Five Novel Deleterious Variants in *FANCA, FANCF* and *FANCG* Identified in Pakistani Fanconi Anemia Families Using Exome Sequencing

Ashley S. Thompson
National Cancer Institute Division of Cancer Epidemiology and Genetics https://orcid.org/0000-0003-1663-0317
Nusrat Saba
Institute of Biomedical and Genetic Engineering
Lisa J. McReynolds
National Cancer Institute Division of Cancer Epidemiology and Genetics
Saeeda Munir
Institute of Biomedical and Genetic Engineering
Parvez Ahmed
Quaid-i-Azam International Hospital
Sumaira Sajjad
Institute of Biomedical and Genetic Engineering
Kristine Jones
National Cancer Institute Division of Cancer Epidemiology and Genetics
Meredith Yeager
National Cancer Institute Division of Cancer Epidemiology and Genetics
Frank X. Donovan
National Human Genome Research Institute
Settara C. Chandrasekharappa
National Human Genome Research Institute
Blanche P. Alter
National Cancer Institute Division of Cancer Epidemiology and Genetics
Sharon A. Savage
National Cancer Institute Division of Cancer Epidemiology and Genetics
Sadia Rehman (sadiawasiq@gmail.com)
Institute of Biomedical and Genetic Engineering

Research article

Keywords: Fanconi anemia, hematopoietic cell transplantation, genetic testing, inherited bone marrow failure syndrome

Posted Date: October 7th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-52108/v1

Abstract

Background Fanconi anemia (FA), a cancer-prone inherited bone marrow failure syndrome associated with characteristic dysmorphology is primarily caused by autosomal recessive inheritance of pathogenic germline variants in any of 22 different DNA repair genes. Pathogenic variants in *FANCA* are the most frequent cause, followed by *FANCC* and *FANCG.* There are limited data on the specific molecular causes of FA in different ethnic groups**.**

Methods We evaluated 19 patients with FA undergoing hematopoietic cell transplantation evaluation, from 17 families in Pakistan with exome sequencing and copy number variant analysis. To accompany these efforts, we reviewed the literature and curated a list of variants reported in patients with FA from South Asia and the Middle East.

Results The genetic causes for disease were identified in 14 families: 7 *FANCA,* 2 *FANCC*, 1 *FANCF*, 2 *FANCG,* and 2 *FANCL*. Homozygous and compound heterozygous variants were present in 12 and 2 families, respectively. Nine families carried variants previously reported as pathogenic, including two families with the South Asian *FANCL* founder variant. We also identified five novel likely deleterious variants in *FANCA*, *FANCF*, and *FANCG* in affected patients.

Conclusions Our study supports the importance of determining the genomic landscape of FA in diverse populations, in order to improve understanding of FA etiology and assist in the counseling of families.

Background

Fanconi anemia (FA [MIM:227650]) is a cancer-prone inherited bone marrow failure syndrome associated with radial ray abnormalities, characteristic facies, and other medical problems (1, 2). Approximately 5% of patients with FA have the VACTERL-H phenotype (Vertebral anomalies, Anal atresia, Cardiac anomalies, Tracheoesophageal fistula, Esophageal atresia, Renal structural anomalies, Limb anomalies [primarily radii and/or thumbs], and Hydrocephalus) (3). Additional FA phenotypic findings are associated with Pigmentation of the skin, small Head, small Eyes, central Nervous system anomalies (excluding hydrocephalus), Otologic anomalies, and Short stature (PHENOS) (4). Patients with FA have exceedingly high risks of head and neck squamous cell carcinoma (HNSCC) and leukemia compared with the general population (5). FA-associated bone marrow failure (BMF) frequently requires hematopoietic cell transplantation (HCT).

The laboratory diagnosis of FA is based on increased chromosome breakage in lymphocytes or fibroblasts after culture with clastogens such as mitomycin C (MMC) or diepoxybutane (DEB) and can be confirmed by germline genetic testing (6). The majority of individuals with FA have autosomal recessive inheritance of pathogenic germline variants in specific DNA repair genes (7). *FANCB* and *FANCO/RAD51C* are the exceptions, respectively inherited in an X-linked and autosomal dominant pattern. *FANCA* accounts for approximately 65% of cases, while *FANCC* and *FANCG* account for an additional 20% of cases in individuals of European ancestry (8). FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, and FANCM comprise the core FA complex (upstream) in conjunction with FANCT/UBE2T which is responsible for recognizing interstrand crosslinks and activating the FA DNA repair pathway through interaction with FANCD2 and FANCI (the ID complex). The downstream pathway facilitates DNA repair by homologous recombination performed by FANCD1/BRCA2, FANCJ/BRIP1, FANCN/PALB2, FANCO/RAD51C, FANCP/SLX4, FANCQ/ERCC4, FANCR/RAD51, FANCS/BRCA1, FANCU/XRCC1, FANCV/REV7, and FANCW/RFWD3 (9–11).

There have been a limited number of reports on the genetic etiology of FA in populations from South Asia and the Middle East. Such studies have identified several novel disease-causing germline genetic variants, including the first reports of FA caused by pathogenic variants in *FANCO/ RAD51C* or *FANCE* (12–40), and highlight the importance of germline genetic studies of FA in underrepresented regions. In this report, we evaluated the genetic causes of FA in 19 patients from 17 unrelated families being considered for HCT in Pakistan.

Methods

Study subjects

This project was approved by the ethical review committee of the Institute of Biomedical and Genetic Engineering (IBGE, Islamabad, Pakistan). Individuals with FA and their first-degree relatives were evaluated by their referring physicians. Chromosome breakage using MMC on primary lymphocytes was performed at the Armed Forces Institute of Bone Marrow Transplant Center (ABMTC, Rawalpindi, Pakistan) (41). All patients had chromosome breakage results consistent with FA and severe BMF necessitating evaluation for HCT. De-identified blood-derived DNA samples from 17 families, including 19 patients with FA and 33 relatives, were sent to the National Cancer Institute's Cancer Genomics Research Laboratory (NCI CGR, Gaithersburg, MD, USA) for sequencing and genotyping; a list of all study participants and their disease status is in Supplementary Table 1.

Sequence Analysis

Whole exome sequencing (WES) was performed at NCI's CGR as previously described (42). Variants were filtered based on their presence as homozygous or biallelic in previously identified FA-associated genes. The pattern of inheritance was considered when DNA from family members was available. Variant annotation was performed using ANNOVAR (43) and the computational resources of the National Institutes of Health (NIH) High Performance Computing (HPC) Biowulf cluster (http://hpc.nih.gov). Additional filters applied include a Genome Aggregation Database (gnomAD) minor allele frequency (MAF) < 1% in all populations, bioinformatic prediction tools (Criteria for deleterious variants: MetaSVM > 0, REVEL ≥ 0.5, and CADD phred > 20), and clinical significance indicators (ClinVar and InterVar) (44–49). Potential splice site variants were assessed using Human Splicing Finder (50). We used BAM-matcher and vcftools to assess potential consanguinity as described (51–53).

Copy number variations (CNV) were detected using VarSeq™ v2.1 (VS-CNV), which analyzes changes in WES coverage between the sample and controls (54), and the detected CNVs were visually evaluated using GenomeBrowse® (Golden Helix, Inc., Bozeman, MT) (55, 56). Homozygous or heterozygous deletions were filtered based on p-values (< 0.001) and annotated using ClinVar (48). Z-scores and ratio values were assessed when validating genotypes. Suspected CNV events underwent validation using targeted whole gene sequencing as described (57).

DNA from families 3-FA, 8-FA, 9-FA, 12-FA, 14-FA, and 16-FA was also sequenced using a targeted custom capture design (Roche, Inc) for next generation sequencing (NGS) designed to include all known FA genes and FA candidate genes, including all intronic regions and 5 kb upstream and downstream of each gene, as previously described (57). Sequence reads were aligned to human genome build 19 (GRCh37) using Burrows-Wheeler Alignment tool and variants were called using HaplotypeCaller and the GATK best practices pipeline for germline variants (58–60). SNVs and indels were annotated using ANNOVAR (43) CNVs were detected from the targeted NGS reads and annotated using Nexus Copy Number version 10.0 (BioDiscovery, Inc.).

Additionally, DNA from family 17-FA was sequenced using PacBio® long-range sequencing technology with custom IDT xGen® Lockdown® Probes designed to capture all intronic and exonic regions of *FANCA*. The manufacturer's PacBio® protocols were followed for shearing genomic DNA, end repair, ligation of linear barcoded adapters, amplification, sample pooling, and capturing using IDT xGen® Lockdown® Probes. Libraries were prepared using SMRTbell® protocol for primer annealing, polymerase binding, and sequencing on the Sequel system. Circular consensus reads were generated using default parameters (3 passes, 0.99 accuracy) and demultiplexed according to parameters for symmetrical barcodes. Sequence reads were aligned to human genome build 19 (GRCh37) and structural variants were called using pbsv default parameters.

Variant Curation: FA in South Asia and the Middle East

A comprehensive literature review was performed using the The National Library of Medicine's PubMed database using the following search terms combined with countries in South Asia and the Middle East, (*e.g.* "Fanconi anemia and Pakistan" or "Fanconi and India") to curate previously published FA gene variants reported in patients from the following regions: Afghanistan, Bangladesh, Egypt, India, Iran, Iraq, Israel, Jordan, Lebanon, Nepal, Oman, Pakistan, Saudi Arabia, Syria, Turkey,

and Yemen. Large cohort studies and case reports which only consisted of phenotypic data and did not report patients' specific genotypes were excluded. Studies which only reported FA subtypes by complementation testing were also excluded unless further sequencing efforts revealed the specific variant(s) in the patient(s).

Results

Patient characteristics

There were 19 patients with FA (16 males and 3 females) from 17 unrelated families evaluated in this study (Fig. 1). The majority of families (12/17, 70%) were from Northern or Central Punjab. Other families were from Southern Punjab, Islamabad, Khyber Pakhtunkhwa, and Azad Kashmir. The median age at FA diagnosis was 7 years (range 4–12) and while all were evaluated for HCT, only 5 patients underwent matched sibling HCT. Eight of the 19 patients were deceased at the time of this study. The median age at death was 8.5 years (range 4–13). Pathogenic variants relevant to FA were identified in 14 families with *FANCA* being the most common (7/14, 50%). Homozygous variants in FA-associated genes were identified in 12 of the 14 solved families (86%) and 2 probands had compound heterozygous variants. Physical and genetic findings are listed in Tables 1 and 2, respectively.

Table 1
Clinical features of study participants with Fanconi anemia.

Patient ID†	Gender	Age at Diagnosis (Years), Vital Status	Age at and Cause of Death	Skin	Hand	Renal	Other
1-FA	M	7, deceased	9, NR			Ectopic kidneys	
3-FA	M	5, alive	–	Café au lait macules	Thumb malformation		Short stature
4-FA	M	7, alive	–				
5-FA	M	11, deceased	12, NR				Abnormal left leg growth, short stature
6-FA	M	9, deceased	10, hemorrhage NOS	Hyperpigmentation	Polydactyly		
7-FA	M	7, deceased	8, NR				
8-FA	M	4, deceased	4, brain hemorrhage		Absent radii and metacarpal bones		High arched palate, skeletal malformations NOS
9-FA	F	10, alive	–				
10-FA	M	5, deceased	6, hemorrhagic stroke		Polydactyly		
12-FA	M	11, alive	–				
14-FA	M	12, deceased	13, NR				
16-FA	F	9, alive	–		Bilateral thumb malformations		
17_01-FA	M	6, alive	–		Thumb malformation		Immunodeficiency
17_02-FA	M	10, alive	–		Thumb malformation		Immunodeficiency
18-FA	M	5, alive	–		Polydactyly, small right hand	Ectopic kidneys	Short stature
19-FA	F	8, unknown	–				
20-FA	M	5, alive	–				
21_01-FA	M	8, deceased	13, NR				Numerous rosettes with central eosinophilic material on bone marrow biopsy

† All participants had bone marrow failure and were being considered or underwent hematopoietic cell transplantation. Empty boxes indicate where clinical information was not available.

Patient ID†	Gender	Age at Diagnosis (Years), Vital Status	Age at and Cause of Death	Skin	Hand	Renal	Other
21_02-FA	M	5, alive	–				Numerous rosettes with central eosinophilic material on bone marrow biopsy

† All participants had bone marrow failure and were being considered or underwent hematopoietic cell transplantation. Empty boxes indicate where clinical information was not available.

Table 2
Germline genetic variants identified in study participants with Fanconi anemia.

Gene	Patient ID	Genotype	Variant	Previously reported	ClinVar	gnomAD % MAF‡ All; South Asian ancestry (44)	Consequence of Variant§ (45–47)
FANCA	1-FA	Homozygous	c.3788_3790delTCT, p.Phe1263del	(12)	Pathogenic	0.009929; 0	In-frame deletion
	3-FA	Compound Heterozygous	g.89871674-89880557del	(14)	NR		Exons 4–7 Deletion
			g.89861527-89863726del		NR		Exon 11 Deletion
	4-FA	Homozygous	c.37dupC, p.Gln13Profs*24		NR		Frameshift
	9-FA	Homozygous	g.89856782-89874222del	(15)	NR		Exons 7–14 Deletion
	17_01-FA and 17_02-FA	Compound Heterozygous	c.2749C > T, p.Arg917*	(14)	Pathogenic		Stop-gain
			g.89847600-89853759del	(17)	NR		Exons 15–17 Deletion
	19-FA	Homozygous	c.4070C > A, p.Ala1357Asp		NR		Missense, MetaSVM 0.9149, REVEL 0.702, CADD phred 24.1
	20-FA	Homozygous	c.1541C > A, p.Ala514Asp		NR	**0.000398; 0.003266**	Missense, MetaSVM 0.876, REVEL 0.851, CADD phred 27.1
FANCC	5-FA	Homozygous	c.1642C > T, p.Arg548*	(16)	Pathogenic	0.004954; 0.01307	Stop-gain
	8-FA	Homozygous	c.1642C > T, p.Arg548*				
FANCF	10-FA	Homozygous	c.785T > G, p.Leu262*		NR	0.0007955; 0.003266	Stop-gain
FANCG	18-FA	Homozygous	c.710C > G, p.Ser237*		NR		Stop-gain
	21_01-FA and 21_02-FA	Homozygous	c.1471_1473delAAAinsG, p.Lys491Glyfs*9	(17)	NR	**0.0003976; 0.003266**	Frameshift
FANCL	6-FA	Homozygous	c.1092G > A, p.Trp341_Lys364del	(18)	NR	**0.001994; 0.01634**	Exon 13 Skipping
	7-FA	Homozygous	c.1092G > A, p.Trp341_Lys364del				

‡ % MAF in bold indicate these variants were only reported in South Asian populations. Blank cells indicate the variant was not present in gnomAD.

§ *in silico* tools were used to predict the pathogenicity of missense variants.

FANCA

Patient 1-FA presented at seven years of age with aplastic anemia which progressed to severe BMF. He had ectopic kidneys, but no other phenotypic features were reported. WES revealed a homozygous in-frame deletion in exon 38 of *FANCA* (c.3788_3790delTCT, p.Phe1263del, NC_000016.9:g.89807250_89807252delAGA, rs397507553, ClinVar:41003). Both parents were unaffected carriers and his sibling was wild-type. The c.3788_3790delTCT variant is the most frequently reported *FANCA* variant and has been observed in multiple populations throughout the world, including FA patients from Pakistan (12, 19, 61), with a particularly high prevalence in Spain and Brazil (12, 13).

Patient 3-FA had an abnormal thumb (Fig. 3A) and café au lait spots noted at birth. He also had short stature and low gonadotrophin hormone levels. FA was diagnosed by chromosome breakage on primary lymphocytes after he presented with neutropenia that progressed to severe BMF at 5 years of age. He underwent successful HLA-matched sibling donor HCT at the age of 6 years. We identified two deletions in *FANCA* (NC_000016.9:g.89871674_89880557del, and NC_000016.9:89861527_89863726del) affecting exons 4–7 and 11, respectively. These two deletions have been previously reported in Indian FA patient (14). The exon 11 deletion was paternally inherited, while the deletion of exons 4–7 was maternally inherited. Validation by targeted sequencing methods determined that one unaffected sibling did not carry either deletion. Another unaffected sibling was predicted to be a carrier of the exon 4–7 deletion by VS-CNV but there was insufficient DNA for sequencing validation.

Patient 4-FA presented with severe bone marrow failure at the age of 7 years. Hemophagocytosis was reported on his bone marrow biopsy but no other phenotypic information was available. He underwent successful HLA-matched sibling HCT. A homozygous *FANCA* frameshift variant in exon 1 (c.37dupC, p.Gln13Profs*24, NC_000016.9:g.89882986dupG) was identified by WES. One unaffected sibling was wild-type and the other was a carrier, but parental DNA was not available for analysis.

Patient 9-FA presented with neutropenia which progressed to severe BMF and was diagnosed with FA at 10 years of age. WES revealed and targeted whole gene sequencing validated a large homozygous deletion of exons 7–14 (NC_000016.9:g.89856782_89874222del). Her unaffected sister was a heterozygous carrier. This specific deletion has not been previously reported, but similar large deletions in *FANCA* have been reported (62).

Affected brothers 17_01-FA and 17_02-FA both presented with abnormal thumbs at birth (Fig. 3B). Small ear canals were also noted in 17_01-FA (Fig. 3C). At the ages of 6 and 10 years, respectively, they presented with severe BMF and immunodeficiency leading to an FA diagnosis. Biallelic variants in *FANCA* were identified by various NGS methods in both siblings (c.2749C > T, p.Arg917*, NC_000016.9:g.89831327G > A and NC_000016.9:g.89847600-89853759del). WES revealed a maternally-inherited nonsense variant in exon 28 which has been previously identified in an Indian patient with FA and other populations (rs1060501880, ClinVar:408188) (14, 61). A large deletion of exons 15–17 (NC_000016.9:g.89847600-89853759del) was detected by targeted PacBio® long-range sequencing in both affected siblings and has been previously reported in other patients with FA (17). This deletion was not detected in DNA from father's peripheral blood, but relatedness analyses confirmed paternity with large regions of homozygosity being consistent with offspring from a consanguineous relationship between third-degree relatives. Additionally, analyses of single nucleotide polymorphisms (SNP) in the *FANCA* locus provided evidence for a possible a genotypic reversion in the paternal hematopoietic stem cells or paternal inheritance as a result of gonadal mosaicism. Both such occurrences have been previously reported in patients with FA (6, 63–65).

Patient 19-FA presented at 8 years of age with neutropenia that progressed to severe BMF by age 9 years. A homozygous *FANCA* missense variant in exon 41 (c.4070C > A, p.Ala1357Asp, NC_000016.9:g.89805638G > T) was identified by WES. Her unaffected brother is a heterozygous carrier, but parental DNA was not available. *FANCA* p.Ala1357Asp is not present in gnomAD and is predicted deleterious by *in silico* tools (MetaSVM score = 0.915, REVEL = 0.702, CADD phred = 24.1).

Patient 20-FA presented with moderate aplastic anemia that progressed to severe BMF by 5 years of age. We identified a homozygous missense variant in *FANCA* (c.1541C > A, p.Ala514Asp, NC_000016.9:g.89849440G > T, rs1432656621). His

unaffected sibling is a heterozygous carrier. Although not previously reported, this missense variant is rare in gnomAD at a MAF of 0.0003977% and is predicted deleterious by *in silico* tools (MetaSVM score = 0.876, REVEL = 0.851, CADD phred = 27.1).

FANCC

Two unrelated probands, 5-FA and 8-FA, were homozygous for the same *FANCC* variant (c.1642C > T, p.Arg548*, NC_000009.11:g.97864024G > A). Relatedness analyses determined that these probands were from distinct families. *FANCC* p.Arg548* (rs104886457, ClinVar:12047) has been previously reported in two FA patients from Pakistan (16).

Patient 5-FA presented with moderate aplastic anemia and progressed to severe BMF at the age of 11 years. He also had short stature and abnormal left leg growth. He had two brothers and one sister who died due to similar complications but without a diagnosis. His two surviving unaffected siblings and parents are all heterozygous carriers.

Patient 8-FA presented with moderate aplastic anemia and progressed to severe BMF by 4 years of age. Skin hyperpigmentation, bone deformities including the absence of metacarpals, thumbs, and radii, and a high arched palate were also reported. He died at the age of 4 years due to a brain hemorrhage before HLA-matched sibling HCT could be performed. One unaffected sibling is a carrier, but parental DNA was not available.

FANCF

Patient 10-FA was homozygous for nonsense variant in *FANCF* (c.785T > G, p.Leu262*, NC_000011.9:g.22646572A > C, rs368067979). He was diagnosed with FA at 6 years of age when aplastic anemia progressed to severe BMF. He also had polydactyly and died from a hemorrhagic stroke shortly after his FA diagnosis. Parental DNA was not available and the sibling available for testing was not a carrier.

FANCG

Patient 18-FA had an extra digit, a small right hand, short stature, and ectopic kidneys. He was diagnosed with FA at 5 years of age and underwent HCT from his HLA-matched sister for severe BMF and is doing well. A homozygous nonsense variant in exon 6 of *FANCG* was identified by WES (c.710C > G, p.Ser237*, NC_000009.11:g.35077035G > C). His sibling is wild-type at this locus. The only parent who was available for testing was heterozygous for this loss of function variant.

Proband 21_01-FA and his brother 21_02-FA were diagnosed with FA at the ages of 8 and 5 years, respectively. 21_01-FA had pancytopenia that progressed rapidly following his diagnosis with FA and he died due to a brain hemorrhage. Currently, 21_02-FA does not have cytopenias. The bone marrow of both FA-affected brothers was reported to have numerous rosettes with central eosinophilic material surrounded by small cells seen in a background of fibrosis. The affected brothers have a homozygous frameshift variant in exon 11 of *FANCG* (c.1471_1473delAAAinsG, p.Lys491Glyfs*9, NC_000009.11:g.35075283_35075285delTTTinsC, rs1018027137). One of their unaffected siblings was heterozygous for this variant. This variant has been previously reported in a heterozygous patient with FA (17).

FANCL

Unrelated probands 6-FA and 7-FA were both homozygous for a recently identified *FANCL* South Asian founder variant. Donovan *et al.* established this single nucleotide variation (NC_000002.11:g.58387243C > T) induces aberrant mRNA splicing to skip exon 13 (c.1021_1092del, p.Trp341_Lys364del, rs577063114), resulting in a 24 amino acid deletion from the RING domain of FANCL (18). The genomAD MAF of this variant is 0.001994% in all populations and 0.01634% in South Asian populations.

A pairwise comparison between cases 6-FA and 7-FA was performed to assess potential relationships. A genotype comparison on approximately 7300 common SNPs between the probands and their siblings showed no indication of relatedness between families 6-FA and 7-FA. Parental sequencing data was not available.

Patient 6-FA had an extra thumb and areas of skin hyperpigmentation. He presented with aplastic anemia that progressed to severe BMF at the age of 9 years. Although he was treated with androgens while awaiting an HLA sibling matched HCT, he died

at 10 years of age due to an unspecified hemorrhage. Patient 7-FA was diagnosed with FA after presenting with severe BMF at the age of 7 years old. He died shortly after his diagnosis at the age of 8 years due to an unreported cause.

Gene Unknown Families

Rare heterozygous variants in FA pathway genes identified in probands 12-FA, 14-FA, and 16-FA are reported in Supplementary Table 2. These individuals had chromosome breakage testing consistent with FA. Proband 12-FA had no dysmorphology but was diagnosed with FA at the age of 11 years after presenting with BMF and underwent a successful HLA-matched sibling HCT. No rare deleterious variants were identified in the 22 FA-associated genes.

FA was diagnosed in proband 14-FA at 12 year of age and he died 1 year after diagnosis due to an unreported cause. 14-FA was a heterozygous carrier for a variant of uncertain significance (VUS) in *FANCN* and a likely benign *FANCO* variant.

Bilateral thumb malformations were noted at birth in proband 16-FA. She was diagnosed with aplastic anemia at age 9. She underwent a successful HLA-matched sibling HCT. Heterozygous VUS were present in *FANCA*, *FANCD2*, *FANCI*, and *FANCP*. The *FANCP* variant (c.2209C > T, p.Arg737Cys, NC_000016.9:g.3642818G > A, rs140706384) may be deleterious as it has a REVEL score of 0.449 and CADD score of 26.2, but the MetaSVM score was predicted as tolerated; additional functional studies are required to determine potential pathogenicity. There were no other deleterious variants or large CNV events detected in *FANCP*, so this patient remains gene unknown.

Variants in patients with FA from South Asia and the Middle East

Figure 2A is a heat map showing the distribution of studies published from South Asia and the Middle East on the genetic etiology of FA. Reports from Turkey, India, and Pakistan were the most common, followed by Iran and Saudi Arabia. Figure 2B describes the variants reported to cause FA in multiple populations in South Asia and the Middle East. The majority of genetic studies of patients with FA from South Asia and the Middle East focused on targeted sequencing efforts for *FANCA* (10 out of 29 studies).

The majority of reported variants occurred in *FANCA* and were private to their respective populations. All large deletions, SNPs, and small insertion/deletion variants reported in patients with FA in these regions can be found in Supplementary Tables 3, 4, and 5, respectively. The only large deletions reported were in *FANCA*, similar to our findings and consistent with others. We also identified the recently identified *FANCL* founder mutation (c.1092G > A, p.Trp341_Lys364del) in 2 families from Pakistan (18).

Discussion

Identification of the genetic causes of rare diseases such as FA is important to verify diagnoses, improve clinical management, allow for appropriate genetic counseling, and understand the underlying pathobiology of these disorders. The genetic cause of FA was identified in 16 patients from 14 families in this study; three families remain molecularly undiagnosed. *FANCA* was the most commonly affected gene with pathogenic *FANCA* variants present in 50% of the families. The other pathogenic variants were present in *FANCC, FANCG*, or *FANCL* (2 families each), and *FANCF* (1 family). Only 6 of the 14 variants identified in our study were present in any gnomAD population and 3 of these were solely in South Asian populations (Table 2).

Homozygosity for pathogenic variants was present in 12 of the 14 families. Only two families had compound heterozygous inheritance (both in *FANCA*). Two families were homozygous for the *FANCL* founder variant (18). We were unable to evaluate consanguinity in 14 of the families in this study due to lack of parental DNA samples. However, the presence of homozygous pathogenic variants in our data is consistent with prior studies reporting an approximately 70% rate of consanguineous marriages in Pakistan (66–68). Nine of the 14 pathogenic variants identified have been previously reported (Table 2) (12, 14–18).

Our study was limited only to patients with FA who also had BMF severe enough to warrant HCT evaluation. Nevertheless, this investigation sheds further light onto the type and frequencies of germline pathogenic variants associated with FA in the Pakistani population. The patients included in this study are likely a very small minority of FA cases in Pakistan. It is possible

that in the absence of overt dysmorphology, the diagnosis of FA may be delayed or missed because chronic malnutrition and its complications may confound rare disease diagnoses (66).

Conclusion

The majority of studies performed on the genetics of FA in South Asia and the Middle East included targeted sequencing efforts for *FANCA* with *FANCC*, *FANCG*, and *FANCE* studied less frequently, but are more common than other subtypes. *FANCA*, *FANCC*, and *FANCG* account for upwards of 85% of FA cases in those of European descent, but it is imperative to understand the genetic landscape of FA across all 22 subtypes when assessing FA in diverse populations as the prevalence of subtypes may vary in South Asia and the Middle East from those of European ancestry. While large cohort studies are not common in these regions for identifying specific genotypes, there have been several studies published describing the phenotypes of pediatric patients presenting with aplastic anemia, MDS, and/or AML in combination with other FA clinical features (i.e. VACTERL-H and PHENOS) and/or positive chromosome breakage testing. Without performing genetic testing on suspected FA patients, the genetic heterogeneity of FA in these populations could remain under-reported. Large genotype-phenotype studies of patients with FA around the world are required to better understand the genetic variation in diverse populations in order to uncover disease etiology, improve diagnostics and patient management, as well as provide genetic counseling.

Abbreviations

ABBREVIATION	FULL TERM/PHRASE
FA	Fanconi anemia
VACTERL-H	Vertebral anomalies, Anal atresia, Cardiac anomalies, Tracheoesophageal fistula, Esophageal atresia, Renal structural anomalies, Limb anomalies [primarily radii and/or thumbs], and Hydrocephalus
PHENOS	Pigmentation of the skin, small Head, small Eyes, central Nervous system anomalies (excluding hydrocephalus), Otologic anomalies, and Short stature
HNSCC	Head and neck squamous cell carcinoma
BMF	Bone marrow failure
HCT	Hematopoietic cell transplantation
MMC	Mitomycin C
DEB	Diepoxybutane
IBGE	Institute of Biomedical and Genetic Engineering
ABMTC	Armed Forces Institute of Bone Marrow Transplant Center
NCI CGR	National Cancer Institute's Cancer Genomics Research Laboratory
WES	Whole exome sequencing
NIH	National Institutes of Health
HPC	High Performance Computing
gnomAD	Genome Aggregation Database
MAF	Minor allele frequency
CNV	Copy number variant/variation
VS-CNV	VarSeq™ Copy Number Variation Caller
NGS	Next generation sequencing
VUS	Variant of uncertain significance
MDS	Myelodysplastic syndrome
AML	Acute myeloid leukemia
SNP	Single nucleotide polymorphism

Declarations

ETHICS APPROVAL AND CONSENT TO PARTICIPATE This project was approved by the ethical review committee of the Institute of Biomedical and Genetic Engineering (IBGE, Islamabad, Pakistan). Written informed consent was obtained from all study participants or the parent or guardian of participants who were under the age of 18 years old.

CONSENT FOR PUBLICATION: Written informed consent for the publication of identifying images or other personal or clinical details of participants that have the potential to compromise anonymity was obtained from all study participants or the parent or guardian of participants who were under the age of 18 years old.

AVAILABILITY OF DATA AND MATERIALS The deidentified genomic data generated in this study are available upon reasonable request from Dr. Sadia Rehman (sadiawasiq@gmail.com, Institute of Biomedical and Genetic Engineering, Islamabad, Pakistan) through collaboration agreements.

COMPETING INTERESTS There are no competing interest to report.

FUNDING This study was supported by Institute of Biomedical and Genetic Engineering (IBGE) for sample collection and DNA preparation, while the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute supported for whole exome sequencing (WES) and data analysis.

AUTHORS' CONTRIBUTIONS Data acquisition: NS, SM, PA, SS, SR; Data analysis and interpretation: AST, LJM, KJ, MY, FXD, SCC, BPA, SAS, SR; Drafted manuscript: AST, SAS; All authors approved of the final manuscript.

ACKNOWLEDGEMENTS The authors would like to acknowledge all the patients and their family members for providing samples and clinical details. Additional acknowledgement is extended to the laboratory technicians of IBGE for their help during sample collection and DNA extraction. The authors also acknowledge the DCEG CGR Sequencing Group for their work on this project. Kedest Teshome, PacBio® sequencing. Herb Higson, Illumina® whole exome sequencing and targeted capture sequencing. Mingyi Wang, Dongjing Wu, and Wen Luo performed standard data analysis for Illumina® sequencing efforts. This research was supported by the Institute of Biomedical and Genetic Engineering (IBGE) and Intramural Research Program of the National Cancer Institute, Division of Cancer Epidemiology and Genetics.

References

1. Kottemann MC, Smogorzewska A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. Nature. 2013;493(7432):356–63.
2. Fiesco-Roa MO, Giri N, McReynolds LJ, Best AF, Alter BP. Genotype-phenotype associations in Fanconi anemia: A literature review. Blood Rev. 2019;37:100589.
3. Alter BP, Rosenberg PS. VACTERL-H Association and Fanconi Anemia. Mol Syndromol. 2013;4(1–2):87–93.
4. Alter BP, Giri N. Thinking of VACTERL-H? Rule out Fanconi Anemia according to PHENOS. Am J Med Genet A. 2016;170(6):1520–4.
5. Alter BP, Giri N, Savage SA, Rosenberg PS. Cancer in the National Cancer Institute inherited bone marrow failure syndrome cohort after fifteen years of follow-up. Haematologica. 2018;103(1):30–9.
6. Fargo JH, Rochowski A, Giri N, Savage SA, Olson SB, Alter BP. Comparison of chromosome breakage in non-mosaic and mosaic patients with Fanconi anemia, relatives, and patients with other inherited bone marrow failure syndromes. Cytogenet Genome Res. 2014;144(1):15–27.
7. Knies K, Inano S, Ramirez MJ, Ishiai M, Surralles J, Takata M, et al. Biallelic mutations in the ubiquitin ligase RFWD3 cause Fanconi anemia. J Clin Invest. 2017;127(8):3013–27.
8. Wang AT, Smogorzewska A. SnapShot. Fanconi anemia and associated proteins. Cell. 2015;160(1–2):354–e1.
9. Kitao H, Takata M. Fanconi anemia: a disorder defective in the DNA damage response. Int J Hematol. 2011;93(4):417–24.
10. Rodriguez A, D'Andrea A. Fanconi anemia pathway. Curr Biol. 2017;27(18):R986-R8.
11. Inano S, Sato K, Katsuki Y, Kobayashi W, Tanaka H, Nakajima K, et al. RFWD3-Mediated Ubiquitination Promotes Timely Removal of Both RPA and RAD51 from DNA Damage Sites to Facilitate Homologous Recombination. Mol Cell. 2017;66(5):622–34. e8.
12. Castella M, Pujol R, Callen E, Trujillo JP, Casado JA, Gille H, et al. Origin, functional role, and clinical impact of Fanconi anemia FANCA mutations. Blood. 2011;117(14):3759–69.
13. Levran O, Erlich T, Magdalena N, Gregory JJ, Batish SD, Verlander PC, et al. Sequence variation in the Fanconi anemia gene FAA. Proc Natl Acad Sci U S A. 1997;94(24):13051–6.
14. Solanki A, Mohanty P, Shukla P, Rao A, Ghosh K, Vundinti BR. FANCA Gene Mutations with 8 Novel Molecular Changes in Indian Fanconi Anemia Patients. PLoS One. 2016;11(1):e0147016.
15. Esmail Nia G, Fadaee M, Royer R, Najmabadi H, Akbari MR. Profiling Fanconi Anemia Gene Mutations among Iranian Patients. Arch Iran Med. 2016;19(4):236–40.

16. Aftab I, Iram S, Khaliq S, Israr M, Ali N, Jahan S, et al. Analysis of FANCC gene mutations (IVS4 + 4A > T, del322G, and R548X)in patients with Fanconi anemia in Pakistan. Turk J Med Sci. 2017;47(2):391–8.
17. Gille JJ, Floor K, Kerkhoven L, Ameziane N, Joenje H, de Winter JP. Diagnosis of Fanconi Anemia: Mutation Analysis by Multiplex Ligation-Dependent Probe Amplification and PCR-Based Sanger Sequencing. Anemia. 2012;2012:603253.
18. Donovan FX, Solanki A, Mori M, Chavan N, George M, C SK, et al. A founder variant in the South Asian population leads to a high prevalence of FANCL Fanconi anemia cases in India. Hum Mutat. 2019.
19. Shahid M, Firasat S, Satti HS, Satti TM, Ghafoor T, Sharif I, et al. Screening of the FANCA gene mutational hotspots in the Pakistani fanconi anemia patients revealed 19 sequence variations. Congenit Anom (Kyoto). 2019.
20. Aymun U, Iram S, Aftab I, Khaliq S, Nadir A, Nisar A, et al. Screening for mutations in two exons of FANCG gene in Pakistani population. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2017;161(2):158–63.
21. Vaz F, Hanenberg H, Schuster B, Barker K, Wiek C, Erven V, et al. Mutation of the RAD51C gene in a Fanconi anemia-like disorder. Nat Genet. 2010;42(5):406–9.
22. Shukla P, Rao A, Ghosh K, Vundinti BR. Identification of a novel large intragenic deletion in a family with Fanconi anemia: first molecular report from India and review of literature. Gene. 2013;518(2):470–5.
23. Moghadam AA, Mahjoubi F, Reisi N, Vosough P. Investigation of FANCA gene in Fanconi anaemia patients in Iran. Indian J Med Res. 2016;143(2):184–96.
24. Balta G, de Winter JP, Kayserili H, Pronk JC, Joenje H. Fanconi anemia A due to a novel frameshift mutation in hotspot motifs: lack of FANCA protein. Hum Mutat. 2000;15(6):578.
25. Shamseldin HE, Elfaki M, Alkuraya FS. Exome sequencing reveals a novel Fanconi group defined by XRCC2 mutation. J Med Genet. 2012;49(3):184–6.
26. Ghazwani Y, AlBalwi M, Al-Abdulkareem I, Al-Dress M, Alharbi T, Alsudairy R, et al. Clinical characteristics and genetic subtypes of Fanconi anemia in Saudi patients. Cancer Genet. 2016;209(4):171–6.
27. Zareifar S, Dastsooz H, Shahriari M, Faghihi MA, Shekarkhar G, Bordbar M, et al. A novel frame-shift deletion in FANCF gene causing autosomal recessive Fanconi anemia: a case report. BMC Med Genet. 2019;20(1):122.
28. Solanki A, Kumar Selvaa C, Sheth F, Radhakrishnan N, Kalra M, Vundinti BR. Characterization of two novel FANCG mutations in Indian Fanconi anemia patients. Leuk Res. 2017;53:50–6.
29. Tamary H, Bar-Yam R, Shalmon L, Rachavi G, Krostichevsky M, Elhasid R, et al. Fanconi anaemia group A (FANCA) mutations in Israeli non-Ashkenazi Jewish patients. Br J Haematol. 2000;111(1):338–43.
30. Tamary H, Dgany O, Toledano H, Shalev Z, Krasnov T, Shalmon L, et al. Molecular characterization of three novel Fanconi anemia mutations in Israeli Arabs. Eur J Haematol. 2004;72(5):330–5.
31. Koc A, Pronk JC, Alikasifoglu M, Joenje H, Altay C. Variable pathogenicity of exon 43del (FAA) in four Fanconi anaemia patients within a consanguineous family. Br J Haematol. 1999;104(1):127–30.
32. Wijker M, Morgan NV, Herterich S, van Berkel CG, Tipping AJ, Gross HJ, et al. Heterogeneous spectrum of mutations in the Fanconi anaemia group A gene. Eur J Hum Genet. 1999;7(1):52–9.
33. Kalb R, Neveling K, Hoehn H, Schneider H, Linka Y, Batish SD, et al. Hypomorphic mutations in the gene encoding a key Fanconi anemia protein, FANCD2, sustain a significant group of FA-D2 patients with severe phenotype. Am J Hum Genet. 2007;80(5):895–910.
34. Salem AM, El-Bassyouni HT, El-Kamah GY, Zarouk WA, Eid MM, Mosaad RM, et al. Screening for common mutations in four FANCA gene exons in Egyptian Fanconi anemia patients. Middle East Journal of Medical Genetics. 2014;3(1):24–30.
35. Vundinti BR. Chromosomal instability and molecular mutations in multi spectrum disease of Fanconi anemia. Mol Cytogenet. 2014;7(Suppl 1 Proceedings of the International Conference on Human):I47.
36. Wegner RD, Henrichs I, Joenje H, Schroeder-Kurth T. Fanconi anemia complementation group E: clinical and cytogenetic data of the first patient. Clin Genet. 1996;50(6):479–82.
37. Waisfisz Q, Saar K, Morgan NV, Altay C, Leegwater PA, de Winter JP, et al. The Fanconi anemia group E gene, FANCE, maps to chromosome 6p. Am J Hum Genet. 1999;64(5):1400–5.

38. de Winter JP, Leveille F, van Berkel CG, Rooimans MA, van Der Weel L, Steltenpool J, et al. Isolation of a cDNA representing the Fanconi anemia complementation group E gene. Am J Hum Genet. 2000;67(5):1306–8.
39. Aslan D, Ameziane N, De Winter JP. Molecular diagnosis of Fanconi anemia with next-generation sequencing in a case with subtle signs and a negative chromosomal breakage test. Turk J Pediatr. 2015;57(3):282–5.
40. Dorsman JC, Levitus M, Rockx D, Rooimans MA, Oostra AB, Haitjema A, et al. Identification of the Fanconi anemia complementation group I gene, FANCI. Cell Oncol. 2007;29(3):211–8.
41. Cervenka J, Arthur D, Yasis C. Mitomycin C test for diagnostic differentiation of idiopathic aplastic anemia and Fanconi anemia. Pediatrics. 1981;67(1):119–27.
42. Ballew BJ, Yeager M, Jacobs K, Giri N, Boland J, Burdett L, et al. Germline mutations of regulator of telomere elongation helicase 1, RTEL1, in Dyskeratosis congenita. Hum Genet. 2013;132(4):473–80.
43. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.
44. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv. 2019:531210.
45. Kim S, Jhong JH, Lee J, Koo JY. Meta-analytic support vector machine for integrating multiple omics data. BioData Min. 2017;10:2.
46. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet. 2016;99(4):877–85.
47. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46(3):310–5.
48. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res. 2018;46(D1):D1062-D7.
49. Li Q, Wang K, InterVar. Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. Am J Hum Genet. 2017;100(2):267–80.
50. Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 2009;37(9):e67.
51. Wang PP, Parker WT, Branford S, Schreiber AW. BAM-matcher: a tool for rapid NGS sample matching. Bioinformatics. 2016;32(17):2699–701.
52. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010;42(7):565–9.
53. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010;26(22):2867–73.
54. Fortier N, Rudy G, Scherer A. Detection of CNVs in NGS Data Using VS-CNV. Methods Mol Biol. 2018;1833:115–27.
55. Golden Helix I. VarSeq™ v2.1. Bozeman. MT: Golden Helix, Inc.; 2018.
56. Golden Helix I. GenomeBrowse® v2.1. Bozeman. MT: Golden Helix, Inc.; 2018.
57. Chandrasekharappa SC, Lach FP, Kimble DC, Kamat A, Teer JK, Donovan FX, et al. Massively parallel sequencing, aCGH, and RNA-Seq technologies provide a comprehensive molecular diagnosis of Fanconi anemia. Blood. 2013;121(22):e138-48.
58. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–60.
59. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43(5):491–8.
60. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics. 2013;43:11. 0 1 – 0 33.

61. Kimble DC, Lach FP, Gregg SQ, Donovan FX, Flynn EK, Kamat A, et al. A comprehensive approach to identification of pathogenic FANCA variants in Fanconi anemia patients and their families. Hum Mutat. 2018;39(2):237–54.
62. Ameziane N, Sie D, Dentro S, Ariyurek Y, Kerkhoven L, Joenje H, et al. Diagnosis of fanconi anemia: mutation analysis by next-generation sequencing. Anemia. 2012;2012:132856.
63. Gregory JJ Jr, Wagner JE, Verlander PC, Levran O, Batish SD, Eide CR, et al. Somatic mosaicism in Fanconi anemia: evidence of genotypic reversion in lymphohematopoietic stem cells. Proc Natl Acad Sci U S A. 2001;98(5):2532–7.
64. Gross M, Hanenberg H, Lobitz S, Friedl R, Herterich S, Dietrich R, et al. Reverse mosaicism in Fanconi anemia: natural gene therapy via molecular self-correction. Cytogenet Genome Res. 2002;98(2–3):126–35.
65. Krausz C, Riera-Escamilla A, Chianese C, Moreno-Mendoza D, Ars E, Rajmil O, et al. From exome analysis in idiopathic azoospermia to the identification of a high-risk subgroup for occult Fanconi anemia. Genet Med. 2019;21(1):189–94.
66. National Institute of Population Studies II. Pakistan Demographic and Health Survey. National Institute of Population Studies, ICF International; 2013.
67. Hussain R, Bittles AH. The prevalence and demographic characteristics of consanguineous marriages in Pakistan. J Biosoc Sci. 1998;30(2):261–75.
68. Mobarak AM, Chaudhry T, Brown J, Zelenska T, Khan MN, Chaudry S, et al. Estimating the Health and Socioeconomic Effects of Cousin Marriage in South Asia. J Biosoc Sci. 2019;51(3):418–35.

Figures

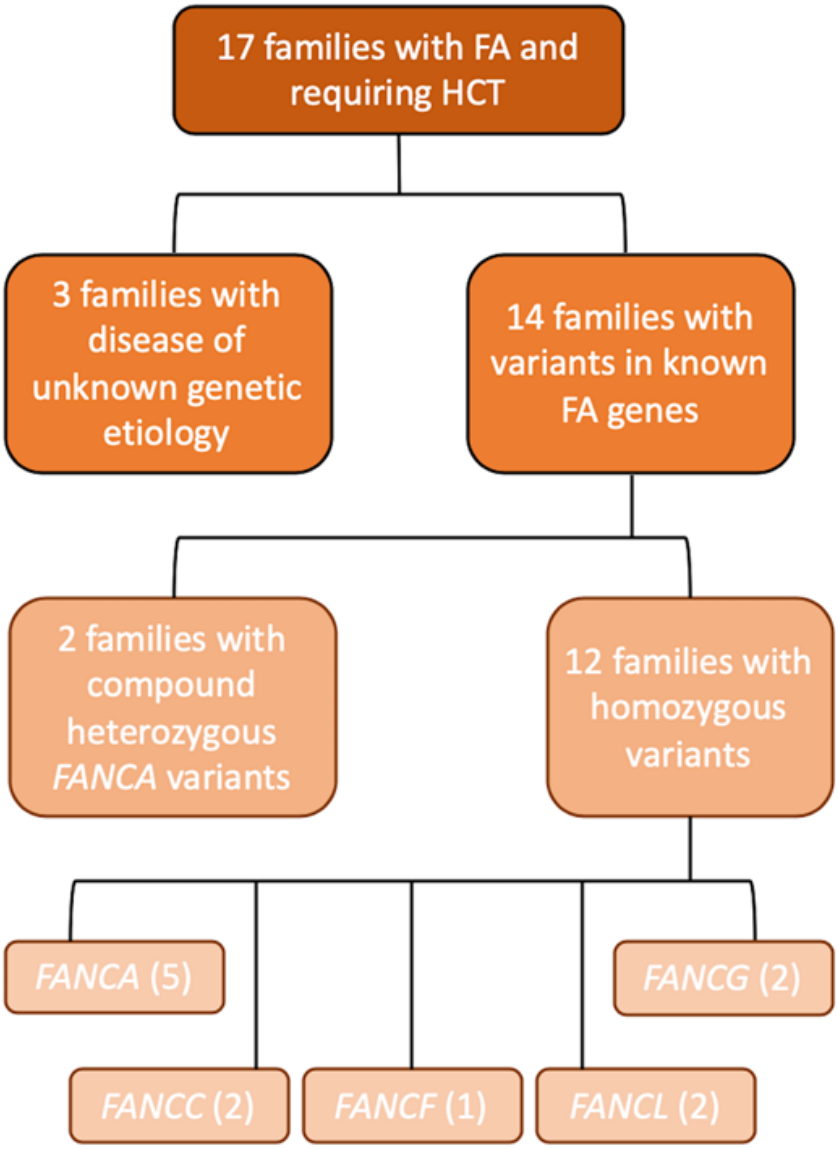

Figure 1

A flowchart of the genetic findings in 19 patients with FA from 17 unrelated families evaluated in this study.

A

B

Gene	Variant	Reported Populations	gnomAD % MAF All	Consequence of Variant
FANCA	c.4261-2A>C	Iran[15], Israel[30]	0.0004008	Exon 42 Skipping
	c.3761_3762delAG, p.Glu1254Glyfs*23	India[14], Turkey[13]	NR	Frameshift
	c.3788_3790delTCT, p.Phe1263del	Iran[15], Pakistan[12,19]	0.009929	Small deletion
	Exon 21 deletion	India[14], Iran[17]	NR	
	Exon 30 deletion	India[14], Iran[15]	NR	
	Exon 31 deletion	India[14], Pakistan[32]	NR	
	Exon 43 deletion	Iran[32], Turkey[31,32]	NR	
	Exon 1-22 deletion	India[14], Iran[15]	NR	
FANCL	c.1092G>A, p.Trp341_Lys364del	India, Pakistan[18]	0.001994	Exon 13 Skipping

Figure 2

(A) A heat map showing the distribution of studies published from South Asia and the Middle East on the genetic etiology of FA. This heat map was generated by the primary author using a tool available in Microsoft PowerPoint (v16.40). (B) The variants reported to cause FA in multiple populations in South Asia and the Middle East. Abbreviations: MAF, minor allele frequency; NR, Not reported. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

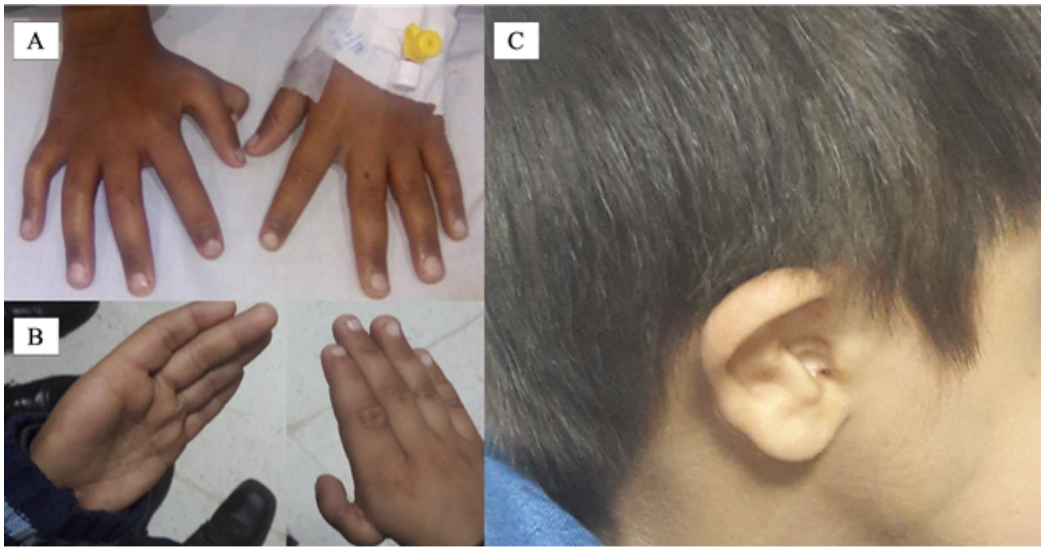

Figure 3

(A) An abnormal right thumb noted at birth in 3-FA, (B) an absent left thumb and abnormal right thumb in patient 17_01-FA, and (C) a small ear canal seen in Patient 17_01-FA.