Università degli Studi di Trieste

PhD program in MOLECULAR MEDICINE

Ph.D Thesis

Identification of putative interactors of Fanconi anaemia proteins by yeast two-hybrid system: characterization of two novel genes highly expressed during spermatogenesis.

Dr. Fabio Morgese

Anno Accademico 2006-2008 (XXI ciclo)

UNIVERSITÀ DEGLI STUDI DI TRIESTE

UNIVERSITÀ DEGLI STUDI DI TRIESTE IRCSS BURLO GAROFOLO

XXI CICLO DEL

DOTTORATO DI RICERCA IN

MEDICINA MOLECOLARE

IDENTIFICATION OF PUTATIVE INTERACTORS OF FANCONI ANAEMIA PROTEINS BY YEAST TWO-HYBRID SYSTEM: CHARACTERIZATION OF TWO NOVEL GENES HIGHLY EXPRESSED DURING SPERMATOGENESIS

SETTORE SCIENTIFICO-DISCIPLINARE: BIO 10

DOTTORANDO

Dr. **Fabio Morgese**

COORDINATORE DEL COLLEGIO DEI DOCENTI

CHIAR.MO PROF. *Giannino Del Sal,Università degli studi di Trieste*

RELATORE

CHIAR.MO PROF. *Anna Savoia, Università degli studi di Trieste*

ANNO ACCADEMICO 2007/2008

Relatore

Chiar.mo prof. *Anna Savoia*, Università degli studi di Trieste, Dipartimento di scienze della riproduzione. Via dell' Istria 65/1, 34137 Trieste.

External advisor

Chiar.mo prof. *Fulvio della Ragione*, Seconda Università degli Studi di Napoli, Dipartimento di biochimica e biofisica "F. Cedrangolo". Via L. De Crecchio n.7, 80138 Napoli

COMMISSIONE D'ESAME:

Presidente effettivo

Chiar.mo prof. *Guido Tarone*, Università di Torino, Dipartimento di genetica, biologia e biochimica. Via Nizza 52, 10126 Torino (TO)

Componente effettivo

Chiar.mo prof. *Silvia Giordano*, Università di Torino, Dipartimento di scienze oncologiche. Strada provinciale 142 di Piovesi, 10060 Candiolo (TO)

Componente effettivo

Chiar.mo prof. *Guidalberto Manfioletti*, Università degli studi di Trieste, Dipartimento di scienze della vita, edif. C11. Via L. Giorgieri 1, 34127 Trieste

Presidente supplente

Chiar.mo prof. *Franco Vittur*, Università degli studi di Trieste, Dipartimento di. scienze della vita. Via L. Giorgieri 1, 34127 Trieste

Componente supplente

Chiar.mo prof. *Stefano Piccolo*, Università di Padova, Biotecnologie mediche. Viale colombo 3, 35126, Padova (PD)

Componente supplente

Chiar.mo prof. *Stefano Gustinchic*, SISSA, settore di neurobiologia, Area science park. Strada statale 14, km 163.5, 34012 Basovizza (TS)

This thesis is dedicated to the ones I love.

It is dedicated to my beautiful wife, for her restless support in difficult moments and for believing in me without "if" and without "but". It is dedicated to my parents, for giving me the chance to realize my dreams believing in what I do. To my mother, for being an example of a woman who never gives up, and to my father, who would certainly have been proud of this work. To my sister, because I know she loves me. This thesis is also dedicated to all the ones who cared about me in these long years, particularly to my friends of a life.

Table of contents

ABSTRACT

Fanconi Anaemia (FA) is a rare human genetic disease characterized by bone marrow failure, malformations, chromosomal instability and cancer susceptibility. Thirteen genes belonging to a common pathway have been identified, but their function is still unclear even if evidence indicates a role in DNA-repair. In the attempt to gain new insights on FA-BRCA pathway, this work aimed at finding and characterizing novel putative interactors of the FA proteins. Using the yeast two-hybrid system, we screened a cDNA library of human testis and rescued two clones. Clone 54, which encoded for a putative ubiquitin-conjugating enzyme E2 (UBE2U), was first found to interact with FANCD2 and then with FANCL (E3 ubiquitin-ligase of the FA pathway), FANCC, FANCE and FANCF by direct interaction mating in yeast. Different assays indicated that the expression of this gene is limited to mouse and human testis (specifically in spermatocytes and spermatides). Interestingly, even mouse Fancd2 showed a high expression level in these two cell types, supporting the hypothesis of an interaction between the two proteins and a role of the FA-BRCA pathway during spermatogenesis. In order to confirm the binding between UBE2U and FANCD2, we transiently transfected cell lines with a tagged UBE2U. However, since we failed to detect the protein at any level, we tried to validate the interaction using mouse testis extract. Using the specific antibody we generated, we were however not able to confirm the binding, but, before excluding definitively the interaction, we should further investigate using more suitable antibodies. Clone 4, encoding for a novel putative exonuclease (ISG20L2), was instead found to interact with the C-terminus of FANCG. Though it was ubiquitously present at low levels in all the cells tested, it showed a stronger expression in mouse testis. In transiently transfected cells, ISG20L2 was detected primarily in nucleoli by immunofluorescence, but it was revealed also in the cytoplasmic fraction by western blot. Both nuclear and cytoplasmic distributions of the protein were confirmed at endogenous levels, after production of a specific antibody. Coimmunoprecipitation studies between ISG20L2 and FANCG did not confirm their interaction, but this might be in agreement with a recent report for ISG20L2 as a nucleolar exoribonuclease, not directly involved with DNA-repair.

Chapter 1.
INTRODUCTION

1.1 Fanconi anaemia: an historical perspective.

Shortly after beginning his paediatric training at the Kinderspital in Zurich in 1920, Guido Fanconi encountered an unusual disorder that he described in a case report published in 1927 (Fanconi 1927). In a family with five children, three brothers had died of a severe condition that resembled pernicious anaemia, a special kind of anaemia characterized by macrocytic red blood cells, increased haemolysis and low levels of serum vitamin B12. In all three boys the disease had become manifest between the ages of five and seven, had had a fatal outcome and was associated with a congenitally small head (microcephaly), good intelligence, intensive brown skin pigmentation, cutaneous haemorrhage, hypoplasia of the testes, squint and very lively tendon reflexes. The blood picture was typical of pernicious anaemia, but with no signs of increased haemolysis. On the basis of a fourth case reported by Uehlinger (Uehlinger 1929), Fanconi realized in 1929 that the disorder affected not only erythropoiesis but all haematopoietic cell lines. Further cases made it clear to him that the haematological manifestations were merely the usually fatal late symptoms of a highly complex pathological process in a disorder often associated with various physical abnormalities. In 1931 the famous haematologist Otto Naegeli proposed that the syndrome should be called Fanconi anaemia (FA). Feat of Fanconi was to recognize that this strange collection of symptoms was a disease in its own right and this was made possible only by his precise observation and meticulous documentation. In the 23-page original paper of 1927, he filled 12 pages just with the symptoms of his three patients (Fanconi 1927). In the following decades there was much speculation about the aetiology of the disease and its pattern of inheritance. The topic occupied Fanconi for 40 years. He did not initially believe the widely held view that it was an autosomal recessive disorder. It seemed unlikely to him that the incredible multiplicity and variability of Fanconi anaemia could be explained by the mutation of a single gene. Moreover, the first epidemiological studies of Weicker and Fichsel showed a ratio of about 1:1 between diseased and healthy children in affected families, a far cry from the ratio of 1:4 typical for autosomal recessive inheritance (Fanconi 1967; Fanconi 1970). As this is a rare disease, if there were an autosomal recessive inheritance pattern then parental consanguinity, that is, descent from the same ancestor, should be very common, which was not the case in experience of Fanconi. Also, Fanconi could not explain the phenomenon that had been described by Imerslund, O'Neill and

Varadi, who had reported two families in which mother and child suffered from Fanconi anaemia (Imerslund 1953). Therefore, in 1964 Fanconi published his hypothesis that Fanconi anaemia was caused by a chromosomal translocation (Fanconi 1964). This hypothesis was supported by the data he had so carefully collected, and also had the backing of geneticists, particularly as it was already known that patients with Fanconi anaemia had a normal number of chromosomes. At this time high-resolution technology for the examination of chromosomes was not widely available. Scientists all over the world began to search for this ominous translocation. They did not find it, but the additional thrust given to chromosome research bore fruit nevertheless. In 1964 and 1965 several research groups were able to show that Fanconi anaemia patients suffered from chromosomal instability that occurred spontaneously but that could also be induced, particularly by DNA cross-linking substances: for example, mitomycin C (MMC) or diepoxybutane (DEB) (Schroeder, Anschutz et al. 1964; Schmid, Scharer et al. 1965; Schroeder 1966; Schroeder 1966; Swift and Hirschhorn 1966; Auerbach and Wolman 1976). Therefore, endogenous and exogenous factors can lead to chromosome breaks and their consequences at any time during prenatal and postnatal development. In eyes of Guido Fanconi this, together with the assumption that there are probably many intermediate stages between maximum chromosome fragility and maximum stability, explained the incredible phenotypic variability between different patients with Fanconi anaemia and their susceptibility to cancer. He was particularly struck by the high incidence of acute myeloid leukaemia (AML) in patients with Fanconi anaemia, but various other cancers, particularly of epithelial origin, had also found unusually often in patients with the syndrome (Fanconi 1967).

Fanconi anaemia is, despite doubts of Fanconi, a recessive disorder. Heterozygotes are not affected, but it seems that certain heterozygote mutations in individual Fanconi anaemia genes might occasionally also be associated with the development of certain cancers (Rischewski, Clausen et al. 2000; Condie, Powles et al. 2002; Liede, Karlan et al. 2004; Tischkowitz, Morgan et al. 2004). The best known example is *BRCA2* (also known as *FANCD1*). On one hand, homozygote mutations in this gene are the molecular genetic correlate of the FA-D1 subtype of Fanconi anaemia. On the other hand, heterozygosity for certain mutations in this gene is found in patients with certain forms of hereditary breast, ovarian and pancreatic cancer (Wooster, Bignell et al. 1995; Lancaster, Wooster et al. 1996; Stratton and Wooster 1996; van der Heijden, Yeo et al. 2003; van der Heijden, Brody et al. 2004). On the other hand, it is currently thought that the Fanconi anaemia pathway might

also be influenced by epigenetic factors (Dhillon, Shahid et al. 2004; Narayan, Arias-Pulido et al. 2004) and this was something Fanconi had already suspected.

Since the identification of Fanconi anaemia as a distinct clinical entity, several milestones stand out in its subsequent characterization. The first is the discovery that cultured lymphocites from FA patients show spontaneous chromosomal instability (Schroeder, Anschutz et al. 1964). The second is the finding that FA cells are hyper-responsive to the chromosome breaking and anti-proliferative effects of cross-linking agents such as mitomycin C (MMC) and diepoxybutane (DEB) (Sasaki and Tonomura 1973; Auerbach and Wolman 1976; Ishida and Buchwald 1982). This hallmark of FA cells allowed to develop a specific diagnostic test for the disease, overcoming most of the difficulties related to the extreme variability of the clinical features. The third milestone comes with the finding that a high degree of genetic heterogeneity underlies the disease, as shown by cell-fusion and complementation studies (Strathdee 1992; H Joenje 1997; Joenje H 2000). The fourth milestone is the design of a complementation cloning method, directly derived from the complementation assay, that enabled the cloning of the first FA genes (Strathdee 1992). The last milestone in FA research field can be considered the purification of the 'FA-core complex' (Meetei, de Winter et al. 2003; Meetei, Sechi et al. 2003), that significantly contributed to discovery the current thirteen FA proteins interaction network.

1.2 Fanconi anaemia: the disease.

1.2.1: Inherited bone marrow failure syndromes, caretaker gene diseases and clinical features of Fanconi anaemia.

Fanconi anaemia (FA) is a rare autosomal and X-linked human recessive disease. It is highly heterogeneous, caused by mutations in at least 13 genes (*FANCA*, *FANCB*, *FANCC*, *FANCD1*, *FANCD2*, *FANCE*, *FANCF*, *FANCG*, *FANCI*, *FANCJ*, *FANCL*, *FANCM* and *FANCN*) each of which associated to different 'complementation groups' (FA-A, FA-B, FA-C, FA-D1, FA-D2, FA-E, FA-F, FA-G, FA-I, FA-J, FA-L, FA-M and FA-N). It is characterized by bone marrow failure, various developmental abnormalities, genomic instability and a high incidence of malignancies. Due to its clinical and cellular features, FA is considered a member of at least two classes of cancer predisposition syndromes: inherited

bone marrow failure syndromes (IBMFSs) and caretaker gene diseases (or DNA-repair diseases). IBMFSs are characterized by impaired haematopoiesis and aplastic anaemia that lead to bone marrow failure and cancer predisposition. Most inherited marrow failure syndromes are also associated with a range of congenital anomalies. Amegakaryocytic thrombocytopenia, Diamond-Blackfan anaemia, dyskeratosis congenita, severe congenital neutropenia, Shwachman-Diamond syndrome are some examples of IBMFS. Hereditary diseases that instead feature genomic instability in combination with a strong propensity to cancer are called caretaker gene diseases (or DNA-repair diseases). Xeroderma pigmentosum, hereditary non-polyposis colorectal cancer, Bloom syndrome, Werner syndrome, Nijmegen breakage syndrome, ataxia telangiectasia-like disorder and ataxia telangiectasia are examples of caretaker gene diseases. Their common feature is an impaired capacity to maintain genomic integrity, which results in the accelerated accumulation of key genetic changes that promote cellular transformation and neoplasia. Cancer predisposition in these two classes of diseases is an indirect result of the primary genetic defect.

The prevalence of FA is estimated to be 1 to 5 per million, and heterozygous carrier frequency is estimated to be 1 in 300, although the true frequency is probably higher (Auerbach AD 2001; Joenje and Patel 2001). FA has been reported in many ethnic groups (Alter 1994) and founder mutations have been described in Afrikaaners (Rosendorff, Bernstein et al. 1987) and Ashkenazy Jews (Verlander, Kaporis et al. 1995) amongst others. Patients with FA show extreme clinical heterogeneity (Alter 1994; Auerbach AD 2001). The median age at diagnosis is 6.5 years for male patients and 8 years for female patients, although the age at diagnosis ranges from birth to 48 years. Survival was 19 years in cases reported between 1981 and 1990 and the median survival age has improved to 30 years in patients reported between 1991 and 2000 (Alter 2003).

The most important clinical features of FA are haematological and these are responsible for the greatest morbidity and mortality in homozygotes. At birth, the blood count is usually normal and macrocytosis is often the first detected abnormality. This is followed by thrombocytopenia and anaemia, and pancytopenia typically presents between the ages of 5 and 10 years, with the median age of onset being 7 years (Auerbach and Allen 1991; Butturini, Gale et al. 1994). Patients with FA develop bone marrow failure typically during the first decade of life. The actuarial risk of developing bone marrow failure is 90% by 40 years of age. A retrospective study of 145 FA cases published in 2004 showed that certain congenital abnormalities were potential risk indicators for the development of bone marrow failure. Patients with abnormal radii had a 5.5 times increased risk of developing

bone marrow failure compared with those cases with normal radii and the risk also increased with the number of heart, kidney, head, hearing and developmental abnormalities present (Rosenberg, Huang et al. 2004). In those children without congenital abnormalities the development of haematological abnormalities can be the first presenting feature of FA and can occasionally be the presenting feature in adulthood (Liu, Auerbach et al. 1991; Cavenagh, Richardson et al. 1996).

The non-haematological clinical phenotype in FA is highly variable and individuals can have a wide multiplicity of clinical abnormalities (table 1) (Auerbach 1999; De Kerviler, Guermazi et al. 2000). Generalized skin hyperpigmentation, 'cafe´ au lait' spots and areas of hypopigmentation are often present and may sometimes be the only feature present. Skeletal abnormalities commonly include radial ray defects such as hypoplasia of the thumbs and radial hypoplasia; other skeletal defects that may occur include congenital hip dislocation, scoliosis and vertebral anomalies. Around one-third of FA patients have renal anomalies including unilateral renal aplasia, renal hypoplasia, horseshoe kidneys or double ureters. FA is associated with altered growth both in utero and postnatally; low-birth weight is common and the median height of FA individuals lies around the fifth percentile. This can sometimes be related to growth hormone deficiency or hypothyroidism (Wajnrajch, Gertner et al. 2001). Microphthalmia, microcephaly, conductive deafness and developmental delay are all often present. Males have a high incidence of genital abnormalities such as hypogenitalia, undescended testes and hypospadias with infertility being the norm, although there have been reports of males with FA fathering children (Liu, Auerbach et al. 1991). The FA phenotype can differ considerably within families; a report of four FA patients from two related consanguineous families who all had the same FANCA mutation demonstrated a wide variation in birth weight, skin pigmentation and the severity of skeletal, renal and genital abnormalities (Koc, Pronk et al. 1999). Moreover, up to one-third of FA cases do not have any obvious congenital abnormalities, and are only diagnosed when another sibling is affected or when they develop haematological problems (Giampietro, Verlander et al. 1997).

Abnormality	**Frequency (%)**
Skeletal (radial ray, hip, vertebral scoliosis, etc..)	71
Skin pigmentation ('cafè au lait', hyper- and hypopigmentation)	64
Short stature	63
Eyes (microphtalmia)	38

Renal and urinary tract	34
Male genital	20
Mental retardation	16
Gastrointestinal	14
Cardiac	13
Hearing	11
Central nervous system	8
No abnormalities	30

Table 1: Frequency of abnormalities in Fanconi anaemia

1.2.2: Cancer in FA.

FA patients have a high risk of leukaemia and solid tumours. A literature review of over 1300 reported cases from 1927 to 2001 found that 9% had leukaemia, mainly acute myeloid leukaemia (AML), 7% had myelodysplastic syndrome (MDS), 5% had solid tumours and 3% had liver tumours (Alter 2003). The distribution of the solid tumours was very unusual compared to the general population, since more than half were of the head and neck, oesophagus or vulva. The liver tumours were generally not malignant, and were thought likely to be related to prolonged androgen therapy used in FA patients treatment. A retrospective cohort study of 145 FA patients published in 2003 showed that nine had developed AML and 14 had solid tumours, most of which were of the head and neck, oesophagus, vulva or cervix (Rosenberg, Greene et al. 2003). The ratio of observed to expected cancers was 50 for all cancers, 48 for solid tumours and 785 for leukaemia. The cumulative incidence to age 48 was 10% for leukaemia and 29% for a solid tumour. A larger study of 754 patients in the International Fanconi Anemia Registry (IFAR) found neoplasms in 23%, with haematological neoplasms (mainly AML or MDS) in 15.9% and non-haematological tumours in 10.5%, with about 3% of patients having more than one neoplasm during their lifetime (Kutler, Singh et al. 2003). The most common non-haematological tumours were squamous cell carcinomas (SCC) of the head and neck (HNSCC), vulva and cervix. The cumulative incidence of haematological and non-haematological neoplasms was 33 and 28% respectively by age 40. The standardized incidence ratio for HNSCC in the IFAR registry was 500, with a median patient age of 31 years (Kutler, Auerbach et al. 2003).

These patients had poor tolerance for radiotherapy and chemotherapy, and the tumours were aggressive. Examples have been reported of young cases of apparently sporadic HNSCC with no known risk factors, who had a severe reaction to radiotherapy and who were later diagnosed as FA (Alter, Joenje et al. 2005). There is also evidence that conditioning for stem cell transplants in FA patients is associated with an increased risk of SCC of the head, neck and oesophagus (Rosenberg, Socie et al. 2005).

Clinical and molecular characterizations of FA-D1 families with biallelic mutations in *BRCA2* have revealed an even stronger cancer-prone phenotype in at least 16 confirmed D1 kindreds reported up to 2006 (Howlett, Taniguchi et al. 2002; Offit, Levran et al. 2003; Hirsch, Shimamura et al. 2004; Wagner, Tolar et al. 2004; Reid, Renwick et al. 2005). Early onset leukaemia was common, with a median onset at 2.2 years of age compared to 13.4 years for all other FA patients in the IFAR registry (Wagner, Tolar et al. 2004); this was mainly AML, but there were also cases of T- and B-cell acute lymphoblastic leukaemia. There were nine cases with brain tumours (mainly medulloblastomas), and five with Wilms tumours (summarized in (Reid, Renwick et al. 2005)). It was proposed initially that FA-D1 would occur in individuals who had at least one partially functional *BRCA2* allele (Howlett, Taniguchi et al. 2002). This would be consistent with the fact that complete knockout of *brca2* is an early embryonic lethal in the mouse (Ludwig, Chapman et al. 1997; Sharan, Morimatsu et al. 1997), whereas some mice homozygous for a C-terminal truncating brca2 mutation are viable but succumb to lymphomas (Connor, Bertwistle et al. 1997; Friedman, Thistlethwaite et al. 1998). The frequency of *BRCA2* mutation carriers in the general population is estimated to be around 1 in 1000 (Thompson D 2004), which corresponds to a mutant allele frequency of 0.0005 and a predicted homozygote frequency in the absence of selection of 1 in 4 million. Considered the incidence of FA in most populations as around 1 in 250,000 and FA-D1 accounts for about 3.3% of all FA cases (Levitus, Rooimans et al. 2004), this corresponds to an incidence of FA-D1 of 0.53 in 4 million, which is about half of the predicted incidence. Thus, it is possible that only a subset of mutant *BRCA2* alleles is compatible with survival to term in humans. However, an assessment of the location of *BRCA2* mutations in FA-D1 patients in comparison with *BRCA2* carriers does not support a bias towards carboxy-terminal mutations in FA-D1 (Reid, Renwick et al. 2005). Also, the nature of the mutations in many of the reported FA-D1 cases would predict early truncation of the BRCA2 protein expressed from both alleles, and its absence has been confirmed in some cases by Western blot analysis (Hirsch, Shimamura et al. 2004).

Generally, the association of haematopoietic stem cell failure and a high risk of cancer with a chromosomal instability syndrome is not surprising. Excessive chromosome breakage might be expected to lead to unrepaired DNA-damage and apoptosis, or to mutations which could confer a selective growth advantage to a progenitor cell. Why cells of myeloid origin are particularly affected, leading to MDS or AML, is however not known; perhaps myeloid precursors are particularly susceptible to forms of DNA-damage that are recognized or repaired by the FA pathway. The unusual spectrum of solid tumours, with a preponderance of squamous cell carcinomas of the oral cavity and vulva or cervix is also unexplained. Evidence both for and against an involvement of human papilloma virus (HPV) in SCCs from FA patients has been presented (Kutler, Wreesmann et al. 2003; van Zeeburg, Snijders et al. 2004). A possible route to carcinogenesis is also an interaction of the environmental toxins to which the mucous membranes of the oral cavity and genital areas are exposed with a constitutional deficiency of the FA patient to detoxify such agents or to repair DNA-damage consequent to the exposure. Moreover, the highly cancer-prone phenotype of FA-D1 patients is not unexpected given the central role occupied by BRCA2 in homologous recombination repair. However, the high incidence of brain tumours in these patients is not characteristic of FA patients in general, or of *BRCA2* mutation carriers, and remains to be explained, as does the association with familial Wilms tumour (Reid, Renwick et al. 2005).

1.2.3: Diagnosis and assessment.

Because FA has such a varied phenotype, it is difficult to diagnose solely on the basis of clinical features and if the diagnosis is suspected, it must be confirmed primarily by chromosome breakage studies and subsequently by molecular analysis. FA cells show increased spontaneous chromosome breakage, but this is highly variable and overlaps with the range for normal cells. Adding DNA cross-linking agents, such as mitomycin C (MMC) or diepoxybutane (DEB), increases the number of chromosome breaks with distinct ranges for normal and FA cells. This provides the basis for a widely used standard diagnostic test (Auerbach and Wolman 1976; Auerbach 1993). In this test, peripheral blood T-lymphocytes are stimulated to divide with a mitogen and then exposed to the DNA cross-linking agent such as DEB or MMC. Metaphase spreads of FA cells have increased numbers of chromosomal breaks per cell and an increased fraction of cells with breaks (Fig.1). Chromosome breakage studies can be carried out on amniotic cells, chorionic villus cells or

foetal blood and therefore it can be used also for prenatal diagnosis (Auerbach, Sagi et al. 1985). Although MMC/DEB testing is highly specific for FA, interpretation is complicated in cases of somatic mosaicism (Dokal, Chase et al. 1996), where a haematopoietic stem cell has undergone gene correction and its progeny have repopulated the bone marrow. In these cases, there may be a normal chromosomal stability response in lymphocytes derived from the corrected stem cell so, if mosaicism is suspected, skin fibroblasts should be used for the breakage assay (Alter, Joenje et al. 2005). Another diagnostic technique with comparable accuracy to chromosome breakage studies is based on flow cytometric analysis of cells exposed to DNA cross-linking agents. It measures the prolonged progression through and arrest within the G2-phase, which is another characteristic of FA cells (Schindler, Kubbies et al. 1985; Berger, Le Coniat et al. 1993; Seyschab, Friedl et al. 1995). Such an approach has the advantage that it is less time-consuming, does not require cytogenetic expertise and it may also sometimes be helpful when there are discrepancies on DEB testing (Toraldo, Canino et al. 1998). However, it is not reliable in cases with concurrent myelodysplasia or leukaemia (Berger, Le Coniat et al. 1993; Seyschab, Friedl et al. 1995). Serum α-fetoprotein (AFP) levels are consistently elevated in FA patients irrespective of whether liver abnormalities are present and this could be used as a fast and cheap screening test in the sizeable group of individuals with early onset leukaemia or cancer, or other FA-like features, such as short stature, thumb abnormalities, skin pigmentation, or macrocytic anaemia (Cassinat, Guardiola et al. 2000). However, diagnostic precision varies with the type of AFP assay technique used so new AFP assays must first be carefully validated (Cassinat, Darsin et al. 2001).

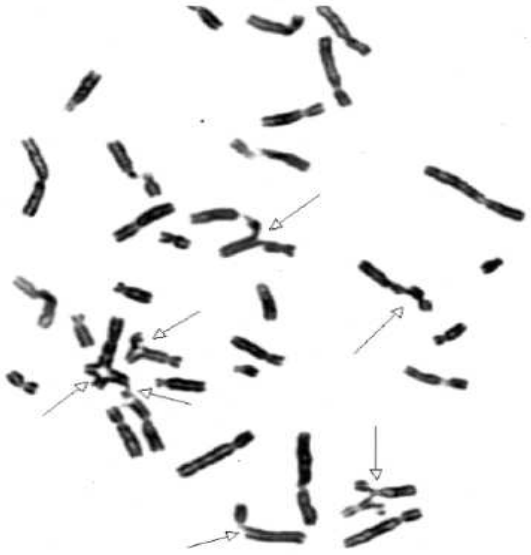

Figure 1:**Diepoxybutane-induced chromosomal breakage in a metaphase lymphocyte from a patient with Fanconi anaemia**. Chromosomes with breaks and fusions are indicated with arrows

1.2.4: Screening of FA complementation groups: methods and its importance.

As described, one of the most striking features of FA is its high heterogeneity. To answer whether one or more genes caused FA, a complementation analysis method was developed in 1992 in the laboratory of M. Buchwald (Strathdee 1992), introducing also the designation of 'complementation groups' for FA patients. In particular, the methodology took advantage of cells hybridization techniques and MMC or DEB sensitivity of FA cells. In this approach, a hybrid cell line generated by fusing lymphoblasts from unrelated FA patients showed, after cross-linking treatment with MMC, either sensitive traits or resistance (complementation of the defect). In the first case, patients were classified to belong to the same 'complementation group', having defect in the same gene and consequently not being able to revert the phenotype in the hybrid cell line. In the second case, they belonged to different complementation groups, having defects in different genes and therefore being able to compensate (complement) each other in the hybrid cell line (Fig. 2). As mentioned, this discovery is one of the milestones in the FA research field since it shed light upon the genetic heterogeneity of the disease and it allowed the cloning of several FA genes, as we will discuss later. In addition, it enabled the determination of complementation groups for each patient, which has showed to be increasingly important (Shimamura and D'Andrea 2003). In fact, some patients with other rare chromosome breakage syndrome, such as Nijmegen breakage syndrome, can have a positive DEB test. Additionally, some relation between genotype and phenotype has been reported (Nakanishi, Taniguchi et al. 2002); for example, patients in complementation group C (FA-C) had a significantly poorer survival than patients in group A (FA-A) and G (FA-G) (Kutler, Singh et al. 2003); FA-D1 subtype, as previously mentioned, is instead associated with increased predisposition to medulloblastoma, Wilms tumour, acute leukaemia in early childhood and is clinically different from other subtypes (Offit, Levran et al. 2003; Hirsch, Shimamura et al. 2004; Wagner, Tolar et al. 2004). Moreover, family members of patients with FA-D1 may be carriers of mutations in *BRCA2* and may be predisposed to cancers (Wagner, Tolar et al. 2004). Therefore, to confirm the diagnosis, to distinguish FA from other chromosome

breakage disorders and to manage each patient with FA and family better, FA subtyping is very important and should be performed routinely. However, complementation analysis is laborious and time consuming, therefore not suited for rapid screenings. As a valid alternative to this method, in 2002 was reported an assay based on the utilization of retroviruses (Hanenberg, Batish et al. 2002). In this approach, retroviruses expressing FA cDNAs were used to correct the phenotype of T-cells from FA patients, and this proved to be a rapid and accurate manner to determine complementation groups. Another important technique to subtype FA patients exploited nowadays is FANCD2 immunoblotting. It is based on the observation that, as we will discuss afterwards, the FA pathway can be divided in three levels and has a crucial point in the monoubiquitination of FANCD2 protein (Shimamura, Montes de Oca et al. 2002). This assay can be used in conjunction with retroviral techniques or direct gene sequencing to provide a rapid diagnostic and subtyping test (Shimamura and D'Andrea 2003).

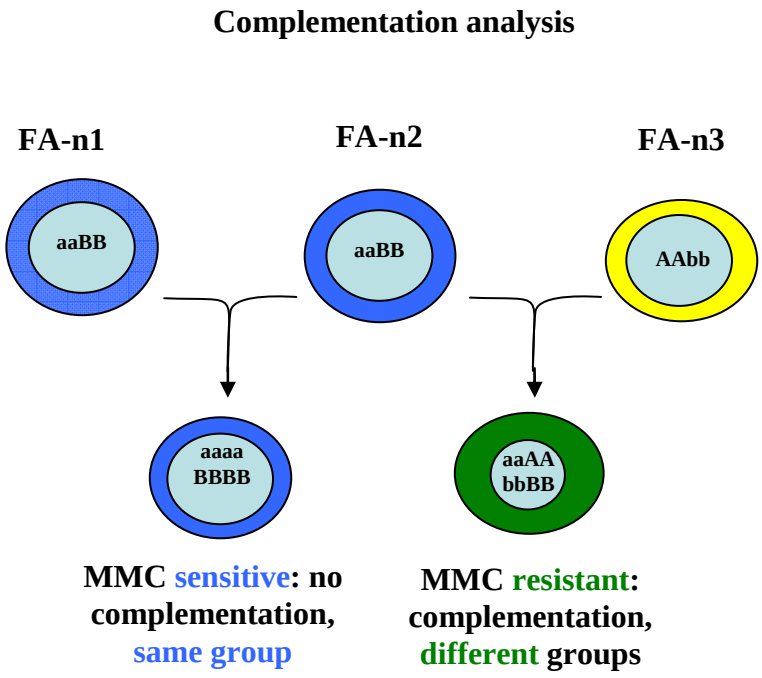

Figure 2. Complementation analysis revealed genetic heterogeneity among FA patients. MMC sensitive Epstein Barr virus immortalized B-lymphoblasts from different FA patients are fused to form a hybrid cell line. The resulting hybrid is either MMC sensitive or MMC resistant. Complementation of the MMC sensitive phenotype is only seen when the two cell lines have a defect in different genes (indicated with letters in lower case) and in that situation the patients are said to belong to different 'complementation groups'

1.2.5: Haematological management and haematopoietic stem cell transplantation.

If there is no haematopoietic defect at time of diagnosis, haematological monitoring may be required once per year, but as the patient becomes older and develops haematological complications, haematologists play an increasingly central role. Many patients who develop bone marrow failure initially respond to treatment with androgens and haematopoietic growth factors (Rackoff 1999). Oral oxymetholone is the usual recommended androgen and trilineage responses can be seen in approximately 60% of cases. However, oxymetholone has side effects which include masculinization, acne, hyperactivity, growth spurt followed by premature closure of the epiphyses resulting in short stature, deranged liver enzymes, hepatic adenomas and potential risk of hepatic adenocarcinomas. Therefore, close monitoring, particularly of liver function (liver enzymes, bilirubin, AFP every 2–3 months and yearly liver ultrasound scans), is mandatory whilst on androgen therapy. Growth factors such as G-CSF and GM-CSF can improve haematopoiesis either in conjunction with androgens (particularly if the patient is neutropenic) or if androgens have failed. They are usually avoided in patients with clonal cytogenetic abnormalities and are discontinued if such abnormalities develop while on therapy because of the potential risk of inducing or promoting leukaemia.

Eventually, most patients become refractory to oxymetholone therapy and the natural history of FA ends in lethal bone marrow failure. The definitive treatment currently available to restore normal haematopoiesis when marrow failure or clonal haematopoietic abnormality occurs is haematopoietic stem cell transplantation (HSCT). As expected, HSCT is however unable to modify non-haematopoietic defects and abnormalities occurring in FA patients. Donor stem cells can be from bone marrow, peripheral blood or umbilical cord blood. The very high sensitivity to oxidative agents and DNA cross-linking agents of FA cells confers very peculiar tissue fragility. Then, when these patients undergo HSCT, they experience unusual and severe toxicities including cardiac myopathy, gastrointestinal haemorrhage, haemorrhagic ystitis and radiation skin burns. These toxicities mean that the classical myeloablative conditioning regimen must be avoided, as demonstrated by Gluckman and colleagues (Berger, Bernheim et al. 1980; Gluckman, Devergie et al. 1980; Gluckman, Devergie et al. 1983), on account of the toxicity induced by high-dose cyclophosphamide (CY) or high dose irradiation. In the same way, the occurrence of either severe acute or

chronic graft versus host disease (GVHD) leads to greater tissue damage than that occurs in non-FA patients. Thus, for decades, the development of adapted low-intensity conditioning regimen and selection of graft associated to the lowest GVHD risk were the main goals of physicians.

In 1998, a group from St. Louis described the long-term follow-up of 50 FA patients transplanted from a sibling donor after low-dose CY (20 mg/kg total dose) and thoraco-abdominal irradiation (500 cGy) as conditioning regimen (Socie, Devergie et al. 1998). With this treatment, 74 and 58 % of patients survived at 54 and 100 months, respectively, after transplantation. The occurrence of chronic GVHD and the use of more than 20 pre-transplant transfusions appeared to have significant adverse impact on survival by multivariate analysis. During the same period, a group from Seattle reported similar results using reduced intensity conditioning regimen without any irradiation. In addition, in this study the incidence of both acute and chronic GVHD appeared to be lower than in the French study (Flowers, Zanis et al. 1996). Then, a conditioning regimen avoiding any radiation and including anti-thymoglobulins was developed. However, such a conditioning regimen could not prevent any solid tumor occurrence because of higher risk of squamous cell carcinoma in all FA patients, including non-transplanted ones. Essentially, as summarized by Gluckman and Wagner in a recent review (Gluckman and Wagner 2008), the best conditioning regimen must prevent rejection in population of patients who have received multiple transfusions, must limit early and late toxicities and must minimize the risk of GVHD. In addition, it seemed very important to plan transplantation before the appearance of abnormal clonality by at least a yearly bone marrow examination. Because of the strong side effects induced *in vitro* by alkylating agents such as CY on FA cell lines, it appeared interesting to develop a conditioning regimen without this type of chemotherapy. In 1997, a group from Israel published the first case report using a fludarabine-based conditioning regimen for an FA child in leukaemic transformation (Kapelushnik, Or et al. 1997). They were followed by many groups all over the world, and currently, fludarabine-based conditioning regimens are considered as the gold standard for FA patient transplantation. In 2007, Wagner and colleagues (Wagner, Eapen et al. 2007) reported the long-term results in 98 FA patients who underwent unrelated bone marrow transplant from 1990 to 2003 in centers collaborating with the National Marrow Donor Program. They showed higher probabilities of neutrophil and platelet recovery after fludarabine-containing regimens than non-fludarabine containing regimens did. Similarly, transplantation related mortality (TRM) and 3-year adjusted overall survival rates appeared better after fludarabine based conditioning regimens than non-

fludarabine based regimens (65 vs 24% and 13 vs 52%, respectively). The other significant poor prognostic factors for mortality in this paper were patient age more than 10 years, cytomegalovirus (CMV) seropositivity and more than 20 blood product transfusions before transplantation. The authors emphasized practice changes including fludarabine-containing conditioning regimens to obtain strong T-cell depletion and early transplantation in FA patients exhibiting marrow failure. Recently, Yabe et al. (Yabe, Yabe et al. 2007) performed *in vitro* experiments that demonstrated the rationale for using fludarabine in these patients. The difference in breakage frequencies between FA and non-FA patient lymphocytes for cultures treated with fludarabine was not statistically significant.

Nevertheless, some teams still remained satisfied by the results obtained with the 'old' conditioning regimen based on CY, thoraco-abdominal radiation and antithymocyte globulin. For example, Cincinnati Children's Hospital Medical Center reported 89% of actuarial survival at 10 years among 35 FA patients who received matched sibling donor HSCT in a single center experience (Farzin, Davies et al. 2007). In the same way, the experience reported in Saudi Arabia appeared interesting after conditioning containing CY, antithymocyte globulin and TBI at 450 cGy for FA patients with myelodysplasia (Ayas, Al-Jefri et al. 2008). Finally, drugs associated with fludarabine and addition of ionizing irradiations still remain to be discussed, according to donor type (either related or unrelated) and/or according to marrow clonal abnormality occurrence before transplantation. In 2006, Bitan et al. (Bitan, Or et al. 2006) reported excellent results regarding seven patients who underwent HSCT from either matched sibling or matched unrelated donor after fludarabine-based reduced-intensity conditioning regimen without radiation. In this short cohort, all seven patients were alive with 100% performance status. At the time of publication, no patient had developed secondary malignancy. Other groups still use the fludarabine-based conditioning regimen containing low-dose radiation (200–400 cGy) when patients are transplanted from unrelated donors (Locatelli, Zecca et al. 2007). For patients without matched sibling or unrelated donor, unrelated partially matched cord blood might represent an acceptable alternative.

For those FA patients who survive the haematological complications, follow-up surveillance for solid malignancies is increasingly important (Alter 2003). A study of secondary cancers in 700 aplastic anaemia patients who received an allogeneic HSCT highlighted FA as an independent risk factor for developing malignancy, with a predicted risk of 42% at 20 years after transplant (Deeg, Socie et al. 1996). A retrospective study of 37 FA patients who received an allogeneic HSCT reported a high risk of secondary

malignancies, notably head and neck carcinomas, which had a cumulative incidence of 53% 10 years after SCT and occurred exclusively in FA patients who had developed acute or chronic GVHD (Guardiola, Socie et al. 2004). Surveillance for solid tumours has been advocated; in females this would be at an annual gynaecological follow-up to detect cervical and vulval tumours and both sexes should be kept under surveillance for oropharyngeal, oesophageal and hepatic tumours (Alter, Greene et al. 2003; Lowy and Gillison 2003). However, at present, it is unknown how effective such surveillance measures would be in FA patients. Cancer management based on the current evidence from FA patients tends to maximize surgical intervention whilst keeping chemotherapy and radiotherapy use to a minimum.

1.2.6: Gene therapy.

The serious haematological complications requiring HSCT mean that FA would be an ideal candidate disease for gene therapy, also because unaffected matched sibling donors are unavailable to the majority of the patients and the results of alternative donor transplantation remain unsatisfactory (Wagner, Eapen et al. 2007). Furthermore, the incidence of somatic mosaicism in FA patients suggests that FA cells corrected using gene therapy may have a selective advantage over defective cells. Gene therapy programmes were initiated in the 1990s but very few patients were treated with this approach and in no case has been reported a long-term haematopoietic reconstitution. The major problem is the low efficiency of transfection of the target cells, which in turn depends on the scarcity of stem cells (Kelly, Radtke et al. 2007) and on the viral vectors. At present, γ-retroviral (Jacome, Navarro et al. 2006) and lentiviral vectors (Yamada, Ramezani et al. 2003) seem to provide encouraging results in transfecting FA bone marrow cells, but a fully convincing strategy capable of expanding the cellular target still seems difficult to find. The newest approach based on the use of agents against proinflammatory cytokine TNF-α (Etanercept) is now currently under evaluation. The rationale of this strategy is that FA haematopoietic cells display excess apoptosis in response to TNF-α, which acts via increased production of reactive oxygen species, which in turn are harmful to FA marrow cells (Dufour and Svahn 2008).

1.3 Genetics and molecular aspects of Fanconi anaemia

1.3.1: Cloning Fanconi anaemia genes: glimpses of a common functional pathway.

The extreme rarity of the disease, together with its genetic heterogeneity, have long been obstacles to the cloning of FA genes until 1992. In that year, as anticipated, the laboratory of Manuel Buchwald developed a cloning method, directly derived from the complementation assay used to assign complementation groups to FA patients (Strathdee 1992). The functional complementation cloning system relied on the capacity of a cDNA library plasmid, that expressed a normal copy of the defective FA gene, to correct the MMC-sensitive phenotype of a FA lymphoblast cell line (Fig. 3). The success of this approach depended on several factors. Firstly, the cDNA library used to initially transfects the FA cells contained the relevant cDNA with its open reading frame intact. Secondly, the FA protein was not toxic when overexpressed. Thirdly, the cell line could be transfected (since human lymphoblasts are typically reluctant to take up exogenous DNA). Lastly, the transfected cell line did not spontaneously revert to MMC resistant at a significant rate. Despite these limitations, the method enabled the cloning of five FA genes: FANCC (Strathdee 1992), FANCA (Lo Ten Foe JR 1996), FANCG (de Winter JP 1998), FANCF (de Winter JP 2000) and FANCE (de Winter JP 2000). Additionally, FANCA was independently identified by positional cloning (Consortium 1996). A third method which combined features of both the complementation and positional cloning approaches led instead to the cloning of FANCD2 (Timmers C 2001).

The six genes identified up to 2001 resulted scattered throughout the genome and did not share any common feature. Database searches for all the predicted proteins revealed that significant conservation only occurred between vertebrate sequences, with the notable exception of FANCD2, that showed highly significant alignment scores with sequences found in lower organisms like *Drosophila melanogaster*, *Caenorhabditis elegans* and *Arabidopsis thaliana*. The relatively high degree of evolutionary conservation of FANCD2 indicated that this protein might act in a conserved pathway and other FA genes had probably been superimposed later during the evolution. The lack of significant functional motifs in any of the FA proteins discovered precluded predictions about their molecular functions, but FANCD2 provided a first clue for a DNA-repair defect in FA cells. In fact,

upon DNA-damage, this protein co-localized in nuclear foci with the breast cancer susceptibility gene BRCA1 and the key protein for homologous recombination RAD51 (Garcia-Higuera, Taniguchi et al. 2001; Taniguchi, Garcia-Higuera et al. 2002). Next, the discovery of *FANCD1* gene in 2002 by direct sequencing of *BRCA2*, definitely stimulated the interest for FA. Bi-allelic mutations in *BRCA2* were found in fact to cause FA while mono-allelic mutations in *BRCA2* cause susceptibility to breast and ovarian cancers (Howlett, Taniguchi et al. 2002). BRCA2 was known to be functionally involved in homologous recombination repair (Moynahan, Pierce et al. 2001) and this finding highlighted the connections between FA, cancer and DNA-repair.

The remaining FA genes have been cloned in recent years, mainly through proteins association studies, by using the biochemical technique of immunoprecipitation coupled with mass spectrometry (Fig. 4). In particular, in 2003, by immunoprecipitating FANCA in HeLa cells with a specific anti-FANCA antibody, the isolated immunocomplex was analyzed by mass spectrometry (Meetei, de Winter et al. 2003; Meetei, Sechi et al. 2003). Nine components were unambiguously identified into the protein complex, afterwards called 'FA-core complex': the five expected proteins FANCA, FANCC, FANCG, FANCE, FANCF and four novel proteins named FAAP43, FAAP95, FAAP100 and FAAP250 (where 'FAAP' stands for FANCA-associated polypeptides and the numbers refer to their molecular weights). A tenth component, FAAP24, was found later in the immunoprecipitate of a FAAP250 antibody (Ciccia, Ling et al. 2007). Three FAAPs (FAAP 43, FAAP95 and FAAP250) were shown to be defective in FA patients. FAAP95 was the protein mutated in complementation group B patients (Meetei, Levitus et al. 2004), whereas FAAP43 was defective in the new isolated complementation group L (Meetei, de Winter et al. 2003) and FAAP250 was defective in patients of new complementation group M (Meetei, Medhurst et al. 2005). The proteins were named FANCB, FANCL and FANCM respectively and the discovery of three new FA genes validated the protein complex purification as a new approach for identification of disease genes. FANCN (or partner and localizer of BRCA2 (PALB2))was later identified using the same strategy (Xia, Sheng et al. 2006; Reid, Schindler et al. 2007; Xia, Dorsman et al. 2007). The two components FAAP 24 and FAAP 100, for which no corresponding mutations have been found in FA patients so far (Ciccia, Ling et al. 2007; Ling 2007), are likely to have pivotal roles nonetheless. In fact, normal cells that are inactivated by small interfering RNA (siRNA) depletion or gene knockout for these two genes show phenotypes that are characteristic of FA patients derived cells. FA patients carrying mutations in these two genes may well exist, but may be so rare that they are

currently absent from the available FA-patient repositories. Finally, the last two FA genes identified were *FANCJ*, encoding for the DNA helicase BRIP1 (BRCA1-interacting protein), also called BACH1 (BRCA1-associated C-terminal helicase 1), and *FANCI*, a paralogue of *FANCD2*. *FANCJ* was independently cloned by two groups in 2005: the first by using positional cloning (Levitus, Waisfisz et al. 2005) and the second by using genetic mapping, mutations identification and western blot data (Levran, Attwooll et al. 2005). Genome-wide search, databases analysis, biochemical studies and sequence analysis allowed instead to three groups the independent identification of *FANCI* (Dorsman, Levitus et al. 2007; Sims AE 2007; Smogorzewska, Matsuoka et al. 2007) as the thirteenth gene involved in FA and as the second FA protein to be monoubiquitinated. Relevant informations concerning all the 13 gene so far identified are reported in table 2.

Figure 3. Outline of the complementation cloning strategy.MMC-sensitive lymphoblasts froman FA patient are transfected with an episomal cDNA expression library containing a hygromycin selection marker. Cells are cultured in the presence of hygromycin to obtain cells that carry a cDNA construct. Selected cells are then cultured in medium with a MMC dose that kills the uncorrected FA cells and selects for the cells in which the FA defect is complemented by the cDNA construct. Subsequently, the episomal vector is extracted from the cells and the cDNA insert is sequenced. Finally, proof of identity for the insert representing the disease gene should come from the finding of pathogenic mutations in the patient cell line.

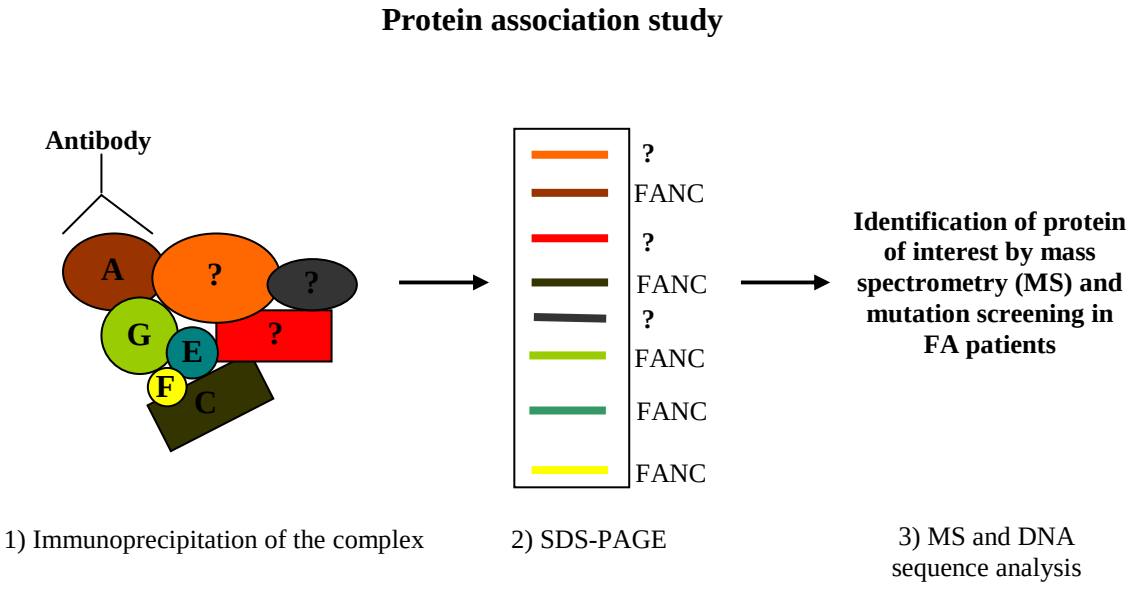

Figure 4. Identification of FA genes by protein association studies. An antibody against FANCA is added to a HeLa cell nuclear extract to precipitate FANCA and the proteins that bind to FANCA. The precipitated proteins are separated on an SDS-PAGE gel and the individual proteins are excised fromthe gel and analyzed bymass spectrometry. Subsequently, the genes encoding the novel proteins in the precipitate are analyzed for pathogenic mutations in FA patients with no mutations in any of the known FAgenes.

In the late 1990s, after cloning the first FA genes, yeast two-hybrid assays and co-immunoprecipitation studies revealed direct protein interactions between FANCA and FANCG, and between FANCC and FANCE (Kupfer, Naf et al. 1997) (Garcia-Higuera, Kuang et al. 1999) (Kruyt, Abou-Zahr et al. 1999) (Waisfisz, de Winter et al. 1999) (Medhurst, Huber et al. 2001). In particular, FANCA, FANCC, FANCF and FANCG all seemed to co-purify when recovered from nuclear extracts (Garcia-Higuera, Kuang et al. 1999). These interactions were thought to have a functional basis, since the whole nuclear complex seemed to be disrupted in cell lines from complementation groups A, C, E, F and G. Expression of the complementing (wild-type) FA protein in these cell lines resulted in a reconstitution of the nuclear complex and correction of the cellular phenotype (Garcia-Higuera, Kuang et al. 1999; Kruyt, Abou-Zahr et al. 1999; de Winter, van der Weel et al. 2000). Cell lines from complementation group D seemed instead to have an intact nuclear complex (de Winter JP 2000; Garcia-Higuera, Taniguchi et al. 2001; Timmers C 2001), which suggested that FANCD2 and the putative FANCD1 proteins acted downstream of the other known FA proteins. FANCD2 protein was detected in two isoforms: FANCD2 short (-S) and FANCD2 long (-L), where the long version consisted of the addition of a single molecule of Ubiquitin on lysine 561, a conserved amino-acid residue in FANCD2 homologues from lower organisms. Mutation of this residue resulted in loss of ubiqutination and failure to complement sensitivity to cross-linkers in FA-D2 cell line (Garcia-Higuera, Taniguchi et al. 2001). This modification step was observed in HeLa cells upon exposure to various DNA-damaging agents and during genome replication. However, the link to the other FA proteins emerged from the crucial observation that FA cell lines from complementation groups A, C, E, F and G expressed only unmodified FANCD2, while the monoubiquitinated form of FANCD2 was restored upon complementation of these cell lines with their respective cDNAs. These results not only established a link between FA proteins and DNA-damage, but also functionally connected the FA proteins function in a linear DNA-damage response pathway in which the first five FA proteins acted upstream of FANCD2. The discovery of FANCD2 monoubiquitination raised also the question of the identity of the E3 ubiquitin-ligase involved in the process. The obvious candidates were the five FA proteins required for FANCD2 monoubiquitination, but none of them contained E3

ubiquitin-ligase motifs. At cellular level, monoubiquitinated FANCD2 was observed to be targeted to discrete nuclear foci where it co-localized and co-purified with BRCA1 (Garcia-Higuera, Taniguchi et al. 2001), a protein implicated in several DNA-damage response pathways (Scully, Chen et al. 1997). Nuclear foci appeared in cells after DNA-damage and in those cells undergoing DNA replication. Noteworthy, BRCA1 had an amino-terminal ring-finger motif, known to function as an E3 ubiquitin-ligase (Lorick, Jensen et al. 1999; Joazeiro and Weissman 2000), therefore this protein was initially speculated to have a potential role in FANCD2 monoubiquitination, until the putative E3 of the pathway was identified.

As we will see in the next section, biochemical works carried on in recent years demonstrated that there actually is a 'FA-core complex' and that it purifies as part of a larger multiprotein complex with BLM (Meetei, Sechi et al. 2003; Ling 2007; Ciccia A 2007), the helicase mutated in Bloom syndrome (Yin, Sobeck et al. 2005), another genetic disease characterized by genomic instability and cancer predisposition.

FA genes	Prevalence in FA patients (estimated percentage)	Chromosomal location	Protein size (kDa)	Requirement for FANCD2/FANCI ubiquitinations	Conservation
FANCA	66%	16q24.3	163	+	Vertebrate
FANCB	~2%	Xp22.31	95	+	Vertebrate
FANCC	10%	9q22.3	63	+	Vertebrate,
FANCD1	~2%	13q12-13	380	-	Vertebrate, worm
FANCD2	~2%	3q25.3	155, 162	+	Vertebrate, worm, insect, slime mould
FANCE	~2%	6p21-22	60	+	Vertebrate
FANCF	~2%	11p15	42	+	Vertebrate
FANCG	9%	9p13	68	+	Vertebrate
FANCI	~2%	15q25-26	140, 147	+	Vertebrate, worm, insect, slime mould
FANCJ	~2%	17q22-24	140	-	Vertebrate, invertebrate, yeast
FANCL	<0.2%	2p16.1	43	+	Vertebrate, insect, slime mould
FANCM	<0.2%	14q21.3	250	+	Vertebrate, invertebrate, yeast, archea
FANCN	~2%	16p12.1	140	-	Vertebrate

Table 2 Characteristics of the Fanconi anaemia genes (Adapted by permission from Macmillan Publishers Ltd: Nat Rev Genet(Wang 2007)).

1.3.2: The FA-BRCA pathway.

As reported, patients from different complementation groups display a diverse spectrum of clinical phenotypes (Levitus 2006), but cells from all subtypes of FA are hypersensitive to DNA interstrand cross-linking agents (DEB and MMC). In addition, biochemical evidences show that FA proteins directly interact. FA proteins are thus considered to cooperate in a common DNA-damage response or DNA-repair pathway called 'FA pathway'. Since the connection with BRCA proteins was elucidated, the pathway is also referred to as the 'FA-BRCA' pathway (D'Andrea AD 2003) or the 'FA-BRCA network' (Venkitaraman 2004). The progressive characterization of the thirteen proteins so far discovered has showed that they act at least at three stages in this pathway. If we consider the monoubiquitinations of FANCD2 and FANCI as an assay, the FA proteins can be functionally divided into three groups (Fig.5). Group I consisting of eight FA proteins (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL and FANCM) plus FAAP 24 and FAAP100, all of which are subunit of the 'FA-core complex'. Cells that are deficient in any one of these proteins are defective in monoubiquitination of FANCD2 and FANCI. As a consequence, both FANCD2 and FANCI act downstream of the FA-core complex and form the group II proteins. It consists exactly of the two paralogues FANCD2 and FANCI which interact to form the FA 'ID' complex (Smogorzewska, Matsuoka et al. 2007). These two proteins are interdependent for their monoubuitination: FANCD2 deficient cells are defective in FANCI monoubiquitination whereas FANCI deficient cells are defective in FANCD2 monoubiquitination (Sims AE 2007; Smogorzewska, Matsuoka et al. 2007). The third group, group III, consists of FANCD1 (BRCA2), FANCN (also known as 'partner and localizer of BRCA2', PALB2) and FANCJ (also known as BRIP1 (BRCA1-interacting protein), or as BACH1 (BRCA1-associated C-terminal helicase 1)). Cells that are defective in any of the members of this last group have normal levels of FANCD2 monoubiquitination, indicating that the corresponding proteins function either downstream of the ID complex or in a parallel pathway.

Group I: the FA core complex

The protein complex constituted by FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, FANCM, FAAP24 and FAAP100 is also known as ‘FA-core complex’. Several independent works revealed that all ten components of the FA-core complex are indispensable for FANCD2 monoubiquitination in response to DNA-damage or during DNA replication (Garcia-Higuera, Taniguchi et al. 2001; Meetei, de Winter et al. 2003; Meetei, Levitus et al. 2004; Ling 2007; Ciccia A 2007). Furthermore, the subsequently discovered FANCI protein (Dorsman, Levitus et al. 2007; Sims AE 2007; Smogorzewska, Matsuoka et al. 2007) was also shown to be monoubiquitinated and its monoubiquitination also depends on the other FA proteins of the core complex (Sims AE 2007; Smogorzewska, Matsuoka et al. 2007).

The FA-core complex is necessary for monoubiquitination of FANCD2 and it is also required for monoubiquitination of FANCI, because monoubiquitination of FANCI depends on the monoubiquitination of FANCD2 (Sims AE 2007; Smogorzewska, Matsuoka et al. 2007). Monoubiquitination consists in a covalent attachment of a single molecule of ubiquitin on a lysine residue of a target protein. This is a three steps reaction catalyzed by a ubiquitin-activating enzyme (E1), a ubiquitin-conjugating enzyme (E2) and a ubiquitin-ligase (E3). Monoubiquitnation can change a subcellular localisation of a protein and/or its activity with biologically relevant consequences. The contributions of some subunits to monoubiquitination is probably indirect, through their participation in the correct assembly of the complex (Meetei, Levitus et al. 2004; Ling 2007), but one subunit, FANCL, is likely to be the catalytic subunit of the complex that monoubiquitinates FANCD2 and FANCI. FANCL contains a PHD finger or ring-finger-type ubiquitin-ligase motif and three WD40 repeats (Meetei, de Winter et al. 2003). Recombinant proteins containing these motifs have E3 ubiquitin-ligase activity. Moreover, point mutations that inactivate the E3 ligase activity fail to rescue the defective FANCD2 monoubiquitination in cells that are deficient for FANCL (Meetei, de Winter et al. 2003; Gurtan, Stuckert et al. 2006). In addition, another subunit that might have a more direct role in FANCD2 monoubiquitination is probably FANCE, given that it can form a stable complex with FANCD2 *in vitro* (Pace P 2002). It is therefore plausible that the core complex works in a concerted manner in the ubiquitination reaction, with FANCL interacting with ubiquitin-conjugating enzyme E2 and FANCE recruiting FANCD2.

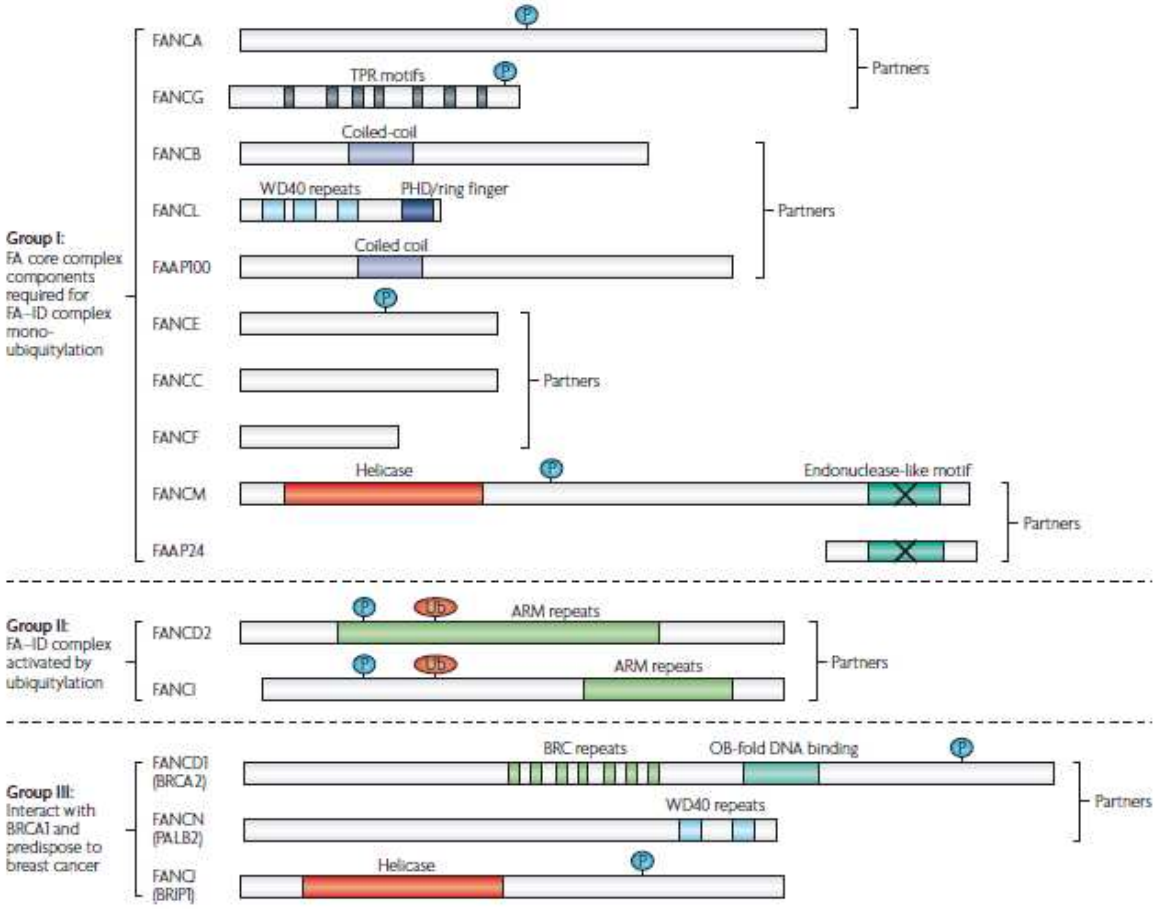

Figure 5: **Classification and domain structures of Fanconi anaemia (FA) proteins.** The FA proteins are classified into three groups on the basis of their roles in the monoubiquitination of FA proteins FANCD2 and FANCI. The brackets indicate that certain FA proteins can form complexes or subcomplexes (such as the FA core complex and the ID complex) in which they work together as partners; inactivation of one such FA protein often affects the stability and nuclear localization of its partner proteins. ARM, armadillo repeat; BRC, the internal repeat domains of BRCA2; BRCA, breast cancer susceptibility; FAAP, FA-associated protein; P, phosphorylation; TPR, tetratricopeptide repeat motifs; Ub, ubiquitin; WD40, a repeat motif found in β-transducin and other proteins. (Reprinted by permission from Macmillan Publishers Ltd: Nat Rev Genet (Wang 2007)copyright (2007).

In the core complex, there are also two components with DNA-interacting domains and activities: FANCM and FAAP24. The protein FANCM contains a DEAH-helicase domain and an ERCC4-like endonuclease domain (Meetei, Medhurst et al. 2005) whereas FAAP24 possesses a similar ERCC4-like endonuclease domain. The protein FAAP24 was shown to interact directly with FANCM and could target FANCM to single-stranded DNA (ssDNA) intermediates generated during repair (Ciccia A 2007). The heterodimer FANCM-FAAP24 is similar to a family of structure specific endonucleases and they are considered as paralogues, since they are the only human proteins with a tandem helicase-endonuclease

domain structure. Their common ancestor might have resembled the archeal DNA-repair protein Hef, which contains the same two domains and show both helicase and endonuclease activities (Komori, Fujikane et al. 2002; Komori, Hidaka et al. 2004). Hef functions as a homodimer and can resolve stalled replication forks, however, so far, no endonuclease activity has been detected for either recombinant FANCM-FAAP24 proteins or for the purified FA-core complex. Careful examination of the FANCM sequence revealed that its endonuclease domain is degenerate at several residues that are supposed to be essential for the endonuclease activity (Meetei, Medhurst et al. 2005). Additionally, the endonuclease domain of FANCM is absent in orthologues from many lower eukaryotes such as *Drosophila melanogaster*, *Caenorhabditis elegans* and yeast (Meetei, Medhurst et al. 2005; Mosedale G 2005). Notably, FAAP24 has no orthologues in these species. These results implied that the endonuclease function of FANCM might have been lost during the evolution. Furthermore, differently from its archeal and yeast orthologues (Hef and PH1) which do have helicase activity towards several DNA substrates (Komori, Fujikane et al. 2002; Komori, Hidaka et al. 2004; Prakash R 2005), FANCM does not seem to have helicase activity *in vitro* either (Meetei, Medhurst et al. 2005). Therefore, the mechanism of FANCM is probably different from its orthologues. Many proteins with helicase domains lack helicase activity but show a DNA translocase activity: examples of such proteins include ATP-dependent chromatin-remodelling enzymes. FANCM does indeed have an ATP-dependent DNA translocase activity, possibly to work as an engine to translocate the core complex along DNA, an important step in DNA-repair. At this stage, FANCM might possess other ATP-dependent DNA-processing activities that remain to be revealed. Interestingly, certain structures that are generated during replication, such as ssDNA and Y-shaped DNA, do not activate FANCD2 monoubiquitination, although the affinity of these same DNAs for FANCM-FAAP24 is higher than that of duplex DNA (Ciccia A 2007).

The activation of the FA-BRCA network depends on ATR (Andreassen, D'Andrea et al. 2004; Pichierri and Rosselli 2004), a checkpoint kinase that is activated by replication stress. Several components of the FA-core complex are hyperphosphorylated in response to DNA damage, including FANCM, FANCA, FANCE and FANCG (Yamashita, Kupfer et al. 1998; Qiao, Mi et al. 2004; Meetei, Medhurst et al. 2005; Matsuoka, Ballif et al. 2007; Wang, Kennedy et al. 2007). The phosphorilation of these proteins involves not only ATR, but also its downstream kinase checkpoint kinase 1 (CHK1) and possibly other kinases. These FA proteins might serve as signal transducers through which ATR regulates the

activity of the core complex, although this possibility remains to be demonstrated yet experimentally.

Group II: the ID complex

This group consists of the two paralogues FANCD2 and FANCI and several lines of evidence suggest that they are coordinately regulated and work together in an 'ID' complex. They co-immunoprecipitate with each other from cell lysates, both proteins are monoubiquitinated in response to DNA damage or replication cues, the monoubiquitination for both proteins depends on the FA core complex and their de-ubiquitination is catalysed by the same enzyme, ubiquitin-specific protease 1 (USP1) (Nijman, Huang et al. 2005; Smogorzewska, Matsuoka et al. 2007). Additionally, both can be phosphorylated by ATR or ataxia telangiectasia mutated (ATM) (Taniguchi, Garcia-Higuera et al. 2002; Andreassen, D'Andrea et al. 2004; Smogorzewska, Matsuoka et al. 2007) and when cells are exposed to DNA-damaging reagents, both proteins are redistributed to the chromatin and nuclei foci, which are believed to be the sites of DNA-repair (Garcia-Higuera, Taniguchi et al. 2001; Sims AE 2007; Smogorzewska, Matsuoka et al. 2007). Finally, each depends on the presence of its partner for stability and ubiquitination.

A key question for this complex is therefore how it functions. Possible clues for this mechanism is that monoubiquitinated FANCD2 preferentially associates with chromatin while non ubiquitinated FANCD2 is enriched in soluble cell extracts (Meetei, Yan et al. 2004; Wang, Andreassen et al. 2004). At this purpose, an elegant study using chicken DT40 cells showed that monoubiquitination is crucial in targeting FANCD2 to chromatin (Matsushita, Kitao et al. 2005). In this work, a FANCD2 mutant with its ubiquitination site replaced by a non ubiquitinable residue (K563R; K563 in chicken is equivalent to K561 in human FANCD2) was fused to a single ubiquitin at its C terminus. When this FANCD2–ubiquitin fusion protein was introduced to FANCD2-deficient cells, it associated with chromatin and corrected the cellular hypersensitivity to the DNA-crosslinking drug cisplatin. However, when the same FANCD2–K563R construct was fused to a ubiquitin point mutant (I44A) that was defective in interacting with ubiquitin-binding proteins, the fusion protein failed to localize to chromatin and also had a reduced ability to complement the drug sensitivity of the same cells. These results suggest that ubiquitin serves as a chromatin-localization signal for FANCD2, and also implies the existence of a chromatin associated ubiquitin-binding protein that recognizes this signal and recruits monoubiquitinated FANCD2. When the FANCD2–K563R mutant was fused to histone H2B, the fusion protein

was constitutively associated with chromatin and could rescue the cisplatin sensitivity. Therefore, the chromatin association of FANCD2 can bypass the requirement of monoubiquitination, indicating that the sole function of monoubiquitination is to target FANCD2 to the chromatin. At the cellular level, monoubiquitinated FANCD2 and FANCI are concentrated at nuclear foci, which are believed to be the sites of DNA-repair. Interestingly, the formation of these foci depends not only on the FA-core complex, but also on BRCA1(Garcia-Higuera, Taniguchi et al. 2001). It was shown that FANCD2 is relocated to sites of stalled replication forks, and that this relocation depends on the FA-core complex, ATR and BRCA1, as well as a new player, γH2AX (Bogliolo, Lyakhovich et al. 2007). γH2AX is a phosphorylated form of a histone H2A variant, H2AX, which specifically associates with damaged DNA regions and is required for the subsequent recruitment of many DNA-repair molecules, including BRCA1. Cells that are deficient for γH2AX are hypersensitive to ICL drugs, but their FANCD2 monoubiquitination is normal (Bogliolo, Lyakhovich et al. 2007). γH2AX coimmunoprecipitates with monoubiquitinated FANCD2, suggesting that γH2AX has a role in recruitment or retention of the ID complex near damaged DNA (Bogliolo, Lyakhovich et al. 2007).

At this point, one of the key questions is what does the ID complex do after it is recruited to DNA-damage sites. One study suggested that FANCD2 recruits BRCA2 to the damaged sites to promote homologous recombination repair (Wang, Andreassen et al. 2004). However, homologous recombination is only modestly affected in cells that are deficient for the FA-core complex or FANCD2 (Yamamoto, Ishiai et al. 2003; Niedzwiedz, Mosedale et al. 2004; Nakanishi, Yang et al. 2005), in contrast to BRCA2-deficient cells, in which homologous recombination is severely affected. These findings argue against a significant involvement of FANCD2 in BRCA2 recruitment during homologous recombination. Other evidence suggests that FA proteins also promote translesion synthesis (TLS) (Niedzwiedz, Mosedale et al. 2004; Nojima, Hochegger et al. 2005). For example, DT40 cells in which translesion polymerases REV1 (or REV1L) and REV3 (or POLZ) are inactive, are highly sensitive to DNA-crosslinking drugs; genetic analysis shows that these polymerases work in the same pathway as FA proteins to repair ICLs. The ID complex might have a similar role to proliferating cell nuclear antigen (PCNA) during polymerase switching in response to DNA-damage (McHugh and Sarkar 2006; Shen, Jun et al. 2006). Monoubiquitinated PCNA interacts with ubiquitin-binding domains of translesion polymerases and recruits them, allowing replicative polymerases to be replaced and the damaged sites to be bypassed

(Kannouche, Wing et al. 2004). By analogy, the monoubiquitinated ID complex might interact with and recruit translesion polymerases through a similar mechanism.

Finally, results from several studies have pointed out that the FA ID complex is functionally more important than the FA-core complex. The clinical phenotypes of FA patients with FANCD2 mutations are in effect generally more severe than those of patients who are defective in FA-core complex proteins, including more frequent malformations and earlier manifestations of haematological malignancies (Kalb, Neveling et al. 2007). Also, all FANCD2 mutations identified so far are hypomorphic, and the patient-derived cells all express residual levels of both monoubiquitinated and non-ubiquitinated FANCD2 (Kalb, Neveling et al. 2007), indicating that FANCD2 is probably essential during human development and that its complete inactivation could be lethal. By contrast, patients who are defective in FA-core complex components frequently carry nullizygous mutations, indicating that they are non-essential. Additionally, FANCD2 knockout mice display certain phenotypes such as small eyes, perinatal lethality and epithelial cancers, which are not observed in mice in which genes encoding FA-core complex components are disrupted (Houghtaling, Timmers et al. 2003). Furthermore, the majority of FA-core complex genes are completely absent in many eukaryotic species that do contain orthologues of FANCD2 (and FANCI). These differences suggest that ubiquitination of FANCD2 by the core complex is required for some but not all of the functions of FANCD2, and that the non-ubiquitinated FANCD2 probably has indispensable roles in some developmental processes and tumour suppression. In addition to its participation in the FA pathway, FANCD2 has been suggested to participate in a separate signalling pathway that is activated by ATM in response to ionizing radiation (which causes double stranded breakages DSBs) (Taniguchi, Garcia-Higuera et al. 2002). In this pathway, ATM phosphorylates FANCD2 on serine 222, leading to activation of the S-phase checkpoint. Thus, the FA ID complex participates in at least two discrete pathways: one that responds to ATR and the FA-core complex, and the other one that is regulated by ATM.

Group III: cancer and DNA repair

This group consists of three proteins: FANCD1 (BRCA2), FANCJ and FANCN. In contrast to proteins from the other two groups, cells lacking group III proteins have normal levels of monoubiquitinated FANCD2 in response to DNA-damage, suggesting that these proteins function either downstream of the ID complex or in a separate pathway.

The first important feature of group III genes is their connection to breast cancer. Homozygous mutations at these loci cause FA while heterozygous mutations predispose female carriers to breast cancer. As previously mentioned, the first evidence for this came with the identification of the *FANCD1* gene as *BRCA2* (Howlett, Taniguchi et al. 2002). *BRCA2* and *BRCA1* are two major tumour-suppressor genes, mutations in which are linked to familial breast and ovarian cancers in women. Although *BRCA1* is not mutated in FA patients, the BRCA1-interacting protein BRIP1, also known as BRCA1-associated C-terminal helicase (BACH1), is identical to another member of group III, FANCJ (Levitus, Waisfisz et al. 2005; Litman, Peng et al. 2005; Godthelp, Wiegant et al. 2006). In addition, BRCA1 associates with BRCA2 to form a complex that is distinct from the BRCA1–BRIP1 complex, and BRCA1 is required for both BRCA2 and BRIP1 to relocate to the sites of DNA-damage (Greenberg, Sobhian et al. 2006). Together, these data underscore an important connection between FA and breast cancer, and suggest that FA could be an attractive model to identify more breast cancer susceptibility genes and examine how they function. In this regard, BRCA1 was shown to form a new complex with receptor-associated protein 80 (RAP80, also known as UIMC1) (Kim, Chen et al. 2007; Sobhian, Shao et al. 2007; Wang, Matsuoka et al. 2007) and abraxas. Both proteins are phosphorylation substrates of ATR or ATM, and are required for resistance to DNA-damage. Notably, RAP80 contains ubiquitin-interacting motifs, and is required for optimal accumulation of BRCA1 to the sites of DNA-damage.

Group III proteins form stable complexes also with other DNA-repair proteins. BRCA2 and FANCN have been purified in a complex with the recombinase RAD51. FANCJ has been found in a separate complex with topoisomerase II binding protein 1 (TOPBP1) and the mismatch repair proteins mutL homologue 1 (MLH1) (Greenberg, Sobhian et al. 2006) and post-meiotic segregation increased 2 (PMS2) (Peng, Litman et al. 2007) (table 3). Both complexes also contain BRCA1 and its partner, breast cancer-associated ring domain 1 (BARD1). Current evidence suggests that the FANCD1 (BRCA2)-FANCN complex participates in homologous recombination repair of DNA-damage. FANCD1 (BRCA2) interacts directly with RAD51, and this interaction might nucleate the formation of RAD51 oligomers on ssDNA, a key step in homologous recombination repair (Sharan, Morimatsu et al. 1997; Pellegrini, Yu et al. 2002; Yang, Jeffrey et al. 2002). In BRCA2-deficient cells, the efficiency of homologous recombination repair is strongly reduced (Moynahan, Pierce et al. 2001), and RAD51 fails to localize to nuclear foci in response to DNA-damage (Yuan, Lee et al. 1999). FANCN (or PALB2) is an essential component of the BRCA2–RAD51

complex, and inactivation of FANCN results in instability and mislocalization of BRCA2, as well as defective homologous recombination repair (Gurtan, Stuckert et al. 2006).

The DNA-repair pathway that is mediated by the FANCJ-associated complex remains unclear. Knockdown experiments in mammalian cells suggest that FANCJ is a downstream effector of BRCA1 and mediates homologous recombination repair (Litman, Peng et al. 2005), but gene-knockout studies in chicken DT40 cells suggest that FANCJ acts independently of BRCA1 and is dispensable for homologous recombination (Bridge, Vandenberg et al. 2005). A complementation study using FANCJ patient derived cells showed that the interaction between FANCJ and BRCA1 is dispensable for the function of FANCJ in the FA pathway, whereas its interaction with MLH1 and PMS2 is essential (Peng, Litman et al. 2007). However, it is not obvious how MLH1 and PMS2 mediate the action of FANCJ, as cells that are deficient in these proteins do not display FA-like phenotypes such as MMC hypersensitivity. Notably, FANCJ contains a DEAH–box DNA helicase domain and a 5′-to-3′ helicase activity (Cantor, Drapkin et al. 2004). Considering that FA proteins function in the same pathway as translesion polymerases (Niedzwiedz, Mosedale et al. 2004), it is possible that FANCJ promotes translesion bypass using a similar mechanism to what has been proposed for the human translesion polymerase POLQ (Seki, Marini et al. 2003) and its *D. melanogaster* orthologue MUS308 (Shima, Munroe et al. 2004). Both of these proteins contain tandem helicase and polymerase domains. MUS308 mutant cells are highly sensitive to ICL drugs, and mice with mutations in POLQ also seem to have a potential defect in homologous recombination or ICL repair (Shima, Munroe et al. 2004). It has been proposed that the helicase domain removes the lagging strand at stalled replication forks and creates a loading zone for the polymerases (Guy and Bolt 2005). FANCJ might promote loading of a translesion synthesis polymerase to stalled replication forks by displacing unwanted DNA or proteins in a similar way.

Proteins (disease)	**Properties or domain**	**Function**
	Upstream of FA core complex	
ATR (Seckel syndrome)	PI-3 kinase	Phosphorilates FA core and ID complex
ATRIP		A partner of ATR
RPA	OB-fold	Bind ssDNA; activates ATR
CHK1	Kinase	A downstream kinase of ATR
HCLK2		Associates with ATR; stabilizes CHK1
NBS1 (Nijmegen	FHA; BRCT	Associates with MRE11 and

breakage syndrome)		RAD50; required for ATR and ATM activation
ATM (ataxia telangiectasia)	PI-3 kinase	Phosphorilates FANCD2 and possibly FANCI
	FA core complex partners	
BLM (Bloom syndrome)	Helicase	Stabilizes and restarts replication forks
Topoisomerase IIIα	Type I topoisomerase	An essential partner of BLM
BLAP75/RMI1	OB-fold	Binds DNA; an essential partner for topoisomerase IIIα
RPA	OB-fold	Facilitates BLM helicase activity
UBE2T	E2	Ubiquitin-conjugating enzyme
USP1	De-ubiquitination enzyme	Removes ubiquitin from the ID complex; regulated by DNA damage and replication
	Downstream FA protein partners	
γH2AX	Histone H2A variant	Phosphorilated following DNA damage; marks regions of DNA damage; recruits or retains DNA-repair molecules such as BRCA1 and FANCD2
BRCA1 (breast cancer)	E3 ring finger; BRCT	Ubiquitin ligase; binds phosphorilated ATM or ATR sites; associates with BRCA2 and FANCJ; recruits FANCD2, BRCA2 and FANCJ to damage sites
BARD1	Ring finger	Binds BRCA1; essential for the ubiquitin ligase activity of BRCA1
RAD51	ATPase	Required for homologous recombination; associates with BRCA2 and PALB2
MLH1 (colorectal cancer)	ATPase	Associates with FANCJ, BRCA1 and BLM; participates in mismatch repair
PMS2 (colorectal cancer)	ATPase	Associates with FANCJ and BRCA1; a partner of MLH1, involved in mismatch repair
TOPBP1	BRCT	Associates with FANCJ and BRCA1; activates ATR

Table 3: Proteins that crosstalk with FA-BRCA pathway

ATM, ataxia telangiectasia mutated; ATR, ataxia telangiectasia and Rad3-related protein; ATRIP, ATR interacting protein; BARD1, BRCA1-associated RING domain 1; BLAP75, BLM-associated protein, 75 kDa; BLM, Bloom syndrome protein; BRCA, breast cancer susceptibility; CHK1, checkpoint kinase 1; FA, Fanconi anaemia; FANC, Fanconi anaemia complementation group; γH2AX, phosphorylated form of histone H2A variant H2AX; HCLK2, human homologue of the worm biological clock 2 protein; MLH1, mutL homologue 1, colon cancer, nonpolyposis type 2; MRE11, meiotic recombination 11 homologue A; NBS1, Nijmegen breakage syndrome protein; PMS2, post-meiotic segregation increased 2; RMI1, RecQ-mediated genome instability 1 homologue; RPA, replication protein A; TOPBP1, topoisomerase (DNA) II binding protein 1;

UBE2T; ubiquitin-conjugating enzyme E2T; USP1, ubiquitin-specific protease 1. (Adapted by permission from Macmillan Publishers Ltd: Nat Rev Genet(Wang 2007).

1.3.3: Crosstalk of the FA-BRCA pathway with other DNA-repair proteins.

The FA proteins interact with a range of proteins that are known to be involved in the DNA-damage response (table 3). Mutations in some of these proteins also cause diseases that are similar to FA, including genomic instability and cancer. One group includes ATR (Andreassen, D'Andrea et al. 2004; Pichierri and Rosselli 2004), its downstream kinase CHK1 (Wang, Kennedy et al. 2007) and molecules that participate in ATR activation such as RPA (Andreassen, D'Andrea et al. 2004), ATRIP (Sobeck, Stone et al. 2006), NBS1 (a component of the MRE11–RAD50–NBS1 complex) (Nakanishi, Taniguchi et al. 2002; Pichierri and Rosselli 2004; Stiff, Reis et al. 2005) and HCLK2 (Collis, Barber et al. 2007). These molecules activate the FA–BRCA network by facilitating phosphorylation of the FA core complex, the ID complex, BRCA1, BRCA2, PALB2 and γH2AX11, (Yamashita, Kupfer et al. 1998; Qiao, Mi et al. 2004; Ho, Margossian et al. 2006; Matsuoka, Ballif et al. 2007; Wang, Kennedy et al. 2007). ATM can also be included in this group, however, it seems to be involved in activation of the intra-S phase checkpoint, but is not involved in FANCD2 monoubiquitylation or repair of ICLs (Taniguchi, Garcia-Higuera et al. 2002). Because ATR is usually activated by collapsed replication forks during DNA synthesis whereas ATM responds mainly to DSBs, the fact that the FA–BRCA network is activated by ATR but not ATM argues that the main function of this network is to recover stalled replication forks.

A second group consists of proteins that work with the core complex, including BLM helicase and its DNA-processing partners, Topoisomerase IIIα, the BLM-associated polypeptide BLAP75 (or RMI1) and RPA. One important function of the BLM complex is to stabilize the stalled replication forks and allow a subsequent homologous recombination-dependent process to restart replication (Cobb, Bjergbaek et al. 2003; Bjergbaek, Cobb et al. 2005). The direct interaction between the BLM and the FA-core complex suggests that they work together to help restart arrested replication forks (Thompson, Hinz et al. 2005). Other proteins of this group include UBE2T, which has been reported as a possible E2 for FANCD2 monoubiquitination (Machida, Machida et al. 2006), and USP1, the de-ubiquitination enzyme for FANCD2 and FANCI (Nijman, Huang et al. 2005). USP1 protein levels are regulated during cell-cycle progression and by DNA-damage signals, suggesting

that cells can modulate the amount of the monoubiquitylated ID complex, which reflects the status of the FA–BRCA network, through regulating USP1 (Huang, Nijman et al. 2006).

A third group consists of proteins that work with downstream FA proteins. Among them are γH2AX and BRCA1, which seem to be essential for the recruitment of the ID complex, FANCJ and BRCA2 to damaged DNA regions. This group also includes proteins with DNA-processing activities, such as recombinase RAD51, which associates with BRCA2 and PALB2, and TOPBP1, MLH1 and PMS2, which interact with FANCJ. The translesion polymerases REV1 and REV3 can also be included in this group as they work in the same pathway as FA proteins, although so far they have not been shown to interact physically with any FA proteins.

1.3.4: FA-BRCA network response following DNA damage: a current model.

Several models have been proposed to describe how the FA-BRCA pathway promotes repair of ICLs. One of the simpler models, described in reference by (Kennedy and D'Andrea 2005), is reported in figure 6. In this model, FA proteins work with nucleotide excision repair, translesion synthesis and homologous recombination proteins to remove an ICL at a stalled replication fork. Briefly, FANCM and FAAP24 might target the core complex to the stalled replication fork, where it can be activated by ATR through phosphorylation. The BLM complex is also recruited and prevents the stalled fork from collapse. The phosphorylated FA core complex then activates the ubiquitin-ligase activity of FANCL, which monoubiquitinates the ID complex (which can possibly also be phosphorylated by ATR before its ubiquitination). The ICL is unhooked by MUS81–EME1 and XPF–ERCC1 endonucleases, creating a DSB in the DNA (Kuraoka, Kobertz et al. 2000; Hanada, Budzowska et al. 2006). ATR also phosphorylates γH2AX, which is incorporated into the chromatin that surrounds the damaged DNA region. The monoubiquitinated ID complex is then recruited to and retained at the site of DNA-damage by γH2AX and BRCA1. This allows subsequent entry of the BRCA2–PALB2–RAD51 complex to the damaged site, thereby promoting a homologous recombination-dependent restart of the stalled fork. The ID complex might also recruit REV1 and REV3 translesion polymerases in a similar mechanism to that by which monoubiquitinated PCNA allows damaged DNA to be bypassed. FANCJ might use its helicase activity to clear the DNA template and promote the loading of translesion polymerase.

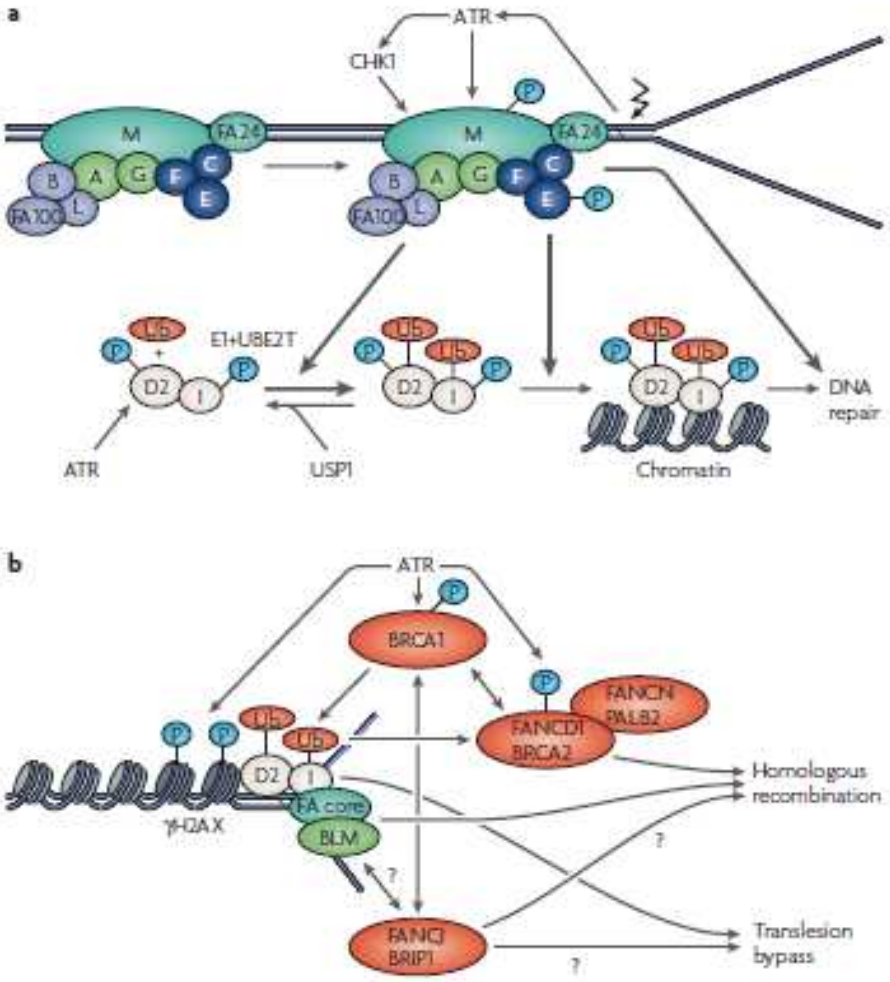

Figure 6: Signalling through the Fanconi anaemia (FA)–breast cancersusceptibility (BRCA) network in response to DNA damage. a | A model for the upstream FA–BRCA network. Briefly, a stalled replication fork activates ataxia telangiectasia and Rad3-related protein (ATR) and its downstream kinase checkpoint kinase 1 (CHK1). They activate the FA core (composed of FA proteins A, B, C, E, F, G and L)and ID complex (composed of FA proteins D2 and I) by phosphorylation. The core complex translocates and is recruited to sites of DNA damage by FA proteins FANCM(M in the figure) and FAAP24 (FA24 in the figure), where it monoubiquitylates the FA–ID complex, and might also help to recruit the ID complex to the chromatin. FANCM and FAAP24 might also directly process the DNA as part of a DNA-repair reaction. The ubiquitin-activating enzyme (E1) and a conjugating enzyme (UBE2T)work together with the FA core complex (E3) to monoubiquitylate the ID complex. A de-ubiquitylation eyzyme, USP1, removes the monoubiquitin from the ID complex to‘switch off’ the network after the damage is repaired. **b:** A model for the downstream FA–BRCA network. ATR phosphorylates the histone H2A variant H2AX (making□H2AX), along with BRCA1 and BRCA2, which might allow various repair complexes to be recruited to the damaged loci through protein–protein interactions. The BRCT domains in BRCA1 recognize phosphorylated ATR or ataxia telangiectasia mutated(ATM) sites, and might be involved in the recruitment of the repair proteins. A complex of BRCA2 and partner and localizer of BRCA2 (PALB2) can promote a homologous recombination-dependent restart of the replication fork, whereas FANCJ(also called BRCA1-interacting protein 1 (BRIP1)) probably promotes

translesion bypass. The BLM helicase complex could work together with the FA core complex to stabilize and re-start the stalled replication forks. P, phosphorylation; Ub ubiquitylation. . (Reprinted by permission from Macmillan Publishers Ltd:: Nat Rev Genet(Wang 2007)copyright (2007).

1.3.5: Other branches of the FA-BRCA pathway.

In addition to defective DNA--repair, a number of other possible causes of the cellular and clinical abnormalities found in FA have been also recognized and these are summarized in this section.

Oxygen sensitivity.

Several works have shown that FA cells exhibit increased oxygen sensitivity, and chromosome breakage is reduced in the presence of low ambient oxygen concentrations (Joenje, Arwert et al. 1981; Joenje 1989). Many DNA cross-linking agents produce reactive oxygen species (ROS) and it is possible that the sensitivity of FA cells to cross-linkers is due to an impaired ability to counteract ROS (Pagano and Youssoufian 2003). Studying apoptosis in FA-C cells, Clarke and Philpott showed that they behaved like normal cells when exposed to MMC at 5% oxygen concentration and were hypersensitive to MMC at 20% oxygen, implying that it is the ROS generated by MMC, rather than the DNA cross-link formation, that causes toxicity in FA cells (Clarke, Philpott et al. 1997). Other lines of evidence come from yeast-two hybrid studies, which have shown that the FANCC protein interacts with NADPH cytochrome p450 reductase (Kruyt, Hoshino et al. 1998) and FANCG with cytochrome p450 2E1 (CYP2E1) (Futaki, Igarashi et al. 2002), both enzymes that produce ROS.

Apoptosis and telomere maintenance.

Apoptosis is abnormally regulated in FA cells, although the exact nature of the abnormality remains unclear. In one study, the treatment of four FA lymphoblastoid cell lines with MMC led to a higher level of apoptosis (Kruyt, Dijkmans et al. 1996), but other studies have shown higher levels of spontaneous apoptosis, lower levels after c-irradiation, and no difference after MMC exposure (Rey, Scott et al. 1994; Rosselli, Ridet et al. 1995; Ridet, Guillouf et al. 1997). Increased apoptosis may be related to the inability of FA cells to repair damage or to defective FA protein interactions with other proteins in the apoptosis pathways (Cumming, Lightfoot et al. 2001; Pang, Keeble et al. 2001). Accelerated telomere shortening has been detected in FA as well (Leteurtre, Li et al. 1999; Hanson, Mathew et al.

2001; Li, Leteurtre et al. 2003), but this is also found in other forms of aplastic anaemia (Ball, Gibson et al. 1998) and in MDS (Boultwood, Fidler et al. 1997). This could be explained by haematopoietic stem cells undergoing a higher than normal number of cell divisions to generate mature end cells. However, there is evidence that, in addition to replicative shortening, there is also a higher rate of breakage at telomeric sequences in FA cells, suggesting a defect in telomere maintenance (Callen, Samper et al. 2002).

Haematopoiesis.

Defective haematopoiesis in FA has been demonstrated using *in vitro* bone marrow culture assays (Stark, Thierry et al. 1993) and FA cells show altered levels of certain growth factors, such as reduced interleukin (IL)-6 (Rosselli, Sanceau et al. 1992), granulocyte-macrophage colony-stimulating factor (GM-CSF) (Stark, Andre et al. 1993), IL-1b (de Cremoux, Gluckman et al. 1996) and increased tumour necrosis factor-a (TNF-α) (Schultz and Shahidi 1993; Rosselli, Sanceau et al. 1994). These may affect haematopoiesis by altering the bone marrow microenvironment, leading to deregulation of cellular homeostasis, differentiation and response to stress (Lensch, Rathbun et al. 1999). Compared with other tissues, haematopoietic tissue is particularly sensitive to DNA-damage caused by radiation or cytotoxic therapy. However, while other DNA-repair disorders such as Xeroderma Pigmentosum or Ataxia telangiectasia have an increased risk of certain malignancies, they do not have the striking incidence of bone marrow failure seen in FA. This suggests that the FA proteins play a specific critical role in the maintenance of haematopoietic stem cells and it is likely that a combination of different functions of the FA-BRCA pathway are implicated. The initial decline in haematopoietic stem cells leading to bone marrow failure could be the result of defective telomere maintenance and high rates of apoptosis whereupon selective pressures. As a result of an altered signalling environment against a background of genome instability, this could lead to the emergence of apoptosis-resistant mutant clones, which in turn may ultimately lead to the development of AML (Lensch, Rathbun et al. 1999). It is clear that as well as interacting in a complex, at least some of the individual FA proteins interact with proteins in other pathways, and are responsible for the cellular characteristics described above (Reuter, Medhurst et al. 2003). For example, FANCC interacts independently with a number of other proteins including hsp70 leading to suppression of the kinase activity of PKR and thereby modulating apoptotic signalling (Pang, Christianson et al. 2002). FANCC also interacts with STAT1, which is involved in the response signals to interferons (Pang, Christianson et al. 2002).

1.3.6: Disease models.

Generation of various knockout murine models for *Fanca* (Cheng, van de Vrugt et al. 2000; Wong, Alon et al. 2003), *Fancc* (Chen, Tomkins et al. 1996; Whitney, Royle et al. 1996), *Fancg* (Yang, Kuang et al. 2001; Koomen, Cheng et al. 2002), *Fancd2* (Houghtaling, Timmers et al. 2003), *Fanca-Fancc* double (Noll, Battaile et al. 2002), *Fancl*/*Pog*(Agoulnik, Lu et al. 2002) and *Fancd1/Brca2* (Navarro, Meza et al. 2006) have been reported. These FA mice display chromosome instability, defective germ cell development, reduced fertility and sensitivity to DNA cross-linking agents. However, they do not develop spontaneous bone marrow failure, except for mouse models of the FA-D1 group harboring homozygous hypomorphic mutations in *Fancd1/Brca2* (Navarro, Meza et al. 2006). Treatment with MMC causes bone marrow failure in *Fancc* knockout mice (Carreau, Gan et al. 1998). The reason for the absence of spontaneous bone marrow failure in FA mice is still unknown. *Fancd2*, *Fanca* and *Fancc* mice are reported to develop tumours including adenocarcinoma, lymphoma, sarcoma and ovarian granulosa cell tumour. In addition, many *Brca1*- and *Brca2*-deficient mouse models with cancer susceptibility have been generated (reviewed in reference (Moynahan 2002)). However, despite the lacking of spontaneous bone marrow failure, FA mice are still regarded as useful tools to test the applicability of new FA treatment modalities. They are in fact essential for gene therapy development (Gush, Fu et al. 2000; Noll, Bateman et al. 2001; Yamada, Olsen et al. 2001; Galimi, Noll et al. 2002; Rio, Segovia et al. 2002; Haneline, Li et al. 2003; Yamada, Ramezani et al. 2003), *ex vivo* manipulation of hematopoietic stem/progenitor cells (Li, Le Beau et al. 2005) and cytokine treatment (Carreau, Liu et al. 2002).

Drosophila, Caenorhabditis elegans, Xenopus and zebrafish homologs of FA genes have been recently identified. The FA gene network is conserved from zebrafish (Titus, Selvig et al. 2006) and *Xenopus* (Sobeck, Stone et al. 2006) to human, while only *FANCD2, FANCL* and *FANCM* have *Drosophila* homologs (Marek and Bale 2006), and *FANCD1/BRCA2, FANCD2, FANCJ* and *FANCM* have homologs in *C. elegans* (Collis, Barber et al. 2006). *Drosophila* and *C. elegans* models can therefore be useful for testing the functions of key proteins in the FA-BRCA pathway. As only FANCD2, FANCL and FANCM are present in *Drosophila*, these proteins may constitute the minimal FA-BRCA pathway machinery. In turn, *C. elegans* may enable a better understanding of the roles of

FANCD1 (BRCA2) and FANCJ in the context of the minimal machinery constituted by FANCD2 and FANCM. As with vertebrate mutants, *Drosophila fancd2* and *fancl* mutants, as well as *C. elegans fancd2* mutants, show hypersensitivity to DNA cross-linking agents (Collis, Barber et al. 2006; Marek and Bale 2006). *Drosophila* and *C. elegans* models may, therefore, be useful to dissect the roles and regulations of the FA-BRCA pathway in a less complex network. The zebrafish model may prove a valuable tool with which to investigate the impact of FA-BRCA pathway failure on development, as *fancd2*-deficient zebrafish embryos develop similar defects to those found in children with FA, including shortened body length, microcephaly and microphthalmia (Liu, Howlett et al. 2003). The *Xenopus* model mostly constitutes a powerful tool with which to investigate the regulation of the FA pathway *in vitro* through cell-free assay systems using replicating egg extracts (Sobeck, Stone et al. 2006). Such assays could also be exploited for screening drugs that modulate the FA-BRCA pathway.

1.3.7: Disease targets and ligands.

As described, the FA-BRCA pathway is required for cellular resistance to DNA cross-linking agents, including widely used anti-cancer drugs such as cisplatin, MMC and melphalan. As such, inhibition of this pathway will lead to sensitization of tumour cells to these drugs, therefore the FA-BRCA pathway is an attractive target for developing small molecule inhibitors that may be clinically useful as chemo-sensitizers in the treatment of cancer. The FA-BRCA pathway involves formation of a multi-subunit protein complex bearing E3 ligase activity, several enzymes (ubiquitin ligase and conjugating enzyme, de-ubiquitinating enzyme, kinase, ATPase/DNA translocase and ATPase/helicase) and many protein–protein or protein–DNA interactions. All of these components are potential targets for small molecule inhibitors of the pathway. Drug development is still in the early stages, and therefore not many informations are available. A high-throughput cell-based screening assay for small molecule inhibitors of the FA-BRCA pathway has, however, been developed in the laboratories of Alan D'Andrea (Chirnomas, Taniguchi et al. 2006). In this screen, FANCD2 nuclear foci formation was utilized as readout of the inhibition of DNA-damage induction. A partial result focusing on one inhibitor, curcumin (diferuloylmethane), has been published (Chirnomas, Taniguchi et al. 2006). Curcumin inhibits FANCD2 monoubiquitination and nuclear foci formation, although its exact target in the pathway has

not been yet identified (Chirnomas, Taniguchi et al. 2006). However, curcumin sensitizes an ovarian cancer cell line to cisplatin in an FA-BRCA pathway-dependent manner, suggesting that curcumin sensitization of cisplatin mostly occurs through FA-BRCA pathway inhibition (Chirnomas, Taniguchi et al. 2006). In order to establish curcumin as a useful cisplatin chemo-sensitizer, a detailed analysis of combinations of curcumin with cisplatin, *in vivo* studies using mouse models and identification of the target of curcumin in the pathway are still required.

More recently, a work published in 2008 shows instead how phenylbutyrate may potentially have therapeutic utility as a cisplatin sensitizer in head and neck cancer (Burkitt and Ljungman 2008). In this study, the authors report that phenylbutyrate inhibits the FA-BRCA pathway through the down regulation of BRCA1 as well as through an FA-BRCA-independent mechanism. This potentially relevant finding could help overcome one of the clinical limitations of cisplatin treatment, since these tumours, initially responsive to this treatment, later acquire resistance (Taniguchi, Tischkowitz et al. 2003).

Elucidating the mechanisms of action of candidate inhibitors identified in the above-mentioned screens seems to be crucial for development of specific and effective inhibitors of the FA-BRCA pathway and for further understanding the regulation of the pathway itself. Precise analyses of the effects of each drug on individual steps in the FA-BRCA pathway will be required to identify their specific targets. These analyses will include *in vitro* ATR kinase assay, *in vitro* UBE2T and FANCL autoubiquitination assays, assessment of FA-core complex formation and evaluation of nuclear foci formation of DNA-repair proteins including BRCA1. A better understanding of the regulatory mechanisms, as well as the identification of new partners of the FA-BRCA pathway, is also crucial for the identification and development of inhibitors. Finally, although multiple groups have been working on structural analysis of FA proteins, so far few FA protein crystal structures have been reported (BRCA2/FANCD1 (Yang, Jeffrey et al. 2002)and FANCF (Kowal, Gurtan et al. 2007)), but also such studies could provide useful information for drug design as well as for elucidation of the targets of the candidate inhibitors.

1.4 Protein-protein interactions and yeast two-hybrid systems.

1.4.1: Signalling networks and how to dissect them.

Signalling networks are key elements in all major aspects of cellular life, playing a major role in inter- and intra- cellular communications. They are involved in diverse processes such as cell-cycle progression, cellular metabolism, cell-cell communication and appropriate response to the cellular environment. The latter comprises a whole range of networks that are involved in regulation of cell development, differentiation, proliferation, apoptosis, and immunological responses. The key mechanism involves the transductions of extra-cellular signal across the cell surface to different effectors in the cytosol and the nucleus (Kolch 2000). Deregulation of these pathways is often associated with malignant diseases (Hahn and Weinberg 2002). Due to the multitude of signalling cascades, it is of utmost importance to decipher the proteins playing key parts in these networks, and the means by which signals are transmitted by these molecules. The means of analysing signalling networks have undergone dramatic changes over the last decade. There has been a general shift in paradigm from dedicating a work of a lifetime to the analysis of a single protein to the analysis of cellular and biochemical processes and networks. This step has been enabled by a dramatic improvement in technologies and experimental strategies mainly in the fields of genomics and proteomics. Fifteen years ago, biochemical and cell-biological approaches, together with genetic screens in model organisms (mainly the fruit fly *Drosophila melanogaster* and the nematode worm *Caenorhabditis elegans*) have dominated the scene to map the function of cellular signalling pathways. The elucidation of signalling networks was traditionally done by identifying the different protein components one by one. The genetic approaches usually designed as variations of epistasis or rescue experiments provided information on the hierarchical relationship of the gene under study to the newly identified genes by showing that one is dependent on (or independent of) the other for particular biological functions. While powerful in mapping interdependencies, molecular mechanisms remained open. In contrast, the biochemical approaches could reveal mechanisms, but unfortunately they were largely confined to known interactions. Pathway mapping was much of a trial and error process driven by educated guesses. Many of the best guesses were based on a synthesis of the genetic and biochemical information showing that in principle such a combination could be very powerful. However, since the participating

genes or proteins were identified and characterized one at a time, either approach could take many years of work with slow progress. The sequencing of the human genome, as well as the genome of important model organisms such as mouse (*Mus musculus*), *D. melanogaster*, *C. elegans*, and the budding yeast *Saccharomyces cerevisiae* has tremendously increased the possibilities of network mapping. DNA microarray analysis of mRNA expression for the first time permitted global genome wide analysis and quickly became a routine platform that is in widespread use today. Proteomics had a slower start due to the increased complexity and widely different physicochemical properties of proteins as compared to nucleic acids, which make analysis much more demanding. Nevertheless, the pace of progress has accelerated enormously in the last few years and proteomic approaches have gained a high popularity especially for the analysis of protein-protein interactions and signalling pathways (Zhu and Snyder 2002; Zhu, Bilgin et al. 2003). A variety of proteomic techniques has been implemented in the search for understanding the molecular mechanisms of signalling and the underlying rationale is simple. The transduction of a signal occurs by binding of ligands to receptors, recruitment of adaptor proteins, enzymatic reactions such as phosphorylation and ubiquitination, subcellular redistribution and assembly (or disassembly) of molecular machines that cooperate in functional pathways. These processes have in common that they invariably involve protein-protein interactions, and it seems logical to presume that by tracking the protein interactions, one can track the flux of the signal and hence reconstruct the pathways. As protein-protein interactions are so important, there are a multitude of genetics and biochemical methods to detect them: yeast two-hybrid system, two-dimensional gel electrophoresis, mass spectrometry, *in vitro* pull-down assays, immunoprecipitation, protein microarray, fluorescence resonance energy transfer (FRET), RNA interference (RNAi) knock-down, cDNA library screening. Each of the mentioned approaches has its own strengths and weaknesses, especially with regard to the sensitivity and specificity of the method. However, the first systematic approach used for searching for potential interaction partners for a selected protein was the yeast two-hybrid system (Fields and Song 1989), and a large amount of our current knowledge of protein-protein interactions involved in signalling events has been derived from experiments exploiting this technique.

1.4.2: Yeast two-hybrid system.

The yeast two-hybrid system was originally developed by Fields and Song in the 1989. Protein-protein interactions between two proteins had generally been studied using biochemical techniques such as cross-linking, co-immunoprecipitation and co-fractionation by chromatography. Fields and Song generated a novel genetic system to study these interactions, exploiting the yeast *Saccharomyces cerevisiae* and the properties of its transcriptional factor GAL4. The GAL4 protein is composed of two separable and functionally distinct domains: an N-terminal domain which binds to specific DNA sequences (binding domain/BD), and a C-terminal domain containing acidic regions, necessary to activate transcription (activator domain/AD). A system with two hybrid proteins, where the GAL4 DNA-binding domain (BD) was fused to a protein "X" (bait) and a GAL4 activating region (AD) to a protein "Y" (prey), was set up. The principle was that, if proteins X and Y can form a protein– protein complex so that the complex brings the two domains of GAL4 into proximity, transcription of a reporter gene regulated by GAL4 is initiated (Fig. 7). They tested the system constructing separate plasmids carrying the sequences coding for (a) the GAL4 BD and (b) the GAL4 AD, fused in frame to DNA sequences coding for the protein products of the yeast genes *SNF1* and *SNF4* respectively. SNF1 and SNF4 proteins were known to interact (personal communication to the authors)(Fields and Song 1989). They introduced these two plasmids into a yeast strain containing a reporter gene, the *LacZ* gene from *E.Coli*, fused to the *GAL1* promoter. The resulting β-galactosidase activity was detected by the formation of blue yeast colonies on medium supplemented with the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) (Fields and Song 1989). The two-hybrid system immediately appeared to have the potential to identify ideally any two interacting proteins. Further development of the system allowed investigators to use a protein of interest as bait, to identify cellular proteins that interact (Chien, Bartel et al. 1991), especially since cDNA libraries become largely available. Generally, the DNA sequence coding for the product of a gene of interest (X) can be fused to *GAL4BD* and introduced on a selectable plasmid into an appropriate yeast strain. There it will bind to the *GAL1* promoter and function as the bait. The yeast cells are transformed with a cDNA library constructed in a GAL4AD vector encoding for prey proteins (Y). The transformed cells are plated onto medium that selects for the presence of both plasmids and are then screened for β-galactosidase activity. Interaction between the GAL4AD-Y fusion protein (prey) and the bait provided by the GAL4BD-X fusion protein activates transcription of *GAL1-lacZ* and

results in a blue yeast colony. The gene coding for the interacting protein can be recovered from the blue yeast colony.

Several versions of the basic two-hybrid system have been described in literature (Chien, Bartel et al. 1991; Durfee, Becherer et al. 1993; Vojtek, Hollenberg et al. 1993). One of the most used versions employs the selectable reporter gene *GAL1-HIS3*. In this system, the binding domain plasmid carries the yeast *TRP1* gene and a sequence derived from X fused to GAL4BD. The activating domain plasmid carries the yeast *LEU2* gene and a sequence from a cDNA library fused to GAL4AD. These plasmids are transformed into a yeast strain which requires tryptophan, leucine, and histidine (*trp1*, *leu2*, *his3*) and contains a functional copy of *HIS3* controlled by the *GAL1* promoter (*GAL1-HIS3*). If the binding domain and activating domain fusion proteins interact, the *GAL1-HIS3* reporter gene will be activated and the yeast cell will grow on Synthetic Dropout (SD) medium lacking the amino acids Trp, Leu, and His. Thus there is positive selection for the presence of activating domain plasmids encoding proteins that interact with the BD-X fusion protein. When compared with biochemical methods, the major distinction of two-hybrid method is that the readout is essentially generated by phenotypes that can be easily scored (cell growth, colour, fluorescence). Furthermore, two-hybrid interactors can be unambiguously identified by DNA sequencing of the plasmid encoding the prey protein, taking advantage of well-established tools for DNA analysis. It is this ability, to identify both an interacting protein and the gene which encodes for it, that makes the two-hybrid system so powerful.

Requirements of the two-hybrid system

The steps involved in accomplishing a two-hybrid screen with a specific protein gene (*X*) are as follows: (1) Construction of the BD-X fusion plasmid. (2) Transformation of the BD-X fusion plasmid into yeast. (3) Testing the BD-X fusion plasmid in yeast for auto-activation of reporter genes. (4) Testing the BD-X fusion plasmid for expression of X protein. (5) Construction of the AD-cDNA fusion plasmid library. (6) Transformation of the AD-cDNA fusion plasmid library into the yeast strain carrying the BD-X fusion plasmid and selection for the activation of *GAL1-HIS3*. (7) Screening putative positives for the activation of other reporter genes. (8) Identification of false positives. (9) Analysis of true positives. (10) DNA sequence analysis of AD-cDNA fusions. (11) Deletion mapping of interacting domains. These procedures require some expertise in DNA manipulation and yeasts handling. All of the following operations are necessary to ensure that the outcome is a collection of clones with the desired properties.

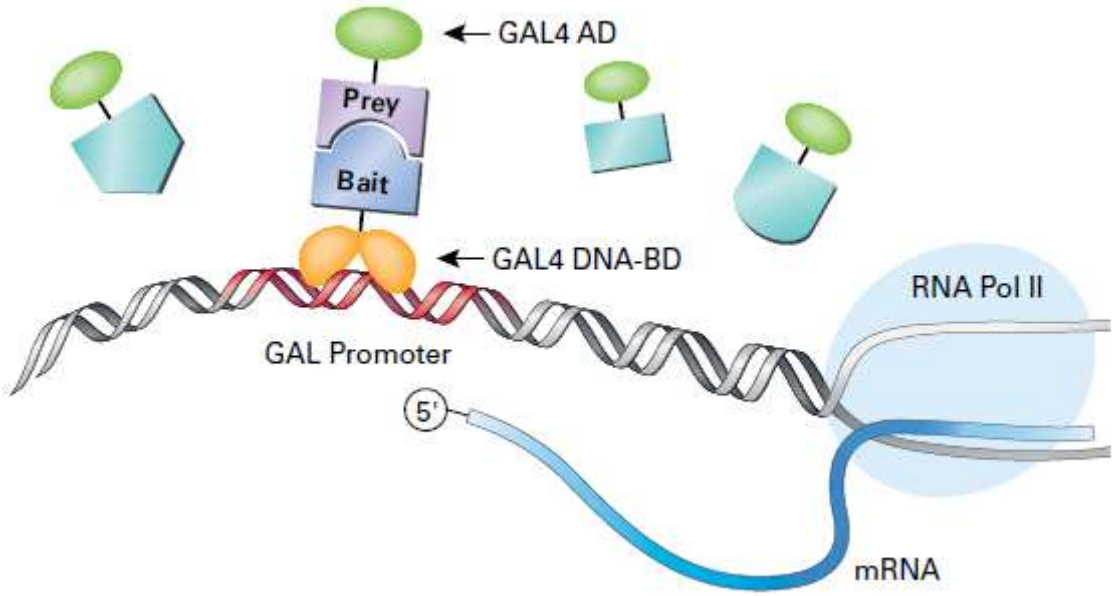

Figure 7. The two-hybrid principle. Two proteins are expressed separately, one (a bait protein) fused to the GAL4 DNA-binding domain (BD) and the other (a prey protein) fused to the GAL4 transciptional activation domain (AD).

Strong and weak points of yeast two-hybrid system

The two-hybrid system is a proven, invaluable tool of cell biology. It has a high sensitivity, enabling detection of weak and transient interactions. The experimental setup is relatively straightforward even if time consuming. It enables simultaneous detection and characterization of protein–protein interaction using a single protocol. It can be used to identify novel proteins interacting with the bait protein, without the need of any *a priori* knowledge. Since the experiment is undertaken in an *in vivo* system (e.g. yeast), proteins of interest are likely to fold appropriately and thus the screen is more likely to detect genuine interacting proteins. However, the two-hybrid system also has its specific limitations. Both false positives and false negatives can occur and are considered to be the most serious technical problems. False positivity can arise for a variety of reasons: proteins interacting with the DNA upstream of the reporter gene or with proteins that interact with the promoter sequence are prone to detection as false-positive in the two-hybrid system. For example, the original two-hybrid system was designed to activate transcription through RNA polymerase II. Therefore, it is problematic to identify novel proteins which interact with the RNA polymerase II activators (as bait). False-negativity can also be caused by a range of reasons: some proteins fused with BD or AD cannot localize in the yeast nucleus, cannot fold

properly, are not functional when expressed as a fusion protein, are toxic to the host, sometimes the appropriate posttranslational modification does not take place, or the potential interacting protein is not sufficiently represented in the library. Therefore, other independent methods (e.g. techniques based on biological function) must be used in parallel to confirm and verify the hits detected by the two-hybrid system

1.4.3: Evolution of two-hybrid system.

Over the years, scientists modified the traditional two-hybrid system technique. Whilst these derivatives all use similar principles, the various modifications enabled 'fine-tuning' of the approach for specific purposes.

1) A one-hybrid system was developed to study protein–DNA interaction, where the protein binds to AD and the DNA fragment of interest is cloned upstream of the reporter gene. If the protein binds to the specific DNA (promoter) of interest, transcription of the reporter gene is initiated (Meng and Wolfe 2006).

2) In order to characterize protein function, sometimes it is desirable to select for disruption of a particular interaction by introduction of point mutations, deletions or through the use of protein or pharmacological inhibitors. Thus, reverse and counter-selection two-hybrid systems were developed. In these systems, the wild-type protein which can interact with the bait, causes sensitivity to selection reagents (e.g. cytotoxic compounds) and the cells die. Only cells expressing proteins harboring the mutations disrupting the interaction will survive (Shih, Goldman et al. 1996; Vidal, Brachmann et al. 1996). Therefore, this approach can be used to identify novel mutants of the prey that no longer bind to the other protein of interest (the bait). Similarly, chemical compound libraries can also be screened for entities with the ability to block particular protein–protein interactions.

3) One of the major limitations of the traditional two-hybrid system has always been that some proteins, especially components of signal transduction pathways in higher eukaryotes, need post-translational modification for their ability to interact with others and thus they require modifying enzymes which are not present in yeast. The three-hybrid system was developed and applied to facilitate the post-translational modification of the prey proteins, by co-transfecting the necessary enzyme into the system, to ensure that prey proteins are functional in the host. This technique enables rapid mapping of modifications required for a certain inter-protein interaction. The most common such modification is

protein phosphorylation, which is most often induced by specific stimuli and requires the expression of particular kinases. An example of the application of the Y3H strategy was published by Osborne and colleagues. in their library screening study to identify proteins which specifically interact with an immunoreceptor tyrosine-based activation motif (ITAMs)-containing IgE receptor-derived, phosphorylated bait (Osborne, Dalton et al. 1995). In the bait, the gamma subunit of the high-affinity IgE receptor, FcεRI, was used to isolate a novel SH2-containing family member (interactions between FcεRI cytoplasmic tail and the Syk or Lyn SH2 domains), which requires the phosphorylationof the ITAMs by tyrosine kinases. A plasmid encoding the tyrosine kinase was introduced together with the bait and the prey. This method was initially used for characterizing interactions that are mediated by tyrosine phosphorylation, but can be adapted to other post-translational modifications. Another use of the three-hybrid system is detection of weak interactions between multiple proteins. In most cases, proteins bind to a number of other proteins and form large, multi-component complexes containing both weak and strong interactions. In order to identify novel proteins that weakly interact with a protein of interest, a known interacting protein can be co-expressed and this may supply a bridge and thus strengthen the interaction for the novel proteins with lower affinity interaction (Tomashek, Sonnenburg et al. 1996; Tirode, Malaguti et al. 1997).

4) In addition to the classical *Saccharomyces cerevisiae* system, other hosts, such as *E. coli*, have also been used. This was proposed to have multiple advantages over the yeast system, such as fast growth, higher transformation efficiency, nuclear localization not required, domains with eukaryotic activation domains do not activate *E. coli* transcription, and fewer indirect interactions involving bridging by endogenous proteins (review (Hu 2001)). A recent study exploiting the *E. coli* two-hybrid system was based on the twin-arginine translocation (Tat) pathway to identify interacting proteins in this pathway. Two reporter systems via the Tat pathway were used, one based on growth on selective media (maltose based) and the other enzymatic assay using a chromogenic substrate. Compared to other studies, the development of this *E. coli* two-hybrid system improved the accuracy of proteome-wide two-hybrid analyses (Strauch and Georgiou 2007).

5) In the original two-hybrid system, interacting fusion proteins (both prey and bait) need to be transported into the nucleus in order to activate the transcription of the reporter. This limits the interactions detected by the system. For instance, using full-length trans-membrane proteins as baits is problematic, due to misfolding or lack of localization in the nucleus. Some strategies to allow the two-hybrid system to take place in the cytoplasm and

membrane were developed to circumvent this pitfall. Instead of using nuclear transcription factors to construct two fusion proteins, *β*-galactosidase was split into two fragments and re-constituted through a bait/prey interaction, and the *β*-galactosidase activity served as a measurement of the strength of the interaction of the bait and the prey. Thus the protein–protein interaction could be studied in the cytoplasmic milieu (Rossi, Blakely et al. 2000). In another approach, the Ras-controlled signalling cascade on the *Saccharomyces cerevisiae* plasma membrane, which harbors temperature-sensitive Ras Guanine Exchange factors (GEF) Cdc25-2, was used. GEFs stimulate the transition of Ras between an inactive GDPbound form and an active GTP-bound form. Sos protein is one of the mammalian GEFs. If the Sos is recruited to the membrane, it will stimulate the transition of the Ras, hence the initiation of the signaling cascade, and allow growth at the non-permissive temperature (37 ◦C). Therefore, a bait (X) fused with Sos can identify membrane proteins (Y, prey) which interact with the bait. The readout of this system is based on the survival of *Saccharomyces cerevisiae* and the ability to grow when the temperature is increased from 25 ◦C to 37 ◦C (Aronheim, Engelberg et al. 1994; Aronheim, Zandi et al. 1997).

AIMS OF THE STUDY

The FA-BRCA pathway consists of at least 13 proteins, playing a role not yet elucidated in DNA-damage response. Additionally, FA proteins could also be involved in other mechanisms as oxidative stress, haematopoiesis and apoptosis. During the last few years, an intricate network of interacting proteins has been emerging, suggesting that many members potentially involved in this pathway are still missing. Consequently, identification of the FA interactors is regarded as essential achievement for the understanding of molecular mechanisms responsible for the disease.

In this perspective, the main goals of this PhD project were the identification and the characterization of novel genes involved in the FA-BRCA pathway. For this purpose, the work of this thesis was subdivided in two parts.

First, we have exploited the systematic approach of the yeast two-hybrid system to isolate potential interactors of FA proteins, focusing our attention on two novel proteins, one (UBE2U) binding FANCD2 and the other one (ISG20L2) binding FANCG. Then, we have characterized these two genes using different approaches, such as bioinformatics, biochemistry, molecular biology and cellular biology. Our attempts to functionally link these two proteins to the FA-BRCA pathway have so far failed. However, their function should be further investigated to exclude any potential connection with the FA proteins, at least during spermatogenesis, where the role of FA-BRCA pathway has never been studied, despite the high expression levels of its components in testis.

Chapter 2
MATERIALS AND METHODS

Plasmids

The coding sequences of *FANCA*, *FANCC*, *FANCD2*, *FANCE*, *FANCF*, *FANCG* full-length and their fragments were amplified by polymerase chain reaction (PCR) with *Pfu* DNA-polymerase (Promega) from the pREP4 expression plasmids in which they were originally cloned (IMAGE consortium). The *FANCL* cDNA was obtained by reverse transcription (RT)-PCR on CaCo2 mRNA (primers) and cloned into TOPO TA cloning (Invitrogen). Full length cDNAs and partially overlapping cDNA fragments of FA genes were subcloned in frame with the GAL4 DNA-Binding Domain into the yeast two-hybrid vector pGBKT7 (Clontech) and in frame with the GAL4 DNA-Activating Domain into the vector pGADT7 (Clontech)(Table 4). All constructs were verified by direct DNA sequence analysis by using T7 Sequencing Primer (5'–TAATACGACTCACTATAGGGC–3'), 3' DNA-BD Sequencing Primer (5'–ATCATAAATCATAAGAAATTCGCC–3') and 3' AD Sequencing Primer (5'–AGATGGTGCACGATGCACAG–3') The full-length cDNAs of clone 54 (short and long) were obtained by (RT)-PCR on human testis mRNA and subcloned into yeast vectors pGBKT7 and pGADT7 (Clontech). Clone 54 was subcloned into mammalian expression vector pcDNA 3.1 (Invitrogen) in fusion with either FLAG, or HA, or Myc tag Coding sequence of ISG20L2 was obtained by (RT)–PCR from HEK293 cell lines mRNA and it was cloned into TOPO TA cloning. Plasmid constructs were generated by subcloning cDNA from TOPO TA cloning by proper enzymatic digestion of the plasmid.

Bacterial strains and plasmids transformation

Escherichia Coli strains DH5α (Invitrogen) and Top10F (Invitrogen) were used in the construction and propagation of all plasmid constructs. For each transformation, 1-50 μg of DNA was added to 25-50 μl of chemically competent cells. Incubated on ice for 30 min, the cells were next subjected to heat shock at 37 or 42 °C for 1 min and next incubated on ice for 2 min. The cells were recovered in 1 ml Luria-Bertani (LB: 1% bacto-tryptone, 1% NaCl and 0.5% Bacto-yeast extract) or SOC (2% bacto-tryptone, 0.5% bacto-yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 20mM glucose) broth and then incubated for 40 min at 37°C with shaking (200-250 rpm). Cells were plated on LB-agar (18g/L) containing appropriate antibiotics (ampicillin or kanamycin 100μg/ml) with an incubation at 37°C overnight.

cDNA	Forward (5'-3')	Riverse (5'-3')
FANCA	CCGGAATTTAAGGCCATGTCCGACTCG	GCGTGTCGACCCGTCAGAAGAGATGAGG
FANCA1	CCGGAATTCAAGGCCATGTCCGACTCG	GCGTGTCGACCCACTCTCTCTGCATCTG
FANCA2	CCGGAATTCACTCACAAGATCGTGAGG	GCGTGTCGACAGAGTTGGGTTCTGCCCT
FANCA3	CCGGAATTCAAAGTCCCTGACTCCCGT	GCGTGTCGACATCTGAAAGAGCATCAGC
FANCA4	GCTCCATGGTGGACACACAGAACCTTC	CCGGAATTCTAGCTCCTCTCTCTCGCA
FANCA5	CCGGAATTCCAAGAAGGCCGGCAG TTT	GCGTGTCGACCCGTCAGAAGAGATG AGG
FANCC	CCGGGGATCCGACATCACCTTTTCGCT	CAAGTCTAGAAGGCACGCGTCGACACGC
FANCC1	CCGGGGATCCGACATCACCTTTTCGCT	GCGTGTCGACAGTGGCTATGATTTCCAG
FANCC2	CCGGGGATCCACCTTCCCAGCCTTGAA	CAAGTCTAGAAGGCACGCGTCGACACGC
FANCD2	CCGGAATTCGTGCACAAGACATTGGTC	GCGTGTCGACTGGGGTCTAATCAGAGTC
FANCD21	CCGGAATTCGTGCACAAGACATTGGTC	GCGTGTCGACAATACAGCTCTGACCGCT
FANCD22	CCGGAATTCGAGCTTCGGGAGAAGTTG	GCGTGTCGACAATACAGCTCTGACCGCT
FANCD23	CCGGAATTCAGTTTGACCCAAGAGAGA	GCGTGTCGACGATGAACTTCGTCACAAG
FANCD24	CCGGAATTCAGAGGCCAGCTAAACAAG	GCGTGTCGACTGCTGTGCCAGGCTCAAT
FANCD25	CCGGAATTCGAGGAGATTGCTGGTGTT	GCGTGTCGACTGGGGTCTAATCAGAGTC
FANCE	CCGGAATTCCGGCACCCGTGCCCCGGC	GCGTGTCGACGGTGATGGAGCACCAAGA
FANCE1	CCGGAATTCCGGCACCCGTGCCCCGGC	GCGTGTCGACCAAGGTCTTCAGCAGCTG
FANCE2	CCGGAATTCCCTATTAAGGACCAGCCT	GCGTGTCGACGGTGATGGAGCACCAAGA
FANCF	CCGGGGATCCTCATGGAATCCCTTCTG	TCTGTATAGGCAATGCTGGTCGACACGC
FANCG	CCGGAATTCGGCCCAGCCTCGGCCACC	GCGTGTCGACCTCTTCAAAACGTGGCAG
FANCG1	CCGGAATTCGGCCCAGCCTCGGCCACC	GCGTGTCGACAGTGCCACAATGAAGGGG
FANCG2	CCGGAATTCTCTAGGCTCTATCAGCAA	GCGTGTCGACCTCTTCAAAACGTGGCAG
FANCL	CCGGAATTCAGAGCTTTTCTGTGTTTCTC	GCGTGTCGACCTTATTTCAGTGTTTCCTTCC
Ube2u	CTACAGATGTTGCTTTCTTAC	CATGAAATGTGGATCTCCAACGTG
	GTAGCAACCTCAGTATGTAC	GCCAGGTCCTGAGCATCTCCCGG
UBE2U	ATGCACGGCA GAGCTTACCG	TTACCTGTCGTATGTAGTAAATCC
	ACAGAATTCA GTTTGGCAGGT	CATTGAGCAG CATCTTACTT GCC

Table 4: Primers used in this study

Isolation of plasmid DNA from E. coli.

Large-scale (midi-prep) and mini-prep plasmid DNA preparations were carried out using the QIAGEN MIDI and MINI prep kits, respectively. Both procedures are based on the alkaline lysis method, but use a support column to purify isolated plasmid DNA. The lysis is obtained with a 1% SDS, 0.2 N NaOH solution; the alkaline mixture ruptures the cells and the detergent breaks apart the lipid membranes and solubilizes the cellular proteins. The NaOH also denatures the DNA, providing single strands. The lysate is supplemented with acetic acid and potassium acetate (3 M K+ and 5 M acetate). The acetic acid neutralizes the pH, allowing the DNA strands to renature. The potassium acetate also precipitates the SDS from solution, along with the cellular debris. The *E. coli* chromosomal DNA, a partially

renatured tangle at this step, is also trapped in the precipitate. The plasmid DNA remains in solution. The centrifuged samples are then applied to kit spin columns, the plasmidic DNA binds to the silica matrix of the columns allowing repeated washes with an ethanol-based buffer. The purified DNA was eluted with dH_2O or Tris-HCl 10 mM, EDTA 1 mM pH 7.5 (TE) and was always checked by enzymatic digestion with the appropriate enzymes.

Quantification of plasmid DNA

DNA concentrations were determined for 1:10 to 1:100 dilutions of stocks, according to the following formula: absorbance of one A_{260} unit indicates a DNA concentration of approximately 50 μg/ml. Plasmid DNA concentration was determined by measuring the absorbance at 260 nm (*A*260) in a spectrophotometer using a quartz cuvette. For DNA quantification, *A*260 readings should lie between 0.1 and 1.0. An absorbance of 1 unit at 260 nm corresponds to 50 μg plasmid DNA per ml, considering the standard path length of 1 cm. This relationship is only valid for measurements made at neutral pH, therefore, samples were diluted in a low-salt buffer with neutral pH (e.g. Tris·Cl, pH 7.0).

Ligation reaction

The ligation reactions were generally carried out as follows:

X ng vector DNA : Y ng insert DNA = 1: 2-5

1 μl 10× ligation buffer (New England Biolabs)

0.5 μl T4 DNA Ligase (400 U/μl) (New England Biolabs)

dH_20 to 10 μl

Ligation was carried out for 1 h at room temperature; 1/10 to whole reaction volume was used for transformation of chemical competent *E.coli* cells.

Yeast two hybrid system

Yeast two-hybrid analyses were performed according to the manufacturer's instructions for MATCHMAKER 2 Hybrid System (Clontech).

Host strains

System 2 features the yeast strain AH109 (*MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2 : : GAL1*$_{UAS}$*-GAL1*$_{TATA}$*-HIS3, GAL2*$_{UAS}$*-GAL2*$_{TATA}$*-ADE2, URA3 : : MEL1*$_{UAS}$*-MEL1* $_{TATA}$*-lacZ)* using three reporter genes (ADE2, HIS3, and MEL1) under the control of distinct GAL4 upstream activating sequences (UASs) and TATA boxes. Strain Y187 (MATα, ura3-52, his3-200, ade2-101,

trp1-901, leu2-3, 112, gal4Δ, met–, gal80Δ, URA3 : : GAL1UAS-GAL1TATA-lacZ) contains the lacZ reporter gene under control of the GAL1 UAS. Unless stated otherwise, AH109 was used as the host strain for pGBKT7 (Clontech) constructs (baits) whereas Y187 was used as the host strain for pGAD (Clontech) constructs (preys).

Vectors

pGBKT7 vector (Clontech) encodes for the DNA Binding Domain (BD) of yeast transcriptional factor GAL4 and carries yeast nutritional marker for media lacking of tryptophan (i.e. SD-W). pGBKT7 confers bacteria resistance to kanamycin. pGADT7 vector (Clontech) encodes for the Activating Domain (AD) of yeast transcriptional factor GAL4 and carries yeast nutritional marker for media lacking of leucine (i.e. SD-L). pGADT7 confers bacteria resistance to ampicillin. pGBKT7-53 (Clontech) encodes for murine p53 fused with GAL4 DNA-BD while pGADT7-T (Clontech) encodes for large T-antigen of SV40 fused with GAL4 AD: p53 and large T-antigen interact in a yeast two-hybrid assay (Li & Fields, 1993; Iwabuchi *et al.*, 1993) and they were used as positive control in interaction matings. pGBKT7-Lam (Clontech) encodes a fusion of the DNA-BD with human lamin C and provided a control for a fortuitous interaction between an unrelated protein and either the pGADT7-T control or AD/library plasmid.

Yeast strain maintenance, recovery from frozen stocks, and routine culturing

Yeast strains were stored in YPD medium (20 g/L Difco peptone, 10 g/L Yeast extract, Clontech) with 25% glycerol at –70°C. Transformed yeast strains were stored in the appropriate SD dropout medium (Clontech) to keep selective pressure on the plasmid. To prepare glycerol stock cultures of yeast, a sterile inoculation loop was used to scrape an isolated colony from the agar plate. Cells were resuspended in 200–500 µl of YPD medium (or the appropriate SD medium) in a 1.5-ml microcentrifuge tube and vortexed vigorously to thoroughly disperse the cells. Sterile 50% glycerol was added to a final concentration of 25% and the vial was shaken before freezing at –70°C.A small portion of the frozen glycerol stock streaked onto a YPD (or appropriate SD) agar plate (YPD or SD, 20g/L Agar) and incubated at 30°C until yeast colonies reach ~2 mm in diameter (3–5 days), was used as working stock. Fresh working plates were streaked from the frozen stock at 1–2-month intervals. To prepare liquid overnight cultures, fresh (<1-month old) colonies were used from the working stock plate. One large colony was used (2–3-mm diameter) per 5 ml of medium. If colonies were small, or if it was necessary to inoculate a larger volume, several colonies

were used. Cells in the medium were dispersed by vigorously vortexing for ~1 min the suspension and subsequently incubated at 30°C for 16–18 hr with shaking at 230–270 rpm (OD_{600} > 1.5). For mid-log phase culture, the overnight culture was transferred into fresh medium so as to produce an OD_{600} = 0.2–0.3 and incubate at 30°C for 3–5 hr with shaking (230–250 rpm). This produced a culture with an OD_{600} ~0.4–0.6.

PEG/LiAc yeast transformation procedure

One-two weeks old colonies were inoculated into 1 ml of YPDA (YPD, adenine hemisulfate 0,003%) or proper SD and vortexed vigorously to break up any clumps. Cells were transferred to a flask containing 50 ml of YPDA or suitable SD and incubated at 30°C for 16–18 hr with shaking (250 rpm) to stationary phase (OD_{600}>1.5). The overnight culture was then transferred into 300 ml of YPDA (enough to produce an OD_{600} = 0.2–0.3) and incubated at 30°C for 3 hr with shaking (230–270 rpm) as the OD_{600} was 0.5 ± 0.1. Cells were then placed in 50-ml tubes and centrifuged at 1,000 x g for 5 min at room temperature. Supernatant was discarded and cell pellets were resuspended by vortexing in 25-50 ml of sterile TE (Tris-base 10 mM, EDTA 1 mM, pH 7.5) or H_2O. Pooled cells were centrifuged at 1,000 x g for 5 min at room temperature. Decanted the supernatant, cell pellets were re-suspended in 1.5 ml of freshly prepared, sterile TE/LiAc 100mM pH 7.5. For each transformation, 0,1 mg of herring testes carrier DNA (Clontech), DNA-BD/bait 0.1 mg and/or AD/library 0.1 mg were mixed to 0,1 ml of yeast competent cells by vortexing. Next, 0,6 ml of sterile PEG/TE-LiAc solution (PEG 4000 40%, Sigma) was added and the solution was mixed by vortexing at high speed. The mix was incubated at 30°C for 30 min with shaking (200 rpm). Seventy ul of DMSO (Sigma) were added and mixed well to each transformation solution. Cells were incubated for 15 min at 42°C in a water bath and then chilled on ice for 1–2 min. After centrifugation for 5 sec at room temperature at 13000 rpm, the supernatant was removed and cells were re-suspended in 0,5 ml of TE and plated on proper SD agar plates. To obtain a plate with well separated colonies, 100 µl of a 1:1000, 1:100, and 1:10 dilution were spread on 100-mm SD agar plates. These also provided controls for (co)transformation efficiency. Plates were incubated up-side-down at 30°C until colonies appeared (generally, 2–4 days). The largest colonies were picked and re-streaked on the same selection medium for master plates.

Autoactivation tests

In tests for auto-activation of reporter genes, all the DNA-BD and AD fusion constructs were independently transformed into strain AH109. The transformants were assayed for *MEL1* activation by plating on SD/–Trp/X-α-Gal and SD/–Leu/X-α-Gal respectively and subsequently checked for coloration. Positive and negative controls were carried out in parallel. Transformants giving white colonies were selected while the ones giving blue colonies were discarded.

Plating and screening transformation mixtures

Transformations of library screenings were plated on SD/–Ade/–His/–Leu/–Trp/X-α-Gal medium to screen for *ADE2, HIS3,* and *MEL1* expression. Library transformations were plated also on SD/–Leu/–Trp medium to select the DNA-BD and AD vectors. Small-scale transformations were plated on 100-mm plates while large- and library-scale transformations were plated on 150-mm plates and incubated at 30°C until colonies appeared. When screening an AD/library, the transformation efficiency and the number of clones screened were estimated as follows. Colonies (cfu) growing on the dilution plate that has 30–300 cfu were counted and :

cfu * total suspension vol. (µl)/ Vol.plated (µl)*dilution factor*DNA used (µg)= cfu/µg DNA represented the transformation efficiency, while

cfu/µg * µg of library plasmid used = No. of clones screened

Colony-lift filter assay

For each plate of transformants to be assayed, a sterile Whatman #5 filter was pre-soaked by placing it in 5 ml of Z buffer/X-gal solution ($Na_2HPO4*7H_2O$ 16.1 g/L NaH_2PO4*H_2O 5.50 g/L KCl 0.75 g/LMgSO4*7H2O 0.246 g/L, 0.27 ml β-mercaptoethanol (Sigma), 0.33 mg/ml X-gal) in a clean 100- or 150-mm plate. A clean, dry filter was placed over the surface of the plate of colonies to be assayed and it was gently rubbed to help colonies cling to the filter. In order to orient the filter to the agar, holes were poked through the filter into the agar in three or more asymmetric locations. Once the filter was evenly wetted, it was carefully lifted off the agar plate with forceps and transferred into a pool of liquid nitrogen (colonies facing up). The filter was completely submerged until completely frozen (~10 sec). It was then removed from the liquid nitrogen and allowed to thaw at room

temperature. This freeze/thaw treatment was to permeabilize the cells. The filter, colony side up, was thus placed on the pre-soaked filter avoiding trapping air bubbles and incubated at 30°C (or room temperature). It was checked periodically for the appearance of blue colonies (30 min - 6 hours). The β-galactosidase-producing colonies were identified by aligning the filter to the agar plate using the orienting marks. The corresponding positive colonies were picked from the original plates to fresh medium.

Plasmid isolation from yeast

For each plasmid purification, a large (2–4-mm) and fresh (2–4-day-old) yeast colony was inoculated into 0.5 ml of the appropriate SD liquid medium and vortexed vigorously to completely break up the colony and resuspend the cells. Cells were incubated at 30ºC overnight with shaking at 230–250 rpm. The day after, cells were spin down by centrifuging at 14,000 rpm for 5 min. The supernatant was poured off and the pellet was resuspended in the residual liquid (total volume ~50 µl).Ten µl of lyticase (Sigma) solution (5 units/µl in TE buffer) was added to each tube and the cells were thoroughly resuspended by vortexing or by repeatedly pipetting up and down. Tubes were incubated at 37ºC for 30–60 min with shaking at 200-250 rpm. Ten µl of 20% SDS was added to each tube and the mixture was vortexed vigorously for 1 min. To ensure the complete lysis of the cells, one cycle of freezing and thawing at –20ºC followed by mixing was carried out. The volume of the sample was brought up to 200 µl in TE buffer (pH 7.0) and 200 µl of phenol:chloroform:isoamyl alcohol (25:24:1) (JT Baker) was added to the sample. The mixture was vortexed at the highest speed for 5 min and subsequently centrifuged at 14,000 rpm for 10 min. The aqueous phase was transferred to a fresh tube. Eight µl of 10 M ammonium acetate and 500 µl of 100% ethanol (JT Baker) were added. DNA precipitation was enhanced by placing the solution at -70ºC for 1 hour. The suspension was centrifuged at 14,000 rpm for 10 min, the supernatant was discarded and the pellet was air-dried and next resuspended in 20 µl of H2O.

Rescue AD/Library plasmids via transformation of E. coli.

Yeast-purified plasmids DNA were transformed into *E. coli* as described and to select for transformants containing only the AD/library plasmid, bacteria were plated on LB medium containing ampicillin 100µg/ml.

Re-testing protein interactions in yeast

Protein interactions in yeast were confirmed by both co-transformation and yeast interaction mating.

1. Cotransformation

The DNA-BD/bait and AD/library plasmid were transformed together into AH109 using the small scale PEG/LiAc transformation procedure. Transformants were plated on SD/–Ade/–His/–Leu/–Trp and incubated at 30ºC for 3-5 days. Colonies were subjected to colony-lift filter assay.

2. Yeast Mating (direct interaction mating)

AH109 was transformed with the AD/library plasmid (prey) and selected on SD/–Leu while Y187 was transformed with the BD/bait plasmid and selected on SD/–Trp plates. For each candidate AD/library plasmid to be tested, several yeast matings were set up as control. For each mating experiment, one large and fresh colony of each type has been picked from the working stock plates and placed in one 1.5-ml microcentrifuge tube containing 0.5 ml of YPD medium. Cells have been completely resuspended by vortexing and then incubated at 30°C overnight (20–24 hr) with shaking at 200 rpm. Aliquots of 100-ml of the mating culture have been spread on the SD minimal media using double dropout (SD/–Leu/–Trp) to select for diploids and both plasmids and quadruple dropout (SD/–Ade/–His/–Leu/–Trp/X-α-gal) to select for diploids in which a positive two-hybrid interaction occurred.

Preparation of yeast cultures for protein extraction

For each transformed yeast strain assayed in a western blot, a 5-ml overnight culture in the appropriate SD selection medium was prepared as described previously, except single isolated colony was used. As a negative control, a 10-ml culture of an untransformed yeast colony in YPD medium was prepared. For each clone to be assayed (and the negative control), aliquots of 50 ml of YPD medium were separately inoculated with the entire overnight culture at 30°C with shaking (220–250 rpm) until the OD_{600} reached 0.4–0.6. The OD_{600} (of a 1-ml sample) was multiplied by the culture volume to obtain the total number of OD_{600} units. The culture was quickly chilled by pouring it into a prechilled 100-ml centrifuge tube halfway filled with ice and the tube was immediately placed in a prechilled rotor and centrifuged at 1000 x g for 5 min at 4°C. Supernatant was poured off and the cell pellet was resuspended in 50 ml of ice-cold H2O. The pellet was recovered by centrifugation at 1,000 x g for 5 min at 4°C and immediately frozen by placing the tube in liquid nitrogen. Cells were stored at -80ºC.

Yeast protein extraction (Urea/SDS method)

Complete cracking buffer (Urea 8 M, SDS 5% w/v, Tris-HCl [pH6.8] 40 mM, EDTA 0.1 mM, Bromophenol blue 0.4 mg/ml, β-mercaptoethanol 10%, protease inhibitor solution (Sigma))was pre-warmed at 60°C and 100 µl of cracking buffer per 7.5 OD_{600} units of cells was used. Cell pellets were quickly thawed by separately resuspending each one in the prewarmed cracking buffer and putting them for 1 min at 60°C. Because the initial excess PMSF in the cracking buffer degrades quickly, additional aliquots of the PMSF stock solution to the samples were added after every 7 min, approximately, until they are placed on dry ice or stored at –80°C. Each cell suspension was transferred to a 1.5-ml screw-cap microcentrifuge tube containing 80 µl of glass beads (Sigma) per 7.5 OD_{600} units of cells. Samples were heated at 70°C for 10 min and then mixed vigorously by vortexing for 1 min. Debris and unbroken cells were pelleted in a microcentrifuge at 14,000 rpm for 5 min at 4°C and supernatants were transferred to fresh 1.5-ml tubes and placed on ice (first supernatants). Pellets were treated as follows: tubes were placed in a 100°C water bath for 3–5 min and subsequently vortexed vigorously for 1 min. Debris and unbroken cells were pelleted in a microcentrifuge at 14,000 rpm for 5 min at 4°C and each supernatant (second supernatant) was combined with the corresponding first supernatant. Samples were boiled briefly and immediately loaded onto a gel or stored in a –70°C freezer until ready for the western blotting analysis.

Proteins expression in yeast

Western blots of transformants were performed as described below and the blots were probed with antibodies directed against the c-Myc and HA epitope tags. Untransformed yeast were used as a control.

Sequence AD/Library inserts

DNA isolated from E.coli was used and inserts in the positive AD/library plasmids were sequenced using the 3' AD Sequencing Primer and T7 Sequencing Primer provided with the system. The presence of an open reading frame (ORF) fused to the GAL4 AD sequence was verified and the sequence was compared to those in GenBank, EMBL, or other databases.

Additional two-hybrid tests

For each interaction confirmed in the yeast direct interaction mating, the library insert was transferred from the AD to the DNA-BD vector and vice versa subcloning with suited enzymatic digestions and ligation. If restriction sites were not compatible between the prey and the vector, cDNA library fragments were first amplified by PCR and then subcloned by enzymatic digestion and ligation. Next, the two-hybrid assay was repeated.

Cell cultures, treatment and isolation

Human embryonic kidney (HEK293) and cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Gibco, Invitrogen), supplemented with 10% foetal bovine serum (FBS) (Euroclone, Celbio), 100 U/ml penicillin (Celbio), and 100 μg/ml streptomycin (Celbio), at 37 °C in 5% CO_2. Cells were routinely maintained in T-75 cell culture flasks (Corning, Costar) or 100 mm tissue culture dishes (Corning, Costar).Treatment with N-acetyl-L-leucyl-L-leucyl-L-norleucinal, (LLnL), was done by adding the proteasome inhibitor at a final concentration of 10 μM in the culture medium 8-16 h prior harvesting. HEK293 cells were transfected with LipofectAMINE2000 (Invitrogen) while HeLa cells were transfected with PolyFect transfection reagent (Qiagen), accordingly to the manufacturers instructions. Testes from 18-20 days old CD1 (Charles River) mice were used to obtain pachytene spermatocytes by elutriation technique. Spermatogonia were obtained from mice 8-9 days old and round spermatids from 30 days old testes. Spermatozoa were obtained from the cauda epididymidis of adult (4-12 months old) mice.

Immunofluorescence

Cells were grown on four well chamber slides (Becton Dickson) and transiently transfected as described. After 24-72 hours, cells were washed with phosphate buffered saline (PBS, Gibco, Invitrogen) and fixed with 4% paraformaldehyde (PFA, Sigma) in PBS for 20 min. Next, cells were permeabilized with 2% Triton X-100 (Roche) in PBS for 10 min and treated with blocking buffer (5% FCS, 0.2% Tween-20 (Sigma) in PBS for 30 min. Cells were subsequently incubated with anti-FLAG (Sigma) at 1/500 dilution or anti-HA (Santa Cruz) at 1/100 dilution in blocking buffer for 2 h at room temperature. After washing five times with 0.2% Tween-20 in PBS, cells were incubated in blocking buffer with either FITC- or TRITC- conjugated anti-mouse or anti-rabbit antibody (Alexa Fluor, Molecular Probes), 1/250 dilution, at room temperature for 1h. After five more washes of 0.2% Tween-20 in PBS, chamber walls were removed and slides were mounted in vectashield with 4',6-

diamidino-2-phenylindole (DAPI) (Vector laboratories). The cells were examined under a fluorescence microscope (Zeiss, Axiosckop 2).

Preparation of whole-cell and nuclear/cytoplasmic lysates.

Total cell extracts from 1 to 2 X 10^7 were prepared by lysis in 1 ml of ice-cold RIPA buffer (20-50 mM Tris-HCL, pH 7.5-8.0, 150 mM NaCl, 1% Igepal CA-630, 0.5-1% sodium dodecylsulfate (SDS) with complete protease inhibitor cocktail (Sigma) and PMSF). The cell suspensions were disrupted with a sonicator for 10 sec (twice, with a 30 sec interval) and subsequently incubated in ice for 20 min mixing occasionally. Lysates were cleared by centrifugation at 15000 X g for 15 min at 4ºC and supernatants were recovered and transferred into a new tube. For extraction of proteins from mouse testis, the tissue has been homogenized into a ice-cold pestle with ice-cold RIPA buffer and complete protease inhibitors cocktail. After incubation of 30 min in ice, mixing occasionally, lysates were cleared by centrifugation at 15000 X g for 15 min at 4ºC. Supernatants were recovered and transferred into a new tube. For nuclear and cytoplasmic extracts, NE-PER extraction reagents kit (Pierce) was used, following the manufacturer's instructions.

Immunoprecipitation.

Protein samples of 0.750-1.500 mg were immunoprecipitated by overnight incubation at 4ºC with 1-2 μg of the designed antibody on a rotating wheel. Protein A or protein G (Sigma Aldrich) resin was blocked for 1 hour with BSA 1mg/ml in RIPA buffer prior incubation with the immunoprecipitation solution for 1 hour. Next, the resin was washed four times with RIPA buffer and protein samples were eluted with loading sample buffer (LSB, 50 mM Tris-HCl, pH 6.8, 2% sodium dodecylsulfate [SDS], 5% β-mercaptoethanol, 2.5% glycerol and 0.5% bromophenol blue) and denatured at 100 °C for 5 min. Immunoprecipitations with with untransfected cells, pre-immune anti-sera and non-related antibodies were carried out as controls.

In vitro transcription and translation.

Transcription and translation *in vitro* were carried out by using TNT® T7 Quick Coupled Transcription/Translation System (Promega) following manufacturer's instructions.

Briefly, 40µl of TNT® Quick Master Mix, 1ul of Methionine, 1mM, 2µl of plasmid DNA template(s)
(0.5µg/µl) and 7 µl of nuclease-free were mixed and incubated at 30ºC for 60-90 min. Empty vectors or plasmid vectors containing unrelated cDNA were used as negative controls. The results of translation were analysed by western blotting.

Protein quantification.

Protein concentrations were determined using the Bradford assay (Bradford 1976). Protein samples were mixed with Bradford reagent (Bio-Rad) and the absorbance at 595 nm was measured in a spectrophotometer (Varian). Protein absorbances were converted to mg/ml concentrations using a standard curve constructed by measuring the absorbance of a range of bovine serum albumin (BSA) concentrations.

SDS/PAGE and Western blotting.

Protein samples were resuspended in loading sample buffer (20 mM Tris-HCl, pH 6.8, 2% sodium dodecylsulfate [SDS], 5% β-mercaptoethanol, 2.5% glycerol and 2.5% bromophenol blue) and denatured at 100 °C for 5 min. Thirty to seventy-five µg protein was analyzed by SDS polyacrylamide gel electrophoresis (SDS-PAGE) on 10, 12.5 or 15% (w/v) polyacrilamide gels, using the Bio-Rad maxi and mini gel equipment, as detailed in the manufacturer instructions. Protein molecular weight markers used to establish the apparent molecular weights of proteins resolved on the SDS polyacrilamide gels were: New England Biolabs (175-6.5 kDa range), Benchmark Protein ladder (Invitrogen, 220-10 kDa range) and See blue Plus 2 pre-stained (Invitrogen, 198-3 kDa range). All of the mini gels were run at 100 V while all of the maxi-gels were run at 120 V with samples in the stacking gel and 200 V with samples in the resolving gel for an appropriate length of time, using SDS running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS). Proteins were electroblotted onto polyvinylidene difluoride (PVDF) membranes (Immobilon-P, Millipore) or Nitrocellulose membrane (Millipore) using a wet blotter (Bio-Rad) at 100 V for 1 h or at 40 V over night in blotting buffer (25 mM Tris, 192 mM glycine, 20% methanol). Membranes were blocked with 5% non-fat milk (BioRad) in TBS, pH 7.6, 0.5-1.0% Tween-20 (TBS-T) for 1 h. The membranes were then incubated with polyclonal rabbit antibodies or monoclonal mouse antibodies for 1 h, using the recommended dilution for each antibody as described in the antibody section. After 4 washes in TBS-T for 5-10 minutes, the membranes were incubated with a horseradish peroxidase (HRP)-conjugated anti-mouse or anti-rabbit antibody

(1:10,000) (Santa Cruz Biotechnology), and the immunoblots were visualised using ECL detection kits, with enhanced chemiluminescence (Pierce).

Antibodies

Antibodies used include the following: rabbit anti-FANCA (Covance) ; monoclonal and polyclonal anti-FANCD2 (FI-17; Santa Cruz Biotechnology and Novus Biotechnology respectively); rabbit polyclonal anti-*fancd2* (kind gift of dr.Alan D'Andrea), rabbit polyclonal anti-FANCG (Fanconi Anemia Research Fund); monoclonal anti-FLAG (M2; Sigma- Aldrich), monoclonal anti-cMyc (9E10; Santa Cruz Biotechnology), and monoclonal anti-HA (12CA5; Roche Diagnostics). A mouse ERK antibody (1:500) (Santacruz Biotechnology) was used as the control for equal loading of total extracts. Mouse anti-HDAC (Santacruz Biotechnology) or anti-Sp1 antibody (Santa Cruz Biotechnology) for equal loading of nuclear lysates. Rabbit antibodies against clone 54 murine and against ISG20L2 human were generated by PRIMM, after immunization with peptides specifically designed by the company.

RNA isolation and cDNA preparation.

Total RNA was extracted from cell lines by the guanidinium isothiocyanate method, using Trizol reagent (Invitrogen). The RNA concentrations were determined by absorbency at 260 nm, and their quality was verified by the integrity of the 18S and 28S rRNAs after ethidium bromide staining of the total RNA samples subjected to 1% agarose gel electrophoresis. Synthesis of cDNA from total RNA (4 µg) was performed using the Super Script II First Strand Kit (Invitrogen), with random hexamers, according to the manufacturer instructions.

Bioinformatic analysis.

The human sequence database (http://genome.ucsc.edu/cgi-bin/hgGateway) was screened by using as initial query the sequence the human proteins and the BLAT program. The Genscan and Gene Predictions were corrected by searching the nonredundant sequence database and the human EST database (http://www.ncbi.nlm.nih.gov/) using BLAST program. The mouse sequence database were screened by using the human cDNAs and the

BLAT program. The retrieved putative sequences were then ascertained by searching the nonredundant sequence and the human EST database. To characterize the genomic structure, both human and mouse cDNAs were used to screen the respective genomic DNAs at the Human Genome Browser Gateway and in the High Troughput Sequences Database.

Northen blot analysis.

A membrane carrying a panel of human polyA+ RNA was hybridized with ^{32}P-labeled clone 54 and ISG20L2 human probes according to the manufacturers instructions (Clontech).

RNA in situ hybridization.

The clone 54 antisense probe was obtained by linearization of cDNA in pCRII-TOPO vector (TOPO TA cloning, Invitrogen) with EcoRV (New England Biolabs) and transcription with SP6 RNA polymerase, while digestion of the same plasmid with BamHI (New England Biolabs) and transcription with T7 RNA polymerase generated the sense control probe. The ISG20L2 antisense probe was obtained by linearization of the cDNA pCRII-TOPO vector with XhoI and transcription with T3 RNA polymerase, and the sense control probe was transcribed with T7 RNA polymerase after digestion of the plasmid with NotI. Embryos, at different developmental stages, were harvested from CD1 pregnant female mice and fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) overnight at 4ºC.Whole-mount in situ hybridization analysis was performed on embryos at E9.5 and E10.5. Embryos were treated with proteinase K (10 Ag/ ml) for either 7 (E9.5) or 10 min (E10.5). Hybridization with the digoxigenin-labeled probes (as reported below) was performed overnight at 68ºC in 50% formamide, 5_ standard saline citrate buffer (SSC; pH 4.5), 1% sodium dodecyl sulfate (SDS), 50 Ag/ml yeast RNA, and 50 Ag/ml heparin and the digoxigenin-labeled probes (1 Ag/ml). The following day, after two washes in a solution containing 50% formamide, 4_ SSC, and 1% SDS at 68ºC, embryos were washed twice in the solution containing 50% formamide and 2_ SSC at 65ºC and five times in 100 mM maleic acid, 150 mM NaCl, 0.1% Tween 20, pH 7.5 (MABT) and finally treated for 2 h in 1% blocking reagent (Roche) in MABT with 2% sheep serum. Embryos were incubated overnight with alkaline phosphatase (AP)-labeled anti-digoxigenin antibody (1:2000; Roche) in 1% blocking reagent in MABT at 4ºC. After extensive washes in Tris buffer saline with 0.1% Tween 20 (TBST) embryos were exposed to the substrate for AP, nitroblue tetrazolium–5-bromo-4-chloro-3-inodoyl phosphate (NBT-BCIP; Sigma). Staining was

blocked by washes with PBS, pH 5.5, followed by postfixation in 4% paraformaldehyde for 1 h. Hybridized embryos were photographed at the stereoscope using a digital camera (Leica DC500). Embryos at E14.5, neonatal mice, and adult tissues were analyzed by in situ hybridization of sectioned tissue. To this purpose, embryos and dissected organs were cryoprotected by treatment with 30% sucrose in PBS and embedded in 7.5% gelatin. Twenty-micrometer cryosections were collected on Superfrost Plus slides and postfixed with 4% paraformaldehyde in PBS for 15 min. After being bleached with 6% H2O2 in PBS with 0.1% Tween 20, sections were treated with either 1 (embryonic tissues) or 5 Ag/ml (neonatal and adult tissue) proteinase K for 15 min. After washes with 2 mg/ml glycine and postfixation with 4% paraformaldehyde and 0.2% glutaraldehyde sections were prehybridized with 50% formamide, 5_ SSC (pH 4.5), 1% SDS, 50 Ag/ml yeast RNA, and 50 Ag/ml heparin. Hybridization was performed overnight at 65jC. Hybridized sections were washed with 50% formamide, 4_ SSC, and 1% SDS at 65ºC and with 50% formamide and 2_ SSC at 60ºC. Sections were blocked for 1 h with 1% Blocking Reagent (Roche) in MABT containing 10% sheep serum and incubated with AP-labeled anti-digoxigenin antibody(1:2000; Roche) in 1% Blocking Reagent in MABT overnight at 4ºC. After extensive washing with TBST, sections were exposed to NBT-BCIP (Sigma). Reaction was blocked by washes with PBS, pH 5.5, followed by postfixation in 4% paraformaldehyde for 20 min. Slides were coverslipped with 70% glycerol in PBS and photographed (AxioCam digital camera; Carl Zeiss, Gottingen, Germany) using a microscope with Nomarski optics (Axioplan; Zeiss).

Tissue preparation.

Freshly frozen, postfixed sections were used for all of the experiments. At least one litter was used for each age group from which data are reported, with at least two embryos examined at each age. Pregnant CD1 mice were decapitated. For embryos at E14.5 and the entire uterus were freshly frozen and embedded in Optimal Cutting Temperature (OCT) compound (Roche Products, UK) for cryostat sectioning. Mouse embryos older than E14.5 were dissected out quickly and staged according to their limb bud morphology, before being frozen and embedded in OCT. Embryonic trunk sagittal and transverse sections (14 μm) were prepared to detail the expression patterns. Sections were mounted onto 3-aminopropyltriethoxy-silane-coated microscope slides and stored at -80 °C until use in the hybridization procedure.

Polymerase chain reaction (PCR)

The general PCR conditions were as follow: (*Pfu* polymerase)

1 µl cDNA template

5 µl Pfu buffer

20 pmol forward primer

20 pmol reverse primer

1 µl Pfu polymerase (3 U/µl)

0.2 mM of each dNTP

dH_20 to 50 µl

The PCR reaction was performed with the following cycling parameters:

1 cycle	1 min at 95 °C
30 cycles	30 sec at 95 °C
	30 sec at the annealing temperature
	2 min at 72 °C (1min per each kb of DNA to amplify)
1 cycle	10 min at 72 °C

Restriction enzyme digestion of DNA

The analytical DNA digestions were performed in a final volume of 20-50 µl, at 37 °C for 1-2 h. All of the restriction enzymes were from New England Biolabs, and the digestions were performed in the appropriate buffers, supplied by the manufacturer with the enzyme. All of the digestions were analysed by agarose gel electrophoresis. Each analytical digestion used 500 ng plasmid DNA, with 3 units of enzyme per µg plasmid DNA.

Agarose gel electrophoresis

Agarose gels (1% w/v in TAE, 40 mM Tris-acetate, pH 7.5, 2 mM EDTA) were prepared and supplemented with ethidium bromide (1 µg/ml). The percentage of the agarose in gels was determined depending on the size of the DNA fragments to be resolved. Gels were generally run at 120 V in 1× TAE buffer, and DNA was visualized on a UV transilluminator.

Isolation of DNA from agarose gels

Following agarose gel electrophoresis, the DNA-containing gel slices were excised under UV light. The DNA was extracted from these gel slices using QiaexII Gel Extraction Kit (QIAGEN), following the protocol supplied by the manufacturer. Briefly, the gel slices containing the DNA were dissolved in sodium perchlorate, with the buffer composition adjusted to allow binding of the released DNA on to a silica support. The bound DNA was then washed from the agarose and the electrophoresis buffer with alcohol-containing buffer. Purified DNA was eluted from the resin using 20 μl dH_20.

DNA sequence analysis

For DNA sequence analysis, plasmids or PCR were processed by the DNA sequencing core at TIGEM. Sequencing reactions were carried out using Taq DNA polymerase and fluorescently labelled dideoxy terminators, and then run on our automated sequencers. Control reactions are carried out to insure high quality control. Single run sequencing reactions can yield 600 to 700 bases routinely, depending on the type of template and primer used. Sequences obtained with an automated sequencer begin roughly 15 to 20 bases after the primer sequence.

Chapter 3

RESULTS

3.1 Identification of putative FA interactors using a yeast two-hybrid system approach.

3.1.1: Library screening.

The first part of this PhD project was aimed at the identification of new interactors of FA proteins by using the systematic approach of the yeast two-hybrid screening. At this purpose were generated 23 bait constructs of 7 FA genes (partial and full length), fused with the GAL4 DNA binding domain (BD) in pGBKT7 vector (table 5). Since functional domains of the FA proteins employed were not known, the length of each partial cDNA fragment was arbitrarily chosen to be approximately 1000 bp, so as to maximize the chances of expression in yeast. Fragments of cDNA were designed to partially overlap each other, amplified by PCR and subcloned into the bait vector. Once transformed in AH109, each construct was tested for:

- Toxicity effects in yeast (by monitoring the culture growth rate)
- Self-activation of reporter genes (by plating them on selective media: SD-WH, SD-WA, SD-WHA in presence of X-α-Gal)
- Protein production and stability (by western blotting analysis)

No baits were discarded for toxicity effects. FANCG, $FANCG_1$ $FANCA_1$ and $FANCA_2$ were rejected for auto-activation of reporter genes. FANCA and $FANCC_2$ were discarded as no protein was detected while FANCD2 was expressed in low amount. Of the 16 baits suited for yeast two-hybrid assay, we used 8 constructs, $FANCA_3$, $FANCA_5$, $FANCD2_1$, $FANCD2_2$, $FANCD2_3$, $FANCD2_4$, $FANCD2_5$ and $FANCG_2$ to screen a commercial cDNA library. They included all the FANCD2 baits, since FANCD2 is considered a fundamental 'crossroad' for FA-BRCA pathway (in review (Wang 2007)), the only suitable fragment of FANCG since it is crucial for the FA complex assembly (Garcia-Higuera, Kuang et al. 1999; Medhurst, Huber et al. 2001) as well as two of the three suited FANCA baits because FANCA (as well as FANCD2 (Garcia-Higuera, Taniguchi et al. 2001)) binds directly BRCA1, connecting FA pathway with DNA-repair (Folias, Matkovic et al. 2002). A human testis library fused with the GAL4 activation domain (AD) into the pACT2 vector was used for the screening. The rationale behind this choice is that FA genes are expressed in testis at high levels (Chen, Tomkins et al. 1996; Lo Ten Foe JR 1996; Whitney, Royle et al. 1996) and testis manifests pathological phenotypes in both patients and animal models (Chen, Tomkins et al. 1996; Whitney, Royle et al. 1996; Cheng, van de Vrugt et al. 2000; Yang, Kuang et al. 2001;

Koomen, Cheng et al. 2002; Houghtaling, Timmers et al. 2003; Wong, Alon et al. 2003). Moreover, FA-BRCA pathway in this tissue was still uncharacterized.

We selected and analysed 167 clones by direct sequencing (table 6). Databases queries allowed the identification of 60 putative interactors, as they belong to known or predicted proteins, and exclusion of 107 clones as genomic, Alu or out of frame sequences, 5' and 3' untranslated regions.

Construct	FA protein	Insert (protein amminoacids)
pGBKT7-FANCA	FANCA	FULL LENGTH (1-1456)
pGBKT7- FANCA$_1$		1-355
pGBKT7- FANCA$_2$		291-647
pGBKT7- FANCA$_3$		594-948
pGBKT7- FANCA$_4$		910-1256
pGBKT7- FANCA$_5$		1210-1456
pGBKT7- FANCC	FANCC	FULL LENGTH (1-558)
pGBKT7- FANCC$_1$		1-325
pGBKT7- FANCC$_2$		246-558
pGBKT7-FANCD2	FANCD2	FULL LENGTH (1-1468)
pGBKT7-FANCD2$_1$		1-359
pGBKT7-FANCD2$_2$		321-671
pGBKT7- FANC D2$_3$		615-961
pGBKT7- FANC D2$_4$		927-1270
pGBKT7- FANC D2$_5$		1233-1468
pGBKT7- FANC E	FANC E	FULL LENGTH (1-537)
pGBKT7- FANCE$_1$		1-308
pGBKT7- FANCE$_2$		249-537
pGBKT7- FANCF	FANCF	FULL LENGTH (1-374)
pGBKT7- FANCG	FANCG	FULL LENGTH (1-622)
pGBKT7- FANCG$_1$		1-351
pGBKT7- FANCG$_2$		287-622
pGBKT7- FANCL	FANCL	FULL LENGTH

Table 5: FA bait plasmids generated for this work and their length (in aa). In red, baits giving auto-activation. In brown, baits producing low or no amount of protein. In black, baits suited but not used for yeast two-hybrid screening. In green, baits used to screen a cDNA testis library.

BAITS	No of clones after			
	Library screening	Sequencing analysis	Direct mating BD-bait/AD-prey	Direct mating AD-bait/BD-prey
pGBKT7-FANCA$_3$	1	1		
pGBKT7-FANCA$_5$	2	2		
pGBKT7-FANCG$_2$	134	44	**7**	**clone 4 (ISG20L2)**
pGBKT7-FANCD2$_1$	0	0		
pGBKT7-FANCD2$_2$	0	0		
pGBKT7-FANCD2$_3$	29	12	**3**	**clone 54 (UBE2U)** **clone 62 (CRIPS1)**
pGBKT7-FANCD2$_4$	1	1		
pGBKT7-FANCD2$_5$	0	0		

Table 6: Library screening and yeast interaction mating results.

3.1.2: Direct interaction matings by co-transformations.

In order to investigate the interactions found, the 60 preys were tested in a direct interaction assay with their respective baits, allowing us to confirm 10 bindings (table 6). Next, preys and baits were switched into the vectors and tested by direct interaction matings in this new combination. All of them gave no auto-activation of reporter genes and each interaction was tested in three independent analyses. Three clones resulted positive, as confirmed also by X-α-Gal tests. Specifically, the clone 4 was found to interact with FANCG$_2$, while clones 54 and 62 bound FANCD2$_3$ (table 6). To validate these interactions in yeast, full length cDNA of clones 4, 54 and 62 were obtained by RT-PCR and cloned into pGBKT7 and pGADT7 vectors, in frame with BD and AD, respectively. After auto-activation tests and protein production detection, they were used for yeast direct interaction mating as described previously. All interactions, between full length FANCD2 and clones 54 and 62 and between FANCG$_2$ (the only FANCG portion without auto-activaction) and clone

4, were confirmed. Since clone 62 represented the CRISP2 (cysteine-rich secretory protein 2) gene, known to be specifically expressed in testis without any evident association with FA pathway, we decided to focus our attention on the other two clones, because potentially connected to the FA-BRCA pathway. In fact, accordingly to databases analysis and bioinformatic tools, clone 4 represented a novel *ISG20L2* gene encoding for a protein with a putative conserved domain of 'EXOIII' family, which includes a variety of exonuclease proteins, therefore possibly linked to DNA-repair. Clone 54 was a novel gene as well (*UBE2U*), for the synthesis of a predicted ubiquitin-conjugating enzyme E2 and as mentioned, monoubiquitination of FANCD2 is a crucial step in the FA-BRCA pathway. While the putative E3 in the pathway was known as the FANCL product, the E2 enzyme involved in the ubiquitination modification was unknown at the time of this study. Then, the work of this thesis focused on the characterization of these two novel genes, including expression studies and validation of the interactions with the respective proteins.

3.2 UBE2U: a novel putative Ubiquitin-conjugating enzyme E2.

3.2.1: Cloning of human and murine clone 54: UBE2U.

Clone 54 had a putative open reading frame of 226 amino acids. The predicted protein showed high homology to the members of the ubiquitin-conjugating enzyme family E2 and it is reported with the provisional name of UBE2U in database. The human sequence databases were screened using as initial query the sequence of clone 54 and BLAST and BLAT programs. The screening identified ESTs that allowed the extension of the transcript at both ends and the identification of putative alternative forms. Combining results from EST database and exon-intron genomic structure, at least three transcripts of approximately 1.6, 3.3 and 3.8 kb were expected, all resulting in alternative splicing (Fig. 10). Both transcripts of 3.3 and 3.8, share the same open reading frame for 226 residues (isoform A, GenBank Accesion number NP_689702). They differ at the 3' UTR in that a splicing event between exons 8a and 8b removes intronic bp in transcript of 3.3 kb (Fig.10). The transcript of 1.6 kb derives from the assembly of another two exons: the 9 and 10. It encodes for a 317 amino acids protein (isoform B), consisting of 91 additional amino at the COOH-terminus of

isoform A (Fig. 8). The murine putative *Ube2u* gene was also cloned by bioinformatic tools, using the human cDNA and protein sequences, as queries. The BLAST and BLAT programs allowed us to retrieve predicted genes (chr4_19.128 and chr4_19.129) on syntenic regions of chromosome 4, the cDNA of which was blasted against the non redundant sequence and the human EST databases to reassure the sequence. The mouse transcript was of 1552 bp, resulting in a protein of 352 amino acids (Fig. 8), consistent with a predicted sequence also obtained by automated computational analysis (GenBank Accession Number XP_355502). Contrarily to what observed in the human *UBE2U* gene, there was no evidence for alternatively spliced forms.

3.2.2: Homology, genomic organization and alternative splicing of the UBE2U gene.

The alignment between the human and the mouse proteins is shown in Fig. 8, where we also included the Macaca (GenBank Accession Number BAB69736) and rat (as determined by bioinformatic analysis) genes. The alignment shows that the homology between the orthologs is conserved in correspondence of human isoform A, while it is almost absent in the COOH-terminal region, containing the 91 human additional residues that characterize isoform B. In this work we refer to UBE2U as isoform A unless stated otherwise. The human and mouse genomic structures were determined by searching the respective genome database using as a template the cDNAs in a BLAT analysis. Positive match were obtained in the syntenic region of human 1 and mouse 4 chromosomes. Ten exons with the appropriate consensus sequences from intronic splicing were identified in both species. The size of all exons perfectly matched, suggesting that the genomic structures between the two genes are well conserved. Comparison of the genomic sequence of the human and mouse genes in a mVista search allowed the identification of conserved regions only corresponding to exons 1 to 8.

3.2.3: Interaction of clone 54 (UBE2U) and other FA proteins in yeast.

In order to determine whether UBE2U is part of the FA-BRCA pathway, we analyzed interactions between UBE2U and other FA proteins by direct yeast interaction matings. In these experiments, cDNA full length of human UBE2U was used as bait and FA proteins (or

partial proteins) as preys and vice-versa (table 7). In these studies, not only we did confirm the interaction of UBE2U with FANCD2$_3$, but we also found that UBE2U interacts with FANCC, FANCF and FANCL.

Figure 8. Alignment of human UBE2U with UBE2U of mouse, rat and macaca origin. In yellow is highlighted the ubiquitin-conjugating domain (UBC) while with the square is reported the cysteine active residue of the catalytic site. The isoform A of the human protein is highlighted in green.

Additionally, UBE2U appeared to interact with itself as well, giving self-dimerization. These interactions were found using both, UBE2U fused to BD with FA cDNAs fused to AD and vice versa (using UBE2U fused to AD with FA cDNAs fused to BD). Particularly worth of note is the interaction of UBE2U with FANCL, since FANCL is considered the putative ubiquitin-ligase E3 of the FA-BRCA pathway *in vivo* (Meetei, de Winter et al. 2003). We also detected interactions between BD-UBE2U and AD-FANCD2, as well as between AD-UBE2U and BD-FANCE. Interestingly, we found AD-FANCL interacting with BD-

FANCD2$_1$ and BD-FANCD2$_3$, as well as with BD-FANCE. Finally, BD-FANCL interacted with AD-FANCF and with itself (Table 7 and Fig.9).

pGBKT7 (BD)	pGADT7 (AD)								
	FANCD2	FANCD2$_1$	FANCD2$_2$	FANCD2$_3$	FANCC	FANCE	FANCF	FANCL	UBE2U
FANCD2	nd							-	-
FANCD2$_1$								+	-
FANCD2$_2$								-	-
FANCD2$_3$								+	+
FANCC								-	+
FANCE								+	+
FANCF								-	+
FANCL	-	-	-	-	-	-	+	+	+
UBE2U	+	-	-	+	+	-	+	+	+

Table 7: Direct interaction matings of UBE2U and FA proteins. Direct interaction mating assays carried out between UBE2U and FA proteins. BD: proteins fused with the GAL4 binding domain; AD: proteins fused with GAL4 activation domain. Interactions (+) found in both two-hybrid protein combinations are in green, whereas those found in only one combination are in yellow. (-): no interaction; nd: interactions not checked. FANCD2$_4$ and FANCD2$_5$ were not used since as preys (with AD) they auto-activated reporter genes

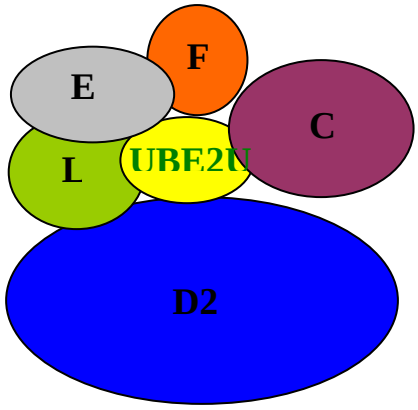

Figure 9: Interaction between UBE2U and FA proteins in yeast. Schematic representation of interactions found between UBE2U and FA proteins investigated in yeast direct interaction mating assays

3.2.4: Northern blot and alternative splicing of UBE2U.

As a first step toward the characterization of UBE2U, Northern blot analysis of human tissues was performed using a specific probe (Fig. 10). Consistent with bioinformatics data, a highly expressed transcript of 1.6 kb, as well as transcripts of major sizes, were detected in human testis. The gene was not expressed in thymus, spleen, small intestine, colon, prostate, uterus, and peripheral blood leukocytes. In order to confirm the presence of the three 1.6, 3.3, and 3.8 kb transcripts, we amplified the putative cDNAs using specific oligonucleotides (1F/10R, 1F/8aR and 1F/8bR). No product was obtained from different cell lines, suggesting that the gene is not widely expressed (data not shown).

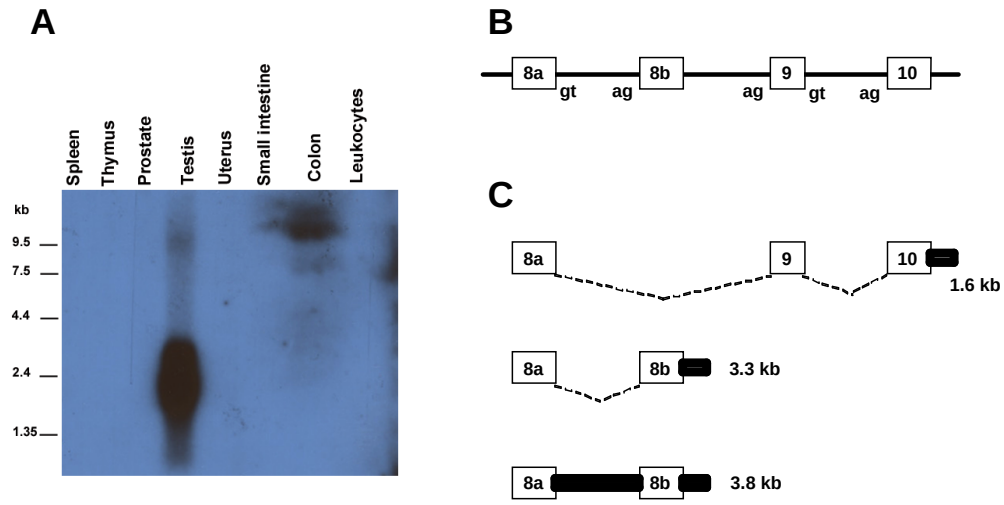

Figure 10: Transcripts analysis for *UBE2U*. A: Northern blot analysis of *UBE2U* on human different tissues. **B**: Schematic representation of genomic structure of human *UBE2U* from exon 8a to exon 10. **C**: Alternative splicing of human *UBE2U* leading to three different transcript.

3.2.5: Ube2u is expressed in testis.

To understand at which developmental stages and in which tissues the gene could be active, expression studies by RT-PCR and *in situ* hybridization (ISH) were carried out in mouse. RT-PCR analysis revealed that the expression of this gene was restricted to testis, as reported in figure 11. During mouse embryonic development, in fact, the *Ube2u* gene is not expressed in eyes (E14.5), mesencephalon (E14.5), telencephalon (E14.5, E16.5), kidney (E16.5), heart (E14.5) and lung (E14.5, E16.5). In adult mouse, the gene is not transcribed in

peripheral blood, bone marrow, fibroblasts and cochlea. Moreover, no expression was found in mouse melanoma cell lines B16 or in embryonic stem (ES) cells.

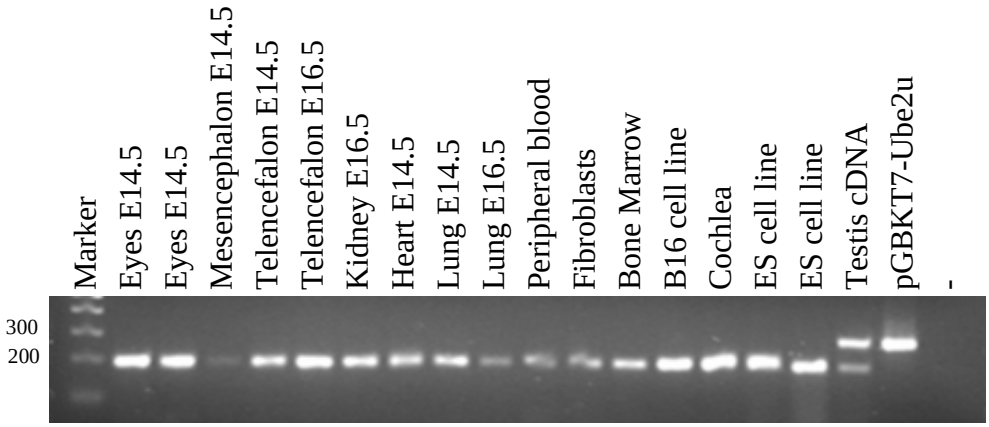

Figure 11: *Ube2u* is expressed in testis. RT-PCR analysis of *Ube2u* in cell lines, embryonic and adult tissues from mouse. Actin (200bp) has been used as internal control of RT-PCR. (-), –RT. Marker: 2-Log DNA Ladder.

3.2.6: Ube2u is expressed in meiotic cells and its expression overlaps with Fancd2 expression.

In order to further characterize the expression of the *Ube2u* gene in mouse testis, we studied adult sections using RNA *in situ* hybridization. The expression of the *Ube2u* mRNA varied with the seminiferous cycle in different developmental stages of mouse testis. Whereas no hybridization could be seen in seminiferous tubules in 2 weeks mouse testis, strong signals were detected in seminiferous tubules of 3, 5 and 12 weeks mice (Fig.12). The presence of *Ube2u* mRNA, as indicated by brown staining, was limited to central layers of epithelium cross section of seminiferous tubule, where spermatocytes (meiotic cells) and round spermatides are located. A weak or no signal was instead found in correspondence of spermatogonia (outer layer) or elongated spermatides (inner, lumen layer). Interestingly, even the *Fancd2* gene resulted expressed mainly in round and elongated spermatides (Fig. 12), supporting thus the hypothesis of interaction between the two products

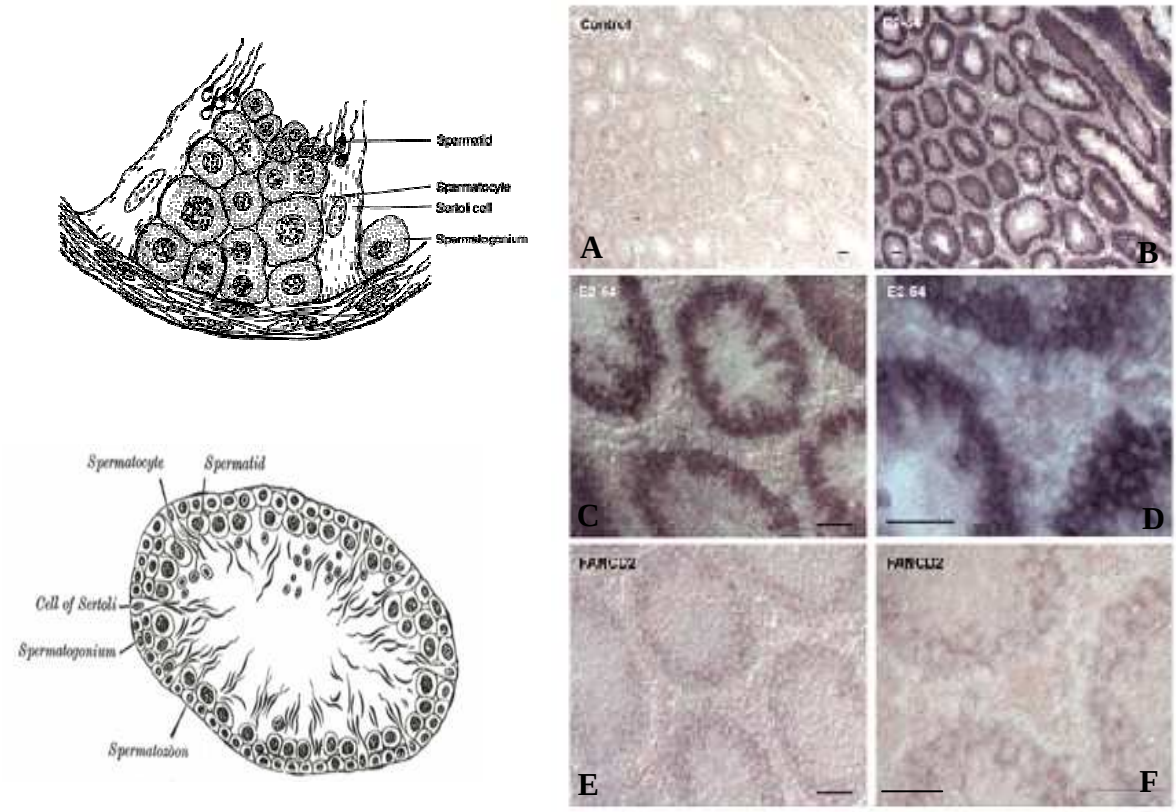

Figure 12: ***RNA in situ hybridization analysis of Ube2u and Fancd2*****.** RNA in situ hybridization on adult testis (A,B,C,D,E and F)sections. Epithelium of seminiferous tubules shows various stages of male germ cells development: from the outer layers to the inner layers we progressively find more differentiated and mature cells (from spermatogonia to spermatides). Antisense probes of *Ube2u* detected transcript in correspondence of spermatocytes and spermatides (B, as well as C and D at higher magnifications). Antisense probes of *Fancd2* (E and F higher magnification) detected transcripts in spermatides but at low levels also in spermatocytes, showing with Ube2u. Sense probes used as controls do not show any hybridization signals (A).

In order to confirm this expression pattern, we decided to carry out RT-PCR analysis on mRNA samples of cells at different stages during the spermatogenesis, from spermatogonia to mature spermatozoa. As shown in figure13, a PCR product of the expected size was found in whole testis extracts and interstitial cells. It was also detected in spermatocytes, secondary spermatocytes and spermatides but not in spermatogonia or in spermatozoa.

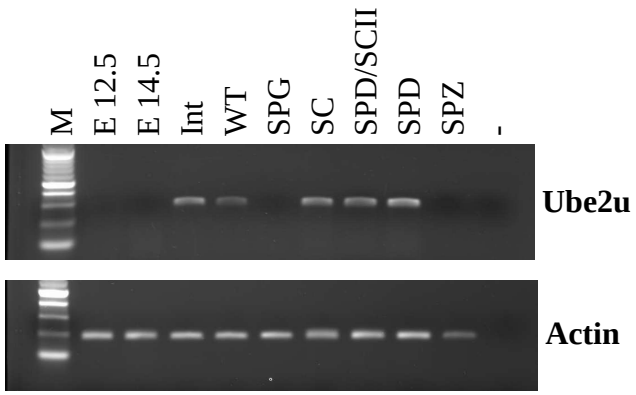

Figure 13: RT-PCR analysis on cells during spermatogenesis. A RT-PCR product of 318 bp indicated the Ube2u gene. RT-PCR of actin (200bp) has been used as loading control. M, 2-Log DNA ladder, E12.5 and E14.5; Int, interstitial cells; WT, whole testis, SPG, spermatogonia, SC, spermatocytes, SPD, spermatides, SCII, secondary spermatocytes, SPZ, spermatozoa; (-), -RT.

3.2.7: Western blotting analysis of Ube2u in mouse testis.

Western blotting analysis was carried out on protein samples from cells of mouse spermatogenesis. These experiments exploited two antibodies (YFE05-1 and YFE05-2) directed against the murine Ube2u protein produced in our laboratory (M&M next section) (Fig. 14). Samples analyzed included whole testis and interstitial cells protein extracts of 4-12 months old mice, together with spermatogones, spermatocytes spermatides and spermatozoa. In agreement with RT-PCR data, the protein was identified in whole testis and in interstitial cells. Among cells of the different stages of spermatogenesis the protein was detected in spermatocytes, spermatides and spermatozoa but not in spermatogonia. Accordingly to the amount of ERK1/2 used as samples loading control, the highest expression level was detected in spermatocytes. These data are consistent with those obtained by RT-PCR except for the presence of a low amount of protein in spermatozoa, where the *Ube2u* transcript was not detected.

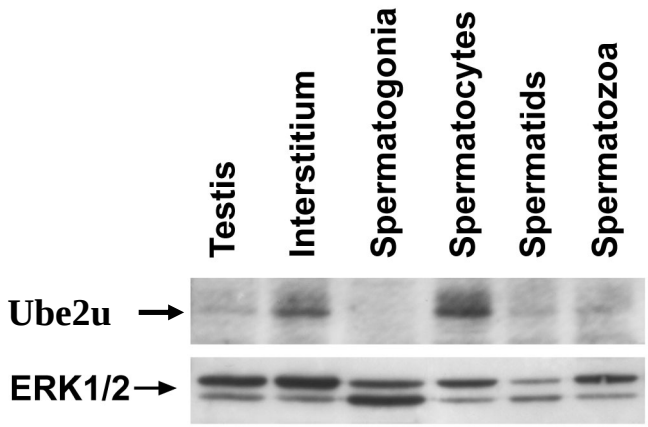

Figure 14: Western blotting analysis on mouse protein extracts of cells during spermatogenesis. YFE05-1 has been used as specific anti-Ube2u antibody while anti-ERK1/2 has been used to detect ERK1 and ERK2 as loading control. Protein is detected at high levels in interstitial cells and in spermatocytes.

3.2.8: No UBE2U expression in transient transfections.

In order to further characterize the novel putative gene at the protein level and to validate the interaction between human UBE2U and FANCD2 found in yeast, we performed experiments of protein cellular localization and co-immunoprecipitation in mammalian cells. However, due to the limitations of expression pattern of UBE2U, no cellular models were available for these studies (we did not detect expression in any cell line investigated). Additionally, a specific antibody directed against the human protein was not available; therefore analyses could not be performed at endogenous level. For these reasons, we carried out transient transfections in mammalian cell lines. The human cDNA of UBE2U was subcloned into expression vectors pcDNA3.1 fused in frame with either N-ter FLAG- or myc- or HA- tags and transiently transfected in HeLa, Cos-7 and HEK293 cell lines. Even using different protocols to increase efficiency of transfection, we were not able to detect the protein at any level in any of the cell lines used by both immunofluorescence and western blotting techniques. In order to verify whether the proteasome mediated the degradation of UBE2U, we treated HEK293 cells with *N*-acetyl-L-leucinyl-L-leucinal-L-norleucinal (LLnL) and again we detected no expressed protein, as showed in fig. 15A, whereas a signal is detected loading HA-UBE2U transcribed and translated (tnt) *in vitro*. Since butyrate is also a molecule able to induce the transcription, we treated cells with this compound alone and in

combination with LLnL as well. Again, no fused HA protein was detected (Fig. 15B). Next, in order to exploit the transcription and translation (tnt) *in vitro* to validate the interaction between UBE2U and FANCD2, we tested whether antibodies against HA were able to immunoprecipitate HA-UBE2U produced in rabbit reticulocytes (Fig. 15C). Although the assays were repeated several times, no immunoprecipates were detected, suggesting that the antibody anti-HA was not able to immunoprecipitate the protein from rabbit retyculocites. Moreover, we could not exclude that, despite using protease inhibitors, the protein was not stable and thus degraded *in vitro*.

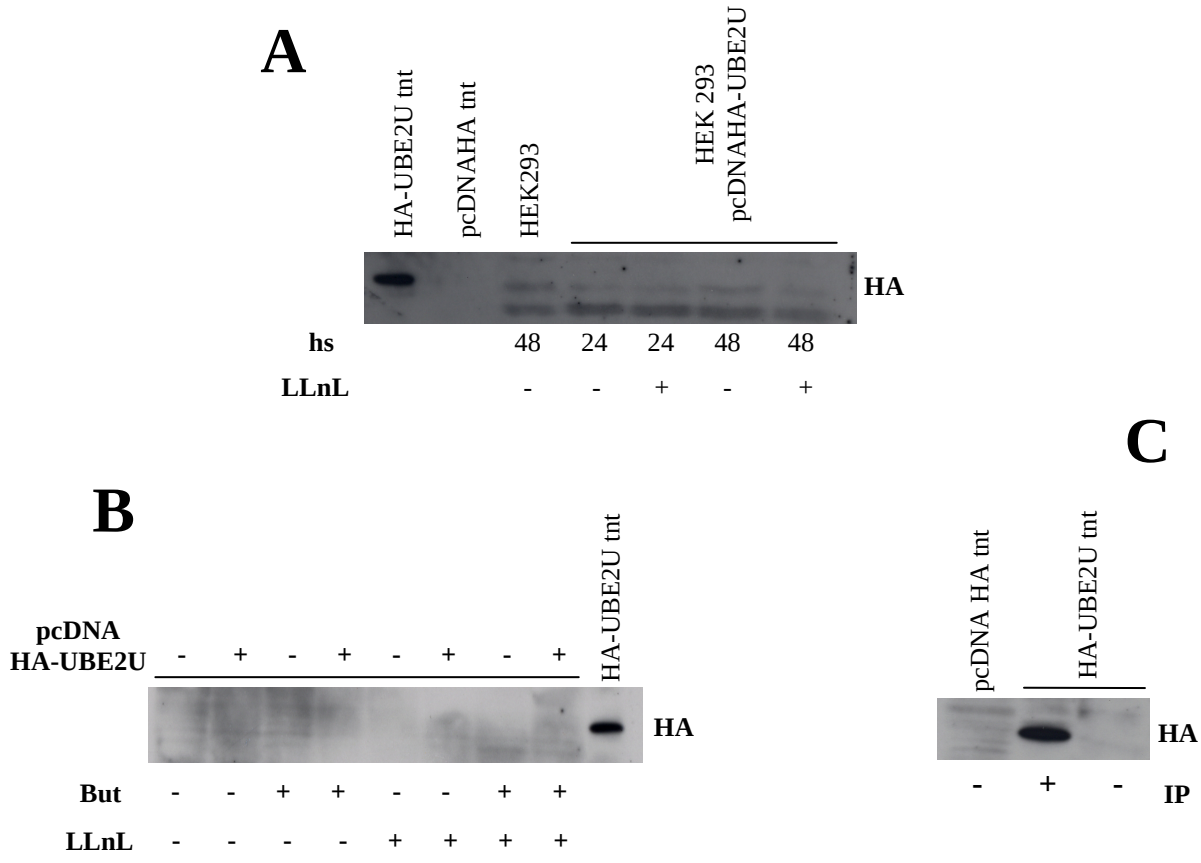

Figure15 Western blotting analyses of HEK293 cells transiently transfected with pcDNA3.1HA-UBE2U and treated with LLnL and/or butyrate. **A**: HEK293 cells were transfected with pcDNA3.1HA-UBE2U and treated with LLnL 50 μM for 10 hours. Cells were harvested 24 or 48 hours (hs) after transfection.. HA-UBE2U tnt has been obtained by transcription and translation (tnt) of pcDNA3.1HA-UBE2U in rabbit reticulocytes (1/10 of reaction loaded). **B**: HEK293 cells transiently transfected with pcDNA3.1HA-UBE2U and treated with LLnL 50 uM (for 10 hours) and/or butyrate 2mM (for 24 hours). Cells were harvested 48 hours after transfection. **C**: Immunoprecipitation of HA-UBE2U tnt using anti-HA antibodies. No immunoprecipitates are detected. Tnt, transcripted and translated.

3.2.9: No interaction between Ueb2u and Fancd2 in mouse testis protein extracts.

In order to investigate whether UEB2U is able to bind FANCD2, we took advantage of the high level of both proteins in mouse testis and tried to confirm their interaction by using mouse testis extracts. Thus, we produced two rabbit polyclonal antibodies (two different sera) directed against the same amminoterminal portion of the murine protein. Immune antisera were purified with protein A beads and tested against whole testis protein extracts. In experiments of western blotting, an expected band of 42 kDa was specifically detected (theoretical molecular weight of UBE2U = 41.9 kDa) by the two antibodies used (YFE05-1 and YFE05-2) as shown in fig. 16. In filters incubated with corresponding pre-immune antisera the bands were not detected, supporting the specificity of the antibodies generated. Thanks to the availability of an antibobody against mouse Fancd2 (kindly donated by dr Alan D'Andrea), we set up experiments of immunoprecipitation using antibodies against both Ube2u and Fancd2. As shown in fig. 17 Fancd2 was clearly immunoprecipitated by its relative antisera. However, no specific band of Ube2u was detected in the immuniprecipitates at least under our experimental conditions. On the other hand, the immunoprecipitation experiments performed with our anti-Ube2U antibodies suggest that the antisera are not able to specifically precipitate the antigen (data not shown). Future immunochemical experiments and the development of additional reagents are needed to further investigate the putative interaction between the two proteins

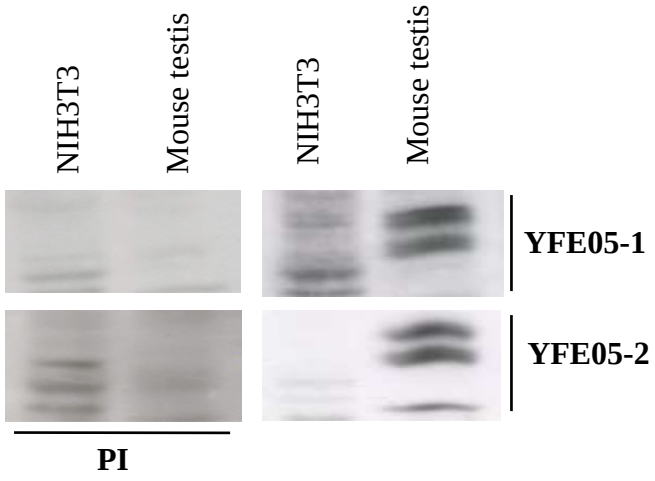

Figure 16: Western blot using the two specific antibodies YFE05-1 and YFE05-2, generated against murine Ube2u. Protein extracts from mouse testis were probed with immune and pre-immune (PI) antisera. Protein extracts from NIH3T3 cells were used as negative control.

.

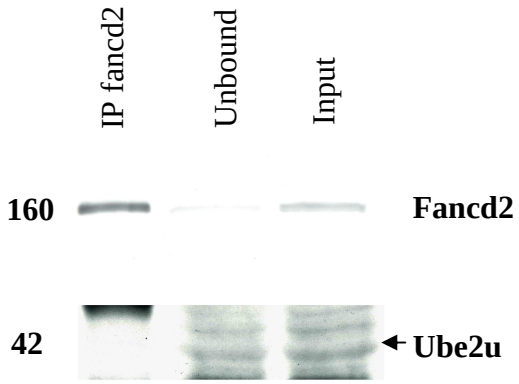

Figure 17: Coimmunoprecipitation of Fancd2 and Ube2u. Protein extracts from mouse testis were immunoprecipitated with anti Fancd2 and immunoblotted for Fancd2 and Ube2u (YFE05-1).

3.3 ISG20L2: a novel putative exonuclease.

3.3.1: Cloning of human and murine clone 4.

The potential interactor of FANCG (clone 4) had a partial open reading frame that resulted to belong to the *ISG20L2* (interferon stimulated exonuclease gene 20-like 2) gene localized on chromosome 1q23.1. Two chromosomal regions on 6p22.3 and within intron 12 of the USP28 gene on chromosome 11q23.2 also share a collinear high homology to the ISG20L2 cDNA, suggesting that two pseudogenes are also present in the genome. The human transcript of 2,0 kb encoded for a predicted protein of 353 amino acids characterized by aputative conserved domain of the EXOIII family (Fig. 18), which includes a variety of exonuclease proteins, such as ribonuclease T and the alpha and epsilon subunits of DNA polymerase III. The murine ortholog gene *Isg20l2* is localized on the syntenic region of chromosome 3, whereas its pseudogene is on chromosome 6. The mouse transcript of 2,0 kb encodes for a protein of 368 amino acids sharing a 72% of residues identity with the human product. Alignment of different species shows that the gene is conserved throughout evolution (Fig. 18).

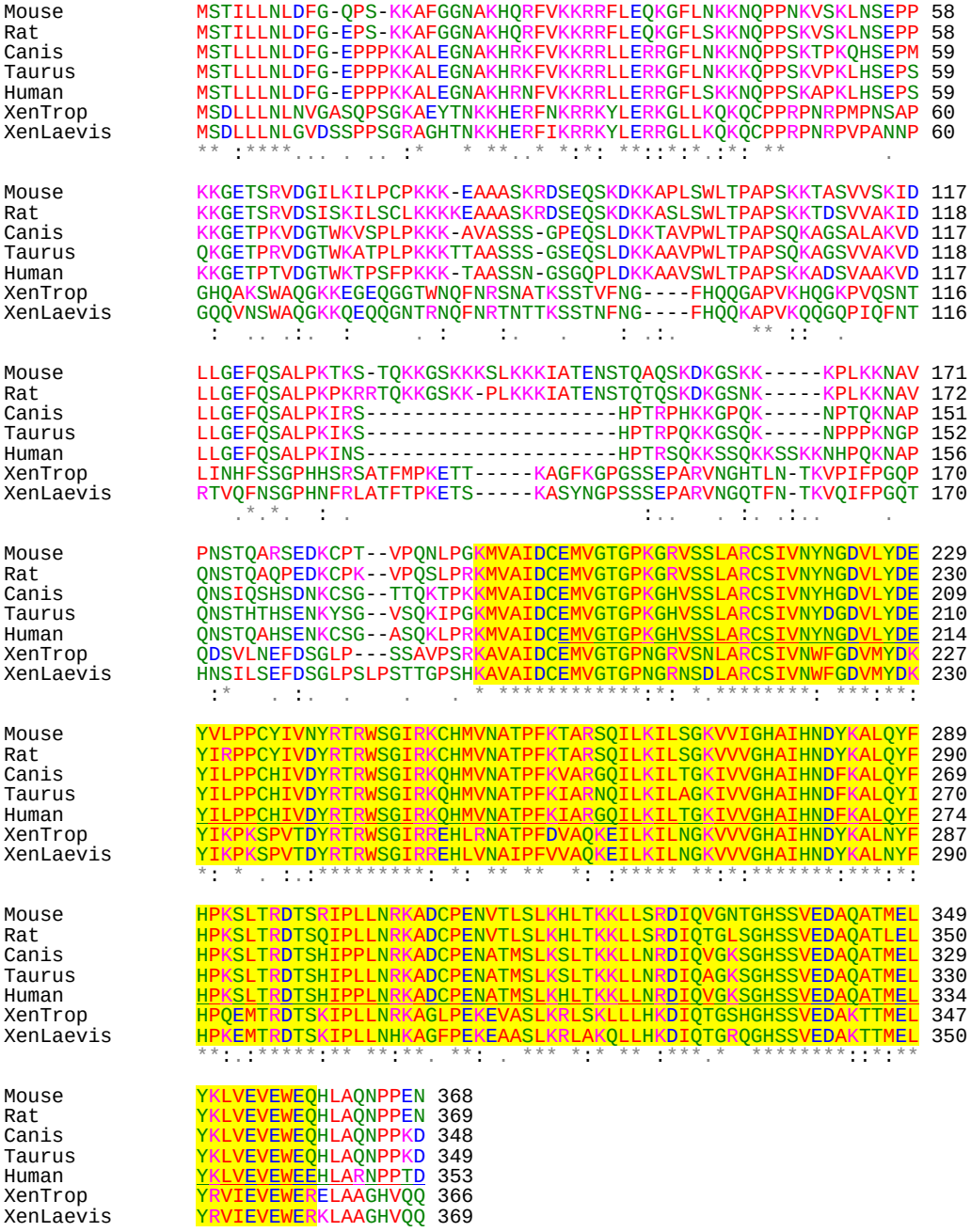

Figure 18: Alignment of ISG20L2 human protein with orthologs of other species. The conserved EXOIII domain is highlighted in yellow.

3.3.2: High expression of ISG20L2 in testis.

The full length mouse transcript was consistent with Northern blot analysis on mouse tissues, showing a single transcript of 2,0 kb expressed in testis but not in heart, brain, spleen, lung, liver, skeletal muscle, and kidney (data not shown). Northern blot analysis carried out on human tissues showed a single transcript of 2,0 kb in testis, while no bands were detected in spleen, thymus, prostate, uterus, small intestine colon and leukocyte (Fig. 19A). RT-PCR performed on human cell lines revealed a band of 1263 bp, proving expression of the gene in lymphoblasts, HeLa, HEK293, CaCo2 and fibroblasts, as showed in fig. 19B. On the basis of northern blot data from mouse tissues, we investigated the expression of *Isg20l2* in testis by RNA *in situ* hybridization and high levels of expression were detected in spermatides (data not shown).

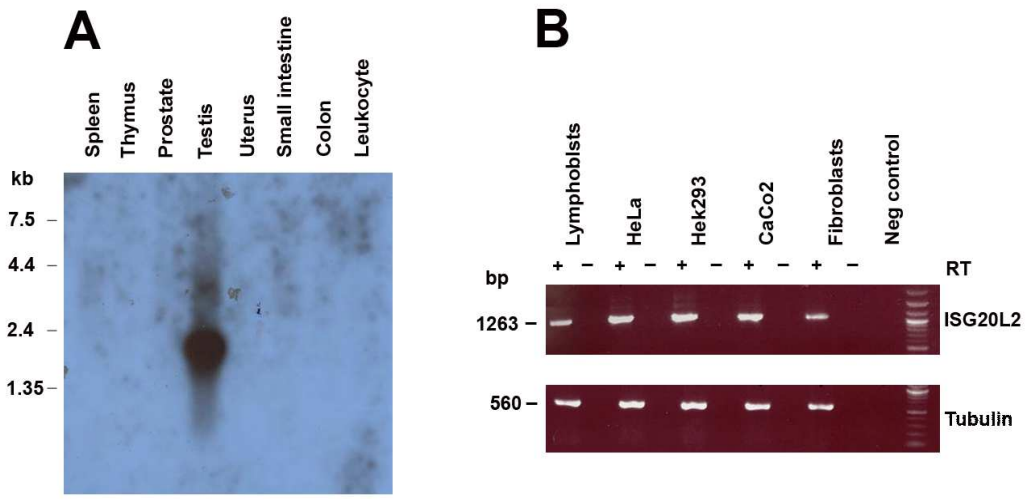

Figure 19: Expression analysis of human *ISG20L2*. A) Northern blot analysis of *ISG20L2* in several human tissues. **B)** RT-PCR analysis of *ISG20L2* in several human cell lines.

3.3.3: Biochemical analysis of ISG20L2.

With the aim of characterizing this novel gene, immunofluorescence, western blotting and co-immunoprecipitation analysis were set up. The coding region of human ISG20L2 was subcloned into pcDNA3.1 in frame with FLAG or HA tags and transiently transfected in HeLa and HEK293 cell lines. The cellular localization of the recombinant protein was

studied by immunofluorescence and western blot analysis of both cytoplasmic and nuclear extracts. The interaction between FANCG and IGS20L2 was investigated by co-immunoprecipitation studies. In addition, we exploited also a specific antibody generated against the human protein ISG20L2 as we report below.

HeLa cells transfected with pcDNA3.1FLAG-ISG20L2 were analysed by immunofluorescence 48 hours after transfection and the signal of the exogenous protein was detected mainly in the nucleolus (Fig. 20).

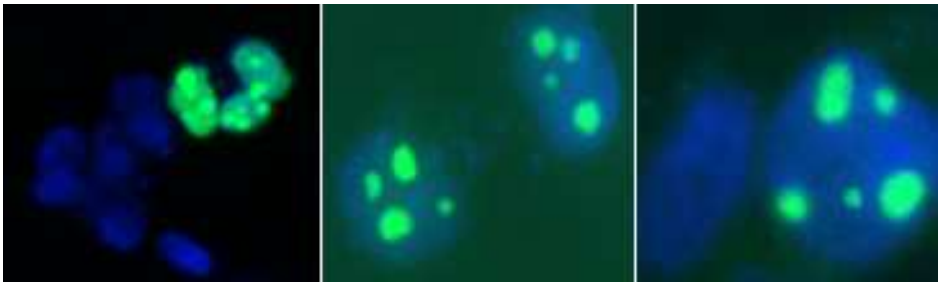

Figure 20: Immunofluorescence analysis of HeLa cells transfected with pcDNA3.1FLAG-ISG20L2. The recombinant protein localizes in nucleoli (darker zones of DAPI staining) in three different imagines chosen as examples from indipendent transfections.

Expression of the ISG20L2 protein was investigated by mainly using antibodies generated against the human protein. We tested 4 different anti-sera produced against the C- and N- termini of the human protein ('PEA05-1' and 'PEA05-2' against N-terminus, 'QEA05-1' and 'QEA05-2' against C-terminus) initially by western blotting analysis of FLAG-ISG20L2. First, rabbit reticulocytes were exploited for the transcription and translation *in vitro* of ISG20L2 subcloned into pcDNA3.1FLAG. As showed in fig. 21(A), QEA05-1 was able to detect a faint band at the expected molecular weight of 43 kDa, while the other anti-sera tested, PEA05-1, PEA05-2 and QEA05-2 did not recognize any specific band (data not shown). Next, transient transfections of HEK 293 cells were carried out with pcDNA3.1FLAG-ISG20L2. Nuclear and cytoplasmic protein extracts were probed using an anti-FLAG and the specific anti-ISG20L2 antibodies. As shown in fig. 21(B) QEA05-1 was able to detect a specific band, while the other anti-sera tested gave not clear outcomes (data not shown). Bands of the expected molecular size were identified in the nuclear as well as in the cytoplasmic fraction by using QEA05-1. Signals were detected in both non-transfected and transiently transfected cells. Then, the signals were compared to those obtained with

anti-FLAG antibody showing a perfect match. As a result, QEA05-1 seemed able to detect the over-expressed, but also the endogenous protein in HEK293 cells and it was chosen for the following experiments.

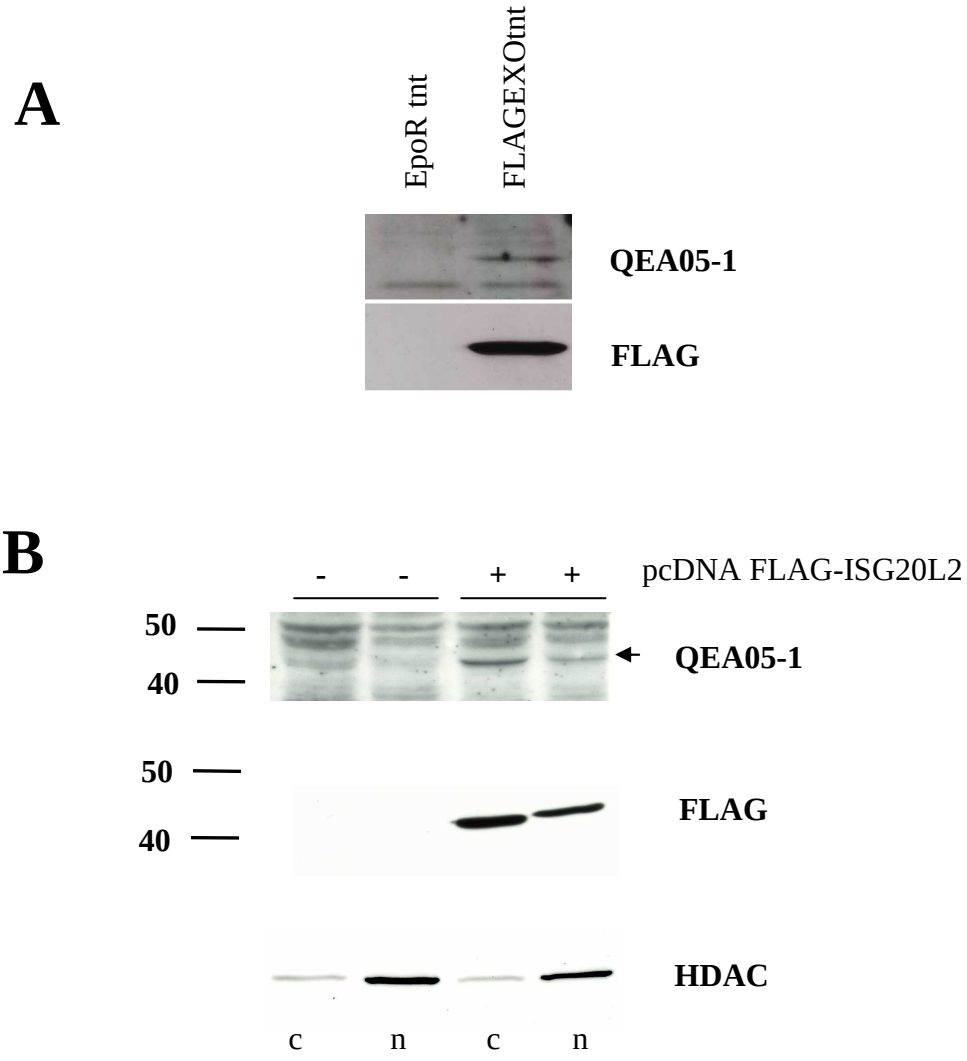

Figure 21: Testing QEA05-1. **A**: Anti-ISG20L2 antibody QEA05-1 was tested on transcripted and translated *in vitro* (tnt) FLAG-ISG20L2. Anti-FLAG was used to check for protein expression, while we used a plasmid containing an unrelated cDNA (pcDNA3.1HA-EpoR) as negative control. **B: Western blot analysis of nuclear and cytoplasmic extracts of HEK293 cells transfected with pcDNA3.1FLAG-ISG20L2.** Cytoplasmic and nuclear extracts of HEK293 cells untransfected and transfected with pcDNA3.1 FLAG-ISG20L2 are shown. Anti-Histone deacetylase (HDAC) antibody has been used as nuclear marker. Additionally, in these experiments, filters hybridized with anti-ISG20L2 (QEA05-1) were re-probed with anti-FLAG antibody and vice-versa (filters hybridized with anti-FLAG antibody were re-probed, after stripping, with anti-ISGL20L2).

To further assess the specificity of this antibody, proteins from nuclei and cytoplasms of HEK293 (untransfected and transiently transfected with pcDNA3.1FLAG-ISG20L2),

Panc-1, HepG2 and EPN cells were separated by electrophoresis and analysed by western blot. QEA05-1 and its pre-immune antiserum were used to hybridize two filters. Specific ISG20L2 bands of the expected molecular weight were detected on the filter probed with antibody QEA05-1 whereas the same bands were absent in the filter hybridized with pre-immune antiserum. The specificity of this antibody was further confirmed by stripping the filters and re-probing them with anti-FLAG antibody, which detected bands of the same size of those obtained with QEA05-1 antibody in transfected cells. The expression study of of the endogenous protein was further extended to nuclear and cytoplasmic fractions from other cell lines. Protein extracts from lymphoblast, EPN (epithelial prostatic, human), HeLa, HepG2 (Human hepatocellular liver carcinoma), HEK293, K562 (erythroleukaemia, human) LAN-5 (neuroblastoma) and Panc1 (human pancreatic carcinoma, epithelial-like) cell lines were analysed. As showed in fig. 22, the ISG20L2 was expressed in both compartments of all the cell lines investigated.

In order to confirm the interactions observed in yeast two-hybrid screening studies, we performed also a series of coimmunoprecipitation experiments. In these investigations, we co-transfected pcDNA3.1FLAG-FANCG and pcDNA3.1HA-ISG20L2 in HEK293 cells. When we immunoprecipited both the proteins by the specific anti-tag antibodies, we were unable to demonstrate in the immunoprecipitate material the presence of the other protein, thus suggesting the absence of interaction between ISG20L2 and FANCG (data not reported). However, due to the intrinsic constraints of the approach (i.e. the binding between the two proteins might mask the epitopes recognized by the antibodies), we believe that it is important to repeat the experiments employing additional antisera directed towards further protein epitopes.

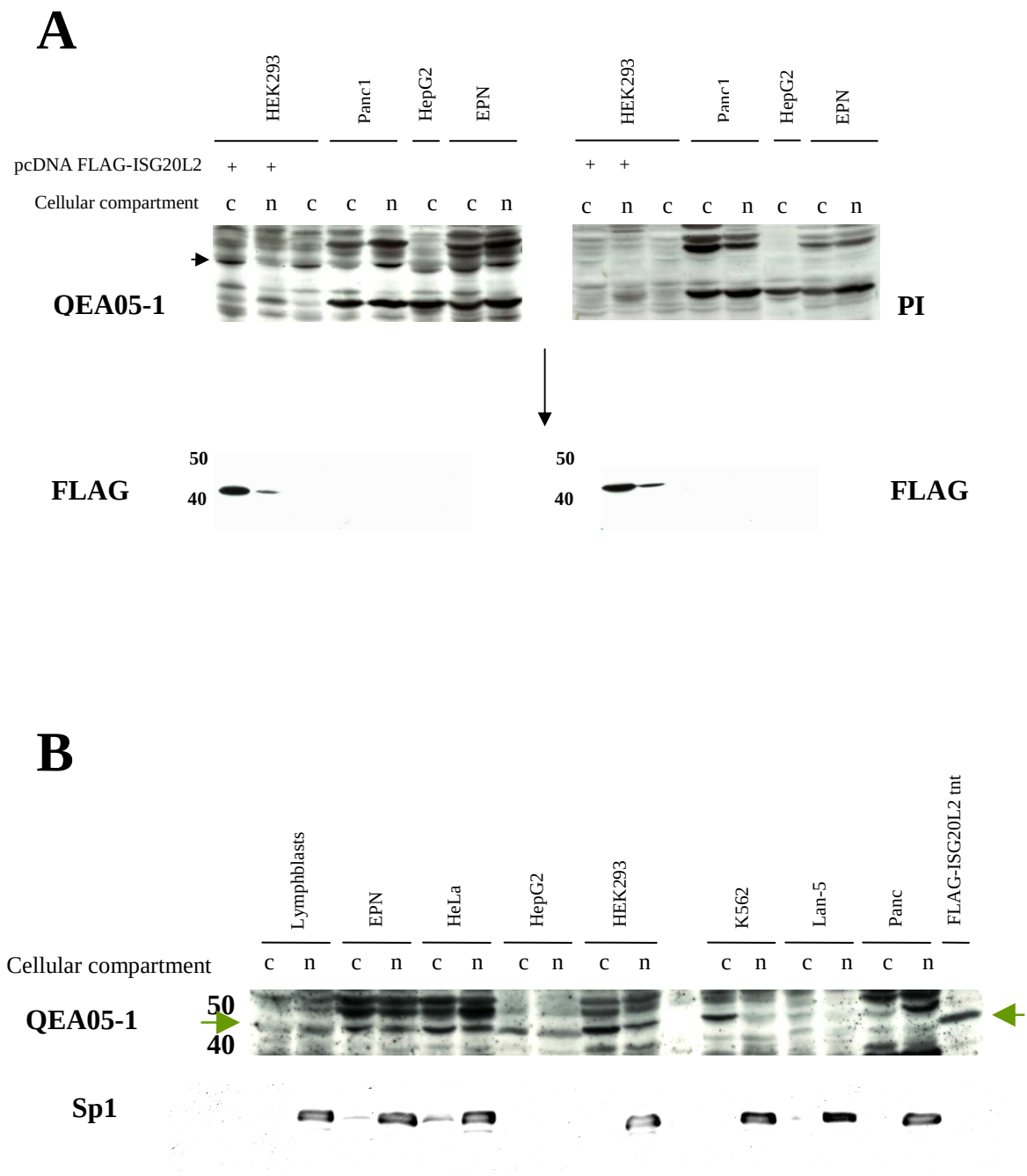

Figure 22 A: **Western blot analysis of nuclear (n) and cytoplasmic (c) compartment of several cell lines**. Filters were first separately probed with anti-ISG20L2 specific antibody (QEAO5-1) and its pre-immune antiserum (QEA05-1 PI). Next, filters were reprobed with anti-FLAG antibody. Only HEK293 cells transiently transfected with pcDNA3.1FLAG-ISG20L2 showed the signals. **B: Western blot analysis of nuclear (n) and cytoplasmic (c) compartment of several human cell lines**. Specific bands of ISG20L2 are present in both the compartments analyzed in all the cell lines investigated. Sp1 has been used as nuclear marker

Chapter 4
DISCUSSION

4.1 Screening a cDNA library of human testis.

The discovery and characterization of the thirteen FA genes have revealed the connection between FA, the DNA-damage response and cancer. Especially in the last decade, studies in the FA research field disclosed a crowded network of interacting proteins, but hinted also that other players of this network are still missing. In addition, the pathogenetic mechanism underlying the disease is still unclear. As described in the introduction of this PhD thesis work, understanding the molecular biology of FA is nowadays crucial not only for haematologists and oncologists, who manage patients with FA or with bone marrow failure syndromes or with cancers, but also for scientists interested in DNA-repair, ubiquitin biology, fertility, drugs sensitivity and resistance.

The principal aim of the present work was to identify novel interactors of FA proteins so as to contribute to unravelling the FA-BRCA pathway. Starting point of this project was the large-scale approach of the yeast two-hybrid system, a powerful methodology to detect protein-protein interactions. When we started our studies, only nine of the thirteen FA genes were cloned. We chose as baits the most characterized genes not only as full length cDNAs, but also as partially overlapping cDNA fragments mainly for two reasons. First, one limitation of protein expression in yeast two-hybrid systems is that the yeast translational machinery is not always able to biosynthesize large proteins. Second, in case of interactions, we would also have had insights into functional domains which were and still are largely unknown for FA proteins. For these reasons, we arbitrarily chose to clone cDNA fragments of average size of approximately 1000 bp. As reported, we chose to screen a human testis library because the only phenotype shared between FA patients and knocked out mice was hypogonadism and reduced fertility (Chen, Tomkins et al. 1996; Whitney, Royle et al. 1996; Cheng, van de Vrugt et al. 2000; Yang, Kuang et al. 2001; Koomen, Cheng et al. 2002; Houghtaling, Timmers et al. 2003; Wong, Alon et al. 2003), suggesting a crucial role of the FA-BRCA pathway in gametogenesis. Moreover, consistently with these data, FA genes are expressed at high levels in testis (Chen, Tomkins et al. 1996; Lo Ten Foe JR 1996; Whitney, Royle et al. 1996). Finally, the FA-BRCA pathway was completely uncharacterized in that tissue and the hypothesis of a role in crossing-over events was particularly suggestive. Among 23 bait-constructs generated, a few of them were excluded because of their auto-activation of reporter genes or of their inefficiency in synthesizing proteins in yeast. Then we selected 8 baits. Five including all fragments of FANCD2, because the monoubiquitination of this protein is a crucial step in the FA-BRCA pathway (in review (Wang 2007)) and

because it directly binds BRCA1 (Taniguchi, Garcia-Higuera et al. 2002) connecting FA-BRCA pathways with DNA-repair. Two baits were selected for two portions of FANCA because binding to several proteins of the FA-core complex and directly interacting with BRCA1 as well (Folias, Matkovic et al. 2002), and one bait for the only suitable fragment of FANCG, being this protein a scaffold for the FA-core complex, not dispensable for its assembly (Garcia-Higuera, Kuang et al. 1999; Medhurst, Huber et al. 2001). After eight independent library screenings, 167 clones were obtained and further investigated to confirm interaction in a tiered approach. Disappointedly, 107 of them (64%) resulted to be frame shifted sequences and 5' and 3' untranslated regions (UTR) or even genomic sequences and ALU sequences, representing artefact protein products. Of the 60 clones selected, 44 interacted with the N-terminal part of FANCG whereas 12 interacted with $FANCD2_3$ and 1 with $FANCD2_4$. Just one clone was found to interact with $FANCA_3$ while another two were bound to $FANCA_5$. The relatively high number of clones rescued using $FANCG_2$ as bait could be explained considering the four tetratricopeptide repeat (TPR) domains present in the partial protein (of the seven contained in the full length). The tetratricopeptide repeat (TPR) is a degenerate 34 amino acid sequence identified in a wide variety of proteins, present in tandem arrays of 3–16 motifs. It forms scaffolds to mediate protein–protein interactions and the assembly of multiprotein complexes (Das, Cohen et al. 1998). In $FANCG_2$ these repeats probably facilitated the interaction with numerous preys. Notably, all clones rescued from the screenings were found only once and none of them bound to any known interaction. Although these observations should have alerted us on the limitations of our analyses, we justified the results taking into consideration three aspects. First, using partial length cDNAs as baits we could have destroyed potential binding domains, which had not been well defined even for known FA proteins. Second, we were dealing with a tissue, as the testis, where the FA-BRCA pathway was never investigated. Finally, the FA-core complex assembly is dynamic and it could involve several determinant missing players. Further investigations were carried out to verify interactions with yeast direct interaction matings, leading us to assess that 50 clones resulted false positives. The high number of false positives detected was probably due to different cellular environments present in strains used for the large-scale screening (AH109 mated with Y187, diploid) and the direct interaction mating (AH109 alone), as well as to different culture conditions between the two procedures. The remaining 10 clones were tested after switching vectors, confirming the interactions of clone 4 with $FANCG_2$, as well as clones 54 and 62 with $FANCD2_3$. Although again this outcome is of difficult interpretation, we can consider that in this system, baits are hybrid proteins fused

with the AD or the BD domains of GAL4 (accordingly to the vector in which they are cloned in). Large-scale screening was performed using BD-baits interacting with an AD-prey library. The interactions found were afterwards examined for the opposite combination: AD-bait with BD-prey. Therefore, accordingly to the domain (AD or BD) a protein is fused with, the conformation can vary, leading to formation or disruption of different functional interacting domains.

The last tiered step towards the selection of potential interactors of the FA proteins allowed us to thus isolate 3 novel candidates that were first characterized using bioinformatics tools. Clone 4 encoded for ISG20L2, a protein with a predicted exonuclease domain 'EXOIII', clone 54 encoded for a predicted ubiquitin-conjugating enzyme E2 (UBE2U), while clone 62 corresponded to CRISP2 (cysteine-rich secretory protein 2). These findings provided the basis for the second part of this PhD project, aimed at further characterizing two novel interactors, UBE2U and ISG20L2, potentially associated with the FA-BRCA pathway. The first, because the ubiquitin-conjugating enzyme E2 of the FA-BRCA pathway was unknown and the second, because an exonuclease could have been implicated in DNA-damage repair. CRISP2, a testis specific protein of the sperm acrosome (Busso, Cohen et al. 2005), was not investigated because unlikely associated with FA.

4.2 Two-hybrid system in FA: data from literature.

An extensive literature search for applications of yeast two-hybrid system in FA allowed us to find that, until 2002, four of sixteen proteins described to interact with FA proteins had been identified using this technique. Through several library screenings, GRP94 and FAZF were found to bind FANCC (Hoshino, Wang et al. 1998; Hoatlin, Zhi et al. 1999) while SNX5 and CYP2E1 where found to interact with FANCA and FANCG, respectively (Otsuki, Kajigaya et al. 1999; Futaki, Igarashi et al. 2002). However, after these first reports, the interactions have not been further confirmed or investigated. The largest-scale screening was next reported by Reuter and colleagues (2003), who described 69 novel potential FANCA, FANCC and FANCG interacting proteins. These proteins were collected from experiments performed in four different laboratories, using five libraries derived from various tissues. All candidates were tested in the yeast two-hybrid system, and many were additionally evaluated by *in vitro* and *in vivo* co-immunoprecipitation and co-localization

experiments. Their studies provided new evidences for a connection of the FA-BRCA pathway with a range of cellular functions linked to haematopoiesis, cancer, and detoxification (Reuter, Medhurst et al. 2003). Authors indicated new possible pathways to look at, together with new proteins putatively involved in the FA-BRCA pathway, but again there are no new relevant updates or reports regarding these candidates. Consistently with also our outcomes, the yeast two-hybrid system therefore appears to be not the most suitable strategy to isolate new FA proteins interactors from large-scale library screenings. More successful was instead the application of direct interaction mating assays to detect interactions between candidates or to map binding domains. In 2002, Folias and colleagues reported the interaction between the N-terminal portion of FANCA and BRCA1, hypothesizing, for the first time in FA research, a link to DNA-repair process (Folias, Matkovic et al. 2002). This connection was later confirmed also in the work reported by Zhu and Dutta (Zhu and Dutta 2006), in which they showed that rereplication activated the FA-BRCA pathway, both for the activation of a G_2/M checkpoint and for DNA-repair processes, involving FANCA, BRCA1 and ATR. Also in 2003, Gordon and Buchwald investigated the architecture of the FA proteins complex by studying direct interaction mating in yeast two-hybrid system (Gordon and Buchwald 2003). They identified binding domains between FANCA/FANCG, FANCC/FANCE and FANCF/FANCG. Additionally, they detected a direct interaction between FANCE and FANCD2, providing a direct link between the FA protein core complex and its downstream target. Again in 2003, Hussain and colleagues tested the interactions between BRCA2 (FANCD1) or its effector RAD51 and the proteins FANCA, FANCC and FANCG. They found that FANCG co-localized in nuclear foci with both BRCA2 (FANCD1) and RAD51 after DNA damage, binding to two separate sites flanking the BRC repeats of BRCA2 (FANCD1) (Hussain, Witt et al. 2003). Later, the same group showed that FANCD2 bound a highly conserved C-terminal site of BRCA2 that also bound FANCG/XRCC9 (Hussain, Wilson et al. 2004). Since 2004, there have been no new significant data regarding FA proteins studied with the yeast two-hybrid system, until the last year, when a novel interactor of FANCD2, Tip60, was identified screening a human foetal library from brain (Hejna, Holtorf et al. 2008). Tip60 is a histone acetyltransferase, member of a conserved chromatin remodelling complex (Doyon and Cote 2004; Doyon, Selleck et al. 2004). It is able to acetylate not only histones, but also several target proteins involved in DNA-repair or checkpoint response to DNA-damage, including p53 (Kusch, Florens et al. 2004) and ATM (Sun, Jiang et al. 2005). Tip 60 also co-localizes with γH2AX in DNA damage-induced foci (Sun, Jiang et al. 2005) and acetylates H2AX thereby regulating its

ubiquitination by UBC13 during double strand breakage repair (Ikura, Tashiro et al. 2007). Again in 2008, Tremblay and colleagues reported the use of the yeast direct two-hybrid assay to check for connections between FA and HSC self-renewal pathways (Tremblay et al 2008). HES1 was found to be a novel interactor of the FA-core complex, and, analogously to other subunits of the complex, it is required for cellular resistance to MMC, FANCD2 monoubiquitination and proper stability of several FA-core complex components (Tremblay, Huang et al. 2008). HES1 is a downstream effector of Notch1, involved in increasing HSC self-renewal, reducing HSC cycling and preserving the long-term reconstitution ability of primitive haematopoietic cells (Varnum-Finney, Xu et al. 2000; Stier, Cheng et al. 2002; Kunisato, Chiba et al. 2003; Yu, Alder et al. 2006).

As a result, after almost ten years of work done by several groups, the search for new interactors of the FA proteins by screening libraries with the yeast two-hybrid approach has not provided us with many functional insights, in spite of all the efforts carried out to achieve this aim. This is likely due to different aspects related to both the yeast model and partly the characteristics of the FA proteins. In fact, proteins produced in the yeast two-hybrid system are fused with BD or AD and could not localize in the yeast nucleus, or could not fold properly, or could not be functional when expressed as a fusion protein, or could be toxic to the host, or, sometimes, the appropriate post-translational modification could not take place, or the potential interacting protein could not be sufficiently represented in the library. On the other hand, aside from FANCJ and FANCM, there are no orthologues of FA proteins in yeast and the lacking of a complete FA machinery, together with the dynamic assembly of the core complex, could influence outcomes from this system. At this purpose, the most relevant biochemical discoveries in FA have been achieved investigating direct interactions between candidate proteins, as previously enlisted, or using immunochemical analysis coupled with mass spectrometry (Meetei, de Winter et al. 2003; Meetei, Sechi et al. 2003), as described in the introduction section of this thesis.

Nevertheless, since the three clones (4, 54 and 62) we identified were confirmed by direct interaction mating in both combinations of the bait/prey constructs, they have been regarded as good candidates for further studies.

4.3 Characterization of UBE2U as a potential new interactor of FANCD2.

UBE2U was considered the most promising interactor found, because it bound to FANCD2 and showed high homology to the members of the ubiquitin-conjugating enzyme family E2. FANCD2, together with FANCI, is the monoubiquitinated protein of the FA-BRCA pathway. Ubiquitination is a post-translational modification that consists in a covalent binding of a molecule of ubiquitin to the ε-amine of lysine residues of target proteins. It requires a series of ATP-dependent sequential steps, catalyzed by E1 (ubiquitin-activating), E2 (ubiquitin-conjugating) and E3 (ubiquitin-ligating) enzymes (reviewed here (Hochstrasser 2006)). Notably, while the putative E3 ligase *in vivo* of the FA-BRCA pathway was known (FANCL), the ubiquitin-conjugating enzyme E2 was still unknown when we isolated this clone. The identifying characteristic of an E2 is a 14-16 kDa core (UBC domain) that is 35% conserved among family members. This portion of the enzyme contains an active site cysteine and substitution of this residue abolishes E2 activity (reviewed here (Passmore and Barford 2004)). E2 proteins have various individual and redundant roles within the cell and have important functions in target degradation (proteasomal pathway), and target modification for regulatory purposes (non-proteasomal pathways) such as in FA-BRCA pathway. There are several different E2 enzymes (over 30 in humans), which are broadly grouped into four classes: class I consists of only the UBC domain, whereas classes II, III, and IV have a C-terminal, N-terminal and both N- and C-terminals extensions, respectively, in addition to the UBC domain. These extensions appear to be important for some subfamily function, including E2 localization and protein-protein interaction (reviewed here (Glickman and Ciechanover 2002)).

Human UBE2U belongs to the class II and has the catalytic cysteine (Cys89) while mouse isoform does not. Murine Ube2u could represent an E2 variant (UEV), that lacking the catalytic cysteine, forms heterodimers with E2s containing functional active sites (in review here (Hochstrasser 2006)). As a consequence of alternative splicing, two isoforms are expressed in human: isoform A, (227 aa residues long) and isoform B (321 aa residues long). In mouse, instead, there is only the longest isoform of the transcript of *Ube2u*. Isoform A results more conserved among orthologues than the carboxyterminal extension characterizing isoform B, suggesting that probably isoform A is the functionally active one in human. In particular, within this extension of the mouse Ube2u protein, there is a poly-Glu

stretch (residues 268 to 297) interrupted by a Lys at residue 291. The functional significance of this motif is unknown but there is no conservation among the species for this zone of the protein. Northern blotting analysis on human and mouse tissues detected a high level of expression only in testis. This expression pattern was confirmed by RT-PCR on different mouse embryonic and adult tissues and cell lines. Moreover, RNA *in situ* hybridization revealed that spermatocytes and spermatides were the cells with the highest levels of expression and this pattern was confirmed during spermatogenesis at the RNA and protein levels. Spermatogenesis is a process in which spermatogonia differentiate into mature spermatozoa and consists of different phases, including the two meiotic divisions that take place in primary and secondary spermatocytes. *Ube2u* resulted expressed during these phases and its expression pattern overlapped with that of *Fancd2*, which during meiosis, is known to localize in synaptonemal complexes of meiotic chromosomes together with BRCA1 (Garcia-Higuera, Taniguchi et al. 2001). At this stage of characterization, we hypothesized that Ube2u could have been involved in the regulation of Fancd2 during meiosis, therefore we moved further to confirm the connection with the FA-BRCA pathway. First, we determined whether UBE2U bound with other FA proteins and whether FA proteins bound each other in yeast. In addition to the previous with FANCD2, interactions between UBE2U and FANCC, FANCF and FANCL were detected. Being FANCL the E3 enzyme, it was tested against FANCD2 finding no interaction. However, FANCL was found to interact with BD-$FANCD2_1$ and BD-$FANCD2_3$. We detected also an interaction between BD-UBE2U and AD-FANCD2. The interaction between FANCE and FANCD2 was instead not confirmed (Pace P 2002; Gordon, Alon et al. 2005). However, we could not exclude that, despite the high sensitivity of the direct interaction mating, the low level of expression of FANCD2 full length has influenced this outcome. Once ascertained the interactions of UBE2U with FANCD2 and other FA proteins in yeast, we investigated these interactions in mammalian cells. Transient transfections of several cell lines were carried out with several constructs. Expression of the protein was investigated by immunofluorescence and western blotting analysis without, however, detecting the protein, even after inhibition of the proteasome. Now, since UBE2U has a restricted pattern of expression, we can hypothesize that it might be toxic to the cells not constitutively expressing it and that the control might be at level of mRNA (also by 5' and 3' UTRs). Only using translation and transcription *in vitro* we were able to produce small amounts of the tagged protein, the immunprecipitation of which, however, failed.

The inability to express UBE2U had a relevant impact on this PhD project, since it prevented us from studying the human protein. Moreover, since it is not endogenously expressed in all mammalian cell lines tested, alternative approaches like RNA interference were unfeasible. However, murine *Fancd2* and *Ube2u* showed an overlapping expression pattern in testis, therefore we exploited the mouse as a model to verify the interactions between the two proteins, after production of antibodies directed against the murine Ube2u. As described, they allowed us to detect the protein in spermatocytes, spermatides, and spermatozoa, but not to immunoprecipitate Ube2u. In the final part of this work, thanks to a suited antibody anti-Fancd2 (kindly donated by Dr Alan D'Andrea), we checked for interactions between Fancd2 and Ube2u, however again without success. Remarkably, although during our studies the ubiquitin-conjugating enzyme E2 of FANCD2 and FANCI was recognized as UBE2T (Machida, Machida et al. 2006), UBE2U should be further investigated to rule out any potential implication with the FA-BRCA network, which could enrol different players when meiosis occurs in testis, especially since UBE2T has not been studied yet in this tissue.

4.4 ISG20L2: another potential player of the FA pathway.

ISG20L2 was another gene studied in this project because it was found to interact with the C-terminal end of FANCG. The protein had a predicted conserved exonuclease domain EXOIII, a characteristic of a variety of exonuclease proteins (ribonuclease T, alpha and epsilon subunits of DNA polymerase III). Ribonuclease T is responsible for the end-turnover of tRNA, and removes the terminal AMP residue from uncharged tRNA. DNA polymerase III is a complex, multichain enzyme responsible for most of the replicative synthesis in bacteria, and exhibits 3' to 5' exonuclease activity. This clone was selected because its potential exonuclease activity could have linked it to the DNA-repair. We found ISG20L2 to be expressed in human testis, mainly in spermatides, but not in other tissues. However, its expression was detected in all cell lines tested. Moreover, immunofluorescence analysis localized the exogenous protein in nucleoli, in agreement with a proteomic study published in 2002 (Andersen, Lyon et al. 2002), but at the present there are no evidences for connections between the FA-BRCA pathway and the nucleolus compartment. The nucleolus is an organelle fundamental for ribosome biogenesis, even if several studies suggest that it

may also play a role in other mechanisms such as RNA transport, RNA modification, and cell cycle regulation (Schneiter, Kadowaki et al. 1995; Pederson 1998; Olson, Dundr et al. 2000; Ideue, Azad et al. 2004). In this regard, Gleizes and co-workers (2007) reported that ribosome biogenesis and in particular rRNA processing was altered in Diamond-Blackfan anemia, an inherited bone marrow failure syndrome due to mutations in at least 7 genes encoding for ribosomal proteins (Choesmel, Bacqueville et al. 2007). Moreover, there is evidence for a role of ribosome biogenesis in cancer (Amsterdam, Sadler et al. 2004; Pandolfi 2004). Therefore, we tested the interaction of ISG20L2 with FANCG using ectopically expressed proteins (FLAG-FANCG and HA-ISG20L2), but as mentioned, we were not able to confirm the direct binding, suggesting we should try further to define the best procedures changing experimental conditions or reagents such as epitope tags, incubation conditions and antibodies. In this regard, since there were no available antibodies against ISG20L2, we produced antibodies directed against both the C- and the N- termini of the human protein. Using different controls, such as FLAG-ISG20L2 over-expressed in HEK293 cells or transcripted and translated *in vitro*, we defined the conditions to detect the endogenous protein. ISG20L2 was expressed in all cell lines tested. Moreover, in addition to an expected nuclear localization, both endogenous and overexpressed proteins were also found in the cytoplasm. Previous immunofluorescence studies did not reveal any sign of the protein in that compartment, likely because its concentration is below the threshold level of detection for the technique.

During the last year, a work was published by Diaz and co-workers (Coute, Kindbeiter et al. 2008) showing that the recombinant tagged ISG20L2 protein is a component of nucleoli with a 3’ to 5’ exoribonuclease activity. Whereas its N-terminus binds to several ribosomal proteins, the C-terminal end is crucial for the localization in the nucleolus. When ISG20L2 is overexpressed, the 12S precursor of 5.8S rRNA decreases; conversely, it increases when ISG20L2 is silenced. All together these data indicated that ISG20L2 is involved in the biogenesis of the large ribosomal subunit and in the processing of 12S pre-rRNA to the mature 5.8S rRNA. In this context, although our results do not provide new insights into the functional role of ISG20L2, we extended our knowledge towards the characterization of the new gene. We produced and validate the first antibody, allowing us to reveal that the endogenous protein is not restricted to the nucleolus but it also localizes in the cytoplasm, rising new questions about its putative role in this cellular compartment. ISG20L2 contains an exonuclease domain conserved among members of the DEDDh group of the DEDD superfamily of exonuclease plethora. This group contains

several RNases, as well DNases, which are involved in various biological functions. More specifically, ISG20L2 is a member of the REX4 subgroup, which is composed of other three members in vertebrate (REX4, ISG20, and IS20L1). They all accumulate within nucleolus but their specific role in this nuclear compartment remains to be determined. As our knowledge on the REX4 family is very limited, any conclusions on the potential role of ISG20L2 within cytoplasm would be very speculative. However, since exonucleases are often multifunctional factors, we cannot exclude it might be involved in processing of cytoplasmic RNAs. Among the many different proteins identified as associated with the N-terminal half of ISG20L2, whereas 18 are ribosomal components, another 5 are involved in mRNA splicing and trafficking. (Counté et al, 2008). Therefore, it is plausible that ISG20L2, apart from an involvement in ribosome biogenesis and RNA metabolism within nucleoli, may play an additional role in the turn-over of mRNAs in the cytoplasm.

Chapter 5
CONCLUSIONS

Despite the identification of many genes involved in FA-BRCA pathway, a clear picture of how this pathway could explain the major phenotypic features of FA remains elusive. Their functional characterization has been leading to connect the FA proteins to the complex network responsible for DNA-damage response, even if many aspects remain to be elucidated. This PhD thesis project started and developed from this background with the aim to find new genes involved in the FA-BRCA pathway. By using the systematic approach of the yeast two-hybrid system, we identified two novel genes potentially involved in the FA-BRCA pathway.

The first encoded for a putative ubiquitin-conjugating enzyme UBE2U and interacted with one of the two monoubiquitinated members of the pathway, FANCD2. The other one, ISG20L2, encoded for a putative exonuclease and interacted with FANCG. Both were regarded as good candidates of the pathway and further investigated. The interactions between these proteins were at first validated in yeast, but afterwards they were not confirmed in mammalian cells. We concluded that the yeast two-hybrid system is not the most suitable tool for a large-scale screening of interactors in FA. Indeed, a careful examination of literature highlighted that this strategy was not particularly successful, especially when using a protein of the FA-core complex as bait. On the other hand, studies on FA protein associations based on characterization of immunocomplexes are being proved to be effective and will represent the most suitable technique in this field for the near future.

The highest expression of UBE2U and ISG20L2 in testis, resulted at first to be promising because it would lead us to provide insights into the role of the FA-BRCA pathway in testis. Indeed the FA pathway was never investigated during spermatogenesis, even though FA genes show the highest expression during this process and the only phenotype FA knock-out mice share with patients is infertility. However, a series of impediments (lack of over-expression, quality of antibodies produced) prevented us from gaining information into the field. For this reason, we cannot exclude that UBE2U and ISG20L2 might be functionally related to the FA-BRCA pathway at least in testis and further investigations are needed to definitively exclude these interactions. Moreover, the enzymatic characterizations of UBE2U as ubiquitin-conjugating enzyme E2 and ISG20L2 as a potential cytoplasmic ribonuclease should be determinant to shed light upon the functional role of these two genes.

ACKNOWLEDGEMENTS

My deepest appreciations go to prof. Anna Savoia for giving me the important chance of this training, for her constructive criticism, for her trust and for encouraging me in dire times. I owe also a great debt to prof. Fulvio Della Ragione for having me as a full member of his lab during my PhD training and for his support during these years. Big thanks go also to his lab, “FDR’s lab”, for their company and for their help. Special thanks go to my friends and colleagues at TiGeM for their uniqueness. Least but not last, huge thanks go to dr. Mariateresa Di Stazio, dr. Annamonica Bianco and dr. Valeria Conti for their help, support, friendship and understanding: without their contributions, this work could not have been done.

REFERENCES

Agoulnik, A. I., B. Lu, et al. (2002). "A novel gene, Pog, is necessary for primordial germ cell proliferation in the mouse and underlies the germ cell deficient mutation, gcd." Hum Mol Genet **11**(24): 3047-53.

Alter, B. (2003). Inherited bone marrow failure syndromes. Hematology of Infancy and Childhood. O. S. Nathan DG, Ginsburg D, Look AT, Saunders. **1**.

Alter, B. P. (1994). Clinical features of Fanconi anaemia. Aplastic Anaemia: Acquired and Inherited. N. S. Y. B. P. Alter, Saunders, Philadelphia**:** 275–308.

Alter, B. P. (2003). "Cancer in Fanconi anemia, 1927-2001." Cancer **97**(2): 425-40.

Alter, B. P., M. H. Greene, et al. (2003). "Cancer in Fanconi anemia." Blood **101**(5): 2072.

Alter, B. P., H. Joenje, et al. (2005). "Fanconi anemia: adult head and neck cancer and hematopoietic mosaicism." Arch Otolaryngol Head Neck Surg **131**(7): 635-9.

Amsterdam, A., K. C. Sadler, et al. (2004). "Many ribosomal protein genes are cancer genes in zebrafish." PLoS Biol **2**(5): E139.

Andersen, J. S., C. E. Lyon, et al. (2002). "Directed proteomic analysis of the human nucleolus." Curr Biol **12**(1): 1-11.

Andreassen, P. R., A. D. D'Andrea, et al. (2004). "ATR couples FANCD2 monoubiquitination to the DNA-damage response." Genes Dev **18**(16): 1958-63.

Aronheim, A., D. Engelberg, et al. (1994). "Membrane targeting of the nucleotide exchange factor Sos is sufficient for activating the Ras signaling pathway." Cell **78**(6): 949-61.

Aronheim, A., E. Zandi, et al. (1997). "Isolation of an AP-1 repressor by a novel method for detecting protein-protein interactions." Mol Cell Biol **17**(6): 3094-102.

Auerbach, A., Buchwald, M. & Joenje, H. (1999). Fanconi anaemia. The Genetic Basis of Human Cancer. B. V. K. W. Kinzler. New York, McGraw Hill Inc**:** 317-332.

Auerbach, A. D. (1993). "Fanconi anemia diagnosis and the diepoxybutane (DEB) test." Exp Hematol **21**(6): 731-3.

Auerbach, A. D. and R. G. Allen (1991). "Leukemia and preleukemia in Fanconi anemia patients. A review of the literature and report of the International Fanconi Anemia Registry." Cancer Genet Cytogenet **51**(1): 1-12.

Auerbach AD, B. M., Joenje H. (2001). Fanconi anemia. The Metabolic and Molecular Bases of Inherited Disease. N. Y. McGraw-Hill. **1:** 753-768.

Auerbach, A. D., M. Sagi, et al. (1985). "Fanconi anemia: prenatal diagnosis in 30 fetuses at risk." Pediatrics **76**(5): 794-800.

Auerbach, A. D. and S. R. Wolman (1976). "Susceptibility of Fanconi's anaemia fibroblasts to chromosome damage by carcinogens." Nature **261**(5560): 494-6.

Ayas, M., A. Al-Jefri, et al. (2008). "Allogeneic stem cell transplantation in Fanconi anemia patients presenting with myelodysplasia and/or clonal abnormality: update on the Saudi experience." Bone Marrow Transplant **41**(3): 261-5.

Ball, S. E., F. M. Gibson, et al. (1998). "Progressive telomere shortening in aplastic anemia." Blood **91**(10): 3582-92.

Berger, R., A. Bernheim, et al. (1980). "In vitro effect of cyclophosphamide metabolites on chromosomes of Fanconi anaemia patients." Br J Haematol **45**(4): 565-8.

Berger, R., M. Le Coniat, et al. (1993). "Fanconi anemia. Chromosome breakage and cell cycle studies." Cancer Genet Cytogenet **69**(1): 13-6.

Bitan, M., R. Or, et al. (2006). "Fludarabine-based reduced intensity conditioning for stem cell transplantation of Fanconi anemia patients from fully matched related and unrelated donors." Biol Blood Marrow Transplant **12**(7): 712-8.

Bjergbaek, L., J. A. Cobb, et al. (2005). "Mechanistically distinct roles for Sgs1p in checkpoint activation and replication fork maintenance." Embo J **24**(2): 405-17.

Bogliolo, M., A. Lyakhovich, et al. (2007). "Histone H2AX and Fanconi anemia FANCD2 function in the same pathway to maintain chromosome stability." Embo J **26**(5): 1340-51.

Boultwood, J., C. Fidler, et al. (1997). "Telomere length in myelodysplastic syndromes." Am J Hematol **56**(4): 266-71.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Anal Biochem **72**: 248-54.

Bridge, W. L., C. J. Vandenberg, et al. (2005). "The BRIP1 helicase functions independently of BRCA1 in the Fanconi anemia pathway for DNA crosslink repair." Nat Genet **37**(9): 953-7.

Burkitt, K. and M. Ljungman (2008). "Phenylbutyrate interferes with the Fanconi anemia and BRCA pathway and sensitizes head and neck cancer cells to cisplatin." Mol Cancer **7**: 24.

Busso, D., D. J. Cohen, et al. (2005). "Human testicular protein TPX1/CRISP-2: localization in spermatozoa, fate after capacitation and relevance for gamete interaction." Mol Hum Reprod **11**(4): 299-305.

Butturini, A., R. P. Gale, et al. (1994). "Hematologic abnormalities in Fanconi anemia: an International Fanconi Anemia Registry study." Blood **84**(5): 1650-5.

Callen, E., E. Samper, et al. (2002). "Breaks at telomeres and TRF2-independent end fusions in Fanconi anemia." Hum Mol Genet **11**(4): 439-44.

Cantor, S., R. Drapkin, et al. (2004). "The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations." Proc Natl Acad Sci U S A **101**(8): 2357-62.

Carreau, M., O. I. Gan, et al. (1998). "Bone marrow failure in the Fanconi anemia group C mouse model after DNA damage." Blood **91**(8): 2737-44.

Carreau, M., L. Liu, et al. (2002). "Short-term granulocyte colony-stimulating factor and erythropoietin treatment enhances hematopoiesis and survival in the mitomycin C-conditioned Fancc(-/-) mouse model, while long-term treatment is ineffective." Blood **100**(4): 1499-501.

Cassinat, B., D. Darsin, et al. (2001). "Intermethod discordance for alpha-fetoprotein measurements in Fanconi anemia." Clin Chem **47**(8): 1405-9.

Cassinat, B., P. Guardiola, et al. (2000). "Constitutive elevation of serum alpha-fetoprotein in Fanconi anemia." Blood **96**(3): 859-63.

Cavenagh, J. D., D. S. Richardson, et al. (1996). "Fanconi's anaemia presenting as acute myeloid leukaemia in adulthood." Br J Haematol **94**(1): 126-8.

Chen, M., D. J. Tomkins, et al. (1996). "Inactivation of Fac in mice produces inducible chromosomal instability and reduced fertility reminiscent of Fanconi anaemia." Nat Genet **12**(4): 448-51.

Cheng, N. C., H. J. van de Vrugt, et al. (2000). "Mice with a targeted disruption of the Fanconi anemia homolog Fanca." Hum Mol Genet **9**(12): 1805-11.

Chien, C. T., P. L. Bartel, et al. (1991). "The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest." Proc Natl Acad Sci U S A **88**(21): 9578-82.

Chirnomas, D., T. Taniguchi, et al. (2006). "Chemosensitization to cisplatin by inhibitors of the Fanconi anemia/BRCA pathway." Mol Cancer Ther **5**(4): 952-61.

Choesmel, V., D. Bacqueville, et al. (2007). "Impaired ribosome biogenesis in Diamond-Blackfan anemia." Blood **109**(3): 1275-83.

Ciccia A, L. C., Coulthard R, Yan Z, Xue Y, Meetei AR, Laghmani el H, Joenje H, McDonald N, de Winter JP, Wang W, West SC. (2007). "Identification of FAAP24,

a Fanconi anemia core complex protein that interacts with FANCM." Mol Cell. **25(3)**: 331-43. .
Ciccia, A., C. Ling, et al. (2007). "Identification of FAAP24, a Fanconi anemia core complex protein that interacts with FANCM." Mol Cell **25**(3): 331-43.
Clarke, A. A., N. J. Philpott, et al. (1997). "The sensitivity of Fanconi anaemia group C cells to apoptosis induced by mitomycin C is due to oxygen radical generation, not DNA crosslinking." Br J Haematol **96**(2): 240-7.
Cobb, J. A., L. Bjergbaek, et al. (2003). "DNA polymerase stabilization at stalled replication forks requires Mec1 and the RecQ helicase Sgs1." Embo J **22**(16): 4325-36.
Collis, S. J., L. J. Barber, et al. (2007). "HCLK2 is essential for the mammalian S-phase checkpoint and impacts on Chk1 stability." Nat Cell Biol **9**(4): 391-401.
Collis, S. J., L. J. Barber, et al. (2006). "C. elegans FANCD2 responds to replication stress and functions in interstrand cross-link repair." DNA Repair (Amst) **5**(11): 1398-406.
Condie, A., R. L. Powles, et al. (2002). "Analysis of the Fanconi anaemia complementation group A gene in acute myeloid leukaemia." Leuk Lymphoma **43**(9): 1849-53.
Connor, F., D. Bertwistle, et al. (1997). "Tumorigenesis and a DNA repair defect in mice with a truncating Brca2 mutation." Nat Genet **17**(4): 423-30.
Consortium, F. A. B. C. (1996). "Positional cloning of the Fanconi anaemia group A gene." Nature Genet. **14**: 324-328.
Coute, Y., K. Kindbeiter, et al. (2008). "ISG20L2, a novel vertebrate nucleolar exoribonuclease involved in ribosome biogenesis." Mol Cell Proteomics **7**(3): 546-59.
Cumming, R. C., J. Lightfoot, et al. (2001). "Fanconi anemia group C protein prevents apoptosis in hematopoietic cells through redox regulation of GSTP1." Nat Med **7**(7): 814-20.
D'Andrea AD, G. M. (2003). "The Fanconi anaemia/BRCA pathway." Nat Rev Cancer **3**: 23-34.
Das, A. K., P. W. Cohen, et al. (1998). "The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions." Embo J **17**(5): 1192-9.
de Cremoux, P., E. Gluckman, et al. (1996). "Decreased IL-1 beta and TNF alpha secretion in long-term bone marrow culture supernatant from Fanconi's anaemia patients." Eur J Haematol **57**(3): 202-7.
De Kerviler, E., A. Guermazi, et al. (2000). "The clinical and radiological features of Fanconi's anaemia." Clin Radiol **55**(5): 340-5.
de Winter JP, L. F., van Berkel CG, Rooimans MA, van Der Weel L, Steltenpool J, Demuth I, Morgan NV, Alon N, Bosnoyan-Collins L, Lightfoot J, Leegwater PA, Waisfisz Q, Komatsu K, Arwert F, Pronk JC, Mathew CG, Digweed M, Buchwald M, Joenje H (2000). "Isolation of a cDNA representing the Fanconi anemia complementation group E gene." Am J Hum Genet. **67(5)**: 1306-8.
de Winter JP, R. M., van Der Weel L, van Berkel CG, Alon N, Bosnoyan-Collins L, de Groot J, Zhi Y, Waisfisz Q, Pronk JC, Arwert F, Mathew CG, Scheper RJ, Hoatlin ME, Buchwald M, Joenje H (2000). "The Fanconi anaemia gene FANCF encodes a novel protein with homology to ROM." Nat Genet. **24(1)**: 15-6.
de Winter, J. P., L. van der Weel, et al. (2000). "The Fanconi anemia protein FANCF forms a nuclear complex with FANCA, FANCC and FANCG." Hum Mol Genet **9**(18): 2665-74.
de Winter JP, W. Q., Rooimans MA, van Berkel CG, Bosnoyan-Collins L, Alon N, Carreau M, Bender O, Demuth I, Schindler D, Pronk JC, Arwert F, Hoehn H, Digweed M, Buchwald M, Joenje H. (1998). "The Fanconi anaemia group G gene FANCG is identical with XRCC9." Nat Genetic **20(3)**: 281-3.

Deeg, H. J., G. Socie, et al. (1996). "Malignancies after marrow transplantation for aplastic anemia and fanconi anemia: a joint Seattle and Paris analysis of results in 700 patients." Blood **87**(1): 386-92.
Dhillon, V. S., M. Shahid, et al. (2004). "CpG methylation of the FHIT, FANCF, cyclin-D2, BRCA2 and RUNX3 genes in Granulosa cell tumors (GCTs) of ovarian origin." Mol Cancer **3**: 33.
Dokal, I., A. Chase, et al. (1996). "Positive diepoxybutane test in only one of two brothers found to be compound heterozygotes for Fanconi's anaemia complementation group C mutations." Br J Haematol **93**(4): 813-6.
Dorsman, J. C., M. Levitus, et al. (2007). "Identification of the Fanconi anemia complementation group I gene, FANCI." Cell Oncol **29**(3): 211-8.
Doyon, Y. and J. Cote (2004). "The highly conserved and multifunctional NuA4 HAT complex." Curr Opin Genet Dev **14**(2): 147-54.
Doyon, Y., W. Selleck, et al. (2004). "Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans." Mol Cell Biol **24**(5): 1884-96.
Dufour, C. and J. Svahn (2008). "Fanconi anaemia: new strategies." Bone Marrow Transplant **41 Suppl 2**: S90-5.
Durfee, T., K. Becherer, et al. (1993). "The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit." Genes Dev **7**(4): 555-69.
Fanconi, G. (1927). "Familiaere infantile perniziosaartige Anaemie (pernizioeses blutbild and konstitution)." Jahrbuch Kinderheild **117**: 257-280.
Fanconi, G. (1964). "[Hypothesis of Chromosomal Translocation as a Genetic Interpretation of Fanconi's Familial Constitutional Panmyelopathy.]." Helv Paediatr Acta **19**: 29-33.
Fanconi, G. (1967). "Familial constitutional panmyelocytopathy, Fanconi's anemia (F.A.). I. Clinical aspects." Semin Hematol **4**(3): 233-40.
Fanconi, G. (1970). "Der Wandel der Medizin - wie ich ihn erlebte."
Farzin, A., S. M. Davies, et al. (2007). "Matched sibling donor haematopoietic stem cell transplantation in Fanconi anaemia: an update of the Cincinnati Children's experience." Br J Haematol **136**(4): 633-40.
Fields, S. and O. Song (1989). "A novel genetic system to detect protein-protein interactions." Nature **340**(6230): 245-6.
Flowers, M. E., J. Zanis, et al. (1996). "Marrow transplantation for Fanconi anaemia: conditioning with reduced doses of cyclophosphamide without radiation." Br J Haematol **92**(3): 699-706.
Folias, A., M. Matkovic, et al. (2002). "BRCA1 interacts directly with the Fanconi anemia protein FANCA." Hum Mol Genet **11**(21): 2591-7.
Friedman, L. S., F. C. Thistlethwaite, et al. (1998). "Thymic lymphomas in mice with a truncating mutation in Brca2." Cancer Res **58**(7): 1338-43.
Futaki, M., T. Igarashi, et al. (2002). "The FANCG Fanconi anemia protein interacts with CYP2E1: possible role in protection against oxidative DNA damage." Carcinogenesis **23**(1): 67-72.
Galimi, F., M. Noll, et al. (2002). "Gene therapy of Fanconi anemia: preclinical efficacy using lentiviral vectors." Blood **100**(8): 2732-6.
Garcia-Higuera, I., Y. Kuang, et al. (1999). "Fanconi anemia proteins FANCA, FANCC, and FANCG/XRCC9 interact in a functional nuclear complex." Mol Cell Biol **19**(7): 4866-73.
Garcia-Higuera, I., T. Taniguchi, et al. (2001). "Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway." Mol Cell **7**(2): 249-62.

Giampietro, P. F., P. C. Verlander, et al. (1997). "Diagnosis of Fanconi anemia in patients without congenital malformations: an international Fanconi Anemia Registry Study." Am J Med Genet **68**(1): 58-61.

Glickman, M. H. and A. Ciechanover (2002). "The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction." Physiol Rev **82**(2): 373-428.

Gluckman, E., A. Devergie, et al. (1983). "Radiosensitivity in Fanconi anaemia: application to the conditioning regimen for bone marrow transplantation." Br J Haematol **54**(3): 431-40.

Gluckman, E., A. Devergie, et al. (1980). "Bone marrow transplantation in Fanconi anaemia." Br J Haematol **45**(4): 557-64.

Gluckman, E. and J. E. Wagner (2008). "Hematopoietic stem cell transplantation in childhood inherited bone marrow failure syndrome." Bone Marrow Transplant **41**(2): 127-32.

Godthelp, B. C., W. W. Wiegant, et al. (2006). "Inducibility of nuclear Rad51 foci after DNA damage distinguishes all Fanconi anemia complementation groups from D1/BRCA2." Mutat Res **594**(1-2): 39-48.

Gordon, S. M., N. Alon, et al. (2005). "FANCC, FANCE, and FANCD2 form a ternary complex essential to the integrity of the Fanconi anemia DNA damage response pathway." J Biol Chem **280**(43): 36118-25.

Gordon, S. M. and M. Buchwald (2003). "Fanconi anemia protein complex: mapping protein interactions in the yeast 2- and 3-hybrid systems." Blood **102**(1): 136-41.

Greenberg, R. A., B. Sobhian, et al. (2006). "Multifactorial contributions to an acute DNA damage response by BRCA1/BARD1-containing complexes." Genes Dev **20**(1): 34-46.

Guardiola, P., G. Socie, et al. (2004). "Acute graft-versus-host disease in patients with Fanconi anemia or acquired aplastic anemia undergoing bone marrow transplantation from HLA-identical sibling donors: risk factors and influence on outcome." Blood **103**(1): 73-7.

Gurtan, A. M., P. Stuckert, et al. (2006). "The WD40 repeats of FANCL are required for Fanconi anemia core complex assembly." J Biol Chem **281**(16): 10896-905.

Gush, K. A., K. L. Fu, et al. (2000). "Phenotypic correction of Fanconi anemia group C knockout mice." Blood **95**(2): 700-4.

Guy, C. P. and E. L. Bolt (2005). "Archaeal Hel308 helicase targets replication forks in vivo and in vitro and unwinds lagging strands." Nucleic Acids Res **33**(11): 3678-90.

H Joenje, A. B. O., M Wijker, F M di Summa, C G van Berkel, M A Rooimans, W Ebell, M van Weel, J C Pronk, M Buchwald, and F Arwert (1997). "Evidence for at least eight Fanconi anemia genes." Am J Hum Genet **61(4)**: 940-944.

Hahn, W. C. and R. A. Weinberg (2002). "Rules for making human tumor cells." N Engl J Med **347**(20): 1593-603.

Hanada, K., M. Budzowska, et al. (2006). "The structure-specific endonuclease Mus81-Eme1 promotes conversion of interstrand DNA crosslinks into double-strands breaks." Embo J **25**(20): 4921-32.

Haneline, L. S., X. Li, et al. (2003). "Retroviral-mediated expression of recombinant Fancc enhances the repopulating ability of Fancc-/- hematopoietic stem cells and decreases the risk of clonal evolution." Blood **101**(4): 1299-307.

Hanenberg, H., S. D. Batish, et al. (2002). "Phenotypic correction of primary Fanconi anemia T cells with retroviral vectors as a diagnostic tool." Exp Hematol **30**(5): 410-20.

Hanson, H., C. G. Mathew, et al. (2001). "Telomere shortening in Fanconi anaemia demonstrated by a direct FISH approach." Cytogenet Cell Genet **93**(3-4): 203-6.

Hejna, J., M. Holtorf, et al. (2008). "Tip60 is required for DNA interstrand cross-link repair in the Fanconi anemia pathway." J Biol Chem **283**(15): 9844-51.

Hirsch, B., A. Shimamura, et al. (2004). "Association of biallelic BRCA2/FANCD1 mutations with spontaneous chromosomal instability and solid tumors of childhood." Blood **103**(7): 2554-9.

Ho, G. P., S. Margossian, et al. (2006). "Phosphorylation of FANCD2 on two novel sites is required for mitomycin C resistance." Mol Cell Biol **26**(18): 7005-15.

Hoatlin, M. E., Y. Zhi, et al. (1999). "A novel BTB/POZ transcriptional repressor protein interacts with the Fanconi anemia group C protein and PLZF." Blood **94**(11): 3737-47.

Hochstrasser, M. (2006). "Lingering mysteries of ubiquitin-chain assembly." Cell **124**(1): 27-34.

Hoshino, T., J. Wang, et al. (1998). "Molecular chaperone GRP94 binds to the Fanconi anemia group C protein and regulates its intracellular expression." Blood **91**(11): 4379-86.

Houghtaling, S., C. Timmers, et al. (2003). "Epithelial cancer in Fanconi anemia complementation group D2 (Fancd2) knockout mice." Genes Dev **17**(16): 2021-35.

Howlett, N. G., T. Taniguchi, et al. (2002). "Biallelic inactivation of BRCA2 in Fanconi anemia." Science **297**(5581): 606-9.

Hu, J. C. (2001). "Model systems: Studying molecular recognition using bacterial n-hybrid systems." Trends Microbiol **9**(5): 219-22.

Huang, T. T., S. M. Nijman, et al. (2006). "Regulation of monoubiquitinated PCNA by DUB autocleavage." Nat Cell Biol **8**(4): 339-47.

Hussain, S., J. B. Wilson, et al. (2004). "Direct interaction of FANCD2 with BRCA2 in DNA damage response pathways." Hum Mol Genet **13**(12): 1241-8.

Hussain, S., E. Witt, et al. (2003). "Direct interaction of the Fanconi anaemia protein FANCG with BRCA2/FANCD1." Hum Mol Genet **12**(19): 2503-10.

Ideue, T., A. K. Azad, et al. (2004). "The nucleolus is involved in mRNA export from the nucleus in fission yeast." J Cell Sci **117**(Pt 14): 2887-95.

Ikura, T., S. Tashiro, et al. (2007). "DNA damage-dependent acetylation and ubiquitination of H2AX enhances chromatin dynamics." Mol Cell Biol **27**(20): 7028-40.

Imerslund, O. (1953). "[Hypoplastic anemia with multiple abnormalities (Fanconi's anemia).]." Nord Med **50**(38): 1301-4.

Ishida, R. and M. Buchwald (1982). "Susceptibility of Fanconi's anemia lymphoblasts to DNA-cross-linking and alkylating agents." Cancer Res **42**(10): 4000-6.

Jacome, A., S. Navarro, et al. (2006). "A simplified approach to improve the efficiency and safety of ex vivo hematopoietic gene therapy in fanconi anemia patients." Hum Gene Ther **17**(2): 245-50.

Joazeiro, C. A. and A. M. Weissman (2000). "RING finger proteins: mediators of ubiquitin ligase activity." Cell **102**(5): 549-52.

Joenje, H., F. Arwert, et al. (1981). "Oxygen-dependence of chromosomal aberrations in Fanconi's anaemia." Nature **290**(5802): 142-3.

Joenje H, L. M., Waisfisz Q, D'Andrea A, Garcia-Higuera I, Pearson T, van Berkel CG, Rooimans MA, Morgan N, Mathew CG, Arwert F (2000). "Complementation analysis in Fanconi anemia: assignment of the reference FA-H patient to group A." Am J Hum Genet. **67(3)**: 759-762.

Joenje, H. and K. J. Patel (2001). "The emerging genetic and molecular basis of Fanconi anaemia." Nat Rev Genet **2**(6): 446-57.

Joenje, H. G., J.J.P. (1989). Oxygen metabolism and chromosomal breakage in Fanconi anaemia. Fanconi Anaemia: Clinical, Cytogenetic and Experimental Aspects of Fanconi Anaemia. G. O. A. D. A. T.M. Schroeder-Kurth. Berlin, Springer-Verlag**:** 174-182.

Kalb, R., K. Neveling, et al. (2007). "Hypomorphic mutations in the gene encoding a key Fanconi anemia protein, FANCD2, sustain a significant group of FA-D2 patients with severe phenotype." Am J Hum Genet **80**(5): 895-910.

Kannouche, P. L., J. Wing, et al. (2004). "Interaction of human DNA polymerase eta with monoubiquitinated PCNA: a possible mechanism for the polymerase switch in response to DNA damage." Mol Cell **14**(4): 491-500.

Kapelushnik, J., R. Or, et al. (1997). "A fludarabine-based protocol for bone marrow transplantation in Fanconi's anemia." Bone Marrow Transplant **20**(12): 1109-10.

Kelly, P. F., S. Radtke, et al. (2007). "Stem cell collection and gene transfer in Fanconi anemia." Mol Ther **15**(1): 211-9.

Kennedy, R. D. and A. D. D'Andrea (2005). "The Fanconi Anemia/BRCA pathway: new faces in the crowd." Genes Dev **19**(24): 2925-40.

Kim, H., J. Chen, et al. (2007). "Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response." Science **316**(5828): 1202-5.

Koc, A., J. C. Pronk, et al. (1999). "Variable pathogenicity of exon 43del (FAA) in four Fanconi anaemia patients within a consanguineous family." Br J Haematol **104**(1): 127-30.

Kolch, W. (2000). "Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by protein interactions." Biochem J **351 Pt 2**: 289-305.

Komori, K., R. Fujikane, et al. (2002). "Novel endonuclease in Archaea cleaving DNA with various branched structure." Genes Genet Syst **77**(4): 227-41.

Komori, K., M. Hidaka, et al. (2004). "Cooperation of the N-terminal Helicase and C-terminal endonuclease activities of Archaeal Hef protein in processing stalled replication forks." J Biol Chem **279**(51): 53175-85.

Koomen, M., N. C. Cheng, et al. (2002). "Reduced fertility and hypersensitivity to mitomycin C characterize Fancg/Xrcc9 null mice." Hum Mol Genet **11**(3): 273-81.

Kowal, P., A. M. Gurtan, et al. (2007). "Structural determinants of human FANCF protein that function in the assembly of a DNA damage signaling complex." J Biol Chem **282**(3): 2047-55.

Kruyt, F. A., F. Abou-Zahr, et al. (1999). "Resistance to mitomycin C requires direct interaction between the Fanconi anemia proteins FANCA and FANCG in the nucleus through an arginine-rich domain." J Biol Chem **274**(48): 34212-8.

Kruyt, F. A., L. M. Dijkmans, 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, F. A., T. Hoshino, 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): 3050-6.

Kunisato, A., S. Chiba, et al. (2003). "HES-1 preserves purified hematopoietic stem cells ex vivo and accumulates side population cells in vivo." Blood **101**(5): 1777-83.

Kupfer, G. M., D. Naf, et al. (1997). "The Fanconi anaemia proteins, FAA and FAC, interact to form a nuclear complex." Nat Genet **17**(4): 487-90.

Kuraoka, I., W. R. Kobertz, et al. (2000). "Repair of an interstrand DNA cross-link initiated by ERCC1-XPF repair/recombination nuclease." J Biol Chem **275**(34): 26632-6.

Kusch, T., L. Florens, et al. (2004). "Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions." Science **306**(5704): 2084-7.

Kutler, D. I., A. D. Auerbach, et al. (2003). "High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia." Arch Otolaryngol Head Neck Surg **129**(1): 106-12.

Kutler, D. I., B. Singh, et al. (2003). "A 20-year perspective on the International Fanconi Anemia Registry (IFAR)." Blood **101**(4): 1249-56.

Kutler, D. I., V. B. Wreesmann, et al. (2003). "Human papillomavirus DNA and p53 polymorphisms in squamous cell carcinomas from Fanconi anemia patients." J Natl Cancer Inst **95**(22): 1718-21.

Lancaster, J. M., R. Wooster, et al. (1996). "BRCA2 mutations in primary breast and ovarian cancers." Nat Genet **13**(2): 238-40.

Lensch, M. W., R. K. Rathbun, et al. (1999). "Selective pressure as an essential force in molecular evolution of myeloid leukemic clones: a view from the window of Fanconi anemia." Leukemia **13**(11): 1784-9.

Leteurtre, F., X. Li, et al. (1999). "Accelerated telomere shortening and telomerase activation in Fanconi's anaemia." Br J Haematol **105**(4): 883-93.

Levitus, M., Joenje, H. & de Winter, J. P. (2006). "The Fanconi anemia pathway of genomic maintenance." Cell Oncol. **28**: 3–29.

Levitus, M., M. A. Rooimans, et al. (2004). "Heterogeneity in Fanconi anemia: evidence for 2 new genetic subtypes." Blood **103**(7): 2498-503.

Levitus, M., Q. Waisfisz, et al. (2005). "The DNA helicase BRIP1 is defective in Fanconi anemia complementation group J." Nat Genet **37**(9): 934-5.

Levran, O., C. Attwooll, et al. (2005). "The BRCA1-interacting helicase BRIP1 is deficient in Fanconi anemia." Nat Genet **37**(9): 931-3.

Li, X., M. M. Le Beau, et al. (2005). "Ex vivo culture of Fancc-/- stem/progenitor cells predisposes cells to undergo apoptosis, and surviving stem/progenitor cells display cytogenetic abnormalities and an increased risk of malignancy." Blood **105**(9): 3465-71.

Li, X., F. Leteurtre, et al. (2003). "Abnormal telomere metabolism in Fanconi's anaemia correlates with genomic instability and the probability of developing severe aplastic anaemia." Br J Haematol **120**(5): 836-45.

Liede, A., B. Y. Karlan, et al. (2004). "Cancer risks for male carriers of germline mutations in BRCA1 or BRCA2: a review of the literature." J Clin Oncol **22**(4): 735-42.

Ling, C. e. a. (2007). "FAAP100 is essential for activation of the Fanconi anemia-associated DNA damage response pathway." EMBO J. **26**: 2104-2114.

Litman, R., M. Peng, et al. (2005). "BACH1 is critical for homologous recombination and appears to be the Fanconi anemia gene product FANCJ." Cancer Cell **8**(3): 255-65.

Liu, J. M., A. D. Auerbach, et al. (1991). "Fanconi anemia presenting unexpectedly in an adult kindred with no dysmorphic features." Am J Med **91**(5): 555-7.

Liu, T. X., N. G. Howlett, et al. (2003). "Knockdown of zebrafish Fancd2 causes developmental abnormalities via p53-dependent apoptosis." Dev Cell **5**(6): 903-14.

Lo Ten Foe JR, R. M., Bosnoyan-Collins L, Alon N, Wijker M, Parker L, Lightfoot J, Carreau M, Callen DF, Savoia A, Cheng NC, van Berkel CG, Strunk MH, Gille JJ, Pals G, Kruyt FA, Pronk JC, Arwert F, Buchwald M, Joenje H. (1996). "Expression cloning of a cDNA for the major Fanconi anaemia gene, FAA." Nat Genet. **14(3)**: 320-3.

Locatelli, F., M. Zecca, et al. (2007). "The outcome of children with Fanconi anemia given hematopoietic stem cell transplantation and the influence of fludarabine in the conditioning regimen: a report from the Italian pediatric group." Haematologica **92**(10): 1381-8.

Lorick, K. L., J. P. Jensen, et al. (1999). "RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination." Proc Natl Acad Sci U S A **96**(20): 11364-9.

Lowy, D. R. and M. L. Gillison (2003). "A new link between Fanconi anemia and human papillomavirus-associated malignancies." J Natl Cancer Inst **95**(22): 1648-50.

Ludwig, T., D. L. Chapman, et al. (1997). "Targeted mutations of breast cancer susceptibility gene homologs in mice: lethal phenotypes of Brca1, Brca2, Brca1/Brca2, Brca1/p53, and Brca2/p53 nullizygous embryos." Genes Dev **11**(10): 1226-41.

Machida, Y. J., Y. Machida, et al. (2006). "UBE2T is the E2 in the Fanconi anemia pathway and undergoes negative autoregulation." Mol Cell **23**(4): 589-96.
Marek, L. R. and A. E. Bale (2006). "Drosophila homologs of FANCD2 and FANCL function in DNA repair." DNA Repair (Amst) **5**(11): 1317-26.
Matsuoka, S., B. A. Ballif, et al. (2007). "ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage." Science **316**(5828): 1160-6.
Matsushita, N., H. Kitao, et al. (2005). "A FancD2-monoubiquitin fusion reveals hidden functions of Fanconi anemia core complex in DNA repair." Mol Cell **19**(6): 841-7.
McHugh, P. J. and S. Sarkar (2006). "DNA interstrand cross-link repair in the cell cycle: a critical role for polymerase zeta in G1 phase." Cell Cycle **5**(10): 1044-7.
Medhurst, A. L., P. A. Huber, et al. (2001). "Direct interactions of the five known Fanconi anaemia proteins suggest a common functional pathway." Hum Mol Genet **10**(4): 423-9.
Meetei, A. R., J. P. de Winter, et al. (2003). "A novel ubiquitin ligase is deficient in Fanconi anemia." Nat Genet **35**(2): 165-70.
Meetei, A. R., M. Levitus, et al. (2004). "X-linked inheritance of Fanconi anemia complementation group B." Nat Genet **36**(11): 1219-24.
Meetei, A. R., A. L. Medhurst, et al. (2005). "A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M." Nat Genet **37**(9): 958-63.
Meetei, A. R., S. Sechi, et al. (2003). "A multiprotein nuclear complex connects Fanconi anemia and Bloom syndrome." Mol Cell Biol **23**(10): 3417-26.
Meetei, A. R., Z. Yan, et al. (2004). "FANCL replaces BRCA1 as the likely ubiquitin ligase responsible for FANCD2 monoubiquitination." Cell Cycle **3**(2): 179-81.
Meng, X. and S. A. Wolfe (2006). "Identifying DNA sequences recognized by a transcription factor using a bacterial one-hybrid system." Nat Protoc **1**(1): 30-45.
Mosedale G, N. W., Alpi A, Perrina F, Pereira-Leal JB, Johnson M, Langevin F, Pace P, Patel KJ. (2005). "The vertebrate Hef ortholog is a component of the Fanconi anemia tumor-suppressor pathway." Nat Struct Mol Biol. **12(9)**: 763-71.
Moynahan, M. E. (2002). "The cancer connection: BRCA1 and BRCA2 tumor suppression in mice and humans." Oncogene **21**(58): 8994-9007.
Moynahan, M. E., A. J. Pierce, et al. (2001). "BRCA2 is required for homology-directed repair of chromosomal breaks." Mol Cell **7**(2): 263-72.
Nakanishi, K., T. Taniguchi, et al. (2002). "Interaction of FANCD2 and NBS1 in the DNA damage response." Nat Cell Biol **4**(12): 913-20.
Nakanishi, K., Y. G. Yang, et al. (2005). "Human Fanconi anemia monoubiquitination pathway promotes homologous DNA repair." Proc Natl Acad Sci U S A **102**(4): 1110-5.
Narayan, G., H. Arias-Pulido, et al. (2004). "Promoter hypermethylation of FANCF: disruption of Fanconi Anemia-BRCA pathway in cervical cancer." Cancer Res **64**(9): 2994-7.
Navarro, S., N. W. Meza, et al. (2006). "Hematopoietic dysfunction in a mouse model for Fanconi anemia group D1." Mol Ther **14**(4): 525-35.
Niedzwiedz, W., G. Mosedale, et al. (2004). "The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair." Mol Cell **15**(4): 607-20.
Nijman, S. M., T. T. Huang, et al. (2005). "The deubiquitinating enzyme USP1 regulates the Fanconi anemia pathway." Mol Cell **17**(3): 331-9.
Nojima, K., H. Hochegger, et al. (2005). "Multiple repair pathways mediate tolerance to chemotherapeutic cross-linking agents in vertebrate cells." Cancer Res **65**(24): 11704-11.

Noll, M., R. L. Bateman, et al. (2001). "Preclinical protocol for in vivo selection of hematopoietic stem cells corrected by gene therapy in Fanconi anemia group C." Mol Ther **3**(1): 14-23.

Noll, M., K. P. Battaile, et al. (2002). "Fanconi anemia group A and C double-mutant mice: functional evidence for a multi-protein Fanconi anemia complex." Exp Hematol **30**(7): 679-88.

Offit, K., O. Levran, et al. (2003). "Shared genetic susceptibility to breast cancer, brain tumors, and Fanconi anemia." J Natl Cancer Inst **95**(20): 1548-51.

Olson, M. O., M. Dundr, et al. (2000). "The nucleolus: an old factory with unexpected capabilities." Trends Cell Biol **10**(5): 189-96.

Osborne, M. A., S. Dalton, et al. (1995). "The yeast tribrid system--genetic detection of trans-phosphorylated ITAM-SH2-interactions." Biotechnology (N Y) **13**(13): 1474-8.

Otsuki, T., S. Kajigaya, et al. (1999). "SNX5, a new member of the sorting nexin family, binds to the Fanconi anemia complementation group A protein." Biochem Biophys Res Commun **265**(3): 630-5.

Pace P, J. M., Tan WM, Mosedale G, Sng C, Hoatlin M, de Winter J, Joenje H, Gergely F, Patel KJ. (2002). "FANCE: the link between Fanconi anaemia complex assembly and activity." EMBO J. **21(13)**: 3414-23.

Pagano, G. and H. Youssoufian (2003). "Fanconi anaemia proteins: major roles in cell protection against oxidative damage." Bioessays **25**(6): 589-95.

Pandolfi, P. P. (2004). "Aberrant mRNA translation in cancer pathogenesis: an old concept revisited comes finally of age." Oncogene **23**(18): 3134-7.

Pang, Q., T. A. Christianson, 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." J Biol Chem **277**(51): 49638-43.

Pang, Q., W. Keeble, et al. (2001). "FANCC interacts with Hsp70 to protect hematopoietic cells from IFN-gamma/TNF-alpha-mediated cytotoxicity." Embo J **20**(16): 4478-89.

Passmore, L. A. and D. Barford (2004). "Getting into position: the catalytic mechanisms of protein ubiquitylation." Biochem J **379**(Pt 3): 513-25.

Pederson, T. (1998). "The plurifunctional nucleolus." Nucleic Acids Res **26**(17): 3871-6.

Pellegrini, L., D. S. Yu, et al. (2002). "Insights into DNA recombination from the structure of a RAD51-BRCA2 complex." Nature **420**(6913): 287-93.

Peng, M., R. Litman, et al. (2007). "The FANCJ/MutLalpha interaction is required for correction of the cross-link response in FA-J cells." Embo J **26**(13): 3238-49.

Pichierri, P. and F. Rosselli (2004). "The DNA crosslink-induced S-phase checkpoint depends on ATR-CHK1 and ATR-NBS1-FANCD2 pathways." Embo J **23**(5): 1178-87.

Prakash R, K. L., Van Komen S, Anke Schürer K, Kramer W, Sung P. (2005). "Saccharomyces cerevisiae MPH1 gene, required for homologous recombination-mediated mutation avoidance, encodes a 3' to 5' DNA helicase." J Biol Chem **280(9)**: 7854-60

Qiao, F., J. Mi, et al. (2004). "Phosphorylation of fanconi anemia (FA) complementation group G protein, FANCG, at serine 7 is important for function of the FA pathway." J Biol Chem **279**(44): 46035-45.

Rackoff, W. D. (1999). Treatment of bone marrow failure. Fanconi Anaemia Standards for Clinical Care. J. Owen, Fanconi Anemia Research Fund, Inc.**:** 9-20.

Reid, S., A. Renwick, et al. (2005). "Biallelic BRCA2 mutations are associated with multiple malignancies in childhood including familial Wilms tumour." J Med Genet **42**(2): 147-51.

Reid, S., D. Schindler, et al. (2007). "Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer." Nat Genet **39**(2): 162-4.
Reuter, T. Y., A. L. Medhurst, et al. (2003). "Yeast two-hybrid screens imply involvement of Fanconi anemia proteins in transcription regulation, cell signaling, oxidative metabolism, and cellular transport." Exp Cell Res **289**(2): 211-21.
Rey, J. P., R. Scott, et al. (1994). "Apoptosis is not involved in the hypersensitivity of Fanconi anemia cells to mitomycin C." Cancer Genet Cytogenet **75**(1): 67-71.
Ridet, A., C. Guillouf, et al. (1997). "Deregulated apoptosis is a hallmark of the Fanconi anemia syndrome." Cancer Res **57**(9): 1722-30.
Rio, P., J. C. Segovia, et al. (2002). "In vitro phenotypic correction of hematopoietic progenitors from Fanconi anemia group A knockout mice." Blood **100**(6): 2032-9.
Rischewski, J. R., H. Clausen, et al. (2000). "A heterozygous frameshift mutation in the Fanconi anemia C gene in familial T-ALL and secondary malignancy." Klin Padiatr **212**(4): 174-6.
Rosenberg, P. S., M. H. Greene, et al. (2003). "Cancer incidence in persons with Fanconi anemia." Blood **101**(3): 822-6.
Rosenberg, P. S., Y. Huang, et al. (2004). "Individualized risks of first adverse events in patients with Fanconi anemia." Blood **104**(2): 350-5.
Rosenberg, P. S., G. Socie, et al. (2005). "Risk of head and neck squamous cell cancer and death in patients with Fanconi anemia who did and did not receive transplants." Blood **105**(1): 67-73.
Rosendorff, J., R. Bernstein, et al. (1987). "Fanconi anemia: another disease of unusually high prevalence in the Afrikaans population of South Africa." Am J Med Genet **27**(4): 793-7.
Rosselli, F., A. Ridet, et al. (1995). "p53-dependent pathway of radio-induced apoptosis is altered in Fanconi anemia." Oncogene **10**(1): 9-17.
Rosselli, F., J. Sanceau, et al. (1994). "Abnormal lymphokine production: a novel feature of the genetic disease Fanconi anemia. II. In vitro and in vivo spontaneous overproduction of tumor necrosis factor alpha." Blood **83**(5): 1216-25.
Rosselli, F., J. Sanceau, et al. (1992). "Abnormal lymphokine production: a novel feature of the genetic disease Fanconi anemia. I. Involvement of interleukin-6." Hum Genet **89**(1): 42-8.
Rossi, F. M., B. T. Blakely, et al. (2000). "Monitoring protein-protein interactions in live mammalian cells by beta-galactosidase complementation." Methods Enzymol **328**: 231-51.
Sasaki, M. S. and A. Tonomura (1973). "A high susceptibility of Fanconi's anemia to chromosome breakage by DNA cross-linking agents." Cancer Res **33**(8): 1829-36.
Schindler, D., M. Kubbies, et al. (1985). "Presymptomatic diagnosis of Fanconi's anaemia." Lancet **1**(8434): 937.
Schmid, W., K. Scharer, et al. (1965). "[Chromosal fragility in familial panmyelopathy (Fanconi type)]." Schweiz Med Wochenschr **95**(43): 1461-4.
Schneiter, R., T. Kadowaki, et al. (1995). "mRNA transport in yeast: time to reinvestigate the functions of the nucleolus." Mol Biol Cell **6**(4): 357-70.
Schroeder, T. M. (1966). "[Cytogenetic and cytologic findings in enzymopenic panmyelopathies and pancytopenias. Familial myelopathy of Fanconi, glutathione-reductase deficiency anemia and megaloblastic B12 deficiency anemia]." Humangenetik **2**(3): 287-316.
Schroeder, T. M. (1966). "[Cytogenetic finding and etiology of Fanconi's anemia. A case of Fanconi's anemia without hexokinase deficiency]." Humangenetik **3**(1): 76-81.
Schroeder, T. M., F. Anschutz, et al. (1964). "[Spontaneous chromosome aberrations in familial panmyelopathy]." Humangenetik **1**(2): 194-6.

Schultz, J. C. and N. T. Shahidi (1993). "Tumor necrosis factor-alpha overproduction in Fanconi's anemia." Am J Hematol **42**(2): 196-201.

Scully, R., J. Chen, et al. (1997). "Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage." Cell **90**(3): 425-35.

Seki, M., F. Marini, et al. (2003). "POLQ (Pol theta), a DNA polymerase and DNA-dependent ATPase in human cells." Nucleic Acids Res **31**(21): 6117-26.

Seyschab, H., R. Friedl, et al. (1995). "Comparative evaluation of diepoxybutane sensitivity and cell cycle blockage in the diagnosis of Fanconi anemia." Blood **85**(8): 2233-7.

Sharan, S. K., M. Morimatsu, et al. (1997). "Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2." Nature **386**(6627): 804-10.

Shen, X., S. Jun, et al. (2006). "REV3 and REV1 play major roles in recombination-independent repair of DNA interstrand cross-links mediated by monoubiquitinated proliferating cell nuclear antigen (PCNA)." J Biol Chem **281**(20): 13869-72.

Shih, H. M., P. S. Goldman, et al. (1996). "A positive genetic selection for disrupting protein-protein interactions: identification of CREB mutations that prevent association with the coactivator CBP." Proc Natl Acad Sci U S A **93**(24): 13896-901.

Shima, N., R. J. Munroe, et al. (2004). "The mouse genomic instability mutation chaos1 is an allele of Polq that exhibits genetic interaction with Atm." Mol Cell Biol **24**(23): 10381-9.

Shimamura, A. and A. D. D'Andrea (2003). "Subtyping of Fanconi anemia patients: implications for clinical management." Blood **102**(9): 3459.

Shimamura, A., R. Montes de Oca, et al. (2002). "A novel diagnostic screen for defects in the Fanconi anemia pathway." Blood **100**(13): 4649-54.

Sims AE, S. E., Sims RJ 3rd, Arita AG, Lach FP, Landers T, Wurm M, Freund M, Neveling K, Hanenberg H, Auerbach AD, Huang TT. (2007). "FANCI is a second monoubiquitinated member of the Fanconi anemia pathway." Nat Struct Mol Biol. **14(6)**: 564-7.

Smogorzewska, A., S. Matsuoka, et al. (2007). "Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair." Cell **129**(2): 289-301.

Sobeck, A., S. Stone, et al. (2006). "Fanconi anemia proteins are required to prevent accumulation of replication-associated DNA double-strand breaks." Mol Cell Biol **26**(2): 425-37.

Sobhian, B., G. Shao, et al. (2007). "RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites." Science **316**(5828): 1198-202.

Socie, G., A. Devergie, et al. (1998). "Transplantation for Fanconi's anaemia: long-term follow-up of fifty patients transplanted from a sibling donor after low-dose cyclophosphamide and thoraco-abdominal irradiation for conditioning." Br J Haematol **103**(1): 249-55.

Stark, R., C. Andre, et al. (1993). "The expression of cytokine and cytokine receptor genes in long-term bone marrow culture in congenital and acquired bone marrow hypoplasias." Br J Haematol **83**(4): 560-6.

Stark, R., D. Thierry, et al. (1993). "Long-term bone marrow culture in Fanconi's anaemia." Br J Haematol **83**(4): 554-9.

Stier, S., T. Cheng, et al. (2002). "Notch1 activation increases hematopoietic stem cell self-renewal in vivo and favors lymphoid over myeloid lineage outcome." Blood **99**(7): 2369-78.

Stiff, T., C. Reis, et al. (2005). "Nbs1 is required for ATR-dependent phosphorylation events." Embo J **24**(1): 199-208.

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

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

Stratton, M. R. and R. Wooster (1996). "Hereditary predisposition to breast cancer." Curr Opin Genet Dev **6**(1): 93-7.

Strauch, E. M. and G. Georgiou (2007). "A bacterial two-hybrid system based on the twin-arginine transporter pathway of E. coli." Protein Sci **16**(5): 1001-8.

Sun, Y., X. Jiang, et al. (2005). "A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM." Proc Natl Acad Sci U S A **102**(37): 13182-7.

Swift, M. R. and K. Hirschhorn (1966). "Fanconi's anemia. Inherited susceptibility to chromosome breakage in various tissues." Ann Intern Med **65**(3): 496-503.

Taniguchi, T., I. Garcia-Higuera, et al. (2002). "S-phase-specific interaction of the Fanconi anemia protein, FANCD2, with BRCA1 and RAD51." Blood **100**(7): 2414-20.

Taniguchi, T., I. Garcia-Higuera, et al. (2002). "Convergence of the fanconi anemia and ataxia telangiectasia signaling pathways." Cell **109**(4): 459-72.

Taniguchi, T., M. Tischkowitz, et al. (2003). "Disruption of the Fanconi anemia-BRCA pathway in cisplatin-sensitive ovarian tumors." Nat Med **9**(5): 568-74.

Thompson D, E. D. (2004). The BRCA1 and BRCA2 genes. Genetic Predisposition to Cancer. E. D. Eeles RA, Ponder BAJ, Eng C. London, Arnold**:** 256-276.

Thompson, L. H., J. M. Hinz, et al. (2005). "How Fanconi anemia proteins promote the four Rs: replication, recombination, repair, and recovery." Environ Mol Mutagen **45**(2-3): 128-42.

Timmers C, T. T., Hejna J, Reifsteck C, Lucas L, Bruun D, Thayer M, Cox B, Olson S, D'Andrea AD, Moses R, Grompe M. (2001). "Positional cloning of a novel Fanconi anemia gene, FANCD2." Mol Cell. **7(2)**: 241-248.

Tirode, F., C. Malaguti, et al. (1997). "A conditionally expressed third partner stabilizes or prevents the formation of a transcriptional activator in a three-hybrid system." J Biol Chem **272**(37): 22995-9.

Tischkowitz, M. D., N. V. Morgan, et al. (2004). "Deletion and reduced expression of the Fanconi anemia FANCA gene in sporadic acute myeloid leukemia." Leukemia **18**(3): 420-5.

Titus, T. A., D. R. Selvig, et al. (2006). "The Fanconi anemia gene network is conserved from zebrafish to human." Gene **371**(2): 211-23.

Tomashek, J. J., J. L. Sonnenburg, et al. (1996). "Resolution of subunit interactions and cytoplasmic subcomplexes of the yeast vacuolar proton-translocating ATPase." J Biol Chem **271**(17): 10397-404.

Toraldo, R., G. Canino, et al. (1998). "Variable response to the diepoxybutane test in two dizygotic twins with Fanconi's anemia and flow cytometry for diagnosis confirmation." Pediatr Hematol Oncol **15**(1): 45-54.

Tremblay, C. S., F. F. Huang, et al. (2008). "HES1 is a novel interactor of the Fanconi anemia core complex." Blood **112**(5): 2062-70.

Uehlinger, E. (1929). "Konstitutionelle infantile (perniciosaartige) Anämie." Klin. Wschr. **82**: 1501-1503.

van der Heijden, M. S., J. R. Brody, et al. (2004). "Functional defects in the fanconi anemia pathway in pancreatic cancer cells." Am J Pathol **165**(2): 651-7.

van der Heijden, M. S., C. J. Yeo, et al. (2003). "Fanconi anemia gene mutations in young-onset pancreatic cancer." Cancer Res **63**(10): 2585-8.

van Zeeburg, H. J., P. J. Snijders, et al. (2004). "Re: Human papillomavirus DNA and p53 polymorphisms in squamous cell carcinomas from Fanconi anemia patients." J Natl Cancer Inst **96**(12): 968; author reply 968-9.

Varnum-Finney, B., L. Xu, et al. (2000). "Pluripotent, cytokine-dependent, hematopoietic stem cells are immortalized by constitutive Notch1 signaling." Nat Med **6**(11): 1278-81.

Venkitaraman, A. (2004). "Tracing the network connecting BRCA and Fanconi anaemia proteins." Nat Rev Cancer. **4**: 266-276.

Verlander, P. C., A. Kaporis, et al. (1995). "Carrier frequency of the IVS4 + 4 A-->T mutation of the Fanconi anemia gene FAC in the Ashkenazi Jewish population." Blood **86**(11): 4034-8.

Vidal, M., R. K. Brachmann, et al. (1996). "Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions." Proc Natl Acad Sci U S A **93**(19): 10315-20.

Vojtek, A. B., S. M. Hollenberg, et al. (1993). "Mammalian Ras interacts directly with the serine/threonine kinase Raf." Cell **74**(1): 205-14.

Wagner, J. E., M. Eapen, et al. (2007). "Unrelated donor bone marrow transplantation for the treatment of Fanconi anemia." Blood **109**(5): 2256-62.

Wagner, J. E., J. Tolar, et al. (2004). "Germline mutations in BRCA2: shared genetic susceptibility to breast cancer, early onset leukemia, and Fanconi anemia." Blood **103**(8): 3226-9.

Waisfisz, Q., J. P. de Winter, et al. (1999). "A physical complex of the Fanconi anemia proteins FANCG/XRCC9 and FANCA." Proc Natl Acad Sci U S A **96**(18): 10320-5.

Wajnrajch, M. P., J. M. Gertner, et al. (2001). "Evaluation of growth and hormonal status in patients referred to the International Fanconi Anemia Registry." Pediatrics **107**(4): 744-54.

Wang, B., S. Matsuoka, et al. (2007). "Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response." Science **316**(5828): 1194-8.

Wang, W. (2007). "Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins." Nat Rev Genet **8**(10): 735-48.

Wang, X., P. R. Andreassen, et al. (2004). "Functional interaction of monoubiquitinated FANCD2 and BRCA2/FANCD1 in chromatin." Mol Cell Biol **24**(13): 5850-62.

Wang, X., R. D. Kennedy, et al. (2007). "Chk1-mediated phosphorylation of FANCE is required for the Fanconi anemia/BRCA pathway." Mol Cell Biol **27**(8): 3098-108.

Whitney, M. A., G. Royle, et al. (1996). "Germ cell defects and hematopoietic hypersensitivity to gamma-interferon in mice with a targeted disruption of the Fanconi anemia C gene." Blood **88**(1): 49-58.

Wong, J. C., N. Alon, et al. (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." Hum Mol Genet **12**(16): 2063-76.

Wooster, R., G. Bignell, et al. (1995). "Identification of the breast cancer susceptibility gene BRCA2." Nature **378**(6559): 789-92.

Xia, B., J. C. Dorsman, et al. (2007). "Fanconi anemia is associated with a defect in the BRCA2 partner PALB2." Nat Genet **39**(2): 159-61.

Xia, B., Q. Sheng, et al. (2006). "Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2." Mol Cell **22**(6): 719-29.

Yabe, M., H. Yabe, et al. (2007). "In vitro effect of fludarabine, cyclophosphamide, and cytosine arabinoside on chromosome breakage in Fanconi anemia patients: relevance to stem cell transplantation." Int J Hematol **85**(4): 354-61.

Yamada, K., J. C. Olsen, et al. (2001). "Functional correction of fanconi anemia group C hematopoietic cells by the use of a novel lentiviral vector." Mol Ther **3**(4): 485-90.

Yamada, K., A. Ramezani, et al. (2003). "Phenotype correction of Fanconi anemia group A hematopoietic stem cells using lentiviral vector." Mol Ther **8**(4): 600-10.

Yamamoto, K., M. Ishiai, et al. (2003). "Fanconi anemia FANCG protein in mitigating radiation- and enzyme-induced DNA double-strand breaks by homologous recombination in vertebrate cells." Mol Cell Biol **23**(15): 5421-30.

Yamashita, T., G. M. Kupfer, et al. (1998). "The fanconi anemia pathway requires FAA phosphorylation and FAA/FAC nuclear accumulation." Proc Natl Acad Sci U S A **95**(22): 13085-90.

Yang, H., P. D. Jeffrey, et al. (2002). "BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure." Science **297**(5588): 1837-48.

Yang, Y., Y. Kuang, et al. (2001). "Targeted disruption of the murine Fanconi anemia gene, Fancg/Xrcc9." Blood **98**(12): 3435-40.

Yin, J., A. Sobeck, et al. (2005). "BLAP75, an essential component of Bloom's syndrome protein complexes that maintain genome integrity." Embo J **24**(7): 1465-76.

Yu, X., J. K. Alder, et al. (2006). "HES1 inhibits cycling of hematopoietic progenitor cells via DNA binding." Stem Cells **24**(4): 876-88.

Yuan, S. S., S. Y. Lee, et al. (1999). "BRCA2 is required for ionizing radiation-induced assembly of Rad51 complex in vivo." Cancer Res **59**(15): 3547-51.

Zhu, H., M. Bilgin, et al. (2003). "Proteomics." Annu Rev Biochem **72**: 783-812.

Zhu, H. and M. Snyder (2002). ""Omic" approaches for unraveling signaling networks." Curr Opin Cell Biol **14**(2): 173-9.

Zhu, W. and A. Dutta (2006). "An ATR- and BRCA1-mediated Fanconi anemia pathway is required for activating the G2/M checkpoint and DNA damage repair upon rereplication." Mol Cell Biol **26**(12): 4601-11.