CHAPTER ONE:

INTRODUCTION

1.1 MOLECULAR MECHANISMS OF CANCER

1.1.1 Oncogenes

Cancer results from genetic mutations which disrupt the balance of cellular regulation, and is characterized by uncontrolled growth of mutated cells. Principal insights into the molecular mechanisms involved in cancer began with the identification of genes known as oncogenes; activated when mutated or overexpressed, oncogenes cause uncontrolled growth of cells, or reactivation of quiescent cells. These mutations are dominant, and generally are acquired somatically.

Studies with animal tumour viruses, particularly retroviruses, were instrumental in the discovery of oncogenes. Tumourigenic retroviruses can transform cells by modifying the activity of endogenous genes in various ways: direct mutation through proviral insertion, activation of a gene by the promoter and enhancer carried by a viral long terminal repeat (LTR), or through a non-essential, cell-derived, and mutated oncogene carried by the virus. For example, in tumours arising in mice infected with mouse mammary tumour virus (MMTV), the provirus often has affected the mouse oncogene *Wnt-1* (Nusse, 1984; Nusse, 1991). On the other hand, a transforming retrovirus such as the Rous sarcoma virus carries an oncogene (in this case *v-SRC*) derived from a normal, non-transforming chicken gene (*c-SRC*), which became part of the viral genome at some point in the past (Stehelin, 1976; Rous, 1983).

DNA tumour viruses transform cells through the interaction of a viral gene product with the host cell. For instance, the polyomavirus simian virus 40 (SV40) transforms through the action of its large T antigen (tumour antigen) gene product, which interacts with endogenous, host-derived regulatory proteins such as the retinoblastoma protein (RB1) and p53 (Lane and Crawford, 1979). Studies of viral-induced tumourigenesis contributed to the

realization that misregulation of either endogenous or foreign genes can result in uncontrolled cell growth.

1.1.2 Tumour Suppressors

The existence of negative factors in human cancer was postulated before the discovery of oncogenes, but the principal lines of evidence demonstrating the existence of tumour suppressor genes (TSGs) emerged later. Studies showed that fusions of tumourigenic and non-tumourigenic cells resulted in a non-tumourigenic cell, and that suppression of tumourigenesis depended on the presence of certain chromosomes (Harris, 1969). Working from a very different perspective, Alfred Knudson developed the two-hit model of carcinogenesis while studying familial retinoblastoma. Using statistical methods, he predicted that two mutations or "hits" were required for tumourigenesis in this syndrome, suggesting that retinoblastoma resulted not from the presence of an oncogene, but from the loss of what is now known as a tumour-suppressor gene (Knudson, 1971). Thus, in familial retinoblastoma, the first hit is an inherited germline mutation of one TSG allele. Loss of the second allele – or loss of heterozygosity (LOH) – by somatic mutation or chromosomal rearrangement then leads to tumourigenesis (reviewed in Knudson, 2000). In the case of sporadic retinoblastoma, both mutations are somatically acquired. This two-hit model is now applied to tumour-suppressor–related carcinogenesis in general.

More recently, TSGs have been divided into two broad classes based on their mechanism of action: "gatekeepers", or genes which inhibit the growth of tumours or promote apoptosis, and "caretakers", which regulate cellular processes that repair genetic lesions and maintain the overall genetic integrity of each cell (Kinzler and Vogelstein, 1997). The role of caretakers in tumourigenesis is less direct than that of gatekeepers: loss of genetic integrity in a cell which has lost both copies of a caretaker gene leads to the mutation of additional genes, which leads to cancer. This mechanism has been used

to explain why the mutation of some tumour-suppressors is rate-limiting for tumourigenesis, while mutations in others are not (Kinzler and Vogelstein, 1997). In the case of retinoblastoma, the gene involved, *RB1*, acts as a gatekeeper gene; it is directly involved in cell-cycle progression, and mutation of the second allele of *RB1* is the rate-limiting step in carcinogenesis (Knudson, 1971). On the other hand, inherited mutations in caretaker tumour suppressors lead to cancer by predisposing to secondary mutations through genomic instability. These secondary mutations are the rate-limiting step in carcinogenesis. Regardless of which type of gene is mutated, most studies agree that a number of mutations are generally required for progression to tumourigenesis (Knudson, 2001).

Cancer is one of the leading causes of death in the Western world (Centers-for-Disease-Control, 2000; Cancer-Research-UK, 2001). While much has been done to advance prevention, diagnosis, and treatment of various forms of cancer, there is still much work to be done toward understanding the process of tumourigenesis at the molecular level and the functions played by the various genes implicated in familiar cancer syndromes. The use of model systems, especially the mouse, has been instrumental in our understanding of cancer to date.

1.1.3 Mice as Models for Cancer

Before the rediscovery of Mendel's laws in 1900, mice were not extensively used in biological research, though mouse fanciers collected and bred the animals. Inbred strains of mice (generated by performing at least twenty generations of brother-sister matings) were developed to allow scientific study of animals with genetically homogeneous backgrounds. During the generation of the first inbred strains, it was noticed that some strains showed a high incidence of certain cancers – for example, the 129 strain was prone to testicular cancer (Stevens, 1970). Studies involving mutagenic agents identified a variety of chemical and physical methods which could be used to

reliably induce or enhance tumourigenesis in these mouse strains, though little was known about the specific genetic targets of these agents. To move beyond these molecularly uncharacterized and random methods of mutagenesis, techniques to stably and precisely introduce specific mutations into the mouse genome were needed. Transgenic mouse and embryonic stem (ES) cell technologies have met this need.

1.1.3.1 Transgenic and embryonic stem cell technologies

Transgenesis, the stable introduction of an exogenous gene into the germline of an organism (usually by pronuclear injection), was first demonstrated in the 1980s (Gordon and Ruddle, 1981). The first transgenic mouse models of cancer overexpressed viral oncogenes. More recently, models have been generated in which the activity of a transgene can be controlled temporally and spatially. This has enabled the production of mouse models that more accurately reflect the cascade of genetic events that characterize human malignancies (for a recent review, see Thompson, 2004).

ES cell technology provides a mechanism for the study of endogenous gene function by precisely targeting mutations into a gene or region of choice (Bradley, 1992; Bradley, 1998). This technology arose from two lines of research conducted in the 1980s.

First, it was demonstrated that pluripotent cells, later named ES cells, could be isolated from day 3.5 embryos and cultured (Evans and Kaufman, 1981). When reintroduced into blastocyst-stage embryos and implanted into host pseudo-pregnant females, these cells could repopulate the embryo, including the germ cell lineage, resulting in chimæric mice (Bradley, 1984). In these experiments, the host blastocysts were derived from unpigmented albino mice and the pluripotent cells from pigmented black agouti mice, meaning that the resulting chimæras (and the percentage of chimærism) were easily distinguishable. When chimæras are mated to albino mice, germline transmission of genetic material from the ES cells can be assessed by coat

colour: resulting progeny consist of albino mice and black agouti mice, half of which carry the targeted mutation (Bradley, 1984; Ramirez-Solis, 1993).

Second, it was demonstrated that ES cells could be modified in culture before being reintroduced into embryos to generate chimæras (Robertson, 1986; Thomas, 1986). The introduction of mutations or genomic changes into ES cells is generally accomplished through homologous recombination of a transfected targeting vector and the endogenous target genomic region (although other methods, such as retroviral-mediated insertion to generate mutations, also exist). Two types of targeting vectors are used (Figure 1.1a) (reviewed in Hasty, 2000). Insertion vectors carry one region of homology, require one homologous crossover event, and result in genomic integration of the entire targeting vector and a duplication of the region of homology. Replacement vectors carry two regions of homology and require two homologous recombination events for integration. Successful targeting of these vectors results in a specific change – be it a deletion, insertion, or mutation – of a region of the genome. Often, a vector will add a selectable marker to the successfully targeted area, for selection of targeted cells in culture. To date, targeting vectors and ES cell technology have been used to generate mutations in over 3,000 mouse genes (BioMedNet, 2003). ES cell technology can be used to mimic virtually any change in the genome, from single base pair mutations (Hasty, 1991) to large chromosomal deletions and inversions (Ramirez-Solis, 1995; Zheng, 1999a; Zheng, 1999b), and to generate a mouse model carrying the change.

1.1.3.2 Conditional Mutations and the Cre-loxP system

The coupling of site-specific recombinase systems with ES cell targeting vector technology has further refined the art of generating mouse models (reviewed by Kwan, 2002). A conditional targeting vector is very similar to a standard replacement vector, except that instead of replacing a region of the gene of interest, recognition sites for site-specific recombinases are inserted around the region (Figure 1.1b). Recombinase is then used to catalyze the

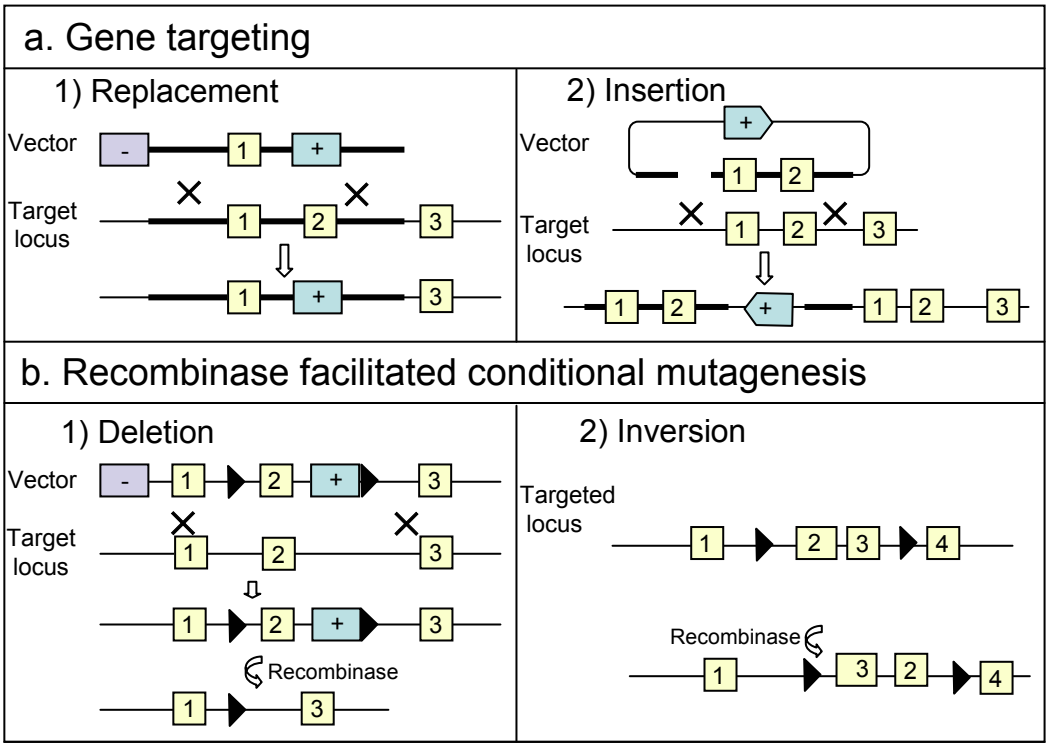


Figure 1.1: Gene-targeting vectors and recombinase-mediated mutagenesis. a. Replacement and insertion targeting vectors. **b.** Conditional targeting vectors, showing that deletion or inversion of the area will occur depending on the orientation of the recombinase recognition sites (black triangles). Figure taken from (Thompson, 2004).

deletion or inversion of the flanked area. The most commonly used recombinase system is the Cre/*loxP* system from bacteriophage P1. The *loxP* (*locus of crossover, P1*) recombinase recognition site is a 34 bp directional sequence recognized by the cyclization recombination (Cre) protein (Sauer and Henderson, 1988). The orientation of *loxP* sites is important; Cre-mediated recombination of two *loxP* sites flanking a region deletes the region if the *loxP* sites are in the same orientation, but inverts it if they are in opposite orientations. The Flp/frt recombination system, derived from the yeast *Saccharomyces cerevisiae,* has a similar mechanism of action, but is less widely used at present (Dymecki, 1996).

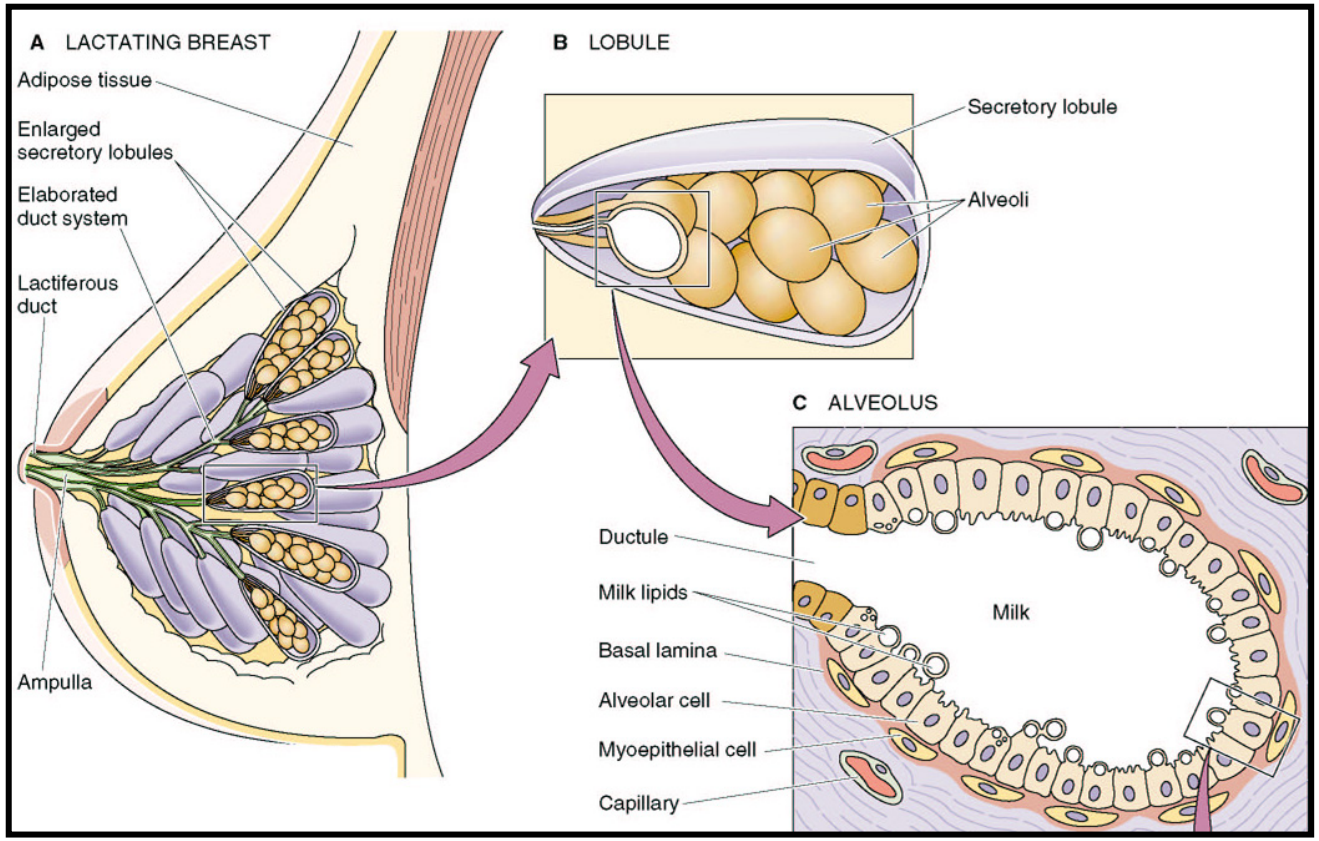
Such recombinase systems have proved to be very useful, both as an aid to circumventing embryonic lethality of a homozygous knockout and in generating more accurate models of human diseases. For example, when modeling human cancers, the optimal mouse model should reflect the biological, genetic, aetiological, and therapeutic aspects of the human cancer it models. High penetrance and short latency of tumourigenesis are desirable, because of the short mouse lifespan (Hann and Balmain, 2001). A common criticism of mouse models of human familial lesions is that the tumour spectra do not always mirror those of the human (Jacks, 1996), but some of the recently-reported conditional mutations are more faithful models of the human situation than are standard knockouts of the same genes (Shibata, 1997; Giovannini, 2000). The use of tissue-specific or adenoviral mechanisms of Cre delivery makes it possible to investigate the effects of a specific genetic lesion only in relevant tissues, or at a particular time. Furthermore, as recombinase efficiency is rarely 100%, conditional mutants provide a mechanism for modeling both the random nature of mutagenesis in human cancers and the microenvironment of a mutant cell amongst normal ones.

1.2 BREAST CANCER

1.2.1 Human breast cancer: a brief overview

One in nine women in the United Kingdom is predicted to develop breast cancer within her lifetime (Cancer-Research-UK, 2003). When considered in light of the impact that this disease has, not only on patients, but also on their family and friends, it is a very relevant topic of study and understandably the focus of much research. The mammary gland has a unique physiology in that it may undergo several rounds of growth, terminal differentiation, and regression during development and multiple pregnancies. The primary breast architecture is laid down during development, and further development occurs during puberty. Hormonal signals during pregnancy trigger large amounts of growth and differentiation to facilitate milk production; cessation of breast-feeding at weaning signals the regression of much of this growth (Figure 1.2d). The breast itself consists of around twenty lobes, each of which has a branching structure of ducts leading to ductules leading to lobules. The lobules contain the alveoli (or acini), where milk-producing cells are located. Epidermal cells lining the alveoli are hormonally stimulated to produce and secrete milk proteins. Milk then travels down the duct system to the nipple (Figure 1.2).

The two major forms of breast carcinoma are classified as ductal and lobular carcinomas, although many texts suggest that the majority of breast carcinomas arise from the terminal ductule-lobule units, regardless of their classification (Aldaz, 2002; Bulpaep, 2003). Others suggest that cancer arises from either mammary stem cells or ductal progenitor cells, as the majority of resulting cancers are ductal in nature (Medina, 2002). This stem cell theory may be supported by a recent finding that only a subset of cells within a breast tumour, which can be segregated using cell-surface markers, are tumourigenic when injected into nude mice (Al-Hajj, 2003).



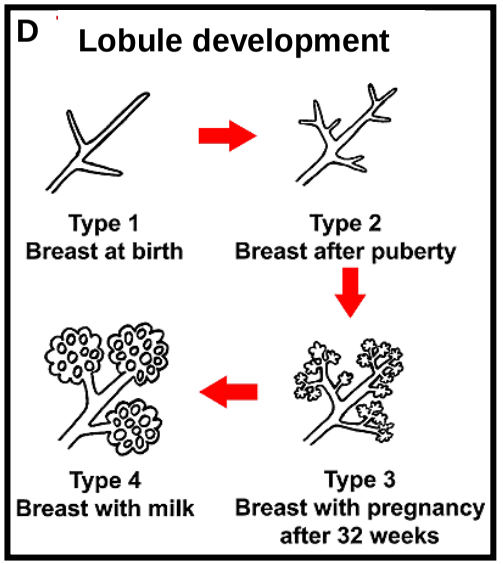


Figure 1.2: Architecture of the human breast, milk production, and breast development.
a. Schematic showing the branching duct-to-lobule structure of the breast during lactation. **b.** The lobule contains the alveoli (also called acini), where the milk-secreting cells are located. **c.** The epidermal cells lining the alveoli secrete the milk. **d.** Development of branching duct-lobule structure; pregnancy triggers the branching and differentiation of ducts and lobules in preparation for milk production.
Figures taken from (Bulpaep, 2003; Brind, 2002).

1.2.2 Hereditary breast cancer and the familial predisposition gene *BRCA1*

That breast cancer might have a hereditary component became clear as early as the mid-nineteenth century. While mutations in familial predisposition genes account for less than 10% of all human breast cancer cases, an estimated 40% of early-onset (before the age of 30) cases are attributable to such mutations (Claus, 1991). The normal lifetime risk of breast cancer by the age of 70 for women in the UK is ~6.7%, but carriers of a mutated predisposition gene have a increased lifetime risk of breast cancer of ~65% (estimates range from 35.3% to 70%) by age 70 (Easton, 1993a; Antoniou, 2002; Antoniou, 2003; Cancer-Research-UK, 2003; King, 2003). A large-scale study undertaken by the Centers for Disease Control suggested an autosomal dominant form of inheritance for familial breast cancers (Claus, 1991).

In 1990, Mary-Claire King's group at the University of California at Berkeley reported a linkage between a region on human chromosome 17q21 with the causal mutation carried by a number of breast-cancer families in which early-onset breast cancer was common (Hall, 1990). In the four years between the appearance of this paper and the successful cloning of the gene known as *BRCA1* (*Breast Cancer 1*), a host of studies were published which provide a picture of positional cloning methods in the pre-genome–sequence era. The region on 17q was the target for the development of radiation hybrid maps (Abel, 1993; Black, 1993) and the assembly of high densities of markers for analysis (Anderson, 1993). Yeast artificial chromosome (YAC) and P1 contigs were used for physical mapping and identification of candidate genes (Albertsen, 1994). Numerous groups reported collections of breast-cancer families linked to the slowly narrowing target region (Devilee, 1993; Easton, 1993b; Spurr, 1993). In 1994, *BRCA1* was identified and its sequence published (Miki, 1994). Confirmation from other groups that this newly identified gene was indeed mutated in families with breast or breast and ovarian cancers swiftly followed (Friedman, 1994; Futreal, 1994). Evidence that *BRCA1* is a tumour suppressor gene emerged even before it was cloned.

In a study of breast and ovarian tumours from families whose disease was linked to chromosome 17q, a majority had lost heterozygosity of the wildtype chromosome in the region where *BRCA1* was believed to be located (the second hit, according to the Knudson hypothesis (Knudson, 1971)) (Smith, 1992).

BRCA1 spans approximately 100 kilobases (kb) of genomic sequence, and is composed of 23 exons which encode a 7.8 kb mRNA. Exon 1 is non-coding, and the region originally identified as exon 4 is an Alu repeat not generally included in the transcript (Miki, 1994). The 220 kiloDalton (kDa) protein shows a predominantly nuclear localization, and forms nuclear "dots," or foci, during S phase of the cell cycle and following DNA damage (Chen, 1995; Scully, 1996; Scully, 1997b). Although initial characterization indicated that the protein had no significant homology to any other sequences in the databases, it does have two features: an N-terminal zinc (Zn)-finger domain, known as a RING-finger, and a C-terminal region with two tandem repeats of a small domain known as a BRCA1 C-terminal, or BRCT, motif (Koonin, 1996). BRCT motifs are also found in other proteins involved in DNA repair or cell-cycle control such as p53 binding-protein 1, XRCC1, and RAD9 (Koonin, 1996; Bork, 1997; Callebaut and Mornon, 1997).

1.2.3 Other familial breast-cancer predisposition genes

While the hunt for *BRCA1* was ongoing, one group noticed that the susceptibility gene in some of the breast-cancer families they analyzed mapped to a different region. This region, located on chromosome 13q, was proposed to harbour a second predisposition gene, *BRCA2* (*Breast Cancer 2*) (Wooster, 1994). Positional cloning techniques were used initially to locate the *BRCA2* gene, but *BRCA2* was ultimately identified using newly-released sequence data from the Human Genome Project. Only fifteen months after publishing the original linkage paper, the same group confirmed that mutations in *BRCA2* were indeed present in the breast-cancer families linked to chromosome 13q (Wooster, 1995). Like BRCA1, the BRCA2 protein was

not similar to other proteins in the databases and had neither a RING domain nor BRCT repeats (Tavtigian, 1996).

More recently, a mutation in a third gene, *CHEK2*, was linked with a small number of familial breast cancer cases (Meijers-Heijboer, 2002; Vahteristo, 2002). Other genes linked to a higher incidence of breast cancer include *p53* and *ATM*, the gene mutated in the human disorder ataxia-telangiectasia. It is likely that other predisposition genes exist, but are either rare or show low penetrance (Antoniou, 2002; Wooster and Weber, 2003).

1.3 CHARACTERISTICS OF *BRCA1*-RELATED HUMAN TUMOURS

1.3.1 Cancer-related *BRCA1* mutations and the risk they confer

According to two recent analyses of *BRCA1*-related breast-cancer families, *BRCA1* mutation carriers have a ~65% overall risk of breast cancer by the age of 70, a 14- to 30-fold increased risk of breast cancer (depending on age) relative to that of non-carriers (Antoniou, 2003; King, 2003). *BRCA1* mutations also confer an increased risk of other cancers, most notably ovarian cancer (~40% overall risk by the age of 70). This will be discussed further in section 1.3.3.

Cancer-related *BRCA1* mutations include small insertions or deletions which cause frameshifts, and nonsense or missense mutations; premature truncation of the mutant protein is common (Friedman, 1994; Gayther, 1995). For a detailed and up-to-date listing of *BRCA1* mutations, see the Breast Cancer Information Core website (BIC, 2003)). Generally, *BRCA1* does not have mutation “hotspots” (unlike the TSG *p53* (Walker, 1999)), although some mutations are more commonly observed, such as the missense mutation C61G, a mutation in the highly-conserved RING domain, or the frameshift mutations 186delAG (in exon 2) and 5382insC, located in one of the BRCT repeats (BIC, 2003). Intronic or exonic mutations which change the normal splice pattern have also been identified (Gayther, 1995; Xu, 1997b). Cancer-related mutations occur all throughout *BRCA1*, from the 5’ end down to a

mutation which gives rise to a protein lacking only the last few amino acids (BIC, 2003). While there does not appear to be a bias in mutation position in regards to tumourigenesis in general, the frequency of ovarian cancers relative to breast cancers is higher in families carrying mutations in the middle of the gene (mostly in exon 11) than at either of the ends (Gayther, 1995; Thompson and Easton, 2002).

1.3.2 *BRCA1* and sporadic breast and ovarian cancers

A number of studies have shown that BRCA1 protein and/or mRNA appears to be down-regulated in some sporadic breast cancers, regardless of their stage of progression (Thompson, 1995; Magdinier, 1998; Rio, 1999; Baldassarre, 2003). The exact cause of this loss of expression is unknown. A percentage of tumours with decreased *BRCA1* expression have undergone LOH at the *BRCA1* locus, but others have not – and some samples with LOH express *BRCA1* at normal levels (Thompson, 1995; Sourvinos and Spandidos, 1998; Rio, 1999; Staff, 2003). In general, LOH at the *BRCA1* locus in sporadic breast cancers does not appear to be critical for tumourigenesis, which is not surprising, as LOH of *BRCA1* in sporadic tumours would not be expected to have a detrimental effect unless one allele was previously modified or mutated (Futreal, 1994; Merajver, 1995). While promoter hypermethylation could provide such a modification, only a small subset of the sporadic breast cancers studied to date show aberrant methylation of the *BRCA1* promoter region (Magdinier, 1998; Rice, 1998). In familial breast cancers, a small study has demonstrated that promoter methylation rarely serves as the second *BRCA1* hit, with LOH the more common mechanism (Esteller, 2001). Although loss of *BRCA1* expression may contribute to sporadic breast cancer, the mechanism behind this downregulation, or the roles it plays in tumourigenesis, have yet to be determined.

1.3.3 *BRCA1*-related, *BRCA2*-related, and sporadic breast cancers

The phenotypic consequences of a *BRCA1* mutation differ from those of a corresponding *BRCA2* mutation (Tables 1.1 and 1.2). The overall risk of breast cancer for carriers of *BRCA2* mutations has been estimated at 45–74% by the age of 70, but their risk for ovarian cancer (~11% overall risk by the age of 70) is lower than that of carriers of *BRCA1* mutations (Antoniou, 2003; King, 2003). Male carriers of *BRCA2* mutations have a ~7% overall risk of breast cancer by the age of 80; mutations in *BRCA2* are predicted to account for approximately 10% of all male breast cancers (Thompson and Easton, 2001). Male breast cancer is not common in *BRCA1* families (Antoniou, 2003). Small increased risks for other cancers have been reported for both genes (Table 1.1) (Consortium, 1999; Brose, 2002).

BRCA1-related human breast tumours are also pathologically distinct from both *BRCA2*-related or sporadic breast tumours (Lakhani, 1998; Armes, 1999). Table 1.2 reports some of the differences between *BRCA1*- or *BRCA2*-related and sporadic breast tumours, including common secondary mutations, responsiveness to hormones, and *p53* mutation status. *p53* mutations are observed more frequently and tend to occur at less commonly-mutated sites in *BRCA1* and *BRCA2* breast tumours than in sporadic breast tumours (Ramus, 1999; Greenblatt, 2001). *BRCA1*-related ovarian cancers are also more likely to carry a mutation in *p53*, although the overall *p53* mutation spectrum in these tumours is very similar to that of sporadic ovarian cancers (Buller, 2001). Additionally, while the pathology of *BRCA1*-related and sporadic breast tumours differ, the pathology of *BRCA1*-related and sporadic ovarian cancers is very similar (Rubin, 1996).

1.4 *Brca1* – the mouse homologue of *BRCA1*

The murine homologue of *BRCA1* (*Brca1*) has 23 coding exons which encode a ~7.2 kb mRNA and a protein of 1812 amino acids (aa) (Abel, 1995; Bennett, 1995; Lane, 1995; Sharan, 1995; Schrock, 1996a). Its location on mouse chromosome 11 correlates with earlier studies which revealed a large linkage

Table 1.1: Cancer risks for carriers of deleterious mutations in *BRCA1* or *BRCA2.*

Characteristic	*BRCA1*	*BRCA2*	References
Risk of breast cancer (by age 70)	~65%	45%[$]	(Johannsson, 1997; Armes, 1999; Lakhani, 2002; King, 2003)
Risk of ovarian cancer (by age 70)	~40%	~11%	(Johannsson, 1997; Armes, 1999; Lakhani, 2002; King, 2003)
Risk of male breast cancer (by age 80)	Not generally observed	7%	(Ramus, 1999; Greenblatt, 2001)
Increased risk of other cancers[*]	Colon, pancreatic, gastric, fallopian tube (female)	Pancreatic, gall bladder, stomach, malignant melanoma, prostate (male)	(Consortium, 1999; Brose, 2002; Thompson and Easton, 2002; Edwards, 2003)

* Observed incidence of other cancers was low; generally the risk relative to that of non-carriers was 2- to 4-fold higher.
[$] Estimated at 74% in (King, 2003).

Table 1.2: *BRCA1* and *BRCA2*–related breast tumours: characteristics compared to sporadic breast tumours.

Characteristic	*BRCA1*	*BRCA2*	Sporadic	References
Loss of oestrogen receptor (ER) activity	90%	34%	35%	(Johannsson, 1997; Armes, 1999; Lakhani, 2002)
Loss of progesterone receptor (PR) activity	79% (or higher)	45%	41%	(Johannsson, 1997; Armes, 1999; Lakhani, 2002)
p53 mutations	More common			(Ramus, 1999; Greenblatt, 2001)
p53 mutated at "hotspots *"	67%		82%	(Greenblatt, 2001)
Multiple *p53* mutations	16%		4%	(Greenblatt, 2001)
Loss of *ErbB2/neu/Her-2* expression	97%	97%	85%	(Lakhani, 2002), supported by (Johannsson, 1997; Armes, 1999)
Loss of Cyclin D expression	Lost in 6 of 9 cases (small study)	Lost in 5 or 9 cases (small study)	Retained in 19 control cases (small study)	(Armes, 1999)
Amount of LOH	Increased			(Tirkkonen, 1997)
LOH hotspots	5q, 4q		8p, 16q	(Tirkkonen, 1997)
Aneuploidy	More common			(Johannsson, 1997)

*p53 codons frequently mutated in human cancers (Walker, 1999).
LOH=loss of heterozygosity.

group of over 50 centiMorgans (cM) shared between human chromosome 17, mouse chromosome 11, and rat chromosome 10 (Remmers, 1992; Yamada, 1994). Overall, the BRCA1 mouse-human protein identity is 57%, but there are two regions of very high homology: the N-terminal region (97% similar; within the RING motif, the identity is 100%) and the C-terminal BRCT repeats (83% identical), which underscores the importance of these two domains (Figure 1.3) (Sharan, 1995). Other mammalian *BRCA1* homologues from the rat (Bennett, 1999) and dog (Szabo, 1996) also demonstrate this low overall conservation but high conservation at the two terminal domains (Szabo, 1996).

One feature of *BRCA1* which differs between the mouse and human homologues is the 5' untranslated region (5' UTR). While both genes are TATA-less (Rice, 1998) and appear to share a bidirectional promoter with another, head-to-head oriented gene, the human *BRCA1* 5' region is more complex (Figure 1.4). Mouse *Brca1* lies head-to-head with a gene called *Nbr1* (*Neighbour of Brca1 1*) (Chambers and Solomon, 1996; Dimitrov, 2001), but the human 5' region contains, in addition to a head-to-head copy of *NBR1* (*Neighbour of BRCA1 1*, also known as *M17S2*, *Membrane component, Chromosome 17, Surface marker 2*) (Campbell, 1994), two differentially spliced copies of exon 1 of *BRCA1* (Xu, 1995), a second gene called *NBR2* (*Neighbour of BRCA1 2*) (Xu, 1997a), and a few *BRCA1* pseudo-exons. This additional complexity likely results from a partial duplication of the region (Brown, 1996). The rat 5'UTR is very similar to that of the mouse (Bennett, 1999).

1.5 FUNCTIONAL CHARACTERISTICS OF *BRCA1*

1.5.1 Tissue expression profile of *BRCA1* in mice and humans

BRCA1 is expressed in many human tissues, including breast and ovary, thymus, kidney, and testis (Miki, 1994). In human testes, both *BRCA1* and *BRCA2* are highly expressed in zygotene and pachytene spermatocytes (Scully, 1997c; Chen, 1998). Murine *Brca1* transcript is also detected in testis

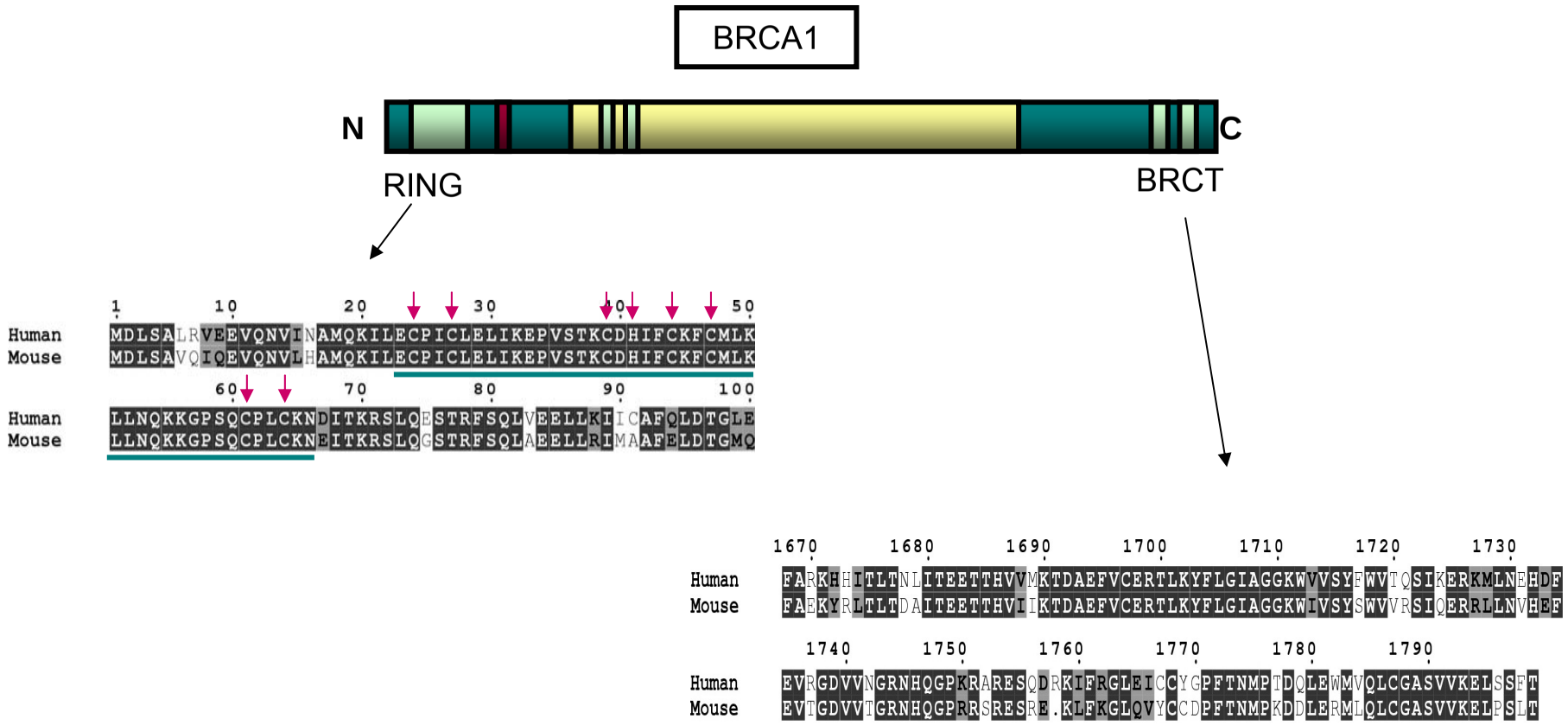


Figure 1.3: Conservation of mouse and human BRCA1 proteins. Schematic of BRCA1 showing the conserved RING domain and BRCT repeats. The yellow box represents exon 11, the largest coding exon. Boxes show the human-mouse protein alignments for the indicated regions. For the RING domain, the key cysteine (C) and histidine (H) residues are indicated by red arrows and the canonical RING motif is underlined in green. Identical residues are highlighted in dark grey, similar ones in light grey. ClustalW and ESPript (Blosum62 matrix) were used to produce the alignments.

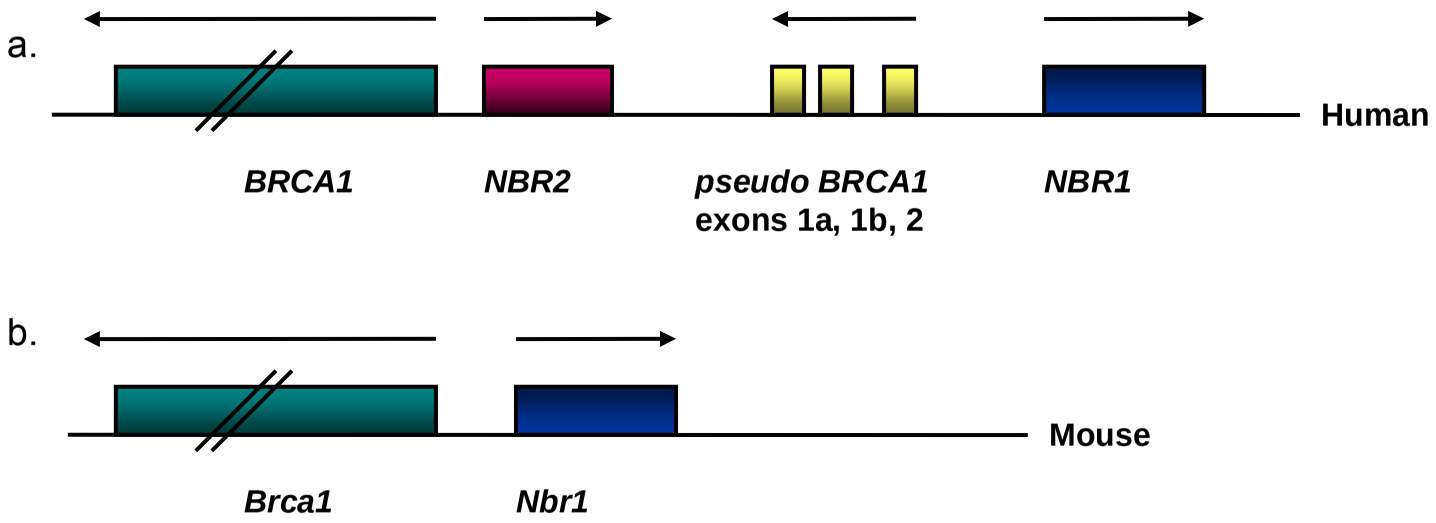


Figure 1.4: Organization of the genomic region upstream of mouse and human *BRCA1*.
a. The genomic region upstream of human *BRCA1* contains two other genes and three *BRCA1* pseudo-exons. **b**. The murine upstream genomic region is less complex. Information derived from mouse and human Ensembl and genome sequences. Arrows indicate the direction of transcription of each gene. The *NBR* genes do not share sequence homology.
Abbreviations: *NBR2*=human *Neighbour of BRCA1 2* gene, *NBR1*=human *Neighbour of BRCA1 1* gene, *Nbr1*=mouse *Neighbour of Brca1 1* gene.

(Lane, 1995; Marquis, 1995). Consistent with the human expression profile, *Brca1* is highly expressed in pachytene spermatocytes and round spermatids, and expression of the transcript increases as the testes develop (Zabludoff, 1996), suggesting that BRCA1 plays a role in meiosis. Supportingly, male infertility has been observed in *Brca1* mice homozygous for either an allele lacking all of exon 11 or one which carries a truncated protein. In both models, spermatogenesis arrested during the prophase stage of meiosis I, although one mutant was only examined in the context of a $p53^{+/-}$ or $p53^{-/-}$ background (Cressman, 1999a; Ludwig, 2001; Xu, 2003).

In the mouse, *Brca1* is widely expressed during development, but in adult animals it appears to be expressed mainly in proliferating cell types involved in differentiation (Marquis, 1995). *Brca1* is expressed in the epithelial cells of the breast, and its expression is increased during pregnancy and lactation, especially in the rapidly growing and differentiating terminal end buds and alveoli. Increased expression of *Brca1* in breast tissue can be induced by oestrogens and progesterone, although this induction may be an indirect effect, as *Brca1* appears to be expressed in growing cells and hormone signaling results in increased proliferation of breast tissue (Lane, 1995; Marquis, 1995). This supposition is supported by *in vitro* studies using human oestrogen-responsive breast cancer cell lines, which demonstrate that upregulation of *BRCA1* expression is delayed by nearly 24 hours following oestrogen stimulation, suggesting that oestrogen does not directly upregulate *BRCA1* (Gudas, 1995; Spillman and Bowcock, 1996; Marks, 1997). The role of BRCA1 in differentiation has been suggested by *in vitro* studies, as a mammary epithelial call line can be induced to differentiate by ectopically expressing *BRCA1* in conjunction with hormonal triggers (Kubista, 2002).

1.5.2 Expression of alternative forms of *BRCA1*

Screening of *BRCA1*-related breast tumours has identified several splice aberrations which appear to be associated with tumourigenesis (Xu, 1997b). Analyses of *BRCA1* transcripts in normal cells has demonstrated that *BRCA1* may normally be expressed in more than one form. Two alternative forms of

BRCA1 have been described which lack all (∆*X.11*) or most (∆*X.11b*) of exon 11, the largest exon (Thakur, 1997; Wilson, 1997). Both of these isoforms were identified by reverse transcription of human cellular RNA followed by amplification by the polymerase chain reaction (RT-PCR) using *BRCA1*-specific primers. The mouse has a similar (single) natural ∆*X.11* splice isoform, and this conservation suggests that the full-length and ∆*X.11* forms of *BRCA1* may both be biologically relevant, perhaps with different roles in the cell (Xu, 1999c; Bachelier, 2000).

These ∆*X.11* splice isoforms will be discussed throughout this chapter, as they have proved useful in studying the functions of BRCA1. When expressed at physiological levels, BRCA1 ∆X.11 proteins form S phase and damage-induced nuclear foci as does full-length BRCA1 (Xu, 1999c), and are recognized by BRCA1 antibodies raised to the N- or C-terminal ends of the protein. However, a glance at Figure 1.5 (exon 11 is depicted in yellow) shows that exon 11 codes for the part of the BRCA1 protein thought to be important for interaction with proteins such as RAD51 and BRCA2, and thus this isoform will likely not participate in RAD51/BRCA2-related functions, thought to be important for DNA repair. This supposition is supported by experiments using mouse embryonic fibroblasts (MEFs) generated from mice expressing only the *Brca1* ∆*X.11* isoform. These cell lines have a defective G2-M checkpoint and are more likely to have extra centrosomes than wildtype MEFs (Xu, 1999c). Furthermore, ectopic expression of *BRCA1* ∆*X.11* in a mammary epithelial cell line will not induce differentiation while full-length *BRCA1* will (Kubista, 2002).

Besides the ∆*X.11* splice isoforms, human *BRCA1* also has two alternative first exons, exon 1a and 1b (Xu, 1995). Other reports suggest that a third first exon may exist, a truncated form of 1a called 1a' (Hsu, 2001; Jakubowska, 2001). Murine *Brca1* appears to have only one exon 1. In both human and mouse, exon 2 contains the translational start site, so the alternative use of first exons in human cells is postulated to have some regulatory role. It has further been proposed that human *BRCA1* is expressed from two alternative

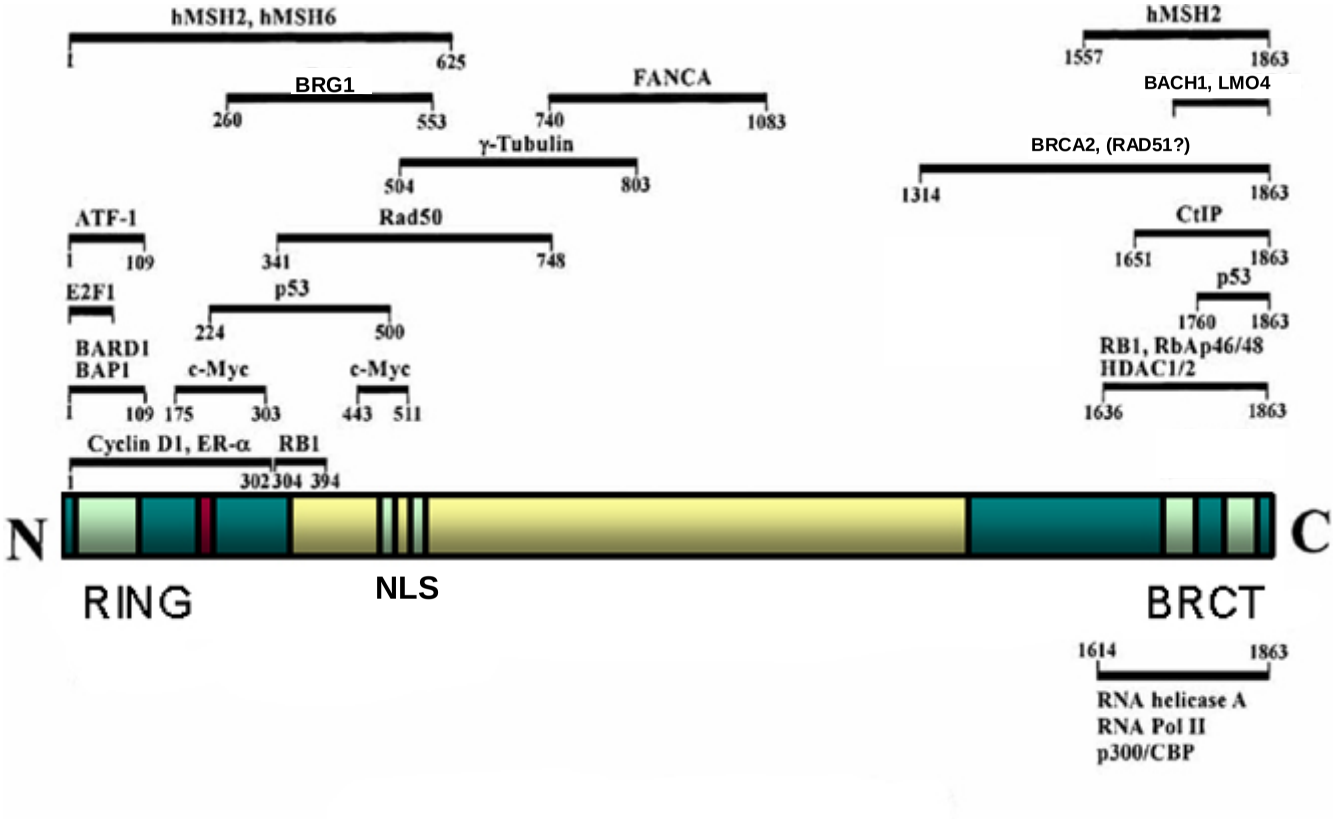


Figure 1.5: BRCA1 interacts with many other proteins at various sites. Schematic showing where BRCA1 is thought to interact with other proteins. This list is not comprehensive. The part of the protein coded by exon 11 is depicted in yellow.
Abbreviations: ATF-1=activating transcription factor-1, BACH1=BRCA1-Associated C-terminal Helicase 1, BAP1= BRCA1-associated protein 1, BARD1=BRCA1-associated RING domain protein-1, BRCT=BRCA1 C-terminal repeat, BRG1= *Brahma*-related gene 1, CBP=p265 CREB binding protein, CtIP=CtBP (C-terminal binding protein)-interacting protein, ER-α= oestrogen receptor-α, FANCA=Fanconi anaemia protein A, HDAC1/2=histone deacetylase-1/2, LMO4=LIM domain-only 4, hMSH2, hMSH6=MutS homologues 2 and 6, NLS=nuclear localization signal, RB1=Retinoblastoma protein, RbAp46/48= retinoblastoma-associated protein 46/48 kDa, RNA Pol II=RNA Polymerase II, RING=zinc finger RING domain.
Adapted from (Rosen, 2003).

promoters (Xu, 1997c), but this finding has been contested by another group (Suen and Goss, 2001).

1.5.3 BRCA1 and the cell cycle

BRCA1 protein and mRNA levels are dynamic. While *BRCA1* appears to be ubiquitously expressed in growing cells (Marquis, 1995; Chen, 1996; Ruffner and Verma, 1997), maximum expression levels occur at the G1-S boundary and during S phase (Gudas, 1996). In a similar manner, BRCA1 protein levels are highest during S phase. BRCA1 protein is also phosphorylated at different levels; hyperphosphorylation occurs during the G1-S transition, and the modification remains throughout M phase; partial dephosphorylation occurs in early G1 (Ruffner and Verma, 1997; Scully, 1997b). BRCA1 protein is mainly nuclear, with a diffuse staining pattern, but is often observed in S phase nuclear "foci" which persist until G2 (Scully, 1996; Scully, 1997c). The biological relevance of these foci is still unknown. Following DNA damage, there is at least a transient upregulation of BRCA1 protein and mRNA levels, and the protein becomes hyperphosphorylated (more so than in S phase) and localizes to damage-induced nuclear foci (Scully, 1997b; MacLachlan, 2000a).

The relationship of BRCA1 levels to cell cycle phases led to the suggestion that BRCA1 might be involved in the G1-S and/or G2-M cell-cycle checkpoints, a supposition now supported by several lines of evidence. There have been reports that BRCA1 may be involved in the S phase transition, as well; one group studying the HCC1937 cancer cell line reported a defective S phase checkpoint in these cells following gamma (γ)-irradiation (Xu, 2001a). Restoration of the defective checkpoint occurred upon transient expression of *BRCA1* (Xu, 2001a; Xu, 2002). However, another group studying the HCC1937 cell line reported a normal S phase checkpoint following γ-irradiation (Scully, 1999), suggesting that further study is needed to determine the role of BRCA1 in this checkpoint. The evidence linking BRCA1 to the G1-S and G2-M checkpoints is more compelling.

1.5.3.1 BRCA1 and the G1-S checkpoint

The G1-S cell-cycle checkpoint exists to prevent the replication of damaged DNA. Progression through this checkpoint requires the kinase/cyclin pairs CDK4/Cyclin D and CDK2/Cyclin E. The CDK2 and CDK4 kinases regulate proteins involved in S phase promotion, including inactivating the suppressor protein RB1 and activating CDC45. CDK2/Cyclin E appears to be a main target for the damage-induced G1-S checkpoint; following damage, the kinases ATM and ATR (through CHK1 and CHK2) destroy the phosphatase CDC25A, which normally activates cyclin D and cyclin E (this a rapid response through to result from proteasome-mediated degradation). Additionally, the p53 protein is stabilized (reviewed in Iliakis, 2003). p53 stabilization has various effects, including upregulation of the cell-cycle related gene $p21^{Waf1/Cip1}$ (referred to here as *p21*). As p21 is a potent suppressor of CDK2, stabilization of p53 is a second pathway for blocking progression into S phase (el-Deiry, 1993; Harper, 1993).

Overexpression of *BRCA1* has been shown to result following DNA damage (Clarkin, 2000; MacLachlan, 2000a). Overexpression of *BRCA1* in cultured cells has also been shown to cause growth suppression in conjunction with an arrest in the G1 phase of the cell cycle (Somasundaram, 1997; Aprelikova, 1999), as well as slowed development of MCF-7 (a human breast cancer cell line)–derived tumours in nude mice (Holt, 1996). However, overexpression of *BRCA1* does not suppress the growth of cells lacking either RB1 or p21 (Somasundaram, 1997; Aprelikova, 1999). Since BRCA1 can upregulate *p21* (Somasundaram, 1997), this suggests that overexpression of *BRCA1* inhibits S phase progression through its effect on p21, which normally suppresses the activity of CDK2. If this is the case, cells lacking p21 or RB1 (which is normally activated by CDK2 to stimulate progression into S phase) would then be expected to be insensitive to *BRCA1* expression levels (see Figure 1.6 for a simplified diagram) (Zhang, 1998; Aprelikova, 1999; MacLachlan, 2000b). Overexpression of *BRCA1* may also aid in stabilization of the p53 protein; two different regions of BRCA1 appear to interact with p53 (see Figure 1.5) (Somasundaram, 1999). Either stabilization of p53 or upregulation of *p21* should result in a G1-S block, as discussed above.

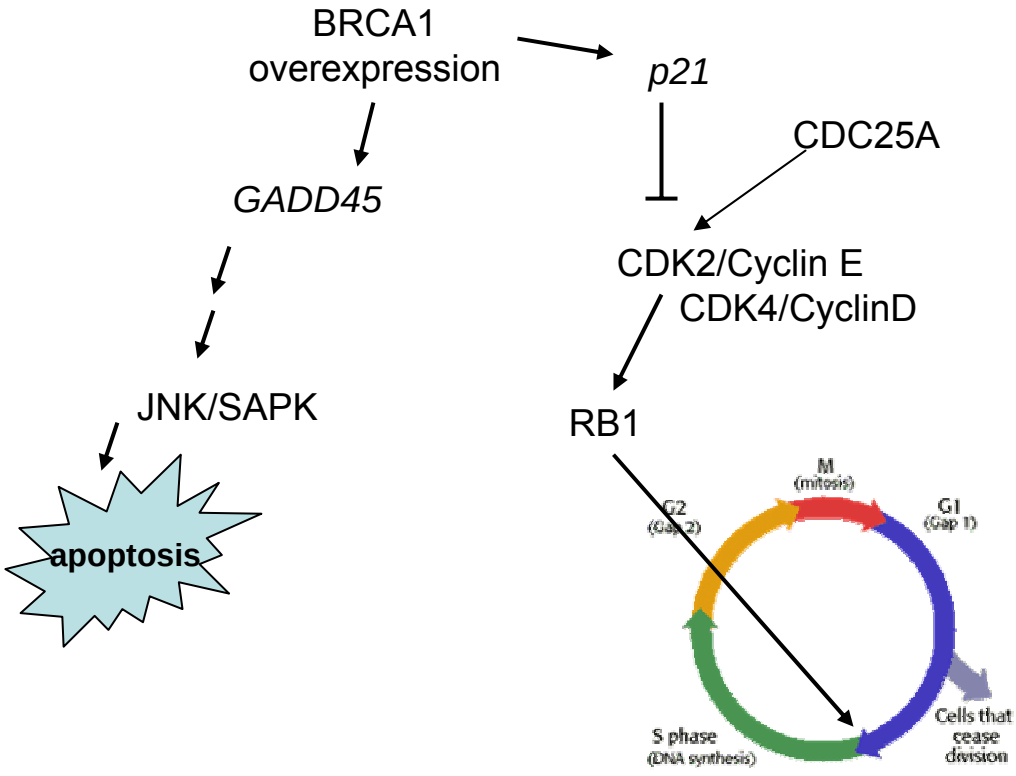


Figure 1.6: Consequences of BRCA1 overexpression at the G1/S checkpoint. BRCA1 transcriptionally upregulates *p21*. The protein product of *p21* represses the kinase activity of CDK2/CyclinE. CDK2 and CDK4 normally phosphorylate RB1 to allow progression into S phase, but repression of CDK2 activity by p21 stops progression. BRCA1 also transcriptionally upregulates *GADD45*, which is thought to lead to a JNK/SAPK-mediated apoptosis response.
Abbreviations: GADD45=Growth Arrest and DNA Damage 45, CDK=Cyclin-dependant kinase 2, RB1=Retinoblastoma protein, JNK=c-Jun N-terminal kinase, SAPK=stress-activated protein kinase.

The effect of *BRCA1* overexpression may have a second consequence, mediated through its effect on the *Growth Arrest and DNA Damage-inducible* gene *GADD45* (Harkin, 1999). *GADD45* was originally identified in a screen for genes upregulated in response to exposure to ultraviolet (UV) light, and was subsequently shown to be upregulated upon exposure to several other forms of DNA damage including hydrogen peroxide (H_2O_2), X-rays, mitomycin C (MMC), and hydroxyurea (HU) (Fornace, 1989; Papathanasiou, 1991). BRCA1 appears to interact both with the *GADD45* promoter and a region in intron 3 of *GADD45* (Harkin, 1999; Jin, 2000). However, the exact mechanism of these interactions is not fully understood, as BRCA1 might not bind directly to *GADD45* (Jin, 2000). Instead, two other proteins may serve as a link between *GADD45* and BRCA1. One is ZBRK1, a BRCA1-dependent corepressor which binds to intron 3 of *GADD45* (Zheng, 2000), another is the oncogene *C-MYC*, which has been shown to attenuate the induction of *GADD45* following UV or methyl methanesulfonate (MMS) treatments (Amundson, 1998). BRCA1, which has been shown to interact with C-MYC and represses its transactivation capabilities (likely by sequestration), is postulated to relieve this repression, and thus indirectly result in upregulation of *GADD45* following damage (Wang, 1998; Mullan, 2001). p53 is also involved in the regulation of *GADD45*, but appears to be involved primarily in the response of *GADD45* to ionizing radiation and to interact with intron 3 of the gene (Kastan, 1992).

Upregulation of *GADD45* by BRCA1 appears to be able to trigger a c-Jun N-terminal kinase/stress-activated protein kinase (JNK/SAPK)–mediated apoptotic response (Figure 1.6). Apoptosis is delayed from the onset of *BRCA1* overexpression, which suggests that cells arrest and then undergo apoptosis (Harkin, 1999; MacLachlan, 2000b). Other groups argue that overexpression of *BRCA1* does not result in apoptosis, although cell-cycle arrest still occurs (Randrianarison, 2001). This discrepancy may be due to the times at which the amount of apoptosis was measured, or to different levels of *BRCA1* overexpression in the various experiments (MacLachlan, 2000b). Although the relationship of BRCA1 overexpression to the induction

of apoptosis requires further work, the involvement of BRCA1 in the G1-S checkpoint is well-supported by various experiments.

1.5.3.2 BRCA1 and the G2-M checkpoint

The G2-M checkpoint delays entry into mitosis if the genome is damaged. CDC2/Cyclin B is the key kinase/cyclin complex involved in the transition into M phase. CDC2 is inhibited by phosphorylation added by the kinases WEEI and MYTI earlier in the cell cycle. Dephosphorylation of CDC2 (likely by the dephosphorylase CDC25C, although other proteins may also be involved) is both a key step in progression into mitosis and a target for inhibition during checkpoint control. The kinases ATM and ATR are again important in triggering the checkpoint; ATM activates CHK2 and ATR activates CHK1, both of which inhibit CDC25C activity to block progression into mitosis (CHK1 also activates WEEI, a direct inhibitor of CDC2) (See Figure 1.7 for a simplified diagram).

Both human (HCC1937) and murine cell lines which express only mutated BRCA1 protein fail to arrest at the G2-M boundary following γ-irradiation (Foray, 1999; Xu, 1999c; Xu, 2001a; Yarden, 2002), suggesting that BRCA1 normally plays a role in regulation of this checkpoint. Recent experiments suggest that the effect of BRCA1 is linked to the kinases CHK1 and ATM: In the absence of BRCA1, CHK1 is not activated, resulting in the deregulation of CDC2/Cyclin B and loss of control over the progression into M phase (Yarden, 2002; Yamane, 2003). Additionally, while transient expression of *BRCA1* in HCC1937 cells restores the G2-M checkpoint, transient expression of a *BRCA1* gene with a mutation in a site generally phosphorylated by the ATM kinase following DNA damage, the checkpoint is not restored, suggesting that phosphorylation of BRCA1 by ATM may be necessary for checkpoint control (Xu, 2001a).

A more recent study has also provided evidence that BRCA1 is involved in the G2-M transition and checkpoint, via an interaction with the Aurora-A kinase, which directly phosphorylates BRCA1. In a mutant mouse cell line expressing only the Δ*X.11* form of *Brca1* (this cell line lacks the G2-M checkpoint

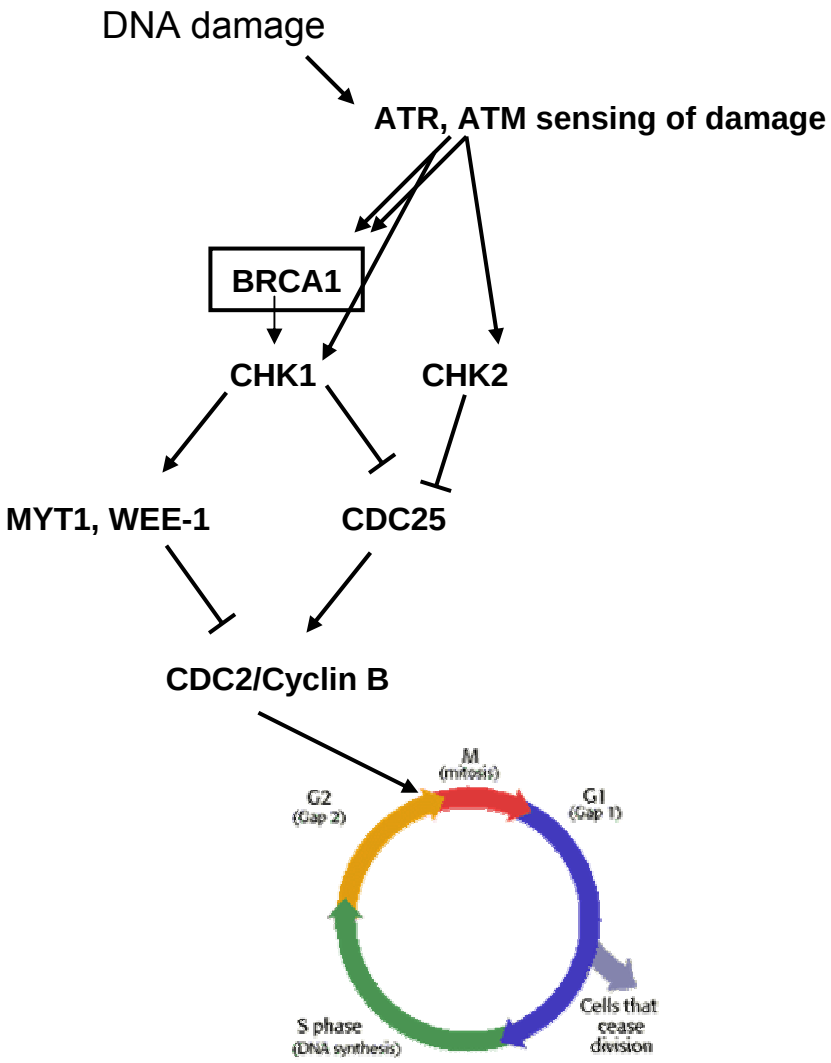


Figure 1.7: BRCA1 at the G2/M checkpoint.
Simplified schematic showing normal arrest at the G2-M boundary following DNA damage when BRCA1 is present. Loss of BRCA1 prevents activation of the CHK1 kinase, leading to deregulation of CDC2/Cyclin B and loss of G2-M checkpoint control. Following DNA damage, BRCA1-deficient cells fail to arrest at this boundary. Abbreviations: CHK1, MYT1, and WEE-1=cell-cycle dependent kinases, CDC25= cyclin-dependent kinase 25, CDC2=cell division cycle 2. Figure adapted from (Yarden, 2002) and (Iliakis, 2003).

following treatment with γ-irradiation (Xu, 1999c)), transient expression of wildtype *Brca1* restores the γ-irradiation-induced G2-M checkpoint. However, expression of a *Brca1* transgene carrying a mutation in the residue phosphorylated by Aurora-A abrogated the irradiation-induced checkpoint (Ouchi, 2004). Loss of Aurora-A itself also abrogates the G2-M checkpoint (Hirota, 2003). In summary, in both mouse and human cell lines, several experiments have provided evidence for the importance of BRCA1 in the G2-M checkpoint, although further experiments are needed to define the exact mechanism of its action.

1.5.3.3 Microarray experiments and expression of cell-cycle genes

Microarray analyses provide evidence that overexpressing *BRCA1* influences the expression profile of a number of genes, including ones involved in cell-cycle checkpoint control (Atalay, 2002; Welcsh, 2002). This type of experiment is likely to be relevant to the DNA repair phenotypes of BRCA1, as overexpression of *BRCA1* appears to occur following DNA damage (Clarkin, 2000; MacLachlan, 2000a). Microarray analyses of *BRCA1*-related tumours indicate that profiles of tumour RNA can be used to categorize breast tumours as sporadic or *BRCA1*- or *BRCA2*-related, a potentially useful diagnostic tool (Berns, 2001; Hedenfalk, 2001; Hedenfalk, 2003). However, in both types of study, the biological relevance of the genes identified is not completely clear. A number of genes which have not been previously studied in relation to BRCA1 have been identified, and additional experiments must be performed in order to clarify their involvement with BRCA1. Additionally, the small sample size in most studies makes meaningful statistical comparisons between tumour types difficult at present. Microarray technology has a very promising future, and it is expected that further studies will prove very useful in furthering the understanding of BRCA1-related biology and tumourigenesis, including its effects on cell cycle–related genes.

1.5.4 Transcriptional regulation of BRCA1

BRCA1 expression, transcription, and mRNA and protein levels appear to be influenced by a number of factors (as discussed in the previous section). At

the expression level, both positive and negative factors have been proposed to alter *BRCA1* expression levels. These include the non-histone chromatin protein HMGA1, which has been shown to bind to the promoter region of *BRCA1* and downregulate its expression, resulting in a lower amount of both mRNA and protein expression. This effect of the HMGA1 protein is confirmed by murine *Hmga1*$^{-/-}$ knockout ES cells, which have higher levels of *Brca1* mRNA expression than do wildtype cells (Baldassarre, 2003).

BRCA1 expression may be upregulated through the action of E2F1 (E2F-transcription factor-1) and RB1. The RB1 protein binds to and sequesters E2F transcription factors; normally, phosphorylation of RB1 at the G1-S boundary inactivates RB1 to allow E2F-mediated upregulation of genes needed to pass through the cycle boundary (Stevaux and Dyson, 2002). The promoters of both human and mouse *BRCA1* contain E2F1 binding sites (to which recombinant E2F1 will bind *in vitro*). Overexpression of *E2F1*, either in transgenic mice or *in vitro* in a human cell line, results in upregulation of *Brca1* mRNA expression. *Brca1* expression is also upregulated in *Rb1*$^{-/-}$ MEFs, perhaps because E2F1 is expected to be active because of loss of sequestration. Ectopic expression of the *Rb1* gene in *Rb1*$^{-/-}$ MEFs restores the expression level of *Brca1* (Wang, 2000a).

p53 also appears to participate in the regulation of BRCA1 (MacLachlan, 2000a). Following DNA damage, BRCA1 mRNA and protein levels are rapidly upregulated (by 15 minutes post-damage), but fall again by 4-12 hours post-damage (Andres, 1998; Clarkin, 2000). This reactive downregulation appears to be caused by p53 binding to the promoter of *BRCA1*. In cancer cell lines which lack p53, damage-induced upregulation of *BRCA1* occurs, but suppression after the initial burst of expression does not occur (MacLachlan, 2000a). In addition, overexpression of *BRCA1* may stabilize the p53 protein; two different regions of BRCA1 appear to interact with p53 (see Figure 1.5 and section 1.5.3.1). Taken together, these results suggest that BRCA1 and p53 may be involved in a type of feedback loop (Somasundaram, 1999). These data provide evidence that BRCA1 and p53 participate in at least one common pathway. However, the fact that *BRCA1*-related tumours frequently

carry *p53* mutations (Table 1.2) suggests that these two proteins have additional roles in separate pathways.

1.6 INTERACTIONS WITH OTHER PROTEINS

Numerous experiments have provided evidence that BRCA1 interacts with many proteins involved in gene regulation, the cell cycle, and/or DNA repair (Figure 1.5). In fact, one group has asserted that BRCA1 associates with a large number of DNA-repair proteins in a BASC (BRCA1-associated genome surveillance complex). Characterization of BASC components has identified such proteins as the RAD50-MRE11-NBS1 complex thought to play a role in double-strand break repair (DSBR), meiotic recombination and maintenance of telomeres, as well as components of the mismatch-repair system and the Bloom's Syndrome protein (BLM), which confers cancer predisposition through a high propensity for sister-chromatid exchanges and mitotic recombination with resultant LOH (Ellis, 1995; Wang, 2000b; Thompson and Schild, 2002).

One caveat to the large number of published interactions is that a diverse array of methods have been used to demonstrate these interactions, and not all partnerships may prove to be biologically relevant when more stringently investigated. Some well-characterized interactions are given as examples here; further cases are discussed throughout this introduction. BRCA1 appears to interact with BRCA2 (Chen, 1998). Both BRCA1 and BRCA2 interact with RAD51 (although this may be an indirect interaction in the case of BRCA1), the mammalian homologue of the bacterial RecA protein which mediates strand-exchange during recombination (Shinohara, 1993; Scully, 1997c). All three co-localize to S phase and DNA damage-induced nuclear foci, and may act in a common pathway during crossing over in meiosis (Chen, 1998; Chen, 1999). BRCA1 also interacts with RAD50 and the Nijmegen breakage syndrome protein NBS1, both part of the RAD50-MRE11-NBS1 complex (Varon, 1998; Wang, 2000b). RAD50 and BRCA1 also co-localize in some damage-induced nuclear foci (Haber, 1998; Petrini, 1999; Zhong, 1999; Wang, 2000b), as does RAD51. However, while RAD50 or

RAD51 alone may co-localize with BRCA1 in foci, the former two proteins are rarely found in the same focus (Maser, 1997; Zhong, 1999). Interestingly, homozygous mouse knockouts of *Brca1*, *Brca2*, *Rad50*, or *Rad51* all exhibit early embryonic lethality with similar phenotypic profiles (Lim and Hasty, 1996; Ludwig, 1997; Sharan, 1997; Luo, 1999).

1.6.1 The RING domain

The N-terminus of BRCA1 contains a cysteine-rich, Zn-finger motif with a C3-H-C4 configuration (C=cysteine and H=histidine). This motif is known as a RING finger, after RING1 (Really Interesting New Gene 1), the first novel protein identified which carried the motif (Freemont, 1991). Originally, this Zn-finger motif was thought to be involved in binding DNA (Lovering, 1993), but proteins containing the RING-finger are now known to be one of two major types of E3 ubiquitin ligases, involved in ubiquitination of proteins targeted for destruction through the proteasome-mediated degradation pathway (Pickart, 2001; Ravasi, 2003; Semple, 2003). This pathway consists of three steps (Figure 1.8). An E1 enzyme activates the small protein ubiquitin and interacts with one of several E2s, which carry the activated ubiquitin and interface in turn with an E3 protein. E3s catalyze the transfer of ubiquitin from an E2 to a protein substrate. E3s recognize only one substrate, or several closely-related substrates. Generally, a chain of several ubiquitin molecules (polyubiquitination) is assembled on proteins prior to their degradation by the 26S proteasome (Weissman, 2001).

1.6.2 BARD1 is an important RING-binding partner of BRCA1

The RING domain of BRCA1 is the interaction site for BARD1 (BRCA1-Associated RING Domain 1), an important protein partner of BRCA1. BARD1 resembles BRCA1, in that it has an N-terminal RING domain and two C-terminal BRCT repeats, but is smaller (777 aa), and also contains ankyrin repeats (Wu, 1996). The mouse homologue of BARD1, Bard1, shares a 77% overall identity with its human counterpart, with high conservation at the RING domain, BRCT repeats, and ankyrin repeats. Though the higher overall

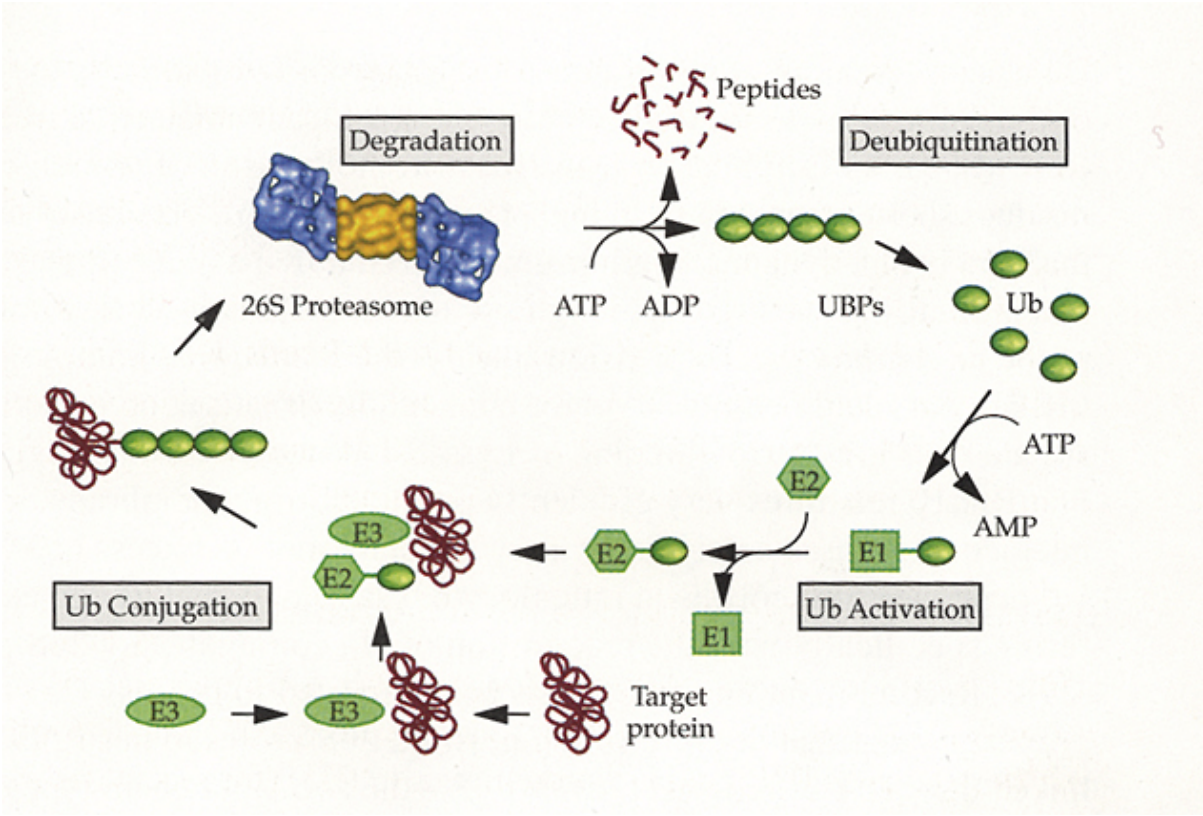


Figure 1.8: The proteasome-mediated ubiquitination pathway of protein degradation. Ubiquitin molecules (small green ovals) are activated by the E1 enzyme, then activated ubiquitin is passed to an E2 enzyme. E3 ubiquitin ligases interface with a single protein target (or closely-related group of targets), and catalyze transfer of the activated ubiquitin molecule from the E2 enzyme to the target protein. A polyubiquitin chain marks the target protein for proteasome-mediated degradation. Recycling of ubiquitin molecules follows protein degradation. Ub=ubiquitin.
Figure drawn by Colin Gordon (see Gordon in references).

homology seems to indicate that BARD1 is more highly conserved than BRCA1, BARD1 is much smaller than BRCA1. If the highly homologous BARD1 regions are not counted, the overall BARD1-Bard1 identity is only 61%, much closer to the 57% overall identity of BRCA1 and Brca1 (Ayi, 1998). Mouse Bard1 and Brca1 also interact via their RING domains (Ayi, 1998).

Polypeptides containing the BRCA1 or BARD1 RING domains have been used to demonstrate that both proteins will form homodimers *in vitro*, although heterodimerization appears to be preferred (Brzovic, 1998; Meza, 1999). Recently, it was shown that BARD1 and BRCA1 RING domains will assemble into dimers and tetramers, but are more commonly observed in supramolecular structures containing 12 RING domains. Whether these larger assemblies of BRCA1 and BARD1 are normally present in cells is unclear, as full-length BRCA1 was not used in these experiments, and the reactions were performed *in vitro* using purified proteins. These studies also indicated that BRCA1 or BARD1 alone will form 12-RING bodies, but the cancer-related C64G mutation in the BRCA1 RING domain abolishes the ability to form 12-RING bodies (Kentsis, 2002).

1.6.3 E3 ubiquitin ligase activity of BRCA1/BARD1

Alone, both BRCA1 and BARD1 do have some E3 ubiquitin ligase activity, but it is low compared to the activity of BRCA1 and BARD1 together. Additionally, the BRCA1/BARD1 12-RING bodies described above are a more effective E3 ligase than BRCA1/BARD1 dimers or tetramers (as measured by an *in vitro* ubiquitination assay using polyubiquitin-specific antibodies), possibly because additional E2s can be conjugated to the 12-RING bodies (Hashizume, 2001; Kentsis, 2002). The BRCA1/BARD1 E3 does have characteristics of normal E3 ligases: its activity is dependent on the presence of an E2 enzyme (in their case, Ubc5c (Hashizume, 2001; Mallery, 2002)), as well as ATP, ubiquitin, and a functional E1 enzyme (Kentsis, 2002). The RING domain of BRCA1 is also the site for interaction with a ubiquitin hydrolase, or deligase, called

BAP1 (BRCA1 Associated Protein 1), but very little is known about the biological consequences of this interaction (Jensen and Rauscher, 1999).

1.6.4 Targets of the BRCA1/BARD1 E3 ubiquitin ligase

The biological targets of the E3 activity of BRCA1/BARD1 are not well-characterized as yet, though *in vitro* studies have demonstrated that the two can monoubiquitinate histone proteins such as H2AX, H2A, H2B, H3, and H4 (Chen, 2002; Mallery, 2002). The biological consequences of this monoubiquitination are not fully understood, but as monoubiquitinated proteins do not seem to be targeted to the 26S proteasome for degradation, this modification may be a way of promoting protein-protein interactions between the modified histones and chromatin-remodeling complexes (Jason, 2002).

Until recently, the Fanconi anaemia protein FANCD2 was considered a likely target for the BRCA1/BARD1 E3 ligase. Fanconi anaemia, a human syndrome characterized by cancer predisposition, is caused by defects in one of a number of *FANC* genes (D'Andrea and Grompe, 2003). BRCA1 has been shown to interact with at least two of the FANC proteins, FANCA and FANCD1 (also known as BRCA2) (Folias, 2002; Howlett, 2002). Mutation of a *FANC* gene confers susceptibility to DNA damage. Following DNA damage, FANCD2 is activated by monoubiquitination and co-localizes with BRCA1 in foci at sites of DNA damage (Garcia-Higuera, 2001; D'Andrea and Grompe, 2003). In HCC1937 cells (a human breast cancer cell line which carries only a mutated version of BRCA1 (Tomlinson, 1998)), there is a decrease in monoubiquitinated FANCD2 following γ-irradiation, which suggests that BRCA1 may play some role in this modification (Garcia-Higuera, 2001). However, it has been shown that, at least *in vitro*, BRCA1/BARD1 is unable to ubiquitinate FANCD2 (Vandenberg, 2003). More recently, a new candidate E3 has been identified, the product of the gene *PHF9*, which is mutated in some Fanconi anaemia patients and may represent a new *FANC* gene (Meetei, 2003). BRCA1 may still play some role in the ubiquitination of FANCD2 following DNA damage, but more work is needed to define this role.

Recent studies of the effects of DNA damage on RNA Polymerase II (RNA Pol II), the polymerase responsible for transcription of coding genes to mRNA, demonstrate that following UV exposure or α-amanitin treatment (to inhibit RNA Pol II activity), a percentage of the large subunit of RNA Pol II is ubiquitinated and degraded (Bregman, 1996; Nguyen, 1996; Ratner, 1998). There has been much speculation that BRCA1/BARD1 is the E3 ligase which ubiquitinates RNA Pol II, though no data have yet been published to support this claim. Recent papers have introduced other candidates: the von Hippel-Lindau protein (pVHL) ubiquitinates the phosphorylated fraction of RNA Pol II large subunit in a UV-dependent manner (Kuznetsova, 2003), as does the human protein NEDD4L, both *in vitro* and *in vivo*. Additionally, the yeast homologue of NEDD4L, Rsp5, ubiquitinates yeast RNA Pol II (Beaudenon, 1999). Although the existence of these alternative candidates does not exclude the possibility that BRCA1/BARD1 can ubiquitinate RNA Pol II, it does suggest that other E3 ligases may have the same ability.

In fact, the only *in vivo* activity of the BRCA1/BARD1 E3 identified to date appears to be auto-polyubiquitination; that is, heterodimerization appears to catalyze the assembly of polyubiquitin chains on BRCA1 and BARD1 themselves (Chen, 2002). Polyubiquitination of BRCA1/BARD1 appears to increase the E3 ligase activity of the heterodimer (Mallery, 2002). This auto-polyubiquitination seems to contradict two reports that BRCA1 and BARD1 stabilize one another (Joukov, 2001b; McCarthy, 2003), since polyubiquitinated proteins are normally targeted for destruction, but a recent paper shows that the polyubiquitin chains assembled on BRCA1/BARD1 have an unconventional Lys-6 linkage configuration, different from the commonly-observed Lys-48 linkage of polyubiquitin chains on proteasome-targeted proteins (Wu-Baer, 2003). Furthermore, a recent *in vitro* study demonstrated that polyubiquitinated BRCA1/BARD1 is de-ubiquitinated – but not degraded – by the 26S proteasome (Nishikawa, 2004). While little is known about alternative consequences of protein ubiquitination, proteasome targeting is not the sole reason for ubiquitin modification, and the ubiquitin chains on BRCA1/BARD1 may serve to stabilize the proteins, direct them to other

proteases, or alter their conformation to facilitate interactions with other proteins (Chen, 2002; Aguilar and Wendland, 2003; Schnell and Hicke, 2003).

1.6.5 BARD1 alone

Several studies indicate that the majority of cellular BARD1 is associated with BRCA1 (Yu and Baer, 2000; Joukov, 2001b). Biochemical fractionation experiments show that a percentage of BARD1 does not co-purify with BRCA1, which suggests that BARD1 may have BRCA1-independent cellular functions (Chiba and Parvin, 2002), although few studies of these putative independent functions have been published. Based on its interaction with BRCA1, it is hypothesized that BARD1 might be a tumour suppressor itself. However, screening panels of tumours (mainly breast tumours) for BARD1 mutations has indicated that if BARD1 is involved in tumourigenesis, its involvement is either rare or confined to cancers which have not been investigated (Thai, 1998; Ghimenti, 2002; Ishitobi, 2003). A mouse knockout of *Bard1* has been generated; it is phenotypically identical to both *Brca1* knockout and *Brca1/Bard1* double knockout mice, although the embryonic lethality of both models precludes extensive investigation (McCarthy, 2003). Mouse models will be discussed further in section 1.9.

1.7 NUCLEAR LOCALIZATION: NLS, NES, AND BARD1

As described in section 1.5.2, both mouse and human *BRCA1* express natural splice isoforms which lack all or most of exon 11, the largest exon. Transiently overexpressed human BRCA1 ΔX.11 or ΔX.11b proteins are localized in the cytoplasm, unlike the full-length form which is found in the nucleus and cytoplasm (Thakur, 1997; Wilson, 1997). Examination of exon 11 following these overexpression experiments revealed the presence of two nuclear localization sequences (NLSs) at the 5' end (Thakur, 1997; Wilson, 1997) (Figure 1.9).

Murine *Brca1* also has NLSs in exon 11, identical in sequence to those of human *BRCA1* (Figure 1.9) (Xu, 1999c; Bachelier, 2000). Transient

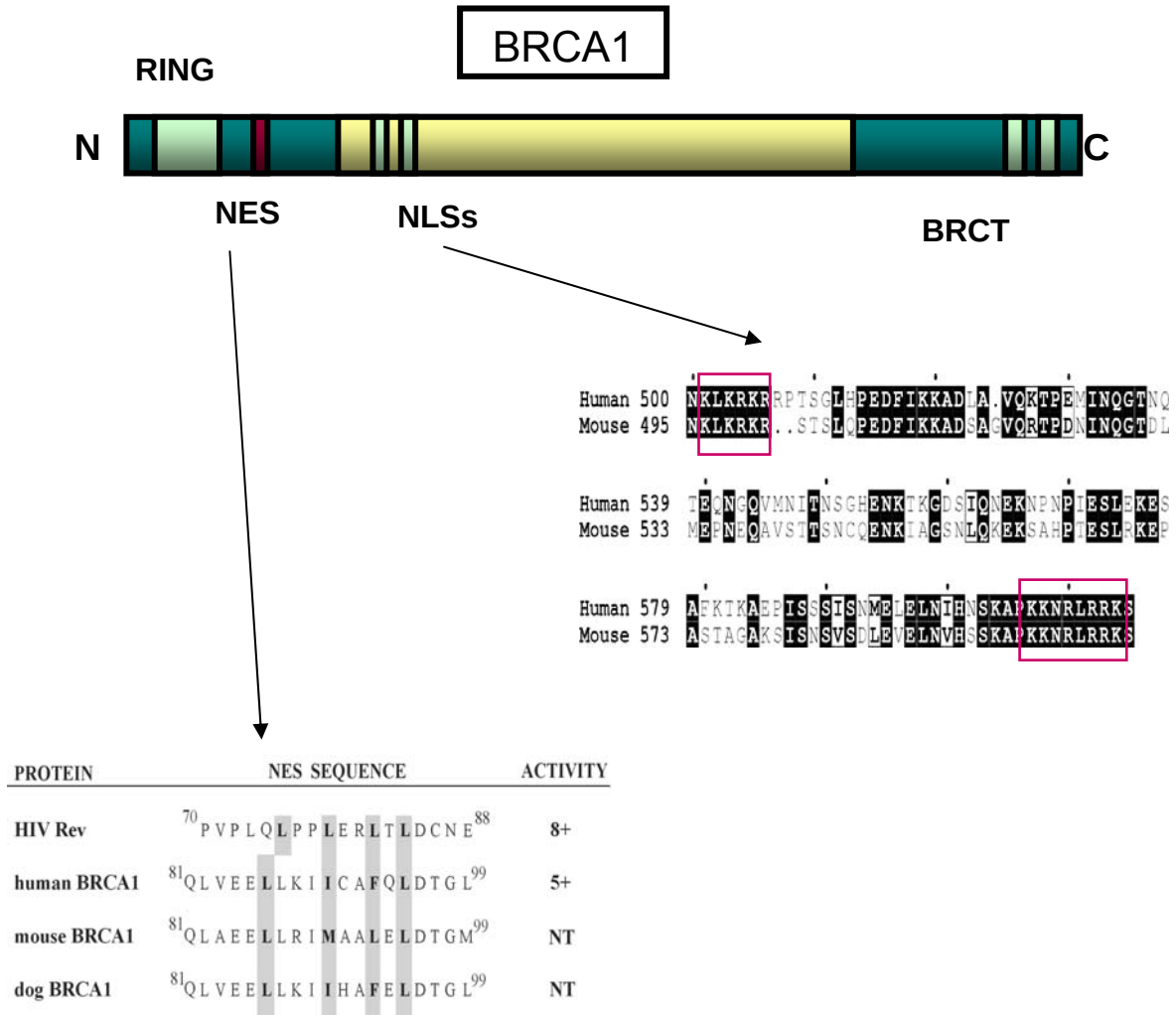


Figure 1.9: Conservation of the NLS and NES signals. BRCA1 protein schematic showing the nuclear localization sequences (NLSs) and the nuclear export sequence (NES). The alignment of the mouse and human NESs with the HIV Rev protein is taken from (Rodriguez, 2000). Mouse and human NLS alignments (red box) demonstrate that these domains are absolutely conserved. The yellow box indicates the part of the protein coded by exon 11. Both the RING domain and the BRCT repeats are indicated in green.
Abbreviations: HIV=human immunodeficiency virus, NES=nuclear export signal, NLSs=nuclear localization signals.

overexpression of *Brca1 ΔX.11 in vitro* revealed that the resulting Brca1 ΔX.11 protein was found in the cytoplasm, similar to what is observed in human cells overexpressing *BRCA1 ΔX.11*. In contrast, in a MEF cell line homozygous for the *Brca1 ΔX.11* isoform, in which the Brca1 ΔX.11 protein is expressed at physiological levels, Brca1 ΔX.11 protein is found in the nucleus, and even forms S phase nuclear foci (Xu, 1999c; Bachelier, 2000).

A nuclear export signal (NES) in exon 6 of *BRCA1* has also been described. This signal was identified based on its similarity to the export signal of the human immunodeficiency virus (HIV) Rev protein, and is well-conserved between species (Figure 1.9). When joined to a reporter gene, the *BRCA1* NES will stimulate nuclear export of the reporter protein (Rodriguez and Henderson, 2000). Overexpression of a mutant *BRCA1* transgene either lacking the NLS or carrying a mutated NES (with two key residues mutated) results in almost exclusively nuclear localization of BRCA1 (Rodriguez and Henderson, 2000). However, if the NES and NLSs are the only factors involved in localization, then a mutant transgene which retains both NES and NLSs should have the same localization pattern as full-length BRCA1. This is not the case. Fabbro *et al.* demonstrated that the majority of cells overexpressing a *BRCA1* transgene which lacks the 5' RING domain (but includes both NLSs and the NES) only showed cytoplasmic localization of BRCA1 (BRCA1 was seen in both the cytoplasm and nucleus in some cells), in contrast to cells expressing full-length *BRCA1*, in which the protein was detected in the nucleus and cytoplasm (Figure 1.10).

Based on these data, Fabbro *et al.* postulated that the binding of another protein to the BRCA1 RING domain might mask the NES (Fabbro, 2002). While testing the effect of RING-binding proteins on the localization of BRCA1, they found that in cells co-transfected with a full-length *BARD1* transgene and a *BRCA1* transgene carrying a mutated NLSs, BRCA1 was located predominantly in the nucleus. When a *BARD1* transgene was co-transfected with a wildtype *BRCA1* transgene, more BRCA1 was localized in the nucleus than when *BARD1* was not overexpressed (Figure 1.10). This increase in nuclear localization was abrogated on deletion of the RING

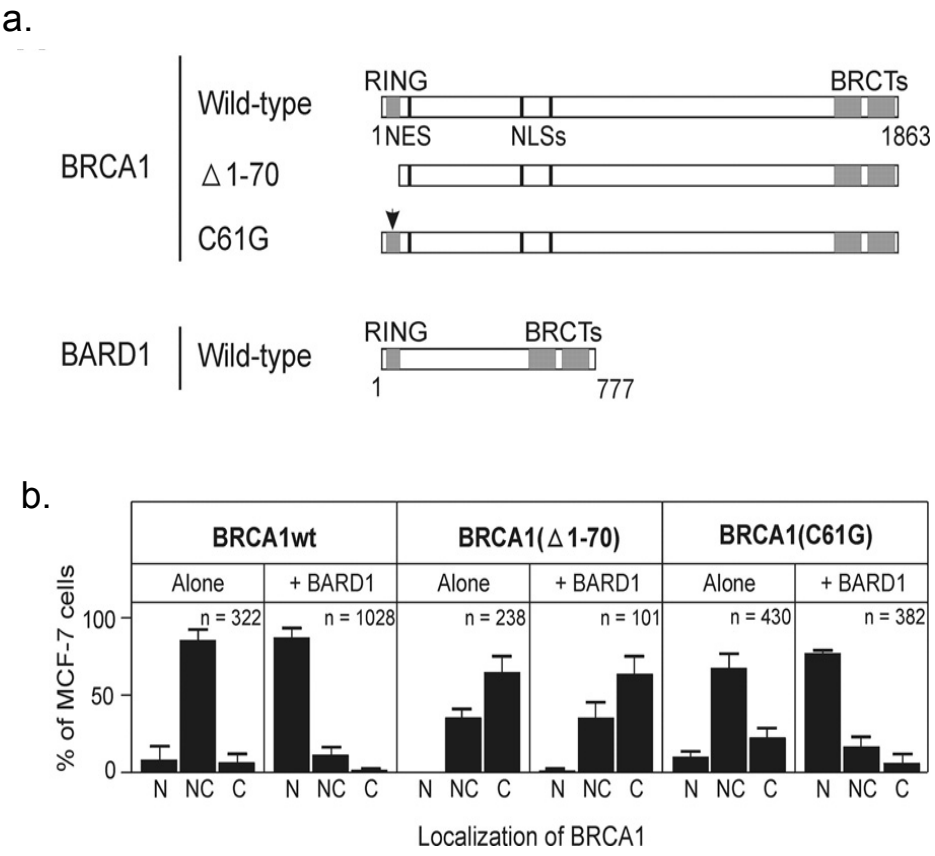


Figure 1.10: Nuclear localization of BRCA1 in the presence or absence of localization signals and BARD1. a. BRCA1 and BARD1 transgenes used in this study. **b.** Percentage of transfected cells showing nuclear (N), nuclear and cytoplasmic (NC), or cytoplasmic (C) localization of BRCA1 protein, as detected by immunofluorescence. Transgenes were transfected with or without a BARD1 transgene ("+BARD1" or "Alone"). The first panel shows the localization of wildtype BRCA1, the second of a transgene lacking the RING domain at the N-terminus (Δ1-70), and the third of a transgene with a cancer-related mutation in the BRCA1 RING domain which still allows binding to BARD1 (C61G).
Figures taken from (Fabbro, 2002).

domain of either protein (Fabbro, 2002). The published BRCA1-BARD1 RING-interaction structure confirms that the NES is buried when BARD1 binds to BRCA1 (Brzovic, 2001b). The hypothesis that BARD1 is a nuclear chaperone and/or nuclear retention partner of BRCA1 may help to explain why overexpressed BRCA1 ΔX.11 protein is located in the cytoplasm, while Brca1 ΔX.11 protein expressed at endogenous levels has the expected nuclear-and-cytoplasmic localization pattern: endogenous levels of BARD1 may be insufficient to cope with the overexpressed transgene (Thakur, 1997; Wilson, 1997; Xu, 1999c; Bachelier, 2000). It cannot be ruled out that the mutated proteins arising from the transgenes used by Fabbro *et al.* may not have folded properly, but the proteins were detected at the expected levels, and N-terminal truncations of BRCA1 have been used in *in vitro* studies in the past, most notably during confirmation of the BRCA1-BARD1 interaction (Wu, 1996). Additionally, it is possible that other proteins which interact with BRCA1 at other domains may be able to serve as chaperones. A very recent paper has also provided evidence that intact BRCA1 is necessary for entry of BARD1 into the nucleus. As is the case for BRCA1, binding of BARD1 to BRCA1 at the RING domain buries the BARD1 NES, resulting in retention of the BARD1 protein in the nucleus. These authors have also provided evidence that increasing the cytoplasmic BARD1 fraction may lead to increased levels of apoptosis (Rodriguez, 2004).

The role of BARD1 in the nuclear localization of BRCA1 does not mean that the NLSs are unimportant or non-functional; in fact, they appear to augment BARD1-mediated import of BRCA1. In the absence of the RING domain, full-length BRCA1 is localized in both the nucleus and the cytoplasm, which indicates that the NLSs are able to influence nuclear import of BRCA1 (Fabbro, 2002).

1.8 BRCA1 IS A TRANSCRIPTION FACTOR

1.8.1 Transactivation *in vitro*

BRCA1 was originally thought to be a potential transcription factor, based on its N-terminal Zn-finger domain and acidic C-terminus (Miki, 1994). The C-terminus of BRCA1 (including the BRCT repeats) is able to function as a transactivation domain. A fusion of this region with the GAL4 DNA-binding domain will activate a reporter gene fused to the GAL4 activation domain (Chapman and Verma, 1996; Monteiro, 1996). However, *in vitro* studies show that a BRCA1 polypeptide containing the RING domain does not bind to double- or single-stranded DNA-cellulose columns. This BRCA1 polypeptide is likely to be folded correctly, as it will heterodimerize with a BARD1 RING domain polypeptide (Meza, 1999). However, residues 452-1079 in the middle of BRCA1 will bind DNA. This reaction does not appear to depend on sequence specificity, but branched substrates are preferred over linear ones, and longer sequences (over 500 bp) are preferred over shorter ones (Paull, 2001). The BRCT repeats at the C-terminus also appear to be able to bind to linear DNA in a sequence-independent manner (Yamane and Tsuruo, 1999). BRCA1 may protect bound DNA from the exonuclease activity of recombinant MRE1, whether MRE1 is part of a recombinant RAD50-MRE1-NBS1 complex or alone. This apparent protection suggests that BRCA1 binds to DNA as part of a repair process, not a transactivation process (Paull, 2001). In contrast to data which suggest that BRCA1 binds to DNA in a sequence-independent manner, a recent study demonstrates that BRCA1 appears to exist in a transcriptional regulatory complex in the cell which binds to a specific sequence motif found in such genes as *GADD45, STAT5A*, and the gene coding for Cyclin B1 (*CCNB1*). A reporter gene carrying this sequence motif was upregulated upon transfection of a *BRCA1* transgene, but expressed at a higher level following transfection of a *BRCA1* transgene carrying a cancer-related mutation (Cable, 2003).

1.8.2 Transcription factor activity *in vivo*

BRCA1 also appears to act as a transcription factor *in vivo*, although in many cases, its influence may depend on other proteins. For example, BRCA1 may act as a p53-dependent transcriptional activator, upregulating such p53-responsive genes as *BCL-associated X (BAX), Mouse Double-Minute 2 human homologue* (*HDM2*), and *p21* (Ouchi, 1998; Zhang, 1998).

BRCA1 also upregulates *p21* independently of p53 (Somasundaram, 1997), an influence which appears to be enhanced further upon overexpression of the androgen receptor (AR) (Park, 2000; Yeh, 2000) and repressed by overexpression of the transcriptional co-repressors CtBP/CtIP (Li, 1999). BRCA1 is thought to physically interact with both CtIP and the AR (Figure 1.5) (Wong, 1998; Li, 1999; Park, 2000; Yeh, 2000). Upregulation of *p21* expression is expected to have an effect on cell-cycle regulation, as p21 inhibits the kinase CDK2 (el-Deiry, 1993; Harper, 1993), which phosphorylates the protein RB1, a key event in the G1-S transition (see Figure 1.6). Indeed, overexpression of BRCA1 *in vitro* appears to result in dephosphorylation of RB1 (Somasundaram, 1997; MacLachlan, 2000b).

BRCA1 also may inhibit oestrogen-mediated signaling through the oestrogen receptor ER-α (Fan, 1999). Considering the high rate of loss of activity of oestrogen receptors in mouse and human *BRCA1*-related tumours (Table 1.2), this suggests that the ER has roles in other cellular pathways, otherwise loss of ER activity would not be expected to contribute to *BRCA1*-related tumourigenesis (Loman, 1998; Lakhani, 2002). The interaction of BRCA1 and the ER is likely to be complex, as it appears to be abolished by other proteins such as p300, and does not happen at all in some cell lines (Fan, 2001; Fan, 2002).

1.8.3 Other indications that BRCA1 may be linked to transcriptional control

Several publications provide evidence that BRCA1 is linked to the JAK/STAT (Janus kinase/signal transducer and activator of transcription) pathway, an important set of transcriptional regulators involved in the response to cytokines or growth factors (Aaronson and Horvath, 2002). BRCA1 may directly interact with JAK1 and JAK2, and overexpression of *BRCA1* has been shown to result in constitutive activation of STAT3 in a prostate cancer cell line (Gao, 2001), as well as upregulation of *JAK1* and *STAT1* in a human embryonic kidney cell line (Welcsh, 2002).

Biochemical fractionation experiments demonstrate that BRCA1 and BARD1 co-purify with the RNA Pol II holoenzyme (Scully, 1997a; Neish, 1998; Chiba and Parvin, 2002), a polymerase complex including RNA Pol II and basal transcription factors such as TFIIH. Depending on the purification method, this complex may also include DNA repair factors such as RAD51, the Ku heterodimer, and replication protein A (Maldonado, 1996), but not sequence-specific transcription factors (Ossipow, 1995). BRCA1 and BARD1 have been shown to interface with the holoenzyme though an interaction with RNA Helicase A (Anderson, 1998). However, recent work has indicated that BRCA1 appears to associate with the hyper-phosphorylated form of RNA Pol II, the processive form of the enzyme, suggesting that BRCA1 may have a role in transcriptional elongation rather than (or in addition to) initiation (Krum, 2003). Further studies are still needed to define the consequences of the interaction of BRCA1 and RNA Pol II.

BRCA1 and BARD1 may also be involved in a DNA-damage–induced transcription block. BARD1 interacts with CstF-50, a subunit of the cleavage-stimulation-factor complex which plays a role in the 3' end cleavage of mRNA precursors prior to polyA tail synthesis (Kleiman and Manley, 1999). UV light exposure or a DNA replication block induced by HU treatment results in a temporary block in cleavage, during which the amount of cellular CstF50/BRCA1/BARD1 complex increases. BARD1 is likely to be necessary

for this interaction, as no damage-induced elongation block occurs if BARD1 is mutated (Kleiman and Manley, 2001).

A few experiments have suggested links between BRCA1 and chromatin proteins and/or the chromatin remodeling complex. BRCA1 appears to interact with a component of the SWI/SNF chromatin-remodeling complex (BRG1), suggesting a possible mechanism for transcriptional activation (Bochar, 2000), and it also interacts with a DEAH-type helicase called BACH (BRCA1-associated C-terminal helicase), which may play a role in DSBR (Cantor, 2001). The phosphorylated version of the chromatin protein H2AX co-localizes with BRCA1 after DNA damage (Chen, 2002). Yeast two-hybrid screens have provided limited evidence for an interaction between BRCA1 and the histone deacetylase proteins HDAC1 and HDAC2 (Yarden and Brody, 1999). As discussed earlier, HMGA1, a non-histone chromatin protein, may downregulate BRCA1 (Baldassarre, 2003). There is also evidence that BRCA1 may have a role in the organization of the non-coding RNA *Xist* on the inactivated X-chromosome (Ganesan, 2002). While the interaction of BRCA1 with chromatin remodeling proteins may be an indication of a physical linkage of BRCA1 with chromatin remodeling, the potential functional consequences of these interactions require further study.

1.9 BRCA1 MOUSE KNOCKOUT MODELS

1.9.1 Knockout alleles of *Brca1*

Although much has been learned about *BRCA1* through studying human cell lines, tumours, and breast cancer families, modeling *BRCA1*-related breast cancer in the mouse was expected to reveal the functions of the gene and allow tumourigenicity to be carefully studied. To date, several *Brca1* mouse knockout models have been generated (Table 1.3). For those knockouts which generate functionally null alleles, embryonic lethality occurs in homozygous mutants between embryonic day (E) 5.5- E10.5. Heterozygotes are normal and healthy with no increased predisposition to cancer, and double-targeted ES cells cannot be generated (Hakem, 1996; Liu, 1996;

Table 1.3: Mouse *Brca1* knockouts.

#	Description of allele	Phenotype of homozygous mutation	Secondary mutations or backgrounds	Other phenotypes/ Results of tumourigenesis studies	References
1	Deletion of exon 2 (part of the RING domain)	Embryonic lethality by E5.5-E9.5	*Brca2*$^{-/-}$ or *Bard1*$^{-/-}$ background: no change in *Brca1*$^{-/-}$ phenotype. *p53*$^{-/-}$background: *Brca1*$^{-/-}$ embryonic development extended by one day.	*Brca1*$^{+/-}$: no increased predisposition to tumourigenesis compared to wildtype mice.	(Ludwig, 1997; McCarthy, 2003)
2	Deletion of exons 5 and 6	Embryonic lethality by E7.5	*p21*$^{-/-}$ or *p53*$^{-/-}$backgrounds: either extended *Brca1*$^{-/-}$ embryonic development by one day.	*Brca1*$^{+/-}$: no increased predisposition to tumourigenesis compared to wildtype mice.	(Hakem, 1996; Hakem, 1997)
3	Conditional, *loxP*-flanked deletion of exons 5 and 6	Viable mice	T-cell specific Cre.	Decreased proliferation of T-cells in *Brca1*$^{c/c}$; T-cell Cre mice.	(Mak, 2000)
4	Deletion of part of 5' end of exon 11 and part of intron 10 (ΔX.11)	Embryonic lethality by E8.5		*Brca1*$^{+/-}$: no increased predisposition to tumourigenesis compared to wildtype mice.	(Liu, 1996)
5	Deletion of part of 5' end of exon 11 and part of intron 10 (ΔX.11)	Embryonic lethality between E10.5- E 13.5	*p53*$^{-/-}$ background (for tumourigenesis studies).	*Brca1*$^{+/-}$: no increased predisposition to tumourigenesis compared to wildtype mice. *Brca1*$^{+/-}$; *p53*$^{-/-}$: tumour latency and spectrum not different from *p53*$^{-/-}$ mice; 20 of 22 tumours screened retained a wildtype allele of *Brca1*.	(Gowen, 1996; Cressman, 1999)
6	Deletion of part of 5' end of exon 11 and part of intron 10 (ΔX.11)	Embryonic lethality between E7.5- E9.5		*Brca1*$^{+/-}$: no increased predisposition to tumourigenesis compared to wildtype mice.	(Shen, 1998)
7	Deletion of exon 11 (Cre-excision of *loxP*-flanked exon in ES cells; ΔX.11)	Embryonic lethality between E12.5-E18.5, although ~2% of homozygotes survive to adulthood.	*p53*$^{+/-}$ or *p53*$^{-/-}$ background (for tumourigenesis studies).	18/66 (27%) *Brca1*$^{-/-}$; *p53*$^{+/-}$ mice developed thymic lymphomas by 28 weeks; all had lost wildtype allele of *p53*. 14/14 (100%) *Brca1*$^{-/-}$; *p53*$^{-/-}$ developed thymic lymphomas by 15 weeks.	(Xu, 1999b; Xu, 2001; Bachelier, 2003)

Table 1.3: Mouse *Brca1* knockouts, continued.

Abbreviations: Apc^{min}: adenomatous polyposis coli (multiple intestinal neoplasia); *K-14: Keratin-14*; *K-5: Keratin-5*; MMTV: mouse mammary tumour virus-long terminal repeat promoter; *WAP*: whey acidic protein, gene promoter; c: conditional allele; E: embryonic day.

#	Description of allele	Phenotype of homozygous mutation	Secondary mutations or backgrounds	Other phenotypes/ Results of tumourigenesis studies	References
8	Conditional, *loxP*-flanked deletion of exon 11	Viable mice	MMTV-Cre, *WAP*-Cre, or *K5*-Cre transgenes. $p53^{+/-}$ or *K5-E2F1* backgrounds (for tumourigenesis studies).	$Brca1^{c/c}$; Cre or $Brca1^{-/c}$; Cre: 35 of 150 mice developed tumours by 2 years. $Brca1^{c/c}$; $p53^{+/-}$; Cre or $Brca1^{-/c}$; $p53^{+/-}$; Cre: almost all of 56 animals developed tumours by 15 months. Most had lost the second *p53* allele. $Brca1^{-/c}$; *K5*-Cre: 13 of 18 mice developed tumours, mostly oral epithelial or inner ear canal by 20 months. $Brca1^{-/c}$; *K5*-Cre; *K5-E2F1*: tumourigenesis accelerated over latter class. Skin or reproductive tract tumours (only 5 mice).	(Xu, 1999a; Brodie, 2001; Berton, 2003)]
9	Exon 11 truncation mutation (*tr*)	Viable mice, males infertile.	$p53^{+/-}$ or $p53^{-/-}$ backgrounds (for tumourigenesis studies). Background strain determined percentage of viability.	$Brca1^{tr/tr}$: wide range of tumours developed in 76 of 86 mice, including 12 mammary tumours. Mean latency ~17 months. $Brca1^{tr/tr}$; $p53^{+/-}$ or $Brca1^{tr/tr}$; $p53^{-/-}$: possible acceleration of *p53*-related tumourigenesis, very small cohort.	(Ludwig, 2001)
10	Conditional, *loxP*-flanked deletion of exons 9-13	Viable	*p53* conditional background (*loxP*-flanked deletion of exons 2-10) and *K14*-Cre transgene.	$Brca1^{c/c}$; *K-14*-Cre; $p53^{c/c}$: 11 of 11 mice developed tumours, mean latency 6 months, no wildtype *p53* or *Brca1* detected. $Brca1^{+/c}$; *K-14*-Cre; $p53^{c/c}$: 8 of 8 mice developed tumours, mean latency 11 months. No wildtype *p53* detected, but all tumours had at least one wildtype *Brca1* allele.	(Jonkers, 2003)
11	Deletion of exons 20-24 (last BRCT repeat)	Embryonic lethality by E10.5	$p53^{+/-}$, $p53^{-/-}$, or $Apc^{min/+}$ backgrounds (for tumourigenesis studies).	$Brca1^{+/-}$: no increased predisposition to tumourigenesis compared to wildtype mice. $Brca1^{+/-}$; $p53^{+/-}$ or $Brca1^{+/-}$; $p53^{-/-}$: tumorigenesis not accelerated compared to $p53^{+/-}$ or $p53^{-/-}$ alone. $Brca1^{+/-}$; $Apc^{+/min}$: tumorigenesis not accelerated compared to $Apc^{+/min}$ alone.	(Hohenstein, 2001)

Ludwig, 1997; Shen, 1998; Hohenstein, 2001). Embryonic lethality appears to result from growth suppression and not from apoptosis (Hakem, 1996; Liu, 1996; Xu, 1999c), although embryos homozygous for one mutation, which survive until E10.5, exhibit apoptosis at day E9.5 (Hohenstein, 2001).

In general, the time of embryonic arrest appears to reflect the region of the gene that is deleted. However, even careful characterization of growth arrest in embryos homozygous for the same knockout allele revealed that the time varied (Hakem, 1996; Liu, 1996; Ludwig, 1997). *Brca1*$^{-/-}$ blastocyst outgrowth in culture was generally poor (Liu, 1996; Ludwig, 1997; Shen, 1998), and these blastocysts were hypersensitive to γ-irradiation (Shen, 1998). Tumourigenesis studies of *Brca1*$^{+/-}$ mice on a *p53*$^{-/-}$ or *p53*$^{+/-}$ background revealed that even when mice were given whole-body γ-irradiation, tumour latency was not accelerated compared to similarly-treated mice lacking only *p53*. Tumours resulting in these γ-irradiated, *Brca1*$^{+/-}$, *p53*$^{+/-}$ or *Brca1*$^{+/-}$, *p53*$^{-/-}$ mice lacked both copies of *p53*, but generally retained a wildtype allele of *Brca1* (Cressman, 1999a; Cressman, 1999b; Hohenstein, 2001). Spectral karyotyping (SKY) analysis, in which each chromosome of a metaphase spread is "painted" with a different colour (Liyanage, 1996; Schrock, 1996b), of E9.5 MEFs from one null knockout (Table 1.3 #6) revealed that *Brca1*$^{-/-}$ MEFs had rearranged, abnormal karyotypes which were exacerbated on a *p53*$^{-/-}$ background (Shen, 1998). More details on the genomic changes in *Brca1*-associated tumourigenesis in the mouse will be discussed in section 1.9.5.

1.9.2 Double knockout models

To try and functionally rescue *Brca1*$^{-/-}$ embryos, several groups have crossed mice carrying *Brca1* knockout alleles onto other knockout backgrounds. While partial rescue of embryonic lethality is achieved on a *p53*$^{-/-}$ or *p21*$^{-/-}$ background, this rescue only extends embryonic development for one additional day (Hakem, 1997; Ludwig, 1997). The mouse knockout of *Bard1* is phenotypically very similar to that of *Brca1*, and the interdependence of these two genes is supported by the fact that

Brca1$^{-/-}$*, Bard1*$^{-/-}$ embryos are indistinguishable from *Brca1*$^{-/-}$ embryos (McCarthy, 2003). *Brca1*$^{-/-}$*, Brca2*$^{-/-}$ mice also have the same phenotype as *Brca1*$^{-/-}$ mice (Ludwig, 1997).

1.9.3 A humanized model of *Brca1*

Despite the fact that the overall identity of the mouse and human BRCA1 proteins is only 57%, human *BRCA1* is able to functionally rescue loss of *Brca1* in the mouse (Lane, 2000; Chandler, 2001). Chandler *et al.* generated transgenic mice carrying a bacterial artificial chromosome (BAC) which included the entire human *BRCA1* gene. These mice were mated to mice heterozygous for a *Brca1* knockout allele (Table 1.3 #11). Backcrossing to the *Brca1*$^{+/-}$ mice resulted in viable *Brca1*$^{-/-}$ mice which also carried the BAC. These mice were normal and healthy, and had no increased incidence of tumours up to 18 months, at time of publication (Chandler, 2001). This BAC transgenic model was modified slightly in a second paper, in which the human BAC used to rescue the *Brca1*$^{-/-}$ genotype carried a biologically relevant human mutation, introduced into the BAC by recombineering (Copeland, 2001), to generate a "humanized" mouse model. The mutation (T64G) mimics a common, cancer-related, RING-domain mutation, and was unable to rescue the *Brca1*$^{-/-}$ mice; *Brca1*$^{-/-}$ embryos with or without the BAC were indistinguishable. The "humanized" model was additionally used to determine that the mutation in the BAC-borne *BRCA1* gene causes aberrant splicing, resulting in premature termination of the BRCA1 protein. This BAC-rescue model demonstrates how such a "humanized" mouse model may provide valuable information about the molecular consequences of human mutations (Yang, 2003).

1.9.4 Alternative models suggest that loss of *Brca1* may not be sufficient for tumourigenesis

The goal of generating mouse models of *BRCA1*-related tumourigenesis was not met by the knockouts described above; thus, alternative alleles of *Brca1* were generated to try and develop such a model. Xu *et al.* generated a

conditional *Brca1* allele in which exon 11 was flanked by *loxP* sites (Table 1.3 #7). Cre-mediated deletion of this allele in ES cells generated a *Brca1* $\Delta X.11$ mutation. Despite the fact that the same group had previously generated an exon 11 knockout model which exhibited early embryonic lethality (Table 1.3 #6), mice homozygous for this newer, Cre-excised *Brca1* $\Delta X.11$ allele developed until E12.5-E18.5 (Shen, 1998; Xu, 1999c). The explanation for the difference in onset of growth arrest was that a *Neomycin phosphotransferase* (*Neo*) cassette had been left in intron 10 in the previous knockout, and this resulted in premature truncation of all *Brca1* transcripts (Brodie and Deng, 2001). *Brca1*$^{\Delta X.11/\Delta X.11}$ embryos still arrested but MEFs could be generated more easily from the older embryos. *Brca1*$^{\Delta X.11/\Delta X.11}$ MEFs were hypersensitive to γ-irradiation, lacked a G2-M checkpoint, and 25% of them had more than two centrosomes (Xu, 1999c). This indication of a potential role for BRCA1 in centrosome biology is supported by more recent evidence suggesting that BRCA1 interacts with α, β, and γ-tubulin, and co-localizes with tubulin at the centrosomes. Exon 11 appears to be important for this interaction, which is consistent with the phenotype of the Brca1$^{\Delta X.11/\Delta X.11}$ MEFs (Hsu and White, 1998; Deng, 2002; Lotti, 2002). More recent studies using this allele have demonstrated that *Brca1*$^{\Delta X.11/\Delta X.11}$ mice on a *p53*$^{+/-}$ background are viable, although they are prone to thymic lymphomas early in life (18 of 66 mice died by 28 weeks of age). Lymphoma formation appears to depend on the *p53* mutation, as all tumours investigated had lost the wildtype *p53* allele. *Brca1*$^{\Delta X.11/\Delta X.11}$ mice on a *p53*$^{-/-}$ background also developed thymic lymphomas (14 of 14 by ~15 weeks of age); this is more quickly than would be expected in *p53*$^{-/-}$ mice (additionally, *p53*$^{-/-}$ mice develop other types of tumour (Donehower, 1992)). This appears to indicate a role for mutated *Brca1* as an accelerant for *p53*-related tumourigenesis (Xu, 2001b; Bachelier, 2003).

The *Brca1*$^{\Delta X.11}$ conditional allele (*co*) was also used to generate conditional *Brca1* mice (one null allele and one conditional $\Delta X.11$ allele: *Brca1*$^{-/co}$), which were crossed to mice carrying a Cre transgene driven by one of two breast-specific promoters (Table 1.3 #8). These promoters were from whey acidic protein (WAP), a milk protein expressed in mammary epithelium during

pregnancy and lactation (Piletz and Ganschow, 1981; Robinson, 1995), and the MMTV-LTR (mouse mammary tumour virus LTR), which has been shown not to be breast-specific, but is expressed in breast epithelium and ductal cells (Wagner, 1997). Conditional mice carrying a Cre transgene showed developmental abnormalities of the mammary gland, but only a few breast tumours were observed in these animals after a long latency. Complete loss of *Brca1* did not appear to have occurred in most tumours. The addition of a $p53^{+/-}$ or $p53^{-/-}$ background to the conditional system accelerated tumourigenesis, but while tumours lacked both copies of *p53* and had other genetic alterations (described further in section 1.9.5), most retained the $Brca1^{co}$ allele, still functioning as a wildtype allele (Xu, 1999b), indicating that the loss of *Brca1* alone was not sufficient for tumourigenesis (Brodie, 2001; Weaver, 2002).

The same $Brca1^{co/-}$ conditional mice have also been used in conjunction with a Cre transgene driven by the bovine *keratin 5* (*K5*) promoter (Berton, 2003). In mice, this promoter is active in epithelial tissues of the oral and sinus cavities, esophagus, bladder, prostate, and vagina, as well as in the basal layer of the epidermis (Ramirez, 1994). Berton *et al.* observed that 72% (13 of 18) *K5*-Cre, $Brca1^{co/-}$ mice developed tumours by ~22 months of age, mainly in the inner ear canal or oral cavity. In a smaller study, tumourigenesis was accelerated by overexpression of *E2F1*, although the resulting tumours mainly occurred in the epidermis (Berton, 2003). The reasons for the effect of *E2F1* were not fully explored, but the link between E2F1 and BRCA1 was described earlier, in section 1.5.4. *Brca1* is overexpressed in *K5-E2F1* transgenic mice, likely due to the E2F1-responsive site in the *Brca1* promoter (Wang, 2000a; Berton, 2003). However, with or without *E2F1* overexpression, none of the mice developed mammary tumours, and tumourigenesis occurred only after a long latency.

Ludwig *et al.* generated a *Brca1* truncation mutation which mimics a mutation observed in human *BRCA1*-related breast tumours (Table 1.3 #9). Mice homozygous for this mutation were viable, but on a mixed 129Sv and C57BL/6 background, only 4% of the mice recovered from a heterozygous

intercross were homozygous for the mutation (Ludwig, 2001). Backcrossing to 129Sv mice, or outcrossing to the MF1 strain corrected the percentage to the expected 25%, although homozygous mutant males were infertile. Homozygous animals developed a variety of tumours with a mean latency of 17 months. While some animals did develop breast tumours, a wide variety of other tumour types were also observed. All tumours appeared to have secondary mutations as demonstrated by a change in gene product levels. This, and the long tumour latency, strongly suggested that while loss of *Brca1* could contribute to tumourigenesis, it was not in itself sufficient for tumourigenesis. Further, the variety of tumours suggested that Brca1 may not be a tissue-specific tumour suppressor in the mouse (Ludwig, 2001).

Jonkers *et al.* utilized mice co-conditional for both *Brca1* (exons 9-13 flanked by *loxP* sites) and *p53* (exons 2-10 flanked by *loxP* sites) (*Brca1*$^{c/c}$*, p53* $^{c/c}$ – see also Table 1.3 #10), which also carried a Cre transgene driven by the *Keratin-14* promoter (*K14*-Cre), active in skin and breast epithelia (Jonkers and Berns, 2003). Normally, tumourigenesis studies using a *p53*$^{-/-}$ background are compromised by the tumours these mice develop at an early age; in this study the *p53* mutation was conditional to reduce this problem. Of the 11 tumours analyzed from *K14*-Cre*, p53*$^{c/c}$, *Brca1*$^{c/c}$ mice, all had undergone recombination of both copies of *p53* and both copies of *Brca1*. In *K14*-Cre*, p53*$^{c/c}$, *Brca1*$^{+/c}$ mice, median tumour latency was roughly two times longer (330 days, versus 180 days), but of the 8 tumours analyzed, all had lost both copies of *p53* and retained at least one wildtype copy of *Brca1*. Breast and skin tumours developed in equal numbers (Jonkers and Berns, 2003). The conclusions were similar to those of the previous studies: *Brca1* loss was neither necessary nor sufficient for tumourigenesis, and the tissue-specificity of human *BRCA1*-related cancer was not mimicked.

1.9.5 Additional alterations in *Brca1*-related mouse tumours

To date, only a small number of murine *Brca1*-related breast tumours have been generated, and only limited data detailing the genetic changes in these tumours has been published. However, two groups have reported on the

presence or absence of a limited number of protein products in the mammary tumours from two different mouse models described above: the *Brca1*$^{co/-}$ mouse model, and the exon 11 truncation model (Table 1.3 #8 and 9) (Brodie, 2001; Ludwig, 2001; Weaver, 2002). In the former, a few of the tumours analyzed were from mice which also had a *p53*$^{+/-}$ background. Table 1.4 profiles some of the findings from these analyses. It is interesting to note that for some factors, such as the loss of the oestrogen and progesterone receptors, the mouse models mimic the characteristics of human *BRCA1*-related tumours. However, for some genes, such as Cyclin D1, the protein product is generally lost in human *BRCA1*-related tumours but not in murine ones. For other proteins such as ErbB2, the two mouse models did not agree, perhaps because the mice carry different alleles (Johannsson, 1997; Armes, 1999; Brodie, 2001; Ludwig, 2001; Lakhani, 2002; Weaver, 2002). The level of aneuploidy is generally higher in human *BRCA1*-related breast tumours compared to sporadic breast tumours (Johannsson, 1997). While structural and genomic abnormalities clearly occur in *Brca1*-related mouse tumours, no publication has related the amount of aneuploidy to that of sporadic control tumours (Xu, 1999b; Weaver, 2002).

These studies demonstrate that, whatever the role of *BRCA1* in human cancer, loss of *Brca1* in the mouse is not in itself sufficient for tumourigenesis. Moreover, the tissue specificity of human tumours is not fully recapitulated in the mouse. The reason for this difference (or the reason for the tissue specificity of *BRCA1*-related human tumours) is not clear. However, the choice of promoter for Cre expression in the conditional mouse models may be in part responsible. Arguments have been made against using milk-protein promoters for expression of breast-cancer genes, as these promoters are hormonally regulated, so would likely not be expressed in mammary stem cells. They also tend not to be expressed in ductal cells (a site of normal *Brca1* expression), which are thought to be a common site of tumourigenesis (Marquis, 1995; Rijnkels and Rosen, 2001). On the other hand, more widely-expressed (or ubiquitous) promoters may result in unwanted effects on other tissues. Promoters specific to mammary stem cells would be ideal, but none have yet been identified (Rijnkels and Rosen, 2001; Medina, 2002).

Table 1.4: Characteristics of mammary tumours in *Brca1* knockout mice, compared to human *BRCA1*-related tumours.

Models used are described in Table 1.3 (ΔX.11 conditional model, see Table 1.3 #8; tumours here were from *Brca1*$^{c/c}$; MMTV-Cre (n=6) or *Brca1*$^{c/c}$; MMTV-Cre; *p53*$^{+/-}$ (n=3) animals; *Brca1* truncation model, see Table 1.3 #9; all tumours from *Brca1*$^{tr/tr}$ animals).

Protein/characteristic	ΔX.11 *Brca1* conditional knockout model (n=9)	*Brca1* exon 11 truncation model (n=10)	Similar to human *BRCA1*-related tumours?	References
Loss of oestrogen receptor (ER-α) expression	7	5 (+4 downregulated)	Yes	(Armes, 1999; Brodie, 2001; Ludwig, 2001; Lakhani, 2002)
Loss of progesterone receptor (PR) expression	Not measured	7	Yes	(Armes, 1999; Ludwig, 2001; Lakhani, 2002)
Loss of ErbB2/neu expression	2	6 (+1 downregulated)	Mouse models not consistent; human tumours generally have lost expression.	(Armes, 1999; Brodie, 2001; Ludwig, 2001; Lakhani, 2002)
Expression of Cyclin D1 detected	9	10	No	(Armes, 1999; Brodie, 2001; Ludwig, 2001)

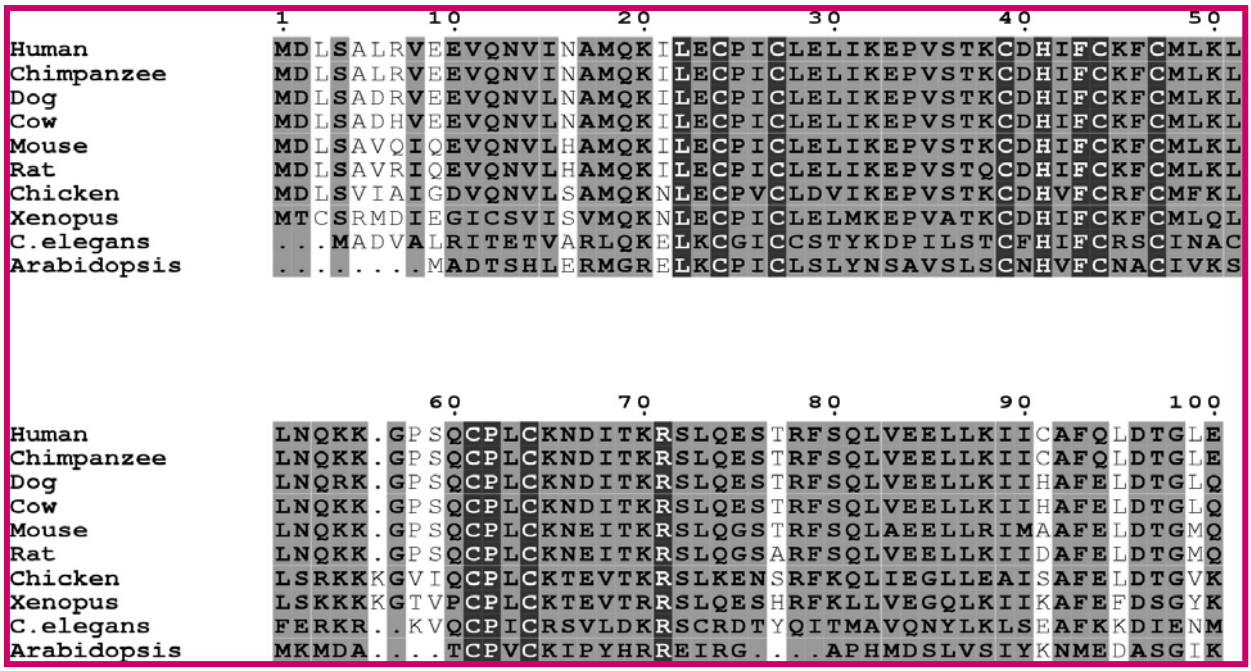
However, while *BRCA1*-related breast cancer may be difficult to model using knockout alleles of *Brca1* in mice, the underlying mechanisms for tumourigenesis – *i.e*., the caretaker roles of BRCA1 in cell cycle control and the response to DNA damage – can be very adequately investigated using these mice. In fact, it is here that the underlying similarity between not only the mouse and human, but also more evolutionarily distant homologues, is most compelling.

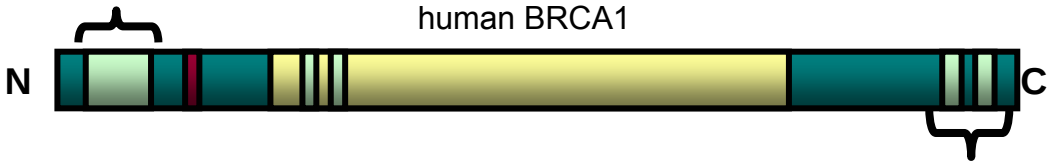
1.10 NON-MAMMALIAN HOMOLOGUES OF *BRCA1*

Homologues of human *BRCA1* were first thought to exist only in other mammals, but *BRCA1* homologues have now been identified in chicken (Orelli, 2001), *Xenopus* (frog) (Joukov, 2001b), *C. elegans* (Boulton, 2004), *Arabidopsis thaliana*, and *Oryza sativa* (rice) (Lafarge and Montane, 2003). All these homologues were identified by their well-conserved RING domain and BRCT repeats (Figure 1.11). In fact, in *C. elegans* or the plants, these domains comprise the majority of the protein (Lafarge and Montane, 2003; Boulton, 2004). While the overall similarity between the *Xenopus* or *C. elegans* homologues and the human BRCA1 protein is not as high as that of the mouse-human similarity, an alignment of the RING domains from several homologues from human to *Arabidopsis* shows that the key C and H residues of the RING motif are absolutely conserved (Figure 1.11). A *BARD1* homologue has also been identified in *Xenopus*, chicken, and *C. elegans* (Joukov, 2001b; Orelli, 2001; Boulton, 2004).

The conservation of BRCA1-related phenotypes amongst the different species is striking. In *C. elegans*, RNA-interference (RNAi, which decreases the expression of the target gene) of *BRC-1* or *BRD-1* (the *C. elegans BRCA1* and *BARD1* homologues) results in cell-cycle checkpoint–independent apoptosis, which is increased in response to γ-irradiation. BRC-1 also forms damage-induced nuclear foci (Boulton, 2004). Antisense-mediated depletion of x*BRCA1* or x*BARD1*, the *Xenopus* homologues of *BRCA1 and BARD1*, results in severe developmental defects in later-stage embryos, non-viable frogs, and higher levels of aneuploidy in cells (Joukov, 2001b). *Arabidopsis*

a.





b.

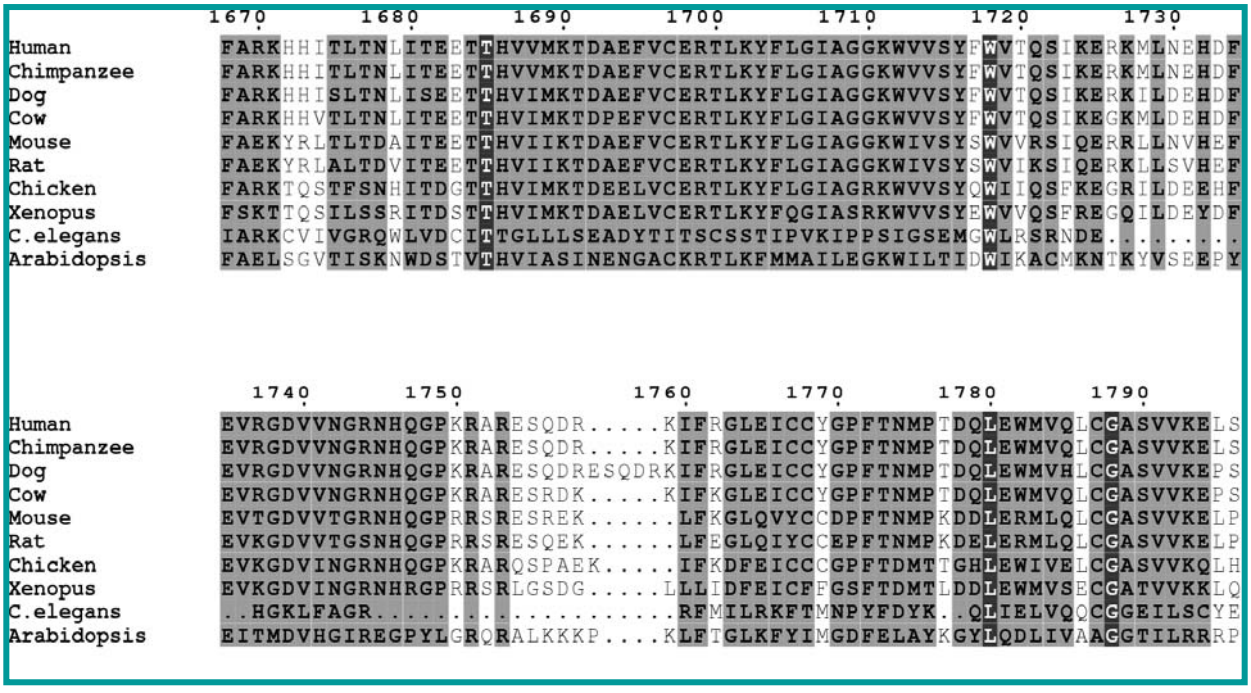


Figure 1.11: Multiple alignments showing conservation of the RING domain and BRCT repeats. a. RING-domain alignment for indicated species. Identical residues are in dark grey, similar ones in light grey. Note that the key cysteine (C) and histidine (H) residues of the zinc-finger are absolutely conserved. **b.** BRCT repeats, aligned as above.
ClustalW and ESPript (Blosum62 matrix) were used to produce alignments.

thaliana BRCA1 is upregulated in plantlets in response to γ-irradiation and appears to be upregulated at the same time as the *Arabidopsis Rad51* gene (Lafarge and Montane, 2003). On the whole, these phenotypes are strongly reminiscent of those seen in *Brca1* knockout mice or in *BRCA1*-deficient mouse or human cell lines.

1.11 DNA REPAIR

1.11.1 Focal indication of a role for BRCA1 in the response to DNA damage

BRCA1 clearly plays roles in various types of DNA damage repair. Following exposure to γ-irradiation (which mainly causes double-stranded breaks), MMC (a DNA cross-linking agent), UV light, or a HU-induced DNA replication block, normal S phase BRCA1 foci disappear (Scully, 1997b; Zhong, 1999; Wu, 2000). BRCA1 reappears in foci later as soon as an hour after damage, and these damage-induced foci may persist for 8-12 hours post-damage.

Damage-induced foci do not have the same composition as BRCA1 S phase foci; after damage, BRCA1 is hyperphosphorylated on various residues and the foci appear to include a different subset of proteins (Scully, 1997b; Wang, 2000b). While proteins like RAD51, BRCA2, and BARD1 localize with BRCA1 in both S phase and damage-induced foci, PCNA (proliferating cell nuclear antigen), which forms BRCA1-independent foci during S phase, co-localizes with BRCA1 following damage, presumably at replication forks (Scully, 1997b; Scully, 1997c; Chen, 1998; Wang, 2000b). Other proteins which co-localize with BRCA1 following DNA damage include the RAD50-MRE11-NBS1 complex involved in DSBR, the histone protein H2AX (Celeste, 2002), FANCD2 (D'Andrea and Grompe, 2003), BLM, and other proteins which comprise the BASC (Wang, 2000b). In mouse $Brca1^{\Delta X.11/\Delta X.11}$ MEFs, which lack the Rad51-interaction domain coded by exon 11, Rad51 foci do not form following γ-irradiation (Huber, 2001). However, the mutant form of *BRCA1* in the human cancer cell line HCC1937 does carry the RAD51-interaction domain, and RAD51 foci form normally following γ-irradiation (Zhong, 1999).

RAD50 foci also appear to form normally in cells lacking BRCA1 (Wang, 2000b; Wu, 2000).

1.11.2 Damage-induced phosphorylation of BRCA1

BRCA1 is phosphorylated normally during certain parts of the cell cycle, but phosphorylation also occurs following DNA damage. Some studies have indicated that damage-induced phosphorylation may depend on factors such as cell cycle stage and the dose of damage received (Scully, 1997b; Okada and Ouchi, 2003). Several kinases may phosphorylate BRCA1 in response to DNA damage, the two key ones being ATM (Ataxia telangiectasia mutated) and ATR (ATM and RAD3-related), both members of the phosphatidylinositol 3-kinase (PI3K) family. These two kinases phosphorylate many cell-cycle proteins and are important both in normal checkpoint control and the response to DNA damage (Shiloh, 2001). In humans, the recessive disorder ataxia-telangiectasia (AT) results from a loss of *ATM*, and is characterized by neuronal degeneration (cerebellar ataxia), sterility, and a greatly increased cancer risk. Cell lines from AT patients show chromosomal breakage and telomere instability and are hypersensitive to ionizing radiation (but not to base-damaging agents such as UV light) (reviewed in Thompson and Schild, 2002). No human disorder has been attributed to mutations in the *ATR* gene. The corresponding mouse knockout models mimic the human conditions fairly well; homozygous *Atm* knockout mice are infertile, hypersensitive to ionizing radiation, and succumb early in life to thymic lymphomas (Barlow, 1996; Xu, 1996), while homozygous *Atr* knockout embryos die early in embryonic development, before E7.5 (Brown and Baltimore, 2000).

Evidence that ATM and ATR phosphorylate BRCA1 has come from a number of experiments, including investigating the phosphorylation status of endogenous BRCA1 before and after DNA damage in AT cells or in cells constitutively expressing a dominant-negative form of ATR (Cortez, 1999; Gatei, 2001). Phosphorylation of overexpressed, tagged BRCA1 protein in such cells has also been monitored using phosphorylation-specific BRCA1 antibodies; these studies helped to define the residues targeted by each

kinase (Tibbetts, 2000; Gatei, 2001). The roles of ATM and ATR are not absolutely delineated, but the current understanding is that ATM appears to phosphorylate BRCA1 following γ-irradiation, but not after UV exposure (Cortez, 1999; Gatei, 2000; Tibbetts, 2000). ATR phosphorylates BRCA1 following UV exposure, and also may phosphorylate BRCA1 to some extent following γ-irradiation (Tibbetts, 2000; Gatei, 2001). The actual mechanisms are more complex than this summary indicates, as kinases downstream from ATM, or possibly independent kinases may also phosphorylate BRCA1 (Ruffner and Verma, 1997; Altiok, 1999; Lee, 2000; Foray, 2002). However, the PI3K kinase DNA-PKcs (DNA protein kinase, catalytic subunit), which plays a major role in both non-homologous end joining (NHEJ) and V(D)J (variable(diverse)joining) recombination of T- and B-cell receptor genes, does not appear to phosphorylate BRCA1 following DNA damage (Scully, 1997b).

Both ATM and ATR may have other links with BRCA1; ATM and BRCA1 have been shown to interact by co-immunoprecipitation (ATM is part of the BASC), and co-localize to some damage-induced nuclear foci (Cortez, 1999; Tibbetts, 2000; Wang, 2000b). In addition, BRCA1 may be necessary for a certain subset of ATM and ATR phosphorylation activities, as abnormal phosphorylation of targets of these kinases in HCC1937 cells, which lack wildtype BRCA1, is observed. Phosphorylation of some of these targets returns to normal upon expression of wildtype BRCA1 (Foray, 2003).

The clearest conserved phenotype of BRCA1 and its homologues is a role in the response to DNA damage. Since DNA damage can be broadly classified into double-strand breaks and damage to bases (Figure 1.12), these two categories will be considered separately. Evidence indicates that BRCA1 plays roles in repairing both types of damage.

1.11.3 Double-strand break repair

1.11.3.1 An overview of DNA double-strand break repair

DNA double-strand breaks (DSBs) may be caused such agents as γ-irradiation, free radical attack, or strand crosslinks. These lesions are

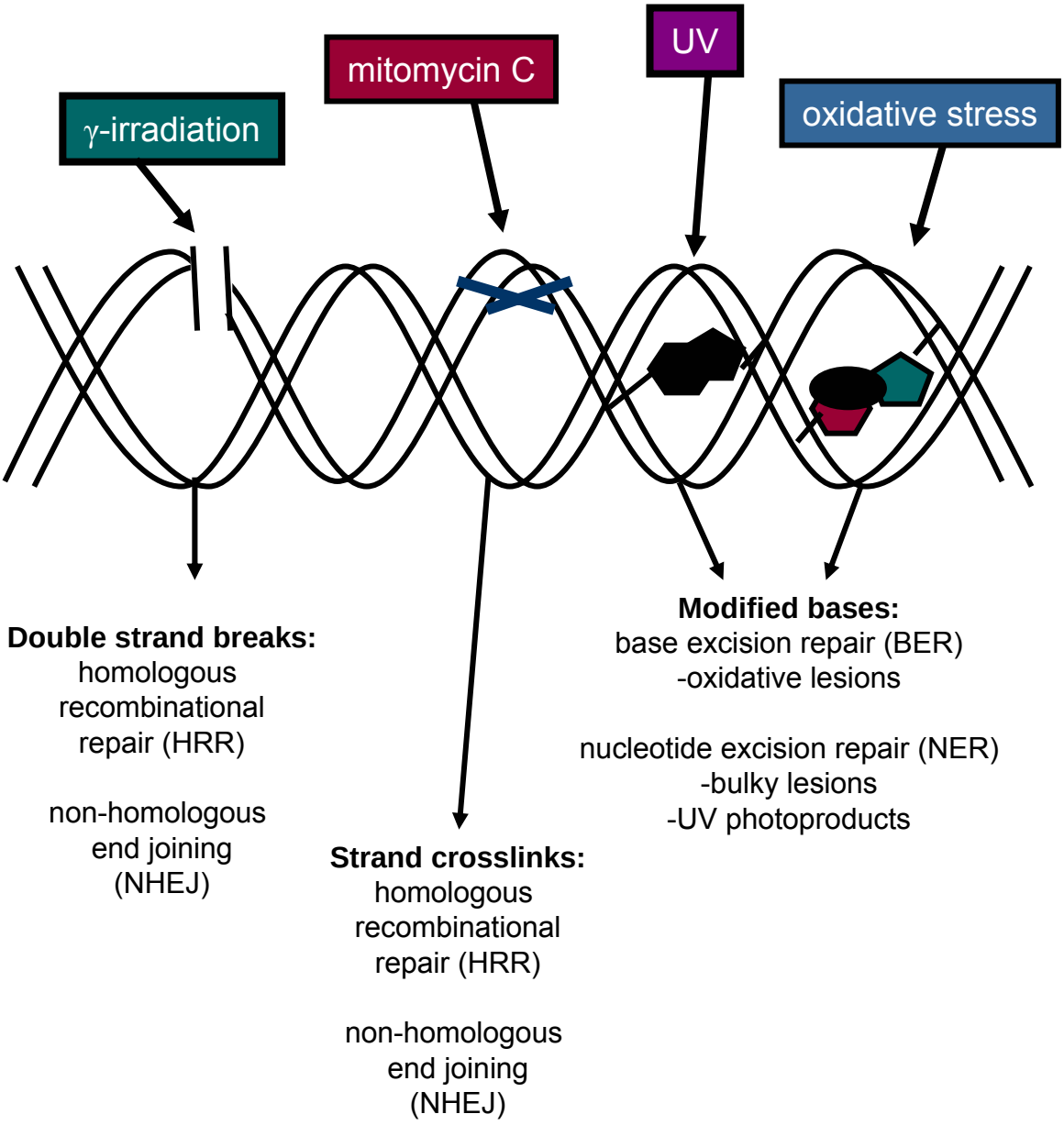


Figure 1.12: Representative types of DNA damage and primary repair pathways for each. Schematic showing common mutagens used in damage-repair research, their characteristic lesions, and the pathways used to repair them.
UV=ultraviolet light irradiation.

repaired by two major pathways: homologous recombinational repair (HRR) and non-homologous end joining (NHEJ) (Figure 1.13). HRR utilizes a homologous chromosome or sister chromatid to precisely repair the chromosome, while NHEJ simply rejoins the break in a sequence-independent manner. HRR is a high-fidelity repair process, while NHEJ often is not; exonuclease activity at break ends can cause a loss of genetic information, and the mechanism of NHEJ itself generally results in a small insertion or deletion at the break site. Many of the proteins which play key roles in detection and repair of strand breaks were first discovered through yeast screens for repair-deficient strains, and a large number of these proteins have homologues in higher eukaryotic organisms (reviewed in Chu, 1997; Thompson and Schild, 2001).

1.11.3.1.1 The balance between NHEJ and HRR

Mammalian cells were once thought to primarily repair DSBs through NHEJ, but recent evidence suggests that HRR is involved in repairing at least 30-50% of DSBs in mammals (Liang, 1998). The impetus to use HRR or NHEJ for repair appears to depend on several factors, including the stage of the cell cycle. Studies using *Rad54*$^{-/-}$ (HRR deficient) or *Ku70*$^{-/-}$ (NHEJ deficient) chicken DT40 cells showed that *Rad54*$^{-/-}$ cells are γ-irradiation sensitive in late S and G2 phases, while *Ku70*$^{-/-}$ cells are γ-irradiation sensitive in G1 and early S phases. Cells lacking both proteins are more sensitive to γ-irradiation than the single mutants, which suggests that these pathways do not fully complement one another (Takata, 1998; Wang, 2001b). This may be due to a dependence on the availability of an appropriate substrate for HRR, as the use of HRR in G1 would result in LOH of the repaired area. This is supported by evidence showing that in mouse ES cells, a sister chromatid is used more often than a homologue for HRR (Johnson and Jasin, 2000).

Kinetically, double-strand break repair (DSBR) occurs in a biphasic manner – there is a “fast” component of repair over the first 30-60 minutes which mends some 80% of breaks, and a “slow” phase which works on the remaining breaks over the next 24 hours (Figure 1.14). Repair of genomic DNA is generally monitored using pulsed-field gel electrophoresis, which allows

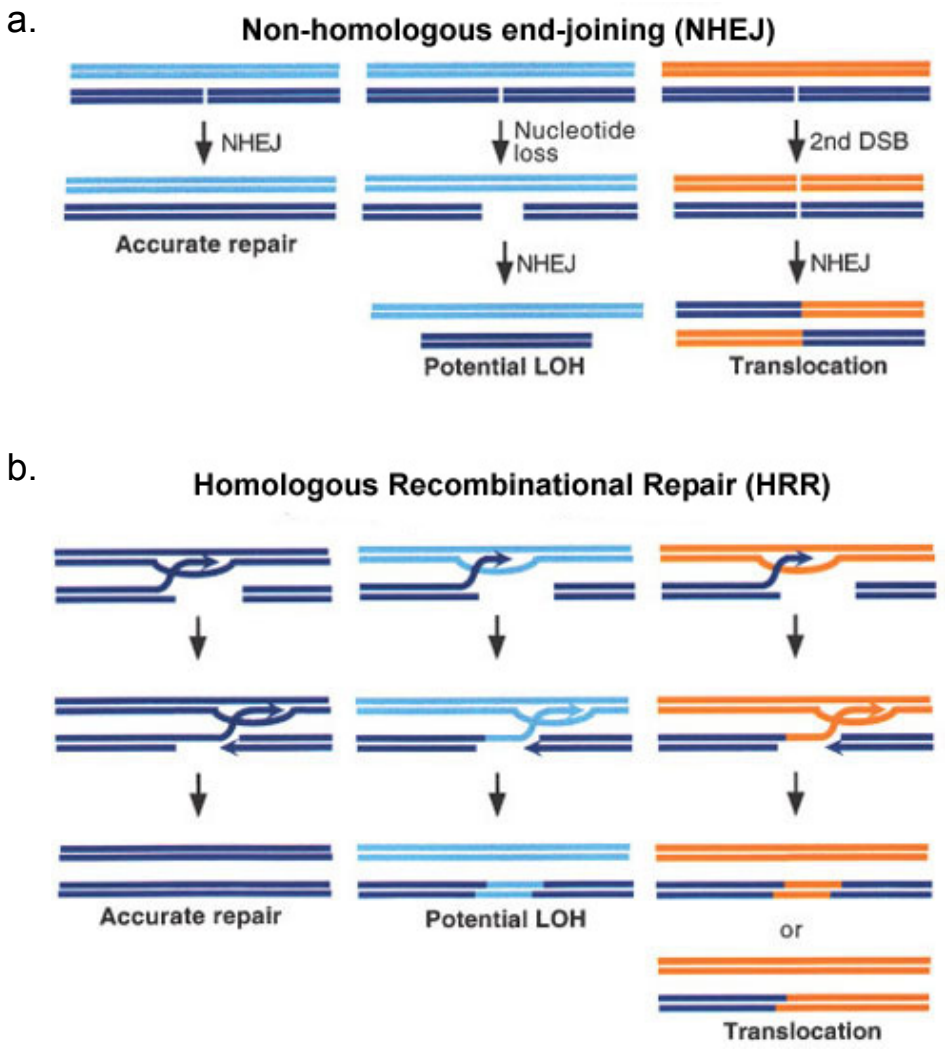


Figure 1.13: Non-homologous end joining (NHEJ) and Homologous Recombinational Repair (HRR) of double-strand breaks. a. NHEJ rejoins a double-strand break in a sequence-independent manner. Three possible outcomes, including loss of genetic information, are shown. **b.** HRR uses a homologous chromosome or sister chromatid to accurately repair a double-strand break. Depending on the substrate used for repair, loss of heterozygosity (LOH) may occur, but generally a net loss of genetic information does not occur. Figure modified from (Ferguson and Alt. Oncogene 20:5572, 2001.

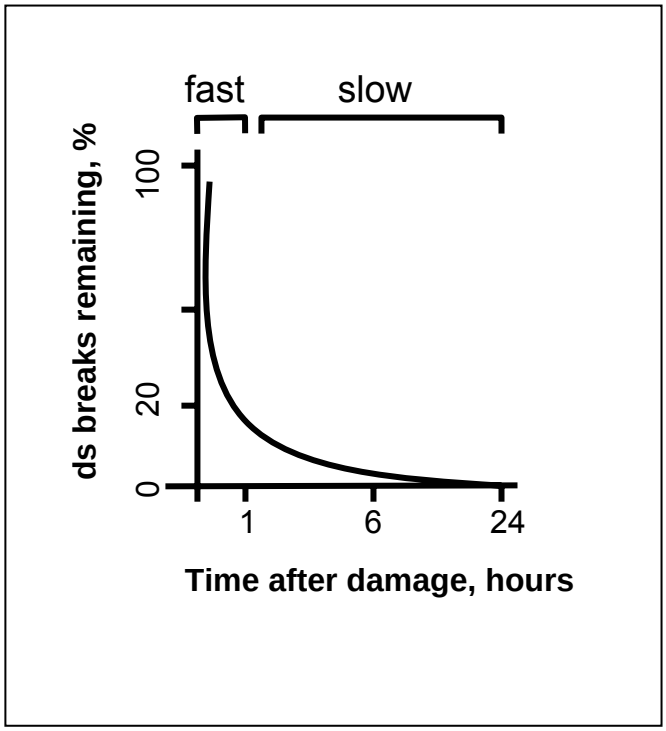


Figure 1.14: **Double-strand break repair kinetics.**
Typical double-strand break repair kinetics showing “fast” repair over the first hour post-damage and “slow” repair following.
ds: double-strand.

separation of very large pieces of DNA; repaired DNA runs more slowly than broken DNA, allowing the percentage of faster-running, broken DNA to be measured over time. Generally, even the "slow" phase is virtually complete by six hours post-damage. Most evidence indicates that "fast" repair is done by NHEJ and the "slow" repair by HRR (reviewed in Biedermann, 1991; Iliakis, 1991; DiBiase, 2000). For example, cells lacking DNA-PKcs (a key protein in NHEJ), are hypersensitive to γ-irradiation. Their "fast" and "slow" repair kinetics do not change, but a much smaller proportion of breaks are repaired by the "fast," or NHEJ, component of repair. The γ-irradiation hypersensitivity of these cells is hypothesized to result from the increased fraction of breaks left for the "slow" component to repair, despite the fact that the "slow" component is working normally (Iliakis, 1991; DiBiase, 2000). These kinetics may help explain the presence of BRCA1 foci in normal cells after most DNA repair has already taken place. While most DSBs are mended in the first couple of hours following DNA damage, the "slow" component of repair may continue to work on breaks for some hours past that time.

1.11.3.1.2 Early Cellular Responses to DSBs

The cellular response to double-strand breaks begins with recognition or detection of lesions, followed by the triggering of downstream repair/reaction events. The kinases ATM and ATR (discussed in section 1.11.2) appear to be key players in the recognition of strand breaks; they phosphorylate and activate an overlapping but distinct set of targets, triggering repair or other downstream processes in response to breaks (reviewed in Jackson, 2002). The global importance of ATM is underscored by the fact that human A-T cells are deficient in ionizing-radiation-induced G1-S, intra-S, and G2-M phase checkpoints (Lavin and Shiloh, 1997). ATM is activated minutes after damage, and can itself bind DNA ends (reviewed in Thompson and Schild, 2001). Other damage-recognition factors may also exist.

Other early events in repair are likely to include changes in histone proteins to help relax chromatin structure around the break to allow for access by repair proteins. In mammals, this includes the phosphorylation of the histone protein H2AX, an early event following exposure to damaging agents (Rogakou,

1998). A an increase in cellular deoxribonucleotides used for repair-related synthesis is also likely to occur (Tanaka, 2000).

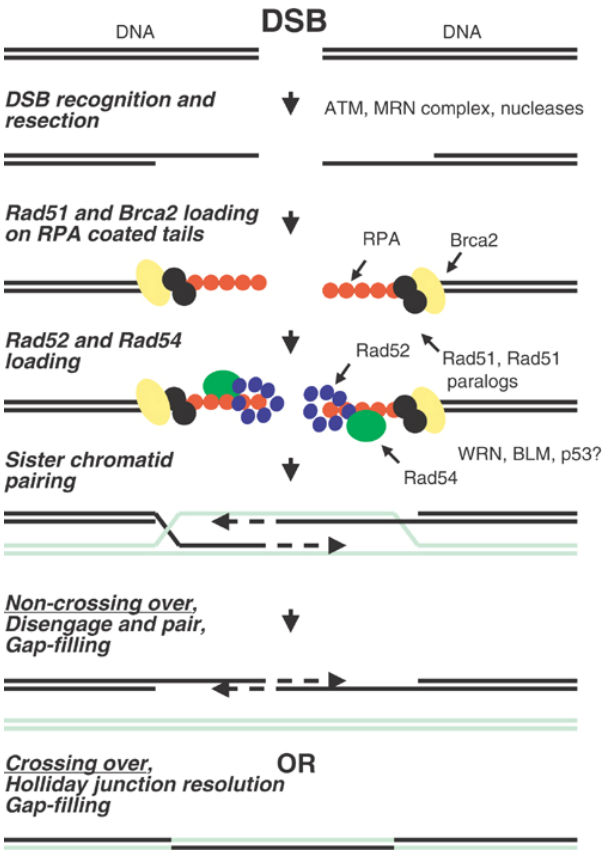
In addition to triggering repair-related pathways in response to DNA damage, the ATM and ATR kinases may also induce cell-cycle arrest or delay (the key cell-cycle checkpoint proteins CHK1 and CHK2 are phosphorylated and activated by ATR and ATM, respectively) to allow additional time for repair. If the damage is overwhelming an apoptotic pathway may be triggered, possibly by ATM and/or ATR: both directly phosphorylate the p53 protein, and ATM may also indirectly induce accumulation of p73 to induce apoptosis (reviewed in Dasika, 1999; Bernstein, 2002; Thompson and Schild, 2002; Iliakis, 2003).

1.11.3.1.3 Proteins Involved in Homologous Recombinational Repair

Homologous recombination repair is thought to begin with resection of the DSB into a single-stranded 3' overhang which invades a double-stranded homologous region (see Figure 1.15a for an overview of the process). Resection may involve the mammalian RAD50-MRE11-NBS1 complex - RAD50 is thought to bind DNA, and the complex has both exo- and endo-nuclease activities *in vitro*, as well as helicase activity. However, the MRE11 nuclease tends to generate 5' overhangs, leading some to suggest that this complex may play more of a an "organizer" role in break repair (reviewed in Thompson and Schild, 2001).

A key protein in HRR is RAD51, a mammalian orthologue of the bacterial RecA protein. RAD51 forms a filament on ssDNA overhangs, and its ability to hydrolyze ATP appears to be necessary for recombination (Thompson and Schild, 2001; Jackson, 2002; Thompson and Schild, 2002). Initially, the ssDNA region is likely to be coated with the heterotrimeric Replication Protein A (RPA), a protein with high affinity for ssDNA which is involved in replication, DNA repair, and recombination (Wold, 1997). Binding of RPA to the ssDNA region is important for the formation of an even coating of RAD51 along the ssDNA, as the helix-destabilizing properties of RPA minimize secondary structure, especially when the region of ssDNA is long (Treuner, 1996; Sung, 2003). Since RPA has a high affinity for ssDNA, RAD51 alone displaces RPA

a.



b.

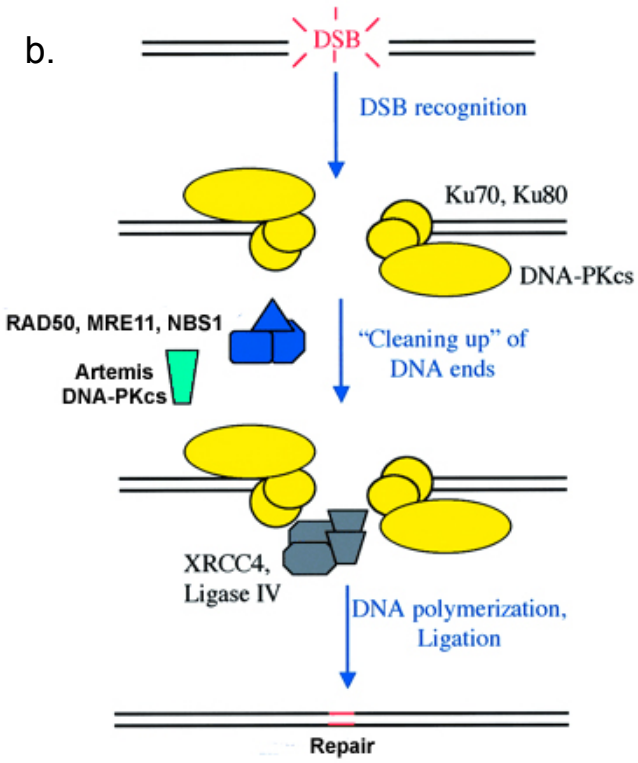


Figure 1.15: Proteins and mechanisms involved in HRR and NHEJ double-strand break repair. a. Homologous recombinational repair, diagram showing key proteins and mechanisms involved. Note the dual possible outcomes of the Holiday junction following strand invasion. Figure taken from (Valerie, 2003).
b. Proteins and mechanism (simplified) of NHEJ, showing recognition of the break site, binding of the DSB end by the Ku proteins, end modification, and religation of the break site. Figure modified from (Jackson, 2002).

very slowly. RAD52 interacts with both RPA and RAD51, and greatly accelerates the displacement of RPA by RAD51 (Sugiyama and Kowalczykowski, 2002; Symington, 2002; Kantake, 2003), although BRCA2 (which also interacts directly with RAD51) also appears to aid in loading or organization of RAD51 on ssDNA (Yang, 2002). RAD52 is thought to be one of the key factors in the "decision" to repair a break by HRR or NHEJ, as it forms a multimeric complex which will bind to double-stranded ends, and may compete with the NHEJ-related Ku heterodimer to bind free DNA ends (reviewed in Symington, 2002). Loss of RAD52 in yeast results in severe HRR defects, although loss of Rad52 in mice is not a lethal event (unlike loss of Rad51) (Lim and Hasty, 1996; Rijkers, 1998).

Once the RAD51-ssDNA nucleoprotein filament is formed and a homologous dsDNA section has been identified, strand invasion occurs, with displacement of one strand appearing as a D-loop. RAD52 and RAD54 assist in the invasion of the RAD51-coated ssDNA into a homologous section of dsDNA; RAD54 may increase the efficiency of ss/ds DNA pairing and appears to have helicase activity. A DNA polymerase extends the 3' terminus of the invading strand, copying the information from the homologous partner, and the break is ligated by DNA ligase I (Thompson and Schild, 2001). Migration and resolvation of the Holiday junction may then occur, although, in mammalian cells, an alternative non-crossover pathway may predominate in which the Holiday junction disengages (illustrated in Figure 1.15a). This mechanism would protect against LOH at the region, but further experimental work is needed to define how frequently this pathway is used and if novel proteins are involved in this mechanism (Johnson and Jasin, 2000).

The mechanism of HRR is more complex than this summary indicates. In addition to RAD51, there are several mammalian RAD51 paralogues (RAD51B, C, and D, XRCC2, and XRCC3) which appear to be involved in strand invasion and junction resolvation, as well as a number of other proteins which are likely to be involved either in regulation or mechanism (reviewed in Thompson and Schild, 2001; Thompson and Schild, 2002).

1.11.3.1.4 Proteins Involved in Non-homologous end-joining

Several of the core proteins involved in mammalian NHEJ also mediate V(D)J recombination in immune cells; the results of a protein deficiency in one of these proteins often leads to severe immunodeficiency disorders as well as problems with repair of DSBs (Lieber, 2003).

Key elements in NHEJ include the Ku heterodimer (Ku70/Ku80), which forms a ring-type structure and threads onto free DNA ends (see Figure 1.15b for an overview of the NHEJ process). As mentioned above, the Ku proteins may compete with RAD52 to bind to free ends. The Ku heterodimer is likely to be the localization signal for DNA-PKcs, a key kinase in NHEJ and in V(D)J recombination which is appears to be activated by the presence of DNA breaks. The importance of this protein is underscored by the fact that the scid (severe combined immunodeficiency) mutation in mice and humans results from lack of DNA-PKcs (Hendrickson, 1991; Thompson and Schild, 2001). Once activated, the targets of DNA-PKcs are likely to include XRCC4, which interacts with and stimulates the activity of DNA ligase IV, responsible for ligation of the DSB (reviewed in Dasika, 1999).

As in HRR, processing of the break region generally occurs. Break sites which cannot be directly re-ligated are often sites for limited addition/deletion of bases. Additionally, NHEJ often occurs at an area of microhomology (1-4 identical bases). As a result of microhomology joining, gaps or overhangs are often left which must be filled in or removed (Lieber, 2003). The RAD50-MRE11-NBS1 complex may be responsible for NHEJ-related processing; not only does the this complex possesses exo- and endonuclease activity, but yeast strains which lack these proteins (Rad50, Mre11, or Xrs2 – the third yeast protein) are deficient in NHEJ (Jackson, 2002). However, a second processing complex in vertebrates has been more recently identified, composed of DNA-PKcs and the protein Artemis. Artemis mutations are found in a subset of human scid patients who have an increased sensitivity to ionizing radiation (Moshous, 2001). The DNA-PKcs/Artemis complex has endonucleolytic activity at 3’ and 5’ overhangs, as well as the ability to open hairpins (important for V(D)J recombination) (Ma, 2002; Lieber, 2003).

Despite their importance in mammalian cells, Artemis and DNA-PKcs have only been identified in vertebrates to date, in contrast to the Ku proteins, which have homologues in all eukaryotes examined (Jackson, 2002; Lieber, 2003). It is not clear at present whether this indicates that other kinases fill the role of DNA-PKcs in non-vertebrates. Scid mice are not only deficient in V(D)J recombination, they are hypersensitive to γ-irradiation, which argues that DNA-PKcs is necessary for NHEJ in vertebrates (Berton, 2003). It is possible that in non-vertebrates, the RAD50-MRE11-NBS1-type complex is responsible for end-processing of double-strand break ends prior to NHEJ, while in vertebrates, the DNA-PKcs/Artemis complex can fill this role in addition to its vital role in processing intermediates during V(D)J recombination (Lieber, 2003).

The actual process of HRR or NHEJ is more complex – and is likely to involve more complex interactions than these summaries indicate. Additionally, unanswered questions remain, including whether a polymerase is involved in NHEJ or not (Lieber, 2003). In recent years, the involvement of a large number of other proteins involved in the process or regulation of DSBR has been shown. A number of these proteins are implicated in human cancer syndromes, such as the Bloom's Syndrome helicase BLM, the Fanconi anemia FANC proteins, BRCA2, and BRCA1 (which will be discussed further in the next sections) (reviewed in Thompson and Schild, 2002). The study of these proteins in mice is compounded by the fact that mouse knockouts of the genes are often embryonic lethal (*Rad50, Rad51, Rad51b, Rad51d, Atr, Xrcc4, Brca1*, *Brca2*, and *DNA ligase IV*), precluding extensive studies of the effects of loss of the gene product on the whole organism (reviewed in Dasika, 1999; Thompson and Schild, 2001). Other knockouts, such as those of *Ku*, *Atm*, *Rad54,* or DNA-PKcs/scid, are viable, but the mice or cells are often hypersensitive to double-strand break-inducing agents, have immune defects, or, in the case of Ku knockout mice, show signs of premature aging (Hendrickson, 1991; Barlow, 1996; Xu, 1996; Thompson and Schild, 2001; Jackson, 2002).

1.11.3.2 BRCA1 and Homologous Recombinational Repair (HRR)

Evidence that BRCA1 is involved in HRR of DSBs comes from several sources, including its interactions with proteins involved in HRR, the altered repair kinetics of cells lacking BRCA1, and a decreased ability of *Brca1*$^{-/-}$ mouse ES cells to successfully integrate a targeting cassette or repair a break via homologous recombination.

As described earlier, BRCA1 interacts with numerous proteins implicated in homologous recombination, including RAD51 (Scully, 1997c), ATM (Cortez, 1999; Gatei, 2000; Tibbetts, 2000), BLM (Wang, 2000b), and two components of the RAD50-MRE11-NBS1 complex: the Nijmegen breakage syndrome protein NBS1, and RAD50 (Varon, 1998; Wang, 2000b).

Several groups have examined the response of the HCC1937 human cell line (which carries only a C-terminal truncated version of BRCA1) to γ-irradiation (Table 1.5, experiments 1-3). In all three cases, "fast" repair of breaks during the first hour after damage was equally efficient in HCC1937 cells and controls (Abbott, 1999; Foray, 1999; Scully, 1999). Two groups showed that the remaining breaks were competently repaired only by the control cells; at either 6 or 24 hours post-damage, HCC1937 cells had not finished DSBR (Foray, 1999; Scully, 1999). However, a third group demonstrated that all the γ-irradiation-induced breaks in HCC1937 cells and controls were repaired by four hours post-damage. In this case, the majority of DSBs were repaired within the first two hours, which may have left too few breaks to allow an accurate measurement of the rate of the "slow" component (Abbott, 1999). HCC1937 cells were also hypersensitive to γ-irradiation when compared to controls (Abbott, 1999; Foray, 1999). These three studies indicate that cells lacking BRCA1 may have a defect in the "slow", or homologous recombinational, component of repair – albeit using a cell line which is known to have a number of mutations besides the one in *BRCA1* (Tomlinson, 1998).

The existence of a mouse *Brca1*$^{-/-}$ ES cell line provides a simple way of measuring homologous recombination efficiency, namely gene targeting. Two conventional gene-targeting vectors integrated correctly into *Brca1*$^{-/-}$ cells at a

Table 1.5: Repair response of HCC1937* cells to double-strand breaks.

#	Cell line(s) used	γ-irradiation dose	Repair, one hour after damage	Repair, six hours after damage	Repair, 24 hours after damage	γ-irradiation hypersensitivity?+	Reference
1	HCC1937+Ad*BRCA1-GFP* (adenovirally expressed at near-endogenous levels)	8 Gy	50%	Nearly complete	Complete	No, compared to HCC1937 control	(Scully, 1999)
	HCC1937+Ad*GFP* (empty adenoviral vector)		50%	~60%	~75%	Yes	
2	HCC1937	30 Gy	~50%		58%	Yes	(Foray, 1999)
	Lymphoblast cell line (wildtype)		~50%		Nearly complete	No	
3	HCC1937	10 Gy	Most (no numbers given)	No breaks observed after 4 hours		Yes	(Abbott, 1999)
	HCC1937 + *BRCA1* lacking aa 702-843 (this mutant does not cause growth suppression when overexpressed)		Most (no numbers given)	No breaks observed after 4 hours		No	
4	HCC1937	40 Gy	92%	98%	Complete	Not measured	(Wang, 2001)
	HCC1937 + 30 μM wortmannin (to inhibit NHEJ)		78%	83%	89%	Not measured	

* The HCC1937 human cell line carries only a C-terminal mutated version of BRCA1.
+ γ-irradiation hypersensitivity compared to control cell line shown for each experiment unless otherwise noted. Gy=gray (dose of γ-irradiation).

significantly lower rate than into wildtype cells (an estimated 13-fold reduction, once corrected for the 4-fold increase in random integration) (Moynahan, 1999). The use of a cassette containing an I-SceI site (cleavage by the nuclease SceI at an I-SceI site generates a DSB) designed to distinguish between HRR or NHEJ repair events revealed a 5- to 6-fold decrease in the amount of HRR, and a 1.5-fold increase in NHEJ (Moynahan, 1999). A third study using a similar I-SceI-containing reporter cassette again demonstrated a significant decrease in HRR efficiency in *Brca1*$^{-/-}$ cells (Moynahan, 2001). All of these experiments utilized ES cells homozygous for a *Brca1* Δ*X.11* allele (Table 1.3 #5) (Gowen, 1996).

Moynahan *et al.* also showed that the homologous recombination efficiency of a *Brca1*$^{+/-}$ ES cell line is identical to that of wildtype cells (Moynahan, 2001). Similar studies using human cancer cell lines demonstrated a heterozygous effect; that is, a heterozygous cell line showed an intermediate phenotype as compared to wildtype and *BRCA1*-deficient cell lines (Abbott, 1999; Foray, 1999; Baldeyron, 2002). A caveat must be added to these results: the *Brca1*$^{\Delta X.11/\Delta X.11}$ ES cell line used in these studies was the only double-targeted line ever recovered by this group (Gowen, 1996), and the addition of a *Brca1*-containing transgene could not fully rescue either the HRR defect or hypersensitivity to MMC. However, retargeting a wildtype allele back into the *Brca1* locus fully rescued hypersensitivity to MMS. This indicates that recovery of double-targeted ES cells is a rare event, but rescue of the phenotype by retargeting suggests that there is not a secondary mutation in this cell line which affects the *Brca1*-related phenotype (Moynahan, 2001).

Overall, these data from both human and mouse cell lines indicates that BRCA1 does indeed play a role in HRR.

1.11.3.3 BRCA1 and non-homologous end joining (NHEJ)

While a good deal of evidence argues that BRCA1 is involved in HRR, evidence related to its role in NHEJ is not as clear or consistent.

The HCC1937 cell line exhibited normal “fast” repair in all three studies described above (Table 1.5, experiments 1-3). Two additional studies using wortmannin, an inhibitor of the NHEJ-related DNA-PKcs kinase, further suggest that HCC1937 cells do not have an NHEJ defect. HCC1937 cell extracts were as efficient as control cell extracts in mediating end rejoining in an *in vitro* assay (Merel, 2002). Additionally, the repair kinetics of HCC1937 cells following γ-irradiation in the presence or absence of wortmannin showed that wortmannin inhibition of DNA-PKcs (and thus NHEJ) meant that cells repaired fewer breaks in the first hour (78% vs. 92%), but that “slow” repair occurred with unchanged kinetics (Table 1.5, experiment 4). DSBR was virtually complete by 24 hours, regardless of the presence of wortmannin, likely because the “fast” component of repair was not fully inhibited and still repaired the majority of breaks (Wang, 2001a). In summary, the experiments in Table 1.5 suggest that HCC1937 cells do not have a deficiency in NHEJ. Indeed, a recent experiment using the HCC1937 cell line indicates that there may be an increase in NHEJ in these cells (using an assay for random plasmid integration) which is restored to wildtype levels on expression of a *BRCA1* transgene, suggesting that BRCA1 may normally function in suppression of NHEJ in favor of HRR (Zhang, 2004).

NHEJ proteins are instrumental in V(D)J recombination of immune cells, as evidenced by the phenotype of scid mice, which lack the instrumental DNA-PKcs kinase involved in NHEJ and do not develop mature T or B cells (Blunt, 1995; Kirchgessner, 1995). V(D)J recombination in T and B cells is slightly different: the genes used to generate the T- or B-cell receptors differ, but the same set of proteins carry out the mechanics of recombination (Gellert, 2002). A conditional *Brca1* mouse was generated which carried one null and one *loxP*-flanked version of exons 5 and 6 of *Brca1* (*Brca1*$^{co/-}$) and a Cre transgene driven by a T-cell specific promoter from the tyrosine kinase gene *Lck* (Table 1.3 #3) (Hakem, 1996; Mak, 2000). The authors speculated that if *Brca1*-deficient cells had a defect in NHEJ, these mice would have a reduced or absent number of mature T-cells. *Brca1*$^{co/-}$, *Lck*-Cre mice did have a 90% reduction in T-cell numbers, but V(D)J recombination appeared to be unaffected (Mak, 2000). Additionally, a second group studying the

development of lymphomas in *Brca1*$^{\Delta X.11/\Delta X.11}$*, p53*$^{-/-}$ mice (Table 1.3 #7) have shown that although these mice develop tumours at an early age, mature T and B cells are not depleted, and V(D)J recombination appears to occur normally (Xu, 2001b; Bachelier, 2003).

However, one group has persistently documented a decreased efficiency of NHEJ in *Brca1*$^{-/-}$*, p53*$^{-/-}$ MEFs generated from E 9.5 *p53*$^{-/-}$*, Brca1*$^{\Delta X.11/\Delta X.11}$ embryos (*Brca1* allele from Table 1.3 #4). In a cell-free end-joining assay, extracts from these *Brca1*$^{-/-}$*, p53*$^{-/-}$ MEFs were less efficient at end-rejoining than an extract from *p53*$^{-/-}$ MEFs, and end-joining could be impaired in a wildtype cell extract by addition of antibodies against Brca1 (Zhong, 2002a). The same MEFs have been used in a variety of other NHEJ assays: an I-SceI reporter-cassette assay, an assay to monitor the re-annealing of a linearized plasmid, and a retroviral infection assay (Zhong, 2002b). Retroviral infection of cells defective in NHEJ is a cytotoxic event; NHEJ appears to mediate circularization of non-integrated virus, and an inability to circularize non-integrated copies of virus may result in cell death triggered by the presence of excess DNA free ends (Daniel, 1999; Daniel, 2001; Li, 2001). Regardless of the assay, *Brca1*$^{-/-}$*, p53*$^{-/-}$ MEFs showed a significant decrease in NHEJ activity when compared to *p53*$^{-/-}$ MEFs, although some assays indicated that the defect may be in precise end-joining, not overall end-joining (Zhong, 2002a; Zhong, 2002b). While critics may point out that measuring NHEJ efficiency is not as straightforward as measuring HRR efficiency (Ferguson and Alt, 2001), a variety of assays have been performed. One caveat to these experiments is that there is a possibility of additional mutations in the *Brca1*$^{-/-}$*, p53*$^{-/-}$ MEF line. Given the still-conflicting evidence, it is fair to say that more studies on the role of BRCA1 in NHEJ are still required.

1.11.4 Repair of Mutated Bases

1.11.4.1 An overview of base repair

Mutated bases, such as pyrimidine dimers from UV exposure, oxidative lesions such as 8-oxo-guanine from oxygen free-radical exposure, or replication errors, are repaired by three different pathways – base excision

repair (BER), nucleotide excision repair (NER), and mismatch repair (MMR). Although some of the proteins involved in MMR co-immunoprecipitate with BRCA1 (Wang, 2000b), experiments involving BRCA1 to date have focused mainly on BER and NER. These two pathways share some substrates, but differ in their method of action. NER uses a core set of proteins to recognize many lesions, and is generally involved in the repair of bulky lesions such as UV-induced photoproducts. In contrast, BER uses lesion-specific proteins to recognize damage, and tends to repair oxidative lesions. NER is further subdivided into transcription coupled repair (TCR), which preferentially and rapidly repairs the transcribed strand of active genes, and global genomic repair (GGR), which repairs the remainder of the genome and is slower than TCR (Figure 1.16). Recent evidence shows that BER also has a TCR component (reviewed in Svejstrup, 2002).

Repair of mutated bases is an important process for maintaining genome stability, and mutations in repair pathways manifest in clinical syndromes such as Xeroderma Pigmentosum (XP) and Cockayne's Syndrome (CS). XP patients have an overall NER deficiency (except for XPC types, which have only a mutation in GGR, and are competent for TCR), and consequently have a very high incidence of skin cancers from failure to repair UV-induced damage. CS patients, on the other hand, have a specific deficiency in the TCR component of NER, but since they can still repair lesions through GGR, they do not have an increased incidence of skin cancer from UV exposure. However, they do have other, severe, symptoms, likely because of a lack of repair of oxidative lesions on the transcribed strand (TCR-BER) (reviewed in de Boer and Hoeijmakers, 2000; Svejstrup, 2002).

1.11.4.2 BRCA1 and Base Repair

Experimental data implicating BRCA1 in the repair of mutated bases is not as abundant as the DSBR data, but there is evidence for its involvement. BRCA1 is upregulated, hyperphosphorylated, and located in damage-induced foci following UV exposure (Scully, 1997b; Clarkin, 2000; Okada and Ouchi, 2003). Upregulation of BRCA1 appears to result in upregulation of the genes

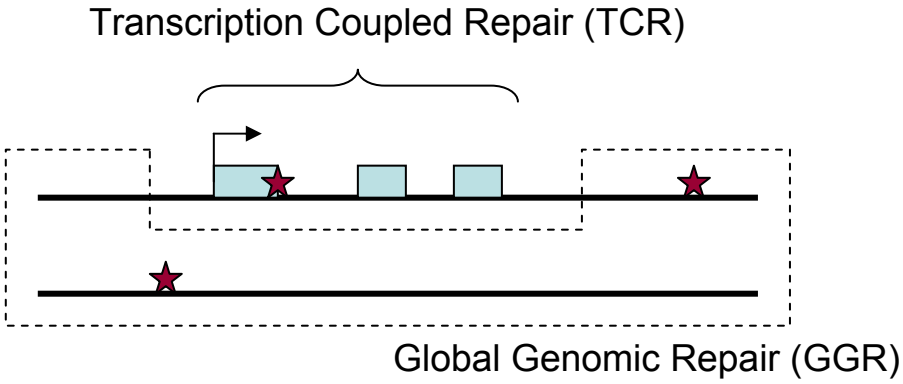


Figure 1.16: Base repair on transcribed or non-transcribed strands of DNA. Schematic showing that mutations (red stars) in actively transcribed genes are repaired by transcription-coupled repair (TCR), while those elsewhere are repaired by the global genomic repair (GGR) pathway.

p21 and *GADD45*, as described in section 1.5.3.1 (Somasundaram, 1997; Amundson, 1998; Harkin, 1999; MacLachlan, 2000b).

The role of BRCA1 in NER has been investigated using tetracycline-controlled overexpression of *BRCA1* in $p53^{+/+}$ or $p53^{-/-}$ human cell lines (Harkin, 1999). In $p53^{-/-}$ cells, UV-induced lesions are repaired efficiently by TCR, but cannot be repaired by GGR. However, when *BRCA1* expression was induced in $p53^{-/-}$ cells, UV-induced lesions on the non-transcribed strand were repaired efficiently. Induction of *BRCA1* in $p53^{+/+}$ cells did not significantly change the amount of repair on either strand. Induction of BRCA1 expression in $p53^{-/-}$ cells led to normal upregulation of the *GADD45* gene following UV exposure. While these results are interesting, this study was based on overexpressing *BRCA1*, and the role of endogenous levels of BRCA1 in this process should still be investigated, especially in light of a recent study which shows that while *BRCA1* is indeed upregulated shortly after UV exposure, expression levels drop again about an hour after UV exposure. Constant expression of *BRCA1* after damage may not accurately model the consequences of this downregulation (Harkin, 1999; Okada and Ouchi, 2003). A second group has shown that mouse $Brca1^{\Delta X.11/\Delta X.11}$, $p53^{-/-}$ MEFs may be hypersensitive to UV exposure when compared to $Brca1^{+/\Delta X.11}$, $p53^{-/-}$ MEFs (allele from Table 1.3 #5); the difference in sensitivity between the two genotypes was statistically significant only at some of the UV doses tested (Cressman, 1999a).

Base mutation can also occur following exposure to oxidative stresses, such as oxygen free-radicals. Common oxidative lesions include thymine glycols and 8-oxo-guanine (Collins, 1999). HCC1937 cells appear to be deficient in TCR of an 8-oxo-guanine lesion transfected into the cells on a plasmid, but are competent to repair the same lesion on the non-transcribed strand (Le Page, 2000b). 8-oxo-guanine lesions are generally repaired by BER, but repair of the mutation in actively transcribed genes occurs through a TCR pathway (Le Page, 2000a). $Brca1^{\Delta X.11/\Delta X.11}$ ES cells (Table 1.3 #5) appear to be hypersensitive to oxidative damage from H_2O_2 exposure, and some evidence indicates that this may result from a deficiency in TCR (Gowen, 1998; Gowen, 2003). $Brca1^{\Delta X.11/\Delta X.11}$, $p53^{-/-}$ MEFs are also hypersensitive to

H_2O_2 exposure when compared to *Brca1*$^{+/\Delta X.11}$, *p53*$^{-/-}$ MEFs (allele from Table 1.3 #5) (Cressman, 1999a).

Evidence from these studies indicates that BRCA1 plays a role in the repair of base damage, but more work is necessary before this role can be precisely defined. In particular, it will be important to investigate the role of BRCA1 in repair when *p53* is not also mutated. Recent evidence suggests that p53 may regulate the level of *BRCA1* expression following DNA damage. The absence of both p53 and BRCA1 may cause synergistic effects which would confound the conclusions of some of these studies (Somasundaram, 1999; MacLachlan, 2000a).

1.11.5 In summary

BRCA1 clearly plays a role in DNA repair. However, the results of the variety of assays (using a myriad of cell lines and BRCA1 mutations) used to reach this conclusion are not without contradiction, and more work is needed to clearly define what roles BRCA1 plays in the repair of base lesions and DNA DSBs.

1.12 THE AIMS OF THIS PROJECT

1.12.1 Existing murine alleles of *Brca1*

A number of mouse *Brca1* knockout alleles and mouse models had been generated when this study began, most designed to investigate the consequences of the loss of *Brca1* on the tumourigenic process. The main findings were that homozygous mutants were embryonic lethal while heterozygotes were normal and had no increased predisposition to tumours (Hakem, 1996; Liu, 1996; Ludwig, 1997; Shen, 1998; Hohenstein, 2001). Embryonic lethality in homozygous mutant ES cells or early embryos compared to tumourigenesis in mature breast and ovarian tissues in human carriers of *BRCA1* mutations can best be described as paradoxical: it seems contradictory but is compatible with the definition of *BRCA1* as a caretaker

tumour-suppressor gene. In mature cells of *BRCA1* mutation carriers, the second *BRCA1* mutation is acquired somatically, probably through LOH. This results in genomic instability leading to mutations in other genes which in turn leads to tumourigenesis. In homozygous *Brca1* mutant embryos, the second *Brca1* hit has already occurred, and the resulting genomic instability is likely incompatible with the massive amounts of growth and differentiation needed to generate a viable mouse. SKY analysis done on embryos homozygous for a *Brca1* Δ*X.11* allele supports this idea by showing that these embryos do indeed have an increased number of chromosomal defects. The addition of a *p53* mutation, which increases the time before $Brca1^{-/-}$ embryos undergo growth arrest, exacerbates the extent of genomic rearrangement. This indicates that the *p53* mutation does not mitigate the *Brca1*-deficient phenotype, but more likely allows damaged cells to bypass a cell-cycle checkpoint for genomic integrity (Shen, 1998).

1.12.2 A conditional *Brca1* ES cell system

These previously-generated mouse models were useful in helping to define the role of *BRCA1* in genomic stability. However, even the conditional mutations of *Brca1*, which were just emerging as this project began, had limited use in revealing the genetic or biochemical pathways behind the caretaker role of *Brca1*. Instead of looking solely at the mouse model, it seemed practical to try and address functional questions in ES cell lines, where additional genetic manipulations could be carried out using familiar and well-tested techniques. As in embryos, *Brca1* appears to be necessary for ES cell viability.

In order to study *Brca1* in ES cells, a conditional *Brca1* system was generated, consisting of one knockout allele and one conditional allele ($Brca1^{co/-}$) (Figure 1.17). These alleles target exon 2 of *Brca1*, both because it contains the translational start site, and because a previous *Brca1* exon 2 knockout allele behaves as a null allele (Table 1.3 #1) (Ludwig, 1997). Targeting exon 2 represented a departure from the numerous groups who produced and were studying exon 11 knockouts. Since *Brca1* Δ*X.11* is a

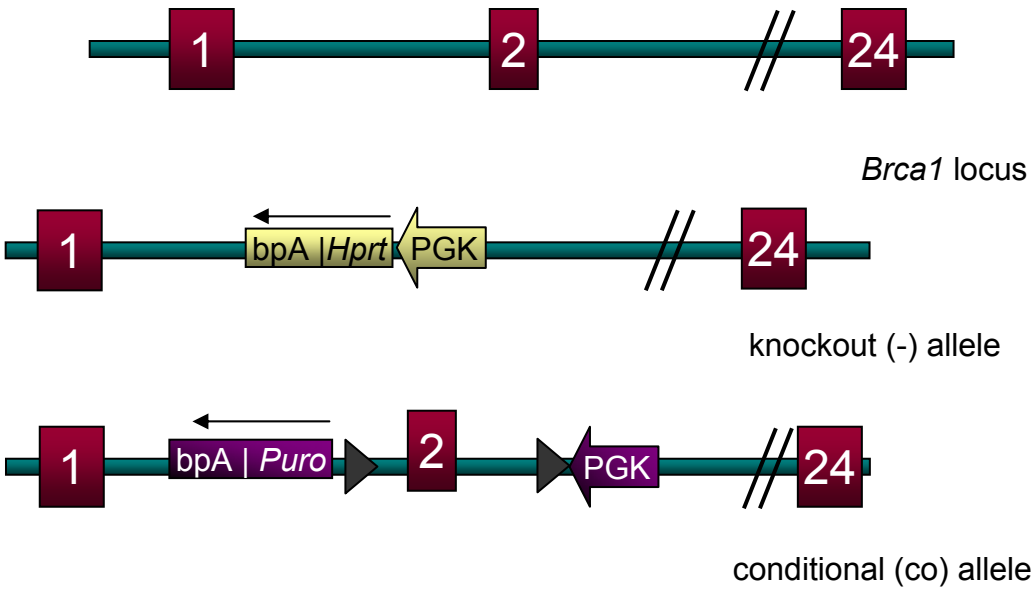


Figure 1.17: Overview of the *Brca1* conditional ES cell alleles. Knockout (-) and conditional knockout (co) alleles of *Brca1*, both targeting exon 2. The knockout allele carries a *Hypoxanthine phosphoribosyltransferase* (*Hprt)* mini-gene which confers HAT resistance, and the conditional allele carries a bipartite *Puromycin* (*Puro)* gene designed to allow puromycin-mediated selection of the allele following Cre-mediated excision. Grey triangles represent *loxP* sites.

natural splice isoform, predicted to share at least some function with full-length *Brca1*, its presence in the ES cells might confuse the findings of a functional screen.

1.12.2.1 A gene trap suppressor screen

Brca1$^{co/-}$ ES cells were primarily generated for use in a genome-wide gene trap screen for suppressors of *Brca1*. There is some support for the notion that suppressors of *Brca1* exist: one mouse study demonstrated that mice homozygous for a truncated Brca1 protein were born at the expected Mendelian ratios only on certain strain backgrounds (Ludwig, 2001). Although a report exists describing a woman who is homozygous for a cancer-related *BRCA1* mutation (Boyd, 1995), this finding has been disputed and attributed to a PCR error (Kuschel, 2001).

The suppressor screen consists of two steps; genome-wide mutagenesis and subsequent selection for viable cells carrying functional suppressors. Before recombinase-mediated deletion of the second copy of *Brca1*, a genome-wide mutagen in the form of a retrovirally-delivered gene trap is introduced into conditional ES cells. This gene trap carries a splice acceptor upstream of an antibiotic resistance gene (*β-geo*, a fusion of the *Neo* gene, which encodes resistance to the drug G418, and *β-galactosidase*) which lacks a translational start site. Integration of the gene trap into the intron of a gene is expected to result in splicing of *β-geo* into the transcript, mutating the gene by truncation, tagging the truncated gene with the inserted trap, and allowing selection and/or screening of trapped cell lines (Figure 1.18).

Following gene trapping, the conditional allele undergoes Cre-mediated deletion. Deletion of the conditional allele is a selectable event, as the conditional allele was designed with a split *puromycin N-acetyltransferase* (*Puro*) selection cassette around exon 2 (Figure 1.17). ES cells which lack both copies of *Brca1* are expected to be non-viable, but if the gene trap cassette traps a suppressor of *Brca1*, cells carrying that trap should be viable. The screening criteria are highly stringent, for antibiotic resistance markers

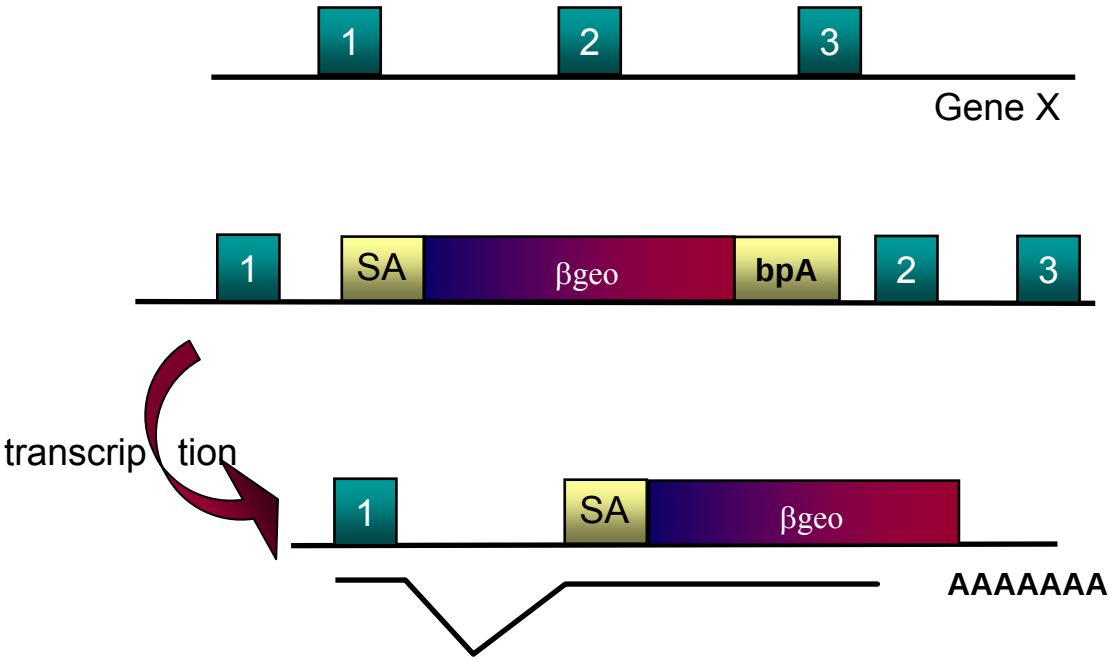


Figure 1.18: Gene trap mutagenesis. A splice-acceptor (SA) *β-geo* (a fusion between the *Neomycin phosphotransferase* (*neo*) gene which codes for resistance to G418 and *β-galactosidase*) gene trap is randomly integrated into the genome. Inclusion of the gene trap cassette in an intron results in splicing of *β-geo* into the transcript, both tagging (with *Neo*) and mutating (through truncation) the trapped gene.

are carried by both alleles of *Brca1* and by the gene trap, and only functional suppressors should result in viable cells (Figure 1.19).

1.12.2.2 Trapping recessive suppressors

One drawback to this screen as described is that it would only be expected to trap dominant suppressors. Therefore, a modification was made to allow screening for recessive genes. A colleague in the lab, Guangbin Luo, has generated a mouse knockout model of Bloom's Syndrome (*Blm*$^{-/-}$). Bloom's Syndrome is a rare recessive syndrome which results from mutation of the RecQ helicase homologue BLM. Mutation of the BLM helicase results in an increased frequency of mitotic recombination and LOH, both in human patients and in a mouse *Blm*$^{-/-}$ model. Guangbin determined that the mitotic recombinational rate in *Blm*$^{-/-}$ ES cells was approximately 20-fold higher than that of wildtype cells (Luo, 2000). This increase in the rate of mitotic recombination can be exploited for screening for recessive mutations: given sufficient doubling times in culture, cells carrying one copy of a gene trap should undergo LOH at that locus. Half of such events should result in homozygosity of the gene trap at the given locus. By these means, recessive suppressors can be trapped in essentially the same screen as the one for dominant suppressors.

The main condition for either suppressor screen was that both knockout alleles of *Brca1* had to behave as null alleles (once fully deleted), in order to allow for selection by cell viability.

1.12.3 Tumourigenesis studies

Besides being used for the suppressor screen, both knockout alleles of *Brca1* were used to generate mice for tumourigenesis studies. The standard knockout was used in conjunction with a Bloom's Syndrome knockout background. It was thought that using this background might accelerate loss of the wildtype copy of the *Brca1* gene and subsequently accelerate tumourigenesis. The conditional allele was used in a tumourigenesis study in

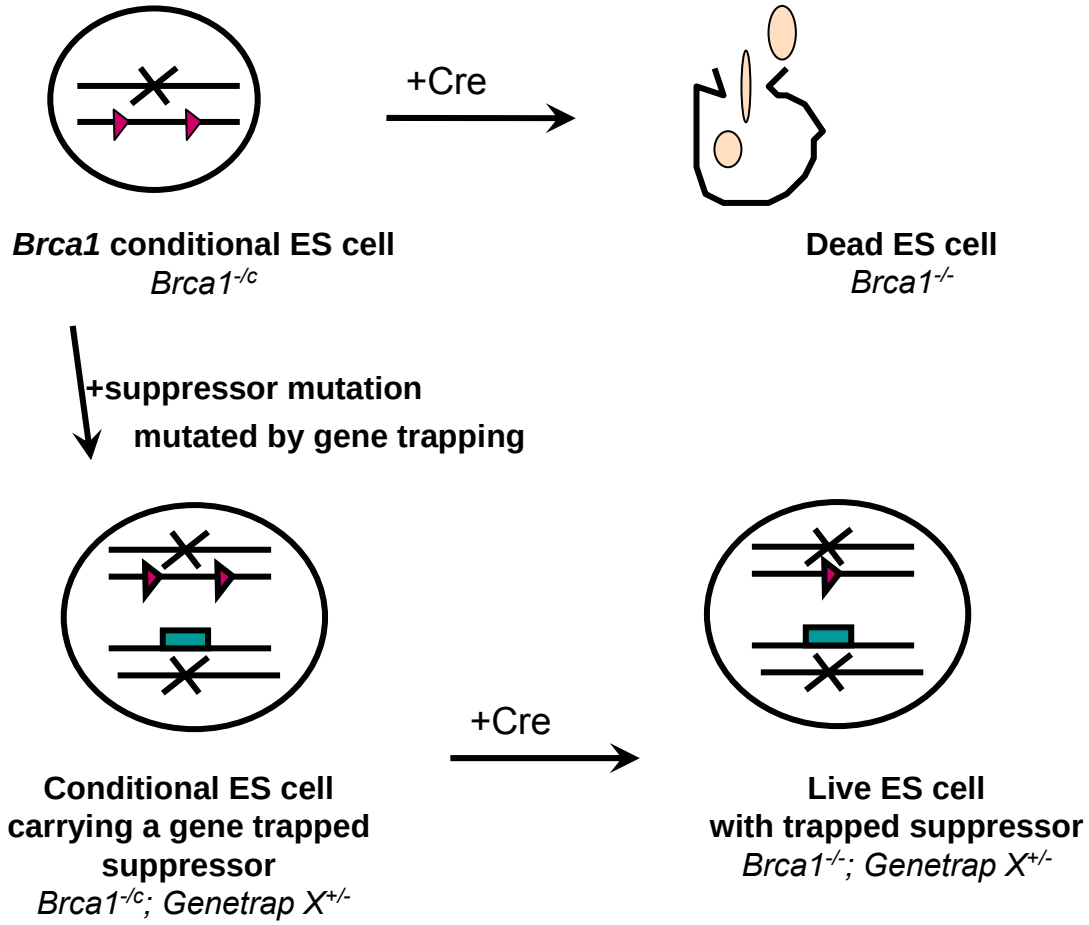


Figure 1.19: General overview of the gene trap suppressor screen in conditional *Brca1* ES cells. Complete loss of *Brca1* in ES cells is expected to be a lethal event. A conditional ES cell system, with one knockout and one *loxP*-flanked (red triangles) conditional allele, is subjected to genome-wide mutagenesis though gene trapping, then the conditional allele is excised by Cre. Only cells carrying a suppressor mutation will be viable following this loss of *Brca1*.

conjunction with Cre transgenes driven either by a breast-specific or ubiquitously-expressed promoter.

1.12.4 An overview of the chapters in this work

1.12.4.1 Generation of knockout alleles

Chapter 3 discusses the first goal of this project, which was to generate the two *Brca1* knockout alleles and target them into ES cells. Unexpectedly, the conditional allele generated in this study did not behave as a null allele following Cre-mediated deletion. This precluded the use of the conditional ES cells in a suppressor screen, but provided a new tool for studying *Brca1* function, as ES cells carrying two copies of this recombined conditional allele were viable. As this allele was predicted to give rise to a protein which lacks the N-terminal RING domain, it was named *gollum* (*gol*). Chapter 4 describes the generation of mice from ES cells carrying knockout alleles of *Brca1* and the results of the tumourigenesis studies performed using these mice.

1.12.4.2 DNA damage and the gol allele

The second aim of this work, discussed in Chapter 5, was to determine the response of *gol/gol* and *+/gol* ES cells to various forms of DNA damage. A large body of experiments provides evidence that *Brca1* is involved in DNA repair, and it was expected that having a mutant allele which specifically lacked one part of the protein would be useful in determining the role that domain had in the response to DNA damage. Immunolocalization was also performed and it was determined that the protein produced from the *gol* allele is able to localize to the nucleus and forms both S phase and DNA damage-induced nuclear foci. The localization of the mutant protein was particularly important in light of the discovery that Bard1, which interacts with Brca1 at the RING domain, is a nuclear chaperone and retention protein for Brca1, as a RING-less version of Brca1 might be not be expected to localize to the nucleus.

1.12.4.3 Molecular characteristics of the gol allele

The viability of the *gol* allele was unexpected, especially in light of the fact that the standard knockout generated in this study deletes the same exon. The third goal of this study, described in Chapter 6, was to molecularly characterize the *gol* allele, including determining the nature of the protein produced from the *gol* allele and investigating potential changes in *Brca1* RNA or protein levels in cells carrying this allele. It was also of interest to determine if Bard1 was able to bind to Brca1, in light of the findings that Brca1 localized to the nucleus in *gol/gol* cells. The findings suggest that the *gol* allele may be a useful tool for many further experiments.