Fanconi anemia: A model disease for studies on human genetics and advanced therapeutics

Massimo Bogliolo and Jordi Surrallés*

Genome Instability and DNA Repair Group, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain and Centre for Biomedical Network Research on Rare Diseases (CIBERER).

*Correspondence should be addressed to Prof. Dr. Jordi Surrallés, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Campus de Bellaterra S/N, Bellaterra (Barcelona), Spain. Tel: +345811830. Fax: +34935812387; e-mail: jordi.surralles@uab.cat
Abstract

Fanconi anemia (FA) is characterized by bone marrow failure, defects in development, and chromosome fragility. We review the recent discovery of FA genes and efforts to develop genetic therapies for FA. Because genetic evidence excludes *FANCM* as an FA gene, 14 genes remain *bona fide* FA genes and 2 (FANCO and FANCS) cause an FA like syndrome. Monoallelic mutations in 6 *FANC* genes predispose to breast and ovarian cancer. FANC proteins repair stalled DNA replication forks by unhooking DNA interstrand cross-links and promoting homologous recombination. The genetic characterization of patients with FA is essential for developing therapies, including hematopoietic stem cell transplantation from a savior sibling donor after embryo selection, untargeted gene therapy, targeted gene therapy, or genome editing using genetic recombination or engineered nucleases. Newly acquired knowledge about FA promises to provide a cure in the near future.
Introduction: Fanconi anemia and Fanconi anemia-like genes

Fanconi anemia (FA), which affects approximately 1–3 of 500,000 newborns, causes bone marrow failure (BMF), developmental defects, and cancer predisposition. Hallmarks of FA are chromosome fragility and hypersensitivity to drugs that induce DNA interstrand cross-links (ICLs). Numerous other physiological and cellular abnormalities likely contribute to pathogenesis (Figure 1). The current decade has been prolific for the discovery of novel FA genes. Thus, seventeen genes are associated to FA, including the recently discovered genes FANCO/RAD51C[1], FANCP/SLX4[2,3], FANCO/ERCC4[4], and FANCS/BRCA1[5]. However, following the stringent criteria of at least 2 patients with BMF and a positive chromosome fragility test, only 14 met the criteria for bona fide FA genes (FANCA, B, C, D1, D2, E, F, G, I, J, L, N, P, Q). FANCO and S are FA-like genes because they cause a chromosome fragility syndrome with FA-related malformations but without BMF (Figure 2). FANCM should also be excluded from the list of FA genes, because the only patient known to carry biallelic mutations in FANCM[6], also carries biallelic pathogenic FANCA mutations, and her brother was subtyped FANCA [7]. Further, enforced expression of FANCM failed to complement the genetic defect of this patient’s cells [6]. Moreover, pathogenic FANCM variants are more common than originally predicted in some populations, and two Finns with homozygous loss-of-function FANCM mutations exhibit normal hematology [8]. Thus, we recommend excluding FANCM as an FA gene, although, together with FAAP100, FAAP25, and other FA-core complex interacting proteins, FANCM is involved in the FA ICL repair (ICLR).
pathway (see below). Similarly, whole genome exon sequencing (WES) detected biallelic XRCC2 mutations in a consanguineous FA family [9]; however, because of the lack of genetic complementation data or any other functional evidence of a causative role of this homozygous mutation in disease, XRCC2 should not be considered an FA gene. Further, this patient may harbor mutations in known FA genes that are not easily detected by WES, such as large deletions or deep intronic mutations.

The genetic heterogeneity and the number of private and founder mutations makes the mutational analysis of FA patients extremely difficult [10-12]. However, the implementation of next-generation sequencing (NGS) technologies, including WES or targeted sequencing of FA genes, together with high-resolution methods to detect large deletions, such as comparative genome hybridization arrays, single-nucleotide polymorphism arrays, and targeted Multiplex Ligation-dependent Probe Amplification will facilitate subtyping and mutational analysis of new patients with FA [4,13-17].

FA genes that predispose to breast and ovarian cancer

FANCD1/BRCA2, FANCS/BRCA1, FANCJ/BRIP1, FANCM, FANCN/PALB2, and FANCO/RAD51Care major breast and ovarian cancer susceptibility genes in carriers with monoallelic mutations (Figure 1) highlighting the fundamental link between FA and familial breast and ovarian cancer (FBOC). Rad51C mutations influence ovarian cancers more than breast cancer [18,19], and are linked to other
tumors such as head and neck cancer [20,21]. RAD51C and FANCM were initially associated to FA before they were candidates for FBOC in monoallelic carriers [22,23] highlighting the role of FA research in molecular oncology.

Historically, only FA genes (see below; Figure 2) with a direct role in the homologous recombination repair (HRR) branch of the FA pathway are linked to FBOC [24,25]. The two recently identified FA genes upstream of those encoding HRR components, \textit{FANCP/SLX4} and \textit{FANCQ/ERCC4}, were also excluded as major breast cancer susceptibility genes in Italian, German, Spanish, Estonian, Jewish, and non-Jewish American populations [26-33]. However, a pathogenic mutation in the HRR upstream gene \textit{FANCM} is associated with breast cancer susceptibility in the Finnish population [23], suggesting a core complex-independent role (see below) of FANCM. In fact, even that FANCM is not essential for Rad51 foci formation and HRR, the camptothecin sensitivity of FANCM cells is shared with and FANCD1 and FANCN cells, linking FANCM to the branch of the FA pathway connected to HRR[7].

\textbf{The Fanconi anemia pathway: find, unhook, bypass, and recombine}

ICLs are highly damaging, because they impede transcription and replication-fork progression. Since they affect both DNA strands, ICLs complicate error-free DNA repair, because an undamaged DNA template is not available. The FA DNA repair pathway coordinates reactions that remove ICL damage to restore genome
integrity (for more detailed reviews, see [34,35]. Eight FA proteins (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, and FANCM) form a nuclear complex (FANCore) whose ubiquitin E3 ligase function is activated by blocked DNA replication forks (Figure 3). The activated FANCore complex monoubiquitinates the FANCD2–FANCI heterodimer (ID complex) [36,37] and the activated ID complex relocates to the damaged DNA in an ATR and BRCA1-dependent manner [36-38]. The ID complex promotes nucleolytic cleavage of the 3´ and 5´ sites of DNA to unhook the ICL and successively induces trans-lesion polymerases Rev1 and pol[39-42]. These reactions extend the leading DNA strand above and past the unhooked ICL to produce a substrate that is processed by successive HRR reactions [34].

In the last few years, major advances in understanding the unhooking step of ICL repair came from the discovery of the FA genes SLX4/FANCP and ERCC4/FANCO. Current models predict that DNA integrity is not restored without cleaving the 3´ and 5´ sequences flanking the lesion. Several nucleases contribute to ICLR, such as XPF-ERCC1, MUS81-EME1, SLX1, SNM1A, and SNM1B [43-45]. However, the lack of FA patients with mutations in the genes encoding these nucleases hindered the identification of the main FA/BRCA pathway nuclease. BTBD12/SLX4 provides a platform for several endonucleases involved in ICL repair, including ERCC4-ERCC1, MUS81-EME1, and SLX1 that dock to cleave DNA flaps, replication forks, and Holliday junctions [46-48]. SLX4 is mutated in patients with bona fide FA (FANCP) [2,20]. In contrast to FA mouse models, the Slx4 knockout (KO) mouse exhibits an FA phenotype with developmental defects
and cytopenia[49]. The N-terminal segment of SLX4/FANCP harbors ubiquitin zinc finger (UBZ) domains, suggesting that SLX4 is recruited to DNA damage via interaction with ubiquitinated proteins involved in the DNA damage response (DDR). SLX4 recruitment in chicken DT-40 cells may depend on FANCD2 monoubiquitination but with uncertain relevance to mammals [50]. SLX4 serves as a scaffold that organizes specific nucleases for transport to DNA lesions and regulates nuclease activity [50,51]. SLX4 contains SUMO-interacting motifs (SIMs) required for binding sumoylated DNA repair proteins, and SLX4 acts directly or indirectly as a SUMO E3 ligase, and its SUMO-related functions are not required for ICL repair but for a general response to replication stress. Accordingly, mutations of SLX4 SIMs do not produce ICLs hypersensitivity but cause common fragile site instability and increased mitotic catastrophe [52,53].

ERCC4 mutations were identified using WES and Sanger sequencing in two unrelated and unassigned FA patients [4]. Both patients had characteristic FA symptoms, including BMF, chromosome fragility, and FA-related birth defects; ERCC4 was therefore renamed FANCQ [4]. ERCC4–ERCC1 is a heterodimeric endonuclease discovered as an essential component of the nucleotide excision repair (NER) system. The catalytic subunit ERCC4 was originally renamed XPF, because it is mutated in patients with xeroderma pigmentosum (XP), complementation group F [54]. In contrast to other XP-related proteins, defects in XPF sensitize cells to UV light and ICLs, indicating an NER-independent DNA repair role [55,56]. The ERCC4/FANCQ mutations uncouple NER and ICLR functions as follows: mutations that inhibit ICLR but not NER activity of
ERCC4/FANCQ caused FA. Moreover, FA patients lacked XP symptoms, consistent with functional NER despite ERCC4/FANCQ defects [4].

The Fanconi anemia-associated nuclease 1 (FAN1) was identified in 2010 and immediately stole the scene as the main candidate FA nuclease [57-60]. FAN1 is a structure-specific nuclease that, when depleted, causes sensitivity to ICLs. Moreover, FAN1 is recruited to damaged DNA via the interaction through its N-terminal UBZ domain with monoubiquitinated FANCD2. Recruitment of FAN1 through FANCD2 monoubiquitination is consistent with the inhibition of nucleolytic incisions near the ICL in vitro upon FANCD2 depletion [41]. Although FAN1 interacts with FANCD2 and contributes to the FA/BRCA pathway of ICLR [61-64], FAN1 deficiency is not associated with an FA phenotype [65]. Instead, biallelic lack-of-function point mutations in FAN1 cause karyomegalic interstitial nephritis, linking chronic kidney failure to ICLR [66]. Further, in contrast to ERCC4-ERCC1, FAN1 activity (similarly to MUS81-EME1 activity) is not required for nucleolytic incisions near an ICL [67]. Therefore, the role of FAN1 nuclease activity in the FA pathway is still under discussion.

**Fanconi anemia and endogenous aldehydes**

Groundbreaking series of studies led by Dr. KJ Patel’s team shed a light on understanding the source of DNA damage that cause FA as they genetically demonstrate that the FA DNA repair pathway counteracts the genotoxicity of endogenous aldehydes [68]. Aldh2, which mediates the metabolism of
acetaldehyde, is essential for embryonic development of FA mice, consistent with findings using DT40 cells, as chicken ALDH5 and the FA pathway mediate synthetic lethality [68,69]. Aldh2+/−Fancd2+/−KO mice are not viable if the mother lacks functional Aldh2. When the mother is Aldh2+/−, Aldh+/−Fancd2+/− offspring are viable, demonstrating that maternal aldehyde catabolism rescues embryo lethality.

HSCs of viable Aldh+/−Fancd2+/− mice spontaneously suffer increased DNA damage, and the mice develop acute leukemia and aplastic anemia. Further, there is a 600-fold reduction in the HSC population in mice deficient in the FA DNA repair pathway and the detoxification of acetaldehyde [70]. When similar experiments are performed by knocking out Fanca, which is mutated in the majority of FA patients, the phenotype is even worse. Aldh+/−Fanca−/− embryos do not develop if the mothers are Aldh2+/−, and the embryos of Aldh2+/− mothers develop but die before birth. When Aldh+/−Fanca−/− embryos are transferred to Aldh2+/+ mothers they result in viable offspring but neonates have low numbers of hematopoietic stem and progenitor cells, indicating that fetal Aldh2 is essential for proper hematopoiesis[71]. These findings are relevant to Asian populations, particularly that of Japan with a 40% carrier frequency of a dominant-negative ALDH2 allele. In Japanese patients with FA that bear the dominant-negative ALDH2 allele, the FA phenotype is more severe with earlier onset of BMF and increased FA-related birth defects[72], which provides strong evidence that mouse data may apply to humans. It remains to be determined whether targeting aldehyde metabolism ameliorates BMF and cancer predisposition of patients with FA [73-75]. ALDH2-deficient Japanese who consume alcohol are at higher risk of developing esophageal cancer, macrocytosis, and macrocytic anemia and oral
microorganisms produce high levels of acetaldehyde in saliva[76], suggesting a modality to prevent BMF as well as head and neck cancer in FA patients.

**Novel therapies: from genes to patients**

The discovery of innovative therapies highlights the pioneering role of FA translational research in the history of medicine. Eliane Gluckman, one of the leading hematologists worldwide, performed the first umbilical cord-blood (UCB) transplant in Paris to cure a patient with FA [77]. Cord blood banking is now performed worldwide, and more than 30,000 patients with blood disorders and other diseases benefited from this source of blood progenitor cells [78]. The outcome of transplanting patients with FA using HLA-matched unrelated HSC donors doubled because of drugs such as fludarabine and improved protocols, approaching the excellent survival rates using HLA-matched donor siblings [79,80]. The first preimplantation genetic diagnosis combined with HLA-matching, generated a savior baby to cure a sibling with FA using UBC transplantation [81]. Hundreds of children with numerous blood disorders were subsequently cured [82,83]. However, FA families should be informed of its low success rate (<5% of babies born per in vitro fertilization cycle) due to Mendelian restrictions and high aneuploidy rates associated with advanced maternal age [84].

Unfortunately, HSC donors are not available for all patients with FA, and HSCT increases further their high cancer risk. To overcome these limitations, FA gene therapy clinical trials are in progress [85]. Difficulties in collecting sufficient blood
progenitor cells from patients with FA and inefficient transduction protocols with first-generation retroviral vectors led to unsuccessful initial clinical trials [86-88]. Improvements may come from a safer and more efficient lentiviral vector expressing human FANCA using the weak PGK promoter [89,90] developed by Bueren’s laboratory, which was designed as an orphan drug by the European Medicines Agency. Drugs such as plerixafor efficiently mobilize HSCs for apheretic collection and this approach can be useful in patients with FA. Studies of mosaic patients and mice with FA after ex vivo gene therapy indicate a survival advantage of genetically corrected cells in vivo, even in the absence of myeloablative conditioning regimens [91,92]. These measures, together with the expected lack of graft-versus-host disease after gene therapy, may cure BMF of patients with FA and prevent HSCT-related cancers.

FA research also played a pioneering role in the field of regenerative medicine (Figure 4). Disease-free blood progenitor cells were first generated from the skin of a patient with FA via induced pluripotent stem (IPS) cells [93]. Ultimately, sufficient IPS cell-derived HSCs must be generated for autotransplantation. Because FA fibroblasts are difficult to reprogram, this study uncovered a novel role of the FA pathway in cell reprogramming. Consequently, correcting FA genes restores the reprogramming efficiency of FA fibroblasts to IPS cells [93-98]. Reprogramming induces the DDR [99] and activates the FA pathway [94], leading to P53-mediated apoptosis and low reprogramming efficiency [96]. This is partially circumvented by preventing reactive oxygen species (ROS)-mediated DNA damage [94] or by suppressing P53 during reprogramming using RNA interference [96] or human
papillomavirus P53-repressing E6 protein [100]. Safe and controlled FANCA gene correction was achieved using integration-free genome editing by genetic recombination with helper-dependent adenoviral vectors [96] or by targeting FANCA insertion into the safe locus AAVS1[97] using engineered nucleases (“safe harbor” strategy) [101]. Genome editing using CRISPR/Cas9-engineered nucleases corrects FANCA mutations in human FA fibroblasts [102]. Although clinical translation of gene-corrected IPS cells and genome editing with engineered nucleases is difficult in the short term, successful engraftment of iPSC-derived and gene-corrected blood progenitor cells may cure FA and other blood disorders characterized by low numbers of bone marrow HSCs [97,101]. Therefore, editing fibroblasts and IPS-cell genomes may soon translate directly to HSCs from FA patients (Figure 4a). FA-IPS cells offer a novel tool to model FA physiology and pathogenesis and provide a cell platform for drug screening [96]. Finally, therapies designed to enhance the correct mRNA processing at a mutant TT splice donor in FANCC using suppressor U1 snRNAs, suggests that correcting pathological mRNA processing at specific mutant splice sites might apply to FA complementation groups in a mutation-specific fashion [103]. Therefore, our better understanding of the molecular genetic defects of FA and the use of gene correction strategies may contribute to futures treatments for this devastating disease.

Acknowledgments
This review article is dedicated to the memory of Prof. Johan de Winter (VUMC-Amsterdam) whose contribution to the genetics of Fanconi anemia will last forever. Surrallés’ laboratory is currently funded by the Generalitat de Catalunya (SGR0489-2009; SGR317-2014), the ICREA-Academia program, the Marató de TV3 (project 464/C/2012), the Spanish Ministry of Science and Innovation (projects CB06/07/0023 and SAF2012-31881), the European Commission (EUROFANCOLEN project HEALTH-F5-2012-305421) and the European Regional Development FEDER Funds. CIBERER is an initiative of the Instituto de Salud Carlos III, Spain.

Conflicts interest

The authors declare no conflicts of interest
References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest


**this paper first describes homozygous missense mutations in the RAD51C gene in a consanguineous family with an FA-like phenotype.


**Articles 2 and 3 independently described several FA individuals with pathogenic biallelic mutations in SLX4


**This study showed that a subset of mutations in ERCC4 specifically disrupts the function of XPF in ICLR causing FA subtype Q.


**This paper described that biallelic mutations in BRCA1 cause an FA-like phenotype with chromosome fragility and malformation but without BMF.


**This paper described that biallelic mutations in BRCA1 cause an FA-like phenotype with chromosome fragility and malformation but without BMF.


*This paper described non-FA individuals with biallelic loss-of-function mutations in FANCM, definitively excluding FANCM as an FA causing gene.


**Once implicated in FA, RAD51C was described in this study to cause breast and ovarian cancer predisposition in monoallelic mutation carriers**

*This paper implicated FANCM in breast cancer predisposition*


*This study is of particular interest in the field as it described the first FA mouse model that recapitulates the FA clinical phenotype*


**This paper presented a forks collision model of ICLR pathway at a single nucleotide resolution**


**In this seminal paper the authors identified, for the first time, naturally generated aldehydes as a possible physiological source of DNA damage, linking alcohol metabolism to the FA clinical phenotype.**


*This study shows in adult FA mice that hematopoietic stem and progenitor cells are highly sensitive to aldehydes providing a mechanism for BMF in FA patients.*


73. Chen CH, Ferreira JC, Gross ER, Mochly-Rosen D: Targeting aldehyde dehydrogenase 2: new therapeutic opportunities. *Physiol Rev* 2014, **94**:1-34.


**Raya and co-workers showed in this paper that on correction of the genetic defect, somatic cells from FA patients can be reprogrammed to iPSC cells to generate disease-free haematopoietic progenitors.


**Liu and co-worker generated integration-free FA IPS cells to model FA pathogenesis and therapeutics. See annotation to reference 102.


**See annotation to reference 102.


*References 96, 97 and 102 first employed genome editing technologies to genetically correct FA in human FA fibroblasts offering novel therapeutic strategies to cure the disease*

Figure 1: FA clinical and cellular phenotypes. (a) A wide range of clinical features, all of them of incomplete penetrance, characterizes the FA clinical phenotype. The main FA clinical features can be organized in four fundamental categories: abnormal embryo development (birth defects, microcephaly), bone marrow failure, reproductive and endocrine defects and cancer predisposition including solid tumors and leukemia. In some genetic subtypes the disease links to breast and ovarian cancer predisposition in monoallelic carriers and some adult FA patients (b) The hallmark of FA cells is a DNA repair defect that causes cellular sensitivity to ICL-inducing agents, chromosome instability and cell cycle alterations. FA cells have several other phenotypic abnormalities most probably related the DNA repair defect such as the oxygen and aldehydes sensitivity, increased cell death and cell reprogramming defects. Additionally, FA cells overproduce proinflammatory cytokines that are known to be proapoptotic for HSC. All these characteristics are probably related to each other and their combination, together with stochastic factors, account for the clinical phenotype of FA patients.

Figure 2: FA related genes. Biallelic mutations in 14 genes (FANCA, B, C, D1/BRCA2, D2, E, F, G, I, J/BRIP1, L, N/PALB2, P/SLX4 and Q/ERCC4; marked in purple) cause FA, characterized by the triad of bone marrow failure, chromosome fragility and malformations. Biallelic mutations in two genes (FANCD1/Rad51C and S/BRIP1; marked in red) cause an FA-like syndrome without bone marrow failure. There are a number of FA-associated genes not linked to a FA disease phenotype in biallelic mutation carriers including FANCM and this gene is therefore excluded from the list of FA genes. Six of the above genes (FANCD1, J, M, N, O, and S; marked in dark blue) are associated to familiar breast and ovarian cancer in monoallelic mutation carriers highlighting the fundamental link between FA and cancer predisposition in the general population.

Figure 3: The FA/BRCA DNA ICL repair pathway: FANCore complex is activated upon stalled replication forks and monoubiquitinates the ID complex. Activated ID complex relocates to chromatin and promotes SLX4 and ERCC4 activities in ICL unhooking and subsequent translesion synthesis. In the final step of this process, genome integrity is restored by homologous recombination repair upon the action of downstream FA proteins. Proteins coded by bona fide FA genes are showed in blue while the products of FA-like genes are showed in green. The FA associated proteins whose genes are not mutated in any FA or FA-like patient are shown in red.

Figure 4: Cell and gene therapy strategies in Fanconi anemia. (a) Gene corrected blood progenitors can be generated after genetic correction of patient-derived hematopoietic stem cells. This is classically known as ex vivo gene therapy. Alternatively, blood progenitors can be generated from disease-free IPS cells after reprogramming gene-corrected skin fibroblasts. Corrected blood progenitors would finally be use to cure the disease by autotransplantation. (b) There are several strategies for FA gene correction. Ongoing FA gene therapy clinical trials are based on untargeted gene therapy using lentiviral vectors carrying a wild type copy of the FANCA gene. Novel genetic tools allow targeted correction of the endogenous FANCA mutation by homologous recombination by the use of either helper dependent adenoviral vectors (HDAdV) or engineered nucleases such as CRISPR/Cas9. Alternatively, genetic correction can be done by targeting the insertion of the wild type FANCA gene into a safe harbor locus with the help of engineered nucleases such as TALEN.
Figure 4 in TIFF format
Click here to download high resolution image