UNDERSTANDING MOLECULAR MECHANISMS OF FANCONI ANAEMIA

GAURAV JOSHI

PhD THESIS

2024

SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCE AND TECHNOLOGY, THIRUVANATHAPURAM

An Institution of National Importance established by an Act of the Indian Parliament (Act No.52 of 1980)

Dept. of Science and Technology, Govt. of India

www.sctimst.ac.in

UNDERSTANDING MOLECULAR MECHANISMS OF FANCONI ANAEMIA

GAURAV JOSHI

PhD THESIS

2024

SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCE AND TECHNOLOGY, THIRUVANATHAPURAM

An Institution of National Importance established by an Act of the Indian Parliament (Act No.52 of 1980)

Dept. of Science and Technology, Govt. of India

www.sctimst.ac.in

DECLARATION BY THE STUDENT

CERTIFICATE

I, **Gaurav Joshi** hereby certify that I had personally carried out the work depicted in the thesis titled, "**Understanding molecular mechanisms of Fanconi anaemia".**

No part of this thesis has been submitted for the award of any other degree or diploma prior to this date.

Signature: 

Name of the Candidate: Gaurav Joshi

Reg No: 2017/PHD(CMC,V)/11

Date: 11.02.2024

Institute for Stem Cell Science and Regenerative Medicine

An autonomous Institutee of
Department of Biotechnology,
Government of India.

inStem
CSCR
CENTER FOR STEM CELL RESEARCH
(a unit of inStem, Bengaluru)
Christian Medical College Campus
Bagayam, Vellore - 632002.

Center for Stem Cell Research

(A unit of inStem, Bengaluru in Collaboration with DBT and CMC, Vellore)

NOT TO BE MINISTERED UNTO BUT TO MINISTER
CHRISTIAN MEDICAL COLLEGE
VELLORE
INDIA

CERTIFICATE

Name of the Guide: **Dr. R.V. Shaji**

Division/Department: **Centre for Stem Cell Research (a Unit of inStem, Bengaluru)**

This is to certificate that **Gaurav Joshi,** Center for Stem Cell Research of Christian Medical College has fulfilled the requirement prescribed for the Ph.D degree of the Sree Chitra Tirunal Institute for Medical Science and Technology, Thiruvananthapuram.

The thesis entitled, **"Understanding molecular mechanisms of Fanconi anaemia"**, was carried out under my direct supervision. No part of the thesis was submitted for the award of any degree or diploma prior this date.

*Clearance was obtained from the Institutional review board, the institutional bio safety committee, the Institutional committee for stem cell research and therapy and Institutional Animal ethics committee for carrying out the study.

Signature of the guide
Dr. R.V. Shaji
Professor, Department of Haematology,
Adjunct Scientist, Centre for Stem Cell Research (A Unit of inStem, Bengaluru),
Christian Medical College- Vellore

Date:

Christian Medical College Campus,
Bagayam, Vellore - 632002. INDIA.

Phone : +91 - 416 - 228 5101 / 5107
Fax : +91 - 416 - 228 5103
E-mail : cscr@cmcvellore.ac.in

APPROVAL OF THE THESIS

The thesis entitled

UNDERSTANDING MOLECULAR MECHANISMS OF FANCONI ANAEMIA

Submitted by

GAURAV JOSHI

for the degree of

Doctor of Philosophy

of

Sree Chitra Tirunal Institute for Medical Sciences and Technology,

Trivandrum

Is evaluated and approved by

.....	.....
Dr.Shaji RV	(Name & Signature of the
(Name & Signature of the Guide)	external examiner

ACKNOWLEDGEMENTS

I extend my gratitude to my supervisor, **Dr. R.V. Shaji**, for his guidance and expertise and introducing me to Fanconi Anaemia. Your constant support and encouragement have been instrumental in shaping this research. I am grateful for the countless hours spent in discussions, reviewing drafts, and providing constructive feedback, all of which have significantly enriched the quality of this thesis.

I am deeply indebted to the esteemed members of the Doctoral Advisory Committee (DAC). **Dr. Biju George**, your insightful feedback and valuable suggestions have significantly enriched this work. Your extensive knowledge and dedication to research have been a source of inspiration. I would like to express my gratitude for your clinical expertise in the subject, which has significantly contributed to my better understanding of the disease. **Dr. Poonkuzhali Balasubramanian**, your valuable inputs during the DAC meetings have been invaluable in refining the scope of this thesis.

I would like to express my sincere gratitude to **Dr. Alok Srivastava, Head of the Department, Centre for Stem Cell Research (a Unit of inStem, Bengaluru)** for his invaluable support throughout the course of my thesis work. I am deeply grateful for his insightful feedback, which has helped me refine my ideas and approach, ensuring that this thesis meets the highest academic standards. I owe sincere gratitude to **Dr. Uday Kulkarni** for his invaluable help throughout my research. He has been instrumental in helping me with the collection of clinical data. I would also thank HOD, **Dr. Aby Abraham** for his support throughout my research. I also thank **Dr. Anu Korula, Dr. Fouzia**, **Dr. Sharon,** and **Dr. Sushil** for their support in the present study. I also thank **Dr. Eunice Sindhuvi** for her support through my research, especially in providing Aplastic anaemia sample data.

Heartfelt thanks to my best friend **Aswin**, an amazing roommate for 10 years whose presence and friendship made every moment enjoyable. I want to express my sincere gratitude for your unwavering support during the frustrating times throughout this journey. Your presence and encouragement have been a constant source of strength for me, and I cannot thank you enough for being by my side when things got tough. I

also acknowledge your invaluable contribution to the statistical analysis of my thesis. I am deeply grateful to **Sonam** for her unwavering support and invaluable assistance throughout this journey. Her knowledge and approachable nature have made her my go-to person for scientific doubts, and her belief in my abilities has been a driving force that kept me going, even in challenging times. Her friendship and guidance have shaped me academically and personally, making this academic journey so much easier and enjoyable. I deeply thank **Esther** for her invaluable scientific support, friendship, and constant encouragement throughout this journey. Her expertise and genuine presence have made this experience truly enriching. Grateful to **Krittika** for enriching discussions. Your insights and support made this journey meaningful. You are an exceptional friend and an inspiration throughout this endeavour. I extend my heartfelt gratitude to **Prasanth** for his invaluable experimental assistance and being an incredible friend. I also would like to thank **Dr. Abhirup** for being helpful with sharing cloning techniques and being a good friend. Also, I would like to thank **Dr. Vasanth** for the technical expertise and being a helpful colleague. I am extremely thankful to **Kirti** for her invaluable assistance in conducting IMF experiments and engaging in enriching scientific discussions.

I'm thankful to **Dr. Dinesh Babu**, for our helpful discussions and teamwork. Your presence in the lab has created a friendly atmosphere of learning and collaboration. My heartfelt thanks to the dedicated and skilled technician **Sweety** for patient diagnosis. I also thank **Chitra, Dhavapriya, Praveena, Jothi,** and **Rakshini**. Your technical expertise and support in the laboratory have been indispensable to the success of this research. A special mention of gratitude to our lab manager, **Chudar**, for efficiently managing the lab. Your organizational skills have been invaluable. I would also like to express my appreciation to the core staff at CSCR, **Dr. Sandya Rani**, **Abdul**, and **Joel**, for their assistance and collaboration. Your support in imaging and flow acquisition for this research is deeply appreciated. I would like to extend my heartfelt appreciation to **Tamil Vanan J.**, for his invaluable assistance and support in facilitating access to the extensive resources available in the library, which greatly enriched the depth of this thesis.

I'd like to express my sincere thanks to **Dr. Nancy** for her exceptional teaching on ficoll separation, the first technique that I learned, and for expanding my understanding of Fanconi anaemia. Her dedication and expertise have been instrumental in broadening my knowledge in these crucial areas of research. I also want to thank **Dr. Madhavi** and **Dr. Phaneendra** for their valuable insights and guidance on bioinformatics analysis. I thank **Revathi** for helping with my administrative and financial aspects. I also thank **Sam Prabhakar** and **Rekha Singh** for their help in acquiring materials and reagents for the experiments. I thank **Dr. Nithya** for scientific discussions and suggestions. Also, a big thank you to my lab seniors **Dr. Sumitha, Dr. Thiyagaraj, Dr. Kannan, Dr. Musheer, Dr. Aneesha** for sharing their experience and colleagues **Smitha, Karthik, Anurag** and **Debanjan** for being there.

I would like to extend my gratitude to **Dr. Santhosh Kumar** and **Ms. Radha** for their coordination in the thesis submission process. Their assistance and support have been crucial in ensuring a smooth and efficient submission experience.

I also thank **Dr. Ram Prasad, Dr. Sakthivel, Dr. Murugan, Dr. Vivek, Dr. Thenral and Kaviyarasan** from MedGenome Labs for Exome sequencing and data analysis. I extend my thanks to **Madavan** and **Hina** from Theomics for their help in bioinformatics analysis.

I would also like to thank all the patients and their families for providing samples for the study.

Furthermore, I am immensely thankful to the **WELLCOME-DBT Alliance** for their generous support in financing my research. I also thank **UGC** for my fellowship. I also take this opportunity to thank **DBT-CTEP** for providing travel grants for attending and presenting my work at different international conferences.

I am also thankful to **Anand Pai** and **Suchitra Pai** for treating me as their family and being supportive. I am deeply grateful to my sister, **Dr. Anjali**, for being a constant pillar of strength in my life and for offering unwavering emotional support. Her presence and encouragement have been a source of solace during challenging times, and her guidance has helped me navigate through various hurdles. I express my heartfelt gratitude to my parents for everything they have done for me throughout my life. Their unwavering love, support, and sacrifices have shaped me into the person I

am today. No words can fully express the depth of my appreciation for their love and care. I am blessed to have such amazing parents who have always been there for me, and I will forever cherish the bond we share. Thank you, **Mommy** and **Papa.**

Lastly, I want to thank **CMC** for providing me an opportunity to work a such a prestigious institution. Finally, I would like to acknowledge all the individuals who may not be named here but have contributed in various ways to the completion of this thesis. Your collective efforts and support have made this endeavor possible. Thank you all for being an integral part of my research journey, and for making it a truly enriching and memorable experience. Finally, I am grateful to God for bestowing upon me the strength and guidance that have helped me stay grounded and maintain my sanity.

Gaurav Joshi

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

S.No	ABBREVIATION	FULL FORM
1	2A peptide	2A self cleaving peptide
2	AA	Aplastic Anaemia
3	ABI	Applied Biosystems
4	ACMG	American College of Medical Genetics and Genomics
5	AF488	Alexa Fluor 488
6	ALDH2	Aldehyde Dehydrogenase 2
7	ALL	Acute Lymphoblastic Leukaemia
8	ALT	Alternative Lengthening of Telomeres
9	AML	Acute Myeloid Leukaemia
10	AMP	Adenosine Monophosphate
11	AP	Alkaline Phosphatase
12	ATP	Adenosine Triphosphate
13	BED	Browser Extensible Data
14	BFU-E	Burst forming unit erythroid
15	BMF	Bone Marrow Failure
16	BMFS	Bone Marrow Failure Syndrome
17	BMI	Body Mass Index
18	bp	Base Pair
19	BSA	Bovine serum albumin
20	BWA	Burrows Wheeler Aligner
21	Cas9	CRISPR associated protein 9
22	CBA	Chromosome Breakage Analysis
23	CBC	Complete Blood Cell Count
24	cDNA	Complementary DNA
25	CFSs	Common Fragile Sites
26	CFU-E	Colony forming unit erythroid
27	CFU-G	Colony forming unit granulocyte
28	CFU-GEMM	Colony forming unit granulocyte, erythroid, monocyte, megakaryocyte
29	CFU-GM	Colony forming unit granulocyte monocyte
30	CFUM	Colony forming unit monocyte
31	CMV	Cytomegalovirus
32	CNV	Copy Number Variant
33	CO_2	Carbon Dioxide

34	CRISPR-cas9	Clustered Regularly Interspaced Short Palindromic Repeats CRISPR associated protein 9
35	DAPI	4',6-diamidino-2-phenylindole
36	dbSNP	The Single Nucleotide Polymorphism Database (a public database of genetic variations)
37	DEB	Diepoxybutane
38	DMEM	Dulbecco's Modified Eagle Medium
39	DNA	Deoxyribonucleic Acid
40	DNase I	Deoxyribonuclease I
41	DOX	Doxycycline
42	DRP1	Dynamin Related Protein 1
43	DSB	DNA Double Strand Break
44	E3	Ubiquitin Ligase
45	EdU	5-ethynyl-2'-deoxyuridine.
46	EF1a	Elongation Factor 1a
47	EMT	Epithelial Mesenchymal Transition
48	ESC	Embryonic Stem Cells
49	EVE	Evolutionary model of variant effect
50	EVS	Exome Variant Server (a database of human genetic variations)
51	ExomeDepth	A bioinformatics tool for detecting copy number variants from exome sequencing data
52	FA	Fanconi Anaemia
53	FA-iPSCs	Fanconi anaemia induced pluripotent stem cells
54	FBS	Foetal Bovine Serum
55	Fe-S	Iron sulphur
56	G4-G	G quadruplexes
57	GATK	Genome Analysis Toolkit
58	gnomAD	Genome Aggregation Database (a database of human genetic variations)
59	GRCh37/hg19	Genome Reference Consortium human reference genome build 37 / human genome build 19
60	HCT	Haematopoietic Stem Cell Transplantation
61	HEK	Human Embryonic Kidney
62	HGMD	Human Gene Mutation Database (a database of human germline mutations associated with inherited diseases)
63	hiPSCs	Human Induced Pluripotent Stem Cells
64	HIV-1	Human Immunodeficiency Virus Type 1
65	HPV	Human Papilloma Virus
66	HR	Homologous Recombination
67	HSCs	Haematopoietic Stem Cells

68	HSCT	Haematopoietic Stem Cell Transplantation
69	HSPCs	Haematopoietic Stem and Progenitor Cells
70	HSV-1	Herpes Simplex Virus 1
71	HU	Hydroxyurea
72	IBMFS	Inherited Bone Marrow Failure Syndrome
73	ICLs	Interstrand Crosslinks
74	IDT	Integrated DNA Technologies (a biotechnology company)
75	IFN-γ	Interferon gamma
76	IL1-β	Interleukin beta
77	IMDM	Iscove's Modified Dulbecco's Medium
78	iPSCs	Induced Pluripotent Stem Cells
79	IRES	Internal Ribosome Entry Site
80	LB	Lysogeny broth (a common growth medium for bacteria)
81	LC3	Microtubule associated Protein Light Chain 3
82	LP	Lipid Peroxidation
83	LRT	Likelihood Ratio Test
84	LTR	Long Terminal Repeats
85	LV	Lentiviral Vectors
86	MD	Mitochondrial Disease
87	mdDNA	Mitochondrial DNA
88	MDS	Myelodysplastic Syndrome
89	MLPA	Multiplex Ligation dependent Probe Amplification
90	MLS	Mitochondrial Localization Signal
91	MMC	Mitomycin C
92	MRI	Magnetic Resonance Imaging
93	mtDNA	Mitochondrial DNA
94	mtROS	Mitochondrial ROS
95	NAC	N-acetyl-cysteine
96	NANOG	Homeobox protein NANOG
97	NEB	New England Biolabs
98	NER	Nucleotide Excision Repair
99	NGS	NextGeneration Sequencing
100	NHEJ	Nonhomologous End Joining
101	NLS	Nuclear Localization Signals
102	NSAA	Non-Severe Aplastic Anaemia
103	OCT4	Octamer binding transcription factor 4
104	p53	Tumour Protein 53
105	PAX6	Paired box 6
106	PBMNCs	Peripheral Blood Mononuclear Cells

107	PBS	Phosphate Buffered Saline
108	PCR	Polymerase Chain Reaction
109	PHA	Phytohemagglutinin
110	PHENOS	Skin Pigmentation, small Head, small Eyes, Nervous system, Otology, and Short stature
111	PKR	eukaryotic translation initiation factor 2alpha kinase 2
112	PKR	Protein Kinase R
113	PolyPhen-2	Polymorphism Phenotyping v2
114	PRDX3	Peroxiredoxin3
115	Ptet	TetOff promoter based on the Escherichia coli Tet repressor protein (TetR)
116	qPCR	Quantitative polymerase chain reaction
117	rDNA	Ribosomal DNA
118	ROS	Reactive Oxygen Species
119	RPM	Revolutions per minute
120	RPMI	Roswell Park Memorial Institute
121	RPS27L	Ribosomal Protein S27like
122	RT	Room temperature
123	rTA	Reverse tetracycline controlled transactivator
124	rtTA	Reverse tetracycline controlled transactivator
125	SAA	Severe Aplastic Anaemia
126	SALSA	Synthetic Amplification of the Ligated Samples
127	SCC	Squamous Cell Carcinoma
128	SCEs	Sister Chromatid Exchanges
129	SDS	PAGE Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
130	SEM	Standard error of the mean
131	SFFV	Spleen Focus Forming Virus
132	SIFT	Sorting Intolerant From Tolerant
133	SNV	Single Nucleotide Variant
134	SOD1	Superoxide Dismutase 1
135	SOX17	SRY box transcription factor 17
136	SOX2	SRY related HMG box
137	SSEA4	Stage specific embryonic antigen 4
138	TCA	Tricarboxylic Acid
139	T-CELLS	T-lymphocytes
140	TCGA	The Cancer Genome Atlas
141	TGF-beta	Transforming Growth Factor beta
142	TGF-β	Transforming Growth Factor Beta
143	TLRs	Toll like Receptors

144	TNF	Tumour Necrosis Factor
145	TNF-α	Tumour Necrosis Factor Alpha
146	TRA-I-60	Tumour Related Antigen 160
147	TRA-I-81	Tumour Related Antigen 181
148	TRE	Tet Response Element
149	UTR	Untranslated Region
150	VACTERL	Vertebral, Anal, Cardiac, Tracheoesophageal fistula, Oesophageal atresia, Renal, upper Limb and Hydrocephalus
151	VCF	Variant Call Format
152	VSAA	Very Severe Aplastic Anaemia
153	WT	Wildtype

UNDERSTANDING MOLECULAR MECHANISMS OF FANCONI ANAEMIA

SYNOPSIS SUBMITTED BY

Gaurav Joshi

Reg no: (2017/Ph.D.(CMC,V)/11)

TO

SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND TECHNOLOGY, TRIVANDRUM

SYNOPSIS

Background of the study:

Fanconi anemia (FA) is a rare inherited bone marrow failure disease caused by mutations in any of the 22 genes of the FA DNA repair pathway. The estimated incidence of FA is 1 per 360,000 live births. FA commonly presents as bone marrow failure leading to pancytopenia and congenital physical abnormalities. Around 20% of patients develop acute myeloid leukemia (AML), 30% develop solid tumors, and 40% develop myelodysplastic syndrome (MDS) by the age of 50. FA patients do not tolerate standard doses of chemotherapy drugs used in other hematological diseases or conditioning regimens for curative allogeneic hematopoietic cell transplantation (allo-HCT).

Accurate laboratory diagnosis of FA is essential as it can be challenging to differentiate idiopathic aplastic anemia (AA) from patients with FA without physical abnormalities. Chromosomal breakage analysis (CBA) is the most widely used diagnostic test for FA. Although increased G2/M cell-cycle arrest in FA cells treated with ICL agents can be used as a marker of FA, some rare hematological diseases can also cause chromosomal breakage, leading to false positives. However, defective FANCD2 ubiquitination (FANCD2-Ub) analysis is a robust test for diagnosing FA, which is not routinely used. Somatic mosaicism in 25% of FA patients can cause false-negative results in blood cells, so non-hematopoietic cells such as fibroblasts are preferred for performing these tests.

Identifying defective genes and pathogenic variants is important for carrier detection and prenatal diagnosis of FA in affected families and genotype-phenotype correlation in patients. Whole-exome sequencing (WES) studies have been conducted to determine the frequencies of defective genes and the spectrum of mutations in populations. Strong associations have been established between malignancies and biallelic pathogenic variants in FANCD1/BRCA2 and FANCN/PALB2 and monoallelic variants in FANCS/BRCA1, FANCJ/BRIP1, and FANCO/RAD51C.

The identification of new FA genes has enhanced our knowledge of the canonical FA pathway and revealed potential targets for intervention. However, the

impact of non-ICL repair functions of some FA genes on disease progression is becoming increasingly apparent. It is important to investigate alternative pathways that can modify the cellular and clinical phenotypes of FA. These modifiers may be influenced by environmental exposures and targeted by drugs, offering clinical benefits to FA patients. FA pathway can also affect disease severity. FA patients with biallelic mutations in FANCA, the most commonly affected gene in FA, typically have a more severe clinical phenotype compared to patients with mutations in other FA genes. However, within each complementation group, there can be significant variability in disease severity, suggesting the influence of other genetic modifiers. Genotyping a large number of patients from different populations, particularly those with high consanguinity rates, helps to better understand the genotype-phenotype correlation.

Studies conducted on FA patients' primary cells and FA mouse models suggest that the fate of hematopoietic progenitors is determined during fetal liver hematopoiesis. Unfortunately, it is technically and ethically impossible to trace earlier developmental events in humans. Additionally, the prevailing hypocellularity in the bone marrow presents a significant challenge to studying FA in primary patient cells. Transformed FA cell lines have been generated to study FA cellular phenotypes, but they do not precisely recapitulate the disease phenotypes, which could be due to transformation-related artifacts. Primary patient fibroblasts are not useful for studying hematopoietic stem and progenitor defects in FA leading to BMF. While mice models are commonly used to study FA pathogenesis, they provide a limited approach as they do not mimic the exact human hematopoietic phenotype. Induced pluripotent stem cells (iPSCs) are a type of stem cell that can be generated from adult cells, such as skin cells, by reprogramming them to an embryonic-like state using transcription factors OCT4, SOX2, KLF4, L-MYC (OSKM). Induced pluripotent stem cells (iPSCs) and human embryonic stem cells (hESCs) have the capability of indefinite in vitro expansion and can differentiate into almost all cell types, making them excellent tools for generating disease-relevant cell types to model FA. However, cellular reprogramming of FA somatic cells to iPSCs has extremely low efficiency due to elevated levels of DNA damage during reprogramming.

Despite the defective reprogramming in FA-pathway deficient cells, stable patient-specific iPSCs have been generated using alternative reprogramming approaches such as complementation of patient cells and in situ gene correction of patient cells before reprogramming. High-risk human papillomavirus E6 protein-mediated p53 degradation has also been used for the reprogramming of FA deficient patient cells. However, while FA somatic cells can be reprogrammed to iPSCs by stable inhibition of p53, these colonies could not be stabilized into cell lines, suggesting that p53-independent mechanisms may underlie the proliferation defect of FA iPSCs.

To date, *FANCA* and *FANCI*, to a certain extent, *FANCC* are the only three FA complementation groups studied using iPSCs. iPSCs generated by gene complementation or in situ gene correction cannot be used to study disease phenotype. Therefore, inducible complementation systems have been used to study the FA disease phenotype, where cells were reprogrammed in the presence of an active FA pathway. iPSCs can differentiate into all cell types in the body and can be used to model disease in a laboratory setting. Induced pluripotent stem cells (iPSCs) are a promising tool to model Fanconi anemia (FA), a rare genetic disorder characterized by bone marrow failure, cancer susceptibility, and developmental abnormalities. However, cellular reprogramming of FA somatic cells to iPSCs is challenging due to elevated levels of DNA damage. Alternative reprogramming approaches such as complementation of patient cells or in situ gene correction have been used to generate stable patient-specific iPSCs. FANCA and FANCC are the only two FA complementation groups studied using iPSCs to date. Inducible complementation systems have been used to study disease phenotype, where cells were reprogrammed in the presence of an active FA pathway using lentiviral vectors.

Objectives of the study:

1. Genotyping of Fanconi anaemia patients.
2. Genotype-phenotype association in patients with FA.
3. FA disease modelling using iPSCs to understand phenotypic heterogeneity.

Experimental methods used in the study:

1. FANCD2 Ub and chromosome breakage analysis (CBA) in peripheral blood T-cells and fibroblasts from FA patients.
2. Next-generation sequencing (NGS) of patients' DNA samples and bioinformatics analysis for identification of disease-causing variants in FA genes.
3. Multiplex ligation-dependent probe amplification and gene dosage PCR for confirmation of deletions identified by NGS.
4. Long-amplicon next-generation sequencing and bioinformatic analysis to detect mutations in *FANCA* and *FANCG* genes.
5. Bioinformatic analysis of NGS data to identify disease modifiers.
6. Generation of tightly regulated doxycycline (dox) inducible lentiviral vectors for complementation of patient fibroblasts.
7. Generation of iPSCs from FA patient fibroblasts by OSKM delivery using Sendai viral and episomal vectors.
8. Characterization of iPSC lines for their genomic integrity and pluripotency using the following methods:
    a. Karyotyping to confirm that no karyotypic abnormalities were acquired during reprogramming.
    b. Immunofluorescence for pluripotency markers OCT4, SSEA4, NANOG, and TRA-1-81.
    c. Flow analysis for TRA-1-60.
    d. qPCR analysis for pluripotency markers OCT4, NANOG, SOX2, GDF3 and DNMT3B.
    e. Trilineage differentiation of iPSCs into 3 germ layers viz ectoderm, endoderm and mesoderm and immunofluorescence for PAX6, SOX17, and BRACHYURY respectively.
9. Functional analysis of FA-iPSCs in the presence and absence of dox to study the effect of loss of FA pathway by various techniques such as alkaline phosphatase (AP) staining, cell cycle analysis, yH2AX-FANCD2 colocalization immunofluorescence, FANCD2 ubiquitination analysis.

10. Haematopoietic stem and progenitor cell (HSPC) differentiation and methylcellulose assay to determine colony formation potential in the presence and absence of FA pathway.

Major findings:

1. A comparison of CBA and FANCD2-Ub analysis performed in a large number of FA patients in this study confirmed that FANCD2-Ub analysis, which is currently not being used for diagnosis, is also suitable for FA diagnosis. We found increased sensitivity of CBA scores in FA diagnosis when a cut-off of 15 (arrived at using the ROC curve and Youden's index) was used. However, a randomized comparative analysis is required to confirm this cut-off.
2. Although NGS has many advantages, it has limitations in detecting CNVs. To address this challenge, robust bioinformatics methods are necessary for efficient detection of deletions. In our study, we used ExomeDepth for CNV analysis, and implemented filters to discard false positives, achieving 100% accuracy in identifying deletions in our patients. Our results suggest that the use of improved bioinformatics methods can enhance the efficiency and accuracy of CNV detection. *FANCA* exhibits high genetic heterogeneity and is the most frequently mutated gene in FA, with frequencies ranging from 60% to 80% across different populations. However, our study of FA patients in India revealed that *FANCA* (58.2%), *FANCL* (19.2%), and *FANCG* (13%) were the most commonly mutated genes in our cohort. Although we identified a large number of patients (~20%) with homozygous *FANCL* c.1092G>A;p.K364=, they presented with diverse phenotypes. The high consanguinity rate in our patient population led to over 87% of patients being homozygous for pathogenic variants in FA genes and 45% of patients being born from consanguineous marriages. By leveraging the presence of a *FANCL* founder variant and the high frequency of pathogenic variants in *FANCA* and *FANCG*, we established a new, faster, and cost-effective molecular diagnosis strategy for Indian FA patients, which can diagnose up to 90% of FA patients. Overall, this algorithm could expedite the diagnosis of FA and serve as a cost-effective alternative to whole exome sequencing.

3. We performed genotype-phenotype correlation analysis for the largest groups in our cohort, namely patients with *FANCA*, *FANCL*, and *FANCG* mutations. Patients with *FANCA* and *FANCG* complementation groups showed significantly higher chromosome breakage analysis (CBA) scores (p=0.002) compared to those with *FANCL* mutations. Moreover, patients with *FANCG* mutations had a significantly higher prevalence of kidney anomalies than those with *FANCL* and *FANCA* mutations (p=0.01). However, we did not find any other significant associations between genotypes and observed phenotypes in FA patients, except for a trend (p=0.07) towards a higher occurrence of myelodysplastic syndrome/acute myeloid leukemia (MDS/AML) in patients with *FANCA* and *FANCG* defects compared to *FANCL*.
4. As expected, we found that haematological malignancies were more frequently observed in patients with mutations in exons 1-19 (p=0.0006). Interestingly, we observed an extremely high occurrence of MDS/AML in patients with mutations in exons 27-30 (p=0.0006). We further investigated whether a specific type of mutation was linked to haematological malignancy. We found that missense mutations had a significantly higher propensity (p=0.03) to lead to haematological malignancy than splicing, deletion, and frameshift mutations. Our findings suggest that genotype-phenotype correlations could provide valuable insights into the clinical features of FA patients and guide personalized treatment approaches.
5. We performed an in-depth analysis of NGS data from FA patients to identify disease modifiers. We identified variants in genes involved in crucial pathways such as regulation of hematopoiesis, ribosome biogenesis, DNA damage response, immune response, inflammatory pathways, tumor suppressor genes, and telomere biology. Disruptions of these pathways have significant implications for the development and progression of FA.
6. While previous studies have successfully generated iPSCs with defects in *FANCA* through inducible complementation, patient-specific iPSCs with mutations in other FA complementation groups have yet to be investigated. To address this, we obtained fibroblasts from patients with mutations in *FANCA,*

FANCC, FANCF, FANCL, FANCI, and *FANCT* genes. To achieve inducible complementation, we transduced the cells with doxycycline (dox)-inducible lentiviral vectors that expressed wild-type cDNAs for these genes. The vectors included both the FA gene and mCherry, linked by a P2A self-cleaving peptide, which enabled us to monitor the dox-inducible expression of FA genes.

7. In our attempt to induce doxycycline (dox)-inducible FANCA transgene expression, we initially used the pInducer20-FANCA vector system. However, we observed high levels of background expression in fibroblasts even in the absence of dox, as evidenced by the presence of FANCD2 monoubiquitination. This leakiness could be attributed to the vector's transgenes being cloned in 5' to 3' LTR, where the 5' LTR acts as a strong promoter compared to the 3' LTR, resulting in transgene expression even in the absence of dox. Therefore, the pInducer20 vector system was deemed unsuitable for studying cellular phenotypes of FA. To address this issue, we switched to the pSJL225 vector system that utilizes a 3rd generation rtTA (Tet-On 3G) and TRE3G promoter for expression of FA transgenes including *FANCA, FANCC, FANCF, FANCI, FANCL,* and *FANCT*. This system incorporates an optimized rtTA variant called Tet-On 3G, which exhibits significantly reduced background expression in the absence of the inducer while displaying enhanced expression in its presence, resulting in tighter control of gene expression. The lack of mCherry expression in the absence of dox induction indicates that the vector's regulatory elements are effectively preventing any basal or background expression of mCherry, thus demonstrating tight control over gene expression.
8. Clones were selected based on their mCherry expression upon dox induction and its subsequent loss and cell death upon dox withdrawal. Pluripotency was confirmed by the expression of markers OCT4, NANOG, SSEA4, and TRA-1-81 in the undifferentiated state. The cells could differentiate into ectoderm, endoderm, and mesoderm, and expression of PAX6, SOX17, and BRACHYURY was analyzed.
9. We used a vector that co-expressed mCherry with the FA transgene to monitor the effects of the FA pathway on cell survival. The iPSCs were cultured without

doxycycline until mCherry expression was completely lost. *FANCF* and *FANCI* complemented clones were cultured without doxycycline for 10-15 days. We found that dox withdrawal resulting in the absence of an active FA pathway did not affect the pluripotency of any of the complemented iPSCs as observed by TRA-1-60 staining before and after dox withdrawal. Interestingly, *FANCF* and *FANCI* iPSCs did not show cell death after dox withdrawal. However, loss of mCherry expression and progressive cell death was observed upon dox withdrawal in *FANCA, FANCC, FANCL*, and *FANCT* complemented clones, as determined by flow analysis and AP staining. This suggests that the observed cell death upon dox withdrawal in these clones was not due to the loss of pluripotency.

10. Treatment with hydroxyurea (HU) induced monoubiquitination of FANCD2 in iPSCs generated for *FANCA, FANCC, FANCI, FANCL and FANCT*, and the detection of monoubiquitinated FANCD2 was only observed in cells treated with HU, indicating that the activation of the FA pathway requires the presence of a DNA damaging agent. However, cell death after dox withdrawal occurred without HU treatment, indicating that endogenous metabolites that induce interstrand cross-links (ICLs) may not be cleared in the absence of the FA pathway. This suggests the existence of additional intricate mechanisms that govern the FA pathway and its potential association with cell death, which may have important implications for understanding the bone marrow failure observed in FA patients. Furthermore, it is important to note that the cell death observed due to G2/M arrest and the lack of FANCD2 ubiquitination occurs through different mechanisms in cells without an active FA pathway.
11. The iPSC clones derived from *FANCF* and *FANCI* defective fibroblasts did not undergo cell death after dox withdrawal. FANCD2 monoubiquitination was observed in *FANCF* iPSCs even after dox withdrawal. Interestingly, low levels of *FANCF* expression were observed in these iPSC clones, which was similar to those in control cells, and may have contributed to the lack of cell death upon dox withdrawal. We generated several FANCI iPSC clones, and two out of five clones did not show FANCD2Ub even after dox induction. Interestingly,

these cells died after blasticidin selection, suggesting that lentiviral integration did not occur in these cells. This suggests that *FANCI* iPSCs can be generated without complementation. Furthermore, *FANCI* iPSC clones did not show cell death upon dox withdrawal, but there was a loss of FANCD2Ub and FANCI expression after dox withdrawal, as observed by Western blot analysis. These findings indicate that *FANCI* expression is not necessary for reprogramming or iPSC survival, and FANCD2Ub is not linked to the survival of FA-iPSCs.

12. FA-iPSCs generated by inducible complementation of cells with defective FA pathway exhibit G2/M arrest. Cell cycle analysis was performed on *FANCA, FANCC, FANCF, FANCI, FANCL, and FANCT* clones using EdU and DAPI to monitor the cell cycle of dox- cultures. It was observed that these clones exhibited cell-cycle arrest in the G2/M phase compared with dox+ cultures. G2/M arrest was significantly higher in dox-cells even in the absence of mitomycin C (MMC), suggesting intracellular DNA-damaging agents produced in the absence of the FA pathway. *FANCA* defective cells showed the highest G2/M arrest (48.9%) among all the complementation groups without MMC treatment. After treatment with 1ng/ml MMC, G2/M arrest increased in all defective cells, with FANCA showing the highest increase (54.2%). *FANCF* clones did not display significantly elevated levels of G2/M arrest after dox withdrawal, possibly due to the basal level of *FANCF* expression observed in these clones. Similarly, *FANCI* clones that did not die after dox withdrawal did not show increased G2/M arrest after dox withdrawal even after MMC treatment. These observations suggest that the FA pathway is essential for maintaining genomic integrity and preventing DNA damage-induced G2/M arrest in FA-iPSCs. Abnormal results obtained in *FANCF* and *FANCI* iPSCs need further validation.
13. Haematopoietic lineage differentiation of patient-derived iPSCs from the 6 complementation groups was carried out with and without dox induction using a modified protocol. In dox- conditions, mesoderm colonies were smaller and contained fewer cells compared to dox+ and control cells. Dox- cells showed progressive cell death and senescence, while dox+ cells showed good

proliferation. Hematopoietic cells were identified using CD34, CD45, CD43, and CD235a markers, and the number of HSPCs generated from FA-iPSC clones in dox+ conditions was higher than in dox- conditions. To examine the hallmarks of HSPC dysfunction in FA, we differentiated iPSCs with or without doxycycline to maintain or deactivate the inducible FA transgene, respectively, during assays of haematopoietic function. Colony formation assays (CFU) were performed on day 13 of differentiation. FA pathway deficient HSPCs (dox-) exhibited significantly reduced clonogenicity compared to FA proficient HSPCs (dox+ cells), indicating an impaired hematopoietic function in FA.

Significance and implications of the study:

The NGS analysis conducted to identify disease-causing variants in the Indian population is one of the largest of its kind, with an extremely high detection rate for disease-causing variants. It was found that the most prevalent genotypes in the Indian population are *FANCA*, *FANCL,* and *FANCG*. These genotypes should be prioritized for disease-causing variant detection in FA. Interestingly, *FANCL* mutation was present at an extremely high frequency in the Indian patients indicating the presence of the founder effect as previously reported. However, an unbiased approach was used for the recruitment of FA patients. An algorithm for robust and cost-effective diagnosis is also proposed in this study which could detect disease-causing variants in nearly ~90% of the patients.

The genotype-phenotype association study provided insights into the impact of FA genotypes on the phenotypic characteristics of the patients and demonstrated a significant association between certain genotypes and malignant transformation. Moreover, the study identified disease modifiers that could potentially explain the phenotypic heterogeneity observed in FA patients.

This is one of the largest studies that generated FA-iPSCs from 6 complementation groups. A tightly regulated dox inducible vector was generated that could efficiently regulate transgene expression. The phenotypic heterogeneity between different complementation groups was explained by disease modelling using patient-specific FA-iPSCs. Defects in most of the complementation groups exhibited bone marrow failure-like phenotype. These iPSCs exhibited all the cellular phenotypes of FA and

could be used as a perfect model to identify potential druggable targets which could improve treatment outcomes.

1. INTRODUCTION

1.1. Background

Fanconi Anaemia (FA) is an autosomal recessive genetic disorder characterized by chromosomal instability and defective DNA repair mechanisms. It was first identified in 1927 by Dr. Guido Fanconi and has since been the subject of extensive scientific research. This rare condition primarily affects the haematopoietic system, resulting in bone marrow failure, congenital anomalies, and an elevated risk of developing haematologic malignancies and solid tumours (Niraj et al., 2019).

The genetic basis of FA involves mutations in any of the 22 currently known FA genes, which encode proteins that participate in the FA pathway. This pathway is vital for the repair of DNA damage, particularly during processes such as DNA interstrand crosslink repair and homologous recombination (Prakash et al., 2015; Wang et al., 2015). Dysfunction of this pathway impairs the cell's ability to repair DNA lesions, leading to genomic instability and increased susceptibility to mutagenesis. The clinical presentation of FA varies widely, with affected individuals exhibiting distinct phenotypic features. Common physical characteristics may include short stature, skeletal abnormalities, microcephaly, café-au-lait spots, and radial ray defects. These clinical manifestations can aid in the initial diagnosis of the disease (Altintas et al., 2023).

Diagnosing FA requires multiple tests, such as chromosomal breakage analysis by treating the patients' blood cells with diepoxybutane (DEB) or mitomycin C (MMC) and genetic testing to identify pathogenic mutations in FA pathway genes

(George, Solanki, Chavan, Rajendran, Raj, Mohan, Nemani, Kanvinde, Munirathnam, Rao, Radhakrishnan, Lashkari, et al., 2021; Joshi et al., 2023). Prenatal testing and carrier screening are also available for families with a history of FA. The study of FA has significantly contributed to our understanding of DNA repair pathways and cancer genetics. Insights gained from FA research have not only advanced our knowledge of the molecular mechanisms involved in DNA repair but have also shed light on the pathogenesis of various malignancies, leading to potential therapeutic targets for cancer treatment (Nalepa and Clapp, 2018).

Genotype-phenotype correlation in FA is a subject of extensive investigation to comprehend the intricate relationship between specific gene mutations and the resulting clinical manifestations. While some general trends have been identified, the correlation between genotype and phenotype in FA is far from straightforward. Certain mutations in FA genes, such as *FANCA* and *FANCC*, have been associated with more severe clinical presentations, characterized by early-onset bone marrow failure and a higher predisposition to cancer (Alter, Giri, Savage and Philip S Rosenberg, 2018). On the other hand, mutations in genes like *FANCD2* have been linked to milder phenotypes and a lower risk of developing cancer (Nalepa and Clapp, 2018). However, numerous exceptions exist, and additional factors, such as the type and location of mutations, compound heterozygosity, modifier genes, and environmental influences, further contribute to the phenotypic variability observed in FA patients (Cheung and Taniguchi, 2017). Continued research and advancements in genomics are crucial to elucidating these complexities and providing more accurate genotype-phenotype correlations for improved patient care and management. Recent studies have shed light

on the impact of specific mutations within FA genes on the clinical spectrum of FA, further highlighting the intricacies of the genotype-phenotype correlation. For instance, in a study by (Faivre et al., 2000) the authors found that patients with certain types of *FANCA* mutations were more likely to develop severe bone marrow failure at an early age, while patients with other mutations exhibited a milder form of the disease (Maria Castella, Pujol, Callén, Trujillo, Casado, Gille, Lach, Auerbach, Schindler, Benítez, Porto, Ferro, Muñoz, et al., 2011). Similarly, in a comprehensive analysis of *FANCC* mutations, identified specific mutations associated with distinct clinical outcomes, including variations in haematological parameters and the development of haematopoietic malignancies (Faivre et al., 2000).

FA leads to a wide range of clinical manifestations, including bone marrow failure, increased risk of cancer, and developmental abnormalities. Several studies have shown that the TGF-beta pathway, a crucial cellular signalling pathway, interacts with the FA pathway (Kim et al., 2015; L Liu et al., 2014; Pal et al., 2017). Dysregulation of TGF-beta signalling can influence FA phenotypes and potentially serve as a disease modifier in FA. Aldehydes, highly reactive molecules generated during cellular metabolism and exposure to environmental toxins, can cause DNA damage and interfere with the normal functioning of the FA pathway, leading to increased genomic instability (Voulgaridou et al., 2011). Aldehydes have been implicated in the pathogenesis of FA and may contribute to the clinical features observed in FA patients, such as bone marrow failure and cancer predisposition (Van Wassenhove et al., 2016).

Moreover, ribosome defects, referring to abnormalities in ribosome biogenesis and function, can impact protein synthesis and cellular homeostasis. Recent research has identified connections between ribosome defects and the FA pathway, suggesting that these defects may contribute to the pathophysiology of FA and potentially serve as disease modifiers (Gueiderikh et al., 2022; Samuel B Sondalle et al., 2019).

While the insights gained from murine models of FA are valuable, their ability to fully recapitulate the human disease phenotype solely through FANC gene mutations is a significant challenge (Parmar et al., 2009). It is possible that mice are more susceptible to DNA damage and have alternative regulatory mechanisms for FANC proteins compared to humans, further limiting the usefulness of murine FA models for understanding FA pathophysiology and developing novel treatments (Bakker, De Winter, et al., 2013). Additionally, the diverse nature of mutations in various types of FA, including point mutations, small insertions/deletions, splicing mutations, and large intragenic deletions, poses a challenge in accurately reproducing all human mutations in mice through targeted gene knock-ins/outs. Therefore, it remains difficult to precisely replicate the full spectrum of human mutations in mouse models.

Human embryonic stem cells (hESCs) and more recently, induced pluripotent stem cells (iPSCs), exhibit a remarkable capacity for unlimited in vitro expansion while maintaining their potential to differentiate into various cell types representing the diverse tissues found in the developing embryo. Generating induced pluripotent stem cells (iPSCs) for modelling FA has been challenging due to the need for an intact FA DNA repair pathway during reprogramming and pluripotency maintenance. To

overcome this hurdle, researchers introduced wild-type cDNA of the defective FA pathway gene into patients' somatic cells before reprogramming (Müller et al., 2012). However, permanent repair of FA pathway mutations in iPSCs can hinder the derivation of FA-deficient haematopoietic stem and progenitor cells (HSPCs) for disease modelling (Raya et al., 2009). Despite this limitation, alternative reprogramming approaches, such as complementation and CRISPR-Cas9-mediated gene correction, have been used successfully. In some cases, the FA fibroblasts were reprogrammed using doxycycline-inducible lentiviral vectors, activating the FA pathway (Bharathan et al., 2017; Müller et al., 2012). Patient-specific iPSCs have been generated for various FA complementation groups, but it is crucial to generate FA-iPSCs for other groups to understand the role of each protein in the FA pathway, considering their noncanonical functions that influence the FA phenotype in patients.

1.2 Rationale and Hypothesis

FA is a genetically diverse disorder, and about 40% of FA cases show no physical abnormalities, making diagnosis challenging. Additionally, not all characteristic abnormalities are present in every FA patient, contributing to the variable clinical presentation. The extent of bone marrow failure (BMF) in FA varies from mild to severe cytopenias. Some cases may require haematopoietic stem cell transplantation (HCT) to address the haematological deficiencies, while others may progress to myelodysplastic syndrome (MDS) or acute myeloid leukaemia (AML), which are known complications of FA.

Cohort studies investigating the relationship between genotypes and phenotypes in FA have yielded contradictory findings due to differences in populations and the genetic heterogeneity of FA. The complexity of genetic mutations associated with FA makes it challenging to establish clear genotype-phenotype associations. Transformation to haematological malignancy is well-recognized in FA, adding to the disease's clinical complexity and risk profile.

Despite the existence of only one study on the Indian population (George, Solanki, Chavan, Rajendran, Raj, Mohan, Nemani, Kanvinde, Munirathnam, Rao, Radhakrishnan, Lashkari, et al., 2021), a clear association of the most prevalent genotypes (*FANCA*, *FANCL*, and *FANCG*) with specific phenotypic manifestations in FA patients has not been thoroughly explained. More research is needed to understand the underlying mechanisms and genetic factors that contribute to the variable clinical presentations and outcomes of FA in different populations.

Despite extensive research, the precise mechanisms underlying the clinical manifestations of FA remain poorly defined. Notably, different clinical phenotypes have been observed for specific gene mutations, further adding to the complexity of the disease. Approximately 90% of FA patients present with aplastic anaemia (AA), which can be classified as non-severe aplastic anaemia (NSAA) in 20-30% of cases or severe aplastic anaemia (SAA) or very severe aplastic anaemia (VSAA) in 40-50% of cases. Moreover, a subset of patients, around 10-15%, develop leukaemia, while approximately 28% present with solid tumours. Additionally, patients exhibit a wide range of physical abnormalities, contributing to the morphological diversity seen in

FA. Strikingly, even among patients with similar genetic mutations, variable phenotypes are observed, adding to the complexity of the disease.

The pathogenesis and anomalies in FA may arise from various factors, including non-canonical functions of FA genes or the influence of non-FA genes. These factors could interact and contribute to the broad spectrum of clinical presentations in FA. To comprehend the significant variability in clinical phenotypes, it becomes essential to explore changes in different pathways that impact these manifestations. Understanding the intricate interplay between these pathways will shed light on the underlying mechanisms and the reasons behind the heterogeneous clinical course of FA. Further research in this direction is crucial for advancing our knowledge of the disease and developing more targeted and effective treatment strategies.

Generation of induced pluripotent stem cells (iPSCs) from FA patient cells has been challenging due to its extreme inefficiency, likely caused by elevated levels of DNA damage that trigger apoptosis. Inhibition of p53 has allowed for the derivation of iPSC colonies from FA patient cells, but unfortunately, these colonies fail to grow into stable cell lines, as p53-independent signalling limits the self-renewal of FA iPSCs. Alternative reprogramming approaches have been used to successfully derive FA patient iPSCs, but they encountered difficulties in differentiating these iPSCs into haematopoietic stem and progenitor cells (HSPCs).

An *in-situ* gene correction method has shown promise in reprogramming FA cells into iPSCs, but the corrected cells cannot be effectively used as a model to study the diseased phenotype of FA. As an alternative, a conditional approach was employed

in the laboratory, where iPSCs could be derived from FA-proficient cells in the presence of doxycycline. By using this approach, researchers have studied isogenic FA-proficient and deficient cell lines after reprogramming, allowing them to better investigate the cellular and molecular differences between these cells and gain insights into the pathophysiology of FA.

1.2. Objectives of the study

1. *Genotyping of FA patients.*
2. *Genotype-phenotype association in patients with FA.*
3. *FA disease modelling using iPSCs to understand phenotypic heterogeneity.*

1.3. Brief overview of the thesis chapters

1.3.1. Literature review

Chapter 2 is a literature review which elaborates on FA disease, the FA pathway, the genotypes identified, and studies conducted for genotype-phenotype correlation and genetic modifiers of the disease. This chapter also focussed on the models to study FA phenotype and the use of Induced Pluripotent Stem Cells (iPSCs) in FA disease modelling. In this chapter advancements in the development of lentiviral vectors and inducible lentiviral vectors have also been discussed.

1.3.2. Materials and Methods

This section encompasses the various diagnostic techniques employed for the assessment of FA patients. These techniques include exome sequencing, the utilization of bioinformatics pipelines for the identification of disease-causing variants, chromosome breakage analysis (CBA), FANCD2-Ub analysis, Multiplex Ligation-dependent Probe Amplification (MLPA) for *FANCA*, lentiviral complementation of FA transgenes utilising both constitute and inducible promoters, reprogramming through episomal and Sendai virus methods, characterization of FA-induced pluripotent stem cells (FA-iPSCs), functional assays for studying FA defects in iPSCs, and methods for evaluating bone marrow failure (BMF) phenotypes, such as Alkaline phosphatase (AP) staining after doxycycline withdrawal, and the differentiation of Haematopoietic Stem and Progenitor Cells (HSPCs).

1.3.3. Results

The results are divided into 3 sections. The first section includes genotyping of FA patients using exome sequencing to reveal the most prominent genotypes in the Indian population. This section also compares the widely used method CBA with FANCD2-Ub analysis. We also propose an algorithm which is robust and cost-effective for FA diagnosis. In the second section, we present genotype-phenotype correlation based on defects in the 22 FA genes. Genotype-phenotype was further explored by the identification of other defective genes which may have a role in FA pathophysiology. The third section includes the modelling of FA using patient-specific iPSCs to understand the phenotypic heterogeneity at the cellular level. Through the

comparison of multiple complementation groups, it becomes possible to assess and compare their respective defects.

1.3.4. Discussion

This section provides a comprehensive explanation of the major findings obtained from the study. The results are thoroughly compared with existing scientific data, highlighting their significance in the context of FA research.

1.3.5. Summary and Conclusion

This chapter presents a concise summary of the key observations derived from the doctoral research conducted. The significant findings and contributions made during the course of this study are highlighted, providing a comprehensive overview of the research outcomes. Moreover, the chapter also outlines the ongoing work and future directions, emphasizing potential avenues for further investigation and exploration in the field of study.

2. REVIEW OF LITERATURE

2.1. Background

Fanconi anaemia (FA) is a rare inherited bone marrow failure syndrome (IBMFS) that presents with pancytopenia, increased susceptibility to malignancies, and distinct physical abnormalities/congenital malformations (Neveling et al., 2009). FA is caused by pathogenic variants in one of the 22 genes in the Fanconi Anaemia (FA)-DNA repair pathway. The diagnosis of Fanconi anaemia is typically established during childhood, although in some cases, individuals may not be diagnosed until adulthood due to the variability of disease manifestations and delays in diagnosis (Alter et al., 2005). Increased surveillance for organ dysfunction and the use of reduced doses of chemotherapy and/or radiation are typically recommended for individuals with Fanconi anaemia (FA) before chemotherapy for leukaemia with underlying FA or haematopoietic cell transplantation (Dietz et al., 2017). To avoid transplantation with an undetected and asymptomatic sibling with FA, FA must be ruled out and excluded from the sibling donors for haematopoietic cell transplantation (HCT).

2.2. Epidemiology

Despite being a rare condition, FA is considered the most common among IBMFSs. The estimated incidence of FA is approximately 1 in 300,000 live births, while the prevalence ranges from 1 to 9 cases per million individuals (Bagby, 2016). However, it's important to note that the carrier frequency of FA can vary depending

on the specific population under study (Bagby, 2016). FA has been documented in diverse racial and ethnic groups, encompassing a wide range of populations worldwide.

While the majority of individuals are typically diagnosed with FA between the ages of 6 and 9 when bone marrow failure becomes apparent, there is a growing trend of earlier diagnoses due to increased awareness of the variable clinical presentations and improved recognition of the disease. A significant proportion, up to 9%, of individuals with FA are diagnosed after the age of 16, usually when they present with malignancies (Risitano et al., 2016). Certain populations exhibit a higher incidence of FA due to the presence of specific founder mutations. For instance, Spanish Romani populations have a higher prevalence of the *FANCA* c.295C>T mutation, while Ashkenazi Jews commonly carry the *FANCC* IVS4 + 4A>T mutation. Patients of North African and Middle Eastern descent often exhibit exon 15 deletions, whereas the South African Afrikaner population is characterized by a large intragenic deletion. Additionally, the *FANCL* c.1092G>A mutation is frequently observed in the Indian population, among other variations (Amouri, Talmoudi, Messaoud, Catherine D. d'Enghien, et al., 2014; Bagby, 2016; Callen, 2005; George, Solanki, Chavan, Rajendran, Raj, Mohan, Nemani, Kanvinde, Munirathnam, Rao, Radhakrishnan, Lashkari, et al., 2021; Joshi et al., 2023; Kutler and Auerbach, 2004; Magdalena et al., 2005; Risitano et al., 2016; Rosenberg et al., 2011; Strom et al., 2004; Tamary et al., 2000; Tipping et al., 2001; Yagasaki et al., 2003). The carrier frequency of specific founder mutations in FA varies across different populations. In the general population, the carrier frequency is approximately 1 in 190 individuals. However, in the South

African Afrikaans population, the carrier frequency is higher at 1 in 100. Similarly, among Ashkenazi Jews, the carrier frequency is also 1 in 100. The carrier frequency is even more elevated in Spanish Romani populations, reaching up to 1 in 64 individuals (Callen, 2005; Rosenberg et al., 2011; Strom et al., 2004; Tipping et al., 2001). Founder mutations have also been identified in individuals from diverse regions, including Tunisia, Japan, Korea, and Brazil (Amouri, Talmoudi, Messaoud, Catherine D. d'Enghien, et al., 2014; Magdalena et al., 2005; Yagasaki et al., 2003).

2.3. Pathophysiology

FA is a hereditary disorder characterized by impaired DNA damage repair mechanisms within cells. This deficiency results in genomic instability, disrupted regulation of the cell cycle, and ultimately, cell death. FA is attributed to an inherent inability to effectively repair DNA interstrand crosslinks (ICLs), which are particularly harmful forms of DNA damage where the two strands of DNA are abnormally fused. This impairment in repairing ICLs contributes significantly to the pathogenesis of FA (Niraj et al., 2019). The disrupted repair of interstrand crosslinks (ICLs) results in genomic instability, irregular regulation of the cell cycle, and eventual cell death. DNA damage can arise during fetal development, leading to congenital anomalies. Additionally, throughout childhood and adulthood, the presence of DNA damage is linked to bone marrow failure (BMF), organ damage, and an increased risk of developing cancer (Nicholas E. Mamrak et al., 2017; Nalepa and Clapp, 2014; Niraj et al., 2019; Tercanli et al., 2001). FA patients display a haematopoietic stem and progenitor cell (HSPC) defect prior to the onset of clinical bone marrow failure (BMF),

indicating an early manifestation of the disease. Furthermore, activation of the p53/p21 axis has been observed in human foetal liver samples during the phase of extensive expansion of the haematopoietic stem cell (HSC) pool (Ceccaldi, Parmar, Mouly, Delord, Jung Min Kim, et al., 2012).

2.4. The FA pathway

2.4.1. Canonical role of FA proteins

ICLs are toxic lesions that hamper the DNA interstrand separation upon replication and when left unresolved they produce a clastogenic effect on cells leading to genomic instability (Deans and West, 2011). At the site of interruption, the replication fork triggers a plethora of events that initiate the recruitment of FA proteins. The canonical FA pathway is a tightly regulated mechanism. FA gene-encoded proteins coordinate with nucleotide excision repair (NER) and homologous recombination (HR) in the multistep pathway to repair the ICLs (Niraj et al., 2019). ICL repair is undertaken by a highly sophisticated and coordinated network of proteins of BER, NER, and HR mechanisms that prevents the activation of NHEJ machinery (Renaudin and Rosselli, 2020). The categorization of FA genes is primarily based on their association with the monoubiquitination process involving two FA proteins, namely FANCD2 and FANCI, which form a heterodimer known as the FANCI/D2 complex. This categorization is established based on a combination of molecular and clinical factors. Nine FA genes encode proteins that comprise the multi-subunit ubiquitin ligase FA core complex that is responsible for the monoubiquitination

reaction (Gueiderikh et al., 2022). Initialization of the FA pathway is implicated by the localization of the cargo protein FANCM along with its partners at the site of ICL. The FA core complex is a multicomponent complex that contains nine of the known FA proteins: FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, FANCT and FANCM. The FA core complex facilitates the assembly of a ubiquitin E3-ligase machinery that transiently interacts with the E2 enzyme FANCT. This interaction leads to the monoubiquitination of the FANCD2/FANCI (ID2) complex. Subsequently, an endonuclease complex (FANCP and FANCQ) is recruited, which removes abnormal crosslinks and induces a DNA double-strand break (DSB). The repair of this DSB is then mediated by FANCD1 (also known as *BRCA2*, a gene associated with breast cancer susceptibility), FANCN, FANCJ, and FANCO. These proteins operate through the DNA repair pathway, employing nucleotide excision repair and homologous recombination mechanisms during the S or G2 phases of the cell cycle, thereby restoring DNA integrity. Finally, the deubiquitinase USP1 deubiquitinates the ID2 complex **(Figure 2.1)**. Because the majority of FA patients carry mutations in core complex proteins, FANCI/D2 monoubiquitination is impaired in greater than 90% of FA patients (Bogliolo et al., 2020; George, Solanki, Chavan, Rajendran, Raj, Mohan, Nemani, Kanvinde, Munirathnam, Rao, Radhakrishnan, Lashkari, et al., 2021; Joshi et al., 2023). While FANCD2 monoubiquitination plays a crucial role in interstrand crosslink (ICL) repair, it alone does not serve as a reliable indicator of an efficient Fanconi Anaemia (FA) pathway. Higher levels of monoubiquitinated FANCD2 within a cell do not necessarily enhance ICL repair and can even hinder the process in certain cases. This concept is supported by studies that

have demonstrated the necessity of USP1-mediated FANCD2 de-ubiquitination for the formation of FANCD2 foci and the development of resistance to ICL-inducing agents (Liang et al., 2014; Oestergaard et al., 2007). Furthermore, Usp1-deficient mice have an FA-like phenotype (Kim et al., 2009; Parmar et al., 2010). Why de-ubiquitination of FANCD2 is important for ICL repair is unclear, although it might prevent the accumulation of this protein at genomic sites where it cannot function (Oestergaard et al., 2007). Furthermore, studies suggest that disrupting the proper equilibrium between monoubiquitinated and non-ubiquitinated FANCD2 by impairing USP1 can lead to aberrantly low phosphorylation of FANCI. This imbalance in phosphorylation has the potential to impede the function of the monoubiquitinated FANCI/D2 complex. (Cheung et al., 2017).

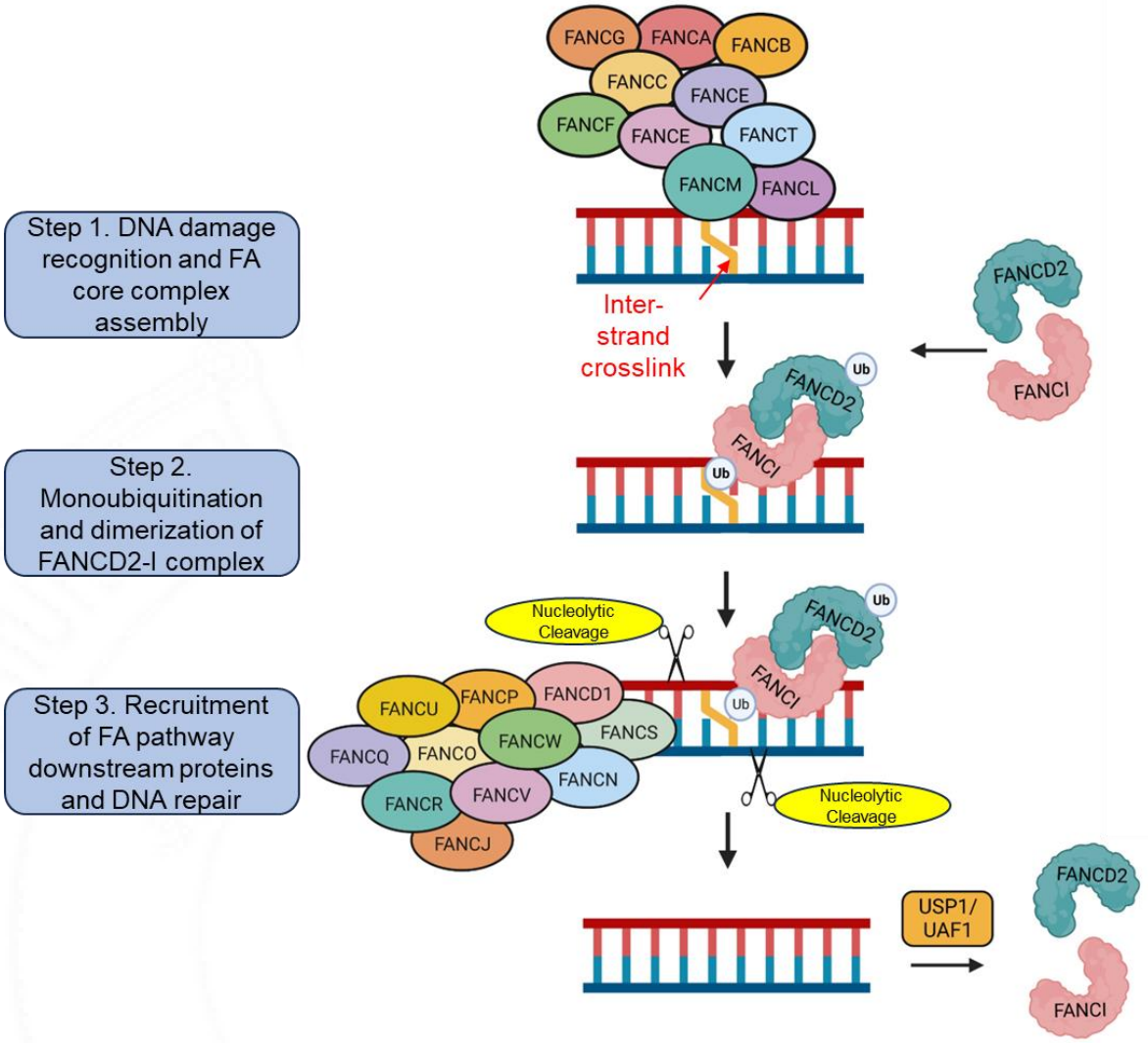


Figure 2.1: Fanconi Anaemia DNA repair pathway. 1. Interstrand crosslinks recognition and FA core complex recruitment. 2. Dimerization and monoubiquitination of FANCD2 and FANCI proteins. 3. Monoubiquitinated FANCD2 recruits downstream repair proteins that include endonucleases (FANCP and FANCQ) to remove ICL and other proteins for error-free homologous recombination (HR) mediated repair.

2.4.2. Non-canonical role of FA proteins

The current understanding highlights that although all FA proteins contribute to genome stability, several of them possess additional functions, some of which operate independently of other FA proteins or extend beyond the nuclear envelope. (Pang, 2001; Pang, Christianson, et al., 2001; Pang et al., 2000; Pang, Keeble, et al., 2001) and some are functionally independent of the DNA damage response (Garbati et al., 2016; Sumpter et al., 2016; Sumpter and Levine, 2017). Furthermore, important endogenous factors that suppress FA HSC function are emerging and include proteins normally involved in the inflammatory response (Fagerlie et al., 2001; Pang, 2001; Pang et al., 2002; Pang, Keeble, et al., 2001; Walter et al., 2015; H Zhang et al., 2016), pathways functionally linked to or activated by inflammation, including oxidative stress (Bijangi-Vishehsaraei et al., 2005; Cumming et al., 2001; Futaki, 2002; Gille et al., 1987; Kontou et al., 2003; Li and Pang, 2014; Pagano et al., 2005; Ruppitsch et al., 1998; Saadatzadeh et al., 2004), mitophagy (Sumpter, Sirasanagandla, Álvaro F. Fernández, et al., 2016), and the production of endogenous aldehydes (Hira et al., 2013; Langevin et al., 2011; Pontel et al., 2015; Rosado et al., 2011; Tacconi et al., 2017). **Table 2.1** enumerates some of the non-canonical functions associated with FA proteins.

Considering the diverse functions mentioned earlier, it is not unexpected that FA mutations in haematopoietic stem cells (HSCs) would lead to several cellular consequences. These include the loss of quiescence (Ho et al., 2017; Lier et al., 2014; Walter et al., 2015), heightened apoptosis upon exposure to inflammatory cytokines or redox stress (Beere, 2001; Chen et al., 2010; Cumming et al., 2001), and ultimately,

a progressive decline in stem cell population under inflammatory stress conditions **(Figure 2.2).**

Table 2.1. Non-canonical functions of FA genes

Fanconi anaemia proteins	Biochemical function	Expected effects
FANCD2 and FANCA	In response to oxidative stress, FANCD2 and FANCA form a complex with BRG1 within promoters of antioxidant genes.	Enhance antioxidant defenses
FANCD2	Binds to FOXO3A and reduces reactive oxygen species (ROS) accumulation and enhances antioxidant gene expression.	Reduces the accumulation of ROS and enhances cellular resistance to oxidative stress
FANCG	Binds to and stabilizes mitochondrial PRDX3. Loss of FANCA or FANCC also destabilizes PRDX3.	Enhances resistance to resistance to H_2O_2 and mitomycin C
FANCC	Binds GSTP1 and activates its activity in response to apoptotic stimuli.	Prevents apoptosis in growth factor-deprived haematopoietic cells
FANCN	Binds to KEAP1	Enhances the oxidative stress response
FANCC, -A, -F, -L, -D1, -D2, and -S	Clear damaged mitochondria (mitophagy). FANCA and FANCC interact with Parkin and translocate to damaged mitochondria. Knockdown of FANCC, -F, or -L leads to defective selective autophagy.	Decrease mitochondrial ROS production Reduce activation of inflammasomes
FANCD2	Localizes to mitochondria and interacts with mitochondrial protein (Atad3).	Stabilizes mitochondria, enhances cisplatin resistance, and suppresses apoptosis
FANCP	Using different domains, FANCP interacts with XPF and ERCC1 to repair interstrand crosslinks and interacts with MUS81 to resist TOP1 inhibitors.	Mediates resistance to both cross-linking agents and topoisomerase I inhibitors
FANCD2 and FANCA	Suppress transforming growth factor beta signaling in haematopoietic stem and progenitor cells (HSPCs) in the ground state and during inflammatory stress and enhance expression of genes involved in homologous recombination.	Enhance HSPC survival and function in the face of inflammation and cross-linking agents
FANCC	Binds hsp70 and suppresses the kinase activity of PKR independently of the core complex.	Enhances survival of cells exposed to inflammatory cytokines
FANCC	Binds to and facilitates activation of STAT1 in response to haematopoietic growth factors.	Facilitates haematopoietic growth factor signaling
FANCC and FANCA	Suppress Toll-like receptor (TLR)-induced tumor necrosis factor-alpha (*TNFα*) and interleukin 1 beta (*IL-1β*) expression in macrophages.	Prevent overproduction of inflammatory cytokines in macrophages
FANCC	Complexes with CtBP1 and suppresses *DKK1* (a *WNT* suppressor) gene expression.	Facilitates WNT signaling and haematopoietic stem cell self-renewal
FANCL	Ubiquitinates beta-catenin, enhancing its nuclear function.	Enhances pluripotency of HSPCs
FANCP	Suppresses accumulation of cytoplasmic DNA and the consequent activation of the interferon pathway.	Suppresses cGAS-STING and reduces pro-inflammatory cytokine production induced by replication stress

FANCD2 and FANCI	Associate with the spliceosomal protein SF3B1 independently of the core complex and influence SF3B1 trafficking.	Enhance stem cell function, coordinate DNA replication and co-transcriptional processes, and likely enhance erythropoiesis

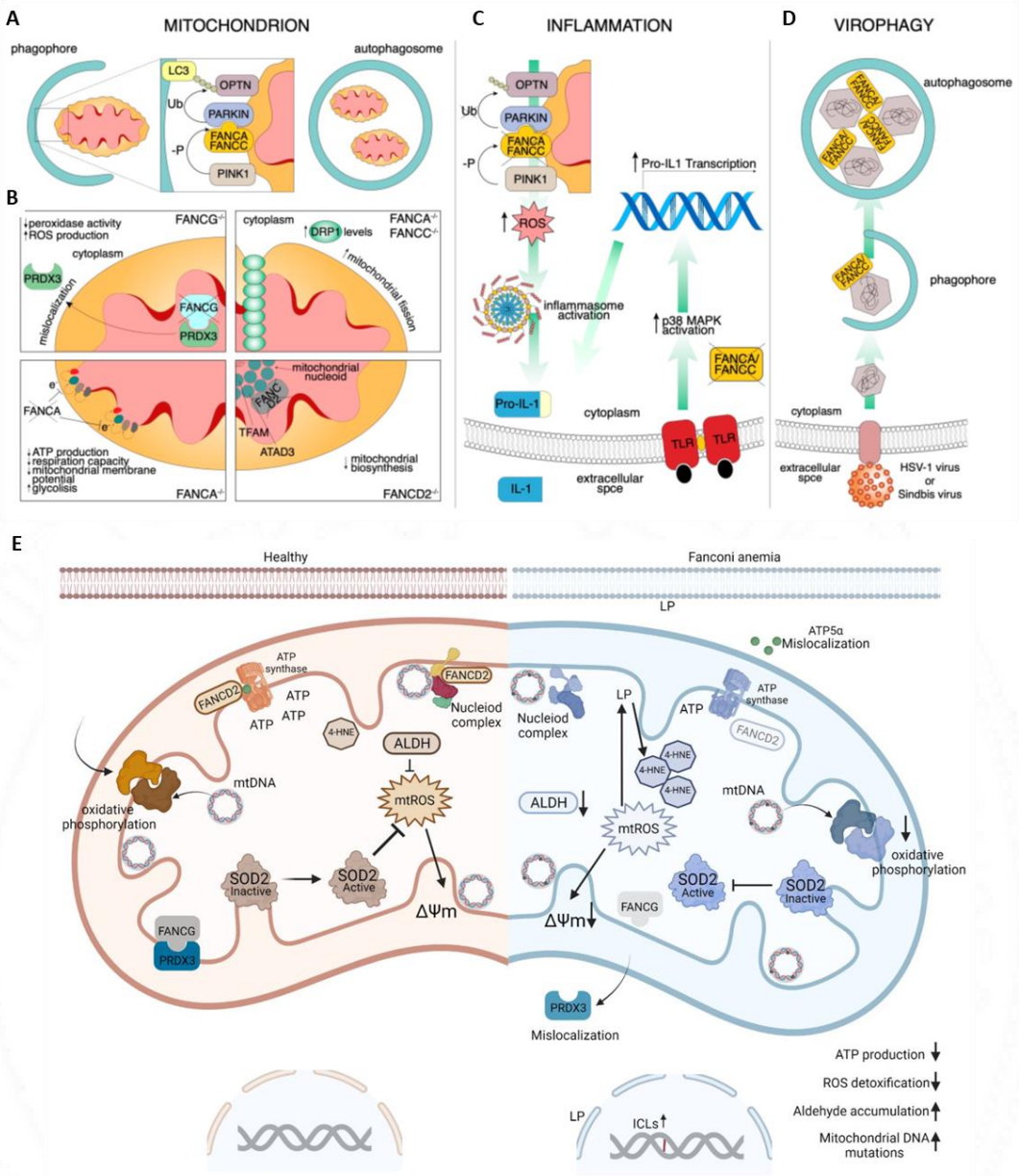


Figure 2.2: Schematic of the noncanonical FA pathways. **(A)** During mitophagy, damaged mitochondria are targeted by the phagophore. The E3 ligase PARKIN ubiquitinates the mitophagy receptor to identify the substrate, enabling LC3-mediated autophagosome recognition. FANCC aids PARKIN in localizing to mitochondria. **(B)** FA protein-related mitochondrial functions. (Top left) FANCG-deficient cells display PRDX3 mislocalisation. (Top right) *FANCA* and *FANCC* downregulated cells show increased DRP1 levels resulting in an increment of mitochondrial fission. (Bottom left) FANCA deficiency impacts negatively on mitochondrial electron transport chain efficiency, reducing ATP production and respiration capacity and altering the mitochondrial membrane potential. This inevitably leads to a shift from aerobic to

glycolytic metabolism. (Bottom right) FANCD2 mitochondrial nucleoid complex interaction regulates mitochondrial biosynthesis. **(C)** FANCA and FANCC knockdown increase the pro-inflammatory secretion of IL-1β in a dual convergent manner. On the left, impaired mitophagy stimulates inflammasome activation through ROS overproduction and, consequently, IL-1β is secreted. On the right, exogenously stimulated Toll-like receptors activate a signal transduction cascade that culminates with Pro-IL-1β transcription. **(D)** Both FANCA and FANCC are required for Sindbis and HSV-1 nucleocapsid recognition to control viral infection by virophagy. **(E)** Schematic of Fanconi Anaemia (FA) protein function in healthy (brown, left) versus FA-deficient (blue, right) mitochondria. Mitochondrial ROS (mtROS) are continuously generated by metabolic and inflammatory reactions. Mitochondrial DNA (mtDNA) primarily encodes mitochondrial proteins, including those involved in respiration and oxidative phosphorylation. Increased mtROS production leads to increased mtDNA mutations and lipid peroxidation (LP). FANCD2 localizes, in part, to mitochondria in a process facilitated by the ATPase Family AAA Domain-Containing Protein 3A (ATAD3) member (yellow barbell) of the mitochondrial nucleoid complex, which is essential for transcription and translation of mitochondrial proteins; FANCD2 deficiency destabilizes the nucleoid complex. FANCD2 also interacts with the ATP5α subunit (green dots) of ATP synthase. FAND2 knockdown results in the mislocalization of ATP5α and diminishes ATP production. In contrast, FANCG binds to PRDX3 on the inner mitochondrial membrane, and FANCG-mutant fibroblasts harbour mislocalized PRDX3, which results in reactive oxygen species (ROS) accumulation. Similar to peroxidase peroxiredoxin-3 (PRDX3), superoxide dismutase 2 (SOD2) has antioxidant activities in the mitochondria and is activated in the presence of ROS. FANCA-deficient cells have shown a decrease in SOD2 activity. Uncontrolled accumulation of ROS results in decreased mitochondrial transmembrane potential ($\Delta\Psi m$) and thus ATP synthesis. Schematic of FA protein function in healthy (brown, left) versus FA-deficient (blue, right) mitochondria. Ref: (Chihanga et al., 2022; Milletti Giacomo et al., 2020)

2.4.2.1. Inflammation: a damaging stressor in Fanconi anaemia haematopoietic cells

Due to the well-established myelosuppressive effects of interferon-gamma (IFNγ) in the development of acquired aplastic anaemia (Chen et al., 2015; Lin et al., 2014; Nakao et al., 1992), initial in vitro investigations were carried out on Fanconi Anaemia (FA) cells to explore the possibility of their heightened sensitivity towards IFNγ and other known inhibitory cytokines, such as tumour necrosis factor (TNF). These studies yielded positive results (Fagerlie et al., 2001; Grompe et al., 1995; Whitney et al., 1996) and led to the discovery of non-canonical signalling pathways in which FA proteins directly participated (Pang, 2001; Pang, Keeble, et al., 2001). More recent studies have unequivocally demonstrated that murine Fanconi Anaemia (FA) stem cells exhibit unique vulnerability to various inflammatory stresses (Hu et al., 2013; Walter et al., 2015) In a study by (Garbati et al., 2013) BMF was prevented by inhibiting the activity of interleukin beta (IL-1β), an inflammatory cytokine known to suppress the proliferation of FA haematopoietic stem cells (HSCs) (Hu et al., 2013).

Numerous inflammatory cytokines and adhesion molecules participate in the inflammatory response, including those that can enhance or inhibit the replicative capacity of haematopoietic stem and progenitor cells (HSPCs). Therefore, it is crucial to determine whether the 'replicative stress' induced in the FA HSPC pool contributes to cell damage or if the suppressive, pro-apoptotic, and differentiation-inducing effects of inflammatory cytokines cause cell death. But the inflammatory response and its association with oxidative stress (to which FA cells are also hypersensitive (Du et al., 2012; Li et al., 2010; Pang and Andreassen, 2009; Sejas et al., 2007; Zhang et al., 2004,

2007) provide opportunities for pharmacological intervention either in the ground state or during episodes of overt inflammation.

The inflammatory vulnerabilities observed in FA haematopoietic stem cells (HSCs) are likely attributed to a range of intricate indirect effects. One such effect involves the interaction between wild-type *FANCC* and *HSP70*, which suppresses the activation of eukaryotic translation initiation factor 2-alpha kinase 2 (PKR) in normal cells exposed to IFNγ along with double-stranded RNA or IFN combined with TNF. The interaction involves the canonical substrate-binding domain of HSP70, and loss of the FANCC/ HSP70 interaction results in the hyperactivation of PKR in FA cells (Pang, 2001; Pang et al., 2002; Pang, Keeble, et al., 2001). Notably, mutations in *FANCA*, *FANCG*, and *FANCC* have been found to augment the binding interaction between FANCC and PKR, leading to hyperactivation of PKR (Zhang et al., 2004). The recently described differential FANCA-binding functions of HSP90 and HSP70 and the functional consequences of such differential binding on the canonical function of various FANCA mutant substrates (Karras et al., 2017) add to the complexity of the future *in vivo* studies now required to confirm that these effects are truly independent of induced DNA damage. Furthermore, the induction of ALDH and HSP70 in response to specific viral infections, along with their collaborative role in controlling such infections (Lin et al., 2011), highlights the importance of investigating the interdependent relationships between ALDH, heat shock proteins, IFN-dependent pathways, and FA proteins in FA cells. These relationships may potentially be disrupted and warrant further formal examination.

2.4.2.2. Macrophage dysfunction in FA

An additional challenge related to the abnormal inflammatory response observed in FA is the excessive production of cytokines by FA mononuclear phagocytes, which are derived from FA HSCs, upon activation of Toll-like receptors (TLRs) (Brégnard et al., 2016; Dao et al., 2012; Fang et al., 2010; Garbati et al., 2013, 2016; Huard et al., 2014; Karras et al., 2017; Kumari et al., 2014; Lin et al., 2011; Matsunawa et al., 2014; Pang and Andreassen, 2009; Sejas et al., 2007; Svahn et al., 2015; Vanderwerf et al., 2009; Zhang et al., 2004, 2007). So, the pharmacological suppression of induced cytokine production or function may also reduce the stress on the HSPC pool. Preliminary in vitro evidence of these phenomena has been previously reported (Svahn et al., 2015). However, the range of non-canonical signalling defects in FA is expanding, involving interdependent pathways that present a challenge in selecting a specific molecular focus. Environmental cues that induce inflammation, for instance, often lead to oxidative stress (Brüne et al., 2013; Latorre et al., 2014; Lugrin et al., 2014; Salzano et al., 2014) and increased endogenous aldehyde levels (Cohen and Denkers, 2015; Hill et al., 1998; Ito et al., 2014). Any one of these stressors can excessively suppress FA HSPCs and induce hyperactive inflammatory cytokine production in more well-differentiated myeloid cells.

2.4.2.3. Mitophagy defects

Some FA proteins now provide novel functional insights into long-recognized ties between inflammation and oxidative stress. Many of the non-canonical functions outlined in **Table 2.1** have to do with controlling oxidative stress. For example, FANCC and FANCA normally participate in mitophagic responses by binding to

Parkin (which itself is known to play a role in removing damaged mitochondria) (Pickrell and Youle, 2015), thereby clearing damaged mitochondria and reducing reactive oxygen species. Knockdown of *FANCC*, *FANCF*, or *FANCL* also leads to defective mitophagy (Orvedahl et al., 2011). Importantly, the function of FANCC in Parkin-mediated mitophagy is independent of its role in genomic DNA damage repair (Sumpter, Sirasanagandla, Álvaro F. Fernández, et al., 2016). This point was recently demonstrated by using a *FANCC* mutant which does not rectify the dysfunction induced by FA mutations in the canonical pathway but does correct the Parkin-mediated mitophagic function of the protein (Sumpter, Sirasanagandla, Álvaro F. Fernández, et al., 2016). This same *FANCC* mutant (c.67delG, also known as 322delG) also rescues defective STAT1 activation in *FANCC*-deficient cells without correcting the canonical pathway (Pang, Christianson, et al., 2001; Pang et al., 2000) and is reportedly associated with a milder phenotype (Yamashita et al., 1996), suggesting that parallel non-canonical pathways contribute to the progression of BMF in FA.

2.4.2.4. FA and Replication Stress

Apart from their involvement in interstrand crosslink (ICL) repair, FA proteins also play a significant role in safeguarding genome stability against various replication stressors. These stressors encompass endogenous sources, such as oncogenes and aldehydes, as well as damaging agents like hydroxyurea (Niraj et al., 2019). Interestingly, certain FA proteins have been found to safeguard specific regions of the genome known as common fragile sites (CFSs). Notably, the monoubiquitinated forms

of FANCI and FANCD2 play a role in maintaining the integrity of CFSs FRA3B and FRA16D (Howlett et al., 2005), which are located within the large tumour suppressor genes *FHIT* and *WWOX*. Furthermore, the functions of FA proteins appear to be influenced by the levels of replication stress, as suggested by several studies. Under physiological conditions, FANCD2, independently of FANCI, interacts with and recruits the BLM helicase complex to restart stalled replication forks (Chaudhury et al., 2013). Additionally, in a FA pathway-independent manner, FANCD2 and FANCI are observed to associate with the replicative helicase MCM2–7 complex in response to replication stress mediated by ATR (Lossaint et al., 2013). At high levels of replication stress (induced by pharmacological treatment), FANCD2, FANCI, and the FA core complex proteins function synergistically to confer fork stability and promote replication restart (Michl et al., 2016).

Among the plethora of genome-protective mechanisms in which FA proteins take part, FANCM and its ability to act as a docking site for both FA core complex and BLM are essential to prevent sister chromatid exchanges (SCEs) and to suppress alternative lengthening of telomeres (ALT). The former mechanism often arises from unresolved D-loops that bypass Holliday junction intermediates, while the latter represents a telomere maintenance mechanism employed by cancers that do not reactivate telomerase expression (Johnson, 2000). In terms of mechanisms, the recruitment of FANCM and BLM complexes at lesion sites facilitates the dissociation of Rad-51-formed D-loops through their inherent helicase activity (Deans and West, 2009; Rosado et al., 2009).

Moreover, it has been demonstrated that FANCM can independently promote branch migration of Holliday junctions, likely impacting D-loops resolution (Gari et al., 2008). FANCM and BLM complex in cooperation with BRCA1 also play an important role in reducing and resolving replication stress that arises spontaneously within ALT telomeres (Pan et al., 2017). In a comprehensive study, (Pan et al., 2017) observed that the depletion of FANCM led to the activation of telomeric CHK1 signalling and the recruitment of BLM, RAD51, and BRCA1 at ALT telomeres. Concurrently, a recent investigation revealed an additional direct impact of the FANCM and BLM complex in suppressing ALT activity through their capacity to remodel replication forks (Lu et al., 2019). Both studies capitalized on the inherent vulnerability of ALT cancer cells to replication stress by either depleting FANCM or inhibiting the FANCM-BLM interaction pharmacologically, resulting in synthetic lethality. Another helicase known for its role in maintaining genome stability is FANCJ. In physiological conditions, guanine-rich sequences form four-stranded structures (G-quadruplexes, G4) that interfere with DNA replication, repair and RNA transcription (Maizels, 2008). To prevent that, FANCJ recognizes G4 through a specific binding motif and mediates multiple rounds of unfolding through its helicase activity (Wu et al., 2008; Wu and Spies, 2016).

2.4.2.5. Dysregulation of Mitochondrial Metabolism in FA-Deficient Cells

FA pathway deficiency not only leads to defects in interstrand crosslink (ICL) repair but also contributes to dysregulated cellular metabolism, making metabolic abnormalities a significant factor in the observed clinical phenotypes. This

dysregulation becomes apparent through alterations in mitochondrial structure and function, resulting in reduced energy production, increased oxidative stress, and impaired mitophagy. Moreover, mutations in FA proteins result in an elevated aldehyde load, causing aldehyde-induced damage, which, in turn, reduces the cellular capacity for aldehyde detoxification. This metabolic imbalance leads to the hyperproduction of inflammatory cytokines and increased sensitivity to their effects (Abad et al., 2021; Vanderwerf et al., 2009). The metabolic reprogramming observed in individuals with FA is associated with a range of phenotypes that cannot be fully explained by deficient DNA repair alone. These phenotypes include short stature, insulin resistance, thyroid dysfunction, abnormal body mass index (BMI), and dyslipidemia (Zhao Marion G. et al., 2018).

Metabolism, a dynamic process, plays a vital role in sustaining cellular viability. It encompasses various functions, including the maintenance of membrane potentials, the generation of metabolic energy through nutrient oxidation (catabolism) to support cell maintenance and repair, as well as the ATP-driven synthesis of complex macromolecules (anabolism) necessary for cell proliferation. Additionally, metabolism facilitates tissue-specific activities such as muscle contraction and the generation of action potentials in the brain. Cancer cells frequently exhibit alterations in nutrient uptake and utilization, with a notable dependence on glutamine (Lane et al., 2020) and many show a strong dependence on glutamine (DeBerardinis and Cheng, 2010; Yuneva et al., 2007). However, in FA-associated cancers, mitochondrial defects lead to reduced oxygen consumption, diminished tricarboxylic acid (TCA) cycle activity, and decreased reliance on glutaminolysis and glutamine oxidation (Ruppitsch

et al., 1997). Nonetheless, the amido nitrogen of glutamine is crucial for nucleobase formation and supporting proliferation (Abad et al., 2021). Notably, FA cells lacking FANCC display age-associated alterations in several amino acids, including glutamine, while aspartate and glutamate are associated with cancer (Nepal et al., 2018). Therefore, unless FA cells can upregulate glutamine synthetase, their proliferation becomes strictly dependent on exogenous glutamine. Uncontrolled proliferation, a hallmark of cancer, is sustained by the activation of pathways that couple energy production with biosynthesis. In this context, specific metabolic reprogramming mechanisms frequently observed in cancer, such as the generation of reactive oxygen species (ROS), mitochondrial DNA (mtDNA) mutations, aldehyde clearance, and lipid metabolism have been explored. A well-known example of metabolic reprogramming that occurs in malignant cells but is largely absent in non-transformed cells is the "Warburg effect." This phenomenon involves a preference for lactic fermentation over glucose oxidation, even under normoxic conditions, resulting in lactate accumulation and secretion (Jang et al., 2013; Vander Heiden et al., 2009; Warburg, 1924). The molecular mechanisms underlying the Warburg effect continue to be refined and updated. Mitochondrial dysfunction due to mtDNA mutations, the accumulation of ROS, and the release of oncometabolites into the cytosol (known as retrograde signalling) have been demonstrated to trigger cytosolic signalling pathways that promote neoplastic transformation (Hsu et al., 2016).

2.4.2.5.1. Mitochondrial Activities Play a Role in Oncogenesis

Mitochondria, the cellular organelles responsible for oxygen-dependent energy metabolism, play a crucial role in oncogenesis. They serve as both the main source and target of reactive oxygen species (ROS) formation. Under normal conditions, the transport of ADP, phosphate, and protons across the inner membrane accelerates electron transport, and the majority of the oxygen consumed by the respiratory electron chain is reduced to water. Mitochondrial DNA (mtDNA) encodes certain respiratory enzyme subunits necessary for oxidative phosphorylation, as well as rRNA and tRNA essential for mitochondrial protein synthesis. The remaining mitochondrial proteins are encoded by nuclear DNA (Wallace, 2012). Mutations in mtDNA are believed to significantly enhance oxidative phosphorylation, support neoplastic transformation, and meet the sustained bioenergetic demands of cancer cells (Park et al., 2009). A meta-analysis of 20 different cancer types from 859 non-FA patients revealed that 66% of them harboured mtDNA mutations, including adult leukaemia (9/24, 38%) and head and neck cancers (337/467, 72%) (Lee, 2014). Although the exact role of these mutations in cancer and metastasis is still unclear, they are thought to increase energy metabolism, ROS generation, and cell survival support. Numerous studies have established connections between FA gene products and mitochondrial dysfunction (Cappelli et al., 2017; Cuccarolo et al., 2012; Kumari et al., 2014; Milletti et al., 2020; Mukhopadhyay et al., 2006; Pallardó et al., 2010; Solanki et al., 2020). Oxidative stress-induced mitochondrial dysfunction, coupled with decreased scavenging of endogenous aldehydes (Langevin et al., 2011), increased lipid peroxidation (Marnett, 2000) and impaired ATP production (Cappelli et al., 2017; Kumari et al., 2014; Pagano et al., 2014), have emerged as metabolic phenotypic hallmarks of FA (Kumari et al.,

2014). Further comprehensive metabolic studies are required to determine compensatory reprogramming strategies in FA, such as increased lactic fermentation, which provides metabolic energy and intermediates for proliferative anabolism. Additionally, such studies can shed light on the creation of tumour-supporting microenvironments through acidification and negative regulation of immune cell function (Tu et al., 2021).

Recently, FA has been recognized as a mitochondrial disease (MD) due to the connections between FA proteins and impaired mitochondrial activities (Pagano et al., 2021). MDs are a group of disorders associated with mutations in nuclear and mitochondrial DNA, resulting in impaired oxidative phosphorylation. These disorders exhibit various clinical pathologies, including short stature, exercise intolerance, and hypertrophic or dilated cardiomyopathy (DiMauro and Hirano, 2009). While FA shares limited clinical similarities with MDs apart from short stature, both conditions feature mitochondrial dysfunction as a hallmark (Pagano et al., 2021).

2.4.2.5.2. Oxidative stress

Oxidative stress occurs when there is an imbalance favouring the production of reactive oxygen species (ROS) over the antioxidant defence mechanisms. The majority of ROS are generated during mitochondrial respiration. In FA cells, such as FANCA mutant fibroblasts, electron transport in Complexes I and III is impaired. This leads to alterations in the ATP/AMP ratio, resulting in decreased respiratory capacity, mitochondrial membrane potential, and oxygen uptake (Cappelli et al., 2013; Kruyt et al., 1998; Kumari et al., 2014; Pagano et al., 2014; Ravera et al., 2013) Additionally,

essential enzymes involved in energy production pathways may become inactivated. FANCA, FANCC, and FANCG have been found to interact with cytochrome P450-redox-related activities and respond to oxidative damage (Futaki, 2002; Kruyt et al., 1998). Moreover, the FANCG protein has been identified in mitochondria, with its interaction observed with the mitochondrial peroxidase peroxiredoxin3 (PRDX3) (Mukhopadhyay et al., 2006). However, in FA cells, PRDX3 is mislocalised, and thioredoxin-dependent peroxidase activity is significantly disrupted. Recently, a mitochondrial localization signal (MLS) has been identified on FANCG, and a single nucleotide change (C.65G>C) in eight FA patients results in the conversion of arginine at position 22 of the MLS to proline (p.Arg22Pro) (K et al., 2020). The mutant protein (R22P) fails to localize to the mitochondria and provide protection against oxidative stress. Interestingly, this mutant still participates in the formation of the FA core complex in the nucleus and exhibits resistance to interstrand crosslink (ICL) agents. Furthermore, in stable cells expressing the R22P mutant, there is a deficiency of iron in the FA protein FANCJ, an iron-sulphur (Fe-S)-containing helicase involved in DNA repair (K et al., 2020). This suggests that oxidative stress-mediated mitochondrial dysfunction directly contributes to defective FANCJ, leading to genomic instability.

2.4.2.5.3. Reactive oxygen species (ROS) may be a cause or consequence of mitochondrial abnormalities in FA.

Reactive oxygen species (ROS) are generated by several endogenous sources that require oxygen for their production (Ray et al., 2012). Hydrogen peroxide (H_2O_2) is produced through various oxygen-dependent oxidation reactions and the dismutation

of superoxide (Wang et al., 2018). Oxygen radicals, including superoxide, can oxidize macromolecules such as lipids and proteins, as well as form adducts between DNA strands (Wang, 2008). Antioxidant defence mechanisms, such as the upregulation of NADPH quinone oxidoreductase-1 and Redox factor-1 (Ref-1), help counteract the negative effects of ROS (Ray et al., 2012; Zahra et al., 2021). Oxidative stress is defined by the imbalance between oxidants and antioxidants (Marnett, 2000; Park et al., 2004). The mitochondrial respiratory chain is a major source of ROS production, occurring as a byproduct of ATP production facilitated by NADH oxidases (Beswick et al., 2001). Other endogenous enzymatic reactions outside the mitochondria, such as prostaglandin synthesis, phagocytosis, and the cytochrome P450 system, can also generate ROS (Snezhkina et al., 2019). Additionally, exogenous agents like metals, therapeutic agents, radiation, and environmental toxins contribute to ROS production (JI et al., 2013; Snezhkina et al., 2019).

In FA-deficient cells, it has been observed that mitochondria shift to a semi-resting state where ATP production is compromised, and the rate of oxygen consumption decreases due to impaired Complex I activity (Kumari et al., 2014; Ravera et al., 2013). Deficiencies in FANCA, FANCC, and FANCD2 have been associated with defects in mitochondrial respiratory chain Complex I activity, resulting in reduced ATP production (Ravera et al., 2013). Impaired oxygen consumption and decreased mitochondrial membrane potential, due to low ATP production, are phenotypes observed in FANCA, FANCC, and FANCD2 deficient cells (Kumari et al., 2014). Overexpression of superoxide dismutase 1 (SOD1) rescued oxygen uptake and respiration capacity in FA cells. Certain mitochondrial enzymes responsible for

ROS clearance exhibited an inability to respond to H_2O_2 in FA cells, suggesting impaired mitochondrial detoxification mechanisms (Kumari et al., 2014). In pathological conditions like FA, ROS accumulation and the failure to detoxify them further impair mitochondrial activities (Degan et al., 1995; Du et al., 2008; Park et al., 2004). The accumulation of ROS in FA-deficient cells leads to a disruption of the transmembrane potential, ultimately resulting in lower ATP levels compared to genetically corrected cells. Abnormal mitochondrial morphology, including membrane thinning or rupture, abnormal shapes, and impaired mitophagy, can contribute to the low transmembrane potential and ROS overproduction in FA cells. Treatment with H_2O_2 further increased mitochondrial fragmentation in these cells. However, the addition of the ROS scavenger N-acetyl-cysteine (NAC) reduced ROS production in FA-deficient cells (Kumari et al., 2014). Crucially, the presence of ROS scavengers was found to reduce the sensitivity of FA-deficient cells to crosslinking agents, including mitomycin C. Pretreatment of FA-deficient cells with NAC significantly improved ATP production, restored oxygen consumption, and increased resistance to mitomycin C, although abnormal mitochondrial morphologies were not rescued (Kumari et al., 2014). These findings suggest that some of the ROS-related phenotypes induced by mitochondrial dysfunction in FA-deficient cells can be reversible (Kumari et al., 2014; Ravera et al., 2013).

2.4.2.5.4. Increased ROS Exacerbate DNA Damage in FA

ROS are produced by various endogenous sources that rely on oxygen, and they play a role in organismal defence against pathogens (D'Autréaux and Toledano, 2007;

Leto et al., 2009; Shen and Li, 2010; Yang et al., 2011). However, at high concentrations, ROS can cause oxidative damage to DNA, proteins, and lipids, leading to the production of pro-inflammatory cytokines, apoptosis, autophagy, and necrosis. Elevated ROS levels also contribute to chronic inflammation, which can promote cancer initiation and increase cancer risk (Balkwill and Mantovani, 2001; Korniluk et al., 2017). Chronic inflammation and persistent infections, such as persistent HPV infection, can induce the production of chemokines and cytokines, which further promote cellular transformation, bypass tumour suppressor activity, and stimulate proliferation and angiogenesis (Hudson et al., 1999; Komori et al., 1993; Raman et al., 2007; Strieter et al., 1995). ROS stimulate the production of pro-inflammatory chemokines and cytokines, creating a positive feedback loop that results in further ROS accumulation and an inflammatory oncogenic environment (Jaiswal et al., 2000). Notably, ROS-induced lipid peroxidation generates aldehydes, particularly malondialdehyde (MDA) and 4-hydroxynonenal (4HNE) **(Figure 2)**, which can stimulate interstrand crosslinks (ICLs) and hyper-mutagenicity (Shen and Li, 2010; Yau, 1979). The increased production of endogenous aldehydes via ROS and lipid peroxidation may interact with classical ICL repair defects, contributing to the development of structural variants and chromosomal abnormalities characteristic of FA (Auerbach, 2009; Yang et al., 2021).

In FA cells, damaged mitochondria are more prone to rupture, triggering apoptosis even at lower stress levels compared to normal cells. Proteins such as FANCA, FANCC, FANCD2, FANCF, FANCL, BRCA1/FANCS, and BRCA2/FANCD1 are involved in Parkin-mediated mitophagy, revealing a non-

canonical role of FA proteins in disease pathologies (Solanki et al., 2020; Sumpter, Sirasanagandla, Álvaro F. Fernández, et al., 2016). Defective mitophagy leads to the accumulation of damaged mitochondria and increased intracellular oxidative stress (Drake et al., 2017). Interestingly, FANCC has been found to be essential for host immunity against herpes simplex virus type 1 and Sindbis virus in murine embryonic fibroblasts (MEFs), indicating its involvement in virophagy and autophagy (Sumpter, Sirasanagandla, Álvaro F. Fernández, et al., 2016). Overall, these findings suggest that genetic defects in FA lead to the accumulation of ROS, mitochondrial abnormalities, and impaired antioxidant defences, further compromising cellular respiration and ATP synthesis (Hadjur et al., 2001).

2.4.2.6. Metabolic dysfunction in FA

The mitochondria play a vital role in generating ATP through aerobic metabolism, and in FA cells, normal metabolism is disrupted. To compensate for the energetic defect, complementary pathways are likely involved. In a pioneering study on FA metabolism, glycolysis was identified as the primary energy source when aerobic metabolism was impaired due to dysfunctional mitochondrial electron transport complexes (Cappelli et al., 2017). However, glycolysis alone is insufficient to meet the energy demands of FA cells. The differentiation of haematopoietic stem cells (HSCs) into lymphocytes requires significant energy metabolism. The defective metabolic maturation observed in the bone marrow during the transition from a quiescent state may, at least in part, be explained by the impaired energy metabolism in FA cells.

2.4.2.6.1. Aldehyde metabolism and FA

FA patients typically do not have a history of exposure to external agents that induce interstrand crosslinks (ICLs). This raises the question of whether there are endogenous sources of ICLs in FA pathogenesis. Aldehydes, which are produced as byproducts of normal cellular metabolism, have been identified as capable of inducing ICLs (Voulgaridou et al., 2011). Studies using a chicken cell culture system found that cells deficient in various FA genes were sensitive to formaldehyde, further supporting the involvement of aldehydes in FA (Ridpath et al., 2007). The first in vivo evidence suggesting the role of aldehydes in FA pathogenesis focused on acetaldehyde, which is metabolized by the enzyme aldehyde dehydrogenase 2 (ALDH2) to acetate. Mice deficient in both ALDH2 and Fancd2 exhibited a more severe FA phenotype, including aplastic anaemia, leukaemia, and congenital defects, compared to Fancd2-deficient mice with a milder phenotype (Langevin et al., 2011). A similar study using Fanca-deficient mice confirmed several aspects of this interaction between impaired acetaldehyde metabolism and FA pathways, suggesting its generalizability to other FA genes (Oberbeck et al., 2014).

Subsequent to the acetaldehyde studies, an *in vivo* study involving formaldehyde provided further insights. Formaldehyde is not only produced as a byproduct of cellular demethylation reactions but is also found in various external sources, suggesting that lifestyle modifications may help reduce exposure. Mice deficient in both alcohol dehydrogenase 5 (ADH5), the major enzyme responsible for formaldehyde clearance, and Fancd2 (Adh5 -/- Fancd2 -/-) displayed more severe

impairment in the bone marrow compartment and increased dysfunction of the liver and kidneys compared to single knockout animals (Pontel et al., 2015).

These findings suggest that the accumulation of endogenous aldehydes, resulting from the disruption of key enzymes in aldehyde metabolism, can exacerbate certain FA phenotypes. However, establishing interstrand crosslinks (ICLs) as the pathogenic lesion resulting from excess aldehydes remains a challenging task, as both acetaldehyde and formaldehyde have the capacity to induce various types of DNA lesions (Voulgaridou et al., 2011).

Significantly, the findings from mouse studies were subsequently corroborated by investigations involving human subjects. In a study conducted on Japanese FA patients (Hira et al., 2013), the presence of the ALDH2E^{504K} allele in a heterozygous state was linked to a faster progression towards aplastic anaemia. However, the dominant-negative ALDH2 allele did not seem to influence the development of acute myeloid leukaemia (AML) or myelodysplastic syndrome (MDS). On the other hand, individuals homozygous for the ALDH2^{E504K} allele (identified in three FA patients in the study) exhibited a remarkably rapid progression to MDS and bone marrow failure. It is yet to be determined whether the presence of the ALDH2^{E504K} allele, within the context of FA, confers an increased risk of solid tumours.

The discovery that aldehyde metabolism can influence the FA phenotype carries significant clinical implications. Firstly, determining the ALDH2 genotype in FA patients, particularly those of East Asian descent, could help assess their susceptibility to bone marrow failure and potential development of malignancy. Secondly, recognizing the presence of aldehydes in the environment has led to an

increased understanding of the importance of lifestyle modifications in managing FA. For instance, patients with FA should be advised to avoid alcohol and tobacco. Thirdly, the aforementioned studies suggest the potential use of ALDH2 agonists in FA patients to delay the onset of bone marrow failure and potentially reduce the risk of malignancy (Chen et al., 2014). Nearly a decade ago, a specific agonist called Alda-1 was discovered as one potential option. (Chen et al., 2008).

There is potential for pharmacological modification of aldehyde metabolism in the treatment of FA without the need for new drug development. In a recent study, metformin, a widely used diabetes drug, was found to offer protection against chromosomal breakage and the development of malignancy in an FA mouse model (Q-S Zhang et al., 2016). The researchers discovered that metformin's protective effect may be attributed to its ability to scavenge aldehydes. This unexpected mechanism of action suggests the possibility of screening existing drug libraries for compounds with aldehyde detoxifying/scavenging activity. By identifying such drugs that already have established pharmacokinetic and safety data, it could accelerate therapeutic opportunities for FA patients.

2.4.2.7. TGFβ signalling and FA

The recognition of abnormal cytokine production in FA has been well-established for a while (Dufour, 2003; Rosselli et al., 1992). However, hyperactive transforming growth factor beta (TGFβ) signalling has also been specifically identified in FA patients (Korthof et al., 2013). Despite this discovery, it remains challenging to

determine from this study alone whether the increased TGFβ signalling actively contributes to the progression of FA.

A study conducted on FANCA−/− fibroblast line by (H Zhang et al., 2016) addressed this issue by using an shRNA library screen to identify genes that, when knocked down, could rescue the sensitivity of FANCA-deficient fibroblasts to mitomycin C (MMC). Among the top hits from this screen were genes involved in TGFβ signalling. Subsequent experiments involving the administration of a TGFβ neutralizing antibody to Fancd2-deficient mice demonstrated the rescue of haematopoietic stem cell function and prevention of bone marrow failure. This rescue mechanism was found to involve the upregulation of homologous recombination (HR) and the suppression of non-homologous end joining (NHEJ). The involvement of TGFβ signalling in regulating DNA repair processes has been previously reported in contexts unrelated to FA, as have the effects of suppressing NHEJ in FA cells (Kim et al., 2015; L Liu et al., 2014; Pal et al., 2017). These findings highlight the potential of targeted inhibition of modifier pathways to overcome the DNA repair defect in FA.

Numerous drugs targeting TGFβ signalling have been developed and are currently being evaluated in clinical trials, particularly for their potential as oncology drugs (Akhurst and Hata, 2012). Utilizing drugs targeting TGFβ signalling would offer a relatively expedient approach compared to developing new drugs from scratch. In the context of FA, these drugs hold clinical potential for several reasons. Firstly, they have the potential to slow down the progression of the disease, particularly in relation to bone marrow failure. Secondly, TGFβ inhibitors may enhance the ex vivo expansion of stem cells derived from FA patients, overcoming limitations in replicative potential

and facilitating gene therapy and gene editing approaches (Habi et al., 2005)). Thirdly, targeting the TGFβ pathway may provide new opportunities for managing malignancy development in FA patients. However, the use of TGFβ inhibitors in this context is not without complexities. TGFβ exhibits pleiotropic effects in cancer biology, acting as a tumour suppressor in the early stages of cancer development and a tumour promoter in later stages (Roberts and Wakefield, 2003; Wakefield, 2002). Consequently, administering TGFβ inhibitors in FA patients may have unintended consequences in terms of cancer development and progression. While TGFβ inhibition appears to have a favourable impact on the bone marrow failure component of FA, further studies are needed to investigate the intricate interplay between TGFβ and FA pathways in carcinogenesis.

2.4.2.8. Autophagy and FA

Cells possess an inherent capability to eliminate unwanted structures through degradation processes. One specific type of degradation is mitophagy, which involves the autophagic removal of damaged mitochondria. Another form of degradation, known as virophagy, pertains to the elimination of viruses. In an investigation conducted by (Sumpter, Sirasanagandla, Álvaro F. Fernández, et al., 2016), it was discovered that mice lacking FANCC, and cells derived from FA-C patients exhibited defects in both mitophagy and virophagy. Importantly, these deficiencies were independent of FANCC's role in DNA repair. Furthermore, the impairment of mitophagy was observed upon the knockdown of various other FA genes, including *FANCA*, *FANCF*, *FANCL*, *FANCD1*, *FANCD2*, and *FANCS*, indicating that this

defect could contribute to disease progression in multiple complementation groups. Notably, these findings were partly corroborated by a separate study conducted by (Shyamsunder et al., 2016).

The impaired virophagy observed in FA has significant implications, particularly in terms of susceptibility to viral infections, including those that have been implicated in tumour formation, such as Human Papilloma Virus (HPV). The relationship between HPV and cancer development in FA patients remains a topic of debate (Alter et al., 2013; Sauter et al., 2015). However, the findings discussed in this study offer support for the existing guidelines recommending vaccination of FA patients against HPV. The defective virophagy observed in FA provides a rationale for this preventive measure, aiming to reduce the risk of viral-associated complications in this patient population.

The impaired process of mitophagy in FA patients can contribute to the generation of reactive oxygen species (ROS), leading to two important outcomes. Firstly, it can activate a cytokine-driven inflammatory state, which contributes to the development of bone marrow failure. Secondly, it can cause genomic damage, further increasing the risk of malignancy in FA patients. The association between oxidative stress and FA has been recognized in previous studies (Du et al., 2008), and the use of antioxidants as a potential treatment for FA has been proposed (Zhang et al., 2010). However, there is currently a lack of clear evidence demonstrating the benefits of antioxidant therapy in FA patients, and concerns about potential toxicities associated with antioxidant use should be taken into account. Nonetheless, considering the role of FA proteins in mitophagy, it may be worthwhile to conduct more controlled trials

to evaluate the use of antioxidants in FA patients. Similarly, modulators of cytokine signalling could be explored as potential therapeutic options. In line with the mitophagy study, previous research has also reported an abnormal inflammasome response in Fancc-/- mice (Garbati et al., 2013), further supporting the potential relevance of targeting cytokine signalling in FA.

Although the identification of non-canonical FA gene functions is intriguing, it is important to note that their existence alone does not definitively establish their role in driving the progression of FA, even if they offer explanations for certain FA cellular phenotypes. Further investigations and studies are required to fully understand the contributions of impaired autophagy to FA and to determine its significance in the disease.

2.4.3 Bone marrow failure (BMF) in Fanconi Anaemia (FA): Involvement of nucleolar stress and perturbed ribosome biogenesis

2.4.3.1. Implications of nucleolar Stress and ribosome biogenesis

In order to gain deeper insights into the biochemical and molecular origins of BMF in FA it is imperative to underscore recent publications that have revealed surprising connections between the FANC/BRCA pathway and the homeostasis of nucleolar and ribosome biogenesis. Significantly, the impairment of *FANCA* function, detected in more than 75% of FA patients globally (with approximately 60-70% experiencing *FANCA* loss and an additional 10-15% facing a lesser extent of *FANCG* loss), results in nucleolar stress and the erroneous localization of nucleolar proteins.

These findings shed light on crucial mechanisms underlying the development of BMF in FA. (Gueiderikh et al., 2021).

FANCA was found to coimmunoprecipitate with Nucleolin and Nucleophosmin 1 (NPM1), two crucial nucleolar proteins. This observation indicates FANCA's involvement in the aggregation of proteins within the nucleolus, a membrane-less condensate formed around ribosomal DNA (rDNA) repeats. The structural and genetic integrity of this nucleolar condensate heavily relies on the proper functioning of the FANC/BRCA pathway (Gueiderikh et al., 2021; Samuel B. Sondalle et al., 2019). Previous studies have demonstrated that the interaction between FANCC and FANCA plays a role in stabilizing NMP1 proteins in the cytoplasm, particularly when their Nuclear Localization Signals (NLS) are mutated (Du et al., 2010). Around 30% of sporadic AML cases have been found to exhibit these mutations. (Falini et al., 2020; Zarka et al., 2020). Consequently, the FANC core complex interacts with both wild-type (WT) and mutated NPM1, playing a role in stabilizing NPM1 and regulating its subcellular localization. The loss of function in *FANCA* and *FANCI* impacts rDNA transcription and pre-rRNA processing. Additionally, FANCA and FANCI, along with FANCG, are essential for sustaining the overall protein synthesis rate (Gueiderikh et al., 2021; Samuel B. Sondalle et al., 2019).

Analysis using mass spectrometry demonstrated that among the FANC/BRCA pathway proteins, only FANCA, FANCI, and FANCD2 were identified in the 40S, 60S, and 80S ribosome fractions. Interestingly, in the polysome fraction, FANCD2 and FANCI were the sole proteins from the FANC/BRCA pathway that were detected

(Gueiderikh et al., 2021). The coimmunoprecipitation of FANCD2 and FANCI with ribosomal protein S27-like (RPS27L) was observed, and it was found that reduced levels of RPS27L affect the stability of FANCD2 and FANCI, resulting in impaired repair of DNA interstrand crosslinks (ICLs) (Sun et al., 2020). The initial single-cell transcriptome analysis of primary HSPCs (haematopoietic stem and progenitor cells) derived from both healthy donors and FA patients revealed that a significant portion of BM-derived FA cells exhibited increased expression levels of MYC and its target genes. However, in a smaller fraction of cells, this upregulation was balanced by heightened activation of p53- and TGFβ-associated intracellular signalling pathways (Renaudin and Rosselli, 2021; Rodríguez et al., 2021). MYC is a well-established master controller of ribosome biogenesis and protein synthesis (van Riggelen et al., 2010). Consequently, several mRNAs encoding ribosomal proteins and translation factors, such as RPL19, RPL13, eIF4A1, eIF4E, eIF4B, RPL5, and RPL27A, exhibit disrupted expression in HSPCs of FA patients (Rodríguez et al., 2021). Examining the transcriptional profiles in patient-derived models of haematopoietic differentiation and analysing the proteomic profile of BM samples obtained from FA patients has unveiled enriched signatures indicating changes in ribosome biogenesis and translation in cells lacking FANCA (Hou et al., 2020; Marion et al., 2020).

According to reports, the deficiency of FANCA results in imbalances in the levels of ribosomal proteins and translation factors associated with ribosome subunits, monosomes, and polysomes. This includes proteins like eIF4G1, eIF4E, eIF5A, and RPL22A (Gueiderikh et al., 2021). Furthermore, it has been observed that BRCA1/FANCS coexists with RNA Pol I at the rDNA repeat and forms a

coimmunoprecipitated complex with the RNA Pol I holoenzyme, which plays a crucial role in rDNA transcription. In the absence of BRCA1, rDNA transcription is reduced (Johnston et al., 2016). Therefore, these changes in ribosome biogenesis might be directly linked to at least four proteins encoded by genes defining three different modules of the FANC/BRCA pathway, namely, *FANCA*, *FANCG*, *FANCI*, and *BRCA1*/*FANCS*.

2.4.4. Inheritance and type of mutations

FA is characterized by mutations in at least 22 different genes, referred to as *FANCA* to *FANCW* (Niraj et al., 2019). While most cases of FA follow an autosomal recessive pattern of inheritance, certain FA-associated genes may exhibit other inheritance patterns. The commonly mutated genes in FA include *FANCA*, *FANCC*, and *FANCG*, and some FA-related genes were previously identified under different names in other contexts (e.g., *FANCS*, also known as *BRCA1*). Pathogenic variants in FA genes can involve point mutations, large deletions, or gene duplications (Chandrasekharappa et al., 2013). There have been reports of genotype-phenotype correlations for specific FA genes (Demuth et al., 2000; David I. Kutler, Singh, et al., 2003).

In most cases, FA genes follow an autosomal recessive inheritance pattern, meaning that the presence of disease requires the loss of function in both alleles. This can occur through homozygous pathogenic variants or compound heterozygous variants within a single gene. However, there are exceptions, such as *FANCR*, which follows an autosomal dominant pattern, and *FANCB*, which is X-linked recessive.

Individuals who possess heterozygous variants in FA genes, excluding FANCB and FANCR, are regarded as unaffected carriers of the condition.

2.4.5. Genomic instability

FA proteins play a crucial role in maintaining the stability of the genome. When the function of FA genes is lost, the repair of DNA interstrand crosslinks (ICLs) is disrupted, leading to a hindrance in normal DNA replication and transcription. This interference occurs as a result of impaired strand separation, replication fork stalling, and compromised DNA integrity (Duxin and Walter, 2015). ICLs can be generated from various sources, including endogenous aldehydes (such as lipid peroxidation byproducts), exogenous aldehydes (formed after alcohol consumption), as well as exposure to radiation and DNA alkylating agents used in cancer chemotherapy. The faulty repair of ICLs contributes to genomic instability, abnormal regulation of the cell cycle, and cell death. Furthermore, FA proteins interact with other pathways involved in DNA damage response and participate in stress response mechanisms, including oxidative stress (e.g., oxidative stress pathways).

2.4.6. Phenotype in FA

FA proteins are present in all tissues, as documented by (Uhlén et al., 2015). However, the consequences of losing the FA pathway are specific to each organ. The underlying mechanisms are not well understood, partially due to the incomplete representation of FA phenotypes in most mouse models (Parmar et al., 2009). Progressive bone marrow failure is a prominent characteristic of FA. It occurs when

haematopoietic stem cells become depleted due to various stressors, including elevated TNFα levels (Li et al., 2007), aldehyde toxicity (Langevin et al., 2011) and/or overactive p53 signalling (Ceccaldi, Parmar, Mouly, Delord, Jung Min Kim, et al., 2012). This failure in the bone marrow can subsequently lead to the development of myelodysplastic syndrome (MDS) and acute myeloid leukaemia (AML) (Li et al., 2007; Romick-Rosendale et al., 2013; Savage and Walsh, 2018).

Individuals with FA mutations have a significantly increased risk of developing acute myeloid leukaemia (AML) by 700-fold and myelodysplastic syndrome (MDS) by 6000-fold, compared to the general population (Blanche P. Alter, 2014). Interestingly, acquired mutations and epigenetic silencing of FA genes have also been detected in sporadic cases of AML, although at lower frequencies (Cerami et al., 2012; Hess et al., 2008; Tischkowitz et al., 2003, 2004). Currently, haematopoietic stem cell transplantation (HSCT) is the only curative treatment option for FA-related haematologic malignancies. Modified HSCT conditioning regimens with reduced alkylating agent doses have improved survival rates for FA patients with haematological malignancies, although outcomes are still not as optimal as those observed in non-FA individuals undergoing transplantation for marrow failure. It's important to note that patients who survive after successful HSCT continue to be at a high risk of developing non-haematological malignancies (Alter, 2003; Kottemann and Smogorzewska, 2013; David I. Kutler, Auerbach, et al., 2003).

In contrast to the bone marrow failure observed in the haematopoietic system, Fanconi anaemia (FA) does not show a similar decline in stem cell fitness in other hyper-proliferative organs like the skin or mucosa. However, individuals affected by

FA are prone to developing early-onset and highly aggressive squamous cell carcinomas (SCCs) derived from keratinocytes. These SCCs primarily affect the head and neck, oesophagus, anogenital tract, and skin (David I. Kutler, Auerbach, et al., 2003; David I. Kutler, Singh, et al., 2003; Kutler et al., 2016; Rosenberg et al., 2008; Ruiz-Torres et al., 2021). The risk of developing SCCs in FA patients is significantly elevated compared to the general population, with a 700-fold increased risk for head and neck SCC, a 2000-fold increased risk for oesophageal SCC, and a 4000-fold increased risk for vulvar SCC (Alter et al., 2013; David I. Kutler, Auerbach, et al., 2003; Romick-Rosendale et al., 2013; Velleuer and Dietrich, 2014). Due to the heightened sensitivity of FA-deficient cells to DNA damage, especially crosslinks, conventional chemotherapy and radiation treatments pose significant challenges in the management of these patients (Khoury et al., 2018). As a result, surgical resection is often the preferred treatment option, albeit with high tumour recurrence rates and poor long-term survival outcomes.

Limited studies have been conducted on epidermal carcinogenesis in individuals with Fanconi anaemia (FA). However, analysis of data from The Cancer Genome Atlas (TCGA) has revealed that approximately 18% of sporadic head and neck squamous cell carcinomas (SCCs) exhibit point mutations and other genetic variations in FA genes, including deletions (Nalepa and Clapp, 2018). These findings suggest a selective pressure for the loss of the FA pathway during the development of cancer. Recent transcriptome-wide association studies have further supported these results, identifying reduced expression of FANCA in sporadic skin SCC (Lane et al., 2020). Cell lines derived from head and neck SCCs, which are FA-proficient but have

undergone knockdown of *FANCA*, *FANCD2*, or *FANCJ*, exhibit specific FA-related defects, a shift towards an epithelial-mesenchymal transition (EMT)-like state, and a more invasive phenotype (DeBerardinis and Cheng, 2010). These observations suggest that both constitutional FA pathway loss in FA patients and somatic acquisition of FA pathway alterations in the general population contribute to the early and aggressive development of SCCs in the skin and mucosa, although the underlying mechanisms remain poorly understood. Furthermore, the skin of individuals with FA and laboratory models of FA demonstrates defects in cell-cell and cell-substrate adhesion complexes in keratinocytes, along with EMT-like phenotypes (Yuneva et al., 2007). These findings, combined with the susceptibility of FA patients to skin blistering, suggest that the FA pathway plays a role in maintaining the integrity of the epidermis (Yuneva et al., 2007).

The physical characteristics associated with Fanconi anaemia (FA) exhibit significant heterogeneity and affect multiple systems, providing important clues for testing and early diagnosis (Shimamura and Alter, 2010). Classical congenital abnormalities observed in FA patients include those found in the VACTERL-H (Vertebral, Anal, Cardiac, Tracheo-oesophageal fistula, Oesophageal atresia, Renal, upper Limb and Hydrocephalus) association (Solomon et al., 2012). In a comprehensive literature review of FA cases, approximately 12% of FA patients met the criteria for VACTERL-H association (presence of at least 3 of the 8 common features), which is consistent with previous studies reporting a range of 5% to 30% (Alter and Giri, 2016; Faivre et al., 2005). Other common abnormalities observed in FA patients have been categorized under the acronym PHENOS (skin Pigmentation,

small Head, small Eyes, Nervous system, Otology, and Short stature) (Alter and Giri, 2016). In the aforementioned literature review, approximately 9% of FA patients exhibited ≥4 out of the 6 PHENOS features (Fiesco-Roa et al., 2019). According to (Shimamura and Alter, 2010), the most commonly observed abnormalities seen in individuals with Fanconi anaemia (FA) comprise short stature, alterations in skin pigmentation, malformations of the upper limbs, abnormalities in male genitalia, microcephaly (characterized by a small head), and manifestations affecting the eyes and kidneys. These features, with the exception of male genitalia anomalies, are encompassed within the VACTERL-H or PHENOS classifications. It is important to note that while the majority of FA patients will present with at least one abnormality, approximately 25-40% may exhibit no evident abnormalities, indicating that the absence of these features does not rule out an FA diagnosis (Alter and Giri, 2016; Faivre et al., 2005). **Table 2.2** provides a useful guide for evaluating patients whose physical appearance suggests a possible diagnosis of FA. The combination of abnormalities listed in **Table 2.2** can assist in further assessing potential FA cases.

Table 2.2. Manifestations that are indicators for Fanconi anaemia screening.

Organ, system, or feature	Abnormality
Height	Short stature
Head	Microcephaly
Central nervous system	Small pituitary and stalk interruption; agenesis of the corpus callosum; cerebellar hypoplasia; hydrocephalus; dilated ventricles; developmental delay
Eyes	Microphthalmia; epicanthal folds; almond-shaped fissures; ptosis; strabismus; cataracts
Otology	Hearing loss (conductive, sensorineural, or mixed); abnormal pinna; atretic, narrow canal; and abnormal middle ear bones
Facial	FA facies; triangular face; micrognathia; pointed chin; mid-face hypoplasia; facial nerve palsy; microsomia; hypertelorism; hypotelorism; cleft palate
Heart	Patent ductus arteriosus; atrial septal defect; ventricular septal defect; coarctation; situs inversus; truncus arteriosus
Gastrointestinal	Tracheoesophageal fistula Atresias: oesophageal, duodenal, jejunal Anal malformations: imperforate or bifurcated anus Annular pancreas Intestinal malrotation
Renal	Horseshoe, ectopic, hypoplastic, dysplastic, absent, hydronephrosis, hydroureter
Male Genitalia	Undescended, small or absent testis; microphalus; hypospadias; micropenis; absent testis; infertility
Female Genitalia	Hypoplastic, absent or bicornuate uterus; gonadal dysgenesis; small ovaries; rectovaginal fistula; vaginal atresia; late menarche; early menopause; infertility
Upper limb	Thumb: absent, hypoplastic, triphalangeal, polydactyly Radius: absent, hypoplastic

	Thenar-eminence: hypoplastic, absent Others: absent first metacarpal, clinodactyly Ulna: short, dysplastic
Lower limb	Hips: congenital dislocation/dysplasia, malrotation Feet: toe syndactyly, abnormal toes, club feet
Vertebral	Web, hemivertebrae; Klippel-Feil; scoliosis; kyphosis; coccygeal aplasia
Skin	Café au lait macules; generalized hypo- or hyperpigmentation
Bone marrow failure	Anaemia; leukopenia; thrombocytopenia; aplastic anaemia; myelodysplastic syndrome
Leukaemia	Mainly acute myeloid leukaemia
Squamous cell carcinoma	Head and neck; oesophageal; anogenital (including vulvar, skin)
Other cancers	Skin basal cell carcinoma; medulloblastoma; neuroblastoma; Wilms' tumour; breast; lung

2.4.7. Genotype/Phenotype associations in Fanconi anaemia

FA is a disease characterized by genetic and clinical diversity. In certain cases, knowing the specific gene and variant(s) involved becomes crucial for identifying potential risks and gaining insights into the clinical course. Medical management for most individuals with FA is tailored to their specific clinical presentation. However, for those with variants in genes leading to altered phenotypes, identifying the genotype becomes vital for appropriate medical care and prognostic purposes, especially as certain genes with FA-like phenotypes may not exhibit classic FA symptoms. Genotype-to-phenotype correlations are often based on a limited number of cases, and there have been instances where outliers to the typical phenotype have been observed. Below, are several FA variants for which sufficient information is available.

2.4.7.1. FANCA

Based on research conducted by (Faivre et al., 2000), individuals carrying homozygous null variants in the FANCA gene tend to develop anaemia at an earlier age and have a greater likelihood of developing leukaemia compared to those with FANCA variants that retain some level of function. Nevertheless, a distinct examination carried out by (Maria Castella, Pujol, Callén, Trujillo, Casado, Gille, Lach, Auerbach, Schindler, Benítez, Porto, Ferro, Muñoz, et al., 2011) revealed that the age of onset of anaemia and the incidence of leukaemia were unaffected in patients with homozygous null *FANCA* variants or in those expressing an abnormal form of the protein. Certain variants, including p.His913Pro and p.Arg951Gln/Trp, have been

linked to a delayed onset of the disease and a slower progression of haematological symptoms (Bottega et al., 2018).

2.4.7.2. FANCB

Males carrying a truncating variant in the *FANCB* gene often exhibit obvious characteristics that align with VACTERL-H (McCauley et al., 2011). However, milder manifestations have been observed in patients with missense variants or somatic mosaicism (Asur et al., 2018; Jung et al., 2020). On the other hand, female carriers of *FANCB* do not seem to show any related disease manifestations (McCauley et al., 2011).

2.4.7.3. FANCC

The International Fanconi Anaemia Registry (IFAR) observed that individuals with *FANCC* gene variants experienced an earlier onset of bone marrow failure and had lower survival rates compared to those with *FANCA* or *FANCG* variants (David I. Kutler, Singh, et al., 2003). However, the European FA Research Group's findings contradicted this, as they described the *FANCC* group having the least severe haematologic course and fewer somatic abnormalities compared to *FANCA* and *FANCG* (Faivre et al., 2000). Multiple variants in the *FANCC* gene have been linked to specific phenotypes. Variants located in exon 15 were associated with the development of blood abnormalities at an earlier age, more congenital abnormalities, and poorer survival when compared to variants in exon 2 (Gillio et al., 1997; David I. Kutler, Singh, et al., 2003). Additionally, the variant c.456+4A>T was found to be

related to a more severe disease presentation in Ashkenazi Jewish individuals (Gillio et al., 1997; Yamashita et al., 1996). However, this specific variant has also been documented in other populations (AFTAB et al., 2017; Rodriguez et al., 2005), where it might not be correlated with a severe phenotype in certain groups. Several studies suggest that the c.67delG founder variant is generally associated with milder symptoms, but there have been exceptions (de Vries et al., 2012; Gillio et al., 1997; Yamashita et al., 1996). In the Saudi population, a study reported that the founder variant c.165+1G>T might be associated with a milder form of the disease (Hartmann et al., 2010).

2.4.7.4. FANCD1/BRCA2

In 2002, a study revealed that individuals with FA and pathogenic biallelic variants of the *BRCA2* gene could develop leukaemia, acute myeloid leukaemia (AML), or acute lymphoblastic leukaemia (ALL) at a significantly earlier age than expected (Howlett et al., 2002). Moreover, they faced a heightened risk of developing brain tumours (e.g., medulloblastoma, glioblastoma multiforme, astrocytoma) and kidney tumours (e.g., Wilms tumour), which are not commonly associated with FA (Alter et al., 2006; Offit, 2006). Therefore, patients with biallelic *FANCD1/BRCA2* variants should consider additional screening with brain magnetic resonance imaging (MRI) and kidney ultrasound (Wagner, 2004). While certain studies conducted in this population have indicated a severe phenotype, characterized by the presence of multiple congenital abnormalities and an elevated 97% risk of developing any type of malignancy by the age of 5.2 years (Alter et al., 2006), there have also been

documented reports of older individuals exhibiting milder manifestations or experiencing a later onset of the disease (Degrolard-Courcet et al., 2014; Rickman and Smogorzewska, 2019).

2.4.7.5. FANCG

The European FA Research Group reported that individuals with pathogenic variants in this gene experience more severe cytopenia and a higher incidence of leukaemia compared to patients with variants in other FA genes (Faivre et al., 2000). However, this pattern was not observed in the data collected by the IFAR (David I. Kutler, Singh, et al., 2003).

2.4.7.6. FANCM

The *FANCM* phenotype requires further investigation for better understanding. Originally suggested as a core complex gene in 2005, *FANCM* was associated with a Fanconi Anaemia (FA) phenotype in a family with siblings affected by the condition (Meetei et al., 2005). However, later studies revealed biallelic *FANCA* variants in those affected siblings, which raised questions about status of *FANCM* as a canonical FA gene (Singh et al., 2009). Following that, biallelic loss of function *FANCM* variants have been detected in individuals diagnosed with FA. However, some researchers propose that these variants might give rise to an alternate phenotype characterized by an early-onset cancer syndrome rather than the classical FA phenotype, as these particular cohorts did not exhibit bone marrow failure and congenital anomalies (Bogliolo et al., 2018; Catucci et al., 2018). A patient with compound heterozygous

FANCM variants was reported in 2014. These individual exhibited characteristics such as chromosome fragility, anomalies in the right hand's thumb and thenar eminence, and experienced bone marrow failure (Chang et al., 2014). Additionally, there have been reports of early-onset breast cancer and reduced fertility, including primary ovarian insufficiency and mild to severe spermatogenesis in two families, in presumed biallelic carriers without an overt phenotype (Catucci et al., 2018; Fouquet et al., 2017; Yin et al., 2019).

2.4.7.7. FANCN/PALB2

Typically, individuals with variants in the *FANCN*/*PALB2* gene exhibit a more severe clinical presentation. Similar to the *FANCD1*/*BRCA2* phenotype, these individuals tend to develop solid tumours and leukaemia at an earlier age compared to those with variants in other FA genes (Reid et al., 2007). Commonly observed tumours associated with *FANCN/PALB2* variants include medulloblastoma, Wilms tumour, acute myeloid leukaemia (AML), and neuroblastoma (Ghazwani et al., 2016; Reid et al., 2007; Serra et al., 2012). In the absence of consensus guidelines, it may be appropriate to consider implementing cancer surveillance recommendations designed for patients with biallelic *FANCD1/BRCA2* variants for individuals with *FANCN/PALB2* variants. However, it is worth noting that there have been reports of phenotypes that fall outside this typical spectrum (Byrd et al., 2016), indicating that as additional cases emerge over time, the phenotypic range associated with *FANCN/PALB2*-associated FA may expand further.

2.4.7.8. FANCO/RAD51C

There have been reports of two families with an FA-like disorder and biallelic variants in *FANCO/RAD51C* (Jacquinet et al., 2018; Vaz et al., 2010). In both families, affected individuals presented significant congenital anomalies, some of which are atypical in classical FA, such as palate anomalies, holoprosencephaly, and overlapping fingers. FA diagnoses were confirmed by demonstrating hypersensitivity to diepoxybutane (DEB) and mitomycin C (MMC), along with increased radial breaks. However, the risk of haematologic features and squamous cell tumours in these cases remains unknown.

2.4.7.9. FANCR/RAD51

The clinical presentation of *FANCR* exemplifies the heterogeneous nature of FA. A patient with *FANCR* was identified to have microcephaly, slow growth, limb abnormalities, and severe learning impairment (Ameziane et al., 2015). Chromosome breakage assays were positive, and cells were sensitive to multiple chemotherapeutic agents, with intact FANCD2 monoubiquitination, indicating a downstream mutation in a gene related to this modification. Notably, bone marrow failure was not part of the patient's clinical presentation, leading to a classification as FA-like despite meeting FA criteria based on chromosome breakage.

2.4.7.10. FANCS/BRCA1

The first confirmed cases of biallelic *FANCS*/*BRCA1* variants were reported in two women, one with ovarian carcinoma and severe cisplatin treatment toxicity and the other with ductal breast carcinoma (Domchek et al., 2013; Sawyer et al., 2015).

Both individuals had short stature, microcephaly, dysmorphology, and intellectual or developmental disability. In another publication, two families with four children were reported to have chromosome breakage consistent with FA and homozygous truncating *BRCA1* variants. Among the four children, one developed T-cell acute lymphocytic leukaemia at 5 years and another developed neuroblastoma at 2 years. The other two children remained free of cancer at the ages of 5 and 15.5 years, respectively (Seo et al., 2018). Another case of *FANCS/BRCA1* was reported in a 2.5-year-old female with similar clinical features but without a history of cancer (Freire et al., 2018).

2.4.7.11. FANCQ/ERCC4

Biallelic variants in *FANCQ*/*ERCC4* have been linked not only to the FA phenotype but also to autosomal recessive Cockayne syndrome, xeroderma pigmentosum, and a single case of XFE progeroid syndrome. The phenotypes can be diverse, depending on how gene function is affected (Bogliolo et al., 2013; Kashiyama et al., 2013; Niedernhofer et al., 2006; Sijbers et al., 1996).

2.4.7.12. FANCT/UBE2T

Unlike most recently identified FA genes, which function downstream of FANCI/D2 monoubiquitination, *FANCT* (*UBE2T*) acts upstream and interacts with the FA core complex component *FANCL*. Biallelic mutations in *UBE2T* were first reported in two Japanese patients, both experiencing bone marrow failure in the first decade of life (Hira et al., 2015). One of these patients developed myelodysplastic

syndrome (MDS) progressing to acute myeloid leukaemia (AML). The third patient was reported in two separate studies and displayed multiple congenital abnormalities, positive chromosome breakage assay but no history of malignancy (Rickman et al., 2015; Virts et al., 2015). The third patient exhibited somatic mosaicism, with relatively normal *UBE2T* transcript levels and FA pathway function in the haematopoietic compartment, despite compound heterozygous mutations in the *UBE2T* gene. As a result, *UBE2T* was not identified as the causal gene based on peripheral blood analysis alone.

2.4.7.13. FANCU/XRCC2

Mutations in *XRCC2* were identified in a 2.5-year-old patient with congenital abnormalities suggestive of FA (Shamseldin et al., 2012). Chromosome breakage tests were positive, but peripheral blood counts were normal, and aplastic anaemia was absent. Complementation studies later established *XRCC2* as *FANCU* (Park et al., 2016), acting downstream of FANCI/D2 monoubiquitination and mediating HR by promoting *RAD51* binding to double-strand DNA breaks. (Suwaki et al., 2011).

2.4.7.14. FANCV/REV7

Biallelic mutations in *REV7*, a subunit of the DNA polymerase ζ complex, were found to cause a classical presentation of FA, including aplastic anaemia, in an eight-year-old female (Bluteau et al., 2016). Impaired translesion synthesis was suggested as the cause of the disease (Lee et al., 2014). *REV7* also participates in the regulation of non-homologous end joining (NHEJ) and homologous recombination (HR)

pathways. It specifically acts to suppress HR while promoting NHEJ (Boersma et al., 2015; Xu et al., 2015). However, the precise mechanisms through which the loss of this function contributes to the FA phenotype remain unclear.

2.5. Bone marrow failure (BMF)

In Fanconi anaemia (FA), BMF, which stands for bone marrow failure, refers to a deficiency or impairment of haematopoietic stem cells (HSCs). This condition leads to bone marrow hypoplasia, resulting in single or multiple cytopenias and eventually aplasia with pancytopenia. BMF arises due to the premature and selective attrition of CD34+ HSCs, which can be observed even before the onset of cytopenias. The exact mechanisms responsible for stem cell loss are not fully understood, but they involve defective DNA repair processes leading to increased DNA damage and cell cycle arrest. Additionally, heightened levels of reactive oxygen species (ROS) and circulating inflammatory cytokines, as well as excessive damage from reactive aldehydes, play a role when the intact FA repair pathways are absent (Brosh et al., 2017; Garaycoechea and Patel, 2014; Saadatzadeh et al., 2004). BMF is further influenced by chronic activation of haematopoietic stem cells (HSCs) due to the impact of persistent stress, dysregulated oxidative stress, and exposure to inflammatory cytokines within the bone marrow microenvironment (Brosh et al., 2017).

During development, endogenous aldehydes have been observed to impact haematopoietic stem cells (HSCs) (Ridpath et al., 2007). To highlight the significance of these endogenous aldehydes, a study involving 64 patients with FA who were deficient in acetaldehyde dehydrogenase 2 reported an accelerated progression of BMF

(Hira et al., 2013). Additional factors contributing to HSC attrition include elevated levels of inflammatory cytokines, impaired redox signalling, reduced heat shock protein response, and abnormal telomere shortening (Briot et al., 2008; Cumming et al., 2001; Fagerlie et al., 2001; Haneline et al., 1998; Li and Youssoufian, 1997; Pang, 2001; Pang, Christianson, et al., 2001; Rathbun et al., 2000; Uziel et al., 2008). According to one study, BMF was observed in 80 % of 754 patients enrolled in an FA registry at the time of enrollment, and the cumulative incidence reached 90 % by the age of 40 years.

2.7. Leukaemia and malignancies

2.7.1. Haematological malignancies

The bone marrow characteristics observed in individuals with FA may resemble those found in other causes of bone marrow failure (BMF), such as aplastic anaemia or myelodysplastic syndromes (MDS). In patients diagnosed during infancy due to congenital anomalies, screening bone marrow biopsies are typically normocellular. However, as cytopenias develop, the bone marrow may show significant hypocellularity that exceeds the extent of cytopenias. While the bone marrow might exhibit erythroid dysplasia, such as hyperplastic erythroblast islands and megaloblastic features, it should not be interpreted as MDS unless other MDS-related features are present, such as increased blasts or cytogenetic abnormalities. Conversely, the presence of dysplasia in the myeloid series, elevated myeloblasts, or dysmegakaryopoiesis should be considered evidence of clonal abnormalities

consistent with MDS (Park et al., 2015; Soulier, 2011). MDS and leukemia are prevalent conditions in FA, with MDS or acute myeloid leukemia (AML) sometimes being the initial presenting findings. Lymphoid malignancies also occur, but they are less common in most FA subtypes. Compared to the general population, the risk of MDS in FA patients is estimated to be increased by 6000-fold, while the risk of AML is increased by 700-fold. The lower prevalence of leukaemia (3.1 %) is likely due to the early use of haematopoietic cell transplantation (HCT) in patients who develop MDS (Blanche P. Alter, 2014). According to a US National Cancer Institute (NCI) registry, MDS was reported in 16 % of 163 patients after a 15-year follow-up, with a cumulative incidence of MDS reaching 80 % by age 60 in patients who did not undergo HCT (Alter, Giri, Savage and Philip S Rosenberg, 2018).

Patients with biallelic mutations in *FANCD1*/*BRCA2* face a particularly high risk of leukaemia, with an 80 % cumulative incidence by the age of 10 (Alter et al., 2006). The majority of these patients develop acute myeloid leukaemia (AML), though a few may experience T-cell acute lymphoblastic leukaemia (ALL). Individuals with *FANCD1*/*BRCA2* mutations impacting the IVS7 site are particularly susceptible to an early onset of leukaemia, with the majority developing acute myeloid leukaemia (AML) by the age of three years.

In patients with FA who develop MDS or AML, karyotypic abnormalities are frequently observed. These abnormalities may include chromosome 1p translocations, monosomy 7, and chromosome 3q gains (Soulier, 2011). A review of 46 AML cases in FA patients revealed common cytogenetic abnormalities such as chromosomal gains in 1q, 3q, or 13q, along with loss of chromosome 7q (Rochowski et al., 2012). Notably,

cytogenetic lesions more commonly associated with de novo AML, such as t(8;21), trisomy 8, and inv(16), were not found in any of the patients with FA. Additionally, a study involving 53 patients identified 18 cases with 3q amplification, which was associated with shorter survival and an increased risk of AML development (Tönnies et al., 2003). Close monitoring of bone marrow and blood counts is crucial for patients with FA and any cytogenetic abnormalities. While some cytogenetic abnormalities may remain stable or become undetectable over time, the loss of part or all of chromosome 7 is more likely to indicate malignant progression.

2.7.2. Solid tumors

Individuals with FA have a higher incidence of solid tumour types, and these tumours tend to present at younger ages compared to unaffected individuals. Among the most common tumours are squamous cell cancers (SCC) of the skin, head/neck/tongue, as well as skin basal cell carcinoma, and anogenital cancers. In many cases of FA, a malignancy serves as the initial indication, often remaining undiagnosed until adulthood. While solid tumours are relatively rare in children with FA (except for patients with FANCD1/BRCA2 mutations), the risk increases with age, especially in those over 30 years old. The incidence of solid tumours is on the rise as individuals with FA have longer lifespans, partly due to the broader use of haematopoietic cell transplantation (HCT) to treat BMF and the increased risk of long-term malignancies associated with HCT.

According to the US NCI registry, 20 % of individuals with FA develop solid tumours by the age of 60, with SCC of the skin and head/neck/tongue, skin basal cell

carcinoma, and anogenital cancers being the most prevalent (Alter et al., 2018). A study following 754 patients for over 20 years found that 28 % developed solid tumours by the age of 40 (David I. Kutler, Singh, et al., 2003). The most common solid tumours were SCC of the head, neck, oesophagus, anus, and urogenital region, accounting for 49 %, followed by liver tumours (23 %), kidney (8 %), brain tumours (6 %), breast (4 %), and others like germ cell tumours and sarcomas. Similar findings have been reported in other study cohorts (Risitano et al., 2016; Rosenberg et al., 2003).

2.8. Diagnosis

2.8.1. Laboratory tests

2.8.1.1. Preliminary tests

Fanconi anaemia is a condition characterized by insufficient production of various blood cell types in the bone marrow. To diagnose this condition, a complete blood cell count (CBC) is typically performed to assess the levels of different blood cells. Additionally, a reticulocyte count is conducted to evaluate the bone marrow's ability to produce red blood cells. Confirmation of the presence of Fanconi anaemia involves examining the cells obtained from bone marrow aspiration. Since aplastic anaemia shares similarities with Fanconi anaemia, the diagnosis process also involves screening for other developmental abnormalities. X-ray examinations are employed to detect skeletal malformations, while ultrasound imaging or magnetic resonance

imaging (MRI) is utilized to visualize the formation and positioning of the kidneys and heart (Mehta and Ebens, 1993).

2.9.1.2 Chromosomal Breakage Analysis

The first test to be conducted in individuals suspected of having Fanconi anaemia is the chromosome breakage test. This assay is typically carried out in a clinical cytogenetics laboratory using a sample of the patient's peripheral blood. Lymphocytes isolated from the blood sample are subjected to DNA cross-linking agents commonly utilized for Fanconi anaemia testing, including diepoxybutane (DEB) and mitomycin C (MMC). The chromosomes are then examined for any evidence of chromosomal breakage (Auerbach, 2009; M. Castella et al., 2011). Patients without Fanconi anaemia typically exhibit only a few chromosome breaks or rearrangements. In contrast, individuals with Fanconi anaemia usually display multiple chromosomal breaks and rearrangements per cell, including complex rearrangements like radial figures. The test results report should include breakage and rearrangement rates, as well as the distribution of chromosomal breakage among cells or the average number of aberrations per cell with and without radial figures, following the guidelines of the American College of Medical Genetics and Genomics (Esplin, 2020). To ensure accuracy, it is recommended that all tests include at least two separate cultures. These cultures can involve samples treated with different concentrations of MMC or a combination of MMC and DEB. In cases where low white blood cell counts hinder setting up two cultures, a second specimen from the patient should be obtained, if possible, to verify the findings obtained from the initial culture.

2.8.1.3 Cell Cycle Analysis

While the chromosome breakage test is the primary method for diagnosing Fanconi anaemia, some laboratories may opt for cell cycle kinetics analysis instead of chromosome breakage assessment. In this approach, peripheral blood lymphocytes are treated with mitogen and DNA cross-linking agents, and their cell cycle progression is observed (Poot et al., 1994; Seyschab et al., 1995). In the absence of any DNA damage, normal lymphocytes progress through all phases of the cell cycle without experiencing significant delays. However, cells with DNA damage, such as those in Fanconi anaemia, stop at the S/G2 phase of the cycle to repair the damage before proceeding to the M phase. In individuals with Fanconi anaemia, a higher percentage of cells (generally 40% or more) will be arrested during the S/G2 phase compared to individuals without the condition. Although cell cycle analysis is not commonly used in clinical settings, the principles and flow chart utilized for the chromosome breakage test should be applied.

2.8.1.4 Immunoblotting of FANCD2 Ubiquitination

FANCD2 immunoblotting is based on the observation that during DNA damage, a complex of upstream FA gene products facilitates the ubiquitination of FANCD2, resulting in the generation of a long ubiquitinated form (D2-L). This ubiquitinated form can be differentiated from the shorter non-ubiquitinated form (D2-S) through Western blot analysis using a specific antibody targeting FANCD2. For this test, cells isolated from peripheral blood or skin fibroblasts of Fanconi anaemia patients are lysed for Western blot analysis. Fibroblasts are treated with MMC for 16

hours, harvested, and used for preparing cell lysates. The proteins are separated using a 7% tris-glycine SDS-PAGE gel and transferred onto a polyvinylidene difluoride membrane. An anti-FANCD2 antibody is used to probe for the non-ubiquitinated and ubiquitinated forms of FANCD2 (Joshi et al., 2023; Pilonetto et al., 2017; Shimamura, de Oca, John L. Svenson, et al., 2002). In control samples, a normal distribution of FANCD2 proteins with its long ubiquitinated form (FANCD2L) and a short unmodified non-ubiquitinated band (FANCD2S) will be observed. However, this test may not detect gene defects downstream of FANCD2.

2.8.1.5 Complementation assay

Complementation testing may also entail the transfer of complementary DNA (cDNA) for known FA genes into patient cells, followed by evaluating whether it corrects the FA cellular phenotype. FA patient-derived cells are transduced with retroviruses (Hanenberg et al., 2002) or lentiviruses carrying cDNAs that complement different genes or subtypes of FA. The expression of these cDNAs is facilitated by tissue-specific constitutive promoters located at the 5' end. The cells transduced with the complementing FA subtype cDNA will exhibit a phenotypic reversion to wildtype (Yamada et al., 2001). While complementation analysis is no longer routinely performed, it can be valuable in certain cases. It helps define the functional subtype (complementation group) of FA when no variants in FA genes are detected by NGS, when a variant of uncertain significance is found in an FA gene, or when variants are identified in multiple distinct FA genes. Complementation analysis can be useful in attributing causality to a specific gene variant (Mehta and Ebens, 1993).

2.8.2 Genetic Testing

Next-generation sequencing (NGS) is essential for confirming the diagnosis of FA and identifying the specific genetic defect. NGS testing is necessary for all individuals with positive chromosome breakage analysis (CBA) and FANCD2-Ub- results, as well as for patients with physical abnormalities characteristic of FA (Joshi et al., 2023). Additionally, NGS testing should be performed in patients with extremely low or zero CBA scores, as T-cells from some FA individuals may not exhibit high CBA results. Moreover, testing is essential for patients with FANCD2-Ub+ results in both T-cells and fibroblasts, as mutations downstream of the FA pathway do not affect FANCD2 monoubiquitination (Joshi et al., 2023).

NGS testing serves multiple purposes, including identifying specific genetic abnormalities, personalizing care based on genotype/phenotype correlations (e.g., cancer screening in individuals with mutations associated with an increased risk of solid tumours), and excluding other chromosomal breakage disorders (Mehta and Ebens, 1993). It is also crucial for evaluating family members as potential transplant donors and providing appropriate management for affected relatives. Additionally, NGS testing plays a vital role in prenatal testing and genetic counselling, especially since heterozygous carriers will not exhibit abnormal chromosomal breakage testing (David I. Kutler, Singh, et al., 2003). For family members of an individual with a known FA mutation, individual gene testing, such as Sanger sequencing, can be performed instead of NGS testing.

With the vast number of missense variants identified, classifying the disease relevance of genetic variants, especially those causing missense mutations, can be

challenging. While numerous pathogenicity prediction programs are in use, many of these variants' pathogenicity remains unknown. EVE pathogenicity prediction tool has proven reliable in effectively classifying missense variants in FA. EVE scores above 0.6 demonstrate a strong correlation with functional analysis, such as complementation analysis. These findings suggest that EVE pathogenicity prediction is highly efficient and can replace functional analysis of variants, such as complementation analysis, in a significant number of cases (Frazer, Notin, Dias, Gomez, Joseph K. Min, et al., 2021; Joshi et al., 2023).

2.9. Models to Study FA

Mice with targeted single deletions in various genes, such as *Fanca*, *Fancc*, *Fancd2*, and *Fancg*, display reduced long-term HSC repopulating activity and germ cell loss. They also show increased cellular sensitivity to DNA interstrand crosslinks and oxidative stress. However, these mice do not exhibit the full clinical characteristics of FA, including marrow aplasia, haematological abnormalities, and early life tumorigenesis (Chen et al., 1996; Noll et al., 2002; Pulliam-Leath et al., 2010; Suzuki et al., 1997; Wong, 2003; Zhang et al., 2010). Cells cultured from all FA mouse models demonstrate an accumulation of chromosomal aberrations when exposed to DNA cross-linking agents, indicating some level of functional conservation of the FA DNA repair pathway across species. The cells found in the spleens of the mutant mice display a significant susceptibility to the buildup of unrepaired chromosomal aberrations after being exposed to DNA cross-linking agents, along with an unusual sensitivity to IFN-γ.

Particularly, Fancc-/- mice exhibit extreme sensitivity to the DNA cross-linking agent, Mitomycin C, leading to bone marrow failure within 3-8 weeks. The data suggest that the loss of single genes in the FA pathway in mice does not compromise short-term survival but impairs the capacity to repair damage induced by environmental insults or DNA damaging agents. This implies that the loss of additional genes may be necessary to fully replicate the characteristics of human FA. To address this, several double mutant mouse models have been created to investigate processes that could enhance the development of FA **(Table 2.3)**. For instance, Fancc-/- mice do not develop bone marrow hypocellularity, but the double mutants of Fancc-/- and Sod1-/- exhibit this feature along with anaemia and leukopenia, providing evidence that oxidative stress contributes to bone marrow failure in FA (Hadjur et al., 2001). Other double mutants of Fancd2-/- and Aldh2-/- show unusual sensitivity to endogenous aldehydes in utero (Garaycoechea et al., 2012; Langevin et al., 2011). When postnatal double-deficient mice are exposed to ethanol, which serves as a source of exogenous acetaldehyde, it leads to a rapid onset of bone marrow failure syndrome (BMFS) and triggers the spontaneous development of acute leukaemia. These findings suggest that the FA pathway plays a crucial role in countering the toxicity induced by acetaldehyde.

Another promising model is the *Btbd12* knockout mouse, the ortholog of *Slx4* (*Fancp*), which mimics many features of FA, including peripheral cytopenia, reduced fertility, dysmorphic features, ocular abnormalities, hydrocephalus, chromosomal instability, accumulation of damaged chromosomes, hypersensitivity to DNA crosslinking agents, and abnormal lymphopoiesis (Crossan et al., 2011).

While data from these models are intriguing, the need to create double knockouts to partially recapitulate the phenotype observed in humans, resulting solely from mutations in the FANC genes, remains a significant challenge. Murine FA models may not be optimal tools to understand the pathophysiology of FA and develop novel treatments due to potential differences in the susceptibility of mice to sustain and retain DNA damage or the presence of alternate regulatory mechanisms for FANC proteins in humans. Additionally, the heterogeneous nature of mutations in various types of FA, including point mutations, small insertions/deletions, splicing mutations, and large intragenic deletions, makes it difficult to precisely replicate all human mutations through targeted gene knock-ins/outs in the mouse system.

Table 2.3. Summary of key murine models of inherited bone marrow failure syndrome.

Gene names	Affected systems	References
Fanca	Homozygotes displayed FA-like phenotypes including growth retardation, microphthalmia, craniofacial malformations, and hypogonadism. Homozygous females demonstrate premature reproductive senescence and an increased incidence of ovarian cysts. Homozygous males exhibit an elevated frequency of mispaired meiotic chromosomes and increased apoptosis in germ cells, implicating a role for Fanca in meiotic recombination. Fancc-/- Fanca-/- display the same phenotype as the single mutants suggesting that these two genes are epistatic.	(Noll et al., 2002; Wong, 2003)
Fancc	Homozygotes do not show developmental abnormalities or haematological defects till 9-12 months of age. Male and female mutant mice have reduced numbers of germ cells and females have markedly impaired fertility. The CFC capacity of haematopoietic progenitors is abnormal and the cells are hypersensitive to gamma-interferon. Fancc-/-Tert-/- double mutant mice have exacerbate telomere attrition when murine bone marrow cells experience high cell turnover after serial transplantation and an increase in the incidence of telomere sister chromatid exchange. Fancc-/-Fancg-/- double-mutant mice develop spontaneous haematologic sequelae, including bone marrow failure, acute myeloid leukemia, myelodysplasia, and complex random chromosomal abnormalities.	(Chen et al., 1996; Pulliam-Leath et al., 2010)
Fancd1	Homozygous null mutants are embryonic lethal with abnormalities including growth retardation, neural tube defects, and mesoderm abnormalities; conditional mutations cause genetic instability and enhanced tumor formation; mutants with truncated BRCA2 protein survive, are small, infertile, show improper tissue differentiation, and develop lymphomas and carcinomas	(Suzuki et al., 1997)
Fancd2	Homozygous mutant mice exhibit meiotic defects and germ cell loss. In addition, mutant mice display perinatal lethality, susceptibility to epithelial cancer, and microphthalmia. Homozygous mice have smaller haematopoietic stem cell pools and reduced lymphoid progenitor frequency. Fancd2-/-Aldh2-/- double homozygous mice are unusually sensitive to ethanol exposure in utero, and ethanol consumption by postnatal double-deficient mice rapidly precipitates bone marrow failure and spontaneously developed acute leukemia. Aged Aldh2-/-Fancd2-/- mutant mice	(Garaycoechea et al., 2012; Langevin et al., 2011; Zhang et al., 2010)

	which do not develop leukemia, spontaneously develop aplastic anaemia, with concomitant accumulation of damaged DNA within the haematopoietic stem and progenitor cell pool.	
Fancg	Females and males homozygous for targeted null mutations exhibit hypogonadism and reduced fertility. Cytogenetic analysis shows somatic chromosome aberrations occurrence at a higher spontaneous rate. Cells are also more sensitive to mitomycin C.	(Yang et al., 2001)
Fanci	These mice show craniofacial, vision, and eye abnormalities.	(Eppig et al., 2015)
Fancn	Homozygotes display embryonic lethality with impaired inner cell mass proliferation, impaired gastrulation, absence of the amnion, somites and tail bud, and general improper organogenesis.	(Rantakari et al., 2010)
Fancm	Homozygotes exhibit reduced female transmission, hypogonadism, premature death and increased incidence of tumors.	(Bakker et al., 2009)
Fancp	Homozygotes display exhibit preweaning lethality, reduced fertility, abnormal eye morphology, abnormal skeletal morphology, hydrocephalus, chromosomal instability, early cellular replicative senescence, and abnormal lymphopoiesis. Mutant mice are characterized by blood cytopenia, premature senescence, accumulation of damaged chromosomes, and hypersensitivity to DNA cross-linking agents.	(Crossan et al., 2011)
Fanco	Mice homozygous for a null mutation display embryonic lethality. Mice carrying a null and a hypomorphic allele have partial penetrance of male and female infertility due to defects in meiosis.	(Eppig et al., 2015)
Fancs	Homozygous null mutants are embryonic lethal with abnormalities including growth retardation, neural tube defects, and mesoderm abnormalities; conditional mutations cause genetic instability and enhanced tumor formation; mutants with truncated BRCA1 protein survive, have a kinky tail, pigmentation anomalies, male infertility, and increased tumor incidence.	(Gowen et al., 1996)

The series of studies mentioned so far strongly indicate that animal models of BMFS only serve as approximate representations of the underlying causes in humans. Consequently, there exists a clear and largely unmet necessity to establish alternative disease models to enhance our understanding of the mechanisms behind BMFS and to explore potential management approaches. Despite significant efforts, the past decades have been marked by concerning failures in developing new treatments. The crux of this predicament lies in the fact that animal disease models do not faithfully mimic human disease, leading to either ineffective drugs or unacceptable toxicities when tested in humans. For more successful drug discovery, utilizing human models affected by the disease would be preferable. However, obtaining affected tissues is often limited to deceased patients, which renders them unsuitable for testing drug toxicity and efficacy since they represent the end stages of the disease. This is particularly challenging for inherited BMFS, as ex vivo expansion of HSC from aplastic bone marrow proves to be difficult (Walasek et al., 2012).

2.10. iPSCs for Disease Modelling

Human embryonic stem cells (hESCs) and more recently, induced pluripotent stem cells (iPSCs), have the unique ability to undergo indefinite in vitro expansion while retaining their capacity to differentiate into cell types characteristic of various tissues in the developing embryo. The reprogramming of somatic cells to pluripotency by introducing four transcription factors expressed by ESC was initially reported in 2006 (Takahashi et al., 2007). These reprogrammed cells were termed "induced

pluripotent stem cells" or iPSCs by Shinya Yamanaka. Subsequently, human iPSCs (hiPSCs) were also reported in the following year by Yamanaka's group and another research team (Takahashi and Yamanaka, 2006; Yu et al., 2007). The pluripotent nature of iPSCs offers particular interest for clinical research because they can be derived from individual patients and can be disease-specific. This characteristic makes iPSCs a promising platform for studying the pathophysiology of specific diseases and for testing the effectiveness and toxicity of drugs in the relevant cells. Despite the need to address potential challenges such as the acquisition of de novo genetic alterations during reprogramming, the possible retention of certain epigenetic marks from the somatic parent cells, and aberrant DNA methylation acquired during the reprogramming process that may influence the cells' differentiation capabilities, there is a growing recognition of the value of iPSC-based disease modelling as a tool to advance drug discovery (Gore et al., 2011; Hussein et al., 2011; Kim et al., 2011; Nishizawa et al., 2016).

The differentiation of pluripotent stem cells into patient-specific haematopoietic cell types forms the foundation for modelling inherited BMFS. However, the main obstacle lies in generating fully functional haematopoietic stem cells (HSCs). While several studies have described the successful generation of haematopoietic progenitors, further differentiation in both in vitro and in vivo experimental models often reveals a bias toward myeloid differentiation. Achieving high long-term engraftment efficiency and multi-lineage differentiation of functional HSCs in the bone marrow of immune-compromised mice remains a challenging task

(Ledran et al., 2008; Narayan et al., 2006; Risueño et al., 2012; Suzuki et al., 2013; Tian et al., 2006; Vodyanik et al., 2006; Wang et al., 2005). Despite the challenges, several patient-specific iPSC lines from individuals with inherited BMFS have been successfully established. Potential applications of iPSCs for modelling blood disorders are shown in **Figure 2.3.**

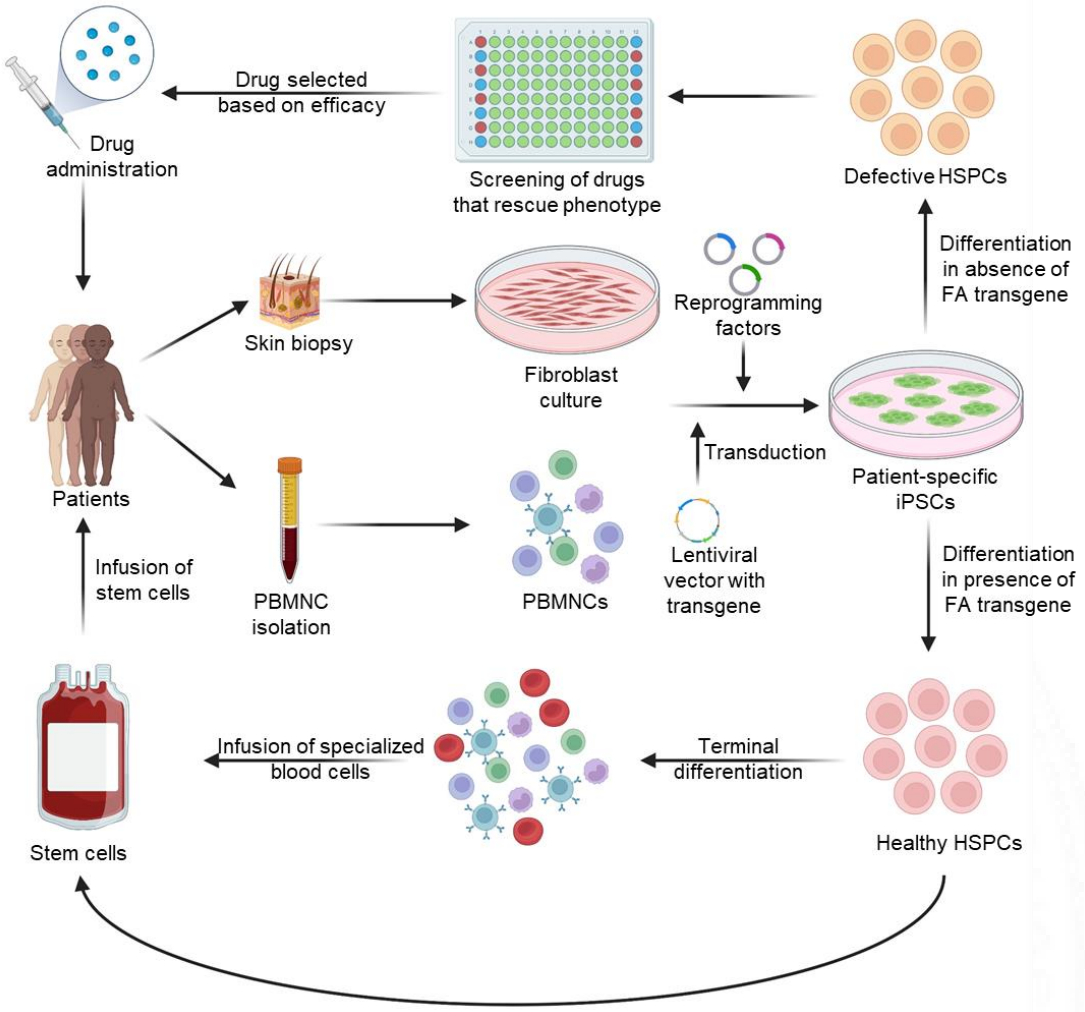


Figure 2.3: Potential applications of hiPSCs for disease modelling of genetic haematological disorders and drug discovery. Patient-specific iPSCs offer a powerful tool for understanding disease mechanisms and conducting drug screening. The use of inducible vectors for transgene expression enables the creation of patient-specific isogenic iPSC lines, which serve as ideal models for unraveling disease mechanisms. Moreover, iPSCs can serve as a valuable cell source for large-scale cell production, facilitating cell-based therapy. (Abbreviations: iPSC - induced pluripotent stem cell; HSPC - haematopoietic stem and progenitor cell.)

2.11. Disease Modelling Using FA-iPSCs.

The rarity of Fanconi Anaemia (FA), the low cellularity of patient bone marrow, and limited access to certain tissues make studying FA in primary patient cells challenging and often impractical. Transformed FA cell lines have been practical surrogates, but they may not faithfully recapitulate FA disease phenotypes due to transformation-related artifacts (Donahue et al., 2003; Schwaiger et al., 1982). Although primary patient fibroblasts are useful in studying DNA damage repair in FA, there is limited understanding of stem cell defects in FA, despite the development of multiple mouse genetic models of FA (these models do not develop anaemia except hypomorphic *Fancd1* mutation and *Btbd12* deficient mouse model) (Crossan et al., 2011; Navarro et al., 2006)).

The modelling of Fanconi Anaemia (FA) has presented challenges because effective reprogramming and induction, as well as the maintenance of pluripotency, necessitate an intact FA DNA repair pathway. The challenge in reprogramming FA cells could be overcome by introducing a wild-type cDNA of the defective FA pathway gene into the patient's somatic cells to restore the active FA pathway before reprogramming (Müller et al., 2012). However, the rescue of FA genetic defects by introducing a complementing cDNA in FA somatic cells can rescue reprogramming, but permanent repair of FA pathway mutations in iPSCs precludes the downstream derivation of FA-deficient HSPCs for disease modelling (Raya et al., 2009). Despite this limitation, patient-specific iPSCs have been generated using alternative

reprogramming approaches, including complementation and CRISPR-Cas9 mediated gene correction of patients' cells before reprogramming (Müller et al., 2012; Osborn et al., 2016; Raya et al., 2009; Yung et al., 2013). Other approaches, such as HDAdV-mediated targeted correction of FANCA defective cells before reprogramming, have been employed (GH Liu et al., 2014). However, the iPSCs derived using this approach underwent in-situ gene correction, rendering them unsuitable for studying disease phenotypes. Inhibition of p53 by HPV E6 protein has also been used for reprogramming FA patient cells into iPSCs, but these colonies could not be maintained as cell lines (Chlon et al., 2014).

iPSCs have been generated from the fibroblasts of FA patients with defects in *FANCA*, *FANCC*, and *FANCI* genes (Bharathan et al., 2017; Chlon et al., 2016; G-H Liu et al., 2014; Müller et al., 2012; Raya et al., 2009; Yung et al., 2013). However, the FA-iPSCs generated by gene complementation or gene correction could not be used for studying the disease phenotype. To address this, an inducible complementation system involving doxycycline-inducible lentiviral vectors to express normal cDNAs of the defective gene in the patient's fibroblasts has been employed (Bharathan et al., 2017; Chlon et al., 2016). In such cases, the FA fibroblasts could be reprogrammed successfully by doxycycline supplementation that activates the FA pathway. The iPSCs and differentiated haematopoietic progenitors could be used to study disease mechanisms after the inducer's withdrawal (Chlon et al., 2016; Marion et al., 2020). The inducible complementation system to generate iPSCs has been reported only for the FANCA complementation group. It is crucial to generate FA-

iPSCs for other complementation groups to understand the role of each protein in the FA pathway, especially since most proteins in this pathway have noncanonical functions that influence the FA phenotype in patients. Increasing evidence suggests that the noncanonical functions of the FA pathway proteins play a significant role in FA pathogenesis (Milletti et al., 2020).

2.12. Lentiviral Vectors for Lentiviral Transgene Expression

The first generation of lentiviral vectors (LVs) relied on three plasmids, each encoding specific components: the Envelope (Env) gene, all lentiviral open reading frames (ORFs) except Env, and the expression cassette associated with two viral long terminal repeats (LTRs) and sequences necessary for viral RNA export, genome packaging, and reverse transcription (Durand and Cimarelli, 2011; Naldini et al., 1996). Substantial modifications were subsequently introduced in later LV generations to improve both gene transfer performance and safety. In the second generation, most retroviral "accessory" genes were removed, and in the third generation, the tat gene segment was also eliminated, with the rev gene expressed on a fourth plasmid (Dull et al., 1998; Durand and Cimarelli, 2011; Zufferey et al., 1999). Despite these improvements, certain safety issues, such as the risk of insertional mutagenesis, remained associated with LV-mediated gene transfer. To mitigate this risk, self-inactivated vectors (SIN) were engineered by deleting the enhancer/promoter sequence of the U3 region from the 3′LTR (Miyoshi et al., 1998). This deletion reduces the activation of nearby genes at the vector integration site in the genome (Durand and

Cimarelli, 2011). Initially, transgene expression was controlled by the LTRs, but later vector generations and SIN vectors incorporated different internal promoters and regulatory sequences to express the transgene of interest (Benabdellah et al., 2016; Durand and Cimarelli, 2011). To express multiple transgenes from a single vector, various promoters can be combined. Nevertheless, this approach is constrained by the packaging capacity of LVs, which is approximately 10 kb, and can potentially result in promoter interferences (Benabdellah et al., 2016; Park et al., 2019; Yu et al., 2003). To overcome this limitation, IRES (internal ribosome entry site) sequences or 2A peptide sequences have been introduced in LVs. These elements induce internal cap-independent translation and ribosomal skipping mechanisms, respectively, allowing multiple transgenes to be expressed using only one promoter (multicistronic vectors).

It is important to note that many promoters commonly used in lentiviral vectors (LVs), such as the cytomegalovirus (CMV) minimal promoter, Elongation factor 1a (EF1a) promoter, or spleen focus-forming virus (SFFV) promoter, are constitutive. As a result, they do not allow for regulated transgene expression, both in terms of quantity and timing, which is essential for various applications, particularly in clinical settings.

Regulated expression of the transgene under specific conditions, such as exposure to exogenous or endogenous signal molecules or during specific time periods, would offer several advantages. Firstly, it may reduce the likelihood of inducing immune responses against the transgenic protein. Secondly, a temporary pattern of promoter activity could decrease the potential activation of proto-oncogenes near the inserted vector, thus lowering the risk of malignancies. Lastly, the ability to

not only switch the transgene expression ON or OFF but also modulate the dosage could be critical in adjusting it in response to variations in certain host parameters. Therefore, the development of more sophisticated technologies is necessary to create LVs capable of continuously monitoring physiological parameters and adjusting the output response through transgene expression to, for example, restore homeostasis.

To achieve precise regulation, inducible lentiviral (LV) systems must meet several criteria. On one hand, the inducible promoter should have low background activity in the non-induced state to avoid unintended transgene leakage, which could have serious consequences. On the other hand, under inducing conditions, the promoter activity must be sufficiently high to enable the transgene to perform its functions effectively. Within this range, the promoter activity can be fine-tuned depending on the amount of inducer, allowing for adjustable transgene expression intensity. These regulatory properties must be compatible with the host cells, ensuring that the induction does not interfere with endogenous regulatory networks and cell physiology (Kelm et al., 2004). Additionally, the inducible system must be designed to accommodate potential limitations in the viral vector's packaging capacity.

Inducible systems can utilize promoters that respond to exogenous stimuli, such as small molecules (e.g., antibiotics) or temperature variations, as well as those that sense endogenous molecules, such as pathological markers (e.g., pro-inflammatory cytokines). Alternatively, the expression level of the transgene can be indirectly adjusted by modifying the inserted transgene loci, such as through transgene excision. Overall, LVs with inducible transgene expression offer vast possibilities for

both basic and translational research and are expected to gain widespread use in the future.

Various small molecules have been explored as potential candidates to selectively activate an inducible promoter, thereby controlling transgene transcription. Among these candidates, antibiotics have emerged as promising options due to their clinical approval, minimal or negligible side effects, desired pharmacokinetic profiles, and lack of cross-talk with endogenous signalling cascades. This makes them compatible with cell networks. The most widely used system based on this approach is the Tet-system, which is activated by tetracycline or its analogue doxycycline and is compatible with lentiviral gene transfer technologies.

The Tet-system was first described in 1992 (Gossen and Bujard, 1992) and consists of two key elements: (i) the Escherichia coli Tet repressor protein (TetR), which binds to the Tet operator (TetO) and represses transcription, and (ii) tetracycline or its analogue, doxycycline. Upon the addition of the drug, a conformational change occurs within the TetR, disrupting its binding to TetO, and subsequently allowing transcription to proceed.

To induce gene expression when bound to a minimal promoter sequence derived from the human CMV promoter (Ptet) placed downstream of TetO in the absence of doxycycline, the activation domain of the herpes simplex virus VP16 protein is fused to the TetR, creating the tetracycline-controlled trans-activator (tTA) (Tet-Off system). A modification of four amino acids within this activator has also been introduced, resulting in a reverse tTA (rtTA), which requires tetracycline or

doxycycline to bind to Ptet and enable transgene transcription (Tet-On system). The choice between these systems depends on the specific application and whether the system will be primarily studied in the induced or repressed state. For gene therapy and in vivo applications, the Tet-On system is more suitable, as the Tet-Off system relies on drug clearance rates, and maintaining the off-state would require continuous doxycycline administration to suppress gene expression.

The Tet system offers numerous advantages for achieving inducible gene expression. It allows for dose-dependent and fully reversible regulation. Additionally, it utilizes well-tolerated antibiotic drugs that can be easily administered, either through intravenous injection, adding them to drinking water, or incorporating them into the diet (Markusic and Seppen, 2010) , without causing disruptions to host physiology. However, their use is still associated with certain side effects, such as the development of antibiotic resistance (Chait et al., 2016). Another concern related to this system is the immunogenic and toxic potential of tTA and rtTA, which likely stems from the VP16 domain. Over the years, several improvements have been made to enhance the system's performance and reduce side effects, especially by further engineering rtTA. For example, the rtTA2s-M2 transactivator was obtained through mutagenesis and no longer contains the VP16 domain. This construct exhibits improved stability, specificity, and inducibility (Koponen et al., 2003; Urlinger et al., 2000). In addition to optimizing the trans-activator, significant attention has been given to improving the tet-responsive promoters, leading to variants with reduced leakage and a higher dynamic range of expression. This has been achieved by truncating the CMV minimal

promoter in Ptet, using the HIV-1 long terminal repeat promoter instead of CMV in Ptet, or introducing random mutations in Ptet (Baron et al., 1995; Loew et al., 2006, 2010).

Over time, the Tet-system has been successfully integrated into lentiviral vectors (LVs) to regulate transgene expression in a wide range of cells (Koponen et al., 2003; Vieyra and Goodell, 2007; Yang et al., 2014). Initially, a "two-vectors" approach was developed, with one vector encoding the tetracycline response element fused to a promoter followed by the gene of interest, and a second vector encoding the transcriptional transactivator or trans-repressor. However, this strategy posed a challenge in selecting cells expressing the trans-activator or trans-repressor, leading to heterogeneity in cellular responses in the absence of selection. Additionally, the integration sites could influence the expression of the trans-activator or trans-repressor and the inducibility of the transgene, affecting overall regulation performance. The use of two LVs, each carrying a part of the regulation system, also raised concerns about the risk of insertional mutagenesis, which could be particularly problematic in terms of biosafety.

To address these challenges, the "two-vectors" approach has been replaced with more sophisticated "all-in-one" lentiviral vectors (LVs) that incorporate both the tetracycline regulatory element and the inducible expression cassette within a single construct. Initially, the rtTA or rTA transactivators were controlled by weak constitutive promoters, such as the phosphoglycerate kinase 1 promoter (Barde et al., 2006) or the elongation factor 1a short promoter (Giry-Laterrière et al., 2011) , with

the inducible transgene positioned downstream of a Tet response element (TRE). Different construct configurations were tested, with the TRE promoter inserted either upstream (Benabdellah et al., 2016; De Groote et al., 2016) or downstream (Kafri et al., 2000) of the constitutive promoter. Comparative analyses of various promoter architectures showed that a head-to-head orientation yielded superior results compared to tail-to-tail or head-tail organizations (Park et al., 2019). However, interferences between the constitutive promoter driving rtTA or rTA and the TRE promoter were still observed, leading to increased background levels. To address this, auto-regulated LVs were engineered. In such constructs, the rtTA or rTA were controlled by the TRE promoter and placed after IRES or T2A sequences (Centlivre et al., 2010; Huang et al., 2015; Ogueta et al., 2001). Other successful auto-regulated configurations involved placing two CMV promoters under the regulation of the same TRE (Yang et al., 2014). In these auto-regulated constructs, the leakage of rtTA or rTA expression was sufficient to amplify the system. Furthermore, the background expression levels were significantly decreased, and the kinetics of induction were improved. Tet auto-regulated LVs have exhibited functional efficacy in vivo, maintaining expression levels even after multiple rounds of doxycycline induction, thus showcasing their potential for future clinical applications (Centlivre et al., 2010). The Tet-regulation system has already been successfully used to transduce numerous cell lines and primary cells from both rodents and humans (Alexeyev et al., 2010; Benabdellah et al., 2011, 2016; Hoyng et al., 2014; Reiser et al., 2000).

3. MATERIALS AND METHODS

3.1. Chromosome breakage analysis (CBA)

T-cells were activated by incubating 0.5mL whole blood with 4.5 mL of RPMI 1640 medium containing 10% fetal bovine serum (FBS) and 60 μg/mL phytohemagglutinin (PHA) (Thermo Fisher Scientific) at 37ºC/5% CO_2. After 24 hours of incubation, the culture was supplemented with mitomycin C (MMC) at concentrations of 50 and 100 ng/mL. Following an additional 72 hours, the cells were subjected to metaphase arrest by treating them with 0.02 μg/mL colcemid for 1 hour. Subsequently, 40 well-spread metaphases exhibiting normal ploidy were selected for analysis. Ambiguous structures were carefully excluded from the analysis to ensure the accuracy and reliability of results. A conservative linear discriminant function, which ensured both 100% sensitivity and 100% specificity, was employed to calculate the CBA scores (Kuffel et al., 1997).

3.2. Isolation and culture of human dermal fibroblasts

Human dermal fibroblasts were isolated and cultured following a well-established protocol as previously described (Kotton et al., 2010). In brief, 1 mm^3 skin biopsies obtained from the patients were first washed with phosphate-buffered saline (PBS; HyClone). Subsequently, the biopsies were placed in a 15 mL conical flask and incubated overnight at 37°C/5% CO_2 in 1 mL of a skin biopsy digestion medium. This

medium consisted of DMEM supplemented with 20% FBS, 0.25% collagenase type I (ThermoFisher Scientific), 0.05% DNase I (Merck), and antibiotics (100 U/mL penicillin and 100 µg/mL streptomycin). The following day, the tube was briefly vortexed, and 7 mL of DMEM supplemented with 20% FBS was added. The dissociated cells were subsequently plated in a T25 culture flask and incubated at 37ºC (5% CO_2) for 3 days. On the fourth day, the medium was replaced with DMEM supplemented with 10% FBS and antibiotics. The cells were then cultured, with a media change performed on alternate days, until they reached approximate confluency of 80%. After reaching the desired confluency, the cells were passaged at a ratio of 1:4 using 0.05% Trypsin–EDTA (Thermo Fisher Scientific). Following two passages, the cells were cryopreserved for future use.

3.3. Western blot analysis in the PHA-stimulated T-cells, dermal fibroblasts and iPSCs

Peripheral Blood Mononuclear Cells (PBMNCs) obtained from a 5 mL blood sample were isolated by density gradient separation using Ficoll–Hypaque solution (Hyclone). Subsequently, the PBMNCs were cultured in 5 mL of RPMI medium supplemented with 10% FBS and 60 µg/mL phytohemagglutinin (PHA). Following 4 days of culture, the cells were harvested and whole-cell lysates were prepared for western blot analysis. Fibroblasts were isolated from skin biopsies of patients using the above protocol. 2×10^6 fibroblasts were treated with 0.25 µg/µL MMC for 16 h. Following the MMC treatment, the cells were harvested using trypsin–EDTA

treatment for the preparation of whole cell lysates. To prepare lysates from iPSCs, 0.5-2x10^6 cells grown on cell culture plates were treated with 2 mM hydroxyurea (HU) for 24 h. Following HU treatment, the cells were harvested by TrypLE treatment for whole-cell lysate preparation. Protein separation was performed using a 7% tris–glycine SDS-PAGE gel, and subsequently, the separated proteins were transferred onto a polyvinylidene difluoride membrane. The list of antibodies used in this study is shown in **Table 3.1**. Following incubation with the appropriate enzyme-conjugated secondary antibody, the detection was performed using an enhanced luminol-based chemiluminescence substrate (Westar Supernova, Cyanagen), and the images were captured using FluorChem M (ProteinSimple) or Chemi-Doc (Bio-Rad).

Table 3.1. Antibodies used for characterization of iPSCs and western blot analysis.

Purpose	Antibody	Company Cat#
Pluripotency markers	Rabbit anti-OCT4A	Cell Signaling Technology, Cat#2840
	Rabbit anti-NANOG	Cell Signaling Technology, Cat# 4903
	Mouse anti-SSEA4	Cell Signaling Technology, Cat# 4755
	Mouse anti-Tra-I-81	Cell Signaling Technology, Cat# 4745
Trilineage differentiation	Rabbit anti-PAX6	Cell Signaling Technology, Cat# 60433
	Rabbit anti-Brachyury	Cell Signaling Technology, Cat# 81694
	Rabbit anti-SOX17	Cell Signaling Technology, Cat# 81778
Western blot analysis	Rabbit anti-FANCA	Cell Signaling Technology, Cat# 14657
	Rabbit anti-FANCC	Antibodies.com, Cat# A8967
	Mouse anti-FANCD2	Santa Cruz Biotechnology, Cat# sc-20022
	Mouse anti-FANCF	Santa Cruz Biotechnology, Cat# sc-271952
	Rabbit anti-FANCI	Merck, Cat# ABE1817
	Mouse anti-FANCL	Santa Cruz Biotechnology, Cat# sc-137067
	Rabbit anti-FANCT/UBE2T	Proteintech, Cat# 10105-2-AP
	Mouse anti-Actin	BD bioscience, Cat# BD-612656
	Rabbit anti-Vinculin	Cell Signaling Technology, Cat# 13901s
ϒ-H2AX-FANCD2 colocalization immunofluorescence	Anti-γ-H2AXAlexa Fluor® 488	Merck millipore, Cat# 05-636-AF488
	Anti-FANCD2 DyLight 550	Novus Biologicals Cat# NB100-182R
HSPC differentiation antibodies	Anti-CD34APC-cy7	Biolegend, Cat# 343514
	Anti-CD43FITC	BD Biosciences, Cat# 555475
	Anti-CD45BV421	BD Biosciences, Cat# 563879
	Anti-CD235aPE	BD Biosciences, Cat# 555570
	Anti-CD309PE	BD Biosciences, Cat# 560494
	Anti-CD31FITC	BD Biosciences, Cat# 555445
Secondary antibodies	Goat anti-rabbit IgG Alexa-flour 488	Invitrogen, Cat# A11034
	Goat anti-mouse IgG Alexa-flour 594	Invitrogen, Cat# A11032
	Anti-mouse IgG HRP-Linked Antibody	Cell Signaling Technology, Cat# 7076s
	Goat anti-Rabbit IgG (H+L) Secondary Antibody, HRP	ThermoFisher Scientific, Cat# 32460

3.4. Exome sequencing

Exon capture was performed using different panels for the respective sample sets. SureSelect Human All Exon V5 from Agilent was utilized for 35 samples, xGen Exome Research Panel v1 from IDT for another 35 samples, SureSelect Human All Exon V5+UTR from Agilent for 42 samples, SureSelect Human All Exon V6+UTR from Agilent for additional 35 samples and a focused exome sequencing panel for 6 samples. All of these panels have 100% coverage of the coding sequence of FA genes. Sequencing was performed on an Illumina HiSeq X system, generating 2×150 bp sequence reads with 8–10 GB of data per sample.

3.5. Bioinformatics analysis of exome sequencing data and variant prioritisation

To remove the adapter sequences from the raw reads, Fastq-mcf (ea-utils-1.1.2-806) was employed. Bedtools-2.17 was utilized for calculating coverage metrics. The identification of germline variants was performed using a GATK Best Practices bioinformatics analysis, utilizing Sentieon (v201808.01). The sequences obtained underwent quality checks and were aligned to the human reference genome (GRCh37/hg19) using the Sentieon aligner. Subsequently, the alignment files were then processed using Sentieon to remove duplicates, recalibrate, and re-align indels. Gene annotation of the variants was performed using the Ensembl Variant Effect Predictor, utilizing the human gene model (release 91) as a reference. Single-

nucleotide variants (SNVs) with an overall population allele frequency greater than 3% in databases such as 1000Genome Phase 3 (Auton et al., 2015), gnomAD (v2.1), EVS, dbSNP (v151), as well as MedGenome's Indian population database (Wall et al., 2019) were excluded from the analysis. Additionally, deep intronic and intergenic variants were excluded for further analysis. The variants were prioritized sequentially, based on their reported status in disease databases (HGMD, ClinVar) as well as utilizing multiple in-silico prediction tools including PolyPhen-2, SIFT, MutationTaster2, LRT, EVE (Dong et al., 2015). Copy number variants (CNVs) were detected from the targeted sequence data using ExomeDepth (v1.1.10). This algorithm identifies CNVs by comparing the read depths of the test data with a matched aggregate reference dataset. In all samples, CNVs were evaluated to identify deletions and duplications involving one or more exons. CNVs with Bayes factors greater than 15 were considered, while those with low statistical confidence and the potential to be false positives due to the presence of pseudogenes in the genome were filtered out. The prioritized variants were classified based on the guidelines provided by the American College of Medical Genetics (ACMG). The classification of variants was performed using a combination of online tools, including Varsome and EVE (Frazer, Notin, Dias, Gomez, Joseph K. Min, et al., 2021), a recent computational method. In addition, VarSeq 2.2.0 (Golden Helix Inc., Bozeman, MT, United States), a clinical genomics interpretation and reporting platform, was employed to analyze SNVs and CNVs in the 6 samples. This additional analysis was conducted specifically for cases

in which our pipeline detected heterozygous mutations, requiring further investigation to determine disease-causing genotypes.

Table 3.2. List of primers used in the study.

Purpose	Target Gene	Primer sequence (5'-3')
Primer for cDNA synthesis of *FANCL* gene	*FANCL*	Reverse TTAAGTTTCCAGCTCTTCAC
Primer for cDNA synthesis of *FANCF* gene	*FANCF*	Reverse ATACTTTGGACACACGAAGG
Primer for cDNA synthesis of *FANCT* gene	*FANCT*	Reverse TTTCAGGTTTAAAAGATTTC
Primer for cDNA synthesis of *FANCC* gene	*FANCC*	Reverse ACCTTGAGATTACACTAACAAT
Gibson cloning of *FANCL* gene into pSJL225 vector	*FANCL*	Forward aacttcagcctgctgaagcaggctggagacgtggaggagaaccctggacccATGGCGGTGACGGAAGCGAG Reverse ctctgatcttttattctagcTCAGTGTTTCCTTCCAGACATTTTTAAGGTAATTGG
Gibson cloning *FANCF* gene into pSJL225 vector	*FANCF*	Forward aacttcagcctgctgaagcaggctggagacgtggaggagaaccctggacccATGGAATCCCTTCTGCAGCAC Reverse ctctgatcttttattctagcCTATACAGAACTGAGGCCTGC
Gibson cloning *FANCT* gene into pSJL225 vector	*FANCT*	Forward aacttcagcctgctgaagcaggctggagacgtggaggagaaccctggacccATGCAGAGAGCTTCACGTCTG Reverse ctctgatcttttattctagcCTAAACATCAGGATGAAATTTCTTTTCTATGC
Gibson cloning *FANCC* gene into pSJL225 vector	*FANCC*	Forward aacttcagcctgctgaagcaggctggagacgtggaggagaaccctggacccATGGCTCAAGATTCAGTAGATCTTTC Reverse ctctgatcttttattctagcCTAGACTTGAGTTCGCAGCTC
Gibson cloning *FANCA* gene fragment 1 into pSJL225 vector	*FANCA*	Forward aacttcagcctgctgaagcaggctggagacgtggaggagaaccctggacccATGTCCGACTCGTGGGTCCC Reverse aagacgtcagCAGGAGGTCCACAGCCATGTG
Gibson cloning *FANCA* gene fragment 2 into pSJL225 vector	*FANCA*	Forward ggacctcctgCTGACGTCTTTCTGTCAGAAC Reverse ctctgatcttttattctagcTTAGAAGAGATGAGGCTCCTG
Gibson cloning *FANCI* gene into pSJL225 vector	*FANCI*	Forward aacttcagcctgctgaagcaggctggagacgtggaggagaaccctggacccATGGACCAGAAGATTTTATC Reverse ctctgatcttttattctagcTTATTTTTTCCTTTTCTTCTTGG
Gibson cloning *FANCF* gene into pCWcas9 vector	*FANCF*	Forward aacttcagcctgctgaagcaggctggagacgtggaggagaaccctggacccATGGAATCCCTTCTGCAGCAC Reverse ctctgatcttttattctagcCTATACAGAACTGAGGCCTGC
Gibson cloning *FANCI* gene into pCWcas9 vector	*FANCI*	Forward tggctagcGCGGAGTTCTGTGATATGAGCAACA Reverse ccggatccCCCAAAGTTCACATTAACTCAGGCATTTC

Gibson cloning *FANCL* gene into pCWcas9 vector	*FANCL*	Forward cagatcgcctggagaattggGTTTCTCCGGACTTCGAG Reverse aaggcgcaaccccaaccccgCGAAATGTTGTATTCTTATTTCAGTG
qPCR for pluripotency	*OCT3/4*	Forward CCTCACTTCACTGCACTGTA Reverse CAGGTTTTCTTTCCCTAGCT
qPCR for pluripotency	*SOX2*	Forward CCCAGCAGACTTCACATGT Reverse CCTCCCATTTCCCTCGTTTT
qPCR for pluripotency	*NANOG*	Forward GCTTGCCTTGCTTTGAAGCA Reverse TTCTTGACTGGGACCTTGTC
qPCR for pluripotency	*DNMT3A*	Forward ATAAGTCGAAGGTGCGTCGT Reverse GGCAACATCTGAAGCCATTT
qpCR for housekeeping gene	*β-ACTIN*	Forward GACGACATGGAGAAAATCTG Reverse ATGATCTGGGTCATCTTCTC
Primer for detection of mutation (c.1092G>A;p.K364=) in *FANCL*	*FANCL*	Forward CATCCCTTACTCTTGCAAAACA Reverse TCACCTAGGAAATCTAGAAAAGGA
Gene dosage PCR	*UBE2T*	Forward FAM-TTGATAATCTACCAGAGGCT Reverse GGCCAGTTTACTCCCAGACA
Long amplicons for NGS Fragment 1	*FANCA*	Forward TGGTCTCTTCAGGACCAACC Reverse TGCGATAAGCCAAGATAGCA
Long amplicons for NGS Fragment 2		Forward GATCTGGATGGAGGCAACAGAGT Reverse CCCAACAAGAGATGACCGGATAC
Long amplicons for NGS Fragment 3		Forward GGGCTTTGTTTGAGGAAGTCTGTT Reverse CGAGAAGGTGAGCTTTCTGTACCA
Long amplicons for NGS Fragment 4		Forward TACCCCTAAGGATCCCAAAAGGAT Reverse AGAAGGCTCCATGCGTCTAA
Long amplicons for NGS Fragment 5		Forward GCCTCGACTGGTCTAGAAGTCC

Long amplicons for NGS Fragment 6		Reverse TGCCCAGGATCTACTAGGCCATTT
		Forward AGGGGCAGGGTAAACAATGTGAGT
		Reverse CCCTCAAGTACCACATGACCAAAC
Long amplicons for NGS Fragment 7	*FANCG*	Forward GCTCGACAGTGAGCAGAGAAAGGAT
		Reverse CTTTCACCCTGTACCCACACAGACA
Gene dosage PCR for the housekeeping gene	*Albumin*	Forward FAM-CCAGAGATTTCCCAAAGCTGA
		Forward GGACAGACGAAAGCACAGAAG

3.6. Detection of deletions using multiplex ligation-dependent probe amplification (MLPA) and gene dosage analysis

Deletions in *FANCA* were detected using MLPA. SALSA MLPA P031 and P032 Probe Mixes along with the SALSA MLPA EK1 reagent kit (MRC Holland) following the manufacturer's protocol. The analysis of MLPA results was performed using Coffalyser software (MRC Holland). To confirm the deletion identified in the *UBE2T* gene by CNV analysis, a gene dosage quantitative multiplex fluorescent-PCR was performed, as previously described (Mayuranathan et al., 2012). Briefly, a multiplex PCR was performed amplifying a 327 bp fragment from exon 7 of the *UBE2T* gene, along with a short fragment of the human albumin gene (Mayuranathan et al., 2012). The forward primers used in the multiplex PCR were fluorescently labeled **(Table 3.2)**. The resulting PCR products were then analysed by capillary electrophoresis. For accurate gene dosage analysis, the PCR was performed using 200ng of DNA and 20 cycles with 0.4 μM labeled primers. The amplified products were then separated by capillary electrophoresis using an ABI-3130 Genetic Analyser (Applied Biosystems). The resulting data were analysed using GeneMapper software version 4.0 (Applied Biosystems). To quantify the gene dosage, the peak heights of the *UBE2T* amplicon were first intranormalized. This was achieved by dividing the peak heights of the *UBE2T* amplicon by the peak heights of the Albumin gene amplicon. Subsequently, internormalization was performed by dividing the intranormalized peak height of the UBE2T amplicon from the patient sample with that

of the control sample. This allowed for a relative comparison of the gene dosage between the patient sample and the control sample.

3.7. Long-amplicon next-generation sequencing

The *FANCA* gene was amplified into 6 fragments, ranging from 7.15 to 15.25 kb, and the *FANCG* gene as a single fragment of 6.73kb from the patients' DNA samples. The amplification was carried out using GoTaq Long PCR Master Mix (Promega), following the manufacturer's protocol. Primers were designed to encompass all the exons and splice sites of the genes using Primer3 (v. 0.4.0) software (http://bioinfo.ut.ee/primer3-0.4.0/). Primer details are provided in **Table 3.2**. The amplified DNA fragments (amplicons) were then analysed by 1% agarose gel electrophoresis. Based on the ethidium bromide staining intensities of the PCR amplicons, equal amounts of amplicons were pooled and subjected to shearing. Subsequently, the libraries were prepared using TruSeq Nano DNA Library Preparation Kit from Illumina. The resulting libraries were sequenced on an Illumina NextSeq-500 sequencer, generating 2×150 bp reads at an approximate sequencing depth of 100×. The sequence data obtained were subjected to quality analysis. A minimum of 75% of the sequenced bases (Q30) were processed to generate FastQ files. To analyze the NGS data (FastQ) a pipeline developed in Galaxy (galaxyproject.org) was utilized. The brief outline of the pipeline for analysis of NGS data involved the following steps. First, a browser extensible data (BED) file was created using BED Tools to define the target region. The raw reads were then processed to trim the

adapters. Subsequently, the trimmed reads were aligned against the BED file using BWA (Galaxy version 0.7.17.1). Post-processing was performed, including the removal of PCR duplicates (Kumaran et al., 2019). Variant calling was performed using FreeBayes, and then the variants were normalized using VCF ALLELIC PRIMITIVES (Zook et al., 2019). Furthermore, the VCF files were further annotated using SnpSift (ref seq-hg19). Finally, variants were also annotated using wANNOVAR to determine the functional consequences of the mutations.

3.8. Complementation plasmids

pLX301-*FANCA*, pLX301-*FANCG*, and pLX301-*FANCC* were generated by Gateway cloning using the entry vectors pENTR223.1-*FANCA* (TransOMIC Technologies), pDONR221-*FANCG* (DNASU), and pDONR221-*FANCC* (DNASU), and the destination vector pLX301 (a kind gift from David Root, Addgene ID: 25895). Gateway cloning was carried out using the LR Clonase Enzyme mix (Thermo Scientific). *FANCL* and *FANCF* cDNAs were generated from a normal individual's blood cells using PrimeScript 1st Strand cDNA Synthesis Kit (TaKaRa) and gene-specific primers **(Table 3.2)** and amplified using PrimeSTAR GXL PCR master mix (TaKaRa) with primer sequences suitable for Gibson cloning **(Table 3.2)**. For Gibson cloning the primers were designed using NEB builder (https://nebuilder.neb.com/#!/) with overhangs complementary to the target plasmid (pCWcas9 and pSJL225). The amplified insert's 5' and 3' ends had an overlap length of 20bp with the vector's 5' and 3' ends. The inserts were then cloned into pCW-Cas9 (a kind gift from Eric Lander and

David Sabatini, Addgene ID: 50661) by Gibson assembly using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs) after removing the Cas9 gene with *BamH*I and *Nhe*I enzymes.

The vector pSJL225 expresses mCherry after doxycycline induction and was a kind gift from St. Jude Children's Research Hospital. *FANCA* and *FANCC* cDNAs were amplified from pENTR223.1-*FANCA* (TransOMIC Technologies) and pDONR221-*FANCC* (DNASU), respectively. *FANCL*, *FANCT (UBE2T)*, and *FANCF* cDNAs were amplified from blood cells of a normal individual using PrimeScript 1st Strand cDNA Synthesis Kit (TaKaRa) and PrimeSTAR GXL PCR master mix (TaKaRa). pSJL225-*FANCA*, pSJL225-*FANCC*, pSJL225-*FANCF*, and pSJL225-*FANCT* were generated by Gibson assembly using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs) after digesting pSJL225 (all-in-one tet-on inducible vector) with *Esp3I* and *BlpI* enzymes. The products generated after Gibson assembly and Gateway cloning were transformed into XL1-Blue competent cells as per the standard transformation protocol. The transformed product was then plated onto suitable antibiotic-containing LB agar plates and incubated overnight at 37^{o}C. The bacterial colonies obtained were then inoculated in LB broth with suitable antibiotics overnight in a shaker incubator at 37^{o}C and 220 rpm. The overnight cultures were then centrifuged at 4000 rpm and the pellet obtained was used for plasmid isolation as per FAVORGEN (FavorPrep plasmid extraction kit) protocol. Plasmids extracted were digested along with the original target vector using HindIII to confirm integration of insert. The insert in plasmids with correct digestion patterns was

sequenced using Sanger sequencing. Plasmids that had the correct cDNA sequence of FA transgenes by Sanger sequencing were used for the preparation of lentiviruses **(Figure 3.1)**.

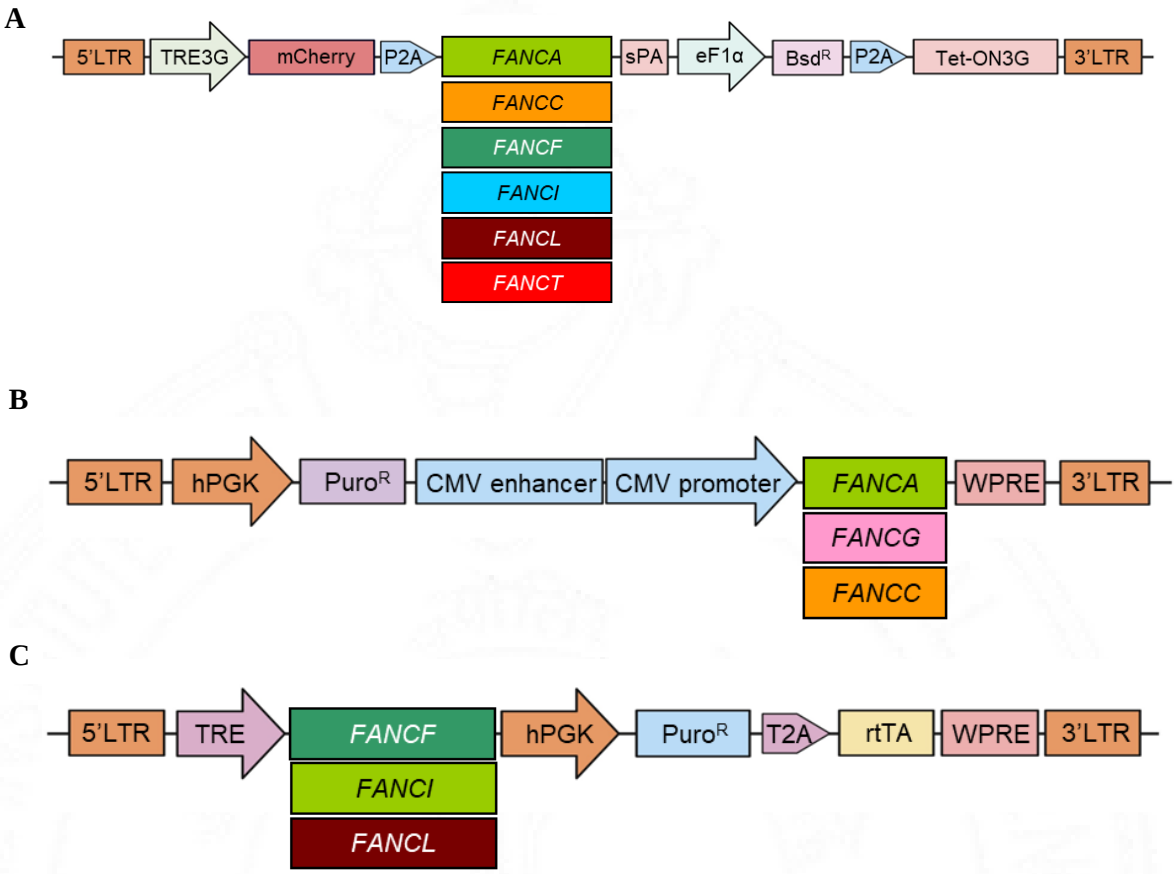


Figure 3.1: Plasmids used for lentiviral complementation of FA transgenes. **(A)** Tightly regulated doxycycline-inducible lentiviral vectors (pSJL225) for expression of *FANCA*, *FANCC*, *FANF*, *FANCI*, *FANCL*, and *FANCT* transgenes. **(B)** Lentiviral vector for constitutive expression *FANCA*, *FANCC*, and *FANCG* transgenes. **(C)** Doxycycline inducible lentiviral vector for expression of *FANCF*, *FANCI*, and *FANCL* transgenes.

3.9. Lentivirus preparation and transduction of fibroblasts

Lentiviruses were generated using a second-generation packaging system, which involved the envelope plasmid pMD2.G and the packaging plasmid psPAX2 (obtained as gifts from Didier Trono, Addgene IDs: 12256 and 12260, respectively). The lentiviral expression, envelope, and packaging plasmids were mixed in a ratio of 2:1:1 and transfected into HEK 293T cells using the TransIT-293 transfection reagent (Mirus Bio), following the manufacturer's protocol. The supernatants containing the lentiviral particles were collected at 48, 60, and 72 hours post-transfection and were concentrated using Lenti-X Concentrator (Clontech Laboratories), according to the manufacturer's protocol. To transduce fibroblasts, once they reached 70% confluency, an appropriate volume of lentiviruses in 2 mL of medium containing 6 µg/mL polybrene (Sigma–Aldrich). After 24 h, the medium was replaced. Following a total of 72 h post-transduction, antibiotics were added for the selection of transduced cells. The cells were then cultured in a medium containing the appropriate until the selection was complete.

3.10. Reprogramming plasmids

Plasmids used in reprogramming were a kind gift from Shinya Yamanaka. The reprogramming plasmids used were pCXLE-hOCT3/4-shp53-F (Addgene ID 27077), pCXLE-hUL (Addgene ID 27080) and pCXLE-hSK (Addgene ID 27078) were a kind gift from Shinya Yamanaka.

3.11. Reprogramming using Y4 episomal vectors

The patient cells were complemented with the normal cDNAs and cultured in the presence of doxycycline and nucleofected with pCXLE-hOCT3/4-shp53-F, pCXLE-hUL and pCXLE-hSK plasmids in the ratio 1:1:1 (1µg each) into 0.5 $x10^6$ patient fibroblasts using neon transfection system. The conditions used were 1650 V, 10 ms, and 3 pulses. The cells were then seeded in one well of 6 well dishes with 2mL fibroblast medium (DMEM with 10% FBS) without antibiotics. After 48 hours of transfection, medium change was done with fibroblast medium (DMEM with 10% FBS and antibiotics (100 U/mL penicillin and 100 µg/mL streptomycin)) containing 100ng/mL doxycycline. After 48 hours of transfection with Y4 plasmids, cells were continuously supplemented with 100ng/mL doxycycline in the culture. Seven days later, the cells were trypsinized and seeded at a ratio of 1:12 into matrigel coated 6 well dishes with TeSR™-E7™ (STEMCELL technologies) medium. Medium change was done every day until the emergence of iPSC colonies was observed. Colonies were picked up and passaged following standard protocol **(Figure 3.2)**.

3.12. Culture of induced pluripotent stem cells (iPSCs)

Human iPSC lines were maintained on Matrigel (BD) coated plates in mTESR plus (STEMCELL technologies). The medium was changed every alternate day. Cells were passaged once they reach ~80% confluence, at a ratio of 1:4 to 1:6 every 3 to 4 days using standard passaging techniques with 0.5mM EDTA (Gibco).

3.13. Karyotyping

Briefly, iPSC colonies in culture were exposed to 200 μg/ml colcemid for 30 min and harvested as a single cell suspension with 0.05% trypsin. The cell pellet was treated with 0.075 M KCl solution for 12 min at 37 °C and then fixed using modified Carnoy's fixative (methanol and acetic acid in the ratio 3:1), followed by centrifugation at 1000 rpm for 10 min at room temperature. The fixed cells resuspended in 5 ml of modified Carnoy's fixative were spread on a slide and stained according to standard cytogenetics protocols. Images were acquired using the AxioImager A1 (Carl Zeiss, Germany) and analysed using Ikaros Software (Metasystems, Germany). The size of the structural abnormality detected was > 3–10 mb and the band level was 400 bphs.

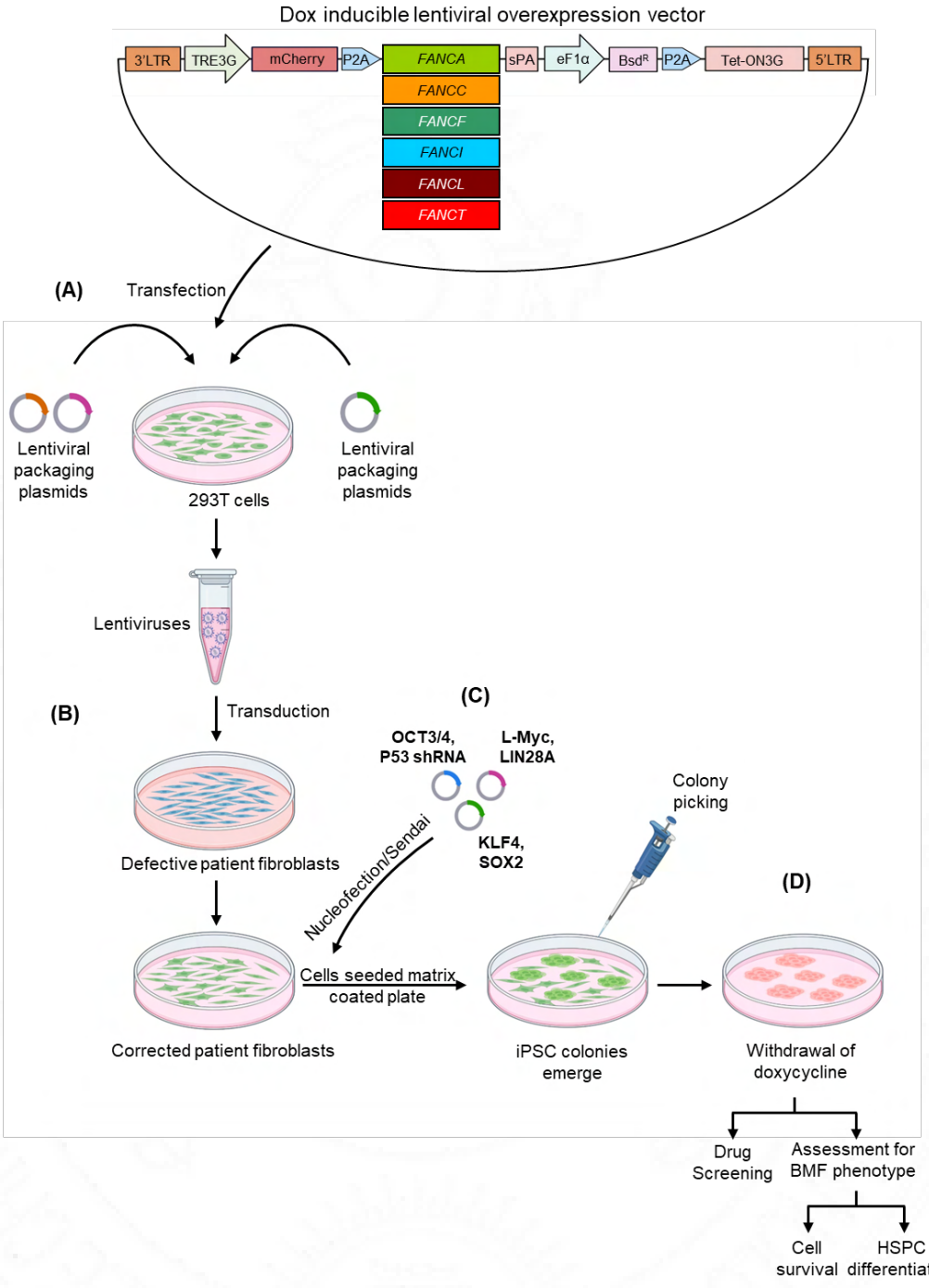


Figure 3.2: Schematic representation of the generation of iPSCs from FA patient fibroblasts. **(A)** Transfection of dox inducible lentiviral vector along with lentiviral packaging plasmids in 293T cells. **(B)** Transduction of FA pathway defective patient fibroblasts. **(C)** Nucleofection of FA pathway corrected patient fibroblasts with Yamanaka reprogramming factors or transduction with Sendai reprogramming kit. **(D)**

Withdrawal of doxycycline to study the FA pathway defects in iPSCs using various techniques.

3.14. Immunofluorescence

The iPSCs were analysed for the pluripotency markers OCT4, NANOG, SSEA4, and TRA-I-81. The fixing of iPSC colonies was performed at room temperature (RT) for 15 minutes with 4% paraformaldehyde solution in PBS. The cells were then permeabilized for 15 minutes at RT with PBS containing 0.25% Triton X-100 and then blocked with blocking buffer (PBS containing 1% bovine serum albumin (BSA)) for 30 min at room temperature. The fixed and permeabilized cells were then probed with primary antibodies in the blocking buffer for 16 h at 4 °C followed by incubation with suitable fluorescently labelled secondary antibodies for 2h at RT **(Table 3.1)**. Nuclear staining was performed using DAPI (1:400) in 1X PBS. Images were captured using a fluorescent microscope (DS6000, Leica Microsystems) with appropriate filters. For γH2AX-FANCD2 colocalization analysis, the iPSC colonies were cultured on a confocal dish and treated with 60 nM mitomycin C (MMC) for 12 h. After MMC treatment, the cells were fixed for 15 minutes with methanol, permeabilized for 15 minutes with PBS containing 0.25% Triton X-100 and blocked for 30 minutes with 1% BSA in PBS. The colonies were stained with fluorescently labelled primary antibodies (anti-γH2AX-AF488 and anti-FANCD2-dylight-550 antibodies) in PBS with 1% BSA **(Table 3.1)**. Further for nuclear staining, the colonies were incubated with DAPI (1:400) for 10 min at room temperature. Images were

captured in a confocal laser scanning microscope (FV1000, Olympus) with appropriate filters and a 100× objective.

3.15. Trilineage differentiation

In vitro, the differentiation of iPSCs into three germ layers mesoderm, endoderm, and ectoderm was carried out using STEMdiff™ Trilineage Differentiation Kit (STEMCELL Technologies) protocol. The differentiation potential of iPSCs to the three germ layers was analysed by immunofluorescence (PAX6 for ectoderm, BRACHYURY for mesoderm, and SOX17 for endoderm). Antibodies used for immunofluorescence are listed in **(Table 3.1)**.

3.16. Cell cycle analysis

The iPSC lines from patients with *FANCA*, *FANCC*, *FANCF*, *FANCI*, *FANCL*, and *FANCT* cultured in the presence and absence of DOX were analysed using ClickIT EdU flow cytometry assay kit (Life Technologies). iPSCs from all the complementation groups at ~70% confluency were treated with EdU for 1 hour at a final concentration of 100 μM. Single cells obtained after the treatment with TrypLE Express (Life Technologies) were fixed for 15 minutes with 4% paraformaldehyde and permeabilized for 15 minutes using a buffer containing saponin (provided in the kit). Subsequently, the cells were treated with ClickIT reaction buffer containing Alexaflour 647 Azide for 30 min at room temperature and finally stained with DAPI at 4 °C for 1 hour. Finally, the cells were resuspended in PBS, acquired on CytoFLEX

LX Flow Cytometer, and cell cycle analysis was performed using FlowJo, LLC software.

3.17. qPCR analysis for the transgene and pluripotency gene expression

RNA was extracted from iPSCs clones using NucleoZOL (Macherey-Nagel). According to the manufacturer's instructions, reverse transcription was performed with PrimeScript™ RT Reagent Kit (TaKaRa) with 1 μg of total RNA. SYBR Premix Ex Taq II (Takara Bio) with specific primers listed in **Table 3.2**. was used for quantitative RT-PCR and analysed with QuantStudio™ 5 Real-Time PCR System (ThermoFisher Scientific) real-time PCR systems. Relative gene expression values were determined using the ΔΔCT method with the β-Actin gene as the housekeeping gene. The list of primers is given in **Table 3.2**.

3.18. Alkaline phosphatase staining

The FA-iPSC lines cultured in the presence or absence of DOX were analysed by alkaline phosphatase staining. The staining was performed using Leukocyte alkaline phosphatase kit (86R, Sigma) following the manufacturer's protocol. Briefly, the colonies were fixed using a fixative containing PBS, acetone, and formaldehyde at a 25:65:10 ratio for 20 s. The fixed colonies were rinsed with distilled water and then incubated with a dye mix containing SODIUM NITRITE solution, FRV alkaline solution, and NAPHTHOL AS-BI at a 1:1:1 ratio for 30 min in the dark. Following incubation, the whole well of the culture plate was imaged using a digital camera.

3.19. Differentiation of iPSCs into haematopoietic progenitors

FA-iPSCs were differentiated for 12 days using the STEMdiff™ Haematopoietic Kit (05310, STEMCELL Technologies, Inc.). iPSCs were passaged using 1X EDTA and 80-100 clumps were seeded per well of 12 well plates. The cells were maintained in the mTESR^{TM} Plus medium with and without doxycycline. On D0 of differentiation, haematopoietic differentiation medium A (SDA) was added to promote mesodermal differentiation and a half medium A change was done on D2. On D3 of differentiation, 60,000 cells were replated in tissue culture-treated plates with SDA. On day 4 of differentiation medium was completely changed to medium B (SDB), followed by half-medium change till day 11. Haematopoietic stem and progenitor cell markers CD34, CD45, CD43, and CD235a were analysed in suspension cells on day 12 of differentiation **(Table 3.1)**.

3.20. Colony formation assay

A haematopoietic colony formation assay was performed using MethoCult (STEMCELL technologies) as per the manufacturer's protocol. Briefly, 14000 cells were resuspended in Iscove's Modified Dulbecco's Medium (IMDM) with 2% FBS and mixed with 3 ml aliquot of Methocult H4435 and Methocult H4636. Following the resuspension of cells, equal volumes of MethoCult with cells were added to individual 30 mm non-treated dishes in duplicate. To achieve a consistent cell distribution, the dishes were gently rotated, ensuring the uniform dispersion of the cell

suspension throughout each dish. The duplicate plates were then kept in a 100 mm plate along with a water dish containing 3 ml sterile water and incubated at 37^{o}C and 5% CO_2 for 14 days. The colonies obtained (CFU-E, BFU-E, CFU-G, CFU-M, CFU-GM, and CFU-GEMM) were counted and imaged using the EVOS FL Auto imaging system (Thermo Fisher Scientific).

3.21. Statistics

Pearson's chi-square test was used to compare group differences in discrete variables. The data analyses were conducted using Microsoft Excel, and GraphPad Prism. An unpaired Student t-test was applied, with a significance cutoff of 0.05, unless specified otherwise. The results are reported as the mean ± standard error of the mean (SEM), unless stated otherwise.

4. RESULTS

4.1. Genotyping of Fanconi Anaemia (FA) patients.

Fanconi Anaemia (FA) is a rare genetic bone marrow failure disorder, with an estimated incidence of 1 per 360,000 live births (Nicholas E Mamrak et al., 2017). This disease is caused by genetic alterations in any of the 22 genes of the FA DNA repair pathway (Bogliolo et al., 2020). The pathway consists of core complex proteins encoded by the FA upstream pathway genes (*FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL*, and *FANCM*), responsible for the monoubiquitination of FANCD2/I complex. This complex then recruits proteins encoded by the downstream pathway genes (*BRCA2, BRIP1, PALB2, RAD51C, SLX4, ERCC4, RAD51, BRCA1, XRCC2, REV7,* and *RFWD3*) to facilitate the repair of DNA interstrand crosslinks (ICLs). FA is primarily an autosomal recessive disorder, except for *FANCB*, which exhibits X-linked inheritance, and *FANCR*, which demonstrates autosomal dominant inheritance. (Niraj et al., 2019; Toksoy et al., 2020).

Bone marrow failure, which leads to pancytopenia, is the prevailing phenotype in FA, observed in approximately 90% of patients. The onset and duration of this symptom can vary among individuals, resulting in variable latency periods before its manifestation (Dokal, 2006). Around 75% of FA patients exhibit characteristic congenital physical abnormalities. These distinctive physical features are commonly observed in individuals with FA (Alter and Giri, 2016). Approximately 20% of FA

patients develop acute myeloid leukaemia (AML), while approximately 30% develop solid tumours. By the age of 50, around 40% of individuals with FA experience myelodysplastic syndrome (MDS). These conditions represent additional risks and complications associated with FA (Blanche P Alter, 2014). In certain cases, malignancies may be present in patients before the underlying FA is diagnosed (Alter, Giri, Savage and Philip S. Rosenberg, 2018). Furthermore, FA patients have been found to have reduced tolerance to standard doses of DNA-damaging chemotherapy drugs commonly used in the treatment of other haematological diseases and as part of the conditioning regimen for curative allogeneic hematopoietic cell transplantation (allo-HCT). Distinguishing idiopathic aplastic anaemia (AA) from patients with FA who do not present physical abnormalities can be challenging. Therefore, obtaining an accurate laboratory diagnosis of FA becomes crucial (Chirnomas and Kupfer, 2013; Dokal, 2006).

Chromosomal breakage analysis (CBA) is the predominant and extensively utilized diagnostic test for FA (Oostra et al., 2012). However, some rare diseases such as Nijmegen breakage syndrome (Gennery et al., 2004; Nakanishi et al., 2002; New et al., 2005), Roberts syndrome, and Warsaw Breakage Syndrome (van der Lelij et al., 2010) can also cause chromosomal breakage and pose false positives. While increased G2/M cell-cycle arrest in FA cells treated with interstrand crosslink (ICL) agents can serve as a marker of FA (Schindler et al., 2007), it is also observed in AML cells (Ceccaldi et al., 2011). Complementation analysis, a method that involves the use of viral vectors to express wild-type cDNAs of FA proteins in FA cells and subsequently

correct the cellular phenotype, can be a time-consuming process, (Chandra et al., 2005) taking approximately 4–5 weeks to establish a diagnosis. Pathogenic variants leading to defects in the FA core complex proteins, observed in over 90% of FA patients (Nicholas E Mamrak et al., 2017; Niraj et al., 2019), result in the failure to monoubiquitinate the short form of FANCD2 (FANCD2S) to its active long-form (FANCD2L) (Niraj et al., 2019). While the analysis of defective FANCD2 ubiquitination (FANCD2-Ub) is a reliable test (Shimamura, de Oca, John L Svenson, et al., 2002), it is not commonly employed for the diagnosis of FA. As somatic mosaicism is observed in around 25% of FA patients, relying solely on blood cell analysis such as CBA and FANCD2-Ub analysis can result in false-negative outcomes (Lo Ten Foe et al., 1997). To overcome this limitation and enhance the accuracy of these tests, non-hematopoietic cells, such as fibroblasts, are preferred for conducting these analyses (Fargo et al., 2014).

The identification of defective genes and pathogenic variants is essential for detecting carriers and facilitating prenatal diagnosis of FA in affected families, as well as establishing a correlation between the genotype and phenotype in patients. A few whole exome sequencing (WES) studies have been carried out to determine the frequencies of defective genes and the spectrum of mutations in populations (Bogliolo et al., 2020; Donovan et al., 2020; George, Solanki, Chavan, Rajendran, Raj, Mohan, Nemani, Kanvinde, Munirathnam, Rao, Radhakrishnan, Prasada, et al., 2021; Nie et al., 2020). Strong associations between malignancies and biallelic pathogenic variants in *FANCD1*/*BRCA2* and *FANCN*/*PALB2* (Blanche P Alter, 2014; David I Kutler et al.,

2003; Reid et al., 2007; Wagner et al., 2004; Xia et al., 2007) and monoallelic variants in *FANCS*/*BRCA1*, *FANCJ*/*BRIP1* and *FANCO*/*RAD51C* (Niraj et al., 2019; Van der Heijden et al., 2004) have been established. Genotyping a large number of patients from different populations, especially those with high consanguinity rates, helps better comprehend the genotype-phenotype correlation (Blanche P Alter, 2014; Maria Castella, Pujol, Callén, Trujillo, Casado, Gille, Lach, Auerbach, Schindler, Benítez, Porto, Ferro, Arturo, et al., 2011; Faivre et al., 2000; Neveling et al., 2009).

This chapter describes an approach to establishing a comprehensive diagnosis of FA. In accordance with predefined inclusion criteria, individuals who exhibited clinical symptoms consistent with FA were identified and enrolled. The selection criteria encompassed various factors such as haematological abnormalities, growth retardation, physical malformations, and a positive family history of FA. By employing such stringent criteria, a cohort of patients truly representative of FA was ensured, maximizing the relevance and reliability of subsequent analyses. To ascertain the presence of FA and further characterize its genetic basis, several diagnostic tests were employed. The chromosome breakage test, a hallmark examination for FA diagnosis, was performed on patient samples. This test involved subjecting patient cells to agents that induce DNA damage, thereby revealing the characteristic hypersensitivity of FA cells to such stressors. By quantifying the extent of chromosome breakage and aberrations, the presence of FA and its subtypes could be identified, providing crucial insights into the genomic instability associated with the disease. In addition to the chromosome breakage test, the FANCD2-Ub test was

conducted. This assay aimed to evaluate the activation of the Fanconi Anaemia pathway by assessing the ubiquitination status of the FANCD2 protein, a key component of the pathway. Defects in FANCD2-Ub serve as a crucial diagnostic marker for FA. To further elucidate the genetic landscape of FA in the selected patients, exome sequencing was employed. By sequencing and analysing these regions, potentially pathogenic variants and mutations responsible for FA could be identified, shedding light on the specific genes and genetic mechanisms involved in the disease. Lastly, complementation analysis, a functional assay, was performed to validate the identified genetic variants and determine the specific complementation group associated with each patient. This analysis involved introducing patient cells to genetically complementing cell lines, assessing their ability to restore normal cellular function and overcome the characteristic cellular defects observed in FA. By matching patients to specific complementation groups, a more precise genotype-phenotype correlation could be established, facilitating personalized treatment strategies and genetic counselling.

4.1.1 Clinical Phenotypes of FA Patients

A total of 153 patients (61 females and 92 males) were chosen for inclusion in the molecular characterization of FA. Notably, a substantial proportion of the patient cohort (77.8%) were from the southern states of India, underscoring the regional representation within the study. Specifically, the majority of patients originated from Andhra Pradesh (28.1%), followed closely by Tamil Nadu (24.8%) and Kerala

(20.9%). This distribution highlights the prevalence of FA in the southern regions of India and provides valuable insight into the geographical distribution and potential genetic factors contributing to the disease within this population. **(Figure 4.1.1)**.

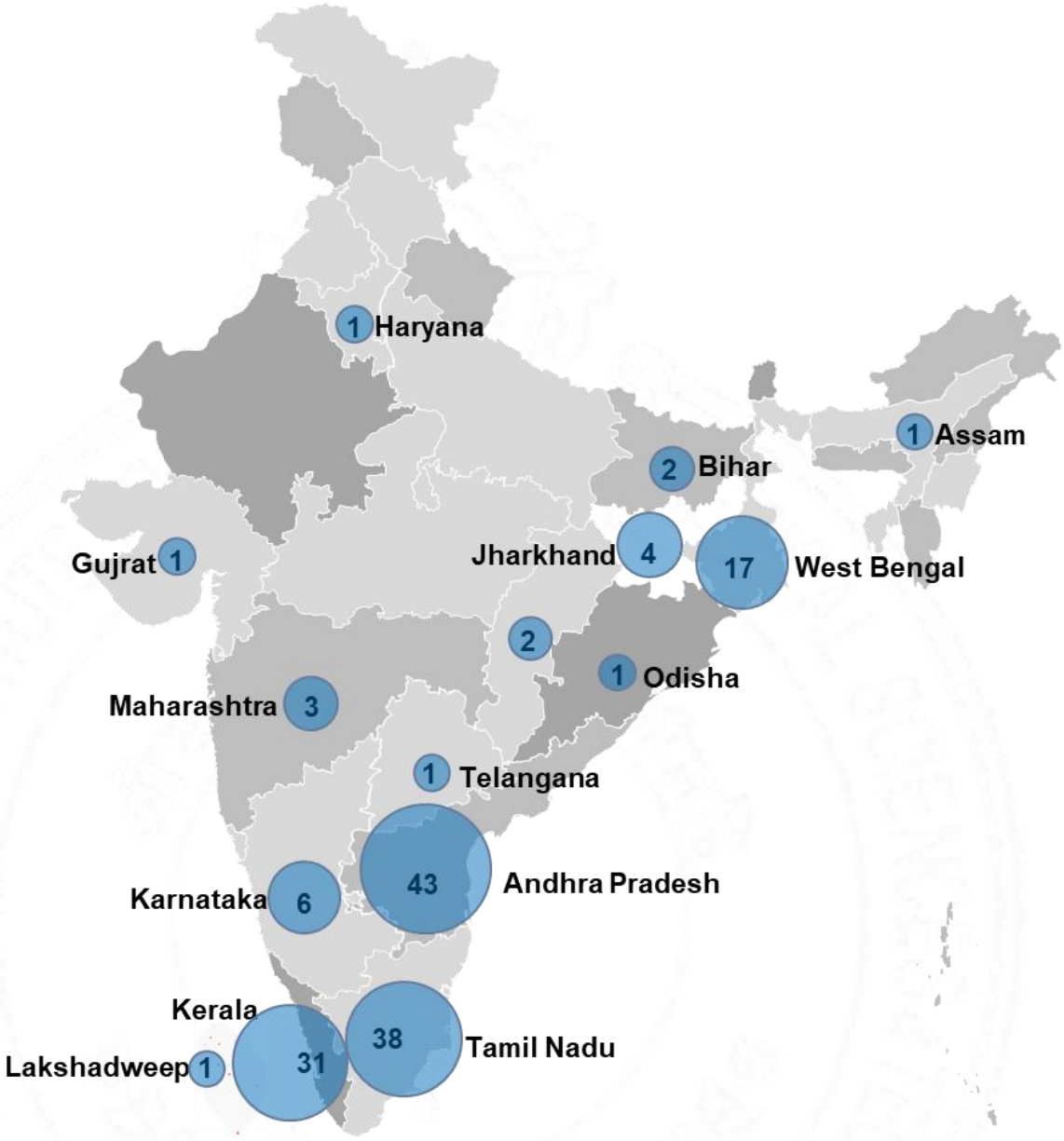


Figure 4.1.1. Geographical distribution of FA patients included in the study.

In our cohort, the median age at the time of diagnosis was found to be 9 years, encompassing a wide range from infancy (0 years) to adulthood (45 years). The predominant phenotype observed among the patients was cytopenia, as evidenced by abnormal peripheral blood cell counts, which was identified in an overwhelming majority of cases, specifically 140 patients (98.6%). Additionally, 126 patients (88.8%) exhibited the characteristic feature of hypocellular bone marrow, further contributing to the comprehensive clinical picture of FA in our study cohort. These findings underscore the haematological manifestations that frequently accompany FA, highlighting the importance of close monitoring and early detection of these clinical indicators in the diagnostic evaluation of FA patients. Distinctive physical abnormalities associated with Fanconi Anaemia (FA) were apparent in a significant majority of the patients, with 146 individuals (95.4%) displaying these characteristic features. Among these patients, skin pigmentation anomalies were the most prevalent, observed in 125 out of 153 cases (81.7%). Additionally, radial ray abnormalities, indicative of skeletal deformities, were present in 71 patients (46.4%), while both short stature and microcephaly were identified in 44 individuals each (28.7%). These physical manifestations serve as important diagnostic indicators for FA, reflecting the systemic impact of the underlying genetic abnormalities. **(Table A1)**. Among the patient cohort, a notable occurrence of haematological malignancies was observed in 29 individuals (18.9%). The median age of these patients at the time of diagnosis was 18 years, with a range spanning from 7 to 45 years. Specifically, within this subset, one patient presented with Acute Myeloid Leukaemia (AML), while 23 patients (15%)

were initially diagnosed with Myelodysplastic Syndrome (MDS). Notably, during the follow-up period, five patients initially diagnosed with MDS experienced disease progression and subsequently developed AML. These findings emphasize the heightened risk of haematological malignancies in FA patients and underscore the importance of ongoing monitoring and timely interventions to mitigate the progression and impact of these malignancies in affected individuals. A significant proportion of the patient cohort, comprising 97 individuals (63.4%), underwent treatment with androgen therapy, specifically utilizing medications such as danazol and stanozolol. Furthermore, 43 patients (28.1%) received an allogeneic hematopoietic stem cell transplant (allo-HCT) as part of their treatment regimen. Allo-HCT, a well-established therapeutic approach, involved the infusion of healthy donor stem cells to replace the diseased bone marrow and reconstitute the hematopoietic system. This procedure aimed to provide patients with a healthy source of blood cells and potentially halt the progression of FA-related complications. These treatment modalities, androgen therapy, and allo-HCT, represent distinct approaches employed to address the diverse clinical manifestations and management of FA patients. The selection of treatment strategies is typically based on individual patient characteristics, disease severity, and consideration of potential risks and benefits. **(Table A1)**. A notable portion of the patient cohort, specifically twenty individuals (13.1%), had a familial history of FA within their immediate families. This indicates that these patients had one or more family members who had previously been diagnosed with FA. The presence of familial cases highlights the hereditary nature of FA and suggests a potential genetic

predisposition within these families. The recognition of familial cases is vital for understanding the inheritance patterns, facilitating genetic counselling, and providing valuable insights into the underlying genetic mechanisms contributing to the development of FA.

4.1.2. Chromosome Breakage and FANCD2-Ub Analysis in FA Patients: Comparative Analysis

To evaluate and compare the sensitivity of diagnostic tests in differentiating FA cases from non-FA cases, we conducted complementary assays, namely Chromosome Breakage Analysis (CBA) and FANCD2-Ubiquitination analysis. The representative images of these tests are depicted in (**Figure 4.1.2A**) for CBA and **(Figure 4.1.2B)** for FANCD2-Ub analysis. CBA, a widely recognized diagnostic tool for FA, assesses the propensity of patient cells to exhibit chromosomal breakage and aberrations when subjected to DNA-damaging agents. The extent of chromosome breakage observed in FA cases is typically higher compared to non-FA cases, highlighting the diagnostic sensitivity of this assay.

On the other hand, FANCD2-Ub analysis provides insight into the functional activation of the FA pathway. The absence or impairment of FANCD2 ubiquitination is indicative of FA, as this process is disrupted in FA patients. This analysis serves as an additional measure to distinguish FA cases from non-FA cases, offering complementary information to aid in accurate diagnosis. By comparing the results of

these two tests, we aimed to determine their respective sensitivities in identifying FA cases.

CBA was conducted on peripheral blood (PB) samples obtained from 149 out of the 153 FA patients in our study. Unfortunately, four cases could not be included in the analysis due to the unavailability of samples. Consistent with expectations, the FA patients demonstrated significantly higher CBA scores, with a median score of 55 (ranging from 0 to 232.4), as compared to the 72 idiopathic AA patients. These idiopathic AA patients, who did not exhibit the characteristic physical abnormalities associated with FA, showed a median CBA score of 4.1 (ranging from 4.1 to 98.9). Furthermore, CBA scores were notably lower among the 83 normal individuals included in the study, with a median score of 0 (ranging from 0 to 8.2). These findings substantiate the sensitivity of CBA as a diagnostic tool, as it effectively distinguished FA patients from both idiopathic AA patients and normal individuals, emphasizing the significance of chromosomal breakage patterns in FA diagnosis and providing valuable insights for accurate patient stratification and management. **(Figure 4.1.2C)**.

Table 4.1.1. FANCD2-Ub status, genotypes, and clinical characteristics of the patients with very low CBA scores (0-10)

ID	CBA#	FANCD2 Ub		DNA change	Gene	Zygosity	Skeletal abnormalities	Facial dysmorphism	Skin changes	Other abnormalities
		PB	Fib							
FA-07	0	FANCD2-Ub-	NA	NC_000002.11:g.58387243C>T	*FANCL*	Hom	NA	MCH, frontal blossing	HP, HPOP	NA
FA-29	0	FANCD2-Ub+	FANCD2-Ub+	NC_000003.11:g.10076177A>C	*FANCD2*	Het	NA	NA	NA	HSK, kidney defects
FA-39	0	FANCD2-Ub+	FANCD2-Ub+	NC_000009.11:g.97873815G>A	*FANCC*	Het	Thenar hypoplasia	NA	HP	NA
FA-523*	0	FANCD2-Ub+	FANCD2-Ub-	NC_000002.11:g.58387243C>T	*FANCL*	Hom	NA	MCH	CAL, HP	NA
FA-525	0	NA	FANCD2-Ub-	NC_000002.11:g.58387243C>T	*FANCL*	Hom	Syndactyly, Thumb placed distally	NA	HP, HPOP	NA
FA-528	0	FANCD2-Ub+	FANCD2-Ub+	NC_000017.10:g.79516335G>T	*C17orf70*	Het	SS	MCH	NA	NA
FA-532	0	NA	FANCD2-Ub-	NC_000002.11:g.58387243C>T	*FANCL*	Hom	NA	DF	HP, HPOP	NA
FA-535	0	FANCD2-Ub-	NA	NC_000002.11:g.58387243C>T	*FANCL*	Hom	SS	MGH	HP	NA
FA-538	0	FANCD2-Ub-	NA	NC_000002.11:g.58387243C>T	*FANCL*	Hom	NA	MCH	CAL	NA
FA-541	0	NA	FANCD2-Ub-	NC_000006.11:g.35423588G>A	*FANCE*	Het	NA	NA	NA	NA
FA-561	0	FANCD2-Ub-	NA	NC_000016.9:g.89833551_89833647del	*FANCA*	Hom	NA	Hypertelorism	HP	NA

FA 33/22	0	NA	NA	NC_000002.11:g.58387243C>T	*FANCL*	Hom	Nail dystropy	NA	Patent Ductus Arteriosus	NA
FA-30/19	0	NA	FANCD2-Ub-	NC_000002.11:g.58387243C>T	*FANCL*	Hom	NA	MCH	HP, HPOP	NA
O-123	0	FANCD2-Ub-	FANCD2-Ub-	NC_000002.11:g.58387243C>T	*FANCL*	Hom	BT	MCH	POH, HPOP, CAL	NA
FA-554	4.1	FANCD2-Ub-	NA	NC_000016.9:g.89857810C>G NC_000016.9:g.89803957_89806507del	*FANCA*	CH	Bilateral clinodactyly	NA	CAL, HP, Pallor	NA
FA-650/18	4.1	FANCD2-Ub-	NA	NC_000009.11:g.35079204dup	*FANCG*	Hom	SS	NA	HP	NA
FA01/19 *	8.2	FANCD2-Ub+	FANCD2-Ub-	NC_000002.11:g.58387243C>T	*FANCL*	Hom	NA	NA	CAL, HP	NA

SS: short stature; DF: dysmorphic face; HP: hyperpigmentation; MCH: microcephaly; HPOP: hypopigmentation; MGH: micrognathia; CAL: café au lait; HSK: horse-shoe kidney.

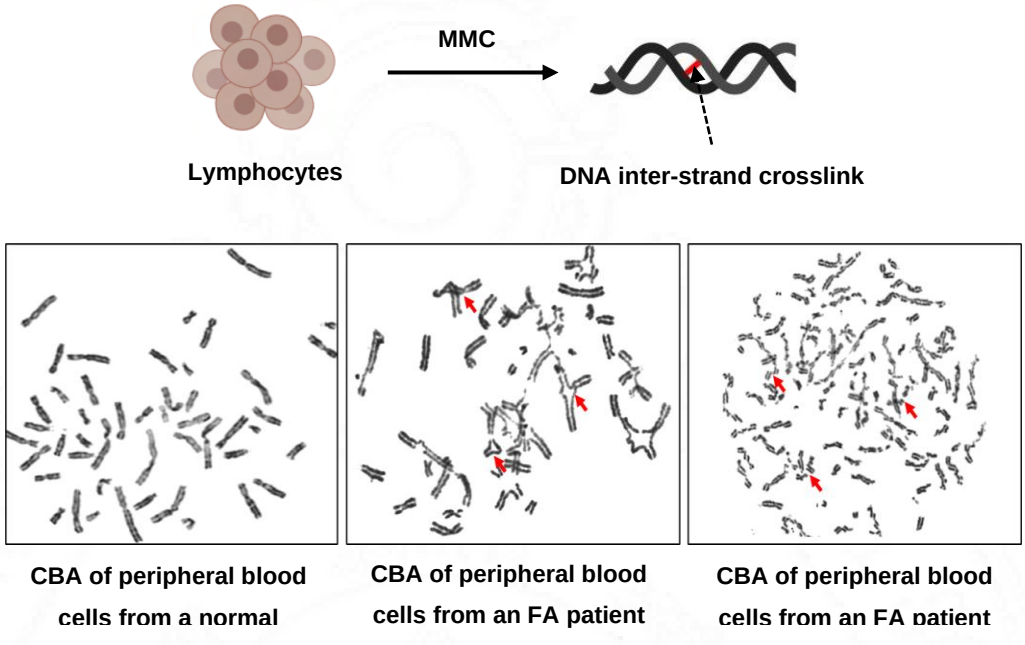
* Mosaicism detected by comparing the FANCD2-Ub analysis in the PB cells and fibroblasts.

CBA score: chromosome breakage analysis score

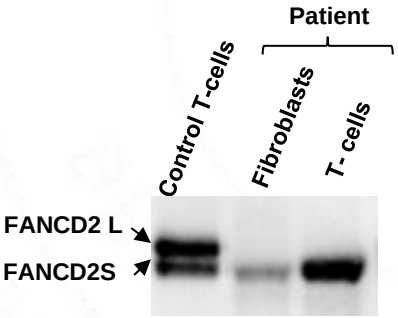
Among the 149 FA patients who underwent CBA, a subset of 17 individuals (11.4%) exhibited remarkably low CBA scores (median: 0; range: 0–8.2). Notably, these scores overlapped with those observed in both normal controls and AA patients. This finding indicates a potential challenge in accurately distinguishing these FA patients with very low CBA scores from individuals without FA or those with idiopathic AA, solely based on CBA results. The presence of overlapping scores emphasizes the need for comprehensive diagnostic approaches that incorporate additional tests and clinical evaluation to ensure accurate classification and management of these patients. **(Figure 4.1.2C** and **Table A1)**. Further analysis was conducted on these patients due to the presence of FA-related physical abnormalities in addition to pancytopenia. FANCD2-Ub analysis was conducted for a total of 142 FA patients in our study. The test was performed using different sample types depending on availability. FANCD2-Ub analysis was carried out on peripheral blood (PB) cells of 54 patients, fibroblast cells of 33 patients, and both PB cells and fibroblast cells of 55 patients. **(Figure 4.1.2B** and **Figure 4.1.2D)**. Lack of FANCD2 ubiquitination (FANCD2-Ub-) was detected in a significant proportion of patients within each group. Specifically, among the patients analysed using PB cells, FANCD2-Ub- was observed in 52 out of 54 cases (96.3%). Similarly, in the group analysed using fibroblast cells, FANCD2-Ub- was found in 31 out of 33 patients (93.9%). In the subset of patients for whom both PB cells and fibroblast cells were examined, 52 out of 55 individuals (94.5%) exhibited FANCD2-Ub-. These results highlight the functional impairment of the FA

pathway, specifically affecting FANCD2 ubiquitination, across different cell types in the majority of patients. To investigate the presence of mosaicism among the patients, we conducted a comparative analysis of FANCD2-Ubiquitination (FANCD2-Ub) in both T cells and fibroblasts from 55 patients. Remarkably, only 3 out of the 55 patients exhibited FANCD2-Ub+ in their peripheral blood (PB) cells, indicating normal FANCD2 ubiquitination, while simultaneously demonstrating FANCD2-Ub- in their fibroblast cells, suggesting a lack of FANCD2 ubiquitination. This discrepancy in FANCD2-Ub status between PB cells and fibroblasts highlights the presence of mosaicism in these particular patients **(Figure 4.1.2D)**. Among these patients, one individual was lost to follow-up, and two patients presented with Myelodysplastic Syndrome (MDS). These two patients were specifically evaluated for FA due to their marginally elevated CBA scores and the presence of FA-associated physical abnormalities. In our study, a significant proportion of patients from the Indian population, specifically 135 out of 142 individuals (95.1%), demonstrated a lack of FANCD2 monoubiquitination (FANCD2-Ub-) **(Figure 4.1.2D)**. This observation strongly suggests that FA cases within the Indian population predominantly arise due to pathogenic variants in the upstream pathway genes of FA. NGS analysis confirmed the presence of pathogenic variants in the upstream genes of the FA pathway, including *FANCA, FANCC, FANCE, FANCF, FANCG, FANCI, FANCL*, and *FANCT*, in the patients exhibiting FANCD2-Ub-.

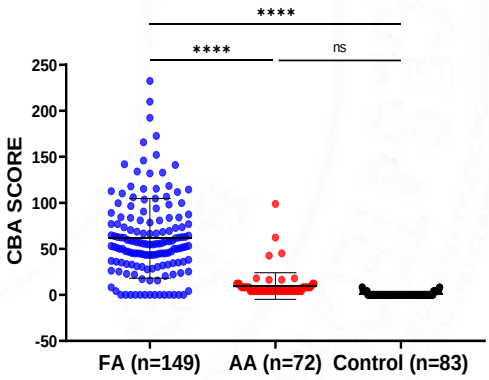
A


MMC
Lymphocytes
DNA inter-strand crosslink
CBA of peripheral blood cells from a normal
CBA of peripheral blood cells from an FA patient
CBA of peripheral blood cells from an FA patient

B


Patient
Control T-cells
Fibroblasts
T- cells
FANCD2 L
FANCD2S

C


****
****
ns
CBA SCORE
250
200
150
100
50
0
-50
FA (n=149)
AA (n=72)
Control (n=83)

D

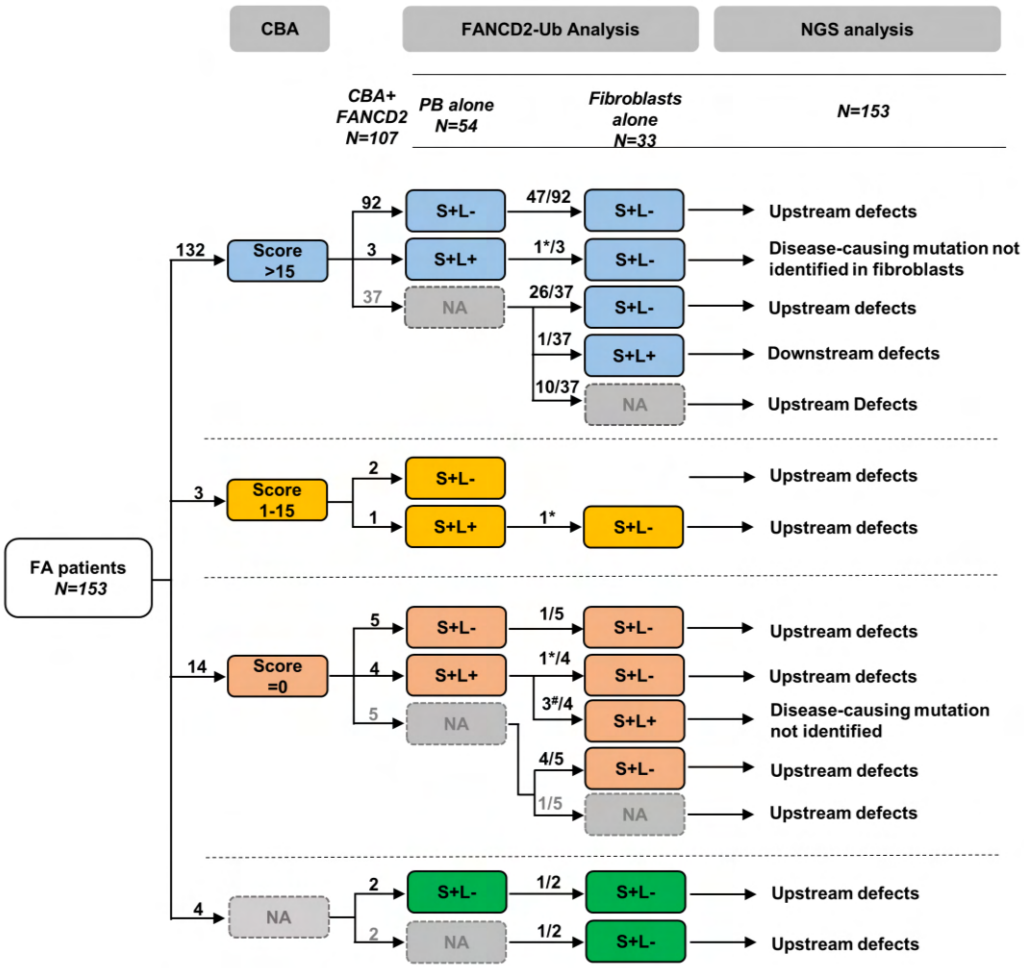


Figure 4.1.2. Chromosome breakage assay (CBA) and FANCD2-Ub analysis in 149 FA patients. **(A)** Representative microscopy images of chromosomes showing higher numbers of abnormalities in the T-cells of FA patients than in normal controls. **(B)** Representative western blot analysis for the detection of FANCD2-S and FANCD2-L isoforms. T -cells and fibroblasts from an FA patient have only FANCD2-S, whereas both the -S and -L forms are present in the cells of a normal control subject. **(C)** Comparison of CBA scores in FA and aplastic anaemia (AA) patients and normal controls. **(D)** Comparison of CBA scores and FANCD2-Ub results in the 153 FA patients recruited in this study. The interpretation and primary diagnosis based on the results of CBAs and FANCD2-Ub analysis before

performing exome sequencing are shown on the right-hand side. *Mosaic patients; # only heterozygous variants identified by NGS; NA: not analysed.

We performed a comparative analysis of the sensitivities of Chromosome Breakage Analysis (CBA) and FANCD2-Ubiquitination (FANCD2-Ub) analysis in the peripheral blood cells of 107 patients who underwent both tests. Among these patients, CBA scores were found to be greater than 15 in 95 individuals (88.7%), indicating a positive result. Conversely, FANCD2-Ub analysis revealed a lack of FANCD2 ubiquitination (FANCD2-Ub-) in 99 patients (92.5%). Notably, among the patients with CBA scores greater than 15, three individuals (2.8%) displayed a unique pattern of CBA>15 and FANCD2-Ub+, suggesting the presence of downstream pathogenic variants. These findings highlight the ability of CBA and FANCD2-Ub analysis to identify different aspects of the underlying genetic defects in FA patients, with CBA scores providing insight into chromosomal breakage patterns and FANCD2-Ub analysis indicating the functionality of the FA pathway. The detection of patients with CBA>15 and FANCD2-Ub+ emphasizes the importance of considering downstream variants and expands our understanding of the molecular heterogeneity of FA. In two patients, NGS analysis confirmed the existence of pathogenic variants in *FANCJ*, while in one patient, it identified pathogenic variants in *BRCA2*. Among the 17 patients with low CBA scores (<15), seven individuals were found to be FANCD2-Ub-. Interestingly, among the five patients with low CBA scores and FANCD2-Ub+ in the peripheral blood, two of them exhibited FANCD2-Ub- in the fibroblast cells **(Table 4.1.1)**.

The identification of mosaicism adds a layer of complexity to the genetic profile of these patients and highlights the importance of analysing both the cell

types viz blood and dermal fibroblasts for comprehensive diagnostic evaluation. The sensitivity analysis of FANCD2-Ub and CBA was conducted in non-mosaic cases of FA. In this evaluation, the FANCD2-Ub analysis demonstrated an overall sensitivity of 97.83%, affirming its exceptional accuracy in detecting FA cases. Likewise, the application of CBA on peripheral blood samples exhibited a sensitivity of 90.21%. These findings were compared with the NGS results, which were considered as the gold standard for analysis **(Table 4.1.2)**.

Table 4.1.2. Calculation of sensitivity of CBA and FANCD2-Ub analysis with NGS as gold standard

		CBA		
		Disease		
Test	**Present**	**n**	**Absent**	**n**
Positive	True Positive	129	False positive	3
Negative	False Negative	14	True Negative	3
Sensitivity	90.21%			
Specificity	50%			
		FANCD2 Ub analysis		
		Disease		
Test	**Present**	**n**	**Absent**	**n**
Positive	True Positive	135	False positive	4
Negative	False Negative	3	True Negative	3
Sensitivity	97.83%			
Specificity	43%			

CBA and FANCD2-Ub analysis are compared to NGS

4.1.3. Detection of Pathogenic Variants by Exome Sequencing: Unveiling the Genetic Landscape of FA.

NGS was conducted for all 153 FA patients enrolled in this study. The DNA used for sequencing analysis was extracted from two different sources: peripheral blood mononuclear cells (PBMNCs) from 74 patients and fibroblasts from 79 patients. The bioinformatics pipeline for identifying the pathogenic variants is shown in **(Figure 4.1.3)**. Genetic alterations were identified in the FA pathway genes of 122 individuals, representing approximately 79.7% of the patient cohort. These genetic variations were observed either in a homozygous state or as compound heterozygous genotypes. **(Table A2)**. Of the remaining 31 patients, 16 individuals who did not exhibit any single nucleotide variants (SNVs) and 15 patients who were found to be heterozygous were further analysed for the presence of copy number variations (CNVs). To detect CNVs, we utilized ExomeDepth (Plagnol et al., 2012), a computational tool that compares the test exome reads to a reference dataset obtained from the same batch. This approach allows for the normalization of read depths, enabling the identification of potential CNVs.

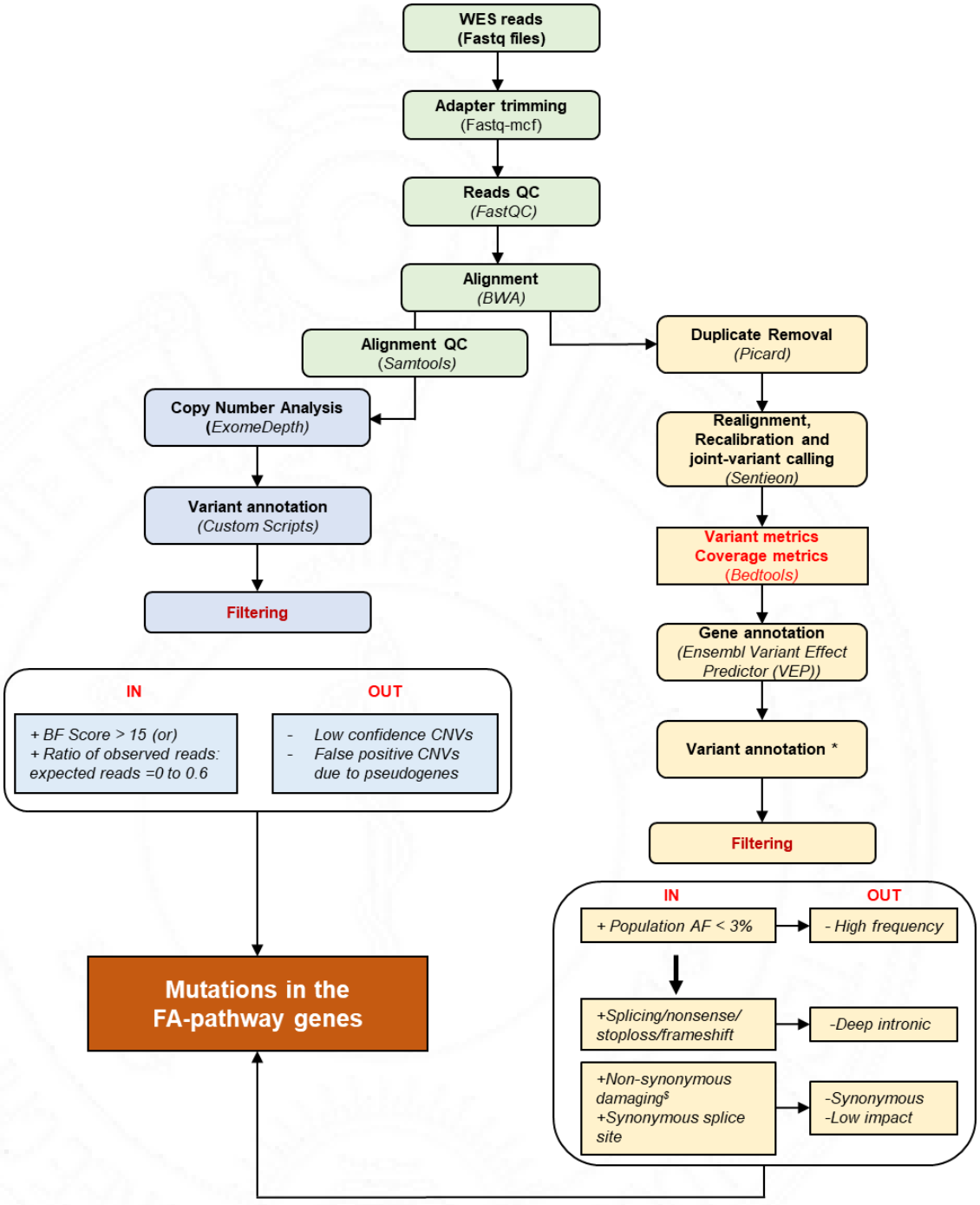
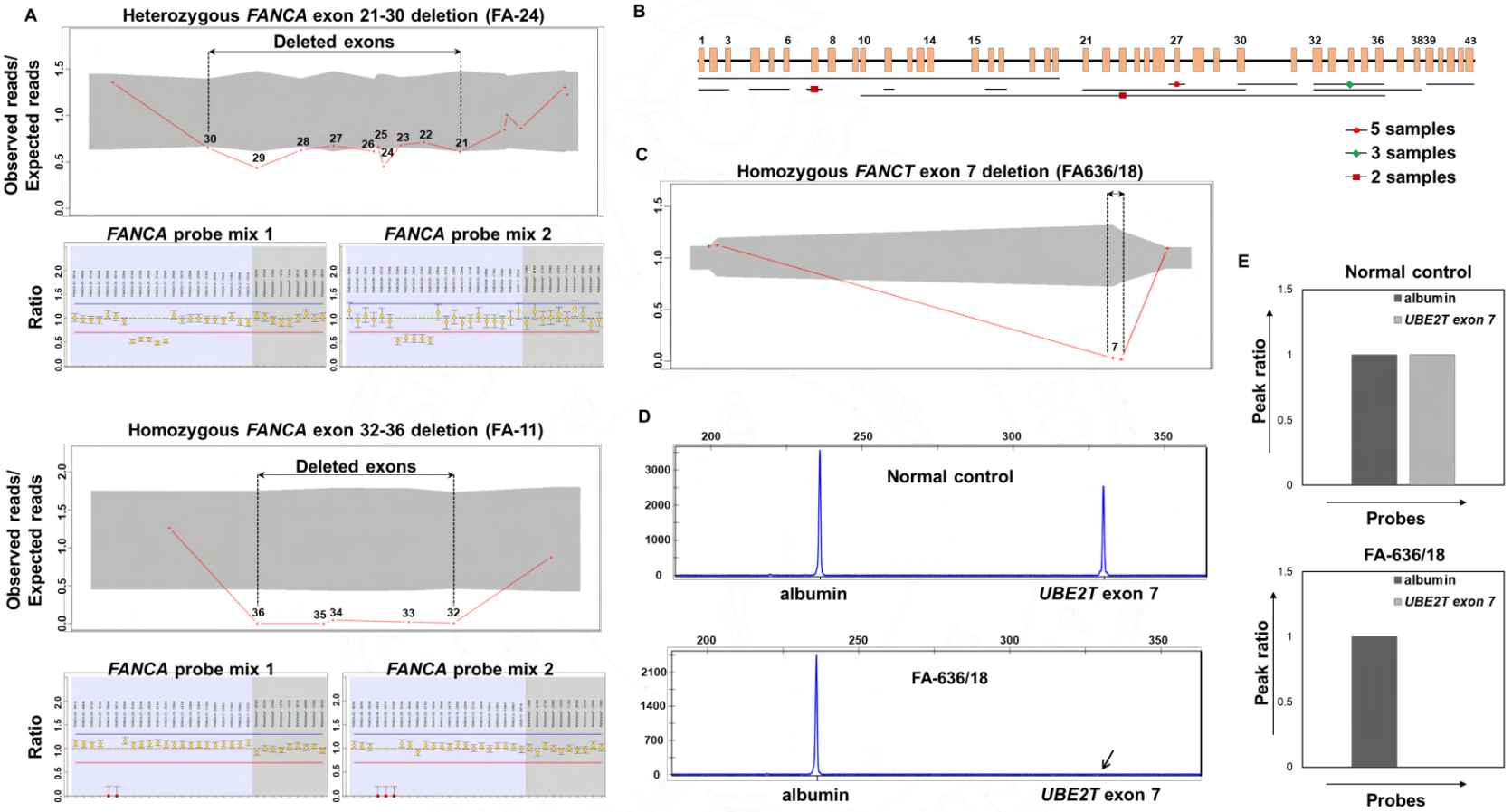


Figure 4.1.3. Pipeline for the bioinformatics analysis of WES reads. *ClinVar, SwissVar, OMIM, HGMD, gnomAD, 1000 Genome Phase3, dbSNP, EVS, 1000

Japanese Genome and internal Indian population databases were used for variant annotations. $-Nonsynonymous variants analysed by EVE algorithm.

By employing ExomeDepth (Plagnol et al., 2012), we aimed to uncover any large-scale genomic alterations, such as duplications or deletions, that may contribute to the pathogenicity of FA **(Figure 4.1.4A and Figure 4.1.4C)**. CNVs were predicted in a total of 23 patients using ExomeDepth, as illustrated in **Figure 4.1.4B** and detailed in **Table 4.1.3**. Among these patients, seven individuals exhibited heterozygous deletions in the *FANCA* gene, indicating the loss of one allele of the gene. Furthermore, one patient demonstrated compound heterozygosity, with two distinct *FANCA* deletions present on different alleles. Additionally, 13 patients displayed homozygous deletions in the *FANCA* gene, suggesting the absence of both alleles. Lastly, two patients exhibited homozygous deletions in the *FANCT* gene. MLPA analysis confirmed the presence and zygosity of the *FANCA* deletions in the 21 patients who were predicted to have deletions by ExomeDepth **(Figure 4.1.4A** and **Figure 4.1.5)**. The predicted homozygous *FANCT* exon 7 deletion was confirmed by a quantitative PCR using fluorescently labelled primers **(Table 3.2)** and capillary electrophoresis **(Figure 4.1.4D and Figure 4.1.4E)**. By integrating both single nucleotide variants (SNVs) and copy number variations (CNVs), we successfully identified disease-associated genotypes in 146 out of 153 (95.4%) patients through exome sequencing, as documented in **Table A2.**

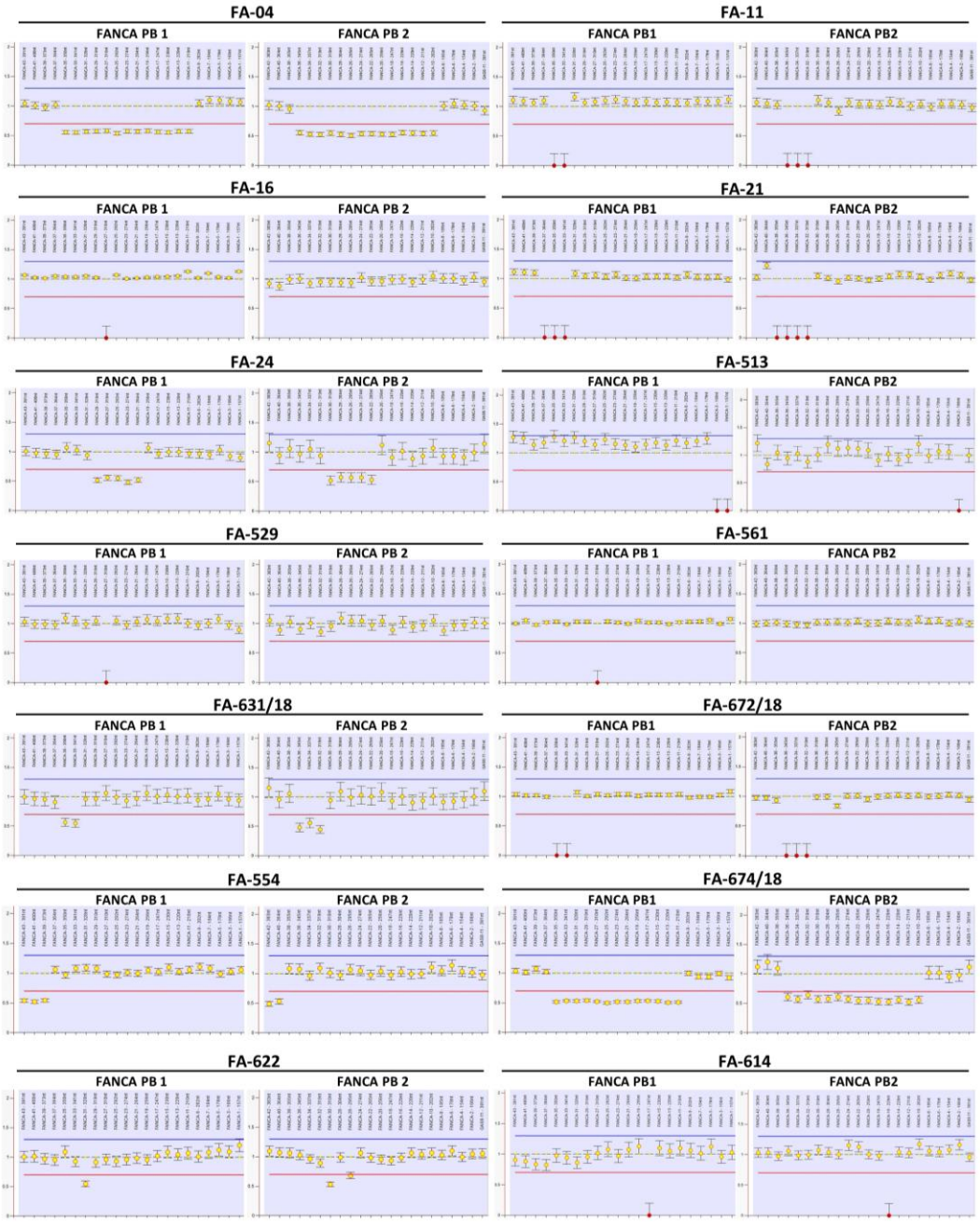


SCTIMST, TRIVANDRUM

Figure 4.1.4. Detection of large deletions in the FA pathway genes. **(A)** Representative results of two *FANCA* deletions detected by bioinformatics analysis in two different patients and confirmation of these deletions by MLPA. FA-24 is heterozygous for a deletion from exon 21 to exon 30, and FA-11 is homozygous for a deletion from exon 32 to exon 36. Probe mixes 1 and 2 cover all the exons of *FANCA*. **(B)** Position of the different *FANCA* deletions identified in this study **(C)** Detection of *FANCT(UBE2T)* deletion by bioinformatics analysis. **(D-E)** Genomic quantitative-PCR for gene dosage analysis to calculate the copy number of exon 7 of the *FANCT* gene. The peak heights of the amplified product of the exon 7 of the *UBE2T* gene and the control albumin gene are shown. Arrow indicates the absence of exome 7 peak in the patient. **(E)** Graph showing the gene dosage of *FANCT*.

Table 4.1.3. Deletions identified in the *FANCA* and *FANCT* genes of FA patients.

Patient ID	Gene	Deleted exons	Zygosity
FA04	*FANCA*	Exons 10-36	Heterozygous
FA11	*FANCA*	Exons 32-36	Homozygous
FA16	*FANCA*	Exon 27	Homozygous
FA21	*FANCA*	Exons 32-38	Homozygous
FA24	*FANCA*	Exons 21-30	Heterozygous
FA513	*FANCA*	Exons 1-3	Homozygous
FA529	*FANCA*	Exon 27	Homozygous
FA561	*FANCA*	Exon 27	Homozygous
FA631/18	*FANCA*	Exons 32-36	Heterozygous
FA636/18	*FANCT*	Exon 7	Homozygous
FA672/18	*FANCA*	Exons 32-36	Homozygous
FA554	*FANCA*	Exons 39-43	Heterozygous
FA674/18	*FANCA*	Exons 10-36	Heterozygous
FA592	*FANCA*	Exons 1-20 and Exon 27	Compound Heterozygous
FA614	*FANCA*	Exons 16-17	Homozygous
FA622	*FANCA*	Exons 30-31	Heterozygous
FA17/19	*FANCA*	Exon 7	Homozygous
FA18/19	*FANCA*	Exon 7	Homozygous
O-117	*FANCA*	Exons 4-6	Homozygous
O-126	*FANCA*	Exon 11	Homozygous
FA-08/19	*FANCA*	Exon 27	Homozygous
FA-30/21	*FANCT*	Exon 7	Homozygous
FA 14/22	*FANCA*	Exon 4-5	Heterozygous


FA-04
FANCA PB 1
FANCA PB 2
FA-11
FANCA PB1
FANCA PB2
FA-16
FANCA PB 1
FANCA PB 2
FA-21
FANCA PB1
FANCA PB2
FA-24
FANCA PB 1
FANCA PB 2
FA-513
FANCA PB 1
FANCA PB2
FA-529
FANCA PB 1
FANCA PB 2
FA-561
FANCA PB 1
FANCA PB2
FA-631/18
FANCA PB 1
FANCA PB 2
FA-672/18
FANCA PB1
FANCA PB2
FA-554
FANCA PB 1
FANCA PB 2
FA-674/18
FANCA PB1
FANCA PB2
FA-622
FANCA PB 1
FANCA PB 2
FA-614
FANCA PB1
FANCA PB2

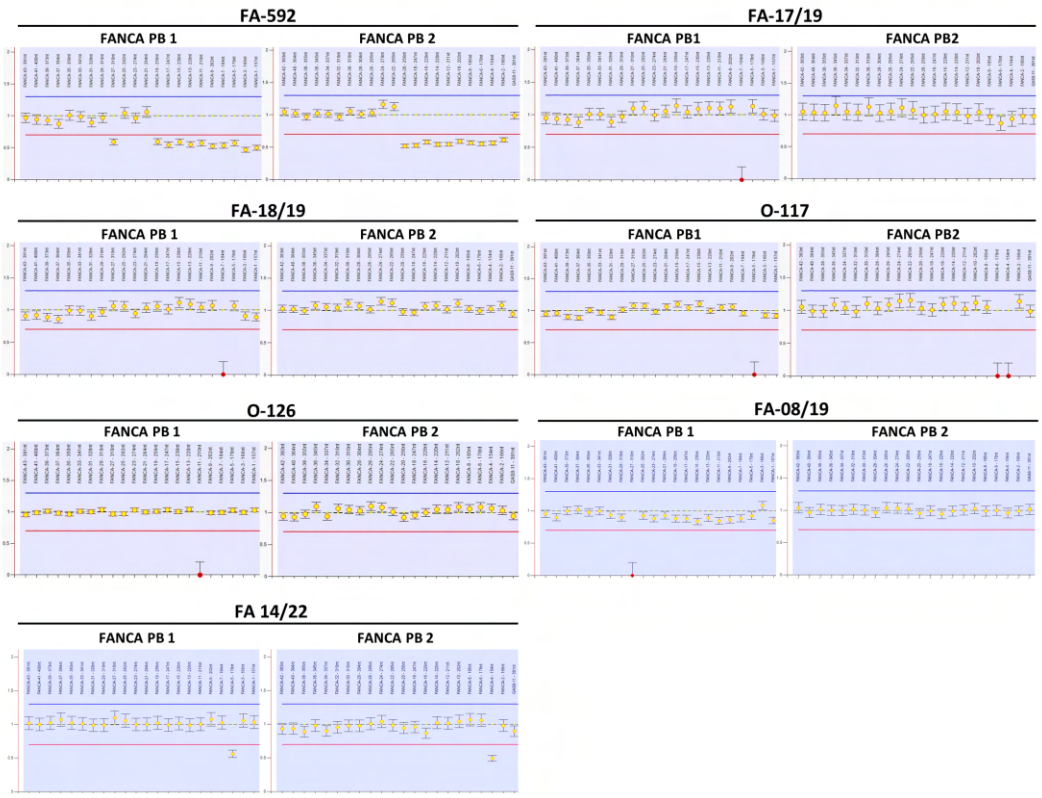


Figure 4.1.5. MLPA analysis to detect *FANCA* gene deletions detected in 21 patients.

Consistent with previous reports, our findings revealed that FA cases within the Indian population were predominantly attributed to pathogenic variants occurring in the upstream genes, including *FANCA, FANCC, FANCG, FANCL,* and *FANCT*. The comprehensive analysis of both single nucleotide variants (SNVs) and copy number variations (CNVs) enabled us to gain a deeper understanding of the genetic landscape of FA in this population. Pathogenic variants were found in *FANCD2* and the downstream genes *FANCJ* and *BRCA2*. In our study, we identified a total of 101 distinct variants across 10 genes associated with the FA pathway.

Notably, 51 of these variants, accounting for 50.5% of the total, were novel and had not been previously reported. **(Table A2)**. Among the identified variants, a total of 127 patients exhibited a homozygous genotype, with 111 patients carrying single nucleotide variants (SNVs) and 16 patients having deletions. Additionally, 19 patients displayed a compound heterozygous genotype, including 11 patients with two different SNVs, 7 patients with a combination of SNVs and large deletions, and 1 patient with two different large deletions. However, in 7 patients, the FA genotypes could not be definitively established. In 2 of these cases, a heterozygous pathogenic variant was identified, suggesting a potential contribution to the disease. In 4 patients, the identified variants were deemed likely benign according to the American College of Medical Genetics and Genomics (ACMG) classification. Furthermore, in 1 patient, a variant of uncertain significance (VUS) was identified, necessitating further evaluation to determine its clinical significance **(Table 4.1.4).** To ensure a comprehensive analysis of the genetic variations in the 7 patients with inconclusive

FA genotypes, we employed Golden Helix VarSeq 2.2.0 (Golden Helix Inc., Bozeman, MT, United States). This clinical genomics interpretation and reporting platform allowed us to conduct a more detailed examination of the samples, considering potential challenges such as low read counts and filtration strategies in our initial pipeline. By leveraging the capabilities of VarSeq, we aimed to identify any additional single nucleotide variants (SNVs) and copy number variations (CNVs) that may have been missed during the initial analysis. We could not detect any additional variants in these samples. Out of the 7 patients who underwent further analysis using Golden Helix VarSeq, four of them exhibited CBA scores of zero. Among these four patients, three showed FANCD2-Ub+ in both peripheral blood (PB) and fibroblasts, while one displayed FANCD2-Ub- specifically in fibroblasts. Among the remaining three patients who had CBA scores above 20, two showed FANCD2-Ub- exclusively in PB, and one displayed FANCD2-Ub- in both PB and fibroblasts. **(Table 4.1.4)**. After excluding the variants found in multiple families, we identified a total of 101 unique variants within the FA pathway genes. These variants encompassed various types of genetic alterations, including 20 (19.8%) splicing variants, 26 (25.7%) missense variants, 25 (24.8%) frameshift variants, 14 (13.9%) nonsense variants, 1 (1%) stop loss variant, and 15 (14.9%) large deletions. **(Figure 4.1.6)**. In contrast to previous studies, which predominantly reported compound heterozygous pathogenic variants in FA patients (Bogliolo et al., 2020; Chang et al., 2014), our findings revealed that 87% of the patients in our study were homozygous for the identified pathogenic variants

(Figure 4.1.6). This higher prevalence of homozygosity may be attributed to the elevated rate of consanguinity within the population under investigation.

Table 4.1.4. FANCD2-Ub status, CBA, pathogenicity prediction and clinical phenotype of 7 FA patients with heterozygous mutations

FA ID	Gene	Mutation	VarSome results	EVE results	Type of mutations	CBA score	FANCD2 ubiquitination		Physical anomalies
							T-cells	Fibroblasts	
FA-29	*FANCD2*	NC_000003.11:g.10076177A>C	Likely benign	Pathogenic	Missense	0	FANCD2-Ub+	FANCD2-Ub+	HSK, kidney defects
FA-518	*FANCC*	NC_000009.11:g.97873815G>A	Likely benign	Benign	Missense	0	FANCD2-Ub+	FANCD2-Ub+	Thenar hypoplasia, HP
FA-519	*FANCD2*	NC_000003.11:g.10083368C>T	Pathogenic	NA (nonsense variant)	Nonsense	62.7	FANCD2-Ub-	FANCD2-Ub-	SS, Absent simian crease, TH, TF, HP
FA-528	*C17orf70*	NC_000017.10:g.79516335G>T	Likely benign	NA	Missense	0	FANCD2-Ub+	FANCD2-Ub+	SS, MCH
FA-541	*FANCE*	NC_000006.11:g.35423588G>A	Likely benign	VUS	Missense	0	NA	FANCD2-Ub-	NA
FA-32/19	*FANCA*	NC_000016.9:g.89869731A>T	VUS	Pathogenic	Missense	22.1	FANCD2-Ub-	NA	Musculoskeletal abnormalities, Hypertelorism, MCH, HAP, CAL
FA-07/20	*FANCD2*	Deletion exon 22-23	Pathogenic	NA	Deletion	30.3	FANCD2-Ub-	NA	Clubbing in both hands, HAP, MCH, COP

HSK: horse-shoe kidney; AE: abnormal eyes; MCH: microcephaly; HP: hyperpigmentation; SS: short stature; TH: thumb hypoplasia; TF: triangular facies; COP: Circum oral pigmentation

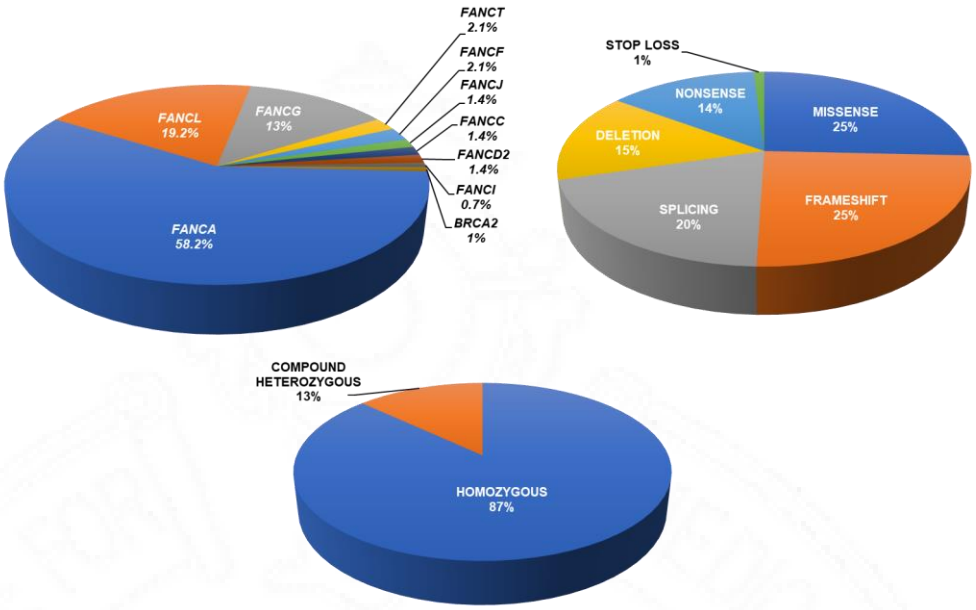


Figure 4.1.6. Genotyping of 146 FA patients. Left: The percentages of the defective genes in homozygous and compound heterozygous states. Right: The percentages of different types of mutations identified. Bottom: Zygosity of mutations identified.

4.1.4. Highly frequent pathogenic variants in the Indian Population.

We observed that *FANCA* was the gene with the highest frequency of mutations, accounting for 58.2% of the total mutations identified. This finding is consistent with previous studies conducted in diverse populations, where *FANCA* mutations have also been reported to be prevalent, ranging from 60% to 80% (Bogliolo et al., 2020; Esmail Nia et al., 2016; Nie et al., 2020) **(Table A2, Figure 4.1.6)**. Among the cohort of patients included in the study, a total of 85 individuals were found to have pathogenic variants in the *FANCA* gene. In these patients, a total of 68 distinct pathogenic variants were identified, with 33 of them being novel, previously unreported variants. The majority of the patients (70 out of 85) exhibited homozygous pathogenic variants in *FANCA*, indicating the presence of two identical pathogenic variants in the gene. Additionally, 15 patients displayed compound heterozygous mutations, characterized by the presence of two different pathogenic variants in the gene. In our study, we detected a total of 14 distinct deletions in 21 patients affecting the *FANCA* gene. One notable finding was the presence of a high frequency (3.4%) of a specific deletion involving exon 27 of FANCA, which aligns with the findings of a previous study conducted in the Indian population (George, Solanki, Chavan, Rajendran, Raj, Mohan, Nemani, Kanvinde, Munirathnam, Rao, Radhakrishnan, Prasada, et al., 2021) **(Figure 4.1.4, Table 4.1.3)**. This particular deletion event seems to recur frequently among the patients analyzed in both studies, suggesting its potential relevance as a recurrent genetic alteration in individuals with Fanconi Anemia. These observations further support the significance of examining specific genomic regions,

such as exon 27 of *FANCA*, for the detection of clinically relevant deletions and underline the importance of comprehensive genetic testing. In contrast to other populations where *FANCC* is reported to have a frequency of 10-15%, we observed a relatively low frequency (1.4%) of *FANCC* pathogenic variants in our patient cohort (George, Solanki, Chavan, Rajendran, Raj, Mohan, Nemani, Kanvinde, Munirathnam, Rao, Radhakrishnan, Prasada, et al., 2021; Nie et al., 2020). On the other hand, the frequency of *FANCG* pathogenic variants in our study (13%) was comparable to that reported in other populations, which range from 9-12% (Bogliolo et al., 2020; Esmail Nia et al., 2016). Rare pathogenic variants in FA genes were identified in our patient cohort, including *FANCT/UBE2T* mutations in 3 patients, *FANCI* mutations in 1 patient, *FANCJ/BRIP1* mutations in 2 patients, *FANCF* mutations in 2 patients, *FANCD2* mutations in 1 patient, and *FANCD1/BRCA2* mutation in 1 patient **(Table A2, Figure 4.1.6)**.

4.1.5. Identification of a Founder Variant in the *FANCL* Gene.

Although FANCL-related FA is considered a rare subtype, our study identified a relatively higher prevalence of FANCL pathogenic variants in our patient cohort. Out of the analysed patients, 28 individuals (19.2%) were found to have FANCL pathogenic variants, which is higher than the reported prevalence of 0.2-0.4% in FA cases. These findings suggest that FANCL-related FA may have a higher frequency in our studied population than previously recognized. **(Table A2)**. A synonymous splicing variant, c.1092G>A; p.K364=, in the *FANCL* gene, was identified in a

homozygous state in 28 (19.2%) patients. Subsequent Sanger sequencing analysis of the PCR-amplified *FANCL* cDNA from a patient with this variant confirmed the previously reported exon 13 skipping (Donovan et al., 2020). This finding provides further evidence of the functional impact of this variant on the splicing process in the *FANCL* gene. **(Figure 4.1.7A-C)**. Lentiviral transduction of wild-type *FANCL* cDNA successfully restored FANCD2-Ub in the fibroblasts of a patient carrying the pathogenic variant. This functional rescue experiment provides strong evidence that the identified variant in the *FANCL* gene leads to impaired FANCD2 ubiquitination and underscores the importance of *FANCL* in the FA pathway. **(Figure 4.1.7D)**. All the patients carrying this pathogenic variant were from South Indian states, with 11 patients from Andhra Pradesh, 9 from Kerala, 6 from Tamil Nadu, 1 from Maharashtra, and 1 from Karnataka. This geographical distribution highlights the prevalence of this variant in specific regions of South India, suggesting a potential regional founder effect or shared ancestry among these individuals **(Figure 4.1.1)**.

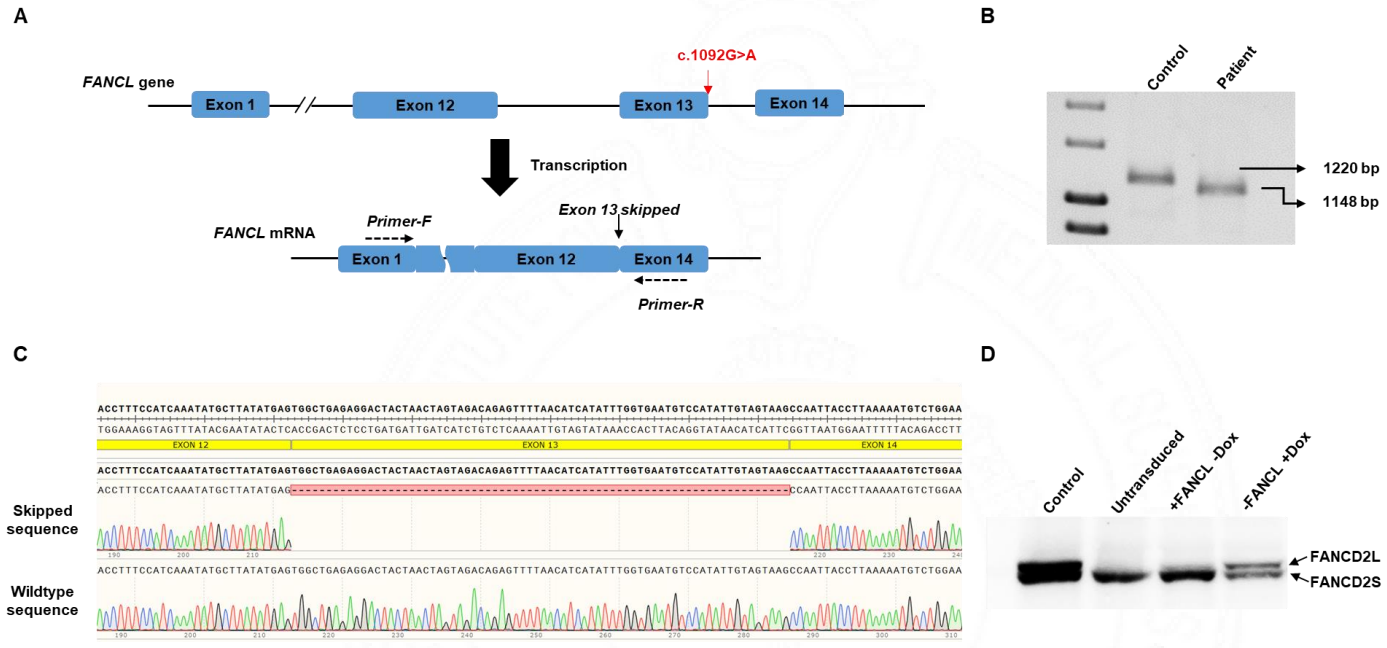


Figure 4.1.7. Molecular analysis of *FANCL* (c.1092G>A;p.K364=) identified in our cohort of FA patients. **(A)** Diagrammatic representation of the location of the mutation in the *FANCL* gene, exon 13 skipping in *FANCL* mRNA, the position of the primers used for amplification of the cDNA and Sanger sequencing are shown. **(B)** The PCR amplification of *FANCL* cDNAs from a normal control and an FA patient with the c.1092G>A;p.K364= mutation showing the difference in molecular weights of the amplified products. **(C)** Sanger sequencing results showing skipping of exon 13 of *FANCL* in the patient. **(D)** Western blot analysis showing the restoration of FANCD2 Ub after lentiviral complementation of a 'patient's fibroblasts with wild-type *FANCL* cDNA. The results were compared with a

normal control and untransduced fibroblasts from a patient with *FANCL* (c.1092G>A;p.K364=) mutation. The vector used for complementation analysis was pCW-*FANCL*, which has doxycycline-inducible *FANCL* expression

Although this pathogenic variant has been reported previously in 12 Indian FA patients (Donovan et al., 2020), our comprehensive study, encompassing representative samples from various regions of the country, provides a more accurate assessment of its frequency among Indian FA patients. By analysing a larger and more diverse patient cohort, we could determine the exact prevalence and impact of this variant within the Indian FA patients. We also identified an additional pathogenic variant in the *FANCL* gene: a nonsense variant c.997C>T; p.Gln333* detected in a patient in the compound heterozygous state with the FANCL c.1092G>A; p.K364= variant. **(Table A2)**. Other highly frequent pathogenic variants included the *FANCG* c.2786A>C variant (n=5) and the *FANCG* c.1761-2A>C variant (n=5). Additionally, the *FANCA* gene exhibited a high frequency of pathogenic variants, including the *FANCA* c.3066+1G>T variant (n=4), the *FANCA* c.319delG variant (n=4), and the *FANCA* c.826+2T>C variant (n=4) **(Table A2)**.

4.1.6. Evaluating the Pathogenicity of Missense Variants.

We identified a total of 26 missense variants in the genotyped cohort of 153 Fanconi Anaemia (FA) patients. To assess the pathogenicity of these variants, we employed a comprehensive approach utilizing the ACMG guidelines (Laboratories et al., 2018), the ClinVar database (Laboratories et al., 2018), and the VarSome variant discovery tool (Kopanos et al., 2019). These tools incorporate multiple pathogenicity prediction methods to classify the variants as pathogenic, likely pathogenic, or variants of unknown significance (VUS). By employing this rigorous analysis, we aimed to

accurately determine the functional impact of these missense variants and their relevance to the development of FA in the patient population. We identified 6 pathogenic variants by ACMG guidelines, 8 by ClinVar, and 7 by Varsome (**Table 4.1.5**). In addition to the aforementioned methods, we further analysed the missense variants using the evolutionary model of variant effect (EVE) (Frazer, Notin, Dias, Gomez, Joseph K Min, et al., 2021) tool available at (https://evemodel.org/). This tool utilizes evolutionary conservation information to predict the pathogenicity of variants. Among the 26 missense variants identified in our patient cohort, the EVE analysis classified 21 variants as pathogenic, 2 variants as variants of unknown significance (VUS), 1 variant as benign, and 2 variants did not have any EVE scores available. By incorporating the EVE tool into our analysis, we gained additional insights into the potential functional impact of these missense variants in the context of evolutionary conservation. (**Table 4.1.5).**

Table 4.1.5. Determination of pathogenicity of missense mutations

Sample ID	Gene	cDNA change	Amino acid change	Zygosity	ACMG	ClinVar	Varsome	EVE prediction	Final verdict[#]
FA-02	*FANCJ (BRIP1)*	NM_032043.3:c.1878 A>T	p.Glu626Asp	Hom	Likely pathogenic	VUS	Pathogenic	Pathogenic	Pathogenic
FA-02/21	*FANCA*	NM_000135.4:c.3788 T>C	p.Phe1263Ser	CH	Likely pathogenic	VUS	Likely pathogenic	VUS	VUS/Likely Pathogenic
FA-02/21	*FANCA*	NM_000135.4:c.1540 G>A	p.Ala514Thr	CH	VUS	NA	VUS	Pathogenic	Pathogenic
FA-03	*FANCA*	NM_000135.4:c.2786 A>C	p.Tyr929Ser	CH	Likely pathogenic	VUS	VUS	Pathogenic	Pathogenic*
FA-05	*FANCA*	NM_000135.4:c.1304 G>A	p.Arg435His	Hom	Pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic
FA-06/20	*FANCA*	NM_000135.4:c.4198 C>T	p.Arg1400 Cys	Hom	Likely pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic
FA-12	*FANCA*	NM_000135.4:c.1303 C>T	p.Arg435Cys	CH	Pathogenic	Pathogenic	Likely pathogenic	Pathogenic	Pathogenic
FA-18	*FANCA*	NM_000135.4:c.2852 G>C	p.Arg951Pro	Hom	Pathogenic	NA	Likely pathogenic	Pathogenic	Pathogenic
FA-18/21	*FANCA*	NM_000135.4:c.2290 C>T	p.Arg764Trp	Hom	VUS	Pathogenic	Pathogenic	Pathogenic	Pathogenic
FA-21/21	*FANCA*	NM_000135.4:c.3239 G>A	p.Arg1080 Gln	Hom	Likely pathogenic	Likely pathogenic	Likely pathogenic	Pathogenic	Pathogenic
FA-31/21	*FANCA*	NM_000135.4:c.1430 T>C	p.Leu477Ser	Hom	VUS	NA	Pathogenic	Pathogenic	Pathogenic
FA-38	*FANCG*	NM_004629.2:c.425 T>C	p.Leu142Pro	CH	VUS	NA	VUS	Pathogenic	Pathogenic
FA-5/21	*UBE2T/FANCT*	NM_014176.4:c.232 A>C	p.Asn78His	Hom	VUS	NA	VUS	Pathogenic	Pathogenic
FA-527	*FANCC*	NM_000136.3:c.1585 A>C	p.Thr529Pro	Hom	VUS	VUS	VUS	Pathogenic	Pathogenic*

FA-533	*FANCA*	NM_000135.4:c.3934 G>A	p.Asp1312 Asn	CH	Likely pathogenic	NA	VUS	Pathogenic	Pathogenic

4.1.7. Complementation analysis for pathogenicity prediction of variants of uncertain significance.

The validation of variant pathogenicity can be effectively achieved through complementation analysis, wherein wild-type cDNA is transferred into FA cells using lentiviral-mediated gene transfer, resulting in the restoration of cellular phenotypes. This method is considered reliable for confirming the pathogenicity of the variants (Chandra et al., 2005). Following the antibiotic selection of the fibroblasts transduced with lentiviral vectors carrying wild-type cDNAs, the cells were exposed to MMC and subsequently assessed for their FANCD2-Ub status. In order to validate the complementation analysis, we initially examined the fibroblasts of 13 patients who harboured pathogenic null variants in the genes associated with the upstream pathway of FA **(Table 4.1.6)**. Remarkably, we observed a complete restoration of FANCD2 ubiquitination (FANCD2-Ub) in all of these patients, indicating successful complementation of the FA upstream pathway genes **(Figure 4.1.8, Table 4.1.6)**. Subsequently, we conducted complementation analysis in fibroblasts from 7 patients who had variants of uncertain significance (VUS), or likely pathogenic variants based on the ACMG classification. These variants were identified in a homozygous or compound heterozygous state. The patients included 4 with *FANCA* variants, 1 with *FANCG* variant, 1 with *FANCC* variant, and 1 with *FANCF* variant. **(Table 4.1.6)**. All these patients showed restoration of FANCD2-Ub after complementation.

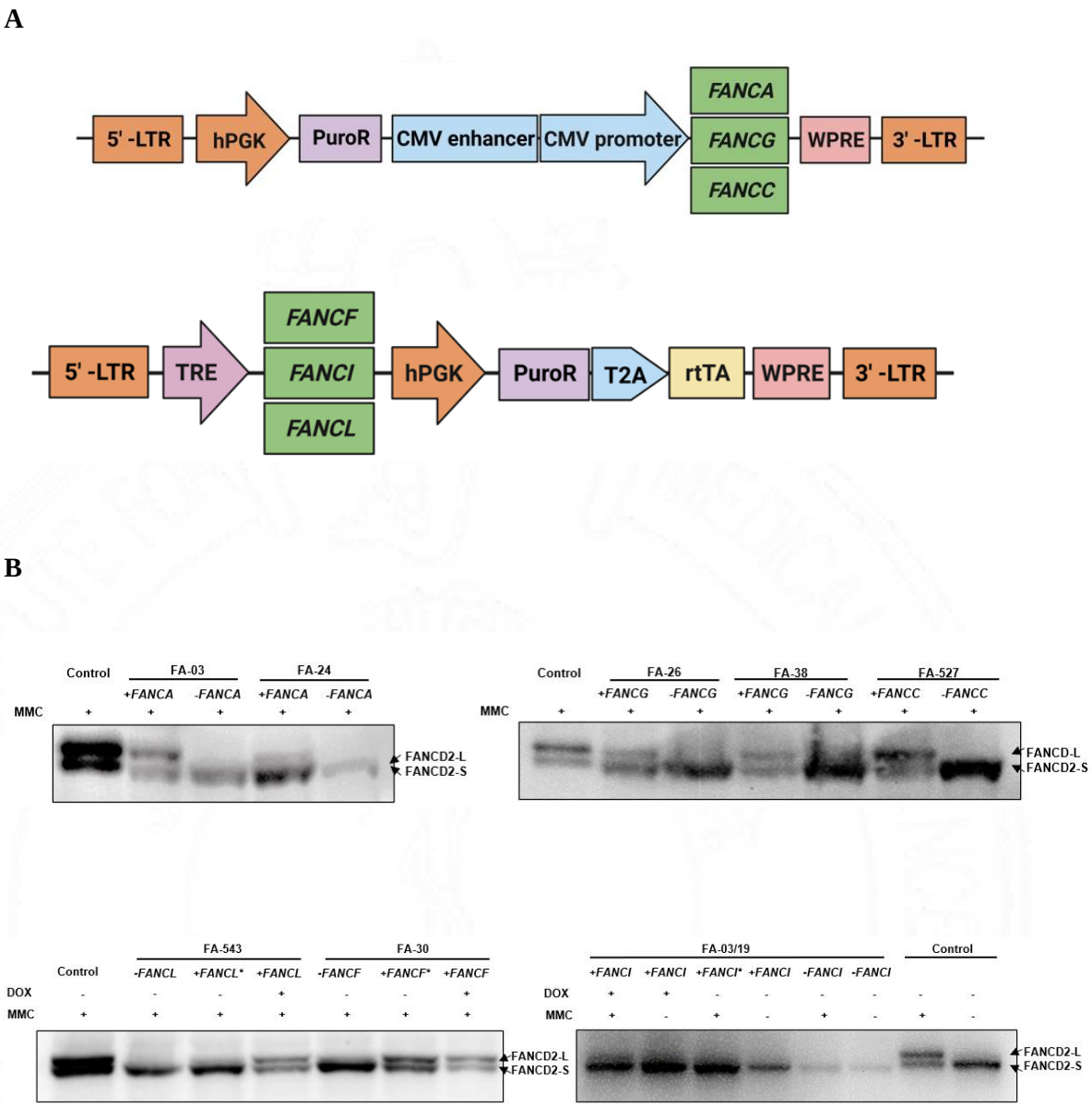


Figure 4.1.8. Lentiviral complementation analysis (A) Lentiviral constitutive expression vectors for complementation analysis of *FANCA*, *FANCG* and *FANCC* and doxycycline-inducible expression vectors for *FANCF*, *FANCI* and *FANCL*. B) FANCD2 western blot results after complementation of *FANCA*, *FANCG*, *FANCC*, *FANCF*, *FANCI* and *FANCL* genes in the fibroblasts with mutations in these genes. *Leaky expression vector that exhibits transgene expression in the absence of doxycycline (DOX). FA-03, FA-24, FA-26, FA-38, FA-527, FA-543, FA-30, FA-03/19 are patients IDs. hPGK: human polyglycerate kinase promoter, PuroR: puromycin resistance gene, CMV: cytomegalovirus, WPRE: woodchuck hepatitis virus posttranscriptional regulatory element, LTR: long terminal repeat, TRE:

tetracycline response element, T2A: self-cleaving 2A peptide, rtTA: reverse tetracycline-controlled transactivator.

Table 4.1.6. The genotypes of the FA patients selected for complementation analysis.

ID	Gene	Zygosity	DNA change	Type of mutation	Varsome	EVE
FA-01	*FANCG*	Hom	NC_000009.11:g.35074215T>G	Splice site	Pathogenic	NA
FA-03	*FANCA*	CH	NC_000016.9:g.89813237A>C	Splice site	Pathogenic	NA
			NC_000016.9:g.89828423T>G	Missense	VUS	Pathogenic
FA-03/19	*FANCI*	CH	NC_000015.9:g.89804822del	Frameshift deletion	Pathogenic	NA
			NC_000015.9:g.89848828_89848832del	Intronic splice acceptor variant	VUS	NA
FA-04	*FANCA*	CH	NC_000016.9:g.89882396T>C	Splice site	Pathogenic	NA
			NC_000016.9:g.89811367_89865640del	Deletion	Pathogenic	NA
FA-06	*FANCA*	Hom	NC_000016.9:g.89828423T>G	Missense	Likely pathogenic	Pathogenic
FA-12	*FANCA*	CH	NC_000016.9:g.89809286_89809290dup	Frameshift insertion	Likely pathogenic	NA
			NC_000016.9:g.89857867G>A	Missense	Likely pathogenic	Pathogenic
FA-18	*FANCA*	Hom	NC_000016.9:g.89828357C>G	Missense	Likely pathogenic	Pathogenic
FA-24	*FANCA*	CH	NC_000016.9:g.89809302C>T	Nonsense	Pathogenic	NA
			NC_000016.9:g.89828358_89842224del	Deletion	Pathogenic	NA
FA-26	*FANCG*	Hom	NC_000009.11:g.35076427_35076431del	Splice site	VUS	NA
FA-30	*FANCF*	Hom	NC_000011.9:g.22646233T>C	Stop loss	VUS	NA
FA-31	*FANCA*	Hom	NC_000016.9:g.89882944C>G	Splice site	Pathogenic	NA
FA-36	*FANCG*	Hom	NC_000009.11:g.35075650_35075651del	Frameshift deletion	Pathogenic	NA
FA-38	*FANCG*	CH	NC_000009.11:g.35078223A>G	Missense	VUS	Pathogenic
			NC_000009.11:g.35076856_35076857del	Frameshift deletion	VUS	NA

FA-40	*FANCA*	Hom	NC_000016.9:g.89877448del	Frameshift deletion	Pathogenic	NA
FA-517	*FANCC*	Hom	NC_000009.11:g.97873912C>A	Nonsense	Pathogenic	NA
FA-527	*FANCC*	Hom	NC_000009.11:g.97864081T>G	Missense	VUS	Pathogenic
FA-531	*FANCG*	Hom	NC_000009.11:g.35076026C>G	Splice site	Pathogenic	NA
FA-533	*FANCA*	CH	NC_000016.9:g.89858399_89858400del	Frameshift deletion	VUS	NA
			NC_000016.9:g.89806402C>T	Missense	VUS	Pathogenic
FA-543	*FANCL*	Hom	NC_000002.11:g.58387243C>T	Exonic splice donor variant	Pathogenic	NA
FA-641/18	*FANCF*	Hom	NC_000011.9:g.22647316A>C	Missense	VUS	VUS

NA- not applicable
VUS- Variants of uncertain significance

4.1.8. Developing an Efficient Molecular Diagnosis Strategy for FA.

Our study revealed that a significant proportion, 56.8%, of the patients in our cohort had single nucleotide variants (SNVs) in the *FANCA* and *FANCG* genes. To efficiently detect pathogenic variants in these genes, we developed a specialized method called long amplicon next-generation sequencing (LA-NGS). In the LA-NGS approach, we designed six long amplicons to cover the entire *FANCA* gene and one long amplicon to encompass the *FANCG* gene. These long amplicons were generated through a long-range polymerase chain reaction. Subsequently, the amplified PCR products from all the long amplicons were pooled together in a single tube. The pooled amplicon mixture was then subjected to next-generation sequencing (NGS), which generated high-throughput sequencing data encompassing the entire targeted regions of *FANCA* and *FANCG*. The resulting NGS data were subsequently analysed using bioinformatics tools and pipelines specifically designed for variant calling and annotation. This LA-NGS method enabled us to efficiently screen for and identify pathogenic variants in the *FANCA* and *FANCG* genes in a cost-effective and time-efficient manner, providing a comprehensive analysis of these genes in our patient cohort. (**Figures 4.1.9A** and **4.1.9B)**. To assess the reliability and accuracy of our developed long amplicon next-generation sequencing (LA-NGS) method in detecting single nucleotide variants (SNVs), we performed validation experiments using DNA samples from 24 patients who were previously confirmed to have SNVs in the *FANCA* and *FANCG* genes **(Table 4.1.7)**.

Table 4.1.7. Mutations identified in *FANCA* and *FANCG* genes by LA-NGS

FA ID	Gene	Mutation	Type of mutations	Site of mutation
FA-31	*FANCA*	NC_000016.9:g.89882944C>G	Splice site	Intron 1
FA-521	*FANCA*	NC_000016.9:g.89877481T>G	Splice site	Intron 3
FA-500/ FA-05	*FANCA*	NC_000016.9:g.89857866C>T	Missense	Exon 14
FA-503/ FA-15	*FANCA*	NC_000016.9:g.89851261C>T	Splice site	Intron 15
FA-505/ FA-09	*FANCA*	NC_000016.9:g.89877340_89877341del	Splice site	Intron 4
		NC_000016.9:g.89877448del	Frameshift deletion	Exon 4
FA-515/ FA-03	*FANCA*	NC_000016.9:g.89813237A>C	Splice site	Intron 34
		NC_000016.9:g.89828423T>G	Missense	Exon 29
		NC_000016.9:g.89882965;C>T	Missense	Exon 1
FA-551/ FA-18	*FANCA*	NC_000016.9:g.89828357C>G	Missense	Exon 29
FA-555/ FA-516	*FANCA*	NC_000016.9:g.89825022del	Frameshift insertion	Exon 30
FA-568	*FANCA*	NC_000016.9:g.89828423T>G	Missense	Exon 29
FA-574/ FA-511	*FANCA*	NC_000016.9:g.89831327G>A	Nonsense	Exon 28
FA-20/19	*FANCA*	NC_000016.9:g.89877481T>G	Splice site	Intron 3
FA-27/19	*FANCA*	NC_000016.9:g.89866011A>G	Splice site	Intron 9
FA-36/19	*FANCG*	NC_000009.11:g.35074215T>G	Splice site	Intron 13
FA-37/19	*FANCG*	NC_000009.11:g.35074215T>G	Splice site	Intron 13
FA-573	*FANCA*	NC_000016.9:g.89877448del	Frameshift Deletion	Exon 4
		NC_000016.9:g.89883022A>T	Missense	Exon 1
FA-580	*FANCG*	NC_000009.11:g.35078222_35078223delinsAGCAGTT	Frameshift insertion deletion	Exon 4
FA-581	*FANCA*	NC_000016.9:g.89809216_89809219del	Frameshift Deletion	Exon 37
FA-584	*FANCA*	NC_000016.9:g.89877327_89877344del	Frameshift Deletion	Exon 4, Intron 4
FA-588	*FANCG*	NC_000009.11:g.35074215T>G	Splice site	Intron 13
FA-591	*FANCA*	NC_000016.9:g.89866011A>G	Splice site	Intron 9
FA-593	*FANCA*	NC_000016.9:g.89828358G>A	Missense	Exon 29
FA-598	*FANCG*	NC_000009.11:g.35075059G>A	Nonsense	Exon 12
FA-599	*FANCA*	NC_000016.9:g.89877327_89877344del	Frameshift Deletion	Exon 4, Intron 4
FA-622	*FANCA*	NC_000016.9:g.89816231_89816232del	Frameshift deletion	Exon 32

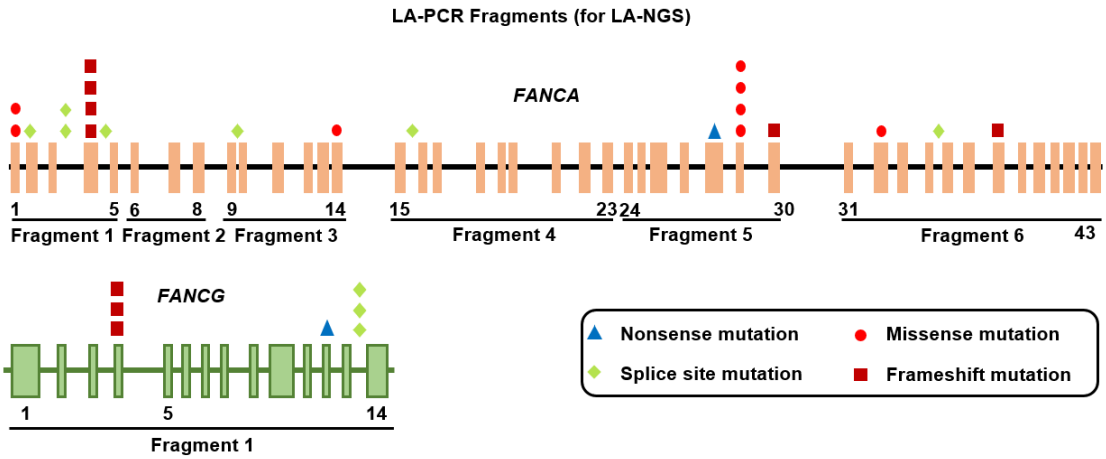
The successful detection of the known SNVs in *FANCA* and *FANCG* in all 24 patient samples validated the reliability and effectiveness of our LA-NGS approach. This method offers several advantages over current molecular diagnostic strategies. One notable advantage is its cost-effectiveness, making it a more economical choice for genetic analysis compared to other approaches.

By specifically targeting the *FANCA* and *FANCG* genes with long amplicons, this technique minimizes unnecessary sequencing of non-relevant regions, resulting in cost savings without compromising the detection of pathogenic variants. Additionally, the LA-NGS method offers a faster turnaround time compared to exome sequencing, which involves sequencing the entire exome. By focusing on the specific target genes of interest, the sequencing process is streamlined, reducing the time required for sample preparation and data analysis. This accelerated workflow enables more efficient genetic testing and timely reporting of results, which is crucial for clinical decision-making.

Furthermore, the LA-NGS method requires less extensive bioinformatics analysis compared to exome sequencing. Since *FANCA* and *FANCG* target regions are specifically amplified and sequenced, the data analysis is focused on a smaller subset of genomic regions, simplifying the bioinformatics pipeline. This streamlined analysis reduces computational requirements and facilitates data interpretation, making the LA-NGS method more accessible to laboratories with limited computational resources. Overall, the combination of cost-effectiveness, faster turnaround time, and simplified bioinformatics analysis makes the LA-NGS method a valuable tool for genetic

analysis, particularly for identifying pathogenic variants in the *FANCA* and *FANCG* genes.

A



B

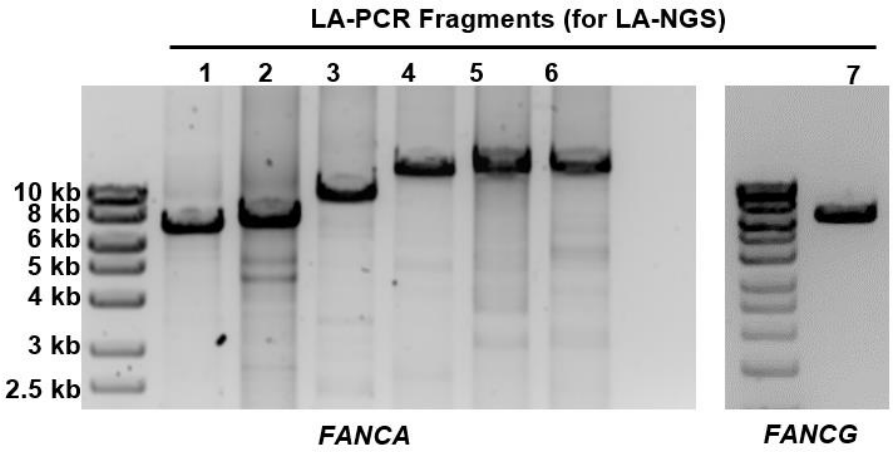
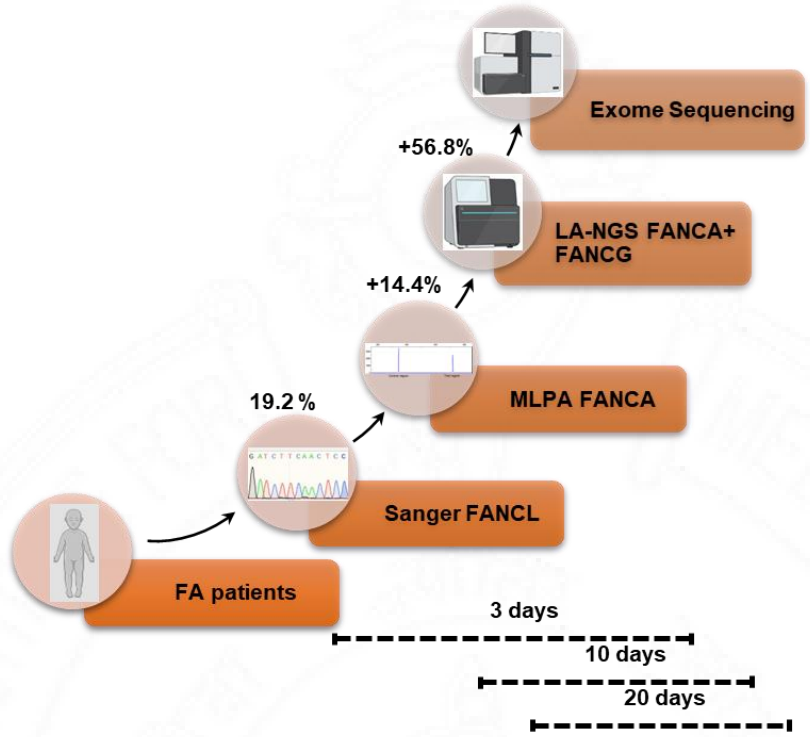


Figure 4.1.9. Long amplicon next-generation sequencing to identify mutations in *FANCA* and *FANCG* gene **(A)** Diagrammatic representation of the regions of *FANCA* and *FANCG* amplified for LA-PCR-NGS. **(B)** Agarose gel images showing the amplified *FANCA* and *FANCG* gene fragments. *FANCA*: fragment 1- 7126 bp; fragment 2- 8595 bp; fragment 3- 9726 bp; fragment 4- 13739 bp; fragment 5- 13483 bp; fragment 6- 15253 bp; and *FANCG*: fragment 7- 6760 bp.

Considering the high prevalence of the *FANCL* c.1092G>A;p.K364= pathogenic variant in 19.2% of FA patients, Sanger sequencing becomes a suitable initial test for genotyping Indian FA patients. This method enables the detection of this specific variant, allowing for rapid identification of affected individuals. Additionally, MLPA (Multiplex Ligation-dependent Probe Amplification) can be employed to detect *FANCA* deletions, which account for 14.4% of the overall pathogenic variants. Both Sanger sequencing and MLPA provide relatively quick results, typically within 48 hours. For those patients who test negative for the identified pathogenic variants through Sanger sequencing and MLPA, the next step involves utilizing long amplicon next-generation sequencing (LA-NGS). This approach is particularly effective in detecting single nucleotide variants (SNVs) in the *FANCA* and *FANCG* genes, which represent approximately 57% of the FA pathogenic variants in the Indian population. By applying LA-NGS, a comprehensive analysis can be conducted to identify these SNVs, aiding in the molecular diagnosis of a substantial portion of FA patients.

By following this algorithmic approach, incorporating Sanger sequencing, MLPA, and LA-NGS, it becomes possible to achieve molecular diagnosis in approximately 90% of FA patients within the Indian population **(Figure 4.1.10A)**. This streamlined algorithm was tested in 27 new FA patients with a median CBA score of 66.8 (0 to 115) and offers a balanced combination of accuracy, efficiency, and timely results, contributing to improved diagnostic capabilities **(Figure 4.1.10B)**.

A



B

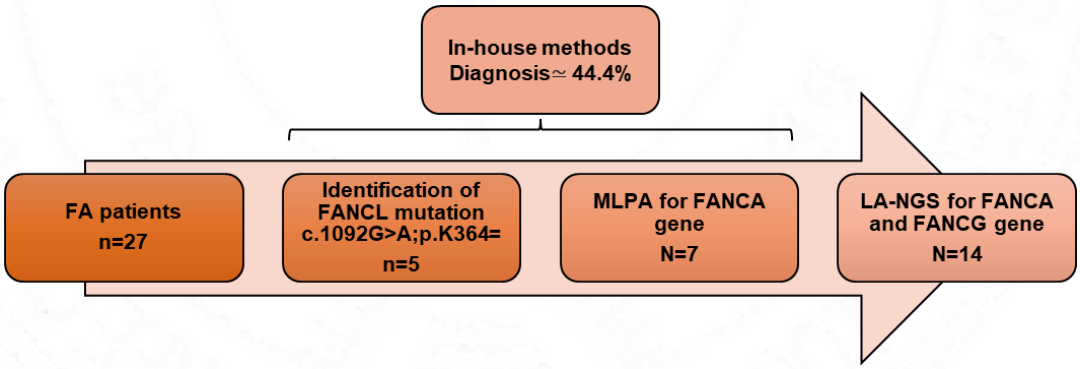


Figure 4.1.10. Diagnostic strategies used for the molecular diagnosis of FA in the Indian population **(A)** Algorithm for the molecular diagnosis of FA. **(B)** The new algorithm tested in 27 FA patients.

4.2 Genotype-phenotype Association in Patients with Fanconi Anaemia (FA).

Fanconi anaemia (FA) is a disease characterized by genetic and clinical variability. Our understanding of the FA pathway in molecular biology is constantly expanding, leading to fascinating discoveries about the mechanisms underlying the repair of interstrand crosslinks (ICLs). Ongoing research into this field has unveiled intriguing insights into the intricate processes involved. Furthermore, the identification of novel FA genes presents exciting prospects for therapeutic interventions, particularly through gene therapy and gene editing techniques. These approaches hold promise for complementing or rectifying specific gene defects associated with FA. In certain instances, knowledge of the specific gene and variant(s) can help identify potential risk factors and gain insights into the clinical progression of the disease. The medical management approach for most FA patients is determined based on their clinical presentation. However, for individuals with variants in genes that modify the clinical phenotypes, identifying the comprehensive genotypes becomes crucial for accurate medical management and prognosis of the disease, especially when these genes manifest FA-like symptoms distinct from the classic FA phenotype. It is important to note that the available genotype/phenotype information is often limited to a small number of cases, and there have been instances of atypical phenotypes that deviate from the conventional understanding.

While drug-based treatments are being explored, a definitive therapeutic strategy to fully restore the functionality of the FA pathway has not yet emerged. Although our current knowledge of the molecular biology of FA has provided valuable insights, further research is necessary to develop a comprehensive understanding of the intricate workings of this pathway. Such advancements will pave the way for the emergence of effective drug-based therapies capable of restoring FA pathway function. In this chapter, our analysis focuses on the exome sequencing data of FA patients who exhibit mutations in FA pathway genes, as outlined in Chapter 1. As previously mentioned, it has been observed that the FA phenotype may arise from mutations in genes other than FA genes. Therefore, our analysis aimed to explore the presence of additional gene mutations that could contribute to the manifestation of the FA phenotype in these patients. By examining the exome sequencing data, we aim to uncover potential genetic factors beyond the FA pathway genes that may play a role in the development of the particular FA phenotype.

4.2.1. Genotype-phenotype correlation

The median age at diagnosis in the patients with *FANCL* (median 6; range 1-24) ($p<0.0001$) and *FANCG* (median 7; range 1-30) ($p<0.001$) mutations was significantly lower compared to patients with *FANCA* (median 12; range 0-39) mutations **(Figure 4.2.1A)**. Genotype-phenotype correlation was performed for patients with *FANCA*, *FANCL*, and *FANCG* mutations as they were the largest group in the cohort.

A B

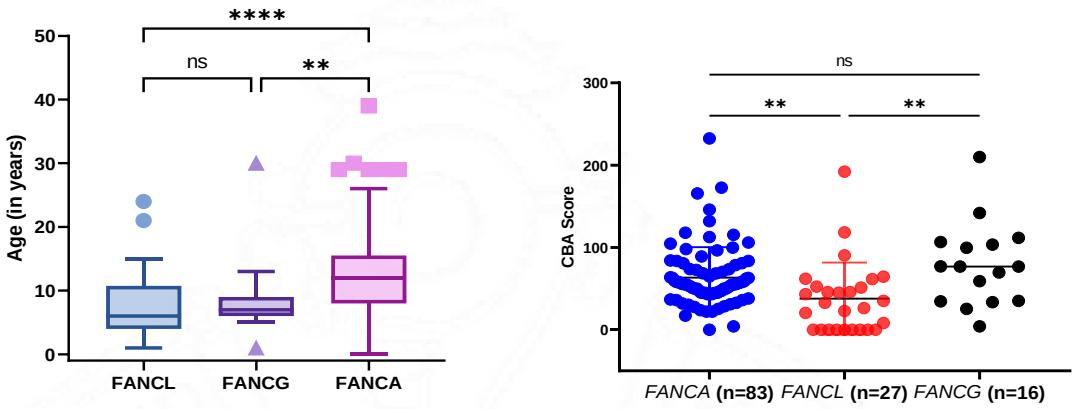


Figure 4.2.1. Age of disease presentation and association with CBA scores in predominant FA genetic subtypes. **(A)** Age of disease presentation in predominant genetic subtypes. **(B)** Association of CBA scores in predominant genetic subtypes.

4.2.2. Physical abnormalities

FANCG patients had a significantly high prevalence of kidney anomalies (8/18; 44.4%) compared to patients with *FANCL* (2/22; 11.1%) and *FANCA* (12/65; 18.4%) mutations (p=0.01). No association was found between pigmentation defects and the complementation groups (p=0.08) *FANCA* (71/86; 82%), *FANCL* (27/28; 96.4%), and *FANCG* (14/19; 73.6%). There was no correlation between radial ray abnormalities and defective genes (p=0.08) *FANCA* (37/73; 50.6%), *FANCL* (8/24; 33.3%), and *FANCG* (11/16; 68.5%). There was no correlation between microcephaly and defective genes (p=0.1) *FANCA* (19/86; 22%), *FANCL* (11/28; 39.3%), and *FANCG* (6/19; 31.6%). There was no correlation between eye abnormalities and defective genes (p=0.7) *FANCA* (19/86; 22%), *FANCL* (11/28; 39.3%), and *FANCG* (6/19; 31.6%). There was no correlation between short stature and defective genes (p=0.6) *FANCA* (22/86; 25.6%), *FANCL* (8/28; 28.6%), and *FANCG* (7/19; 36.8%) **(Table 4.2.1)**.

Table 4.2.1. Genotype phenotype association in FA patients

Clinical characteristics	Complementation group (% (n/total))			p-value
	FANCA	FANCL	FANCG	
Physical deformities				
Kidney anomalies	18.4% (12/65)	11.1% (2/22)	44.4% (8/18)	**0.01**
Pigmentation defects	82% (71/86)	96.4 (27/28)	73.6% (14/19)	0.08
Radial ray anomalies	50.6% (37/73)	33.3% (8/24)	68.5% (11/16)	0.08
Microcephaly	22% (19/86)	39.3% (11/28)	31.6% (6/19)	0.1
Eye anomalies	22% (19/86)	39.3% (11/28)	31.6% (6/19)	0.7
Short stature	25.6% (22/86)	28.6% (8/28)	36.8% (7/19)	0.6
Haematological parameters				
Type of aplastic anaemia	13.9% (12/86)	17.8% (5/28)	26.3% (5/19)	0.4
Occurrence of MDS/AML	22.1% (19/86)	3.5% (1/28)	15.7% (3/19)	0.07

Table 4.2.2. Involvement of FANCA exon mutations and type of mutations (all genotypes) in malignant transformation

Haematological parameters				
FANCA exons	**Function**	**n/total**	**Percentage**	**p-value**
Exon 1-19	BRCA1 interaction domain	8/19	42.10%	6.00E-04
Exon 24	Not defined	1/19	5.26%	
Exon 25	Not defined	1/19	5.26%	
Exon 27-30	Altinas et al.	8/19	42.10%	
Intron 31	Not defined	1/19	5.26%	
Type of mutations				
Type of mutation	**Mutation effect**	**Haematological malignancy (n/total)**	**Percentage**	**p-value**
Missense	Hypomorphic	11/19	37.90%	0.03
Splicing	Null	6/29	20.70%	
Deletion	Null	5/29	17.20%	
Frameshift	Null	2/29	6.90%	

4.2.3. Haematological parameters

Aplastic anaemia was characterised into non-severe aplastic anaemia (NSAA) and severe aplastic anaemia (SAA) (includes both very severe aplastic anaemia (VSAA) and severe aplastic anaemia (SAA)). There was no association between severe aplastic anaemia (SAA) and defective genes (P=0.4) *FANCA* (12/86; 13.9%), *FANCL* (5/28; 17.8%) and *FANCG* (5/19; 26.3%). There was a trend (p=0.07) towards higher occurrence of MDS/AML in patients with *FANCA* (19/86; 22.1%) and *FANCG* (3/19; 15.8%) defects compared to *FANCL* (1/28; 3.6%). As most of the MDS/AML (19/29; 65.5%) cases had mutations in the *FANCA* gene, we investigated the regions in the *FANCA* gene that may have an association with this transformation. *FANCA* exons were grouped as exons 1-19 (*BRCA1* interaction domain) (Altintas et al., 2023), exon 24, exon 25, exons 27-30 (Altintas et al., 2023) and intron 31. As expected, haematological malignancies were more observed in patients with mutations in exons 1-19 (8/19; 42.1%) (p=0.0006). Interestingly, there was an extremely high occurrence of MDS/AML in patients with mutations in exons 27-30 (8/19; 42.1%) (p=0.0006) **(Table 4.2.2)**.

Further investigation was performed to identify if a particular type of mutation is linked to transformation to haematological malignancy. We identified that missense (11/29; 37.9%) mutations have a significantly higher propensity (p=0.03) **(Table 4.2.2)** to develop haematological malignancy compared to splicing (6/29; 20.7%), deletion (5/29; 17.2%) and frameshift (2/29; 6.9%) mutations. Patients belonging to the *FANCA* complementation group exhibited significantly higher CBA scores, with a

median score of 55.4 (range: 0-232.4), while patients in the FANCG complementation group had a median score of 76.8 (range: 0-209.8). These scores were notably higher compared to patients in the *FANCL* complementation group, who had a median score of 32.8 (range: 0-192.2). The observed differences in CBA scores among the three complementation groups were statistically significant (p=0.002) **(Figure 4.2.1B)**. The variation in CBA scores further underscores the importance of complementation analysis in identifying specific genetic defects and subtypes within the broader spectrum of FA.

4.2.4. Integrating scRNA data to enhance exome data filtration for FA, HA, and AA Samples Strategy for NGS data analysis.

The analysis involved the examination of whole exome sequencing data from 140 FA patients in order to identify disease modifiers. Additionally, targeted panel sequencing data was available for 48 samples of haemolytic anaemia (AA) and 52 samples of aplastic anaemia (AA), which served as control groups. The targeted panel used in for clinical exome analysis encompassed a comprehensive set of 5226 genes. From this initial gene list, we further narrowed down the selection by including only those genes that exhibited at least one variant present across all 240 samples. After applying this filter, the final gene list consisted of 5115 genes, which met the criteria of having variants across the entire sample set. The gene list underwent additional filtering to include only homozygous variants while excluding variants located in the 5' untranslated region (5'UTR), 3' untranslated region (3'UTR), synonymous variants,

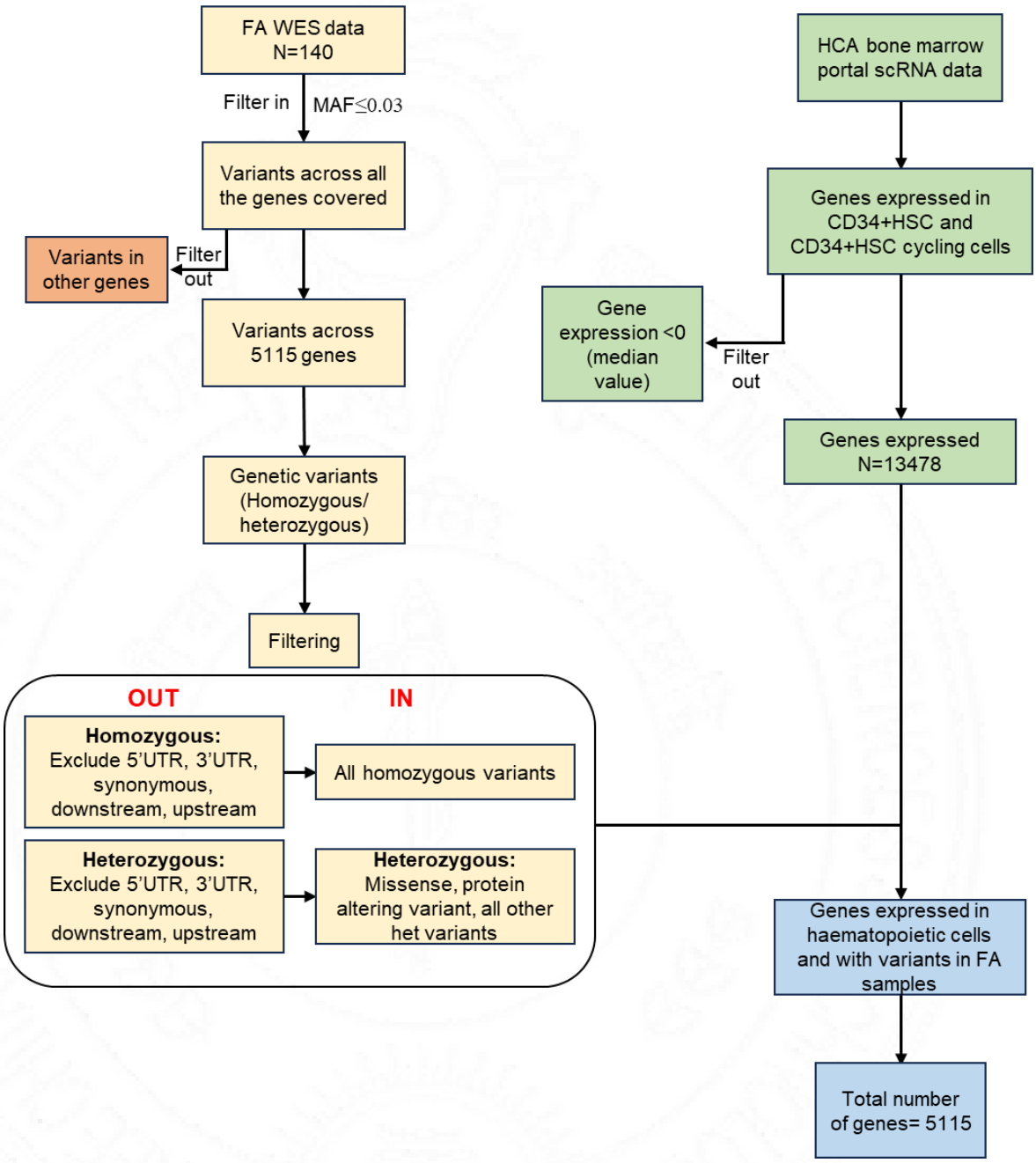
downstream variants, and upstream variants. The same filtering criteria were also applied to heterozygous variants **(Figure 4.2.2A)**. For the samples for which whole exome sequencing was carried out, variants within the selected 5115 genes were considered. The PCA plot analysis revealed that the majority of the HA samples exhibited tight clustering, indicating lower genotypic heterogeneity within this group **(Figure 4.2.2B)**. In contrast, the FA samples, although clustered together, did not form a compact cluster, indicating greater genotypic heterogeneity among these samples. The observed clustering pattern suggests that the HA samples share more similar genetic characteristics, while the FA samples exhibit a broader range of genetic variations. This finding underscores the presence of significant genotypic diversity within the FA samples, which may contribute to the phenotypic variability observed in this condition.

4.2.5. Single-cell RNA sequencing (scRNA) to identify the target genes for variant analysis.

Differentially expressed genes from scRNA-seq data were obtained from http://www.altanalyze.org/ICGS/HCA/splash.php. The dataset comprised 35 distinct cell types representing various hematopoietic lineages. For each cell type in this dataset, there were 8 replicates, consisting of 4 samples from male bone marrow and 4 samples from female bone marrow. To identify relevant genes, those with a median expression greater than 0 across all 8 samples (4 males and 4 females) in both CD34+HSC and CD34+HSC cycling cells were included in the analysis. In total, there

were 13,478 genes found to be expressed in CD34+HSC and CD34+HSC cycling cells. scRNA gene expression analysis was also performed for 22 FA genes across all the 35 sample types. Since the scRNA sequencing was conducted on healthy individuals who were not exposed to interstrand crosslinking agents, it resulted in moderate expression levels of most Fanconi Anaemia (FA) genes. However, among the FA genes analysed, *MAD2L2*, *RAD51C*, and *BRCA2* exhibited higher expression across all cell types studied. These particular genes stood out with elevated expression levels, indicating their potential importance in DNA repair and maintenance of genome stability even in the absence of exposure to interstrand crosslinking agents **(Figure 4.2.2C)**.

A


FA WES data
N=140
Filter in
MAF≤0.03
Variants across all the genes covered
Variants in other genes
Filter out
Variants across 5115 genes
Genetic variants (Homozygous/ heterozygous)
Filtering
OUT
IN
Homozygous:
Exclude 5'UTR, 3'UTR, synonymous, downstream, upstream
All homozygous variants
Heterozygous:
Exclude 5'UTR, 3'UTR, synonymous, downstream, upstream
Heterozygous:
Missense, protein altering variant, all other het variants
HCA bone marrow portal scRNA data
Genes expressed in CD34+HSC and CD34+HSC cycling cells
Gene expression <0 (median value)
Filter out
Genes expressed N=13478
Genes expressed in haematopoietic cells and with variants in FA samples
Total number of genes= 5115

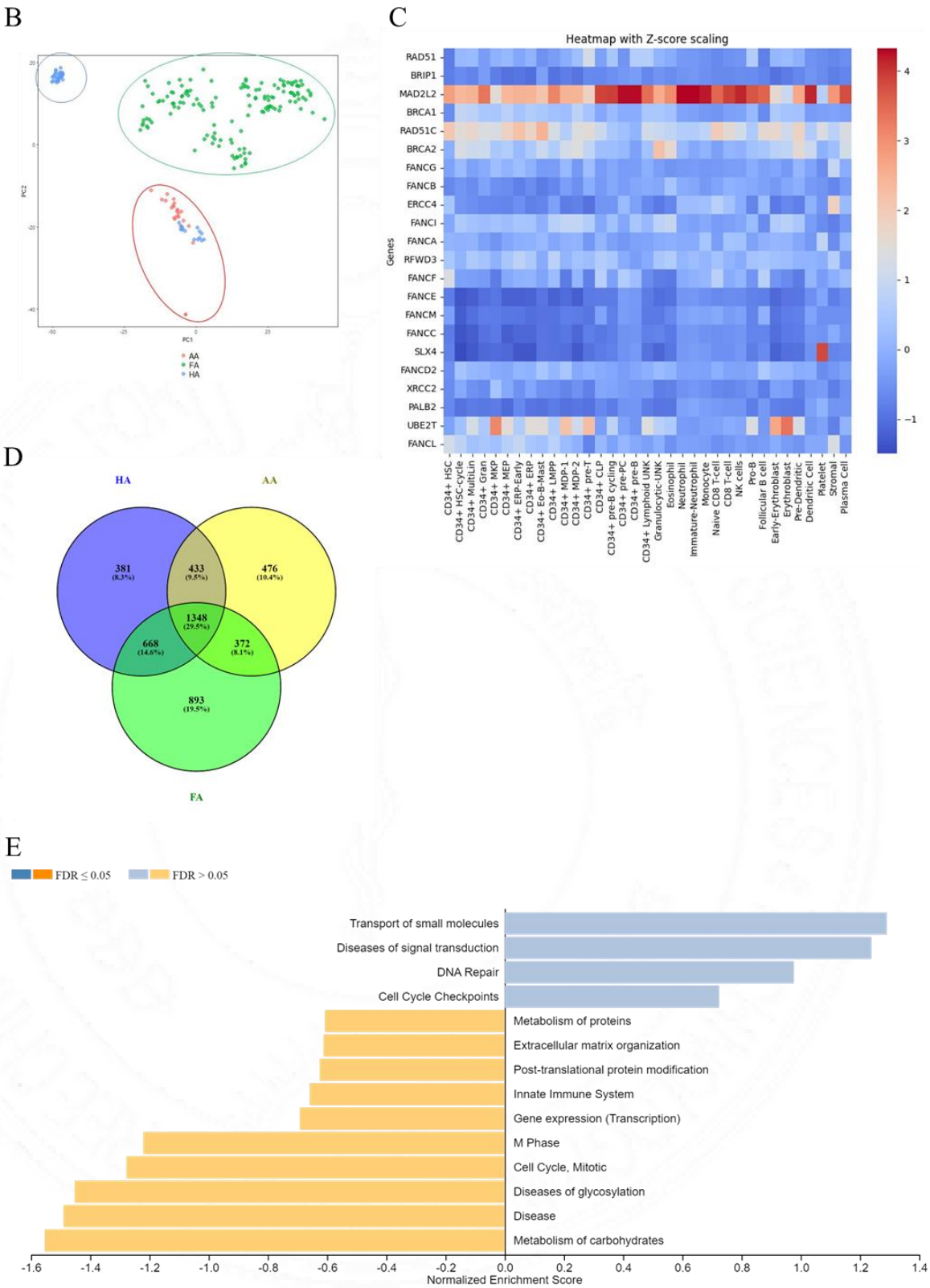

B
PC2
PC1
AA
FA
HA
C
Heatmap with Z-score scaling
Genes
RAD51
BRIP1
MAD2L2
BRCA1
RAD51C
BRCA2
FANCG
FANCB
ERCC4
FANCI
FANCA
RFWD3
FANCF
FANCE
FANCM
FANCC
SLX4
FANCD2
XRCC2
PALB2
UBE2T
FANCL
CD34+ HSC
CD34+ HSC-cycle
CD34+ MultiLin
CD34+ Gran
CD34+ MKP
CD34+ MEP
CD34+ ERP-Early
CD34+ ERP
CD34+ Eo-B-Mast
CD34+ LMPP
CD34+ MDP-1
CD34+ MDP-2
CD34+ preT
CD34+ CLP
CD34+ preB cycling
CD34+ pre-PC
CD34+ pre-B
CD34+ Lymphoid UNK
Granulocytic-UNK
Eosinophil
Neutrophil
Immature-Neutrophil
Monocyte
Naive CD8 T-cell
CD8 T-cell
NK cells
Pro-B
Follicular B cell
Early-Erythroblast
Erythroblast
Pre-Dendritic
Dendritic Cell
Platelet
Stromal
Plasma Cell
D
HA
AA
FA
381
(8.3%)
433
(9.5%)
476
(10.4%)
1348
(29.5%)
668
(14.6%)
372
(8.1%)
893
(19.5%)
E
FDR ≤ 0.05
FDR > 0.05
Transport of small molecules
Diseases of signal transduction
DNA Repair
Cell Cycle Checkpoints
Metabolism of proteins
Extracellular matrix organization
Post-translational protein modification
Innate Immune System
Gene expression (Transcription)
M Phase
Cell Cycle, Mitotic
Diseases of glycosylation
Disease
Metabolism of carbohydrates
Normalized Enrichment Score

Figure 4.2.2. Analysis of scRNA and exome sequencing data **(A)** Integrative bioinformatics pipeline for selecting haematopoietic-expressed genes and FA-associated variations from WES and scRNA Data. **(B)** Genetic variant based PCA plot reveals distinct clustering of Fanconi anaemia (FA), Haemolytic anaemia (HA), and aplastic anaemia (AA) samples. **(C)** Expression of 22 FA pathway genes in 35 haematopoietic lineages based on scRNA sequencing data analysis. **(D)** Identification of CD34+ HSC and cycling-HSC specific genes with FA-associated variants using venn diagram. (E) Pathways obtained after GSEA analysis using WEB-based GEne SeT AnaLysis Toolkit (https://www.webgestalt.org/).

4.2.6. Pathway analysis of CD34+ HSCs and CD34+ cycling HSCs expressed genes exclusively mutated in FA samples.

In the subsequent analysis, the focus was on genes that demonstrated significant expression in both CD34+ HSCs and CD34+ cycling HSCs. Additionally, these genes were required to have variants present across the 140 FA samples while showing no variants in the HA and AA samples. By applying these filtering criteria, a subset of genes meeting these conditions was selected for further investigation. This selection aimed to identify genes that exhibited differential expression patterns between the CD34+HSC and CD34+HSC cycling cells and harboured variants specifically associated with FA, indicating their potential involvement in the FA disease phenotype. These genes were considered as candidates for further analysis to explore their potential role as disease modifiers or contributors to the pathogenesis of FA. To ensure rigorous analysis, we employed stringent criteria of only selecting only genes with pathogenic variants by both SIFT and POLYPHEN. There were a total of 893 genes with pathogenic variants and exclusively mutated in FA samples **(Figure 4.2.2D)**. To increase the stringency of the criteria further, we exclusively chose variants that exhibited homozygous or compound heterozygous patterns in genes following autosomal recessive and X-linked recessive inheritance, as well as all variants present in genes with autosomal dominant inheritance. We identified a total of 120 genes for which Gene Set Enrichment Analysis (GSEA) analysis was performed. GSEA was performed based on the pathogenic variant burden on a gene. As anticipated, we observed a considerable variant burden in DNA repair and cell

cycle checkpoint genes. Surprisingly, we also noticed a lesser but still notable variant burden in genes related to metabolism and the immune system, both of which are known to be dysregulated in FA **(Figure 4.2.2E)**.

4.2.7. Patients with FA have additional unrecognised inherited bone marrow failure syndromes.

We filtered out FA patients exome sequencing data based on the variants present in 104 genes involved in Haematopoiesis, DNA damage repair, Ribosome biology disorders and Telomere biology disorders (McReynolds et al., 2022). Our aim was to elucidate the precise genetic changes occurring within the population of FA patients and their potential association with aberrations in these crucial cellular processes.

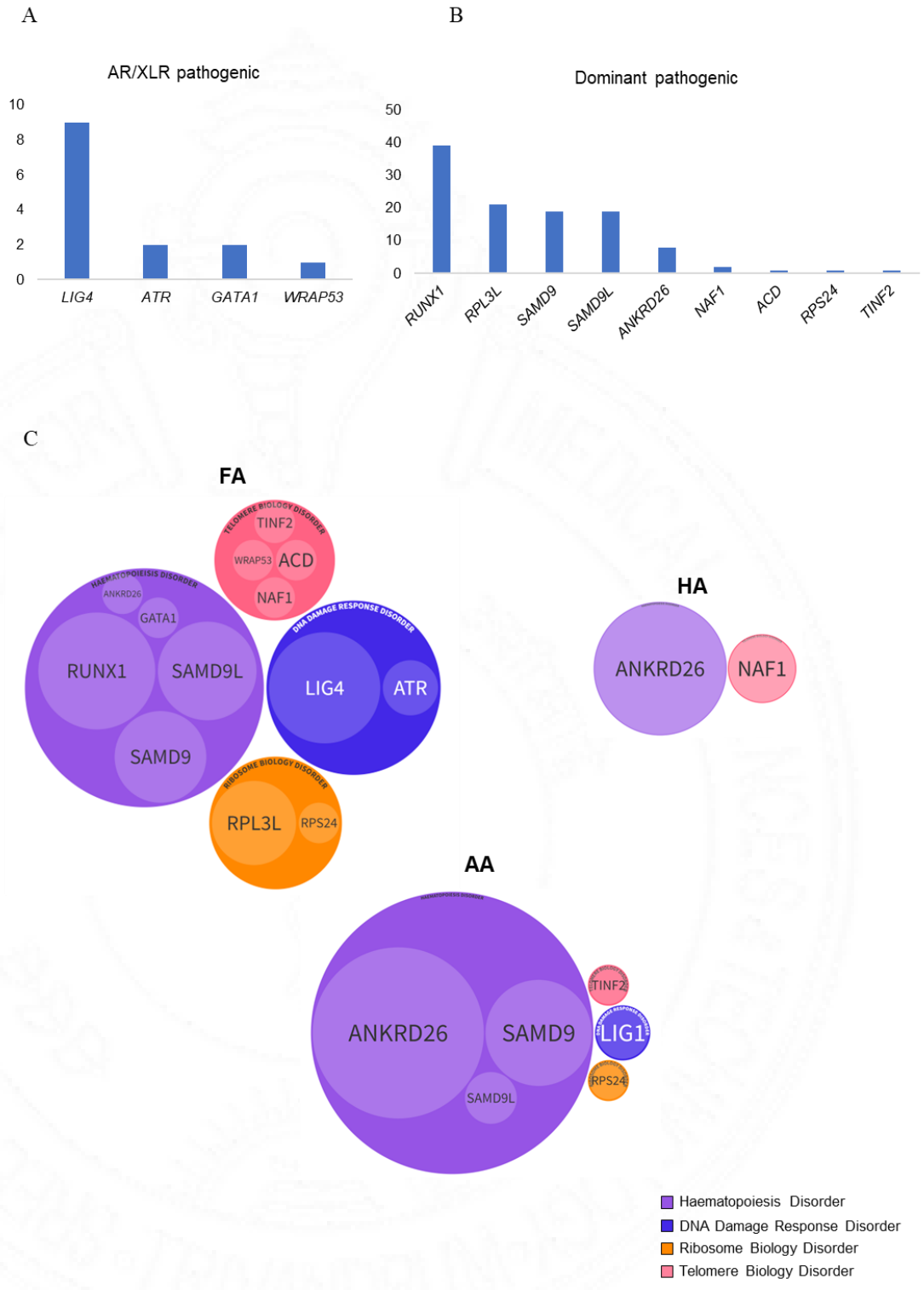
Upon analysing the filtered data, we anticipated discovering a significant proportion of variants in the genes directly associated with the FA pathway. We segregated genes based on their inheritance into two groups i.e., autosomal dominant and autosomal recessive with X-linked recessive. Remarkably, during our investigation of autosomal and X-linked recessive inheritance genes, we encountered two genes linked to DNA repair, specifically *LIG4* (n=9; 6.4%) and *ATR* (n=2; 1.4%), all of which exhibited pathogenic variants. Additionally, we also found pathogenic variants in gene *GATA1* (n=2; 1.4%) which is involved in haematopoiesis, that may lead to BMF. Moreover, we observed pathogenic variants in the gene *WRAP53* (n=1; 0.7%) which plays a crucial role in telomere biology **(Figure 4.2.3A)**. We then

focussed on genes that have dominant inheritance. Interestingly most of the pathogenic variants were present in genes involved in Haematopoiesis (*RUNX1* (n=10; 7.1%), *SAMD9* (n=6; 4.3%), *SAMD9L* (n=7; 5%) and *ANKRD26* (n=1; 0.7%)) followed by Ribosome biology genes (*RPL3L* (n=5; 3.6%) and *RPS24* (n=1; 0.7%)) and Telomere biology genes (*NAF1* (n=1; 0.7%), *ACD* (n=1; 0.7%) and *TINF2* (n=1; 0.7%)) **(Figure 4.2.3B)**.

Overall, considering variants in both autosomal dominant, autosomal recessive and X-linked recessive genes, we identified pathogenic variants in genes involved in maintenance of haematopoiesis (n=25; 17.8%). Of particular interest were the intriguing and unanticipated discoveries related to the presence of variants in genes associated with ribosome biology (n=6; 4.3%). The ribosome, a cellular organelle responsible for protein synthesis, is generally not considered a primary target in the context of FA. Nevertheless, our analysis uncovered variants in genes that play essential roles in ribosome biology. This discovery suggests a potential connection between ribosome dysfunction and the pathogenesis of FA or its associated symptoms. Pathogenic variants were also present in DNA damage response genes (n=11; 7.8%) and in telomere biology genes (n=4; 2.8%) of the FA patients **(Figure 4.2.3C and Figure 4.2.3D)**. In patients with other blood disorders such as HA that do not exhibit bone marrow failure, only a small number of patients had presence of pathogenic variants in gene involved in haematopoiesis and telomere biology **(Figure 4.2.3C)**.

These novel findings related to ribosome biology warrant further investigation and could provide valuable insights into previously unexplored mechanisms

underlying FA. Understanding the involvement of ribosome biology in FA pathogenesis could open up new avenues for therapeutic interventions or diagnostic approaches. Future studies could explore the functional consequences of these variants and their impact on ribosome function and overall cellular homeostasis.


A
AR/XLR pathogenic
10
8
6
4
2
0
LIG4
ATR
GATA1
WRAP53
B
Dominant pathogenic
50
40
30
20
10
0
RUNX1
RPL3L
SAMD9
SAMD9L
ANKRD26
NAF1
ACD
RPS24
TINF2
C
FA
TELOMERE BIOLOGY DISORDER
TINF2
WRAP53
ACD
NAF1
HAEMATOPOIESIS DISORDER
ANKRD26
GATA1
RUNX1
SAMD9L
SAMD9
DNA DAMAGE RESPONSE DISORDER
LIG4
ATR
RIBOSOME BIOLOGY DISORDER
RPL3L
RPS24
HA
ANKRD26
NAF1
AA
ANKRD26
SAMD9
SAMD9L
TINF2
LIG1
RPS24
Haematopoiesis Disorder
DNA Damage Response Disorder
Ribosome Biology Disorder
Telomere Biology Disorder

D

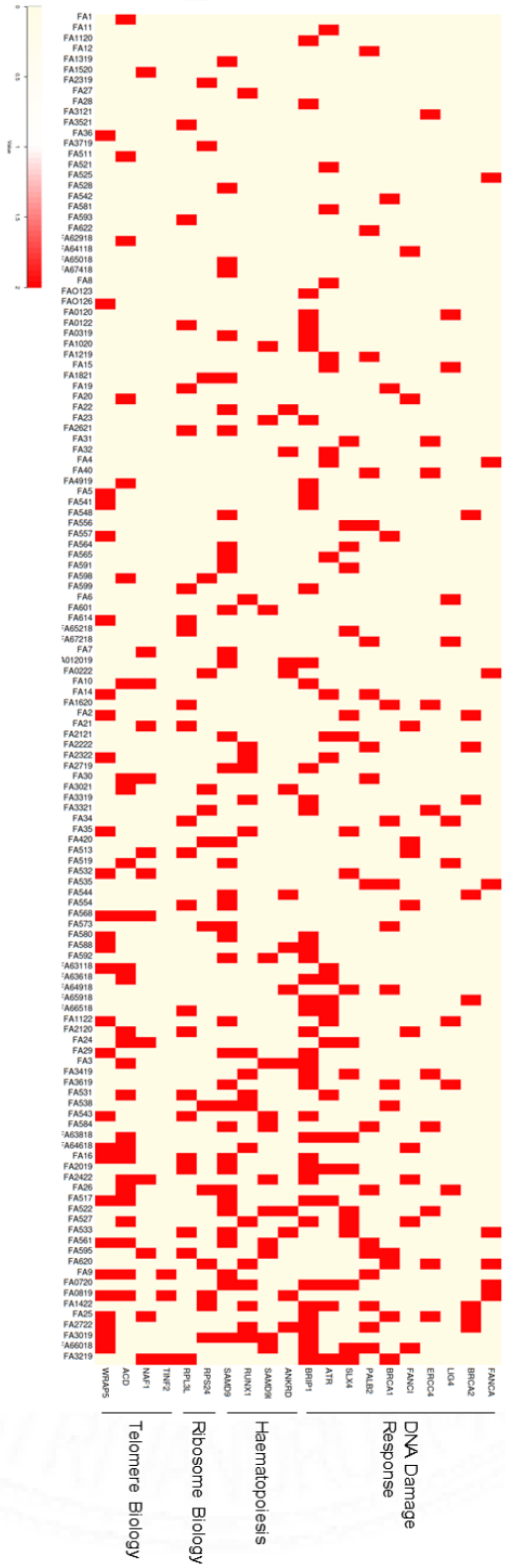

WRAP5
ACD
NAF1
TINF2
RPL3L
RPS24
SAMD9
RUNX1
SAMD9L
ANKRD
BRIP1
ATR
SLX4
PALB2
BRCA1
FANCI
ERCC4
LIG4
BRCA2
FANCA
Telomere Biology
Ribosome Biology
Haematopoiesis
DNA Damage Response

Figure 4.2.3: Genes defective in FA patients. **(A)** Defective genes with autosomal recessive and X-linked inheritance. **(B)** Defective genes with dominant inheritance. **(C)** Graphical representation of the number of patients defective in genes involved in maintenance of haematopoiesis, ribosome biogenesis, DNA damage response and telomere biology. **(D)** Pathogenic variants and genes associated with DNA Damage Repair in individual FA patients. AR: autosomal recessive, XLR:X-linked recessive, AD: Autosomal dominant

4.3 FA disease modelling using iPSCs to understand phenotypic heterogeneity.

Investigating the pathogenesis of FA in human hematopoietic progenitor cells (HSPCs) is challenging due to prevailing hypocellularity in the bone marrow of the patients. Transformed cell lines derived from FA patients (Donahue et al., 2003; Kruyt et al., 1996) have been used to study the cellular and molecular mechanisms underlying the disease, but they may not fully replicate the diverse disease phenotypes observed in patients (GH Liu et al., 2014). Moreover, studies in FA mouse models and human liver samples (obtained between 14 and 18 weeks of gestation from medical abortion after FA prenatal diagnosis) suggest that the fate of FA HSPCs is already determined during foetal liver haematopoiesis (Ceccaldi, Parmar, Mouly, Delord, Jung Min Kim, et al., 2012; Kamimae-Lanning et al., 2013). However, tracing earlier developmental events in humans is technically and ethically impossible. Mice models provide a limited approach to comprehending FA pathogenesis as they do not mimic the exact human haematopoietic phenotype (Bakker, de Winter, et al., 2013). Additionally, the knockout of single FA genes in mice typically does not produce HSC failure unless they experience replicative stress, which reinforces the need for human-based systems in which to validate observations made in mouse models (van de Vrugt and Joenje, 2014).

In this chapter, we explore the potential of induced pluripotent stem cells (iPSCs) as a tool for modelling and studying the pathogenesis of Fanconi Anaemia

(FA). However, iPSC-based FA modelling has proven challenging due to the requirement for an intact FA DNA repair pathway for effective reprogramming and induction and maintenance of pluripotency. The challenge in reprogramming FA cells could be overcome by introducing a wild-type cDNA of the defective FA pathway gene in the patients' somatic cells to restore the active FA pathway before reprogramming (Müller et al., 2012). However, permanent repair of FA pathway mutations in iPSCs can prevent the downstream derivation of FA-deficient HSPCs for disease modelling (Raya et al., 2009). Despite this limitation, patient-specific iPSCs have been generated using alternative reprogramming approaches, including complementation and CRISPR-Cas9 mediated gene correction (Osborn et al., 2016) of patients' cells before reprogramming (Müller et al., 2012; Raya et al., 2009). Inhibition of p53 by HPV E6 protein has also been used for reprogramming FA patient cells into iPSCs, but these colonies could not be maintained as cell lines (Chlon et al., 2014). One of the approaches includes the complementation of FA transgene with doxycycline-inducible lentiviral vectors (Bharathan et al., 2017; Chlon et al., 2016). In such cases, the FA fibroblasts could be reprogrammed successfully by doxycycline supplementation that activates the FA pathway. The iPSCs and differentiated haematopoietic progenitors could be used to study disease mechanisms after the inducer's withdrawal (Chlon et al., 2016; Marion et al., 2020). iPSCs have been generated from the fibroblasts of FA patients with defects in *FANCA, FANCC,* and *FANCI* genes (Bharathan et al., 2017; Chlon et al., 2016; G-H Liu et al., 2014; Müller et al., 2012; Raya et al., 2009; Yung et al., 2013). The inducible complementation

system to generate iPSCs has been reported only for the FANCA complementation group. It is crucial to generate FA -iPSCs for other complementation groups to understand the role of each protein in the FA pathway, especially since most proteins in this pathway have noncanonical functions that influence the FA phenotype in patients. Increasing evidence suggests that the noncanonical functions of the FA pathway proteins play a significant role in FA pathogenesis (Milletti et al., 2020).

Therefore, in this study, we describe the generation of patient-derived iPSCs using a tightly regulated doxycycline-inducible system for 6 FA complementation groups (*FANCA, -C, -F, -I, -L, -T*) and describe the cellular phenotypes resulting from the loss of the FA pathway.

4.3.1. Generation of a "non-leaky" doxycycline-inducible vector

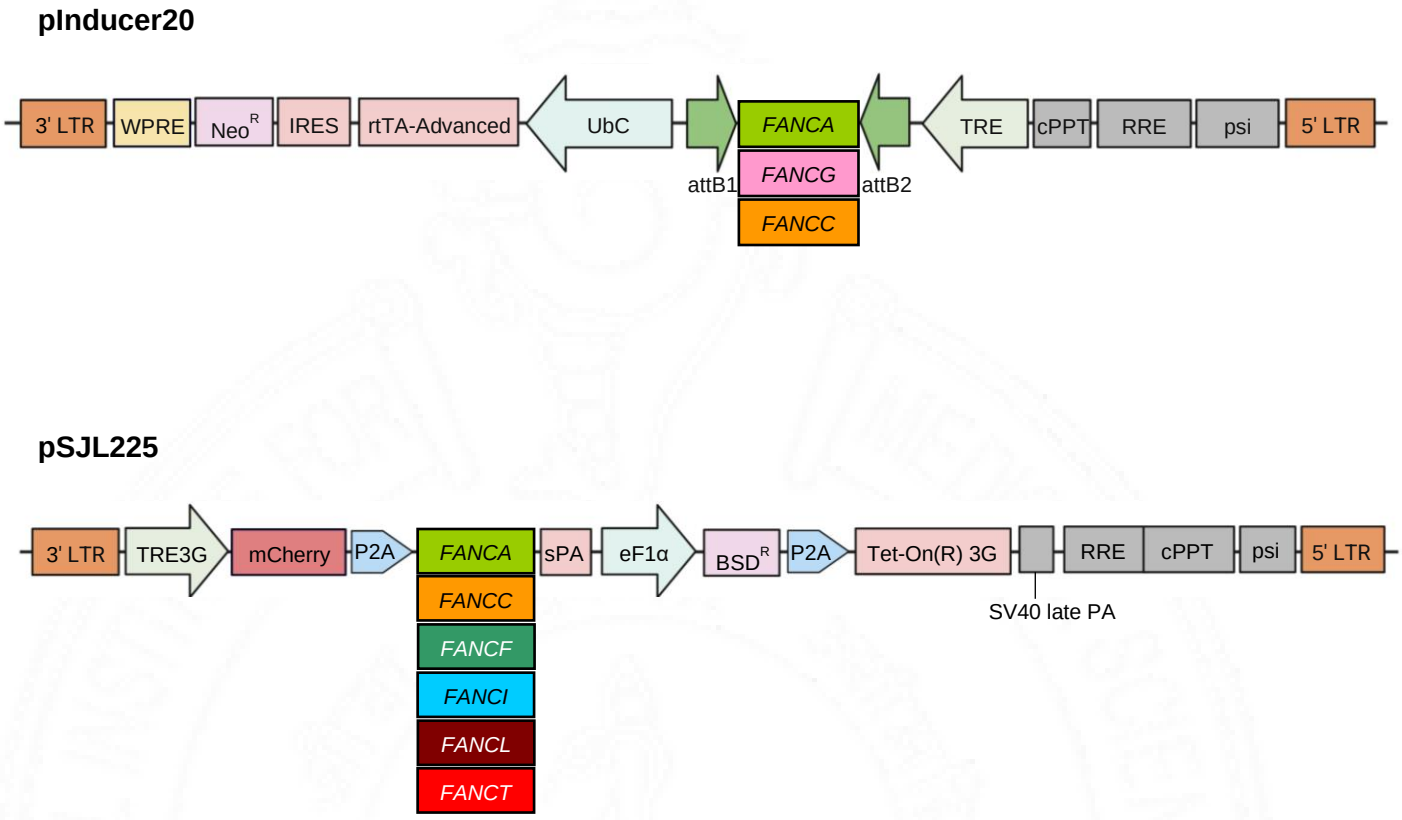
For inducible complementation, we initially used the pInducer20-FANCA vector **(Figure 4.3.1A)** (Meerbrey et al., 2011) to achieve doxycycline-inducible expression of the *FANCA* transgene in patient fibroblasts. Although this vector has been widely used by several research groups, we observed leaky expression of the transgene, as evidenced by the presence of *FANCA* transgene expression and FANCD2 monoubiquitination in the transduced cells even in the absence of doxycycline **(Figure 4.3.1B)**. pSJL225 doxycycline-inducible lentiviral vector system (a gift from St. Jude Children's Research Hospital, Memphis) harbours modified reverse tetracycline-controlled transactivator (rtTA) Tet-On 3G and tetracycline response element (TRE) promoter sequences for the expression of transgenes **(Figure 4.3.1A)**. The DNA

sequences of these elements are being constantly modified to generate robust doxycycline-inducible vectors for efficient and tight expression of transgenes (T. Das et al., 2016). Tet-On 3G sequence **(Figure 4.3.1C)** has three amino acid changes; p.S12G which increases doxycycline sensitivity and p.S67F and p.K71R which could further increase the sensitivity.

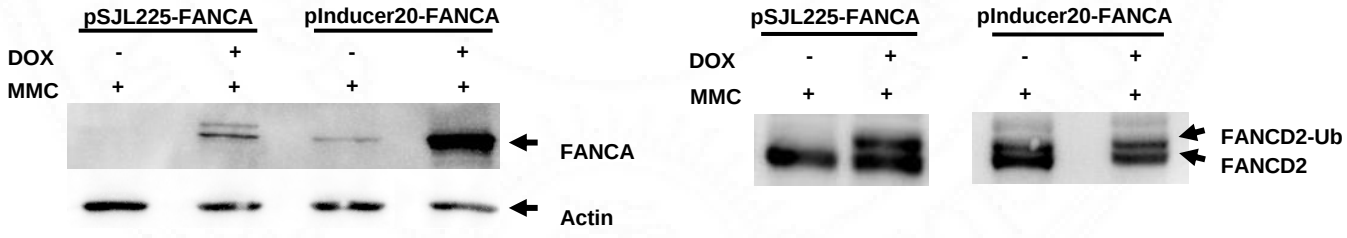
The inducible promoter TRE3G provides for very low basal expression and high maximal expression after induction (Loew et al., 2010). TRE and TRE3G consist of 7 repeats of a 19 bp *tet* operator sequence located upstream of a minimal CMV promoter, referred to as the tetracycline response element (TRE). Although the sequences of *tet* operator repeats are identical in all Tet-On generations, the junction sequences of TRE3G have been altered to an even spacing, and the central portions are randomized. Additionally, elements from the minimal CMV promoter have been mutated to consensus **(Figure 4.3.1D)**. Because TRE3G lacks binding sites for endogenous mammalian transcription factors and is inactive in the absence of a transactivator protein, basal expression from TRE3G is the lowest of any TRE-containing promoter available so far. Another major difference between pSJL225 and pInducer20 vectors is the direction of the expression cassettes for the constitutive expression of rTtA and the inducible expression of the transgene **(Figure 4.3.1A)**. Both the expression cassettes are in the 5'LTR to 3' LTR direction in the pInducer20 and in the 3’LTR to 5’LTR direction in the pSJ225 vector. 5’ LTR contains CMV promoter sequence, and we hypothesize that it could override the inducible expression in the absence of doxycycline **(Figure 4.3.1D)**. These lentiviral vectors were designed

to express the FANC gene and mCherry via a P2A self-cleaving peptide to monitor the doxycycline-inducible FANC complementation gene expression **(Figure 4.3.1A)**. The lack of mCherry expression in pSJL225 in the absence of doxycycline induction suggests that the vector's regulatory elements are effectively preventing any basal or background expression of mCherry, indicating tight control over gene expression.

A


pInducer20
3' LTR
WPRE
Neo^R
IRES
rtTA-Advanced
UbC
attB1
FANCA
FANCG
FANCC
attB2
TRE
cPPT
RRE
psi
5' LTR
pSJL225
3' LTR
TRE3G
mCherry
P2A
FANCA
FANCC
FANCF
FANCI
FANCL
FANCT
sPA
eF1α
BSD^R
P2A
Tet-On(R) 3G
SV40 late PA
RRE
cPPT
psi
5' LTR

B


pSJL225-FANCA
pInducer20-FANCA
DOX - + - +
MMC + + + +
FANCA
Actin
pSJL225-FANCA
pInducer20-FANCA
DOX - + - +
MMC + + + +
FANCD2-Ub
FANCD2

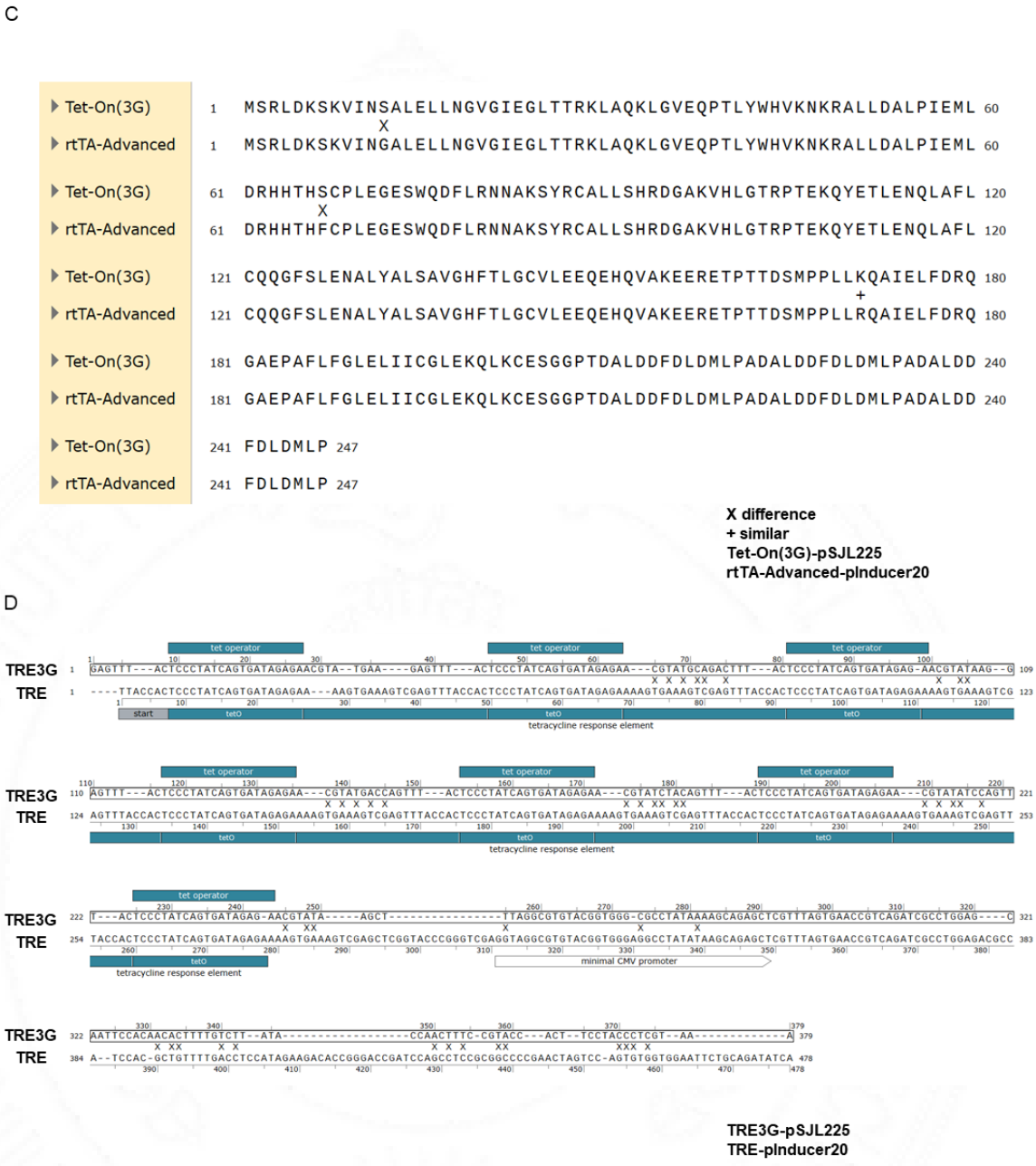


Figure 4.3.1. Doxycycline inducible lentiviral vectors for FA transgene expression. (**A)** Illustration of pInducer 20 and pSJL225 highlighting the essential components of the lentiviral vectors. (**B)** Western blot analysis demonstrating the expression of FANCA and FANCD2-ubiquitination in the transduced cells cultured in the presence and the absence of doxycycline. **(C)** Comparison of the reverse tetracycline-responsive transcriptional activator in pSJL225 and pInducer20, highlighting the crucial amino acid changes responsible for enhanced sensitivity. (**D)** Analysis of the Tet-responsive

promoter sequence variations between pSJL225 and pInducer20, highlighting the sequence differences that contribute to the tight regulation of transgene expression.

4.3.2. Generation of lentiviral vectors with doxycycline-inducible complementation of *FANCA, FANCC, FANCF, FANCI, FANCL,* and *FANCT* in patients' fibroblasts

We derived fibroblasts from the skin biopsies of the patients with mutations in *FANCA*, *FANCC, FANCF*, *FANCI, FANCL,* and *FANCT* genes **(Figure 4.3.2A-C)**. Normal cDNAs of these genes were cloned into pSJL225 doxycycline-inducible lentiviral vector **(Figure 4.3.1A)**. Patients' fibroblasts were transduced with pSJL225-*FANCA*, -*FANCC*, -*FANCF*, *FANCI*, -*FANCL* and -*FANCT* (**Figure 4.3.2D**). The blasticidin-selected transduced cells were used for the experiments. Western blot analysis confirmed the absence of protein expression by these mutations **(Figure 4.3.2E)**.

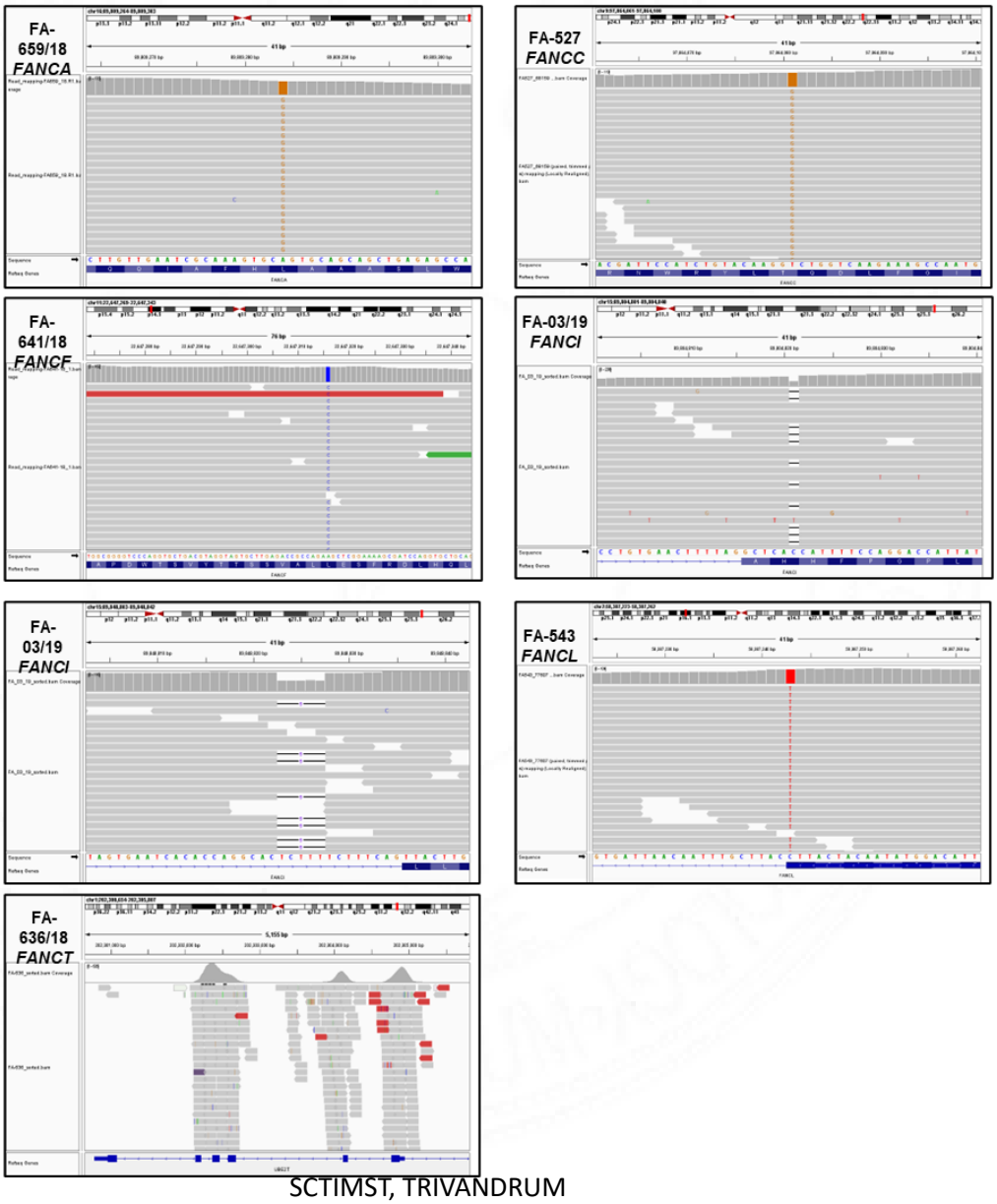
Since the patients' mutations were present in the genes in the FA upstream pathway, which is involved in the monoubiquitination of FANCD2 (FANCD2-Ub) during ICL repair, we carried out FANCD2-Ub analysis in the transduced fibroblasts to confirm successful complementation and FA-pathway activation. The transduced fibroblasts were cultured in the presence and absence of doxycycline and, subsequently, were treated with mitomycin C (MMC) to induce interstrand crosslinks (ICLs). The results showed that there was FA transgene expression and restoration of FANCD2-Ub in all 6 patients' fibroblasts after doxycycline supplementation **(Figure 4.3.2E)**.

The transduced cells demonstrated robust transgene expression, as measured by the increase in mCherry expression detected by flow cytometry upon doxycycline induction **(Figure 4.3.2E**, **Figure 4.3.2F).** In the absence of doxycycline, no detectable mCherry expression **(Figure 4.3.2F)** was observed, confirming tight regulation of transgene expression. Our results showed that the doxycycline-inducible expression lentiviral vectors that we generated for the expression of FANC genes do not exhibit "leaky" transgene expression in the transduced fibroblasts compared to the previously used doxycycline-inducible lentiviral expression vectors (Marion et al., 2020).

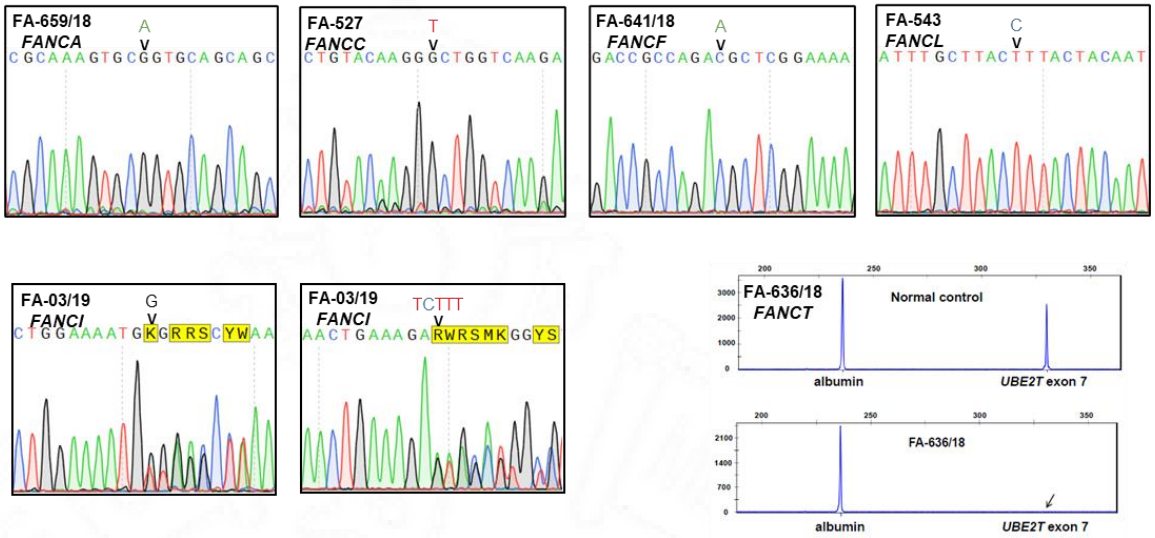
A

Sample ID	Gene	Type of mutation	cDNA change	Amino acid change	Zygosity
FA-659/18	*FANCA*	Missense	NM_000135.4:c.3689T>C	p.Leu1230Pro	Homozygous
FA-527	*FANCC*	Missense	NM_000136.3:c.1585A>C	p.Thr529Pro	Homozygous
FA-641/18	*FANCF*	Missense	NM_022725.4:c.41T>G	p.Leu14Arg	Homozygous
FA-03/19	*FANCI*	Frameshift deletion	NM_001113378.2:c.295delC	p.His99IlefsTer10	Compound heterozygous
		Intronic splice acceptor variant	NM_001113378.2:c.3256-8_3256-4delTCTTT	-	
FA-543	*FANCL*	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	Homozygous
FA-636/18	*FANCT*	Deletion	NM_014176:c.469_*178del	-	Homozygous

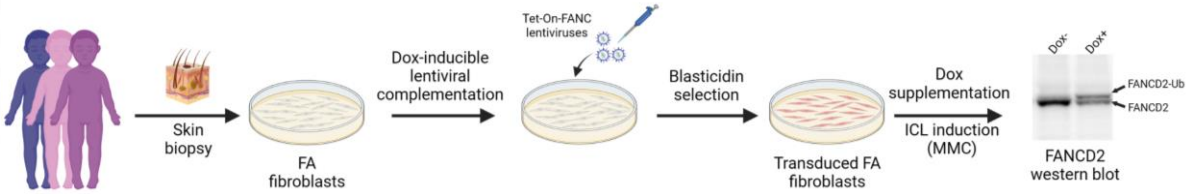
B



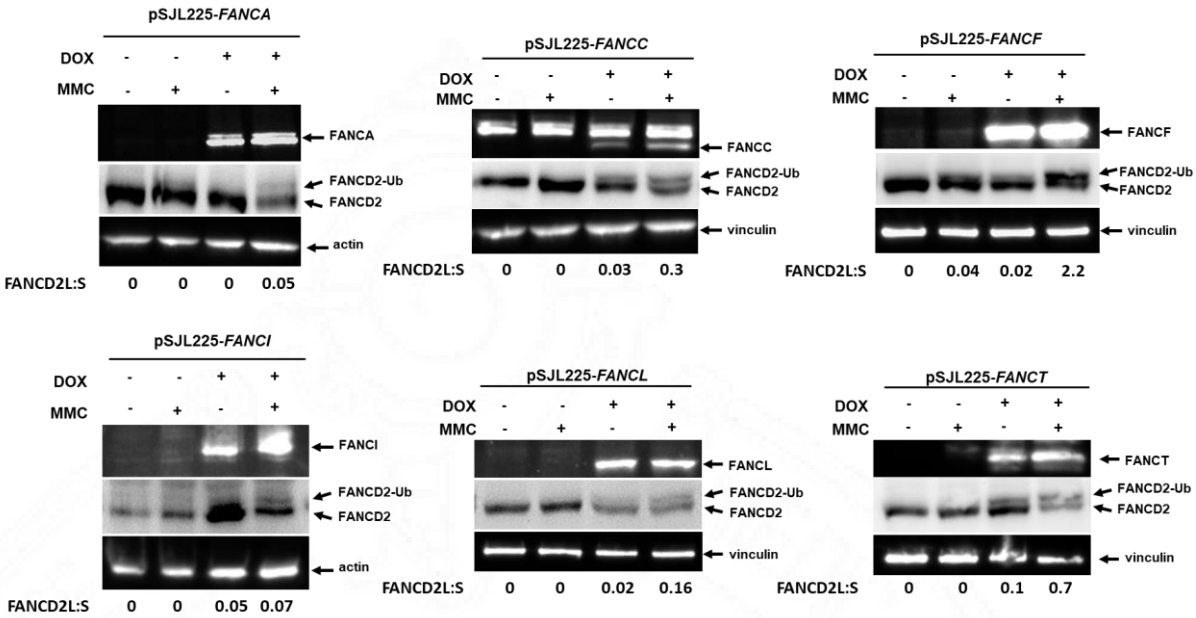
C


FA-659/18
FANCA
FA-527
FANCC
FA-641/18
FANCF
FA-543
FANCL
FA-03/19
FANCI
FA-03/19
FANCI
FA-636/18
FANCT
Normal control
albumin
UBE2T exon 7
FA-636/18
albumin
UBE2T exon 7

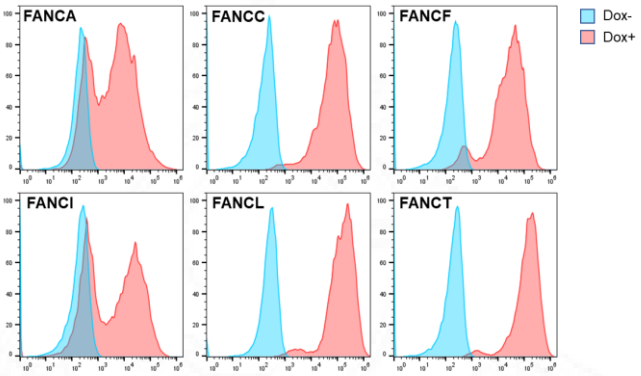
D


Skin biopsy
FA fibroblasts
Dox-inducible lentiviral complementation
Tet-On-FANC lentiviruses
Blasticidin selection
Transduced FA fibroblasts
Dox supplementation
ICL induction (MMC)
Dox-
Dox+
FANCD2-Ub
FANCD2
FANCD2 western blot

E


pSJL225-FANCA
DOX - - + +
MMC - + - +
FANCA
FANCD2-Ub
FANCD2
actin
FANCD2L:S 0 0 0 0.05
pSJL225-FANCC
DOX - - + +
MMC - + - +
FANCC
FANCD2-Ub
FANCD2
vinculin
FANCD2L:S 0 0 0.03 0.3
pSJL225-FANCF
DOX - - + +
MMC - + - +
FANCF
FANCD2-Ub
FANCD2
vinculin
FANCD2L:S 0 0.04 0.02 2.2
pSJL225-FANCI
DOX - - + +
MMC - + - +
FANCI
FANCD2-Ub
FANCD2
actin
FANCD2L:S 0 0 0.05 0.07
pSJL225-FANCL
DOX - - + +
MMC - + - +
FANCL
FANCD2-Ub
FANCD2
vinculin
FANCD2L:S 0 0 0.02 0.16
pSJL225-FANCT
DOX - - + +
MMC - + - +
FANCT
FANCD2-Ub
FANCD2
vinculin
FANCD2L:S 0 0 0.1 0.7

F



G

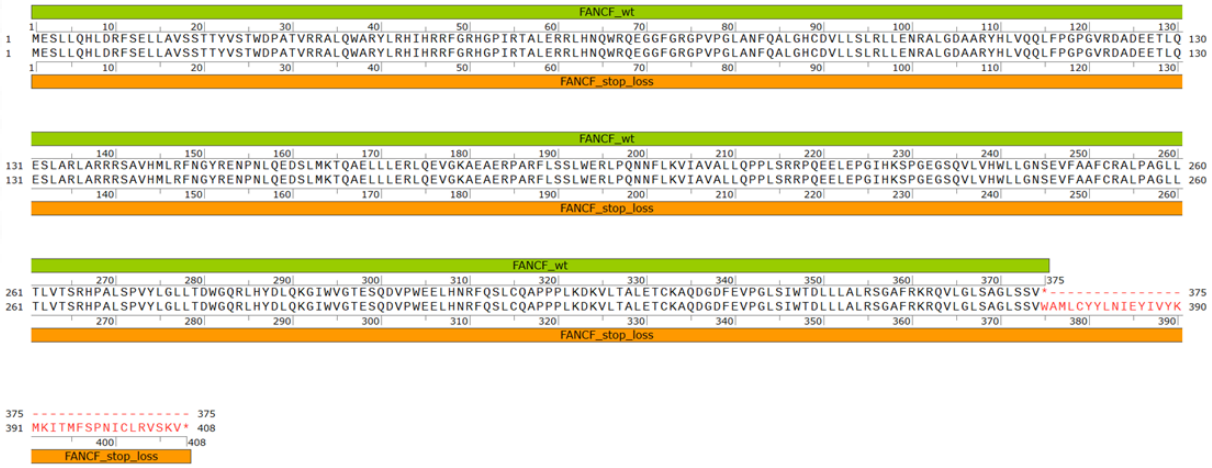


Figure 4.3.2. Genotypes of FA fibroblasts used for complementation using dox-inducible lentiviral vectors. **(A)** Tabular representation of FA patients' genotype. **(B)** Integrative Genomics Viewer (IGV) images of the mutations identified in the patients. **(C)** Sanger sequencing and gene dosage analysis to detect mutations identified by exome sequencing. **(D)** Schematic representation of inducible expression normal cDNAs of *FANC* genes in the fibroblasts of FA patients. The fibroblasts derived from the skin biopsy were transduced with the dox-inducible lentiviral vectors. The transduced cells were selected with blasticidin, treated with MMC to induce ICLs and analyzed for the restoration of FANCD2 ubiquitination in the dox supplemented fibroblasts. **(E)** Doxycycline induced FA transgene expression and restoration of FANCD2 monoubiquitination in patient fibroblasts after complementation with wt doxycycline-inducible expression of lentiviral FA transgene. **(F)** Flow analysis of mCherry expression in patient fibroblasts transduced with lentiviral FA transgene before and after doxycycline induction. **(G)** Protein sequence alignment between wild

type FANCF and mutated FANCF protein sequence shows an increase of 33 amino acids due to stop loss mutation. *Incomplete blasticidin selection

When treated with an ICL agent in the presence of doxycycline, the cells restored the active FA pathway by reconstituting the FA core complex responsible for FANCD2 monoubiquitination. Unexpectedly, there was a faint FANCF expression in the transduced cells cultured without doxycycline before and after MMC treatment **(Figure 4.3.2E)**. This occurrence could potentially be attributed to a stop loss mutation, leading to the production of a larger FANCF protein (46.3 kDa) **(Figure 4.3.2G)** with diminished functionality. The presence of faint FANCD2 monoubiquitination in FANCF-/- fibroblasts following lentiviral transduction further confirmed these findings **(Figure 4.3.2E)**. Interestingly, ICL induction by MMC treatment was not required for the *FANCC, FANCF, FANCI, FANCL,* and *FANCT* complemented cells to restore FANCD2-Ub, while the *FANCA* transduced could restore FANCD2-Ub only after MMC treatment. This difference may be due to the differential sensitivity of the FA-pathway genes to endogenous ICL agents.

4.3.3. Reprogramming of *FANCA, FANCC, FANCF, FANCI, FANCL,* and *FANCT* mutant patient-derived fibroblasts.

As the functional FA pathway is essential for successful cellular reprogramming, fibroblasts derived from FA patients fail to reprogram to generate iPSCs unless the fibroblasts were complemented with normal cDNAs of the genes mutated in the FA fibroblasts (Bharathan et al., 2017; Chlon et al., 2016; Raya et al., 2009). Therefore, we complemented the patients' fibroblasts with the doxycycline-

inducible lentiviral vectors to express the normal sequences of the FA genes and reprogramming was performed in the presence of doxycycline in the medium.

The fibroblasts were nucleofected with episomal plasmids (Okita et al., 2013) or transduced with Sendai viruses (Fusaki et al., 2009) for expressing reprogramming factors. We obtained various numbers of iPSC colonies formed after reprogramming, ranging from 3 to 45, from the transduced fibroblasts **(Table 4.3.1).** The efficiency of reprogramming using episomal vectors ranged from 0.0008 to 0.03%. We picked 3 to 6 iPSC colonies for each complemented gene and cultured them for >15 passages to generate stable iPSC lines. Thirty iPSC clones co-expressed mCherry with the *FANC* transgene from the lentiviral vector **(Figure 4.3.3A, Figure 4.3.3B; Table 4.3.1).** We generated 47 stable iPSC lines: 9 for *FANCC*, 15 for *FANCT*, 11 for *FANCF*, 5 for *FANCI*, 2 for *FANCL*, and 5 for *FANCA*.

After 15 passages to stabilize the iPSC lines, they were analysed for the cellular phenotypes in the presence and absence of doxycycline. As the vector used for complementation co-expressed mCherry with the FA transgene, we monitored mCherry expression and cell survival in the absence of an active FA pathway. For further experiments, we selected 1 iPSC clone generated from *FANCA* and *FANCI* complemented, and 2 iPSC clones from other complemented fibroblasts. The clones were chosen based on their ability to express mCherry upon doxycycline induction and exhibit cell death upon doxycycline withdrawal. The iPSCs were cultured without doxycycline for either 5 days (*FANCA*, *FANCC*, and *FANCT* complemented clones) or 8 days (*FANCL* complemented clone) until mCherry expression was completely

lost, while *FANCF* and *FANCI* complemented clones could be cultured without doxycycline for more than 10-15 days. Doxycycline withdrawal that results in the loss of active FA pathway did not affect pluripotency of any of the complemented iPSCs **(Figure 4.3.3C)**, suggesting that the observed cell death **(Figure 4.3.3D)** after doxycycline withdrawal in *FANCA*, *FANCC*, *FANCL*, and *FANCT* complemented clones was not due to the loss of pluripotency. However, *FANCF* and *FANCI* iPSCs did not show cell death after doxycycline withdrawal **(Figure 4.3.3D)**. The absence of cell death in FANCF-induced pluripotent stem cells (iPSCs) upon doxycycline withdrawal could be explained by the faint FANCF expression in the fibroblasts used for reprogramming. Conversely, the survival of *FANCI* clones without cell death after doxycycline withdrawal may suggest that *FANCI* is not a critical factor for iPSCs' viability. In the FA-iPSCs derived from *FANCA*, *FANCC*, *FANCI*, *FANCL*, and *FANCT* complementation groups, the presence of monoubiquitinated FANCD2 was exclusively detected following treatment with hydroxyurea (HU) **(Figure 4.3.3E)**. This finding strongly suggests that the Fanconi Anaemia (FA) pathway becomes active specifically in response to DNA damage**.** However, the cell death after doxycycline withdrawal occurred without HU treatment indicating that endogenous ICL-inducing metabolites might not be cleared in the absence of the FA pathway.

After treating with HU and MMC to induce ICLs, we observed FANCD2-Ub and FANCD2-γH2AX co-localization foci in the iPSCs cultured in the presence of doxycycline but not in the iPSCs cultured in the absence of doxycycline **(Figure 4.3.3E, Figure 4.3.3F).** Importantly, monoubiquitinated FANCD2 and FANCD2 foci

could be detected in the cells cultured in the presence of doxycycline only after treatment with hydroxyurea and MMC respectively, indicating that the FA pathway is only activated in the presence of DNA damaging agent in iPSCs.

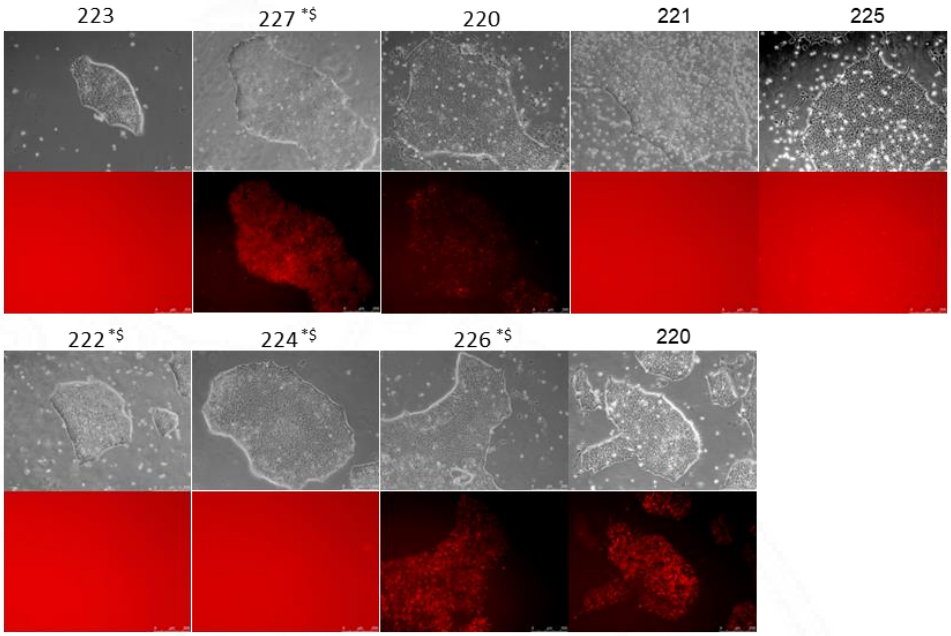
Table 4.3.1. Mode of reprogramming and mCherry expression in iPSC clones generated from FA patients' fibroblasts.

Complementation group	Reprogramming method	Total number of colonies obtained	mCherry + colonies	No of colonies picked up for culture and passaging	Trilineage differentiation	Pluripotency immunofluorescence	Karyotyping	Cell cycle analysis	Tra-I-60 (after dox withdrawal)	AP staining	ϒH2AX-FANCD2 colocalization
FANCA	Episomal	9	0	2	0	0	0	0	0	0	0
	Sendai	3	3	3	1	1	1	1	1	1	1
FANCC	Episomal	52	5	9	2	2	2	2	2	1	2
FANCF	Episomal	12	8	8	0	0	0	0	0	0	0
	Sendai	No data available	3	3	2	2	2	2	2	1	2
FANCI	Episomal	35	1	5	1	1	1	1	1	1	1
FANCL	Episomal	3	2	2	2	2	2	2	2	1	0
FANCT	Episomal	27	8	11	2	2	2	2	2	1	2
	Sendai	No data available	NA	4	0	0	0	0	0	0	0

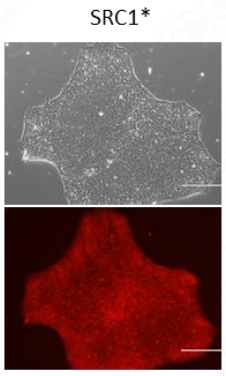
Note: A-absent, P-present and ND- not done

A

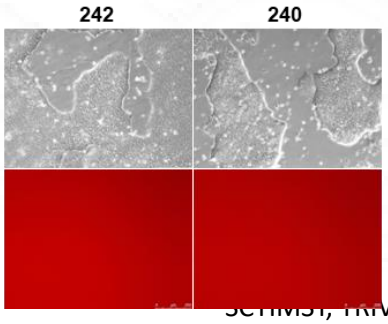
FA641/18 (*FANCF*)



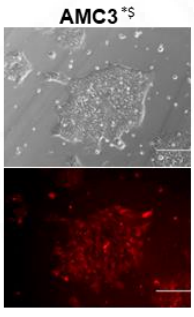
FA543 (*FANCL*)



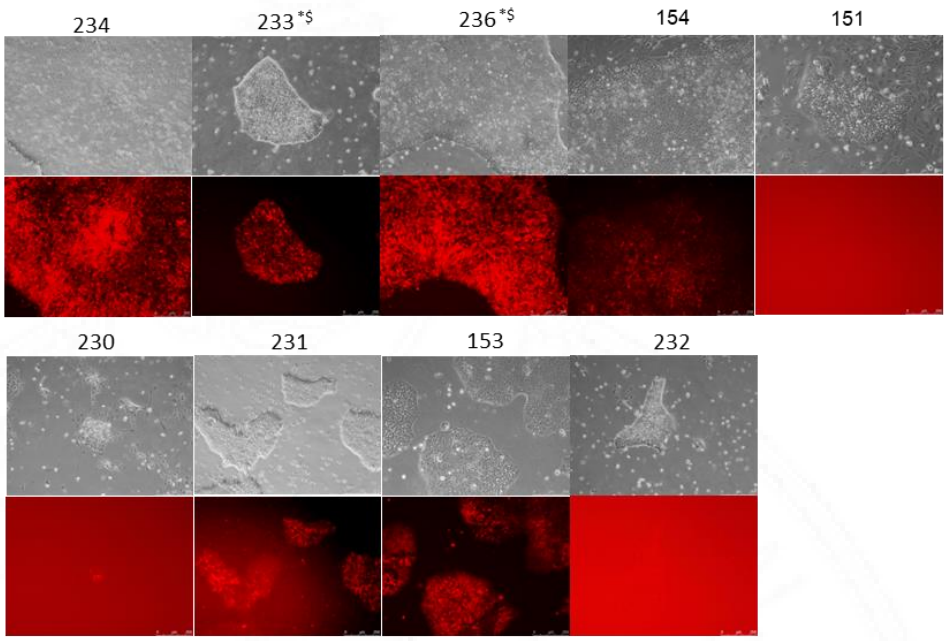
FA-24 (*FANCA*)



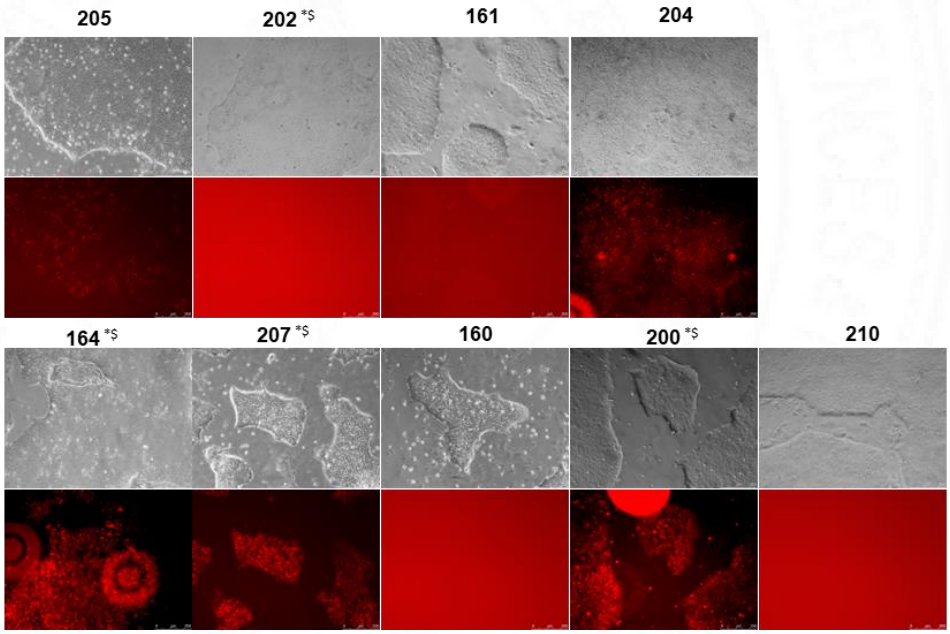
FA-03/09 (*FANCI*)



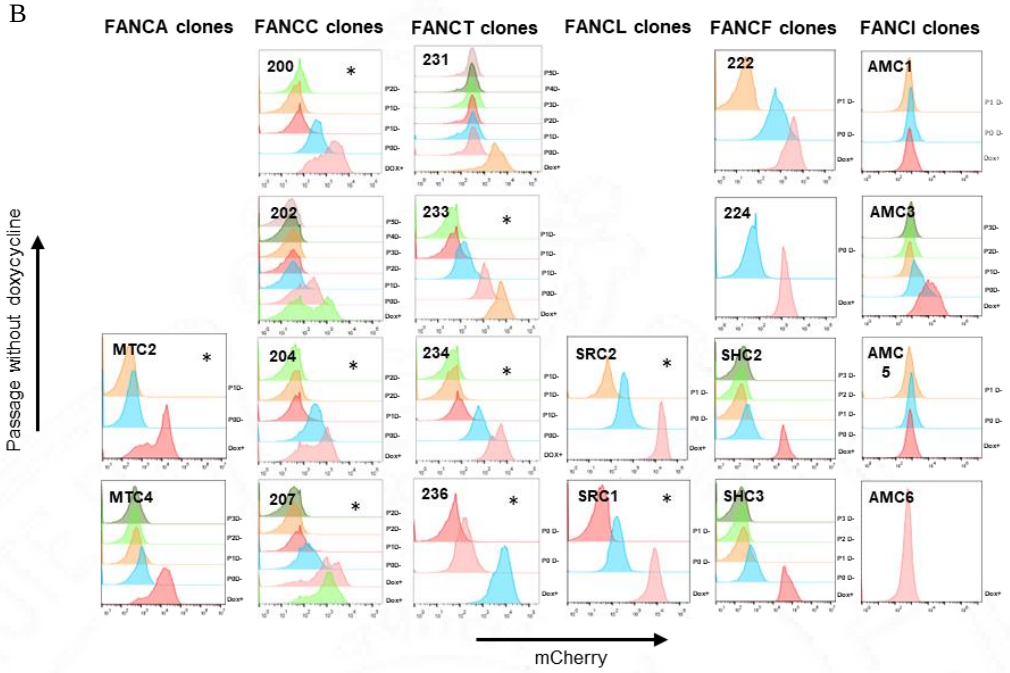
FA636 (*FANCT*)



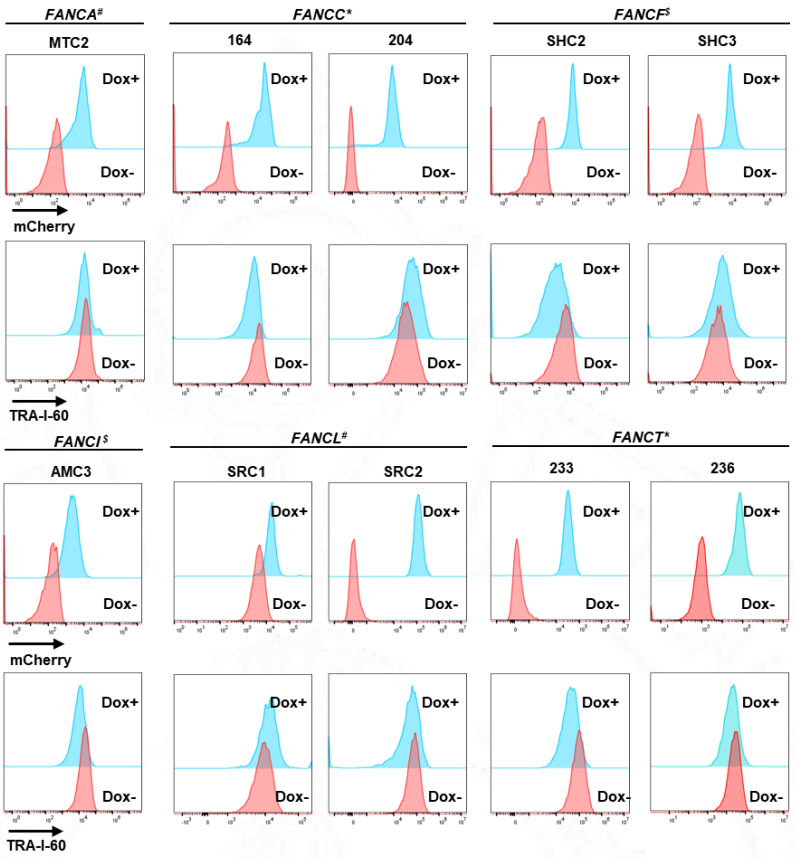
FA527 (*FANCC*)



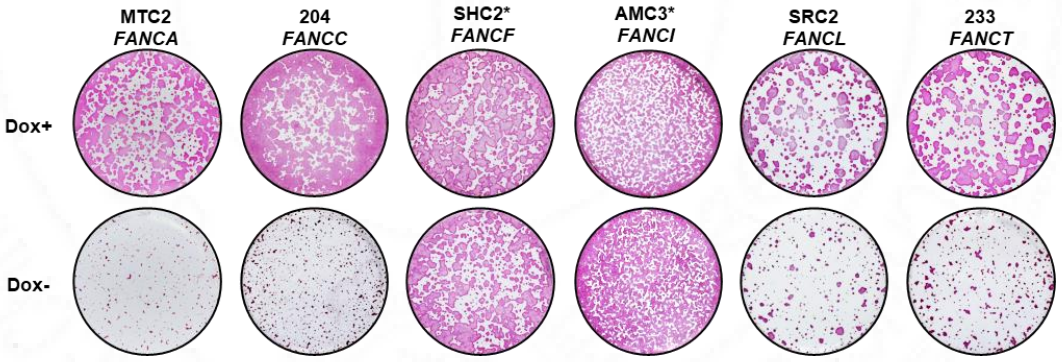
B


FANCA clones
FANCC clones
FANCT clones
FANCL clones
FANCF clones
FANCI clones
200
231
222
AMC1
202
233
224
AMC3
MTC2
204
234
SRC2
SHC2
AMC 5
MTC4
207
236
SRC1
SHC3
AMC6
Passage without doxycycline
mCherry

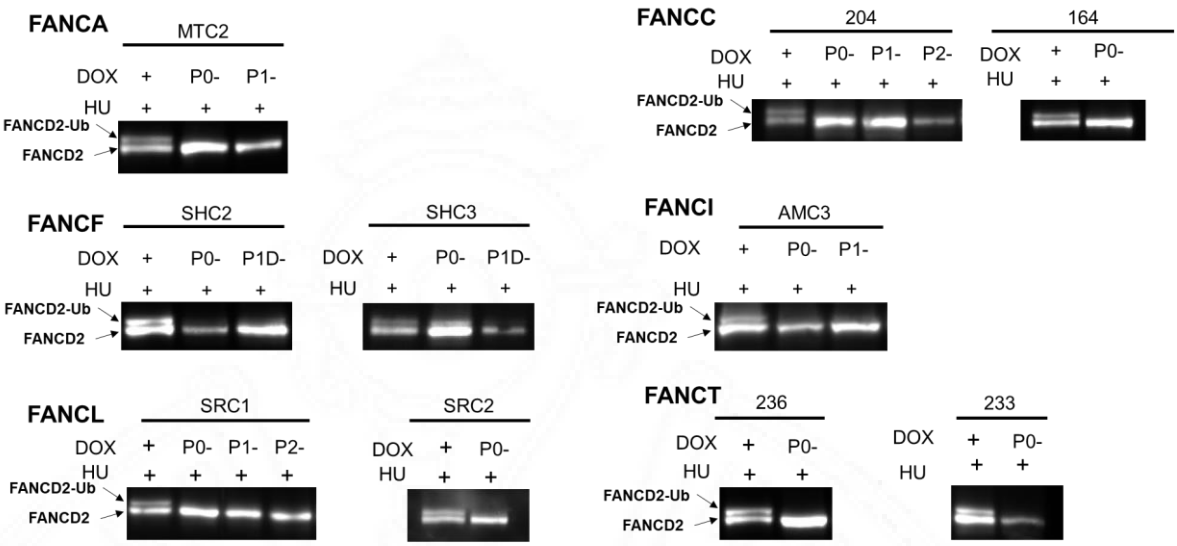
C


FANCA#
MTC2
FANCC*
164
204
FANCF$
SHC2
SHC3
Dox+
Dox-
mCherry
TRA-I-60
FANCI$
AMC3
FANCL#
SRC1
SRC2
FANCT*
233
236
*Cell death starts at day 5 of Dox withdrawal
#Cell death starts at day 8 of Dox withdrawal
$Cells do not die after dox withdrawal

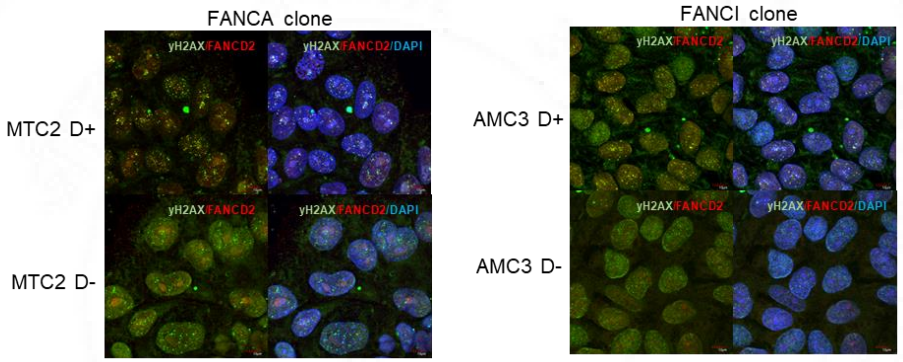
D


MTC2
FANCA
204
FANCC
SHC2*
FANCF
AMC3*
FANCI
SRC2
FANCL
233
FANCT
Dox+
Dox-

E


FANCA
MTC2
DOX + P0- P1-
HU + + +
FANCD2-Ub
FANCD2
FANCC
204
DOX + P0- P1- P2-
HU + + + +
FANCD2-Ub
FANCD2
164
DOX + P0-
HU + +
FANCF
SHC2
DOX + P0- P1D-
HU + + +
FANCD2-Ub
FANCD2
SHC3
DOX + P0- P1D-
HU + + +
FANCI
AMC3
DOX + P0- P1-
HU + + +
FANCD2-Ub
FANCD2
FANCL
SRC1
DOX + P0- P1- P2-
HU + + + +
FANCD2-Ub
FANCD2
SRC2
DOX + P0-
HU + +
FANCT
236
DOX + P0-
HU + +
FANCD2-Ub
FANCD2
233
DOX + P0-
HU + +

F


FANCA clone
MTC2 D+
MTC2 D-
γH2AX FANCD2
γH2AX FANCD2/DAPI
FANCI clone
AMC3 D+
AMC3 D-
γH2AX FANCD2
γH2AX FANCD2/DAPI

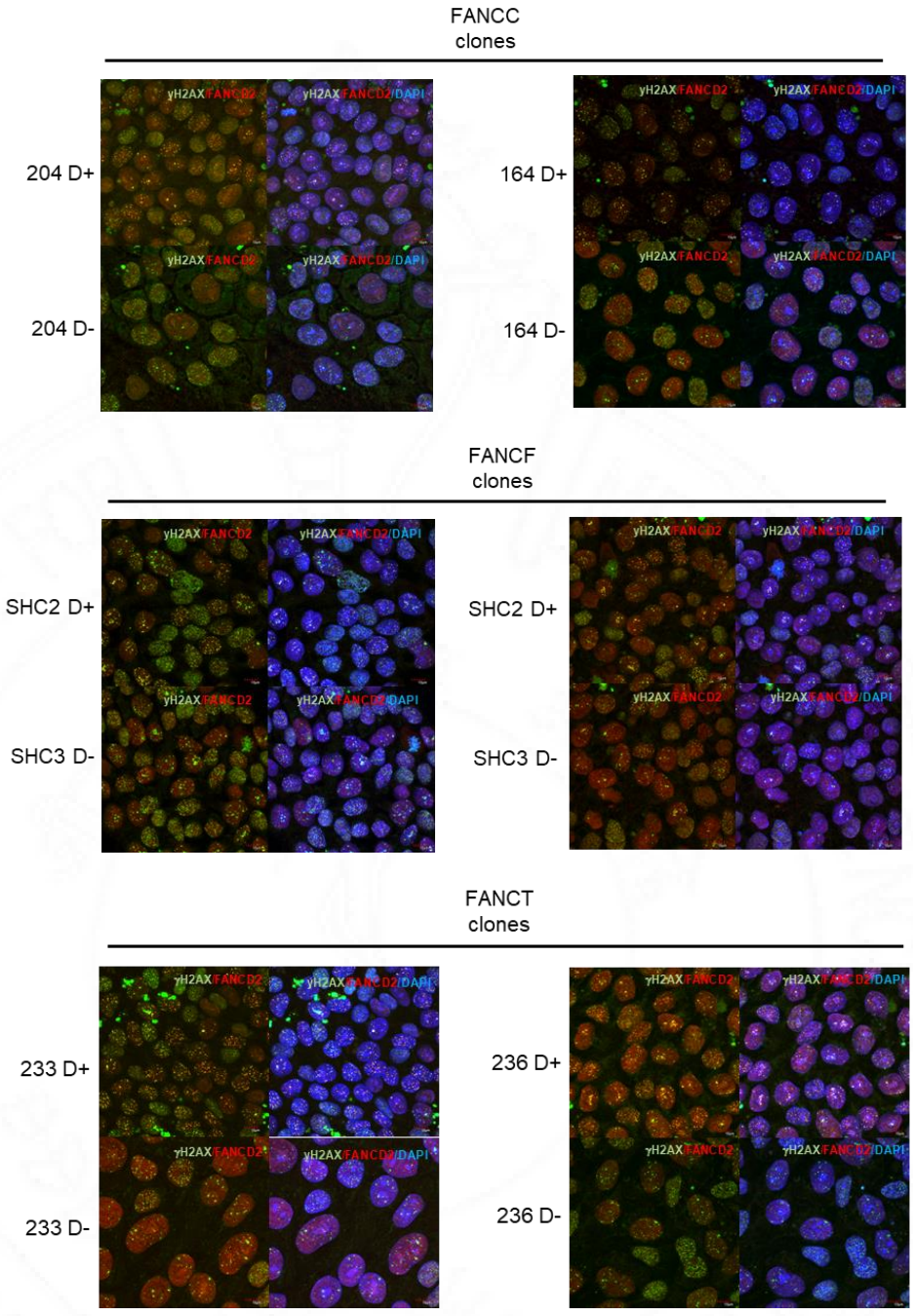


Figure 4.3.3. Analysis of cellular phenotypes and functional pathways in FA-iPSC clones under doxycycline treatment and withdrawal. **(A)** Microscopic images for FA-iPSC clones depicting morphology and mCherry expression in the presence of doxycycline. All images were taken while the cells were in culture with 100ng/ml doxycycline. **(B)** Doxycycline withdrawal in FA-iPSC clones showing reduction of mCherry expression in the clones showing cell death. **(C)** Pluripotency assessment of

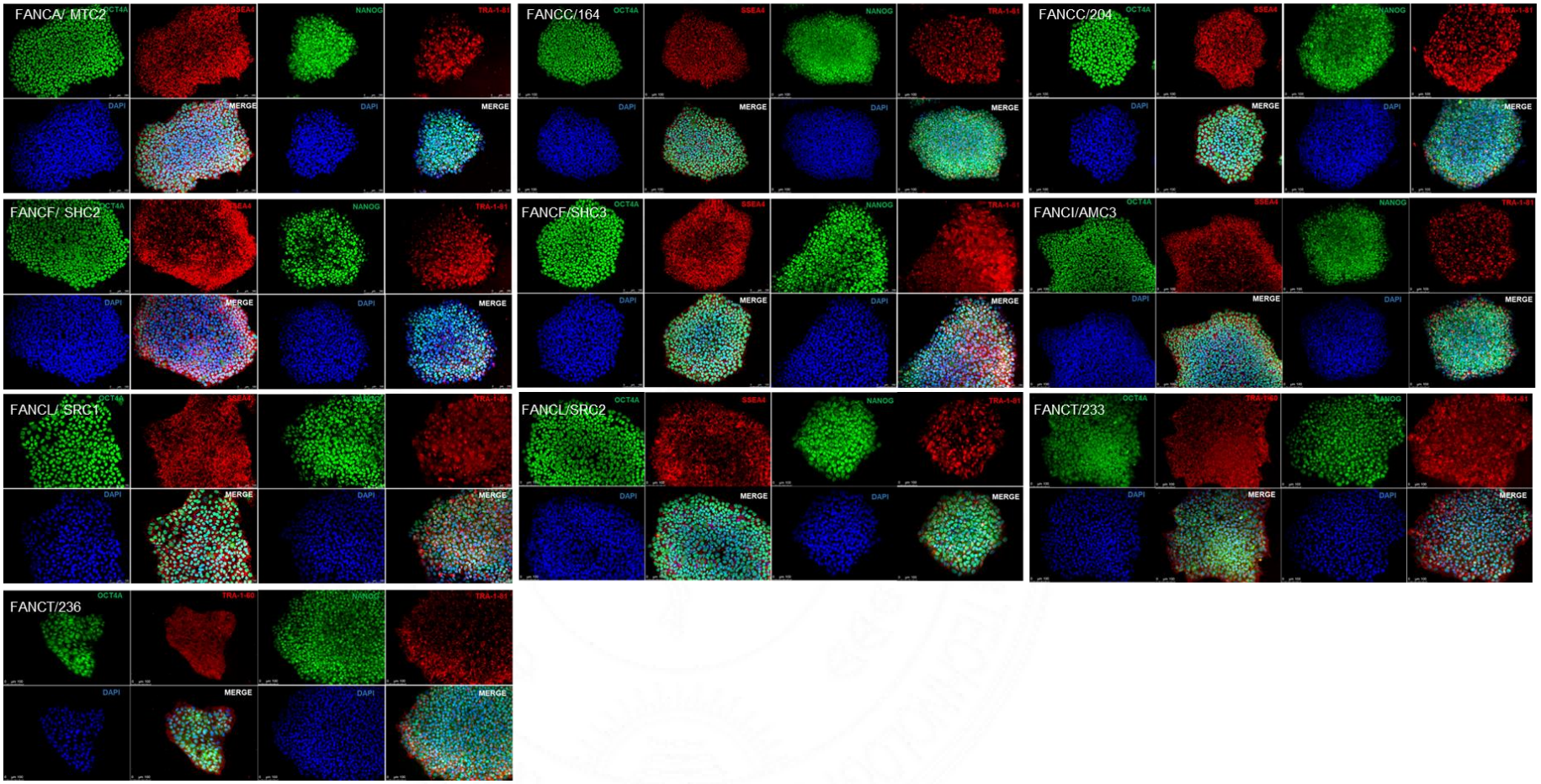
iPSCs with and without doxycycline revealing that FA-iPSCs maintained their pluripotency even in the absence of the FA pathway. **(D)** AP staining after dox withdrawal showing spontaneous cell death in *FANCA*, *FANCC*, *FANCL* and *FANCT* clones. **(E)** FANCD2 monoubiquitination analysis (FANCD2-Ub) after doxycycline withdrawal in the FA-iPSCs. **(F)** FANCD2 foci formation and FANCD2-γH2AX colocalization in cells cultured in the presence of doxycycline indicating activated functional FA pathway (yellow dots) in these cells. No FANCD2 foci or FANCD2-γH2AX colocalization was observed in cells cultured in absence of doxycycline indicating defective FA pathway (green dots). Cells were treated with 60nM MMC for 24 hours.

* Signifies cells death. #No cell death observed after doxycycline withdrawal. γH2AX: AF488 (green) FANCD2: dylight 550 (Red)

4.3.4. Characterisation of FA-iPSCs.

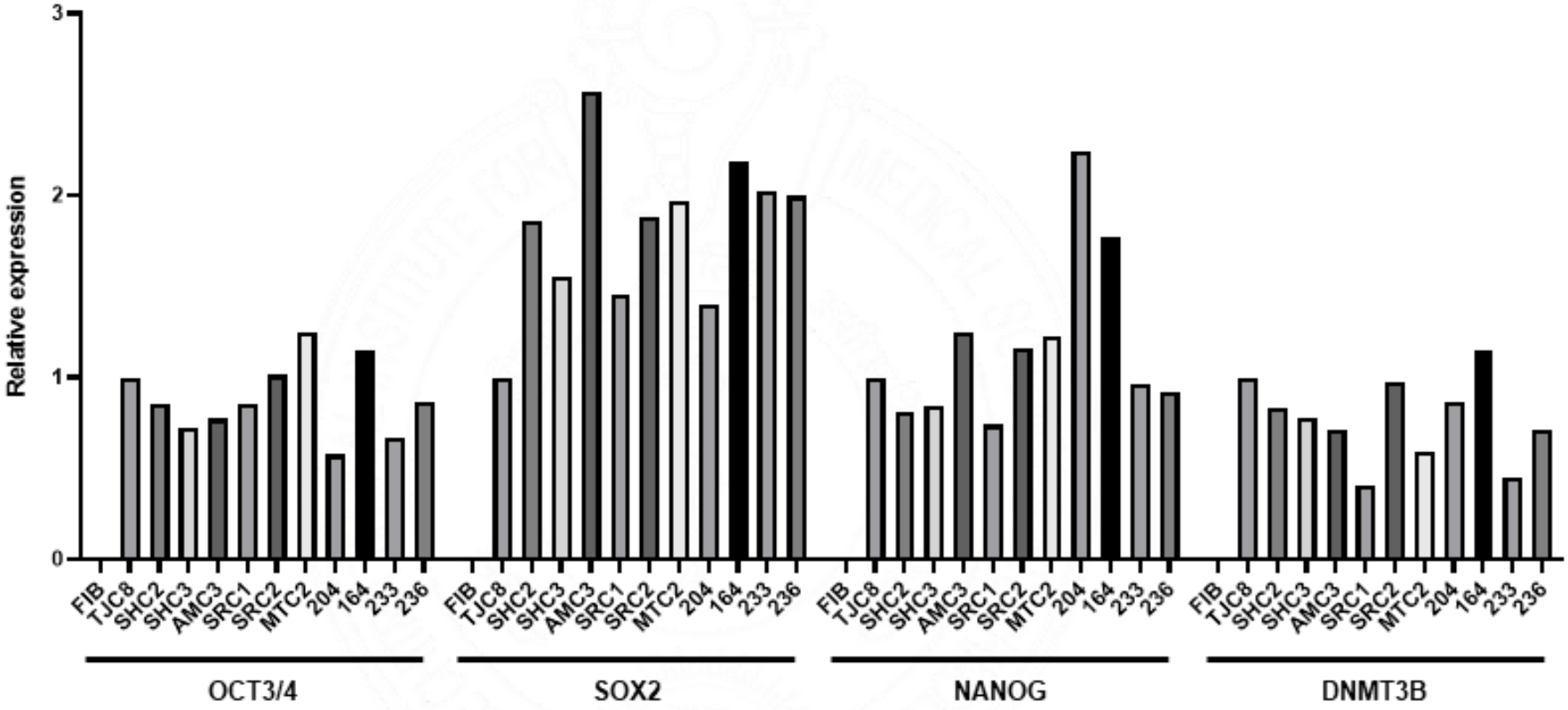
All the iPSC lines were characterized for their expression of pluripotency markers, OCT4, NANOG, SSEA4, and TRA-I-81 in the undifferentiated state **(Figure 4.3.4A)**. The iPSCs also showed pluripotency marker expression of TRA-I-60 by flow analysis **(Figure 4.3.3C)**. The qPCR analysis conducted to evaluate the expression of pluripotency genes OCT3/4, SOX2, NANOG, and DNMT3B demonstrated robust gene expression in the FA-iPSCs when compared to the fibroblast cells from which they were derived **(Figure 4.3.4B).** They also could be differentiated into the cells of three germ layers of ectoderm, endoderm, and mesoderm, and analysed for the expression of PAX6, SOX17, and BRACHYURY, respectively **(Figure 4.3.4C)**. Karyotyping analysis done for the clones did not show any abnormalities except for *FANCL* clones **(Figure 4.3.4D)**.

A

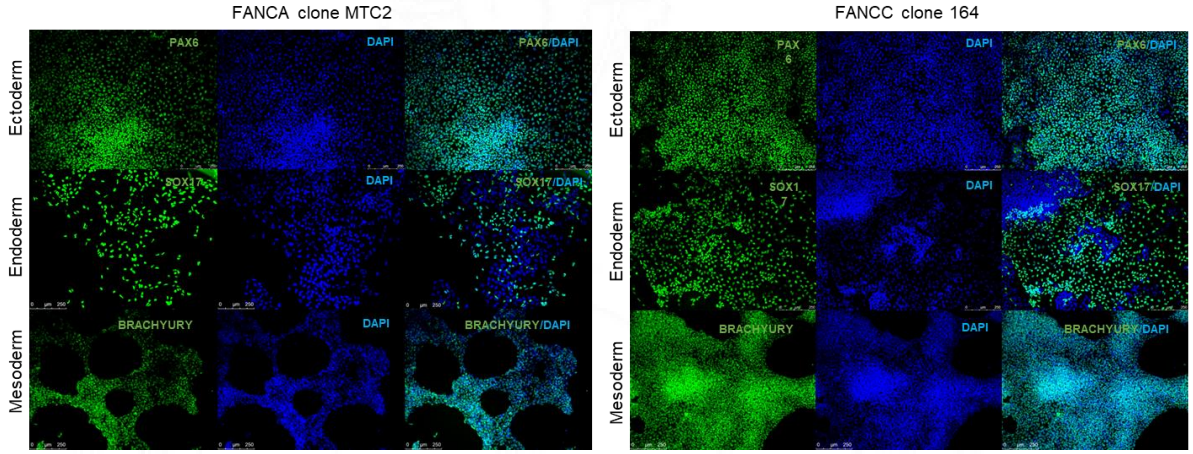


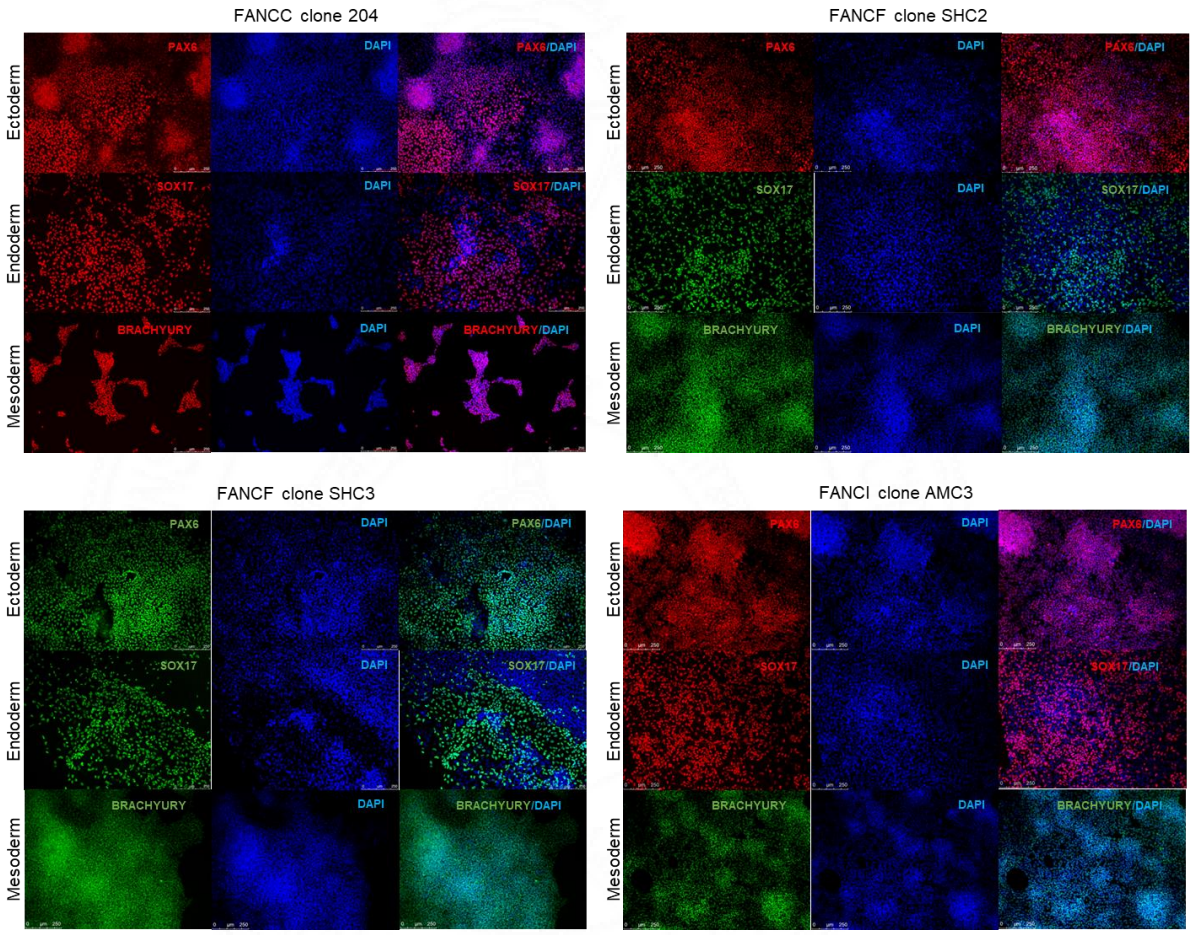
SCTIMST, TRIVANDRUM

B

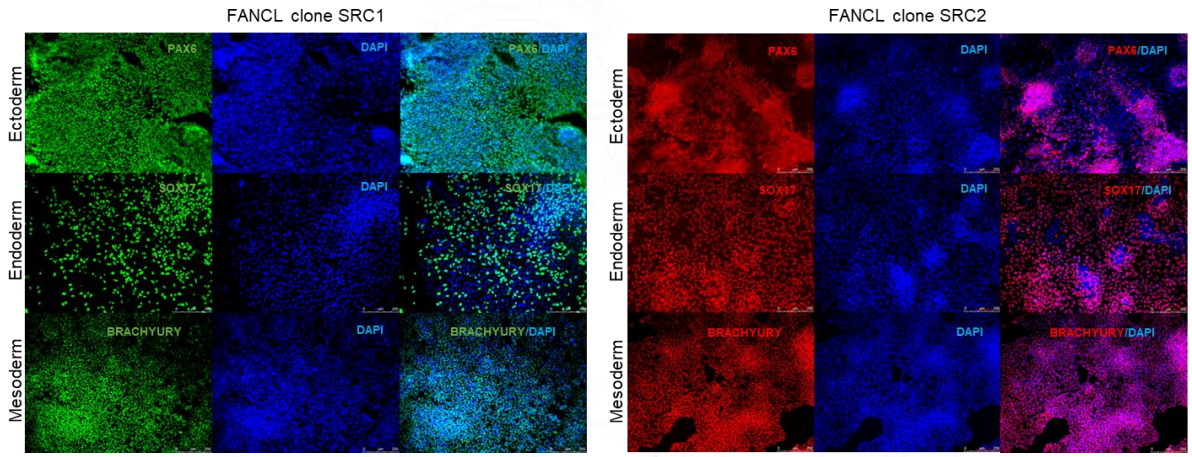

Relative expression
3
2
1
0
FIB
TJC8
SHC2
SHC3
AMC3
SRC1
SRC2
MTC2
204
164
233
236
OCT3/4
SOX2
NANOG
DNMT3B

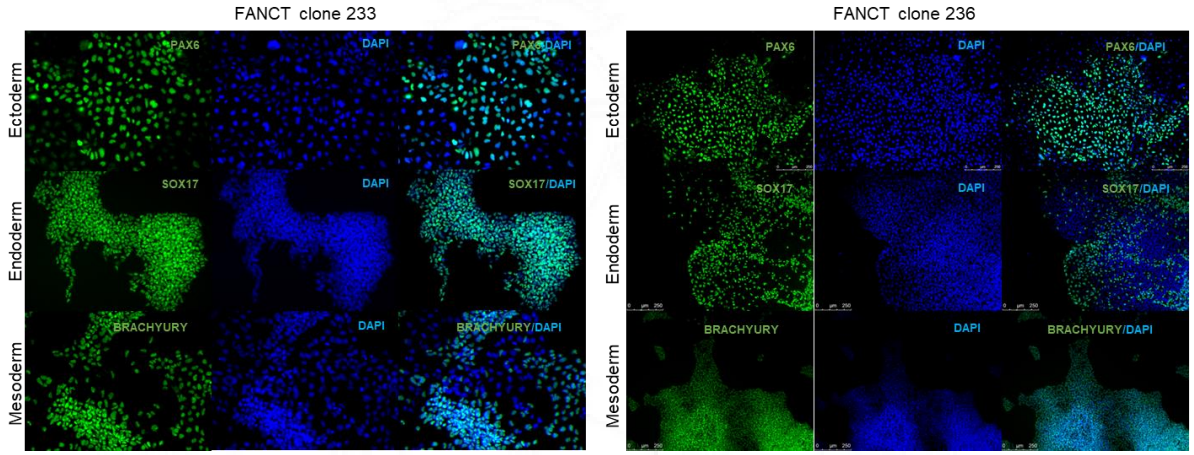
C


FANCA clone MTC2
FANCC clone 164
Ectoderm
Endoderm
Mesoderm
PAX6
DAPI
PAX6/DAPI
SOX17
SOX17/DAPI
BRACHYURY
BRACHYURY/DAPI
PAX 6
SOX1 7



SCTIMST, TRIVANDRUM


FANCL clone SRC1
FANCL clone SRC2
Ectoderm
Endoderm
Mesoderm
PAX6
DAPI
PAX6/DAPI
SOX17
SOX17/DAPI
BRACHYURY
BRACHYURY/DAPI


FANCT clone 233
FANCT clone 236
Ectoderm
Endoderm
Mesoderm
PAX6
DAPI
PAX6/DAPI
SOX17
DAPI
SOX17/DAPI
BRACHYURY
DAPI
BRACHYURY/DAPI

D

Figure 4.3.4. Characterisation of FA-iPSCs **(A)** Immunofluorescence results of pluripotency marker analysis in *FANCA* and *FANCI* clones (1 clone each), as well as *FANCC*, *FANCF*, *FANCL*, and *FANCT* clones (2 clones each) showing high level expression of all the pluripotency markers. **(B)** qPCR analysis for pluripotency markers in FA-iPSCs and fibroblasts compared to TJC8 control iPSC line[$]. **(C)** Trilineage differentiation of FA iPSC clones shows differentiation into three germ layers viz. ectoderm, endoderm and mesoderm. **(D)** Karyotyping analysis for FA-iPSC clones. Except *FANCL* clones, all the clones showed normal karyotype.

[$]in-house generated iPSC control line

4.3.5. iPSC clones derived from *FANCF* and *FANCI* after complementation do not show progressive cell death after doxycycline withdrawal.

Initially, we conducted a doxycycline withdrawal experiment on 1 clone each of *FANCF* and *FANCI* iPSCs. As these clones did not exhibit any signs of cell death upon doxycycline withdrawal **(Figure 4.3.3D)**, we analysed more *FANCF* and *FANCI* clones to investigate whether this lack of cell death was consistent across all clones belonging to these two complementation groups. Despite all the *FANCF* iPSC clones (n=11) displaying positive mCherry expression when cultured in the presence of a high concentration of doxycycline (1000 ng/ml), none of them demonstrated any signs of cell death upon doxycycline withdrawal. However, an intriguing observation was made with the *FANCI* iPSC clones. Out of the 5 clones tested, only 1 clone displayed mCherry expression when cultured in the presence of doxycycline. Sanger sequencing confirmed the presence of mutations in these *FANCF* and *FANCI* iPSC clones **(Figure 4.3.5A, Figure 4.3.5C).** Upon analysing *FANCF* iPSCs, we observed that low level of FANCD2 monoubiquitination after doxycycline withdrawal **(Figure 4.3.3E)**. We measured the expression of *FANCF* in the *FANCF*-complemented iPSC clones, and it was found that all the clones exhibited low-level expression of *FANCF*. The levels of FANCF expression in the cells cultured in the absence of doxycycline were comparable to those in the control cells, which explains why these cells did not undergo cell death upon doxycycline withdrawal **(Figure 4.3.5B)**. Among the 5 iPSC clones derived from FANCI-/- fibroblasts, only 1 clone exhibited mCherry expression

following doxycycline induction **(Figure 4.3.3B)**. To confirm the integration of the transgene in these clones, we subjected these lines to blasticidin treatment, as the complementation vector contained a co-expressed blasticidin resistance gene. The blasticidin selection resulted in cell death in all the clones except AMC3, which showed mCherry expression after doxycycline induction. These results strongly suggest that the four clones lacking mCherry expression did not possess integration of the FANCI complementation vector. These findings indicate that FANCI-/- fibroblasts can be reprogrammed without the need for complementation, supporting previous reports (Osborn et al., 2016). These results also suggest that FANCI is not essential for the reprogramming process. Interestingly, all the five *FANCI* iPSC clones, without FANCI complementation, exhibited the absence of FANCD2 ubiquitination after doxycycline withdrawal **(Figure 4.3.5D)**, highlighting the importance of non-canonical functions of FA pathway proteins in the reprogramming and maintenance of iPSCs. AMC3, FANCI iPSC clone was cultured with and without doxycycline for 8 days, to eliminate any residual FANCI protein in the cells cultured without doxycycline. The cells were then treated with 60ng/ml and 100ng/ml MMC for 16 hours for alkaline phosphatase (AP) staining and cell counting. AP staining after MMC treatment showed that fewer iPSC colonies were present in the plates with the cells cultured without doxycycline compared to those cultured with doxycycline **(Figure 4.3.5E)**. Cell counting also showed a notable difference, as the cell count was higher in doxycycline-cultured cells compared to the doxycycline-free cells **(Figure 4.3.5F)**.

This data also indicates that the absence or presence of FANCD2 ubiquitination may not be linked to the death or survival of iPSCs.

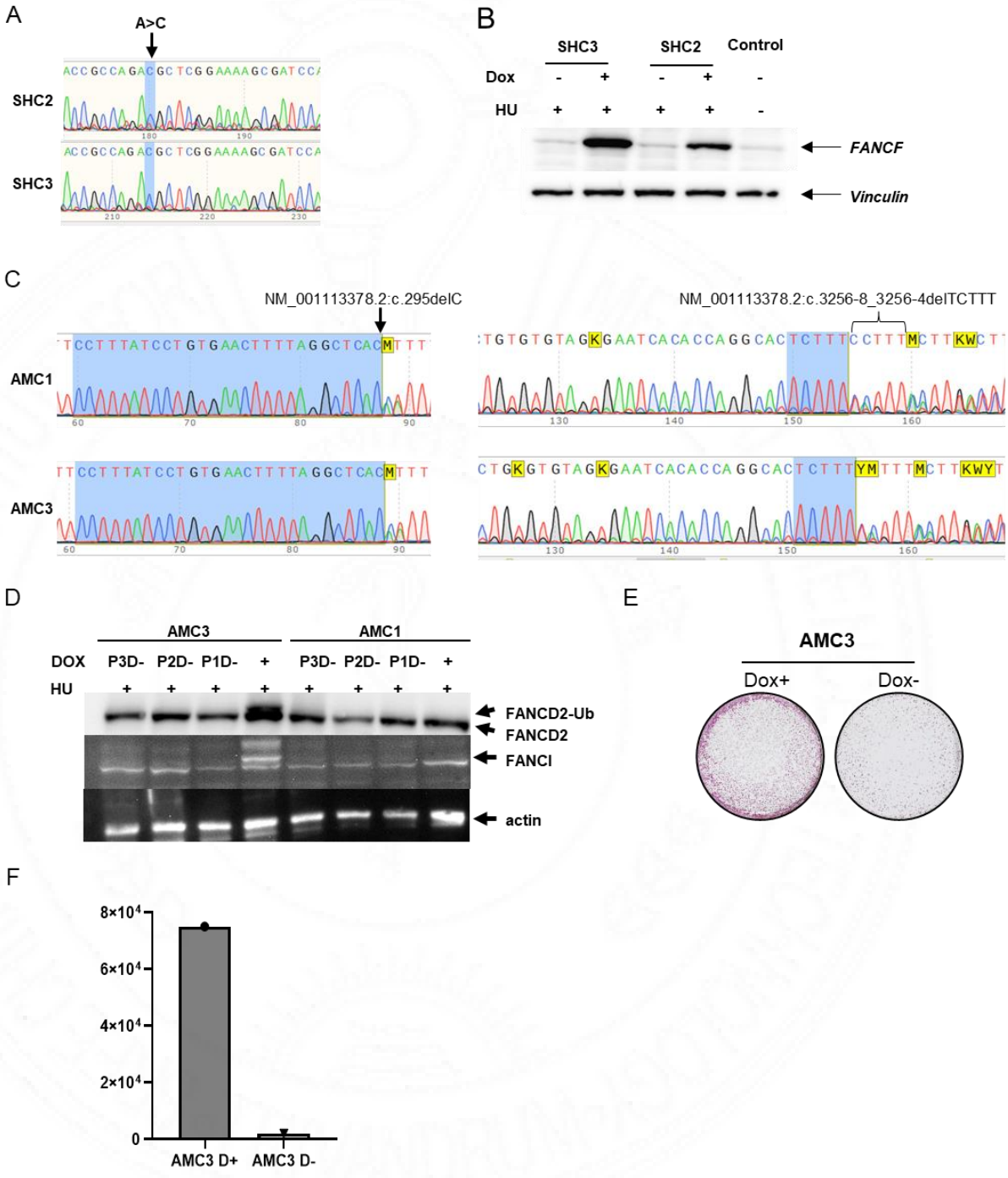
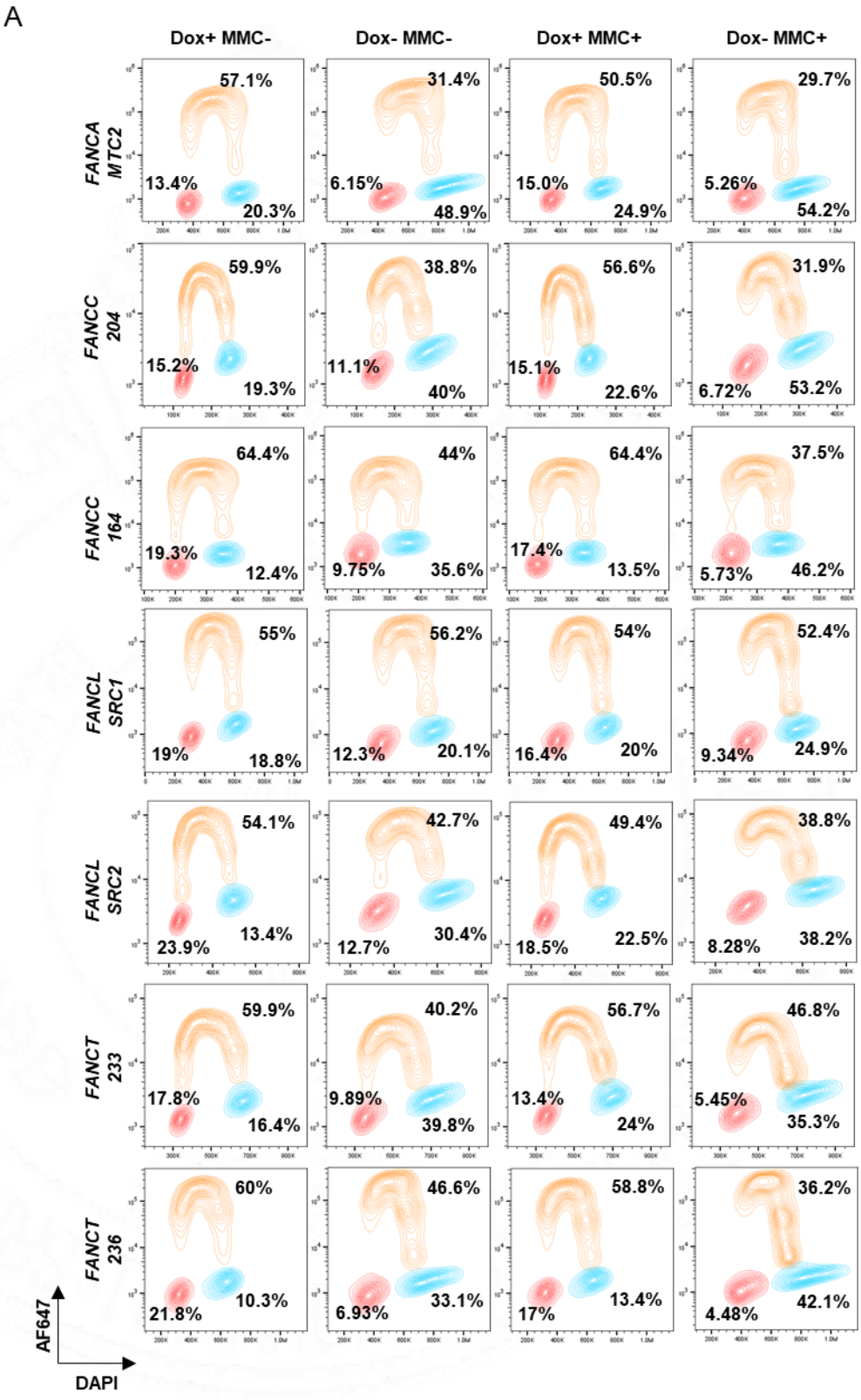


Figure 4.3.5. Analysis of the *FANCF* and *FANCI* iPSC clones that do not show cell death after doxycycline withdrawal. **(A)** Sanger sequencing for *FANCF* mutation in iPSC clones. **(B)** FANCF expression in *FANCF* iPSC clones cultured in the absence of doxycycline, showing that the levels are similar to that of the control iPSC line. **(C)** Sanger sequencing for *FANCI* mutations in iPSC clones confirming the identity of the clones. **(D)** The loss of FANCD2-Ub after doxycycline withdrawal in AMC3 clone but no FANCD2-Ub detected in AMC1 clone even after treatment with a high concentration of doxycycline (1000ng/ml). **(E)** Alkaline phosphatase (AP) staining revealing greater cell death in the doxycycline-free AMC3 iPSC cultures treated with 100ng/ml MMC compared to those cultured with doxycycline. **(F)** Significant decrease in cell count of AMC3 observed after 48 hours of 60ng/ml MMC treatment in doxycycline-free cultures, compared to the iPSC cultures supplemented with doxycycline

4.3.6. Cell cycle analysis

The FA-iPSCs generated by inducible complementation cells exhibit G2/M arrest in the absence of doxycycline (Bharathan et al., 2017). Cell cycle analysis using 5-ethynyl-2'-deoxyuridine (EdU) and 4',6-diamidino-2-phenylindole (DAPI) for 1 clone of *FANCA*, 2 clones of *FANCC*, 1 clone of *FANCL*, and 2 clones of *FANCT* cultured in the absence of doxycycline revealed cell-cycle arrest in the G2/M phase compared with iPSCs cultured in presence of doxycycline **(Figure 4.3.6A)**. In the absence of MMC, G2/M arrest observed in *FANCA* (48.9%) defective cells (doxycycline-) was higher compared to FA-iPSCs with defects in *FANCC* (40%), *FANCL* (20.1-30.4%) and *FANCT* (33.1-39.8%) other complementation groups. After treatment of cells with 1ng/ml MMC, G2/M arrest increased further in defective cells; *FANCA* (54.2%), *FANCC* (53.2%), *FANCL* (24.9-38.2%), and *FANCT* (35.3-42.1%) compared to FA-pathway proficient cells. In the absence of doxycycline, FA-iPSCs exhibited a significantly higher G2/M arrest even without the presence of mitomycin C (MMC), indicating the existence of intracellular DNA-damaging agents generated in the absence of the FA pathway. Furthermore, upon the addition of MMC, the G2/M arrest further increased in cells cultured without doxycycline for *FANCA*, *FANCC*, *FANCL*, and *FANCT*, in comparison to cells cultured with doxycycline. Despite exposure to MMC, *FANCF* clones did not display significantly elevated levels of G2M arrest after doxycycline withdrawal **(Figure 4.3.6B).** This could be attributed to the basal level of *FANCF* expression observed in these clones **(Figure 4.3.5B**). Out of the *FANCI* clones, the single clone that displayed mCherry expression upon doxycycline

induction showed a survival advantage after the withdrawal of doxycycline. Additionally, this clone did not exhibit an increased G2/M arrest after doxycycline withdrawal, even in the presence of MMC treatment (**Figure 4.3.6B)**. However, it is noteworthy that cell death due to G2/M arrest and the lack of FANCD2 ubiquitination are not related and occur through different mechanisms in the cells without an active FA pathway.


A
Dox+ MMC-
Dox- MMC-
Dox+ MMC+
Dox- MMC+
FANCA MTC2
57.1%
13.4%
20.3%
31.4%
6.15%
48.9%
50.5%
15.0%
24.9%
29.7%
5.26%
54.2%
FANCC 204
59.9%
15.2%
19.3%
38.8%
11.1%
40%
56.6%
15.1%
22.6%
31.9%
6.72%
53.2%
FANCC 164
64.4%
19.3%
12.4%
44%
9.75%
35.6%
64.4%
17.4%
13.5%
37.5%
5.73%
46.2%
FANCL SRC1
55%
19%
18.8%
56.2%
12.3%
20.1%
54%
16.4%
20%
52.4%
9.34%
24.9%
FANCL SRC2
54.1%
23.9%
13.4%
42.7%
12.7%
30.4%
49.4%
18.5%
22.5%
38.8%
8.28%
38.2%
FANCT 233
59.9%
17.8%
16.4%
40.2%
9.89%
39.8%
56.7%
13.4%
24%
46.8%
5.45%
35.3%
FANCT 236
60%
21.8%
10.3%
46.6%
6.93%
33.1%
58.8%
17%
13.4%
36.2%
4.48%
42.1%
AF647
DAPI

B

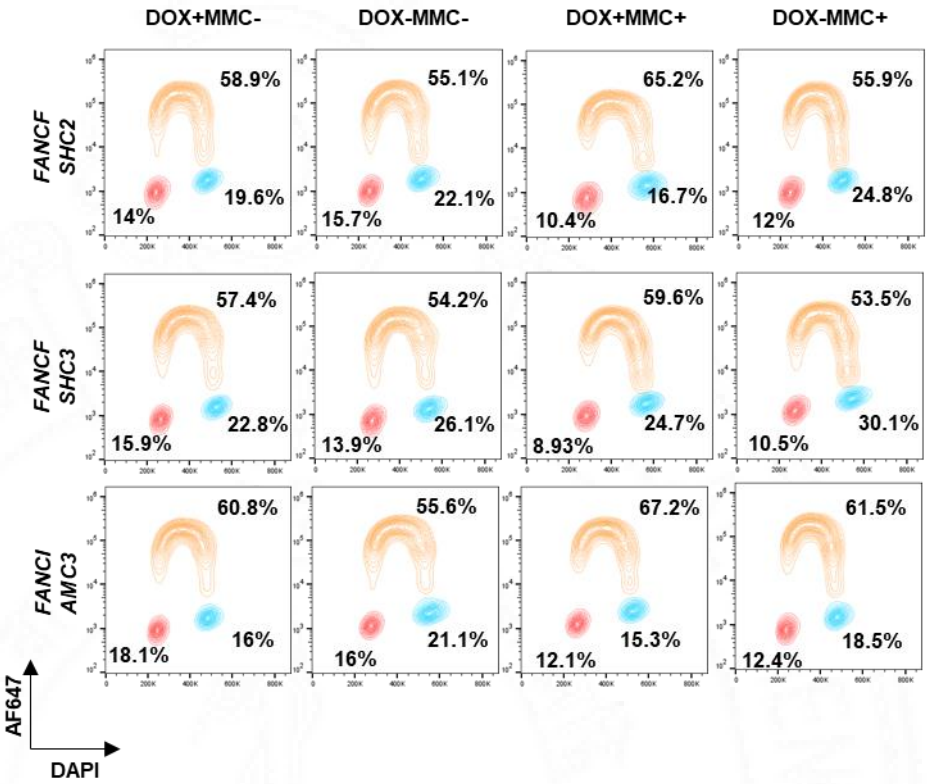


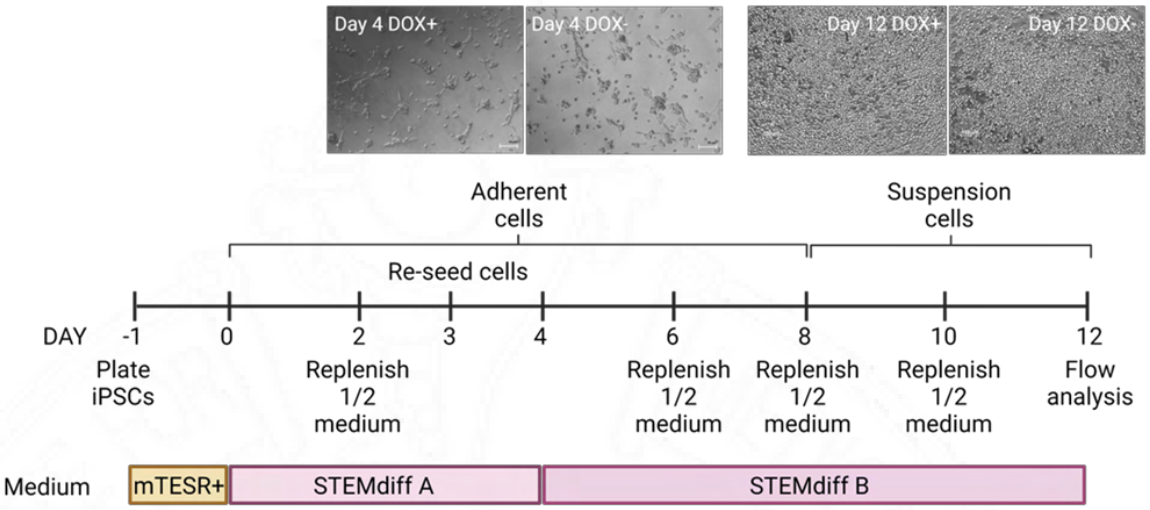
Figure 4.3.6. Cell cycle analysis of FA-iPSCs. **(A)** Cell cycle analysis for *FANCA*, *FANCC*, *FANCL* and *FANCT* clones in the presence of MMC showing increased G2/M arrest when cultured without doxycycline. **(B)** Cell cycle analysis for *FANCF* and *FANCI* clones that do not show increased G2/M arrest in the absence of doxycycline from culture even in the presence of MMC.

4.3.7. Haematopoietic stem and progenitor cell (HSPC) differentiation of FA iPSCs.

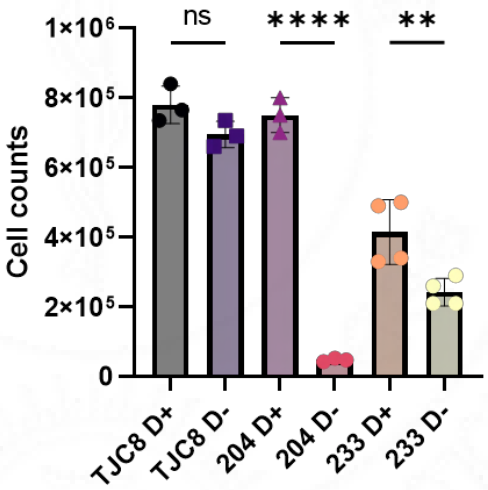
To determine whether our disease model system recapitulates the hallmarks of HSPC dysfunction in FA, we examined the haematopoietic capacity of *FANCT* and *FANCC*-deficient HSPCs. Patient-derived iPSCs generated for these 2 complementation groups (*FANCC* and *FANCT*) were cultured either with or without doxycycline to either maintain or deactivate the inducible FA transgenes, respectively, during the assays of haematopoietic function using a modified protocol **(Figure 4.3.7A)**. In the selected iPSC clones, the FA transgene expression in iPSCs lasted only 5 days after doxycycline withdrawal, resulting in consequent progressive cell death. During the iPSC-HSPC differentiation, in the initial culture conditions that favoured mesodermal specification (medium A, day 0 to 3) **(Figure 4.3.7A)**, an adherent monolayer was rapidly formed in the control, doxycycline + and doxycycline – iPSCs. Subsequently, the cells underwent hematoendothelial differentiation before the eventual release of hematopoietic progenitors as suspension cells (after day 10) **(Figure 4.3.7A**). The cells cultured in the absence of doxycycline conditions showed progressive cell death, whereas the cells cultured in the presence of doxycycline conditions showed good proliferation. In the first phase of differentiation (days 5 to 9), most cells in the culture were non-hematopoietic. From day 10 to 12, as expected, haematopoietic cells were formed with a concomitant decrease in the proportion of non-hematopoietic cells likely resulting from an endothelial-to-haematopoietic transition (EHT) process in culture. The hematopoietic cells were identified by the expression of CD34, CD45, CD43, and CD235a surface markers **(Figure 4.3.7C)**. The number of HSPCs generated from the differentiation of *FANCC* and *FANCT* clones in the presence of doxycycline conditions was significantly higher compared to clones differentiated in the

absence of doxycycline **(Figure 4.3.7B).** Flow analysis of the co-expression of CD34, CD43, CD45, and CD235a did not show any difference between the cells differentiated in the presence of doxycycline and the absence of doxycycline conditions **(Figure 4.3.7C)**. Next, to determine whether our system recapitulates the hallmarks of HSPC dysfunction in FA, we examined the colony-forming potential of FA-deficient HSPCs. We cultured iPSC-derived HSPCs, formed with and without doxycycline, to identify their haematopoietic clonogenic potential. The cells were taken for colony formation assay (CFU) on day 12 of differentiation. In the CFU assay, the *FANCC* and *FANCT* HSPCs generated in the absence of doxycycline showed significantly reduced clonogenicity compared to the FA-proficient HSPCs generated in the presence of doxycycline **(Figure 4.3.7D)**. These results indicate that FA deficiency in HSPCs derived from iPSCs impairs hematopoietic colony formation, recapitulating a known phenotype of HPCs from the bone marrow of human FA patients.

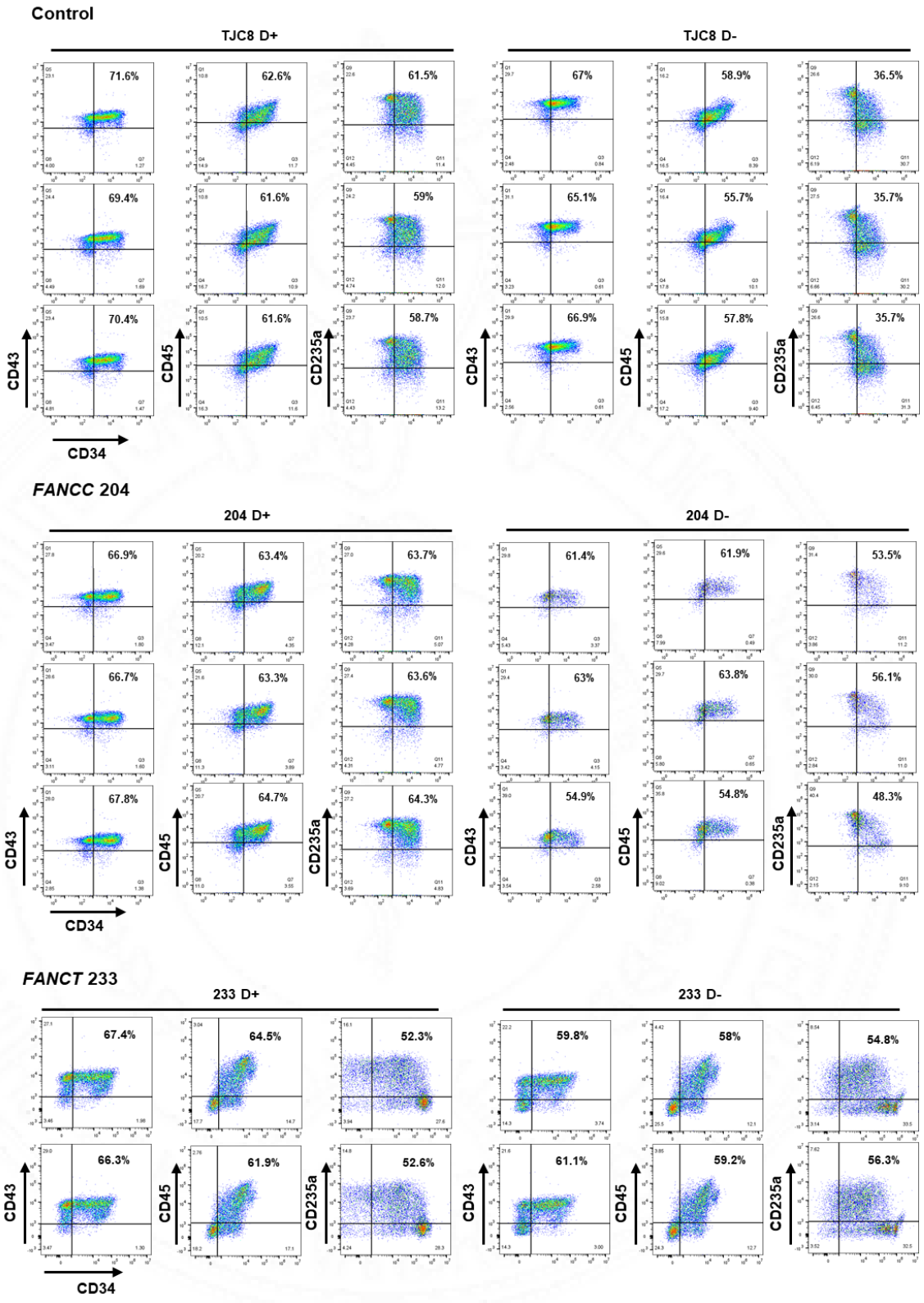
A


Day 4 DOX+
Day 4 DOX-
Day 12 DOX+
Day 12 DOX-
Adherent cells
Suspension cells
Re-seed cells
DAY
-1
0
2
3
4
6
8
10
12
Plate iPSCs
Replenish 1/2 medium
Replenish 1/2 medium
Replenish 1/2 medium
Replenish 1/2 medium
Flow analysis
Medium
mTESR+
STEMdiff A
STEMdiff B

B


ns
****
**
1×10⁶
8×10⁵
6×10⁵
4×10⁵
2×10⁵
0
Cell counts
TJC8 D+
TJC8 D-
204 D+
204 D-
233 D+
233 D-

C


Control
TJC8 D+
TJC8 D-
71.6%
62.6%
61.5%
67%
58.9%
36.5%
69.4%
61.6%
59%
65.1%
55.7%
35.7%
70.4%
61.6%
58.7%
66.9%
57.8%
35.7%
CD43
CD45
CD235a
CD34
FANCC 204
204 D+
204 D-
66.9%
63.4%
63.7%
61.4%
61.9%
53.5%
66.7%
63.3%
63.6%
63%
63.8%
56.1%
67.8%
64.7%
64.3%
54.9%
54.8%
48.3%
CD43
CD45
CD235a
CD34
FANCT 233
233 D+
233 D-
67.4%
64.5%
52.3%
59.8%
58%
54.8%
66.3%
61.9%
52.6%
61.1%
59.2%
56.3%
CD43
CD45
CD235a
CD34

D

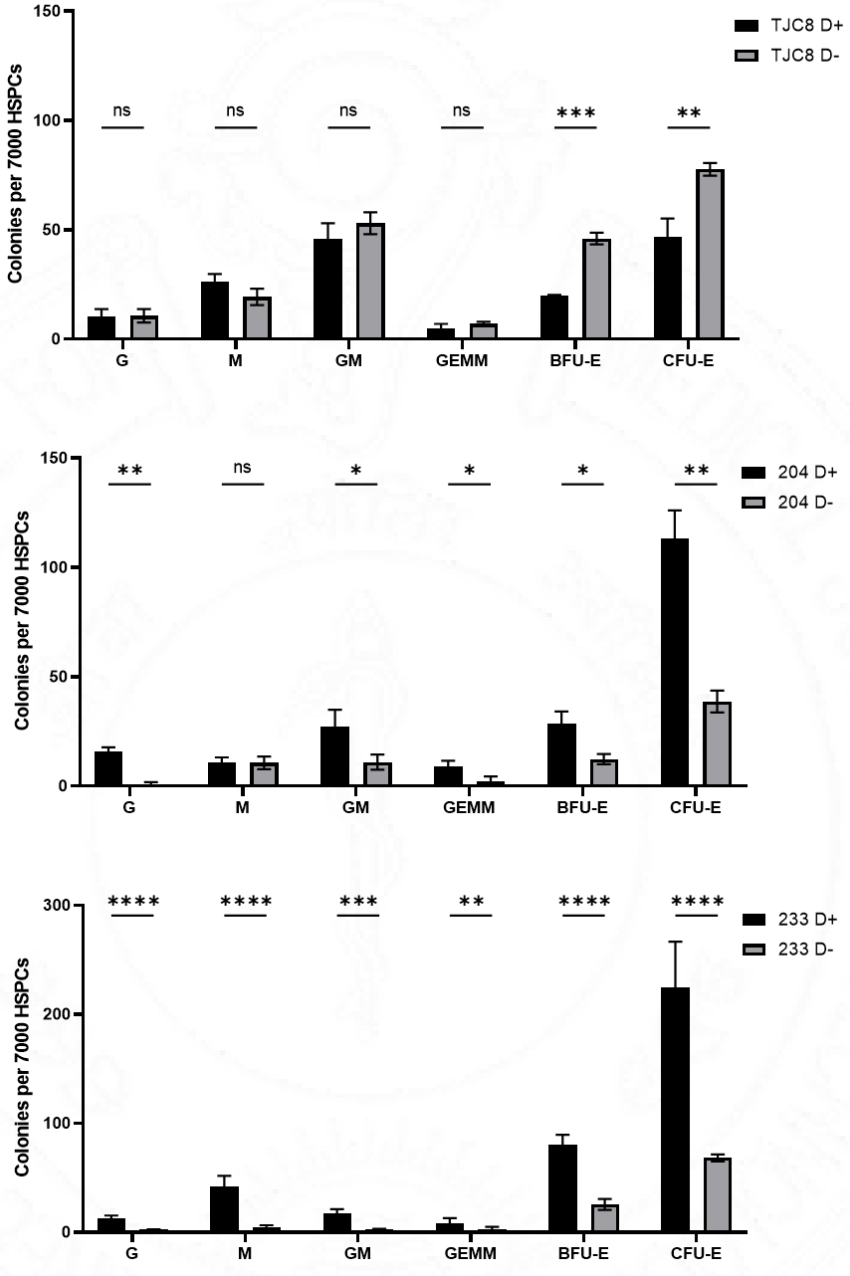


Figure 4.3.7. Differentiation of FA-iPSCs to HSPCs to understand the defects in different haematopoietic lineages. **(A)** Schematic representation of differentiation of FA-iPSCs to HSPCs **(B)** Graph illustrating decreased number of HSPCs formed during differentiation in the absence of doxycycline. **(C)** Flow cytometric analysis of HSPCs

generated from control and FA-iPSCs in the presence and the absence of doxycycline. Suspension cells were analysed for CD34, CD43, CD45 and CD235a. **(D)** The methylcellulose assay to assess colony forming capacity of HSPCs.

5. Discussion

To evaluate the effectiveness of CBA (chromosome breakage analysis) and FANCD2-Ub (FANCD2 monoubiquitination) as diagnostic tests for Fanconi Anaemia (FA), we conducted an extensive study involving 153 Indian FA patients. We aimed to assess the applicability of these tests in the molecular diagnosis of FA. We also employed next-generation sequencing (NGS) techniques along with state-of-the-art bioinformatics tools to genotype the patients. This approach allowed us to identify disease-causing genotypes in an impressive 95.4% of the patients, providing valuable insights into the genetic basis of FA in the Indian population.

During our analysis, we identified the presence of a *FANCL* founder mutation, along with a high frequency of *FANCA* and *FANCG* mutations among the patients. These findings are significant as they contribute to the development of a novel, faster, and cost-effective molecular diagnosis strategy specifically tailored for Indian FA patients. This strategy takes advantage of the identified *FANCL* founder mutation and the increased occurrence of *FANCA* and *FANCG* mutations, streamlining the diagnostic process.

A precise laboratory diagnosis is imperative for effective clinical management of Fanconi Anaemia (FA). While Chromosome Breakage Analysis (CBA) is widely regarded as the "gold standard" test for FA, it is not without its limitations. Challenges

such as laborious standardization and user variability in the interpretation of scores have been associated with CBA. In this study, we conducted a comprehensive comparison between CBA and FANCD2-Ub analysis in a large cohort of FA patients. Our findings reaffirmed that FANCD2-Ub analysis, which is currently not utilized for diagnosis, is indeed suitable for diagnosing FA.

Recognizing the importance of addressing the limitations of CBA, the inclusion of FANCD2-Ub analysis in the diagnostic process can offer additional benefits. By leveraging FANCD2-Ub analysis alongside CBA, we can enhance the accuracy and reliability of FA diagnosis. This expanded approach mitigates the concerns related to standardization and user variability, contributing to a more robust and comprehensive diagnostic framework. Therefore, based on the outcomes of our study, we advocate for the inclusion of FANCD2-Ub analysis as a valuable diagnostic tool for FA. In our study, we made an interesting observation regarding the sensitivity of CBA scores in diagnosing Fanconi Anaemia (FA) when a specific cut-off value was employed. Using receiver operating characteristic (ROC) curve analysis and Youden's index, we determined that a cut-off value of 15 yielded increased sensitivity in FA diagnosis. However, it is important to note that further confirmation of this cut-off value through a randomized comparative analysis is necessary.

Within our cohort, we identified 17 patients with low CBA scores ranging from 0 to 8.2, with a median score of 0. Out of these 17 patients, 12 were selected for FANCD2-Ub analysis in peripheral blood. Strikingly, 7 out of these 12 patients demonstrated defective FANCD2-Ub (58.3%). Subsequent analysis involving

pathogenic variant analysis unequivocally confirmed these 7 patients as FA cases. This finding further strengthens the potential significance of incorporating FANCD2-Ub analysis into the diagnostic process for FA. By detecting defective FANCD2-Ub in a subset of patients with low CBA scores, we were able to accurately identify additional cases of FA that would have otherwise been missed. Additionally, our findings underscore the relevance of FANCD2-Ub analysis in cases where low CBA scores are observed, as it can aid in identifying additional FA cases that might have been overlooked.

Defects in the downstream genes of the Fanconi Anaemia (FA) pathway, which do not impact FANCD2 ubiquitination, are considered to be very rare, accounting for approximately 2% to 6% of FA patients (Nie et al., 2020; Toksoy et al., 2020). In our study, we also observed downstream pathogenic variants in only approximately 2.2% of the Indian patients, further supporting their rarity. Given the low frequency of downstream gene defects, FANCD2-Ub analysis emerges as a reliable diagnostic test for FA. Spontaneous reversal of pathogenic variants occurs in the haematopoietic cells of a subset of FA patients, ranging from 15%–25% (Soulier et al., 2005). To overcome this challenge and ensure accurate diagnosis, primary skin fibroblasts obtained from FA patients are commonly utilized. By analysing these fibroblasts, which are not subject to spontaneous reversion, the diagnosis of FA can be more accurately determined.

In this study, we performed FANCD2-Ub analysis on both T-cells and fibroblasts from a cohort of 55 patients. Surprisingly, we identified only 3 cases (5.5%)

that exhibited mosaicism, where FANCD2-Ub was present in T-cells but absent in fibroblasts. This low incidence of mosaicism in our cohort, which was less than 15%, could be attributed to the fact that the patients were referred from a haematology clinic due to presenting symptoms such as pancytopenia, haematological abnormalities, and physical anomalies. It has been previously reported that mutations in the downstream genes of the FA pathway are found in approximately 2% to 6% of FA patients (Nie et al., 2020; Steinberg-Shemer et al., 2020; Toksoy et al., 2020). We found the mutations in the downstream FA pathway genes in only ~1.7% of the patients. Given the rarity of downstream gene mutations, it becomes evident that FANCD2-Ub analysis offers a diagnostic advantage in the Indian population compared to other populations.

Despite the utilization of various pathogenicity prediction programs, the assessment of disease relevance for genetic variants, particularly those resulting in missense mutations, remains a complex challenge. To date, 6.5 million missense variants have been identified (gnomAD (Karczewski et al., 2020)), and the pathogenicity of >98% of these variants is unknown (Landrum and Kattman, 2018) New experimental techniques enable the assessment of thousands of mutations simultaneously (Esposito et al., 2019; Trenkmann, 2018). Through these high-throughput experiments, researchers can gather data on the effects of these mutations and subsequently analyse and scrutinize the results to assign a clinical interpretation to human variants (Rehm et al., 2015). EVE, a recently reported computational method for the classification of human genetic variants based on evolutionary sequences. Notably, EVE has demonstrated superior performance compared to existing state-of-

the-art computational methods in predicting variant pathogenicity. Intriguingly, it has been found that the accuracy of predictions from EVE is on par with those derived from previously described high-throughput experiments. By leveraging evolutionary sequences, this computational method offers a unique perspective on the functional implications of genetic variants. Its ability to outperform current computational methods indicates its effectiveness in accurately assessing the pathogenicity of variants (Frazer, Notin, Dias, Gomez, Joseph K Min, et al., 2021). The integration of computational methods like EVE into variant classification pipelines has the potential to significantly enhance our ability to interpret genetic variants, particularly missense mutations.

Accurate detection of defective genes and pathogenic variants holds great importance in the fields of genetic counselling and the development of targeted prenatal genetic testing. The identification of these genetic abnormalities enables comprehensive counselling for individuals and families affected by Fanconi Anaemia (FA). In addition to genetic counselling, early molecular diagnosis plays a crucial role in several aspects, including participation in gene therapy for FA. Timely detection of FA allows patients to access appropriate treatments and interventions, contributing to improved patient outcomes (Río et al., 2019). While targeted gene panels have been established for FA (Ameziane et al., 2012; De Rocco et al., 2014; Ghemlas et al., 2015), the use of whole exome sequencing (WES) provides a valuable opportunity for the identification of novel genes associated with the disease. Although there have been limited reports on WES analysis of FA patients, typically involving small cohorts

ranging from 15 to 25 patients (Nie et al., 2020), these studies have demonstrated the potential of WES in expanding our understanding of the genetic landscape of FA. In a recent comprehensive whole exome sequencing (WES) study involving 68 European Fanconi Anaemia (FA) patients, pathogenic variants were successfully identified in 93.3% of the patients (Bogliolo et al., 2020). We performed a WES analysis on a larger cohort of FA patients and identified pathogenic variants with 95.0% genotyping efficiency. However, in the case of seven patients, only heterozygous variants were detected through WES analysis. In such instances, additional investigations such as gene expression and protein analysis could help identify potentially missed pathogenic variants. By examining the expression levels of genes associated with FA and conducting protein analyses, it is possible to gain further insights into the underlying mechanisms and potentially uncover pathogenic variants that were not initially detected by WES alone. The inclusion of larger cohorts in future WES studies holds great potential for identifying rare variants, detecting new genes, and gaining deeper insights into the genetic heterogeneity and complexity of FA.

While next-generation sequencing (NGS) has revolutionized genetic analysis, it has limitations when it comes to detecting copy number variants (CNVs), particularly deletions. To overcome this challenge, the development of robust bioinformatics methods is essential for accurate detection of deletions. A recent study by (Bogliolo et al., 2020) successfully applied a bioinformatics tool using custom scripts to efficiently identify deletions in Fanconi Anaemia (FA) genes. In our study, we employed ExomeDepth (Plagnol et al., 2012) for CNV analysis, which allowed us

to detect deletions in our patients. To ensure high accuracy and reliability, we implemented stringent filters to discard false positives. By applying these filters, we achieved 100% accuracy in detecting deletions in our cohort of FA patients. These findings demonstrate the efficacy of improved bioinformatics approaches in efficiently detecting CNVs, specifically deletions, within the FA genes. Consistent with previous reports (Amouri et al., 2014; Ben Haj Ali et al., 2021; Callén et al., 2004), our findings indicate that deletions affecting the FANCA gene are prevalent among Indian Fanconi Anaemia (FA) patients, accounting for approximately 14.4% of cases. By combining the analysis of both single nucleotide variants (SNVs) and copy number variants (CNVs), we were able to identify disease-associated genotypes in approximately 95% of the patients. These results are in line with a study conducted by (Bogliolo et al., 2020) which also employed a comprehensive analysis of both SNVs and CNVs and reported a high detection rate of pathogenic variants in FA patients.

FANCA has high genetic heterogeneity and is the most commonly mutated FA gene, with frequencies ranging from 60% to 80% in different populations (Bogliolo et al., 2020; Esmail Nia et al., 2016; Nie et al., 2020) However, we found among our cohort of FA patients, the most frequently mutated genes were *FANCA* (58.2%), *FANCL* (19.2%), and *FANCG* (13%). Remarkably, a substantial proportion (~20%) of our patients exhibited homozygosity for the *FANCL* c.1092G>A;p.K364= variant. However, it is worth noting that despite carrying this homozygous variant, these patients presented with diverse phenotypes, highlighting the influence of other factors beyond genetic mutations in shaping the disease manifestation. This observation

suggests the existence of additional genetic modifiers or environmental influences that contribute to the phenotypic heterogeneity observed in FA. Our cohort displayed a notable degree of consanguinity, with over 87% of the patients exhibiting homozygous pathogenic variants in FA genes. A large number of patients, 65 (45%), were born from consanguineous marriages. Due to the high rate of homozygosity, our NGS analysis identified only 101 distinct variants in the 153 cases analysed. Consequently, the number of recurrent pathogenic variants was relatively small, with 20 recurrent variants detected in more than one patient. The frequencies of these recurrent variants ranged from 1.3% to 17.5%, highlighting their varying prevalence within the cohort.

In our study, we identified pathogenic variants in a total of 10 Fanconi Anaemia (FA) genes, namely *FANCA, FANCG, FANCC, FANCD2, FANCL, FANCF, FANCT, FANCI, FANCD1,* and *FANCJ*. These genes emerged as potential candidates for prioritizing the design of the FA genotyping panel and implementing bioinformatics analysis specific to the Indian patient population. Our findings revealed that employing Sanger sequencing to detect the *FANCL* pathogenic variant and multiplex ligation-dependent probe amplification (MLPA) to identify *FANCA* deletions enabled the diagnosis of approximately 33.6% of FA cases. To further enhance diagnostic efficiency, we developed a faster and cost-effective long-range amplicon-based next-generation sequencing (LA-NGS) strategy, which successfully detected point pathogenic variants in the *FANCA* and *FANCG* genes. Notably, these two genes accounted for approximately 56.8% of the genotypes observed in Indian FA patients.

Furthermore, our study identified the presence of a founder variant in the *FANCL* gene and a high frequency of pathogenic variants in the FANCA and FANCG genes among Indian FA patients. These findings were instrumental in establishing a novel and efficient molecular diagnosis strategy, which has the potential to diagnose approximately 90% of FA patients in the Indian population.

In conclusion, our comprehensive study encompassing CBA, FANCD2-Ub analysis, NGS, and advanced bioinformatics tools has yielded significant advancements in the diagnosis of Fanconi Anaemia (FA) within the Indian patient population. Through the identification of disease-causing genotypes and the establishment of a novel and more efficient molecular diagnosis strategy, our research holds the potential to enhance the diagnosis and management of FA in the Indian population.

The study provides valuable insights into the genotype-phenotype correlations in FA patients with different mutations in *FANCA*, *FANCL*, and *FANCG* genes. The findings highlight the importance of genetic analysis in understanding the clinical manifestations and disease progression in FA. The lower median age at diagnosis in patients with *FANCL* and *FANCG* mutations suggests that these mutations might lead to earlier disease onset or more severe symptoms.

The lack of significant correlations between specific gene mutations and physical abnormalities indicates that other factors, such as genetic modifiers or environmental influences, may play a role in shaping the clinical phenotype. Similarly, the absence of a strong association between severe aplastic anaemia and defective genes suggests

that other genetic or environmental factors may contribute to the development of this complication.

The integration of single-cell RNA sequencing data provides a comprehensive view of gene expression patterns in FA patients and healthy individuals.

The findings from our investigation of autosomal and X-linked recessive inheritance genes in FA patients have revealed intriguing and potentially significant insights into the genetic basis of the disease. Notably, we identified two genes linked to DNA repair, *LIG4* and *ATR*, both of which exhibited pathogenic variants in a subset of the FA patients. DNA repair is a critical cellular process, and mutations in genes involved in this pathway can lead to genomic instability, contributing to the development of FA.

In addition to DNA repair genes, we also discovered pathogenic variants in gene *GATA1*, which is known to be involved in haematopoiesis. Interestingly, alterations in *GATA1* have been associated with bone marrow failure, indicating its potential role in the pathogenesis of FA and its connection to the bone marrow phenotype observed in these patients.

Another notable finding was the identification of pathogenic variants in the gene *WRAP53*, a key player in telomere biology. Telomeres play a crucial role in protecting the ends of chromosomes from degradation and maintaining genomic stability. Dysfunction in telomere maintenance has been linked to various bone marrow failure syndromes, and our discovery of a pathogenic variant in WRAP53 suggests a potential involvement of telomere biology in FA pathogenesis.

Furthermore, when we focused on genes with dominant inheritance patterns, we observed that most pathogenic variants were present in genes involved in haematopoiesis. This observation strengthens the connection between haematopoietic abnormalities and FA, further supporting the notion that defects in haematopoietic stem cells contribute to the bone marrow failure observed in FA patients.

Of particular interest was the unanticipated discovery of pathogenic variants in genes associated with ribosome biology in a subset of FA patients. The ribosome is primarily known for its role in protein synthesis, and its connection to FA has not been well-established. Our analysis highlights a potential link between ribosome dysfunction and the pathogenesis of FA or its associated symptoms, opening up a new avenue of research in understanding the molecular mechanisms underlying the disease.

In summary, our study provides valuable insights into the genetic landscape of FA and its potential association with bone marrow failure syndromes. The identification of pathogenic variants in genes involved in DNA repair, haematopoiesis, telomere biology, and ribosome biology expands our understanding of the complex molecular pathways that contribute to the development and progression of FA. Further research is warranted to explore the functional consequences of these genetic variants and their impact on the pathophysiology of FA, which may ultimately lead to the development of targeted therapeutic strategies for this disorder.

The generation of patient-specific induced pluripotent stem cells (iPSCs) holds great promise for disease modelling and understanding genetic disorders. We could successfully generate iPSCs from FA patients by using inducible complementation.

Specifically, iPSCs were derived from fibroblasts with mutations in the *FANCA, FANCC, FANCF, FANCI, FANCL,* and *FANCT* genes. To restore the function of the mutated genes, the fibroblasts were transduced with doxycycline-inducible lentiviral vectors carrying normal cDNAs of the respective FA genes. These vectors allowed for doxycycline-inducible expression of both the FA gene and a fluorescent marker (mCherry), facilitating easy monitoring of gene expression levels. To confirm successful complementation, FANCD2 monoubiquitination (FANCD2-Ub) analysis was performed in the transduced fibroblasts. The cells were cultured with or without doxycycline and treated with an interstrand crosslinking agent (mitomycin C) to induce interstrand crosslinks (ICLs). The results demonstrated that FANCD2-Ub was restored in all patient fibroblasts after doxycycline supplementation. Notably, MMC treatment was not required for most complemented cells, except for FANCA transduced cells, which showed restoration of FANCD2-Ub only after MMC treatment. FANCD2 ubiquitination was not observed in cells cultured without doxycycline.

Furthermore, the transduced cells exhibited transgene expression, as evidenced by mCherry fluorescence upon doxycycline induction. Importantly, there was no detectable leakage of transgene expression in cells cultured without doxycycline, indicating tight control of gene expression by the doxycycline-inducible lentiviral vectors used in this study. The restoration of the active FA pathway was observed in the transduced cells when treated with an ICL agent in the presence of doxycycline, achieved by reconstituting the FA core complex responsible for FANCD2

ubiquitination. Interestingly, FANCD2 monoubiquitination was observed in FANCF-deficient fibroblasts after lentiviral transduction, even in cells cultured without doxycycline after MMC treatment. This finding suggests the existence of potential compensatory mechanisms or alternative pathways in these cells, which warrant further investigation.

The comparison between pSJL225 and pInducer20 vectors showed that pSJL225 utilizes an optimized rtTA variant called Tet-On 3G, resulting in tighter control of gene expression. The optimized rtTA variant contains three amino acid changes that increase doxycycline sensitivity. Additionally, pSJL225 and pInducer20 differ in the direction of expression cassettes, which may contribute to tighter control of gene expression in pSJL225. The lack of mCherry expression in pSJL225 without doxycycline confirms the vector's effective prevention of basal or background expression, further supporting its tight control over gene expression. The withdrawal of doxycycline resulted in the loss of mCherry expression and progressive cell death in iPSC lines complemented with *FANCA*, *FANCC*, *FANCL*, and *FANCT* genes, indicating the importance of an active FA pathway for cell survival. However, iPSC lines complemented with *FANCF* and *FANCI* genes did not show cell death upon doxycycline withdrawal. The iPSC clones were derived from fibroblasts with *FANCF* and *FANCI* mutations. Interestingly, these clones did not undergo cell death when doxycycline was withdrawn, unlike clones with other FA gene mutations. However, Sanger sequencing confirmed the presence of mutations in the iPSC clones. FANCD2 monoubiquitination was unaffected in *FANCF* iPSCs after doxycycline withdrawal,

indicating that the low-level expression of *FANCF* in these clones allowed them to survive. Of *FANCI* iPSC clones, only one clone showed mCherry expression after doxycycline induction. Blasticidin selection resulted in cell death in most clones, suggesting a lack of integration of the pSJL225-*FANCI* complementation vector. These findings suggest that *FANCI* may not be required for reprogramming, but the absence of FANCD2 ubiquitination in all *FANCI* iPSC clones highlights the non-canonical functions of FA pathway proteins in reprogramming and iPSC maintenance. Additionally, the data suggest that the presence or absence of FANCD2 ubiquitination is not directly linked to the survival or death of iPSCs. These observations are of great significance as they reveal the existence of additional intricate mechanisms that govern the FA pathway and its potential association with cell death, which may have important implications for understanding the bone marrow failure observed in FA patients.

Cell cycle analysis was conducted on FA-iPSCs carrying defective FA pathway genes (*FANCA*, *FANCC*, *FANCF*, *FANCI*, *FANCL*, and *FANCT*) to examine their cell cycle progression in the absence of doxycycline. The results showed that cells from *FANCA*, *FANCC*, and *FANCT* displayed a significant arrest at the G2/M phase compared to the cells cultured with doxycycline. This observation suggests that the loss of doxycycline-inducible FA transgene expression in FA-iPSCs leads to disruption in the normal cell cycle progression. Only one clone from *FANCL* demonstrated increased G2/M arrest, possibly due to abnormal karyotypes shared by both clones. However, G2/M arrest was not observed in any clones from *FANCF* and *FANCI* after doxycycline withdrawal. The absence of G2/M arrest in *FANCF* clones

can be attributed to the remaining low expression of FANCF, whereas the lack of G2/M arrest in the *FANCI* clone may be due to compensatory activation of alternative DNA repair mechanisms and/or defects in signaling pathways involved in cell cycle regulation and DNA damage response. Furthermore, as G2/M arrest was observed even without ICL inducer (MMC) we hypothesize that this G2/M arrest might be associated with the presence of endogenous DNA-damaging agents within the cells. The defective FA pathway in FA-iPSCs could contribute to the accumulation of such DNA-damaging agents, resulting in cellular stress and activation of cell cycle checkpoints to halt progression through the cell cycle. These findings highlight the importance of the FA pathway in maintaining proper cell cycle regulation and DNA repair processes. The G2/M arrest observed in FA-iPSCs without doxycycline induction suggests that the FA pathway plays a crucial role in preventing the accumulation of endogenous DNA damage and maintaining genomic stability.

The haematopoietic cells derived from *FANCC* and *FANCT* clones in the presence of doxycycline yielded significantly higher numbers of haematopoietic stem/progenitor cells (HSPCs) compared to cells derived in the absence of doxycycline. Additionally, functional assays were performed to evaluate the haematopoietic colony-forming potential of FA-deficient HSPCs. The clonogenicity of *FANCT* and *FANCC* HSPCs generated in the absence of doxycycline was markedly reduced compared to FA-proficient HSPCs (generated in the presence of doxycycline) in colony formation assays, indicating impaired haematopoietic colony formation in FA-deficient HSPCs derived from iPSCs. These findings demonstrate that the iPSC

model successfully recapitulates the characteristic dysfunction of haematopoietic progenitor cells observed in FA patients.

6. SUMMARY AND CONCLUSIONS

In the first objective to establish comprehensive strategy to characterize FA, compared the effectiveness of two diagnostic tests, chromosome breakage analysis (CBA) and FANCD2-ubiquitination (FANCD2-Ub) analysis. We analysed these tests in 153 Indian FA patients. Exome sequencing (NGS) was used to genotype the patients, which identified disease-causing genotypes in 95.4% of cases. A *FANCL* founder mutation and high frequencies of *FANCA* and *FANCG* mutations were found, leading to the development of a faster and cost-effective molecular diagnosis strategy for Indian FA patients. While CBA is considered the "gold standard," it has limitations due to the standardization challenges and the interpretation variability. The study highlighted the suitability of FANCD2-Ub analysis for diagnosing FA, advocating for its inclusion alongside CBA to enhance the diagnosis accuracy. The study also found that a specific CBA score cut-off of 15 improved sensitivity in FA diagnosis. Additionally, FANCD2-Ub analysis helped identify FA cases with low CBA scores, indicating its importance in such cases.

Downstream gene defects were rare in the cohort, emphasizing the reliability of FANCD2-Ub analysis in diagnosing FA. While less frequent, the detection of mosaicism underscores the significance of conducting FANCD2-Ub analysis for diagnosis. The study emphasized the challenges in assessing the pathogenicity of missense mutations. Computational algorithms like EVE showed promise in accurately predicting variant pathogenicity. Accurate detection of defective genes and

pathogenic variants is crucial for genetic counselling, prenatal testing, and gene therapy for FA. Whole exome sequencing (WES) could expand the knowledge of the genetic landscape of FA in the Indian population. Improved bioinformatics approaches like ExomeDepth facilitated the efficient detection of copy number variants (CNVs), with deletions in *FANCA* being prevalent in Indian FA patients.

FANCA, *FANCL*, and *FANCG* were the most frequently mutated genes in the cohort. Consanguinity contributed to homozygosity for pathogenic variants, and the study identified 10 FA genes as potential candidates for the FA genotyping panel. Overall, the research provides valuable insights into the genetic basis of FA in the Indian population and proposes a more efficient molecular diagnosis strategy.

The research emphasizes the significance of genetic analysis in understanding disease progression and manifestations. Patients with *FANCL* and *FANCG* mutations tend to be diagnosed at a younger age, suggesting potential differences in disease severity or earlier onset of symptoms. Specific gene mutations were not strongly associated with particular physical abnormalities, implying the involvement of other factors like genetic modifiers or environmental influences in shaping the clinical phenotype. By incorporating single-cell RNA sequencing data, the study provides a comprehensive view of gene expression patterns in FA patients and healthy individuals. The discovery of pathogenic variants in genes related to ribosome biology opens new possibilities for understanding FA pathogenesis and identifying potential therapeutic targets related to ribosome dysfunction. The identification of disease-modifying genes underscores the complexity of FA, emphasizing the need for

personalized approaches to treatment. Understanding the genetic landscape of FA may lead to targeted therapies and better disease management. In conclusion, the research in the first objective of the study contributed to the understanding of FA and highlights potential therapeutic targets. However, further research with larger cohorts and functional studies is necessary to validate and expand upon these findings, ultimately leading to improved clinical outcomes for FA patients.

The integration of single-cell RNA sequencing data provided a comprehensive view of gene expression patterns in FA patients and healthy individuals. Furthermore, the investigation of autosomal and X-linked recessive inheritance genes in FA patients revealed significant insights into the genetic basis of the disease. Notably, pathogenic variants were identified in genes involved in DNA repair (*LIG4* and *ATR*), haematopoiesis (*GATA1*), telomere biology (*WRAP53*), and ribosome biology. These findings shed light on potential molecular mechanisms underlying FA and its associated symptoms.

In conclusion, the study enhances our understanding of the genetic landscape of FA and its potential connection to bone marrow failure syndromes. Further research is needed to explore the functional implications of these genetic variants, which may lead to targeted therapeutic strategies for FA.

This study successfully generated induced pluripotent stem cells (iPSCs) from FA patients with mutations in the *FANCA*, *FANCC*, *FANCF*, *FANCI*, *FANCL*, and *FANCT* genes. The iPSCs were created using doxycycline-inducible lentiviral vectors carrying normal cDNAs of the respective FA genes. The restoration of the active FA

pathway was observed in the transduced cells, as evidenced by the reconstitution of FANCD2 monoubiquitination when treated with an interstrand crosslinking agent. The study also showed tight control of gene expression using the doxycycline-inducible lentiviral vectors.

The iPSCs derived from fibroblasts with *FANCF* and *FANCI* mutations showed intriguing characteristics, suggesting potential compensatory mechanisms or alternative pathways in these cells. However, further investigation is required to understand the non-canonical functions of FA pathway proteins in reprogramming and iPSC maintenance. Cell cycle analysis revealed disruptions in normal cell cycle progression in FA-iPSCs when the doxycycline-inducible FA transgene expression was lost. Additionally, hematopoietic cells derived from *FANCC* and *FANCT* iPSCs displayed impaired colony-forming potential, indicating dysfunctional hematopoietic progenitor cells, which is characteristic of Fanconi Anaemia. This study represents one of the most extensive investigations that successfully generated iPSCs for six FA complementation groups and thoroughly examined their respective phenotypes.

Overall, this iPSC model provides valuable insights into the genetic mechanisms and potential therapeutic targets for Fanconi Anaemia. However, further research is needed to fully validate and expand upon these findings, which could lead to improved understanding and treatment of the condition.

7. BIBLIOGRAPHY

Abad E, Samino S, Grodzicki RL, et al. (2021) Identification of metabolic changes leading to cancer susceptibility in Fanconi anemia cells. *Cancer Letters* 503: 185–196. DOI: 10.1016/j.canlet.2020.12.010.

AFTAB I, IRAM S, KHALIQ Saba, et al. (2017) Analysis of FANCC gene mutations (IVS4+4A>T, del322G, and R548X)in patients with Fanconi anemia in Pakistan. *TURKISH JOURNAL OF MEDICAL SCIENCES* 47: 391–398. DOI: 10.3906/sag-1506-53.

Akhurst RJ and Hata A (2012) Targeting the TGFβ signalling pathway in disease. *Nature Reviews Drug Discovery* 11(10): 790–811. DOI: 10.1038/nrd3810.

Alexeyev MF, Fayzulin R, Shokolenko IN, et al. (2010) A retro-lentiviral system for doxycycline-inducible gene expression and gene knockdown in cells with limited proliferative capacity. *Molecular Biology Reports* 37(4): 1987–1991. DOI: 10.1007/s11033-009-9647-7.

Alter BP (2003) Cancer in Fanconi anemia, 1927-2001. *Cancer* 97(2): 425–440. DOI: 10.1002/cncr.11046.

Alter Blanche P (2014) Fanconi anemia and the development of leukemia. *Best practice & research. Clinical haematology* 27(3–4): 214–221. DOI: 10.1016/j.beha.2014.10.002.

Alter Blanche P. (2014) Fanconi anemia and the development of leukemia. *Best Practice & Research Clinical Haematology* 27(3–4): 214–221. DOI: 10.1016/j.beha.2014.10.002.

Alter BP and Giri N (2016) Thinking of VACTERL-H? Rule out Fanconi Anemia according to PHENOS. *American Journal of Medical Genetics Part A* 170(6): 1520–1524. DOI: 10.1002/ajmg.a.37637.

Alter BP, Joenje H, Oostra AB, et al. (2005) Fanconi Anemia: Adult Head and Neck Cancer and Hematopoietic Mosaicism. *Archives of Otolaryngology–Head & Neck Surgery* 131(7): 635–639. DOI: 10.1001/archotol.131.7.635.

Alter BP, Rosenberg PS and Brody LC (2006) Clinical and molecular features associated with biallelic mutations in FANCD1/BRCA2. *Journal of Medical Genetics* 44(1): 1–9. DOI: 10.1136/jmg.2006.043257.

Alter BP, Giri N, Savage SA, et al. (2013) Squamous cell carcinomas in patients with Fanconi anemia and dyskeratosis congenita: A search for human papillomavirus. *International Journal of Cancer* 133(6): 1513–1515. DOI: 10.1002/ijc.28157.

Alter BP, Giri N, Savage SA and Rosenberg Philip S (2018) Cancer in the National Cancer Institute inherited bone marrow failure syndrome cohort after fifteen years of follow-up. *Haematologica* 103(1): 30–39. DOI: 10.3324/haematol.2017.178111.

Alter BP, Giri N, Savage SA and Rosenberg Philip S. (2018) Cancer in the national cancer institute inherited bone marrow failure syndrome cohort after fifteen years of follow-up. *Haematologica* 103(1). Ferrata Storti Foundation: 30–39. DOI: 10.3324/haematol.2017.178111.

Altintas B, Giri N, McReynolds LJ, et al. (2023) Genotype-phenotype and outcome associations in patients with Fanconi anemia: the National Cancer Institute cohort. *Haematologica* 108(1). Ferrata Storti Foundation: 69–82. DOI: 10.3324/haematol.2021.279981.

Ameziane N, Sie D, Dentro S, et al. (2012) Diagnosis of fanconi anemia: Mutation analysis by next-generation sequencing. *Anemia* 2012. DOI: 10.1155/2012/132856.

Ameziane N, May P, Haitjema A, et al. (2015) A novel Fanconi anaemia subtype associated with a dominant-negative mutation in RAD51. *Nature Communications* 6(1): 8829. DOI: 10.1038/ncomms9829.

Amouri A, Talmoudi F, Messaoud O, d'Enghien Catherine D, et al. (2014) High frequency of exon 15 deletion in the FANCA gene in Tunisian patients affected with Fanconi anemia disease: implication for diagnosis. *Molecular genetics & genomic medicine* 2(2): 160–165. DOI: 10.1002/mgg3.55.

Amouri A, Talmoudi F, Messaoud O, d'Enghien Catherine D., et al. (2014) High frequency of exon 15 deletion in the FANCA gene in Tunisian patients affected with Fanconi anemia disease: implication for diagnosis. *Molecular Genetics & Genomic Medicine* 2(2): 160–165. DOI: 10.1002/mgg3.55.

Asur RS, Kimble DC, Lach FP, et al. (2018) Somatic mosaicism of an intragenic FANCB duplication in both fibroblast and peripheral blood cells observed in a Fanconi anemia patient leads to milder phenotype. *Molecular Genetics & Genomic Medicine* 6(1): 77–91. DOI: 10.1002/mgg3.350.

Auerbach AD (2009) Fanconi anemia and its diagnosis. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 668(1–2): 4–10. DOI: 10.1016/j.mrfmmm.2009.01.013.

Auton A, Abecasis GR, Altshuler DM, et al. (2015) A global reference for human genetic variation. *Nature* 526(7571): 68–74. DOI: 10.1038/nature15393.

Bagby GC (2016) Multifunctional Fanconi proteins, inflammation and the Fanconi phenotype. *EBioMedicine* 8: 10–11. DOI: 10.1016/j.ebiom.2016.06.005.

Bakker ST, van de Vrugt HJ, Rooimans MA, et al. (2009) Fancm-deficient mice reveal unique features of Fanconi anemia complementation group M. *Human Molecular Genetics* 18(18): 3484–3495. DOI: 10.1093/hmg/ddp297.

Bakker ST, De Winter JP and Te Riele H (2013) Learning from a paradox: Recent insights into Fanconi anaemia through studying mouse models. *DMM Disease Models and Mechanisms*. DOI: 10.1242/dmm.009795.

Bakker ST, de Winter JP, te Riele Hein, et al. (2013) Learning from a paradox: recent insights into Fanconi anaemia through studying mouse models. *Disease models & mechanisms* 6(1): 40–7. DOI: 10.1242/dmm.009795.

Balkwill F and Mantovani A (2001) Inflammation and cancer: back to Virchow? *The Lancet* 357(9255): 539–545. DOI: 10.1016/S0140-6736(00)04046-0.

Barde I, Zanta-Boussif MA, Paisant S, et al. (2006) Efficient control of gene expression in the hematopoietic system using a single Tet-on inducible lentiviral vector. *Molecular Therapy* 13(2): 382–390. DOI: 10.1016/j.ymthe.2005.09.012.

Baron U, Freundlieb S, Gossen M, et al. (1995) Co-regulation of two gene activities by tetracycline via a bidirectional promoter. *Nucleic Acids Research* 23(17): 3605–3606. DOI: 10.1093/nar/23.17.3605.

Beere HM (2001) Stressed to Death: Regulation of Apoptotic Signaling Pathways by the Heat Shock Proteins. *Science's STKE* 2001(93). DOI: 10.1126/stke.2001.93.re1.

Ben Haj Ali A, Messaoud O, Elouej S, et al. (2021) FANCA Gene Mutations in North African Fanconi Anemia Patients. *Frontiers in genetics* 12: 610050. DOI: 10.3389/fgene.2021.610050.

Benabdellah K, Cobo M, Muñoz P, et al. (2011) Development of an All-in-One Lentiviral Vector System Based on the Original TetR for the Easy Generation of Tet-ON Cell Lines. *PLoS ONE* 6(8): e23734. DOI: 10.1371/journal.pone.0023734.

Benabdellah K, Muñoz P, Cobo M, et al. (2016) Lent-On-Plus Lentiviral vectors for conditional expression in human stem cells. *Scientific Reports* 6(1): 37289. DOI: 10.1038/srep37289.

Beswick RA, Dorrance AM, Leite R, et al. (2001) NADH/NADPH Oxidase and Enhanced Superoxide Production in the Mineralocorticoid Hypertensive Rat. *Hypertension* 38(5): 1107–1111. DOI: 10.1161/hy1101.093423.

Bharathan SP, Nandy K, Palani D, et al. (2017) Generation of an induced pluripotent stem cell line that mimics the disease phenotypes from a patient with Fanconi anemia by conditional complementation. *Stem Cell Research* 20. Elsevier B.V.: 54–57. DOI: 10.1016/j.scr.2017.02.006.

Bijangi-Vishehsaraei K, Saadatzadeh MR, Werne A, et al. (2005) Enhanced TNF-α–induced apoptosis in Fanconi anemia type C–deficient cells is dependent on apoptosis signal-regulating kinase 1. *Blood* 106(13): 4124–4130. DOI: 10.1182/blood-2005-05-2096.

Bluteau D, Masliah-Planchon J, Clairmont C, et al. (2016) Biallelic inactivation of REV7 is associated with Fanconi anemia. *Journal of Clinical Investigation* 126(9): 3580–3584. DOI: 10.1172/JCI88010.

Boersma V, Moatti N, Segura-Bayona S, et al. (2015) MAD2L2 controls DNA repair at telomeres and DNA breaks by inhibiting 5′ end resection. *Nature* 521(7553): 537–540. DOI: 10.1038/nature14216.

Bogliolo M, Schuster B, Stoepker C, et al. (2013) Mutations in ERCC4, Encoding the DNA-Repair Endonuclease XPF, Cause Fanconi Anemia. *The American Journal of Human Genetics* 92(5): 800–806. DOI: 10.1016/j.ajhg.2013.04.002.

Bogliolo M, Bluteau D, Lespinasse J, et al. (2018) Biallelic truncating FANCM mutations cause early-onset cancer but not Fanconi anemia. *Genetics in Medicine* 20(4): 458–463. DOI: 10.1038/gim.2017.124.

Bogliolo M, Pujol R, Aza-Carmona M, et al. (2020) Optimised molecular genetic diagnostics of Fanconi anaemia by whole exome sequencing and functional studies. *Journal of Medical Genetics* 57(4). BMJ Publishing Group: 258–268. DOI: 10.1136/jmedgenet-2019-106249.

Bottega R, Nicchia E, Cappelli E, et al. (2018) Hypomorphic FANCA mutations correlate with mild mitochondrial and clinical phenotype in Fanconi anemia. *Haematologica* 103(3): 417–426. DOI: 10.3324/haematol.2017.176131.

Brégnard C, Guerra J, Déjardin S, et al. (2016) Upregulated LINE-1 Activity in the Fanconi Anemia Cancer Susceptibility Syndrome Leads to Spontaneous Pro-inflammatory Cytokine Production. *EBioMedicine* 8: 184–194. DOI: 10.1016/j.ebiom.2016.05.005.

Briot D, Macé-Aimé G, Subra F, et al. (2008) Aberrant activation of stress-response pathways leads to TNF-α oversecretion in Fanconi anemia. *Blood* 111(4): 1913–1923. DOI: 10.1182/blood-2007-07-099218.

Brosh RM, Bellani M, Liu Y, et al. (2017) Fanconi Anemia: A DNA repair disorder characterized by accelerated decline of the hematopoietic stem cell compartment and other features of aging. *Ageing Research Reviews* 33: 67–75. DOI: 10.1016/j.arr.2016.05.005.

Brüne B, Dehne N, Grossmann N, et al. (2013) Redox Control of Inflammation in Macrophages. *Antioxidants & Redox Signaling* 19(6): 595–637. DOI: 10.1089/ars.2012.4785.

Byrd PJ, Stewart GrantS, Smith A, et al. (2016) A Hypomorphic PALB2 Allele Gives Rise to an Unusual Form of FA-N Associated with Lymphoid Tumour Development. *PLOS Genetics* 12(3): e1005945. DOI: 10.1371/journal.pgen.1005945.

Callen E (2005) A common founder mutation in FANCA underlies the world's highest prevalence of Fanconi anemia in Gypsy families from Spain. *Blood* 105(5): 1946–1949. DOI: 10.1182/blood-2004-07-2588.

Callén E, Tischkowitz MD, Creus A, et al. (2004) Quantitative PCR analysis reveals a high incidence of large intragenic deletions in the FANCA gene in Spanish Fanconi anemia patients. *Cytogenetic and genome research* 104(1–4). Switzerland: 341–345. DOI: 10.1159/000077513.

Cappelli E, Ravera S, Vaccaro D, et al. (2013) Mitochondrial respiratory complex I defects in Fanconi anemia. *Trends in Molecular Medicine* 19(9): 513–514. DOI: 10.1016/j.molmed.2013.07.008.

Cappelli E, Cuccarolo P, Stroppiana G, et al. (2017) Defects in mitochondrial energetic function compels Fanconi Anaemia cells to glycolytic metabolism. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1863(6): 1214–1221. DOI: 10.1016/j.bbadis.2017.03.008.

Castella M., Pujol R, Callen E, et al. (2011) Chromosome fragility in patients with Fanconi anaemia: diagnostic implications and clinical impact. *Journal of Medical Genetics* 48(4): 242–250. DOI: 10.1136/jmg.2010.084210.

Castella Maria, Pujol R, Callén E, Trujillo JP, Casado JA, Gille H, Lach FP, Auerbach AD, Schindler D, Benítez J, Porto B, Ferro T, Muñoz A, et al. (2011) Origin, functional role, and clinical impact of Fanconi anemia FANCA mutations. *Blood* 117(14): 3759–3769. DOI: 10.1182/blood-2010-08-299917.

Castella Maria, Pujol R, Callén E, Trujillo JP, Casado JA, Gille H, Lach FP, Auerbach AD, Schindler D, Benítez J, Porto B, Ferro T, Arturo M, et al. (2011) Origin, functional role, and clinical impact of fanconi anemia fanca mutations. *Blood* 117(14). The American Society of Hematology: 3759–3769. DOI: 10.1182/blood-2010-08-299917.

Catucci I, Osorio A, Arver B, et al. (2018) Individuals with FANCM biallelic mutations do not develop Fanconi anemia, but show risk for breast cancer, chemotherapy toxicity and may display chromosome fragility. *Genetics in Medicine* 20(4): 452–457. DOI: 10.1038/gim.2017.123.

Ceccaldi R, Briot D, Larghero J, et al. (2011) Spontaneous abrogation of the G2 DNA damage checkpoint has clinical benefits but promotes leukemogenesis in Fanconi anemia patients. *Journal of Clinical Investigation* 121(1): 184–194. DOI: 10.1172/JCI43836.

Ceccaldi R, Parmar K, Mouly E, Delord M, Kim Jung Min, et al. (2012) Bone marrow failure in Fanconi anemia is triggered by an exacerbated p53/p21 DNA damage response that impairs hematopoietic stem and progenitor cells. *Cell stem cell* 11(1): 36–49. DOI: 10.1016/j.stem.2012.05.013.

Ceccaldi R, Parmar K, Mouly E, Delord M, Kim Jung Min, et al. (2012) Bone Marrow Failure in Fanconi Anemia Is Triggered by an Exacerbated p53/p21 DNA Damage Response that Impairs Hematopoietic Stem and Progenitor Cells. *Cell Stem Cell* 11(1): 36–49. DOI: 10.1016/j.stem.2012.05.013.

Centlivre M, Zhou X, Pouw SM, et al. (2010) Autoregulatory lentiviral vectors allow multiple cycles of doxycycline-inducible gene expression in human hematopoietic cells in vivo. *Gene Therapy* 17(1): 14–25. DOI: 10.1038/gt.2009.109.

Cerami E, Gao J, Dogrusoz U, et al. (2012) The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discovery* 2(5): 401–404. DOI: 10.1158/2159-8290.CD-12-0095.

Chait R, Palmer AC, Yelin I, et al. (2016) Pervasive selection for and against antibiotic resistance in inhomogeneous multistress environments. *Nature Communications* 7(1): 10333. DOI: 10.1038/ncomms10333.

Chandra S, Levran O, Jurickova I, et al. (2005) A Rapid Method for Retrovirus-Mediated Identification of Complementation Groups in Fanconi Anemia Patients. *Molecular Therapy* 12(5): 976–984. DOI: https://doi.org/10.1016/j.ymthe.2005.04.021.

Chandrasekharappa SC, Lach FP, Kimble DC, et al. (2013) Massively parallel sequencing, aCGH, and RNA-Seq technologies provide a comprehensive molecular diagnosis of Fanconi anemia. *Blood* 121(22): e138–e148. DOI: 10.1182/blood-2012-12-474585.

Chang L, Yuan W, Zeng H, et al. (2014) Whole exome sequencing reveals concomitant mutations of multiple FA genes in individual Fanconi anemia patients. *BMC Medical Genomics* 7(1): 24. DOI: 10.1186/1755-8794-7-24.

Chaudhury I, Sareen A, Raghunandan M, et al. (2013) FANCD2 regulates BLM complex functions independently of FANCI to promote replication fork recovery. *Nucleic Acids Research* 41(13): 6444–6459. DOI: 10.1093/nar/gkt348.

Chen C, Liu Yu, Liu Yang, et al. (2010) Mammalian target of rapamycin activation underlies HSC defects in autoimmune disease and inflammation in mice. *Journal of Clinical Investigation* 120(11): 4091–4101. DOI: 10.1172/JCI43873.

Chen C-H, Budas GR, Churchill EN, et al. (2008) Activation of Aldehyde Dehydrogenase-2 Reduces Ischemic Damage to the Heart. *Science* 321(5895): 1493–1495. DOI: 10.1126/science.1158554.

Chen C-H, Ferreira JCB, Gross ER, et al. (2014) Targeting Aldehyde Dehydrogenase 2: New Therapeutic Opportunities. *Physiological Reviews* 94(1): 1–34. DOI: 10.1152/physrev.00017.2013.

Chen J, Feng X, Desierto MJ, et al. (2015) IFN-γ-mediated hematopoietic cell destruction in murine models of immune-mediated bone marrow failure. *Blood* 126(24): 2621–2631. DOI: 10.1182/blood-2015-06-652453.

Chen M, Tomkins DJ, Auerbach W, et al. (1996) Inactivation of Fac in mice produces inducible chromosomal instability and reduced fertility reminiscent of Fanconi anaemia. *Nature Genetics* 12(4): 448–451. DOI: 10.1038/ng0496-448.

Cheung RS and Taniguchi T (2017) Recent insights into the molecular basis of Fanconi anemia: genes, modifiers, and drivers. *International Journal of Hematology*. Springer Tokyo. DOI: 10.1007/s12185-017-2283-4.

Cheung RS, Castella M, Abeyta A, et al. (2017) Ubiquitination-Linked Phosphorylation of the FANCI S/TQ Cluster Contributes to Activation of the Fanconi Anemia I/D2 Complex. *Cell Reports* 19(12): 2432–2440. DOI: 10.1016/j.celrep.2017.05.081.

Chihanga T, Vicente-Muñoz S, Ruiz-Torres S, et al. (2022) Head and Neck Cancer Susceptibility and Metabolism in Fanconi Anemia. *Cancers* 14(8): 2040. DOI: 10.3390/cancers14082040.

Chirnomas SD and Kupfer GM (2013) The inherited bone marrow failure syndromes. *Pediatric Clinics of North America* 60(6): 1291–1310. DOI: 10.1016/j.pcl.2013.09.007.

Chlon TM, Hoskins EE, Mayhew CN, et al. (2014) High-Risk Human Papillomavirus E6 Protein Promotes Reprogramming of Fanconi Anemia Patient Cells through Repression of p53 but Does Not Allow for Sustained Growth of Induced Pluripotent Stem Cells. *Journal of Virology* 88(19). American Society for Microbiology: 11315–11326. DOI: 10.1128/jvi.01533-14.

Chlon TM, Ruiz-Torres S, Maag L, et al. (2016) Overcoming Pluripotent Stem Cell Dependence on the Repair of Endogenous DNA Damage. *Stem Cell Reports* 6(1). Cell Press: 44–54. DOI: 10.1016/j.stemcr.2015.12.001.

Cohen SB and Denkers EY (2015) Impact of Toxoplasma gondii on Dendritic Cell Subset Function in the Intestinal Mucosa. *The Journal of Immunology* 195(6): 2754–2762. DOI: 10.4049/jimmunol.1501137.

Crossan GP, van der Weyden L, Rosado I V, et al. (2011) Disruption of mouse Slx4, a regulator of structure-specific nucleases, phenocopies Fanconi anemia. *Nature Genetics* 43(2): 147–152. DOI: 10.1038/ng.752.

Cuccarolo P, Viaggi S and Degan P (2012) New insights into redox response modulation in Fanconi's anemia cells by hydrogen peroxide and glutathione depletors. *FEBS Journal* 279(14): 2479–2494. DOI: 10.1111/j.1742-4658.2012.08629.x.

Cumming RC, Lightfoot J, Beard K, et al. (2001) Fanconi anemia group C protein prevents apoptosis in hematopoietic cells through redox regulation of GSTP1. *Nature Medicine* 7(7): 814–820. DOI: 10.1038/89937.

Dao K-HT, Rotelli MD, Petersen CL, et al. (2012) FANCL ubiquitinates β-catenin and enhances its nuclear function. *Blood* 120(2): 323–334. DOI: 10.1182/blood-2011-11-388355.

D'Autréaux B and Toledano MB (2007) ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nature Reviews Molecular Cell Biology* 8(10): 813–824. DOI: 10.1038/nrm2256.

De Groote P, Grootjans S, Lippens S, et al. (2016) Generation of a new Gateway-compatible inducible lentiviral vector platform allowing easy derivation of co-transduced cells. *BioTechniques* 60(5): 252–259. DOI: 10.2144/000114417.

De Rocco D, Bottega R, Cappelli E, et al. (2014) Molecular analysis of Fanconi anemia: The experience of the bone marrow failure study group of the Italian Association of Pediatric Onco-Hematology. *Haematologica* 99(6): 1022–1031. DOI: 10.3324/haematol.2014.104224.

de Vries Y, Lwiwski N, Levitus M, et al. (2012) A Dutch Fanconi Anemia FANCC Founder Mutation in Canadian Manitoba Mennonites. *Anemia* 2012: 1–6. DOI: 10.1155/2012/865170.

Deans AJ and West SC (2009) FANCM Connects the Genome Instability Disorders Bloom's Syndrome and Fanconi Anemia. *Molecular Cell* 36(6): 943–953. DOI: 10.1016/j.molcel.2009.12.006.

Deans AJ and West SC (2011) DNA interstrand crosslink repair and cancer. *Nature Reviews Cancer* 11(7): 467–480. DOI: 10.1038/nrc3088.

DeBerardinis RJ and Cheng T (2010) Q's next: the diverse functions of glutamine in metabolism, cell biology and cancer. *Oncogene* 29(3): 313–324. DOI: 10.1038/onc.2009.358.

Degan P, Bonassi S, Caterina M De, et al. (1995) In vivo accumulation of 8-hydroxy-2'-deoxyguanosine in DNA correlates with release of reactive oxygen species in Fanconi's anaemia families. *Carcinogenesis* 16(4): 735–742. DOI: 10.1093/carcin/16.4.735.

Degrolard-Courcet E, Sokolowska J, Padeano M-M, et al. (2014) Development of primary early-onset colorectal cancers due to biallelic mutations of the FANCD1/BRCA2 gene. *European Journal of Human Genetics* 22(8): 979–987. DOI: 10.1038/ejhg.2013.278.

Demuth I, Wlodarski M, Tipping AJ, et al. (2000) Spectrum of mutations in the Fanconi anaemia group G gene, FANCG/XRCC9. *European Journal of Human Genetics* 8(11): 861–868. DOI: 10.1038/sj.ejhg.5200552.

Dietz AC, Mehta PA, Vlachos A, et al. (2017) Current Knowledge and Priorities for Future Research in Late Effects after Hematopoietic Cell Transplantation for Inherited Bone Marrow Failure Syndromes: Consensus Statement from the Second Pediatric Blood and Marrow Transplant Consortium International Conference on Late Effects after Pediatric Hematopoietic Cell Transplantation. In: *Biology of Blood and Marrow Transplantation*, 1 May 2017, pp. 726–735. Elsevier Inc. DOI: 10.1016/j.bbmt.2017.01.075.

DiMauro S and Hirano M (2009) Pathogenesis and Treatment of Mitochondrial Disorders., pp. 139–170. DOI: 10.1007/978-90-481-2813-6_10.

Dokal I (2006) Fanconi's anaemia and related bone marrow failure syndromes. *British Medical Bulletin* 77–78(1): 37–53. DOI: 10.1093/bmb/ldl007.

Domchek SM, Tang J, Stopfer J, et al. (2013) Biallelic Deleterious BRCA1 Mutations in a Woman with Early-Onset Ovarian Cancer. *Cancer Discovery* 3(4): 399–405. DOI: 10.1158/2159-8290.CD-12-0421.

Donahue SL, Lundberg R, Saplis R, et al. (2003) Deficient Regulation of DNA Double-strand Break Repair in Fanconi Anemia Fibroblasts. *Journal of Biological Chemistry* 278(32): 29487–29495. DOI: 10.1074/jbc.M213251200.

Dong C, Wei P, Jian X, et al. (2015) Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Human molecular genetics* 24(8): 2125–2137. DOI: 10.1093/hmg/ddu733.

Donovan FX, Solanki A, Mori M, et al. (2020) A founder variant in the South Asian population leads to a high prevalence of FANCL Fanconi anemia cases in India. *Human Mutation* 41(1). John Wiley and Sons Inc.: 122–128. DOI: 10.1002/humu.23914.

Drake LE, Springer MZ, Poole LP, et al. (2017) Expanding perspectives on the significance of mitophagy in cancer. *Seminars in Cancer Biology* 47: 110–124. DOI: 10.1016/j.semcancer.2017.04.008.

Du W, Adam Z, Rani R, et al. (2008) Oxidative Stress in Fanconi Anemia Hematopoiesis and Disease Progression. *Antioxidants & Redox Signaling* 10(11): 1909–1921. DOI: 10.1089/ars.2008.2129.

Du W, Li J, Sipple J, et al. (2010) Cytoplasmic FANCA-FANCC Complex Interacts and Stabilizes the Cytoplasm-dislocalized Leukemic Nucleophosmin Protein (NPMc). *Journal of Biological Chemistry* 285(48): 37436–37444. DOI: 10.1074/jbc.M110.113209.

Du W, Rani R, Sipple J, et al. (2012) The FA pathway counteracts oxidative stress through selective protection of antioxidant defense gene promoters. *Blood* 119(18): 4142–4151. DOI: 10.1182/blood-2011-09-381970.

Dufour C (2003) TNF-alpha and IFN-gamma are overexpressed in the bone marrow of Fanconi anemia patients and TNF- suppresses erythropoiesis in vitro. *Blood* 102(6): 2053–2059. DOI: 10.1182/blood-2003-01-0114.

Dull T, Zufferey R, Kelly M, et al. (1998) A Third-Generation Lentivirus Vector with a Conditional Packaging System. *Journal of Virology* 72(11): 8463–8471. DOI: 10.1128/JVI.72.11.8463-8471.1998.

Durand S and Cimarelli A (2011) The Inside Out of Lentiviral Vectors. *Viruses* 3(2): 132–159. DOI: 10.3390/v3020132.

Duxin JP and Walter JC (2015) What is the DNA repair defect underlying Fanconi anemia? *Current Opinion in Cell Biology* 37: 49–60. DOI: 10.1016/j.ceb.2015.09.002.

Eppig JT, Blake JA, Bult CJ, et al. (2015) The Mouse Genome Database (MGD): facilitating mouse as a model for human biology and disease. *Nucleic Acids Research* 43(D1): D726–D736. DOI: 10.1093/nar/gku967.

Esmail Nia G, Fadaee M, Royer R, et al. (2016) Profiling Fanconi Anemia Gene Mutations among Iranian Patients. *Archives of Iranian medicine* 19(4). Iran: 236–240.

Esplin ED (2020) Addendum: American College of Medical Genetics guideline on the cytogenetic evaluation of the individual with developmental delay or mental retardation. *Genetics in Medicine* 22(12): 2128. DOI: 10.1038/s41436-020-0875-5.

Esposito D, Weile J, Shendure J, et al. (2019) MaveDB: An open-source platform to distribute and interpret data from multiplexed assays of variant effect. *Genome Biology* 20(1). Genome Biology: 1–11. DOI: 10.1186/s13059-019-1845-6.

Fagerlie SR, Diaz J, Christianson TA, et al. (2001) Functional correction of FA-C cells with FANCC suppresses the expression of interferon γ–inducible genes. *Blood* 97(10): 3017–3024. DOI: 10.1182/blood.V97.10.3017.

Faivre L, Guardiola P, Lewis C, et al. (2000) Association of complementation group and mutation type with clinical outcome in fanconi anemia. European Fanconi Anemia Research Group. *Blood* 96(13): 4064–70.

Faivre L, Portnoï MF, Pals G, et al. (2005) Should chromosome breakage studies be performed in patients with VACTERL association? *American Journal of Medical Genetics Part A* 137A(1): 55–58. DOI: 10.1002/ajmg.a.30853.

Falini B, Brunetti L, Sportoletti P, et al. (2020) NPM1-mutated acute myeloid leukemia: from bench to bedside. *Blood* 136(15): 1707–1721. DOI: 10.1182/blood.2019004226.

Fang H-Y, Chang C-L, Hsu S-H, et al. (2010) ATPase family AAA domain-containing 3A is a novel anti-apoptotic factor in lung adenocarcinoma cells. *Journal of Cell Science* 123(7): 1171–1180. DOI: 10.1242/jcs.062034.

Fargo JH, Rochowski A, Giri N, et al. (2014) Comparison of chromosome breakage in non-mosaic and mosaic patients with Fanconi anemia, relatives, and patients with other inherited bone marrow failure syndromes. *Cytogenetic and Genome Research* 144(1): 15–27. DOI: 10.1159/000366251.

Fiesco-Roa MO, Giri N, McReynolds LJ, et al. (2019) Genotype-phenotype associations in Fanconi anemia: A literature review. *Blood Reviews*. Churchill Livingstone. DOI: 10.1016/j.blre.2019.100589.

Fouquet B, Pawlikowska P, Caburet S, et al. (2017) A homozygous FANCM mutation underlies a familial case of non-syndromic primary ovarian insufficiency. *eLife* 6. DOI: 10.7554/eLife.30490.

Frazer J, Notin P, Dias M, Gomez A, Min Joseph K., et al. (2021) Disease variant prediction with deep generative models of evolutionary data. *Nature* 599(7883). Nature Research: 91–95. DOI: 10.1038/s41586-021-04043-8.

Frazer J, Notin P, Dias M, Gomez A, Min Joseph K, et al. (2021) Disease variant prediction with deep generative models of evolutionary data. *Nature* 599(November). Springer US. DOI: 10.1038/s41586-021-04043-8.

Freire BL, Homma TK, Funari MFA, et al. (2018) Homozygous loss of function BRCA1 variant causing a Fanconi-anemia-like phenotype, a clinical report and review of previous patients. *European Journal of Medical Genetics* 61(3): 130–133. DOI: 10.1016/j.ejmg.2017.11.003.

Fusaki N, Ban H, Nishiyama A, et al. (2009) Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proceedings of the Japan Academy, Series B* 85(8): 348–362. DOI: 10.2183/pjab.85.348.

Futaki M (2002) The FANCG Fanconi anemia protein interacts with CYP2E1: possible role in protection against oxidative DNA damage. *Carcinogenesis* 23(1): 67–72. DOI: 10.1093/carcin/23.1.67.

Garaycoechea JI and Patel KJ (2014) Why does the bone marrow fail in Fanconi anemia? *Blood* 123(1): 26–34. DOI: 10.1182/blood-2013-09-427740.

Garaycoechea JI, Crossan GP, Langevin F, et al. (2012) Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature* 489(7417): 571–575. DOI: 10.1038/nature11368.

Garbati MR, Hays LE, Keeble W, et al. (2013) FANCA and FANCC modulate TLR and p38 MAPK–dependent expression of IL-1β in macrophages. *Blood* 122(18): 3197–3205. DOI: 10.1182/blood-2013-02-484816.

Garbati MR, Hays LE, Rathbun RK, et al. (2016) Cytokine overproduction and crosslinker hypersensitivity are unlinked in Fanconi anemia macrophages. *Journal of Leukocyte Biology* 99(3): 455–465. DOI: 10.1189/jlb.3A0515-201R.

Gari K, Décaillet C, Stasiak AZ, et al. (2008) The Fanconi Anemia Protein FANCM Can Promote Branch Migration of Holliday Junctions and Replication Forks. *Molecular Cell* 29(1): 141–148. DOI: 10.1016/j.molcel.2007.11.032.

Gennery AR, Slatter MA, Bhattacharya A, et al. (2004) The clinical and biological overlap between Nijmegen Breakage Syndrome and Fanconi anemia. *Clinical Immunology* 113(2): 214–219. DOI: 10.1016/j.clim.2004.03.024.

George M, Solanki A, Chavan N, Rajendran A, Raj R, Mohan S, Nemani S, Kanvinde S, Munirathnam D, Rao S, Radhakrishnan N, Lashkari HP, et al. (2021) A comprehensive molecular study identified 12 complementation groups with 56 novel FANC gene variants in Indian Fanconi anemia subjects. *Human Mutation* 42(12). John Wiley and Sons Inc: 1648–1665. DOI: 10.1002/humu.24286.

George M, Solanki A, Chavan N, Rajendran A, Raj R, Mohan S, Nemani S, Kanvinde S, Munirathnam D, Rao S, Radhakrishnan N, Prasada H, et al. (2021) Comprehensive molecular study identified 12 complementation groups with 56 novel FANC gene variants in Indian Fanconi anemia subjects. *Human mutation*: 0–3. DOI: 10.1002/humu.24286.

Ghazwani Y, AlBalwi M, Al-Abdulkareem I, et al. (2016) Clinical characteristics and genetic subtypes of Fanconi anemia in Saudi patients. *Cancer Genetics* 209(4): 171–176. DOI: 10.1016/j.cancergen.2016.02.003.

Ghemlas I, Li H, Zlateska B, et al. (2015) Improving diagnostic precision, care and syndrome definitions using comprehensive next-generation sequencing for the

inherited bone marrow failure syndromes. *Journal of Medical Genetics* 52(9): 575–584. DOI: 10.1136/jmedgenet-2015-103270.

Gille JohanJP, Wortelboer HeleenM and Joenje H (1987) Antioxidant status of Fanconi anemia fibroblasts. *Human Genetics* 77(1). DOI: 10.1007/BF00284708.

Gillio AP, Verlander PC, Batish SD, et al. (1997) Phenotypic consequences of mutations in the Fanconi anemia FAC gene: an International Fanconi Anemia Registry study. *Blood* 90(1): 105–10.

Giry-Laterrière M, Cherpin O, Kim Y-S, et al. (2011) Polyswitch Lentivectors: "All-in-One" Lentiviral Vectors for Drug-Inducible Gene Expression, Live Selection, and Recombination Cloning. *Human Gene Therapy* 22(10): 1255–1267. DOI: 10.1089/hum.2010.179.

Gore A, Li Z, Fung H-L, et al. (2011) Somatic coding mutations in human induced pluripotent stem cells. *Nature* 471(7336): 63–67. DOI: 10.1038/nature09805.

Gossen M and Bujard H (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences* 89(12): 5547–5551. DOI: 10.1073/pnas.89.12.5547.

Gowen LC, Johnson BL, Latour AM, et al. (1996) Brca1 deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities. *Nature Genetics* 12(2): 191–194. DOI: 10.1038/ng0296-191.

Grompe M, Whitney MA, Low M, et al. (1995) Progenitor cell dysfunction in mice with fanconi anemia group C gene disruption. In: *Blood*, 1995, p. 972. WB SAUNDERS CO INDEPENDENCE SQUARE WEST CURTIS CENTER, STE 300, PHILADELPHIA ….

Gueiderikh A, Maczkowiak-Chartois F, Rouvet G, et al. (2021) Fanconi anemia A protein participates in nucleolar homeostasis maintenance and ribosome biogenesis. *Science Advances* 7(1). DOI: 10.1126/sciadv.abb5414.

Gueiderikh A, Maczkowiak-Chartois F and Rosselli F (2022) A new frontier in Fanconi anemia: From DNA repair to ribosome biogenesis. *Blood Reviews* 52: 100904. DOI: 10.1016/j.blre.2021.100904.

Habi O, Delisle M, Messier N, et al. (2005) Lack of self-renewal capacity in Fancc−/− stem cells after ex vivo expansion. *STEM CELLS* 23(8): 1135–1141. DOI: 10.1634/stemcells.2004-0356.

Hadjur S, Ung K, Wadsworth L, et al. (2001) Defective hematopoiesis and hepatic steatosis in mice with combined deficiencies of the genes encoding Fancc and Cu/Zn superoxide dismutase. *Blood* 98(4): 1003–1011. DOI: 10.1182/blood.V98.4.1003.

Haneline LS, Broxmeyer HE, Cooper S, et al. (1998) Multiple inhibitory cytokines induce deregulated progenitor growth and apoptosis in hematopoietic cells from Fac-/- mice. *Blood* 91(11): 4092–8.

Hanenberg H, Batish SD, Pollok KE, et al. (2002) Phenotypic correction of primary Fanconi anemia T cells with retroviral vectors as a diagnostic tool. *Experimental Hematology* 30(5): 410–420. DOI: 10.1016/S0301-472X(02)00782-8.

Hartmann L, Neveling K, Borkens S, et al. (2010) Correct mRNA Processing at a Mutant TT Splice Donor in FANCC Ameliorates the Clinical Phenotype in Patients and Is Enhanced by Delivery of Suppressor U1 snRNAs. *The American Journal of Human Genetics* 87(4): 480–493. DOI: 10.1016/j.ajhg.2010.08.016.

Hess CJ, Ameziane N, Schuurhuis GJ, et al. (2008) Hypermethylation of the FANCC and FANCL promoter regions in sporadic acute leukaemia. *Cellular oncology :*

the official journal of the International Society for Cellular Oncology 30(4): 299–306. DOI: 10.3233/clo-2008-0426.

Hill GE, Miller JA, Baxter BT, et al. (1998) Association of malondialdehyde–acetaldehyde (MAA) adducted proteins with atherosclerotic-induced vascular inflammatory injury. *Atherosclerosis* 141(1): 107–116. DOI: 10.1016/S0021-9150(98)00153-1.

Hira A, Yabe H, Yoshida K, et al. (2013) Variant ALDH2 is associated with accelerated progression of bone marrow failure in Japanese Fanconi anemia patients. *Blood* 122(18): 3206–3209. DOI: 10.1182/blood-2013-06-507962.

Hira A, Yoshida K, Sato K, et al. (2015) Mutations in the Gene Encoding the E2 Conjugating Enzyme UBE2T Cause Fanconi Anemia. *The American Journal of Human Genetics* 96(6): 1001–1007. DOI: 10.1016/j.ajhg.2015.04.022.

Ho TT, Warr MR, Adelman ER, et al. (2017) Autophagy maintains the metabolism and function of young and old stem cells. *Nature* 543(7644): 205–210. DOI: 10.1038/nature21388.

Hou H, Li D, Gao J, et al. (2020) Proteomic profiling and bioinformatics analysis identify key regulators during the process from fanconi anemia to acute myeloid leukemia. *American journal of translational research* 12(4): 1415–1427.

Howlett NG, Taniguchi T, Olson S, et al. (2002) Biallelic Inactivation of BRCA2 in Fanconi Anemia. *Science* 297(5581): 606–609. DOI: 10.1126/science.1073834.

Howlett NG, Taniguchi T, Durkin SG, et al. (2005) The Fanconi anemia pathway is required for the DNA replication stress response and for the regulation of common fragile site stability. *Human Molecular Genetics* 14(5): 693–701. DOI: 10.1093/hmg/ddi065.

Hoyng SA, Gnavi S, de Winter F, et al. (2014) Developing a potentially immunologically inert tetracycline-regulatable viral vector for gene therapy in the peripheral nerve. *Gene Therapy* 21(6): 549–557. DOI: 10.1038/gt.2014.22.

Hsu C-C, Tseng L-M and Lee H-C (2016) Role of mitochondrial dysfunction in cancer progression. *Experimental Biology and Medicine* 241(12): 1281–1295. DOI: 10.1177/1535370216641787.

Hu L, Huang W, Hjort E, et al. (2013) Increased Fanconi C expression contributes to the emergency granulopoiesis response. *Journal of Clinical Investigation* 123(9): 3952–3966. DOI: 10.1172/JCI69032.

Huang Y, Zhen R, Jiang M, et al. (2015) Development of all-in-one multicistronic Tet-On lentiviral vectors for inducible co-expression of two transgenes. *Biotechnology and Applied Biochemistry* 62(1): 48–54. DOI: 10.1002/bab.1239.

Huard CC, Tremblay CS, Magron A, et al. (2014) The Fanconi anemia pathway has a dual function in Dickkopf-1 transcriptional repression. *Proceedings of the National Academy of Sciences* 111(6): 2152–2157. DOI: 10.1073/pnas.1314226111.

Hudson JD, Shoaibi MA, Maestro R, et al. (1999) A Proinflammatory Cytokine Inhibits P53 Tumor Suppressor Activity. *Journal of Experimental Medicine* 190(10): 1375–1382. DOI: 10.1084/jem.190.10.1375.

Hussein SM, Batada NN, Vuoristo S, et al. (2011) Copy number variation and selection during reprogramming to pluripotency. *Nature* 471(7336): 58–62. DOI: 10.1038/nature09871.

Ito K, Zolfaghari R, Hao L, et al. (2014) Inflammation rapidly modulates the expression of ALDH1A1 (RALDH1) and vimentin in the liver and hepatic

macrophages of rats in vivo. *Nutrition & Metabolism* 11(1): 54. DOI: 10.1186/1743-7075-11-54.

Jacquinet A, Brown L, Sawkins J, et al. (2018) Expanding the FANCO/RAD51C associated phenotype: Cleft lip and palate and lobar holoprosencephaly, two rare findings in Fanconi anemia. *European Journal of Medical Genetics* 61(5): 257–261. DOI: 10.1016/j.ejmg.2017.12.011.

Jaiswal M, LaRusso NF, Burgart LJ, et al. (2000) Inflammatory cytokines induce DNA damage and inhibit DNA repair in cholangiocarcinoma cells by a nitric oxide-dependent mechanism. *Cancer research* 60(1): 184–90.

Jang M, Kim SS and Lee J (2013) Cancer cell metabolism: implications for therapeutic targets. *Experimental & Molecular Medicine* 45(10): e45–e45. DOI: 10.1038/emm.2013.85.

JI K, XING C, JIANG F, et al. (2013) Benzo[a]pyrene induces oxidative stress and endothelial progenitor cell dysfunction via the activation of the NF-κB pathway. *International Journal of Molecular Medicine* 31(4): 922–930. DOI: 10.3892/ijmm.2013.1288.

Johnson RD (2000) Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. *The EMBO Journal* 19(13): 3398–3407. DOI: 10.1093/emboj/19.13.3398.

Johnston R, D'Costa Z, Ray S, et al. (2016) The identification of a novel role for BRCA1 in regulating RNA polymerase I transcription. *Oncotarget* 7(42): 68097–68110. DOI: 10.18632/oncotarget.11770.

Joshi G, Arthur NBJ, Geetha TS, et al. (2023) Comprehensive laboratory diagnosis of Fanconi anaemia: comparison of cellular and molecular analysis. *Journal of Medical Genetics*: jmedgenet-2022-108714. DOI: 10.1136/jmg-2022-108714.

Jung M, Ramanagoudr-Bhojappa R, van Twest S, et al. (2020) Association of clinical severity with FANCB variant type in Fanconi anemia. *Blood* 135(18): 1588–1602. DOI: 10.1182/blood.2019003249.

K JCB, Kapoor BS, Mandal K, et al. (2020) Loss of Mitochondrial Localization of Human FANCG Causes Defective FANCJ Helicase. *Molecular and Cellular Biology* 40(23). DOI: 10.1128/MCB.00306-20.

Kafri T, van Praag H, Gage FH, et al. (2000) Lentiviral Vectors: Regulated Gene Expression. *Molecular Therapy* 1(6): 516–521. DOI: 10.1006/mthe.2000.0083.

Kamimae-Lanning AN, Goloviznina NA and Kurre P (2013) Fetal origins of hematopoietic failure in a murine model of Fanconi anemia. *Blood* 121(11): 2008–2012. DOI: 10.1182/blood-2012-06-439679.

Karczewski KJ, Francioli LC, Tiao G, et al. (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581(7809): 434–443. DOI: 10.1038/s41586-020-2308-7.

Karras GI, Yi S, Sahni N, et al. (2017) HSP90 Shapes the Consequences of Human Genetic Variation. *Cell* 168(5): 856-866.e12. DOI: 10.1016/j.cell.2017.01.023.

Kashiyama K, Nakazawa Y, Pilz DT, et al. (2013) Malfunction of Nuclease ERCC1-XPF Results in Diverse Clinical Manifestations and Causes Cockayne Syndrome, Xeroderma Pigmentosum, and Fanconi Anemia. *The American Journal of Human Genetics* 92(5): 807–819. DOI: 10.1016/j.ajhg.2013.04.007.

Kelm JM, Kramer BP, Gonzalez-Nicolini V, et al. (2004) Synergies of microtissue design, viral transduction and adjustable transgene expression for regenerative medicine. *Biotechnology and Applied Biochemistry* 39(1): 3. DOI: 10.1042/BA20030124.

Khoury R, Sauter S, Butsch Kovacic M, et al. (2018) Risk of Human Papillomavirus Infection in Cancer-Prone Individuals: What We Know. *Viruses* 10(1): 47. DOI: 10.3390/v10010047.

Kim JM, Parmar K, Huang M, et al. (2009) Inactivation of Murine Usp1 Results in Genomic Instability and a Fanconi Anemia Phenotype. *Developmental Cell* 16(2): 314–320. DOI: 10.1016/j.devcel.2009.01.001.

Kim K, Zhao R, Doi A, et al. (2011) Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nature Biotechnology* 29(12): 1117–1119. DOI: 10.1038/nbt.2052.

Kim M-R, Lee J, An YS, et al. (2015) TGFβ1 Protects Cells from γ-IR by Enhancing the Activity of the NHEJ Repair Pathway. *Molecular Cancer Research* 13(2): 319–329. DOI: 10.1158/1541-7786.MCR-14-0098-T.

Komori A, Yatsunami J, Suganuma M, et al. (1993) Tumor necrosis factor acts as a tumor promoter in BALB/3T3 cell transformation. *Cancer research* 53(9): 1982–5.

Kontou M, Adelfalk C, Hirsch-Kauffmann M, et al. (2003) Suboptimal Action of NF-κB in Fanconi Anemia Cells Results from Low Levels of Thioredoxin. *Biological Chemistry* 384(10–11). DOI: 10.1515/BC.2003.166.

Kopanos C, Tsiolkas V, Kouris A, et al. (2019) VarSome: the human genomic variant search engine. *Bioinformatics* 35(11): 1978–1980. DOI: 10.1093/bioinformatics/bty897.

Koponen JK, Kankkonen H, Kannasto J, et al. (2003) Doxycycline-regulated lentiviral vector system with a novel reverse transactivator rtTA2S-M2 shows a tight control of gene expression in vitro and in vivo. *Gene Therapy* 10(6): 459–466. DOI: 10.1038/sj.gt.3301889.

Korniluk A, Koper O, Kemona H, et al. (2017) From inflammation to cancer. *Irish Journal of Medical Science (1971)* 186(1): 57–62. DOI: 10.1007/s11845-016-1464-0.

Korthof ET, Svahn J, de Latour RP, et al. (2013) Immunological profile of Fanconi anemia: A multicentric retrospective analysis of 61 patients. *American Journal of Hematology* 88(6): 472–476. DOI: 10.1002/ajh.23435.

Kottemann MC and Smogorzewska A (2013) Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature* 493(7432): 356–363. DOI: 10.1038/nature11863.

Kotton DN, Mostoslavsky G and Murphy GJ (2010) Human dermal fibroblast isolation. *Center for regenerative medicine*: 10569.

Kruyt FA, Dijkmans LM, van den Berg TK, et al. (1996) Fanconi anemia genes act to suppress a cross-linker-inducible p53-independent apoptosis pathway in lymphoblastoid cell lines. *Blood* 87(3): 938–48.

Kruyt FAE, Hoshino T, Liu JM, et al. (1998) Abnormal Microsomal Detoxification Implicated in Fanconi Anemia Group C by Interaction of the FAC Protein With NADPH Cytochrome P450 Reductase. *Blood* 92(9): 3050–3056. DOI: 10.1182/blood.V92.9.3050.

Kuffel DG, Lindor NM, Litzow MR, et al. (1997) Mitomycin C chromosome stress test to identify hypersensitivity to bifunctional alkylating agents in patients with Fanconi anemia or aplastic anemia. *Mayo Clinic proceedings* 72(6). England: 579–580. DOI: 10.4065/72.6.579.

Kumaran M, Subramanian U and Devarajan B (2019) Performance assessment of variant calling pipelines using human whole exome sequencing and simulated data. *BMC Bioinformatics* 20(1): 342. DOI: 10.1186/s12859-019-2928-9.

Kumari U, Ya Jun W, Huat Bay B, et al. (2014) Evidence of mitochondrial dysfunction and impaired ROS detoxifying machinery in Fanconi Anemia cells. *Oncogene* 33(2): 165–172. DOI: 10.1038/onc.2012.583.

Kutler DI and Auerbach AD (2004) Fanconi anemia in Ashkenazi Jews. *Familial Cancer* 3(3–4): 241–248. DOI: 10.1007/s10689-004-9565-8.

Kutler David I., Singh B, Satagopan J, et al. (2003) A 20-year perspective on the International Fanconi Anemia Registry (IFAR). *Blood* 101(4): 1249–1256. DOI: 10.1182/blood-2002-07-2170.

Kutler David I, Singh B, Satagopan J, et al. (2003) A 20-year perspective on the International Fanconi Anemia Registry (IFAR). *Blood* 101(4). United States: 1249–1256. DOI: 10.1182/blood-2002-07-2170.

Kutler David I., Auerbach AD, Satagopan J, et al. (2003) High Incidence of Head and Neck Squamous Cell Carcinoma in Patients With Fanconi Anemia. *Archives of Otolaryngology–Head & Neck Surgery* 129(1): 106. DOI: 10.1001/archotol.129.1.106.

Kutler DI, Patel KR, Auerbach AD, et al. (2016) Natural history and management of Fanconi anemia patients with head and neck cancer: A 10-year follow-up. *The Laryngoscope* 126(4): 870–879. DOI: 10.1002/lary.25726.

Laboratories KD, Genetics M, Health O, et al. (2018) Standards and guidelines for the interpretation of sequence variants. *Acta Ophthalmologica* 96(S261): 134–134. DOI: 10.1111/aos.13972_502.

Landrum MJ and Kattman BL (2018) ClinVar at five years: Delivering on the promise. *Human Mutation* 39(11): 1623–1630. DOI: 10.1002/humu.23641.

Lane AN, Higashi RM and Fan TW-M (2020) Metabolic reprogramming in tumors: Contributions of the tumor microenvironment. *Genes & Diseases* 7(2): 185–198. DOI: 10.1016/j.gendis.2019.10.007.

Langevin F, Crossan GP, Rosado I V., et al. (2011) Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature* 475(7354): 53–58. DOI: 10.1038/nature10192.

Latorre E, Mendoza C, Layunta E, et al. (2014) TLR2, TLR3, and TLR4 activation specifically alters the oxidative status of intestinal epithelial cells. *Cell Stress and Chaperones* 19(2): 289–293. DOI: 10.1007/s12192-013-0461-8.

Ledran MH, Krassowska A, Armstrong L, et al. (2008) Efficient Hematopoietic Differentiation of Human Embryonic Stem Cells on Stromal Cells Derived from Hematopoietic Niches. *Cell Stem Cell* 3(1): 85–98. DOI: 10.1016/j.stem.2008.06.001.

Lee H-C (2014) Somatic alterations in mitochondrial DNA and mitochondrial dysfunction in gastric cancer progression. *World Journal of Gastroenterology* 20(14): 3950. DOI: 10.3748/wjg.v20.i14.3950.

Lee Y-S, Gregory MT and Yang W (2014) Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. *Proceedings of the National Academy of Sciences* 111(8): 2954–2959. DOI: 10.1073/pnas.1324001111.

Leto TL, Morand S, Hurt D, et al. (2009) Targeting and Regulation of Reactive Oxygen Species Generation by Nox Family NADPH Oxidases. *Antioxidants & Redox Signaling* 11(10): 2607–2619. DOI: 10.1089/ars.2009.2637.

Li J and Pang Q (2014) Oxidative Stress-Associated Protein Tyrosine Kinases and Phosphatases in Fanconi Anemia. *Antioxidants & Redox Signaling* 20(14): 2290–2301. DOI: 10.1089/ars.2013.5715.

Li J, Sejas DP, Zhang X, et al. (2007) TNF-α induces leukemic clonal evolution ex vivo in Fanconi anemia group C murine stem cells. *Journal of Clinical Investigation* 117(11): 3283–3295. DOI: 10.1172/JCI31772.

Li J, Du W, Maynard S, et al. (2010) Oxidative stress–specific interaction between FANCD2 and FOXO3a. *Blood* 115(8): 1545–1548. DOI: 10.1182/blood-2009-07-234385.

Li Y and Youssoufian H (1997) MxA overexpression reveals a common genetic link in four Fanconi anemia complementation groups. *Journal of Clinical Investigation* 100(11): 2873–2880. DOI: 10.1172/JCI119836.

Liang Q, Dexheimer TS, Zhang P, et al. (2014) A selective USP1–UAF1 inhibitor links deubiquitination to DNA damage responses. *Nature Chemical Biology* 10(4): 298–304. DOI: 10.1038/nchembio.1455.

Lier A, Walter D, Geiselhart A, et al. (2014) Stress-induced exit from dormancy alters redox signaling in HSCs, resulting in de novo DNA damage and bone marrow failure in the absence of a functional fanconi anemia signaling pathway. *Experimental Hematology* 42(8): S45. DOI: 10.1016/j.exphem.2014.07.168.

Lin F, Karwan M, Saleh B, et al. (2014) IFN-γ causes aplastic anemia by altering hematopoietic stem/progenitor cell composition and disrupting lineage differentiation. *Blood* 124(25): 3699–3708. DOI: 10.1182/blood-2014-01-549527.

Lin Y-R, Hung H-C, Leu J-H, et al. (2011) The Role of Aldehyde Dehydrogenase and Hsp70 in Suppression of White Spot Syndrome Virus Replication at High

Temperature. *Journal of Virology* 85(7): 3517–3525. DOI: 10.1128/JVI.01973-10.

Liu GH, Suzuki K, Li M, et al. (2014) Modelling Fanconi anemia pathogenesis and therapeutics using integration-free patient-derived iPSCs. *Nature Communications* 5. Nature Publishing Group. DOI: 10.1038/ncomms5330.

Liu G-H, Suzuki K, Li M, et al. (2014) Modelling Fanconi anemia pathogenesis and therapeutics using integration-free patient-derived iPSCs. *Nature Communications* 5(1): 4330. DOI: 10.1038/ncomms5330.

Liu L, Zhou W, Cheng C-T, et al. (2014) TGFβ Induces "BRCAness" and Sensitivity to PARP Inhibition in Breast Cancer by Regulating DNA-Repair Genes. *Molecular Cancer Research* 12(11): 1597–1609. DOI: 10.1158/1541-7786.MCR-14-0201.

Lo Ten Foe JR, Kwee ML, Rooimans MA, et al. (1997) Somatic mosaicism in Fanconi anemia: Molecular basis and clinical significance. *European Journal of Human Genetics* 5(3): 137–148. DOI: 10.1159/000484749.

Loew R, Vigna E and Lindemann D (2006) Retroviral vectors containing Tet-controlled bidirectional transcription units for simultaneous regulation of two gene activities. *Journal of Molecular and Genetic Medicine* 02(01). DOI: 10.4172/1747-0862.1000021.

Loew R, Heinz N, Hampf M, et al. (2010) Improved Tet-responsive promoters with minimized background expression. *BMC Biotechnology* 10(1): 81. DOI: 10.1186/1472-6750-10-81.

Lossaint G, Larroque M, Ribeyre C, et al. (2013) FANCD2 Binds MCM Proteins and Controls Replisome Function upon Activation of S Phase Checkpoint Signaling. *Molecular Cell* 51(5): 678–690. DOI: 10.1016/j.molcel.2013.07.023.

Lu R, O'Rourke JJ, Sobinoff AP, et al. (2019) The FANCM-BLM-TOP3A-RMI complex suppresses alternative lengthening of telomeres (ALT). *Nature Communications* 10(1): 2252. DOI: 10.1038/s41467-019-10180-6.

Lugrin J, Rosenblatt-Velin N, Parapanov R, et al. (2014) The role of oxidative stress during inflammatory processes. *Biological Chemistry* 395(2): 203–230. DOI: 10.1515/hsz-2013-0241.

Magdalena N, Pilonetto DV, Bitencourt MA, et al. (2005) Frequency of Fanconi anemia in Brazil and efficacy of screening for the FANCA 3788-3790del mutation. *Brazilian Journal of Medical and Biological Research* 38(5): 669–673. DOI: 10.1590/S0100-879X2005000500003.

Maizels N (2008) Genomic Stability: FANCJ-Dependent G4 DNA Repair. *Current Biology* 18(14): R613–R614. DOI: 10.1016/j.cub.2008.06.011.

Mamrak Nicholas E., Shimamura A and Howlett NG (2017) Recent discoveries in the molecular pathogenesis of the inherited bone marrow failure syndrome Fanconi anemia. *Blood Reviews*. Churchill Livingstone. DOI: 10.1016/j.blre.2016.10.002.

Mamrak Nicholas E, Shimamura A and Howlett NG (2017) Recent discoveries in the molecular pathogenesis of the inherited bone marrow failure syndrome Fanconi anemia. *Blood reviews* 31(3): 93—99. DOI: 10.1016/j.blre.2016.10.002.

Marion W, Boettcher S, Ruiz-Torres S, et al. (2020) An induced pluripotent stem cell model of Fanconi anemia reveals mechanisms of p53-driven progenitor cell differentiation. *Blood Advances* 4(19). American Society of Hematology: 4679–4692. DOI: 10.1182/bloodadvances.2020001593.

Markusic D and Seppen J (2010) Doxycycline Regulated Lentiviral Vectors., pp. 69–76. DOI: 10.1007/978-1-60761-533-0_4.

Marnett LJ (2000) Oxyradicals and DNA damage. *Carcinogenesis* 21(3): 361–370. DOI: 10.1093/carcin/21.3.361.

Matsunawa M, Yamamoto R, Sanada M, et al. (2014) Haploinsufficiency of Sf3b1 leads to compromised stem cell function but not to myelodysplasia. *Leukemia* 28(9): 1844–1850. DOI: 10.1038/leu.2014.73.

Mayuranathan T, Rayabaram J, Edison ES, et al. (2012) A novel deletion of β-globin promoter causing high HbA2 in an Indian population. *Haematologica* 97(9): 1445–1447. DOI: 10.3324/haematol.2012.062299.

McCauley J, Masand N, McGowan R, et al. (2011) X-linked VACTERL with hydrocephalus syndrome: Further delineation of the phenotype caused by FANCB mutations. *American Journal of Medical Genetics Part A* 155(10): 2370–2380. DOI: 10.1002/ajmg.a.33913.

McReynolds LJ, Rafati M, Wang Y, et al. (2022) Genetic testing in severe aplastic anemia is required for optimal hematopoietic cell transplant outcomes. *Blood* 140(8): 909–921. DOI: 10.1182/blood.2022016508.

Meerbrey KL, Hu G, Kessler JD, et al. (2011) The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proceedings of the National Academy of Sciences* 108(9): 3665–3670. DOI: 10.1073/pnas.1019736108.

Meetei AR, Medhurst AL, Ling C, et al. (2005) A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nature Genetics* 37(9): 958–963. DOI: 10.1038/ng1626.

Mehta PA and Ebens C (1993) *Fanconi Anemia*. University of Washington, Seattle, Seattle (WA). Available at: http://europepmc.org/abstract/MED/20301575.

Michl J, Zimmer J, Buffa FM, et al. (2016) FANCD2 limits replication stress and genome instability in cells lacking BRCA2. *Nature Structural & Molecular Biology* 23(8): 755–757. DOI: 10.1038/nsmb.3252.

Milletti G, Strocchio L, Pagliara D, et al. (2020) Canonical and Noncanonical Roles of Fanconi Anemia Proteins: Implications in Cancer Predisposition. *Cancers* 12(9): 2684. DOI: 10.3390/cancers12092684.

Milletti Giacomo, Luisa Strocchio, Pagliara Daria, et al. (2020) Canonical and Noncanonical Roles of Fanconi Anemia Proteins: Implications in Cancer Predisposition. *Cancers* 12(9): 2684-NA. DOI: 10.3390/cancers12092684.

Miyoshi H, Blömer U, Takahashi M, et al. (1998) Development of a Self-Inactivating Lentivirus Vector. *Journal of Virology* 72(10): 8150–8157. DOI: 10.1128/JVI.72.10.8150-8157.1998.

Mukhopadhyay SS, Leung KS, Hicks MJ, et al. (2006) Defective mitochondrial peroxiredoxin-3 results in sensitivity to oxidative stress in Fanconi anemia. *Journal of Cell Biology* 175(2): 225–235. DOI: 10.1083/jcb.200607061.

Müller LUW, Milsom MD, Harris CE, et al. (2012) Overcoming reprogramming resistance of Fanconi anemia cells. *Blood* 119(23). American Society of Hematology: 5449–5457. DOI: 10.1182/blood-2012-02-408674.

Nakanishi K, Taniguchi T, Ranganathan V, et al. (2002) Interaction of FANCD2 and NBS1 in the DNA damage response. *Nature Cell Biology* 4(12): 913–920. DOI: 10.1038/ncb879.

Nakao S, Yamaguchi M, Shiobara S, et al. (1992) Interferon-gamma gene expression in unstimulated bone marrow mononuclear cells predicts a good response to cyclosporine therapy in aplastic anemia. *Blood* 79(10): 2532–2535. DOI: 10.1182/blood.V79.10.2532.2532.

Naldini L, Blömer U, Gallay P, et al. (1996) In Vivo Gene Delivery and Stable Transduction of Nondividing Cells by a Lentiviral Vector. *Science* 272(5259): 263–267. DOI: 10.1126/science.272.5259.263.

Nalepa G and Clapp DW (2014) Fanconi anemia and the cell cycle: new perspectives on aneuploidy. *F1000Prime Reports* 6. DOI: 10.12703/P6-23.

Nalepa G and Clapp DW (2018) Fanconi anaemia and cancer: An intricate relationship. *Nature Reviews Cancer*. Nature Research. DOI: 10.1038/nrc.2017.116.

Narayan AD, Chase JL, Lewis RL, et al. (2006) Human embryonic stem cell–derived hematopoietic cells are capable of engrafting primary as well as secondary fetal sheep recipients. *Blood* 107(5): 2180–2183. DOI: 10.1182/blood-2005-05-1922.

Navarro S, Meza NW, Quintana-Bustamante O, et al. (2006) Hematopoietic Dysfunction in a Mouse Model for Fanconi Anemia Group D1. *Molecular Therapy* 14(4): 525–535. DOI: 10.1016/j.ymthe.2006.05.018.

Nepal M, Ma C, Xie G, et al. (2018) Fanconi Anemia complementation group C protein in metabolic disorders. *Aging* 10(6): 1506–1522. DOI: 10.18632/aging.101487.

Neveling K, Endt D, Hoehn H, et al. (2009) Genotype-phenotype correlations in Fanconi anemia. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*. DOI: 10.1016/j.mrfmmm.2009.05.006.

New H V., Cale CM, Tischkowitz M, et al. (2005) Nijmegen breakage syndrome diagnosed as Fanconi anaemia. *Pediatric Blood & Cancer* 44(5): 494–499. DOI: 10.1002/pbc.20271.

Nie D, Zhang J, Wang F, et al. (2020) Comprehensive analysis on phenotype and genetic basis of Chinese Fanconi anemia patients: Dismal outcomes call for nationwide studies. *BMC Medical Genetics* 21(1). BioMed Central Ltd. DOI: 10.1186/s12881-020-01057-3.

Niedernhofer LJ, Garinis GA, Raams A, et al. (2006) A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. *Nature* 444(7122): 1038–1043. DOI: 10.1038/nature05456.

Niraj J, Färkkilä A and D'Andrea AD (2019) The fanconi anemia pathway in cancer. *Annual Review of Cancer Biology*. Annual Reviews Inc. DOI: 10.1146/annurev-cancerbio-030617-050422.

Nishizawa M, Chonabayashi K, Nomura M, et al. (2016) Epigenetic Variation between Human Induced Pluripotent Stem Cell Lines Is an Indicator of Differentiation Capacity. *Cell Stem Cell* 19(3): 341–354. DOI: 10.1016/j.stem.2016.06.019.

Noll M, Battaile KP, Bateman R, et al. (2002) Fanconi anemia group A and C double-mutant mice. *Experimental Hematology* 30(7): 679–688. DOI: 10.1016/S0301-472X(02)00838-X.

Oberbeck N, Langevin F, King G, et al. (2014) Maternal Aldehyde Elimination during Pregnancy Preserves the Fetal Genome. *Molecular Cell* 55(6): 807–817. DOI: 10.1016/j.molcel.2014.07.010.

Oestergaard VH, Langevin F, Kuiken HJ, et al. (2007) Deubiquitination of FANCD2 Is Required for DNA Crosslink Repair. *Molecular Cell* 28(5): 798–809. DOI: 10.1016/j.molcel.2007.09.020.

Offit K (2006) BRCA Mutation Frequency and Penetrance: New Data, Old Debate. *JNCI: Journal of the National Cancer Institute* 98(23): 1675–1677. DOI: 10.1093/jnci/djj500.

Ogueta SB, Yao F and Marasco WA (2001) Design and in vitro characterization of a single regulatory module for efficient control of gene expression in both plasmid DNA and a self-inactivating lentiviral vector. *Molecular medicine (Cambridge, Mass.)* 7(8): 569–79.

Okita K, Yamakawa T, Matsumura Y, et al. (2013) An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells. *Stem Cells* 31(3): 458–466. DOI: 10.1002/stem.1293.

Oostra AB, Nieuwint AWM, Joenje H, et al. (2012) Diagnosis of fanconi anemia: chromosomal breakage analysis. *Anemia* 2012: 238731. DOI: 10.1155/2012/238731.

Orvedahl A, Sumpter R, Xiao G, et al. (2011) Image-based genome-wide siRNA screen identifies selective autophagy factors. *Nature* 480(7375): 113–117. DOI: 10.1038/nature10546.

Osborn MJ, Lonetree CL, Webber BR, et al. (2016) CRISPR/Cas9 targeted gene editing and cellular engineering in fanconi anemia. *Stem Cells and Development* 25(20). Mary Ann Liebert Inc.: 1591–1603. DOI: 10.1089/scd.2016.0149.

Pagano G, Degan P, d'Ischia M, et al. (2005) Oxidative stress as a multiple effector in Fanconi anaemia clinical phenotype. *European Journal of Haematology* 75(2): 93–100. DOI: 10.1111/j.1600-0609.2005.00507.x.

Pagano G, Shyamsunder P, Verma RS, et al. (2014) Damaged mitochondria in Fanconi anemia - an isolated event or a general phenomenon? *Oncoscience* 1(4): 287–295. DOI: 10.18632/oncoscience.29.

Pagano G, Tiano L, Pallardó F V., et al. (2021) Re-definition and supporting evidence toward Fanconi Anemia as a mitochondrial disease: Prospects for new design in

clinical management. *Redox Biology* 40: 101860. DOI: 10.1016/j.redox.2021.101860.

Pal D, Pertot A, Shirole NH, et al. (2017) TGF-β reduces DNA ds-break repair mechanisms to heighten genetic diversity and adaptability of CD44+/CD24− cancer cells. *eLife* 6. DOI: 10.7554/eLife.21615.

Pallardó F V., Lloret A, Lebel M, et al. (2010) Mitochondrial dysfunction in some oxidative stress-related genetic diseases: Ataxia-Telangiectasia, Down Syndrome, Fanconi Anaemia and Werner Syndrome. *Biogerontology* 11(4): 401–419. DOI: 10.1007/s10522-010-9269-4.

Pan X, Drosopoulos WC, Sethi L, et al. (2017) FANCM, BRCA1, and BLM cooperatively resolve the replication stress at the ALT telomeres. *Proceedings of the National Academy of Sciences* 114(29). DOI: 10.1073/pnas.1708065114.

Pang Q (2001) FANCC interacts with Hsp70 to protect hematopoietic cells from IFN-gamma/TNF-alpha-mediated cytotoxicity. *The EMBO Journal* 20(16): 4478–4489. DOI: 10.1093/emboj/20.16.4478.

Pang Q and Andreassen PR (2009) Fanconi anemia proteins and endogenous stresses. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 668(1–2): 42–53. DOI: 10.1016/j.mrfmmm.2009.03.013.

Pang Q, Fagerlie S, Christianson TA, et al. (2000) The Fanconi Anemia Protein FANCC Binds to and Facilitates the Activation of STAT1 by Gamma Interferon and Hematopoietic Growth Factors. *Molecular and Cellular Biology* 20(13): 4724–4735. DOI: 10.1128/MCB.20.13.4724-4735.2000.

Pang Q, Keeble W, Diaz J, et al. (2001) Role of double-stranded RNA-dependent protein kinase in mediating hypersensitivity of Fanconi anemia complementation

group C cells to interferon γ, tumor necrosis factor-α, and double-stranded RNA. *Blood* 97(6): 1644–1652. DOI: 10.1182/blood.V97.6.1644.

Pang Q, Christianson TA, Keeble W, et al. (2001) The Fanconi anemia complementation group C gene product: structural evidence of multifunctionality. *Blood* 98(5): 1392–1401. DOI: 10.1182/blood.V98.5.1392.

Pang Q, Christianson TA, Keeble W, et al. (2002) The Anti-apoptotic Function of Hsp70 in the Interferon-inducible Double-stranded RNA-dependent Protein Kinase-mediated Death Signaling Pathway Requires the Fanconi Anemia Protein, FANCC. *Journal of Biological Chemistry* 277(51): 49638–49643. DOI: 10.1074/jbc.M209386200.

Park J, Kim M, Jang W, et al. (2015) Founder Haplotype Analysis of Fanconi Anemia in the Korean Population Finds Common Ancestral Haplotypes for a FANCG Variant. *Annals of Human Genetics* 79(3): 153–161. DOI: 10.1111/ahg.12097.

Park JS, Sharma LK, Li H, et al. (2009) A heteroplasmic, not homoplasmic, mitochondrial DNA mutation promotes tumorigenesis via alteration in reactive oxygen species generation and apoptosis. *Human Molecular Genetics* 18(9): 1578–1589. DOI: 10.1093/hmg/ddp069.

Park J-Y, Virts EL, Jankowska A, et al. (2016) Complementation of hypersensitivity to DNA interstrand crosslinking agents demonstrates that XRCC2 is a Fanconi anaemia gene. *Journal of Medical Genetics* 53(10): 672–680. DOI: 10.1136/jmedgenet-2016-103847.

Park S-J, Ciccone SLM, Beck BD, et al. (2004) Oxidative Stress/Damage Induces Multimerization and Interaction of Fanconi Anemia Proteins. *Journal of Biological Chemistry* 279(29): 30053–30059. DOI: 10.1074/jbc.M403527200.

Park SK, Hwang BJ and Kee Y (2019) Promoter cross-talk affects the inducible expression of intronic shRNAs from the tetracycline response element. *Genes & Genomics* 41(4): 483–490. DOI: 10.1007/s13258-019-00784-z.

Parmar K, D'Andrea A and Niedernhofer LJ (2009) Mouse models of Fanconi anemia. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 668(1–2): 133–140. DOI: 10.1016/j.mrfmmm.2009.03.015.

Parmar K, Kim J, Sykes SM, et al. (2010) Hematopoietic Stem Cell Defects in Mice with Deficiency of Fancd2 or Usp1. *Stem Cells* 28(7): 1186–1195. DOI: 10.1002/stem.437.

Pickrell AM and Youle RJ (2015) The Roles of PINK1, Parkin, and Mitochondrial Fidelity in Parkinson's Disease. *Neuron* 85(2): 257–273. DOI: 10.1016/j.neuron.2014.12.007.

Pilonetto D V., Pereira NF, Bonfim CMS, et al. (2017) A strategy for molecular diagnostics of Fanconi anemia in Brazilian patients. *Molecular Genetics and Genomic Medicine* 5(4). Wiley-Blackwell: 360–372. DOI: 10.1002/mgg3.293.

Plagnol V, Curtis J, Epstein M, et al. (2012) A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics (Oxford, England)* 28(21): 2747–2754. DOI: 10.1093/bioinformatics/bts526.

Pontel LB, Rosado IV, Burgos-Barragan G, et al. (2015) Endogenous Formaldehyde Is a Hematopoietic Stem Cell Genotoxin and Metabolic Carcinogen. *Molecular Cell* 60(1): 177–188. DOI: 10.1016/j.molcel.2015.08.020.

Poot M, Hoehn H, Kubbies M, et al. (1994) Chapter 21 Cell-Cycle Analysis Using Continuous Bromodeoxyuridine Labeling and Hoechst 33358—Ethidium

Bromide Bivariate Flow Cytometry., pp. 327–340. DOI: 10.1016/S0091-679X(08)61726-4.

Prakash R, Zhang Y, Feng W, et al. (2015) Homologous Recombination and Human Health: The Roles of BRCA1, BRCA2, and Associated Proteins. *Cold Spring Harbor Perspectives in Biology* 7(4): a016600. DOI: 10.1101/cshperspect.a016600.

Pulliam-Leath AC, Ciccone SL, Nalepa G, et al. (2010) Genetic disruption of both Fancc and Fancg in mice recapitulates the hematopoietic manifestations of Fanconi anemia. *Blood* 116(16): 2915–2920. DOI: 10.1182/blood-2009-08-240747.

Raman D, Baugher PJ, Thu YM, et al. (2007) Role of chemokines in tumor growth. *Cancer Letters* 256(2): 137–165. DOI: 10.1016/j.canlet.2007.05.013.

Rantakari P, Nikkilä J, Jokela H, et al. (2010) Inactivation of Palb2 gene leads to mesoderm differentiation defect and early embryonic lethality in mice. *Human Molecular Genetics* 19(15): 3021–3029. DOI: 10.1093/hmg/ddq207.

Rathbun RK, Christianson TA, Faulkner GR, et al. (2000) Interferon-gamma-induced apoptotic responses of Fanconi anemia group C hematopoietic progenitor cells involve caspase 8-dependent activation of caspase 3 family members. *Blood* 96(13): 4204–11.

Ravera S, Vaccaro D, Cuccarolo P, et al. (2013) Mitochondrial respiratory chain Complex I defects in Fanconi anemia complementation group A. *Biochimie* 95(10): 1828–1837. DOI: 10.1016/j.biochi.2013.06.006.

Ray PD, Huang B-W and Tsuji Y (2012) Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. *Cellular Signalling* 24(5): 981–990. DOI: 10.1016/j.cellsig.2012.01.008.

Raya Á, Rodríguez-Piz I, Guenechea G, et al. (2009) Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature*. DOI: 10.1038/nature08129.

Rehm HL, Berg JS, Brooks LD, et al. (2015) ClinGen — The Clinical Genome Resource. *New England Journal of Medicine* 372(23): 2235–2242. DOI: 10.1056/NEJMsr1406261.

Reid S, Schindler D, Hanenberg H, et al. (2007) Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nature Genetics* 39(2): 162–164. DOI: 10.1038/ng1947.

Reiser J, Lai Z, Zhang X-Y, et al. (2000) Development of Multigene and Regulated Lentivirus Vectors. *Journal of Virology* 74(22): 10589–10599. DOI: 10.1128/JVI.74.22.10589-10599.2000.

Renaudin X and Rosselli F (2020) The FANC/BRCA Pathway Releases Replication Blockades by Eliminating DNA Interstrand Cross-Links. *Genes* 11(5): 585. DOI: 10.3390/genes11050585.

Renaudin X and Rosselli F (2021) Tipping the Scale: MYC Gains Weight in Fanconi Anemia Bone Marrow Failure Progression. *Cell Stem Cell* 28(1): 8–9. DOI: 10.1016/j.stem.2020.12.013.

Rickman K and Smogorzewska A (2019) Advances in understanding DNA processing and protection at stalled replication forks. *Journal of Cell Biology* 218(4): 1096–1107. DOI: 10.1083/jcb.201809012.

Rickman KA, Lach FP, Abhyankar A, et al. (2015) Deficiency of UBE2T, the E2 Ubiquitin Ligase Necessary for FANCD2 and FANCI Ubiquitination, Causes FA-T Subtype of Fanconi Anemia. *Cell Reports* 12(1): 35–41. DOI: 10.1016/j.celrep.2015.06.014.

Ridpath JR, Nakamura A, Tano K, et al. (2007) Cells Deficient in the FANC/BRCA Pathway Are Hypersensitive to Plasma Levels of Formaldehyde. *Cancer Research* 67(23): 11117–11122. DOI: 10.1158/0008-5472.CAN-07-3028.

Río P, Navarro S, Wang W, et al. (2019) Successful engraftment of gene-corrected hematopoietic stem cells in non-conditioned patients with Fanconi anemia. *Nature Medicine* 25(9). Springer US: 1396–1401. DOI: 10.1038/s41591-019-0550-z.

Risitano AM, Marotta S, Calzone R, et al. (2016) Twenty years of the Italian Fanconi Anemia Registry: where we stand and what remains to be learned. *Haematologica* 101(3): 319–327. DOI: 10.3324/haematol.2015.133520.

Risueño RM, Sachlos E, Lee J-H, et al. (2012) Inability of Human Induced Pluripotent Stem Cell-Hematopoietic Derivatives to Downregulate MicroRNAs In Vivo Reveals a Block in Xenograft Hematopoietic Regeneration. *Stem Cells* 30(2): 131–139. DOI: 10.1002/stem.1684.

Roberts AB and Wakefield LM (2003) The two faces of transforming growth factor β in carcinogenesis. *Proceedings of the National Academy of Sciences* 100(15): 8621–8623. DOI: 10.1073/pnas.1633291100.

Rochowski A, Olson SB, Alonzo TA, et al. (2012) Patients with Fanconi anemia and AML have different cytogenetic clones than de novo cases of AML. *Pediatric Blood & Cancer* 59(5): 922–924. DOI: 10.1002/pbc.24168.

Rodríguez A, Zhang K, Färkkilä A, et al. (2021) MYC Promotes Bone Marrow Stem Cell Dysfunction in Fanconi Anemia. *Cell Stem Cell* 28(1): 33-47.e8. DOI: 10.1016/j.stem.2020.09.004.

Rodriguez DEA, Lima CSP, Lourenço GJ, et al. (2005) Molecular analysis of the most prevalent mutations of the FANCA and FANCC genes in Brazilian patients with

Fanconi anaemia. *Genetics and Molecular Biology* 28(2): 205–209. DOI: 10.1590/S1415-47572005000200004.

Romick-Rosendale LE, Lui VWY, Grandis JR, et al. (2013) The Fanconi anemia pathway: Repairing the link between DNA damage and squamous cell carcinoma. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 743–744: 78–88. DOI: 10.1016/j.mrfmmm.2013.01.001.

Rosado I V., Niedzwiedz W, Alpi AF, et al. (2009) The Walker B motif in avian FANCM is required to limit sister chromatid exchanges but is dispensable for DNA crosslink repair. *Nucleic Acids Research* 37(13): 4360–4370. DOI: 10.1093/nar/gkp365.

Rosado I V, Langevin F, Crossan GP, et al. (2011) Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway. *Nature Structural & Molecular Biology* 18(12): 1432–1434. DOI: 10.1038/nsmb.2173.

Rosenberg PS, Greene MH and Alter BP (2003) Cancer incidence in persons with Fanconi anemia. *Blood* 101(3): 822–826. DOI: 10.1182/blood-2002-05-1498.

Rosenberg PS, Alter BP and Ebell W (2008) Cancer risks in Fanconi anemia: findings from the German Fanconi Anemia Registry. *Haematologica* 93(4): 511–517. DOI: 10.3324/haematol.12234.

Rosenberg PS, Tamary H and Alter BP (2011) How high are carrier frequencies of rare recessive syndromes? Contemporary estimates for Fanconi Anemia in the United States and Israel. *American Journal of Medical Genetics Part A* 155(8): 1877–1883. DOI: 10.1002/ajmg.a.34087.

Rosselli F, Sanceau J, Wietzerbin J, et al. (1992) Abnormal lymphokine production: a novel feature of the genetic disease Fanconi anemia. *Human Genetics* 89(1): 42–48. DOI: 10.1007/BF00207040.

Ruiz-Torres S, Brusadelli MG, Witte DP, et al. (2021) Inherited DNA Repair Defects Disrupt the Structure and Function of Human Skin. *Cell Stem Cell* 28(3): 424-435.e6. DOI: 10.1016/j.stem.2020.10.012.

Ruppitsch W, Meißlitzer C, Weirich-Schwaiger H, et al. (1997) The role of oxygen metabolism for the pathological phenotype of Fanconi anemia. *Human Genetics* 99(6): 710–719. DOI: 10.1007/s004390050437.

Ruppitsch W, Meißlitzer C, Hirsch-Kauffmann M, et al. (1998) Overexpression of thioredoxin in Fanconi anemia fibroblasts prevents the cytotoxic and DNA damaging effect of mitomycin C and diepoxybutane. *FEBS Letters* 422(1): 99–102. DOI: 10.1016/S0014-5793(97)01608-6.

Saadatzadeh MR, Bijangi-Vishehsaraei K, Hong P, et al. (2004) Oxidant Hypersensitivity of Fanconi Anemia Type C-deficient Cells Is Dependent on a Redox-regulated Apoptotic Pathway. *Journal of Biological Chemistry* 279(16): 16805–16812. DOI: 10.1074/jbc.M313721200.

Salzano S, Checconi P, Hanschmann E-M, et al. (2014) Linkage of inflammation and oxidative stress via release of glutathionylated peroxiredoxin-2, which acts as a danger signal. *Proceedings of the National Academy of Sciences* 111(33): 12157–12162. DOI: 10.1073/pnas.1401712111.

Sauter SL, Wells SI, Zhang X, et al. (2015) Oral Human Papillomavirus Is Common in Individuals with Fanconi Anemia. *Cancer Epidemiology, Biomarkers & Prevention* 24(5): 864–872. DOI: 10.1158/1055-9965.EPI-15-0097-T.

Savage SA and Walsh MF (2018) Myelodysplastic Syndrome, Acute Myeloid Leukemia, and Cancer Surveillance in Fanconi Anemia. *Hematology/Oncology Clinics of North America* 32(4): 657–668. DOI: 10.1016/j.hoc.2018.04.002.

Sawyer SL, Tian L, Kähkönen M, et al. (2015) Biallelic Mutations in BRCA1 Cause a New Fanconi Anemia Subtype. *Cancer Discovery* 5(2): 135–142. DOI: 10.1158/2159-8290.CD-14-1156.

Schindler D, Friedl R, Gavvovidis I, et al. (2007) Applications of cell cycle testing in fanconi anemia. *Monographs in Human Genetics* 15(Mmc): 110–130. DOI: 10.1159/000102552.

Schwaiger H, Hirsch-Kauffmann M and Schweiger M (1982) UV-Repair is impaired in fibroblasts from patients with Fanconi's anemia. *Molecular and General Genetics MGG* 185(3): 454–456. DOI: 10.1007/BF00334139.

Sejas DP, Rani R, Qiu Y, et al. (2007) Inflammatory Reactive Oxygen Species-Mediated Hemopoietic Suppression in Fancc-Deficient Mice. *The Journal of Immunology* 178(8): 5277–5287. DOI: 10.4049/jimmunol.178.8.5277.

Seo A, Steinberg-Shemer O, Unal S, et al. (2018) Mechanism for survival of homozygous nonsense mutations in the tumor suppressor gene BRCA1. *Proceedings of the National Academy of Sciences* 115(20): 5241–5246. DOI: 10.1073/pnas.1801796115.

Serra A, Eirich K, Winkler AK, et al. (2012) Shared Copy Number Variation in Simultaneous Nephroblastoma and Neuroblastoma due to Fanconi Anemia. *Molecular Syndromology* 3(3): 120–130. DOI: 10.1159/000341935.

Seyschab H, Friedl R, Sun Y, et al. (1995) Comparative evaluation of diepoxybutane sensitivity and cell cycle blockage in the diagnosis of Fanconi anemia. *Blood* 85(8): 2233–7.

Shamseldin HE, Elfaki M and Alkuraya FS (2012) Exome sequencing reveals a novel Fanconi group defined by XRCC2 mutation. *Journal of Medical Genetics* 49(3): 184–186. DOI: 10.1136/jmedgenet-2011-100585.

Shen X and Li L (2010) Mutagenic repair of DNA interstrand crosslinks. *Environmental and Molecular Mutagenesis*: NA-NA. DOI: 10.1002/em.20558.

Shimamura A and Alter BP (2010) Pathophysiology and management of inherited bone marrow failure syndromes. *Blood Reviews* 24(3): 101–122. DOI: 10.1016/j.blre.2010.03.002.

Shimamura A, de Oca RM, Svenson John L., et al. (2002) A novel diagnostic screen for defects in the Fanconi anemia pathway. *Blood* 100(13): 4649–4654. DOI: 10.1182/blood-2002-05-1399.

Shimamura A, de Oca RM, Svenson John L, et al. (2002) A novel diagnostic screen for defects in the Fanconi anemia pathway. *Blood* 100(13): 4649–4654. DOI: 10.1182/blood-2002-05-1399.

Shyamsunder P, Esner M, Barvalia M, et al. (2016) Impaired mitophagy in Fanconi anemia is dependent on mitochondrial fission. *Oncotarget* 7(36): 58065–58074. DOI: 10.18632/oncotarget.11161.

Sijbers AM, de Laat WL, Ariza RR, et al. (1996) Xeroderma Pigmentosum Group F Caused by a Defect in a Structure-Specific DNA Repair Endonuclease. *Cell* 86(5): 811–822. DOI: 10.1016/S0092-8674(00)80155-5.

Singh TR, Bakker ST, Agarwal S, et al. (2009) Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M. *Blood* 114(1): 174–180. DOI: 10.1182/blood-2009-02-207811.

Snezhkina A V., Kudryavtseva A V., Kardymon OL, et al. (2019) ROS Generation and Antioxidant Defense Systems in Normal and Malignant Cells. *Oxidative Medicine and Cellular Longevity* 2019: 1–17. DOI: 10.1155/2019/6175804.

Solanki A, Rajendran A, Mohan S, et al. (2020) Mitochondrial DNA variations and mitochondrial dysfunction in Fanconi anemia. *PLOS ONE* 15(1): e0227603. DOI: 10.1371/journal.pone.0227603.

Solomon BD, Bear KA, Kimonis V, et al. (2012) Clinical geneticists' views of VACTERL/VATER association. *American Journal of Medical Genetics Part A* 158A(12): 3087–3100. DOI: 10.1002/ajmg.a.35638.

Sondalle Samuel B, Longerich S, Ogawa LM, et al. (2019) Fanconi anemia protein FANCI functions in ribosome biogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 116(7): 2561–2570. DOI: 10.1073/pnas.1811557116.

Sondalle Samuel B., Longerich S, Ogawa LM, et al. (2019) Fanconi anemia protein FANCI functions in ribosome biogenesis. *Proceedings of the National Academy of Sciences* 116(7): 2561–2570. DOI: 10.1073/pnas.1811557116.

Soulier J (2011) Fanconi Anemia. *Hematology* 2011(1): 492–497. DOI: 10.1182/asheducation-2011.1.492.

Soulier J, Leblanc T, Larghero J, et al. (2005) Detection of somatic mosaicism and classification of Fanconi anemia patients by analysis of the FA/BRCA pathway. *Blood* 105(3): 1329–1336. DOI: 10.1182/blood-2004-05-1852.

Steinberg-Shemer O, Goldberg TA, Yacobovich J, et al. (2020) Characterization and genotype-phenotype correlation of patients with Fanconi anemia in a multi-ethnic population. *Haematologica* 105(7): 1825–1834. DOI: 10.3324/haematol.2019.222877.

Strieter RM, Polverini PJ, Arenberg DA, et al. (1995) THE ROLE OF CXC CHEMOKINES AS REGULATORS OF ANGIOGENESIS. *Shock* 4(3): 155–160. DOI: 10.1097/00024382-199509000-00001.

Strom CM, Crossley B, Redman JB, et al. (2004) Molecular screening for diseases frequent in Ashkenazi Jews: Lessons learned from more than 100,000 tests performed in a commercial laboratory. *Genetics in Medicine* 6(3): 145–152. DOI: 10.1097/01.GIM.0000127267.57526.D1.

Sumpter R and Levine B (2017) Emerging functions of the Fanconi anemia pathway at a glance. *Journal of Cell Science* 130(16): 2657–2662. DOI: 10.1242/jcs.204909.

Sumpter R, Sirasanagandla S, Fernández Álvaro F., et al. (2016) Fanconi Anemia Proteins Function in Mitophagy and Immunity. *Cell* 165(4): 867–881. DOI: 10.1016/j.cell.2016.04.006.

Sumpter R, Sirasanagandla S, Fernández Álvaro F., et al. (2016) Fanconi Anemia Proteins Function in Mitophagy and Immunity. *Cell* 165(4). Cell Press: 867–881. DOI: 10.1016/j.cell.2016.04.006.

Sun S, He H, Ma Y, et al. (2020) Inactivation of ribosomal protein S27-like impairs DNA interstrand cross-link repair by destabilization of FANCD2 and FANCI. *Cell Death & Disease* 11(10): 852. DOI: 10.1038/s41419-020-03082-9.

Suwaki N, Klare K and Tarsounas M (2011) RAD51 paralogs: Roles in DNA damage signalling, recombinational repair and tumorigenesis. *Seminars in Cell & Developmental Biology* 22(8): 898–905. DOI: 10.1016/j.semcdb.2011.07.019.

Suzuki A, de la Pompa JL, Hakem R, et al. (1997) Brca2 is required for embryonic cellular proliferation in the mouse. *Genes & Development* 11(10): 1242–1252. DOI: 10.1101/gad.11.10.1242.

Suzuki N, Yamazaki S, Yamaguchi T, et al. (2013) Generation of Engraftable Hematopoietic Stem Cells From Induced Pluripotent Stem Cells by Way of

Teratoma Formation. *Molecular Therapy* 21(7): 1424–1431. DOI: 10.1038/mt.2013.71.

Svahn J, Lanza T, Rathbun K, et al. (2015) p38 mitogen-activated protein kinase inhibition enhances in vitro erythropoiesis of Fanconi anemia, complementation group A–deficient bone marrow cells. *Experimental Hematology* 43(4): 295–299. DOI: 10.1016/j.exphem.2014.11.010.

T. Das A, Tenenbaum L and Berkhout B (2016) Tet-On Systems For Doxycycline-inducible Gene Expression. *Current Gene Therapy* 16(3): 156–167. DOI: 10.2174/1566523216666160524144041.

Tacconi EM, Lai X, Folio C, et al. (2017) BRCA1 and BRCA2 tumor suppressors protect against endogenous acetaldehyde toxicity. *EMBO Molecular Medicine* 9(10): 1398–1414. DOI: 10.15252/emmm.201607446.

Takahashi K and Yamanaka S (2006) Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* 126(4): 663–676. DOI: 10.1016/j.cell.2006.07.024.

Takahashi K, Tanabe K, Ohnuki M, et al. (2007) Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* 131(5): 861–872. DOI: 10.1016/j.cell.2007.11.019.

Tamary H, Bar-Yam R, Shalmon L, et al. (2000) Fanconi anaemia group A (FANCA) mutations in Israeli non-Ashkenazi Jewish patients. *British Journal of Haematology* 111(1): 338–343. DOI: 10.1046/j.1365-2141.2000.02323.x.

Tercanli S, Miny P, Siebert MS, et al. (2001) Fanconi anemia associated with increased nuchal translucency detected by first-trimester ultrasound. *Ultrasound in Obstetrics and Gynecology* 17(2): 160–162. DOI: 10.1046/j.1469-0705.2001.00321.x.

Tian X, Woll PS, Morris JK, et al. (2006) Hematopoietic Engraftment of Human Embryonic Stem Cell-Derived Cells Is Regulated by Recipient Innate Immunity. *STEM CELLS* 24(5): 1370–1380. DOI: 10.1634/stemcells.2005-0340.

Tipping AJ, Pearson T, Morgan N V., et al. (2001) Molecular and genealogical evidence for a founder effect in Fanconi anemia families of the Afrikaner population of South Africa. *Proceedings of the National Academy of Sciences* 98(10): 5734–5739. DOI: 10.1073/pnas.091402398.

Tischkowitz M, Ameziane N, Waisfisz Q, et al. (2003) Bi-allelic silencing of the Fanconi anaemia gene FANCF in acute myeloid leukaemia. *British Journal of Haematology* 123(3): 469–471. DOI: 10.1046/j.1365-2141.2003.04640.x.

Tischkowitz MD, Morgan N V, Grimwade D, et al. (2004) Deletion and reduced expression of the Fanconi anemia FANCA gene in sporadic acute myeloid leukemia. *Leukemia* 18(3): 420–425. DOI: 10.1038/sj.leu.2403280.

Toksoy G, Uludag Alkaya D, Bagirova G, et al. (2020) Clinical and Molecular Characterization of Fanconi Anemia Patients in Turkey. *Molecular Syndromology* 11(4): 183–196. DOI: 10.1159/000509838.

Tönnies H, Huber S, Kühl J-S, et al. (2003) Clonal chromosomal aberrations in bone marrow cells of Fanconi anemia patients: gains of the chromosomal segment 3q26q29 as an adverse risk factor. *Blood* 101(10): 3872–3874. DOI: 10.1182/blood-2002-10-3243.

Trenkmann M (2018) Putting genetic variants to a fitness test. *Nature Reviews Genetics* 19(11): 667. DOI: 10.1038/s41576-018-0056-4.

Tu VY, Ayari A and O'Connor RS (2021) Beyond the Lactate Paradox: How Lactate and Acidity Impact T Cell Therapies against Cancer. *Antibodies* 10(3): 25. DOI: 10.3390/antib10030025.

Uhlén M, Fagerberg L, Hallström BM, et al. (2015) Tissue-based map of the human proteome. *Science* 347(6220). DOI: 10.1126/science.1260419.

Urlinger S, Baron U, Thellmann M, et al. (2000) Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity. *Proceedings of the National Academy of Sciences* 97(14): 7963–7968. DOI: 10.1073/pnas.130192197.

Uziel O, Reshef H, Ravid A, et al. (2008) Oxidative stress causes telomere damage in Fanconi anaemia cells – a possible predisposition for malignant transformation. *British Journal of Haematology* 142(1): 82–93. DOI: 10.1111/j.1365-2141.2008.07137.x.

van de Vrugt H and Joenje H (2014) *Mice with a targeted disruption of the Fanconi anemia homolog Fanca*. Available at: https://www.researchgate.net/publication/12403684.

Van der Heijden MS, Brody JR, Gallmeier E, et al. (2004) Functional Defects in the Fanconi Anemia Pathway in Pancreatic Cancer Cells. *The American Journal of Pathology* 165(2): 651–657. DOI: https://doi.org/10.1016/S0002-9440(10)63329-9.

van der Lelij P, Oostra AB, Rooimans MA, et al. (2010) Diagnostic Overlap between Fanconi Anemia and the Cohesinopathies: Roberts Syndrome and Warsaw Breakage Syndrome. *Anemia* 2010: 1–7. DOI: 10.1155/2010/565268.

van Riggelen J, Yetil A and Felsher DW (2010) MYC as a regulator of ribosome biogenesis and protein synthesis. *Nature Reviews Cancer* 10(4): 301–309. DOI: 10.1038/nrc2819.

Van Wassenhove LD, Mochly-Rosen D and Weinberg KI (2016) Aldehyde dehydrogenase 2 in aplastic anemia, Fanconi anemia and hematopoietic stem

cells. *Molecular Genetics and Metabolism* 119(1–2): 28–36. DOI: 10.1016/j.ymgme.2016.07.004.

Vander Heiden MG, Cantley LC and Thompson CB (2009) Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* 324(5930): 1029–1033. DOI: 10.1126/science.1160809.

Vanderwerf SM, Svahn J, Olson S, et al. (2009) TLR8-dependent TNF-α overexpression in Fanconi anemia group C cells. *Blood* 114(26): 5290–5298. DOI: 10.1182/blood-2009-05-222414.

Vaz F, Hanenberg H, Schuster B, et al. (2010) Mutation of the RAD51C gene in a Fanconi anemia–like disorder. *Nature Genetics* 42(5): 406–409. DOI: 10.1038/ng.570.

Velleuer E and Dietrich R (2014) Fanconi anemia: young patients at high risk for squamous cell carcinoma. *Molecular and Cellular Pediatrics* 1(1): 9. DOI: 10.1186/s40348-014-0009-8.

Vieyra DS and Goodell MA (2007) Pluripotentiality and Conditional Transgene Regulation in Human Embryonic Stem Cells Expressing Insulated Tetracycline-ON Transactivator. *Stem Cells* 25(10): 2559–2566. DOI: 10.1634/stemcells.2007-0248.

Virts EL, Jankowska A, Mackay C, et al. (2015) AluY-mediated germline deletion, duplication and somatic stem cell reversion in UBE2T defines a new subtype of Fanconi anemia. *Human Molecular Genetics* 24(18): 5093–5108. DOI: 10.1093/hmg/ddv227.

Vodyanik MA, Thomson JA and Slukvin II (2006) Leukosialin (CD43) defines hematopoietic progenitors in human embryonic stem cell differentiation cultures. *Blood* 108(6): 2095–2105. DOI: 10.1182/blood-2006-02-003327.

Voulgaridou G-P, Anestopoulos I, Franco R, et al. (2011) DNA damage induced by endogenous aldehydes: Current state of knowledge. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 711(1–2): 13–27. DOI: 10.1016/j.mrfmmm.2011.03.006.

Wagner JE (2004) Germline mutations in BRCA2: shared genetic susceptibility to breast cancer, early onset leukemia, and Fanconi anemia. *Blood* 103(8): 3226–3229. DOI: 10.1182/blood-2003-09-3138.

Wagner JE, Tolar J, Levran O, et al. (2004) Germline mutations in BRCA2: shared genetic susceptibility to breast cancer, early onset leukemia, and Fanconi anemia. *Blood* 103(8): 3226–3229. DOI: 10.1182/blood-2003-09-3138.

Wakefield L (2002) TGF-beta signaling: positive and negative effects on tumorigenesis. *Current Opinion in Genetics & Development* 12(1): 22–29. DOI: 10.1016/S0959-437X(01)00259-3.

Walasek MA, van Os R and de Haan G (2012) Hematopoietic stem cell expansion: challenges and opportunities. *Annals of the New York Academy of Sciences* 1266(1): 138–150. DOI: 10.1111/j.1749-6632.2012.06549.x.

Wall JD, Stawiski EW, Ratan A, et al. (2019) The GenomeAsia 100K Project enables genetic discoveries across Asia. *Nature* 576(7785): 106–111. DOI: 10.1038/s41586-019-1793-z.

Wallace DC (2012) Mitochondria and cancer. *Nature Reviews Cancer* 12(10): 685–698. DOI: 10.1038/nrc3365.

Walter D, Lier A, Geiselhart A, et al. (2015) Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. *Nature* 520(7548): 549–552. DOI: 10.1038/nature14131.

Wang AT, Kim T, Wagner JE, et al. (2015) A Dominant Mutation in Human RAD51 Reveals Its Function in DNA Interstrand Crosslink Repair Independent of Homologous Recombination. *Molecular Cell* 59(3). Cell Press: 478–490. DOI: 10.1016/j.molcel.2015.07.009.

Wang L, Menendez P, Shojaei F, et al. (2005) Generation of hematopoietic repopulating cells from human embryonic stem cells independent of ectopic HOXB4 expression. *Journal of Experimental Medicine* 201(10): 1603–1614. DOI: 10.1084/jem.20041888.

Wang Y (2008) Bulky DNA Lesions Induced by Reactive Oxygen Species. *Chemical Research in Toxicology* 21(2): 276–281. DOI: 10.1021/tx700411g.

Wang Y, Branicky R, Noë A, et al. (2018) Superoxide dismutases: Dual roles in controlling ROS damage and regulating ROS signaling. *Journal of Cell Biology* 217(6): 1915–1928. DOI: 10.1083/jcb.201708007.

Warburg O (1924) Über den Stoffwechsel der Carcinomzelle. *Die Naturwissenschaften* 12(50): 1131–1137. DOI: 10.1007/BF01504608.

Whitney MA, Royle G, Low MJ, et al. (1996) Germ Cell Defects and Hematopoietic Hypersensitivity to y-Interferon in Mice With a Targeted Disruption of the Fanconi Anemia C Gene. *Blood* 88(1): 49–58. Available at: https://www.sciencedirect.com/science/article/pii/S0006497120646031.

Wong JCY (2003) Targeted disruption of exons 1 to 6 of the Fanconi Anemia group A gene leads to growth retardation, strain-specific microphthalmia, meiotic defects and primordial germ cell hypoplasia. *Human Molecular Genetics* 12(16): 2063–2076. DOI: 10.1093/hmg/ddg219.

Wu CG and Spies M (2016) G-quadruplex recognition and remodeling by the FANCJ helicase. *Nucleic Acids Research* 44(18): 8742–8753. DOI: 10.1093/nar/gkw574.

Wu Y, Shin-ya K and Brosh RM (2008) FANCJ Helicase Defective in Fanconia Anemia and Breast Cancer Unwinds G-Quadruplex DNA To Defend Genomic Stability. *Molecular and Cellular Biology* 28(12): 4116–4128. DOI: 10.1128/MCB.02210-07.

Xia B, Dorsman JC, Ameziane N, et al. (2007) Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nature genetics* 39(2). United States: 159–161. DOI: 10.1038/ng1942.

Xu G, Chapman JR, Brandsma I, et al. (2015) REV7 counteracts DNA double-strand break resection and affects PARP inhibition. *Nature* 521(7553): 541–544. DOI: 10.1038/nature14328.

Yagasaki H, Oda T, Adachi D, et al. (2003) Two common founder mutations of the fanconi anemia group g gene FANCG/XRCC9 in the Japanese population. *Human Mutation* 21(5): 555–555. DOI: 10.1002/humu.9142.

Yamada K, Olsen JC, Patel M, et al. (2001) Functional Correction of Fanconi Anemia Group C Hematopoietic Cells by the Use of a Novel Lentiviral Vector. *Molecular Therapy* 3(4): 485–490. DOI: 10.1006/mthe.2001.0287.

Yamashita T, Wu N, Kupfer G, et al. (1996) Clinical variability of Fanconi anemia (type C) results from expression of an amino terminal truncated Fanconi anemia complementation group C polypeptide with partial activity. *Blood* 87(10): 4424–4432. DOI: 10.1182/blood.V87.10.4424.bloodjournal87104424.

Yang H-C, Cheng M-L, Ho H-Y, et al. (2011) The microbicidal and cytoregulatory roles of NADPH oxidases. *Microbes and Infection* 13(2): 109–120. DOI: 10.1016/j.micinf.2010.10.008.

Yang T, Burrows C and Park JH (2014) Development of a doxycycline-inducible lentiviral plasmid with an instant regulatory feature. *Plasmid* 72: 29–35. DOI: 10.1016/j.plasmid.2014.04.001.

Yang Y, Kuang Y, De Oca RM, et al. (2001) Targeted disruption of the murine Fanconi anemia gene,Fancg/Xrcc9. *Blood* 98(12): 3435–3440. DOI: 10.1182/blood.V98.12.3435.

Yang Z, Wu XS, Wei Y, et al. (2021) Transcriptional Silencing of ALDH2 Confers a Dependency on Fanconi Anemia Proteins in Acute Myeloid Leukemia. *Cancer Discovery* 11(9): 2300–2315. DOI: 10.1158/2159-8290.CD-20-1542.

Yau TM (1979) Mutagenicity and cytotoxicity of malonaldehyde in mammalian cells. *Mechanisms of Ageing and Development* 11(2): 137–144. DOI: 10.1016/0047-6374(79)90031-9.

Yin H, Ma H, Hussain S, et al. (2019) A homozygous FANCM frameshift pathogenic variant causes male infertility. *Genetics in Medicine* 21(1): 62–70. DOI: 10.1038/s41436-018-0015-7.

Yu J, Vodyanik MA, Smuga-Otto K, et al. (2007) Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science* 318(5858): 1917–1920. DOI: 10.1126/science.1151526.

Yu X, Zhan X, D'Costa J, et al. (2003) Lentiviral vectors with two independent internal promoters transfer high-level expression of multiple transgenes to human hematopoietic stem-progenitor cells. *Molecular Therapy* 7(6): 827–838. DOI: 10.1016/S1525-0016(03)00104-7.

Yuneva M, Zamboni N, Oefner P, et al. (2007) Deficiency in glutamine but not glucose induces MYC-dependent apoptosis in human cells. *The Journal of Cell Biology* 178(1): 93–105. DOI: 10.1083/jcb.200703099.

Yung SK, Tilgner K, Ledran MH, et al. (2013) Brief report: Human pluripotent stem cell models of fanconi anemia deficiency reveal an important role for fanconi anemia proteins in cellular reprogramming and survival of hematopoietic progenitors. *Stem Cells* 31(5): 1022–1029. DOI: 10.1002/stem.1308.

Zahra KF, Lefter R, Ali A, et al. (2021) The Involvement of the Oxidative Stress Status in Cancer Pathology: A Double View on the Role of the Antioxidants. *Oxidative Medicine and Cellular Longevity* 2021: 1–25. DOI: 10.1155/2021/9965916.

Zarka J, Short NJ, Kanagal-Shamanna R, et al. (2020) Nucleophosmin 1 Mutations in Acute Myeloid Leukemia. *Genes* 11(6): 649. DOI: 10.3390/genes11060649.

Zhang H, Kozono DE, O'Connor KW, et al. (2016) TGF-β Inhibition Rescues Hematopoietic Stem Cell Defects and Bone Marrow Failure in Fanconi Anemia. *Cell Stem Cell* 18(5): 668–681. DOI: 10.1016/j.stem.2016.03.002.

Zhang Q-S, Marquez-Loza L, Eaton L, et al. (2010) Fancd2 −/− mice have hematopoietic defects that can be partially corrected by resveratrol. *Blood* 116(24): 5140–5148. DOI: 10.1182/blood-2010-04-278226.

Zhang Q-S, Tang W, Deater M, et al. (2016) Metformin improves defective hematopoiesis and delays tumor formation in Fanconi anemia mice. *Blood* 128(24): 2774–2784. DOI: 10.1182/blood-2015-11-683490.

Zhang X, Li J, Sejas DP, et al. (2004) The Fanconi Anemia Proteins Functionally Interact with the Protein Kinase Regulated by RNA (PKR). *Journal of Biological Chemistry* 279(42): 43910–43919. DOI: 10.1074/jbc.M403884200.

Zhang X, Sejas DP, Qiu Y, et al. (2007) Inflammatory ROS promote and cooperate with the Fanconi anemia mutation for hematopoietic senescence. *Journal of Cell Science* 120(9): 1572–1583. DOI: 10.1242/jcs.003152.

Zhao Marion G., Sauter Sharon, Kovacic Melinda Butsch, et al. (2018) Lipidomic Profiling Links the Fanconi Anemia Pathway to Glycosphingolipid Metabolism in Head and Neck Cancer Cells. *Clinical cancer research : an official journal of the American Association for Cancer Research* 24(11): 2700–2709. DOI: 10.1158/1078-0432.ccr-17-3686.

Zook JM, McDaniel J, Olson ND, et al. (2019) An open resource for accurately benchmarking small variant and reference calls. *Nature biotechnology* 37(5): 561–566. DOI: 10.1038/s41587-019-0074-6.

Zufferey R, Donello JE, Trono D, et al. (1999) Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element Enhances Expression of Transgenes Delivered by Retroviral Vectors. *Journal of Virology* 73(4): 2886–2892. DOI: 10.1128/JVI.73.4.2886-2892.1999.

8. LIST OF PUBLICATIONS

Joshi G, Arthur NBJ, Geetha TS, Datari PVR, Modak K, Roy D, Chaudhury AD, Sundaraganesan P, Priyanka S, Na F, Ramprasad V, Abraham A, Srivastava VM, Srivastava A, Kulkarni UP, George B, Velayudhan SR. Comprehensive laboratory diagnosis of Fanconi anaemia: comparison of cellular and molecular analysis. J Med Genet. 2023 Aug;60(8):801-809. doi: 10.1136/jmg-2022-108714. Epub 2023 Mar 9. PMID: 36894310.

Raina K, **Joshi G**, Modak K, Premkumar C, Priyanka S, Rajesh P, Velayudhan SR, Thummer RP. Generation and characterization of induced pluripotent stem cell line IITGi001-A derived from adult human primary dermal fibroblasts. Stem Cell Res. 2023 Jun 28;71:103159. doi: 10.1016/j.scr.2023.103159. Epub ahead of print. PMID: 37392703.

Nandy K, Babu D, Rani S, **Joshi G**, Ijee S, George A, Palani D, Premkumar C, Rajesh P, Vijayanand S, David E, Murugesan M, Velayudhan SR. Efficient gene editing in induced pluripotent stem cells enabled by an inducible adenine base editor with tunable expression. Sci Rep. 2023 Dec 11;13(1):21953. doi: 10.1038/s41598-023-42174-2. PMID: 38081875; PMCID: PMC10713686.

9. APPENDICES

Table A1. Spectrum of congenital abnormalities, BMT outcome, CBA and defective genes in 153 FA patients

ID	Age/ Gender	BMT outcome	BMT source	OS	CBA	Type of AA	Gene (mutation type)	Diagnosis	Treatment	Consanguinity	Skeletal abnormalities	Facial dysmorphism	Skin changes	Other abnormalities	Relatives with FA
FA-01/ FA-536	6 / M	NA	NA	NA	34.4	VSAA	FANCG (3' splicing)	NA	Danazol	Yes	NA	NA	HPOP, Leukoplakia	NA	NA
FA-522	5 / M	NA	NA	NA	NA	NSAA	FANCL (Exonic splice variant)	NA	NA	No	Thenar and hypothenar hypoplasia	NA	CAL, HP	NA	NA
FA-02	34 / M	NA	NA	NA	141	NSAA	FANCJ (Missense)	MDS/AML	Danazol	No	NA	MCH	CAL	NA	NA
FA-03/ FA-515	7 / F	NA	NA	NA	62.9	NSAA	FANCA (5' splicing/ missense)	NA	Danazol	NA	TH, TA	DF	HP	NA	NA
FA-03/19	4 / M	NA	NA	NA	66.8	NSAA	FANCI (5' splicing/ missense)	NA	Stanozolol	No	SS	LSE, MOA, MCH, TF	HP, CAL	NA	Sister

FA-04/ FA-509	19 / M	NA	NA	NA	106	NSAA	FANCA (3' splicing/ deletion)	NA	Stanozolol	No	BT, Absent thumb	LSE	HP, HPOP, CAL	NA	NA
FA-05/ FA-500	10 / M	NA	NA	NA	73.6	NSAA	FANCA (Missense)	NA	Stanozolol	Yes	NA	NA	HP	Small kidney (L)	NA
FA-18/21	7 / M	NA	NA	NA	132	NSAA	FANCA (Missense)	MDS	Stanozolol	Yes	SS, TH	EF	HP	NA	NA
FA-675/18	3 / F	Alive	PBSCT/ Haplo	53	192.2	NSAA	FANCL (Exonic splice variant)	NA	Stanozolol/ BMT	No	NA	MCH, EF	NA	NA	NA
FA-08/ FA-552	14 / M	Alive	PBSCT/ MRD	109	53.1	NSAA	FANCA (Frameshift insertion)	DKC	Stanozolol/BMT	Yes	NA	NA	HP, leukoplakia	NA	NA
FA-09/ FA-505	17 / F	Alive	PBSCT/ MRD	111	64.7	NSAA	FANCA (5' splicing/ frameshift deletion)	NA	Danazol/BMT	No	NA	NA	HPOP	NA	NA
FA-10	6 / F	NA	NA	NA	49.9	NSAA	FANCA	NA	Danazol	Yes	Clinodactyly	MCH, MGH, HAP	HP, POH	NA	NA

							(Frameshift deletion)								
FA-11/ FA-524	4 / M	NA	NA	NA	55.6	NSAA	FANCA (Deletion)	NA	Stanozolol	No	SS, Arachinodactyly	DF	HP	NA	NA
FA-12	4 / F	NA	NA	NA	36	NSAA	FANCA (Frameshift insertion/deletion)	NA	Danazol	No	NA	NA	HP, POH	NA	NA
FA-12/19	12 / M	NA	NA	NA	45.1	NSAA	FANCA (Frameshift deletion/ 3' splicing)	NA	Stanozolol	No	Hypothenar, clinodactyly, partial syndactyly	HAP	CAL	HSK	NA
FA-13	30 / M	NA	NA	NA	33.5	NSAA	FANCG (Frameshift deletion)	MDS	Stanozolol	NA	NA	MCH,TF	HPOP, HP	abnormal right kidney	NA
P-603	4 / M	NA	NA	NA	118.3	NSAA	FANCL (Exonic splice variant)	NA	Stanozolol	Yes	Clinodactyly	LSE, HAP, depressed nasal bridge, MCH	Ichthyosis, CAL, HP	NA	NA
FA-31/21	7 / F	NA	NA	NA	145.9	NSAA	FANCA	MDS (Monosomy 7)	Danazol	NA	NA	EF, HAP	NA	NA	NA

							(Missense)								
FA-16	13 / F	NA	NA	NA	63	NSAA	FANCA (Deletion)	NA	Danazol	Yes	SS	MGH	HP	NA	NA
FA-17/19	12 / M	NA	NA	NA	89.1	NSAA	FANCA (Deletion)	NA	Stanozolol	No	NA	MCH	CAL	NA	Brother
FA-565	11 / M	NA	NA	NA	NA	NSAA	FANCA (Nonsense)	MDS	Stanozolol	No	Clinodactyly	MCH, MOA	CAL, Pallor	NA	NA
O-126	11 / M	Alive	PBSCT/ MUD	52	78.6	NSAA	FANCA (Deletion)	MDS-AML	Stanozolol/ BMT	Yes	Simian crease in both hands	LSE, hypertelorism	HP	NA	NA
FA-595	8 / M	NA	NA	NA	90.5	NSAA	FANCL (Exonic splice variant)	NA	Stanozolol	No	NA	NA	CAL, HPOP	NA	Brother
FA-33/19	4 / M	NA	NA	NA	64.5	NSAA	FANCL (Exonic splice variant)	NA	Danazol	No	NA	NA	HP, CAL	Small testes	NA
FA-20	1 / M	NA	NA	NA	25.3	NSAA	FANCG (3' splicing)	NA	Stanozolol	Yes	BT, Polydactyly	AE,MGH	HP	NA	NA
FA-20/19	12 / M	NA	NA	NA	31	NSAA	FANCA	NA	Stanozolol	Yes	NA	DF	HP	Ectopic kidney	NA

							(3' splicing)								
FA-21/ FA-510	15 / M	NA	NA	NA	59	NSAA	FANCA (Deletion)	NA	Stanozolol	Yes	BT	AE	HP	HSK	NA
FA-22	4 / M	Dead	PBSCT/ MRD	0	46.8	NSAA	FANCA (5' splicing)	NA	Stanozolol/ BMT	Yes	Polydactyly, Nail dystrophy	DF	Petechiae, POH, CAL, Leokoplakia	NA	NA
FA-23	6 / F	NA	NA	NA	54.7	NSAA	FANCA (5' splicing)	NA	NA	No	Thumb Abnormalities	MCH	HPOP, HP	NA	NA
FA-24/ FA-502	13 / M	Dead	PBSCT/ MRD	10	57.2	NSAA	FANCA (Nonsense/deletion)	SCC	Stanozolol/ BMT	No	Thenar hypoplasia	NA	CAL, HP	NA	NA
FA-638/18	3 / M	NA	NA	NA	62	NSAA	FANCL (Exonic splice variant)	NA	Stanozolol	No	Bilateral clinodactyly, SST	LSE, MOA	COP, HP, CAL	NA	NA
FA-26/ FA-530	6 / M	NA	NA	NA	76.8	NSAA	FANCG (5' intronic splicing)	NA	Stanozolol	Yes	SS, Polydactyly, Clubbing	MCH	HPOP, HP	NA	NA
FA-27	12 / F	NA	NA	NA	47.4	NSAA	FANCA (5' splicing)	NA	Danazol	Yes	NA	NA	HP	NA	NA

FA-27/19	9 / M	NA	NA	NA	23.7	NSAA	FANCA (5' splicing)	NA	Stanozolol	Yes	Triphalageal thumb	Hypertelorism, HAP	HP	NA	Sister, 1 death at birth
FA-28	12 / F	NA	NA	NA	53.3	NSAA	FANCA (5' splicing)	NA	Danazol	Yes	SS	MCH	HP	NA	NA
FA-29	19 / F	Alive	PBSCT/ MRD	11	0	VSAA	FANCD 2(Missense)	NA	BMT	Yes	NA	NA	NA	HSK	NA
FA-30/ FA-501	7 / M	NA	NA	NA	44.2	NSAA	FANCF (Stop loss)	SCC	Stanazolol	NA	Thenar hypolasia	LSE, MCH	HPOP, CAL, HP	NA	NA
FA-548	4 / M	NA	NA	NA	61.5	NSAA	FANCL (Exonic splice variant)	NA	Stanozolol	No	SS	MCH, DF	HP, HPOP	NA	NA
FA-31/ FA-514	9 / F	NA	NA	NA	27.7	NSAA	FANCA (5' splicing)	NA	Danazol	No	SS	HAP	HP	NA	NA
FA-32	8 / M	NA	NA	NA	55.6	NSAA	FANCA (5' splicing)	NA	Stanozolol	No	SS	LSE	HP, HPOP	NA	NA
FA-33	11 / M	Alive	PBSCT/ Haplo	8	30.3	NSAA	FANCA (Frameshift insertion)	NA	Stanozolol/ BMT	Yes	NA	NA	HP	NA	Sister with AA, brother died

FA-34	9 / M	Alive	PBSCT/ MRD	8	49.9	NSAA	FANCA (5' splicing)	NA	Stanozolol/ BMT	Yes	SS	TF	CAL, HP	NA	NA
FA-35	12 / M	Dead	PBSCT/ MRD	0	36	NSAA	FANCA (Frameshift deletion)	NA	BMT	NA	TH, SD	Depressed nasal bridge, MCH	CAL	NA	NA
FA-36/ FA-526	8 / M	NA	NA	NA	35	NSAA	FANCG (Frameshift deletion)	NA	Stanozolol	Yes	NA	TF, ASE	CAL, HP	Hydronephrosis right kidney	NA
FA-36/19	9 / M	NA	NA	NA	NA	NSAA	FANCG (3' splicing)	MDS/AML	Stanozolol	No	TH	HAP, small mouth and eyes, ASE	HP, CAL	Small kidneys	Sister
FA-37/19	7 / F	NA	NA	NA	110.1	NSAA	FANCG (3' splicing)	MDS	Danazol	No	NA	HAP, small mouth and eyes, ASE	HP, CAL	NA	Brother
FA-38	13 / F	Alive	PBSCT/ MRD	8	69.5	NSAA	FANCG (Missense/frameshift deletion)	NA	BMT	No	Subtle thumb abnormality	NA	HP	fused left renal and right renal fossa	NA
FA-39 /FA-518	13 / M	NA	NA	NA	0	NSAA	FANCC (Missense)	NA	Stanozolol	Yes	Thenar hypoplasia	NA	HP	NA	NA

FP-98-Q-471	1 / M	NA	NA	NA	52.4	NSAA	FANCL (Exonic splice variant)	NA	Stanozolol	Yes	NA	Small eyes,small face	COH, Ash leaf macule	NA	NA
FA-511	6 / M	Alive	PBSCT/ MRD	11	25.3	NSAA	FANCA (Nonsense)	NA	Stanozolol/ BMT	Yes	SST	NA	HP, CAL	NA	Sibling died of FA-AML
FA-513	9 / M	NA	NA	NA	80.9	NSAA	FANCA (Deletion)	NA	Stanozolol	Yes	TH	LSE	HP, CAL	Ectopic left kidney	NA
FA-516	12 / M	NA	NA	NA	31.9	NSAA	FANCA (Frameshift deletion)	NA	Stanozolol	Yes	NA	LSE, MCH	CAL	NA	NA
FA-517	5 / M	Dead	PBSCT/ MRD	8	62.2	NSAA	FANCC (Nonsense)	NA	Danazol/ BMT	Yes	TA, Ectrodactyly, Clinodactyly	HAP, MOA, MCH	CAL	Fused ectopia on left (R to L) kidney	NA
FA-519	8 / M	NA	NA	NA	62.7	SAA	FANCD2 (Nonsense)	NA	Stanozolol	No	SS, Absent simian crease, TH	TF	HP	Ectopic and malrotated left kidney	NA
FA-521	25 / M	NA	NA	NA	58.1	NSAA	FANCA (3' splicing)	NA	Stanozolol	Yes	TH	AE	HP	NA	NA
FA-19	5 / M	NA	NA	NA	51.3	NSAA	FANCL	NA	Stanozolol	NA	BT, SS	MOA, MCH	HP	NA	NA

							(Exonic splice variant)								
FA-564	2 / M	NA	NA	NA	45.8	NSAA	FANCL (Exonic splice variant)	NA	Stanozolol	Yes	SS	MOA	COP, CAL	NA	NA
FA-527	6 / F	Alive	PBSCT/ MRD	117	152	SAA	FANCC (Missense)	NA	Danazol/ BMT	No	SS, Clinodactyly	NA	HP	NA	NA
FA-528	5 / F	Alive	PBSCT/ MRD	80	0	NSAA	C17orf70 (Missense)	NA	BMT	No	SS	MCH	NA	NA	NA
FA-529	30 / F	NA	NA	NA	55	NSAA	FANCA (Deletion)	NA	Danazol	No	SS, BT	MCH	HP	NA	NA
FA-531	6 / M	Dead	PBSCT/ MUD	0	77	NSAA	FANCG (3' splicing)	NA	BMT	Yes	Clinodactyly	NA	HP, CAL	ARP	NA
FA-601	13 / M	Dead	PBSCT/ Haplo	0	45.8	NSAA	FANCL (Exonic splice variant)	NA	Danazol/ BMT	No	NA	Oral leukoplakia, MCH	HP, HPOP, POH	NA	Brother
FA-533	11 / F	Dead	PBSCT/ MUD	1	43.3	NSAA	FANCA (Frameshift deletion/	NA	Danazol/ BMT	No	NA	HAP, MGH	HP	NA	NA

							Missense)								
FA-14/ FA-534	6 / M	NA	NA	NA	45.1	NSAA	FANCL (Exonic splice variant)	NA	Stanozolol	No	NA	NA	HP	NA	NA
FA-25	4 / M	Alive	PBSCT/ MRD	127	43.3	SAA	FANCL (Exonic splice variant)	NA	Stanozolol/ BMT	No	SS	NA	HP	Ectopic kidney	NA
FA-34/19	7 / M	NA	NA	NA	35.1	SAA	FANCL (Exonic splice variant)	NA	Stanozolol	No	NA	NA	HP	NA	NA
FA-544	15 / M	NA	NA	NA	32.8	SAA	FANCA (5' splicing)	NA	Stanozolol	NA	Clinodactyly	LSE, HAP	HPOP, CAL, HP	NA	NA
FA-40	11 / M	NA	NA	NA	51.5	VSAA	FANCA (Frameshift deletion)	AML	NA	Yes	NA	TF	CAL	kidney abnormalities	NA
FA-543	24 / F	NA	NA	NA	32.8	NSAA	FANCL (Exonic splice variant)	NA	Danazol	Yes	Nail dystrophy	HAP	HPOP	NA	NA
FA-549	10 / M	NA	NA	NA	22	NSAA	FANCA (Missense)	NA	Stanozolol	No	BT, Polydactyly	NA	HP	Ectopic kidney	NA

FA-554	4 / M	NA	NA	NA	4.1	NSAA	FANCA (5' splicing/ deletion)	NA	Stanozolol	No	SS, Bilateral clinodactyly	NA	CAL, HP, Pallor	NA	Sister
FA-556	9 / M	NA	NA	NA	17.1	VSAA	FANCA (Frameshift deletion)	NA	Stanozolol	No	NA	NA	HP	Kidney abnormalities	NA
FA-557	12 / M	Alive	PBSCT/ MRD	71	83.6	NSAA	FANCA (5' splicing)	NA	Stanozolol/ BMT	No	NA	NA	HP, HPOP	NA	NA
FA-547	14 / F	NA	NA	NA	73.4	NSAA	FANCA (Nonsense)	RCMD-RS MDS	Danazol	Yes	Skeletal deformity	MOA ,DF, MGH	HPOP	NA	NA
FA18/20 N-914	15 / F	Alive	PBSCT/ MRD	32	26.2	NSAA	FANCL (Exonic splice variant)	DKC	Danazol/BMT	No	Nail dystrophy, Scoliosis	NA	HP, oral leukoplakia, CAL	NA	NA
FA-593	14 / M	NA	NA	NA	112.6	SAA	FANCA (Missense)	MDS	Stanozolol	Yes	TH (R), TA (L)	NA	HP	NA	NA
FA-573	11 / M	NA	NA	NA	87.3	NSAA	FANCA (Frameshift deletion/ Missense)	NA	Stanozolol/ Metformin	No	SS	NA	CAL	Hypogonadism	NA

FA-581	5 / M	Alive	PBSCT/ MRD	25	68.6	NSAA	FANCA (Frameshift deletion)	NA	Stanozolol/ BMT	Yes	SS, TH	MCH, LSE, MOA	POH	Ectopic kidney	NA
FA-588	9 / M	Alive	PBSCT/ MRD	70	103.5	SAA	FANCG (3' splicing)	NA	Stanozolol/ BMT	No	SS, Short neck, thumb abnormalities	MCH, hypertelorism	POH, HP	NA	NA
FA-591	14 / F	NA	NA	NA	45.1	NSAA	FANCA (5' splicing)	NA	Danazol	Yes	Thenar and hypothenar atrophy	NA	HP, CAL	NA	yes
FA-592	13 / F	Alive	PBSCT/ MRD	70	83.6	NSAA	FANCA (Deletion)	NA	Danazol/ BMT	No	SS	NA	POH, CAL	NA	NA
FA-18/19	15 / M	NA	NA	NA	115.3	NSAA	FANCA (Deletion)	MDS	Stanozolol	No	BT, polydactyly	DF, EF, HAP	HP	NA	Brother
FA-49/19	6 / M	NA	NA	NA	22.8	NSAA	FANCL (Exonic splice variant)	DKC	NA	Yes	SS	NA	HP, HPOP, CAL	Hypoplastic kidneys	NA
FA-598	7 / M	Alive	PBSCT/ MRD	68	111.7	SAA	FANCG (Nonsense)	NA	Stanozolol/ BMT	Yes	SS	NA	CAL	Atrophic left kidney	NA
FA-599	15 / F	NA	NA	NA	70.4	NSAA	FANCA (Frameshift	NA	Danazol	Yes	NA	Small and closely set eyes, HAP	CAL, HP	Fused kidneys	NA

							deletion)								
FA-665/18 FA12/22	3 / F	NA	NA	NA	20.5	NSAA	FANCL (Exonic splice variant)	NA	Danazol	Yes	SS, clinodactyly	MCH, LSE, HAP	CAL, COP	Anorectal malformation	NA
FA-614	14 / M	Alive	PBSCT/ MRD	67	99.8	NSAA	FANCA (Deletion)	NA	Stanozolol/ BMT	No	SS, clinodactyly	HAP	HPOP	Left ectopic kidney	NA
FA-622	14 / F	NA	NA	NA	56.5	NSAA	FANCA (Frameshift deletion/ Deletion)	NA	Danazol	No	NA	NA	HP	NA	NA
FA-629/18	8 / M	NA	NA	NA	106.6	NSAA	FANCG (5' intronic splice donor variant)	NA	NA	No	SS, TH	HAP, EF	NA	Absent kidney	NA
FA-15 FA-503	16 / F	Dead	PBSCT/ MRD	1	50.6	NSAA	FANCA (Exonic splice variant)	MDS	Danazol/BMT	NA	SS	MCH	HP	ARP	Sister
FA-636/18	45 / F	NA	NA	NA	66.3	NSAA	FANCT (Deletion)	MDS/AML	NA	No	SS, congenital kyphoscoliosis, SD	MCH, TF	HP	Absent left kidney	NA

FA-637/18	35 / M	Dead	PBSCT/MRD	1	134	NSAA	FANCJ (Missense)	MDS	BMT	No	NA	NA	NA	Ectopic kidney	NA
FA01/19	7 / M	Dead	PBSCT/MRD	0	8.2	NSAA	FANCL (Exonic splice variant)	MDS	BMT	NA	NA	NA	CAL, HP	NA	NA
FA-641/18	29 / M	Alive	PBSCT/MRD	64	93.9	NSAA	FANCF (Missense)	MDS	BMT	No	Hypoplasia of upper and lower limbs	NA	HP, CAL	NA	NA
FA-650/18	5 / F	NA	NA	NA	4.1	NSAA	FANCG (Frameshift insertion)	NA	Danazol	Yes	SS, TA	NA	HP	NA	NA
FA-652/18	5 / F	NA	NA	NA	172.6	SAA	FANCA (Missense)	NA	Danazol	Yes	NA	DF	NA	NA	NA
FA-659/18	0 / M	Alive	PBSCT/MRD	37	44.2	NSAA	FANCA (Missense)	NA	Stanozolol/ BMT	Yes	congenital TH, CTEV	MCH, LSE	POH, CAL	Congenital acyanotic heart disease	NA
FA-660/18	26 / F	NA	NA	NA	43.3	NSAA	FANCA (3' splicing)	NA	NA (Pregnant)	Yes	Polydactyly	HAP	HP	NA	NA
FA-07	3 / M	Alive	PBSCT/MRD	89	0	VSAA	FANCL (Exonic splice variant)	NA	Stanozolol/BMT	NA	NA	MCH, frontal blossing	HP, HPOP	NA	NA

FA-672/18	5 / M	NA	NA	NA	104.6	NSAA	FANCA (Deletion)	NA	Danazol	No	BT	HAP	CAL	NA	NA
FA-674/18	7 / M	NA	NA	NA	96.4	SAA	FANCA (Missense/Deletion)	NA	Stanozolol	No	TH	MCH	Patches over body	kidney abnormalities, ARP	NA
FA-30/19	10 / F	NA	NA	NA	0	NSAA	FANCL (Exonic splice variant)	NA	Danazol	Yes	NA	MCH	HP, HPOP	NA	NA
O-117	15 / F	Alive	PBSCT/ MRD	65	56.5	SAA	FANCA (Deletion)	NA	Danazol/ BMT	Yes	SS, emaciated	MCH, MOA, LSE, HAP, EF	POH, HPOP	NA	NA
FA-523	11 / M	NA	NA	NA	0	NSAA	FANCL (Exonic splice variant)	NA	Stanozolol	No	NA	MCH	CAL, HP	NA	NA
FA-08/19	16 / F	NA	NA	NA	98	NSAA	FANCA (Deletion)	MDS-EB1	Danazol	Yes	SS, Polydactyly	NA	HP	congenital acyanotic heart disease	NA
FA-561	18 / F	Dead	PBSCT/ MRD	0	0	NSAA	FANCA (Deletion)	MDS	Danazol/ BMT	Yes	SS	Hypertelorism	HP	NA	NA
FA-525	11 / M	Alive	PBSCT/ MRD	109	0	NSAA	FANCL (Exonic splice variant)	NA	Stanozolol/ BMT	No	Syndactyly, DPT	NA	HP, HPOP	NA	NA

FA-541	22 / F	Alive	PBSCT/ MRD	83	0	SAA	FANCE (Missense)	NA	BMT	Yes	NA	NA	NA	NA	NA
FA-542	NA / F	NA	NA	NA	NA	NA	FANCA (5' splicing)	NA	NA	NA	NA	NA	NA	NA	NA
FA-23/19	18 / M	NA	NA	NA	63.8	NSAA	FANCA (Exonic splice variant)	MDS	Stanozolol	Yes	SS	EF, HAP	NA	NA	NA
FA-580	7 / F	NA	NA	NA	59	NSAA	FANCG (Frameshift insertion)	NA	Danazol	No	NA	NA	NA	HSK	NA
FA-584	15 / F	NA	NA	NA	43.3	NSAA	FANCA (Frameshift deletion)	NA	Danazol	Yes	NA	NA	Pigmented tongue	NA	NA
FA-646/18	13 / F	Alive	PBSCT/ MRD	54	117.8	NSAA	FANCA (Missense)	NA	Danazol/ BMT	Yes	NA	NA	NA	Anorectal malformation	Sister died with congenital anomaly
FP-23-P-11	16 / M	NA	NA	NA	49.9	NSAA	FANCA (Missense)	NA	Danazol	No	NA	NA	NA	Anorectal malformation	NA
P-177	8 / F	NA	NA	NA	54.7	NSAA	FANCA	NA	Danazol	No	NA	Epicanthus	NA	NA	NA

							(Missense)								
FA-532	7 / F	NA	NA	NA	0	SAA	FANCL (Exonic splice variant)	NA	Danazol	Yes	NA	DF	HP, HPOP	NA	NA
FA-04/20	20 / F	Dead	PBSCT/ MUD	0	38	NSAA	FANCA (Frameshift deletion)	NA	Stanozolol/ BMT	No	NA	NA	NA	NA	NA
FA-06/20	5 / F	NA	NA	NA	45.1	NSAA	FANCA (Missense)	NA	Danazol	Yes	Polydactyly	HAP, EF	NA	NA	NA
FA-10/20	9 / F	NA	NA	NA	72.7	NSAA	FANCA (Missense)	NA	Danazol	Yes	NA	MCH, HAP	HP, CAL	NA	NA
FA-10/21	7 / F	NA	NA	NA	52.2	NSAA	FANCA (Frameshift deletion)	NA	Danazol	Yes	Thenar and hypothenar hypoplasia	MCH, Small and closely set eyes, HAP	CAL	NA	NA
FA-11/20	6 / M	NA	NA	NA	58.1	NSAA	FANCA (Frameshift insertion)	NA	Stanozolol	Yes	SS	MCH, TF, HAP	CAL	NA	NA
FA-15/20	6 / M	NA	NA	NA	76.8	NSAA	FANCG	NA	Stanozolol	No	SS	NA	HP	ARP	NA

							(Frameshift deletion)								
FA-18/ FA-551	22 / M	NA	NA	NA	44.9	VSAA	FANCA (Missense)	MDS/AML	Stanozolol	No	SS	LSE	HP	NA	Sister had AA
FA-21/20	24 / F	Dead	PBSCT/ MRD	0	165.7	NSAA	FANCA (Nonsense)	MDS	BMT	Yes	NA	MCH	HP	NA	NA
FA-5/21	30 / M	NA	NA	NA	15.5	NSAA	FANCT (Missense)	MDS-MLD	Danazol	No	NA	NA	NA	NA	NA
FA-01/20	5 / M	NA	NA	NA	99.6	VSAA	FANCG (Nonsense)	NA	Stanozolol	Yes	Clinodactyly	MCH, MGH	NA	Left ectopic kidney	NA
FA-02/21	20 / M	NA	NA	NA	47.1	NSAA	FANCA (Missense/missense)	NA	Stanozolol	No	SS	NA	NA	NA	NA
FA-06	26 / M	Dead	PBSCT/ MRD	7	58.1	NSAA	FANCA (Missense)	MDS	BMT	NA	SS, TH, Thenar hypoplasia	TF	HPOP, HP	NA	NA
FA-13/19	5 / F	NA	NA	NA	68.4	NSAA	FANCA (Nonsense)	NA	Danazol	No	Polydactyly, Thumb anomalies	MCH, HAP	CAL	NA	NA
FA-631/18	29 / F	NA	NA	NA	59.7	NSAA	FANCA	RCMD-RS MDS	Danazol	No	NA	DF	NA	NA	Sister died during

							(Nonsense)								chemotherapy
FA-21/21	13 / F	Alive	PBSCT/MRD	22	49.9	NSAA	FANCA (Missense)	NA	Danazol/BMT	No	NA	NA	HP	NA	NA
FA-568	29 / M	Dead	PBSCT/MRD	0	28.5	SAA	FANCA (Missense)	MDS	BMT	No	NA	NA	NA	NA	Sibling died
FA-26/21	8 / M	NA	NA	NA	80.7	SAA	FANCA (Frameshift insertion)	NA	NA	Yes	Hypothenar wasting	NA	HP	Anorectal malformation	NA
FA-30/21	39 / F	NA	NA	NA	15.5	NSAA	UBE2T/FANCT (Deletion)	NA	NA	Yes	Polydactyly	NA	HP	HSK	NA
FA-16/20	29 / F	NA	NA	NA	62.2	VSAA	FANCA (5' Splice site)	MDS	Danazol	No	NA	NA	NA	NA	NA
FA-32/19	19 / M	NA	NA	NA	22.1	NSAA	FANCA (Missense)	NA	Stanozolol	Yes	Musculoskeletal abnormalities	Hypertelorism, MCH, HAP	CAL	NA	NA
FA-535	9 / M	NA	NA	NA	0	SAA	FANCL (Exonic splice variant)	NA	Stanozolol	Yes	SS	MGH	HP	NA	NA

FA-33/21	9 / M	NA	NA	NA	44.9	NSAA	FANCA (Frameshift insertion)	NA	NA	Yes	NA	Fanconi facies	CAL	NA	NA
FA-538	4 / M	NA	NA	NA	0	NSAA	FANCL (Exonic splice variant)	NA	Stanozolol	NA	SS, TH	MCH, HAP	CAL	NA	NA
FA-35/21	9 / F	NA	NA	NA	36.9	NSAA	FANCA (Frameshift insertion)	NA	Danazol	Yes	Thenar hypoplasia	NA	NA	NA	NA
FA-649/18	0 / M	NA	NA	NA	114.4	NSAA	BRCA2 (Missense)	NA	NA	No	Thumb Abnormalities	MCH	HP, HPOP	Wilms tumor	Sibling died of FA-AML
FA 07/20	15 / M	NA	NA	NA	30.3	NSAA	NA (Deletion)	AA	Stanozolol	No	Clubbing in both hands	HAP, MCH	COP	NA	NA
FA 01/22	7 / M	NA	NA	NA	132.7	NSAA	FANCD2 (Frameshift deletion)	NA	Danazol	No	Polydactyly	NA	HPOP, POH	Fused ectopic kidney	NA
FA 02/22	39 / M	NA	NA	NA	55.4	NSAA	FANCA (Missense)	NA	NA	Yes	NA	NA	NA	NA	Brother with AA

FA 11/22	5 / F	Dead	PBSCT/ MRD	0	141.8	NSAA	FANCG (Frameshift insertion)	NA	BMT	Yes	Clinodactyly	MCH	NA	Ectopic HSK	NA
FA 14/22	9 / F	NA	NA	NA	84.1	SAA	FANCA (Frameshift deletion/ Deletion)	NA	Danazol	No	BT	NA	POH	NA	NA
FA 23/22	29 / F	NA	NA	NA	232.4	NSAA	FANCA (Missense)	MDS-EB1	Danazol	NA	Thenar Atrophy	NA	CAL	NA	NA
FA 24/22	8 / F	NA	NA	NA	115.1	VSAA	FANCF (Nonsense)	MDS	Danazol	NA	Polydactyly	EF	NA	NA	NA
FA 22/22	6 / F	NA	NA	NA	66.8	NSAA	FANCD2 (Missense)	NA	Danazol	NA	Clinodactyly	Small face, epicanthus HAP	NA	Bilateral crackles	Grandmother has FA
FA 27/22	7 / M	NA	NA	NA	209.8	NSAA	FANCG (Nonsense)	NA	Stanozolol	NA	SS	MCH, hypertelorism	Skin discolouration	NA	NA
O-123	11 / F	NA	NA	NA	0	NSAA	FANCL (Exonic splice variant)	NA	Danazol	No	BT	MCH	POH, HPOP, CAL	NA	NA

FP-97-Q-454/FA 33/22	21 / F	NA	NA	NA	0	NSAA	FANCL (Exonic splice variant)	NA	Danazol	Yes	Nail dystrophy	NA	HPOP, HP, leukoplakia	NA	NA
FP-93-Q-379/FA 25/22	9 / M	NA	NA	NA	84.3	SAA	FANCG (Nonsense)	NA	Stanozolol	No	Polydactyly	Depressed nasal bridge	NA	Patent Ductus Arteriosus	NA

Skeletal abnormalities: SS- Short stature; TH- Thumb hypoplasia; BT- Bifid thumb; SST- Short stubby toes; TA- Thumb aplasia; CTEV- Congenital talipes equinovarus; HSK- Horse shoe kidney; DPT- Distally placed thumb; ARP- absent radial pulse; SD- Sprengel's deformity

Facial Dysmorphism: DF- Dysmorphic Face; MCH- Microcephaly; MGH- Micrognathia; AE- Almond shaped eyes; Low set ears- LSE; High arched palate- HAP; Arched palate- AP; Elfin Facies- EF; TF- Triangular facies; Microopthalamia- MOA; ASE- Antemongoloid slant of eyes.

Skin Changes: HP- Hyperpigmentation; HPOP- Hypopigmentation; CAL- Café au lait; COP- Circum oral pigmentation; POH- Perioral hypopigmentation

Other anomalies: HSK- Horse-shoe kidney

Table A2: Disease associated genotypes identified by exome sequencing in 146 out of 153 FA patients.

Sample ID	**Gene**	**DNA change**	**Type of mutation**	**cDNA change**	**Amino acid change**	**dbSNP ID**	**Zygosity**	**Varsome Results**	**EVE Results**
FA-01	*FANCG*	NC_000009.11:g.35074215T>G	3' splice site mutation	NM_004629.2:c.1761-2A>C	-	rs765150956	Hom	Pathogenic	NA
FA-02	*FANCJ (BRIP1)*	NC_000017.10:g.59857679T>A	Missense	NM_032043.3:c.1878A>T	p.Glu626Asp	rs1567812484	Hom	Pathogenic	Pathogenic

FA-03	*FANCA*	NC_000016.9:g.89813237A>C	5' splice site mutation	NM_000135.4:c.3408+2T>G	-	Novel	CH	VUS	NA
		NC_000016.9:g.89828423T>G	Missense	NM_000135.4:c.2786A>C	p.Tyr929Ser	Novel		VUS	Pathogenic
FA-04	*FANCA*	NC_000016.9:g.89882396T>C	3' splice site mutation	NM_000135.4:c.80-2A>G	-	Novel	CH	VUS	NA
		NC_000016.9:g.89811367_89865640del	Deletion	NM_000135.4:c.827_3626del	-	Novel		Likely pathogenic	NA
FA-05	*FANCA*	NC_000016.9:g.89857866C>T	Missense	NM_000135.4:c.1304G>A	p.Arg435His	rs1060501879	Hom	Pathogenic	Pathogenic
FA-06	*FANCA*	NC_000016.9:g.89828423T>G	Missense	NM_000135.4:c.2786A>C	p.Tyr929Ser	Novel	Hom	Likely pathogenic	Pathogenic
FA-07	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-08	*FANCA*	NC_000016.9:g.89805365dup	Frameshift insertion	NM_000135.4:c.4185dupG	p.Ile1396AspfsTer29	Novel	Hom	Pathogenic	NA
FA-09	*FANCA*	NC_000016.9:g.89877340_89877341del	5' splice site mutation	NM_000135.4:c.426_426+1delAG	K143Rfs*7/5' splice variant	rs763114336	CH	VUS	NA
		NC_000016.9:g.89877448del	Frameshift deletion	NM_000135.4:c.319delG	p.Val107PhefsTer31	rs1411237340		VUS	NA
FA-10	*FANCA*	NC_000016.9:g.89809218_89809219del	Frameshift deletion	NM_000135.4:c.3761_3762delAG	p.Glu1254GlyfsTer23	rs868273545	Hom	Pathogenic	NA
FA-11	*FANCA*	NC_000016.9:g.89811367_89816310del	Deletion	NM_000135.4:c.3067_3626del	-	Novel	Hom	Pathogenic	NA
FA-12	*FANCA*	NC_000016.9:g.89809286_89809290dup	Frameshift insertion	NM_000135.4:c.3690_3694dupGCACT	p.Phe1232CysfsTer17	Novel	CH	Likely pathogenic	NA
		NC_000016.9:g.89857867G>A	Missense	NM_000135.4:c.1303C>T	p.Arg435Cys	rs148473140		Likely pathogenic	Pathogenic
FA-13	*FANCG*	NC_000009.11:g.35077267_35077273del	Frameshift deletion	NM_004629.2:c.637_643delTACCGCC	p.Tyr213LysfsTer6	rs587776640	Hom	Pathogenic	NA
FA-15	*FANCA*	NC_000016.9:g.89851261C>T	Exonic splice donor variant	NM_000135.4:c.1470+1G>A	p.Lys369Lys	rs1555556175	Hom	Pathogenic	NA
FA-16	*FANCA*	NC_000016.9:g.89833551_89833647del	Deletion	NM_000135.4:c.2505_2601del	p.Lys835SerfsTer22	George et al. 2021	Hom	Pathogenic	NA

FA-18	*FANCA*	NC_000016.9:g.89828357C>G	Missense	NM_000135.4:c.2852G>C	p.Arg951Pro	Novel	Hom	Likely pathogenic	Pathogenic
FA-19	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-20	*FANCG*	NC_000009.11:g.35074215T>G	3' splice site mutation	NM_004629.2:c.1761-2A>C	-	rs765150956	Hom	Pathogenic	NA
FA-21	*FANCA*	NC_000016.9:g.89807212_898163110del	Deletion	NM_000135.4:c.3067_3828del	-	Novel	Hom	Pathogenic	NA
FA-22	*FANCA*	NC_000016.9:g.89866011A>G	5' splice site mutation	NM_000135.4:c.826+2T>C	-	Novel	Hom	Pathogenic	NA
FA-23	*FANCA*	NC_000016.9:g.89818545C>A	5' splice site mutation	NM_000135.4:c.3066+1G>T	-	rs587783028	Hom	Pathogenic	NA
FA-24	*FANCA*	NC_000016.9:g.89809302C>T	Nonsense	NM_000135.4:c.3671G>A	p.Trp1224Ter	Novel	CH	Likely pathogenic	NA
		NC_000016.9:g.89828358_898422241del	Deletion	NM_000135.4:c.1827_2852del	-	Castella et al. 2011		Likely pathogenic	NA
FA-25	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-26	*FANCG*	NC_000009.11:g.35076427_35076431del	5' intronic splice donor variant (+2)	NM_004629.2:c.1076+3_1076+7delGAGGT	-	rs780410457	Hom	VUS	NA
FA-27	*FANCA*	NC_000016.9:g.89818545C>A	5' splice site mutation	NM_000135.4:c.3066+1G>T	-	rs587783028	Hom	Pathogenic	NA
FA-28	*FANCA*	NC_000016.9:g.89882944C>G	5' splice site mutation	NM_000135.4:c.79+1G>C	-	rs1483028018	Hom	Pathogenic	NA
FA-30	*FANCF*	NC_000011.9:g.22646233T>C	Stop loss	NM_022725.4:c.1124A>G	p.Ter375TrpextTer33	Novel	Hom	VUS	NA
FA-31	*FANCA*	NC_000016.9:g.89882944C>G	5' splice site mutation	NM_000135.4:c.79+1G>C	-	rs1483028018	Hom	Pathogenic	NA
FA-32	*FANCA*	NC_000016.9:g.89824984C>G	5' splice site mutation	NM_000135.4:c.2981+1G>C	-	Novel	Hom	Pathogenic	NA
FA-33	*FANCA*	NC_000016.9:g.89813247dup	Frameshift insertion	NM_000135.4:c.3401dupT	p.Phe1135LeufsTer80	Novel	Hom	Pathogenic	NA
FA-34	*FANCA*	NC_000016.9:g.89818545C>A	5' splice site mutation	NM_000135.4:c.3066+1G>T	-	rs587783028	Hom	Pathogenic	NA

FA-35	*FANCA*	NC_000016.9:g.89806416del	Frameshift deletion	NM_000135.4:c.3920delA	p.Gln1307ArgfsTer2	rs1228394297	Hom	Pathogenic	NA
FA-36	*FANCG*	NC_000009.11:g.35075650_35075651del	Frameshift deletion	NM_004629.2:c.1246_1247delCT	p.Leu416MetfsTer2	Novel	Hom	Pathogenic	NA
FA-38	*FANCG*	NC_000009.11:g.35078223A>G	Missense	NM_004629.2:c.425T>C	p.Leu142Pro	Novel	CH	VUS	Pathogenic
		NC_000009.11:g.35076856_35076857del	Frameshift deletion	NM_004629.2:c.792_793delAG	p.Arg264SerfsTer24	Novel		VUS	NA
FA-40	*FANCA*	NC_000016.9:g.89877448del	Frameshift deletion	NM_000135.4:c.319delG	p.Val107PhefsTer31	rs1411237340	Hom	Pathogenic	NA
FA-511	*FANCA*	NC_000016.9:g.89831327G>A	Nonsense	NM_000135.4:c.2749C>T	p.Arg917Ter	rs1060501880	Hom	Pathogenic	NA
FA-513	*FANCA*	NC_000016.9:g.89880929_89883055del	Deletion	NM_000135.4:c.-31_283del	-	Savoia et al. 1996	Hom	Pathogenic	NA
FA-516	*FANCA*	NC_000016.9:g.89825022del	Frameshift deletion	NM_000135.4:c.2944delA	p.Thr982ProfsTer7	Novel	Hom	Pathogenic	NA
FA-517	*FANCC*	NC_000009.11:g.97873912C>A	Nonsense	NM_000136.3:c.1162G>T	p.Gly388Ter	rs371897078	Hom	Pathogenic	NA
FA-521	*FANCA*	NC_000016.9:g.89877481T>G	3' splice site mutation	NM_000135.4:c.284-2A>C	-	rs756023006	Hom	Pathogenic	NA
FA-522	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-523	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-525	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-527	*FANCC*	NC_000009.11:g.97864081T>G	Missense	NM_000136.3:c.1585A>C	p.Thr529Pro	rs587778326	Hom	VUS	Pathogenic
FA-529	*FANCA*	NC_000016.9:g.89833551_89833647del	Deletion	NM_000135.4:c.2505_2601del	p.Lys835SerfsTer22	George et al. 2021	Hom	Pathogenic	NA
FA-531	*FANCG*	NC_000009.11:g.35076026C>G	3' splice site mutation	NM_004629.2:c.1077-1G>C	-	Novel	Hom	Pathogenic	NA
FA-532	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA

FA-533	*FANCA*	NC_000016.9:g.89858399_89858400del	Frameshift deletion	NM_000135.4:c.1164_1165delAG	p.Arg388SerfsTer20	Novel	CH	VUS	NA
		NC_000016.9:g.89806402C>T	Missense	NM_000135.4:c.3934G>A	p.Asp1312Asn	Novel		VUS	Pathogenic
FA-534	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-535	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-538	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-542	*FANCA*	NC_000016.9:g.89866011A>G	5' splice site mutation	NM_000135.4:c.826+2T>C	-	Novel	Hom	Pathogenic	NA
FA-543	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-544	*FANCA*	NC_000016.9:g.89858878C>T	5' splice site mutation	NM_000135.4:c.1083+1G>A	-	Novel	Hom	Pathogenic	NA
FA-547	*FANCA*	NC_000016.9:g.89845356G>A	Nonsense	NM_000135.4:c.1771C>T	p.Arg591Ter	rs753980264	Hom	Pathogenic	NA
FA-548	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-549	*FANCA*	NC_000016.9:g.89828423T>G	Missense	NM_000135.4:c.2786A>C	p.Tyr929Ser	Novel	Hom	Likely pathogenic	Pathogenic
FA-554	*FANCA*	NC_000016.9:g.89857810C>G	5’ Splice site mutation	NM_000135.4:c.1359+1G>C	-	rs1555561294	CH	Likely pathogenic	NA
		NC_000016.9:g.89803957_89806507del	Deletion	NM_000135.4:c.3829_*1052del	-	Novel		Likely pathogenic	NA
FA-556	*FANCA*	NC_000016.9:g.89806405del	Frameshift deletion	NM_000135.4:c.3931delA	p.Ser1311ValfsTer52	Novel	Hom	Pathogenic	NA
FA-557	*FANCA*	NC_000016.9:g.89882943A>G	5' splice site mutation	NM_000135.4:c.79+2T>C	-	rs1319795682	Hom	Pathogenic	NA
FA-561	*FANCA*	NC_000016.9:g.89833551_89833647del	Deletion	NM_000135.4:c.2505_2601del	p.Lys835SerfsTer22	George et al. 2021	Hom	Pathogenic	NA
FA-564	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA

FA-565	*FANCA*	NC_000016.9:g.89883009C>T	Nonsense	NM_000135.4:c.15G>A	p.Trp5Ter	Novel	Hom	Pathogenic	NA
FA-568	*FANCA*	NC_000016.9:g.89828423T>G	Missense	NM_000135.4:c.2786A>C	p.Tyr929Ser	Novel	Hom	Likely pathogenic	Pathogenic
FA-629/18	*FANCG*	NC_000009.11:g.35075954C>G	5' intronic splice donor variant (+5)*	NM_004629.2:c.1143+5G>C	-	rs778328620	Hom	Pathogenic	NA
FA-631/18	*FANCA*	NC_000016.9:g.89831438G>A	Nonsense	NM_000135.4:c.2638C>T	p.Arg880Ter	rs762804216	CH	Pathogenic	NA
		NC_000016.9:g.89811367_898163 10del	Deletion	NM_000135.4:c.3067_3626del	-	Novel		Pathogenic	NA
FA-636/18	*UBE2T/FANCT*	NC_000001.10:g.202300964_20230 1088del	Deletion	NM_014176.4:c.470_594del	p.Ala157GlyfsTer18	Novel	Hom	Pathogenic	NA
FA-637/18	*FANCJ (BRIP1)*	NC_000017.10:g.59885995G>A	Missense	NM_032043.3:c.751C>T	p.Arg251Cys	rs752309409	Hom	Likely pathogenic	Pathogenic
FA-638/18	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-641/18	*FANCF*	NC_000011.9:g.22647316A>C	Missense	NM_022725.4:c.41T>G	p.Leu14Arg	Novel	Hom	VUS	VUS
FA-646/18	*FANCA*	NC_000016.9:g.89828357C>T	Missense	NM_000135.4:c.2852G>A	p.Arg951Gln	rs755922289	Hom	Likely pathogenic	Pathogenic
FA-650/18	*FANCG*	NC_000009.11:g.35079204dup	Frameshift insertion	NM_004629.2:c.119dupA	p.Gln41AlafsTer16	Novel	Hom	Pathogenic	NA
FA-652/18	*FANCA*	NC_000016.9:g.89816214G>A	Missense	NM_000135.4:c.3163C>T	p.Arg1055Trp	rs753063086	Hom	Likely pathogenic	Pathogenic
FA-659/18	*FANCA*	NC_000016.9:g.89809284A>G	Missense	NM_000135.4:c.3689T>C	p.Leu1230Pro	Novel	Hom	VUS	Pathogenic
FA-660/18	*FANCA*	NC_000016.9:g.89865641C>T	3' splice site mutation	NM_000135.4:c.827-1G>A	-	rs753728435	Hom	Pathogenic	NA
FA-665/18 FA-12/22	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-672/18	*FANCA*	NC_000016.9:g.89811367_898163 10del	Deletion	NM_000135.4:c.3067_3626del	-	Novel	Hom	Pathogenic	NA

FA-674/18	*FANCA*	NC_000016.9:g.89816214G>A	Missense	NM_000135.4:c.3163C>T	p.Arg1055Trp	rs753063086	CH	Pathogenic	Pathogenic
		NC_000016.9:g.89811367_898656 40del	Deletion	NM_000135.4:c.827_3626del	-	Novel		Pathogenic	NA
FA-675/18	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	CH	Pathogenic	NA
		NC_000002.11:g.58388695G>A	Nonsense	NM_001114636.1:c.997C>T	p.Gln333Ter	rs776298788		Pathogenic	NA
FA18/20	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA01/19	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FP-23-P-11	*FANCA*	NC_000016.9:g.89828423T>G	Missense	NM_000135.4:c.2786A>C	p.Tyr929Ser	Novel	Hom	VUS	Pathogenic
FA-17/19	*FANCA*	NC_000016.9:g.89871689_898718 01del	Deletion	NM_000135.4:c.597_709del	p.Ser199GlyfsTer 24	Esmail nia et al. 2016	Hom	Pathogenic	NA
FA-18/19	*FANCA*	NC_000016.9:g.89871689_898718 01del	Deletion	NM_000135.4:c.597_709del	p.Ser199GlyfsTer 24	Esmail nia et al. 2016	Hom	Pathogenic	NA
FA-20/19	*FANCA*	NC_000016.9:g.89877481T>G	3' splice site mutation	NM_000135.4:c.284-2A>C	-	rs756023006	Hom	Pathogenic	NA
FA-27/19	*FANCA*	NC_000016.9:g.89866011A>G	5' splice site mutation	NM_000135.4:c.826+2T>C	-	Novel	Hom	Pathogenic	NA
FA-30/19	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Likely pathogenic	NA
FA-36/19	*FANCG*	NC_000009.11:g.35074215T>G	3' splice site mutation	NM_004629.2:c.1761-2A>C	-	rs765150956	Hom	Pathogenic	NA
FA-37/19	*FANCG*	NC_000009.11:g.35074215T>G	3' splice site mutation	NM_004629.2:c.1761-2A>C	-	rs765150956	Hom	Pathogenic	NA
FA-49/19	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-573	*FANCA*	NC_000016.9:g.89877448del	Frameshift deletion	NM_000135.4:c.319delG	p.Val107PhefsTer31	rs1411237340	CH	Pathogenic	NA
		NC_000016.9:g.89883022A>T	Missense	NM_000135.4:c.2T>A	p.Met1Lys	rs769479800		Pathogenic	NA
FA-580	*FANCG*	NC_000009.11:g.35078222_350782 23delinsAGCAGTT	Frameshift insertion	NM_004629.2:c.425_426delTGinsAACTGCT	p.Leu142GlnfsTer12	Novel	Hom	Pathogenic	NA

FA-581	*FANCA*	NC_000016.9:g.89809216_89809219del	Frameshift deletion	NM_000135.4:c.3759_3762del	p.Glu1254SerfsTer11	Novel	Hom	Pathogenic	NA
FA-584	*FANCA*	NC_000016.9:g.89877327_89877344del	Frameshift deletion	NM_000135.4:c.423_426+14del	-	Novel	Hom	Pathogenic	NA
FA-588	*FANCG*	NC_000009.11:g.35074215T>G	3' splice site mutation	NM_004629.2:c.1761-2A>C	-	rs765150956	Hom	Pathogenic	NA
FA-591	*FANCA*	NC_000016.9:g.89866011A>G	5' splice site mutation	NM_000135.4:c.826+2T>C	-	Novel	Hom	Pathogenic	NA
FA-595	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-598	*FANCG*	NC_000009.11:g.35075059G>A	Nonsense	NM_004629.2:c.1501C>T	p.Gln501Ter	Novel	Hom	Pathogenic	NA
FA-599	*FANCA*	NC_000016.9:g.89877327_89877344del	Frameshift deletion	NM_000135.4:c.423_426+14del	-	Novel	Hom	Pathogenic	NA
FA-601	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-614	*FANCA*	NC_000016.9:g.89849270_89849510del	Deletion	NM_000135.4:c.1472_1624del	-	Neil V morgan et al. 1999	Hom	Pathogenic	NA
O-117	*FANCA*	NC_000016.9:g.89874703_89877479del	Deletion	NM_000135.4:c.285_596del	-	Savoia et al. 1996	Hom	Pathogenic	NA
O-123	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
O-126	*FANCA*	NC_000016.9:g.89862316_89862426del	Deletion	NM_000135.4:c.894_1004del	p.Trp298_Lys335delinsGly	Solanki et al. 2016	Hom	Pathogenic	NA
P-177	*FANCA*	NC_000016.9:g.89813297C>G	Missense	NM_000135.4:c.3350G>C	p.Arg1117Thr	Novel	Hom	Likely pathogenic	Pathogenic
FA-12/19	*FANCA*	NC_000016.9:g.89877448del	Frameshift deletion	NM_000135.4:c.319delG	p.Val107PhefsTer31	rs1411237340	CH	Pathogenic	NA
		NC_000016.9:g.89882396T>C	3' splice site mutation	NM_000135.4:c.80-2A>G	-	Novel		Pathogenic	NA
FA-622	*FANCA*	NC_000016.9:g.89818552_89825113del	Deletion	NM_000135.4:c.2853_3060del	-	Novel	CH	Pathogenic	NA
		NC_000016.9:g.89816231_89816232del	Frameshift deletion	NM_000135.4:c.3146_3147delTT	p.Phe1049Ter	Novel		Pathogenic	NA

FA-593	*FANCA*	NC_000016.9:g.89828358G>A	Missense	NM_000135.4:c.2851C>T	p.Arg951Trp	rs755546887	Hom	Likely pathogenic	Pathogenic
FA-592	*FANCA*	NC_000016.9:g.89845258_89882945del	Deletion	NM_000135.4:c.79_1777del	-	Castella et al. 2011	CH	Pathogenic	NA
		NC_000016.9:g.89833551_89833647del	Deletion	NM_000135.4:c.2505_2601del	p.Lys835SerfsTer22	George et al. 2021		Pathogenic	NA
FA-03/19	*FANCI*	NC_000015.9:g.89804822del	Frameshift deletion	NM_001113378.2:c.295delC	p.His99IlefsTer10	rs759398314	CH	Pathogenic	NA
		NC_000015.9:g.89848828_89848832del	Intronic splice acceptor variant	NM_001113378.2:c.3256-8_3256-4delTCTTT	-	Novel		VUS	NA
P-603	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-04/20	*FANCA*	NC_000016.9:g.89877448del	Frameshift deletion	NM_000135.4:c.319delG	p.Val107PhefsTer31	rs1411237340	Hom	Pathogenic	NA
FA-06/20	*FANCA*	NC_000016.9:g.89805352G>A	Missense	NM_000135.4:c.4198C>T	p.Arg1400Cys	rs745882980	Hom	Pathogenic	Pathogenic
FA-10/20	*FANCA*	NC_000016.9:g.89805352G>A	Missense	NM_000135.4:c.4198C>T	p.Arg1400Cys	rs745882980	Hom	Pathogenic	Pathogenic
FA-10/21	*FANCA*	NC_000016.9:g.89809216_89809219del	Frameshift deletion	NM_000135.4:c.3759_3762del	p.Glu1254SerfsTer11	Novel	Hom	Pathogenic	NA
FA-11/20	*FANCA*	NC_000016.9:g.89806411_89806427dup	Frameshift insertion	NM_000135.4:c.3909_3925dup	p.Thr1309ArgfsTer6	Novel	Hom	Pathogenic	NA
FA-15/20	*FANCG*	NC_000009.11:g.35077289del	Frameshift deletion	NM_004629.2:c.619del	p.Leu207SerfsTer2	Novel	Hom	Pathogenic	NA
FA-16/20	*FANCA*	NC_000016.9:g.89818545C>A	5' splice site mutation	NM_000135.4:c.3066+1G>T	-	rs587783028	Hom	Pathogenic	NA
FA-21/20	*FANCA*	NC_000016.9:g.89805660C>A	Nonsense	NM_000135.4:c.4048G>T	p.Glu1350Ter	Novel	Hom	Pathogenic	NA
FA-5/21	*UBE2T/FANCT*	NC_000001.10:g.202302631T>G	Missense	NM_014176.4:c.232A>C	p.Asn78His	rs776219033	Hom	VUS	Pathogenic
FA-01/20	*FANCG*	NC_000009.11:g.35074476dup	Nonsense	NM_004629.2:c.1652dup	p.Tyr551Ter	Novel	Hom	Pathogenic	Pathogenic
FA-02/21	*FANCA*	NC_000016.9:g.89807252A>G	Missense	NM_000135.4:c.3788T>C	p.Phe1263Ser	George et al. 2021	CH	Likely pathogenic	VUS

		NC_000016.9:g.89849441C>T	Missense	NM_000135.4:c.1540G>A	p.Ala514Thr	Novel		VUS	Pathogenic
FA-08/19	*FANCA*	NC_000016.9:g.89833551_89833647del	Deletion	NM_000135.4:c.2505_2601del	p.Lys835SerfsTer22	George et al. 2021	Hom	Pathogenic	NA
FA-13/19	*FANCA*	NC_000016.9:g.89831446G>C	Nonsense	NM_000135.4:c.2630C>G	p.Ser877Ter	Solanki et al. 2016	Hom	Likely pathogenic	NA
FA-18/21	*FANCA*	NC_000016.9:g.89836600G>A	Missense	NM_000135.4:c.2290C>T	p.Arg764Trp	De Rocco et al. 2014	Hom	Pathogenic	Pathogenic
FA-21/21	*FANCA*	NC_000016.9:g.89816138C>T	Missense	NM_000135.4:c.3239G>A	p.Arg1080Gln	rs1555538571	Hom	Likely pathogenic	Pathogenic
FA-23/19	*FANCA*	NC_000016.9:g.89836973T>G	Exonic splice donor variant	NM_000135.4:c.2221A>C	p.Arg741=	Novel	Hom	Likely pathogenic	NA
FA-26/21	*FANCA*	NC_000016.9:g.89809211_89809212insCT	Frameshift insertion	NM_000135.4:c.3761_3762insAG	p.Glu1255GlyfsTer12	rs868273545	Hom	Pathogenic	NA
FA-30/21	*UBE2T/FANCT*	NC_000001.10:g.202300964_202301088del	Deletion	NM_014176.4:c.470_594del	p.Ala157GlyfsTer18	Novel	Hom	Pathogenic	NA
FA-31/21	*FANCA*	NC_000016.9:g.89851302A>G	Missense	NM_000135.4:c.1430T>C	p.Leu477Ser	Novel	Hom	Pathogenic	Pathogenic
FA-33/19	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-33/21	*FANCA*	NC_000016.9:g.89809211_89809212insCT	Frameshift insertion	NM_000135.4:c.3761_3762insAG	p.Glu1255GlyfsTer12	rs868273545	Hom	Pathogenic	NA
FA-34/19	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FA-35/21	*FANCA*	NC_000016.9:g.89805365dup	Frameshift insertion	NM_000135.4:c.4185dupG	p.Ile1396AspfsTer29	Novel	Hom	Pathogenic	NA
FA-649/18	*BRCA2*	NC_000013.10:g.32893238G>C	Missense	NM_000059.4:c.92G>C	p.Trp31Ser	Caleca et al. 2018	Hom	Likely pathogenic	NA
FA 01/22	FANCD2	NC_000003.11:g.10088404del	Frameshift deletion	NM_033084.6:c.1275del	p.Leu426Ter	Novel	CH	Likely pathogenic	NA
		NC_000003.11:g.10088407_10088408del	Frameshift deletion	NM_033084.6:c.1278_1278+1del	-	Novel		Likely pathogenic	NA
FA 02/22	FANCA	NC_000016.9:g.89816213C>A	Missense	NM_000135.4:c.3164G>T	p.Arg1055Leu	rs1429943036	Hom	Likely pathogenic	Pathogenic

FA 11/22	FANCG	NC_000009.11:g.35076495_35076496dup	Frameshift insertion	NM_004629.2:c.1011_1012dup	p.Leu338ProfsTer14	Novel	Hom	Likely pathogenic	NA
FA 14/22	*FANCA*	NC_000016.9:g.89877115_89877479del	Deletion	NM_000135.2:c.285_522+1del	-	Novel	CH	Pathogenic	NA
		NC_000016.9:g.89882301_89882304del	Frameshift deletion	NM_000135.4:c.172_175delAATG	p.Asn58Profs*6	Novel		Likely pathogenic	NA
FA 23/22	*FANCA*	NC_000016.9:g.89828423T>G	Missense	NM_000135.4:c.2786A>C	p.Tyr929Ser	Novel	Hom	Likely pathogenic	Pathogenic
FA 24/22	*FANCF*	NC_000011.9:g.22646861G>A	Nonsense	NM_022725.4:c.496C>T	p.Gln166Ter	rs766279442	Hom	Likely pathogenic	NA
FA 22/22	*FANCD2*	NC_000003.11:g.10106113G>A	Missense	NM_033084.6:c.2021G>A	p.Gly674Asp	Novel	Hom	Likely pathogenic	Benign
FA 27/22	*FANCG*	NC_000009.11:g.35075032G>A	Nonsense	NM_004629.2:c.1528C>T	p.Gln510Ter	Novel	Hom	Pathogenic	NA
FP-98-Q-471	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FP-97-Q-454/FA 33/22	*FANCL*	NC_000002.11:g.58387243C>T	Exonic splice donor variant	NM_001114636.1:c.1107G>A	p.Lys369(=)	rs577063114	Hom	Pathogenic	NA
FP-93-Q-379/FA 25/22	*FANCG*	NC_000009.11:g.35075554G>C	Nonsense	NM_004629.2:c.1341C>G	p.Tyr447Ter	rs778986343	Hom	Pathogenic	NA

NA: Not applicable as EVE is available only for missense mutations

* Likely pathogenic by ACMG guidelines but pathogenic by Varsome.