *Variant characterization*

In order to perform the mutational analysis of *FANCA*, all 43 exons and adjacent intronic sequences were sequenced in our case and control datasets. In the 97 BRCAX breast cancer patients (1 per family) and 95 healthy individuals studied, we identified 85 sequence variants (Table 1). Among the twenty-eight exonic variants identified, ten were found only in cases, while 5 others are found exclusively in healthy individuals. A high proportion (10 out of 28) of the exonic sequence variants were not reported in public databases or relevant publications. The c.796A/G (Thr266Ala) and c.1501G/A (Gly501Ser) are the coding variants whose corresponding frequencies show a significant difference of MAF between cases and controls with *p*-values of 0.048 and 0.003, respectively. These variants are over-represented in controls and both non-synonymous substitutions show a high MAF >20% (Table 1). It is important to mention that to further support the frequency of the variants identified in our cohort, additional genotype data from two independent cohorts, both from the French Canadian control populations, were obtained ($n = 254$ and $n = 140$, respectively), for a total of 24 sequence variants. In addition, the four variants displaying a significant *p*-value were genotyped in an additional distinct cohort of 192 control individuals, also of French Canadian origin. We were able to extract information on a total of 28 SNPs displaying a reasonable frequency for a total of 586 new control individuals.

As indicated in Table 1, in all but one instance (rs17226337), the MAF observed in the other control cohorts are similar to those calculated in our initial cohort. This confirms that the frequency of the variants included in Table 1 is a reliable representation of the frequency observed in the French Canadian population not affected with breast cancer. As for the four SNPs displaying a *p*-value <0.05, the c.1627-32T/C (rs17226337) variant which was further genotyped in 192 individuals from the French Canadian population provided by Dr. Vincent Raymond, presents a MAF (0.025) similar to what is seen in the breast cancer series, suggesting that this variant is unlikely to have an effect on breast cancer risk in the French Canadian population.

Intronic analyzes were performed on sequences adjacent to the transcribed DNA portions and the resulting proximal variants are also presented in Table 1. Among these intronic variants, a significant difference in genotype frequencies was observed for a variant located in intron 10 (c.894-8A/G) that was under-represented in the case dataset (AA vs. AG + GG 0.151 95% CI 0.0133–0.702; $p = 0.00634$) and for another variant namely, c.3348 + 18A/G found in intron 33 which was over-represented in cases (AA vs. AG + GG 3.256 95% CI 1.133–9.358; $p = 0.02206$). Borderline significance is estimated for the three following intronic variants: c.1226-80T/C, c.1627-32T/C and c.3240-42G/A.

3.2. *In silico and splicing analyzes*

The two variants located in the proximal promoter region (less than 150 bp from the transcriptional start site) were also evaluated for their potential influence on transcription factor binding sites using the MatInspector program. The 13 bp duplication sequence (c.-42-87ins13) is associated with the appearance of four motifs that are not present on the reference sequence: a motif for the TATA box binding protein (TBP), an estrogen response element (ERE), a motif for the vertebrate steroidogenic factor SF1 and a second motif for the cAMP-responsive element binding protein (CREB). Interestingly the c.-42-76G/C variant could potentially create a binding site for p53 and a motif for a chorion-specific transcription factor with a GCM DNA binding domain (GCMF) (data not shown).

As shown in Table 2, the Thr561, Cys625, Glu1023, Leu1143 and Ala1352 amino acid show a high degree of conservation in lower species, suggesting that these amino acids are under strong functional constraint or may have a specific role on protein conformation, while the other residues leading to amino acid changes are not well conserved in distant species. Analysis of amino acid changes was performed in an attempt to predict the functional consequences of *FANCA* missense variants using the PolyPhen and SIFT softwares as well as the ALIGN-GVGD algorithm implementing an extension of the Grantham difference. As displayed in Table 3, both the SIFT and PolyPhen softwares predicted the Thr561Met, Cys625Ser and Ser1088Phe changes to be damaging for protein

Table 1 – Sequence variations in *FANCA* gene and genotype frequencies in familial breast cancer cases and controls.

SNP	SNP ID[a] (dbSNP ID)	Location	Amino acid change	Series (N =)	Hom	Het	Rare hom	MAF[b]	*p*-value[c]	Other cohorts
1	c.-42-100ins13 (rs36233534)	Promoter	–	Cases (95)	59	31	5	0.216	0.146	
				Controls (93)	48	38	7	0.280		
2	c.-42-76 G/C (NA)	Promoter	–	Cases (95)	95	0	0	0	0.311	
				Controls (93)	92	1	0	0.005		
3	**c.-1C/A (NA)**	**5' UTR**	–	**Cases (96)**	**92**	**4**	**0**	**0.021**	**0.047**	
				Controls (93)	**93**	**0**	**0**	**0**		0.009[f]
4	**c.17 T/A (rs1800282)**	**Exon 1**	**Val6Asp**	**Cases (96)**	**84**	**12**	**0**	**0.063**	**0.888**	
				Controls (93)	**82**	**11**	**0**	**0.059**		
5	**c.343 G/A (NA)**	**Exon 4**	**Gly115Arg**	**Cases (97)**	**96**	**1**	**0**	**0.005**	**0.326**	
				Controls (93)	**93**	**0**	**0**	**0**		
6	c.427-59A/G (rs2074963)	Intron 4	–	Cases (97)	89	8	0	0.041	0.296	
				Controls (93)	81	12	0	0.065		0.042[d]
7	**c.694A/C (rs61757384)**	**Exon 7**	**Arg232Arg**	**Cases (96)**	**94**	**2**	**0**	**0.010**	**0.157**	
				Controls (95)	**95**	**0**	**0**	**0**		
8	**c.695 G/A (NA)**	**Exon 7**	**Arg232Lys**	**Cases (96)**	**96**	**0**	**0**	**0**	**0.314**	
				Controls (95)	**94**	**1**	**0**	**0.005**		
9	c.710-12A/G (rs1800286)	Intron 7	–	Cases (96)	30	45	21	0.453	0.570	
				Controls (91)	32	42	17	0.418		0.417[d], 0.389[e]
10	**c.796A/G (rs7190823)**	**Exon 9**	**Thr266Ala**	**Cases (97)**	**51**	**38**	**8**	**0.278**	**0.048**	
				Controls (94)	**36**	**43**	**15**	**0.388**		0.329[d], 0.357[e]
11	c.894-8A/G (rs1164881)	Intron 10	–	Cases (97)	95	2	0	0.010	0.006	
				Controls (90)	79	11	0	0.061		0.059[f]
12	c.1083 + 120 G/A (rs17226159)	Intron 12	–	Cases (96)	69	23	4	0.161	0.381	
				Controls (93)	72	20	1	0.118		
13	c.1083 + 142 G/A (NA)	Intron 12	–	Cases (96)	94	2	0	0.010	0.159	
				Controls (94)	94	0	0	0		
14	c.1083 + 195C/T (NA)	Intron 12	–	Cases (96)	93	3	0	0.016	0.317	
				Controls (94)	94	1	0	0.005		
15	c.1083 + 222 G/C (rs6500453)	Intron 12	–	Cases (96)	46	38	12	0.323	0.270	
				Controls (95)	38	42	15	0.379		0.425[d]
16	c.1084-49 G/C (rs1800287)	Intron 12	–	Cases (96)	50	38	8	0.281	0.094	
				Controls (95)	38	43	14	0.374		0.425[d]
17	c.1084-30A/G (rs6500452)	Intron 12	–	Cases (96)	60	32	4	0.208	0.070	
				Controls (95)	47	39	9	0.300		0.241[d], 0.271[e]
18	**c.1143 G/T (rs1800331)**	**Exon 13**	**Thr381Thr**	**Cases (97)**	**88**	**9**	**0**	**0.046**	**0.338**	
				Controls (95)	**82**	**13**	**0**	**0.068**		0.042[d]
19	c.1225 + 12A/G (NA)	Intron 13	–	Cases (97)	96	1	0	0.005	0.321	
				Controls (95)	95	0	0	0		
20	c.1225 + 151C/T (rs6500451)	Intron 13	–	Cases (96)	83	13	0	0.068	0.379	
				Controls (95)	86	9	0	0.047		
21	c.1226–80 T/C (rs6500450)	Intron 13	–	Cases (96)	53	35	8	0.266	0.050	
				Controls (95)	39	41	15	0.374		0.458[d]
22	c.1226-75C/G (NA)	Intron 13	–	Cases (96)	96	0	0	0	0.313	
				Controls (95)	94	1	0	0.005		
23	c.1226-20A/G (rs1800330)	Intron 13	–	Cases (97)	57	34	6	0.237	0.151	
				Controls (95)	46	39	10	0.311		0.383[d]
24	**c.1235C/T (rs11646374)**	**Exon 14**	**Ala412Val**	**Cases (97)**	**88**	**9**	**0**	**0.046**	**0.457**	
				Controls (95)	**83**	**12**	**0**	**0.063**		0.053[d], 0.058[e]
25	**c.1501 G/A (rs2239359)**	**Exon 16**	**Gly501Ser**	**Cases (97)**	**55**	**28**	**14**	**0.289**	**0.003**	
				Controls (94)	**33**	**42**	**19**	**0.426**		0.362[d]
26	c.1627–32 T/C (rs17226337)	Intron 17	–	Cases (97)	93	4	0	0.021	0.049	
				Controls (92)	92	0	0	0		0.025[f]
27	c.1627-31C/G (NA)	Intron 17	–	Cases (97)	96	1	0	0.005	0.530	
				Controls (92)	90	2	0	0.011		
28	**c.1679C/T (NA)**	**Exon 18**	**Thr561Met**	**Cases (97)**	**94**	**3**	**0**	**0.015**	**0.086**	
				Controls (94)	**94**	**0**	**0**	**0**		
29	c.1715 + 82 T/C (rs1800335)	Intron 18	–	Cases (97)	50	39	8	0.284	0.135	
				Controls (91)	37	40	14	0.374		
30	c.1776 + 64 T/C (rs2302162)	Intron 19	–	Cases (97)	88	9	0	0.046	0.660	
				Controls (89)	79	10	0	0.056		
31	c.1826 + 15 T/C (rs1800337)	Intron 20	–	Cases (97)	50	39	8	0.284	0.068	
				Controls (89)	34	40	15	0.393		0.425[d]

(continued on next page)

Table 1 (continued)

SNP	SNP ID[a] (dbSNP ID)	Location	Amino acid change	Series (N =)	Hom	Het	Rare hom	MAF[b]	p-value[c]	Other cohorts
32	c.1826 + 29insGT (rs1799742)	Intron 20	–	Cases (97)	88	9	0	0.046	0.365	
				Controls (89)	77	12	0	0.067		
33	**c.1830A/G (rs1800338)**	**Exon 21**	**Ala610Ala**	**Cases (91)**	**88**	**3**	**0**	**0.016**	**0.081**	
				Controls (91)	**91**	**0**	**0**	**0**		
34	**c.1874 G/C (NA)**	**Exon 21**	**Cys625Ser**	**Cases (91)**	**91**	**0**	**0**	**0**	**0.316**	
				Controls (91)	**90**	**1**	**0**	**0.005**		
35	**c.1927C/G (rs17232910)**	**Exon 22**	**Pro643Ala**	**Cases (97)**	**90**	**7**	**0**	**0.036**	**0.192**	
				Controls (93)	**81**	**12**	**0**	**0.065**		
36	c.2014 + 42 G/T (rs1800339)	Intron 22	–	Cases (97)	89	8	0	0.041	0.208	
				Controls (93)	80	13	0	0.070		0.05[d]
37	**c.2151 G/T (rs1131660)**	**Exon 23**	**Met717Ile**	**Cases (97)**	**87**	**10**	**0**	**0.052**	**0.128**	
				Controls (90)	**86**	**4**	**0**	**0.022**		
38	c.2151 + 8 T/C (rs1800340)	Intron 23	–	Cases (97)	56	35	6	0.242	0.130	
				Controls (90)	42	38	10	0.322		
39	c.2222 + 73A/G (rs1800341)	Intron 24	–	Cases (93)	81	12	0	0.065	0.612	
				Controls (86)	77	9	0	0.052		
40	c.2222 + 100A/G (rs886950)	Intron 24	–	Cases (93)	48	38	7	0.280	0.143	
				Controls (86)	35	37	14	0.378		0.433[d]
41	c.2222 + 107 T/C (rs886951)	Intron 24	–	Cases (93)	48	38	7	0.280	0.143	
				Controls (86)	35	37	14	0.378		
42	c.2223-113C/T (rs886952)	Intron 24	–	Cases (94)	48	40	6	0.277	0.121	
				Controls (86)	34	37	15	0.390		0.425[d]
43	c.2316 + 67 G/A (rs62989960)	Intron 25	–	Cases (94)	78	15	1	0.090	0.193	
				Controls (87)	78	9	0	0.052		0.025[d]
44	**c.2426 G/A (rs7195066)**	**Exon 26**	**Gly809Asp**	**Cases (94)**	**52**	**36**	**6**	**0.255**	**0.209**	
				Controls (87)	**40**	**36**	**11**	**0.333**		0.273[d], 0.302[e]
45	**c.2574C/G (rs17233141)**	**Exon 27**	**Ser858Arg**	**Cases (97)**	**97**	**0**	**0**	**0**	**0.078**	
				Controls (95)	**92**	**3**	**0**	**0.016**		
46	**c.2589C/A (rs72807571)**	**Exon 27**	**Gly863Gly**	**Cases (97)**	**96**	**1**	**0**	**0.005**	**0.321**	
				Controls (95)	**95**	**0**	**0**	**0**		
47	c.2779–7 T/C (rs17233253)	Intron 28	–	Cases (96)	87	9	0	0.047	0.548	
				Controls (91)	80	11	0	0.060		
48	c.2852 + 137 T/C (rs12933317)	Intron 29	–	Cases (96)	87	9	0	0.047	0.531	
				Controls (90)	79	11	0	0.061		
49	c.2852 + 314A/T (rs78904586)	Intron 29	–	Cases (97)	79	16	1	0.094	0.284	
				Controls (80)	60	18	2	0.138		
50	**c.2859C/G (NA)**	**Exon 30**	**Asp953Glu**	**Cases (97)**	**97**	**0**	**0**	**0**	**0.311**	
				Controls (95)	**94**	**1**	**0**	**0.005**		
51	**c.2901C/T (rs17226980)**	**Exon 30**	**Ser967Ser**	**Cases (97)**	**89**	**8**	**0**	**0,041**	**0.440**	
				Controls (95)	**84**	**11**	**0**	**0.058**		
52	c.3067-23 G/A (rs17227057)	Intron 31	–	Cases (96)	87	9	0	0.047	0.601	
				Controls (94)	83	11	0	0.059		
53	c.3067-4 T/C (rs17227064)	Intron 31	–	Cases (96)	87	9	0	0.047	0.601	
				Controls (94)	83	11	0	0.059		
54	**c.3069 G/T (NA)**	**Exon 32**	**Glu1023Asp**	**Cases (96)**	**95**	**1**	**0**	**0.005**	**0.321**	
				Controls (94)	**94**	**0**	**0**	**0**		
55	c.3239 + 32indel19 (NA)	Intron 32	–	Cases (96)	96	0	0	0	0.311	
				Controls (94)	93	1	0	0.005		
56	c.3240-42 G/A (rs1800345)	Intron 32	–	Cases (97)	62	31	4	0.201	0.045	
				Controls (91)	45	36	10	0.308		
57	**c.3263C/T (rs17233497)**	**Exon 33**	**Ser1088Phe**	**Cases (97)**	**88**	**9**	**0**	**0.046**	**0.584**	
				Controls (94)	**83**	**11**	**0**	**0.059**		
58	c.3348 + 18A/G (rs1800347)	Intron 33	–	Cases (97)	82	15	0	0.077	0.022	
				Controls (94)	89	4	1	0.032		0.04[f]
59	c.3348 + 25C/T (NA)	Intron 33	–	Cases (97)	97	0	0	0	0.308	
				Controls (94)	93	1	0	0.005		
60	c.3348 + 29C/T (rs1800348)	Intron 33	–	Cases (97)	95	2	0	0.010	0.579	
				Controls (94)	93	1	0	0.005		
61	c.3348 + 40 T/A (NA)	Intron 33	–	Cases (97)	95	2	0	0.010	0.579	
				Controls (94)	93	1	0	0.005		
62	c.3408 + 21C/G (NA)	Intron 34	–	Cases (97)	96	1	0	0.005	0.324	
				Controls (94)	94	0	0	0		

Table 1 (*continued*)

SNP	SNP ID[a] (dbSNP ID)	Location	Amino acid change	Series (N =)	Hom	Het	Rare hom	MAF[b]	*p*-value[c]	Other cohorts
63	c.3408 + 45 G/A (rs1800355)	Intron 34	–	Cases (96)	87	9	0	0.047	0.456	
				Controls (94)	82	12	0	0.064		0.042[d]
64	**c.3427C/G (rs61753269)**	**Exon 35**	**Leu1143Val**	**Cases (96)**	**95**	**1**	**0**	**0.005**	**0.321**	
				Controls (94)	**94**	**0**	**0**	**0**		
65	c.3513 + 62C/T (rs1800356)	Intron 35	–	Cases (96)	82	13	1	0.078	0.538	
				Controls (95)	84	10	1	0.063		
66	**c.3654A/G (rs1800358)**	**Exon 37**	**Pro1218Pro**	**Cases (97)**	**82**	**15**	**0**	**0.077**	**0.888**	
				Controls (95)	**81**	**14**	**0**	**0.074**		
67	**c.3711C/G (NA)**	**Exon 37**	**Val1237Val**	**Cases (97)**	**97**	**0**	**0**	**0**	**0.311**	
				Controls (95)	**94**	**1**	**0**	**0.005**		
68	c.3765 + 37 G/A (rs34420680)	Intron 37	–	Cases (97)	88	9	0	0.046	0.602	
				Controls (95)	84	11	0	0.058		
69	c.3765 + 41 G/A (NA)	Intron 37	–	Cases (97)	97	0	0	0	0.311	
				Controls (94)	94	1	0	0.005		
70	c.3765 + 61A/G (NA)	Intron 37	–	Cases (97)	97	0	0	0	0.311	
				Controls (94)	94	1	0	0.005		
71	**c.3807 G/C (rs11649210)**	**Exon 38**	**Leu1269Leu**	**Cases (96)**	**84**	**12**	**0**	**0.063**	**0.631**	
				Controls (94)	**80**	**14**	**0**	**0.074**		0.042[d]
72	c.3828 + 81 G/T (rs11649162)	Intron 38	–	Cases (96)	88	8	0	0.042	0.319	
				Controls (94)	82	12	0	0.064		0.042[d]
73	c.3828 + 251A/G (rs17227347)	Intron 38	–	Cases (94)	88	6	0	0.032	0.148	
				Controls (94)	92	2	0	0.011		
74	c.3828 + 295C/T (NA)	Intron 38	–	Cases (95)	94	1	0	0.005	0.319	
				Controls (94)	94	0	0	0		
75	c.3828 + 313 G/A (rs17233734)	Intron 38	–	Cases (95)	95	0	0	0	0.313	
				Controls (94)	93	1	0	0.005		
76	c.3829-306 G/A (rs55927037)	Intron 38	–	Cases (95)	94		0	0.005	0.994	
				Controls (94)	93	1	0	0.005		
77	c.3829-225C/T (rs11648689)	Intron 38	–	Cases (95)	88	7	0	0.037	0.310	
				Controls (94)	83	11	0	0.059		
78	c.3829-107A/T (rs11644967)	Intron 38	–	Cases (95)	83	12	0	0.063	0.673	
				Controls (95)	81	14	0	0.074		0.042[d]
79	c.3935-16C/T (rs1061646)	Intron 39	–	Cases (95)	59	32	4	0.211	0.082	
				Controls (91)	45	36	10	0.308		
80	**c.3981C/T (NA)**	**Exon 40**	**His1327His**	**Cases (97)**	**96**	**1**	**0**	**0.005**	**0.324**	
				Controls (94)	**94**	**0**	**0**	**0**		0.4[d]
81	**c.3982A/G (rs9282681)**	**Exon 40**	**Thr1328Ala**	**Cases (97)**	**89**	**8**	**0**	**0.041**	**0.747**	
				Controls (94)	**85**	**9**	**0**	**0.048**		0.05[d]
82	c.4010 + 92 T/C (rs9282682)	Intron 40	–	Cases (97)	83	14	0	0.072	0.953	
				Controls (92)	79	13	0	0.071		
83	**c.4055C/A (NA)**	**Exon 41**	**Ala1352Asp**	**Cases (97)**	**96**	**1**	**0**	**0.005**	**0.329**	
				Controls (92)	**92**	**0**	**0**	**0**		
84	c.4167 + 46C/T (NA)	Intron 41	–	Cases (97)	94	3	0	0.015	0.093	
				Controls (90)	90	0	0	0		
85	c.4260 + 29 T/C (rs1800359)	Intron 42	–	Cases (97)	30	46	21	0.454	0.407	
				Controls (90)	33	38	19	0.422		0.431[d], 0.4[e]

Exonic variants are displayed in bold characters.
a SNP ID are indicated according to the nomenclature guidelines of the Human Genome Variation Society (RefSeq NM_000135.2 corresponding to transcript variant 1). The first base from the ATG codon is counted as +1. dbSNP ID is indicated according to build 129, NA indicating a SNP not found in the database.
b MAF: Minor allele frequency.
c *p*-Value of the common homozygotes versus heterozygotes and rare homozygotes.
d 254 controls from the asthma familial collection of Saguenay–Lac-Saint-Jean (NorthEastern region of the province of Quebec in Canada).
e 140 controls coming from French Canadian population (CARTaGENE).
f 192 controls from French Canadian population selected at random from the CRCHUQ glaucoma DNA bank (Vincent Raymond).

conformation and function. Following analysis with the Align-GVGD program, among the variants found in our case dataset, only 3 have a grade above C0. The Thr561Met variant, found in three cases, falls within the C65 grade which is the most likely deleterious grade. The Glu1023Asp variant (one case) has an intermediate probability to be damaging and the Ala1352Asp variant (one case) falls within the low probability class C15. As for the variants found in healthy control individuals, only the Cys625Ser rare variant is predicted to be deleterious (grade C65). Therefore, the Thr561Met and Cys625Ser amino acid changes are predicted to be damaging for protein function by all three programs and most of the changes leading to

Table 2 – Non-synonymous sequence variants detected in human FANCA protein and residues found in orthologues.

SNP[a]	SNP ID[b]	Amino acid change	*Macaca mulatta*	*Gallus gallus*	*Mus musculus*	*Canis lupus familiaris*	*Loxodonta africana*	*Monodelphis domestica*	*Xenopus ropicalis*
4	c.17 T/A	Val6Asp	N/A	N/A	Ala	Thr	Ala	Ser	N/A
5	c.343 G/A	Gly115Arg	N/A	Lys	Lys	Gln	Arg	Gln	N/A
8	c.695 G/A	Arg232Lys	N/A	Gly	Glu	Gln	Arg	Gly	N/A
10	c.796A/G	Thr266Ala	N/A	Cys	Ala	Ala	N/A	Cys	N/A
24	c.1235C/T	Ala412Val	N/A	Ser	Ala	Thr	Ala	Ala	N/A
25	c.1501 G/A	Gly501Ser	N/A	Thr	Ser	Ser	Glu	Val	Pro
28	c.1679C/T	Thr561Met	N/A	Thr	Thr	Thr	Thr	Thr	Thr
34	c.1874 G/C	Cys625Ser	N/A	Cys	Cys	Cys	Cys	Cys	N/A
35	c.1927C/G	Pro643Ala	N/A	N/A	Ala	Ala	Pro	Thr	N/A
37	c.2151 G/T	Met717Ile	N/A	N/A	Ala	Gln	Glu	Lys	N/A
44	c.2426 G/A	Gly809Asp	N/A	N/A	Ser	Val	Ala	N/A	N/A
45	c.2574C/G	Ser858Arg	N/A	N/A	Asn	Ser	Gly	Asn	N/A
46	c.2859C/G	Asp953Glu	N/A	N/A	Asp	Asp	Asp	Tyr	N/A
54	c.3069 G/T	Glu1023Asp	Glu	Glu	Glu	Glu	Glu	Glu	N/A
57	c.3263C/T	Ser1088Phe	Ser	N/A	Ser	Ser	Thr	Ser	N/A
64	c.3427C/G	Leu1143Val	Leu	Leu	Leu	Ser	Leu	Leu	Val
81	c.3982A/G	Thr1328Ala	Thr	N/A	Thr	Ile	Ile	Ile	Leu
83	c.4055C/A	Ala1352Asp	Ala	N/A	Ala	Ala	Ala	Ala	Asp

N/A: no corresponding residue found in this species.
a According to Table 1.
b According to the nomenclature of the Human Genome Variation Society.

a potential deleterious effect identified by at least one of the programs used, are located in important domains or binding sites such as the BRCA1 binding site region (aa 1–589) or the Leucine zipper domain (aa 1069–1090).

Analysis of *FANCA* cDNA was performed in breast cancer cases and highlighted the expression of four distinct alternative splicing events (ASEs) (Figure 1). These ASEs are observed in the *FANCA* genomic regions of exons 10–11 and 30–31 (BU616925, CN404731, AK301168, and BI908441). The first alternative spliced isoform (designed FANCAins10A) involved the insertion of 128 bp of intronic sequence located between exons 10 and 11 and is expected to result in a premature termination of *FANCA* open reading frame leading to a protein of 297 aa. The other ASE identified in this region, FANCAΔ11, involves the skipping of exon 11 (113 bp), which is also expected to produce a truncated protein of 299 aa. It should be noted that both proteins lack the major C-terminal part of the protein which contains particularly a part of the BRCA1 binding site as well as the whole leucine zipper domain and BRG1 binding site region (Figure 1B). As for the exonic 30–31 region, the deletion of exon 30 (FANCAΔ30) is an in frame deletion of 129 bp, while the FANCAΔ31 involves the skipping of

Table 3 – Non-synonymous amino acid changes identified in FANCA protein and prediction of the substitution on protein function using SIFT, PolyPhen and Align GVGD softwares.

SNP[a]	Amino acid change	Domain or binding sites (bs)	SIFT	PolyPhen	Align GVGD
4	Val6Asp	CENP-E and BRCA1 bs	Tolerated	Possibly damaging	Class C0
5	Gly115Arg	CENP-E and BRCA1 bs	Tolerated	Benign	Class C0
8	Arg232Lys	CENP-E and BRCA1 bs	Tolerated	Benign	Class C0
10	Thr266Ala	CENP-E and BRCA1 bs	Tolerated	Benign	Class C0
24	Ala412Val	BRCA1 bs	Tolerated	Benign	Class C0
25	Gly501Ser	BRCA1 bs	Tolerated	Possibly damaging	Class C0
28	Thr561Met	BRCA1 bs	Not tolerated	Probably damaging	Class C65
34	Cys625Ser	None	Not tolerated	Probably damaging	Class C65
35	Pro643Ala	None	Tolerated	Benign	Class C0
37	Met717Ile	None	Tolerated	Benign	Class C0
44	Gly809Asp	None	Tolerated	Benign	Class C0
45	Ser858Arg	None	Tolerated	Probably damaging	Class C0
46	Asp953Glu	None	Tolerated	Benign	Class C0
54	Glu1023Asp	None	Tolerated	Benign	Class C15
57	Ser1088Phe	Leucine zipper domain	Not tolerated	Possibly damaging	Class C0
64	Leu1143Val	BRG1 bs	Tolerated	Benign	Class C0
81	Thr1328Ala	BRG1 bs	Tolerated	Benign	Class C0
83	Ala1352Asp	BRG1 bs	Not tolerated	Benign	Class C15

a According to Table 1.

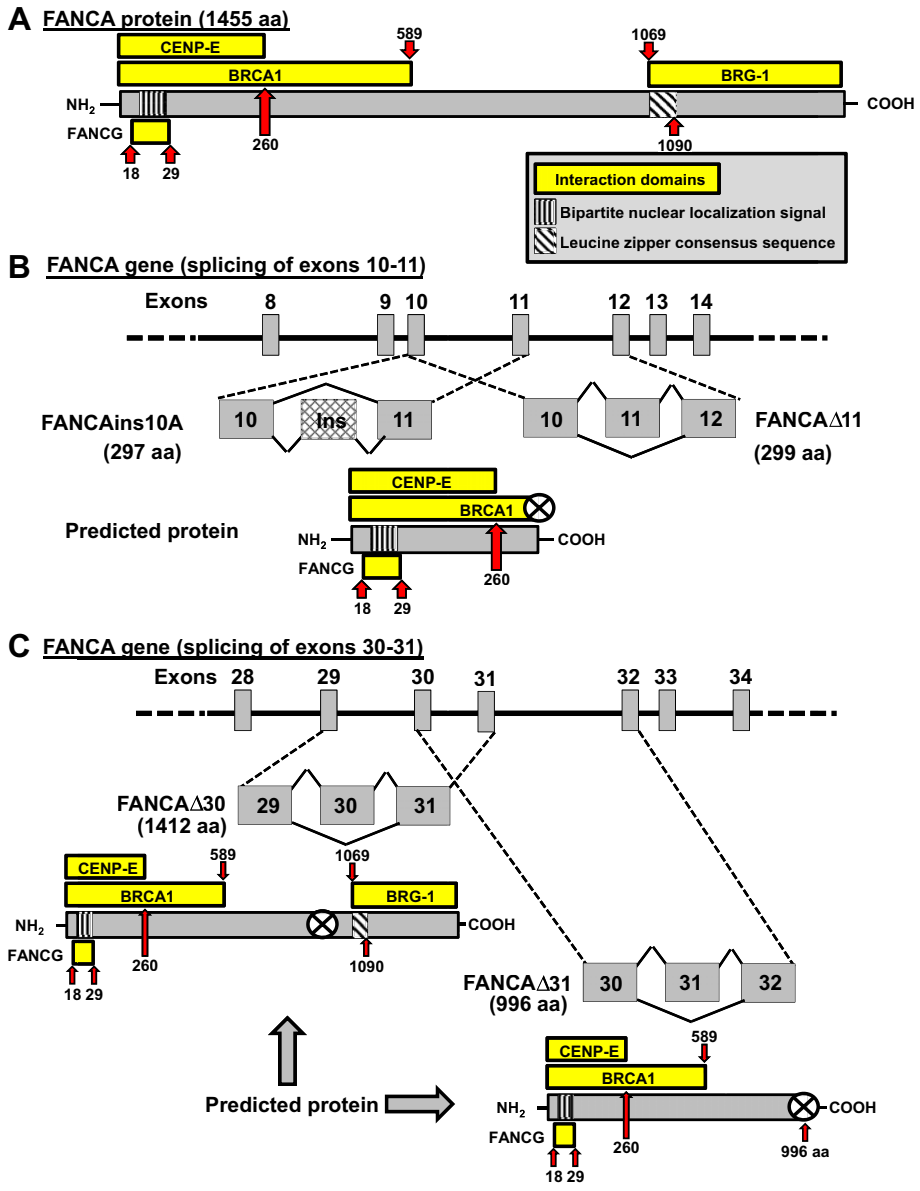

Figure 1 – FANCA interaction domains, and *FANCA* splicing. (A) Schematic representation of FANCA interaction domains based on the literature. (B) *FANCA* exon structure and schematic representation of the putative proteins of the two splicing variants located in the genomic region of exons 9–12, that could be detected via cDNA analyzes: FANCAins10a and FANCAΔ11. (C) *FANCA* exon structure and schematic representation of the putative proteins of the two splicing variants located in the genomic region of exons 29–32, that could be detected via cDNA analyzes: FANCAΔ30 and FANCAΔ31.

exon 31 (85 bp) and could potentially lead to a putative protein of 996 amino acids lacking the BRG1 binding site and the leucine zipper domain (Figure 1C).

Quantitative PCR amplification of the four ASEs described above (FANCAins10A, FANCAΔ11, FANCAΔ30 and FANCAΔ31), performed in eight human breast cancer cell lines (Figure 2) as well as in BRCAX individuals are displayed in Supplemental Figure 1. q-PCR amplification using oligonucleotides specific to each spliced form in the cell lines shows modest expression of FANCAΔ11, FANCAΔ30 and FANCAΔ31 mRNA. However, the variant FANCAins10A exhibits a high mRNA expression level (Figure 2). Relative expression levels of FANCAins10A spliced form are highest in SUM140 cell line and lowest in BT-20 and MDA-231 but no significant variation is observed according to estrogen receptor or differentiation status.

3.3. *Polysome analysis of FANCA ASEs mRNAs*

Given that we could not assume that the spliced forms described above were not subject to mRNA regulation through nonsense mediated decay (NMD), additional experiments have been performed. A series of cell lines from breast cancer patients, in which the alternative form of interest was observed, were treated with puromycin (half of the cells were treated and half were not), an agent known to inhibit NMD. RT-PCR was performed on FANCAins10A, FANCAΔ11, FANCAΔ30 and FANCAΔ31 using specific

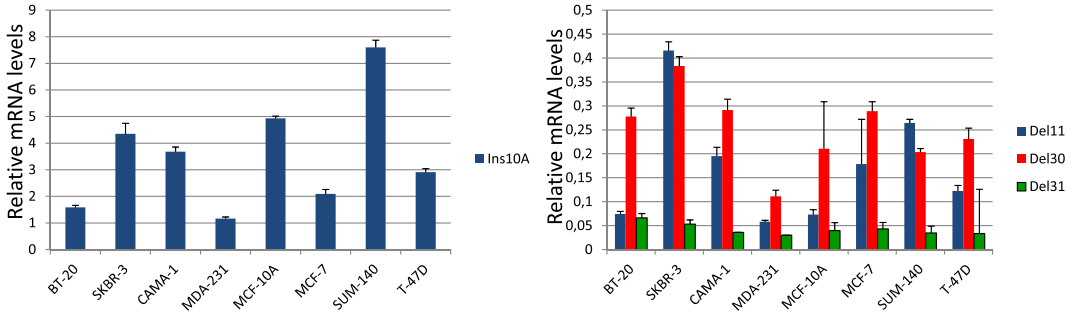

Figure 2 – Expression levels of FANCA Δ11, ins10A, Δ30 and Δ31 spliced forms in cell lines as measured by quantitative real-time PCR experiments. Relative expression levels of FANCA Δ11, ins10A, Δ30 and Δ31 spliced forms were calculated as $E^{(\Delta Ct\ wild\text{-}type\ allele\ -\ \Delta Ct\ splice\ form)}$ in eight breast cancer cell lines where E is primer efficiency.

oligonucleotides to each spliced form, and clearly these forms do not seem to be subject to NMD (data not shown). To further determine whether FANCA splicing variant mRNAs are efficiently translated, polysome profiles were performed in the breast cancer cell line T-47D and mRNA levels were measured using splice variant-specific qRT-PCR primers. Efficiently translated transcripts are associated with polysomes, whereas those degraded by NMD are generally associated with monosomes because they are degraded during the pioneering round of translation, before loading of additional ribosomes. Figure 3 shows the resulting profile with Δ11, ins10A, Δ30 and Δ31 mRNA which are most abundant in polysome fractions 11 to 13 suggesting there are efficiently translated. These spliced mRNAs are thus not degraded by NMD and are associated with translating ribosomes, suggesting efficient production of Δ11, ins10A, Δ30 and Δ31 proteins in T-47D cell line. Interestingly, ribosomal analysis shows that the variant carrying the insertion of the intronic fragment exhibit the highest level of mRNA which is in concordance with the results obtained with quantification of the variants in the different cell lines.

Following identification of ASEs expressed in BRCAX individuals and cancer cell lines, the putative impact of the relevant *FANCA* periexonic and exonic variants (i.e. located in the vicinity of the ASE genomic regions) on mRNA splicing was assessed by *in silico* methods using the HSF web program which allows the identification of enhancer and silencer splicing sites as well as branch point sequences (Table 4). The acceptor and donor motif sequences of exons 10, 10A and 11 have relatively good motif scores, all being over 0.80. Moreover the exon 10A possesses putative branch point sequences

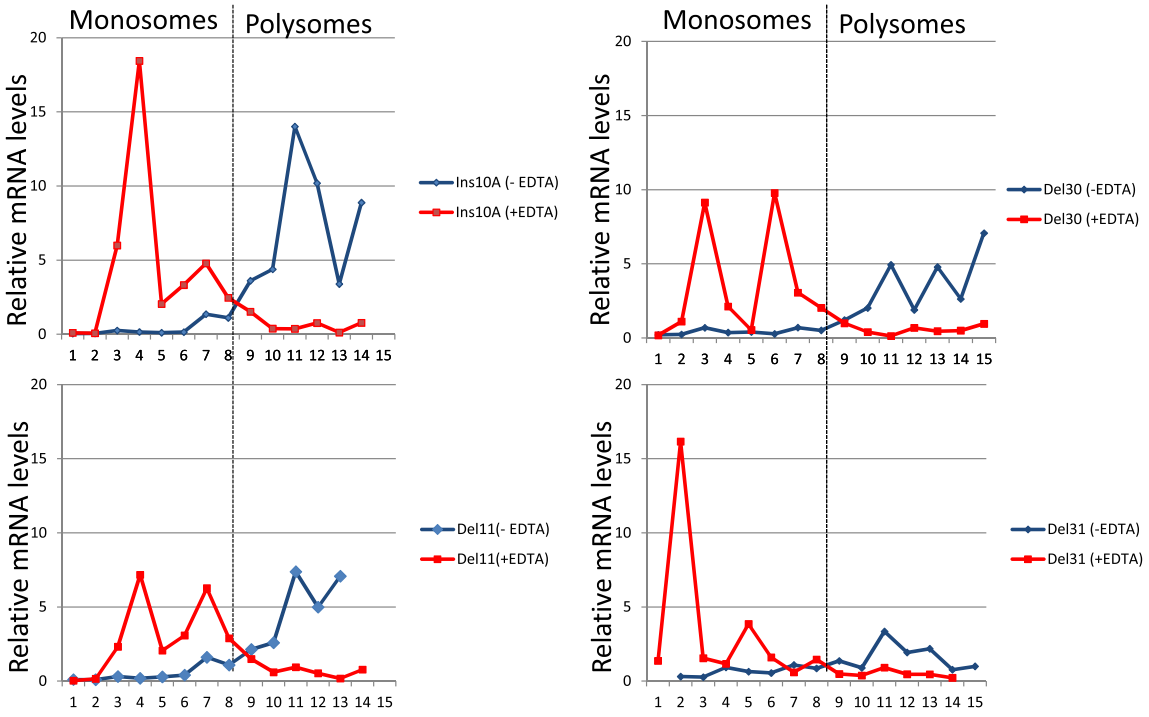

Figure 3 – Polysome analysis of FANCA ASEs mRNAs. Fractionation of monosomes and polysomes (with or without EDTA) as measured by A260. q-RT PCR of FANCA Δ11, ins10A, Δ30 and Δ31 spliced forms was performed on T-47D cells and relative expression values for each fraction were calculated by the equation R = (E)(Ctref−Ct), where Ctref is the average cycle threshold of all the fractions for that variant.

Table 4 – In *silico* analysis of *FANCA* genomic regions and variants potentially involved in the expression of FANCAins10a, FANCAΔ11, FANCAΔ30 and FANCAΔ31 spliced transcripts.

Exon or variant	Location	Motif	Reference score (WT)	Variant score	Potential effect on exon inclusion
Region of exons 10–11					
Exon 10 wt	–	Acceptor site[a]	90.1	–	Positive
		Donor site[a]	86.4	–	Positive
Exon 10A	–	Acceptor site[a]	81.3	–	Positive
		Donor site[a]	86.9	–	Positive
	Exon 10A-80	Branch point site[a]	90.8	–	Positive
	Exon 10A-65	Branch point site[a]	90.7	–	Positive
Exon 11	–	Acceptor site[a]	86.6	–	Positive
		Donor site[a]	88.5	–	Positive
c.894-8A/G	Intron 10	Potential donor site[a]	n.d.	66.6	Not determined
		SF2/ASF (IgM-BRCA1)[a]	n.d.	73.1	Positive (exon 11)
		Enhancer EIE[b]	No value	New site	Positive (exon 11)
		Silencer Motif 2[c]	n.d.	New site (72.3)	Negative (exon 11)
		Silencer IIE[b]	No value	Site broken	Positive (exon 11)
		Silencer IIE[b]	No value	New site	Negative (exon 11)
Region of exons 30–31					
Exon 30	–	Acceptor site[a]	83.3	–	Positive
	–	Donor site[a]	91.8	–	Positive
c.2859 C/G	Exon 30	Potential acceptor site[a]	n.d.	71	Not determined
		SC35[a]	85	Site broken	Negative (exon 30)
		Enhancer PESE octamer[d]	35.8	Site broken	Negative (exon 30)
		Enhancer EIE[b]	No value	Site broken	Negative (exon 30)
		Enhancer 9G8[a]	59.5	Site broken	Negative (exon 30)
		Silencer Motif 1[c]	n.d.	New site (69.7)	Negative (exon 30)
		Silencer Motif 2[c]	n.d.	New site (60.2)	Negative (exon 30)
c.2901 C/T	Exon 30	SRp55[a]	n.d.	New site (76.1)	Positive (exon 30)
		Enhancer PESE octamer[d]	27.8	Site broken	Negative (exon 30)
		Silencer IIE[b]	No value	New site	Negative (exon 30)
Exon 31	–	Acceptor site[a]	92.0	–	Positive
		Donor site[a]	98.8	–	Positive
c.2982-102 G/C	Intron 30	Branch point site[a]	72.2	77.2	Positive (exon 31)
		SRp40[a]	n.d.	New site (85.7)	Positive (exon 31)
		Enhancer PESE octamers[d]	41.1	Site broken	Negative (exon 31)
		Enhancer EIE[b]	No value	Site broken	Negative (exon 31)
		Enhancer 9G8[a]	59.5	Site broken	Negative (exon 31)
		Silencer Motif 3[c]	–	New site (70.0)	Negative (exon 31)
		Silencer Motif 1[c]	65.8	site broken	Positive (exon 31)
		Silencer IIE[b]	No value	Site broken	Positive (exon 31)
		Silencer hnRNP A1[a]	73.3	Site broken	Positive (exon 31)
c.2982-73 G/A	Intron 30	Branch point site[a]	75.7	71.8	Negative (exon 31)
		SF2/ASF (IgM-BRCA1) [a]	84.5	71.5	Negative (exon 31)
		Enhancer PESE Octamer[d]	n.d.	New site (57.6)	Positive (exon 31)
		Enhancer EIE[b]	No value	New site	Positive (exon 31)
		Enhancer 9G8[a]	71.1	Site broken	Negative (exon 31)
		Silencer Motif 1[c]	74.5	Site broken	Positive (exon 31)
		Silencer Motif 2[c]	66.3	Site broken	Positive (exon 31)
		Silencer hnRNP A1[a]	78.6	Site broken	Positive (exon 31)

n.d.: Not detected.
EIE: exon-identity elements.
IIE: Intron-identity elements.
a Based on Human Splicing Finder matrices.
b Zhang C et al., 2008, PNAS 105:5797.
c Sironi M et al., 2004, Nucleic acids research 32:1783.
d Zhang and Chasin 2004, Genes Dev 18:1241.

located 80 and 65 bp upstream of the beginning of exon 10A. The only relevant variant located in intron 10 in the proximity of the acceptor site sequence of exon 11, c.894-8A/G, creates new sites for SF2/ASF (IgM-BRCA1) binding as well as new enhancer EIE and silencer motif 2 binding sites. As for the region containing exons 30–31, the wild type acceptor and donor motif sequences of exons 30 and 31 demonstrate efficient splicing site scores. Two variants located in exon 30, c.2859C/G and c.2901C/T are predicted to have significant effect on binding capacity of several splicing factors. Particularly, c.2859C/G

disrupts motifs for SC35, PESE octamer as well as for enhancer EIE and 9G8. This variant also creates silencer motifs 1 and 2. As for exon 31, the two relevant variants c.2982-102G/C and c.2982-73G/A, located in intron 30 upstream of exon 31, have been evaluated. The c.2982-102G/C variant deletes several consensus motifs such as those for PESE octamer, enhancer EIE and 9G8 as well as silencer motifs 1 and IIE and hnRNP A1 protein (Table 4).

To investigate the potential association between the variants c.2901C/T (rs17226980), c.2982-102G/C (rs12931267) and c.2982-73G/A and the expression of the FANCAΔ30 and FANCAΔ31, real-time PCR was performed using RNA samples from five heterozygotes for c.2901C/T, two heterozygotes for c.2982-102G/C, one heterozygote for c.2982-73G/A and ten wild-type individuals (Supplemental Figure 1). The presence of the Δ30 and Δ31 spliced transcripts were detected in all individuals, including wild-type individuals, supporting the fact that the expression of these variants are not associated with none of these polymorphisms. A similar experiment was performed to analyze the assessment between the c.894-8A/G (rs1164881) variation and the expression of the FANCAΔ11 and FANCAins10A (data not shown) and no significant correlation was observed.

4. Discussion

Over the last decade, several evidences linked the two major breast cancer susceptibility genes *BRCA1* and *BRCA2* to an emerging network of proteins implicated in DNA repair, and whose bi-allelic mutations cause Fanconi Anemia (Wang, 2007). Mono-allelic mutations, like those found in *BRCA1* and *BRCA2/FANCD1*, have been described for *FANCJ*, *FANCN* and *FANCO* in breast cancer susceptibility (Somyajit et al., 2010; Cantor et al., 2001, 2004; Seal et al., 2006; Rahman et al., 2007; Erkko et al., 2007; Pang et al., 2011; Zheng et al., 2010; Meindl et al., 2010), and a recent study conducted with 944 family members being part of the International Fanconi Anemia Registry revealed an increased risk of breast cancer among grandmothers carriers of *FANCC* mutations (Berwick et al., 2007). However, apart from these genes, the involvement of the other *FANC* genes in breast cancer susceptibility remains unclear. Indeed, epidemiological studies focusing on heterozygous mutation carriers have yielded conflicting results (Tischkowitz et al., 2008; Mathew, 2006). Using variant screening and discovery on both genomic and cDNA material of the *FANCA* gene in a cohort of 97 familial breast cancer cases without BRCA1 or BRCA2 mutations as well as among 95 healthy unrelated controls from the same population, we found that: 1) re-sequencing did not identify any deleterious mutations, 2) some variants, and particularly the Gly501Ser non-synonymous change, are associated with a protective effect against breast cancer risk in our cohort and, 3) the *FANCA* gene is subject to multiple alternative splicing events expression. Although extensive promoter screening was beyond the scope of the current study, two variants, including the unreported c.-42-76G/C, were found in the proximal region of the transcriptional starting site. The 13 bp duplication (c.-42-87ins13) has been previously associated with a protective effect in ovarian cancer (OR = 0.72; 95% CI, 0.53–0.99), while no significant effect was seen in breast cancer patients (Thompson et al., 2005). Both promoter variants identified in this study are of interest as even single base changes in promoter sequences can alter regulation of gene expression and contribute to tumorigenesis, particularly if this could affect transcription factor binding site for proteins such as TBP, estrogen receptor, SF1 and CREB as predicted for the c.-42-87ins13 variant (Hasselbach et al., 2005; Najafi and Jangravi, 2010; Zhu et al., 2001; Bond et al., 2004).

Interestingly, among the 85 sequence variants identified in our case and control datasets, the variant showing the most significant *p*-value ($p = 0.003$) is the c.1501G/A (Gly501Ser) non-synonymous change. This variant is over-represented in our control dataset, which is suggestive of a protective effect of the A allele in our cohort. To date, the Gly501Ser variant of *FANCA* has been recently associated with an increased risk of cervical intraepithelial neoplasia grade 3 cancer, with a risk of 1.7 fold (95% CI 1.1–2.6 fold) for the GG genotype when compared to the AA genotype (Wang et al., 2009). Another significant over-representation in the healthy individuals was observed for the c.894-8A/G variation ($p = 0.006$). As predicted by *in silico* analysis, this variant creates a new potential donor site and affects binding site score for SF2/ASF, Enhancer EIE and few silencer proteins. Moreover this variant is located 8 nucleotides upstream of exon 11, which is skipped in the FANCAΔ11 mRNA. Therefore this variant could potentially be involved in the expression of this splicing event. As for the variant c.3348 + 18A/G showing a significant difference of frequency between both series ($p = 0.022$), it creates a new donor site (score = 72.1), a new motif 2 silencer site, and abolishes a EIE site. Among the variants showing a borderline significance, both c.-1C/A and c.1627-32T/C are observed in the case series only. Moreover these variants are carried by the same four individuals likely indicating that both variants are being part of the same allele. The unreported c.-1C/A variation is located in a crucial consensus sequence involved in the initiation of translation. The C nucleotides at position −1 and −2 do not need to be conserved, but contribute to the overall strength (Kozak, 1986). The c.2574C/G variant (Ser858Arg), known as the Indian mutation and located in exon 27, has been previously reported as a disease-causing mutation in Fanconi Anemia (Tamary et al., 2000; Wijker et al., 1999). However this Ser858Arg change is found exclusively in our control dataset, therefore suggesting that this variant at the heterozygous state is non-pathogenic in breast cancer predisposition.

Both the Thr561Met and Cys625Ser changes located within or in the vicinity of the BRCA1 binding site are predicted to be deleterious by all three programs used. Although Thr561Met is found exclusively in three cases, both variants did not show any significant difference in MAF between both cohorts. The Cys625Ser variant has been identified and predicted as pathogenic in the Spanish FA population but this variant was not characterized regarding its effect on protein function (Castella et al., 2011).

Several splicing variants of *FANCA* gene have been identified and partially characterized in the last decade and result from the presence of specific genomic variants. For instance, the c.1567-20A/G (intron 16), c.2278 + 83C/G (intron 28) and c.2222 + 166A/G (intron 24) are known to affect splicing of the FANCA protein (Savino et al., 2003; Bouchlaka et al., 2003).

Several other spliced variants have been reported in the Japanese population (Tachibana et al., 1999; Yagasaki et al., 2004). The significant expression of four splicing variants have been detected in our BRCAX breast cancer individuals and all variants namely FANCAins10A (BU616925), FANCAΔ11 (CN404731), FANCAΔ30 (AK301168) and FANCAΔ31 (BI908441), have been reported in NCBI EST databases but have not been characterized. Of these, only the FANCAΔ30 splice variant results in an in-frame deletion of 43 amino acids located outside known domains, therefore the significance of this deletion is unknown. FANCAins10A, FANCAΔ11 and FANCAΔ30 lead to truncated proteins lacking the BRG-1 binding domain, while the FANCAins10A and FANCAΔ11 truncated proteins have also an excised BRCA-binding domain. Moreover several putative nuclear export sequences located in the region of aa 518, 860 and 1013, are missing in these spliced proteins (Ferrer et al., 2005). Thus, this suggests that interactions with BRCA1 (Folias et al., 2002), the chromatin remodeling BRG1 protein (Otsuki et al., 2001) and with other known protein partners might be affected. No clear polymorphism/mutation seem to be involved in the expression of these spliced transcripts which are not significantly degraded by the NMD mechanism, and are most likely transcribed efficiently as strongly suggested by the polysomal experiments. However, we can not exclude deeper intronic variations affecting intronic splicing enhancers/silencers or creating new cryptic splice sites. Moreover, as reported for some genes related to specific diseases (Ricketts et al., 1987; Naylor et al., 1993; Dietz et al., 1993), the skipping of exons could maintain transcription and translation of a partially functional protein and thus moderate the disease phenotype.

Large intragenic and exonic deletions have been identified in FA patients of different populations such as Afrikaners, Spanish or other European populations (Wijker et al., 1999; Tachibana et al., 1999; Morgan et al., 1999; Tipping et al., 2001; Callén et al., 2004; Centra et al., 1998; Levran et al., 1998; Nakamura et al., 1999). Previous studies have shown that deletions account for 40% of *FANCA* mutations (Morgan et al., 1999). Indeed it is known that the majority of deletions occurred by recombination between two ALU repeats located in *cis* as reported for many other genes (Morgan et al., 1999; Centra et al., 1998; Levran et al., 1998). As demonstrated by Morgan et al (Morgan et al., 1999), recombination hotspots within the *FANCA* gene seem to be located in the genomic regions of introns 17 and 31, which reflect the similar regions of LD block recombination identified in our dataset (data not shown). Regarding the involvement of *FANCA* deletions in breast cancer susceptibility, a novel heterozygous deletion, removing the promoter and 12 exons of the *FANCA* gene, was recently identified in one Finnish breast cancer family (Solyum et al., 2011). However, as reported previously (Seal et al., 2003), MLPA analyzes did not identify any large genomic or exonic deletions within the *FANCA* gene in our French Canadian breast cancer patients.

5. Conclusion

This is the first comprehensive mutation screening of *FANCA* gene in the French Canadian population. Although no deleterious mutation was found, we identify 24 novel polymorphisms not reported in databases including 7 missense variations, as well as four alternative splicing events. It will also be interesting to determine whether inherited polymorphisms in *FANCA* gene resulting in more subtle defects in protein expression or function, can contribute to increased cancer risk or to variable tumor responses to conventional therapies.

Acknowledgments

The authors would like to thank all individuals and families who participated in this study. We thank M. Tranchant, Dr. M. Dumont, and G. Leblanc of the Cancer Genomics Laboratory for sample management and mutation screening. We also thank M. Ouellet, C E Bénard and A.M. Moisan for skillful technical help, Dr. Fabienne Lesueur (IARC, Lyon) for the help with GV-GD analysis and Dr. D. Labuda and C. Moreau at the Centre de Cancérologie Charles Bruneau of Ste-Justine Hospital for help with control DNA samples. We also thank Catherine Laprise (UQAC) for providing genotyping information and Vincent Raymond (CRCHUQ) for providing DNA samples on additional controls from the French Canadian population. The authors also acknowledge the work of Professor Bartha Maria Knoppers and her colleagues regarding ELSI issues, as well as the ethics committees of all participating institutions.

Financial support: This work was supported by the Canadian Institutes of Health Research for the "CIHR Team in Familial Risks of Breast Cancer" program and by the Canadian Breast Cancer Research Alliance, the Fond de la Recherche en Santé du Québec (FRSQ)/Réseau de Médecine Génétique Appliquée (RMGA), the CURE Foundation and Fanconi Canada. N.L. and S.D. hold studentships from Fondation René Bussières and Fondation Desjardins (S.D.), and F.D. held a chercheur-boursier from the Fonds de la Recherche en Santé du Québec (FRSQ).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.molonc.2012.08.002.

Appendix B.

Other members of INHERIT BRCAs involved in clinical aspects of this study:

Paul Bessette: Service de gynécologie, Centre Hospitalier Universitaire de Sherbrooke, Fleurimont, QC, J1H 5N4, Canada.

Jocelyne Chiquette: Clinique des maladies du sein Deschênes-Fabia, Hôpital du Saint-Sacrement, Québec, QC, G1S 4L8, Canada.

Rachel Laframboise: Service de médecine génétique, CHUQ, Pavillon CHUL, Québec, QC, G1V 4G2, Canada.

Jean Lépine: Centre Hospitalier régional de Rimouski, Rimouski, QC, G5L 5T1, Canada.

Bernard Lespérance, Roxane Pichette: Service d'hémato-oncologie, Hôpital du Sacré-Cœur, Montréal, QC, H4J 1C5, Canada.

Marie Plante: Service de gynécologie, CHUQ, L'Hôtel-Dieu de Québec, Québec, QC, G1R 2J6, Canada.

REFERENCES

Akbari, R., Malekzadeh, R., Lepage, P., Roquis, D., Sadjadi, A., Aghcheli, K., Yazdanbod, A., Shakeri, R., Bashiri, J., Sotoudeh, M., Pourshams, A., Ghadirian, P., Narod, S., 2011. Mutations in Fanconi anemia genes and the risk of esophageal cancer. Hum. Genet. 129, 573–582.

Antoniou, A.C., Durocher, F., Smith, P., Simard, J., Easton, D.F., 2006. INHERIT BRCAs program members. BRCA1 and BRCA2 mutation predictions using the BOADICEA and BRCAPRO models and penetrance estimation in high-risk French-Canadian families. Breast Cancer Res. 8 (1), R3. Epub 2005 Dec 12.

Berwick, M., Satagopan, J.M., Ben-Porat, L., Carlson, A., Mah, K., Henry, R., Diotti, R., Milton, K., Pujara, K., Landers, T., Dev Batish, S., Morales, J., Schindler, D., Hanenberg, H., Hromas, R., Levran, O., Auerbach, A.D., 2007. Genetic heterogeneity among Fanconi anemia heterozygotes and risk of cancer. Cancer Res. 67, 9591–9596.

Bond, G.L., Hu, W., Bond, E.E., Robins, H., Lutzker, S.G., Arva, N.C., Bargonetti, J., Bartel, F., Taubert, H., Wuerl, P., Onel, K., Yip, L., Hwang, S.J., Strong, L.C., Lozano, G., Levine, A.J., 2004. A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. Cell 119, 591–602.

Bouchlaka, C., Abdelhak, S., Amouri, A., Ben Abid, H., Hadiji, S., Frikha, M., Ben Othman, T., Amri, F., Ayadi, H., Hachicha, M., Rebaï, A., Saad, A., Dellagi, K., Tunisian Fanconi Anemia Study Group, 2003. Fanconi anemia in Tunisia: high prevalence of group A and identification of new FANCA mutations. J. Hum. Genet. 48, 352–361.

Callén, E., Tischkowitz, M.D., Creus, A., Marcos, R., Bueren, J.A., Casado, J.A., Mathew, C.G., Surrallés, J., 2004. Quantitative PCR analysis reveals a high incidence of large intragenic deletions in the FANCA gene in Spanish Fanconi anemia patients. Cytogenet. Genome Res. 104, 341–345.

Cantor, S.B., Bell, D.W., Ganesan, S., Kass, E.M., Drapkin, R., Grossman, S., Wahrer, D.C., Sgroi, D.C., Lane, W.S., Haber, D.A., Livingston, D.M., 2001. BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. Cell 105, 149–160.

Cantor, S., Drapkin, R., Zhang, F., Lin, Y., Han, J., Pamidi, S., Livingston, D.M., 2004. The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations. Proc. Natl. Acad. Sci. U. S. A. 101, 2357–2362.

Cartegni, L., Wang, J., Zhu, Z., Zhang, M.Q., Krainer, A.R., 2003. ESEfinder: a web resource to identify exonic splicing enhancers. Nucleic Acids Res. 31, 3568–3571.

Castella, M., Pujol, R., Callén, E., Trujillo, J.P., Casado, J.A., Gille, H., Lach, F.P., Auerbach, A.D., Schindler, D., Benítez, J., Porto, B., Ferro, T., Muñoz, A., Sevilla, J., Madero, L., Cela, E., Beléndez, C., de Heredia, C.D., Olivé, T., de Toledo, J.S., Badell, I., Torrent, M., Estella, J., Dasí, A., Rodríguez-Villa, A., Gómez, P., Barbot, J., Tapia, M., Molinés, A., Figuera, A., Bueren, J.A., Surrallés, J., 2011. Origin, functional role, and clinical impact of Fanconi anemia FANCA mutations. Blood 117, 3759–3769.

Centra, M., Memeo, E., d'Apolito, M., Savino, M., Ianzano, L., Notarangelo, A., Liu, J., Doggett, N.A., Zelante, L., Savoia, A., 1998. Fine exon-intron structure of the Fanconi anemia group A (FAA) gene and characterization of two genomic deletions. Genomics 51, 463–467.

De Winter, J.P., van der Weel, L., de Groot, J., Stone, S., Waisfisz, Q., Arwert, F., Scheper, R.J., Kruyt, F.A., Hoatlin, M.E., Joenje, H., 2000. The Fanconi anemia protein FANCF forms a nuclear complex with FANCA, FANCC and FANCG. Hum. Mol. Genet. 9, 2665–2674.

Desjardins, S., Beauparlant, C.J., Labrie, Y., Ouellette, G., Simard, J., BRCAs, I.N.H.E.R.I.T., Durocher, F., 2009. Variations in the Nijmegen Breakage Syndrome gene, NBN/NBS1, and the risk of breast cancer in high-risk non-BRCA1 and BRCA2 French Canadian breast cancer ovarian cancer families. BMC Cancer 9, 181.

Desjardins, S., Belleau, P., Labrie, Y., Ouellette, G., Bessette, P., Chiquette, J., Laframboise, R., Lépine, J., Lespérance, B., Pichette, R., Plante, M., BRCAs, I.N.H.E.R.I.T., Durocher, F., 2008. Genetic variants and haplotype analyses of the ZBRK1/ZNF350 gene in high-risk non BRCA1/2 French Canadian breast and ovarian cancer families. Int. J. Cancer 122, 108–116.

Desmet, F.O., Hamroun, D., Lalande, M., Collod-Béroud, G., Claustres, M., Béroud, C., 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 37, e67.

Dietz, H.C., Valle, D., Francomano, C.A., Kendzior Jr., R.J., Pyeritz, R.E., Cutting, G.R., 1993. The skipping of constitutive exons in vivo induced by nonsense mutations. Science 259, 680–683.

Durocher, F., Labrie, Y., Soucy, P., Sinilnikova, O., Labuda, D., Bessette, P., Chiquette, J., Laframboise, R., Lépine, J., Lespérance, B., Ouellette, G., Pichette, R., Plante, M., Tavtigian, S.V., Simard, J., 2006. Mutation analysis and characterization of ATR sequence variants in breast cancer cases from high-risk French Canadian breast/ovarian cancer families. BMC Cancer 6, 230.

Erkko, H., Xia, B., Nikkila, J., Schleutker, J., Syrjakoski, K., Mannermaa, A., Kallioniemi, A., Pylkas, K., Karppinen, S.M., Rapakko, K., Miron, A., Sheng, Q., Li, G., Mattila, H., Bell, D.W., Haber, D.A., Grip, M., Reiman, M., Jukkola-Vuorinen, A., Mustonen, A., Kere, J., Aaltonen, L.A., Kosma, V.M., Kataja, V., Soini, Y., Drapkin, R.I., Livingston, D.M., Winqvist, R., 2007. A recurrent mutation in PALB2 in Finnish cancer families. Nature 446, 316–319.

Ferrer, M., Rodríguez, J.A., Spierings, E.A., de Winter, J.P., Giaccone, G., Kruyt, F.A., 2005. Identification of multiple nuclear export sequences in Fanconi anemia group A protein that contribute to CRM1-dependent nuclear export. Hum. Mol. Genet. 14, 1271–1281.

Folias, A., Matkovic, M., Bruun, D., Reid, S., Hejna, J., Grompe, M., D'Andrea, A., Moses, R., 2002. BRCA1 interacts directly with the Fanconi anemia protein FANCA. Hum. Mol. Genet. 11, 2591–2597.

Hasselbach, L., Haase, S., Fischer, D., Kolberg, H.C., Stürzbecher, H.W., 2005. Characterisation of the promoter region of the human DNA-repair gene Rad51. Eur. J. Gynaecol. Oncol. 26, 589–598.

Knipscheer, P., Räschle, M., Smogorzewska, A., Enoiu, M., Ho, T.V., Schärer, O.D., Elledge, S.J., Walter, J.C., 2009. The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science 326, 1698–1701.

Kozak, M., 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44, 283–292.

Lensch, M.W., Tischkowitz, M., Christianson, T.A., Reifsteck, C.A., Speckhart, S.A., Jakobs, P.M., O'Dwyer, M.E., Olson, S.B., Le Beau, M.M., Hodgson, S.V., Mathew, C.G., Larson, R.A., Bagby Jr., G.C., 2003. Acquired FANCA dysfunction and cytogenetic instability in adult acute myelogenous leukemia. Blood 102, 7–16.

Levran, O., Diotti, R., Pujara, K., Batish, S.D., Hanenberg, H., Auerbach, A.D., 2005. Spectrum of sequence variations in the FANCA gene: an International Fanconi anemia Registry (IFAR) study. Hum. Mutat. 25, 142–149.

Levran, O., Doggett, N.A., Auerbach, A.D., 1998. Identification of Alu-mediated deletions in the Fanconi anemia gene FAA. Hum. Mutat. 12, 145–152.

Mathe, E., Olivier, M., Kato, S., Ishioka, C., Hainaut, P., Tavtigian, S.V., 2006. Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods. Nucleic Acids Res. 34, 1317–1325.

Mathew, C.G., 2006. Fanconi anemia genes and susceptibility to cancer. Oncogene 25, 5875–5884.

Medhurst, A.L., Huber, P.A., Waisfisz, Q., de Winter, J.P., Mathew, C.G., 2001. Direct interactions of the five known Fanconi anemia proteins suggest a common functional pathway. Hum. Mol. Genet. 10, 423–429.

Medhurst, A.L., Laghmani, el H., Steltenpool, J., Ferrer, M., Fontaine, C., de Groot, J., Rooimans, M.A., Scheper, R.J., Meetei, A.R., Wang, W., Joenje, H., de Winter, J.P., 2006. Evidence for subcomplexes in the Fanconi anemia pathway. Blood 108, 2072–2080.

Meindl, A., Hellebrand, H., Wiek, C., Erven, V., Wappenschmidt, B., Niederacher, D., Freund, M., Lichtner, P., Hartmann, L., Schaal, H., Ramser, J., Honisch, E., Kubisch, C., Wichmann, H.E., Kast, K., Deissler, H., Engel, C., Müller-Myhsok, B., Neveling, K., Kiechle, M., Mathew, C.G., Schindler, D., Schmutzler, R.K., Hanenberg, H., 2010. Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. Nat. Genet. 42, 410–414.

Moldovan, G.L., D'Andrea, A.D., 2009. How the fanconi anemia pathway guards the genome. Annu. Rev. Genet. 43, 223–249.

Morgan, N.V., Tipping, A.J., Joenje, H., Mathew, C.G., 1999. High frequency of large intragenic deletions in the Fanconi anemia group A gene. Am. J. Hum. Genet. 65, 1330–1341.

Najafi, M., Jangravi, Z., 2010. Human PON promoters: from similarity to prediction of polymorphic positions within transcription factor elements. Mini Rev. Med. Chem. 10, 938–945.

Nakamura, A., Matsuura, S., Tauchi, H., Hanada, R., Ohashi, H., Hasagawa, T., Honda, K., Masuno, M., Imaizumi, K., Sugita, K., Ide, T., Komatsu, K., 1999. Four novel mutations of the Fanconi anemia group A gene (FAA) in Japanese patients. J. Hum. Genet. 44, 48–51.

Naylor, J.A., Green, P.M., Rizza, C.R., Giannelli, F., 1993. Analysis of factor VIII mRNA reveals defects in everyone of 28 haemophilia A patients. Hum. Mol. Genet. 2, 11–17.

Ng, P.C., Henikoff, S., 2003. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res. 31, 3812–3814.

Notredame, C., Higgins, D.G., Heringa, J., 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302, 205–217.

Otsuki, T., Furukawa, Y., Ikeda, K., Endo, H., Yamashita, T., Shinohara, A., Iwamatsu, A., Ozawa, K., Liu, J.M., 2001. Fanconi anemia protein, FANCA, associates with BRG1, a component of the human SWI/SNF complex. Hum. Mol. Genet. 10, 2651–2660.

Pang, Z., Yao, L., Zhang, J., Ouyang, T., Li, J., Wang, T., Fan, Z., Fan, T., Lin, B., Xie, Y., 2011 May 20. RAD51C germline mutations in Chinese women with familial breast cancer. Breast Cancer Res. Treat. (Epub ahead of print).

Rahman, N., Seal, S., Thompson, D., Kelly, P., Renwick, A., Elliott, A., Reid, S., Spanova, K., Barfoot, R., Chagtai, T., Jayatilake, H., McGuffog, L., Hanks, S., Evans, D.G., Eccles, D., Breast Cancer Susceptibility Collaboration (UK)Easton, D.F., Stratton, M.R., 2007. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. Nat. Genet. 39, 165–167.

Ramensky, V., Bork, P., Sunyaev, S., 2002. Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 30, 3894–3900.

Raschle, M., Knipscheer, P., Enoiu, M., Angelov, T., Sun, J., Griffith, J.D., Ellenberger, T.E., Schärer, O.D., Walter, J.C., 2008. Cell 134, 969–980.

Reese, M.G., Eeckman, F.H., Kulp, D., Haussler, D., 1997. Improved splice site detection in genie. J. Comput. Biol. 4, 311–323.

Reuter, T., Herterich, S., Bernhard, O., Hoehn, H., Gross, H.J., 2000. Strong FANCA/FANCG but weak FANCA/FANCC interaction in the yeast 2-hybrid system. Blood 95, 719–720.

Ricketts, M.H., Simons, M.J., Parma, J., Mercken, L., Dong, Q., Vassart, G., 1987. A nonsense mutation causes hereditary goitre in the Afrikander cattle and unmasks alternative splicing of thyroglobulin transcripts. Proc. Natl. Acad. Sci. U. S. A. 84, 3181–3184.

Savino, M., Borriello, A., D'Apolito, M., Criscuolo, M., Del Vecchio, M., Bianco, A.M., Di Perna, M., Calzone, R., Nobili, B., Zatterale, A., Zelante, L., Joenje, H., Della Ragione, F., Savoia, A., 2003. Spectrum of FANCA mutations in Italian Fanconi anemia patients: identification of six novel alleles and phenotypic characterization of the S858R variant. Hum. Mutat. 22, 338–339.

Seal, S., Barfoot, R., Jayatilake, H., Smith, P., Renwick, A., Bascombe, L., McGuffog, L., Evans, D.G., Eccles, D., Easton, D.F., Stratton, M.R., Rahman, N., 2003. Breast cancer susceptibility Collaboration. Evaluation of Fanconi anemia genes in familial breast cancer predisposition. Cancer Res. 63, 8596–8599.

Seal, S., Thompson, D., Renwick, A., Elliott, A., Kelly, P., Barfoot, R., Chagtai, T., Jayatilake, H., Ahmed, M., Spanova, K., North, B., McGuffog, L., Evans, D.G., Eccles, D., Breast Cancer Susceptibility Collaboration (UK)Easton, D.F., Stratton, M.R., Rahman, N., 2006. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. Nat. Genet. 38, 1239–1241.

Simard, J., Dumont, M., Moisan, A.M., Gaborieau, V., Malouin, H., Durocher, F., Chiquette, J., Plante, M., Avard, D., Bessette, P., Brousseau, C., Dorval, M., Godard, B., Houde, L., BRCAs, I.N.H.E.R.I.T., Joly, Y., Lajoie, M.A., Leblanc, G., Lépine, J., Lespérance, B., Vézina, H., Parboosingh, J., Pichette, R., Provencher, L., Rhéaume, J., Sinnett, D., Samson, C., Simard, J.C., Tranchant, M., Voyer, P., Easton, D., Tavtigian, S.V., Knoppers, B.M., Laframboise, R., Bridge, P., Goldgar, D., 2007. Evaluation of BRCA1 and BRCA2 mutation prevalence, risk prediction models and a multistep testing approach in French-Canadian families with high risk of breast and ovarian cancer. J. Med. Genet. 44, 107–121.

Smith, P.J., Zhang, C., Wang, J., Chew, S.L., Zhang, M.Q., Krainer, A.R., 2006. An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. Hum. Mol. Genet. 15, 2490–2508.

Solyom, S., Winqvist, R., Nikkilä, J., Rapakko, K., Hirvikoski, P., Kokkonen, H., Pylkäs, K., 2011. Screening for large genomic rearrangements in the FANCA gene reveals extensive deletion in a Finnish breast cancer family. Cancer Lett. 302, 113–118.

Somyajit, K., Subramanya, S., Nagaraju, G., 2010. RAD51C: a novel cancer susceptibility gene is linked to Fanconi anemia and breast cancer. Carcinogenesis 31, 2031–2038.

Sridharan, D., Brown, M., Lambert, W.C., McMahon, L.W., Lambert, M.W., 2003. Nonerythroid alphaII spectrin is required for recruitment of FANCA and XPF to nuclear foci induced by DNA interstrand cross-links. J. Cell. Sci. 116, 823–835.

Stratton, M.R., Rahman, N., 2008. The emerging landscape of breast cancer susceptibility. Nat. Genet. 40, 17–22.

Sunyaev, S., Ramensky, V., Koch, I., Lathe, W., Kondrashov, A.S., Bork, P., 2001. Prediction of deleterious human alleles. Hum. Mol. Genet. 10, 591–597.

Tachibana, A., Kato, T., Ejima, Y., Yamada, T., Shimizu, T., Yang, L., Tsunematsu, Y., Sasaki, M.S., 1999. The FANCA gene in Japanese Fanconi anemia: reports of eight novel mutations and analysis of sequence variability. Hum. Mutat. 13, 237–244.

Tamary, H., Bar-Yam, R., Shalmon, L., Rachavi, G., Krostichevsky, M., Elhasid, R., Barak, Y., Kapelushnik, J., Yaniv, I., Auerbach, A.D., Zaizov, R., 2000. Fanconi anaemia group A (FANCA) mutations in Israeli non-Ashkenazi Jewish patients. Br. J. Haematol. 111, 338–343.

Tavtigian, S.V., Deffenbaugh, A.M., Yin, L., Judkins, T., Scholl, T., Samollow, P.B., de Silva, D., Zharkikh, A., Thomas, A., 2006. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. J. Med. Genet. 43, 295–305.

Thompson, E., Dragovic, R.L., Stephenson, S.A., Eccles, D.M., Campbell, I.G., Dobrovic, A., 2005. A novel duplication polymorphism in the FANCA promoter and its association with breast and ovarian cancer. BMC Cancer 5, 43.

Tipping, A.J., Pearson, T., Morgan, N.V., Gibson, R.A., Kuyt, L.P., Havenga, C., Gluckman, E., Joenje, H., de Ravel, T., Jansen, S., Mathew, C.G., 2001. Molecular and genealogical evidence for a founder effect in Fanconi anemia families of the Afrikaner population of South Africa. Proc. Natl. Acad. Sci. U. S. A 98, 5734–5739.

Tischkowitz, M., Easton, D.F., Ball, J., Hodgson, S.V., Mathew, C.G., 2008. Cancer incidence in relatives of British Fanconi anaemia patients. BMC Cancer 8, 257.

Wang, W., 2007. Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. Nat. Rev. Genet. 8, 735–748.

Wang, S.S., Bratti, M.C., Rodríguez, A.C., Herrero, R., Burk, R.D., Porras, C., González, P., Sherman, M.E., Wacholder, S., Lan, Z.E., Schiffman, M., Chanock, S.J., Hildesheim, A., 2009. Common variants in immune and DNA repair genes and risk for human papillomavirus persistence and progression to cervical cancer. J. Infect. Dis. 199, 20–30.

Wijker, M., Morgan, N.V., Herterich, S., van Berkel, C.G., Tipping, A.J., Gross, H.J., Gille, J.J., Pals, G., Savino, M., Altay, C., Mohan, S., Kokal, I., Cavenagh, J., Marsh, J., van Weel, M., Ortega, J.J., Schuler, D., Samochatova, E., Karwacki, M., Bekassy, A.N., Abecasis, M., Ebell, W., Kwee, M.L., de Ravel, T., Mathew, C.G., 1999. Heterogeneous spectrum of mutations in the Fanconi anemia group A gene. Eur. J. Hum. Genet. 7, 52–59.

Yagasaki, H., Hamanoue, S., Oda, T., Nakahata, T., Asano, S., Yamashita, T., 2004. Identification and characterization of novel mutations of the major Fanconi anemia gene FANCA in the Japanese population. Hum. Mutat. 24, 481–490.

Zheng, Y., Zhang, J., Hope, K., Niu, Q., Huo, D., Olopade, O.I., 2010. Screening RAD51C nucleotide alterations in patients with a family history of breast and ovarian cancer. Breast Cancer Res. Treat. 124, 857–861.

Zhu, Y., Spitz, M.R., Lei, L., Mills, G.B., Wu, X., 2001. A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter enhances lung cancer susceptibility. Cancer Res. 61, 7825–7829.