

**Chiraz Bouchlaka · Sonia Abdelhak
Ahlem Amouri · Hela Ben Abid · Sondes Hadjji
Mounir Frikha · Tarek Ben Othman
Fethi Amri · Hammadi Ayadi · Mongia Hachicha
Ahmed Rebaï · Ali Saad · Koussay Dellagi
The Tunisian Fanconi Anemia Study Group**

Fanconi anemia in Tunisia: high prevalence of group A and identification of new *FANCA* mutations

Received: 3 March 2003 / Accepted: 19 April 2003 / Published online: 24 June 2003
© The Japan Society of Human Genetics and Springer-Verlag 2003

Abstract Fanconi anemia (FA) is a rare autosomal recessive disease characterized by progressive pancytopenia, congenital malformations, and predisposition to acute myeloid leukemia. Fanconi anemia is genetically heterogeneous, with at least eight distinct complementation groups of FA (A, B, C, D1, D2, E, F, and G) having been defined by somatic cell fusion studies. Six genes (*FANCA*, *FANCC*, *FANCD2*, *FANCE*, *FANCG*, and *FANCF*) have been cloned. Mutations of the seventh Fanconi anemia gene, *BRCA2*, have been shown to lead to FAD1 and probably FAB groups. In order to characterize the molecular defects underlying FA in Tunisia, 39 families were genotyped with microsatellite markers linked to known FA gene. Haplotype analysis and homozygosity mapping assigned 43 patients

belonging to 34 families to the FAA group, whereas one family was probably not linked to the *FANCA* gene or to any known FA genes. For patients belonging to the FAA group, screening for mutations revealed four novel mutations: two small homozygous deletions 1693delT and 1751–1754del, which occurred in exon 17 and exon 19, respectively, and two transitions, viz., 513G → A in exon 5 and A → G at position 166 (IVS24+166A → G) of intron 24. Two new polymorphisms were also identified in intron 24 (IVS24–5G/A and IVS24–6C/G).

Keywords Fanconi anemia · Homozygosity mapping · Linkage analysis · Consanguinity · Genetic heterogeneity · Mutation

C. Bouchlaka, S. Abdelhak, A. Amouri, N. Labbane, S. Chakroun, M. Benfadhel, K. Dellagi, L. Aissaoui, R. Belakhal, Z. Belhaj Ali, H. Ben Abid, B. Meddeb, A. Hafsa, M. Elloumi, S. Hadjji, M. Frikha, F. Fakhfakh, H. Ayadi, M. Hachicha, A. Abdelkafi, L. Tordjman, F. Mellouli, T. Ben Othman, S. Ladeb, A. Ben Abdellahim, H. Sennana, H. Elghezal, A. Saad, H. Elomri, A. Laatiri, A. Khelif, S. Ennabli, F. Amri, M. Trudi, and A. Rebaï belong to the Tunisian Fanconi Anemia Study Group

C. Bouchlaka · S. Abdelhak (✉) · A. Amouri
K. Dellagi
Laboratoire d'Immunologie,
Vaccinologie et Génétique Moléculaire,
Institut Pasteur de Tunis, BP 74, 13 Place Pasteur,
1002 Tunis Belvédère, Tunisia
E-mail: sonia.abdelhak@pasteur.rns.tn
Tel.: +216-1-789608
Fax: +216-1-791833

H. Ben Abid
Service d'Hématologie, C.H.U. Aziza Othmana, Tunis, Tunisia

S. Hadjji
Service d'Hématologie, C.H.U. Hédi Chaker, Sfax, Tunisia

M. Frikha
Service de Cancérologie, C.H.U. Habib Bourguiba, Sfax, Tunisia

T. Ben Othman
Centre National de Greffe de Moelle Osseuse, Tunis, Tunisia

F. Amri
Service de Pédiatrie, Hôpital de Kairouan, Kairouan, Tunisia

H. Ayadi
Laboratoire de Génétique Humaine, Faculté de Médecine, Sfax, Tunisia

M. Hachicha
Service de Pédiatrie, C.H.U. Habib Bourguiba, Sfax, Tunisia

A. Rebaï
Centre de Biotechnologie de Sfax, Sfax, Tunisia

A. Saad
Laboratoire de Cytogénétique et de Biologie de la Reproduction, C.H.U. Farhat Hached, Sousse, Tunisia

The Tunisian Fanconi Anemia Study Group
Institut Pasteur de Tunis, Tunis, Tunisia

Introduction

Fanconi anemia (FA; MIM 227650) is a rare autosomal recessive disorder characterized clinically by a progressive bone marrow failure associated with several congenital abnormalities, including skeletal and renal malformations, microcephaly, and abnormal skin pigmentation (Giampietro et al. 1993; Altay et al. 1997). Patients have an increased risk of cancer, most notably of acute myeloid leukemia (Alter 2003). Cells from FA patients grown in vitro exhibit increased spontaneous chromosomal aberrations and are hypersensitive to DNA cross-linking agents such as diepoxybutane (DEB) and mitomycin C (MMC), a characteristic of strong diagnostic value (Berger et al. 1993). FA shows extensive locus heterogeneity, with at least eight complementation groups (FAA–FAG) determined by somatic cell hybridization and corresponding to separate defect genes (Joenje et al. 1997; Timmers et al. 2001). Six genes, viz., *FANCA*, *FANCC*, *FANCD2*, *FANCE*, *FANCF*, and *FANCG*, have been identified (Fanconi Anaemia/Breast Cancer Consortium 1996; Strathdee et al. 1992a, 1992b; Whitney et al. 1995; Hejna et al. 2000; Timmers et al. 2001; Waisfisz et al. 1999; de Winter et al. 1998, 2000a, 2000b; Saar et al. 1998). The products of these six FA genes have no homology with other known proteins or to each other. However, several studies have demonstrated that *FANCA*, *FANCC*, *FANCE*, *FANCF*, and *FANCG* proteins are assembled to form an FA nuclear protein complex (Kupfer et al. 1997; Garcia-Higuera et al. 1999; de Winter et al. 2000c; Siddique et al. 2001; Medhurst et al. 2001; Taniguchi and D'Andrea 2002). This complex is required for the mono-ubiquitination of *FANCD2* on lysine 561 and for the translocation of the *FANCD2* protein to nuclear foci to interact with *BRCA1* in response to DNA damage (Garcia-Higuera et al. 2001).

A recent study has shown that cell lines derived from FAD1 and possibly from FAB patients have biallelic mutations in the *BRCA2* gene and express truncated *BRCA2* protein. This study confirms that *BRCA2* is a FA gene and that mutations in this gene lead to the FA-D1 group (Howlett et al. 2002).

The relative prevalence of each FA subtype widely varies with the geographic or ethnic origin of the studied populations (Whitney et al. 1993; Verlander et al. 1994; Joenje et al. 1995; Joenje 1996; Savoia et al. 1996; Tachibana et al. 1999). *FANCA*, which maps to chromosome 16q24.3 (Pronk et al. 1995), is the most commonly mutated FA gene, accounting for 60%–65% of all cases, but its prevalence widely varies according to ethnical background. The FAA subtype is predominant in North America, South Africa (Afrikaner population), Italy, Japan, and Turkey (Jakobs et al. 1997; Tipping et al. 2001; Savoia et al. 1996; Tachibana et al. 1999; Balta et al. 2000). *FANCA* has an open reading frame of 4.3 kb distributed among 43 exons spanning over 80 kb genomic DNA (Lo Ten Foe et al. 1996; Ianzano et al.

1997). A large spectrum of mutations has been identified in the *FANCA* gene, including microdeletions, large deletions, microinsertions, and point mutations (Levrان et al. 1998; Centra et al. 1998; Morgan et al. 1999; Wijker et al. 1999; Tachibana et al. 1999). One “Tunisian” mutation in exon 10 (890–893del) and two “Moroccan” mutations in exons 24 and 43 (2172–2173 insG, 4275delT) of the *FANCA* gene have been identified in Israeli non-Ashkenazi Jewish patients (Tamary et al. 2000). Except for the afore-mentioned report, FA has not yet been investigated in North African population and particularly in Maghrebian population, which is characterized by its heterogeneous ethnic background and by a high rate of consanguinity.

In order to identify to which complementation group Tunisian patients belong, 39 unrelated families were investigated by using haplotype analysis and homozygosity mapping. Thirty-four families were assigned to the FAA group, whereas one family was probably not linked to *FANCA* gene or to any known FA gene. For the patients who are assigned to the FAA group, we screened the *FANCA* DNA or cDNA by direct sequencing. Four novel mutations in four patients and two new polymorphisms were identified.

Patients and methods

Patients

Thirty-nine unrelated families with a total of 49 FA patients were investigated (Table 1). All patients were diagnosed on the basis of clinical symptoms in combination with hypersensitivity to MMC, as determined in a standard cytogenetic chromosomal breakage test. Nine families were multiplex with at least two affected children; 20 families were simplex with one affected and at least one healthy child, and 10 singleton families had only one affected child. Seven patients, belonging to seven unrelated families, received an allogenic bone marrow transplant. Clinical data were obtained by reviewing the medical records from the referring physicians; some of these data have been reported by Friksa et al. (1998). Malformations were present in 96% of cases and consisted mainly of skeletal malformations. Abnormal skin pigmentation was present in 92% of cases. The ages of patients ranged from 1 year to 30 years. The mean age at diagnosis was 9 years. Profound aplastic anemia developed in 43% of cases. At the time of analysis, 64% of the patients were still living. The mean age of death for patients who had died was 16 years. More than 90% (92%) of FA families were consanguineous, and the degree of kinship was generally first cousins (60%) or second cousins (25%).

Genotyping

After informed consent, blood samples were collected, and genomic DNA was extracted by standard procedures (Sambrook et al. 1989). Microsatellite markers were selected from the Genethon mapping panel (Dib et al. 1996). Primer pairs corresponding to seven polymorphic markers were amplified for linkage analysis at 16q24.3. The markers span an interval of about 10 cM and have the following order from centromere to telomere: D16S520–D16S498–D16S3123–D16S3026–D16S3121–D16S303 (Pronk et al. 1995). D16S3407 is an intragenic marker located 16 kb upstream the 5' untranslated region of the gene (Morgan et al. 1999). D16S303, D16S3121, and D16S3026 have been mapped within the

Table 1 Characteristics of 39 FA Tunisian families. Families were denoted as *TF*, *CF*, and *SF* according to the regions in which they live and corresponding respectively to north, center, and south of Tunisia (+ linked to *FANCA* gene by haplotype analysis and homozygosity mapping, – not linked to *FANCA*, NI non-informative families). Families in *bold type* were selected for mutation screening (for each multiplex family, only one patient was investigated). Mutations identified are given in parenthesis. Carrier haplotypes segregating with 49 patients with the closest markers (D16S3407-D16S303-D16S3121-D16S3026) linked to *FANCA* gene are reported right

Families	Sibs		Geographic origin	Linkage to <i>FANCA</i> gene	Haplotype information
	FA	Healthy			
SFA1^a	3	6	Southern Tunisia	+	2-5-2-3
SFA2^a	2	2	Southern Tunisia	+	2-5-5-3
SFA3^a	2	3	Southern Tunisia	+	2-3-4-5
SFA4^a	2	5	Central Tunisia	+	2-5-2-2
SFA5	1	0	Southern Tunisia	+	2-5-2-2
SFA6^a	2	3	Southern Tunisia	–	2-5-2-2
					1-5-4-4
SFA7	1	0	Southern Tunisia	+	1-5-4-4
SFA8	1	0	Southern Tunisia	+	2-5-2-3
SFA9	1	3	Central Tunisia	+	3-5-2-2
SFA10^b	1	1	Southern Tunisia	NI	3-5-5-5
					2-5-4-5
SFA14	1	0	Southern Tunisia	+	2-5-2-3
SFA15	1	0	Southern Tunisia	+	2-5-2-3
SFA16	1	0	Southern Tunisia	+	2-5-2-3
SFA17^c	1	1	Southern Tunisia	+	2-5-2-2
SFA19	1	0	Southern Tunisia	+	2-5-2-2
SFA21^a	2	1	Central Tunisia	+	2-6-2-2
SFA24^a	2	0	Southern Tunisia	+	2-5-2-3
TFA1	1	1	Northern Tunisia	–	1-6-5-4
					3-2-4-3
TFA2	1	0	Southern Tunisia	+	5-1-5-4
TFA4^{b, c}	1	4	Southern Tunisia	NI	3-5-2-4
					3-5-4-4
TFA5^c	1	4	Southern Tunisia	+	1-5-4-4
TFA6^c	1	4	Southern Tunisia	+	2-5-4-4
TFA8^c	1	1	Southern Tunisia	+	1-4-5-4
TFA9	1	2	Northern Tunisia	+ (1606delT)	2-1-5-3
TFA10^{a, c}	2	3	Central Tunisia	+	2-4-4-4
TFA11	1	1	Northern Tunisia	+ (1751–1754del)	3-5-4-4
TFA13	1	4	Northern Tunisia	+	2-5-2-2
TFA16^b	1	1	Northern Tunisia	+	2-5-4-7
TFA17	1	4	Northern Tunisia	+	2-2-5-8
TFA22	1	0	Northern Tunisia	NI	1-5-5-4
TFA23^c	1	2	Central Tunisia	+	3-5-5-4
TFA29	1	1	Central Tunisia	+ (513G → A)	1-5-5-3
TFA34	1	0	Southern Tunisia	+	2-5-2-4
TFA35	1	6	Southern Tunisia	+	1-7-5-8
CFA1^a	2	2	Central Tunisia	+	3-6-4-3
CFA2	1	7	Central Tunisia	+ (IVS24 + 166A → G)	3-5-5-4
CFA3	1	2	Central Tunisia	+	3-5-5-4
CFA4	1	6	Central Tunisia	+	3-5-5-4
CFA6	1	2	Central Tunisia	+	2-5-4-4

^aMultiplex families

^bNon-consanguineous families

^cFamilies with patients who have received an allogenic bone marrow transplant (for the TFA10 family, only one affected child received an allogenic bone marrow transplant)

critical region of the *FANCA* gene (Fanconi Anaemia/Breast Cancer Consortium 1996). For families not linked to the *FANCA* gene, microsatellite markers corresponding to the *FANCC*, *FANCD2*, *FANCE*, *FANCF*, and *FANCG* genes were tested. Polymerase chain reaction (PCR) was performed in a total volume of 50 µl containing 200 ng genomic DNA, 10 µM each primer, 250 µM each dNTP, 1.5 mM MgCl₂, 1 U *Taq* DNA polymerase. Amplification buffer contained 20 mM TRIS-HCl (pH 8.8) and 50 mM KCl.

The reactions were performed by using a “hot-start” procedure; initial denaturation was for 5 min at 96°C followed by amplification for 40 cycles with denaturation at 94°C for 1 min, annealing for 1 min at the appropriate temperature, and extension at 72°C for 1 min, followed by a final extension at 72°C for 7 min.

The amplified products were run on a 6% polyacrylamide gel and then transferred to a nylon membrane by contact blotting procedure. Amplified fragments were revealed by hybridization at 42°C for 3 h in hybridization buffer with a radioactively labeled poly(AC) probe [d CTP α -³²P] according to Hazan et al. (1992).

Mutation screening

Genomic DNA amplification and direct sequencing

Exons and their flanking regions were amplified from genomic DNA. Primer pairs were generated from intron sequences (Ianzano et al. 1997; Levran et al. 1997; Savino et al. 1997) in order to amplify the 43 investigated exons in 27 PCRs with products ranging from 220 bp to 1340 bp. PCR conditions were as described above. PCR products were purified on QIAquick spin columns (Qiagen) and directly sequenced with the Big Dye terminator kit (Applied Biosystems) with PCR primers on an ABI Prism 377 DNA sequencer (Applied Biosystems).

cDNA isolation, amplification, and direct sequencing

In order to identify the alterations in RNA level, total RNA was isolated from total blood by means of the TRIzol reagent (Gibco BRL Life Technologies). The cDNA first-strand was synthesized by using the ProSTAR first-strand reverse-transcriptase PCR (RT-

PCR) kit (Stratagene). The complete cDNA corresponding to the *FANCA* gene was amplified in 10 overlapping fragments (primer sequences are available upon request). RT-PCR products were purified and directly sequenced.

Linkage analysis

Linkage analysis was performed by using the computer program Genehunter v2.1 (Kruglyak et al. 1996). The frequency of the disease is estimated to be 0.00001, assuming a fully penetrant autosomal recessive mode of inheritance. The allele frequencies of the markers were calculated from a sample of unrelated individuals of the studied population. Multipoint parametric LOD scores were calculated under a hypothesis of genetic homogeneity (LOD) or genetic heterogeneity (HLOD) by using Genehunter. A single locus and a multilocus (up to four consecutive markers) TDT using the tdtx ($x = .2, 3$ or 4) commands of Genehunter were also performed.

Results

Haplotype analysis and homozygosity mapping

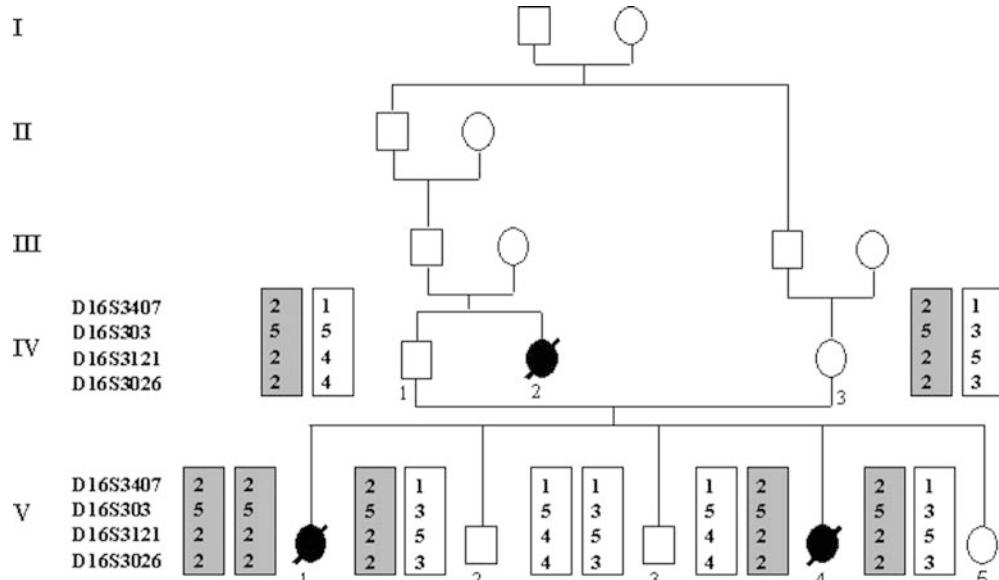
Forty-nine patients, belonging to 39 unrelated families, were subtyped by using linkage analysis with markers linked to previously known FA genes. In view of the higher prevalence of the FAA complementation group, all families were first genotyped with markers overlapping the *FANCA* gene region. For each family, the most likely haplotypes were constructed by visual inspection and by use of Genehunter (genotyping results available on request). Haplotype analysis and homozygosity mapping showed that, except for one multiplex family (SFA6; Fig. 1), all patients belonging to eight multiplex consanguineous families were homozygous by descent for all the markers overlapping the *FANCA* gene region. The affected children of each family had inherited identical haplotypes from their respective parents, whereas the unaffected children had inherited either the maternal or the paternal haplotype carrying the disease or the haplotypes non-transmitted to the affected

children. In these families, segregation of the *FANCA* markers and the *FANCA* gene was observed. Among the 20 simplex families, 17 were consanguineous. Haplotype analysis showed that, except for three families (TFA1, TFA4, and SFA10), all patients belonging to these families were homozygous for the intragenic marker and for at least two other markers closely linked to *FANCA* gene. Furthermore, haplotypes transmitted to affected and to unaffected offspring were different for each family. This result suggested a likely assignment of these simplex families (except for TFA1, TFA4, and SFA10) to complementation group A. TFA4 and SFA10 were non-consanguineous and non-informative families. Among the 10 families with only one affected child, one was non-informative, and patients from the remaining families were homozygous by descent for all markers linked to *FANCA* gene. This result suggested that the *FANCA* gene was likely to be responsible for FA in these families. Multipoint parametric LOD scores were calculated for all families between the *FANCA* gene and the markers overlapping the *FANCA* region. Significant LOD scores were obtained for all markers, with a maximum multipoint LOD score of 8.89 at marker D16S3026 (Table 2), whereas the highest HLOD score

Table 2 Multipoint parametric LOD scores between *FANCA* and the various markers for all the studied families (α proportion of linked families to *FANCA* gene). The intermarker distances are estimated taken from Pronk et al. (1995), Fanconi Anaemia/Breast Cancer Consortium (1996), and Morgan et al. (1999)

Markers	LOD score	(α , HLOD)
D16S3407	$-\infty$	(0.9080, 9.0758)
D16S303	8.612091	(0.9074, 9.8622)
D16S3121	6.234161	(0.8709, 9.5532)
D16S3026	8.890936	(0.9573, 8.9342)
D16S3123	6.572526	(0.8717, 7.6079)
D16S498	5.452369	(0.8728, 6.9302)
D16S520	2.486337	(0.7851, 5.1057)

Fig. 1 Pedigree and haplotype analysis for the SFA6 family with closest markers linked to *FANCA* genes. Each generation is designated by a Roman numeral (I-V), and the members of generations IV and V are designated by Arabic numerals. Affected individual IV-2 has not been investigated (squares males, circles females, slashed symbols deceased individuals, shaded symbols affected individuals)



value of 9.86 was observed at D16S303 ($\alpha=0.9$; α being the proportion of linked families). All HLOD scores were above statistical significance with the smallest value

Table 3 Multipoint parametric LOD scores between *FANCA* and the various markers for the SFA6 family

Markers	LOD score
D16S3407	-∞
D16S303	-2.649551
D16S3121	-4.173705
D16S3026	-1.044777
D16S3123	-0.734242
D16S498	0.457326
D16S520	-0.376292

Fig. 2 Geographic distribution and characteristics of 39 Tunisian FA families. The geographic origins of the patients' ancestors were mainly based on the birth places of the patients' grandparents, as provided by the families participating in this study. Squares Families, asterisks multiplex families, n non-consanguineous families, a families with patients who had received an allogenic bone marrow transplant, filled squares families who carried the first haplotype (2-5-2), hatched squares families who carried the second haplotype (3-5-5-4). For the non-consanguineous families, the paternal and the maternal grandparents of the patients were from the same area

of 5.10 at marker D16S520. When genetic homogeneity was assumed, the LOD scores at these markers dropped, indicating significant genetic heterogeneity (Table 2). These statistical results were not surprising, as the consanguineous family SFA6 excluded linkage to *FANCA* gene (Table 3). Homozygosity mapping excluded linkage with the other FA known genes.

The patient born to the consanguineous family TFA1 was heterozygous for all markers linked to *FANCA* gene and for markers of the *FANCC*, *FANCD2*, *FANCE*, *FANCF*, and *FANCG* genes. The affected and unaffected children did not inherit the same chromosomes from the parents.

In order to identify a possible founder effect in Tunisian families, haplotypes were analyzed in the families consistent with linkage to *FANCA* gene. We found that the more common haplotype 2-5-2 (D16S3407-D16S303-D16S3121) was shared by 12 families, 10 of which were from southern Tunisia, one was from central Tunisia, and one came from northern Tunisia. One other haplotype was observed 3-5-5-4 (D16S3407-D16S303-D16S3121-D16S3026) and was shared by four families from central Tunisia (Table 1, Fig. 2). A single locus and a multilocus TDT were performed for markers D16S3407, D16S303, D16S3121, and D16S3026.

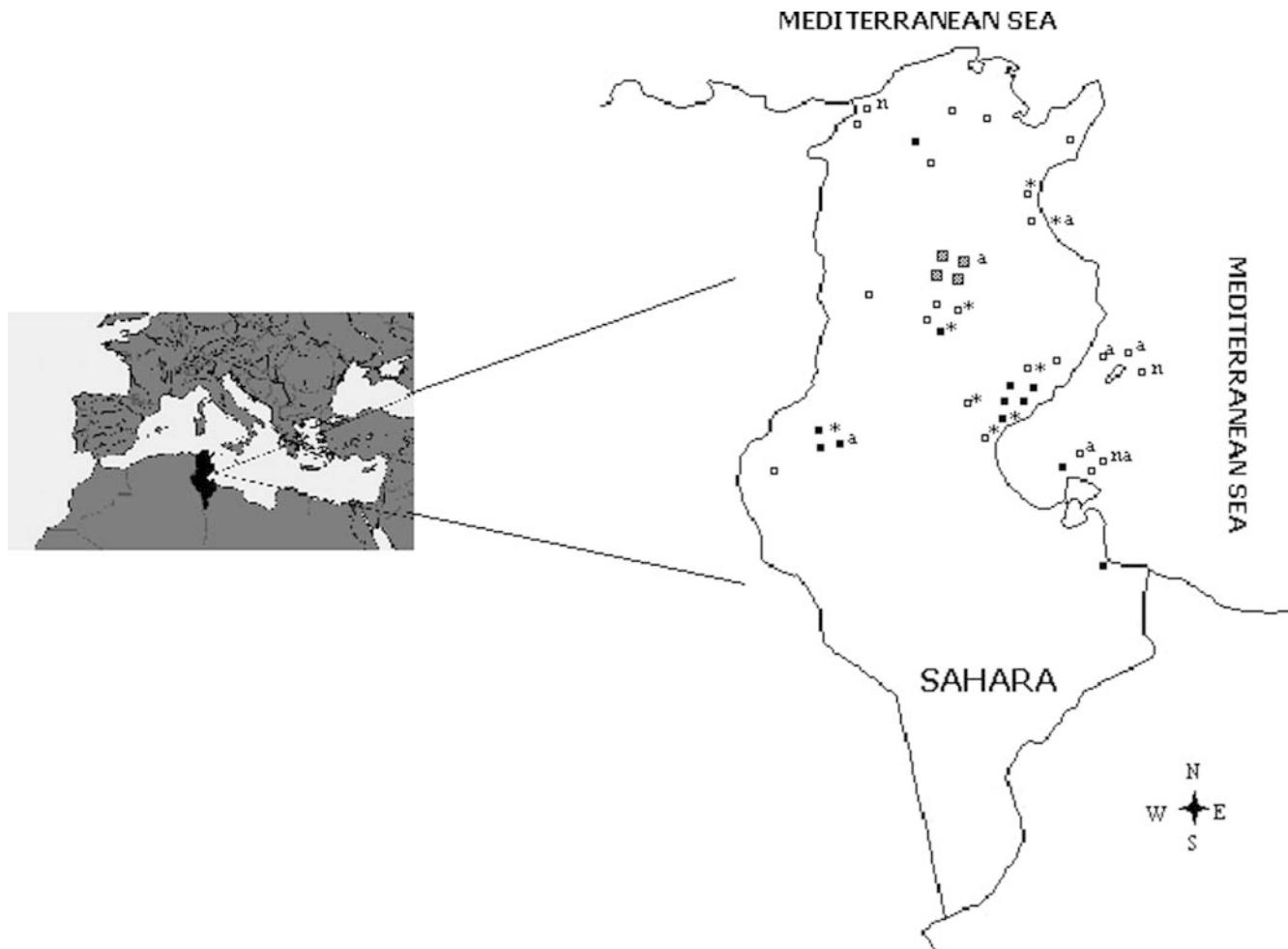


Table 4 Single locus and multilocus TDT for the four proximal markers from the *FANCA* region. *T* and *NT* are respectively the number of transmissions and non-transmissions of the allele or the haplotype. The *TDT value* is thus equal to $(T-NT)^2/(T+NT)$

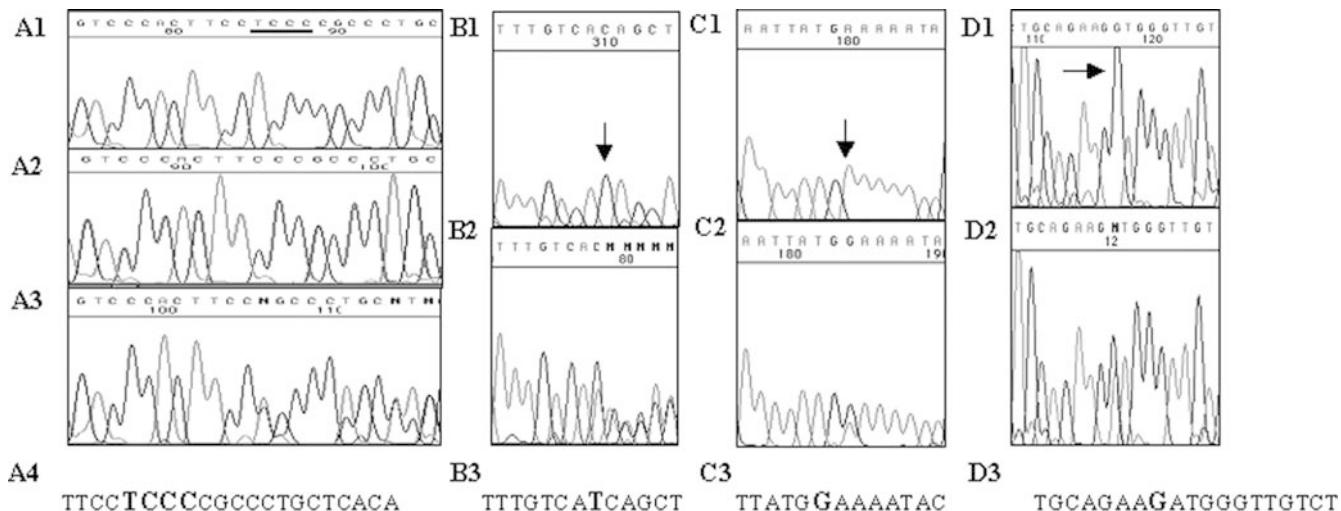
Type of TDT	Allele/haplotype	T	NT	TDT value (χ^2)	P-value
Single locus					
D16S3407 (I)	Allele 2	28	6	14.24	0.000161
D16S303 (II)	Allele 5	13	6	2.58	0.108294
D16S3121 (III)	Allele 2	28	9	9.76	0.001787
D16S3026 (IV)	Allele 2	13	4	4.76	0.029049
Two locus					
I-II	2-5	25	4	5.21	0.000096
II-III	5-2	24	5	12.45	0.000418
III-IV	2-2	14	0	14.00	0.000183
	2-3	12	3	5.40	0.020137
	5-4	12	3	5.40	0.020137
Three locus					
I-II-III	2-5-2	20	4	10.67	0.001091
II-III-IV	5-2-2	10	0	10.00	0.001565
	5-2-3	12	0	12.00	0.000532
Four locus					
I-II-III-IV	2-5-2-2	8	0	8.00	0.004678
	2-5-2-3	12	0	12.00	0.000532
	3-5-5-4	8	1	5.44	0.019631

Different haplotypes were found that were in linkage disequilibrium with the disease locus. Haplotype 2-5-2-3 was particularly interesting with 12 transmissions yielding a TDT value of 12 (Table 4).

Mutation screening

The complete coding sequence of the *FANCA* gene was amplified either from genomic DNA in 26 fragments or

Fig. 3 Sequence analysis of the four mutations identified. *A1* Genomic DNA sequence for the partial exon 19 (*underlined* 4-bp deletion), *A2* 1751–1754del in patient TFA11, *A3* parents are heterozygous carrier for the mutation, *B1* 1606delT in patient TFA9, *B2* parents are heterozygous carrier for the mutation, *C1* 513G → A in patient TFA29, *C2* parents are heterozygous carrier for the mutation, *D1* IVS24 + 166A → G, *D2* parents are heterozygous carrier for the mutation, *A4*, *B3*, *C3*, *D3* DNA sequence flanking each mutation. The deletion consensus sequence CCTG and the direct repeats are *underlined*. The deleted/substituted sequence is in **bold**. The sequence deleted in patient TFA11 (*A4*) is also surrounded by a highly GC-rich region (> 70%), a feature that has frequently been observed at the site of human DNA deletions (Krawczak and Cooper 1991)



from cDNA and screened for mutations by direct sequencing. The search for mutations was started by screening all patients (a total of 43 patients) who were assigned to the FAA group by linkage analysis and homozygosity mapping, for the previously described recurrent mutations (1115–1118del and 3788–3790del) found within *FANCA* in exons 13 and 38, respectively. These mutations were identified in many unrelated patients from various populations (Wijker et al. 1999). We also screened the “Tunisian” (890–893del) and the “Moroccan” (2172–2173insG, 4275delT) mutations occurring in exons 10, 24, and 43, respectively (Tamary et al. 2000). None of these mutations or other mutations was observed in these exons in any of our patients. Screening for mutations in the remaining 38 exons was performed in 16 patients from different families. Patients were selected on the basis of their belonging to the more informative families. A total of four novel homozygous mutations were identified in four unrelated FA patients (Fig. 3).

The first mutation detected in patient TFA11 was a deletion of 4 bp of the sequence (TCCC) in exon 19

(1751–1754del) and created a premature termination codon 20 amino acids downstream (Fig. 3, A2). The second mutation, a deletion of T at position 1606 in exon 17 (1606delT) was found in patient TFA9 in the homozygous state (Fig. 3, B1). This deletion results in a shift of the reading frame and the production of a termination codon 21 amino acids downstream. For patient TFA29, the G to A transition at position 513 (513G → A) led to a substitution of the tryptophan codon with a stop codon (W171X) and, therefore, to a premature truncation of the FANCA protein in exon 5 (Fig. 3, C1). RNA was available from patient CFA2, and sequencing of cDNA revealed an insertion of 166 bp between exons 24 and 25. Sequence analysis of genomic DNA revealed an A to G transition 166 bp downstream from exon 24 (IVS24+166A → G; Fig. 3, D1). This transition results in the use of a downstream cryptic splicing donor site and the insertion of an intronic segment. The score for the cryptic donor splice site was calculated according to the method of Shapiro and Senapathy (1987) and was higher ($s=86$) than the score for the normal signal ($s=78$). The inserted sequence creates a stop codon 40 bp downstream from the tail of exon 24, resulting in a truncated protein of 835 amino acids. Patient CFA2 belonged to the second haplotype, which was shared by three other unrelated patients. Sequencing of exons 24–25 and the intronic region showed that none of these patients carried this mutation, indicating that the second haplotype was not associated with a recurrent mutation and suggesting a mutational heterogeneity affecting Tunisian patients, even those sharing the same haplotype.

For the four mutations found in the four unrelated patients, linkage analysis and haplotype data were consistent with the segregation of the mutation in the parents and the healthy sibs who carried the mutation in a heterozygote state.

In addition to the predicted pathogenic mutations, 11 polymorphisms were detected; nine of these were previously reported and described in the Fanconi Anemia Mutation Database (<http://www.rockefeller.edu/fanconi/mutate>). Two new polymorphisms were identified within intron 24, (IVS24-5G/A and IVS24-6C/G). They might be common variants in our population, because they appeared at homozygous state in the majority of patients and in those of their family members who were studied (80 of 100 chromosomes analyzed).

Discussion

In a total of 39 unrelated families living in various geographic locations in Tunisia, linkage analysis and homozygosity mapping have shown that 94% of these families are likely to belong to the complementation group FAA (Table 1). There is evidence of a high prevalence of this group in Tunisian FA families. A

similar situation has been observed in Italy and Japan (Savino et al. 1997; Tachibana et al. 1999).

For one family (SFA6), haplotype and LOD scores analyses did not show linkage to *FANCA* gene. Homozygosity mapping excluded linkage with the other known FA genes, suggesting that patients born to this consanguineous family could possess mutations in the *BRCA2* gene or may belong to a new complementation group. Nevertheless, because of the hypermutability and the instability of the *FANCA* gene, we cannot exclude the hypothesis of an intrafamilial heterogeneity of mutations. Indeed, in this family, one of the affected children is homozygous for the markers overlapping the *FANCA* gene (including the intragenic marker), and the other is heterozygous for the same markers but does not share the same chromosomes with the unaffected sibs (Fig. 1). This result suggests that the second affected sib may be a compound heterozygous for two different mutant alleles within the *FANCA* gene, one of which would have been inherited from one of the two parents and would correspond to the mutation revealed in the homozygous state in the first affected sib, and the other corresponding to a neomutation. Furthermore, the affected child who is homozygous for the *FANCA* markers carries the most common haplotype (2-5-2) in the homozygous state, whereas the other FA child carries the same haplotype in a heterozygous state. A similar situation of intrafamilial heterogeneity of mutations has been observed for several other disorders within Arab populations (Carrasquillo et al. 1997; Rinat et al. 1999; Ben Arab et al. 2000).

Another patient born to a consanguineous family (TFA1) was heterozygous for all the markers linked to the *FANCA*, *FANCC*, *FANCD2*, *FANCE*, *FANCF*, and *FANCG* genes. Neither the affected nor the unaffected child inherited the same chromosomes from the parents. For this family, the patient may be a compound heterozygous for two different mutant alleles inherited from the two parents and thus not be excluded from any of the tested genes. For the two unclassified patients, TFA1 and SFA6, only a complementation analysis and the screening for mutations will allow molecular diagnosis.

In view of the high prevalence of the FAA group in Tunisia as revealed by linkage analysis, the presence of only two common haplotypes suggests the occurrence of a founder effect for FA in our population, with specific mutations segregating within each haplotype. This situation has been reported in a molecular study of FA families from the Afrikaner population of South Africa, a study showing the presence of five haplotypes associated with four different mutations (Tipping et al. 2001). Somewhat surprisingly, in spite of a high consanguinity rate, our results suggest that patients sharing the same haplotype probably carry different mutations (at least for the second haplotype), providing no evidence of common ancestry. Thus, several independent mutational events have probably affected the *FANCA* gene and are responsible for the disease.

Screening for mutations in the *FANCA* gene has led us to identify four novel mutations in the homozygous state in four unrelated FA patients. These mutations are likely to be pathogenic and would be predicted to generate truncated proteins, although residual function for part of the protein cannot be excluded. Three of these mutations are flanked by motifs (CCTG) and direct repeats (Fig. 3, A4, B3, C3), which have previously been identified as being mutation hot-spot consensus sequences for particularly spontaneous small deletions (<20 bp; Krawczak and Cooper 1991; Smith and Adair 1996). We could amplify *FANCA* cDNA only for patients who were diagnosed at an early age and who did not show profound aplastic anemia. For the others, we failed to amplify the *FANCA* cDNA, although we succeeded in amplifying cDNAs of other genes (e.g., *IFNGR*). The presence of mutations probably reduces transcription of the *FANCA* gene or alters the stability of its mRNA, as reported by others (Wijker et al. 1999). On the other hand, RNA was extracted from 10–20 ml whole blood, and the transcript of *FANCA* gene was reported as being undetectable in unstimulated peripheral blood leukocytes from non-leukemic individuals (J.P. de Winter, unpublished; Joenje and Patel 2001).

We have screened a panel of 16 patients from different geographic areas of Tunisia by direct sequencing of genomic DNA or cDNA to identify the spectrum of mutations in the *FANCA* gene. We have identified only four mutations and 11 polymorphisms, nine of which are common and have been described previously. These data indicate an overall detection rate of 25%; this means that a large proportion of mutations are not detected by current scanning protocols. Several reasons may account for this low detection level. First, FA is genetically a heterogeneous disease, and hence, some of the patients (particularly those belong to simplex or singleton families) may indeed not belong to the FAA group. Second, some patients may carry mutations that are not detected by our screening protocol, such as (1) large intragenic deletions, (2) some mutations that are identified only in cDNA, or (3) mutations localized in either introns or the promoter region.

To our knowledge, this investigation is the first systematic genetic and molecular analysis of FA in the Maghreb region and particularly in Tunisia. The North African population is characterized by a high degree of consanguinity in spite of its various ethnic backgrounds. The rate of consanguineous mating is estimated to 33% in northern Tunisia (Riou et al. 1989). Depending on the area studied, this rate reaches over 60%. This can be explained by the fact that endogamy is culturally favored and serves to keep property within families (Khlat 1997; Steensma et al. 2001); in addition, in some areas, geographic isolation may favor consanguinity. In consequence of the high rate of inbreeding, there is an increase in the prevalence of recessive genetic disorders (Hoodfar and Teebi 1996). This may explain the relatively high incidence of FA in Tunisia (1.4/million/year; M. Frikha, personal communication/in preparation).

The present study has led to several improvements in patient management: (1) genetic counseling in informative families, (2) better identification of bone marrow graft donor, and (3) follow-up of grafted patients by evaluating chimerism (C. Bouchlaka et al., in preparation).

Acknowledgements We acknowledge the participation to this investigation of other members of the Tunisian Fanconi Anemia Study Group; namely Dr. L. Aissaoui, Dr. R. Belakhal, Dr. Z. Belhaj Ali, Dr. B. Meddeb, Dr. A. Hafnia, Dr. M. Elloumi, Dr. F. Fakhfakh PhD, Dr. A. Abdelkafi, Dr. L. Tordjman, Dr. F. Mellouli, Dr. S. Ladeb, Dr. A. Ben Abdellahim, Dr. H. Sennana, Dr. H. Elghezal, Dr. H. Elomri, Dr. A. Laatiri, Dr. A. Khelif, Dr. S. Ennabli, and Dr. M. Trudi. We thank N. Labbane, S. Chakroun, and M. Ben Fadhel, for their technical assistance. We gratefully acknowledge patients and their families for their participation in this study. We thank Hans Joenje for fruitful discussion. This work was supported by funds from the Tunisian State Secretariat for Scientific and Technological Research.

References

- Altay C, Alikasifoglu M, Kara A, Tuncbilek E, Ozbek N, Schroeder-Kurth TM (1997) Analysis of 65 Turkish patients with congenital aplastic anemia (Fanconi anemia and non-Fanconi anemia): Hacettepe experience. Clin Genet 51:296–302
- Alter BP (2003) Cancer in Fanconi anemia, 1927–2001. Cancer 97:425–440
- Balta G, Winter JP de, Kayserili H, Pronk JC, Joenje H (2000) Fanconi anemia A due to a novel frameshift mutation in hotspot motifs: lack of FANCA protein. Hum Mutat 15:578–583
- Ben Arab S, Hmani M, Denoyelle F, Boulila-Elgaied A, Chardonnoux S, Hachicha S, Petit C, Ayadi H (2000) Mutations of GJB2 in three geographic isolates from northern Tunisia: evidence for genetic heterogeneity within isolates. Clin Genet 57:439–443
- Berger R, Le Coniat M, Gendron MC (1993) Fanconi anemia. Chromosome breakage and cell cycle studies. Cancer Genet Cytogenet 69:13–16
- Carrasco MM, Zlotogora J, Barges S, Chakravarti A (1997) Two different connexin 26 mutations in an imbedded kindred segregating non-syndromic recessive deafness: implications for genetic studies in isolated populations. Hum Mol Genet 6:2163–2172
- Centra M, Memeo E, d'Apolito M, Savino M, Ianzano L, Notarangelo A, Liu J, Doggett NA, 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
- Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380:152–154
- Fanconi Anaemia/Breast Cancer Consortium (1996) Positional cloning of the Fanconi anaemia group A gene. Nat Genet 14:324–328
- Frikha M, Mseddi S, Elloumi M, Bouaziz M, Khanfir A, Mnif J, Saad A, Souissi T (1998) La maladie de Fanconi: étude de 43 cas dans le sud Tunisien. Arch Pédiatr 5:1200–1205
- Garcia-Higuera I, Kuang Y, Naf D, Wasik J, D'Andrea AD (1999) Fanconi anemia proteins FANCA, FANCC, and FANCG/XRCC9 interact in a functional nuclear complex. Mol Cell Biol 19:4866–4873
- Garcia-Higuera I, Taniguchi T, Ganesan S, Meyn MS, Timmers C, Hejna J, Grompe M, D'Andrea AD (2001) Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. Mol Cell 7:249–262

- Giampietro PF, Adler-Brecher B, Verlander PC, Pavlakis SG, Davis JG, Auerbach AD (1993) The need for more accurate and timely diagnosis in Fanconi anemia: a report from the International Fanconi Anemia Registry. *Pediatrics* 91:1116–1120
- Hazan J, Dubay C, Pankowiak MP, Becuwe N, Weissenbach J (1992) A genetic linkage map of human chromosome 20 composed entirely of microsatellite markers. *Genomics* 12:183–189
- Hejna JA, Timmers CD, Reifsteck C, Bruun DA, Lucas LW, Jakobs PM, Toth-Fejel S, Unsworth N, Clemens SL, Garcia DK, Naylor SL, Thayer MJ, Olson SB, Grompe M, Moses RE (2000) Localization of the Fanconi anemia complementation group D gene to a 200-kb region on chromosome 3p25.3. *Am J Hum Genet* 66:1540–1551
- Hoodfar E, Teebi AS (1996) Genetic referrals of Middle Eastern origin in a western city: inbreeding and disease profile. *J Med Genet* 33:212–215
- Howlett NG, Taniguchi T, Olson S, Cox B, Waitfisz Q, Die-Smulders C de, Persky N, Grompe M, Joenje H, Pals G, Ikeda H, Fox EA, D'Andrea AD (2002) Biallelic inactivation of BRCA2 in Fanconi anemia. *Science* 297:606–609
- Ianzano L, D'Apolito M, Centra M, Savino M, Levran O, Auerbach AD, Cleton-Jansen AM, Doggett NA, Pronk JC, Tipping AJ, Gibson RA, Mathew CG, Whitmore SA, Apostolou S, Callen DF, Zelante L, Savoia A (1997) The genomic organization of the Fanconi anemia group A (FAA) gene. *Genomics* 41:309–314
- Jakobs PM, Fiddler-Odell E, Reifsteck C, Olson S, Moses RE, Grompe M (1997) Complementation group assignments in Fanconi anemia fibroblast cell lines from north America. *Somat Cell Mol Genet* 23:1–7
- Joenje H (1996) Fanconi anaemia complementation groups in Germany and The Netherlands. European Fanconi Anaemia Research group. *Hum Genet* 97:280–282
- Joenje H, Patel K J (2001) The emerging genetic and molecular basis of Fanconi anaemia. *Nature* 2:446–457
- Joenje H, Lo Ten Foe JR, Oostra AB, Berkel CG van, Rooimans MA, Schroeder-Kurth T, Wegner RD, Gille JJ (1995) Classification of Fanconi anemia patients by complementation analysis: evidence for a fifth genetic subtype. *Blood* 86:2156–2160
- Joenje H, Oostra AB, Wijker M, Summa FM di, Berkel CG van, Rooimans MA, Eboll W, Weel M van, Pronk JC, Buchwald M, Arwert F (1997) Evidence for at least eight Fanconi anemia genes. *Am J Hum Genet* 61:940–944
- Khlat M (1997) Endogamy in the Arab world. In: Teebi AS, Farag TI (eds) *Genetic disorder among Arab populations*. Oxford University Press, New York, pp 63–80
- Krawczak M, Cooper DN (1991) Gene deletions causing human genetic disease: mechanisms of mutagenesis and the role of the local DNA sequence environment. *Hum Genet* 86:425–441
- Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58:1347–1363
- Kupfer GM, Naf D, Suliman A, Pulsipher M, D'Andrea AD (1997) The Fanconi anaemia proteins, FAA and FAC, interact to form a nuclear complex. *Nat Genet* 17:487–490
- Levran O, Erlich T, Magdalena N, Gregory JJ, Batish SD, Verlander PC, Auerbach AD (1997) Sequence variation in the Fanconi anemia gene FAA. *Proc Natl Acad Sci USA* 94:13051–13056
- Levran O, Doggett NA, Auerbach AD (1998) Identification of Alu-mediated deletions in the Fanconi anemia gene FAA. *Hum Mutat* 12:145–152
- Lo Ten Foe JR, Rooimans MA, Bosnayan-Collins L, Alon N, Wijker M, Parker L, Lightfoot J, Carreau M, Callen DF, Savoia A, Cheng NC, van Berkel CG, Strunk MH, Gille JJ, Pals G, Kruyt FA, Pronk JC, Arwert F, Buchwald M, Joenje H (1996) Expression cloning of a cDNA for the major Fanconi anaemia gene, FAA. *Nat Genet* 14:320–323
- Medhurst AL, Huber PA, Waisfisz Q, Winter JP de, Mathew CG (2001) Direct interactions of the five known Fanconi anaemia proteins suggest a common functional pathway. *Hum Mol Genet* 10:423–429
- Morgan NV, Tipping AJ, Joenje H, Mathew CG (1999) High frequency of large intragenic deletions in the Fanconi anemia group A gene. *Am J Hum Genet* 65:1330–1341
- Pronk JC, Gibson RA, Savoia A, Wijker M, Morgan NV, Melchior S, Ford D, Temtamy S, Ortega JJ, Jansen S (1995) Localisation of the Fanconi anaemia complementation group A gene to chromosome 16q24.3. *Nat Genet* 11:338–340
- Rinat C, Wanders RJA, Drukier A, Halle D, Frishberg Y (1999) Primary hyperoxaluria type I: a model for multiple mutations in a monogenic disease within a distinct ethnic group. *J Am Soc Nephrol* 10:2352–2358
- Riou S, El Younsi C, Chaabouni H (1989) Consanguinité dans la population du nord de la Tunisie. *Tun Med* 67:167–172
- Saar K, Schindler D, Wegner RD, Reis A, Wienker TF, Hoehn H, Joenje H, Sperling K, Digweed M (1998) Localisation of a Fanconi anaemia gene to chromosome 9p. *Eur J Hum Genet* 6:501–508
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Laboratory, Cold Spring Harbor, NY
- Savino M, Ianzano L, Strippoli P, Ramenghi U, Arslanian A, Bagnara GP, Joenje H, Zelante L, Savoia A (1997) Mutations of the Fanconi anemia group A gene (FAA) in Italian patients. *Am J Hum Genet* 61:1246–1253
- Savoia A, Zatterale A, Del Principe D, Joenje H (1996) Fanconi anaemia in Italy: high prevalence of complementation group A in two geographic clusters. *Hum Genet* 97:599–603
- Shapiro MB, Senapathy P (1987) RNA splice junctions of different classes of eucaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res* 15:7155–7174
- Siddique MA, Nakanishi K, Taniguchi T, Grompe M, D'Andrea AD (2001) Function of the Fanconi anemia pathway in Fanconi anemia complementation group F and D1 cells. *Exp Hematol* 29:1448–1455
- Smith DG, Adair GM (1996) Characterization of an apparent hotspot for spontaneous mutation in exon 5 of the Chinese hamster HPRT gene. *Mutat Res* 352:87–96
- Steensma DP, Hoyer JD, Fairbanks VF (2001) Hereditary red blood cell disorders in middle eastern patients. *Mayo Clin Proc* 76:285–293
- Strathdee CA, Gavish H, Shannon WR, Buchwald M (1992a) Cloning of cDNAs for Fanconi's anaemia by functional complementation. *Nature* 356:763–767
- Strathdee CA, Duncan AM, Buchwald M (1992b) Evidence for at least four Fanconi anaemia genes including FACC on chromosome 9. *Nat Genet* 1:196–198
- Tachibana A, Kato T, Ejima Y, Yamada T, Shimizu T, Yang L, Tsunematsu Y, Sasaki MS (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 AD, Zaizov R (2000) Fanconi anaemia group A (FANCA) mutations in Israeli non-Ashkenazi Jewish patients. *Br J Haematol* 111:338–343
- Taniguchi T, D'Andrea (2002) The Fanconi anemia protein, FANCE, promotes the nuclear accumulation of FANCC. *Blood* 100:2457–2462
- Timmers C, Taniguchi T, Hejna J, Reifsteck C, Lucas L, Bruun D, Thayer M, Cox B, Olson S, D'Andrea AD, Moses R, Grompe M (2001) Positional cloning of a novel Fanconi anemia gene, FANCD2. *Mol Cell* 7:241–248
- Tipping AJ, Pearson T, Morgan NV, Gibson RA, Kuyt LP, Havenga C, Gluckman E, Joenje H, Ravel T de, Jansen S, Mathew CG (2001) Molecular and genealogical evidence for a founder effect in Fanconi anemia families of the Afrikaner population of South Africa. *Proc Natl Acad Sci USA* 98:5734–5739
- Verlander PC, Lin JD, Udon MU, Zhang Q, Gibson RA, Mathew CG, Auerbach AD (1994) Mutation analysis of the Fanconi anemia gene FACC. *Am J Hum Genet* 54:595–601

- Waisfisz Q, Saar K, Morgan NV, Altay C, Leegwater PA, Winter JP de, Komatsu K, Evans GR, Wegner RD, Reis A, Joenje H, Arwert F, Mathew CG, Pronk JC, Digweed M (1999) The Fanconi anemia group E gene, FANCE, maps to chromosome 6p. *Am J Hum Genet* 64:1400–1405
- Whitney MA, Saito H, Jakobs PM, Gibson RA, Moses RE, Grompe M (1993) A common mutation in the FACC gene causes Fanconi anaemia in Ashkenazi Jews. *Nat Genet* 4:202–205
- Whitney M, Thayer M, Reifsteck C, Olson S, Smith L, Jakobs PM, Leach R, Naylor S, Joenje H, Grompe M (1995) Microcell mediated chromosome transfer maps the Fanconi anaemia group D gene to chromosome 3p. *Nat Genet* 11:341–343
- Wijker M, Morgan NV, Herterich S, Berkel CG van, Tipping AJ, Gross HJ, Gille JJ, Pals G, Savino M, Altay C, Mohan S, Dokal I, Cavenagh J, Marsh J, Weel M van, Ortega JJ, Schuler D, Samochatova E, Karwacki M, Bekassy AN, Abecasis M, Eboll W, Kwee ML, Ravel T de, CG Mathew (1999) Heterogeneous spectrum of mutations in the Fanconi anaemia group A gene. *Eur J Hum Genet* 7:52–59
- Winter JP de, Waisfisz Q, Rooimans MA, Berkel CG van, Bosnayan-Collins L, Alon N, Carreau M, Bender O, Demuth I, Schindler D, Pronk JC, Arwert F, Hoehn H, Digweed M, Buchwald M, Joenje H (1998) The Fanconi anaemia group G gene FANCG is identical with XRCC9. *Nat Genet* 20:281–283
- Winter JP de, Rooimans MA, Weel L van der, Berkel CG van, Alon N, Bosnayan-Collins L, Groot J de, Zhi Y, Waisfisz Q, Pronk JC, Arwert F, Mathew CG, Scheper RJ, Hoatlin ME, Buchwald M, Joenje H (2000a) The Fanconi anaemia gene FANCF encodes a novel protein with homology to ROM. *Nat Genet* 24:15–16
- Winter JP de, Leveille F, Berkel CG van, Rooimans MA, Weel L van der, Steltenpool J, Demuth I, Morgan NV, Alon N, Bosnayan-Collins L, Lightfoot Leegwater PA, Waisfisz Q, Komatsu K, Arwert F, Pronk JC, Mathew CG, Digweed M, Buchwald M, Joenje H (2000b) Isolation of a cDNA representing the Fanconi anemia complementation group E gene. *Am J Hum Genet* 67:1306–1308
- Winter JP de, Weel L van der, Groot J de, Stone S, Waisfisz Q, Arwert F, Scheper RJ, Kruyt FA, Hoatlin ME, Joenje H (2000c) The Fanconi anemia protein FANCF forms a nuclear complex with FANCA, FANCC and FANCG. *Hum Mol Genet* 9:2665–2674