

FANCD2 associated with sporadic breast cancer risk

E.Barroso¹, R.L.Milne², L.P.Fernández¹, P.Zamora³,
J.I.Arias⁴, J.Benítez^{1,2} and G.Ribas^{1,*}

¹Human Genetics Group, ²National Genotyping Centre (CeGen), Human Cancer Genetics Programme, Spanish National Cancer Centre (CNIO), Madrid, Spain, ³Department of Oncology, La Paz Hospital, Madrid, Spain and ⁴Unit of Surgery, Monte Naranco Hospital, Oviedo, Spain

*To whom correspondence should be addressed at: Human Genetics Group, Human Cancer Genetics Program, C/Melchor Fdz Almagro, 3, E-28029, Madrid, Spain. Tel: +34 912246950; Email: gribas@cnio.es

Several components of the Fanconi anaemia (FA) family of proteins allow the formation of the DNA repair complex *foci* formed by proteins such as BRCA1/2 and RAD51. Because the genes that participate in the DNA repair pathway have been described as low-penetrance breast cancer susceptibility genes, we postulated that variants in FA genes could also be associated with sporadic breast cancer risk. We studied seven SNPs in *FANCA*, *FANCL* and *FANCD2* in a total of 897 consecutive and non-related sporadic breast cancer cases and 1033 unaffected controls from the Spanish population. We observed a statistically significant association with sporadic breast cancer for the variant rs2272125 (L1366L) located on *FANCD2* (OR per allele = 1.35; 95% C.I. 1.09–1.67; $P = 0.005$). Both haplotype and diplotype analyses confirmed this association, where one haplotype and pooled diplotypes carrying it were associated with more than 4-fold risk ($P = 0.007$ and $P = 0.006$, respectively). Screening for potential causal variants in *FANCD2* was performed, detecting one in the putative promoter region, which is located in a phylogenetically conserved motif with consensus binding sites for some transcriptional factors, suggesting a functional implication. Our data indicate that a relationship between *FANCD2* and sporadic breast cancer risk may exist.

Introduction

Breast cancer is one of the most common diseases among women in developed countries. Several breast cancer susceptibility genes have been identified, the most important being the high risk tumour suppressor genes *BRCA1* (1) and *BRCA2* (2), the latter also identified as *FANCD1* (3,4). However, under the polygenic model of complex diseases, there are also low risk genes with variants that are common in the general population and that produce a mild susceptibility to the disease. Whether these variants interact together or with other environmental and lifestyle factors in an additive or multiplicative way, is not yet clear (5–8).

Abbreviations: FA, Fanconi anaemia; DSB, double-strand break; SNP, single nucleotide polymorphism. HR, homologous recombination; LD, linkage disequilibrium; MAF, minor allele frequency.

The *BRCA1* and *BRCA2* (*FANCD1*) proteins are implicated in several cellular mechanisms, the most relevant being double-strand breaks (DSBs) by homologous recombination (HR) (9,10). Both *BRCA1* and *BRCA2* interact with the *RAD51* protein, a key component of HR DNA damage repair, to form the nuclear repair *foci* after DNA damage (11–14). Furthermore, the FA protein family has been identified as a component of this DSB HR repair machinery (15–17). To date, 12 complementation groups have been described, for which 11 responsible FA genes have been identified: *FANCA*, -B, -C, -D1, -D2, -E, -F, -G, -J, -L and -M. Eight of these (*FANCA*, -B, -C, -E, -F, -G, -L and -M) form a multi-subunit nuclear FA complex (18–20). Some of the FA proteins are structural proteins in the FA complex. For example, *FANCA* acts as dimer, *FANCG* plays an important role in the maintenance of the complex structure (21,20) and *FANCF* functions as a flexible adaptor protein which is required for the assembly of this complex (22). Other FA proteins have specific functions such as *FANCL*, a ubiquitin ligase (23,24), whose stability and function depends on *FANCB* (25,26). Another example is *FANCD2*, a monoubiquitinated protein, which is activated by *FANCL* and passes from isoform *FANCD2-S* to *FANCD2-L* (27,16). *FANCD2* is located downstream of the FA complex and linked with it through *FANCE* (20). *FANCI*, described previously as helicase *BACH1/BRIP1* (28,29), and *FANCM*, a nuclease that may act as an engine that translocates the FA core complex along DNA, have been most recently identified (30).

The pathway activity that implicates FA proteins is initiated when DNA damage occurs, and *ATM* and *ATR* protein kinases function as DNA damage sensors (31,32). They phosphorylate *FANCD2* which is required for the activation of an S-phase checkpoint (33,34) and for its appropriate monoubiquitination through the FA complex (35). They also phosphorylate several proteins in the FA complex (36) which are necessary for the activation of *FANCL* ubiquitin ligase activity. After monoubiquitination of *FANCD2-S*, *FANCD2-L* translocates to the nuclear repair complex *foci*, helped by *BRCA1* in its formation and accumulation (37), where *BRCA2* and *RAD51* are also located (27). Thus, the monoubiquitination of *FANCD2* is important because it allows its chromatin-associated *foci* assembly (38,39), either in response to DNA damage (40) during the S-phase of the normal cell cycle (41,42), or in replicative stress (43). It has therefore been proposed that *FANCD2* could mediate the *BRCA2*–*RAD51* interaction and also influence the speed and efficiency of DNA repair (17,40,44) although *FANCG* may also be implicated in this process (45). However, an alternative model proposes that *FANCD2* functions in a DNA repair sub-pathway that is independent of DNA cross-over and sister chromatid exchange, where *RAD51* focus formation depends on *BRCA2*, but not *FANCD2* (46). It has been demonstrated that FA proteins are not essential for HR repair machinery. Furthermore, FA cells have a mild defect in the single-strand annealing pathway, which is also

involved in DSB repair. Thus, FA proteins could work in other DNA damage repair pathways of HR (47).

We focused our study on the hypothesis that genes involved in the mechanism of DNA damage repair, acting similarly to tumour suppressors, would act as low susceptibility genes in the development of sporadic breast cancer, as proposed previously (48). In this study, we analysed seven genetic variants in three FA genes: two variants in *FANCA*, three in *FANCL*, and two in *FANCD2*. The first two genes encode proteins in the FA complex and the latter encodes a protein downstream of the FA complex.

Materials and methods

Subjects

This case-control study included a total of 897 consecutive and non-related sporadic breast cancer cases and 1033 healthy controls. The case samples had a mean age of diagnosis of 50 years (range 23–90 years) (Table I). They were collected from seven Spanish public hospitals: 258 (29%) from *Hospital Monte Naranco*, in Oviedo; 164 (18%) from *Fundación Jiménez Díaz*, 132 (15%) from *Hospital La Paz* and 100 (11%) from *Hospital Clínico San Carlos*, all in Madrid; 115 (13%) from *Hospital San Pau*, in Barcelona; 100 (11%) from *Hospital de Salamanca*, in Salamanca; 28 (3%) from *Hospital de Valencia*, in Valencia. Controls (between 22 and 85 years of age) (Table I) were unaffected Spanish women collected from six centres: 462 (45%) from the Menopause Research Centre at the *Instituto Palacios*; 287 (28%) from the National Blood Transfusion Centre; 221 (21%) from *Fundación Jiménez Díaz*; 63 (6%) from *Hospital Clínico San Carlos*, all in Madrid.

Extraction of genomic DNA

Genomic DNA from cases and controls was isolated from peripheral blood lymphocytes using the MagNa Pure LC Instrument (Roche Molecular Biochemicals, Germany) following manufacturer's instructions (49), and a stock solution of DNA at 50 ng/μl was obtained, measured using the PicoGreen dsDNA Quantitation Kit (Molecular Probes, Eugene, OR).

SNP selection

Three public databases available via internet were used to collect information about single nucleotide polymorphisms (SNPs) in the three genes (*FANCA*, *FANCL* and *FANCD2*): NCBI (<http://www.ncbi.nlm.nih.gov>), Ensembl (<http://www.ensembl.org>) and HapMap (<http://www.hapmap.org>). SNPs were considered if they had a minor allele frequency (MAF) greater than or equal to 10%, the exception being putative coding SNPs with MAF $\geq$ 5%. Information on linkage disequilibrium (LD) (blocks within chromosomes that are more probably linked and segregate together than expected at random), was used to select representative SNPs in high LD blocks. Haploview v3.11 (50) was used to select tag-SNPs (a subset of SNPs that represent variation in a gene or chromosomal region) where possible. We selected and successfully genotyped two SNPs on the 3' downstream region of *FANCA* (rs7195066 and rs2239359, spaced 13 Kb apart), three located on introns 1, 7 and 8 of *FANCL* (rs848289, rs1011314 and rs3771224, respectively, spaced over 68 Kb) and two in *FANCD2* (rs722509 and rs2272125, spaced 32 Kb apart), one on intron 19 and the other a synonymous coding change (L1366L) on exon 42. Gene structures and SNP positions are shown in Figure 1A.

Genotyping assays

Genotyping was carried out using the TaqMan platform. SNP assays were designed using Applied Biosystems Assay-by-design and Assay-on-Demand probes (Applied Biosystems, Foster City, CA, USA) (primer and probe information is provided in Supplementary Table I). PCR amplifications were carried out following the manufacturer's standardized protocol. The genotype of each sample was automatically assigned by measuring allele-specific fluorescence in the ABI Prism 7900HT Detection System, using the SDSv2.1 (Sequence Detection Software) for allele discrimination, by Applied Biosystems. Three non-DNA and nine duplicate samples were included in each plate of 384 samples used, as an internal control.

Statistical analysis

Deviations from Hardy-Weinberg equilibrium were tested using the *genhwi* command in STATAv8.0 (51), applying Pearson's Chi-squared test or Fisher's Exact Test where appropriate. Differences in the MAFs between cases and controls were tested using the Chi-squared test. In order to assess associations

Table I. Age categories of control and case samples

Age (years)	Controls		Cases		Total	
	N	(%)	N	(%)	N	(%)
<40	261	(25)	209	(23)	470	(24)
40–44	61	(6)	30	(3)	91	(5)
45–49	90	(9)	45	(5)	135	(7)
50–54	156	(15)	72	(8)	228	(12)
55–59	147	(14)	74	(8)	221	(12)
60–64	91	(9)	72	(8)	163	(8)
65+	22	(2)	112	(13)	134	(7)
Unknown	205	(20)	283	(32)	488	(25)
Total	1033	(100)	897	(100)	1930	(100)

between genotypes, haplotypes (combination of variants along a chromosome) and diplotypes (haplotype combinations in individuals), and breast cancer risk, several analysis were performed. PHASEv2.0 software (52,53) was used to impute haplotypes and compare their frequency distributions in cases and controls. Unconditional logistic regression was applied using STATAv8.0 for the analysis of genotypes, haplotypes and diplotypes, the most likely imputed haplotype being assumed for the latter two. Odds ratios (OR), their 95% confidence intervals (CI) and Wald statistic *P*-values were estimated, using the most frequent genotype, haplotype and diplotype, respectively, among controls as reference. In the analysis of genotypes, dominant, recessive and multiplicative (single parameter) codominant models were assessed against the two parameter codominant model via the likelihood ratio test, and where the latter was not the best fit (nominal *P*-value greater than 0.05), the best fitting model was determined by parsimony. Age was included as a categorical variable (<40, 40–44, 45–49, 50–54, 55–59, 60–64, $\geq$ 65) (Table I). Analysis of haplotypes was repeated using the *haplo.stats* library implemented in R which compares haplotype frequencies in cases and controls in an unbiased way by including haplotype uncertainty in the estimation of ORs. For all statistical tests, results were considered significant for two-sided *P*-values less than 0.05.

Screening for functional variants in the *FANCD2* gene

Haploview v3.11 was used to graphically analyse the extent of LD in the *FANCD2* region (Figure 1B). Selected sequences in *FANCD2* (promoter region and nine exons) were further analysed by PCR, dHPLC and sequencing in samples carrying haplotype H2 in the *FANCD2* gene. Primers in flanking exons, intron-exon boundaries and the promoter region were designed using Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) (54), based on genomic information collected from NCBI (<http://www.ncbi.nlm.nih.gov>). PCR products, obtained under standardized conditions, were analysed by dHPLC using the WAVE DNA-fragment analysis system (Transgenomics Limited, Crewe, UK), and were determined by the WAVEMAKER software from Transgenomics (PCR primers and conditions, and dHPLC elution conditions are showed in Supplementary Table II). Those DNAs with heterozygous peaks detected by dHPLC analysis were purified using the High Pure PCR Product Purification Kit (Roche) and sequenced with the ABI Prism Big DyeTM System and ABI 3700 sequencer (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocol. Additional validation for the identified variants ss49839682, rs17032268, rs9811771 and rs7647987 in *FANCD2* was carried out by dHPLC analysis (conditions shown in Supplementary Table II) in 95 case and 95 control non-H2 carriers. A sample of 551 case and 559 control non-H2 carriers were further genotyped using TaqMan probes (Supplementary Table I) in order to validate the results obtained for ss49839682 and rs9811771. *FANCD2* variants were assessed for functional importance using PupaSNP (<http://pupasnp.bioinfo.ocha.fib.es>), which synthesizes information from: Ensembl v25.34e, dbSNP 122, HapMap release #12, Match v2.2, Transfac Professional v8.3, Haploview v3.0 and PMUT (55). Transfac Professional v8.3 was also applied to screen for transcription factor consensus binding sites using rVista v2.0 (<http://rvista.dcode.org>) (56). Phylogenetic conservation analysis was performed using ECR Browser (<http://ecrbrowser.dcode.org>) (57). Phylogenetic sequence comparisons and construction of phylogenetic trees among mammalian organisms were carried out using Tcoffee (<http://figs-server.cnrs-mrs.fr/Tcoffee/tcoffee.cgi/index.cgi>) (58) (see Figures 2A and B and Supplementary Figure 1).

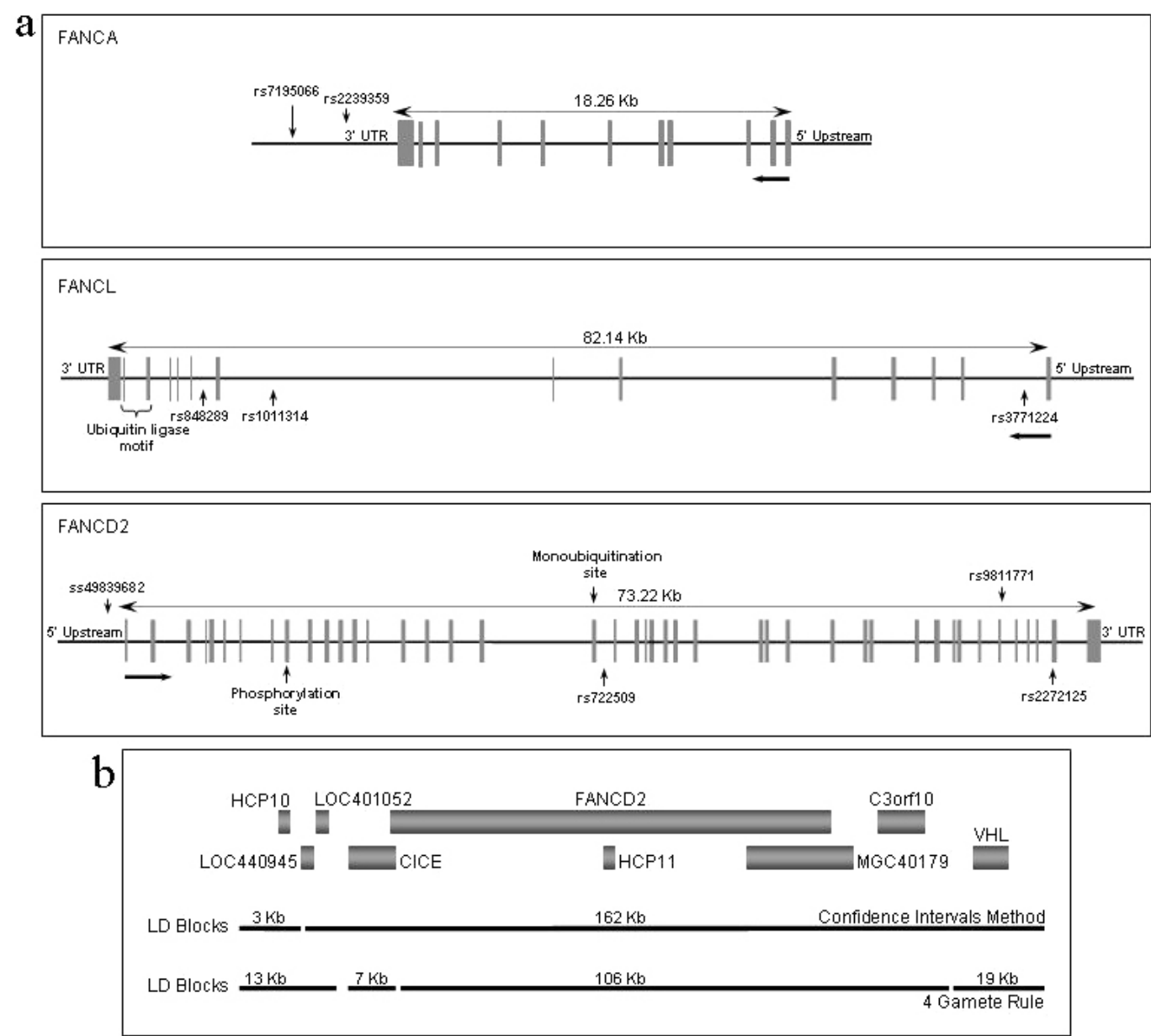


Fig. 1. (A) Gene structure and SNP location. Schematic diagram of the *FANCA*, *FANCL* and *FANCD2* genes. Grey boxes represent exons and solid lines represent untranslated and intron regions, with 5' upstream and 3'-UTR regions noted. Fine and arrowed lines represent gene lengths and solid arrows represent 5'–3' gene orientation. Functional motifs and assayed SNP locations are noted. (B) LD blocks located in the *FANCD2* gene region. Schematic diagram of *FANCD2* and surrounding genes located in the same LD block. Grey boxes represent genes. Solid lines represent LD blocks obtained by the Confidence Intervals Method or the 4 Gamete Rule.

Results

Genotyping

Genotyping call rates were over 97%, and concordance rates for non-DNA and duplicate samples were 100%, for all SNPs genotyped (data are shown in Table II). All seven SNPs complied with Hardy–Weinberg equilibrium among controls.

Analysis of genotypes

A higher MAF was observed among cases than controls for rs2272125 in *FANCD2* (0.13 versus 0.10, respectively, $P = 0.005$). The other six polymorphisms tested did not show any evidence of differences in the allelic frequencies (data not shown). Comparison of genotype frequencies suggested that the risk of developing sporadic breast cancer

associated with rs2272125 followed a codominant model, with the estimated OR for rare CC homozygotes relative to common AA homozygotes ($OR = 2.77$; 95% CI = 1.08–7.13; $P = 0.03$) more than double that for heterozygotes ($OR = 1.28$; 95% CI = 1.02–1.62; $P = 0.04$). The best fitting model was a single parameter codominant one, with an associated multiplicative per-allele OR of 1.35 (95% CI = 1.09–1.67, $P = 0.005$). This association was consistently observed after adjusting for age (Table I) ($OR = 1.32$, $P = 0.03$). We did not find any other associations among the remaining SNPs in this study. Results for all genotyped SNPs are summarized in Table II.

The program Structure v.2.0 was applied to a set of 28 unlinked biallelic markers genotyped in a random subset of 163 cases and 180 controls, finding no evidence of population substructure, with an estimated posterior probability of

Table II. Statistical analysis of genotypes in *FANCA*, *FANCL* and *FANCD2*

Gene	SNP name	Spanish control MAF	Genotype	Controls		Cases		Call rate (%)	OR (95% CI)	P-value
				N	(%)	N	(%)			
<i>FANCA</i>	rs7195066	0.29	CC	269	(49)	274	(50)	99.5	1.00 (Reference)	0.65
			CT	239	(43)	230	(42)		0.94 (0.74–1.21)	
			TT	44	(8)	45	(8)		1.00 (0.64–1.57)	
	rs2239359	0.31	CC	268	(49)	251	(46)	99.3	1.00 (Reference)	0.47
			CT	223	(40)	229	(42)		1.10 (0.85–1.41)	
			TT	60	(11)	64	(12)		1.14 (0.77–1.69)	
<i>FANCL</i>	rs848289	0.30	TT	267	(49)	252	(47)	98.9	1.00 (Reference)	0.60
			TA	236	(43)	238	(44)		1.07 (0.83–1.37)	
			AA	47	(9)	51	(9)		1.15 (0.75–1.77)	
	rs1011314	0.30	GG	276	(50)	257	(47)	99.2	1.00 (Reference)	0.51
			GA	223	(40)	226	(42)		1.09 (0.85–1.40)	
			AA	52	(9)	60	(11)		1.24 (0.42–1.86)	
	rs3771224	0.30	CC	262	(48)	240	(45)	97.1	1.00 (Reference)	0.30
			CT	233	(43)	238	(45)		1.12 (0.87–1.43)	
			TT	46	(9)	52	(10)		1.23 (0.80–1.90)	
<i>FANCD2</i>	rs722509	0.11	TT	739	(79)	650	(77)	97.1	1.00 (Reference)	0.20
			TA	181	(19)	185	(22)		1.16 (0.92–1.46)	
			AA	11	(1)	12	(1)		1.24 (0.54–2.83)	
	rs2272125	0.10	AA	675	(80)	649	(75)	97.4	1.00 (Reference)	0.61
			AC	162	(19)	200	(23)		1.28 (1.02–1.62)	
			CC	6	(1)	16	(2)		2.77 (1.08–7.13)	

Genotype counts, call rates and frequencies in cases and controls with ORs and *P*-values for tests of association between genotypes and sporadic breast cancer.

OR, odds ratio; CI, confidence interval.

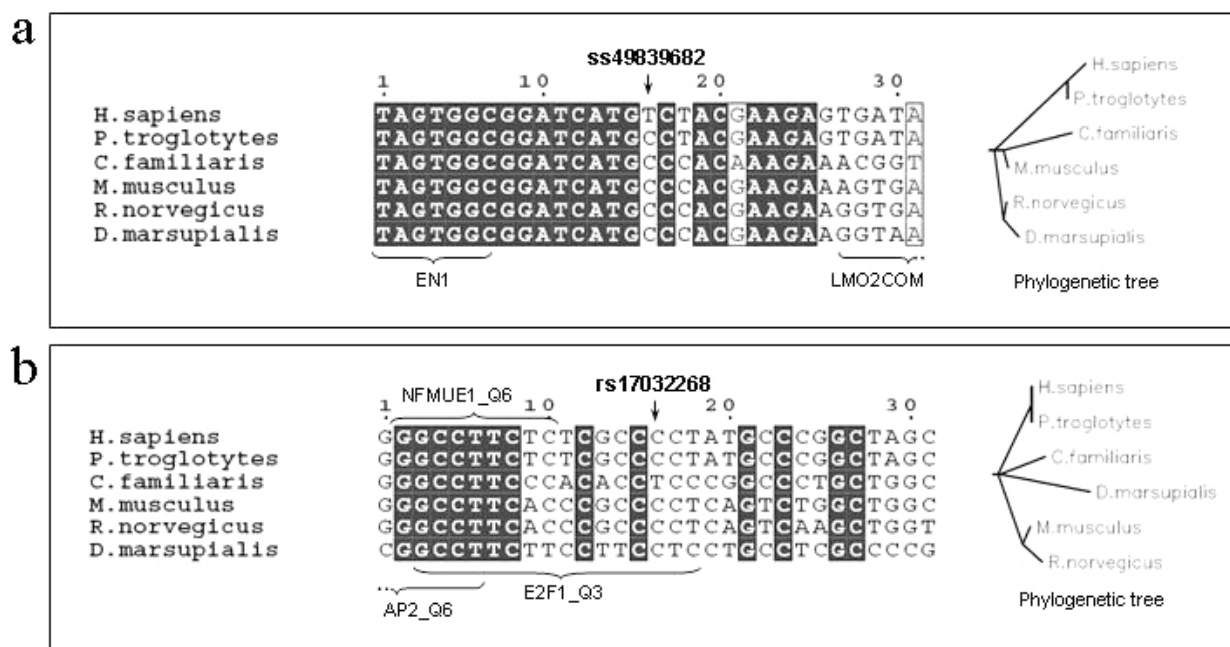


Fig. 2. Phylogenetic analyses of ss49839682 (A) and rs17032268 (B) sequence conservation in mammal organisms. Sequence comparison of surrounding variants and transcription factor consensus binding sites. Phylogenetic trees as measures of divergence are noted. Grey boxes represent highly conserved sequences. SNP sites and transcription factor binding sites are noted.

approximately one for K (number of clusters) = 1, and minimum and maximum ancestry coefficients of 0.45 and 0.55, respectively for K = 2, consistent across independent runs (R.L. Milne unpublished data).

Analysis of haplotypes and diplotypes

We compared the imputed haplotypes and diplotypes formed by the SNPs in *FANCD2* between cases and controls and the

results are presented in Table III. The most common haplotype (H1) found in both cases and controls carried the more common alleles 'T' in rs722509 and 'A' in rs2272125. Comparison of haplotype frequencies between cases and controls suggested that one of the four haplotypes was associated with the risk of breast cancer. This haplotype (H2) carried the combination 'TC' and the estimated OR relative to H1 was 4.46 (95% CI = 1.51–13.21; P = 0.007) assuming the most

Table III. Statistical analysis of haplotypes (a) and diplotypes (b) in the *FANCD2* gene. Bold typeface indicate statistically significant haplotype and diplotype analyses.

	Controls		Cases		OR (95% CI)	P-value
	N	(%)	N	(%)		
(a) Haplotype ^a						
H1	1328	(89)	1340	(87)	1.00 (Reference)	
H2	4	(0)	18	(1)	4.46 (1.51–13.21)	0.007
H3	4	(0)	3	(0)	0.74 (0.17–3.33)	0.70
H4	150	(10)	185	(12)	1.22 (0.97–1.54)	0.09
(b) Diplotype ^b						
D1-1	592	(80)	581	(75)	1.00 (Reference)	
D1-2	4	(1)	14	(2)	3.57 (1.17–10.90)	0.03
D1-3	2	(0)	3	(0)	1.53 (0.25–9.18)	0.64
D1-4	138	(19)	161	(21)	1.19 (0.92–1.53)	0.18
D2-4	0	(0)	4	(1)	—	—
D3-3	1	(0)	0	(0)	—	—
D4-4	6	(1)	10	(1)	1.70 (0.61–4.70)	0.31
D2-	4	(1)	18	(2)	4.59 (1.54–13.63)	0.006

Counts and frequencies in cases and controls with ORs and P-values for tests of association with sporadic breast cancer.
OR, odds ratio; CI, confidence interval.
D2- represent people carrying at least one haplotype H2 (D1-2 and D2-4).
^aH1 = TA; H2 = TC; H3 = AA; H2 = AC.
^bD1-1 = H1/H1; D1-2 = H1/H2; D1-3 = H1/H3; D1-4 = H1/H4; D2-4 = H2/H4; D3-3 = H3/H3; D4-4 = H4/H4.

likely haplotypes were actually observed, and 4.81 (95% CI = 1.63–14.18; *P* = 0.004) allowing for haplotype uncertainty [Table III (a)].
Analysis of diplotypes found that carrying H1 and H2 (diplotype D1–2) was associated with increased risk of breast cancer relative to D1–1 (OR 3.57; 95% CI = 1.17–10.90; *P* = 0.03). Moreover, only four women carried H2 in a diplotype other than D1–2, all in combination with H4 (D2–4) and all four were cases. Pooling all carriers of H2 (D1–2 and D2–4) confirmed that they were at increased risk (OR = 4.59; 95% CI = 1.54–13.63; *P* = 0.006) [Table III (b)].

Screening for putative causal mutations in FANCD2 risk haplotype carriers

rs2272125 is unlikely to directly cause the observed increased breast cancer risk because it is a synonymous coding SNP. Furthermore, we evaluated its potential implication in exon–intron splicing and found no exonic splicing enhancer (ESE) motifs in that region of *FANCD2*. The observed association might be due to a different causal variant in the same LD block as rs2272125, which extends over the whole region including *FANCD2* (over 100kb, see Figure 1B). Because it was unclear where the causative variation could be located, we searched for putative causal variants in the region among the carriers of the risk haplotype H2 (D1–2 and D2–4). We studied the described functional regions in the *FANCD2* protein by DNA sequencing: the phosphorylation motif, located in exon 9, and the monoubiquitination motif, located in exon 19 (Figure 1A). We did not detect any variation from the reference sequence used (NC_000004).

We also sought to characterize the region surrounding rs2272125, (from exon 38 to exon 43), as well as the promoter (1500 bp upstream of ATG) and 3'-UTR regions, the latter two as putative functional domains. Three variants were found in the putative promoter region: rs781871 (MAF not described),

rs17032268 (NCBI MAF = 0.22) and a novel SNP ss49839682 (–544 bp). A transitional change A > G in intron 38, 13 bp from exon 38 (IV38+13), was identified in the public databases as rs9811771, with unknown MAF. Finally, four SNPs in the 3'-UTR region were detected, one (rs7647987) located before the polyA sequence (NCBI MAF = 0.12), and three others located after the polyA sequence: rs3826, rs11716842, and a novel polymorphism ss49839681 (+421 bp from the stop codon), all with unknown MAF (Table IV).
All these variants detected in the H2 carriers were in high LD and, according to PHASEv2.0, the variant-haplotype VH2 (CTGGAG) was observed (see areas highlighted in grey in Table IV). Four individuals carried sections of VH2 and a fifth carried the entire haplotype, together representing 30% of the H2 carriers.
We genotyped four of these eight variants in 190 non-carriers of H2 (95 cases and 95 controls) in order to investigate whether any appeared accumulated in H2 carriers. These variants were: ss49839682, rs17032268, rs9811771 and rs7647987, chosen because our hypothesis was focused on detecting variants with functional implications that also shared the main part of VH2 (Table IV). Only two out of the four variants selected (ss49839682 and rs9811771) appeared to be more frequent in H2 carriers than non-carriers although neither was statistically significant (OR = 2.04; 95% CI = 0.73–5.66; *P* = 0.2; and OR = 1.68; 95% CI = 0.61–4.62; *P* = 0.3, respectively). We subsequently genotyped a sample of 1100 DNAs for these two SNPs, but did not observe statistical significance, although the tendency of the minor allele accumulating in H2 carriers was also observed (Table V).
Analysis of sequence conservation was performed for these four additional *FANCD2* variants in several mammal organisms. Only ss49839682 and rs17032268 (and their surrounding regions) have been phylogenetically maintained in mammals. Furthermore, these two variants located in the promoter region contain several consensus motifs for transcriptional factors (Figure 2A and B, respectively). The variant rs9811771, located near the exon–intron splicing site, was only weakly conserved in *Pan troglodites* and *Canis familiaris* and rs7647987, located at the 3'polyA site, was not conserved at all (see Supplementary Figures 1A and B, respectively).

Discussion

Previous studies have demonstrated that several genes involved in DNA repair, other than *BRCA1* or *BRCA2*, are related to breast cancer susceptibility (59–63). *BRCA2* has been identified as *FANCD1* (3), and several interactions between both *BRCA1* and *BRCA2* and FA proteins have been described (64,45,40,65). It is therefore plausible that alterations in other FA genes might be implicated in breast cancer risk (66). It is known that alterations in FA genes are linked to several types of cancer, included breast cancer, in the general (non-FA) population (67,68). For example, the absence of *FANCF* by aberrant promoter methylation (69,70), and *BACH1/FANCI* mutations (71) have both been found in breast cancer cases. Furthermore, one study found that *FANCD2* is more often expressed in high proliferation rate tissues, such as mammary glands (72), and mouse model studies have shown that a germline disruption of the murine *Fancd2* gene results in ovarian and epithelial breast cancers (73,74). However, despite the evidence in the literature about

Table IV. Variants found by dHPLC and sequencing in samples carrying haplotype H2.

Sample	Promoter			Intron 38		Exon 43		
	rs781871	ss49839682	rs17032268	rs9811771	rs7647987	ss49839681	rs3826	rs11716842
Major allele	T	T	C	A	A	T	T	C
Case-1	T/T	T/C	C/T	A/A	A/A	—	—	—
Case-2	T/A	T/T	C/C	A/A	A/A	—	—	—
Case-3	T/T	T/C	C/T	A/G	A/G	T/A	T/G	C/T
Case-4	T/A	T/T	C/C	A/A	A/A	T/A	T/T	C/T
Case-5	T/T	T/T	C/C	A/A	A/A	T/T	T/T	C/C
Case-6	—	T/T	C/C	A/A	A/A	T/T	T/T	C/C
Case-7	T/T	T/T	C/C	A/G	A/A	T/T	T/T	C/C
Case-8	T/T	T/T	C/C	A/A	A/A	—	—	—
Case-9	—	—	—	A/A	A/A	—	—	—
Case-10	—	T/T	C/C	A/A	A/A	T/T	T/T	C/C
Case-11	—	—	—	A/A	A/A	—	—	—
Case-12	T/A	T/T	C/C	A/A	A/G	T/T	T/T	C/C
Case-13	T/T	T/C	C/T	A/G	A/G	T/T	T/T	C/C
Case-14	T/T	T/T	C/C	A/G	A/G	T/A	T/G	C/C
Control-1	T/T	T/C	C/C	A/A	A/G	T/T	T/T	C/C
Control-2	T/T	T/C	C/T	A/G	A/G	—	—	—
Control-3	T/T	T/T	C/C	A/A	A/A	T/A	T/T	C/C

—, unknown genotype, not amplified.

Samples sharing sections of the same haplotype (VH2) highlighted in grey.

Variants between the dashed lines were selected for further study.

Table V. Observed allele frequencies for four putative functional SNPs, by H2 carrier status

Variant	H2 carriers		190 H2 non-carriers		1100 H2 non-carriers	
	N_{ma}/N_a	MAF	N_{ma}/N_a	MAF	N_{ma}/N_a	MAF
ss49839682	5/30	0.17*	34/380	0.09*	248/2166	0.11*
rs17032268	4/26	0.13	34/346	0.10	—	—
rs9811771	5/34	0.15**	35/376	0.09**	265/2154	0.12**
rs7647987	6/28	0.18	54/298	0.18	—	—

N_{ma} , observed number of minor alleles; N_a , total number of alleles; MAF, minor allele frequency.

* P (comparison of MAFs) = 0.2 and 0.4, using 190 and 1083 non-H2 carrier samples, respectively.

** P (comparison of MAFs) = 0.3 and 0.7, using 188 or 1077 non-H2 carrier samples, respectively.

the potential susceptibility to breast and other cancers associated with the FA family (75), there have been few studies investigating these associations (76).

In this study, we have investigated associations between genetic variants in three genes belonging to the FA family and sporadic breast cancer in the Spanish population. We genotyped seven SNPs in *FANCA*, *FANCL* and *FANCD2*. Whereas no associations with breast cancer risk were found for SNPs in *FANCA* and *FANCL*, we were able to identify an increased risk associated with the minor 'C' allele of rs2272125 in *FANCD2* (OR = 1.35 per allele) (Table II). Both haplotype and diplotype analyses confirmed this association, with carriers of at least one copy of the H2 haplotype at an estimated greater than 4-fold risk of breast cancer ($P = 0.007$ and $P = 0.006$, respectively) (Table III).

Despite being located in a coding region, the variant rs2272125 is of a synonymous nature and it is not located in any ESE motif. A likely explanation for this association is therefore that there is a causative functional change located in close proximity or in LD with this SNP. The region of high

LD that includes rs2272125 is over 100 Kb and spans the entire *FANCD2* gene. Applying two methods of LD inference, the Confidence Intervals Method (77) and the 4 Gametes Rule (78), it was found that this LD block contains *FANCD2* and three other genes of unknown function. *VHL*, a gene close to *FANCD2* which has been implicated in some cancers (79,80), was excluded because it was not included in the LD block according to the latter method (Figure 1B). We therefore focused on *FANCD2*, our hypothesis being that this gene could carry the putative causal variant.

In-depth analysis of several regions of the *FANCD2* gene among H2 carriers allowed us to detect another eight variants, some of them described previously in the SNP databases, but others identified for the first time (ss49839682 and ss49839681) (Table IV). Four of these variants could have functional implications because of their location (promoter, exon–intron splicing site, and 3'-UTR regions). In addition, they seemed to form a variant-haplotype (VH2) within H2 carriers (Table IV). We therefore investigated how frequent these variants were in people not carrying the H2 haplotype. One of the two variants located in the likely promoter region fell in a strongly phylogenetically conserved motif with consensus binding sites for several transcriptional factors (Figure 2A), which made us speculate that it may have functional implications; however, we have been unable to validate statistically this hypothesis (Table V). It is also possible that other variants along *FANCD2* or the immediate surrounding region could be responsible for the increased risk associated with this gene.

Another, less likely possibility, is that the observed association could be due to confounding by unmeasured factors, such as race. Although population stratification was tested for and not detected, this we done using only a subset of the sample and using only 28 markers. Population stratification cannot therefore be entirely discarded, although it appears unlikely to have influenced the results of this study. Alternatively, the observed association could be due to chance. Replication of these findings using samples from other populations is required to address these two concerns.

In conclusion, we found that a synonymous variant (rs2272125) and haplotype (H2) in the *FANCD2* gene are associated with sporadic breast cancer. This is the first study to identify this association. We have also identified a putative functional variant (ss49839682) in the same high LD block, but were not able to conclude that this is causal. Further studies will be required to clarify the full implication of the *FANCD2* protein in sporadic breast cancer risk.

Supplementary material

Supplementary material is available online at <http://www.carcin.oupjournals.org/>.

Acknowledgments

This study was supported by the grant BFI2003-03852. EB and LPF were funded by the *Comunidad Autónoma de Madrid* and by the Spanish Ministry of Science and Technology (MCT), respectively. We would like to thank Álvaro Ruibal (University Hospital, Santiago de Compostela), and Santiago Palacios (*Instituto Palacios*, Madrid) for the use of samples of cases and controls. We would also like to thank Raquel Rodríguez-López for her contribution to the beginning of this work, and Fátima Mercadillo, Alicia Barroso, Emilio González, Jesús López, Victoria Fernández and Rocío Letón for their expert technical skills.

Conflict of Interest Statement: None declared.

References

- Miki,Y., Swensen,J., Shattuck-Eidens,D. *et al.* (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, **266**, 66–71.
- Wooster,R., Bignell,G., Lancaster,J., Swift,S., Seal,S., Mangion,J., Collins,N., Gregory,S., Gumbs,C. and Micklem,G. (1995) Identification of the breast cancer susceptibility gene BRCA2. *Nature*, **378**, 789–792.
- Howlett,N.G., Taniguchi,T., Olson,S. *et al.* (2002) Biallelic inactivation of BRCA2 in Fanconi anemia. *Science*, **297**, 606–609.
- Stewart,G. and Elledge,S.J. (2002) The two faces of BRCA2, a FANCD2 discovery. *Mol. Cell*, **10**, 2–4.
- Risch,N. (2001) The genetic epidemiology of cancer: interpreting family and twin studies and their implications for molecular genetic approaches. *Cancer Epidemiol. Biomarkers Prev.*, **10**, 733–741.
- Antoniou,A.C., Pharoah,P.D., McMullan,G., Day,N.E., Ponder,B.A. and Easton,D. (2001) Evidence for further breast cancer susceptibility genes in addition to BRCA1 and BRCA2 in a population-based study. *Genet. Epidemiol.*, **21**, 1–18.
- Antoniou,A.C. and Easton,D.F. (2003) Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet. Epidemiol.*, **25**, 190–202.
- Comings,D.E., Gade-Andavolu,R., Cone,L.A., Muhleman,D. and MacMurray,J.P. (2003) A multigene test for the risk of sporadic breast carcinoma. *Cancer*, **97**, 2160–2170.
- Liang,F., Han,M., Romanienko,P.J. and Jasin,M. (1998) Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proc. Natl Acad. Sci. USA*, **95**, 5172–5177.
- Moynahan,M.E., Chiu,J.W., Koller,B.H. and Jasin,M. (1999) Brca1 controls homology-directed DNA repair. *Mol. Cell*, **4**, 511–518.
- Marmorstein,L.Y., Ouchi,T. and Aaronson,S.A. (1998) The BRCA2 gene product functionally interacts with p53 and RAD51. *Proc. Natl Acad. Sci. USA*, **95**, 13869–13874.
- Yuan,S.S., Lee,S.Y., Chen,G., Song,M., Tomlinson,G.E. and Lee,E.Y. (1999) BRCA2 is required for ionizing radiation-induced assembly of Rad51 complex *in vivo*. *Cancer Res.*, **59**, 3547–3551.
- Moynahan,M.E., Pierce,A.J. and Jasin,M. (2001) BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol. Cell*, **7**, 263–272.
- Davies,A.A., Masson,J.Y., McIlwraith,M.J., Stasiak,A.Z., Stasiak,A., Venkitaraman,A.R. and West,S.C. (2001) Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol. Cell*, **7**, 273–282.
- Digweed,M., Rothe,S., Demuth,I., Scholz,R., Schindler,D., Stumm,M., Grompe,M., Jordan,A. and Sperling,K. (2002) Attenuation of the formation of DNA-repair foci containing RAD51 in Fanconi anaemia. *Carcinogenesis*, **23**, 1121–1126.
- D'Andrea,A.D. and Grompe,M. (2003) The Fanconi anaemia/BRCA pathway. *Nat. Rev. Cancer*, **3**, 23–34.
- Wang,X. and D'Andrea,A.D. (2004) The interplay of Fanconi anemia proteins in the DNA damage response. *DNA Repair (Amst)*, **3**, 1063–1069.
- Joenje,H., Oostra,A.B., Wijkers,M. *et al.* (1997) Evidence for at least eight Fanconi anemia genes. *Am. J. Hum. Genet.*, **61**, 940–944.
- Medhurst,A.L., Huber,P.A., Waisfisz,Q., de Winter,J.P. and Mathew,C.G. (2001) Direct interactions of the five known Fanconi anaemia proteins suggest a common functional pathway. *Hum. Mol. Genet.*, **10**, 423–429.
- Gordon,S.M. and Buchwald,M. (2003) Fanconi anemia protein complex: mapping protein interactions in the yeast 2- and 3-hybrid systems. *Blood*, **102**, 136–141.
- Nakanishi,K., Moran,A., Hays,T., Kuang,Y., Fox,E., Garneau,D., de Oca,R.M., Grompe,M. and D'Andrea,A.D. (2001) Functional analysis of patient-derived mutations in the Fanconi anemia gene, FANCG/XRCC9. *Exp. Hematol.*, **29**, 842–849.
- Leveille,F., Blom,E., Medhurst,A.L. *et al.* (2004) The Fanconi anemia gene product FANCF is a flexible adaptor protein. *J. Biol. Chem.*, **279**, 39421–39430.
- Meetei,A.R., de Winter,J.P., Medhurst,A.L. *et al.* (2003) A novel ubiquitin ligase is deficient in Fanconi anemia. *Nat. Genet.*, **35**, 165–170.
- Meetei,A.R., Yan,Z. and Wang,W. (2004) FANCL replaces BRCA1 as the likely ubiquitin ligase responsible for FANCD2 monoubiquitination. *Cell Cycle*, **3**, 179–181.
- Meetei,A.R., Levitus,M., Xue,Y. *et al.* (2004) X-linked inheritance of Fanconi anemia complementation group B. *Nat. Genet.*, **36**, 1219–1224.
- Fei,P., Yin,J. and Wang,W. (2005) New advances in the DNA damage response network of Fanconi anemia and BRCA proteins. FAAP95 replaces BRCA2 as the true FANCB protein. *Cell Cycle*, **4**, 80–86.
- Garcia-Higuera,I., Taniguchi,T., Ganesan,S., Meyn,M.S., Timmers,C., Hejna,J., Grompe,M. and D'Andrea,A.D. (2001) Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Mol. Cell*, **7**, 249–262.
- Litman,R., Peng,M., Jin,Z., Zhang,F., Zhang,J., Powell,S., Andreassen,P.R. and Cantor,S.B. (2005) BACH1 is critical for homologous recombination and appears to be the Fanconi anemia gene product FANCF. *Cancer Cell*, **8**, 255–265.
- Bridge,W.L., Vandenberg,C.J., Franklin,R.J. and Hiom,K. (2005) The BRP1 helicase functions independently of BRCA1 in the Fanconi anemia pathway for DNA crosslink repair. *Nat. Genet.*, **37**, 953–957.
- Meetei,A.R., Medhurst,A.L., Ling,C. *et al.* (2005) A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat. Genet.*, **37**, 958–963.
- Morrison,C., Sonoda,E., Takao,N., Shinohara,A., Yamamoto,K. and Takeda,S. (2000) The controlling role of ATM in homologous recombinational repair of DNA damage. *EMBO J.*, **19**, 463–471.
- Cliby,W.A., Roberts,C.J., Cimprich,K.A., Stringer,C.M., Lamb,J.R., Schreiber,S.L. and Friend,S.H. (1998) Overexpression of a kinase-inactive ATR protein causes sensitivity to DNA-damaging agents and defects in cell cycle checkpoints. *EMBO J.*, **17**, 159–169.
- Grompe,M. (2002) FANCD2: a branch-point in DNA damage response? *Nat. Med.*, **8**, 555–556.
- Taniguchi,T., Garcia-Higuera,I., Xu,B., Andreassen,P.R., Gregory,R.C., Kim,S.T., Lane,W.S., Kastan,M.B. and D'Andrea,A.D. (2002) Convergence of the fanconi anemia and ataxia telangiectasia signaling pathways. *Cell*, **109**, 459–472.
- Andreassen,P.R., D'Andrea,A.D. and Taniguchi,T. (2004) ATR couples FANCD2 monoubiquitination to the DNA-damage response. *Genes Dev.*, **18**, 1958–1963.
- Venkitaraman,A.R. (2004) Tracing the network connecting BRCA and Fanconi anaemia proteins. *Nat. Rev. Cancer*, **4**, 266–276.
- Vandenberg,C.J., Gergely,F., Ong,C.Y., Pace,P., Mallery,D.L., Hiom,K. and Patel,K.J. (2003) BRCA1-independent ubiquitination of FANCD2. *Mol. Cell*, **12**, 247–254.
- Montes de Oca,R., Andreassen,P.R., Margossian,S.P., Gregory,R.C., Taniguchi,T., Wang,X., Houghtaling,S., Grompe,M. and D'Andrea,A.D. (2005) Regulated interaction of the Fanconi anemia protein, FANCD2, with chromatin. *Blood*, **105**, 1003–1009.
- Park,W.H., Margossian,S., Horwitz,A.A., Simons,A.M., D'Andrea,A.D. and Parvin,J.D. (2005) Direct DNA binding activity of the Fanconi anemia D2 protein. *J. Biol. Chem.*, **280**, 23593–23598.
- Hussain,S., Wilson,J.B., Medhurst,A.L. *et al.* (2004) Direct interaction of FANCD2 with BRCA2 in DNA damage response pathways. *Hum. Mol. Genet.*, **13**, 1241–1248.

41. Taniguchi,T., Garcia-Higuera,I., Andreassen,P.R., Gregory,R.C., Grompe,M. and D'Andrea,A.D. (2002) S-phase-specific interaction of the Fanconi anemia protein, FANCD2, with BRCA1 and RAD51. *Blood*, **100**, 2414–2420.
42. Rothfuss,A. and Grompe,M. (2004) Repair kinetics of genomic interstrand DNA cross-links: evidence for DNA double-strand break-dependent activation of the Fanconi anemia/BRCA pathway. *Mol. Cell. Biol.*, **24**, 123–134.
43. Howlett,N.G., Taniguchi,T., Durkin,S.G., D'Andrea,A.D. and Glover,T.W. (2005) The Fanconi anemia pathway is required for the DNA replication stress response and for the regulation of common fragile site stability. *Hum. Mol. Genet.*, **14**, 693–701.
44. Wang,X., Andreassen,P.R. and D'Andrea,A.D. (2004) Functional interaction of monoubiquitinated FANCD2 and BRCA2/FANCD1 in chromatin. *Mol. Cell. Biol.*, **24**, 5850–5862.
45. Hussain,S., Witt,E., Huber,P.A., Medhurst,A.L., Ashworth,A. and Mathew,C.G. (2003) Direct interaction of the Fanconi anaemia protein FANCG with BRCA2/FANCD1. *Hum. Mol. Genet.*, **12**, 2503–2510.
46. Ohashi,A., Zdzienicka,M.Z., Chen,J. and Couch,F.J. (2005) Fanconi anemia complementation group D2 (FANCD2) functions independently of BRCA2- and RAD51-associated homologous recombination in response to DNA damage. *J. Biol. Chem.*, **280**, 14877–14883.
47. Nakanishi,K., Yang,Y.G., Pierce,A.J., Taniguchi,T., Digweed,M., D'Andrea,A.D., Wang,Z.Q. and Jasin,M. (2005) Human Fanconi anemia monoubiquitination pathway promotes homologous DNA repair. *Proc. Natl Acad. Sci. USA*, **102**, 1110–1115.
48. Kuschel,B., Auranen,A., McBride,S. *et al.* (2002) Variants in DNA double-strand break repair genes and breast cancer susceptibility. *Hum. Mol. Genet.*, **11**, 1399–1407.
49. Kessler,H.H., Muhlbauer,G., Stelzl,E., Daghofer,E., Santner,B.I. and Marth,E. (2001) Fully automated nucleic acid extraction: MagNA Pure LC. *Clin. Chem.*, **47**, 1124–1126.
50. Barrett,J.C., Fry,B., Maller,J. and Daly,M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.
51. StataCorp. (2003) *Stata Statistical Software: Release 8.0*TX: Stata Corporation, College Station.
52. Stephens,M., Smith,N.J. and Donnelly,P. (2001) A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.*, **68**, 978–989.
53. Stephens,M. and Donnelly,P. (2003) A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am. J. Hum. Genet.*, **73**, 1162–1169.
54. Rozen,S. and Skaletsky,H. (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.*, **132**, 365–386.
55. Conde,L., Vaquerizas,J.M., Santoyo,J., Al-Shahrour,F., Ruiz-Llorente,S., Robledo,M. and Dopazo,J. (2004) PupaSNP Finder: a web tool for finding SNPs with putative effect at transcriptional level. *Nucleic Acids Res.*, **32**, W242–W248.
56. Loots,G.G. and Ovcharenko,I. (2004) rVISTA 2.0: evolutionary analysis of transcription factor binding sites. *Nucleic Acids Res.*, **32**, W217–W221.
57. Ovcharenko,I., Nobrega,M.A., Loots,G.G. and Stubbs,L. (2004) ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. *Nucleic Acids Res.*, **32**, W280–W286.
58. Poirot,O., Suhre,K., Abergel,C., O'Toole,E. and Notredame,C. (2004) 3DCoffee@igs: a web server for combining sequences and structures into a multiple sequence alignment. *Nucleic Acids Res.*, **32**, W37–W40.
59. Easton,D.F. (1994) Cancer risks in A-T heterozygotes. *Int. J. Radiat. Biol.*, **66**, S177–S182.
60. Thorstenson,Y.R., Roxas,A., Kroiss,R. *et al.* (2003) Contributions of ATM mutations to familial breast and ovarian cancer. *Cancer Res.*, **63**, 3325–3333.
61. Rodriguez-Lopez,R., Osorio,A., Sanchez-Pulido,L., De La Hoya,M., Barroso,A., Caldes,T. and Benitez,J. (2003) No mutations in the XRCC2 gene in BRCA1/2-negative high-risk breast cancer families. *Int. J. Cancer*, **103**, 136–137.
62. Smith,T.R., Miller,M.S., Lohman,K., Lange,E.M., Case,L.D., Mohrenweiser,H.W. and Hu,J.J. (2003) Polymorphisms of XRCC1 and XRCC3 genes and susceptibility to breast cancer. *Cancer Lett.*, **190**, 183–190.
63. Meijers-Heijboer,H., van den Ouweland,A., Klijn,J. *et al.* (2002) Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat. Genet.*, **31**, 55–59.
64. Folias,A., Matkovic,M., Bruun,D., Reid,S., Hejna,J., Grompe,M., D'Andrea,A. and Moses,R. (2002) BRCA1 interacts directly with the Fanconi anemia protein FANCA. *Hum. Mol. Genet.*, **11**, 2591–2597.
65. Levran,O., Attwooll,C., Henry,R.T. *et al.* (2005) The BRCA1-interacting helicase BRIP1 is deficient in Fanconi anemia. *Nat. Genet.*, **37**, 931–933.
66. Venkitaraman,A.R. (2002) Connecting Fanconi's anaemia to breast cancer predisposition. *Lancet*, **360**, 1344–1345.
67. Collet-Durel,S., Guitton,N., Nourgalieva,K., Leveque,J., Danic,B. and Chenal,C. (2001) Genomic instability and breast cancer. *Oncol. Rep.*, **8**, 1001–1005.
68. Alter,B.P., Greene,M.H., Velazquez,I. and Rosenberg,P.S. (2003) Cancer in Fanconi anemia. *Blood*, **101**, 2072.
69. Van Der Heijden,M.S., Brody,J.R. and Kern,S.E. (2004) Functional screen of the fanconi anemia pathway in cancer cells by Fancd2 immunoblot. *Cancer Biol. Ther.*, **3**, 534–537.
70. Turner,N., Tutt,A. and Ashworth,A. (2004) Hallmarks of 'BRCAness' in sporadic cancers. *Nat. Rev. Cancer*, **4**, 814–819.
71. Lewis,A.G., Flanagan,J., Marsh,A., Pupo,G.M., Mann,G., Spurdle,A.B., Lindeman,G.J., Visvader,J.E., Brown,M.A. and Chenevix-Trench,G. (2005) Mutation analysis of FANCD2, BRIP1/BACH1, LMO4 and SFN in familial breast cancer. *Breast Cancer Res.*, **7**, R1005–R1016.
72. Holzel,M., van Diest,P.J., Bier,P., Wallisch,M., Hoatlin,M.E., Joenje,H. and de Winter,J.P. (2003) FANCD2 protein is expressed in proliferating cells of human tissues that are cancer-prone in Fanconi anaemia. *J. Pathol.*, **201**, 198–203.
73. Houghtaling,S., Timmers,C., Noll,M., Finegold,M.J., Jones,S.N., Meyn,M.S. and Grompe,M. (2003) Epithelial cancer in Fanconi anemia complementation group D2 (Fancd2) knockout mice. *Genes Dev.*, **17**, 2021–2035.
74. Houghtaling,S., Granville,L., Akkari,Y., Torimaru,Y., Olson,S., Finegold,M. and Grompe,M. (2005) Heterozygosity for p53 (Trp53+/–) accelerates epithelial tumor formation in fanconi anemia complementation group D2 (Fancd2) knockout mice. *Cancer Res.*, **65**, 85–91.
75. Seal,S., Barfoot,R., Jayatilake,H. *et al.* (2003) Evaluation of Fanconi Anemia genes in familial breast cancer predisposition. *Cancer Res.*, **63**, 8596–8599.
76. Thompson,E., Dragovic,R.L., Stephenson,S.A., Eccles,D.M., Campbell,I.G. and Dobrovic,A. (2005) A novel duplication polymorphism in the FANCA promoter and its association with breast and ovarian cancer. *BMC Cancer*, **5**, 43.
77. Gabriel,S.B., Schaffner,S.F., Nguyen,H. *et al.* (2002) The structure of haplotype blocks in the human genome. *Science*, **296**, 2225–2229.
78. Wang,N., Akey,J.M., Zhang,K., Chakraborty,R. and Jin,L. (2002) Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. *Am. J. Hum. Genet.*, **71**, 1227–1234.
79. Kim,W.Y. and Kaelin,W.G. (2004) Role of VHL gene mutation in human cancer. *J. Clin. Oncol.*, **22**, 4991–5004.
80. Otis,C.N., Krebs,P.A., Albuquerque,A., Quezado,M.M., San Juan,X., Sobel,M.E. and Merino,M.J. (2002) Loss of heterozygosity of p53, BRCA1, VHL, and estrogen receptor genes in breast carcinoma: correlation with related protein products and morphologic features. *Int. J. Surg. Pathol.*, **10**, 237–245.

Received November 29, 2005; revised March 21, 2006;
accepted April 19, 2006