Chapter 31. The Fanconi anemia story and the repair of DNA crosslinks 220731cm3

Drugs Against Cancer: Stories of Discovery and the Quest for a Cure

Kurt W. Kohn, MD, PhD
Scientist Emeritus
National Cancer Institute
Bethesda, Maryland
kohnk@nih.gov

CHAPTER 31

The Fanconi anemia story and the repair of DNA crosslinks.

Fanconi anemia: a clue to how DNA crosslinks are repaired.

In 1927, the Swiss pediatrician Guido Fanconi (Figure 31.1), reported an unusual inherited anemia affecting three brothers. The red blood cells of these 3 anemic children were enlarged, leading Fanconi to describe the disease as a form of pernicious anemia. However, the clinical pattern, family inheritance and macrocytic anemia of the children were unusual, and indeed, Fanconi had discovered a new disease, which came to bear his name: Fanconi anemia (Walden and Deans, 2014). This rare genetic disease was to provide a key to unraveling how DNA crosslinks and some other DNA derangements are repaired (Alter et al., 2003; Boisvert and Howlett, 2014; Howlett et al., 2009; Rego et al., 2009; Rosenberg et al., 2003; Vuono et al., 2016).

The first step to this unraveling was made in the early 1960's by several research groups who obtained cells from Fanconi anemia patients and grew them in culture and looked at cells undergoing mitosis, where chromosome structure was clearly seen. They saw chromosome abnormalities in an unusually high frequency (Digweed and Sperling, 1996) (Figure 31.2).

The next step came when researchers tried to find out what might be causing this high frequency of chromosome breaks and abnormalities. A clue had already come from testing a variety of drugs for their ability to cause chromosome damage. Remarkably, chromosome damage in cells from Fanconi anemia patients was most effected by drugs or chemicals that were known to produce DNA inter-strand crosslinks (Weksberg et al., 1979) (Chapters 1 and 3). The sensitivity of Fanconi anemia cells specifically to DNA crosslinkers was observed for cisplatin, carboplatin, nitrogen mustard, cyclophosphamide, and diepoxybutane (Niraj et al., 2019). Apparently, cells had a mechanism to repair DNA crosslinks that was defective in Fanconi anemia cells.

Chromosome abnormalities, seen in the patient's lymphocytes during mitosis, became a criterion for the diagnosis of Fanconi anemia. The chromosome breaks and the underlying DNA damage made the patients prone to developing cancer, with acute myeloid leukemia occurring in about 10%. In addition, newborns bearing mutated Fanconi anemia genes often had congenital abnormalities, presumably because the embryos were defective in their ability to repair certain types of DNA damage that occurs occasionally in all cells.

In addition to the chromosome abnormalities, Fanconi anemia cells were killed by unusually low concentration of DNA crosslinking drugs (Digweed and Sperling, 1996) (Figure 31.3). Researchers surmised that Fanconi anemia cells were defective in ability to repair DNA crosslinks. Indeed, the high sensitivity to treatment with DNA crosslinkers became a diagnostic test for Fanconi anemia. The most reliable DNA crosslinker for the test was diepoxybutane, a bifunctional alkylating agent that has a chemically simple and direct crosslinking mechanism (Auerbach, 1988).

One may wonder why such a complicated DNA repair as the Fanconi system would have evolved specifically to deal with inter-strand crosslink producers rarely found in nature. The answer to this conundrum may be normal metabolic processes that produce rare, but in aggregate many, inter-strand crosslinks. Processes such as lipid peroxidation, histone demethylation, and alcohol metabolism can generate formaldehyde and acetaldehyde byproducts that can react to form inter-strand crosslinks (Niraj et al., 2019).

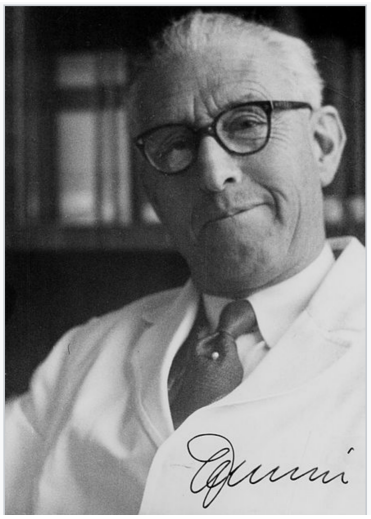

Figure 31.1. Guido Fanconi (1892-1979) was a Swiss pediatrician, regarded as a founder of modern pediatrics. In 1927, he described the hereditary anemia that bears his name. His name in fact became associated with two different diseases: 'Fanconi syndrome' is a disorder of kidney function that must not be confused with 'Fanconi anemia'. (Photo in 1959 by Israeli photographer Ze'ev Aleksandrowicz.)

Figure 31.2. Chromosomes of a Fanconi anemia patient's lymphocyte. The arrows point to chromosome breaks and to an abnormal joining of 2 chromosomes (center). (Photograph by Dr Rolf Wegner, Berlin (Digweed and Sperling, 1996).

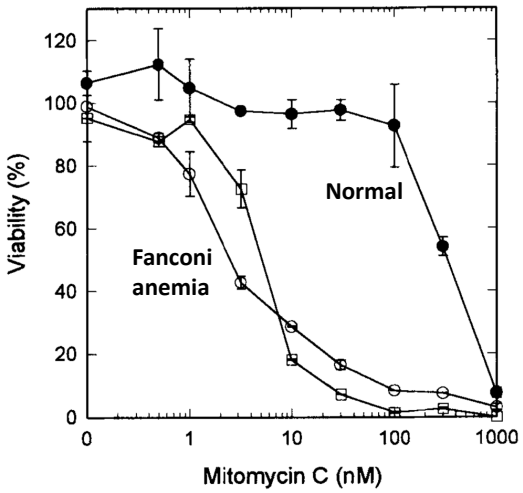

Figure 31.3. Cell lines from two Fanconi anemia (complementation group A) patients (open symbols) were killed by 100-fold lower concentrations of mitomycin than a cell line from a normal person (solid symbols) (Digweed and Sperling, 1996).

The clinical diagnosis of Fanconi anemia was often difficult because symptoms among patients was variable. This problem was solved by testing for abnormally high sensitivity to DNA crosslinkers, such as diepoxybutane or mitomycin. But why was there such variability

in the clinical picture among different patients? Several researchers suspected that each type of clinical pattern was caused by a different gene. In other words, Fanconi anemia might be caused by a defect in any one of two or more genes. This possibility was thoroughly investigated in 1985 by G. Duckworth-Rysiecki, M. Buchwald, and their coworkers using a cell fusion technique (Duckworth-Rysiecki et al., 1985). When they fused together pairs of lymphocytes, each from a different patient, sometimes the response to crosslinking drug became normal and sometimes the high sensitivity remained. They surmised that, when fused cell pairs had normal drug responses, then the disease of the two patients from whom the cells were taken was caused by defects in different genes. That meant that there were at least two different types or "complementation groups" of Fanconi anemia. When cells of different complementation groups were fused together, the resulting cell duos were able to repair the crosslinks and survive normally, but the combination of cells from patients who had the same complementation group did not do so: that was how "complementation group" was defined. In other words, cells of the same complementation group had the same underlying defect, whereas cells of different complementation groups had defects caused by different, independent factors.

By 1992, four different complementation groups of Fanconi anemia had been defined (Strathdee et al., 1992a). The story was becoming complicated. And it didn't stop there. By 1999, eight complementation groups were defined. Thus, there were at least 8 different types of Fanconi anemia, each caused by a mutation in a different gene. Moreover, the proteins encoded by at least three of these genes were found to bind together to form a functional complex (Garcia-Higuera et al., 1999). It began to look like some of the proteins encoded by the Fanconi anemia genes bound to each other and worked together as a multi-protein complex. By 2006, at least 12 Fanconi anemia genes had been discovered, and information began to accrue about how those genes assemble into a complex and how the system functions (Medhurst et al., 2006).

As often happens is a developing field of research, the story gradually became increasingly complicated. As of 2017, 22 distinct Fanconi anemia complementation groups had been found, each of which defined a different gene mutation (Che et al., 2018; Nepal et al., 2017). As many as 13 of these Fanconi genes had rare mutated forms on various chromosomes that were inherited. Family members who carried one of these mutated genes had an increased chance of developing cancer.

Repair of crosslinks that covalently linked the two strands of a DNA double helix must be a lot more complicated than repair of damage to just one strand, and the repair had to be error-free, because DNA crosslinks occurred occasionally even in normal cells. That may be why so many different gene products, *i.e.,* proteins, were required to repair DNA inter-strand crosslinks.

Fanconi anemia resulted when both copies of any one of a person's 22 Fanconi anemia genes were defective, usually because of mutations. Because of the defect, the patient's cells were unable to repair DNA crosslinks. The proteins encoded by the 22 Fanconi anemia genes, together with several other genes, were found to cooperate in the crosslink repair (Nepal et al., 2017). How that repair works is described below.

Inactivating mutations of both copies of one of the Fanconi anemia genes often pushes cells on road to malignancy (Alter et al., 2003). (Those mutations occur fresh in body cells and are relatively common, in contrast to Fanconi anemia inherited from germ cells, which is rare). Although disordered blood cell production and anemia was common in Fanconi anemia in children, cancers of various types usually appeared years later when Fanconi anemia patients had reached adulthood. Cancer treatment in Fanconi anemia patients was difficult, because the patient's normal tissues are sensitive to some of the best anti-cancer drugs.

Fanconi anemia was found to be genetically recessive and occured when both parents has one normal and one mutated gene of a given complementation group. Each child then has a 1-in-4 chance of inheriting the disease. Each child had a 1-in-2 chance of becoming a carrier, like both parents, with a mutated gene paired with a normal gene. Individuals who were carriers of a Fanconi anemia gene mutation had an increased risk of developing cancer sometime during their lives, and the risk varied, depending on which Fanconi gene he/she was carrying (Alter et al., 2003).

Unraveling the roles of the Fanconi anemia ('Fanc') proteins in DNA repair.

Unravelling the molecular details of how the *Fanc* genes and their protein products functioned was important because of their role in repair of DNA damage, particularly the kind of DNA damage caused by some of the most useful cancer chemotherapy drugs: the DNA inter-strand crosslinkers, including nitrogen mustards, cyclophosphamide, platinum complexes, and mitomycin (discussed in Chapters 1 and 3). Moreover, one or another *Fanc* gene was found defective in 40% of cancer cases and put the Fanconi problem high on the list of questions about both cancer cause and opportunities for therapy.

The unravelling problem was hard, because at least 22 *Fanc* anemia genes plus a number of functionally related genes all seemed to work together. How did they all work together to repair DNA crosslinks? The first objective was to clone and determine the nucleotide sequence of each gene and the structure of the proteins encoded by them. But that was only the start of the difficulty. Finding out how all of those proteins worked together to repair DNA crosslinks might have seemed a nearly impossible task. It might have surprised the early researcher who started on this effort that a large part of the story would be revealed within less than 3 decades.

The road to unravelling the complicated story of how the Fanconi system repaired DNA crosslinks began in 1989, when Martin Digweed and Karl Sperling identified an mRNA fraction from non-Fanconi cells that could correct the crosslink repair defect in Fanconi anemia cells (Digweed and Sperling, 1989). The cDNA then served to identify and clone the Fanc genes.

The molecular part of the story began with the cloning of the first *Fanc* gene, which was accomplished in 1992 by Craig Strathdee, Manuel Buchwald and their coworkers at the University of Toronto, Canada (Strathdee et al., 1992b). At that time, 4 Fanconi anemia complementation groups had been defined, and it was known that fusing together cells of different complementation groups corrected the DNA repair defect. But nothing was

known about the genes suspected of being mutated in each of the complementation groups. The Toronto researchers reasoned that the repair defect in cells of a given complementation group might be correctable by a normal version of the defective gene. Although they could not introduce the intact genes, they could get viruses to carry a library of cDNA molecules from normal cells into the recipient cells (each virus particle would carry one of the myriads of normal cDNA molecules). As recipients, they chose cells of complementation group C, because these were highly sensitive to mitomycin, and would give high sensitivity for correction of a DNA repair defect. Some of the cells that survived a good slug of mitomycin might then have harbored a virus that conveyed the correcting DNA sequence. Then it was merely a matter of transferring the viruses that did the correcting back into bacteria for cloning, and *voila!* the cDNA sequence of the complementation group C gene was at hand. It was not quite as simple as that, but that was the essence.

(What is "cDNA"? Each cDNA molecule has a sequence that is complementary to a protein-coding RNA molecule in the cytoplasm. It is produced artificially by means of a reverse transcriptase enzyme that copies protein-coding mRNA sequences found in the cytoplasm into a complementary DNA version.)

The next *Fanc* gene cDNA to be cloned and sequenced was *FancA*, whose gene mutations accounted for 65% of the cases of the disease. This was accomplished in 1996 by an international group led by Hans Joenje and Manuel Buchwald using a method similar to how the Toronto group had cloned and sequences the cDNA of *FancC* (Lo Ten Foe et al., 1996)

The first information about how the *Fanc* genes function came in 1997 from Paul D'Andrea's laboratory at the Dana-Farber Cancer Institute and Harvard Medical School. They found that the FancA and FancC proteins bound to each other and then were able to enter the cell nucleus (Kupfer et al., 1997). This binding was likely important, because a mutant FancC protein from complementation group C patients failed to bind to the FancA protein. A drug that prevented this binding might therefore increase the sensitivity of cancer cells to DNA crosslinking agents.

The researchers went on to show that mutation of any one of several other *Fanc* genes prevented FancA-FancC proteins binding each other and moving into the nucleus (Garcia-Higuera et al., 1999). It seemed that several Fanconi proteins were needed to form a functional multiprotein complex.

The next player to enter the Fanconi dance was FancG. In 1999, Quinten Waisfisz and Hans Joenje at the Free University in Amsterdam, and their colleagues, discovered that FancA, in addition the binding FancC, also bound FancG (Waisfisz et al., 1999). Therefore, FancG seemed to be part of the functional multiprotein complex.

In 2000, the Amsterdam group added another piece to the story by cloning *FancF* and proposing that the Fanconi proteins combine to form a complex that maintains the integrity of the chromosomes in the nucleus (de Winter et al., 2000). By 2001, altogether

six *Fanc* genes had been cloned and several of the proteins were found to bind to each other (Joenje and Patel, 2001). It seemed that the binding together of several Fanc proteins formed a multiprotein complex that entered the nucleus to exert its functions, and the idea emerged that disrupting the complex might become a new cancer therapy.

The next major addition to the story was made in 2002 by K. J. Patel and colleagues at the Universities of Cambridge and Amsterdam, who showed that the multi-protein Fanconi complex activated another Fanconi protein, FancD2, by adding a molecule of the small protein ubiquitin to it (Pace et al., 2002). This was a key discovery, as we shall see.

Another key discovery came from an international group of scientists in 2005 (Meetei et al., 2005), who discovered *FancM*, the gene and protein defective in Fanconi anemia complementation group M. They showed that FancM bound to the other Fanc proteins in the core multiprotein complex and was required for ubiquitylation of FanD2. But more importantly, they found that FancM protein interacted with some abnormal DNA structures. It seemed that FancM's role might recognize and bind DNA crosslink sites and bring other Fanc proteins to the damage site.

By 2006, it became possible to suggest a model of how various Fanconi proteins assemble into complexes and subcomplexes (Medhurst et al., 2006) (Figure 31.4). The model proposed that most of the then-known 11 Fanconi proteins bound to each other to form a multiprotein complex in the cell nucleus and that this complex ubiquitylated FancD2: it stuck a molecule of the small protein, ubiquitin, onto the FancD2 protein. See legend of Figure 31.4 for details.

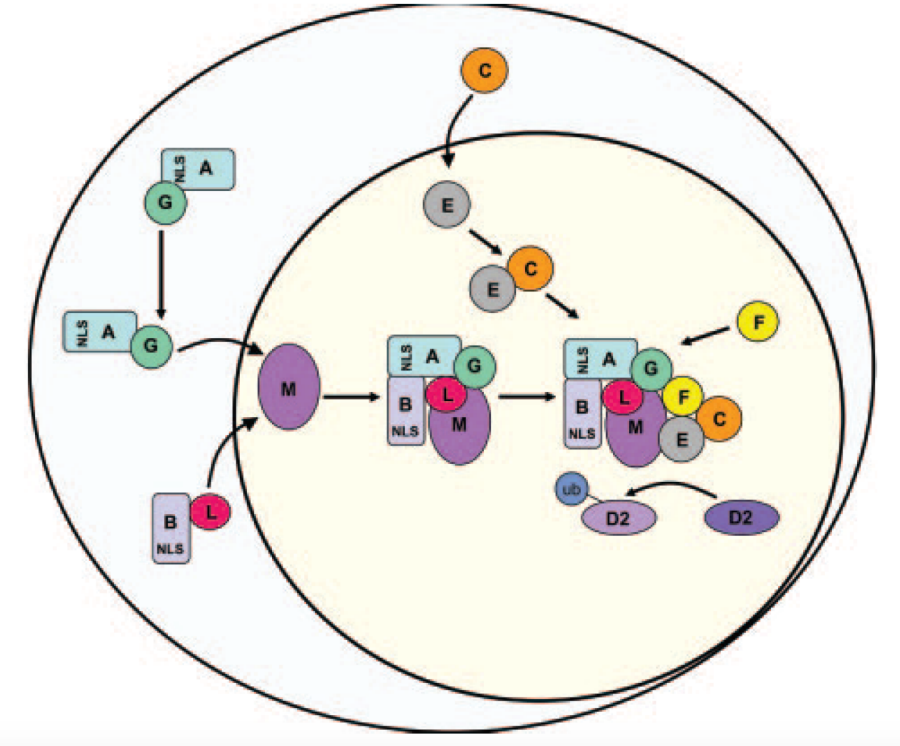

Figure 31.4. A model proposed in 2006 by Medhurst and colleagues of how Fanconi anemia proteins assemble in the cell nucleus at a DNA crosslink site and form a multi-protein complex that ubiquitylates FancD2 (Medhurst et al., 2006). The right side of the diagram proposed that a complex containing Fanc's A, B, C, E, F, G, L, and M ubiquitylates FancD2. The left side of the diagram suggested how the Fanc multiprotein complex forms. The model proposed that FancM, which was thought to recognize and bind to the crosslink site on the DNA, recruits Fanc's A, B, G, and L at the DNA damage site in the nucleus. Fancs A and B had nuclear localization signals in their protein structures, which would carry them, as well as Fancs G and L into the nucleus. Fancs C, E, and F would then somehow become recruited to the complex that then would become capable of ubiquitylating FancD2.

The next important discovery was the gene and protein defective in Fanconi anemia complementation group I, FancI (Smogorzewska et al., 2007). This eye-opening discovery came in 2007 from a research group led by Steve Elledge at Harvard Medical School, that found that FancI protein resembled FancD2 in amino acid sequence. Moreover, the two proteins bound tightly to each other when both of them became ubiqutylated by the core Fanc multiprotein complex. Also, unlike other Fanc proteins, the FancD2-FancI pair was not required for the assembly of the core Fanc complex. The FancD2-FancI pair seemed special, because the other known Fanc proteins had individually unique amino acid sequences, and, unlike FancD2 and FancI, each of them was required for assembly of the core Fanc multiprotein complex.

Another remarkable aspect of the FancD2-FancI dimer was its structure. The structure was determined by crystallography in 2011 by a research group led by Nikola Pavletich at Memorial Sloan-Kettering Cancer Center in New York, together with Steve Elledge's group at Harvard (Joo et al., 2011). The structure of the duo resembled two saxophone-like shapes fitted together, and the ubiquitins were at the junction between the two proteins, evidently serving to lock the protein pair together (Figure 31.5).

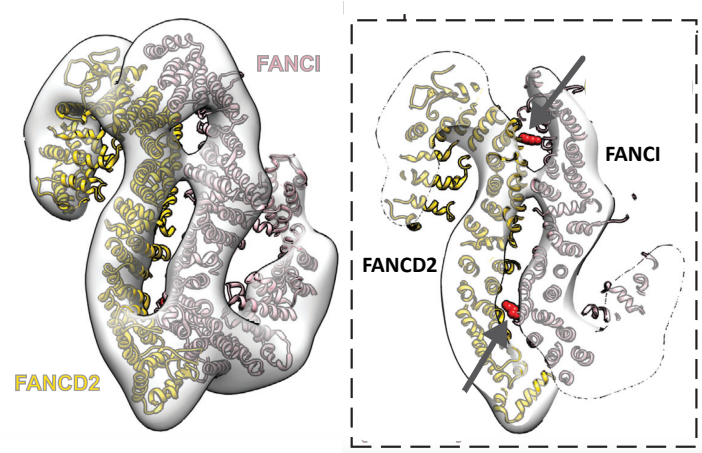

Figure 31.5. The saxophone-like shapes of FancD2 and FancI, showing how the couple nestles together. Each has a ubiquitin attached, and the two ubiquitins are located at the junction between the two Fancs, where they lock the couple together (Joo et al., 2011). The alpha helices of the proteins are shown in yellow. On the right is a section through the dimer, showing the locations of the ubiquitins (red with arrows pointing to them). From (Swuec et al., 2017) with arrows and labels added.

Yet another remarkable finding came in 2007. When cells were treated with a DNA crosslinking drug, FancD2 and FancI entered the cell nucleus, but were not spread all over the nucleus. Instead, the ubiquitylated FancD2 and FancI became concentrated in spots or 'foci' where a DNA crosslink was located (Smogorzewska et al., 2007) (Figure 31.6). It made sense that these and perhaps other Fancs should become localized to places where they were needed to repair the crosslinks. But the fact that the foci were so clearly visible meant that a huge excess of these molecules somehow collected at the site of the crosslink – many more than were needed for the repair. How and why did that happen? As far as I know, this remains an open question. (There is more discussion about nuclear foci in the story of histone gamma-H2Ax in Chapter 28.)

Figure 31.6. FancI and FancD2 both became localized at foci (spots) where there were DNA crosslinks in the nucleus of a cell. The images show a single nucleus of a mitomycin-treated cell. Left: stained with a blue antibody to FancI. Middle: stained with a red antibody to FancD2. Right: merge of the two images. In the merged image, the blue and red spots combine to form white spots, showing that FancI and FancD2 were both present in each spot. From (Smogorzewska et al., 2007).

The Fanc DNA damage response and repair network.

By 2010, it was surmised that most of the known Fanconi anemia proteins (13 of the 22 Fanconi anemia genes were then known) bind together in a large multi-protein core that has essential functions in the DNA crosslink repair pathway (D'Andrea, 2010). Furthermore, FancD2 and FancI were special: they were found to bind to each other and to become activated by becoming ubiquitylated (*i.e.,* by having the small protein, ubiquitin, bound to each of them) (Figure 31.5). The Fanconi core multi-protein complex (or some part of it) was found to carry out this ubiquitylation reaction. The general picture was that FancM detects and binds DNA at the site of a crosslink and then binds and signals the Fanconi core complex to ubiquitylate the FancD2-FancI dimer. The ubiquitylated dimer would then activate downstream proteins that carry out the initial DNA repair steps, but exactly how the repair itself worked was not yet known.

By 2017, 22 *Fanc* genes had been identified and cloned, and a general picture of the roles of the 22 Fanc proteins emerged {Nepal, 2017 #1010) (Figure 31.7). The diagram depicts the multitude of protein species whose functions connect in one way or another to the Fanc DNA repair pathway. The center of the diagram shows the Fanc core multiprotein complex ubiquitylating the FancD2-FancI dimer. FancM (upper left in the diagram) is shown liked to a DNA damage site. The diagram proposed a sequence of steps leading to phosphorylation of the FancD2-FancI dimer by ATM (whose gene is mutated in ataxia telangiectasia; see Chapter 29). The Fanc core multiprotein complex then somehow converted the FancD2-FancI dimer to a ubiquitylated form. The ubiquitylated FancD2-FancI dimer would then be the executor of several outputs of the Fanc pathway (shown along the bottom part of the diagram).

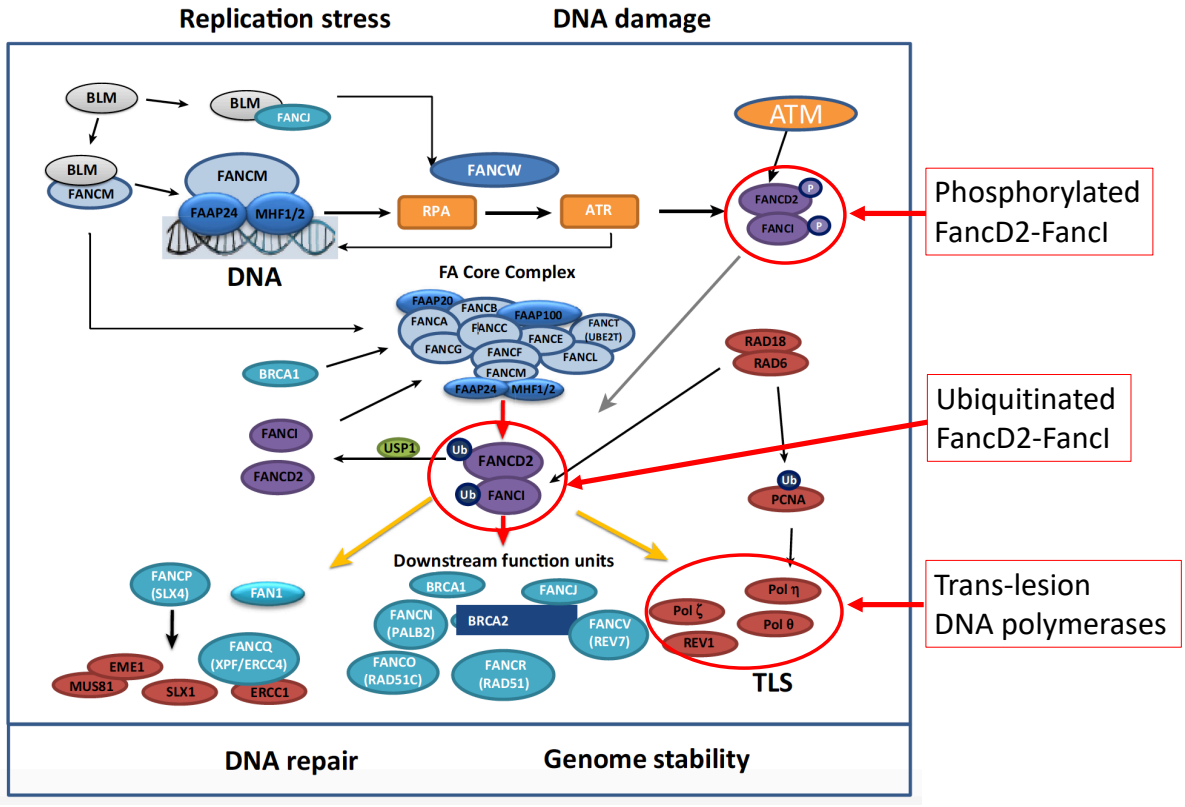

Figure 31.7. Overview of the Fanc DNA repair pathway. (From (Nepal et al., 2017) with labels and ovals added.) Near the upper left of the diagram, FancM (together with two associated protein) recognizes and binds to DNA at a site of damage, such as an inter-strand crosslink. Through a sequence of steps, shown from left to right in the upper part of the diagram, this results in phosphorylation of the FancD2-FancI dimer by ATM (mutated in ataxia telangiectasia; see Chapter 29). The Fanc core multiprotein complex (shown in the middle of the diagram) then causes the FancD2-FancI dimer to become ubiquitylated, which then executes several functions, including stimulation of trans-lesion DNA synthesis.

Thus, ubiquitylated FancD2-FancI dimer is special: it is not part of the core multiprotein complex and is the output or executor of the Fanc pathway. Another Fanc that has special functions in the pathway is FancM, the recognizer of DNA damage sites (top left in Figure 31.7).

All about FancM.

FancM has an unusual story of discovery. Rather than being found as a gene in Fanconi anemia patients, if was found as a previously unknown protein that bound to Fanc proteins, such as FancA (Meetei et al., 2005). FancM was a part of the Fanc pathway, because, in the

absence of FancM, cells did not ubiquitylate FancD2 and had increased sensitivity to being killed by the DNA crosslinker, mitomycin. Mutation of *FancM* was noted to be especially associated with uterine and breast cancers (Basbous and Constantinou, 2019).

Remarkably, *FancM* resembled a gene called *Hef* in ancient microorganisms, the archaea. FancM seemed to have a critical function dating back to the early history of life. What was that critical function? The greatest similarity between FancM and Hef was in a region of amino acid sequence that conferred helicase function (Figure 31.8), which is the ability to unwind the DNA helix. Helicase becomes important during DNA replication and repair, when the growing end of the new strand encounters a break in the strand it is trying to copy. The problem is particularly severe when both template strands have breaks in nearly the same place. This may well have been problematic for organisms since early in evolution. How was the growing strand to find a complementary sequence to copy beyond the break in its template strand? The best solution in diploid organisms (which have a pair of each chromosome) was for the blocked strand to copy the relevant sequences from DNA in the other chromosome of the pair by homologous recombination (Chapter 27A). (The alternative method of double stand break repair, non-homologous end joining Chapter 27B, did not involve FancM.)

We see in Figure 31.8 that, although the FancM helicase and endonuclease functions reside in the same gene product protein in archaea, they are separated in humans and other vertebrates, the helicase function being in FancM and the endonuclease function in ERCC4.

What does FancM actually do in the Fanc DNA repair pathway; how do its DNA damage recognition and DNA unwinding (helicase) abilities come into play? The key may be that FancM can translocate along the DNA duplex; it may be looking for a damage site to lock on to and bring the repair proteins of the FancM pathway to the site (Meetei et al., 2005).

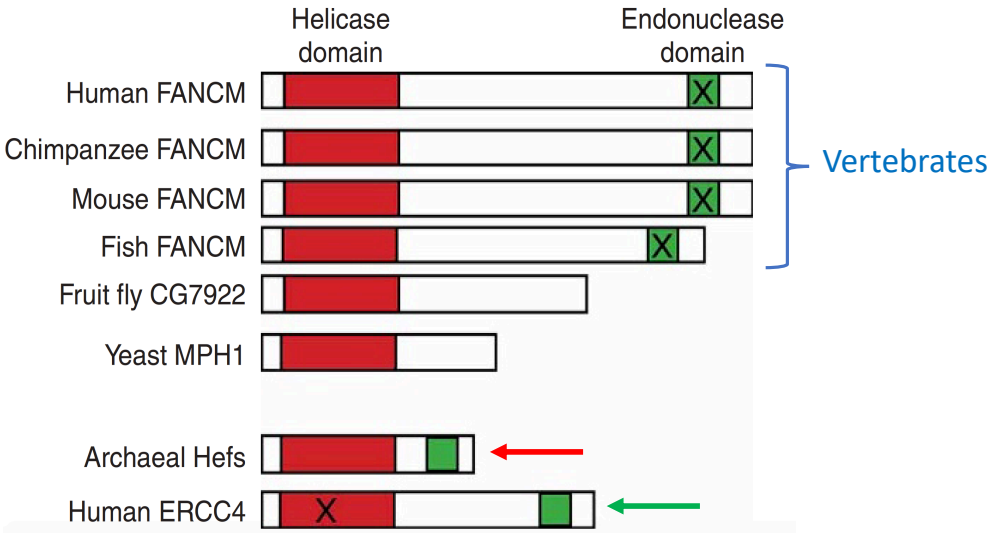

Figure 31.8 The FancM protein in various organisms has a DNA helicase domain (red) and a DNA endonuclease domain (green) in its amino acid sequence. Both domains indicate a role of the FancM protein in DNA repair. However, in some organisms, a domain may be functionally defective (marked by a X), in which case the amino acid sequence is not quite correct for function. Even a defective domain however indicates an evolutionary relationship of the proteins: a common ancestor. The only organisms in which both domains are functional are the ancient Archaea microorganism, whose FancM-like protein is called Hef. Human ERCC4 (green arrow) is a FancM relative that functions as an endonuclease (DNA-cutting) enzyme in DNA repair. FancM in vertebrates from Human to fish has a functional helicase (DNA unwinding) domain, which is important in DNA repair, but its endonuclease function is defective (and provided by ERCC4). (From (Meetei et al., 2005).)

An important finding about inter-strand crosslink repair was that most of the repair occurred at sites where DNA replication was blocked. In other words, sites along the DNA where progressing replication forks have encountered a crosslink and cannot progress further. The repair mechanism is in fact linked to the replication machinery. An experiment demonstrating this, reported in 2013 from Michael Seidman's laboratory at NIH in Baltimore, is shown in Figure 31.9. Under the conditions of the experiment, blocked replication forks coming from one direction (**B**) or from both directions (**C)** occurred with about equal frequencies (the crosslink is indicated by a vertical red bar). The method that revealed growing DNA strands blocked at crosslinks was somewhat complicated but illustrated the remarkable power of the new technology.

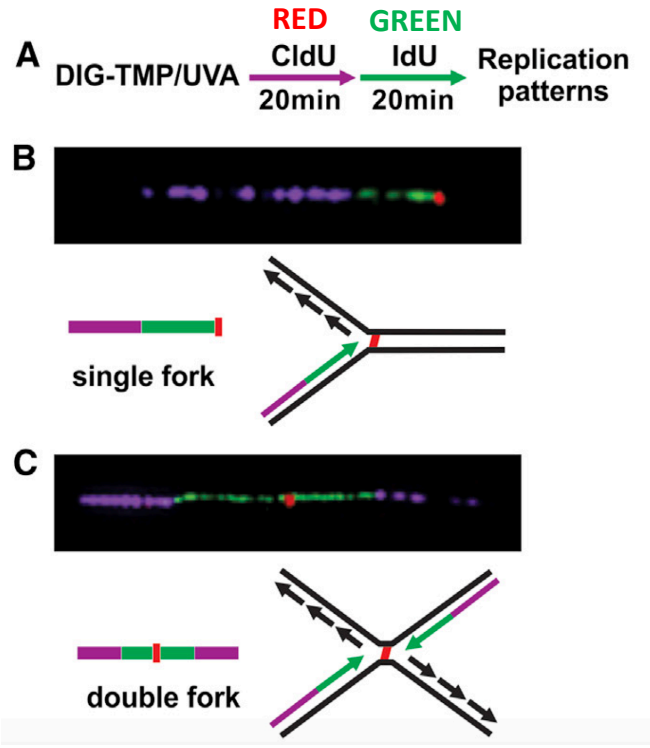

Figure 31.9. An experiment reported in 2013 by Jing Hang and colleagues in Michael Seidman's laboratory at NIH in Baltimore, showing how DNA replication blocks at inter-strand crosslink sites were visualized (Huang et al., 2013). An example of a replication fork coming from one direction or from both directions are shown in **B** and **C**, respectively. In both cases, the replication forks were blocked at the crosslink (vertical red bar in the diagrams). An example of how the replication blocks were visualized are shown above each diagram. The protocol of the experiment is shown in **A**, Inter-strand crosslinks were produced by treating cells with a psoralen compound followed by ultraviolet light (UV) (see section about psoralen in Chapter 1). The psoralen part of the compound had linked to it a part that recognized an antibody that would fluoresce red. The cells were then treated for 20 minutes with chlorodeoxyuridine (CldU), which became incorporated into growing DNA stands; the CldU would be recognized by an antibody that fluoresced red. The CldU was then washed away and replaced by iododeoxyuridine (IdU), which would be recognized by an antibody that fluoresced green. DNA fibers from the cells were then stretched out and treated with the fluorescent antibodies. The more recently replicated DNA glowed green, while the older part of the newly replicated DNA glowed red. We see that the most recently replicated part of the DNA stand (green) stops at the crosslink, where replicated had stopped after proceeding only a short distance (red spot).

In 2013, Kottemann and Smogorzewska (Kottemann and Smogorzewska, 2013) had outined DNA interstrand crosskink repair (Figure 31.10) as going by the following, perhaps

necessarily compicated, steps: First, the core Fanc complex (Figure 31.7), together with accessory proteins, detects the crosslink and becomes activated by phosphorylation mediated by the ATM-related kinase, ATR. The the core complex (perhaps mediated by its FancL component (Garcia et al., 2009)) then adds ubiquitins to the FancI-FancD2 dimer (see Figure 31.7), which induces SLX4, functioning as a scaffold for three nucleases (XPF, MUS81 and SLX1), to bring those nucleases to the crosslink site. The nucleases then cut a DNA strand on either side of the crosslink (red arrows in Figure 31.10). The final steps of the repair were not well understood at the time, although it was known that DNA polymerases that carry out trans-lesion synthesis (TLS) were required (see Figure 31.7). Evidence also indicated that the final process involved homologous recombination with the participation of BRCA2 (Figure 31.11).

In 2015, Xue et al (Xue et al., 2015) depicted the role of FancM in the early steps of inter-strand crosslink repair as shown in (Figure 31.12) and explained in the legend.

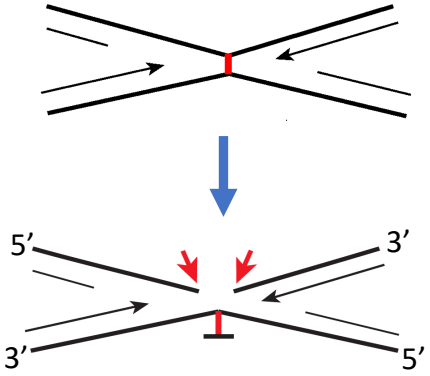

Figure 31.10. First excision step in removing a DNA inter-strand crosslink (ISC) (red bar) (Kottemann and Smogorzewska, 2013). In this representation, DNA replication processes have approached from both directions and have become blocked at the ISC. The Fanc proteins bring a host of DNA repair proteins to the site of the crosslink (Figure 31.7). Additional repair proteins are brought in by FancD2-FancI, after it has become ubiquitylated by the Fanc core complex at the crosslink site. Among the brought-in repair proteins are enzymes that cut the DNA on both sides of the crosslink (red arrows).

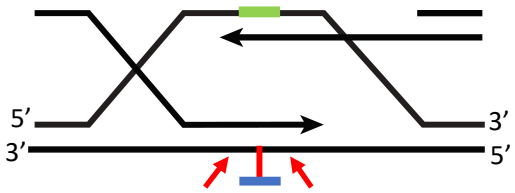

Figure 31.11. The homologous recombination step in DNA inter-strand crosslink repair. The green segment in the middle of the upper stand represents the part of the DNA strand that was healed by a trans-lesion DNA polymerase, which copy a damaged nucleotide in the template by inserting an arbitrary base, usually an A.

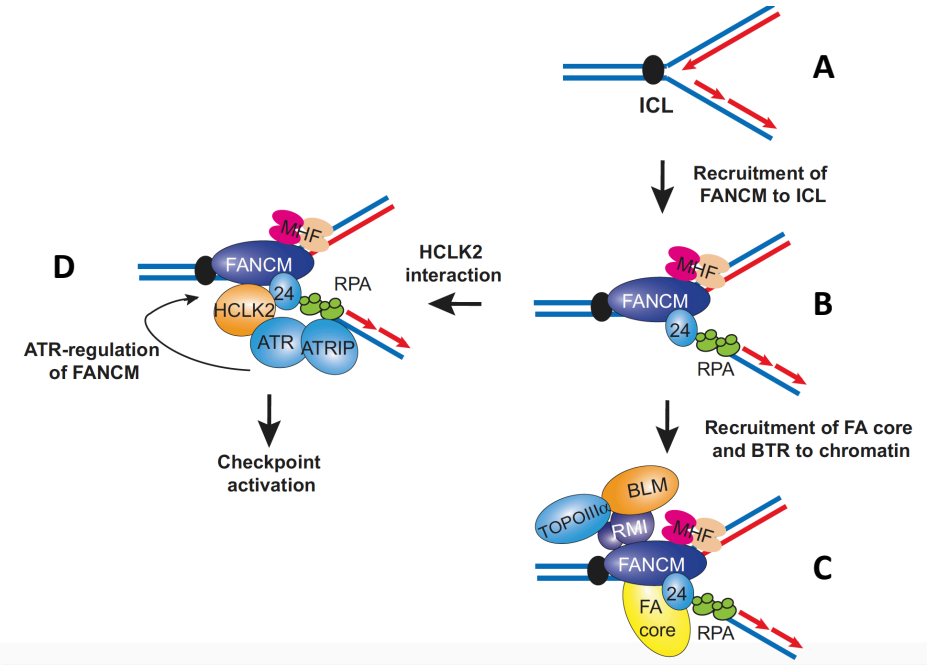

Figure 31.12. The early steps in repair of DNA inter-strand crosslinks, according to a model proposed in 2015 by Xue et al (Xue et al., 2015). First, FancM recognizes and binds to the inter-strand crosslink (with the help of a complex of 2 other proteins, labelled MHF in the diagram) (**A** and **B**). FancM then recruits other proteins required for the repair: the Fanc core complex and 3 other proteins, including the DNA-unwinder BLM (mutated in Bloom's disease) and topoisomerase TopIIIa, which relieves tortional stress in the DNA helix (**C**). In an alternative path, FancM recruits proteins that signal the cell to delay DNA replication and cell division (**D**). The Fanconi anemia core complex in **C** (FancT and FancL (Garcia et al., 2009), within the complex) goes on to ubiquitylate FancD2-FancI (Figure 31.7), which leads to the next phase of the repair. (From (Xue et al., 2015), simplified.)

The final steps of DNA crosslink repair: homologous recombination.

A clue for a role of homologous recombination in the Fanc crosslink repair pathway came as early as 2002 with the discovery by Niall Howlett, Alan D'Andrea and their coworkers at the Dana-Farber Cancer Institute and Harvard Medical School that, surprisingly, the breast cancer-associated genes *BRCA1* and *BRCA2* were an intimate and essential part of the network of *Fanc* genes that repaired DNA crosslinks (Howlett et al., 2002). It was later found that, in fact, the *BRCA1* gene was the same as *FancS*, and the *BRCA2* was the same as *FancD1* (Nepal et al., 2017). BRCA1 and BRCA2 were known to be part of the homologous recombination system that

is required by several DNA repair processes (Chapter 27A). Apparently, some of the *FANC* genes were at a convergence of several DNA repair pathways (Niraj et al., 2019).

A key step then came in 2010, in a paper by Fiona Vaz, Helmut Hanenberg, Detlev Schindler, Christopher Mathew, and their colleagues in London and Germany (Vaz et al., 2010). At the time, there were 13 known Fanconi anemia genes. The authors investigated a Fanconi anemia family where there was no mutation in any of known Fanconi anemia genes. One of the affected family members was a 10-year-old boy whose cultured fibroblast cells had many chromosome rearrangements and whose blood lymphocytes were killed by low concentrations of DNA crosslinking agents (mitomycin and diepoxybutane). The boy, as well as other affected family members, had multiple developmental abnormalities; some died in infancy.

The boy had a mutated gene, but, surprisingly, the mutation was in *Rad51C*, a gene that was involved, together with its close relative *Rad51*, in homologous recombination(Figure 27A). It was later confirmed that mutation of *Rad51C* caused clinical Fanconi anemia. Therefore, *Rad51C* was given the alternative name, *FancO*; the homologous recombination gene, *Rad51C*, was the same as the newly discovered Fanconi anemia gene, *FancO*.

But FancO was not part of the Fanc core complex. Therefore, it was thought to function downstream of the FancD2-FancI ubiquitylation step, perhaps in a recombination process that repairs an intermediate DNA structure. The bottom-line message was that, since mutation of *Rad51C* caused Fanconi anemia, the repair of DNA inter-strand crosslinks may involve DNA recombination. That was consistent with the already mentioned finding that *BRCA1*, *BRCA2*, and *Rad51C,* which were known to function in genetic recombination, were in fact the same as *FancS*, *FancD1*, and *FancO*, respectively.

Fanc DNA repair components function also in other DNA repair pathways.

In 2002, the D'Andrea lab at Harvard had discovered a connection between the Fanconi DNA repair pathway and the repair pathways that are controlled by ATM gene that is mutated in ataxia telangiectasia (see Chapter 29) (Taniguchi et al., 2002). They discovered that FancD2 phosphorylates and thereby activates the ATM kinase enzyme, which is the product of the *ATM* gene.

The remarkable finding, reported from Alan D'Andrea's laboratory linked the Fanconi anemia ('Fanc') DNA repair pathway with DNA repair genes associated with ataxia telangiectasia (Andreassen et al., 2004) (see Chapter 29). They and later researchers found that phosphorylation of FancD2, which switches on the Fanc DNA repair pathway, was carried out by ATR, a kinase related to the *ATM* gene. What's more, phosphorylations by ATM were also implicated in the activation switch. It seemed that phosphorylations by ATM and ATR, as well as ubiquitylations by the Fanc core multiprotein complex, were all required to activate the FancD2-FancI dimer and for the consequent activation of the Fanc

DNA repair pathway. ATM and ATR also signaled to the cell cycle checkpoints to delay the cell cycle in order to give more time for DNA repair.

A connection of Fanconi anemia genes with another DNA repair process was discovered by Amom Meetei, Weidong Wang and their coworkers at the NIH National Institute of Aging and colleagues in the Netherlands (Meetei et al., 2003). They found that FancD2 binds to BLM, the product of the gene, which, when mutated, caused the rare premature aging genetic disorder known as Bloom's syndrome. FancD2 and BLM were in a multiprotein complex different from the Fanconi anemia core complex. The BLM complex functioned to unwind the DNA helix, a step required for many DNA repair pathways, including the Fanconi anemia repair pathway. It was the first time that a biochemical step, namely DNA helix unwinding, was identified as a step in the Fanconi anemia DNA repair pathway; this was an early step in which FancD2 functions as part of the Fanconi core complex.

Yet another connection was that the Fanc complementation group gene, *FancQ*, was the same as the gene for the xeroderma pigmentosum XPF enzyme that function to cleave a DNA strand during repair of UV-induced pyrimidine dimers (see Chapter 22) (Figure 31.7).

Fanc gene mutations in cancers.

Figures 31.13 and 31.14 show how often the Fanconi anemia genes were found to be altered in cancers (Niraj et al., 2019). The alterations were classified as mutations (excluding changes in gene copy numbers), amplifications (increased number of gene copies), or deletions of parts of the gene. We see that the most frequent types of alterations were mutations in some genes and amplifications in some other genes, while deletions were most common in only a few genes.

As many as 40% of the cancers had Fanconi gene alteration. The gene alterations almost always were in the cancers, not in the genome of the patients. Evidently, a Fanconi anemia gene alteration is frequently acquired by the cancer as it develops and may contribute to the development of the malignancy. DNA damage, such as crosslinks, may occur naturally in normal metabolism that can produce small amounts of DNA damaging compounds, such as formaldehyde. The ability of a cell to repair those DNA lesions may therefore be critical. A cell's ability to repair those DNA lesions would be defective if the cell's genome has a mutation or deletion in a Fanconi anemia gene; the consequence would be a step towards malignancy. Why amplification of certain of those genes also may lead to cancer is less clear to me. Amplifications were particularly common in cancers of ovary and breast, while mutations were the dominant alterations in cancers of the uterus (Figure 31.13). It remains to be found out why there is such dependence on cancer type.

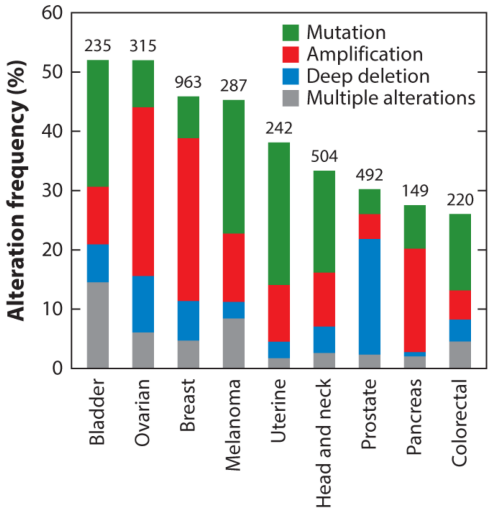

Figure 31.13. The frequency at which Fanconi anemia genes were altered in a total of 3,406 cancers of various types. Mutations, amplifications, or deletions in Fanconi anemia related genes were found in 40% of the cancers. (From Niraj et al (Niraj et al., 2019), who state that the data were generated by The Cancer Genome Atlas and were downloaded from cBioPortal.)

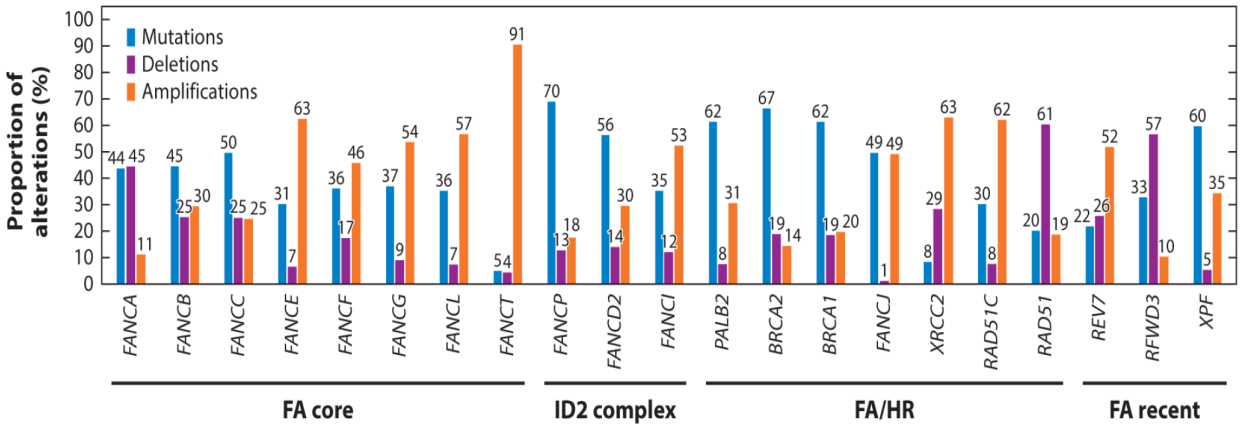

Figure 31.14. Mutations, deletions, and amplifications in Fanconi anemia genes (and several related genes) found in a total of 1,363. For some of the genes the most frequent alterations were mutations; for some genes, amplifications were the most common; for a few genes, including *FancA*, deletions were common. The frequencies of alterations of gene in the tumors were FancA 64%; FancC 12%; FancG 8%; FancD2 4%; Fancs B, F, J, and D1 2% each; Fancs E and I 1% each; all others <1%. FA, Fanconi anemia; ID2, FancD2-FancI; HR, homologous recombination. (From Niraj et al (Niraj et al., 2019), who state that the data were generated by The Cancer Genome Atlas and were downloaded from cBioPortal on May3, 2018.)

Summary of the Fanconi anemia (Fanc) genes and proteins in DNA repair.

The state of knowledge as of 2018 was reviewed by Niraj et al (Niraj et al., 2019). DNA repair by the Fanc proteins primarily targeted inter-strand crosslinks. The repair occurred mainly during DNA replication, when the replication machinery becomes stuck at the crosslink sites on the DNA. Crosslink repair at sites where a crosslink blocks the replication machinery must begin with recognition of the crosslink blockage sites on the DNA. However, proteins of the replication machinery that remain stuck to the crosslink site must be removed in order to make the site accessible to the repair machinery. According to the current picture (Niraj et al., 2019), the cleaning away of replication proteins from the crosslink site is a special job of FancI, which also has other critical functions in the repair process. The cleaned crosslink sites are then recognized by the Fanconi anemia protein, FancM, which binds to the DNA at those places and initiates the repair process by bringing to the site a multiprotein complex consisting of most of the Fanc proteins (the Fanc core complex).

The next step, which actually begins the crosslink repair, was found to be carried out by the combination of two Fanc proteins: FancD2 and FancI (Swuec et al., 2017). (FancD2 was found to be the same as BRCA2). In order to function, however, the FancD2-FancI pair had to be activated by having a molecule of ubiquitin linked to each of them (Figure 31.5). The ubiquitylation that activates the two proteins of the FancD2-FancI dimer was found to be carried out by one of the proteins in the Fanc core complex, namely FancT.

The remarkable way that FancD2 and FancI nestle together and bind to each other by way of their ubiquitins is shown in Figure 31.5. The shapes of the pair was described as saxophone-like, but their shape changes somewhat when they become associated with a DNA damage site; the shape change made them fully active (Niraj et al., 2017).

The picture that emerged was that the DNA at the replication-blocked crosslink site was first cleaned of replication proteins with help of FancI. The crosslinked site then brings in FancM, which then brings the other Fanc core complex of proteins to the crosslink site on the DNA. FancL or FancT within the core complex may then ubiquitylate the two proteins in the FancD2-FancI dimer (which is not itself part of the core complex) (Niraj et al., 2019) (Figure 31.7). The ubiquitylated dimer then proceeds to activate a variety of DNA damage-response processes (Figure 31.7) (Nepal et al., 2017).

The FancD2-FancI dimer was found to receive inputs from several DNA-damage signaling pathways and to respond by transmitting signals to several systems that repair or counteract the effects of the DNA damage (Nepal et al., 2017). The key event, initiated by the DNA damage, was to add a molecule of ubiquitin to both FancD2 and FancI in the dimer. This ubiquitylation activated the dimer to allow it to signal to the down-stream DNA damage-response processes. In brief, the ubiquitylation was attributed to FancT as part of a complex consisting of as many as 12 other Fanc proteins. DNA-binding ability in this big complex was found to reside in FancM, together with 2 other proteins. FancM was found to bind and detect places on DNA that have structural damage, such as crosslinks.

Ubiquitylation of the FancD2-FancI dimer required that both proteins of the dimer be phosphorylated, an action that was carried out by ATM, the kinase that is mutated in ataxia telangiectasia, and/or by the ATM-related kinase, ATR (Chapter 29).

After FancM and its associated proteins have recognized and bound to the DNA crosslink site, the FancD2-FancI duo comes into play. Together with some other proteins, it cuts out a segment of DNA that has the crosslink in it and prepares the cut ends of the DNA for rejoining. Rejoining of the resulting DNA double-strand ends involves homologous recombination with participation of BRCA1, which is the same as FancS (Fanc complementation group S protein) (Niraj et al., 2019). Thus, BRCA1 had a role in both breast cancer and Fanconi anemia.

References

Alter, B.P., Greene, M.H., Velazquez, I., and Rosenberg, P.S. (2003). Cancer in Fanconi anemia. Blood *101*, 2072.

Andreassen, P.R., D'Andrea, A.D., and Taniguchi, T. (2004). ATR couples FANCD2 monoubiquitination to the DNA-damage response. Genes Dev *18*, 1958-1963.

Auerbach, A.D. (1988). A test for Fanconi's anemia. Blood *72*, 366-367.

Basbous, J., and Constantinou, A. (2019). A tumor suppressive DNA translocase named FANCM. Crit Rev Biochem Mol Biol *54*, 27-40.

Boisvert, R.A., and Howlett, N.G. (2014). The Fanconi anemia ID2 complex: dueling saxes at the crossroads. Cell Cycle *13*, 2999-3015.

Che, R., Zhang, J., Nepal, M., Han, B., and Fei, P. (2018). Multifaceted Fanconi Anemia Signaling. Trends Genet *34*, 171-183.

D'Andrea, A.D. (2010). Susceptibility pathways in Fanconi's anemia and breast cancer. The New England journal of medicine *362*, 1909-1919.

de Winter, J.P., van der Weel, L., de Groot, J., Stone, S., Waisfisz, Q., Arwert, F., Scheper, R.J., Kruyt, F.A., Hoatlin, M.E., and Joenje, H. (2000). The Fanconi anemia protein FANCF forms a nuclear complex with FANCA, FANCC and FANCG. Hum Mol Genet *9*, 2665-2674.

Digweed, M., and Sperling, K. (1989). Identification of a HeLa mRNA fraction which can correct the DNA-repair defect in Fanconi anaemia fibroblasts. Mutation research *218*, 171-177.

Digweed, M., and Sperling, K. (1996). Molecular Analysis of Fanconi Anemia. BioEssays : news and reviews in molecular, cellular and developmental biology *18*.

Duckworth-Rysiecki, G., Cornish, K., Clarke, C.A., and Buchwald, M. (1985). Identification of two complementation groups in Fanconi anemia. Somat Cell Mol Genet *11*, 35-41.

Garcia, M.J., Fernandez, V., Osorio, A., Barroso, A., Fernandez, F., Urioste, M., and Benitez, J. (2009). Mutational analysis of FANCL, FANCM and the recently identified FANCI suggests that among the 13 known Fanconi Anemia genes, only FANCD1/BRCA2 plays a major role in high-risk breast cancer predisposition. Carcinogenesis *30*, 1898-1902.

Garcia-Higuera, I., Kuang, Y., Naf, D., Wasik, J., and D'Andrea, A.D. (1999). Fanconi anemia proteins FANCA, FANCC, and FANCG/XRCC9 interact in a functional nuclear complex. Molecular and cellular biology *19*, 4866-4873.

Howlett, N.G., Harney, J.A., Rego, M.A., Kolling, F.W.t., and Glover, T.W. (2009). Functional interaction between the Fanconi Anemia D2 protein and proliferating cell nuclear antigen (PCNA) via a conserved putative PCNA interaction motif. The Journal of biological chemistry *284*, 28935-28942.

Howlett, N.G., Taniguchi, T., Olson, S., Cox, B., Waisfisz, Q., De Die-Smulders, C., Persky, N., Grompe, M., Joenje, H., Pals, G.*, et al.* (2002). Biallelic inactivation of BRCA2 in Fanconi anemia. Science *297*, 606-609.

Huang, J., Liu, S., Bellani, M.A., Thazhathveetil, A.K., Ling, C., de Winter, J.P., Wang, Y., Wang, W., and Seidman, M.M. (2013). The DNA translocase FANCM/MHF promotes replication traverse of DNA interstrand crosslinks. Mol Cell *52*, 434-446.

Joenje, H., and Patel, K.J. (2001). The emerging genetic and molecular basis of Fanconi anaemia. Nat Rev Genet *2*, 446-457.

Joo, W., Xu, G., Persky, N.S., Smogorzewska, A., Rudge, D.G., Buzovetsky, O., Elledge, S.J., and Pavletich, N.P. (2011). Structure of the FANCI-FANCD2 complex: insights into the Fanconi anemia DNA repair pathway. Science *333*, 312-316.

Kottemann, M.C., and Smogorzewska, A. (2013). Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. Nature *493*, 356-363.

Kupfer, G.M., Naf, D., Suliman, A., Pulsipher, M., and D'Andrea, A.D. (1997). The Fanconi anaemia proteins, FAA and FAC, interact to form a nuclear complex. Nat Genet *17*, 487-490.

Lo Ten Foe, J.R., Rooimans, M.A., Bosnoyan-Collins, L., Alon, N., Wijker, M., Parker, L., Lightfoot, J., Carreau, M., Callen, D.F., Savoia, A.*, et al.* (1996). Expression cloning of a cDNA for the major Fanconi anaemia gene, FAA. Nat Genet *14*, 320-323.

Medhurst, A.L., Laghmani el, H., Steltenpool, J., Ferrer, M., Fontaine, C., de Groot, J., Rooimans, M.A., Scheper, R.J., Meetei, A.R., Wang, W.*, et al.* (2006). Evidence for subcomplexes in the Fanconi anemia pathway. Blood *108*, 2072-2080.

Meetei, A.R., Medhurst, A.L., Ling, C., Xue, Y., Singh, T.R., Bier, P., Steltenpool, J., Stone, S., Dokal, I., Mathew, C.G.*, et al.* (2005). A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. Nat Genet *37*, 958-963.

Meetei, A.R., Sechi, S., Wallisch, M., Yang, D., Young, M.K., Joenje, H., Hoatlin, M.E., and Wang, W. (2003). A multiprotein nuclear complex connects Fanconi anemia and Bloom syndrome. Molecular and cellular biology *23*, 3417-3426.

Nepal, M., Che, R., Zhang, J., Ma, C., and Fei, P. (2017). Fanconi Anemia Signaling and Cancer. Trends in cancer *3*, 840-856.

Niraj, J., Caron, M.C., Drapeau, K., Berube, S., Guitton-Sert, L., Coulombe, Y., Couturier, A.M., and Masson, J.Y. (2017). The identification of FANCD2 DNA binding domains reveals nuclear localization sequences. Nucleic acids research *45*, 8341-8357.

Niraj, J., Farkkila, A., and D'Andrea, A.D. (2019). The Fanconi Anemia Pathway in Cancer. Annu Rev Cancer Biol *3*, 457-478.

Pace, P., Johnson, M., Tan, W.M., Mosedale, G., Sng, C., Hoatlin, M., de Winter, J., Joenje, H., Gergely, F., and Patel, K.J. (2002). FANCE: the link between Fanconi anaemia complex assembly and activity. The EMBO journal *21*, 3414-3423.

Rego, M.A., Kolling, F.W.t., and Howlett, N.G. (2009). The Fanconi anemia protein interaction network: casting a wide net. Mutation research *668*, 27-41.

Rosenberg, P.S., Greene, M.H., and Alter, B.P. (2003). Cancer incidence in persons with Fanconi anemia. Blood *101*, 822-826.

Smogorzewska, A., Matsuoka, S., Vinciguerra, P., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Ballif, B.A., Gygi, S.P., Hofmann, K., D'Andrea, A.D.*, et al.* (2007). Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. Cell *129*, 289-301.

Strathdee, C.A., Duncan, A.M., and Buchwald, M. (1992a). Evidence for at least four Fanconi anaemia genes including FACC on chromosome 9. Nat Genet *1*, 196-198.

Strathdee, C.A., Gavish, H., Shannon, W.R., and Buchwald, M. (1992b). Cloning of cDNAs for Fanconi's anaemia by functional complementation. Nature *356*, 763-767.

Swuec, P., Renault, L., Borg, A., Shah, F., Murphy, V.J., van Twest, S., Snijders, A.P., Deans, A.J., and Costa, A. (2017). The FA Core Complex Contains a Homo-dimeric Catalytic Module for the Symmetric Mono-ubiquitination of FANCI-FANCD2. Cell reports *18*, 611-623.

Taniguchi, T., Garcia-Higuera, I., Xu, B., Andreassen, P.R., Gregory, R.C., Kim, S.T., Lane, W.S., Kastan, M.B., and D'Andrea, A.D. (2002). Convergence of the fanconi anemia and ataxia telangiectasia signaling pathways. Cell *109*, 459-472.

Vaz, F., Hanenberg, H., Schuster, B., Barker, K., Wiek, C., Erven, V., Neveling, K., Endt, D., Kesterton, I., Autore, F.*, et al.* (2010). Mutation of the RAD51C gene in a Fanconi anemia-like disorder. Nat Genet *42*, 406-409.

Vuono, E.A., Mukherjee, A., Vierra, D.A., Adroved, M.M., Hodson, C., Deans, A.J., and Howlett, N.G. (2016). The PTEN phosphatase functions cooperatively with the Fanconi anemia proteins in DNA crosslink repair. Sci Rep *6*, 36439.

Waisfisz, Q., de Winter, J.P., Kruyt, F.A., de Groot, J., van der Weel, L., Dijkmans, L.M., Zhi, Y., Arwert, F., Scheper, R.J., Youssoufian, H.*, et al.* (1999). A physical complex of the Fanconi anemia proteins FANCG/XRCC9 and FANCA. Proceedings of the National Academy of Sciences of the United States of America *96*, 10320-10325.

Walden, H., and Deans, A.J. (2014). The Fanconi anemia DNA repair pathway: structural and functional insights into a complex disorder. Annu Rev Biophys *43*, 257-278.

Weksberg, R., Buchwald, M., Sargent, P., Thompson, M.W., and Siminovitch, L. (1979). Specific cellular defects in patients with Fanconi anemia. J Cell Physiol *101*, 311-323.

Xue, X., Sung, P., and Zhao, X. (2015). Functions and regulation of the multitasking FANCM family of DNA motor proteins. Genes Dev *29*, 1777-1788.