Fanconi anemia is characterized by congenital abnormalities, bone marrow failure, and cancer predisposition. To investigate the origin, functional role, and clinical impact of FANCA mutations, we determined a FANCA mutational spectrum with 130 pathogenic alleles. Some of these mutations were further characterized for their distribution in populations, mode of emergence, or functional consequences at cellular and clinical level. The world most frequent FANCA mutation is not the result of a mutational “hot-spot” but results from worldwide dissemination of an ancestral Indo-European mutation. We provide molecular evidence that total absence of FANCA in humans does not reduce embryonic viability, as the observed frequency of mutation carriers in the Gypsy population equals the expected by Hardy-Weinberg equilibrium. We also prove that long distance Alu-Alu recombination can cause Fanconi anemia by originating large interstitial deletions involving FANCA and 2 adjacent genes. Finally, we show that all missense mutations studied lead to an altered FANCA protein that is unable to relocate to the nucleus and activate the FA/BRCA pathway. This may explain the observed lack of correlation between type of FANCA mutation and cellular phenotype or clinical severity in terms of age of onset of hematologic disease or number of malformations.

Introduction

Fanconi anemia (FA) is a rare genetic disease characterized by congenital abnormalities, bone marrow failure, and predisposition to cancer. Fourteen FA complementation groups and their corresponding genes have been identified (FANC-A, -B, -C, -D1/BRCA2, -D2, -E, -F, -G, -I, -J/BRIP1, -I, -M, -N/PALB2, and -O/RAD51C). Proteins encoded by these genes participate in repair of DNA interstrand crosslinks and are essential for accurate processing of stalled replication forks. Four of the FA genes (FANCD1/BRCA2, FANCI/BRIP1, FANCN/PALB2, and FANCO/RAD51C) are known breast cancer susceptibility genes, connecting FA proteins to homologous recombination repair.1,3

The incidence of FA is calculated to be approximately 1 to 5 cases in 1 million births. FA-A is the most frequent complementation group representing approximately two-thirds of the patients in the majority of countries. FA-C and FA-G are also frequent, each accounting for 10% to 15% of the patients, whereas the other groups are rare.4,5 A higher incidence of FA is observed in some particular ethnic groups, such as Spanish Gypsies,6 Ashkenazi Jews,7 and Afrikaaners from South Africa8 because of founder effects associated with a particular mutation.

FA-A is usually associated with private FANCA mutations in individual families. Thus, the number of different pathogenic variants described for the FANCA gene is very high considering the relatively low number of patients.9-14 Mutation type is also heterogeneous, including point mutations, small insertions/deletions, splicing mutations, and large intragenic deletions. Therefore, it is necessary to combine several methodologies for mutation screening. Determination of family pathogenic variants is the ultimate confirmation of diagnosis and can be used as a strategy for subtyping or confirming it, and is necessary for molecular prenatal or preimplantation tests and mutation carrier detection.

Complementation group frequencies of the Spanish FA population have previously been determined.15 As reported, FA-A is the most common complementation group, representing approximately 80% of the patients. To investigate the origin, functional role, and clinical impact of FANCA mutations, we have performed mutation analysis of 67 Spanish FA-A patients previously assigned to complementation group A and ascertained for congenital malformations and clinical course by the Spanish FA Research Network. A complete FANCA mutational spectrum is presented. Some of the
mutations found were further characterized at the molecular and/or functional level. We prove that long distance recombination can result in large interstitial deletions involving FANCA and adjacent genes. We also show that all missense and other nontruncating FANCA mutations lead to a nonfunctional FANCA protein unable to relocate from the cytoplasm to the nuclei, thus explaining the lack of correlation between FANCA mutation and cellular or clinical phenotype in terms of age of onset of hematologic disease and number of malformations.

Methods

Patients and samples

The diagnosis of FA was confirmed on the basis of diepoxybutane (DEB)-induced chromosome fragility tests, as previously described. Clinical data from the FA patients were obtained from their clinicians, including age of onset of hematologic disease and number of congenital malformations as previously described. Classification of patients as having T-cell mosaicism was based on a percentage of aberrant cells less than 50% to 60%. Family informed consent was obtained for all patients included in this study in accordance with the Declaration of Helsinki. This investigation was approved by the Universitat Autònoma de Barcelona University Ethical Committee on Human Research.

Genomic DNA samples from 67 Spanish FA-A patients were obtained, using a standard phenol-chloroform extraction method, from blood, lymphoblastoid cell lines (LCLs), or primary fibroblasts. When necessary, cDNA was synthesized from total RNA using Superscript II RNase H reverse transcriptase (Invitrogen). Thirteen of the patients belong to the Gypsy ethnic group and were directly subtyped by mutation analysis. Of the remaining 54 patients (of white origin), 52 had previously been assigned to FA-A complementation group by retroviral complementation analysis, whereas in 2 of them, the complementation group was identified by PCR and analyzed using denaturing high-performance liquid chromatography and sequencing. Large deletions were amplified by polymerase chain reaction (PCR) and analyzed by denaturing high-performance liquid chromatography and sequencing. Six regions inside TCF25 (transcription factor 25, a transcription silencer) were amplified by quantitative real-time PCR for relative copy number determination. LightCycler 2.0 system (Roche Diagnostics) was used. Amplification reactions were prepared using LightCycler DNA Master SYBR Green I Kit (Roche catalog no. 0300323001), as specified by the manufacturer. Primer sequences and amplification conditions are specified in supplemental Table 2. To amplify the fragment containing the breakpoints, Taq Platinum polymerase (Invitrogen, catalog no. 10966018) was used, following the manufacturer’s instructions. The following primers were used: primer F, 5′-TGGCAG-TAGTTGGAGTACAG-3′; primer R1, 5′-AGACACTCAGTTGAGCCAC-3′; and primer R2, 5′-CTGGACACCTCCTGAGGAGAG-3′. Before sequencing, a 0.9-kb fragment obtained by PCR amplification using primers F and R2 was cloned using TOPO TA system (Invitrogen, catalog no. K4500). Minipreps were performed using NucleoSpin Plasmid QuickPure Kit (Macherey-Nagel, catalog no. 740615). The 0.9-kb fragment was sequenced using universal primers.

FANCD2 monoubiquitination studies, cell fractionation, and Western blot analyses

FANCD2 monoubiquitination was analyzed after treatment of cell lines with 2mM hydroxyurea (HU) for 24 hours. Protein whole cell extracts were prepared using RIPA buffer (Upstate Biotechnology, catalog no. 20-188) essentially as we recently described. Briefly, 5 million cells were washed with phosphate-buffered saline and resuspended in RIPA buffer containing protease and phosphatase inhibitors (CompleteMini-ethylenediaminetetraacetic acid free, Roche Diagnostics, catalog no. 1836170001; and PhosSTOP, Roche Diagnostics, catalog no. 0490683700). Benzonase (Novagen, catalog no. 70746) was added at a final concentration of 25 U/mL, and samples were incubated for 10 minutes at room temperature. Nuclear and cytoplasmic fractions were prepared following the procedure established by Mendez and Stillman with some modifications. Briefly, 5 × 10⁶ cells were resuspended in 200 μL of solution A (10mM N-2-hydroxyethylpiperezine-N′-2-ethanesulfonic acid, 10mM KCl, 1.5mM MgCl₂, 340mM sucrose, 10% glycerol, 1mM dithiothreitol, 0.1% Triton X-100) and incubated for 10 minutes on ice. Cytoplasmic fractions were separated by centrifugation (1300g, 4 minutes). Pellets containing nuclei were washed once in solution A and lysed in 100 μL of solution B (2mM ethylenediaminetetraacetic acid, 0.2mM ethyleneglycoltetraacetic acid, 1mM dithiothreitol, 150mM NaCl, protease and phosphatase inhibitors, and 25 U/mL benzonase). Protein concentration was measured using Bio-Rad Protein Assay (Bio-Rad, catalog no. 500-0001), following the manufacturer’s instructions. Protein extracts were then loaded on an sodium dodecyl sulfate–polyacrylamide gel electrophoresis gel and subjected to standard Western blot procedure followed by immunodetection, as previously described. Primary antibodies used and working dilutions were as follows: anti-FANCA (gift of the Fanconi Anemia Research Center, Research Fund, 1:500), anti-FANCD2 (Abcam, ab2187; 1:2500), anti-ACTIN (Abcam, ab8226; 1:5000), anti-ORC2 (Abcam, ab68348; 1:500), and anti-GAPDH (Abcam, ab9485; 1:2500).

Breakpoint determination of ex1-20del

Microarray comparative genomic hybridization was initially used to narrow down the deletion breakpoint. Two microarrays were used: (1) HumArray3.2, in collaboration with the Array Core Facility of the University of California San Francisco, consisting of 2464 bacterial artificial chromosomes spotted in triplicate with an average spacing between clones of 1.4 Mb; and (2) whole-genome tiling-path bacterial artificial chromosome array, in collaboration with British Columbia Cancer Research Center, consisting of 32 433 bacterial artificial chromosomes. For analysis of homozygosity/loccyzyosity of SNPs, several regions containing polymorphic SNPs were amplified by PCR. Supplemental Table 2 describes primer sequences and conditions used to amplify each SNP-containing region. After amplification, PCR products were purified using NucleoSpin Extract II Kit (Macherey-Nagel, catalog no.740609250) and sequenced. Six regions including 7CF25 (transcription factor 25, a transcription silencer) were amplified by quantitative real-time PCR for relative copy number determination. LightCycler 2.0 system (Roche Diagnostics) was used. Amplification reactions were prepared using LightCycler DNA Master SYBR Green I Kit (Roche catalog no. 0300323001), as specified by the manufacturer’s instructions. Primer sequences and amplification conditions are specified in supplemental Table 3. To amplify the fragment containing the breakpoints, Taq Platinum polymerase (Invitrogen, catalog no. 10966018) was used, following the manufacturer’s instructions. The following primers were used: primer F, 5′-TGGCAG-TAGTTGGAGTACAG-3′; primer R1, 5′-AGACACTCAGTTGAGCCAC-3′; and primer R2, 5′-CTGGACACCTCCTGAGGAGAG-3′. Before sequencing, a 0.9-kb fragment obtained by PCR amplification using primers F and R2 was cloned using TOPO TA system (Invitrogen, catalog no. K4500). Minipreps were performed using NucleoSpin Plasmid QuickPure Kit (Macherey-Nagel, catalog no. 740615). The 0.9-kb fragment was sequenced using universal primers.

Haplotype analysis

Haplotypes associated with the mutation c.3788_3790delTCT in FANCA were determined by analysis of 10 single nucleotide polymorphisms (SNPs) located inside the FANCA region and 4 highly polymorphic microsatellites consisting of a variable number of tandem repeats (VNTRs), flanking FANCA. Regions containing each SNP were amplified by PCR and analyzed using denaturing high-performance liquid chromatography and sequencing. Large deletions were detected using a Multiplex Ligation-Dependent Probe Amplification (MLPA) kit (SALSA MLPA Kit P031/P032-A2 FANCA) developed by MRC-Holland BV.
To induce FANCA foci, cells were treated with 2mM HU for 24 hours. After treatment, 2 × 10⁶ cells from patient-derived LCLs were resuspended in 200 μL of RPMI and extended on poly-lysine coated slides. Cells were incubated in hypotonic solution (phosphate-buffered saline 0.4×) for 1 minute and then fixed with 4% formaldehyde for 10 minutes. Immunodetection was performed as previously described. Antibodies used and working dilutions were as follows: anti-FANCA (Fanconi Anemia Research Fund; 1:500), anti-γ-H2AX (Upstate Biotechnology, DS-636; 1:3000), Alexa 488 antirabbit (Invitrogen, catalog no. A11034; 1:500), and Alexa 568 antimouse (Invitrogen, catalog no. A11031; 1:1000).

### Table 1. Different mutations found in FANCA Spanish population classified by mutation type

<table>
<thead>
<tr>
<th>DNA change</th>
<th>Exon</th>
<th>Protein change</th>
<th>Reference</th>
<th>No. of alleles</th>
<th>No. of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point mutations, 48 (36.9%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.65G&gt;A</td>
<td>1</td>
<td>p.W22X</td>
<td>23</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>c.285C&gt;T</td>
<td>4</td>
<td>p.Q99X</td>
<td>6</td>
<td>24</td>
<td>12</td>
</tr>
<tr>
<td>c.718G&gt;C</td>
<td>21</td>
<td>p.C625S</td>
<td>Novel</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>c.2224C&gt;A</td>
<td>28</td>
<td>p.R880Q</td>
<td>23</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>c.2641C&gt;T</td>
<td>28</td>
<td>p.Q851X</td>
<td>Novel</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Small insertions/ deletions, 54 (41.5%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.50delG</td>
<td>1</td>
<td>p.G17AfxX26</td>
<td>Novel</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>c.233_236delTTGA</td>
<td>3</td>
<td>p.I78TfsX16</td>
<td>Novel</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Splicing mutations, 8 (6.2%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.709 + 5 G-&gt;A</td>
<td>7</td>
<td>D237G + 238ins10 (AFMTRCGFLD)</td>
<td>29</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>c.710-5T&gt;C</td>
<td>8</td>
<td>Altered splicing</td>
<td>Novel</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Large deletions, 20 (15.4%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ex1-3del</td>
<td>—</td>
<td>—</td>
<td>23</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ex1-20del</td>
<td>—</td>
<td>—</td>
<td>Novel</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ex3del</td>
<td>—</td>
<td>—</td>
<td>14</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ex4-5del</td>
<td>—</td>
<td>—</td>
<td>19</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ex15-43del</td>
<td>—</td>
<td>—</td>
<td>Novel</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ex16-17del</td>
<td>—</td>
<td>—</td>
<td>27</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ex18-20del</td>
<td>—</td>
<td>—</td>
<td>23</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ex21-22del</td>
<td>—</td>
<td>—</td>
<td>Novel</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ex21-29del</td>
<td>—</td>
<td>—</td>
<td>Novel</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>ex28-29del</td>
<td>—</td>
<td>—</td>
<td>Novel</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>130</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Statistics

The effect of type of mutation on spontaneous and DEB-induced chromosome fragility test was evaluated using parametric tests (t test), as a normal distribution was observed in both groups. For the analysis of effect of type of mutation on the severity of clinical markers, nonparametric tests were used (U Mann-Whitney). All statistical analyses were performed using SPSS software package.

Results

Spanish FANCA mutational spectrum

A total of 130 mutated alleles have been identified among 67 Spanish FA-A patients, 13 of them belonging to the Gypsy ethnic group. One of the 2 FANCA mutations could not be identified in 4 patients unequivocally subtyped as FA-A. In all of these 4 cases, either the patient was mosaic and/or LCLs from where the DNA was extracted was reverted. Thus, probably the reason why the second mutation could not be detected was that it had been reverted to normal, generating a functional allele. Either way, the overall mutation detection success rate was 97%.

A comprehensive list of all different mutations found is provided in Table 1. A list of all FA-A patients and their corresponding mutations is also provided (supplemental Table 4).

Among 130 mutated alleles, 52 different pathogenic mutations were identified, 20 of them novel. The majority of mutations were point mutations and small insertions/deletions. Many of the latter were most probably caused by DNA polymerase slippage during replication, as they are often (6 of 13) associated with short direct repeats or homonucleotide tracts. Fifteen of the novel mutations were clearly pathogenic (2 nonsense mutations, 8 mutations causing frameshift and premature protein synthesis termination, and 5 large deletions). Six of them affected splicing recognition signals, and their predicted effect (indicated in the table) was confirmed on mRNA from LCLs. One of the mutations affecting splicing was surprisingly deep intronic (c.893/C5/H11022/A). This mutation was seen to activate a cryptic splicing recognition site, resulting in the incorporation of new exonic sequences (data not shown). Finally, the remaining 4 (p.L274P, p.C625S, p.Q742K, and p.L1305F) were missense mutations. Pathogenicity of these mutations was evaluated using 3 different bioinformatic algorithms: Grantham’s scale, POLYPHEN, and SIFT. Although caution should be taken with software predictive programs to evaluate the pathogenicity of missense variants, all 3 algorithms concluded that the variants p.L274P, p.C625S, and p.Q742K would probably compromise protein function. Predicted pathogenicity of mutation p.L1305F was controversial, as the amino acid change is conservative, but the position is conserved through evolution, indicating that this could be an important residue for protein function. The fact that this variant (1) was not found in 3 unrelated patients, (2) is not reported as an SNP, and (3) was not found in more than 100 geographically matched chromosomes (data not shown) probably confirms the pathogenic character of this amino acid change. However, a functional test would be required to be 100% certain of the pathogenic nature of this variant.

A schematic representation of the Spanish FANCA mutational spectrum is shown in Figure 1. As reported before, almost all patients belonging to the Gypsy ethnic group share a common truncating mutation in exon 4 (c.295C/H11022/T, p.Q99X). Only recently, one Gypsy patient was seen to be homozygous for a different mutation (ex4-6del). Mutation c.295C/H11022/T was also detected in homozygosity in 5 Portuguese Gypsy patients analyzed (data not shown), indicating that this mutation has extended throughout the Iberia peninsula. The carrier frequency for this mutation among Spanish Gypsies was previously estimated to be 1 of 67 according to Hardy-Weinberg equilibrium, the highest FA mutation frequency in the world ever reported. To validate this estimation, 135 DNA samples of randomly selected healthy Spanish Gypsy persons were obtained and screened for the c.295C/H11022/T mutation. Two carriers were detected among 135 Gypsies (data not shown), confirming the expected carrier frequency in this population (P < .05).

To our knowledge, this result is the first molecular evidence that FANCA absence has no impact on embryonic viability during development. Although this conclusion was obtained in Spanish Gypsies, it can...
probably be extrapolated to all FAA persons, as this mutation leads to a stop codon and therefore to an N-ter truncation.6

As expected, the mutational spectrum in white Spanish patients is much broader. However, some recurrent mutations were detected: c.3788_3790delTCT in exon 38 (20.7% of patients), c.1115_1118delTTGG in exon 13 (9.4% of patients), and c.3558insG in exon 36 (7.5%). Based on the results of this study, a workflow to optimize mutation screening on Spanish white patients was designed. This strategy consisted of 3 rounds of screens, starting with an MLPA to detect all intragenic large deletions and the sequencing of exons 13, 36, and 38. This first round would detect 58.5% of all Spanish FANCA mutations. Rounds 2 and 3 include sequencing of 15 and 8 exons, respectively. No mutations have been found up until now in the remaining 17 exons of FANCA among Spanish patients; therefore, a fourth round of sequencing would not be required in the great majority of cases. To determine whether this strategy is exportable to other populations, the frequency of mutations in each exon was calculated from the FA Mutation Database. As seen in the lower panel of Figure 2, the first round would also include the screening of the most frequent mutations in other populations worldwide; and together with the next 2 rounds, this strategy would detect close to 90% of all FANCA mutations.

On the origin of the most frequent FANCA mutation in the world: c.3788_3790delTCT

The exon 38 mutation c.3788_3790delTCT is the most frequent FA mutation, not only in Spain (20.7% of all mutated alleles), but also in the rest of the world. This mutation is shared by 80% of patients from La Palma Island (Canary Islands, Southern Spain), suggesting a founder effect that explains the extraordinary high frequency of FA patients in this island (1 in 16 000 births, the highest frequency of FA reported up to now). Similarly, this mutation also accounts for an elevated percentage (51%) of the mutations found in Brazilian patients.12 To determine whether this high frequency is the result of the spreading of a very ancestral mutation or, alternatively, the result of a mutational hotspot associated with the phenomenon of DNA polymerase slippage (this small deletion is associated to a trinucleotide direct repeat), we performed a comprehensive haplotype analysis in a total of 28 patients sharing this mutation but belonging to geographic distant populations (Spain, Brazil, Mexico, Argentina, Peru, United States, United Kingdom, Portugal, Germany, Pakistan, and Nigeria). This analysis was made possible by the elevated number of SNPs in the FANCA gene.12 As shown in Figure 3A, all patients analyzed, except the Nigerian patient, are carriers of the same haplotype, referred to here as haplotype 1, based on the analysis of 10 FANCA intragenic SNPs.12 To rule out the possibility that the same mutation appeared in different persons with the same haplotype, 4 highly polymorphic VNTRs were also analyzed (Figure 3A bottom panel). The 2 microsatellites located at 5’ of FANCA (Figure 3B) were identical in all haplotype 1-carrying patients analyzed, confirming that they all share a common ancestor and therefore ruling out a mutational hotspot.

Breakpoint characterization of ex1-20del

The majority of large intragenic deletions in FANCA are caused by recombination between 2 Alu repeats located in cis as reported for many other genes.25,34,35 However, when one of the breakpoints is telomeric versus exon 1, it is possible that the deletion does not result from intrachromosomal recombination, but from a terminal deletion (with a single breakpoint) as FANCA is very close to 16q telomere. In addition, it is not known whether long distance recombination can also cause FANCA deletions. To further understand the mechanisms of origin of large FANCA deletions, we genetically characterized a large deletion involving the first exon and, therefore, with a putative breakpoint between FANCA and 16q telomere. This analysis was done in the 2 patients with an ex1-20del previously detected by MLPA (Table 1). We sequentially approached and narrowed down the telomeric breakpoint by
several methods. First, microarray comparative genomic hybridization (the ex1-20del patient was included as control in an independent comparative genomic hybridization project) revealed that this FANCA mutation was not caused by a terminal 16q deletion but by an interstitial deletion with 2 breakpoints, the first necessarily between exons 20 and 21 of FANCA and the second within a region of 276 kb at 5’/H11032 of FANCA (supplemental Figure 1). We then applied hemizygosity analysis of SNPs in one of the patients and both parents (supplemental Figure 2), leading to a region of 35 kb containing the second breakpoint inside a gene located at 5’ of FANCA near the 16q telomere, the TCF25 gene. Because no more informative SNPs were found, several regions were then amplified and allele copy number was determined by quantitative real-time PCR (supplemental Figure 3). This led us to map the second breakpoint inside intron 5 of TCF25, a region rich in Alu repeats. At this point, a fragment including the breakpoint was cloned and sequenced, confirming the adjacent presence of intron 20 of FANCA and intron 5 of TCF25 with an ALU repeat in between (Figure 4C). Therefore, the deletion analyzed here spans 112.5 kb and involves 3 genes, FANCA, SPIRE2, and TCF (Figure 4D), revealing that long distance Alu-Alu recombination can occur in cis, leading to the generation of very large pathogenic deletions.

**Functional characterization of missense FANCA mutations**

Functional consequences of missense and other nontruncating mutations (in-frame deletions and insertions) distributed along FANCA were analyzed for protein expression, residual activity, and protein subcellular localization. This study was done with missense FANCA mutations found in 13 Spanish patients where cell lines and phenotypic data were available for functional studies. The list of patient-derived cell lines analyzed and their corresponding mutations is provided in Table 2. As shown in Figure 5A, all cell lines carrying at least one nontruncating mutation (missense or splicing mutation leading to a 10 amino acid insertion) expressed altered FANCA protein. Quantification of protein expression (Figure 5B) revealed a good correlation between gene dosage and...
residual protein level, as cell lines carrying 2 missense mutations expressed a level of FANCA almost equivalent to wild-type, whereas those cell lines with one nontruncating mutation expressed only 30% to 40%. The residual activity of these altered FANCA

Table 2. Patient-derived LCLs used in the functional studies and their corresponding nontruncating FANCA mutations on one or both alleles

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>DNA</th>
<th>Protein</th>
<th>DNA</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>FA5</td>
<td>c.4198C&gt;T (Ex 42)</td>
<td>p.R1800C</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>FA56</td>
<td>c.3913C&gt;T (Ex 39)</td>
<td>p.L1305F</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>FA58</td>
<td>c.821C&gt;T (Ex 9)</td>
<td>p.L274P</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>FA75</td>
<td>c.3788_3790delTCT (Ex 38)</td>
<td>p.F1263del</td>
<td>c.2639G&gt;A (Ex 28)</td>
<td>p.R880Q</td>
</tr>
<tr>
<td>FA82</td>
<td>c.3788_3790delTCT (Ex 38)</td>
<td>p.F1263del</td>
<td>c.709 + 5G&gt;T (Ex 7)</td>
<td>p.D237G + 238ins10 (AFMT25CGFLD)</td>
</tr>
<tr>
<td>FA88</td>
<td>c.709 + 5G&gt;A (Ex 7)</td>
<td>p.D237G + 238ins10 (AFMT25CGFLD)</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>FA91</td>
<td>c.2224C&gt;A (Ex 25)</td>
<td>p.Q742K</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>FA155</td>
<td>c.3788_3790delTCT (Ex 38)</td>
<td>p.F1263del</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>FA170</td>
<td>c.3788_3790delTCT (Ex 38)</td>
<td>p.F1263del</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>FA328</td>
<td>c.2303T&gt;C (Ex 26)</td>
<td>p.L768P</td>
<td>c.2303T&gt;C (Ex 26)</td>
<td>p.L768P</td>
</tr>
</tbody>
</table>

*T indicates truncating mutation.
proteins was then tested by analyzing FANCD2 monoubiquitylation. As shown in Figure 5C, residual monoubiquitylation of FANCD2 was not detected in any of the FA-A cell lines expressing an altered FANCA protein. Finally, subcellular localization of FANCA was also evaluated by Western blot (Figure 5D) and immunohistochemistry (Figure 5E). In all mutated cell lines, the altered FANCA protein was mainly located in the cytoplasm, whereas it is both cytoplasmic and nuclear in wild-type cells. Although we cannot rule out the possibility that a residual proportion of FANCA could still be able to enter into the nucleus of the FA cells, the quantity would be dramatically reduced (5- to 20-fold) compared with the wild-type, according to densitometric analysis (Figure 5D). By immunohistochemistry, FANCA forms nuclear foci colocalizing with γ-H2AX on HU treatment in wild-type cells but, as expected, no FANCA foci were observed in the mutated FA cell lines expressing an altered FANCA protein. This set of data allowed us to conclude that, in all FA cells bearing at least one nontruncating mutation, a rather stable FANCA protein is indeed produced, but it is nonfunctional because of its reduced ability to normally translocate to the nucleus, bind stalled replication forks, and activate the FA/BRCA pathway by FANCD2 monoubiquitylation.

Genotype-phenotype correlations

Previous studies suggest an association of mutation type with clinical outcome in FA in the sense that patients homozygous for null mutations had an earlier onset of anemia compared with patients expressing an altered FANCA protein.36 The authors hypothesized that a milder clinical phenotype would not be unexpected if the altered protein had partial activity. To further investigate this issue, we studied the association between type of mutation, the sensitivity of the peripheral blood T cells of the patient to interstrand crosslink-inducing agents, and severity of clinical phenotype in the Spanish FA-A cohort. Patients were divided into 2 groups, depending on the FANCA protein expression: patient homozygotes for 2 null mutations versus patients with 1 or 2 missense mutations leading to an altered protein. FANCA expression/absence and functionality were previously confirmed in all patients (Figure 5; and data not shown). Because all missense mutations led to altered nonfunctional FANCA protein (Figure 5), all patients expressing a nonfunctional FANCA were included in the same category. Spontaneous and DEB-induced chromosome fragility was evaluated on peripheral blood T cells in all patients, and those showing a T-cell mosaicism were excluded from this genotype-phenotype analysis. As shown in Figure 6A and B, no differences in spontaneous and DEB-induced chromosome fragility were observed between the 2 mutation subtypes, which is consistent with the absence of nuclear FANCA observed in all mutant patient-derived cell lines (Figure 5). We then analyzed the effect of mutation subtype on clinical phenotype that was measured using 2 parameters: number of malformations and age of onset of hematologic disease, following previously established criteria.17,36
As shown in Figure 6C, the age of onset of hematologic disease is virtually identical in the 2 mutation subtypes (6.65 ± 0.47 years vs 6.25 ± 1.11 years). The number of malformations is also similar in the 2 mutation subtypes as shown in Figure 6D (3.25 ± 0.36 vs 3.36 ± 0.53). Thus, all data lead to the conclusion that the severity and evolution of the clinical phenotype are not improved by the residual presence of a non functional FANCA protein in patients bearing nontruncating mutations.

**Discussion**

Mutation analysis of Spanish FA-A patients revealed a wide FANCA mutational spectrum. Fifty-two different mutations were detected among 67 FA patients, 20 of them novel, and reported here for the first time. Small deletion and insertions were common and most probably caused by DNA polymerase slippage during replication as they are often (6 of 13) associated with short direct repeats or homonucleotide tracts as previously reported. Three mutations were recurrently found in exons 13 (c.1115_1118delTTGG), 36 (c.3558insG), and 38 (c.3788_3790delTCT). These mutations represent the 3 most frequent FANCA mutations. As in many other populations, a considerable proportion of large intragenic deletions (15.4%) were found among Spanish mutated alleles. Based on the results of this study, we propose a strategy to screen FANCA mutations that includes serial rounds of PCR/sequencing reactions and an MLPA reaction. Detection of deletions and sequencing of only exons 13, 36, and 38 allow nearly 60% of mutations to be identified in FA-A Spanish patients. At the same time, this strategy is also exportable to other populations as it allows approximately 50% of mutations to be detected with the first round and close to 90% with the other 2 rounds. Thus, assuming that 3 of 4 patients worldwide are FANCA and that the first 3 rounds would detect 4 of 5 mutations, this strategy would allow the direct subtyping and mutation detection of 60% of all FA patients (75% of all Spanish FA patients) in less than a week. Consequently, only a minority of patients would require further subtyping and sequencing efforts. We are currently investigating the applicability of the proposed workflow to Latin American patients as we expect that our strategy will enable pathogenic mutations to be detected in an even higher percentage of patients because of founder effects, as seen in La Palma Island (later in “Discussion”). In conclusion, this scheme will make the turnaround time for diagnosis faster because approximately 2 of 3 mutations are expected to be found in less than a week. This is especially important when urgent prenatal tests are required.

Embryonic lethality has been associated with mutations in some FA genes, in particular, with FANCD1/BRCA2. In Spain, with an estimated frequency of BRCA2 carriers of 1 of 500, the number of FANCD1 patients would be more than 40 fold below that expected by Hardy-Weinberg equilibrium. In contrast, we provide the first molecular evidence that biallelic FANCA mutations do not challenge embryonic development in humans; otherwise, the observed number of the Gypsy mutation carriers would have been lower than
expected. This lack of negative selection could partially explain the predominance of the FA-A complementation group among FA patients. This clearly distinct effect between FANCA and FANC1 mutations on embryo viability and development further supports the idea that FANC1 patients (and probably other downstream patients) partially overlap FA phenotype but are a distinct clinical entity from a genetic, functional, and syndromic perspective.

Mutation c.3788_3790delTCTT in FANCA is the most common FA mutation in Spain and in the rest of the world. It is particularly common in Brazil29 and other populations of white or Latin origin,24 but it has also been detected in other distant populations, such as Nigeria and Pakistan. We also detected a founder effect of this mutation in La Palma Island (Canary Islands). A complete haplotype analysis of patients from inland Spain, La Palma island, Brazil, Mexico, Argentina, Peru, United States, United Kingdom, Portugal, Germany, Pakistan, and Nigeria all bearing this mutation revealed that all patients share a common ancestor with the only exception of the Nigerian patient, whose mutation was previously reported.24 Therefore, even though this small deletion is associated to a short trinucleotide repeat, the high frequency and worldwide distribution of c.3788_3790delTCTT are not the result of a mutational hotspot but caused by a very ancestral mutation that has been disseminated from Indo-European countries to North and South America in the last few centuries. In this spreading process, La Palma Island cluster was probably originated by a founder effect less than 400 years ago, when all ships navigating from Spain to America had to mandatorily transit via this island.

Intragenic large deletions are very common in FANCA. It has been repeatedly demonstrated that the majority of these deletions are the product of a recombination between Alu-repeats in cix.12,25,40 The breakpoints of one of these deletions (ex1-20del) were characterized for this purpose. We found that recombination between very distant Alu repeats (at least 112 kb apart) can also result in large deletions involving adjacent genes (SPIRE2 and TCF25). This result indirectly proves that SPIRE2 and TCF25 are not mutated in any autosomal dominant inherited syndrome, as 3 identified carriers of this deletion are completely normal and healthy. In addition, the 2 FA patients with this deletion are classic FA patients with no signs of any other clinical abnormality unrelated to FA mutations.

According to the Fanconi Anemia Mutation Database, approximately one-fourth of the mutations found in FANCA are missense mutations, which can be found distributed all along the gene. Studies on residual function of the resulting mutant proteins often lead to the discovery of functionally important protein domains. However, apart from NLS and phosphorylation residues,41,42 no additional domains have been so far described for FANCA. The results of this study show that all missense mutations analyzed, independently of their position within the gene, lead to common functional consequences: a mutant protein is expressed at rather normal levels relative to gene dosage, but the FA/BRCA pathway remains inactive because of the impaired nuclear transport of the mutant protein. This is in agreement with previous studies for some particular missense mutations.43,44 Therefore, DNA damage signaling and other possible nuclear FANCA functions are completely abolished in all FANCA mutants irrespective of the mutation type.

A previous study suggests that FA patients homozygous for null mutations had an earlier onset of anemia compared with patients expressing an altered FANCA protein.46 The authors hypothesized that a milder clinical phenotype would not be unexpected if the altered protein had partial activity. In keeping with the aforementioned conclusion, we could not detect any genotype-cell phenotype association in terms of the ability of the FA cell to process stall replication forks and hence avoid chromosomal aberrations (either spontaneous or DEB-induced). Similarly, no clinical differences in terms of age of onset of hematologic disease and number of malformations were found between patients who do not express FANCA and patients who express a mutant form of the protein. The aforementioned study46 includes a high proportion of South African Afrikaner FA patients with a homozygote deletion of ex12:31; therefore, they could not exclude population-specific effects (Christopher Mathew, email, April 22, 2010). Our results, together with the evidence of different clinical phenotypes even in siblings carrying the exact same mutations,46 and the apparent discrepancy between American and European cohorts, suggest that mutation type has little prognostic value in FA-A patients and that many other players, including stochastic factors, genetic background, ethnicity, and environmental exposures, play a critical role in the clinical evolution of the disease. Finally, it is tempting to speculate 2 potential therapeutic consequences of our results. First, patients with missense mutations would probably not generate an immunologic response after gene therapy as they expressed a rather stable protein with similar antigenicity compared with wild-type FANCA, according to our Western blot data. Second, the chances of finding drugs able to reactive the FA pathway are probably higher if the screening is done in FA cells with missense mutations leading to an abnormal protein.

Acknowledgments

We thank all Spanish Fanconi anemia families and their clinicians for their continuous support and for providing samples, Aurora de la Cal and Ana Molina for sampling coordination, Jaume Bertranpetit, Montse Baiget, Francesc Palau and M. Luisa Martínez Frias for providing DNA samples from Gypsy persons, and Alicia Barroso and Glòria Umbert for technical support.

The laboratory of J. Surrallés is funded by the Generalitat de Catalunya (SGR0489-2009), the La Caixa Fundation Oncology Program (BM05-67-0), Fundación Genoma España, the Spanish Ministry of Science and Innovation (projects PIS P106-1099, CB06/07/0023, SAF2006-3440 and SAF2009-11936), the Commission of the European Union (project RISC-RAD F6R-CT-2003-508842), the European Regional Development Funds. Centre for Biomedical Network Research on Rare Diseases is an initiative of the Instituto de Salud Carlos III. J.A.B. is supported by SAF2009-07164, PLE 2009-0100, and VII FWP PERSIST 222878. A.D.A. is funded by the National Institutes of Health Clinical and Translational Science Award SUL1RR024143-04 to The Rockefeller University Hospital.

Authorship

Contribution: M.C. wrote the paper, designed experiments, performed research, and analyzed data; R.P., E. Callén, J.P.T., J.A.C., H.G., and F.P.L. performed research and analyzed data; A.D.A. analyzed data, designed experiments, and wrote the paper; D.S., J. Benitez, B.P., and T.F. performed research, contributed vital new reagents, and wrote the paper; A. Muñoz, J. Sevilla, L.M., E. Cela, C.B., C.D.d.H., T.O., J.S.d.T., I.B., M. Torrent, J.E., A.D., A.R.-V., PG., J. Barbot, M. Tapia, A. Moliniés, and A.F. contributed vital materials, provided clinical data, and analyzed data; J.A.B. analyzed data and wrote the paper; and J. Surrallés coordinated the
study, wrote the paper, performed research, designed the study, analyzed data, and contributed vital new reagents.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

References


Correspondence: Jordi Surrallés, Group of Genome Instability and DNA Repair, Department of Genetics and Microbiology, Universitat Autonoma de Barcelona, Campus de Bellaterra s/n, 08193 Bellaterra, Barcelona, Spain; e-mail: jordi.surralles@uab.es.
Origin, functional role, and clinical impact of Fanconi anemia *FANCA* mutations

Maria Castella, Roser Pujol, Elsa Callén, Juan P. Trujillo, José A. Casado, Hans Gille, Francis P. Lach, Arleen D. Auerbach, Detlev Schindler, Javier Benítez, Beatriz Porto, Teresa Ferro, Arturo Muñoz, Julián Sevilla, Luis Madero, Elena Cela, Cristina Beléndez, Cristina Díaz de Heredia, Teresa Olivé, José Sánchez de Toledo, Isabel Badell,Montserrat Torrent, Jesús Estella, Angeles Dasí, Antonia Rodríguez-Villa, Pedro Gómez, José Barbot, María Tapia, Antonio Molinés, Ángela Figuera, Juan A. Bueren and Jordi Surrallés

Updated information and services can be found at:
http://www.bloodjournal.org/content/117/14/3759.full.html

Articles on similar topics can be found in the following Blood collections
- Clinical Trials and Observations (4622 articles)
- Hematopoiesis and Stem Cells (3452 articles)
- Red Cells, Iron, and Erythropoiesis (818 articles)

Information about reproducing this article in parts or in its entirety may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at:
http://www.bloodjournal.org/site/subscriptions/index.xhtml