MOLECULAR MECHANISM OF NUCLEOTIDE EXCISION REPAIR DEFICIENCY IN NOVEL BREAST CANCER CELL LINES

by

Jennifer M. Johnson

Bachelor of Science, Saint Joseph's University, 1999

Submitted to the Graduate Faculty of

University of Pittsburgh School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2006

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Jennifer M. Johnson

It was defended on

August 18, 2006

and approved by

Billy W. Day, Ph.D.

Stephen G. Grant, Ph.D.

W. Allen Hogge, M.D.

Richard Steinman, M.D., Ph.D.

Donna Beer Stolz, Ph.D.

Dissertation Advisor: Jean J. Latimer, Ph.D.

MOLECULAR MECHANISM OF NUCLEOTIDE EXCISION REPAIR DEFICIENCY IN NOVEL BREAST CANCER CELL LINES

Jennifer M. Johnson

University of Pittsburgh, 2006

Deficiency of DNA repair has been shown to lead to cancer in multiple human disease states. We therefore hypothesized that deficiency of Nucleotide Excision Repair is involved in the etiology of breast cancer. This work was carried out using a novel *in vitro* culture system for both normal and cancerous breast tissue developed in the Latimer laboratory. Primary cultures of nondiseased and tumor tissue were utilized in functional studies, and cell lines generated from these primary cultures were used for expression analysis. Cell lines were generated from the Latimer primary culture system without the use of exogenously transforming agents. Using the Unscheduled DNA Synthesis Assay, we have shown that stage I and stage II breast tumors are deficient in functional NER capacity as compared to nondiseased controls. Analysis of the steady-state mRNA expression levels of 20 genes in the NER pathway using an RNase Protection kit showed that 7 genes (CSB, XPA, XPB, TFIIHp52, TFIIHp44, TFIIHp34, and Cdk7) also showed a loss of expression in early stage tumor cell lines. Microarray analyses corroborated these trends and protein expression was also lost in early stage tumor cell lines for the XPA and CSB genes. In contrast to the loss of DNA repair capacity manifested in early stage tumors, stage III tumors showed an increase in repair capacity. Gain of mRNA expression and protein expression was seen for our candidate genes in late stage tumors. Because loss of expression was reversible and occurred in multiple genes and cell lines, we hypothesized that the mechanism responsible was epigenetic regulation of gene expression. Two early stage cell lines that demonstrated loss of expression of these candidate genes were treated with the

demethylating agent 5-aza-2'deoxycytidine, resulting in the re-expression of 3 genes: TFIIHp52, TFIIHp34, and Cdk7. In contrast, two of the candidate genes that were not reactivated by azacytidine treatment (CSB and XPA) showed a 7-element transcription factor framework that may begin to explain their coordinate regulation. The functional loss of NER capacity in early stage tumors is consistent with a transcriptional regulatory mechanism including both methylation and other transcription factor based functions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

So many people have helped me during the process of creating this dissertation that I barely know where to begin. I would like to thank my family, particularly my husband, Fran, and my son, Francis, for all of their love, support, computer skills, and hugs. Without them, this work would never have been completed. I would also like to thank the many people who have worked in, or rotated through, Dr. Jean Latimer's and Dr. Stephen Grant's laboratories, in particular Crystal Kelly and Amie Benson. All of these people have provided me with help during experiments, with stimulating scientific discussions, and with camaraderie and friendship. They have provided me with boundless enthusiasm when results have been positive, and words of encouragement when problems have arisen. I would like to acknowledge the MSTP program, and my career advisor Dr. Ronald Montelaro, who was always there with gentle encouragement and an open door when I needed it. I would like to thank my dissertation committee members for their scientific input, their career counseling, their enthusiasm, and their hours of time, all of which have shaped my scientific career in countless ways. Finally, I would like to thank Dr. Jean Latimer. She has done all of the above and more in my years in her laboratory. She has devoted, time, space, resources, and inexhaustible energy to both my development as a scientist and this research project. She is a mentor and an inspiration who has shaped both the research that follows and the person that I am today. A page is not enough to enumerate the ways that she has helped me. Thank you all !

1.0 INTRODUCTION

Cancer etiology in adults is a multi-step process involving the accumulation of mutations over decades. These mutations allow for the acquisition of the phenotypes associated with cancer including self-sufficient growth signals, insensitivity to anti-growth signals, evasion of apoptosis, limitless replication potential, sustained angiogenesis, and tissue invasion and metastasis (1). Three types of genes are known to be involved in the etiology of cancer: oncogenes (2), tumor suppressor genes (3), and a subset of tumor suppressor genes known as "mutator" genes, usually involved in DNA metabolism and repair (4). Mutations in DNA repair genes compromise the ability of a cell to deal with genotoxic agents, allowing the persistence of DNA damage that can then be fixed into the genome as mutations during replication.

1.1 DNA DAMAGE

1.1.1 Endogenous DNA Damage

Endogenous forms of damage include the spontaneous alterations in base chemistry. The chemical nature of DNA itself is unstable, with an estimated 18,000 purine residues lost in each cell each day through hydrolysis of the bond connecting the base to the phosphate backbone of DNA (5). This leaves abasic sites where the sugar-phosphate backbone is susceptible to further hydrolysis and subsequent chain breakage. Deamination of cytosine, adenine, guanine, and 5-methylcytosine occurs spontaneously, albeit slowly, at physiologic temperature and pH. The resulting deaminated bases provide a target for DNA repair.

Normal metabolic processes can create reactive oxygen species that can attack DNA. When attack happens at the sugar, fragmentation, base loss, and strand breaks occur. Attack at a base can generated over 80 different products (6). One example is thymine glycol. After attack at the C5-C6 double bond of thymine, a hydroxythymine radical is formed. This hydroxythymine radical can further react with oxygen to make thymine glycol (7). Not only are the strand-breaks generated by reactive oxygen species targets for DNA repair, but the aberrant bases created through attack by reactive oxygen species can be recognized by a number of repair mechanisms.

The final form of endogenous DNA damage is incorporation of incorrect bases during DNA replication, or DNA mismatch. Mismatches arise when DNA polymerase incorporates a noncomplementary base in the daughter strand. While polymerases have a proofreading ability to increase their fidelity, mistakes are made at a frequency of 1 in 10^6 nucleotides (compared to 1

in 10^3 to 1 in 10^5 without proofreading) (8). Incorporation of damaged nucleotides can also occur during synthesis. Mechanisms of DNA repair, particularly mismatch repair, address these mutations.

1.1.2 Exogenous DNA Damage

Exogenous sources of DNA damage are plentiful. They cause crosslinks, bulky adduct formation, and helix destabilization. They can be broken down in terms of chemical agents and radiation. Chemical agents vary widely and include alkylating agents, crosslinking agents, other electrophilic reactants, oxidative agents, and agents that cause strand breaks. Radiation includes both ionizing radiation and UV irradiation.

1.1.3 Chemical Agents that Damage DNA

Alkylating agents are electrophilic compounds capable of attacking the nucleophilic centers of DNA (Table 1.1). They can be either monofunctional (one reactive group) or bifunctional (two reactive groups). Sites for attack exist in all four bases at either ring nitrogens or oxygens. The most frequently found additions occur at N7 of guanine and N3 of adenine (9). Alkylating agents can both be found in the environment and are synthesized in the laboratory for use as chemotherapeutic agents due to their ability to cause DNA-damage. Environmental agents include methyl chloride, streptozocin, and *S*-adenosylmethionine. Chemotherapeutic agents that act as alkylating agents include: nitrogen mustards, melphalan, chlorambucil, cyclophosphamide, cisplatin, carboplatin, mitomycin C, decarbazine, and procarbazine.

DNA crosslinks can be formed by agents either within a strand of DNA (intrastrand) or between two strands (interstrand). Bifunctional alkylating agents (nitrogen mustards, sulfur mustards, cisplatin, mitomycin) can act as crosslinking agents. If they interact with nucleophilic centers located on opposite strands of DNA, they create an interstrand crosslink. Interstrand crosslinks are a complete block to transcription and DNA replication since they prevent strand separation. This has made them useful chemotherapeutic agents, as they stifle the growth of quickly replicating cells. UV and ionizing radiations can also form crosslinks as minor products of irradiation.

Crosslinks can also form between DNA and protein. Aldehydes are efficient mediators of such crosslinks. An example of this is the use of formaldehyde to crosslink DNA with histone proteins in the chromatin immunoprecipitation assay. Aldehydes such as acrolein are known environmental carcinogens present in automotive exhaust and tobacco smoke (10). Crosslinks result in mutations when they are repaired.

The final type of crosslinking agents are those generated by the photoactivation of planar, conjugated compounds. When compounds like psoralen and its derivatives are photoactivated after intercalating DNA, they react with pyrimidines to form monofunctional or bifunctional adducts. Bifunctional adducts can join one pyrimidine to the one below it in the opposite strand, crosslinking the two strands (11). This adduct creates a significant helical distortion, kinking, and even unwinding of the DNA.

Many chemicals can be metabolized into electrophilic reactants (5). Metabolic activation can create powerful carcinogens and mutagens. While metabolism reactions are intended to make xenobiotics more water-soluble to facilitate excretion, a subset of the compounds formed are reactive with DNA. Examples of these agents include *N*-2-acetylaminofluorine (AAF),

which becomes an electrophilic reactant when activated by cytosolic enzymes (12). Benzo[*a*]pyrene, a hydrocarbon present in cigarette smoke as well as petroleum and coal combustion forms a variety of carcinogenic and mutagenic compounds when activated by the so-called phase I metabolizing cytochrome P450 system. The most potent of these, benzo[*a*]pyrene *anti*-diolepoxide (BPDE) binds preferentially to guanine residues in DNA (13).

Aflatoxins are mycotoxins that are ingested by humans in *Aspergillus* -contaminated cereal grains (14). Like the chemicals described above, aflatoxins are also metabolized in the human body into a more active form. The cytochrome P-450 system generates aflatoxin epoxides. Subsequent electrophilic attack preferentially creates guanine adducts that ultimately result in abasic sites (14).

Even some normal hormone metabolites are capable of forming DNA adducts. Estrogens, including the natural hormones estradiol and estrone, are hydroxylated by cytochrome P-450 at several positions (15). Hydroxylation at the 4-position yields catechol estrogens. 4-hydroxylated catechols can be oxidized in the cell to semiquinones and then to DNA-reactive quinones. These quinones react preferentially with the N7 position of purine bases to form bulky adducts. Furthermore, quinines are rich sources of reactive oxygen species which also damage DNA.

Chemicals and enzymes also exist that create strand breaks. These include bleomycin and topoisomerase poisons. Bleomycin is an antitumor antibiotic used as a chemotherapeutic agent. Activated bleomycin contains a bithiazole that binds to DNA and induces hydrogen abstraction, leading to strand breaks and abasic sites (16). Topoisomerases are enzymes in the cell responsible for transiently breaking DNA and moving the ends. They do this while they are covalently linked to the DNA strands. The strands and rejoined as the topoisomerase is released.

Topoisomerase inhibitors like the anti-leukemia agent etoposide “trap” a topoisomerase when it is covalently linked to a broken DNA end (17). Thus, strand breaks with DNA-protein crosslinks are created.

Table 1.1 DNA-damaging agents used in chemotherapy.

Drug Name	**Samples of Cancers it is Used to Treat (18)**	**Method of Damage**
Cyclophosphamide	Breast, malignant lymphomas,	Bifunctional alkylating agent; nitrogen mustard; (inter- and intra-strand crosslinks (19)
Melphalan	Multiple myeloma	Bifunctional alkylating agent; nitrogen mustard (20)
Busulfan	Chronic myelogenous leukemia	Bifunctional alkylating agent; alkyl sulfonate (20)
Chlorambucil	Chronic lymphatic leukemia, malignant lymphomas	Bifunctional alkylating agent; nitrogen mustard (20)
Mitomycin	Adenocarcinoma; stomach, pancreas	Bifunctional alkylating agent (20)
Cisplatin	Ovarian, testicular	Bifunctional alkylating agent (20)
Bleomycin	Head and neck squamous cell carcinoma, lymphoma	Antitumor antibiotic; strand breaks (21)
Doxorubicin	Lymphoma, breast	Anthracycline: intercalates DNA helix to prevent unwinding; topoisomerase inhibitor (22)
5-Fluorouracil	Palliative management of colon, rectum, breast, stomach, pancreatic cancers	Antimetabolite: inhibits DNA synthesis by inhibiting thymidylate synthetase and the normal production of thymidine (23)

1.1.4 Ionizing Radiation

Ionizing radiation creates damage both directly to DNA and by generating areas with high concentrations of hydroxyl radicals (indirect damage) (5). The direct effects of ionizing radiation on DNA result from the absorption of energy by the first layer of tightly-bound water molecules surrounding DNA. This causes the ionization of both bases and sugars. The indirect effect of the generation of reactive oxygen species results in multiple attacks on the DNA bases and the sugar-phosphate backbone. The indirect effect, which constitutes the majority of the damage, results in multiple double-stranded breaks (the most lethal effects of ionizing radiation) as the radicals attack multiple sugars in an adjacent area.

1.1.5 Ultraviolet Radiation

The ultraviolet radiation spectrum is divided into three segments: UV-A (320 to 400 nm), UV-B (295-320 nm) and UV-C (100-295 nm) (5). The ozone layer absorbs and reduces the penetrance of wavelengths less than 300 nm, so solar UV radiation that reaches the earth's surface is comprised mostly of UV-A and UV-B rays. While all three segments of UV radiation can damage DNA, it is UV-C light (close to the absorption peak of DNA at 260 nm) that is the most damaging. Germicidal UV lamps emit UV light at a peak wavelength of 254 nm, and thus provide a laboratory source for investigation of UV-C's effects (24).

UV irradiation causes two main forms of damage: 6-4 photoproducts and cyclobutane pyrimidine dimers. There is an approximate ratio of 3 to 1 of cyclobutane pyrimidine dimers to 6-4 photoproducts in DNA following UV irradiation (25). 6-4 photoproducts, also called

pyrimidine-pyrimidone (6-4) adducts, form when the C6 of the 5' pyrimidine links to the C4 of the 3' pyrimidine. This most frequently occurs with TC pairs, then CC, and finally TT pairs. These products also distort the DNA helix by bringing the pyrimidine planes almost perpendicular to one another. Cyclobutane pyrimidine dimers form as covalent links are forged between adjacent pyrimidines. This linkage creates a kink in the DNA that distorts the helix by bending it approximately 30° towards the major groove and unwinding it by 9° (26). Cyclobutane pyrimidine dimers are extraordinarily stable at physiologic pH and temperature. The rate of formation depends on how frequently two pyridines are located next to one another, with thymine-thymine dimers being the most plentiful. The ratio of the different dimers in irradiated human cells is 68 TT : 13 CT : 16 TC : 3 CC (27).

1.2 DNA DAMAGE RESPONSE

The cell has multiple options from which it can choose to respond to the constant assault on DNA from endogenous and exogenous agents (28). First, it may employ a repair mechanism to remediate the damage. If a base is damaged, mismatched, or incorrectly incorporated, base excision repair, nucleotide excision repair, alternative excision repair, and mismatch repair pathways can deal with the issue. Each of these pathways is responsible for the removal of the inappropriate base and replacement with a correct one. These pathways are specialized to recognize specific types of DNA damage, with some overlap between them. If the sugar-phosphate backbone has been severed, either single-strand break repair or double-strand break repair can be used. Single strand breaks can be repaired by direct reversal by enzymes or

recombinational repair. Double-strand break repair can either be accomplished by recombinational repair (homologous recombination) or non-homologous end joining.

DNA repair may be enhanced in the cell through the use of cell cycle checkpoints and transcriptional regulation (Figure 1.1). In response to some forms of DNA damage, cells are able to arrest their cell cycle, giving the DNA repair mechanisms more time to work. DNA damage can also cause transcriptional and post-transcriptional upregulation of genes or modification of the stability of a variety of proteins that will facilitate or directly participate in DNA repair.

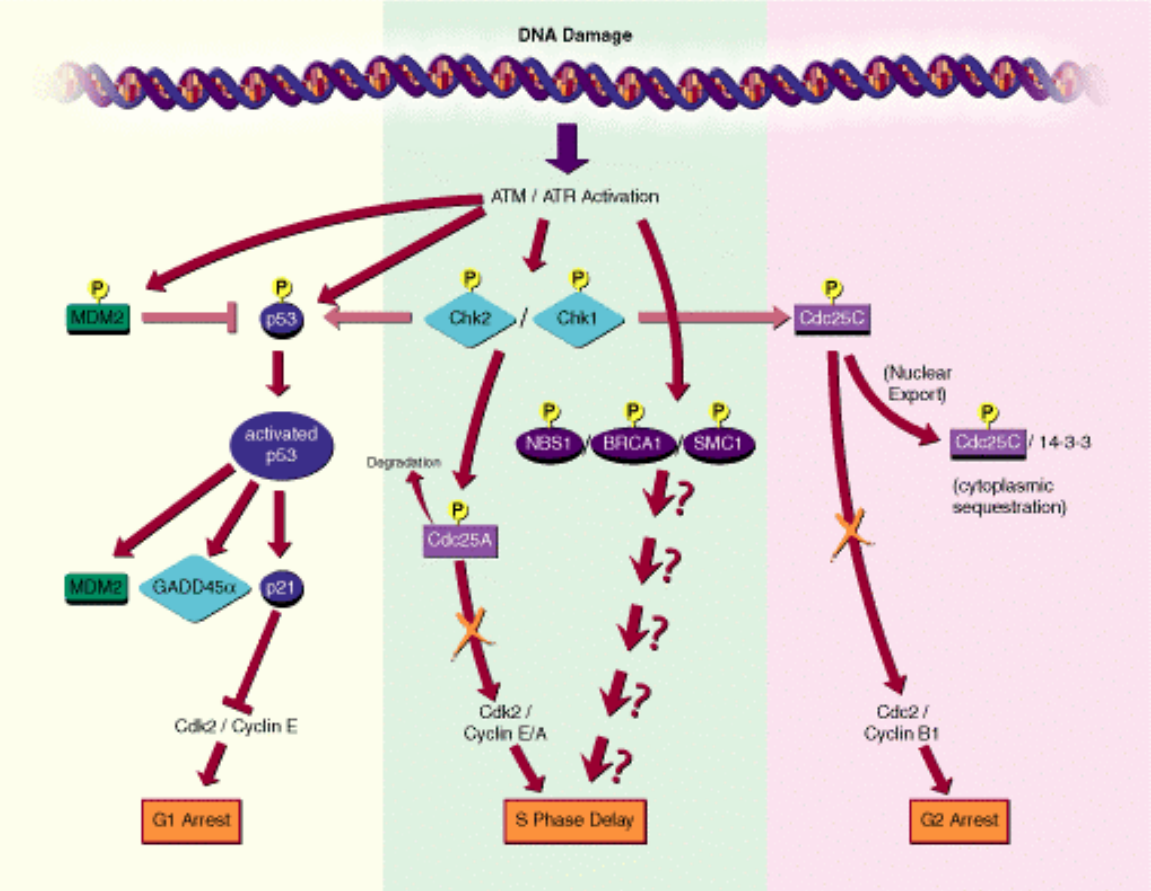


Figure 1.1 Interaction of DNA damage with cell cycle checkpoints.After DNA damage is sensed, cell cycle arrest allows DNA repair and tolerance mechanisms a chance to act. First Printed in R&D Systems 2003 Catalog; reprinted with permission.

If the cell is too extensively damaged, another option open to it is apoptosis. This mechanism of programmed cell death prevents transformation/disease in the multicellular organism in which the cell resides.

The final option left to the cell is tolerance of the damage rather than repair (29). If the damage is encountered during replication, the replication fork may fold back on itself to allow copying of information from the newly replicated template strand rather than from the damaged parental strand. Cells may also invoke recombinational repair as a tolerance mechanism. Both of these mechanisms allow for the use of an undamaged strand of DNA for synthesis. The third method by which damage can be tolerated is translesion synthesis or "error-prone" synthesis. Specialized DNA polymerases with low fidelity are used to continue DNA synthesis through one of these sites of damage, incorporating a base semi-randomly across from the site of damage. In this case, the newly synthesized strand has a higher than normal probability of carrying a mutation. This would be the most undesirable option for a multicellular organism to choose because of its high potential for mutagenesis.

1.2.1 Nucleotide Excision Repair

Nucleotide Excision Repair (NER), also known as long-patch repair, is responsible for the correction of damage that distorts the periodicity of the DNA helix . These include the damage caused by the agents discussed above that create bulky adducts (*e.g.*, alkylating agents, polycyclic carcinogens, epoxides, estrogen-derived quinones), some less bulky helix-distorting lesions (synthetic AP sites, thymine glycol), as well as UV-induced 6-4 photoproducts and cyclobutane pyrimidine dimers.

Following UV irradiation, 6-4 photoproducts are removed very efficiently from mammalian cells at approximately five to ten times the rate that of cyclobutane pyrimidine dimers (25). This is believed to be due to the fact that 6-4 PP are more helix-distorting than cyclobutane pyrimidine dimers (30). In the case of cyclobutane pyrimidine dimers, half of the dimers are resolved in 8-10 hours following a dose of less than 20 Joules/m^2. Complete removal does not take place until up to 20 days after irradiation with 10 J/m^2 (31).

In general, the more the DNA helix is distorted, the more easily it is recognized and remedied by the NER pathway (32). Lesions that cause local unwinding of the helix are the best substrates (33). This pathway can also be recruited for other types of DNA damage lesions that have not been corrected by other DNA repair mechanism, providing redundancy to ensure that DNA repair takes place in the face of overwhelming genotoxic challenge, like chemotherapy. Occasionally, nondamaged DNA can be cleaved and resynthesized, showing that is an intrinsic error rate in recognition of damage by NER (34). Rates for this error have not been published and are estimated to be extremely low.

1.2.2 Transcription Coupled Repair

Two different modes of NER can be distinguished in human cells, one targeting actively transcribed genes and the other repairing nontranscribed sequences (35) (Figure 1.2). It is known that cyclobutane pyrimidine dimers are removed more efficiently from the transcribed strand of genes (36), and that the coding strand of actively transcribed genes is preferentially repaired (37). This mode of repair, called transcription coupled repair, is thought to allow the rapid resumption

of RNA synthesis after UV irradiation. Recognition of damage and initiation of repair in transcription coupled repair takes place by a largely unknown mechanism.

1.2.3 Global Genomic Repair

The second mode of repair, global genomic repair, involves the slower, less efficient repair of bulk DNA. In global genomic repair, the lesion recognition and binding potency of proteins depends on the helix-distorting properties of the lesion (38). Damage that effectively blocks transcription can also be recognized by this pathway; thus, the global genomic repair pathway may overrule transcription coupled repair (39). Global genomic repair involves the activity of genes that recognize helical distortion, such as XPC-hHR23B, DDB1, and DDB2.

Taken together these two modes of repair equal total genomic repair. Both transcription coupled repair and global genomic repair possess their own mechanisms and genes responsible for damage recognition. They then use a common mechanism (and common genes) for unwinding of the DNA helix, incision, excision, synthesis and ligation.

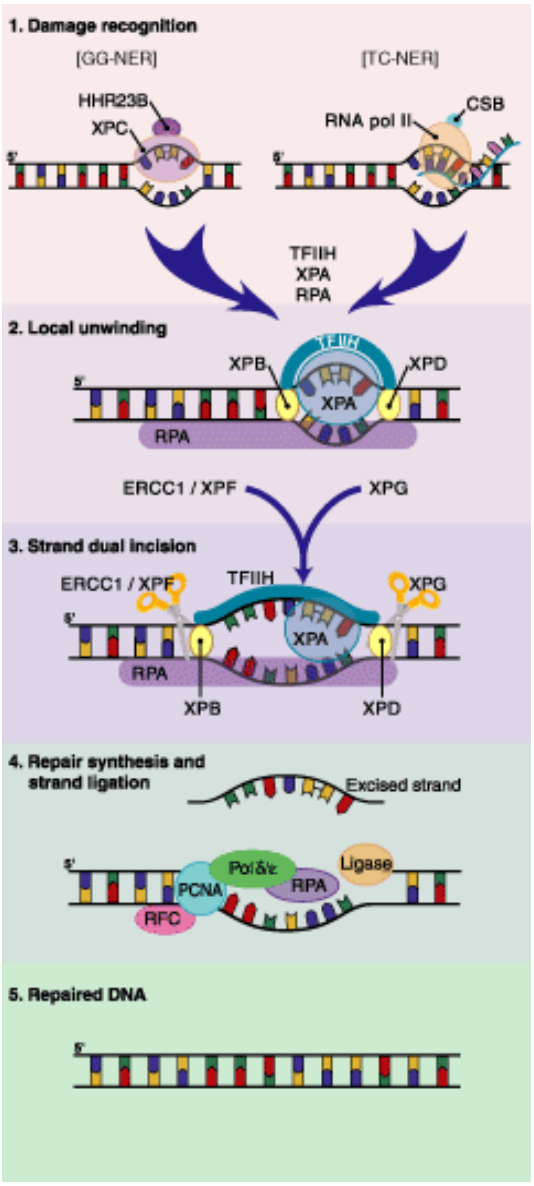


Figure 1.2 Schematic of NER showing the steps of the pathway. Damage is first recognized, usually through its effect on the periodicity of the DNA helix, the area around the damage site is then unwound, and a 27-29 base pair segment containing the damaged site is excised. The gap in DNA is then resynthesized and ligated into place. First Printed in R&D Systems 2003 Catalog; reprinted with permission.

1.2.4 Mechanism of Nucleotide Excision Repair

NER requires the activity of approximately 30 gene products (Table 1.2). These proteins participate in a coordinated fashion to first recognize a site of damage, unwind the DNA around the site of damage, incise the DNA, excise a 27-29 base pair segment of DNA including the lesion, replicate the missing segment of DNA, and ligate the newly-synthesized strand into place. Damage may be recognized by either transcription coupled repair machinery or global genomic repair machinery. Once damage is recognized by one of these two modes, however, the pathway commences through a common set of steps and gene products.

Table 1.2 Genes involved in NER and their association, in mutated form, with human diseases. Genes involved in transcription coupled repair only are shown in pink, genes involved in recognition of damage not related to a stalled RNApol II complex are shown in lavender, and genes used in a common pathway after recognition are shown in yellow. Repair is listed as measured by the unscheduled DNA synthesis assay.

Gene	Function	Associated Human Diseases	Residual Repair-GGR (%Foreskin Fibroblast)
CSA	Damage recognition in transcription coupled repair	Cockayne Syndrome	100%
CSB	Damage recognition in transcription coupled repair	Cockayne Syndrome	100%
XPC	Damage recognition in global genomic repair	Xeroderma Pigmentosum	5-25%
hHR23B	Damage recognition in global genomic repair	None known, mouse knockout is often embryonic lethal	
DDB1	Damage recognition in global genomic repair	Xeroderma Pigmentosum ?	
DDB2	Damage recognition in global genomic repair	Xeroderma Pigmentosum	40-50%
XPA	Initiation of repair and formation of an open repair complex	Xeroderma Pigmentosum	1-2%
RPAp70	Initiation of repair and formation of an open repair complex	None known. Mouse knockout is embryonic lethal	
RPAp32	Initiation of repair and formation of an open repair complex	None known, No mouse knockouts known	
RPAp14	Initiation of repair and formation of an open repair complex	None known, No mouse knockouts known	

Table 1.2 (continued)

Gene	Function	Associated Human Diseases	Residual Repair-GGR (%Foreskin Fibroblast)
TFIIHp62	Member of TFIIH with unknown function; structural ?	None known, No mouse knockouts known. Embryonic lethal in *Drosophila melanogaster*.	
TFIIHp52	Member of TFIIH with unknown function; structural ?	None known, No mouse knockouts known.	
TFIIHp44	Member of TFIIH; DNA-binding and stabilization of the complex	None known, No mouse knockouts known.	
TFIIHp34	Member of TFIIH; DNA-binding and stabilization of the complex	None known, No mouse knockouts known.	
TTD-A	Member of TFIIH; prevention of degradation and concentration *in vivo*	Trichothiodistrophy	15%
XPB	Member of TFIIH; 3' to 5' helicase	Cockayne Syndrome Trichothiodistrophy Xeroderma Pigmentosum	3-7%
XPD	Member of TFIIH; 5' to 3' helicase	Cockayne Syndrome Trichothiodistrophy Xeroderma Pigmentosum	25-50%
CDK7	CTD kinase	None known, No mouse knockouts known. Embryonic lethal in *Drosophila melanogaster*	
CycH	Regulation of CDK7	None known, No mouse knockouts known.	
MNAT1	Regulation of CDK7 and DNA binding	None known, Peri-implantation lethal in mouse knockout.	

Table 1.2 (continued)

Gene	Function	Associated Human Diseases	Residual Repair-GGR (%Foreskin Fibroblast)
XPG	3' endonuclease	Cockayne Syndrome Xeroderma Pigmentosum	5-25%
ERCC1	5' endonuclease with XPF	None known, Mouse knockout is often embryonic lethal	
XPF	5' endonuclease with ERCC1	Xeroderma Pigmentosum	15-30%
PCNA	Forms a ring around DNA where DNA polymerase can dock	None known, No known mouse knockouts	
RFC1	Facilitates binding of PCNA to 3' end of the DNA gap	None known; Mouse RFC is also heteropentameric, but mouse subunits do not directly correlate with human subunits; no known knockouts of mouse subunits	
RFC2	Facilitates binding of PCNA to 3' end of the DNA gap	None known, see RFC1	
RFC3	Facilitates binding of PCNA to 3' end of the DNA gap	None known, see RFC1	
RFC4	Facilitates binding of PCNA to 3' end of the DNA gap	None known, see RFC1	
RFC5	Facilitates binding of PCNA to 3' end of the DNA gap	None known, see RFC1	
PolD1 (δ)	DNA polymerase responsible for gap-filling synthesis	None known, No known mouse knockouts	
PolD2 (δ)	DNA polymerase responsible for gap-filling synthesis	None known, No known mouse knockouts	
LIG1	Ligation of newly synthesized strand	1 human case, 46BR	unknown

1.2.5 Recognition of DNA Damage in Transcription Coupled Repair

General elongation factors can aid in the translocation of a stalled RNApol II complex when it appears at an undamaged DNA site (such as a stall due to a secondary structure), but these elongation factors are ineffective at a site of DNA damage when there is a DNA-distorting adduct. It is at this point that the transcription coupled repair machinery functions to recognize the adduct and repair it. Transcription coupled repair is facilitated by the presence of the gene products CSA (also known as ERCC8) and CSB (also known as ERCC6), whose functions are largely unknown. The current theory for recognition and initiation of repair is that when transcription of genes by RNA polymerase II becomes stalled due to the presence of a lesion in the DNA during the production of mRNA, the gene product CSB is recruited to the lesion. CSB then helps to relieve the stall by one of three proposed mechanisms: (1) remodeling the RNApol II-DNA interface to make the lesion accessible, (2) releasing RNAPII from the damaged site, or (3) promoting damage bypass.

Evidence supporting the idea that CSB remodels the RNApol II-DNA interface comes from their sequence homology. CSB was identified as a member of the SWI2/SNF2 family of ATP-dependent chromatin-remodeling factors. Further *in vitro* experiments have revealed that CSB is a 162 kDa protein capable of remodeling nucleosomes and altering the DNA double helix in an ATP-dependent reaction (40). This ability to move proteins on DNA makes it a candidate for being able to also alter the RNApol II-DNA interface. In theory, CSB could cause retrograde translocation of the RNApol II, allowing access to the site of damage.

Removal of the RNApol II or bypass of the lesion may be accomplished in a similar way. Since CSB is theoretically capable of translocating the RNApol II on DNA, it may force the

RNApol II past the site of damage at the expense of misincorporating or skipping a base. This may allow RNApol II to continue past the damage, or it may cause RNApol II to fall off of the DNA completely. CSB has can interact directly with RNApol II to cause transcriptional elongation, supporting this theory. However, other experiments have now shown that CSB alone is not capable of causing RNApol II dissociation (41).

Another function in which CSB participates is ribosomal RNA synthesis (42). Studies have shown that CSB exists in the nucleolus in a complex with RNApol I, and other genes involved in NER, specifically members of the TFIIH complex and XPG. rRNA synthesis can be restored in CSB-deficient cells in a dose-dependent manner with the addition of CSB. The mechanism by which this occurs is also unknown, but may be similar to that used during stall of an RNApol II complex.

With respect to initiation of repair after recognition of the stalled complex, CSB is able to recruit the TFIIH complex to the lesion. The TFIIH complex is responsible for accomplishing precision unwinding of the DNA helix, as discussed below.

Even less is known about the other gene involved in DNA damage recognition in transcription coupled repair, CSA. CSA interacts with CSB and the p44 subunit of TFIIH. It does not bind to RNApol II nor does it release the stalled RNApol II. After DNA damage, CSA is rapidly translocated to the nucleus in a CSB-dependent manner (43). Sequence homology to the yeast protein Rad28 has identified CSA as a member of the cullin-containing ubiquitin E3 ligases. Cullins are a family of proteins that confer substrate specificity to ubiquitin E3 ligases. However, the phenotype of loss of CSA in mammalian cells differs from that of the loss of Rad28 in *Saccharomyces cerevisiae*. Recovery of RNA synthesis after irradiation does not

depend on Rad28 in yeast as CSA is required for recovery of RNA synthesis in mammalian cells. The exact role of CSA in transcription coupled repair is still unknown.

Should transcription-initiated DNA damage fail, there are two possible fates for a stalled RNApol II complex. If the stalled complex is still bound to the DNA, the large catalytic subunits are targeted for degradation by the proteosome via ubiquitinylation. CSB may play a role in this degradation. Some studies report decreased proteosomal degradation of RNApol II in $CSB^{-/-}$ cells, while others report that the level is equal to that of unaffected cells. The other fate of the stalled complex is recognition and repair through the second mode of NER: global genomic repair.

1.2.6 Recognition of DNA Damage in Global Genomic Repair

The primary mechanism of DNA damage recognition in global genomic repair involves the activity of two gene products: XPC and hHR23B (RAD23B). XPC and hHR23B exist as a protein complex in the mammalian cell. This complex preferentially binds to damaged DNA. This complex binds early in the steps of NER, shown experimentally by an increased in NER efficiency appearing when XPC was allowed time to bind to damaged DNA before other components of the pathway were added. It also binds tightly to DNA. Atomic force spectroscopy has shown that XPC-hHR23B bends when it binds to DNA and binds to the DNA enough to protect it from cleavage by DNAse I. Once XPC-hHR23B binds to the site of damage, there is further alteration of the genome immediately in the area of damage. The binding of XPC-hHR23B and the subsequent bending of the DNA that this causes helps to create enough space for the entry of XPA, RPA, and the TFIIH proteins.

A model has been proposed whereby XPC, hHR23B, replication protein A (RPA), and XPA exist in a preformed complex in the cell. Called the co-operative recognition model, it suggests that recognition of DNA damage occurs through the concerted efforts of all these proteins, each of which contains some ability to select for DNA damage (44). The theory has originated from *in vitro* binding affinity assays conducted with purified human proteins. The footprint generated by XPC binding to DNA is enhanced in the presence of both XPA and RPA proteins, and cooperatively with both. However, XPA-RPA-XPC-DNA complexes have not been found. XPC has also been shown in *in vivo* studies to be present at the site of damage before XPA.

Two other gene products, DDB1 And DDB2 (XPE) may also be involved in damage recognition, but where they fit in to the pathway when compared to XPC-hHR23B is unknown. DDB1 and DDB2 are able to recognize helical distortions such as 6-4 photoproducts and cyclobutane pyrimidine dimers as a heterodimer. The role that these proteins play, however, is unclear. While some *in vitro* experiments with purified components of NER have shown that the addition of DDB proteins does not affect the repair of cyclobutane pyrimidine dimers, others have reported that their addition results in a 10-fold increase in repair (45). How these proteins interact with XPC during recognition is under investigation.

1.2.7 Recognition and Initiation of Repair

Once damage has been recognized, XPA and the RPA complex are recruited to the site in a process Hanawalt and colleagues called “lesion verification” (32). They theorize that XPA and RPA are there to insure that the helical distortion recognized by other factors is not a normal

variation of DNA. XPA is able to bind DNA and may function structurally to link other NER proteins to one other and to DNA. XPA can form complexes with RPA, ERCC1, and TFIIH. The ability of XPA to bind to DNA is enhanced by RPA. Correct location of the members of the NER machinery is required for excision of the correct piece of DNA including the lesion. Therefore, XPA is a required and important component of NER.

RPA was originally identified as a protein essential for DNA replication of simian virus 40 in mammalian cells, as well as for chromosomal replication (46). It is now recognized as a core component of NER. RPA is a complex of three proteins: RPAp70 (RPA1), RPAp32 (RPA2), and RPAp14 (RPA3). RPA is a single-stranded DNA binding protein with DNA binding domains on all three subunits. RPA has been implicated in multiple steps of NER, facilitating damage recognition in conjunction with XPC and XPA, aiding in the stabilization of an open complex with TFIIH, orienting the endonucleases, and stimulating gap-filling synthesis. In NER, it binds to the undamaged strand and facilitates stabilization of the open complex required for repair. It optimally binds to a 30 nucleotide segment of DNA, the size of an open repair bubble. RPA binds to DNA with a defined polarity, with the 5' oriented site binding first (47). This may help to orient the endonucleases utilized in the incision step, with XPF-ERCC1 being associated on the 3'-facing side of the protein, and XPG on the 5' side. Finally, RPA has been implicated in the synthesis step of repair. RPA is capable of stimulating both pol δ and pol ε. This heterotrimer may remain in place after excision to facilitate gap-filling synthesis.

1.2.8 Unwinding

Unwinding of the DNA around the site of damage is required for the excision of the damaged DNA. The basal transcription factor TFIIH performs this function. TFIIH, also known as general transcription factor 2H (GTF2H), was originally identified because of its activity in transcription of class II, or protein coding, genes. It is the only transcription factor known to have multiple enzymatic abilities. With respect to basal transcription, TFIIH phosphorylates RNApol II, stimulating the ability of RNApol II to translocate along the DNA template beyond the promoter region of the gene, a process called promoter clearance. It also may be involved in the pre-initiation step, opening the DNA to allow RNApol II to read the template strand. With respect to DNA repair, TFIIH aids in unwinding the DNA around the site of damage, forming an "open" complex where other repair proteins can act in subsequent steps.

TFIIH is composed of at least 10 proteins: TFIIHp62, p52, p44, p34, TTD-A (p8), XPB (ERCC3, p89), XPD (ERCC2, p80), CDK7, CycH, and MNAT1 (Table 1.3). These can be broken down into three groups: structural proteins responsible for DNA binding and coordination of the complex, helicases responsible for unwinding the DNA, and the CAK subcomplex responsible for phosphorylation of RNApol II and interaction with the cell cycle.

Table 1.3 Members of the TFIIH complex. The helicases are seen in purple, structural proteins are shown in green, and the CAK subcomplex is shown in orange.

Gene	Function
XPB	3' to 5' helicase
XPD	5' to 3' helicase
TFIIHp62	Core structure
TFIIHp52	Core structure
TFIIHp44	Core structure; DNA binding
TFIIHp34	Core structure; DNA binding
TTD-A	Stability and concentration
Cyclin H	CAK subcomplex- regulator of CDK7 function
CDK7	CAK subcomplex- kinase activity
MNAT1	CAK subcomplex- regulator of CDK7 function and DNA binding

Structural members of the TFIIH complex include core proteins TFIIHp62, p52, p44, p34, and TTD-A. These structural members have no known enzymatic functions, but rather bind either to DNA or to one another to position the enzymatic members of the complex for action. p44 and p34 each contain zing-finger motifs. These proteins are believed to function together to position the helicases on the DNA and stabilize the complex in this position. p52 has been shown to directly interact with the XPB helicase, an interaction which is require *in vitro* for function of the helicase (48). This interaction is proposed to facilitate interaction of the XPB helicase with the rest of the TFIIH complex (49).p62 is a structural component of TFIIH that contains a pleckstrin homology domain in its N-terminus. The purpose of this domain is unknown, but it is required for the interaction of the TFIIH complex with the endonuclease XPG during repair (50). TTD-A is responsible for overall TFIIH stability and thereby concentration *in vivo* by protecting the complex from degradation.

The two helicase members of the TFIIH complex are XPB and XPD. They display an opposite polarity, each responsible for unwinding the DNA helix in one direction from the site of damage. XPB is the 3' to 5' helicase and XPD is the 5' to 3' helicase. Each possesses an ATP-ase activity required for unwinding. These helicases thread the DNA through the newly formed NER complex, and their procession arrests when they locate a site of DNA damage. While activity of only the helicase activity of XPB is required for adequate transcription, DNA unwinding by both XPB and XPD is necessary for efficient DNA repair. XPD plays other roles in transcription through its DNA-binding and protein-protein interactions.

The CDK-activating kinase (CAK) subcomplex has a kinase activity used in basal transcription functions, but is also necessary for efficient repair. The subcomplex is comprised of three members: cyclin H, cyclin-dependent kinase 7 (CDK7), and menage a trois 1 (MNAT1). CDK7 provides the kinase activity and Cyclin H provides regulation for its activity. MNAT1 contains a zinc-finger domain that may add to the ability of TFIIH to bind to DNA. CAK can be found in the cell as a preformed complex or in association with TFIIH through XPD.

While CAK is a required for repair, its exact function(s) in NER is unknown. One possibility is that CAK recruits TFIIH away from its basal transcription activities to participate in repair. How this might be accomplished, however, is unknown. Another possibility is that it functions as a link between NER and the cell cycle.

1.2.9 The Link Between Nucleotide Excision Repair and Transcription

Both NER and transcription use a common set of proteins including the TFIIH complex and CSB. Possible recruitment of factors away from one process to participate in the other is a potential site for interplay between the two processes that is under current investigation. TFIIH is involved in the process of transcription initiation and the transition from initiation to elongation. CSB is involved in resolving a RNApol II complex that has become stalled once elongation has begun.

Upon initiation of transcription, the TATA-box binding protein (TBP), transcription factor IIB (TFIIB), transcription factor IIE (TFIIE), transcription factor IIF (TFIIF), and RNApol II all bind together at the promoter site to create the "pre-initiation complex" (Figure 1.3). This complex is held in place until TFIIH also binds. When TFIIH joins the members of the preinitiation complex, it forms the "transcription initiation complex." The CAK subcomplex of TFIIH then phosphorylates multiple members of the transcription initiation complex including the TBP, TFIIE, and TFIIF. Most importantly, TFIIH participates in the phosphorylation of the largest subunit of RNApol II. Once these phosphorylations have been carried out, the transcription factors are released, and RNApol II is able to proceed processively down the DNA. The phosphorylation of the RNApol II subunit is the greatest contributor to this release, but the phosphorylation of the large subunits of TFIIE and TFIIF as well as TBP are also believed to aid in release from initiation by providing a high concentration of phosphorylated proteins in the initiation complex.

Progression does not always proceed smoothly for RNApol II. Obstacles to transcription include DNA-binding proteins, unusual DNA structures, nucleosomes, and DNA damage

lesions. When an obstacle is present, the majority of polymerases will fall off the DNA template and terminate transcription, while others arrest. Slow transcription through an unusual DNA structure is also possible, and is accomplished with the help of the transcription factor TFIIS (51). Even without these obstacles, RNApol II is prone to prolonged pausing and total arrest (52,53). The mechanism by which cells resolve stalled RNA polymerases, whether they are due to DNA damage lesions or to one of the other obstacles listed, is currently thought to involve CSB. There are three mechanisms by which CSB is proposed to resolve a stalled RNApol II complex (54).It is unknown which of these mechanisms is correct at this time. They include: remodeling of the RNApol II-DNA interface to make the lesion more accessible, release of RNApol II from the site of the stalled complex, and the promotion of bypass by RNApol II.

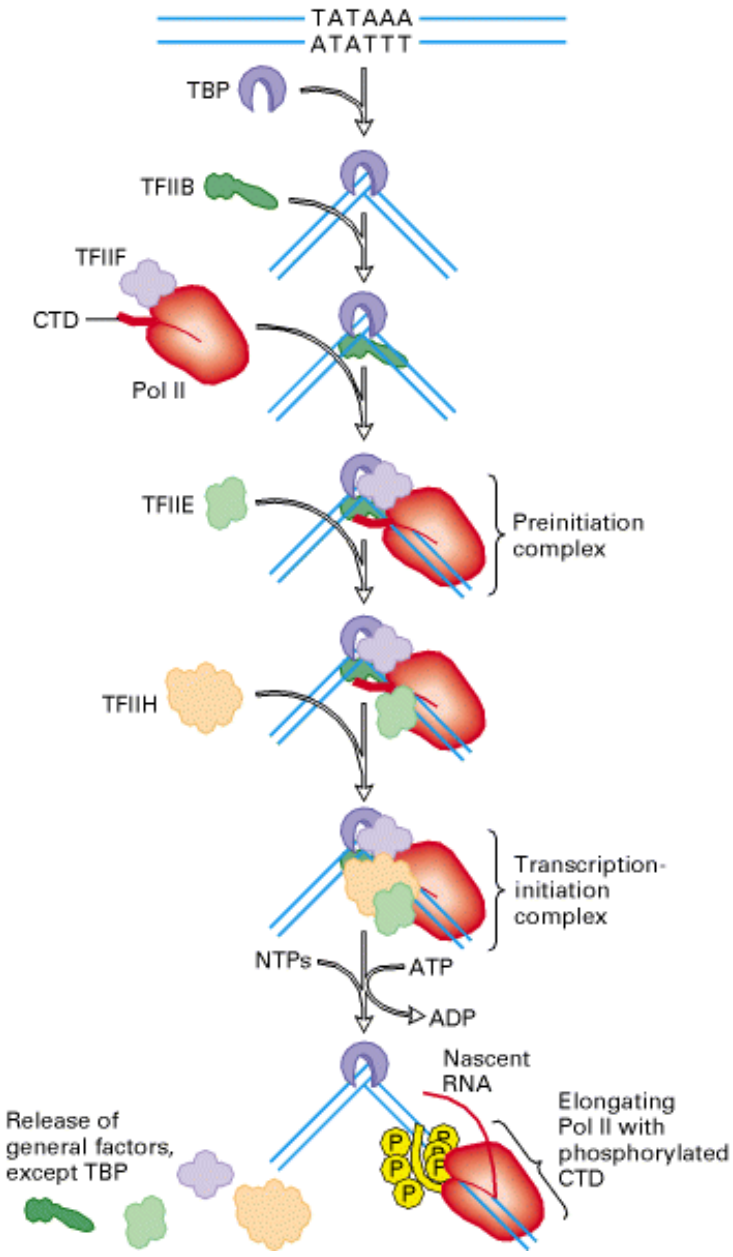


Figure 1.3 TFIIH is recruited to the pre-initiation complex of protein-coding gene transcription. TFIIH is involved in phosphorylation of the carboxyl terminal domain of RNApol II, facilitating its escape from the promoter. It is then released as elongation proceeds. Originally published in (55); reprinted with permission.

1.2.10 The Link Between Nucleotide Excision Repair and the Cell Cycle

TFIIH may play an important role in linking NER and the cell cycle through the CAK subcomplex. Specifically, the CDK7 kinase is proposed to act as a checkpoint between transcription blocked by a DNA damage site and cell cycle control by beginning a signaling cascade to promote cell cycle arrest. CAK has been shown *in vivo* to be capable of phosphorylating CDC2 and CDK2, two cyclin dependent kinases of the cell cycle. Though there is no direct evidence that these two cyclin dependent kinases are the specific mediators of an effect from CDK7 *in vivo*, this experimental observation provides a basis for further investigation.

Another possible mechanism of cell cycle interaction through TFIIH is through TP53. As discussed previously, CSA and CSB are recruited to the site of stalled RNApol II complexes at sites of damage in actively transcribed genes. CSA has been shown to directly physically interact with TFIIHp44 and XPB. Thus, TFIIH is brought to the site of DNA damage. The TFIIH components XPB, XPD, and p62 directly bind to TP53 (56). It has also been shown that TP53 can be phosphorylated *in vitro* by CDK7 (57). Once phosphorylated, TP53 has a greater affinity for p53-reponsive elements, *i.e.,* it has become activated. In this active state, TP53 binds to the TP53-response elements and stimulates the transcription of genes such as p21, which cause G1/S cell cycle arrest (Figure 1.4). This theory is highly speculative and remains to be demonstrated *in vitro* or *in vivo*, although it does fit with the observation that the activity of TP53 as a transcriptional activator requires the accumulation of damage in actively transcribed genes.

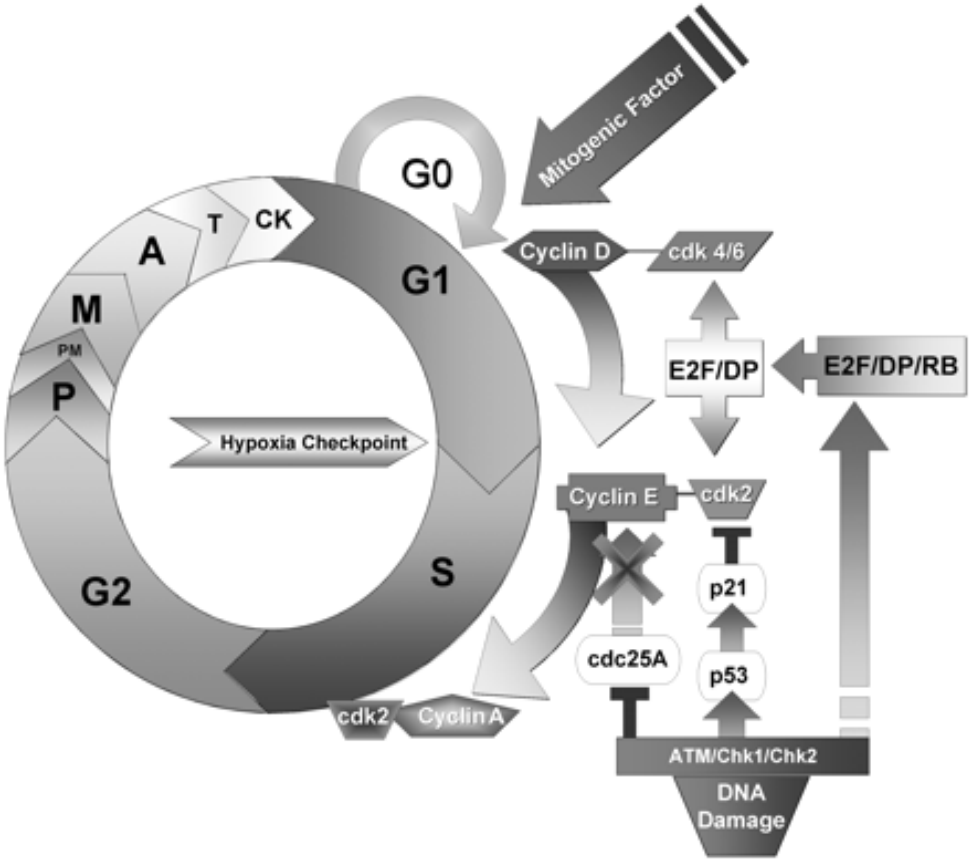


Figure 1.4 Diagram of the interaction of TP53 with the G1/S cell cycle checkpoint (58). Reprinted with permission.

1.2.11 Incision and Excision

Endonucleases are required in NER to cut at sites both 5' and 3' to the damage lesion. The endonuclease responsible for the 5' cut is the structure-specific endonuclease XPG. XPG cuts DNA substrates such as bubbles, splayed arms, and stem-loops. It requires at least one single-stranded arm protruding in either the 3' or 5' direction, and if a bubble substrate is being used, at least 5 nucleotides must be unwound. Consistent with this, XPG cuts at the 3' end of the repair bubble. XPG may also play a structural role in NER. Its physical presence helps to "hold open" the unwound DNA, allowing access to the damage lesion and the surrounding nucleotides.

After the 3' incision has been made by XPG, XPF together with ERCC1 make the 5' incision. ERCC1 and XPF create a stable complex *in vivo*. Their stability is dependent on one another. Together, ERCC1-XPF is another structure-specific endonuclease recognizing stem-loops, bubbles, splay-arms, and flaps (Figure 1.5). The minimum bubble size for ERCC1-XPF detection of a substrate is between 4 and 8 nucleotides. Incisions are always made at the 5' side of a single-stranded arm, consistent with its action as a 5' endonuclease in NER. Affinity cleavage assays have identified XPF as containing the active cleavage site for this holoenzyme, with sequence homology making ERCC1's role consistent with DNA binding (58).

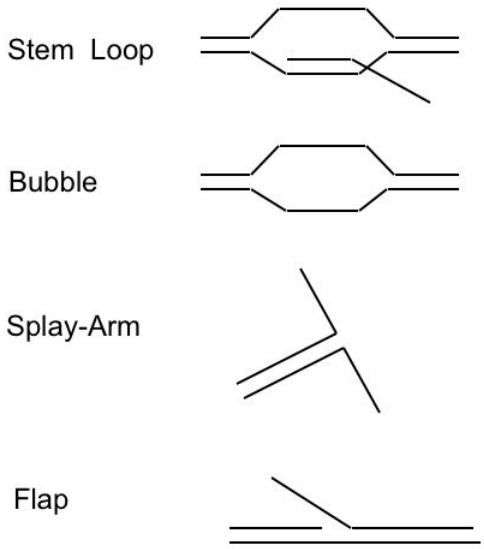


Figure 1.5 Example structures of the branched DNA substrates recognized by the XPG and ERCC1-XPF endonucleases. Recognition occurs where a double-stranded DNA segment is connected to a single-stranded tail.

1.2.12 Synthesis and Ligation

Gap-filling synthesis uses either the Pol δ or Pol ε enzymes. Synthesis also requires the activity of proliferating cell nuclear antigen (PCNA) and replication factor C (RF-C). The five subunit complex of RF-C binds to the 3' terminus of the gap and facilitates the loading of PCNA. PCNA can then form a homotrimeric ring. This ring is a docking position for Pol δ or Pol ε. After synthesis, ligase 1 (LIG1) seals the 5' end of the newly synthesized strand into place.

1.3 MEASUREMENT OF NUCLEOTIDE EXCISION REPAIR

1.3.1 Unscheduled DNA Synthesis Assay

Unscheduled DNA synthesis (UDS) is a laboratory technique used to determine the functional level of total genomic repair within a cell. In this assay, cells are first exposed to genotoxic challenge, usually UV light but also any damage form discussed above that could be recognized by the NER pathway, and then allowed to uptake a radiolabelled nucleotide, ^{3}H thymidine. The cells incorporate the ^{3}H-thymidine wherever they are synthesizing DNA. In unexposed control cells, radiolabel incorporation primarily takes place in cells undergoing semiconservative DNA synthesis in S-phase. Background endogenous DNA damage is also represented in these cells. In exposed cells, incorporation occurs in both in S-phase cells and in all damaged cells, leading to the term "unscheduled" DNA synthesis. The sites of incorporation in damaged, non-S-phase cells are predominantly the sites of gap-filling synthesis undertaken as the last step of NER. Thus, the UDS assay actually measures repair synthesis.

Autoradiography is then used to visualize the locations label incorporation. Cells are fixed and dipped in a photographic emulsion. They are then given time for the ^{3}H-thymidine they have incorporated to expose the photographic emulsion that overlays them, causing the formation of quantifiable silver grains. The emulsion is then developed and the cells stained to allow visualization of the nuclei. This technique is particularly useful because it allows for the determination of the S-phase index of the cell population and discrimination between different types of cells in mixed populations while their DNA repair capacities are quantified.

Data is collected in the UDS assay by counting exposed silver grains located over the nuclei of damaged and control cells after subtracting the local background of grains from each. Cells deficient in NER will have fewer silver grains located over their nuclei as there is less repair synthesis occurring. Control unexposed cells account for endogenous DNA damage, and background counts control for environmental causes of silver grain exposure. Great care must also be taken to include appropriate controls in these experiments to account for interexperimental variation. Traditionally, foreskin fibroblast (FF) cell explants have been utilized as a control (59,60). These cells have a high level of NER, are readily available and culturable as a discarded tissue, and exhibit reproducible levels of repair up to passage 13 (61). Published repair capacities are often expressed as % of the FF control sample. As will be discussed below, the technique of UDS was employed to establish that NER deficiency was the basis of the DNA repair deficiency disorder xeroderma pigmentosum (XP) by Cleaver in 1968 (62).

A new emerging technique for measuring total genomic repair is an antibody-based assay that tracks the presence of damaged DNA sites (63). Antibodies generated to either 6-4 photoproducts, cyclobutane pyrimidine dimers, or benzo[a]pyrene diolepoxide DNA adducts are allowed to bind to their epitopes after cells have been exposed to the appropriate type of damage (64-66). The cells are then monitored over time to watch for the "disappearance" of these damaged sites. The decrease in number of damaged sites over time is thought to be due to repair.

1.3.2 Host Cell Reactivation

The host cell reactivation (HCR) assay specifically provides researchers with a method of investigating the ability of a cell to perform transcription coupled repair (67). This assay indirectly monitors cellular repair of transcriptionally active genes by looking for "reactivation" of a damaged enzymatic marker gene (68). First, a plasmid containing a marker gene such as luciferase is damaged outside of the cell. Then, cells are transfected with this damaged pasmid. It is presumed that some of the damaged reporter plasmid makes its way into the nucleus, where repair occurs, and then gene transcription and translation occur via normal cellular trafficking. Cells are then allowed time to express the reporter enzyme, harvested for protein, and then assayed for the enzymatic activity of the reporter. Two levels of controls are utilized for the assay. The first involves the use of damaged and undamaged versions of the same plasmid to determine the ratio of expression of the damaged (and repaired) plasmid divided by the expression of the undamaged plasmid. In addition, a plasmid distinguishable from the experimental plasmids is also necessary to control for transfection efficiency. In order to make the results of individual experiments comparable to each other, it is helpful if the absolute numbers expressed by the ratio of damaged (and repaired) over undamaged plasmid be divided by similar results derived from a standard cell line run in every experiment, although published reports have not traditionally included this type of control to date. Transient expression plasmid DNA vectors were used by Protic-Sabljic and Kraemer on SV40 transformed human fibroblasts in 1985 (69) and by Athas *et al.* on human lymphocytes in 1991 to look for inherent differences in repair capacity between cancer patients and control nondiseased patients (70).

An advantage of this technique is that it minimizes the cytotoxic effects of damaging agents that might indirectly compromise the repair mechanisms of the cell (71). Damage takes place *in vitro* and can be adapted to investigate specific damaging agents. However, it must be realized that this is an artificial situation, and DNA damage in the human body does not occur extracellularly. Concerns arise from the fact that non-mammalian reporter gene is being expressed in a mammalian cell, although most transfection-based assays utilize non-mammalian genes to minimize backgrounds. Also, if the plasmid that is damaged is capable of replication, translesion synthesis may also be used to deal with the damaged sites.

1.3.3 Measuring Total Genomic Repair and Its Components

Parsing out the relative contributions of transcription coupled repair and global genomic repair in a cell is not currently experimentally feasible. The UDS assay measures total genomic repair. It is capable of detecting all locations of repair synthesis. Since both transcription coupled repair and global genomic repair mechanisms include a step where repair synthesis takes place, this technique does not differentiate between the two. While it is tempting to suggest that the total genomic repair measured by the UDS assay minus the transcription coupled repair capacity measured by HCR would describe the global genomic repair, this fails to account for repair kinetics. Since transcription coupled repair occurs preferentially, it is over-represented in the time frame used for the UDS assay. Likewise, the potential contribution of translesion synthesis to HCR complicates its use as a measure of the transcription coupled repair component.

Another possible suggestion would be the analysis of repair of individual genes, one actively transcribed and one not. This is the technique used by Nouspikel and Hanawalt in their

investigation of "differentiation-associated repair" (72). Based on their observations of specific genes, they have demonstrated that removal of cyclobutane pyrimidine dimers is attenuated in a silenced (nontranscribed) gene but efficient in four actively transcribed genes in differentiated neurons. Before differentiation, these neurons showed no deficits in repair. This phenomenon was therefore named "differentiation-associated repair." While this technique of selecting specific genes to examine gives some indication of relative deficiencies in either mode of repair, it does not reflect repair of the genome as a whole since sequence affects repair efficiency.

1.4 NER DEFICIENCY AND HUMAN DISEASE STATES

Multiple human diseases have been found that result from the mutation of a gene involved in the NER pathway including xeroderma pigmentosum (XP), Cockayne syndrome (CS), and trichothiodistrophy (TTD). Phenotypes arise from the activities of these genes in NER as well as basal transcription.

The elucidation of genes in the NER pathway was done in large part by genetic complementation analyses of these disorders. Cells from an affected patient would be fused with cells from another patient to create a heterodikaryon. The heterodikaryon was then compared to each monokaryon by the UDS assay to see if (1) the defect would be resolved, indicating that a different mutant gene was involved in the pathogenesis of each patient of (2) the defect was still present, indicating that the same gene was mutated in each. This technique identified the 8 complementation groups of XP, the 2 groups of CS, and the discrimination of TTD and the 46BR case (a DNA ligase I mutation) as separate disorders (Table 1.4).

Table 1.4 Complementation groups of the repair deficient disorders xeroderma pigmentosum and Cockayne syndrome.

Complementation Group	**Chromosomal Location of the Gene**	**Repair Capacity of Patients as measured by UDS, relative to Foreskin Fibroblast Control**
XP-A	9q22	1-2%
XP-B	2q14	3-7%
XP-C	3p25	5-25%
XP-D	19q13	25-50%
XP-E	11p11	40-50%
XP-F	16p13	13-50%
XP-G	13q33	5-25%
XP-V	6p12	100%
CS-A	5q12.1	100%
CS-B	10q11.23	100%

1.4.1 Xeroderma Pigmentosum

The classical DNA repair deficiency syndrome XP is perhaps the most well-known disease resulting from the mutation of an NER gene. XP affects 1 in 250,000 in the US with higher frequency in Japan and the Mediterranean areas (up to 1 in 40,000) (73,74). XP patients are primarily characterized by severe UV-sensitivity of the skin, pigmentation abnormalities, and a highly elevated risk of UV-induced skin tumors (75). Often, neurologic degeneration is also a component of their symptoms. XP skin is normal at birth, but then progressive irregular pigmentation occurs with formation of keratoses and basal and squamous cell carcinomas by the third and fourth year of life. Most patients die in the third and fourth decade of life. This disease illustrates one of the most direct links known between the lack of DNA repair and cancer.

Eight different complementation groups of XP have been identified through cell fusion experiments, as described above. The disease is inherited in an autosomal recessive pattern. Severity of disease, in terms of both neurologic degeneration and affect on NER capacity, is determined by the gene that is affected and the location of the mutation within that gene.

1.4.2 Xeroderma Pigmentosum Heterozygotes

There is currently no consensus on whether or not XP heterozygotes have a disease (syndrome) phenotype related to haploinsufficiency of their mutated gene, although the hypoxanthine-guanine phosphoribosyl transferase (HPRT) somatic mutation assay demonstrates a trend towards increased mutation frequency (76-78). Multiple techniques have been used to assess possible repair phenotypes including survival after UV irradiation and NER by both UDS and

HCR. The effect of haploinsufficiency does not seem to be related to complementation group or the technique used to assay it (Table 1.5).

With respect to clinical phenotypes associated with the disease, Swift and Chase performed a study of family members of XP patients versus spousal controls (79). They surveyed 31 families in North American and found an increased rate of lung, stomach, and prostate cancer in blood relatives of XP patients. Four cases of microcephaly and 11 cases of mental retardation were also seen. The complementation group(s) of these patients were unknown. Another study conducted in Britain on an unspecified number of families found 2 cases of skin cancer, 3 stomach cancers, 2 lung cancers, and 1 prostate cancer in parents or grandparents of probands (80).

One would expect, because XP is a classical recessively inherited disorder, that heterozygotes would have no phenotype. The presence of both varying severity of disease phenotypes (increased cancer incidence, microcephaly, mental retardation) and inconsistent reports of repair deficiency by different assays raise the question of their mechanism. What mechanism is responsible for the phenotypes seen if the wild-type NER genes are truly autosomal dominant? Perhaps, in fact, if there is a clear effect, then the gene should be labeled as autosomal co-dominant and that there is haploinsufficiency of the mutated repair gene. This would mean that the "half-dose" of functional protein produced by the single remaining allele is not enough to deal with genotoxic insult. Another possibility is that there are functional polymorphisms present in the nonmutated repair gene. A polymorphism is a difference in genetic sequence among either individuals or groups. There may be a subtle change in function without a dramatic loss of functional protein. If the blood relative from one of these studies had

inherited both the mutant allele and an allele with a polymorphism that correlated with decreased repair capacity, this may result in the disease phenotypes present in this group.

Table 1.5 Experimental phenotypes of xeroderma pigmentosum heterozygotes.

Complementation Group	**Number of Patients Studied**	**Detectable Phenotype in Heterozygotes ?**	**Assay**	**Reference**
XPA	Unspecified	No	UV sensitivity UDS	(81)
XPA	3 sets of parents	No	Clinical history UV sensitivity UDS	(82)
XPA	1	Yes	HCR	(83)
XPC	2	Yes	HCR	(83)
XPD	1	Yes	HCR	(83)
Unknown	4	N	HCR	(84)
Unknown	3	2 Yes	UDS	(85)
Unknown	4	3 Yes	BrdU density gradient	(86)
Unknown	2	No	UDS	(87)
Unknown	1	No	BrdU density gradient	(88)
Various	7 fibroblast cell lines derived from parents	4 Yes	UDS	(89)

1.4.3 Polymorphisms of Nucleotide Excision Repair Genes and Predisposition to Cancer

Multiple molecular epidemiologic studies have focused on looking for polymorphisms in NER genes that can affect repair capacity. The majority of these have focused on the "tobacco-related" cancers (*i.e.*, ones seen with higher prevalence in active smokers and those exposed to secondary tobacco smoke): lung, head and neck squamous cell, prostate, bladder, breast, and esophagus. Because of the repair of bulky adducts formed by carcinogens (notably benzo[*a*]pyrene) found in tobacco smoke is carried out through the NER pathway, smokers are a logical choice of an exposed population. The studies cited showed cancer patients relative to smoker controls for the studies listed in Table 1.6. Because of the association of NER with skin cancers through XP, researchers have also focused on the presence of polymorphisms in UV-induced in skin cancers. Caveats of this analysis include the small study sizes often used. When an association is found, there is often a high variability associated with it. Over 72 different polymorphisms of varying allelic frequencies have been observed in 8 different NER genes (90). Approximately half of the studies listed showed a significant increase in cancer incidence associated with the polymorphisms studied.

Table 1.6 Selected studies of NER gene polymorphisms and risk of cancer as reviewed in Neuman *et al.* (91). The odds ratios represent the likelihood that a sample with the polymorphism of interest has come from a patient with cancer.

Gene	Reference	Cancer Type	N (Cases/Controls)	Odds Ratio (95% Confidence Interval)
XPA	(91)	lung	265/185	0.56 (0.35-0.90)
	(92)	lung	695/695	0.69 (0.53-0.90)
XPC	(93)	bladder	327/537	1.97 (1.10-3.57)
	(94)	breast	166/203	1.47 (1.00-2.16)
XPD	(95)	lung	1092/1240	0.98 (0.8-1.2)
	(96)	head and neck	189/496	1.65 (0.98-2.77)
	(97)	prostate	637/480	1.61 (1.03-2.53)
	(94)	breast	170/181	1.03 (0.70-1.51)
	(98)	esophageal	433/524	1.00 (0.68-1.49)
XPF	(99)	lung	96/94	No association
	(100)	breast	253/268	1.23 (0.72-2.09)
XPG	(101)	lung	310/311	0.54 (0.37-0.80)
	(102)	breast	220/308	1.50 (1.04-2.16)
ERCC1	(103)	lung	1752/1358	1.26 (0.81-1.96)

1.4.4 Cockayne Syndrome

In contrast to XP, the phenotypes associated with CS patients are primarily described as symptoms of premature aging. The neurologic symptoms of CS are often much more debilitating than those seen in XP and are progressive. Common features include: postnatal growth failure, cachetic dwarfism, retinal degeneration, deafness, mental retardation, neurodemyelination, calcifications in the cortex and basal ganglia, osteoporosis, and facial dysmorphism. Patients are sensitive to sunlight, but do not exhibit increased freckling as XP

patients do. In marked contrast to XP, CS patients do not have an increased predisposition to skin and other cancers.

In the US, CS affects under 1 in 250,000 live births (75). Like XP, CS is an autosomal recessive disorder. Also like XP, there are multiple complementation groups that produce the same syndrome phenotype: CSA and CSB. There is also a spectrum of disease severity that can be seen with CS that is not necessarily associated with the gene that is affected (104). Most patients present between the ages of 3 and 5, with the average life span of only 12.5 years. Three types of CS are classified (105):

Type I. ~80% of patients; early to late childhood onset, life expectancy 10-30 years

Type II. Co-natal; evident at birth; life expectancy 6-7 years

Type III. Late onset features; mostly normal growth and development; the most mild form of the disease

Cellular phenotypes associated with CS include: deficient transcription coupled repair (but no compromise in repair as measured by UDS), defective recovery of inhibited RNA synthesis, and increased sensitivity to both UV light and bulky chemicals that interact with DNA. While it is well-established that CS patient cells are defective in transcription coupled repair, the mechanism of pathogenesis is still not understood. CSA and CSB may participate in recognition of the stalled RNA polymerase complex at a lesion site (discussed above) leading to this deficit. The lack of either CSA or CSB also inhibits the cell's ability to bypass a stalled RNA polymerase complex. This role for CSA and CSB may also help to explain some of the phenotypes of CS through a general deficiency of transcriptional machinery, a phenotype CS

shares with Diamond-Blackfan anemia, in which there is a dysfunction of the S19 ribosomal protein (106).

CS cells are also more likely to apoptose in response to UV irradiation or exposure to bulky-adduct forming chemicals (107). Since specific inhibition of RNApol II also can produce this phenotype, the pathogenesis is thought to occur through a transcription-dependent process. The increased apoptosis is response to UV irradiation may provide an explanation as to why there is no increase in cancer in this syndrome. If the cells are more likely to die in response to genotoxic challenge, then there are fewer cells left alive that are genomically unstable.

Mutations in the XP genes XPB, XPD, and XPG cause a combination of the two disorders: XP-CS. Interestingly, all of the genes involved in the combined disorder are ones required for the formation of an "open complex" of unwound DNA during repair. XPB and XPD provide the helicase activity and XPG physically aids in holding open the unwound DNA. As in XP, there are genotype-phenotype correlations, where severity of disease is linked to which gene is affected and where in that gene a mutation falls. Patients exhibit skin cancer, dwarfism, cachexia, retinal degeneration, and progressive neurologic degeneration due to demyelination. The cellular phenotype is like that of XP, with reduced NER as measured by UDS.

1.4.5 Trichothiodistrophy

TTD is a rare autosomal recessive disorder characterized by sulfur-deficient brittle hair, ichthyosis, facial dysmorphism, and mental retardation (108). Eighty percent of cases also exhibit UV sensitivity. Patients do not usually exhibit dry skin, pigmentation abnormalities, or increase in skin cancer incidence. There are three complementation groups for TTD: one

exhibiting a lack of major UV sensitivity called TTD-A, and two with NER deficiencies caused by mutations in their XPB or XPD genes (overlapping with XP) (109,110).

Patients with mutations in TTD-A display a decrease in the cellular concentration of TFIIH. Therefore, a suggested role for the TTDA protein has been the protection of TFIIH proteins from degradation. Patients with XPD and XPB mutations also may have their pathogenesis rooted in transcription deficiencies since both these genes are also involved in the TFIIH complex. TTD cells may also exhibit a deficiency in NER as measured by UDS, however, only certain mutations in these genes are associated with a repair deficiency phenotype.

The majority of TTD patients do not suffer from an increase in skin cancer incidence. However, two patients have now been described that display a combination of the XP and TTD phenotype (111).

1.4.6 UV Sensitivity Syndrome

Patients have been identified who have mild phenotypes associated with both XP and CS. Like XP, these patients have acute recurrent sunburn, slight erythema and dryness, and freckling in sun-exposed skin. However, they also exhibit normal UDS levels. Like CS, these patients have deficient recovery of RNA synthesis after UV exposure and compromised transcription coupled repair. Unlike either CS or XP, these patients have normal growth, development, and lifespan.

Complementation analyses showed that these patients did not belong to any known group of XP or CS patients. The disease is now called UV sensitivity syndrome. No increase in skin cancer is seen in this disease. After complementation analyses failed, a microcell-mediated transfer technique was used to identify a complementing region on chromosome 10. Since the

CSB gene is also located on this chromosome, sequence analysis of CSB was performed. This showed a nonsense mutation in CSB. Complementation of these cells was then accomplished by injection of wild-type CSB cDNA. No explanation has been found for why previous complementation analyses failed. Other patients with UV sensitivity syndrome do not share this mutation in CSB, meaning that there are at least two complementation groups of UV sensitive syndrome and possibly an undiscovered NER gene.

1.4.7 Fanconi Anemia

Fanconi anemia (FA) is a complex autosomal recessive disorder with multiple associated phenotypes. Its incidence is slightly less of that seen in XP and CS at 1 in 360,000 live births. The age at diagnosis can vary with the severity of disease from birth to 48 years (average of 8.3 years). Clinical phenotypes associated with the disease are diverse and include: short stature, hyperpigmentation, dysmorphic features, mental retardation, malformations of the renal and urinary tract, pancytopenia, leukemia (primarily acute myeloid leukemia), myelodysplastic syndrome, and solid tumors. Laboratory diagnosis is made by finding an increased incidence of chromosomal breaks after exposure to a DNA crosslinking agent.

FA patient's cells have many multiple interesting phenotypes. First, they exhibit an increased number of spontaneous chromosomal aberrations in response to crosslinking agents. Cells cannot repair strand breaks and cell death is brought on by these agents. Another abnormality in these cells is a G2 phase of their cell cycle that is almost twice as long as unaffected cells. After treatment with one of these agents, a large portion of cells is trapped in the G2/M cell cycle checkpoint.

While the sensitivity of FA patients to crosslinking agents originally suggested a link to NER, the major pathway by which these lesions are remediated, further evidence now exists to show that this is not the case. FA cells are not sensitive to either ionizing radiation or UV light. Molecular characterization of the disease is starting to reveal an independent mechanism.

There are 12 different complementation groups associated with this disease: FANC-A, -B, -C, -D1, -D2, -E, -F, -G, -I, -J, -L, and –M (112). The genes for all but one of these complementation groups have now been identified. These proteins have little homology to one another and only a subset of them have homologies to any other known genes. Researchers have determined that eight of the genes work together in a complex that has ubiquitin ligase activity. These genes are FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, and FANCM. FANCL is the gene with the ubiquitin ligase activity. The roles of the other genes are unknown, except that their mutations as found in FANC complementation groups result in the failure of the complex to form. Although a gene for the FANC-I complementation group is as yet unidentified, cells from these patients demonstrate a lack of ubiquitinylation activity of the complex, suggesting that it is involved either within the complex or upstream of its effects. FANC-M is a recently identified member of the complex and has a helicase activity, perhaps identifying a role for translocation of the complex along DNA (113).

The target for ubiquitinylation of this complex is FANCD2. Patients in the FANC-D2 complementation group were shown to have functional mutations in their BRCA2 genes (114). BRCA2 functions by interacting with BRCA1 and RAD51, a complex which exists at sites of DNA damage as discussed below. The role of monoubiquitinylation of BRCA2 is also unknown.

FANCF and FANCJ are genes whose products do not seem to participate in the ubiquitinylation reaction common to the rest of the complex members. *FANCF* has a homology to a prokaryotic RNA-binding protein called ROM involved in stabilizing cRNA hairpin sequences. Its activity in human cell is unknown. FANCJ (formerly known as the BRIP1 gene) is involved somewhere downstream of the activity of the FANC complex, since monoubiquitinylation is unaffected in patient cells. This protein has a helicase activity, buts its role in the FA pathway is unknown.

1.4.8 The 46BR Case

One human case of a ligase 1 deficiency was described in 1982, patient 46BR (115). This patient was severely immunodeficient, with stunted growth, facial erythema, and sun sensitivity. She died at 19 years of age due to complications of a pulmonary infection. An SV-40 transformed fibroblast cell line was generated from the patient for further analysis of the cellular phenotype. This cell line showed a hypersensitivity to wide range of DNA damaging agents, including alkylating agents, IR, and UV light. Researchers theorized that the cells might be deficient in a late step of NER because of this array of damaging agents. The UDS assay showed that no deficiency in repair synthesis was present. Further studies, however, showed that the cells were deficient in semi-conservative DNA replication. Taken together, this combination of defects pointed to an inability of the cells to ligate their DNA. In NER, this constitutes the last step of the pathway. Cells would be able to complete gap-filling synthesis but not ligate the newly synthesized strand into place. In semi-conservative DNA synthesis, Okazaki fragments must be ligated together. Deficiencies in strand-break rejoining were believed to be responsible

for the defective response to some agents. Ultimately, it was confirmed that the mutation for the 46 BR case was in the ligase I gene (115).

1.4.9 Mouse Models of DNA Repair Deficient Diseases

A number of NER genes have been knocked out or otherwise altered in the construction of transgenic mice. While many of these mice show defects that mimic their human disease phenotype associated with their complementation group, others do not (116) (Table 1.7). For example, the dermatologic responses of the Xpa knockout mice mimic those of human, but the mice also display an immunodepression not seen in humans. Csa and Csb mice show a spontaneous loss of retinal photoreceptors not observed in the human disease phenotype, and their developmental abnormalities are much more mild. The most striking difference, however, is that Csa, Csb, and Ttda knockout mice all demonstrate an increase in UV-induced tumor formation, a phenotype specifically lacking from the human disease counterparts.

One reason for this difference in phenotype could be the difference that exists between rodents and human with respect to NER (117). Human and rodents respond to the same UV dose with vastly different rates and extents of repair (59). Rodent cells (specifically rat, mouse, and hamster) in particular are deficient in remediation of cyclobutane pyrimidine dimers relative to human cells (118), yet their survival is not compromised. Studies by Bohr and Mellon demonstrated that Chinese hamster ovary cells were proficient in transcription-coupled repair but severely deficient in global genomic repair (37,119). This effect is hypothesized to be due to the lack of expression of a DDB2 gene product. In humans, the protein coded by this gene is involved as an accessory to the recognition of DNA damage by the global genomic repair

pathway. Chinese hamster ovary cells expressing the human DDB2 gene are able to exhibit the same repair characteristics as human cells (120), "recovering" their global genomic repair pathway function. This paradoxical survival of rodent cells secondary to differences in the use of transcription coupled repair versus global genomic repair is commonly referred to as the "repairodox."

The differences noted in these mouse models bring up two important considerations (1) It highlights the need to use a human model system to study the link between repair deficiency and cancer. (2) It brings up the possibility that effects of genes can be tissue specific as well as species specific with different tissues being effected in different species, a consideration which will be discussed further below.

Table 1.7 Summary of mouse knockouts generated for NER genes. Data is taken from the Mouse Mutation Database version 5, which is updated annually.

http://pathcuric1/swmed.edu/research/research

Mouse Knockout	Phenotype	Heterozygotes
Xpa	Develop normally No enhanced mortality until 1.5 years Increased UV-induced skin and corneal cancer Total NER defect Acute photosensitivity Sensitive to 7,12-dimethylbenz[a]anthracene (DMBA) damage Lymphoid and lung tumors Immunosuppression: decreased natural killer cell activity	
Xpb	Non-mendelian inheritance (impaired embryonic development?) UV sensitive cells	
Xpc	Increased UV-induced skin cancer Increased AAF-induced liver and lung tumors No photosensitivity No increase in spontaneous tumorigenesis Defective in apoptosis Defective UDS	Increased UV-induced skin cancer after 1 year of life Increased AAF-induced liver and lung tumors
Xpd	Early embryonic death	
Xpe (DDB2)	Increased UV-induced skin cancer	
Xpf	Severe post-natal growth defects Death at 3 weeks of age Abnormal liver cells- enlarged nuclei Hypersensitivity to DNA-damaging agents	
Xpg	Early postnatal death Degeneration of Purkinje cells in the cerebellum UV sensitivity Defective in NER	

Table 1.7 (continued)

Mouse Knockout	Phenotype	Heterozygotes
Rad23B (hHR23B)	High frequency of embryonic death Early postnatal death Retarded growth Facial dysmorphism Male sterility: impaired testicular development	
Ercc1	Original attempts results in early embryonic death or death shortly after weaning with severe postnatal defects C-terminal truncation mouse: growth failure, liver abnormalities, prematurely senescent cells defective in NER sensitive to killing by UV	
Csa	Increased UV-induced skin cancer UV-sensitive Defective transcription coupled repair Mild growth failure Mild neurologic dysfunction Spontaneous loss of retinal photoreceptors	
Csb	Increased UV-induced skin cancer UV-sensitive Inactivation of transcription coupled repair Inability to result RNA synthesis after exposure Mild growth failure Mild neurologic dysfunction Spontaneous loss of retinal photoreceptors	
Ttda	increased UV-induced skin cancer (de Boer *et al.*. 1998).	

1.5 BREAST CANCER

Other diseases extend the paradigm link that exists between DNA repair and cancer illustrated by XP. One of these, hereditary nonpolyposis colorectal cancer, also known as Lynch syndrome II, involves a loss of post-replicative mismatch repair capacity. Mismatch repair is a form of "post-replication spell checking" that is responsible for catching base mispairings. While DNA polymerases have proofreading ability to increase their fidelity, mistakes are still made at a frequency of about 1 in 10^6 nucleotides. Patients with hereditary nonpolyposis colorectal cancer have an increased risk of breast, colon, endometrial, and ovarian cancers. The tissue specificity that arises from the mutation of a ubiquitously expressed protein (hMLH1 and hMLh2 in this case) is a source of much interest in cancer research. There is as of yet no data to indicate that a single mechanism explains this tissue specificity, rather, it is likely that multiple mechanisms are involved and contribute in different ways to the phenotype. These can include proliferation rates, immune surveillance, predisposition of the tissue for apoptosis, exposure to DNA damaging agents, cyclical acceleration and deceleration of proliferation, and the dependence of the tissue on genes that may have sequences that are particularly susceptible to different forms of damage.

It stands to reason that if loss of mismatch repair can contribute to breast cancer in the context of hereditary nonpolyposis colorectal cancer that other pathways of repair may also play a role in breast cancer etiology as well. Li-Fraumeni syndrome, BRCA1 and 2 mutations, ataxia telangiectasia, and Cowden's syndrome (discussed below) also all show the loss of a DNA repair mechanism and an increased incidence of cancer including cancer of the breast.

1.5.1 Incidence of Breast Cancer

Breast cancer is the most frequent malignancy diagnosed in women in the western world, affecting more than 2.3 million women in the United States in 2003 (121). In 2005, 40,410 were projected to die, 211,240 more diagnosed with an invasive breast cancer, and 58,490 more diagnosed with *in situ* breast cancer (122). The incidence of breast cancer (including both invasive and *in situ* cases) in the US has been increasing since 1980, although the rate of increase has been decreasing though the 1990s (122). Between 1975 and 1980 incidence rates were fairly constant. Compared with this the rate between 1980 and 1987 increased by 4% per year. Between 1987 and 2002, however, the rate of incidence increased by only 0.3% per year. Part of the increase in breast cancer diagnoses beginning in the 1980s can be attributed to the wider use of mammographic screening.

The use of screening mammography has not only increased the number of cancer cases seen due its ability to detect cancers too small to be felt during a physical exam, but also decreased the size, and therefore often the stage, at which most cancers are found. Tumors are being found at their earlier, less aggressive forms, including the non-invasive stage of *in situ* carcinoma. While some small tumors can be of a very aggressive type, already spreading to the rest of the body, the majority of these are classified as either ductal carcinoma *in situ* (DCIS), stage I or stage II (discussed below). These earlier forms of cancer are more treatable, with an increased chance for survival and decreased morbidity. As breast cancer is being "down-staged" by increased screening, there is a greater need for research into these early-stage tumors and their etiology.

1.5.2 Risk Factors

Many factors have been identified for breast cancer that contribute to risk assessment. These factors include both those that increase and decrease one's chance of having the disease. Types of risk factors include lifestyle, gynecologic events, presence of benign breast disease, mammographic phenotype, and environmental factors (123) (Table 1.8, Table 1.9). The following factors are all written from the perspective of whether or not they increase a patient's chance of having breast cancer.

The single greatest risk factor for breast cancer is increased age. The annual incidence of breast cancer in the US in women aged 80-85 years is 15 times greater than that of women aged 30-35 years (124). Risk can also be estimated using the gynecologic history of the patient (125). Nulliparity, later age at first live birth, early menarche, late menopause, and the use of estrogen replacement therapy all contribute to increased risk. Previous breast biopsies that have revealed benign breast disease are also a risk factor. Benign breast disease can be classified as proliferative or nonproliferative disease. Both of these contribute to breast cancer risk, although the magnitude of increased risk associated with nonproliferative disease is much smaller (123). Proliferative disease includes ductal hyperplasia, intraductal papilloma, sclerosing adenosis, and atypical hyperplasia, among others. Nonproliferative lesions include cysts, ectasia, and fibroadenoma. Women with mammographically dense breasts are also at increased risk for breast cancer. This includes the so-called "parenchymal" pattern seen in women with dense breast stroma (126). Part of the increased risk comes from the difficulty with which their mammograms are interpreted. However, evidence also suggests that there is a biologic difference in the breast tissue of women with dense breasts that also contributes to increased

cancer incidence. Mammographic patterns like the parenchymal pattern that are considered high-risk are more prevalent in countries with an increased incidence of breast cancer, suggesting that there may be an association between the two observations (127).

Of the above risks, the only one that is specifically modifiable is the age at first live birth. Fertile women can choose to have children earlier in life (before age 30). Other risks are inherent to the biology of the woman's breast (*i.e.*, density of tissue), her hypothalamic-pituitary-ovarian axis (*i.e.,* age of menarche and menopause) or the human condition (*i.e.*, aging). Risk factors from a woman's gynecologic history are believed to exert their effect through the exposure of the woman to estrogen. Early menarche and late menopause mean that a woman has gone through more menstrual cycles in her lifetime, including more cyclical increases in estrogen levels. Likewise, the use of estrogen replacement therapy after either her ovaries were removed, pharmacologically inactivated, or "shut-off" during the process of menopause means that she had an increase in estrogen exposure. The effect of an early term pregnancy along with mammographic density and benign breast disease may have more to do with the cellular make-up of the breast. Full-term pregnancy confers multiple changes to the breast epithelium including an increased proliferation of the epithelial cells. Benign breast disease and dense breast tissue represent alterations in the cellularity with poorly understood ramifications for the biology of the breast.

Few environmental factors have been definitively shown to contribute to increased risk. The best characterized are ionizing radiation and smoking (128). The total risk attributed to ionizing radiation depends upon the total dose and the time of the exposure and age at exposure, with younger women being at increased risk. This may indicate that there is a sensitive time during the development of the breast where environmental exposures are more damaging. The

effects of smoking can only be observed when smokers are compared to a group of never-smokers who are not passively exposed (128). The effects of both IR and smoking are predominantly believed to be affected through DNA damage. Multiple environmental toxins have been studied, but have not been shown to have a statistically significant effect on increased breast cancer risk. These studies were in large part carried out on white women in industrialized nations where large scale use of the suspected agents was no longer in practice (128). Environmental exposures represent one of the most modifiable forms of risk for breast cancer. Exposures can either be limited or sought out as protective, depending upon their effect.

A family history of breast cancer is also a significant risk factor, particularly for early onset breast cancer (129). The genetics of familial breast cancer are discussed below.

Table 1.8 Selected risk factors for breast cancer and their relative risk from Vogel and Bevers' 2003 *Handbook of Breast Cancer Risk-Assessment*.

Risk Factor	**Comparison Category**	**Risk Category**	**Relative Risk**
Age at menarche	16 years	Younger than 12 years	1.3
Age at menopause	45-54 years	After 55 years	1.5
Age at first live birth	Before 20	Nulliparous or older than 30 years	1.9
Benign breast disease	No biopsy or fine needle aspiration	Any benign disease	1.5
		Proliferative disease	2.0
		Atypical hyperplasia	4.0
Family history of breast cancer	No first degree relative affected	Mother	1.7
		Two first degree relatives	5.0
Obesity	10^{th} percentile	90^{th} percentile	1.2
Alcohol use	Nondrinker	Moderate drinker	1.7
Estrogen replacement Therapy	Never used	Current use $\geq$ 3 years	1.5

Table 1.9 Additional factors also investigated for their effect on breast cancer risk from Gierach and Vogel 2004. Numbers are reported with the perspective of increasing risk of breast cancer.

Risk Factor	**Effect**	**Odds Ratio, Relative Risk, or Hazard Ratio (95% CI)**
Body mass index; Premenopausal	Negative	RR= 0.54 (0.34-0.85)
Body mass index; Postmenopausal	Positive	RR= 1.26 (1.09-1.46) RR= 2.52 (1.62-3.93)
Endogenous hormones	Positive	Increasing quintiles of free estradiol vs. the lowest quintile (case-control of women who developed breast cancer vs. those who had not) RR= 1.38 (0.94-2.03) RR= 1.84 (1.24-2.74) RR= 2.24 (1.53-3.27) RR= 2.58 (1.76-3.78)
Estrogen metabolism	Negative ?	Ratio of 2-hydroxyestrone to 16-hydroxyestone in the highest tertile vs. the lowest two-thirds OR= 0.71 (0.29-1.75) OR premenopausal= 0.58 (0.25-1.34) OR postmenopausal= 1.29 (0.53-3.10)
Age at menarche	Negative	OR premenopausal= 0.91 (0.89-0.93) OR postmenopausal= 1.29 (0.53-3.10)
Age at first live birth	Positive	OR premenopausal= 1.05 (1.05-1.06) OR postmenopausal= 1.03 (1.02-1.04)
Parity	Negative	OR premenopausal= 0.97 (0.94-0.99) OR postmenopausal= 0.88 (0.86-0.90)
Breast-feeding	Negative	Every 12 months of breast feeding OR= 0.96 (0.94-0.97)
Pre-eclampsia	Negative	Ors range from 0.27 (0.08-0.63) to 0.81 (0.61-1.1)
Induced abortion	Null	RR= 1.00 (0.94-1.06)
Bone mineral density	Positive	Highest quartile vs. the lowest quartile RR= 2.7 (1.4-5.3)
Bone fracture	Negative	History of fracture vs. no fracture in the last 5 years OR= 0.80 (0.68-0.94)
TGFβ-1	Positive	Protective effect for women lacking common TGFB-1 genetic polymorphism vs. women with the common variant HR= 0.36 (0.17-0.75)
IGF-I	Positive	Top vs. bottom tertile RR= 2.9 (1.21-6.85)

Table 1.9 (continued)

Risk Factor	**Effect**	**Odds Ratio, Relative Risk, or Hazard Ratio (95% CI)**
Oral contraceptives	Null?	RR= 11.24 (1.15-1.33) RR= 1.11 (0.94-1.32)
Hormone replacement therapy	Positive	HR= 1.26 (1.00-1.59)
Exercise/Physical activity	Negative	ORs range from 0.3-1.6 with an average risk reduction of 30-40%
Alcohol consumption	Positive	12 grams/day vs. nondrinkers RR=1.06 (1.00-1.11)
Smoking	Positive Negative	OR parous, premenopausal= 1.69 (1.13-2.51) OR nulliparous, premenopausal= 7.08 (1.63-30.8) OR postmenopausal= 0.49 (0.27-0.89)
Breast implants	Null?	RR= 0.72 (0.61-0.85)
Diet	Null	Fats compared with equivalent energy intake from carbohydrates for an increment of 5% of energy RR= 1.09 (1.00-1.19) saturated RR= 0.93 (0.84-1.03) monounsaturated RR= 1.05 (0.96-1.16) polyunsaturated
β-Carotene	Negative	OR= 0.41 (0.22-0.79)
Lycopene	Negative	OR= 0.55 (0.29-1.06)
Total carotene	Negative	OR= 0.55 (0.29-1.03)
Folate	Negative	Lowest tenth percentile of folate intake vs. fiftieth percentile RR= 1.21 (0.91-1.61) Among drinkers of greater than 4 grams/day RR= 1.59 (1.05-2.41)
Phytoestrogens	Null	Highest vs. lowest quartile OR= 1.0 (0.80-1.3)
Ionizing radiation	Positive	OR= 1.4 (1.2-1.8) RR= 4.1 (2.5-5.7)
p,p'-bis(4-Chlorophenyl)-1,1-dichloroethane	Null	OR= 1.20 (0.76-1.90)
Chlordane	Null	OR= 0.98 (0.62-1.55)
Dieldrin	Null	OR= 1.37 (0.69-2.72)
Polychlorinated biphenyl congeners	Null	OR= 0.83 (0.54-1.29)
Electromagnetic fields	Null	OR= 1.12 (1.09-1.15)
HIV infection	Null?	OR= 0.18 (0.04-0.76)

1.5.3 General Breast Anatomy

The epithelial component of the breast is a complex system of branching ductwork extending from the nipple and ending at the terminal ductal lobular units (Figure 1.6). Breasts are actually modified sweat glands with interspersed adipose and connective tissue. Ninety percent of breast cancer arises in the mammary epithelium, the layer of epithelial cells lining the lobules and ducts of the breast's architecture. Breast adenocarcinoma can be classified by the location of the cancerous cells, either present in the lobules (15%) or ducts (85%) (122). While the characteristics of lobular carcinomas are often those associated with a better prognosis as compared to ductal carcinomas, the overall survival of patients is the same in the two groups (130). Some studies have shown that the response to chemotoxic chemotherapies is worse in lobular carcinomas (130), yet this does not affect treatment planning at the current time.

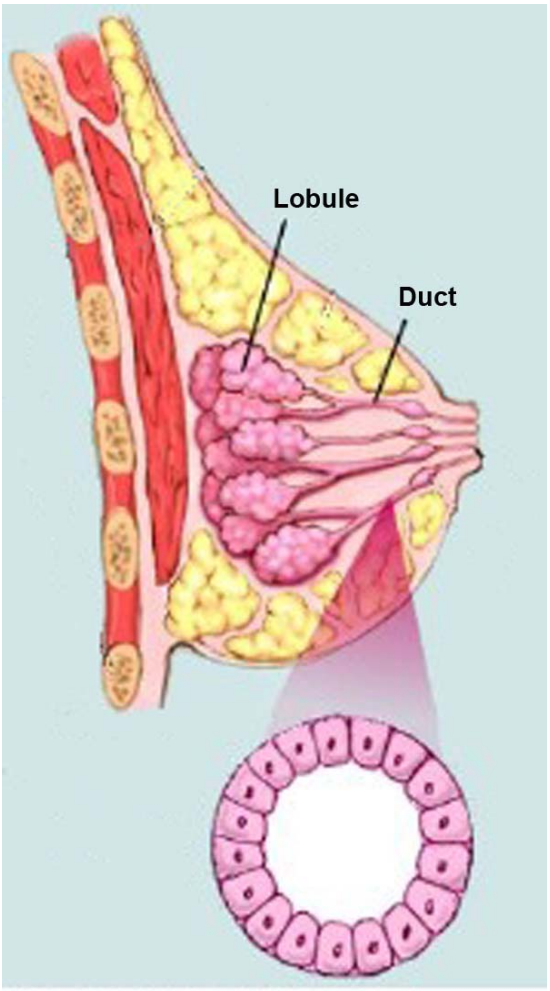


Figure 1.6 Diagram of a human breast showing the progressive branding ducts extending from the nipple back to the terminal ductal lobular units. A cross section of the architecture depicting the epithelial cells lining both the ducts and lobules is shown at the bottom of the figure. Copyright Mary K. Bryson, reprinted with permission.

1.5.4 Staging of Breast Cancer

Breast cancer is staged and graded. Staging refers to the extent of disease within the body, including both the size of the tumor and its spread. Stage is determined using the TNM system devised by the American Joint Committee on Cancer and National Cancer Institute and revised in 2002 (131) (Tables 1.10 and 1.11). In this system T refers to the tumor size, N refers to lymph node involvement, and M refers to the extent of metastases beyond lymph nodes. Tumor size is determined by the tumor's greatest dimension as measured either after the tumor is removed, or, when this is not the case, by radiographic imaging. Lymph node status is either assessed by aspiration of fluid from the lymph node in question, or by surgical removal of the lymph node and histologic assessment for the presence of disease. Finally, distal metastases can be assessed by multiple methods depending upon the organ involved. The level of screening for distal metastases is often driven by symptoms experienced by the patient, but is at the discretion of the medical practitioner.

Table 1.10 Categories of the TNM staging system for breast cancer.

T	**Description**
Tx	Tumor cannot be assessed
T0	No evidence of primary tumor
Tis	Carcinoma *in situ* of any size or Paget's disease without a detectable tumor mass
T1	Tumor ≤ 2 cm
T2	Tumor > 2cm but ≤ 5 cm
T3	Tumor > 5 cm
T4	Tumor of any size with direct spread to the chest wall or skin; any inflammatory carcinoma

N	**Description**
N0	Regional lymph nodes are free of metastasis
N1	Ipsilateral axillary lymph nodes are affected but still movable
N2	Ipsilateral axillary lymph nodes are affected and fixed to one another or to another structure
N3	Ipsilateral internal mammary nodes are affected

M	**Description**
Mx	Cannot be assessed
M0	No distant metastases found
M1	Distant metastases are present

Once the T, N, and M status of a patient is determined, he or she can be staged using the formula in Table 1.11. If one or more of these scores is not completed, staging cannot be appropriately determined.

Table 1.11 The TNM staging schema developed by the American Joint Committee on Cancer and National Cancer Institute as revised in 2002.

Stage	**T (tumor size, invasiveness)**	**N (node involvement)**	**M (extent of metastasis)**
0	*In situ*	N0	M0
I	T1	N0	M0
IIa	T0, T1	N1	M0
	T2	N0	M0
IIb	T2	N1	M0
	T3	N0	M0
IIIa	T0, T1, T2	N2	M0
	T3	N1, N2	M0
IIIb	T4	Any N score	M0
	Any T score	N3	M0
IV	Any T score	Any N score	M1

The stage of a patient is used clinically as a guideline for treatment decisions. It can also be used to assess the severity of the disease, and thus the chances for successful treatment and survival of the patient. The distribution of stage at which cancers are diagnosed in the US is still reported by SEER using the historic staging schema, as described in Table 1.12.

Table 1.12 The stage distribution of breast cancer in US women, all ages, diagnosed between 1996 and 2002 as reported by SEER. The "Unknown" column represents cases that were diagnosed for which the SEER database has no official staging information.

Stage	% of Tumors Diagnosed at this Stage
Localized: confined to the primary tumor site	61%
Regional: spread to the lymph nodes directly beyond the primary site	31%
Distant: cancer that has already metastasized	6%
Unknown	2%

While the majority of women are diagnosed with early stage tumors, there is a relative paucity of information known about the genetics of breast cancer etiology versus the genetics of progression (Figure 1.7).

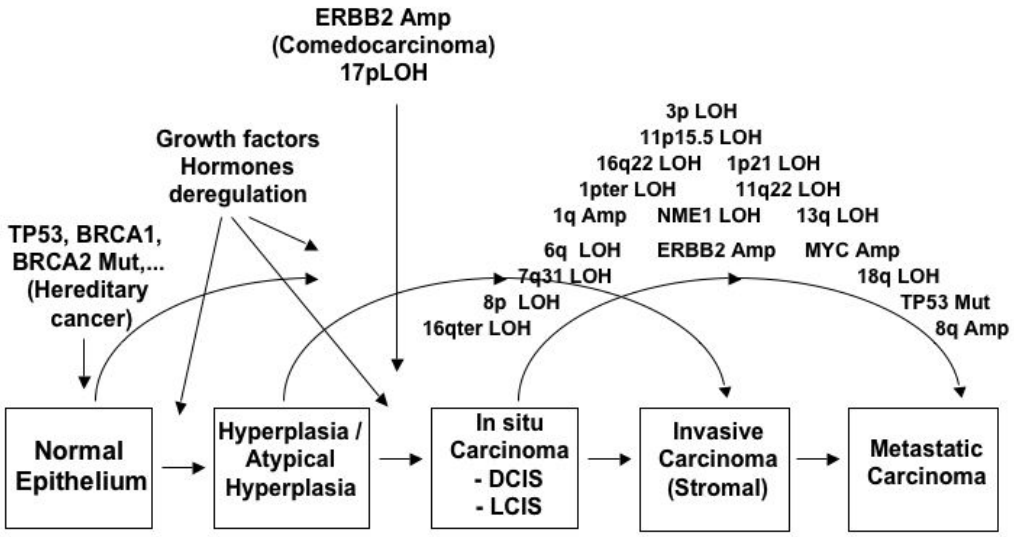


Figure 1.7 Histological and genetic pathway of breast cancer, adapted from Bieche and Lidereau (132); reprinted with permission.

1.5.5 Grading of Breast Cancer

In contrast to staging, grading of cancer is performed by examining at the pathologic qualities of the tissue at the microscopic level. As the American Joint Committee on Cancer's TNM classification system was developed for staging, the modified Scarff-Bloom-Richardson tnm system, also known as a Nottingham score, was developed for pathologic grade (133) (Table 1.13). This system assesses the presence of tubular structures in the breast's architecture, the amount of nuclear pleomorphism (multiple structural forms), and the mitotic index. Nuclei are usually small and are regular in size in the breast epithelium. When the nuclear to cytoplasmic ratio in the cells becomes more variable, with multiple size nuclei being observed on the slide,

this is called nuclear pleomorphism. Each category is given a numeric score as shown in Table 1.10. The score can either be reported as the total score out of 9, or can be broken down into three categories depending upon the convention of the pathologist or the institution where the grading is being performed. Grade 1 is associated with the best prognosis in includes scores of 1-3. Grade 2 is a score of 6 or 7, and Grade 3 is a score of 8 or 9. This pathologic grading system was designed to decrease inter-institutional and inter-individual variation (134).

Table 1.13 The modified Scarff-Bloom-Richardson pathologic grading system for breast cancer.

t (Tubular Structures)	**% of Carcinoma Composed of Tubular Structures**
1	> 75%
2	10 – 75%
3	< 10%

n (Nuclear Pleomorphism)	**Nuclear Pleomorphism**
1	Small uniform nuclei
2	Moderate increase in size and variation of nuclei
3	Marked variation in nuclei

m (Mitotic Index)	**Mitotic Count (per 10 high power fields)**
1	≥ 7
2	8 – 14
3	≥ 15

1.5.6 Receptors Used to Classify Breast Cancer

1.5.6.1 Estrogen Receptor

The estrogen receptor (ER) is an intracellular steroid hormone receptor responsible for a cell's response to the effects of estrogens (135). As is the paradigm for steroid hormone receptors, the estrogen receptor is activated when estrogens pass through the cell membrane and bind to the ER. Once the ligand is bound, the ER dimerizes to become active. The dimerized and ligand-bound ER translocates to the nucleus where it acts as a transcription factor. The complex binds to sequences known as estrogen response elements (EREs) in the control regions of genes. EREs are found in a number of genes that control growth and development, including the progesterone receptor.

Estrogen is a potent mitogen in the normal breast (136). When combined with the action of anterior pituitary hormones, it is also responsible for breast development. Systemically, estrogen stimulates the release of growth hormone. Together estrogen, growth hormone, peptide hormone, and prolactin promote ductal development in the breast.

Human mammary development begins at 5 weeks, with the mammary gland developing from the milk lines, a thickening of the epidermis on the ventral surface of the fetus (137). Basal epithelium and stroma form. In the 15th week, androgens cause the mesenchymal condensation around the epithelial stalk at the chest wall. The growth of solid epithelial columns occurs, each giving rise to a lobe of the mammary gland. The fetal dermis encases these columns giving rise to the surrounding fibrous stroma.

At 7-9 months of gestation, canalization of the epithelial cords in the human mammary gland occur (137). Development of branching lobuloalveolar glandular structures also form at

this time under the influence of HGF and paracrine factors from the adipocytes (peptide and steroid hormones). Prolactin induces secretory activity in the luminal epithelial cells at this stage. The human breast develops completely, in the virgin state unlike that of rodents which require the additional stimulation of copulation and or pregnancy to form lobuloalveolar structures. In postnatal development there is palpable enlargement of the breast bud, which subsides during the first and second weeks of pregnancy as the maternal hormones subside.

The effect of estrogen on breast cancer is one of the most studied areas in research (136). Experiments *in vivo* have shown that cancer cell lines that still express the estrogen receptor have an increased growth rate when estrogen is added to their culture medium. The increased risk of breast cancer shown in patients using estrogen replacement therapy as well as its known effects on proliferation have supported the theory that estrogen itself may contribute to the formation of breast cancer at the molecular level. Other theories maintain that estrogen's proliferative effects may only feed the growth of already genomically unstable cancer cells. This theory posits that the mutagenic effects of estrogen is a result of its mitogenic effects (138). When cells are stimulated by estrogen, they go through their cell cycle at a faster rate, giving them less time to repair damage before S-phase occurs. The cells then replicate their DNA and fix the damage into their genome as mutations. This theory has found support in the effects of radiation poisoning (139) and genotoxic chemotherapeutic agents (18). Rapidly replicating cells (*e.g.*, gastrointestinal track lining, bone marrow) are the ones that initially show damage after a dose of radiation. In genotoxic chemotherapy, the agents target tumors because they are rapidly replicating, with side-effects being the genotoxic events experienced by other quickly replicating tissues (gastrointestinal track lining, bone marrow, and hair follicles). This "mitogen-mutagen" theory assumes that either there is a defect in cell cycle regulation that allows damaged cells to

progress before DNA repair can occur or that there is an increase in damage tolerance systems like translesion (error-prone) synthesis (140).

ER status is one of the most important prognostic factors associated with breast cancer. In women diagnosed between 1992 and 1998, 76% of their tumors were ER + in all age groups (141). With breast cancer in general, older women are more likely to have ER+ tumors than are younger women. Women with lower stage tumors were also more likely to express ER in their tumor (Table 1.14).

Table 1.14 Prevalence of estrogen receptor positive tumors by stage in women diagnosed between 1992 and 1998, all ages (141).

Stage	**Average % tumors positive for ER**
I	81.1
II	72.2
III / IV	66.7

ER negative tumors do not respond to the presence of estrogen (at least via the ER pathway), and utilize growth abilities independent of this mechanism. Patients with ER- tumors have a higher recurrence rate and a decreased overall survival. Part of this may be due to the fact that antiestrogen or organ ablation therapies, which have been used with great success in ER + patients, are not effective in this group. Part may be an ER- tumor is intrinsically more aggressive.

1.5.6.2 Progesterone Receptor

Like the estrogen receptor, the progesterone receptor mediates the body's response to a steroid hormone though dimerization, translocation to the site of action, and action as a transcription factor. Like many other factors including peptides and hormones (inhibin and TGFß, *etc*.) progesterone can contribute to lobuloalveolar development of the breast (136). This action of progesterone requires the presence of the hormone prolactin, and can only take place when preceded by or concurrent with estrogen.

While its role in breast carcinogenesis is less well-understood than ER, it can also be used as a prognostic factor. Overall, 66.5% of patients have PR + tumors (all ages, 1992-1998) ((141)) (Table 1.15). Patients whose tumors have both ER + and PR + have decreased mortality when they are treated with endocrine-based therapies. However, patients whose tumors are ER + and PR - have a worse outcome, and do not have the same decrease in mortality with endocrine-based therapies. Because progesterone receptor is a downstream target of ER, its lack of response to the signals through ER is considered to be a sign of aberrant growth factor signaling within the cell.

Table 1.15 Prevalence of estrogen receptor positive and progesterone receptor positive tumors in women diagnosed between 1992 and 1998, all ages (141).

	PR+	**PR-**
ER+	63.0	12.6
ER-	3.6	20.6

1.5.6.3 Her-2/Neu

Her-2/Neu, also called c-erb-2, is a cell surface receptor tyrosine kinase that is activated by Epidermal Growth Factor (EGF). Like the ER and PR, Her-2/Neu dimerizes and translocates to the nucleus. Unlike them, no natural ligand has yet been found that specifically binds to Her-2/Neu. Dimerization occurs after epidermal growth factor binds to one of the other members of its family of receptors: the epidermal growth factors receptors (142). Her-2/Neu is then phosphorylated by its binding partner. Overexpression of this receptor leads to autophosphorylation and the induction of multiple downstream pathways including phosphatidylinositol 3-kinase, Akt/protein kinase B, and NF-κB (143). The NF-κB pathway is particularly well-characterized, with one of the downstream effects from Her2/neu overexpression being multiple cell survival signals as seen *in vitro*. The increase of cell-survival signals protects these cells from the apoptosis induced by many chemotherapeutics agents *in vitro* (144).

Twenty to thirty percent of invasive breast tumors overexpress Her-2/Neu on their cell surface due to an amplification of the gene, making it a desirable therapeutic target for these patients (143). It has been estimated that up to 60% of cases of ductal carcinoma *in situ*

overexpress this receptor, and nearly 100% of the large-cell comedo-type DCISs (145). Patients who overexpress Her-2/Neu have a lower disease-free survival and overall survival than patients who do not, as well as a decreased time to recurrence (146). This trend, however, is changing with the emergence of therapies to target this receptor.

1.5.7 Molecular Characterization of Breast Cancer

In addition to these well-known markers of breast cancer, many more are currently being found through the use of expression microarrays. Characteristic gene expression patterns in breast cancer have been identified and used to classify tumors into subtypes representing some of the heterogeneity of disease in breast cancer (147). The gene expression patterns defining these groups have been refined multiple times by the authors including data from other large group sets (148). It took 85 microarray profiles to determine these patterns. Researchers have now used 122 samples, with data coming from multiple sources and have confirmed that the results are conserved across multiple cDNA microarray platforms (149). Sorlie *et al.* identified 5 subtypes of breast cancer: basal epithelial-like, ERBB2-overexpressing, normal breast-like, luminal A, and luminal B (Table 1.16).

Table 1.16 Subtypes of breast cancer as defined by microarray expression profiles by Sorlie *et al.*

Subtype	**Characteristics**	**Overall Survival**
Basal epithelial-like	High expression of keratins 5 and 17, laminin, and fatty acid binding protein 7	Worst
ERBB-2 overexpressing	High expression of several genes in the ERBB2 amplicon at 17 q22.24	Poor
Normal breast-like	Highest expression of multiple genes expressed by adipose tissue and nonepithelial cell types; higher expression of basal-type genes and lower expression of luminal genes	Middle
Luminal A	Highest expression of ER α, GATA binding protein 3, X-box binding protein 1, trefoil factor 3, hepatocyte nuclear factor3 α, and estrogen-regulated LIV-1	Best
Luminal B (C)	Lower expression of the genes used to define Luminal A; may be further broken down into tertiles as luminal A, B, and C	Similar to Normal breast

The authors also looked at survival analyses of the different subtypes (147). Included in the analyses were only the 49 women from the original group used to define the subtypes who had locally advanced disease, no distal metastases, and had received uniform treatment with anthracycline and a taxane followed by surgery. Though the authors do not specify stage by the AJCC system, this would correlated to predominantly stage II and possibly some stage III tumors. They saw statistically significant differences in both disease-free survival and overall survival in the subtypes. Overall, the basal-like and ERBB-2 overexpressing groups had the

worst outcomes. These molecular characterizations can provide a list of new potential therapeutic targets, and perhaps the beginning of a way to tailor therapies for women based upon their different subtype.

1.5.8 Racial Differences in Breast Cancer

Further evidence for the heterogeneity of breast cancer can be seen in the racial differences in this disease. Research has shown that overall, African American women have a significantly lower incidence rate of breast cancer than white women. Despite this, African-American women are more likely than white women to die from the disease (121). Breast cancer survival at 5 years after initial diagnosis is only 73.8% among African-American women compared to 81.6% among white women (150). Further, African-American women are less likely to be diagnosed with smaller tumors (less than or equal to 2.0 cm, stages I and II) and more likely to be diagnosed with larger tumors (greater than 2.0 cm) than white women of the same age (150). In the past, these differences have in part been attributed to socioeconomic factors such as the lack of health insurance, lower incomes, and unequal access to medical care, including screening mammography. However, increasing evidence exists to support the hypothesis that there may be additional intrinsic biological differences between the breast tissues of white and African-American women.

One of the biological differences in African American women's breast tumors versus white breast tumors is a difference in the rate of African American women's tumors in expression ER and PR. African American women are more likely to have ER- and PR- tumors compared to white women. Using SEER statistics, Field *et al.* looked at female breast cancers in

the US between 1993 and 1998. During this time period, 59.9% of white women's breast tumors expressed the estrogen receptor, while only 45.6% of African American women's did (150). Twenty five and a half percent of white women's tumors expressed the progesterone receptor, while only 23.7% of African American women's did the same. Statistical significance was not reported for the difference in the two groups for either ER or PR status.

African American women's tumors were as likely to overexpress Her-2/Neu (31.7% versus 29.1%; P = 0.46), but the relationship of this marker to other markers of breast cancer prognosis was changed (151). While in white women Her-2/Neu expression correlated with increased stage, poorer histologic grade, and increased prevalence of lymph node metastases, none of these correlations hold true for African Americans. There were no statistically significant correlations observed in this study between Her-2/Neu status and stage, grade, or other pathologic prognostic factors. This suggests that a biologic difference exists in African American breast cancer in the way it responds to the effects of Her-2/Neu stimulation.

With respect to the subtypes of breast cancer determined by microarray analysis, the distribution of patients in blacks versus whites was different (152). Using 496 cases that were enrolled in the Carolina Breast Cancer Study, Carey *et al.* classified the tumors according to a modified version of the breast cancer subtypes identified by Sorlie *et al.* This work was performed using the same platform, in conjunction with the Sorlie research group, with confirmation of the previous classifications done before proceeding on to analyze racial differences. Their modifications included a subdivision of the HER2-overexpressing group into ER+ and ER-. The differences in race shown here with these multiple marker patterns provides further evidence that there is a biological difference in breast cancer between them (Table 1.17).

Table 1.17 Breast cancer subtype by race as determined by Carey *et al.* in 2006 using microarray subtypes.

Breast Cancer Subtype	**African American**	**Non-African American**
Basal-like	27%	16%
HER2+/ER- (Her-2/Neu overexpressors)	8%	6%
Luminal A	47%	54%
Luminal B	13%	17%
Unclassified	5%	7%

1.5.9 Standards of Care for Breast Cancer

Treatment of breast cancer can be considered as having two main targets: the elimination of gross disease and the elimination of micrometastases. These can be performed with the goal of either prolonging the life of the patient (curative), or easing the patient's pain and suffering. Gross disease indicates accumulations of cells into discrete tumors while micrometastases indicates individual cancer cells present in the body separate from these tumors. To that end, therapy is designed for each woman to both remove or kill her tumor(s) and any other cancer cells that exist in her body based upon the stage of the tumor, its characteristics, and the health of the patient. While many different regimens for breast cancer therapy exist, they can be broken down into three treatment modalities that can be used alone (definitive), in conjunction with one another (adjuvant), or to control symptoms and pain at the end of life (palliative). These modalities are: chemotherapy, surgery, and radiation therapy. In general, surgery is designed to remove gross disease, radiation therapy is used to control local recurrence, and chemotherapy is designed to decrease the existence of micrometastases.

For stage 0 (noninvasive carcinoma) total mastectomy has been the standard of care. However, because equivalent survival rates have been achieved with breast conserving surgery (segmental mastectomy or lumpectomy) combined with radiation therapy, conservative management is now as popular (153). Radiation therapy is used to control local recurrence in this case, decreasing incidence from as high as 25% down to 5-10% at 5 years (154). For lobular carcinoma, breast radiation is given to both breasts, since the risk of ipsilateral and contralateral recurrence is equal (130). Radiation therapy can either be given as an external beam therapy over the course of 6 weeks or through the use of brachytherapy- radioactive seeds placed in the tumor cavity and removed after a short duration of treatment (MammoSite) (155).

For early stage breast cancer (I and II), local treatment includes surgery, radiation therapy, and chemotherapy (156). Surgery can either be a modified radical mastectomy, or a breast conserving procedure coupled with radiation therapy. Axillary node dissection can be used as either a treatment to remove suspect lymph nodes or as part of a staging work-up to determine their status. The more conservative, and more recent, approach of sentinel node biopsy is also used for this purpose, potentially sparing patients significant lymphedema associated with full axillary dissection. In this case, the nodes directly responsible for draining the site of the tumor are identified and removed, sparing the rest of the axilla. The preference for sentinel node biopsy versus axillary dissection is in some cases surgeon- and hospital- specific. In general, sentinel node is the preferred method when determining the involvement of local lymph nodes because of reduced morbidity. Axillary dissection is used when lymph nodes are already suspected to be involved in the disease after either physical exam (a hard mass in the axilla either freely movable or attached to other structures) or after ultrasound examination

showing a mass (157). Radiation therapy is extended to include the axilla in the case of positive nodes (154).

The role of adjuvant chemotherapy in stage I and II patients is highly debated and studied. Chemotoxic chemotherapeutic regimens offer an increased survival to most patients, but also include an increase in morbidity. While it is offered as a standard of care to node positive patients, it is often difficult to determine the benefit for node negative patients. Tumor size and hormone receptor status are the two factors used most often to determine the appropriateness of this therapy. While many have tried, no widely accepted categorization of node negative risk subgroups exists. In general, women with low histologic grade and a tumor of less than 1 cm in size are not given adjuvant chemotoxic chemotherapy because no clear benefit has been demonstrated using any chemotherapeutic regimen. It is clear that more research must be done to identify women for whom chemotoxic chemotherapy is beneficial. The present most commonly used regimen for adjuvant chemotherapy is four cycles of the combined therapy of doxorubicin (Adriamycin) and cyclophosphamide (Cytoxan) (158) (Table 1.18). For women who have local spread of disease (involved lymph nodes) additional chemotherapy may be advised. For these women, they are given 6 cycles of either paclitaxel (Taxol) or docetaxel (Taxotere), both of which are microtubule inhibitors (158).

Table 1.18 Common chemotherapeutic agents used in the treatment of breast cancer(159).

Generic Name	**Common Brand Name**	**Class**	**Mechanism of Action**	**Damage Remediated by NER**
Doxorubicin	Adriamycin	Anthracycline Antibiotic	intercalating agent that blocks DNA synthesis and transcription and can inhibit topoisomerase II causing strand breaks	Yes
Dyclophosphamide	Cytoxan	Alkylating Agent	bis(cloroethyl)amine: forms bulky adducts causing DNA damage, single and double-stranded breaks	Yes
Paclitaxel	Taxol	Antimitotic Agent	arrests cells at G2/M phase of the cell cycle by preventing microtubule depolymerization	No
Docetaxel	Taxotere	Antimitotic Agent	Same as paclitaxel	No

For advanced breast cancer (stage III and IV), preoperative chemotherapy is now becoming standard of care (160). Pre-operative chemotherapy can be used to shrink a large tumor to a more operable mass at the same time local or distal metastases are being treated. Patients then undergo a combination of surgery to remove the breast mass, axillary node dissection, and radiation therapy. Chemotoxic chemotherapy may be continued after the surgery. Therapy at this stage is still meant to be curative, and high rates of local and regional control have been achieved.

Many treatments being used now are based upon the biologic characteristics of individual tumors. These include ER status and the presence of overexpression of Her-2/Neu. These other treatments can be integrated into multimodal care as described below.

Patients who have ER + disease are treated with either a selective estrogen receptor modulator (SERM) such as tamoxifen if they are premenopausal or an aromatase inhibitor if they are postmenopausal. SERMs ideally block the proliferative effects of estrogen in some tissues (including breast) while preserving its actions in others. Specifically, attempts are made to preserve the action of estrone with respect to bone mineralization, serum cholesterol and coagulation factor regulation. For premenopausal women, the ovaries are the primary source of the majority of their circulating estrogen. SERM therapy seeks to allow the protective actions of estrogen throughout the body while blocking its activity in the breast. ER+ patients are offered a 5-year course of the SERM tamoxifen regardless of stage to prevent the effect of estrogen on any micrometastases still present after treatment or, rarely, on newly formed primary tumors. At the present time, a study has been released that shows that the SERM raloxefene is equally effective in preventing breast cancer recurrences when compared head-to-head with tamoxifen (161). Raloxifene is a drug approved for use as a treatment and prevention for post-menopausal osteoporosis in 1997 (162). It acts as an estrogen agonist in the skeleton, on serum lipid metabolism, and on multiple serum coagulation factors. It acts as an agonist in the breast and the uterus. Consistent with this, the study found that there were fewer pathologic fractures and fewer thromboembolic events in the group of women taking raloxefene. This study may lead the way for raloxefene becoming the SERM of choice for breast cancer prevention and treatment.

For women who are postmenopausal, the majority of estrogen comes from the conversion of androgen to estrogen (estrone) in adipose tissues rather than from their ovaries through the

enzyme aromatase. In these women, an aromatase inhibitor like anastrozole (Arimidex) is the therapy of choice (163). Studies are currently being conducted to look at the effectiveness of other SERMs such as raloxefene or the benefit of an additional 5 years of therapy (after the initial 5 years of treatment with a SERM) with either a SERM or aromatase inhibitors depending upon menopausal status.

For women whose tumors overexpress Her-2/Neu, monoclonal antibody therapy is now available. This therapy is based upon a monoclonal antibody, trastuzumab (Herceptin) that is capable of recognizing an epitope of the overexpressed receptor. This antibody works through multiple mechanisms. First, it helps to target the tumor for immune destruction. It also prevents growth of the cancer cells through blocking the signals being sent through the growth factor receptor. Finally, studies have shown synergistic effects between Herceptin and chemotoxic chemotherapies such as cisplatinum through an unknown mechanism. The first studies conducted with Herceptin showed viability as an end-stage treatment for breast tumors overexpressing Her-2/Neu. Clinical trials have been conducted which show excellent results in 5 year survival in metastatic patients where Herceptin is used as a first-line adjuvant therapy with paclitaxel, a chemotoxic chemotherapeutic drug that inhibits microtubules that has already been shown to decrease mortality in this group. Current studies are now being conducted to show Herceptin's benefits as a neo-adjuvant treatment for locally advanced cancer and an adjuvant treatment for earlier stage tumors as well.

1.5.10 Breast Cancer Survival Rates

Breast cancer survival rates are often given in terms of "5-year survival." This refers to the percentage of patients who live at least five years after their cancer is diagnosed. Ten- year survival rates can also be given. Relative survival rates assume that cancer patients will die of other causes and compares the observed survival with that expected for people of their age group without breast cancer. Thus, relative survival denoted the number of deaths that have occurred only due to breast cancer.

The Institute of Medicine has reports both 5 and 10 year relative survival rates for breast cancer based upon the primary tumor at the time of diagnosis (121) (Table 1.19). Survival rates are listed as per the historic staging system as per Table 1.12. Their most recently published data includes women of all races and ages who were diagnosed between 1998 and 2002. Their follow-ups to determine whether the women were still alive include data through 2003.

Table 1.19 Relative breast cancer survival ratesby Stage as Reported SEER. All races and all ages of female breast cancers included in their registries as diagnosed between 1988 and 2002 are included (121).

Stage	5 Year Relative Survival Rate	10 Year Survival Rate
Localized (I)	97.5%	93.5%
Regional (II or III)	79.5%	66.5%
Distant (IV)	24.2%	13.3%
Unstaged	58.1%	45.8%

1.5.11 Hereditary and Familial Breast Cancer

Five to ten percent of breast cancer is considered inherited, with the woman inheriting a single cancer susceptibility gene with a mutation that confers increased risk (123). In addition to this group, another 15 – 20% of cancers are considered to be familial. These cancers arise in the context of a positive family history for breast cancer, or other cancers as discussed below (123). The likely etiology of the tumors is a result of multiple shared genetic and environmental factors over time. The other 70% of cases are considered sporadic with no known common cause.

Hereditary breast cancer is characterized by an earlier age at diagnosis (<50 years), bilateral disease, multiple primary tumors (can be in different tissues), male breast cancer, other cancers or features associated with genetic syndromes, and possibly other family members with a known mutation in a susceptibility gene (123). Hereditary syndromes include: Cowden syndrome, ataxia telangiectasia, Peutz-Jeghers syndrome, and Li-Fraumeni syndrome (Table 1.20). Additionally, mutations in BRCA1 and -2 predispose carriers to breast and other cancers. While these syndromes are usually associated with the germline mutation of the genes associated with them, a *de novo* germline mutation or somatic mutation of any of these genes in the tissue developing a breast tumor is also possible and have been reported in the literature. For example, BRCA1 mRNA and protein are either absent or significantly reduced in 30-40% of sporadic breast cancer cases (164,165).

Cowden's syndrome, also known as multiple hamartoma syndrome, is caused by a mutation in the PTEN gene, a tumor suppressor gene that acts in cell cycle arrest and apoptosis (166). PTEN is a PI3 phosphatase that negatively regulates AKT/PKB signaling pathway. The disease is autosomal dominant with incomplete penetrance affecting 1 in 200,000 in the US.

Twenty to fifty percent of affected females will have a breast adenocarcinoma diagnosed in their lifetime. The primary phenotypes of the disease are mucotaneous (involving the mucosal lining) lesions including hamartomas of the hair follicles, thyroid abnormalities like goiter and adenoma, gastrointestinal lesions such as hamartomatous polyps, macrocephaly, and genitourinary abnormalities such as uterine leiomyoma.

Ataxia telangiectasia is an autosomal recessive disease with prevalence estimated between 1 in 40,000 to 1 in 100,000 live births (167). Heterozygote frequency is estimated at as high as 1 in 100 (167). The ATM gene is a PI3/PI4 kinase involved in cell cycle checkpoint regulation after DNA damage through phosphorylating a number of downstream targets including TP53, BRCA1, CHK2, RAD17, RAD9, and NBS1 (Figure 1.8). Ataxia telangiectasia patients suffer from cerebellar ataxia, immune defects, telangiectases, radiosensitivity, and predisposition to multiple malignancies. These malignancies are usually hematogenous, but may include solid tumors such as breast cancer.

Ataxia telangiectasia heterozygotes are also believed to have a phenotype including radiosensitivity and breast cancer predisposition with an overall relative risk of 2.23 (95% confidence interval of 1.16 to 4.28), with the relative risk going up to 4.94 (95% confidence interval of 1.90 to 12.9) in those younger than 50 years old (168,169). It has been estimated that 8% of all breast cancer cases may be ataxia telangiectasia heterozygotes. The fact that even a haploinsufficiency of the ATM gene can result in increased cancer shows its importance in the formation of these cancers, and the importance of DNA repair in the etiology of these tumors. The radiosensitivity of heterozygotes has created concerns about the use of screening mammography in that exposure to the radiation required to perform the test may be damaging

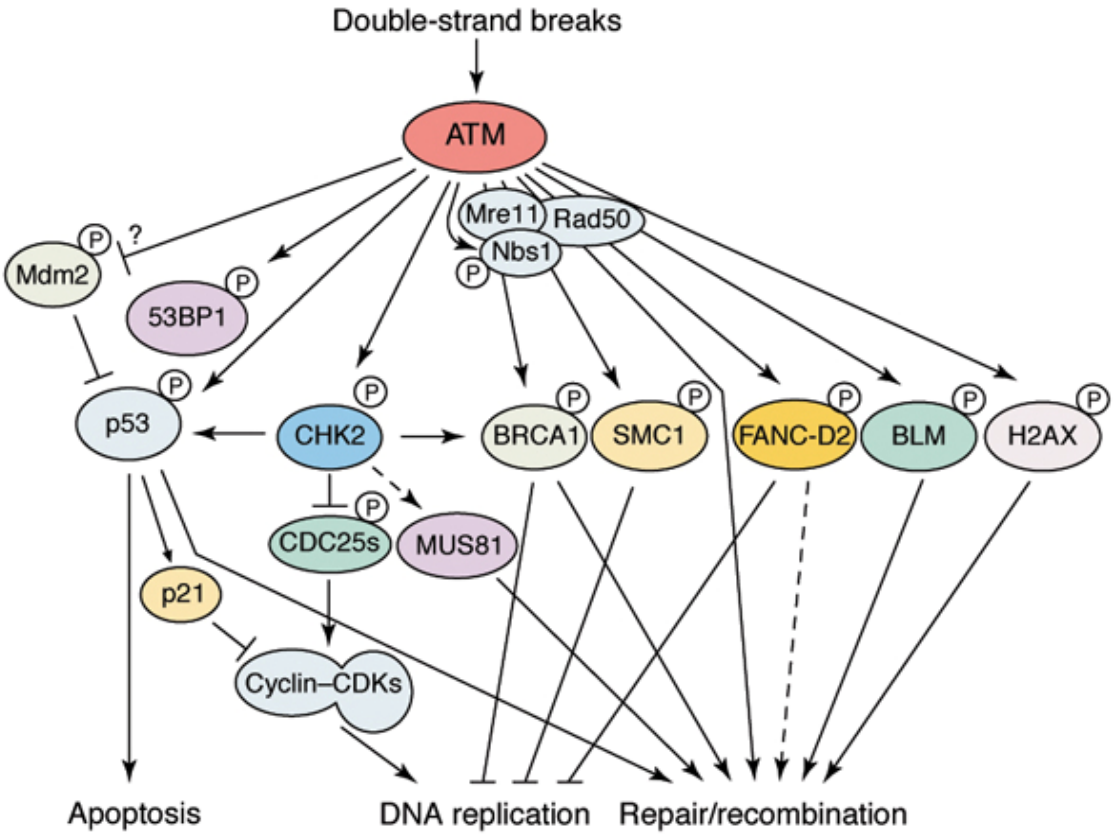


Figure 1.8 Review of the response of the cell to DNA damage. The ATM protein mediates a number of downstream effects, including the halting of DNA replication, initiation of DNA repair, apoptosis (170); reprinted with permission.

enough to these women's breast to cause cancer, but, in the absence of any other reasonably priced imaging tool, this has not changed our policies on screening mammography.

Peutz-Jeghers syndrome is an autosomal dominant disorder present in 1 in 8,300 to 1 in 29,000 live births (171). However, 40% of these cases are *de novo* mutations with no inheritance from a parent. The gene responsible for the disease state is STK11, also known as the LKB1 gene, a serine / threonine kinase tumor suppressor that functions in maintaining cell polarity (172). Patients' symptoms include diffuse hamartomatous polyps of the gastrointestinal tract, melanin spots of the lips and buccal mucosa, a general loss of pigment in adulthood, and multiple malignancies. Frequent sites of tumor include the gastrointestinal tract, breast, thyroid, lung, and uterus. The risk of breast cancer is between 9 and 18 times greater than normal.

Li-Fraumeni syndrome is a highly penetrant, autosomal dominant germline mutation in TP53, a tumor suppressor gene involved in cell cycle checkpoints, apoptosis, and responses to DNA damage (173). TP53 is one of the most common genes mutated in human cancers, and regulates over 100 known genes. It has been found to be mutated in 50% of breast tumors. After genotoxic insult, TP53 can act as a transcription factor, affecting the regulation of multiple downstream targets. One example of its activity is the p21-mediated G1 cell cycle arrest that occurs after DNA damage (Figures 1.4, 1.8). Li-Fraumeni syndrome includes early onset breast cancer (30% of tumors associated with the disease), soft tissue sarcomas, leukemia, brain tumors, and adrenocortical carcinomas. Patients present with multiple primary tumors during childhood. Li-Fraumeni syndrome accounts for less than 1% of breast cancers diagnosed in the US (174).

Li-Fraumeni syndrome patient's fibroblasts have been utilized to demonstrate that a link exists between TP53 expression and NER. Fibroblasts homozygous for loss of TP53 were deficient in their removal of cyclobutane pyrimidine dimers and 6-4 photoproducts from

genomic DNA (global genomic repair) (175,176). These finding were confirmed with bulky lesions generated from cells exposed to BPDE (177). Heterozygotes did not share this deficiency. In contrast, transcription coupled repair was not affected in these cells. Therefore, there is a known interaction between TP53 and the steps of the NER specific to the global genomic repair pathway.

TP53 affects genes of the NER pathway both through transcriptional regulation and through direct protein-protein interaction. The effect of TP53 on global genomic repair was first thought to be mediated solely by DDB2 (p48). DDB2 was shown to be upregulated in UV-irradiated cell in a TP53-dependent manner (178). Later studies conducted by Ford *et al.*, confirmed that there was a link between TP53 and XPC, another gene responsible for recognition of damage in the global genomic repair pathway. Li-Fraumeni syndrome fibroblasts have a decreased expression of the XPC protein that can be restored with regulated expression of TP53 (179). The XPC gene also contains a TP53 responsive element within its promoter. Just as TP53 affects expression of XPC, the protein itself can also interact with the DNA repair proteins XPB and XPD. Wang *et al.*, demonstrated a direct binding of TP53 to both XPB and XPD with subsequent inhibition of their helicase activity (180). In contrast to the positive effect that TP53 exerted on expression of XPC, the direct protein-protein interaction of TP53 to XPB and XPD seems to be negative. The role of TP53 in control of NER is therefore multifaceted and complex.

Table 1.20 Summary of inherited disorders that contribute to breast cancer risk.

Disease	Gene	Function
Cowden's Syndrome (166)	PTEN	Cell cycle arrest Apoptosis
Ataxia Telangiectasia (168)	ATM	DNA damage response Cell Cycle Regulation
Peutz-Jeghers Syndrome (172)	STK11	Serine/threonine kinase tumor suppressor polarity
Li-Fraumeni Syndrome (173)	TP53	DNA damage response Cell Cycle Regulation apoptosis
BRCA1 (181)	BRCA1	DNA damage response Cell cycle regulation
BRCA2 (181)	BRCA2	DNA damage response

1.5.11.1 BRCA1 and BRCA2

The two genes most commonly associated with inherited susceptibility to breast cancer are BRCA1 and BRCA2. Mutations that affect the activity of either of these genes result in an increased lifetime risk of being diagnosed with breast cancer for both men and women. Depending upon the location of the mutation (and its subsequent effect on activity), the risk can range between 50 and 80% lifetime risk of breast cancer (182). In addition to an increased risk in breast cancer, BRCA1 and 2 mutation carriers have an increased risk for ovarian cancer (women) and prostate cancer (men) (182,183). BRCA2 mutation carriers have an increased risk of pancreatic tumors, as well (183). The increase in risk for these others tumors is correlated with the site of mutation. For example, BRCA2 mutation in the center of the gene have a higher risk of ovarian cancer and a lower risk of prostate cancer than mutation in other parts of the gene.

The frequency of BRCA1 mutation carriers in the US varies depending upon which study is cited. Study variability can be due to factors such as the demographics of the population studied or the methods used to detect the mutation. They can range from as many as 1 in 345 to as few as 1 in 1000 (184). Frequency of mutations in BRCA2 in the US, a less frequently studied subject, is estimated at 1 in 1136 (183). The frequency of mutations in these two genes varies greatly with heritage. For example, women who are of Ashkenazi Jewish heritage have an increased chance of carrying a mutation in either BRCA1 or BRCA2 (184) (Table 1.21). African American women have a decreased chance. Researchers at the University of Chicago published data showing the frequency of mutations in either the BRCA1 or 2 genes based upon ethnicity (185). This study included families who came to the University of Chicago for genetic counseling due to a history of breast cancer (Table 1.21).

Table 1.21 Percent of 117 families with a history of breast cancer at the University of Chicago between 1992 and 2003 that had a deleterious mutation in BRCA1 or BRCA2, broken down by race.

Ethnicity	**% BRCA1 mutation**	**% BRCA2 mutation**
White	30.8%	15.4%
African American	16.3%	11.6%
Ashkenazi Jewish	41.4%	27.6%
Hispanic	0%	0%
Asian	0%	0%

The biologic characteristics of breast cancer arising in BRCA1 mutation carriers are slightly different from sporadic tumors (186). These patients are more likely to have an ER-, PR-, or Her-2/Neu- tumor with a higher pathologic grade. BRCA2 mutation carrier tumors are also more likely to have a higher tumor grade, but they are more likely to be ER+ and PR+. While ER, PR, and Her-2/Neu status are useful tools for prognosis in sporadic breast cancer, they are not as reliable in BRCA1 mutation carriers. For currently unknown reasons, the survival of BRCA1 mutation carriers with tumors is often better than these prognostic indicators would lead one to believe. Unfortunately, due to the relatively small number of cases of breast cancer in BRCA2 mutation carriers, not enough data is available to determine whether this lack of correlation of prognostic factors applies to them as well.

Both BRCA1 and BRCA2 are tumor suppressor genes that are inherited in an autosomal dominant pattern. Their tumor suppressor function is believed to come from their activity in DNA repair (181). BRCA1 plays essential roles in double-stranded break repair, specifically in homologous recombination repair and non-homologous end joining. After double-strand breaks occur, BRCA1 is phosphorylated by ATM, ATR, and CHK2 (Figure 1.8). BRCA1 interacts with MRE11, RAD50, and NSB1 in the initial steps of double strand break processing. It can also interact with other proteins to form a "BRCA1-associated genome-surveillance complex (BASC)" where it can influence the choice of repair pathway a cell uses to remediate a lesion (187). BRCA1 causes an increase the in expression of protein of the NER genes DDB2 and XPC in a TP53-dependent manner (discussed above). Reports have shown that $BRCA1^{-/-}$ cells are deficient at repairing oxidative lesions through transcription coupled repair, indicating that there is another way that BRCA1 interacts with the NER pathway beyond the expression of these two

genes (188). Finally, BRCA1 has also been shown to possess an E3 ubiquitin-ligase activity and can ubiquitinylate several targets including gamma tubulin and CtIP among others (189).

BRCA2 is thought to be involved in DNA damage repair pathways through interactions with RAD51 (181). RAD51 is a DNA recombinase involved in repair of double strand breaks and stalled replication forks. Stalled replication forks are a phenomenon that occurs when the replication machinery encounters a lesion with which it cannot deal. If the lesion is in the leading strand template, replication can either (1) become uncoupled, where the lagging strand continues to be synthesized for some distance, or (2) block synthesis of the lagging strand as well. If the lesion blocks both strands (such as an intercalating agent), the replication fork can also stall. After a stall of this nature, the replication machinery can begin to move backwards along its template strands, reannealing to a previous part of the template strands. The DNA that was synthesized from this portion of the template is extruded from the replication machinery and can anneal to itself, forming a short duplex branch of newly synthesized DNA. This structure resembles a Holliday junction, a structure that is naturally formed as an intermediate of homologous recombination. Thus, the remaining steps of this pathway can be used to remediate the block. BRCA2 can function in DNA repair either by binding directly to RAD51 to regulate its function or by binding directly to single stranded DNA. After damage takes place in a cell, BRCA2 has been shown to relocalize to the sites of damage. Its functions after binding there are unknown, however. BRCA2 has also been identified as the gene FANCD1 (114). This gene, identified by its activity in FA, is involved in the repair of stalled replication forks. BRCA2 interacts with FANCD2 and BRCA1 at the site of a stalled transcription fork. Unfortunately, the risk of breast cancer in carriers of germ line mutations in any of the FA genes is still unknown at

this time. Very few numbers of breast cancers have occurred in the cases of FA reported to the International Fanconi Anemia Registry (190).

Our laboratory has shown that haploinsufficiency for BRCA1 does not affect the NER repair capacity as measured by UDS (191). The fact that haploinsufficient tissue does not demonstrate a loss in DNA repair capacity in BRCA1 mutation carriers suggests that loss of NER capacity is not the initial step in carcinogenesis for these patients. However, we have shown an increase in somatic mutation and we are not measuring the predominant effect of the gene- its activity in double-stranded break repair. This phenotype of a lack of repair deficiency in heterozygotes is in contrast to the one known to exist in ATM patients as discussed above. It may highlight a difference in the tumor pathogenesis in BRCA1 mutation carriers as compared to other breast tumors, either sporadic or familial through a different mechanism.

1.6 REGULATION OF GENE EXPRESSION IN CANCER

Cancer in general is characterized by a dysregulation of normal cellular function. As cancer is believed to involve the dysregulation of multiple genes, multiple mechanisms may be involved. This can occur at the level of DNA, RNA, or protein. DNA may be altered in either sequence or structure. mRNA and proteins may be influenced by changes in expression or stability. Control of the transcription of DNA into mRNA is one of the single most important ways that cells determine which genes will be expressed during which periods of their life or in response to certain stimuli and thus represents a way in which the dysregulation of multiple genes can be accomplished in cancer. Transcription of genes can be regulated in a number of ways including

the binding of transcription factors to elements within the regulatory regions of a gene, and control of the accessibility to these control regions of DNA through epigenetic mechanisms. Each of these ways represents a source of dysregulation in cancer. Many oncogenes are transcription factors of elements in signal transduction pathways that regulate transcription factors. Thus, an alteration in one of these oncogenes results in dysregulation of many downstream genes. Similarly, epigenetic effects can either be felt through a change in an oncogene or by affecting a contiguous set of genes or the entire genome.

Elements within the DNA sequence itself that are responsible for controlling expression are called cis-control elements (192). The majority of these reside in the 5' regulatory region of the gene, while others are located in the first intron (noncoding region). The promoter region of genes includes both the 5' region up to and sometimes includes parts of the first intron. Promoters are the central processors of transcriptional control. Promoter elements including the TATA box and GC box are involved in the transcription initiation of many genes at a constitutive low level.

Other cis-acting elements such as enhancers and silencers can be used to confer more refined regulation of gene expression (*e.g.,* tissue specific expression or developmental expression) (192). While the majority of enhancers are involved in the response of the cell to specific stimuli like hormones, enhancers can also be involved in basal level transcription (Basal Level Elements). Basal level transcription is defined as the level of transcription that occurs in the cell without any specific stimulus. Enhancers need not be located in the promoter region and can appear either upstream or downstream of the promoter. Enhancers are usually between 10 and 30 nucleotides in length and can function in either orientation and in any location. Of these

10-30 nucleotides, only a small core is responsible for forming the binding site. Nonconserved sequences intervene between them.

Factors responsible for controlling transcription that are not a part of the DNA molecule itself, but rather bind to specific DNA sequences (*i.e.,* cis-acting elements), are called trans-control elements (192). Trans-control elements can either be present to impact upon the initiation of transcription or to allow it to continue through "roadblock" segments of the DNA sequence where transcription proceeds very slowly. A large number of proteins that function this way, called transcription factors, have been identified. Transcription factors may act as either stimulators of transcription or as repressors.

Sequence comparison has lead to the grouping of transcription factors by the protein-based motifs that they use to bind to DNA (192). These include: zinc-fingers, helix-turn-helices, helix-loop-helices, and leucine zippers. Zinc-fingers are one the most common motifs used by proteins to bind either DNA or RNA. A special class of zinc-fingers are the steroid receptors, so-called because they each bind a particular steroid hormone. As discussed above in reference to the estrogen receptor and progesterone receptor, many of these receptors dimerize once they are bound to ligand. In this form, they bind to their corresponding cis-control-elements to exert transcriptional regulation.

Transcription factors can regulate genes in a variety of ways. They can generate temporal sequences, whereby some genes only turn on after another one has already been transcribed. They can result in complex developmental pathways with some genes only being turned on during a particular phase of development. For example, some genes needed during human embryogenesis are no longer necessary in adults. They can also cause tissue specific expression of genes, allowing for the specific functions of different organs of the body.

Transcription factors can also themselves be regulated in many ways. They can have their own expression regulated by other transcription factors. Like the hormones receptors, they many bind to DNA as dimers. These can be either homo- or hetero-dimers. Different participants in these dimers can act in different ways, some as enhancers and some as repressors. In these cases it can be the ratio of the members of transcription factor complex to one another that ultimately determines their function as a transcription factor.

Over 1454 human genes have now been assigned transcription factor activity out of 38,500 human genes (about 4%) (193) and represent a significant class of genes that can be altered in cancer with wide reaching effects. Most of these have known binding sequences, but approximately 300 have no binding sites yet described for them. These can be subdivided into 150 different families. A family is a group of transcription factors with similar DNA patterns or binding sites with a similar biologic function (Table 1.22). The modeling of transcription factors takes advantage of the similarities that exist between members of the same family, as will be discussed below.

Table 1.22 Selection of known human transcription factors families and their functions. Members of the families were determined from the Genomatix information portal (194). Functions are listed as in Entrez Gene (195).

Transcription Factor Family	**Prevalence**	**Sample Transcription Factors**	**Examples of Functions**
HNF1 Hepatic nuclear factor 1	44.5% vertebrate promoters	HNF1	liver-specific expression of many genes during differentiation and development; frameshift mutations result in Familial Hepatic Adenoma
HOMF Homeodomain	54.0% vertebrate promoters	DLX1	craniofacial patterning and differentiation; survival of inhibitory neurons in the forebrain
		DLX3	Mutation results in richodentoosseous syndrome and amelogenesis imperfecta with taurodontism,
		EN1	pattern formation during development of the central nervous system
		PRRX1	co-activator, enhances serum response factor; growth and differentiation in response to growth factors; establishment of mesodermal muscle types,
		PRRX2	Involved in fetal skin development and wound healing; downregulated in adult skin; possible roles in prevention of scar formation during wound healing
		MSX2	balance between survival and apoptosis of neural crest-derived cells required for proper craniofacial morphogenesis; role in promoting cell growth under certain conditions; target for the RAS signaling pathways
		MEOX2	regulation of vertebrate limb myogenesis neurovascular dysfunction observed in Alzheimer's disease

Table 1.22 (continued)

Transcription Factor Family	Prevalence	Sample Transcription Factors	Examples of Functions
HOXF homeodomain consensus sequence	80.7% vertebrate promoters	BARX1	developing teeth and craniofacial mesenchyme of neural crest origin
		DMBX1	brain and sensory organ development
		HOX…	homeobox genes are found in clusters named A, B, C, and D on four separate chromosomes; Expression is spatially and temporally regulated during embryonic development
		PHOX2A	noradrenergic cell type; expression of tyrosine hydroxylase and dopamine β-hydroxylase; Mutations are associated with autosomal recessive congenital fibrosis of the extraocular muscles
		PITX1	organ development and left-right asymmetry; involved in basal and hormone-regulated activity of prolactin
PAX2 paired box gene 2	34.5% vertebrate promoters	PAX2	target of transcriptional suppression by the tumor suppressor gene WT1; mutations result in optic nerve colobomas and renal hypoplasia
RORA v-ERB and rar-related Orphan Receptor α	57.5% vertebrate promoters	NR4A1	steroid-thyroid hormone-retinoid receptor superfamily; expression is induced by phytohemagglutinin in human lymphocytes and by serum stimulation of arrested fibroblasts; translocation to mitochondria induces apoptosis
		RORA	NR1 subfamily (see below)
		THRA	One of the receptors for thyroid hormone

Table 1.22 (continued)

Transcription Factor Family	**Prevalence**	**Sample Transcription Factors**	**Examples of Functions**
RXRF RXR heterodimer binding sites	66.2% vertebrate promoters	NR1 subfamily	bind as monomers or homodimers to hormone response elements; specific functions unknown, but they control: NM23-2: a nucleoside diphosphate kinase involved in organogenesis and differentiation and NM23-1: the product of a tumor metastasis suppressor candidate gene
		RXRA	Retinoid X receptors; steroid and thyroid hormone receptor superfamily of transcriptional regulators; carry out the effects of retinoids; bind as homodimers or heterodimers
ERE Estrogen receptor	29.7% vertebrate promoters	ESR1 (ERα)	Development of breast epithelium and gastric mucosa; cell cycle progression and cell division
MYT1 Myelin transcription factor 1	68.9% vertebrate promoters	MYT1	Development of the nervous system; binds to the promoter regions of proteolipid proteins in the central nervous system
TBPF TATA binding protein factor	71.0% vertebrate promoters	TBPF	ubiquitous basal level transcription factor

1.6.1 Modeling Transcription Factors

Software programs exist to search DNA sequences for the presence of known transcription factor binding sites. Traditionally, sequence analysis has been done using simple matches to core binding sequences (cis-elements) determined for the transcription factors. This has become more prohibitive with the number of transcription factors know. Computer programs now exist that

not only account for different binding specificities of transcription factors, but can also generate complex models based upon the transcription factors' binding preferences and their spatial organization. These models are called frameworks. Programs that generate these frameworks include: PROMOTER SCAN, FunsiteP, GeneLang, GeneParser, GRAIL, and Frameworker (193). It is hoped that the identification of frameworks in promoters will yield functionally related genes through their coregulation.

1.6.2 Epigenetics

Epigenetic regulation controls the accessibility of a gene to transcription factors and transcriptional machinery. It can be thought of as control "above" the level of the DNA sequence: phenotypic changes without genotypic changes. Even if a transcription factor is present, it will not be able to stimulate or repress transcription if it is not able to bind to its corresponding DNA sequence. Two methods of accomplishing epigenetic regulation are histone modifications and methylation.

1.6.3 Histone Acetylation

Formation of heterochromatin is one way of controlling the expression of genes. DNA present in this conformation is not available to the transcriptional machinery and therefore genes are not expressed. The formation of heterochromatin is accomplished through altering the way that histone proteins interact with DNA. The reversible addition of acetyl groups (histone acetylation), methyl groups (histone methylation), and phosphate groups (histone

phosphorylation) work together or in competition with one another to form a highly selective binding platform for regulatory proteins (Figure 1.9).

When the N-terminal tails of histones H3 and H4 are deacetylated, repressor proteins such as ISWI ATPase binds to them and initiate heterochromatin formation (196). Histones can also be acetylated by the action of histone acetyltransferases. Acetylated histones are associated with actively transcribed regions of DNA. DNA in this open conformation is not necessarily transcriptionally active, but it at least has the potential to be transcribed.

Global transcriptional dysregulation secondary to histone deacetylation is now thought to play a role in the pathogenesis of neurodegenerative disorders like Huntington's disease, where precedent for the use of histone deacetylase inhibitors in treating human disease is now being generated. Evidence includes the presence of CREB-binding protein in aggregates found in the nucleus of affected neurons (197). CREB-binding protein has an acetyltransferase domain that can be inhibited by the polyglutamine peptides such as huntingtin. Polyglutamine can also inhibit the function of other acetyltransferases like p300 and p300/CBP associated factor. There is an overall reduction of the level of histone acetylation in affected neurons that can be restored with histone deacetylase inhibitors like sodium butyrate and suberoylanilide hydroxamic acid (197,198).

1.6.4 Histone Methylation

Histone methylation is slightly less well understood than histone acetylation, but is still known to control the association between histone proteins and DNA. The effect of histone methylation on chromatin configuration is specific to the location of the residue that is methylated. Methylation reactions primarily take place on the lysine residues of the N-terminal protein tail of the histone. Methylation of lysine 9 on H3 is associated with heterochromatin, while methylation of lysine 4 on H3 is found in active chromatin (199). Multiple histone methyltransferases have been identified, each with specificity for a particular lysine of a histone tail.

Methylation of histones may provide a docking point for *de novo* methylation enzymes to bind and restore DNA methylation (196). DNA methylases (discussed below) can bind to methylated histones. This binding may target their activity to the adjacent DNA sequence. It is known that when cells that have been exposed to DNA demethylases and then have that demethylase removed, their DNA becomes methylated again. This suggests that there is some "mark" that is not removed during demethylase treatment that is capable of redirecting DNA methylases when they become available again. As this is an *in vivo* phenomenon, and does not occur with DNA that is not in the presence of its histone proteins, the proposed "mark" is postulated to be present on the histone associated with the methylated DNA region and not the DNA itself. For its ability to regulate the docking of proteins through selective modifications like methylation, the histone tail has been proposed as this "mark."

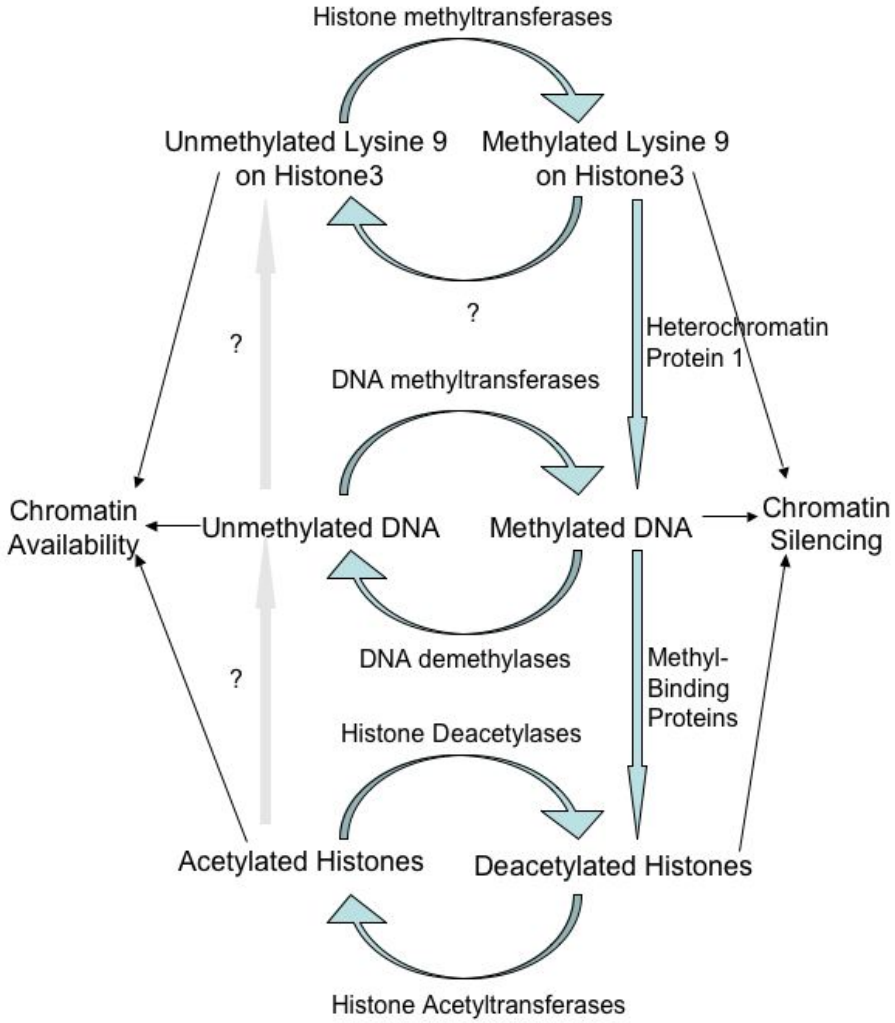


Figure 1.9 Adaptation of the schema by Chiurazzi and Neri regarding the interplay of histone modifications and DNA methylation(196). Reprinted with permission.

1.6.5 DNA Methylation as Regulation of Transcription

In mammalian genomes methylation of position 5 of the pyrimidine ring is used to suppress transcription. Methylation of CpG rich promoters is known to be used to prevent transcriptional initiation of genes on the X-chromosome and imprinted genes. The methyl groups of methylcytosine protrude into the major groove of DNA and both prevent the binding of transcription factors that would normally bind DNA in this region and specifically bind different methyl-binding proteins (200). These methyl binding proteins are associated with gene silencing and chromatin compactions, adding another level of control of gene expression.

Methyl groups are added to CpG sites by at least 3 methyltransferases: DNMT1, DNMT3a, and DNMT3b (196). DNMT1 seems to be more involved in maintenance methylation than DNMT3a and DNMT3b, which are the *de novo* methylation enzymes. The *de novo* methylation enzymes are triggered by an unknown mechanism to establish methylation of a particular region of DNA. Possible signals may involve double-stranded RNA and the small interfering RNAs that they produce. DNMT1 is responsible for ensuring that the methylation state of a gene is maintained after replication. When a CpG dinucleotide is methylated in one strand of DNA (hemimethylation), the complementary dinucleotide is also methylated. Once replication has taken place, these enzymes add methyl groups from a donor like S-adenosylmethionine onto the cytosine of the daughter strand across from other sites of methylation.

While much is known about the mechanism of DNA methylation, relatively little is known about DNA demethylation. It is assumed that because methylation is a reversible chemical process, and because enzymes have been found that carry out both methylation and

demethylation on substrates like histones, that a DNA demethylase enzyme exists. Demethylation can occur by both passive and active processes. Passive demethylation comes from the lack of function of maintenance methylases either through a loss of their expression or inhibition. Active demethylation requires specific enzymatic actions. In contrast to passive demethylation, active demethylation is a targeted phenomenon, used for example during embryogenesis to control larger numbers of genes in the developmental pathway. Passive demethylation has been demonstrated, but active methylation is inferred from the change in methylation pattern that is seen during development.

Three mechanisms have been proposed for demethylation, none of which has wide acceptance (201). (1) The first of these is the direct replacement of the methyl group by a hydrogen atom. One group claimed that Methylated DNA Binding Protein-2 was the enzyme responsible for this, but the claim was disputed when it could not be reproduced in other laboratories (202). (2) In the second mechanism, the bond between the 5-methylcytosine base can be cleaved by a DNA glycosylase and the resulting abasic site repaired with an unmethylated cytosine. Of the three mechanisms, the second has the most experimental evidence to support it. Two mismatch repair DNA glycosylases have been identified (one from chicken embryos and one from human kidney cells) that can act as 5'-methylcytosine-DNA glycosylases *in vitro* (203,204). Their activity towards these substrates is weak compared to their activity towards G/T mismatch substrates *in vitro*, however. (3) The third mechanism requires recognition of the methylated base by NER with subsequent excision of the methylated cytosine and gap filling synthesis replacing the excised fragment with nucleotides including unmethylated cytosine. No experimental evidence exists to support this theory to date.

1.6.6 CpG Islands

In mammals greater than 90% of methylation takes place on the cytosine residues of CpG dinucleotides. The distribution of CpG dinucleotides in mammalian genomes is nonrandom. Because methylated cytosine can be relatively easily deaminated to thymidine, a base normally seen within DNA and thus not recognized by the mismatch repair pathway, methylcytosine is considered mutagenic. Further evidence that CpGs are considered mutational hotspots is that they have been depleted through evolution. The frequency of CpG dinucleotides is approximately 20% of what would be expected with a random distribution of nucleotides in the genome (192).

Regions exist within the genome that are enriched for both C and G nucleotides and CpG dinucleotides. These regions, which can vary in length from 0.5 to 4 kilobases are known are CpG islands. While only 3-5% of cytosines within the human genome are methylated, 70-80% of these methylated cytosines exist within CpG islands (205). These islands are found in the promoters of approximately 40% of genes in the human genome. They are preferentially located at the transcription start sites of housekeeping genes, but can also be associated with tissue specific genes (206,207). Sequence analysis can determine the presence of these CpG islands, and therefore the feasibility of methylation being a factor in transcriptional regulation (208).

In 1987 the first criterion for the recognition of a CpG island were created by Gardiner-Garden and Frommer (209) (Table 1.23). They suggested that a CpG island have: (1) a G+C content > 0.5, (2) a ratio of CpG observed / CpG expected of > 0.60, and (3) would occur in a sequence window of 200 base pairs or greater. Such sequences are now known as Type I CpG islands. The definition was revised in 2002 by Takai and Jones after they completed an analysis

of chromosomes 21 and 22 (210) (Table 1.23). Their modified criterion increased the % GC to ≥ 55%, the observed CG / expected CG ratio to ≥ 0.65, and the minimum length to 500 nucleotides. This definition helped to remove Alu repeat sequences from appearing as CpG islands. Sequences meeting this modified criteria of Takai and Jones are now known as Type II CpG islands. The third and final definition, Type IIa, from Vinogradov combines both of the previous definitions by including the more stringent criterion for % GC and observed CG / expected CG ratio but keeping a minimum length of 200 nucleotides (Table 1.23) (211).

Table 1.23 Criterion required for a sequence to meet one of the three definitions of a CpG island.

	% GC	**Observed CG / Expected CG**	**Length**
Type I (209)	50 %	0.60	200
Type II (210)	55 %	0.65	500
Type IIa (211)	55 %	0.65	200

It has been assumed that the presence of CpG islands can be correlated with control of gene transcription by methylation. However, the correlation between CpGs and transcription might be more related to the overall GC content of a region rather than the presence of a CpG island (211). One line of evidence that suggests it is overall CG content rather than the presence of a CpG islands that correlates with transcriptional regulation is that it is overall CG content rather than CpG dinucleotide frequency that is increased near the transcription start site (212). Since this is the primary regulatory region of the gene, it may be that C content contributes to a

regulation effect at this location. GC content also correlates with the presence of chromatin condensation, which, as discussed above is related to control of gene expression (213).

In a recent paper looking at the relationship between CpG island presence (all 3 types) and GC richness with gene expression levels, patterns were found associated with both (211). It was the GC content and not the CpG island content that was positively associated with the maximum expression levels across all tissues tested. Compared with Type II and Type IIa islands, Type I CpG islands were associated with the breadth of expression across tissues. However, the effects of both of these factors are hypothesized to be mediated by methylation.

1.6.7 Methylation in Cancer

The role of methylation in cancer is the target of much investigation. Because there are so many changes in gene expression present in cancer it is reasonable to assume that there is a global mechanism of gene expression such as methylation at work. Two changes in methylation pattern are both known to exist concurrently in cancer, and specifically in breast cancer: global hypomethylation and regional hypermethylation (214).

1.6.7.1 Hypermethylation in Cancer

Aberrant hypermethylation of the promoters of tumor suppressor genes in cancer has been shown for almost every known tumor suppressor (215). Examples from the list of genes hypermethylated in cancer include BRCA1, hMLH1 (mismatch repair, associated with hereditary nonpolyposis colorectal cancer), and hMSH2 (mismatch repair, associated with hereditary nonpolyposis colorectal cancer) whose silencing may play a role in the transformation process

(216). A study has recently shown that the apparently normal tissue around the site of 10 different sporadic colon cancer tumors has abnormal hypermethylation of the CpG island of the MLH1 gene, involved in mismatch repair (217). Other genes that are hypermethylated in cancer range from steroid receptors to cell adhesion and cell surface molecules (218). Methylation of genes in breast cancer seems to be stage specific, with hypermethylated state either more or less likely for a particular gene depending on stage (219-221).

1.6.7.2 Hypomethylation in Cancer

Although tumor suppressor genes are often hypermethylated, global genomic hypomethylation also manifests in breast and other cancers (222-224). High-performance liquid chromatography analysis of global cytosine methylation revealed a 56% reduction in breast tumors as compared with normal breast tissue (225). The phenomenon has been identified by techniques such as high performance liquid chromatography where the general phenomenon has been measured. Studies now show that, like hypermethylation, there may be specific regions of hypomethylation in the genome (226). Using methylation-specific restriction enzymes to analyze the state of methylation in the genome of cancer cells, no correlation was seen to exist between hypomethylation and stage (223). Global hypomethylation is, however, related to several phenotypes associated with cancer, including the activation of oncogenes and increased genomic instability (227). Functionally, hypomethylation may cause the activation of genes associated with both tumor promotion and with the acquisition of metastatic capabilities. DNA methyltransferase activity is actually increased in some cancer cells, leading to the hypothesis that it is actually an increase in demethylation activity that causes the hypomethylation (228).

1.6.8 Methylation in Xeroderma Pigmentosum

The methylation status of noncancerous skin fibroblasts of XP patients is distinguishable from unaffected controls. The amount of 5'-methylcytosine in XP cells is only 70% of that seen in controls (229). No correlation has been found between the methylation level and the XP complementation groups of the cells. Cells that spontaneously reverted to a normal repair capacity recovered a level of methylation close to that of normal cells, showing that this phenotype was reversible and associated with DNA repair capacity. However, no mechanism has been proposed for aberrant methylation in XP fibroblasts. These results are discordant with the idea that NER is the mechanism responsible for demethylation since a lack of NER function in these XP fibroblasts actually results in hypomethylation. It is interesting, however, that normal XP cells share the phenotype of hypomethylation with cancer cells.

1.6.9 Reactivation of Genes Lost Through Epigenetic Regulation

The possible reactivation of expressed genes lost in carcinogenesis is a topic of research in cancer chemotherapy and treatment of other genetic conditions where increased expression of a gene with lost or reduced expression may be beneficial. This includes genes that have been lost due to methylation of the promoter regions or genes that are located in regions of heterochromatin. Chemicals that have been used in the laboratory to create such effects include demethylating agents and histone deacetylase inhibitors (Table 1.24). These agents, however, have genome-wide effects and require much evaluation to determine their therapeutic potential. Hope comes from experiments like that of Suzuki *et al.* who showed that fewer than 1% of

10,000 evaluated genes changed their regulation after treatment with the demethylating agent 5-azadeoxycytidine opening up at least the possibility that there might not be wide-spread side effects from the global reactivation of genes (230).

Table 1.24 Sample of drugs employed in the laboratory reactivation of human genes.

Drug Class	**Example Drug**
DNA demethylating drugs (DNA methyltransferase inhibitors)	5-azacytidine
	5-azadeoxycytidine
	Zebularine
	Hydralazine
	Procainamide
Histone hyperacetylating drugs (HDAC inhibitors)	Sodium butyrate
	4-phenylbutyrate
	Trichostatin A
	Suberoylanilide hydroxamic acid
	Valproic Acid

1.6.10 Histone Deacetylase Inhibitors

Histone Deacetylase inhibitors are under investigations for use both as anticancer chemotherapies and in the treatment of Huntington's disease. Histone deacetylase inhibitors block cell cycle progression and induce apoptosis. It is unknown if both these effects are mediated through its inhibition of histone acetylation or if there are other targets. Microarray analysis has shown that there is a limited change in gene expression after exposure to a histone deacetylase inhibitor (231).

Trichostatin A is a naturally occurring fungal antibiotic that was one of the first agents identified as a antiproliferative agent (232). Trichostatin A docks at the active site of histone deacetylase, chelating the zinc ion essential for activity. Another histone deacetylase inhibitor with similar action is suberoylanilide hydroxamic acid, a synthetic compound which was originally selected for its ability to induce differentiation in murine leukemic cells (233). It was later determined that this effect was mediated through inhibition of histone deacetylases. Suberoylanilide hydroxamic acid has been shown to inhibit prostate cancer cell growth both *in vivo* and *in vitro* (234).

Valproic acid and sodium butyrate are two more examples of histone deacetylase inhibitors (235). Valproic acid is a small molecular weight carboxylate that causes hyperacetylation of histone tails. It acts by binding to and blocking the catalytic center of histone deacetylases. It is an attractive drug for study because it is already used clinically for the treatment of epileptic disorders. Valproic acid was shown to delay growth of primary breast tumors and lung metastases in mice (236). Sodium butyrate is another inhibitor that works through the same mechanism. It also functions as a differentiating agent and growth inhibitor for many tumor cell lines, including some derived from breast cancer (237).

1.6.11 DNA Demethylating Agents

Because of the known changes that exist in methylation of cancer cells, demethylating agents are being investigated *in vivo* and *in vitro* for their antineoplastic effects. Since hypermethylation of many tumor suppressor genes is seen in multiple cancer types, a variety of demethylating agents are now being sought as a new form of cancer drugs. Enthusiasm has been somewhat stinted when no growth inhibitory properties were seen for these drugs at DNA demethylating doses (238). The exception for this is 5-deoxyazacytidine which has antiproliferative effects at such doses. This has lead to the investigation of demethylating agents not as a single drug therapy, but to potentiate the actions of other therapies. Two types of demethylating agents are available: the nucleoside analogues and the non-nucleoside analogues.

1.6.11.1 Nucleoside Analoges

The nucleoside analogues 5-azacytidine and 5-azadeoxycytidine are potent demethylating agents that have been used to investigate methylation experimentally for years. These analogues are incorporated into DNA during replication in the place of cytosine. When they are incorporated into a site that is targeted for methylation by DNMT1, they covalently bind the methyltransferase and sequester it in this location (239). Since methylation is a possessive procedure, no methylation is carried out downstream of the site of incorporation after the enzyme becomes bound. DNMT1 levels in the cell fall during exposure to either one of these agents (239). 5-azadeoxycytidine was able to restore expression of the estrogen receptor in MDA MB231 breast cancer cells, in which the estrogen receptor was hypermethylated (240). Trichostatin A was also

able to accomplish this reactivation alone, but the two agents worked synergistically when added together (241,242).

In contrast to the traditional nucleoside analogues, zebularine is an analogue that is very stable at physiologic pH and less cytotoxic at effective doses (243). When tested in human bladder cancer cell lines, zebularine, like the nucleoside analogues, was able to cause the expression of the silenced p16 gene, although it was not as potent as 5-azacytidine or 5-azadeoxycytidine. Though 10- to 100- fold less drug was needed to achieve the same effects as the nucleoside analogues, zebularine is much less toxic *in vitro*. The difference in potency between these inhibitors may stem from the fact that critical levels of zebularine triphosphate metabolites must be achieved before they are incorporated into DNA.

1.6.11.2 Non-nucleoside Demethylating Agents

Compared to the nucleoside analogue 5-azadeoxycytidine, the non-nucleoside analogues are far less potent DNA demethylases (244). Two drugs, hydralazine and procainamide, are currently in use clinically to treat different disorders. Hydralazine is a vasodilator that is used as an antihypertensive drug. It was found to decrease the expression of DNMT1 and DNMT3A by Deng *et al.* in 2003 (245). Procainamide is used clinically as an antiarrhythmic that was also shown to inhibit DNA methyltransferase activity (246). Though different cell lines demonstrated different sensitivities to these agents, the lack of potency when compared to 5-azadeoxycytidine was universal. Furthermore, hydralazine has no growth inhibitory effects on cervical, colon, breast, sarcoma, glioma, or head and neck cancer cell lines when used alone (238). These results dampen enthusiasm for use of either of these drugs as single-agent chemotherapies, but still allow the possibility that they could be combined with other treatments for better effect.

1.7 REGULATION OF NUCLEOTIDE EXCISION REPAIR GENE EXPRESSION

1.7.1 Tissue Specificity of Nucleotide Excision Repair

Historically it has be assumed that DNA repair was such an essential component of cellular function that it was expressed constitutively across all tissues, and therefore in all cell types, to allow for the maintenance of genomic integrity. However, the possibility that NER can be regulated was first demonstrated in Dr. Latimer's laboratory by finding that functional NER capacity is present in different levels in various tissues and cell types in mammalian development. Specifically, a study of the repair capacity of the four distinct extraembryonic lineages that comprise the extraembryonic yolk sac, as well as five cell types derived from the fetus itself showed that NER in the mouse is lineage-specific during embryogenesis (247). Our laboratory has now extended this analysis to human somatic tissues. In a comparison of repair capacity carried out using the functional UDS assay, it was shown that neonatal FF, adult breast reduction tissue, and adult peripheral blood lymphocytes (PBLs) all possessed a unique level of DNA repair (140). Further work to be presented in Chapter 2 now shows that normal ovarian epithelium demonstrates a repair capacity similar to that of breast reduction epithelium. These findings highlight the possibility that DNA repair function, and thus genes of the NER pathway, may be subject to regulation.

Further evidence for the tissue specificity of mammalian DNA repair capacities comes from the contrast of phenotypes demonstrated by mouse knockout models of NER genes compared to the human disease phenotypes. While the nature of the disease is often different between species, certain tissues are consistently affected. For example, the eyes of mice are

particularly sensitive to UV-induced retinal tumors, and the $Csa^{-/-}$ and $Csb^{-/-}$ mice display spontaneous loss of photoreceptors. The developmental abnormalities that are present in XP, TTD, and particularly in CS as well as in multiple mice, particularly the Ercc1, Rad23B, and Xpf knockout models, suggest that the central nervous system also contains sensitive tissues. The immune deficiencies experienced by multiple mice also suggest that the lymphoreticular system contains sensitive tissues. These tissues may either be more dependent on repair capacity (*e.g.,* skin, which is under constant UV exposure) or already expressing a low repair capacity constitutively, such that a complete loss is more catastrophic.

1.7.2 Steady State mRNA Expression Studies

Some mRNA expression studies have been completed on genes in the NER pathway. In an attempt to characterize some of the interindividual variation or "range of normal" present in the expression of human repair genes Crallan *et al.* showed that there was little variation in expression in 26 repair genes between 6 normal human urothelial cell cultures (248). The expression levels were analyzed using RNase protection and included 10 genes of the NER pathway as well as base excision repair and mismatch repair. Only the gene XPG, an endonuclease involved in the excision step of NER, showed variation between the cell cultures tested. Expression of all genes tested was detectable. Expression levels of these genes did not correlate with the age of the patients.

Human ovarian cancer cells in culture showed that they can respond to cisplatinum treatment by increasing their mRNA expression of ERCC1 2 to 6 fold over steady state (249). However, no relationship has been reported between ERCC1 mRNA expression and clinical

response to cisplatin chemotherapy in the ovary (250). As NER is the pathway responsible for removing the crosslinks formed in cisplatinum-based chemotherapy, increased ERCC1 mRNA levels been shown to correlate positively with resistance to this therapy in the lung (251).

Expression of ERCC1 and XPB, XPD, and CSB mRNA has been evaluated by slot-blot analysis in both human bone marrow and brain tissues. In bone marrow being harvested for autologous transplantation, over 200-fold inter-individual variation was seen with ERCC1, XPB, and XPD (252). When brain tissue was evaluated, it was seen that XPB mRNA expression was decreased in malignant brain tissues relative to nonmalignant tissue ($p = 0.003$), while CSB (ERCC6) expression remained the same ($p = 0.314$) (253). These studies shows precedence for the alteration of NER gene regulation in cancer. The mechanism of this regulation is unknown.

A multiplex reverse transcriptase-PCR reaction was developed by Wei *et al.* to analyze the mRNA levels of five NER genes: ERCC1, XPB, XPG, CSB, and XPC. This technique was then used to assay expression levels in both a survey of normal tissues, and in the PBLs of lung cancer patients in comparison with matched controls (254-256). Twenty different tissues were assayed in their survey: skin, breast, intestine, liver, testis, ovary, placenta, prostate, adipose tissue, brain, hippocampus, muscle, spleen, lung, stomach, bladder, kidney, colon, phytohemaglutinin-stimulated PBLs, and unstimulated lymphocytes. As the results were generated from only one sample of each type, statistical analyses between them were not presented in this paper. The authors did note, however, a general trend that proliferating tissues (defined as skin, breast, intestine, liver, testis, ovary, placenta, and prostate) had a relatively higher expression of the NER genes tested than in most of the nonproliferating or slowly proliferating tissues (adipose tissue, brain, hippocampus, muscle, spleen, and lung), a finding that is further discussed in Chapter 3 (255). A statistically significant decrease in both XPG and CSB

mRNA expression was seen using this technique in a cohort of lung cancer patients as compared to a control group matched for age, race, sex, and smoking history (256).

1.7.3 Mechanisms of Nucleotide Excision Repair Gene Regulation

Overall, very little is known about the transcriptional regulation of NER genes. What information is available shows that control of expression of genes in the NER pathway is executed by multiple mechanisms. Both transcription factors and epigenetic mechanisms have been implicated with individual NER genes, either in normal cellular activity or in cancer. TP53 can act as a transcription factor for both XPC and DDB2, as discussed above with regard to Li-Fraumeni syndrome in breast cancer. Both of these genes contains a TP53 responsive element in its promoter. In response to DNA damage, TP53 increases the expression of both of these genes.

The mRNA and protein expression of the NER gene XPD has been shown to be regulated by insulin and glucose (257): there is an insulin-dependent increase in XPD mRNA and protein levels. Exposure to glucose acutely can potentiate this action of insulin, but can reduce the response if exposure to glucose is chronic (>3 weeks). The exact mechanism of this regulation is unknown, but has been shown to be dependent on the RAS-signaling pathway, as determined by the ablation of this effect when a RAS inhibitor, maumycin-A, was added. This pathway can activate multiple transcription factors that may then cause the upregulation of XPD mRNA.

Genes involved in NER are also tumor suppressor genes, and therefore may be hypermethylated in cancer. The DDB1, hHR23B, and RPAp32 genes have been shown to be hypermethylated in multiple myeloma cells (258). The interleukin-6-responsive multiple myeloma cell line KAS-6/1 was evaluated by RNase protection for expression of 10 genes in the

NER pathway after exposure to the demethylating agent zebularine. DDB1, hHR23B, and RPAp32 were all increased in expression after treatment. When the chemical was removed from the culture medium and the cells were supplemented with IL-6, the loss of hHR23B expression through methylation was re-established. This work shows precedence for the regulation of NER genes through methylation in cancer.

1.8 OUR HYPOTHESIS AND A GENERAL INTRODUCTION TO THIS WORK

Lack of DNA repair has been shown to lead to cancer in multiple human disease states. We therefore hypothesized that lack of **NER is involved in the etiology of breast cancer**. Using the UDS assay, Latimer *et al.* has shown that 100% of stage I breast tumors are deficient in NER capacity as compared to nondiseased controls. We now show that this is a phenotype shared by most stage II tumors (Chapter 3). In contrast to the loss of DNA repair experienced in early stage tumors, late stage tumors show a no loss of repair or even in some cases an increase in repair capacity over the level of nondiseased breast epithelium (Chapter 3).

Analysis of the mRNA expression levels of 20 genes in the NER pathway using a commercially available RNase protection kit showed that 7 genes (CSB, XPA, XPB, TFIIHp52, TFIIHp44, TFIIHp34, and CDK7) shared the trend seen with functional DNA repair capacity: loss in early stage tumors and gain in late stage tumors (Chapter 5). This trend was seen in multiple normal and multiple tumor cell lines. Because the loss of expression was reversible, occurred in multiple genes in multiple cell lines derived from multiple patients, we hypothesized that the mechanism responsible was epigenetic regulation. Sequence analysis showed that 90%

of the genes analyzed had a CpG island present in their 5' regulatory region, making methylation a plausible candidate for the regulatory mechanism. Two early stage cell lines that had each demonstrated loss of expression of these candidate genes were treated with the demethylating agent 5-aza-2'-deoxycytidine, resulting in the reexpression of 3 of these genes: TFIIHp52, TFIIHp34, and CDK7. Transcription factor analysis was also carried out on the DNA sequences of these candidate genes. Two of these genes that were not reactivated by 5-azacytidine treatment (CSB and XPA) showed a 7-element framework that may begin to explain their coordinate regulation.

In order to accomplish any of this work, however, an *in vitro* system for both normal and cancerous breast tissue had to be established. The Latimer lab has developed a novel primary cell culture system for normal and diseased breast that is shown in Chapter 2. Furthermore, larger numbers of cells are required for molecular biologic analysis. Commercially available cell lines are not available for early stage breast cancer, and cell lines that do exist show either no loss or an increase in DNA repair capacity relative to breast reduction epithelium (Chapter 4). Cell lines were generated from the Latimer primary culture system without the use of exogenous agents and these lines are also discussed in Chapters 4 and 5. These primary cultures and novel cell lines are the framework for all the functional and molecular analysis reported here.

2.0 A NOVEL PARADIGM FOR MAMMARY TISSUE ENGINEERING USING EMBRYONIC STEM CELL BASED MEDIUM

2.1 INTRODUCTION

The breast is a tissue that undergoes dramatic changes in the adult state. Proliferation of epithelial cells followed by involution and apoptosis occurs during every menstrual cycle in which a pregnancy is not established. In addition, vast tissue remodeling occurs during pregnancy, lactation and weaning. Unlike the murine system, these changes have not been well documented in the human breast, with the exception of some important descriptive histological studies (259-261).

Although rodent systems have been useful in investigations of mammary epithelial development, differentiation and malignancy, they may not be representative models of the human breast (137,262). For example, human mammary tissue undergoes virginal development of terminal ductal lobular units (263,264), whereas the rodent terminal ductal lobular units form only with copulation and pregnancy. This difference, combined with the fact that human breast epithelium undergoes involution and tissue remodeling for decades longer than rodents, leads to the inescapable conclusion that studies of human breast development should ideally be performed in a human model system.

Another difference between rodents and humans involves global genomic nucleotide excision repair (117). As discussed in Chapter 1, rodent cells are deficient in remediation of cyclobutane pyrimidine dimers relative to human cells (59,118), yet their survival is not compromised. Studies by Bohr and Mellon demonstrated that Chinese hamster ovary cells were proficient in transcription-coupled repair but severely deficient in global genomic repair (37,119). This paradoxical survival of rodent cells secondary to differences in the use of transcription coupled repair versus global genomic repair makes primary rodent culture systems an inappropriate choice for the study of DNA repair in the human breast.

2.1.1 Human Mammary Epithelial Cell Cultures

Mammary epithelial cells are both challenging to establish in culture and difficult to maintain. Early culture methods included overnight enzymatic disaggregation of normal breast reduction mammoplasty tissue, followed by monolayer culture in specialized media (265-271). These cultures have the advantage of containing pure epithelial cells, however, they survive only a matter of 2-3 weeks or 3-5 passages. Others have isolated purified epithelial cells via Epithelial membrane Antigen-directed cell sorting followed by culture as aggregates or organoids (272). More recently, a method has been developed that involves establishment of floating aggregates of epithelial cells, or "non-adherent mammospheres" (273).

2.1.2 Differentiation of Mammary Cells *In Vitro*

As our knowledge of the paracrine and juxtacrine factors required for epithelial growth and differentiation has deepened, new methods have emerged for the culture of human mammary epithelial cells. These systems demonstrate the role of cell-to-cell signaling from the stromal compartment of the breast and have utilized either a complex substratum like matrigel or an exogenous feeder layer for creating a stromal/paracrine context for isolated epithelial cells in culture (274-277).

Successful culture of isolated breast tumor cells on NIH 3T3 feeder layers (278,279) suggested that juxtacrine and/or paracrine factors also contribute to the establishment of mammary tumor cell lines. Feeder layers have been used extensively for embryonic stem cell culture in order to maintain these totipotent cells in an undifferentiated state. It has been known for some time that the use of feeder layers, together with careful culture in high quality medium, acts to prolong the embryonic phenotype of the cell line specifically in terms of maintaining a high differentiation ability and euploid chromosome content (280).

Our laboratory has developed a method involving physical disaggregation of reduction mammoplasty tissue for establishing primary cultures of nondiseased breast epithelium that initially contain cell types from the stromal, the epithelial, the myoepithelial, and the adipose compartments of the breast. This allows for the free exchange of cell type specific paracrine and juxtacrine factors *in vitro*. These cultures initially manifest simple rounded and attached mammospheres (140,273,281) that later develop into complex branching ductal structures (Figure 2.1).

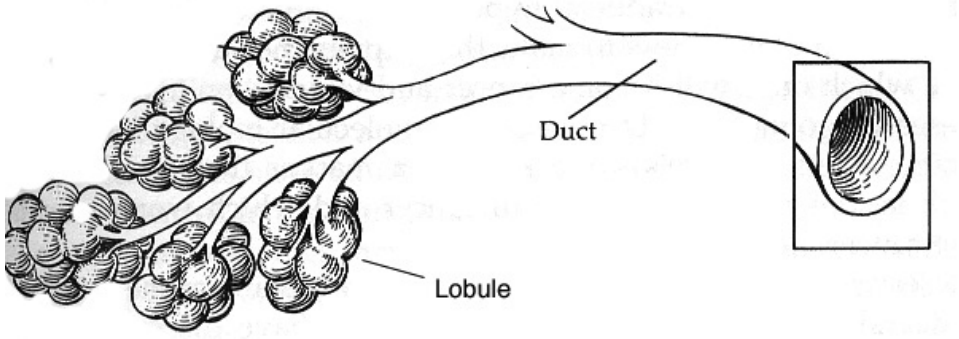


Figure 2.1 Schematic of the ductal architecture of the breast from Dr. Susan Love's Breast BookBook (282), reprinted with permission. The breast is composed of a system of progressively branching system of ducts that end in the terminal ductal lobular units. The terminal ductal lobular units are the milk-producing units of the breast, and the ducts are responsible for its conduction to the nipple where it is expelled. Both the ducts and lobules are lined with epithelial cells and surrounded by a modified form of epithelial cell called the myoepithelial cell.

Our culture medium is adapted from the classic culture medium of the embryonic stem cell literature for which no definite minimal quantities are given, or detailed analyses have been performed; rather, the classic medium contains copious amounts of two types of serum, with the implication that the factors needed to maintain these cells in a proliferative and undifferentiated state are simply not known. We approached the development of our medium with the following rationale: many of the factors necessary to maintain long lived mammary cultures are unknown, and the additives originally designed to maintain embryonic stem cells in a proliferative and

undifferentiated state may also support the continued growth of explanted mammary tissue *in vitro*, especially if the establishment and proliferation of such cultures is dependent on a subpopulation of adult stem cells. The additives in our medium provide nucleosides, amino acids, *etc*. to supplement what might be diminishing pools of such reagents at the critical transition phase between the *in vivo* vascularized state and the *in vitro* environment.

2.2 MATERIALS AND METHODS

2.2.1 Tissue Processing

Nondiseased human breast tissue was obtained from breast reduction mammoplasty surgeries performed at Magee-Womens Hospital under IRB94-108/0504117. A mirror image tissue block was created from half of the fresh tissue block actually processed for culture. Several sections of this "research block" were examined after initial establishment of the culture by a pathologist for every sample cultured.

Fresh breast tissue was stored in refrigerated DMEM (Gibco), 10% fetal calf serum (Hyclone) and 3x antibiotic/antimycotic (Gibco) for up to 12 hours before processing.

Breast tissue was washed with PBS containing 3x antimycotic/antibiotic and minced into very small fragments in MWRIα medium (283). MWRIα medium consists of DMEM (Gibco), 10% Fetal Calf Serum (Hyclone), 10% Newborn Calf Serum (Hyclone), 2.5% Rat Serum (Atlanta Biologicals), 5% nucleosides (Sigma), 5% non-essential amino acids (Sigma), and 5%

of a 0.07% stock of β-mercaptoethanol. Tissue pieces were plated on a thin coat of rBM (Matrigel, Becton Dickinson) in MWRIα and fed daily for a period of 4 days.

2.2.2 Digital Time Lapse Photography

Primary cultures were established on Bioptechs 25mm round rBM coated coverslips precoated with a thin coat of matrigel. After a period of 5-7 days movies were captured using the FCS2 Bioptechs environmentally controlled chamber. Movies were captured using a high resolution Hamamatsu Orca C4742-95 cooled CCD digital camera designed for low light and monochromatic images with a resolution of 1280 x 1024 pixels (12 bits per pixel) and capturing an image every 10 minutes for a total of 8 hours. Individual Tiff images were collected using QED imaging software on a Macintosh G4 computer.

2.2.3 Immunostaining and Confocal Microscopy

Tissue was fixed for 1 hour in 2% paraformaldehyde, rinsed 3x in PBS, then rinsed 3x in PBS containing 0.5% BSA and 0.15% glycine (PBG Buffer). Samples were permeabilized in 0.1% Triton-X-100 in PBG for 30 minutes, and blocked in 20% non-immune goat serum in PBG for 30 min at room temperature. Primary antibodies were diluted in PBG at 1:100. Primary antibodies were:

Cytokeratin 19 (MAB, Clone RCK108) BioGenex

Cytokeratin 14 (MAB Clone LL002) RDI

Cytokeratin 18 (MAB Clone DC10) DakoCytomation

β-Casein: (MAB Clone F20) Abcam

Epithelial Membrane Antigen (MAB Clone E29) DakoCytomation

Rhodamine Phalloidin InVitrogen 1:250 dilution

Primary antibodies or nonimmune isotype control were added to cultures for 1 hour at room temperature. Cultures were washed 5x in PBG, then fluorescently tagged secondary antibodies (Goat anti-mouse Alexa488 (green) 1:500, Molecular Probes, or Goat anti-mouse Cy3 (red) 1:100, Jackson Immunolabs) were added for 1 hr at room temperature. Tissue was washed 3x in PBG, 3x in PBS, then nuclei were stained using 0.01% Hoechst dye (bis Benzimide) in ddH_2O for 30 seconds DRAQ 5 (1:1000 dilution; Biostatus). Following a wash in PBS, tissue was coverslipped using gelvatol (23 g poly(vinyl alcohol) 2000, 50 ml glycerol, and 0.1% sodium azide in 100 ml PBS) and viewed on scanning laser microscopes (Fluoview 1000, Olympus).

2.2.4 Analysis of Nuclear Polarization

Epifluorescence microscopy was performed followed by image deconvolution on chemically stained mammospheres. TMA-DPH was used to stain membranes blue, Syto-16 was used to stain nuclei green and Mitotracker Red was used to stain mitochondria (all from Molecular Probes).

2.2.5 Scanning Electron Microscopy

Primary human mammary epithelial cell cultures were fixed with 2.5% glutaraldehyde in PBS for several hours, washed 3x in PBS then post-fixed in aqueous 1% OsO_4, for 1 hour. Following 3 PBS washes, the cultures were dehydrated through a graded series of 30-100% ethanol, further dehydrated by three additional 15 min washes with absolute ethanol then critical point dried (Emscope CPD 750). Sample coverslips were removed from dishes, mounted onto aluminum stubs then sputter coated with gold/platinum (Cressington 108 Auto). Samples were viewed on a JEOL JSM 6330F scanning electron microscope.

2.2.6 Transmission Electron Microscopy

Cultures were fixed with 2.5% glutaraldehyde in PBS for several hours, washed 3x in PBS then post-fixed in aqueous 1% OsO_4, 1% $K_3Fe(CN)_6$ for 1 hour. Following 3 PBS washes, the cultures were dehydrated through a graded series of 30-100% ethanol, then infiltrated with several changes with Polybed 812 epoxy resin (Polysciences) over 24 hrs. Cultures were embedded by inverting Polybed 812-filled BEEM capsules on top of the cells. Blocks were cured overnight at 37^{o}C, then for two days at 65^{o}C. Monolayers were pulled off the coverslips/dishes and re-embedded for cross section. Ultra thin cross sections (60 nm) were obtained on a Riechart Ultracut E microtome, and post-stained in 4% uranyl acetate for 10 min and 1% lead citrate for 7 min. onto electron microscope film (Kodak, ESTAR thick base). Electron micrographs were digitized on a flatbed scanner at 400 ppi (StudioStar, Agfa).

2.3 RESULTS

2.3.1 Initiation of Culture

We have successfully established primary human mammary epithelial cultures from 38 breast reduction mammoplasty tissue samples. These include tissue from 28 pre-menopausal, 4 peri-menopausal, and 5 post-menopausal patients. Pathological analysis confirmed that all the tissues cultured and utilized in this study were within the range of normal histologies that include hyperplasia and fibrocystic disease. None contained known atypia.

An overview of our tissue engineering system is shown in Figure 2.2. Minced tissue pieces attach to the substratum within 24 - 48 hours, typically rendering the medium acidic and yellow in color. Within 3 days, living attached cells appear as monolayer outgrowths of multiple cell types from freshly severed tissue edges. These attached cells include both epithelial and fibroblastic cells. After 10-11 days in culture a marked increase in both cell populations occurs, due to both cell migration and proliferation, with cells becoming visible microscopically in the areas of the dish not obscured by tissue pieces.

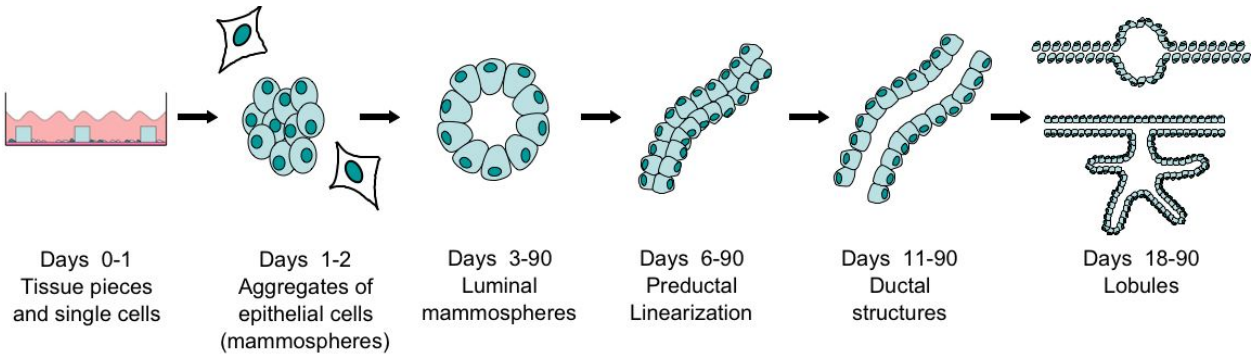


Figure 2.2 Summary schematic of the temporal program manifested by the human mammary tissue engineering system developed in the Latimer laboratory. A thin coat of diluted Matrigel and MWRI medium are utilized in order to achieve long-term organotypic differentiation of ductal structures.

2.3.2 Mammosphere Formation

The term "mammosphere" has previously been used to describe 3-D conformations of mammary epithelial cells in murine, bovine, fur seal, and human systems in which acinar structures have been formed from aggregates of epithelial cells (273,284-287). Depending on the species and the technique used, mammospheres can be formed as attached structures without the presence of a complex substratum (in the case of cape fur seals), embedded in matrigel, or free-floating in the cell culture medium. Mammospheres are defined in this system as 3-dimensional clusters of 20-100 epithelial cells, with only the lowest layer of cells in contact with the matrigel substratum.

The formation of the mammosphere *in vitro* is the first recognizable three-dimensional epithelial architecture formed in our culture system (Figures 2.2, 2.3, and 2.4). Mammospheres have been visible as early as 24 hours after initial establishment of primary cultures. Mammosphere formation is an active process of cellular aggregation and recruitment rather than an artifact of the disaggregation method, *i.e.,* small tissue pieces maintaining their cohesion. *De novo* mammosphere formation was captured in a digital time-lapse movie. This movie shows how a discontinuous monolayer of epithelial cells, initially in extended cytoplasmic contact with each other, aggregate three-dimensionally over the course of eight hours to form a typical mammosphere (Figure 2.3).

Scanning electron microscopy shows the intimate contact between the rounded epithelial cells in mammospheres, often collectively taking on the appearance of having a flattened, coherent proximal cell surfaces (Figure 2.4a,b). Transmission electron microscopy of the lateral surfaces of these cells demonstrates multiple desmosomes linking the surfaces (Figure 2.5).

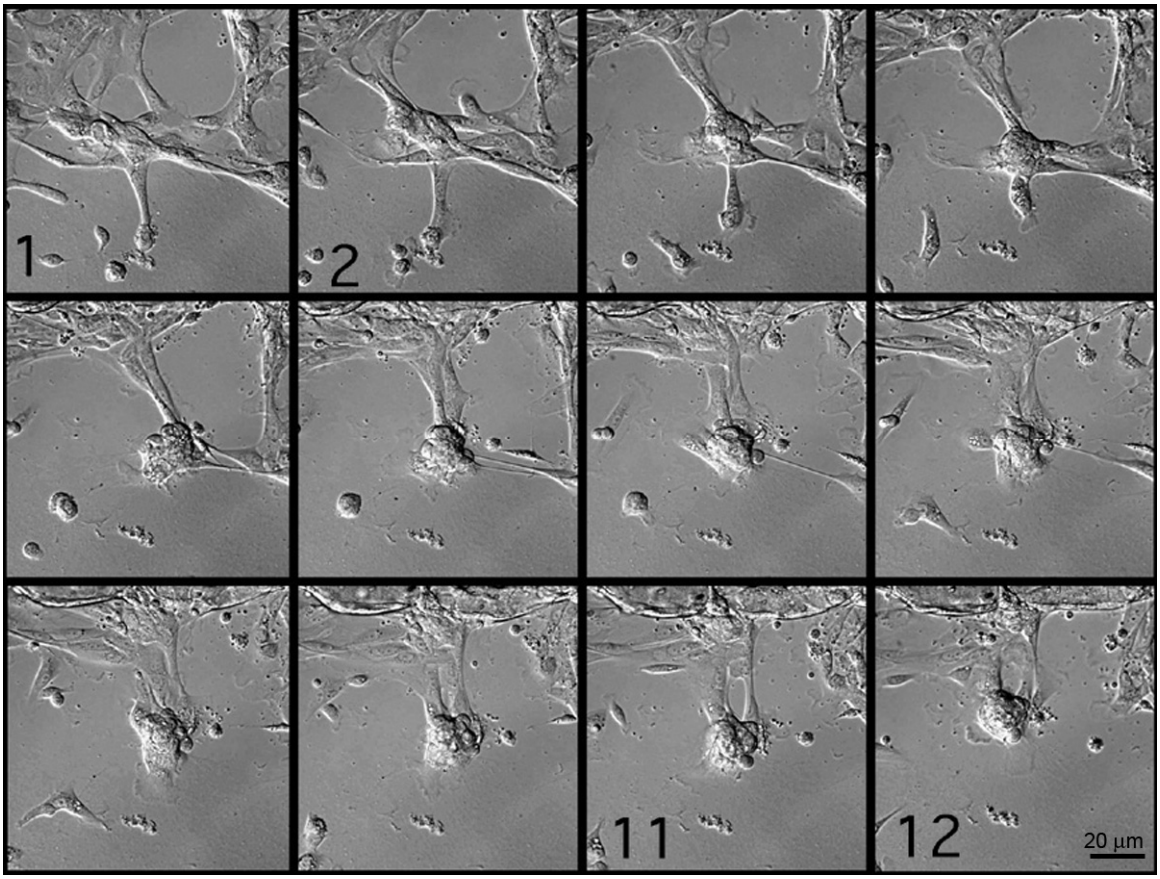


Figure 2.3 Still images captured from a time lapse digital movie filmed of a primary culture of breast reduction mammoplasty tissue. These images show the progressive *de novo* formation of a three dimensional mammosphere from epithelial cells in culture over the course of approximately 8 hours in total. Images shown were captured at 40 minute intervals on a Zeiss Axiovert 100, with DIC (differential interference contrast) optics, 20x objective in conjunction with QED imaging software.

2.3.3 Mammosphere Characterization

Confocal microscopy of saggital sections of mammospheres that were maintained over several weeks showed that 90% of the mammospheres contained internal lumen (Figure 2.4c,e) and that at least 50% were multi-lobulated structures (Figure 2.4h). The size of these structures was variable, ranging from 40 to 50 μm, although much larger structures several hundred microns in diameter and consistent with the sizes of lobules *in vivo* have also been observed.

Cell-type specific expression of epithelial membrane antigen is associated with luminal epithelial cells from the breast and other tissues (270,272), as well as mammary stem cells (273). Confocal microscopy of epithelial membrane antigen-stained mammospheres is shown in Figure 2.4f,g,h. Epithelial cells (based on morphology) stain strongly in the presence of the antibody (Figure 2.4f,h), whereas the surrounding layers of stromal fibroblasts do not Figure 2.4g). A rhodamine-conjugated phalloidin counterstain, a stain specific for filamentous actin which binds every cell regardless of morphology was applied for context.

Luminal mammary epithelial cells are known to specifically express cytokeratins-18 and -19 (270). Mammospheres in our system show clear staining with antibodies against cytokeratin-18 (Figure 4c,d,e) and -19. The mammosphere, shown in an optical section, shows a clear lumen with red cytokeratin-18 staining under confocal microscopy (Figure 2.4c). Figure 2.4d shows the staining in the x/z axis, indicating that the staining is continuous in all planes. The surrounding layers of stromal fibroblasts do not stain with this antibody (Figure 2.4e).

Normal breast epithelial cells have nuclei that are polarized to the basement membrane. We performed epifluorescence microscopy followed by image deconvolution on individual cells

from chemically stained mammospheres (Figure 2.4i). Randomly selected cells from these mammospheres all showed polarized nuclei (green) resting against the cell membrane (blue).

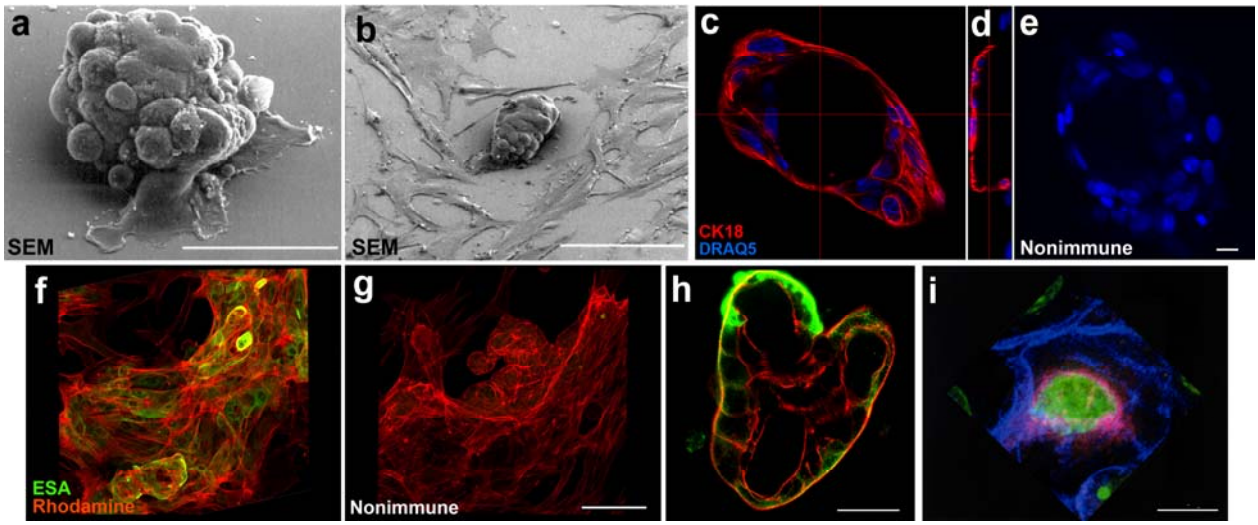


Figure 2.4 Characterization of mammospheres.

(a and b) Scanning electron micrographs of a representative primary culture from breast reduction mammoplasty tissue showing: (a) a typical mammosphere. The flattening and merging of the cells on the top of the mammosphere is suggestive of the presence of gap junctions (desmosomes). The scale bar represents 40 μm.

(b) A second representative mammosphere at lower magnification showing the presence of stromal fibroblasts from the same breast surrounding the mammosphere. The scale bar represents 100 μm. The microscope used for images was a JEOL JSM 6330F Scanning electron microscope at 10 kV.

(c) Photomicrograph of Cytokeratin-18 staining of a luminal mammosphere showing positive staining with a anti-cytokeratin-18 antibody and a goat anti-mouse Cy3 conjugated secondary antibody in a typical cytoplasmic pattern in the x/y axis.

(d) The x/z axis of a confocal stack shows that the cytokeratin-18 staining is evident in all three dimensions.

(e) Photomicrograph of the non-immune staining showing non-specific secondary antibody staining. Indirect immunofluoresence was used in (c), (d), and (e). All 3 micrographs contain nuclei counterstained with the chemical DRAQ 5 (blue). The scale bar represents 10 μm.
(f, g, h) Epithelial membrane Antigen staining of the mammospheres in the context of the stromal fibroblasts. Cells are counterstained with a rhodamine-conjugated phalloidin (F-actin-red). An Olympus Fluoview 1000 scanning laser microscope were utilized (using a 40x and 60x UPlan oil objectives). (f) Confocal photomicrograph showing a three-dimensional reconstruction of a low magnification field of mammospheres intercalated into the stromal fibroblasts stained with epithelial membrane antigen conjugated to Alexa 488 (green). The scale bar represents 50 μm.
(g) Confocal photomicrograph showing a similar field of cells on the opposite chamber of the same slide stained with non-immune serum. The scale bar represents 50 μm.
(h) Confocal photomicrograph of an optical section of a multi-lobulated mammosphere showing a positive stain for epithelial membrane antigen (green). The scale bar represents 20 μm.
(i) A single randomly chosen and representative cell from a mammosphere stained with TMA-DPH (blue membrane stain) and Syto -16 (green nuclear stain) and Mitrotracker Red (red mitochondrial stain), cropped and subsequently deconvolved using Deltavision deconvolution epifluorescence microscopy on a Nikon microscope. This cell shows that mammospheres contain cells with a polarized nucleus. The scale bar represents 5 μm.

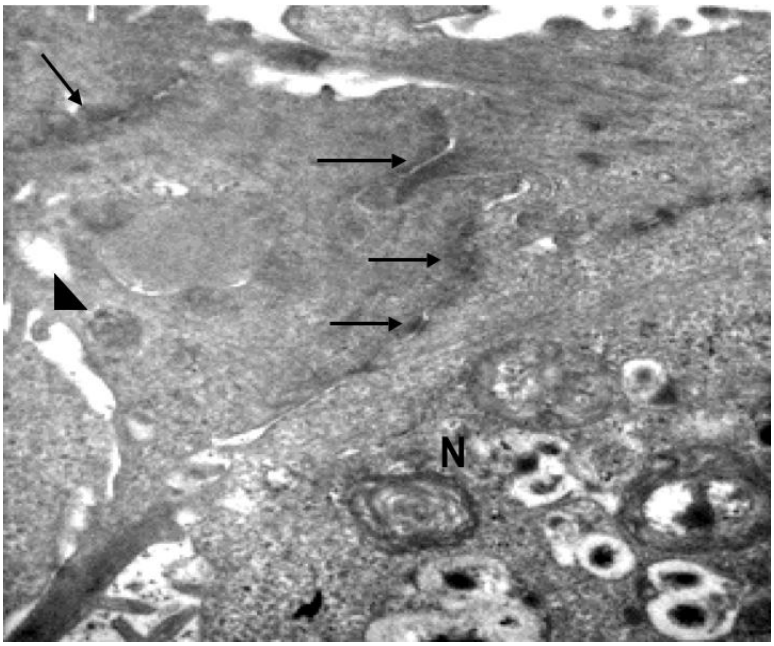


Figure 2.5 Transmission electron microscopy of a three-dimensional mammosphere showing microvili (arrowhead) and desmosomal connections between the cells (arrows). The nucleus is shown in the lower right corner as marked by the N.

2.3.4 *De Novo* Formation of Ductal Structures

Mammospheres, through a process of epithelial cell migration and proliferation, give rise to ductal structures over a period of 2-3 weeks in culture. Eleven to twenty-one days after initiation of human mammary epithelial cell cultures, ductal structures can be seen (Figure 2.2, 2.6). These structures are luminal as shown by the pattern and outline of cytokeratin-19 staining (Figure 6d) and DRAQ 5-stained nuclei in Figure 2.6e.

With greater time in culture (2-5 weeks) linear ductal structures become longer, progressively branching ductal networks (Figure 2.6b). Mature ducts manifest lumen with a diameter consistent with interlobular ducts *in vivo* (40-70 μm). Three dimensional reconstruction of several series of z axis confocal images show that branching structures are generally located 15-30 μm above stromal fibroblasts that have created a complete or partial monolayer on the bottom of the culture dish (Figure 2.6a,b).

Cells consistent with the staining pattern of myoepithelial cells are shown surrounding the inner layer of cells in another patient sample (Figure 2.7). Cytokeratin 14-positive cells are predominantly present on the edges of this structure with epithelium flattened along the exterior of the three dimensional architecture. This phenotype is consistent with myoepithelial cells.

Transmission electron microscopy of one patient sample showed luminal epithelial cells with apical microvilli (Figure 2.8a, b). The lateral surfaces show the presence of desmosomes linking the cells together (Figure 2.8a), as were also seen in the mammospheres (Figure 2.5).

One of the critical features of an organotypic system is the production of proteins that are related to organ function. β-casein is a milk protein that is rarely expressed in any primary epithelial cell culture system (288), although it is expressed in approximately 17% of human

breast tumors (289). In our cultures, this milk protein is expressed in the lumen of both mammospheres and ductal structures (Figure 2.8b).

The length and complexity of a ductal system was captured in a digital movie (Figure 2.9). This movie showed the microscopic examination of a large interconnected complex of ductal units over hundreds of microns. All of the images in this movie were collected at one time on a fixed culture.

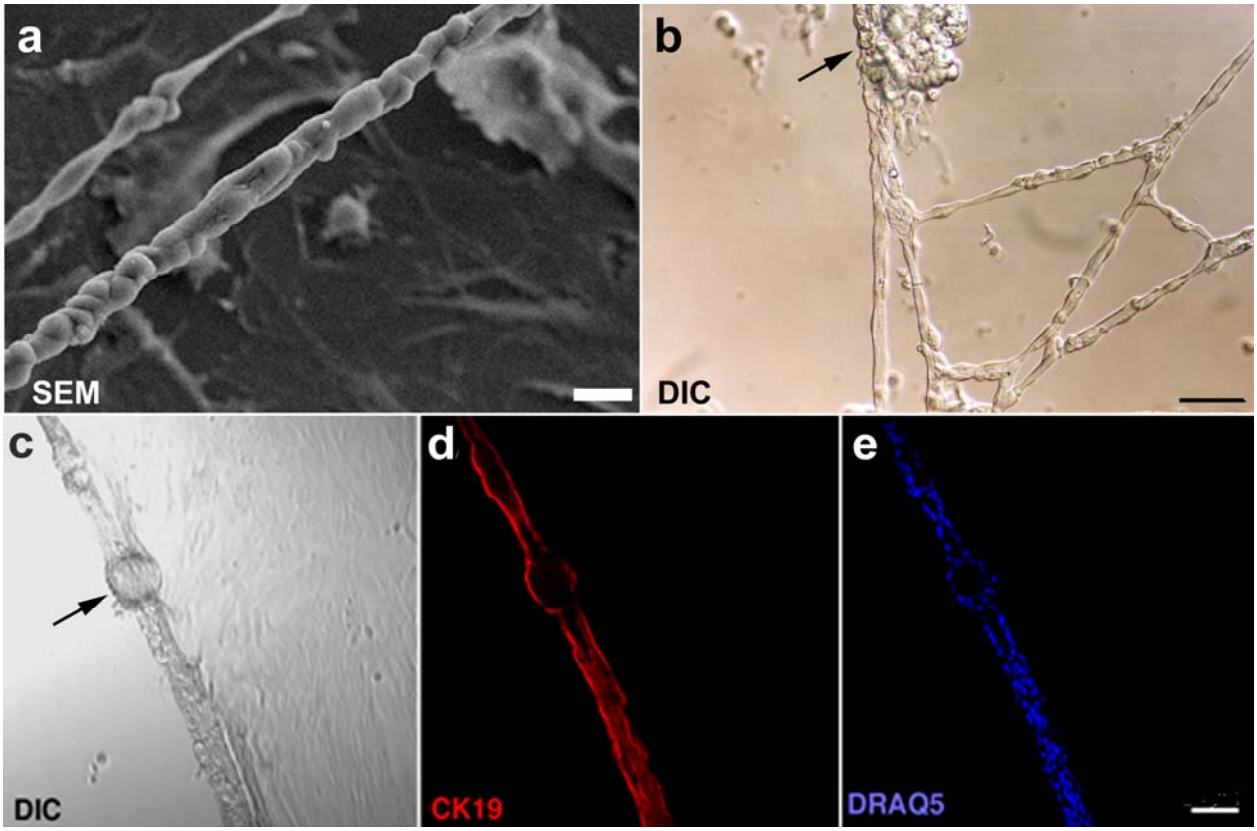


Figure 2.6 Characterization of ducts.

(a) Scanning electron micrograph of a mature duct (foreground) derived from a breast reduction mammoplasty primary culture showing that the ductal structures are suspended above the stromal fibroblast layer on the bottom of the dish. The scale bar represents 100 μm.

(b) Photomicrograph using DIC optics of branching ducts in a 15-day-old primary culture of normal breast reduction. A lobule approximately 30 μm in diameter is shown at the top center of this micrograph (arrow). The microscope used was a Zeiss Axiovert 100, with DIC optics, 20x objective. The scale bar represents 20 μm.

(c) Photomicrograph of a ductal structure containing an intra-duct lobule (arrow) under DIC optics.

(d) Photomicrograph of a confocal section of the same ductal/lobular structure under indirect immunofluoresence using after staining with anti-cytokeratin 19 antibody and a goat anti-mouse secondary antibody conjugated with a Cy3 fluorophore. The luminal nature of the entire structure is evident.

(e) Photomicrograph of a confocal section of the same ductal/lobular structure to show DRAQ 5 nuclear staining. The luminal nature of this structure is clearly shown by the pattern of the stained nuclei. The scale bar represents 20 μm.

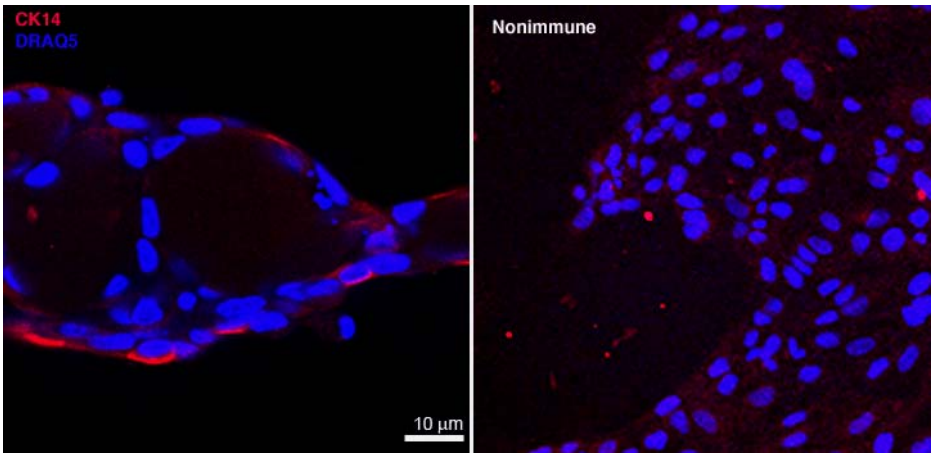


Figure 2.7 Image of an intraductal lobule that demonstrates staining of cytokeratin 14. An Olympus Fluoview 1000 was used to capture this image of an intraductal lobule after staining with anti-cytokeratin 14 antibody and a secondary antibody conjugated with a Cy3 fluorophore. A control without the primary antibody is seen on in the right showing nonspecific staining of the secondary antibody. Nuclei are counterstained with the chemical DRAQ-5 to reveal nuclei.

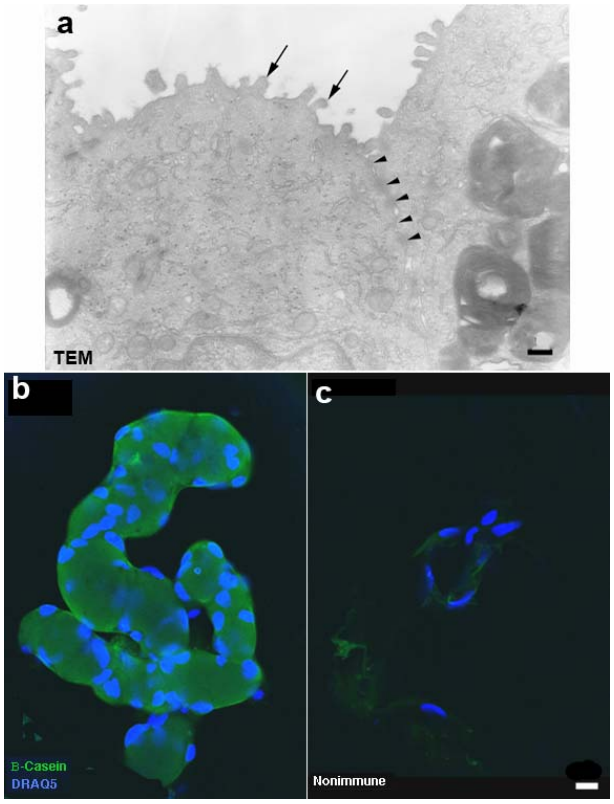


Figure 2.8 Ducts demonstrate apical microvilli and desmosomes and secrete β-casein.

(a) Transmission electron microscopy of a section through a ductal structure showing the presence of features consistent with healthy and functional epithelial cells. Microvilli are present on the apical surfaces of the epithelial cells (arrows). Desmosomes (at the arrowheads) are present at the lateral surfaces connecting the epithelial cells. Sections were photographed using a JEOL JEM 1210 transmission electron microscope at 80 kV. The scale bar represents 200 nm.

(b) Photograph of a tripartite ductal structure attached to the bottom of the dish stained with an anti-β-casein antibody and a secondary goat anti-mouse antibody conjugated to Alexa 488 (green). Nuclei are counterstained with the chemical DRAQ 5 stain (blue). The entire duct stains strongly with this antibody showing that these cells are producing this milk protein.

(c) Photomicrograph of the non-immune control is shown. The scale bar represents 5μm.

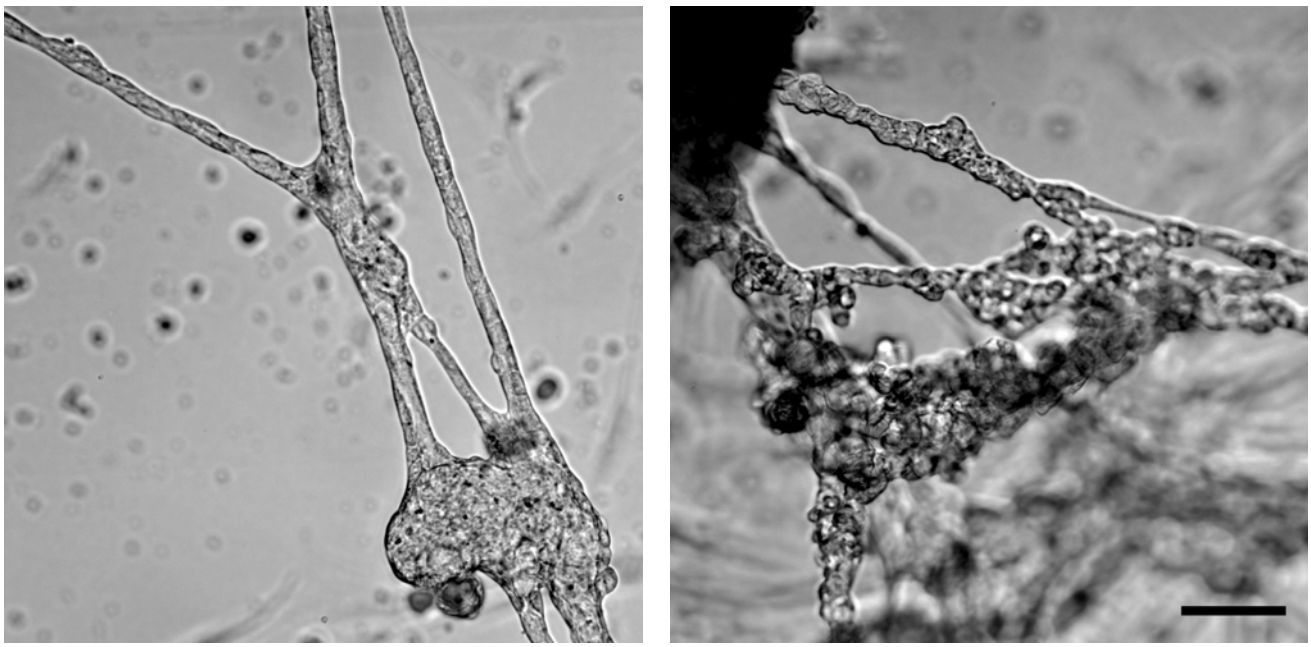


Figure 2.9 Sample DIC images taken from a digital time lapse microscopy movie filmed on a fixed primary culture of breast reduction mammoplasty tissue.This complex structure was suspended above the fibroblast layer formed on the bottom of the dish, seen below the focal level in the images. The microscope used was a Zeiss Axiovert 100, with DIC optics, 20x objective. The scale bar is 20 μm.

2.3.5 Formation of Lobular Structures

The formation of intra-ductal lobular structures (Figure 2.6 c,d,e) and terminal ductal lobular units (Figure 2.10a,b) is also manifested in at least 40% of these long-lived cultures.

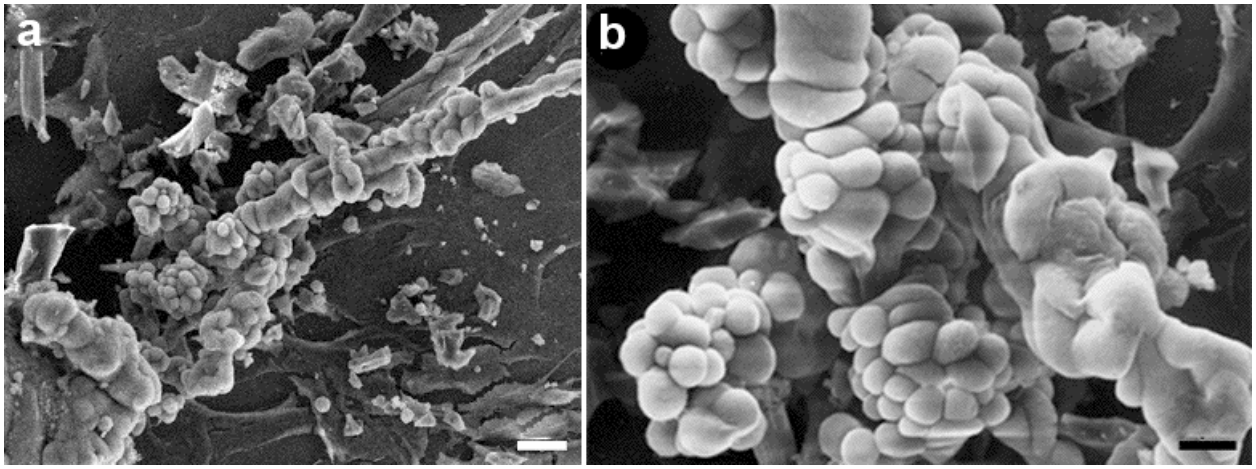


Figure 2.10 Scanning electron micrographs of a terminal ductal lobular unit.

(a) A scanning electron micrograph showing a lower magnification view of a Terminal Ductal Lobular Unit. The ductal structures are visible at the top of this terminal ductal lobular unit (arrowheads). The scale bar represents 100 μm.

(b) A higher magnification image of the same terminal ductal lobular unit. The scale bar in represents 10 μm.

2.4 DISCUSSION

In order to perform *in vitro* experiments on complex biological systems it has traditionally been necessary to simplify them to their most essential components, *i.e.,* single cell types or further, clonal populations of cells. Several cell lines have also been created in an attempt to represent normal human epithelial cultures. Stable transfection of such agents as SV40 large T antigen (290,291), human papilloma virus E6 protein (292) and telomerase (293) or mutagenesis with benzo[*a*]pyrene (294) have been used to immortalize nondiseased epithelial cells from breast reduction mammoplasty tissues, allowing cultures to live long enough that they are experimentally manipulable. These immortalized cell lines may exhibit profound changes from the parental cell-type, however, including alterations in such basic metabolic processes as DNA repair (295,296). Our microarray analysis proves that these alterations have occurred (Chapter 4). The spontaneously-immortalized MCF10 cell lines (MCF10A and MCF10F) were generated from a fibrocystic area of a nondiseased breast (297). As immortalized cell lines in traditional monolayer culture, these MCF10-derived lines do not have the capacity to spontaneously re-establish normal three-dimensional epithelial architecture. As a complement to these simplified systems we have developed a system that contains multiple cell types, a thin coat of matrigel and a novel medium that allows for continuous growth in culture for over 3 months. Our system also has the advantage that the epithelial cells remain optically and physically accessible for the purpose of quantifying changes, for example, after exposures to xenobiotics. Our system therefore combines the physiologic relevance of primary human breast cultures with the long-lived and experimentally manipulable nature of cell lines.

In developing our culture system we avoided the classical method of reducing tissue to single cells, in an effort to maintain the juxtacrine relationships between epithelial cells (a similar physical mincing and direct culture method was used in 1958 to generate the BT20 tumor cell line (298). Although the presence of multiple cell types can create complicated scenarios for molecular analysis, we believe that without the presence of paracrine and juxtacrine factors provided by supporting organ-specific cell types, organotypic differentiation of such a progressive and continuous type would not be possible. While other groups have used feeder layers like xenobiotic 3T3 cells, we propose the use of the naturally matched isogenic stromal cells to promote the long-term growth and development of the epithelial cell culture. These stromal cells provide a more physiologically relevant set of juxtacrine and paracrine factors than albino mouse embryo-derived 3T3 fibroblasts.

The strongest indicator that traditional systems for culturing breast epithelial cells are limited is the low success rate of such cultures overall in the literature, and the relative paucity of laboratories which can generate novel cultures of nondiseased and diseased cells. Recent advances in cell culture techniques has increased the efficiency of generating primary cultures from breast reduction tissue to nearly 100% (299). The success rate for generating cell lines from these primary cultures is still disappointingly low, and estimates of the efficiency have not been published. In contrast, our system has provided both a nearly 100% success rate in both the generation of primary cultures and the generation of cell lines (Chapter 4) from breast reduction mammoplasty tissue.

The same systems that have had success with generating primary cultures for normal tissue, however, are not able to grow diseased tissue. Only a 10-15% success rate has been reported overall for tumor cell lines derived from human breast tumors (299,300). The

generalized lack of tissue-specific gene expression in cell lines derived from tumors, most obviously that of estrogen receptors, and those genes identified in the "basal-epithelial-like" breast cancer subtype defined by microarray, has further underscored the need for improved *in vitro* systems. Again, in contrast to this published success rate, we have been able to culture breast tumor tissue from ductal carcinoma in situ, and stage I through IV invasive ductal carcinoma with efficiency of 85%, far superior to current systems.

It is significant that the primary cultures established in this study have not undergone senescence or programmed cell death but have persisted as primary cultures for weeks to months. This may suggest that our culture medium and the initial natural configuration of the epithelial cells as "mammospheres" retain an adult pluripotent population of cells. Interestingly, the mammosphere configuration immediately after tissue disaggregation has recently been shown in another system proved to contain adult mammary stem cells (273).

Since several laboratories have elegantly demonstrated the concept that form follows function *in vitro* (274-277,301), more laboratories are utilizing complex substrata or feeder layers to support epithelial cell attachment and growth. In the literature, it has been published that a concern with primary cell cultures generated from a mixture of cells from the breast is that they will eventually evolve into cultures with an entirely fibroblastic morphology, or alternatively, that the stromal fibroblasts will "overgrow" the epithelial component of the culture (302). We believe that a complex substratum is necessary to initially promote the attachment of the epithelial cells and that this substratum helps to establish these cells until that they can proliferate and reiterate a ductal system in culture. It is our contention that once a cell population has been grown on tissue culture plastic, one cannot entirely trust that its manifested morphology reflects its true origin or nature.

We also use a high percentage of sera in our novel MWRI medium in order to more closely mimic the *in vivo* situation and maintain the signaling from matrix molecules present in solution in serum. Although it is not always possible to use serum in experimental settings, the establishment of primary cultures has historically involved the use of 20-25% serum. This is presumably used in order to ease the transition from the body where the serum concentration by volume is on average approximately 50%. It is clear that we do not know all the necessary factors for organotypic expression of the identified breast ductal system, making high quality serum the obvious choice for this type of culture.

Ductal structures of the complexity of the ones generated in our laboratory's culture method are unprecedented in the literature. Gudjonsson et al published that they could isolate cells that generated "acinus-like spheres" and "elaborate branching structures resembling uncultured terminal ductal lobular units by both morphology and marker expression" (272). Gomm et al also showed luminal structures that formed after purified populations of epithelial and myoepithelial cells were mixed in culture (303). Dontu et al showed similar structures that were affected by Notch signaling in 2004 (304). While each of these publications demonstrate a cell culture morphology more elaborate than simple mammosphere formation, none of them show the progressive branching pattern or the presence of intra-ductal lobules illustrated here. The distances traveled by ductal structures in the literature are at most equivalent in diameter to one of their formed acinar structures while ours are capable of generating a network that spanning over 100 micrometers and in some cases over the entire culture area of a 12.5 cm^2 dish. Cells participating in the structures labeled as terminal ductal lobular units in these publications also did not show the flattened surface appearance of tight desmosomal connections as seen in our scanning electron micrographs. Finally, these other culture systems did not demonstrate the

functional capacity of casein production in their structures as has been established with the Latimer culture system.

Primary cultures of breast tissue have provided our laboratory and others with more organotypic and tissue-relevant models for studying the stromal/epithelial interactions of the breast. The complex remodeling that occurs each month in breast tissue may only be possible in the presence of this stromal/epithelial cross communication after the inducing paracrine signaling from the ovary and pituitary. Xenograft models have proven that during fetal development, the ductal system can only form in mammals when fat is present in the breast, underscoring the need for multiple cell types to induce differentiation (264). Studies using a stromelysin transgenic mouse model have provided further evidence that stromal/epithelial interactions are necessary for maintaining normal epithelial function (305,306). In the present study, we utilize multiple cell types *in vitro* with the endpoint being ductal and lobular morphological differentiation. A greater number of nondiseased cell lines representing a greater population of subjects would be extremely useful for study in order to encompass some of the natural variability in the human population. In addition, the use of such normal cell lines or long-lived organotypic explants from nondiseased breast tissue would add an *in vitro* epidemiological dimension to pharmacologic and genetic studies that has been largely unexplored in the literature.

2.4.1 Tumor Cell Cultures

As far as the creation of tumor cell lines in concerned, it has been observed by many laboratories that tumors explanted in culture yield cells that contain normal karyotypes after weeks of maintenance (307). The use of serum was often blamed for this phenomenon, with the

hypothesis that it was in some way toxic to tumor cells. Alternatively, it was suggested that the media used were so effective for the culture of normal breast tissue that it caused the normal cells of the culture to overgrow the abnormal ones (299). The normal diploid cells that were found in these cultures were also shown to be cytokeratin 19 positive (308), suggesting that they are not luminal epithelial cells. Thus we offer the alternate explanation that the tumor cells are rather fragile in culture and that in other systems they gradually die, leaving the normal cells that have infiltrated the tumor, such as vascular cells or hematopoetic cells, behind after tumor cell death. These normal cells are then responsible for the normal karyotype and cytokeratin 19 phenotype that has been observed.

In contrast to the historic problem with generating primary cultures from breast tumors, we have experience an 85% success rate using the same culture system as used for the breast reduction tissue. We do not see an outgrowth of normal tissue, but rather a distinct tumor phenotype in this system. Tumor cells harvested from stage I through stage IV invasive breast carcinoma did not form ductal architecture as was shown in nondiseased cultures (Figure 2.11). They lacked collaborative behavior with neighboring cells in the same area of the tissue culture dish. The exception to this was stage 0 ductal carcinoma *in situ*, which formed abnormal-looking open structures.

In general, these tumors cells also took on a more fibroblastic than epithelial appearance despite the fact that they were originally derived from breast epithelium. This fibroblastic morphology is consistent with an epithelial to mesenchymal transition (EMT) implicated in the conversion of early stage tumors into invasive malignancies. Epithelial to mesenchymal transition is a phenomenon most closely understood during embryogenesis. It is a process whereby epithelial cell layers lose their polarity and cell-cell contacts, including through the

much-studied loss of expression of E-cadherin, and undergo remodeling of their cytoskeleton (309). Once these cells have decreased their contacts, they acquire a more motile phenotype. In embryogenesis, this is involved in processes such as gastrulation and the formation of multiple organs and tissues. In the adult, epithelial to mesenchymal transition is implicated in wound healing. The proposal that epithelial mesenchymal transition is involved in cancer came shortly after its discovery in the process of embryogenesis (309). The use of this developmental pathway is believed to allow the luminal epithelial cells to lose their architecture and become more motile, allowing invasive movement. Evidence to support this theory has come from studies which have shown genes that are important in embryonic development are often involved in cancer. Conversely, genes discovered through their role in the oncogenic process are often involved in embryogenesis. The best example of these is E-cadherin, involved in cell-cell contact. Loss of expression of this protein is seen at sites of epithelial to mesenchymal transition both in embryogenesis and in carcinogenesis. It is possible that the more fibroblastic appearance and the lack of cell-to-cell connectivity of tumors in our culture system represent this transition.

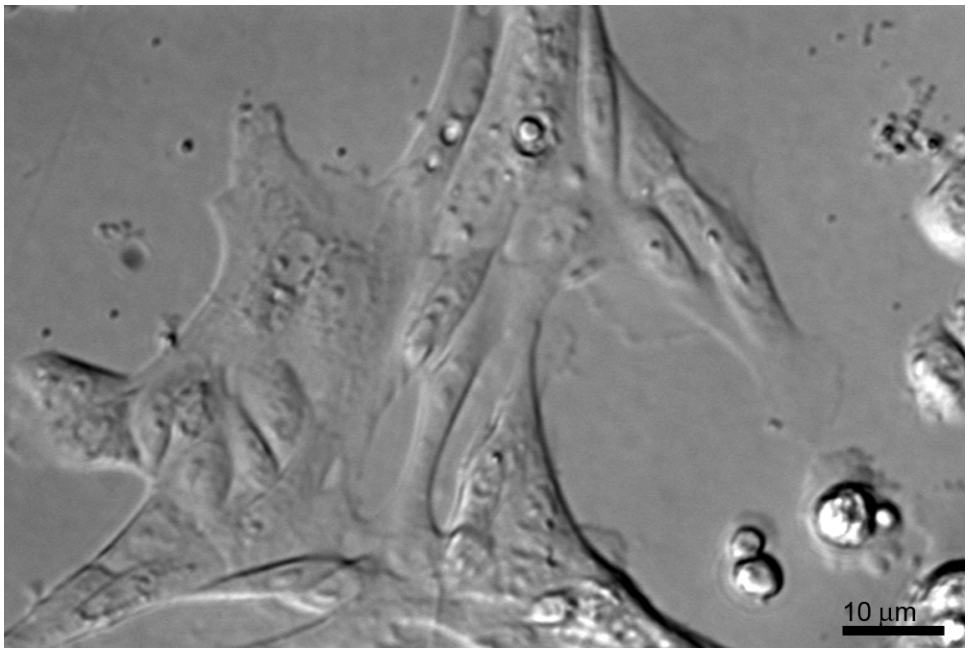


Figure 2.11 DIC photomicrograph of a 4 day-old primary culture generated from a stage II invasive ductal carcinoma in the Latimer lab. The microscope used was a Zeiss Axiovert 100, with DIC optics, 20x objective.

Because this system can culture both tumor and normal cells, it makes an ideal experimental model for the study of transformation. The striking difference in morphology of tumor cells makes it easily identifiable in culture as a distinct phenotype. These architectures, or lack thereof in the tumor samples, remain optically accessible for evaluation. Therefore, this culture system can be used to determine the effects of exposure of breast cells to multiple agents. For example, the effects of suspected environmental carcinogens could be evaluated by looking for either the alteration or destruction of normal architecture within these cultures or the alteration of the time course over which it appears. In the vein, viral transduction of these

primary cultures could possibly be used to evaluate the effects of various suspected oncogenes. Furthermore, the examination of factors involved in *in vivo* breast development can be evaluated in a systematic way. This experimental system allows addition of hormones at known final concentrations in any temporal pattern desired.

The fact that this system is capable of supporting both tumor and normal growth, has a minimum of an 85% success rate with culturing breast tissue (100% with nondiseased breast tissue) makes it an ideal system for more population-based studies of breast cancer. The low rate of success of breast cancer cell culture limits the breast cancer research community to studying a subset of cancer cases which are capable of growth in current culture conditions. This limitation is further discussed Chapter 4 with respect to cell lines. This system will be used in the following chapters for the evaluation of DNA repair capacity in both normal and cancerous tissue, and cell lines generated using this system will be used for the evaluation of mRNA and protein expression as it relates to functional repair capacity.

3.0 FUNCTIONAL ANALYSIS OF THE REPAIR CAPACITY OF NORMAL BREAST TISSUE AND BREAST TUMORS

3.1 INTRODUCTION

Long patch, or Nucleotide Excision Repair (NER), is the primary process by which cyclobutane pyrimidine dimers, 6-4 photoproducts, bulky adducts, and DNA cross-links are removed from the DNA (310). Any damage that generates a helical distortion in the DNA helix is a possible substrate for NER. NER is a complicated process requiring the protein products of 20-30 genes (311). NER involves the recognition of a damage lesion, incisions flanking the lesion on the damaged strand, excision of 27-29 bases including the damage lesion, and replication and ligation to replace the excised information and seal the strand breaks at each end of the newly synthesized region (312).

NER of the total genome can be measured quantitatively using the unscheduled DNA synthesis assay (UDS). The UDS assay involves the measurement of labeled base incorporation into the DNA after *in vitro* exposure to UV light or certain chemicals. The UDS assay is a cell autonomous, functional assay in that it allows one to look at the complex process of NER as a whole, at least as it is expressed in a particular cell type (59,60,62). As applied in our laboratory, this assay predominantly quantifies the repair of UV-induced DNA 6-4 photoproducts, and

elements of both the "global genomic" as well as the "transcription coupled" components of NER contribute to the results (60,310).

The autoradiographic UDS assay requires the analysis of living cells. It has previously been applied primarily to skin fibroblasts and peripheral blood lymphocytes (PBLs) for diagnosis of xeroderma pigmentosum (XP) and other DNA repair diseases impacting specifically on the NER pathway, and remains the only functional assay used to diagnose XP. Classical NER deficiency disorders are characterized by UV sensitivity manifesting mainly in the skin and cornea. In diagnostic studies, a single "normal" sample of the same tissue type is used as a control. Alternatively, foreskin fibroblasts (FF) have historically been used as positive standards in unscheduled DNA synthesis studies, because these cells manifest a consistently high level of repair. Mixtures of FF from several babies have sometimes been used to attempt to account for possible interindividual differences. Studies involving functional assays in general, and specifically functional assays of DNA repair capacity, have been hampered by a technical lack of ability to perform primary explant culture on all cell types. Although repair assays can be performed on established, transformed cell lines, the generation of cell lines from normal adult tissue has proven to be a technical challenge. In addition, during the process of passaging, established cell lines undergo clonal evolution that alters or extinguishes many of the original characteristics of the cells, including their intrinsic repair capacity (313,314). We have developed a novel system for establishing primary cultures of breast epithelial cells and stromal fibroblasts that can also be used to culture a variety of normal tissues, including ovarian epithelium (283). Since most breast tumors arise from epithelial cells, the evaluation of their baseline mutator gene function, *i.e.*, DNA repair capacity, is extremely important, and is now possible in primary epithelial cultures.

It has long been presumed that DNA repair is so essential to the maintenance of genomic integrity that it is constitutively expressed in all cells within an individual. A natural extension of this assumption is that DNA repair capacity is equivalent in all cell-types within an individual. However, the possibility that NER can be regulated is illustrated by the findings that NER is present in different levels in various tissues and cell types in mammalian development. In a previous study of the excision repair capacity of the four distinct extraembryonic lineages that comprise the extraembryonic yolk sac, as well as five cell types derived from the fetus, we have shown that NER in the mouse is lineage-specific during embryogenesis (247). Further analysis has been conducted on a small sampling of human tissues (140), including breast and now ovarian epithelial cells.

Genetic instability is a hallmark of all human cancers (1). It has been hypothesized that genetic instability must arise during carcinogenesis, since otherwise the multistep process would take too long to occur in a human lifetime (315). Previous studies have focused primarily on loss of post-replicative base mismatch repair capacity, as this is the genetic basis of the familial cancer predisposition syndrome, hereditary non-polyposis colorectal cancer (HNPCC) (316). New focus on DNA repair in breast cancer etiology comes from two lines of evidence: (1) The BRCA1 hereditary breast cancer gene has been shown to be involved in DNA double strand break repair, specifically homologous recombination (317,318), and can regulate the expression of two genes involved in NER, XPC and DDB2 when overexpressed in an osteosarcoma cell line (319,320). The data presented here, however, does not support the idea that BRCA1 modulates NER capacity. (2) DNA repair defects have also been identified in the peripheral blood lymphocytes (PBLs) of sporadic breast cancer patients (66,321-324). Loss of NER is the basis of the cancer prone disease XP, which is characterized by a 2000-fold increase in the incidence

of UV-associated skin cancer. With the baseline functional capacity of breast tissue now established, we have investigated the role NER in breast cancer in a BRCA1 mutation carrier and multiple sporadic breast cancer cases.

3.2 MATERIALS AND METHODS

3.2.1 Tissue Procurement and Establishment of Cultures

3.2.1.1 Breast Reduction Tissue

Breast tissue was obtained from patients at Magee-Womens Hospital under Magee-Womens Hospital/University of Pittsburgh IRB # MWH-94-108. For breast reduction mammoplasty tissues, a neighboring piece of mammoplasty tissue (from the same 0.25 cm^2 sample) to that placed into primary culture was fixed and processed in paraffin. Sections were examined by a pathologist to verify the histological normality of the tissue. Breast tissue was rinsed, processed and placed into primary cultures within five hours of surgery. Tissue was mechanically disaggregated and placed on diluted (1:1) Matrigel in a novel tissue culture medium called MWRIα according to a method developed in our laboratory (283) (see also Chapter 2). Primary breast cultures were grown for 7-10 days, imaged using a digital Hammamatsu video camera and then analyzed for NER capacity using the UDS assay. UDS experiments were performed when the epithelial cells were present as "mammospheres" (clusters of epithelial cells numbering from 30-100 cells).

3.2.1.2 Breast Tissue from a BRCA1 Mutation Carrier

The BRCA1 Q1200X mutation carrier examined was a 35.7 year old woman with strong family history of breast cancer recruited into a clinical trial of MRI screening for young woman at high risk for breast cancer with dense breast tissue. Gadolinium enhancement images revealed a small 1 cm lesion in the upper-outer quadrant of the left breast, identified pathologically as an infiltrating ductal carcinoma. The patient underwent a modified radical mastectomy of the left breast and chose to also undergo a contralateral prophylactic total mastectomy. Blood and tissue from this patient were also obtained at Magee-Womens Hospital under Magee-Womens Hospital/University of Pittsburgh IRB # MWH-94-108. Processing of the tissue took place as described above and in Chapter 2 for breast reduction mammoplasty tissue.

3.2.1.3 Breast Tumors

Tissue from invasive breast ductal carcinoma was obtained at Magee-Womens Hospital under Magee-Womens Hospital/University of Pittsburgh IRB # MWH-94-108 from women who were undergoing surgery for removal of a breast tumor. Samples were trimmed within the tumor margin on all sides. One third of each specimen was used to initiate explant culture, one-third was processed for confirmatory histological assessment, and one third was frozen for future studies. Tumor specimens were processed as described above and in Chapter 2 for breast reduction mammoplasty tissue. Breast tumor tissue manifested in culture generally as epithelial cells, but in some cases exhibited a fibroblastic appearance. In contrast to the structures formed in breast reduction mammoplasty culture, architecture were never formed by the invasive breast carcinoma primary explants.

3.2.1.4 Clinical Information Pertaining to Breast Tumor Samples

ER and PR status of the breast tumors were determined clinically using standard peroxidase antibody staining. Seven patients chose not to undergo axillary lymph node dissection to confirm negative lymph node status. Standard inquiries pertaining to family and personal cancer histories were made at Magee-Womens Hospital at the time of diagnosis. None of the subjects reported in this study had significant family histories of breast or related cancers, nor had they ever been diagnosed with another type of cancer themselves.

3.2.1.5 Ovarian Epithelium

Ovarian tissue was obtained at Magee Womens Hospital under Magee-Womens Hospital/University of Pittsburgh IRB # MWH-94-108 from women undergoing hysterectomy for benign uterine conditions. Epithelial cells from the surface of the ovary were removed by direct scraping and placed into culture on diluted (1:1) Matrigel in the novel tissue culture medium MWRIα. Cultures were maintained as described in Chapter 2 for breast reduction mammoplasty tissue.

3.2.1.6 Peripheral Blood Lymphocytes

PBLs were obtained from normal healthy male and female controls ages 20-50 working at Magee-Womens Hospital, Magee-Womens Research Institute or students at the University of Pittsburgh. Lymphocytes were purified using the ficoll gradient method (325) and placed onto a diluted form of Matrigel (BD Biosciences, Bedford, MA) in RPMI medium supplemented with 15% fetal calf serum. It was discovered that peripheral blood lymphocytes adhered to Matrigel-coated chamber slides, so the autoradiographic unscheduled DNA synthesis assay could be performed on these adherent cells (326). Lymphocytes were cultured 5-7 days before performance of the UDS assay. Nineteen out of the thirty-two samples presented here have been previously published as controls in a study on the effects of stress on NER capacity (326).

3.2.1.7 Foreskin Fibroblasts

FF tissue was obtained as discarded tissue from newborn infants after circumcision. These foreskin fibroblasts were converted into primary explants as described in Latimer *et al.* (140,247). Briefly, cells were grown in MEM containing 10% fetal calf serum on uncoated chamber slides (Nalge Nunc International, Naperville, IL). These cells were passaged to promote homogeneity and grown continuously in culture for up to 12 passages. These cultures show constant levels of DNA repair until passage 13 and senesce after approximately 20 passages (61). For UDS experiments, FF were utilized between passages 7 and 10. Four different preparations of FF were used in this study: one consisted of a pool of 3 perinatal circumcisions and the remaining three were made from individual circumcisions.

3.2.2 Cell Lines Used for Comparisons

TK6 lymphoblastoid cells and MDA MB-231 stage IV breast tumor cells were purchased from the American Type Culture Collection (Rockville, MD) and run as controls in each UDS assay. TK6 cells were cultured in RPMI supplemented with 15% fetal calf serum and adhered to Matrigel-coated chamber slides for analysis. MDA MB231 cells were cultured in DMEM with 10% fetal calf serum in uncoated chamber slides.

3.2.3 Unscheduled DNA Synthesis Assay

NER was measured using autoradiography of UDS (59). After a total of seven days in culture, without passaging cultures were irradiated with UV light at 254 nm at a mean fluence of 1.2 Joules/m^2 for 12 seconds in the absence of culture medium, for a total dose of 14 J/m^2. Primary cultures had not reached confluence and were still actively growing at the time the unscheduled DNA synthesis assay was performed. Control established cell lines (MDA MB231 and TK6) were plated subconfluently 1-2 days before the unscheduled DNA synthesis assay to insure that they also were not in a quiescent state brought on by confluence. Careful UV dosimetry was performed using a UV delivery system specifically designed for this assay (327). This machine contains three UV germicidal bulbs placed at a distance of 3 feet (91.4 cm) from an electric turntable where the chamber slides are placed. A 6-inch diameter photographic shutter opens under electronic control to deliver a precisely timed dose of UV light. UV bulbs were warmed up at least one hour before UDS analysis and UV output was checked before each experiment using a Spectroline DM-254X UV meter.

Each sample was represented by at least two chamber slides. One chamber of each 2-chamber slide was shielded from the UV dose to be used as an unirradiated control sample. After UV exposure, all cultures were incubated in medium supplemented with 10 μCi ml [^{3}H]methyl-thymidine (~80 Ci mmol-1) (PerkinElmer Life Sciences, Boston, MA) for 2 hours at 37°C. Labeling medium was then replaced with unlabeled chasing medium containing 10^{-3} M non-radioactive thymidine (Sigma, St. Louis, MO) and incubated for a further 2 hours to clear radioactive label from the intracellular nucleotide pools. After incubation in the post-labeling medium, cells were fixed in 1X SSC, 33% acetic acid in ethanol, followed by 70% ethanol and finally rinsed in 4% perchloric acid overnight at 4°C. All slides were dried and subsequently dipped in photographic emulsion (Kodak type NTB2) and exposed for 10 to 14 days in complete darkness at 4°C.

The length of exposure of emulsion was determined in each experiment by preparing "tester" slides. These are extra slides of the positive controls; for these experiments 2 slides each of foreskin fibroblasts and of the established breast cancer cell line MDA MB231. These slides were dipped in photographic emulsion, dried and packaged separately from the rest of the experiment in a sealed slide box. After 10 days these tester slides were developed and grain counting was performed. If the nuclei over the foreskin fibroblasts averaged 50 or more grains per nucleus, then the rest of the experimental slides were developed (including additional foreskin fibroblasts and MDA MB231 slides). If the grain count was below this level, the remaining slides were left to expose 1-3 days longer before being developed. The tester slides were only used to determine when the exposure time was optimal, since this can vary depending primarily on the age of the radiolabel and the emulsion.

3.2.3.1 Grain Counting Analysis

After photographic development of the emulsion on the slides, the nuclei were stained with Giemsa, then examined at 1000X magnification on a Zeiss Axioskop under oil emersion for grains located immediately over the nuclei of non-S phase cells (S phase cells were distinguished by their high grain counts, at least 10-fold higher than non-S phase, and by a clustered pattern of grains). Local background grain counts were evaluated in each microscopic field, over an area the same size as a representative nucleus, and this total was subtracted from the grain count of each nucleus in that field. The average number of grains per nucleus was quantified for each side of the chamber slide, both unirradiated and irradiated. The final NER value for each slide was calculated by subtracting the unirradiated mean grains per nucleus from the irradiated mean grains per nucleus, after the initial subtraction of local background in each field. NER was ultimately expressed as a percentage of the activity of concurrently analyzed FF. These results were then normalized by comparison to the average for the tissue type control population.

3.2.3.2 Statistical Analysis

To ensure accuracy and guard against transcription errors, raw grain counts from the UDS assay were processed independently in duplicate, once using StatView (version 5.0.1, SAS Institute, Inc., Cary, NC), and once using the Data Analysis Toolpack of the Excel 2001 spreadsheet program (Microsoft Corp., Redmond, WA). The final count from slides of the same cell type within the same experiment and developed the same day were averaged together and expressed as a percentage of concurrently analyzed FF, or, in the case of FF, a percentage of concurrently analyzed MDA MB-231. NER values for FF controls were averaged over all experiments.

Comparisons between different cell lines and cell types were performed using the parametric two-tailed t-test with significance determined at $\alpha < 0.05$. Comparison of UDS values for cells with epithelial and fibroblastic morphologies from the same samples was performed using a paired t-test at $\alpha < 0.05$. The possible effects of donor age, S phase index and clinical variables such as tumor grade, size and hormone receptor status on NER capacity were evaluated using linear regression at $\alpha < 0.05$. Age was evaluated both as a continuous variable and by assignment into two groups, premenopausal (< 50 years old) and peri- or postmenopausal ($\geq$ 50 years old). Estrogen receptor (ER) and progesterone receptor (PR) status were separated into three categories, negative, positive and overexpressed. These data were also evaluated by combining the negative and overexpressed groups into one "abnormal" category. Categorical variables were also compared using the χ^2 test.

3.3 RESULTS

3.3.1 Tissue Specificity of Nucleotide Excision Repair

Breast reduction mammoplasty tissue, FF tissue, and PBLs were analyzed by Latimer et al with the UDS assay (140). This survey of tissues showed distinct repair capacities for each tissue tested. This work has been extended now to include an additional breast reduction culture and 6 ovarian epithelial cultures generated using the same novel tissue engineering method (Figure 3.1). Samples include: 6 lots of FF, 23 breast reduction mammoplasty samples, 32 PBL samples, and 6 ovarian epithelium samples.

FF samples, traditionally used as the positive standard of comparison in diagnosis of XP, were considered to represent 100% repair. PBLs, another cell type that has been often analyzed with the UDS assay, exhibited only 9.0 ± 1.2% (mean ± standard error) of FF NER capacity, demonstrating that there is a greater than 10-fold range in NER capacities amongst normal human cell types (140).

Primary breast reduction mammoplasty cultures (including the additional samples) exhibited an intermediate NER capacity of 25.80 ± 2.34% of FF NER capacity. One additional sample was included in this analysis as compared to the published data of Latimer *et al.* (2003). This sample came from a 19-year-old breast reduction mammoplasty patient. This patient had only slightly lower repair than the point that was published as an outlier in Latimer *et al.* (2003). As with the data set included in the previous publication, this slightly expanded tissue population was significantly different from both FF (higher than breast) ($p= 0.00472$) and from PBLs (lower than breast) ($p= 3.04 \times 10^{-7}$).

Primary ovarian epithelial cultures exhibited a similar repair capacity to that of breast reduction epiethial cells of 21.0 ± 2.5% of FF NER capacity. As with the breast cultures, the ovarian primary cultures were significantly lower than FF ($p= 0.00363$) and higher than PBLs ($p= 0.00269$). However, no difference was observed between the breast and ovarian cultures ($p= 0.178$).

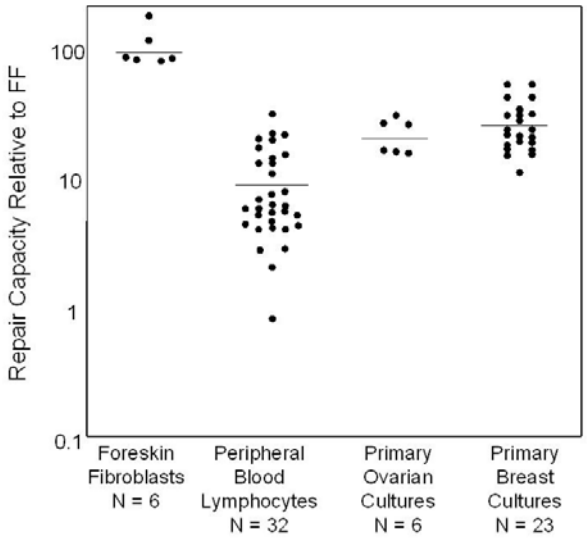


Figure 3.1 NER capacity of different cell types shows tissue specificity. NER capacity of FF (100%, N = 6), primary cultures of breast cells (25.80 ± 2.34%, N = 23), primary cultures of ovarian epithelium (21.0 ± 2.5%, N = 6), and peripheral blood lymphocytes (9.0 ± 1.2%, N = 32) from normal healthy newborns, women, and adults, respectively are shown here on a logarithmic scale. The mean of each data set is shown by the horizontal line.

3.3.1.1 Analysis of Breast Cells by Morphology

We also conducted a study of the effect of morphology on NER capacity (140). Because we used the autoradiographic version of the UDS assay, we were able to analyze the NER capacity of both the fibroblastic and epitheliod cells present in breast reduction mammoplasties generated using the Latimer culture system (Chapter 2) (Figure 3.2). In contrast to breast reduction cultures, ovarian tissue cells were isolated from direct scraping of the epithelial cell lining of the ovary and yielded purely epithelial cell populations. Likewise, FF and PBL cultures yielded cells of only one morphology. For the breast reduction tissue 54% percent of the cells that were cultured (across all samples) were epithelial with the remaining 46% fibroblastic. Three samples yielded only epithelial cells and 1 yielded only fibroblastic cells. Separate quantification of NER in the epithelial and fibroblastic cells from the normal breast revealed no statistically significant difference between the two cell types ($p= 0.621$). The NER level exhibited by the epithelial cells from these samples averaged $24.35 \pm 2.33\%$ of foreskin fibroblasts, whereas the fibroblastic cells from the same cultures exhibited an average NER capacity of $26.31 \pm 2.62\%$. The NER capacities of the epithelial and fibroblastic cells from the same individuals were correlated, though not significantly ($p= 0.0528$), suggesting that genetic factors may modify DNA repair activity similarly in both cell-types.

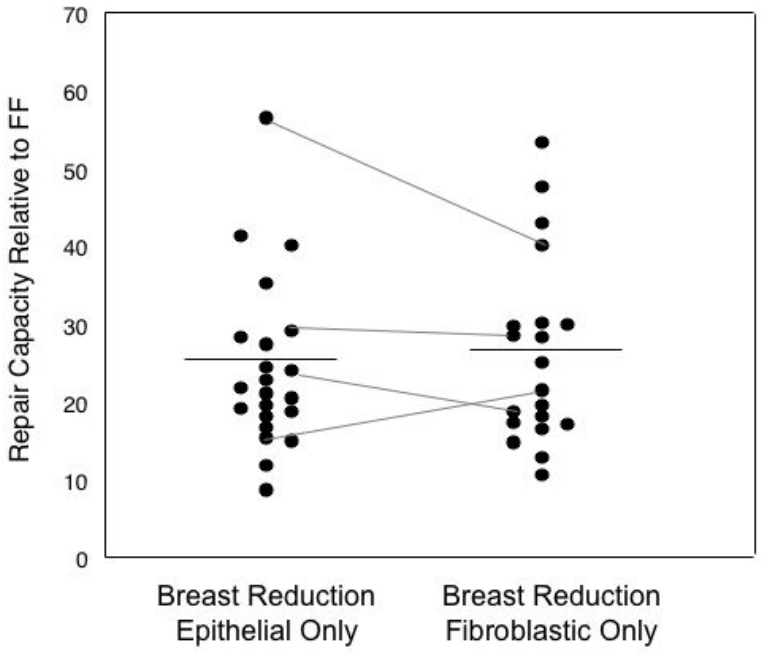


Figure 3.2 Breast reduction epithelial and fibroblastic cells do not differ in their NER capacity. No significant difference was observed in the NER capacities of cells exhibiting an epithelial versus a fibroblastic morphology. The mean of each data set is shown by the horizontal line. Samples from the same patient are shown connected by lines.

3.3.1.2 Effects of Age and Proliferative Index

Two factors have previously been proposed to affect DNA repair capacity: age of donor (328-330), and the mitotic activity of the cell sample (255,331). It has been observed that uninduced human somatic mutation increases with age when measured in a number of different ways

(332,333). It has been suggested that this is consistent with a generalized loss of DNA repair capacity with age (334), and some data on NER in fibroblasts and PBLs seem to support this hypothesis (335,336).

Analyses of data included in Latimer *et al.* (2003) did not support the hypothesis that increased age correlates with decreased repair capacity (140). We reported a significant positive association between NER capacity and age for the PBL data ($p= 0.01$). Breast samples were re-analyzed with the inclusion of the new sample, confirming that there was no association between age and repair capacity in breast ($p= 0.715$). Normal ovarian epithelium samples also did not show a statistically significant trend ($p= 0.6124$). It should be noted, however, that the ovarian epithelium data reflects only 6 data points with an age range of only 12 years (37-49). The average age of the ovarian samples was significantly different than that of the breast samples ($p= 0.0181$) with the ovarians (44.66 years old) being older than the breasts (35.90 years old).

The S phase index, *i.e.*, percentage of cells in the DNA synthesizing phase of the cell cycle, a measure of mitotic activity, has long been assumed to be positively correlated with DNA repair capacity (255,331). This assumption is based on the belief that in more rapidly dividing cells there is a greater requirement for DNA repair. The S phase index can easily be derived from the autoradiographic UDS assay. Cells in S phase incorporate much more label than non-S phase cells, regardless of whether they have been exposed to DNA damaging agents such as UV light. Indeed, the S phase of the cell cycle is the "scheduled" DNA synthesis alluded to by the "unscheduled" DNA synthesis of the UDS assay used to measure NER. Latimer *et al.* had previously found that NER capacity is not correlated with the S phase index amongst mammalian embryonic and extraembryonic lineages (247).

The average S-phase indices of FF, PBLs, breast, and ovarian tissues were compared with the average repair capacities of these tissues (Table 3.1). Repair capacities of each individual sample regardless of tissue of origin were also compared with their S-phase indices. When tissues were represented as a single point, with the average S-phase indices for FF, PBLs, primary ovarian cultures, and primary breast cultures being 25.11%, 4.83%, 19.25%, and 18.51% respectively, there was no significant association between repair capacity and S-phase index ($p= 0.249$). When all data points were included, however, a significant positive trend was seen between repair capacity and S-phase index ($p= 0.000317$) with an R^2 value of 0.182 (Figure 3.3), meaning that only 18% of the data is explained by this trend.

Multivariant analysis was also used to assess the relative contributions of donor age, s-phase index, and tissue type- FF, PBLs, primary ovarian cultures, and primary breast cultures, to the NER capacity. When all of the individual data points from each of these sample groups was considered, age and s-phase index were each significantly associated with NER capacity while there was little evidence for an effect of tissue type. We then attempted to generate a three-factor (age, s-phase index and tissue type) model with each of the sample groups represented as individual points. In this way, we could begin to account for the variability in sample size as we had when analyzing the effects of s-phase index alone. However, there were not enough data points (different sample groups) to carry out this analysis. We therefore carried out this analysis of each sample type with two-factor models: age and s-phase index, age and tissue type, and s-phase index and tissue type. Again, tissue type was not involved in any significant models. Thus, our analyses failed to show any evidence of tissue type systematically effecting NER capacity.

S-phase indices were also compared to repair within the breast and ovarian tissue populations. No significant associations were seen within the either the breast population (p= 0.657), agreeing with what had been published previously, or the ovarian population (p= 0.933). The new control sample from a disease-free breast reduction patient had an S-phase index of 23.1%, within the normal range. The average S-phase index was not significantly different between the breast and ovarian groups (p= 0.819).

Table 3.1 S-Phase indices calculated using the autoradiographic UDS assay in 4 different tissue types.

Tissue	**S-Phase Index**	**Repair Capacity Relative to Foreskin Fibroblast**
Foreskin Fibroblast	25.1%	100%
Peripheral Blood Lymphocytes	4.83%	9.02%
Breast Reduction Mammoplasty	18.5%	25.8%
Ovarian Epithelium	19.2%	21.0%

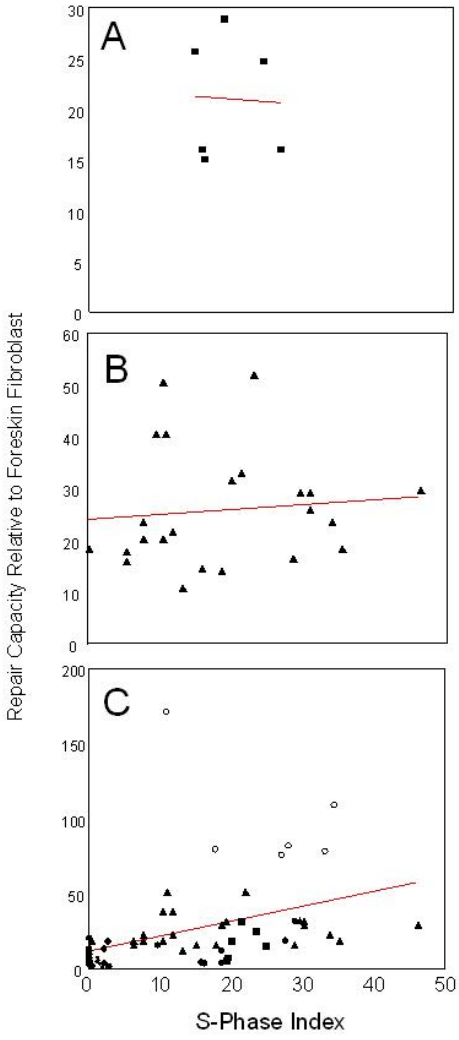


Figure 3.3 The relationship between repair capacity and S-phase indices. A. There was no significant relationship between the S-phase index and repair capacity of the 6 ovarian cultures analyzed. B. There was also no significant relationship observed between S-phase index and repair capacity in 23 breast reduction mammoplasty primary cultures. C. When all data points are considered: 6 ovarian primary cultures (▪), 23 breast reduction primary cultures (▲), 6 FF explants (○), and 32 PBLs (●), the trend becomes significant (p= 0.000317).

3.3.2 Genetic Breast Cancer- Analysis of a BRCA1 Mutation Carrier

In an effort to determine whether haploinsufficiency at the BRCA1 locus was associated with loss of NER, we studied the contralateral breast tissue of a BRCA1 mutation carrier who developed breast cancer.

3.3.2.1 Patient Description

The patient was a 35.7-year-old woman who presented with a very strong family history of breast cancer as depicted in Figure 3.4, and negative physical and mammographic examination. She had extremely dense breast tissue bilaterally by mammography as well as fibrocystic breast tissue by physical examination. She had no previous personal history of breast biopsy or abnormal mammograms. After genetic counseling revealing a strong family history of breast cancer, the patient elected to undergo DNA sequencing of the BRCA1 and BRCA2 genes, which revealed a Q1200X truncation mutation in one of her BRCA1 alleles. The C to T mutation at codon 1200 in exon 11 results in the change of the amino acid glutamine to a stop codon with resulting protein truncation and loss of function. Exon 11 is the largest exon in BRCA1 and has the highest frequency of reported mutations. The Q1200X mutation has been independently observed several times (337).

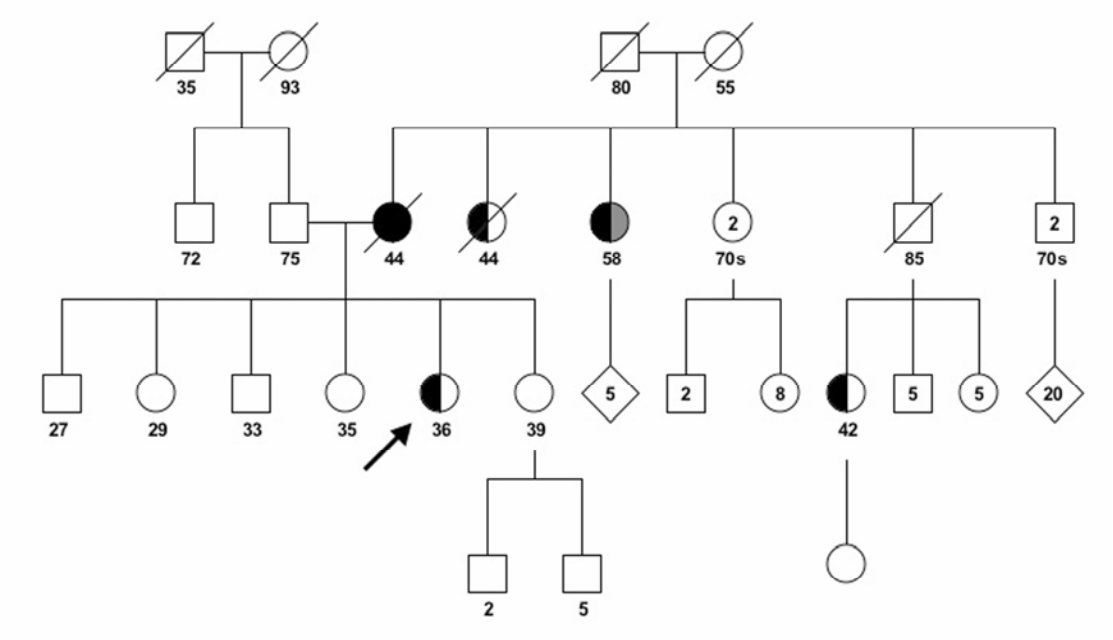


Figure 3.4 Pedigree of the BRCA1 mutation carrier patient (indicated by arrow). The patient, one maternal aunt and one maternal cousin had breast cancer diagnosed at 36, 44 and 41 years old, respectively, as indicated by the half-filled symbols, and her aunt died of the disease. Her cousin underwent lumpectomy followed by chemotherapy, radiotherapy and is presently on tamoxifen. Her mother had breast cancer in both breasts, diagnosed at ages 41 and 42, as indicated by the completely filled symbol. She underwent bilateral mastectomy and hysterectomy followed by chemotherapy and radiotherapy and died of the disease at age 44. A second maternal aunt was diagnosed with colon cancer at age 52 (light half-filled symbol) and breast cancer at age 55 (dark half-filled symbol). Based on this pattern of familial cancer the patient was considered to be at high risk of developing breast cancer and was entered into the low power MRI screening validation and feasibility study. Following her diagnosis, she was confirmed as carrying a Q1200X mutation in the BRCA1 gene.

A bilateral screening mammogram was compared to previous films from another hospital. The breast tissue was described as heterogeneously dense, thus lowering the sensitivity of the screening procedure. There were no masses, significant calcifications or other findings, and the mammogram was interpreted as negative bilaterally. However, when the patient was MRI scanned as previously described (338), with pre- and post-gadolinium enhancement, a 1 cm mass was seen in the upper-outer quadrant of the left breast (shown by ultrasound in Figure 3.5). Another lesion, this one approximately 1.5 cm, was seen in the right breast just above the nipple level medial and close to the chest wall. Core biopsy of the left breast revealed infiltrating ductal carcinoma in 2 of 5 core fragments with high nuclear grade and no lymphatic invasion seen. The core biopsy of the right breast demonstrated benign pathology, specifically, fibrosis with focal ductal epithelial hyperplasia.

The patient chose to undergo left modified radical mastectomy with left axillary lymph node dissection, contralateral prophylactic total mastectomy, and ultimately prophylactic bilateral salpingo-oophorectomy because of her genetic risk status. The pathology in the left breast was consistent with the imaging and core biopsy in size and description. Tumor size was 8 mm in greatest dimension, nuclear grade III, ER/PR and Her2/neu negative, and the nodal status (0/4) was negative (stage TIaN0M0).

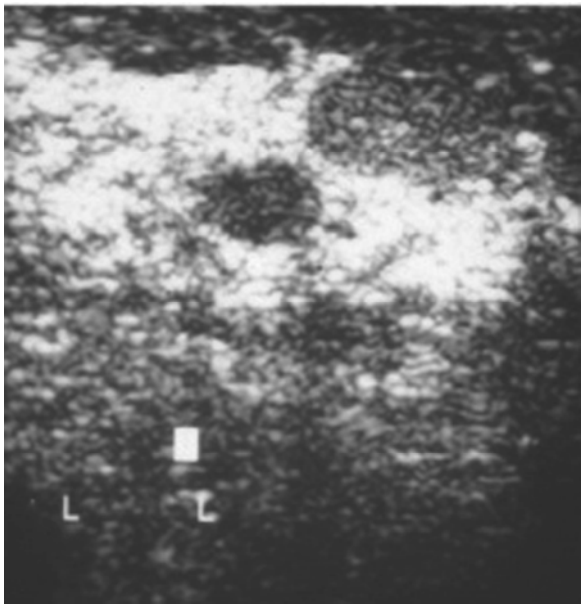


Figure 3.5 Ultrasound of the MRI-detected lesion. Following MRI, the patient was scheduled for ultrasound to identify the questionable lesions seen on MRI for possible core biopsy. Under ultrasound the lesion of concern was identified and biopsied at the 1:00 location in the left breast. Additionally, one lesion seen by MRI in the right breast at the 4:00 location was identified and biopsied.

Normal tissue from this patient, who is both a BRCA1 mutation carrier and has dense breasts, was evaluated to determine whether either of these factors affected *de novo* differentiation in our breast culture system. Four discrete pieces of fresh tissue were provided for live-cell analysis from each of the patient's ipsilateral and contralateral breasts. In the case of the ipsilateral breast, this tissue was provided at increasing distance from the tumor margin in 1 cm increments. All of these normal samples attached and grew in our culture system and were examined for cell-to-cell interactions and morphology over a period of one month. In the context

of breast reduction epithelium tissue samples from 23 patients with no breast disease and 1 with a contralateral sporadic tumor, these patient samples manifested typical mixtures of fibroblastic and epithelial cells. After several days in culture without passaging, the epithelial cells began to self-organize, initially forming three-dimensional mammospheres (Figure 3.6A), and, after 2 weeks in culture, more complex pre-ductal linear columns of epithelial cells (Figure 3.6B). The tissue explants from both breasts showed similar patterns of behavior (Figure 3.6).

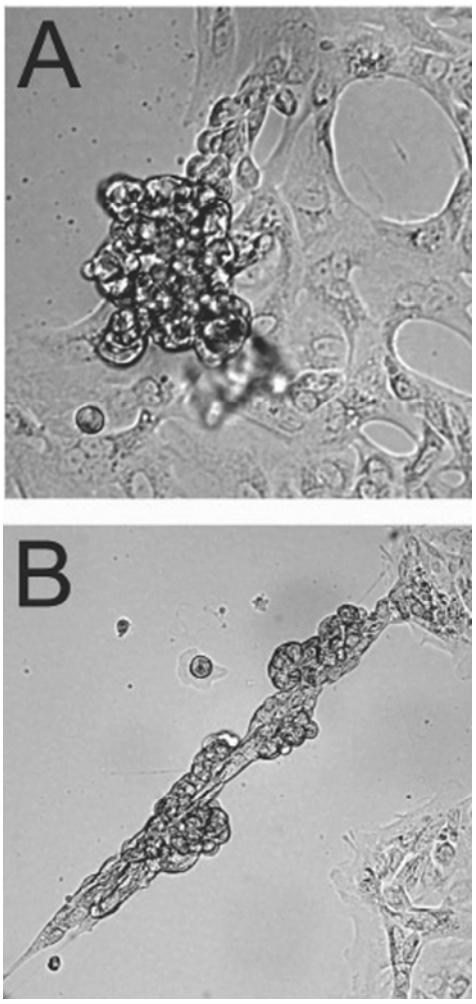


Figure 3.6 Micrographs of nondiseased primary human mammary epithelial cultures from the BRCA1 mutation carrier. A. Contralateral breast - A cluster of epithelial cells called a mammosphere is shown on the left center of the image sitting on a field of fibroblasts. B. Ipsilateral breast - The original tissue block from which this culture was derived was located 4 cm from the infiltrating ductal carcinoma. The structure shown is a cluster of rounded epithelial cells manifesting a column configuration called pre-ductal linearization. Both images were captured under DIC optics on a Zeiss Axiovert 100 microscope at a total of 140x magnification.

3.3.2.2 Cell Growth Kinetics

It has been suggested that the association between breast density and risk of breast cancer is due to increased cell proliferation (339). One measure of cell growth and viability is the S-phase index. As described above, the 23 normal breast reduction samples we observed demonstrated a wide range of proliferation rates, with S-phase index ranging from a low of 0.11% to a high of 46.0% (mean of 18.5 ± 2.47%) (140). We have also analyzed another sample of nondiseased breast reduction epithelium here, obtained from a woman with a previous contralateral invasive breast cancer. The average S-phase index of this sample was 17.0%. The ipsilateral and contralateral tissue samples from the hereditary breast cancer patient exhibited S-phase index of 26.6% and 26.2%, respectively, placing them at slightly over the 70th percentile for growth rate. Thus, all of these breast cancer patient samples appeared to grow well in our system, with S-phase index well within the range of our normal samples. The similarity of the S-phase index values from the two samples from the BRCA1 mutation carrier does not appear to be accidental; the chances of selecting two samples from the normal population with values as close or closer is very small ($p= 0.026$).

3.3.2.3 Functional Analysis of Nucleotide Excision Repair Capacity

PBLs and normal breast epithelial tissue from the hereditary cancer patient were then cultured for performance of the functional UDS assay. As our work has shown that repair capacity is tissue specific, from this poitn forward we will express repair caapcity relative to the tissue specific non-disease controls used in our laboratory. Analysis of cultured blood lymphocytes from the patient established that they had normal NER capacity (99.6% of the average of 33 normal samples) (Figure 3.7). This is well above the cut-off established in our sporadic breast

cancer population, < 65% average normal activity, which when applied to our cases and controls yielded a significant odds ratio of 21.6, as described further below. A trend towards age dependence had been noted in the analysis of the unscheduled DNA synthesis data of the normal controls (p= 0.059) (140); addition of the patient sample supports this trend, but it still fails to reach significance (p= 0.056).

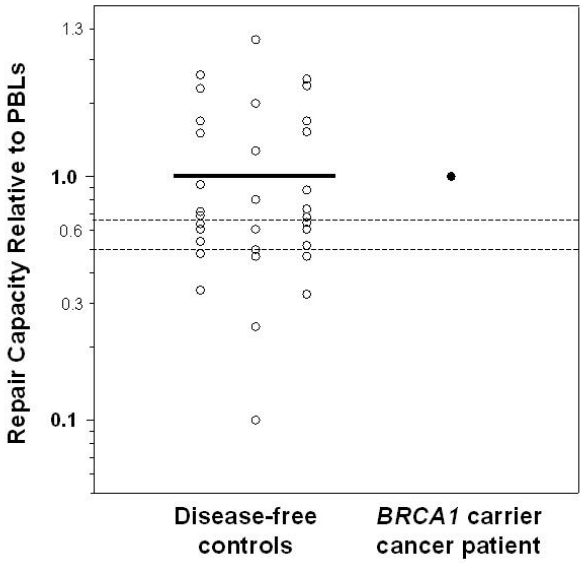


Figure 3.7 Comparison of the NER capacity of a peripheral blood lymphocyte sample from our BRCA1 mutation carrier patient with those of a population of disease-free controls. The dark horizontal line indicates the average for the normal population, while the dotted and thin lines indicate upper limits for residual NER activity in patients with the hereditary NER deficiency disease XP (0.50) and the cut-off established in our breast tissue study that identified tumors with high sensitivity and specificity (0.65).

The functional NER assay was then applied to one sample each from the ipsilateral and contralateral breasts of the patient. The NER capacity of the ipsilateral breast epithelial sample was 1.05 times the average of our population of breast reduction controls, clearly exhibiting no overt DNA repair deficiency (Figure 3.8). The contralateral sample was very similar, with an NER capacity of 1.17 times breast reduction normal. The NER capacity of the contralateral

sample from the sporadic breast cancer patient was 1.62 times the average of the previous breast reduction controls, also in the normal range.

Although the NER values of these two samples from the same patient are similar, they are not close enough to distinguish themselves as coming from the same individual ($p= 0.16$). The NER capacity of the contralateral sample from the sporadic breast cancer patient was 1.62 times the average of the breast reduction controls, also in the normal range. As with the normal breast epithelial cultures examined above, no effects of age or cell proliferation (as represented by the S-phase index) were seen. All of these additional patient samples are consistent with those results.

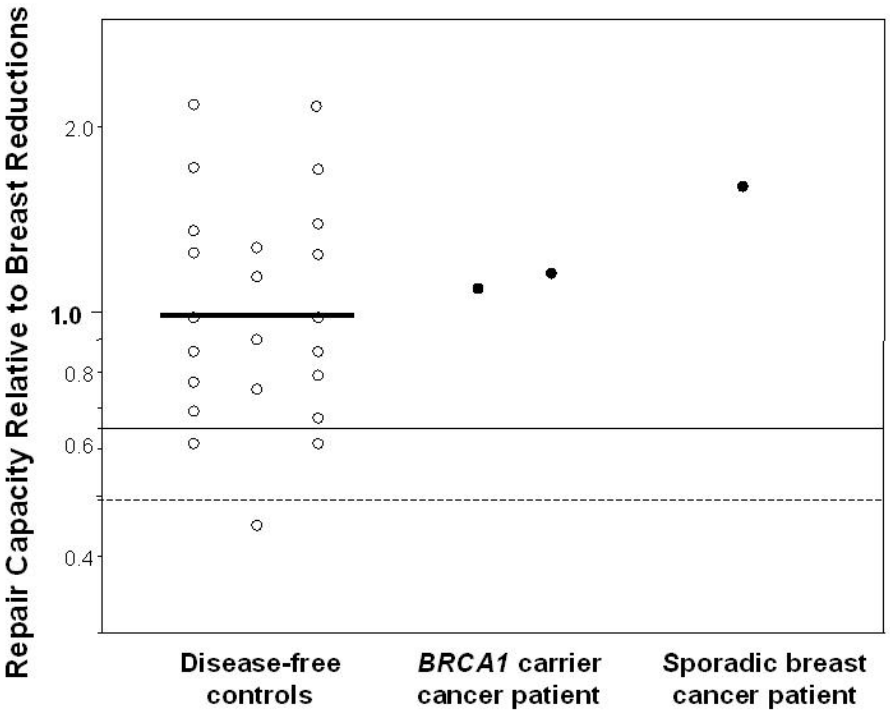


Figure 3.8 Comparison of the NER capacities of two samples of normal breast epithelium from our BRCA1 mutation carrier patient with those of a population of disease-free controls who underwent breast reduction mammoplasty. The dark horizontal line indicates the average for the normal population of breast reduction epithelium, while the dotted line and thin line indicate upper limits for residual NER activity in patients with the hereditary NER deficiency disease XP (0.50) and the cut-off established in our breast tissue study that identified tumors with high sensitivity and specificity (0.65). The patient sample on the left was derived from the ipsilateral (left) breast, while the sample on the right was from the contralateral (right) breast.

3.3.3 Analysis of Sporadic Cancer Cases

3.3.3.1 Early Stage Invasive Ductal Carcinomas

Samples of stage I and stage II breast tumors were obtained at Magee-Womens Hospital. These samples were subsequently fractionated, with a portion frozen, another embedded for confirmatory pathologic analysis, and a third processed for tissue culture according to the Latimer method (Figure 3.9). Within this group the rate of successful culturing for UDS analysis was 85%.

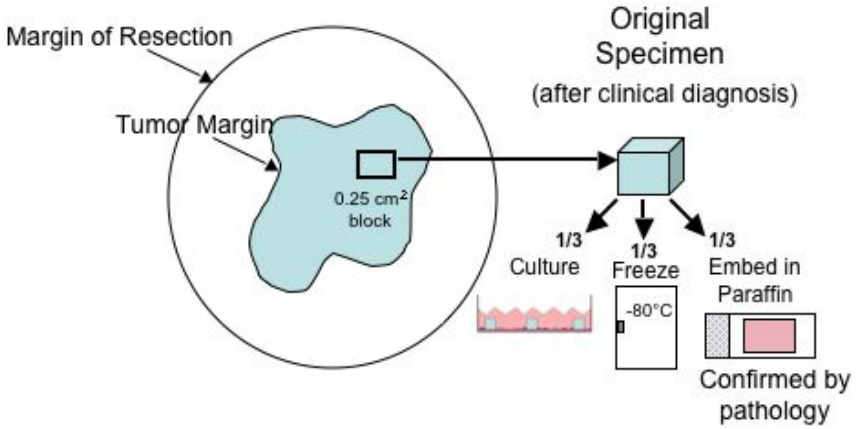


Figure 3.9 Fate of the tissue pieces received from the Tissue Procurement Facility at Magee Womens Hospital of UPMC. Tumor tissue blocks from lumpectomy or partial or complete mastectomy were anonymized and coded by an honest broker and provided to the laboratory within 24 hours of the surgery, but generally in fewer than 5 hours. Tissue is then separated into 3 parts, with one going into culture for analysis, 1 part frozen, and 1 part embedded in paraffin to be confirmed by a pathologist.

Tissue samples from 19 clinically stage I and 24 stage II invasive breast ductal carcinomas histopathologically confirmed to be free of adjacent normal breast epithelium were processed for unscheduled DNA synthesis analysis. Upon culturing, many differences were apparent between the normal and tumor samples. As discussed in Chapter 2, normal breast epithelium formed mammospheres and structures when placed into the Latimer culture system while tumors did not (Figure 3.10). Tumor cells under the same culture conditions in general assumed a more fibroblastic than epithelial appearance, despite the fact that they were derived originally from breast epithelium (Figure 3.10B). The tumor cells also tended to grow as highly

motile singleton cells lacking collaborative behavior with neighboring cells in the same area of the tissue culture dish. In addition, tumor cells in culture were confirmed using parameters used in clinical analysis of fine needle aspiration biopsies, *i.e.,* lack of cohesiveness, enlarged nuclear to cytoplasmic ratio and the presence of prominent nucleoli.

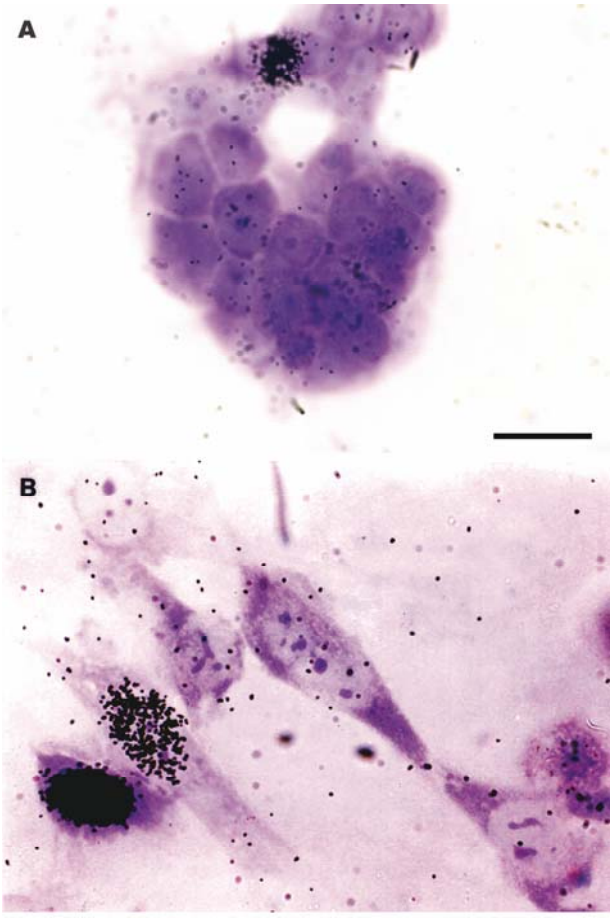


Figure 3.10 UDS analysis of primary nondiseased breast epithelial and tumor cultures. Photomicrographs (1000x bright field magnification) of primary epithelial cultures of A. normal breast reduction cells showing typical epithelial "mammosphere" architecture, and B. stage I breast tumor cells after processing for the UDS assay. Tumor cells manifest classical higher nuclear to cytoplasmic ratios, lack of cohesiveness and epithelial architecture and more activated looking or more prevalent nucleoli. Nuclei with dense silver grains are in S phase. The scale bar represents 4 μm.

3.3.3.2 Patient Demographics

None of the patients in this group reported a significant family history of breast or related cancers. None of these patients received pre-surgical chemo- or radiotherapy. Stage, grade, estrogen receptor status, progesterone receptor status, size of tumor, and age of patient are listed in Table 3.2.

Table 3.2 Patients included in the analysis of early stage tumors by UDS. Positive lymph nodes are expressed as the number of nodes that were seen to contain tumor cells over the total number of nodes investigated. ND = not done.

Patient	Stage	Pathologic Grade	Size, cm	ER	PR	Positive Lymph Nodes
1	1	2	1.2	+	+	0/15
2	1	1	1.5	+	-	ND
3	1	3	1.3	-	-	0/9
4	1	2	1.5	+	+	0/9
5	1	3	1.6	-	+	0/9
6	1	3	1.7	+13	+62	0/19
7	1	3	1.5	-	+	0/39
8	1	2	1.7	+++	-	0/21
9	1	2	2	+	+	0/20
10	1	3	1.5	++	++	0/17
11	1	3	1.2	+	+	ND
12	1	1	1.2	+++	++	ND
13	1	3	1.3	+	+	ND
14	1	3	1.6	+	+	0/21
15	1	2	1.5	+	+	ND
16	1	2	1.2	+	++	0/10
17	1	2	2	+	+	Sentinel 0/1 Axillary 0/4
18	1	2	1.2	+	+	0/6
19	1	3	1.2	-	-	0/19
20	2	3	3	-	-	0/15
21	2	2	3	+	+	6/19
22	2	3	2.8	-	-	0/9
23	2	3	2.9	+	+	0/15
24	2	2	1.8	+	+	1/22
25	2	2	2.2	-	-	1/21
26	2	3	2.3	+	++	5/19
27	2	2	2.2	-	-	0/22
28	2	2	2	+	+	ND
29	2	2	2.3	+	-	0/15
30	2	3	2.8	-	-	1/16
31	2	2	2.3	+/-	+	4/15
32	2	1	3.2	+	+	0/31
33	2	3	4.3	+	+	0/37

Table 3.2 (continued)

Patient	Stage	Pathologic Grade	Size, cm	ER	PR	Positive Lymph Nodes
34	2	3	3.1	+	+	1/13
35	2	2	3.8	+	+	0/17
36	2	2	2.5	++	++	ND
37	2	3	2.5	-	-	0/11
38	2	3	3.5	++	+	0/18
39	2	3	3.2	+	+	0/12
40	2	3	3.5	-	-	0/25
41	2	3	2.5	+	-	2/13
42	2	2	2.4	+	+	0/17
43	2	3	2.4	+	+	0/13

Lymph node status was confirmed in 35 cases by examination of from 6 to 39 axillary lymph nodes, and in 1 case by examination of both the sentinel node and 4 axillary nodes. Magee-Womens Hospital was in the process of developing sentinel node analysis at the time this patient received her surgery. One patient was listed as having no lymph node involvement without a total number of lymph nodes investigated listed (listed as "unspecified" in the pathology report). Seven patients declined axillary excision, so that lymph node status could not be confirmed. Clinically, five were treated as stage I tumors due to their size and the lack of clinical evidence for lymph node metastases by physical exam. These 5 patients were not significantly different from the confirmed lymph node negative population of stage I tumors with regard to tumor size, grade, hormone receptor status or tumor S-phase index, but they were significantly older (mean of 69.6 vs. 52.5 years, $p= 0.030$). Nevertheless, all statistical comparisons conducted on stage I tumor data reflect only the 14 patients with confirmed lymph node negative status; addition of data from the five unconfirmed lymph node patients generally

yielded the same results at a higher level of confidence, reflecting the greater power of the data set from the larger population. Two patients who did not have their axillary nodes evaluated were categorized as stage II because of their tumor size (2 cm and 2.5 cm) and the lack of distal metastases. Again, these data were consistent with what was observed from other stage II tumors. These patients were also older than the average of the stage II tumor samples (82 and 77 versus 53.36 ± 3.1 years). All patients, regardless of whether their lymph nodes were examined, were included in the total analysis of early stage tumors. Based upon tumor size and the lack of distal metastases in these patients even if all untested patients' axillary lymph nodes were positive, they would still at most be considered stage II.

3.3.3.3 Repair Capacity of Early Stage Sporadic Breast Cancers

The UDS assay was used to provide a functional assessment of NER capacity in these tumors relative to their tissue of origin: our established control breast reductions. Figure 3.11 shows the NER capacity measured in breast tumor and normal primary cultures expressed as a percentage of that of the mean of the 23 nondiseased breast reduction controls.

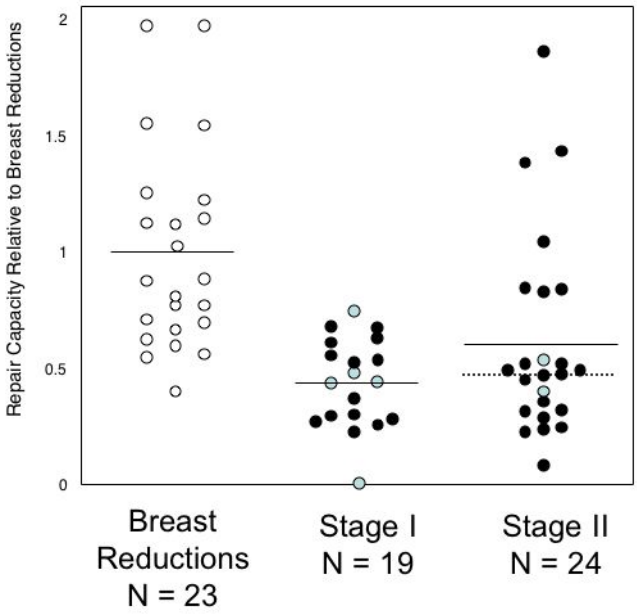


Figure 3.11 Loss of NER capacity in early stage breast tumors. Comparison of NER capacities of primary explant cultures established from normal breast reduction tissue (open circles) and stage I and II ductal carcinoma (filled circles). The patients who chose not to undergo pathologic analysis of their lymph nodes are represented with light blue circles. The average of each group is represented by the horizontal line. The dashed horizontal line for the stage II tumors is the average without the 4 highest repairing tumors.

The NER capacity of all the stage I tumor samples was significantly lower than that of normal breast epithelium, averaging only 44.6 ± 4.42% of normal activity ($p= 4.60 \times 10^{-6}$). Half of all tumor samples had NER capacities lower than the lowest normal epithelial sample. When only the confirmed stage I tumors were considered the average repair relative to breast reduction controls decreased slightly to 44.17 ± 4.53% and maintained its significance ($p= 4.88 \times 10^{-6}$).

The stage II tumors showed a wider distribution of samples as compared to the stage I tumors, with 4 tumors over the average of the breast reduction controls (ranging between 1.05 to 1.87 times the average of normal control). When all of the tumors were included, the stage IIs only exhibited 61.8 ± 8.90% the normal repair capacity (p= 0.00424). Without the four higher repairing stage IIs, the average repair capacity fell to 45.4 ± 4.64% ($p= 6.76 \times 10^{-6}$). The two samples for whom pathologic lymph node analysis was not carried out were still included in analyses of the stage II tumors since a the presence of lymph note metastases in the axilla still would have confirmed them as stage II. The 4 high repairing stage II tumors all may have contained a ductal carcinoma *in situ* component of their tumors, which could drive the repair capacity upward. Pathologic assessment for all stage II tumors is currently under way.

When compared as a whole to the breast reductions, early stage tumors (all sample points) had only 54.17 ± 5.44% of normal repair capacity (p= 0.000102), a trend that was augmented by the omission of the 4 high repair stage IIs (44.99 ± 3.17%, $p= 4.24 \times 10^{-6}$). Stage I tumors did not differ significantly from stage II tumors either when the high stage II tumors were included (p= 0.0931) or not (p= 0.901). This did not change when only the lymph node confirmed patients were included in the stage I group (p= 0.0732 versus 0.720).

3.3.3.4 Late Stage Invasive Ductal Carcinomas

Nine stage III or stage IV tumors were acquired and processed for examination of repair capacity by UDS. None of these patients had received either pre-operative chemo- or radiotherapy. These samples were collected between 1994 and 1996 when pre-operative chemotherapy was not a standard of care for late stage tumors as it is now. Additional samples cannot be obtained because of the change in treatment modality. As was seen with early stage tumors in culture, these samples failed to generate long term architecture. The patient demographics for this group are shown in Table 3.3. Two samples, samples 51 and 52 are from a patient who presented with tumors in both breasts and a scalp metastasis at diagnosis. It is not know whether this was a primary breast tumor with a contralateral metastasis or two independent primary tumors. Stage IV tumors were harvested from the tumors found in breast, not the distal metastases.

Table 3.3 Patients included in the analysis of late stage, chemotherapy-naïve tumors by UDS. Demographics of Stage III and Stage IV chemotherapy-naïve invasive ductal carcinomas analyzed by UDS. Positive lymph nodes are expressed as the number of nodes that were seen to contain tumor cells over the total number of nodes investigated. ND = not done.

Sample	Stage	Grade	Size, cm	ER	PR	Positive Lymph Nodes
44	3	3	7	+	+	2/7
45	3	3	6	-	-	7/11
46	3	3	3.6	-	-	7/17
47	3		5.5	-	+	1/13
48	3	3	3	+	+	8/16
49	4	3	6.5	-	-	9/12
50	4	3	5	-	-	7/12
51	4	3	1.7	-	-	ND
52	4	3	4.8	-	-	ND

3.3.3.5 Repair Capacity of Late Stage Tumors Relative to Early Stage Tumors

When chemotherapy-naïve late stage tumors (III and IV) are compared to early stage tumors (I and II), there is an increase of NER capacity in both the stage III and stage IV groups (Figure 3.12). Stage III tumors are statistically significantly higher in repair than stage I tumors ($p= 0.0303$) and stage II tumors whether the high repair stage II tumors are included ($p= 0.0462$) or not included ($p= 0.0309$). Stage IV tumors are never statistically significantly higher than either the stage I or stage II tumors with or without the high stage IIs ($p= 0.324$, $p= 0.573$, $p= 0.333$). When stage III and stage IV tumors are considered together as a group, the late stage tumors demonstrate significantly higher repair than the stage I ($p= 0.0182$), stage II ($p= 0.0420$), or the combined stage I and IIs ($p= 0.0291$). Again, this is not affected by the omission of the four high repair stage II tumors (stage II $p= 0.0190$, stage I and II $p= 0.0182$).

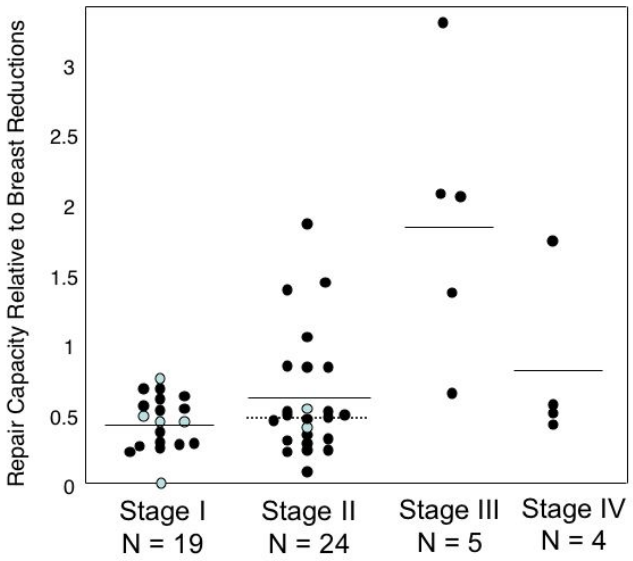


Figure 3.12 NER capacity is stage specific in breast cancer. Comparison of NER capacities of primary explant cultures established from stage I and II ductal carcinoma with stage III and IV ductal carcinoma. The patients who chose not to undergo pathologic analysis of their lymph nodes are represented with light blue circles. The average of each group is represented by the horizontal line. The dashed horizontal line for the stage II tumors is the average without the 4 highest repairing tumors.

3.3.3.6 Repair Capacity of Late Stage Tumors Relative to Breast Reductions

The repair capacity of chemotherapy-naïve stage III tumors does not show a loss of repair relative to breast reduction controls, but rather an increase to 189 ± 44% of breast reduction samples. This increase, however, is not significant due to the high variance within this sample ($p= 0.117$). In contrast to the stage III tumors, the chemotherapy-naïve stage IV tumors still showed a loss of repair compared to breast reduction at 81.5 ± 31.1%, which was also not significant ($p= 0.598$). When all late stage tumors are considered together, there is still an overall gain of function seen to 141 ± 32% of breast reduction, but again there is not statistical significance ($p= 0.250$). Thus, if analysis of repair capacity had been done without regard to the stage of the patient, the significant loss of repair in early stage breast cancer would not have been found.

3.4 DISCUSSION

3.4.1 Tissue Specificity

These data provide further evidence for tissue specific differences in NER capacity, suggesting that the human body is heterogeneous in its ability to deal with certain types of genotoxic insult. Since we have performed our study on primary cultures, the results are as representative as possible of the normal physiological state of the human body. We have also characterized the variability, or "range of normal", in NER capacity for PBLs, breast cells, and to a lesser degree FF and ovarian cells.

Previous studies have provided NER capacity data on human populations, one performed with a transfection-based assay of transcription-coupled repair (70), and another with an antibody-based assessment of global genomic repair (66). The transfection-based report assessing transcription-coupled repair in human lymphocytes was published as a field-test verification of a technique using a damaged reporter plasmid to assess transcription-coupled repair capacity in a type of skin cancer called basal cell carcinoma compared to controls. The control samples represented, in general, middle-aged Caucasians. Their population sample showed a range of phenotypes within this group that varied 3-fold in their ability to repair the damaged plasmid. Thus, they agree that there is a range of normal repair associated with control samples, even though this range is much greater in our data on global genomic repair (15 fold) for human peripheral blood lymphocytes. The further conclusion of this study was that the basal cell carcinoma cases compared to these controls were not significantly different in their transcription-coupled repair.

The second study also used lymphoblastoid cells this time from breast cancer patients and their sisters who did not have breast cancer. The antibody-based assay used is a method of analyzing global genomic repair after a genotoxic insult with benzo[a]pyrene diol-epoxide. This DNA damaging agent is able to form DNA adducts which are recognized and remediated by the NER pathway. An immunofluorescent antibody that binds to benzo[a]pyrene diol-epoxide DNA adducts is added to cells after they have been damaged, and the cells are allowed to repair their DNA. The immunofluoresence of bound antibody is measured before and after repair, with the difference representing the amount of benzo[a]pyrene diol-epoxide DNA adducts that were repaired. This study showed a greater than 4-fold range of repair in their control subjects, more in keeping with the range of transcription-coupled repair in lymphocytes and the 5-fold range

seen in our study of breast epithelium than the 15 fold difference we have observed with human lymphocytes. Nonetheless, all of these studies agree with the idea that there is significant interindividual variation in DNA repair capacity that must be considered. This study of sisters ultimately concluded that repair was decreased in cancer patients relative to controls.

The data presented here on tissue specificity is particularly important with respect to the study of carcinoma. Tissue specificity requires that nondiseased samples of the tissue of interest must be included as a control. For example, studies of breast cancer repair capacity must include breast reduction tissue to give the tumors appropriate context. With the 10-fold difference in repair capacities across different tissues sampled here, the loss or gain of repair in a particular tumor could be completely missed if a relevant tissue-specific control is not used.

Not only is it important to consider that different tissues in the human body have their own intrinsic repair capacity, they also have a degree of interindividual variability. These distributions suggest that a single control "normal" sample is insufficient to provide context for experimental results. Rather, the range of the population should be considered. Breast reduction cultures demonstrated a range of repair with samples showing less than half of the average repair capacity up to twice the average of the repair capacity shown for this tissue. For example, if a tumor with a repair capacity of 10% relative to the FF control was compared to the breast reduction with the lowest repair, it would be considered to have a normal repair (94% relative to this sample). If it was compared to the highest breast reduction, it would be considered to have deficient repair (20% relative this sample). However, if this sample is considered against the entire population of breast reduction controls, we see that it has a repair of 39% of the total population. Thus, the choice for a nondiseased control can have a great impact of your results.

One possible way that researchers have suggested to account for interindividual variation when studying somatic alterations is the use of isogenic, or patient-derived, controls. While this does for the most part account for interindividual variation, is also creates a problem. Since we have shown that there is tissue specificity, this isogenic control should ideally come from the same tissue. This will only be possible if there is a confirmed source of uninvolved or nondiseased tissue in the patient. In the case of the breast, this could be the contralateral breast. For PBLs, however, this would not be possible. Ultimately, it would require duplicate organs to be absolutely sure. The risk of subclinical involvement of the tissue or alterations including the entire organ preclude the assurance of uninvolvement within a single organ.

3.4.2 Putative Factors Affecting DNA Repair

3.4.2.1 Age

In this study we saw no consistent effect within a tissue of donor age, cell proliferation, or cell morphology on NER global genomic repair capacity as measured by the UDS assay. In many types of cancer, including breast cancer, advanced age is the greatest risk factor for development of the disease. This is consistent with a proposal that DNA repair capacities decline with age, allowing greater mutagenesis and therefore carcinogenesis. Previous studies using a transfection-based assay have shown a decline in the transcription-coupled component of NER with donor age in human dermal fibroblasts, lymphocytes, and transformed lymphoblastoid cell lines [46-48]. A more recent study has also shown that global genomic repair capacity is decreased with age in lymphoblastoid cells of normal controls (66). We, however, saw no effect of donor age amongst our breast and ovarian cell cultures, and a significant increase in NER

capacity with age of donor in our peripheral blood lymphocyte samples. Rather than a loss of functional repair with age, as would be expected if the integrity of the NER pathway was compromised over time, this increase of DNA repair in older donors may be associated with the known reduction in hematopoietic progenitor cells that occurs with age, increasingly compromising the diversity of the immune response (340,341). It may be that those cells with intact NER, and by extension competence in other aspects of DNA metabolism (transcription, replication) preferentially persist into advanced age. On the other hand, our assays were performed on an uncharacterized subset of peripheral blood lymphocytes that preferentially adhered to our Matrigel-coated slides and continued to metabolize under these conditions, and we cannot know how representative these cells are of the full lymphocyte repertoire available *in vivo*. A recent study by Goukassian et al (342) that examined the rate of removal of thymine dimers and 6-4 photoproducts in human dermal fibroblasts using adduct-specific antibodies reported data consistent with our study (loss of repair) in similar age groups for breast and peripheral blood lymphocytes. It is possible that the NER associated "age effect" consists of a dramatic post-natal loss of activity that plateaus early in adulthood and is then maintained, or even selectively improved upon, in later years.

3.4.2.2 Cell Morphology and Lineage

The lack of difference in NER capacity between cell types contributing to the breast samples demonstrates that adjoining cells may exhibit similar NER activity despite their different lineages, and also suggests that cell morphology does not significantly affect the unscheduled DNA synthesis assay. It also emphasizes how inter-related these two morphologies of breast

cells are, a point discussed further with respect to epithelial to mesenchymal transitions in Chapter 2.

The high NER capacity exhibited by skin fibroblasts relative to breast reduction fibroblasts (and also by breast tissue, ovarian tissue, and PBLs in general) suggests that NER is regulated at some point by real or potential UV exposure. The difference in skin versus breast fibroblasts in particular suggests that it is the tissue of origin rather than the morphology that is driving the associated DNA repair capacity. One potential reason for the high NER of skin fibroblasts is that this is a tissue that is either exposed to or has a much higher potential of being exposed to UV radiation, thereby requiring a greater capacity for dealing with UV-induced DNA damage.

To further compensate for difference that may exist between cells with respect to their intracellular nucleoside pools, we cultured cells for 7 to 10 days in medium that was supplemented with the nucleosides adenosine, thymidine, cytosine, guanosine, and the ribonucleoside uridine. This ensured that the intracellular nucleoside pools of all cells were filled to capacity and that the biosynthetic pathways of any of these cells did not have to be relied upon to produce nucleosides for either repair synthesis or DNA replication. The possibility that there are different inherent nucleotide pool sizes or different capabilities of the cells to fill their nucleotide pools exists, and is a topic for future investigation.

3.4.2.3 S-Phase Index

Comparing between tissue types, there is a suggestion of a correlation of high cell proliferation with high DNA repair. If all data points are considered individually regardless of tissue of origin, they yield a positive trend with good significance. In the original data published in our tissue specificity study, this trend was not observed. However, when the additional breast sample and the ovarian samples included here are added, significance was seen. This suggests that it is these new data points that may be driving the trend to significance. More cell types will have to be evaluated to determine whether this is a generalizable trend.

No evidence of an effect of cell proliferation on NER capacity was found within samples from the same tissue in the case of breast tissue. Ovarian tissue, however, showed a very slight negative trend, with faster replication cells showing a slight decrease in repair capacity. If confirmed, this result would support the "mitogen-mutagen" theory that hypothesizes that more rapidly proliferating cells are unable to repair their DNA as well as cells that are cycling more slowly. Attempts to rationalize the endocrine theory of cancer with the somatic mutational theory have hypothesized that highly proliferating cells should have lower DNA repair capacity, due to a shortened G1 cell cycle period (343,344), and similar arguments have been made as explanations for the observed relationship between mutation frequency and cell viability (345,346). This theory was developed prior to knowledge about cell cycle checkpoints. The more rapidly proliferating cells referred to in the theory were always transformed, which we now know causes them to lose the ability to pause and repair DNA damage before entering S-phase where they can convert DNA damage into mutations. Our data suggest that it is doubtful that physiologically relevant changes in proliferation would have much overall impact on NER

capacity. However, the effect of proliferation on repair capacity, like repair capacity itself may be different in different tissues.

Our results in the ovarian tissue showing a potential association between rate of proliferation and repair are similar to those of a molecular analysis of expression levels of five NER genes published by Cheng *et al.* (255). They found similar levels of gene expression in stimulated and unstimulated lymphocytes, but were able to distinguish the expression levels of rapidly proliferating tissues (defined as skin, breast, intestine, liver, testis, ovary, placenta, and prostate) from those of slowly or nonproliferating tissues or slowly proliferating tissues (adipose tissue, brain, hippocampus, muscle, spleen, and lung). In general, tissues that were more quickly proliferating had a higher level of gene expression. Analysis of functional repair, however, was never made. Assuming that mRNA levels for these tissues correlates with expression, these results are consistent with the trend we have seen with repair capacity increasing in cells with higher S-phase indices.

3.4.3 Tissue Specificity of NER Capacity and Cancer Risk

The importance of tissue-specificity studies can be illustrated by the fact that although individuals with XP have NER defects in peripheral blood lymphocytes and skin fibroblasts (347,348) they do not show heightened somatic mutation using the allele loss glycophorin A allele loss assay in their hematopoietic bone marrow cells (349), showing that there are phenotypes of XP related to DNA repair that are tissue specific. This brings into question studies whereby PBLs have been used as surrogates for many other types of tissues. Tissue specificity studies need to be performed on the **normal counterparts** of tissues that develop

cancer. Since PBLs are relatively easily obtained, the literature is full of epidemiological studies on PBLs of people with various types of cancer, comparing the repair capacity of their PBLs with those of normal controls. It is now possible, as we have shown, to perform this work directly on the tissue(s) and cell type(s) of interest. A similar approach was taken by Monnat and co-workers (350) who performed the HPRT somatic mutation assay on human kidney epithelial cells, and found a significantly higher baseline mutation frequency than had been established in PBLs. The very low NER activity that we have observed in PBLs probably reflects their terminal differentiation and low potential for transformation. Pertinent to differentiation, in a study by Hanawalt *et al.* differentiated neurons almost completely lose their ability to perform global genomic repair as compared to the undifferentiated neuroteratoma cells from which they were derived (72). The survival of these cells was not compromised, however. Rather, there was a concordant gain in transcription-coupled repair capacity. This study has not yet been extended to demonstrate a generalizable effect of terminal differentiation on somatic mutation rate.

Two previously reported epidemiological studies performed the UDS assay on PBLs from breast cancer patients (323), and from healthy women with first degree relatives with breast cancer (351). In both of these studies, the UDS assay was performed in the presence of hydroxyurea to artificially inhibit spontaneous DNA synthesis (S-phase cells). This allowed for the collection of results to be done by scintillation counting rather than autoradiographic analysis, although we have reservations about this technique.

Measuring NER by scintillation counting is problematic for several reasons. First, it does not allow for the potential influence of proliferation rate on the repair capacity of these cells. A much more physiologic comparison can be made when cells are allowed to proliferate as they do

in the human body. Second, this type of analysis does not allow for analysis of cell morphology, another potential influence on repair capacity- although this is not seen in our study of breast reduction mammoplasties. Finally, inhibition of DNA synthesis by hydroxyurea is not complete. It can allow cells to escape its effects and proceed into S-phase. Should this happen, these S-phase cells would contaminate the results for the sample, incorrectly raising the perceived repair capacity. An average repairing PBL in our system can have as few as 5 silver grains present over the nucleus whereas we estimate that there are about 300 over an S-phase nucleus, although these are impossible to quantify due to the lack of space between grains. Figure 3.10 shows examples of S-phase nuclei as compared to nuclei undergoing repair synthesis in cultured breast reduction and breast tumor tissue. If we had used scintialltion counting, therefore, it is likely that we would not have been able to appreciate these trends in repair. So, only a few S-phase or partial S-phase nuclei would be sufficient to mask the differences observed in our studies.

In any case, these studies concluded that NER capacity was reduced in the PBLs of breast cancer patients and individuals with affected first degree relatives ($p < 0.01$) compared to the lymphocytes of the age-matched control populations. Another, more recent study, using the antibody-based DNA adduct technique discussed above, also used lymphoblastoid cells as a surrogate for breast tissue when comparing sisters discordant for breast cancer (66). This group also found that there was a significant decrease in repair capacity in cancer patients relative to their sister controls. It was suggested from these studies that subtle inherited deficiencies in NER may be a factor in breast cancer etiology (323).

In contrast to the suggestion that this may be an inherited deficiency, our data suggest that nondiseased breast cells themselves have a constituitively lower repair capacity naturally

and not due to an inherited deficiency. Breast may represent a tissue particularly susceptible to damage from UV-mimetic chemicals due to an intrinsically low NER activity.

Our data suggests that breast and ovarian tissue may represent "at risk" or "fragile" tissues because of their constituitive low levels of DNA repair compared to foreskin fibroblasts. Thus, the low repair that is seen in breast and ovarian tissues may contribute to higher rates of mutagenesis, and possibly tumorigenesis. Neither of these tissues are constantly bombarded with UV light, the major source of skin and eye damage that is remediated by NER. They are, however, exposed to other damage such as that taken in from environmental sources: psoralens, bulky adducts formed in coal and petroleum combustion, and catechol estrogens are all potential sources. Over time, the relative inability to breast and ovarian tissues to deal with these insults could provide a pathway to carcinogenesis. Other bodily systems, *i.e.*, cell turnover and immune surveillance, may mediate this, resulting in the known incidence rates of both cancers.

The repair capacities of the nondiseased breast reduction samples are low enough in their repair capacities that they would result in a diagnosis of XP if they were compared directly to FF. The cut-off of 50% FF repair is used to diagnose xeroderma pigmentosum. All of our breast reduction samples and all of our breast tumor samples are below this mark (Figure 3.13). Thus again, the need for a tissue specific control is seen.

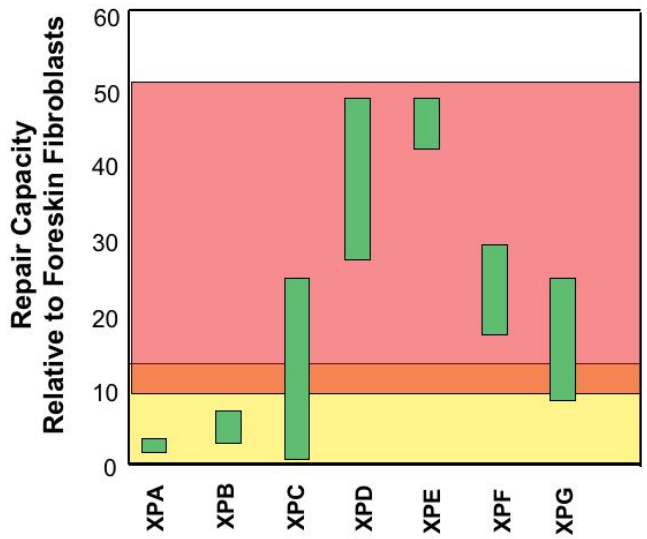


Figure 3.13 The range of NER phenotypes seen for different complementation groups of XP compared the range of repair phenotypes seen for breast reduction tissue and stage I breast tumors. Repair is represented here relative to FF. The repair range for each complementation groups is represented by the vertical green bar. The range of repair (from lowest repairing sample to highest) for breast reductions is shown by the horizontal red stripe and the range of stage I breast cancers is represented by the horizontal yellow stripe. The area of overlap between the two is orange.

Further evidence for the tissue specificity of cancer risk associated with repair capacity comes from transgenic mouse models. These models, in contrast with the human forms of the disease XP, exhibit a number of internal tumors as well as other phenotypes not seen in the human syndromes. A number of these models are exquisitely sensitive to liver damage, resulting

in tumorigenesis. This is not seen in the human disease. The eyes of mouse systems are also particularly susceptible to the disease phenotype: mice generate UV-induced corneal tumors or spontaneous loss of photoreceptors, again phenotypes not associated with loss of NER genes in humans. Different tissues in these mice respond differently to NER loss. While some of the differences between humans and mice may be species-specific, the idea of "fragile" or "at risk" tissues can be seen within the mouse transgenic population as proof of principle.

3.4.4 Normal Repair Capacity in BRCA1 Haploinsufficient Breast Tissue

Since the BRCA1 gene product is known to play a role in DNA recombination repair (317,318), it has been suggested that decreased repair capacity is the basis of the breast cancer predisposition observed in mutation carriers (352-355). Such a cellular phenotype has been difficult to demonstrate, however (356,357). With respect to NER capacity, haploinsufficiency for the BRCA1 gene does not necessarily result in detectable NER deficiency in nondiseased breast cells.

An alternate possibility is that the mutation affects the growth or differentiation of breast epithelial cells in a manner consistent with cancer susceptibility. It has been suggested that dense breast tissue is indicative of generalized hyperproliferation that might promote oncogenesis in accordance with the precepts of the mitogen mutagen theory where rapidly proliferating cells are less efficient in their DNA repair (339). Our findings do not support this theory in that all 8 samples, derived from both the involved and the uninvolved breasts of a hereditary breast cancer patient develop normal epithelial architecture *in vitro*, implying that the epithelial/stromal (paracrine) interactions necessary for the development of this complex

architecture are intact and normal in BRCA1 heterozygotes despite their greater risk of breast cancer. The S-phase index results also indicate that this nondiseased epithelial tissue falls into the typical range of normal for breast reduction control cultures and is demonstrating typical growth in the Latimer culture system.

3.4.5 Sporadic Breast Cancer is Not Associated with a Constituitive Loss of Repair Capacity

The observation that sporadic breast cancer patients have low levels of NER in peripheral lymphocytes suggests that sporadic breast cancer is associated with constitutively low levels of NER as discussed above. In contradiction to this idea, we have presented that the contralateral breast tissue of a woman with sporadic cancer had a repair capacity within the normal range seen with our system. This is similar to the data from the BRCA1 mutation carrier, where the contralateral breast is also within the range of nondiseased breast epithelium. The fact that loss is not seen constituitively in either the sporadic or the mutation carrier case suggests that it is somatic mutations within the breast itself that result in tumors rather than a constituitive loss of repair capacity in all tissues or specifically in the breast tissue.

3.4.6 Loss of Repair Capacity in Early Stage Tumors

Although loss of NER has long been associated with hereditary susceptibility to skin cancer, our study on early stage breast cancers is one of the first studies showing that it may also be involved in the development of internal tumors. Previous reviews have looked at XP patients under the age of 40 and seen a higher incidence of brain sarcomas, other brain neoplasms, and cancers of

the oral cavity. No evidence of an increased incidence of lymphomas or female genital tract tumors was observed (358). While it is well established that NER deficiency is associated with increased susceptibility to specific types of DNA damaging agents, especially UV light, it is not yet clear whether low NER capacity is sufficient to increase the rate of spontaneous mutation in internal organs not exposed to UV. *In vivo* measurements of somatic mutation in blood cells of XP patients have yielded conflicting results (78,349).

NER is a complex pathway requiring the products of over 30 genes, so it represents a large target for genetic or epigenetic modification. Even if exposures are necessary to reveal mutability in NER deficient cells, there is a broad range of lesions repaired by this pathway that provide candidate exposures. Although XP patients are known more for their manifestation of cancers in the skin and other tissues exposed to UV, internal tumors have been seen in this disease, which may represent an exposure to chemical carcinogens whose damage is normally remediated by the NER pathway (358). Their relatively low occurrence as compared to skin cancers may reflect that fact that patients often do not survive long enough for these internal tumors to be manifested (75,359). Transgenic mice with disruption of NER genes (discussed above) also manifest internal tumors (116).

All of our stage I breast tumor samples had NER capacities below the average of normal breast reduction epithelial samples. The proportion of all stage I breast cancer tumors considered to be "deficient" in NER capacity can be estimated from our data in several ways (Figure 3.13). Half of all stage I tumor samples exhibit NER capacities of less than 50% of the mean of the control population, sufficient to result in an XP phenotype (310). The highest NER capacity observed in our stage I patients with confirmed lymph node negative status was approximately 0.68 of normal; one patient with unconfirmed lymph node status was higher than this (0.75). If

0.65 times the average of the NER capacity of the breast reduction samples is established as a cut-off, the sensitivity of detecting tumors based on reduced NER levels alone is 79%, the specificity is 84%, and the odds ratio is 21.6 (95% CI, 10.8 - 43.2).

Thus, a large proportion of breast cancers can be considered to be deficient in NER capacity. Many molecular markers, especially activated oncogenes, have been investigated for their potential role in breast carcinogenesis and as predictive or prognostic factors, however, in contrast to our results, all but one of these markers appear to be present in 30% or less of all breast tumors (360). The exception is inactivation of the DNA damage-inducible gene 14-3-3 σ (360), which has been shown to be lost is 94% of breast cancers observed.

A comparison of nondiseased and tumor NER capacities from the same individuals will be necessary to differentiate between the two possible mechanisms that could be responsible for the observed tumor NER deficiency. First, breast tumors could be preferentially arising in women with constitutively low NER. This possibility is consistent with previous studies that found that DNA repair in general (361), and NER in particular (323,351), were lower in blood samples taken from breast cancer patients than from disease-free controls. This is not consistent with our own samples taken from the contralateral breasts of a sporadic and a hereditary cancer patient (Figure 3.8). The other possibility is that NER deficiency arises somatically during breast carcinogenesis. All of the known NER gene mutations are genetic recessives, and so should act like tumor suppressor genes during carcinogenesis (75).

Alternatively, there could be modulation of NER capacity though alterations in NER gene expression. Mutation in a gene long associated with breast cancer susceptibility, BRCA1, has recently been tied to repair capacity and regulation of expression of NER genes (319). When BRCA1 was overexpressed in osteosarcoma cell lines, it was able to increase the mRNA

expression of DDB2 and XPC and also increase the removal of cyclobutane pyrimidine dimers. This has yet to be shown in breast cells, but it shows precedent for the regulation of this pathway. It has also been shown that viral proteins inhibit NER activity (295,362) creating further precedence for the idea that this pathway can be modulated in cells that are not constitutively repair deficient (see also Chapter 4).

3.4.7 Late Stage Tumors Show a Wide Range of Repair Capacity

The wide range of repair phenotypes shown in late stage breast tumors reflects the heterogeneity of these stage classifications. With respect to AJCC staging guidelines, stage IV disease can reflect primary tumors of all varieties. They can be of any grade or any size. The only requirement is that they have acquired the ability to live outside of the breast (Table 3.4). The samples we harvested for stage IV tumors were taken directly from the breast, representing a population of cells that were "left behind" during the acquisition of traits necessary to leave the breast and survive in distal tissues. Therefore, they do not represent the most advanced form of the tumor in the human body. Consistent with this, preliminary data generated in the Latimer lab suggests that lymph node metastases have a higher repair capacity than their corresponding tumor in the breast. This relationship was significant for one sample and not for another. This is consistent with an increase in genomic instability in the metastasizing tumor population as compared to the original tumor.

Table 3.4 The TNM Staging schema developed by the American Joint Committee on Cancer and National Cancer Institute as revised in 2002.

Stage	**T (tumor size, invasiveness)**	**N (node involvement)**	**M (extent of metastasis)**
0	*In situ*	N0	M0
I	T1	N0	M0
IIa	T0, T1	N1	M0
	T2	N0	M0
IIb	T2	N1	M0
	T3	N0	M0
IIIa	T0, T1, T2	N2	M0
	T3	N1, N2	M0
IIIb	T4	Any N score	M0
	Any T score	N3	M0
IV	Any T score	Any N score	M1

3.4.8 Gain of Repair Capacity and Progression of Cancer

Our data shows that both stage I and stage II tumors demonstrate a low NER capacity, stage III tumors have a high repair capacity, and stage IV tumors have an low/normal repair capacity (an 18.5% loss of repair capacity relative to breast reductions that is not statistically significant). The change in repair phenotype observed between breast tumor stages can be interpreted in multiple ways, depending upon the view of breast cancer progression that is adopted. One theory of breast cancer is the **linear** view that tumors invariably begin as small aggregates of cells that acquire different traits as they progress. This can be thought of as the 1 → 2 → 3 → 4 model. This theory would suggest that early and late stage tumors exist as a continuum. With respect to repair function, it would mean that either tumor progenitor cells or early stage tumor

cells lose their repair capacity leading to the tumor's genetic instability, *i.e.*, the low NER in early stage tumors represents an intrinsic aspect of their etiology. This instability allows the tumor cells to acquire traits through increased mutagenesis, but are able to gain functional repair capacity back as they become stage III and IV tumors.

The second view of breast cancer is a less linear one, or the "1" → 4 model. The same basic traits are acquired by the transforming cells, but not necessarily in a progressive fashion. Some tumors may acquire the ability to home to other tissues and metastasize distally before they have even reached the size of a detectable mass in the breast. This has been seen clinically, with the first clinical presentation of a breast tumor varying from an enlarged lymph node to brain metastases. Small tumors that would be classified as stage I under the AJCC guidelines can still be highly aggressive, seeding cells into the circulatory system even as they are detected in the breast. Therefore, when looking at a stage I tumor, you cannot be sure whether it is a "1" as in 1 → 2 → 3 → 4, or a "1" → 4. With respect to repair capacity, it means that some stage I tumor with low repair can go on to become stage IV tumors without ever going through a phase where they regain their NER capacity seen in our stage III tumors. They may directly acquire the traits necessary to become metastatic through their genomic instability skipping the increase in repair capacity we have observed that goes on in the transition from 2 → 3. Just as we would expect in a heterogeneous disease that can display this range of clinical phenotypes, we also see heterogeneity in DNA repair capacity.

3.4.9 Implications for Chemotherapy

The NER pathway remediates DNA damage caused by a number of genotoxic cancer chemotherapeutic drugs, including cross-linking agents such as cisplatinum and cyclophosphamide (331,362), as well as bulky DNA adducting agents such as N-acetoxy-2-acetylaminofluorene (AAAF) and tamoxifen (363,364). Thus, our NER data suggest that the vast majority of early stage breast cancer should be particularly vulnerable to such agents. It may be useful to consider these findings in the evaluation of clinical trial data, or in the design of new chemotherapeutic regimen, as has been suggested for testicular cancer (365,366). These data also support the idea that early stage tumors can be treated efficaciously by chemotherapy, a practice that has become more common over the last decade.

At the same time, efficiency of repair demonstrated by late stage tumors may help to explain their higher recurrence rates. Few of these individual tumors demonstrated low repair. Micrometastases of tumors that possess the ability to repair genotoxic damage are more likely to persist in the body after a chemotherapy regimen has been completed. Furthermore, the increased success rates of combination treatments with other agents such as the microtubule-inhibiting taxanes could be partly described by this repair phenotype. Initial treatment with genotoxic cyclophosphamide and doxorubicin would provide a lethal dose of genotoxic damage to repair-compromised cells. Repair proficient cells, and deficient cells that had escaped the previous chemotherapeutic assault, would then be destroyed by taxane exposure.

A major area of chemotherapy research right now involves trying to identify which stage I and II tumors will benefit from additional chemotherapy beyond genotoxic chemotherapy. Clinical decisions are often being made in terms of what the patient desires, and what they can

medically tolerate. The benefits for additional treatment with an agent such as a taxane can be small compared to the potential side effects. Development of markers for repair capacity may help to stratify women as to their tumor's repair capacity, and subsequent vulnerability of their tumor to chemotoxic chemotherapy. Therapy could then be *individualized* for each patient based upon the profile of her tumor. Just as endocrine therapies are only given to women whose tumor are ER+ and trastuzumab treatments are reserved only for women who overexpress the Her-2/neu receptor, women with high repair capacity can be given additional non-genotoxic chemotherapies. Women whose repair capacities are low, and who would most likely not benefit greatly from additional therapy would be spared this treatment. This offers the potential for decreasing the morbidity while associated with cancer therapy while increasing cost-effectiveness and efficiency by ensuring only those who will benefit are treated.

4.0 NER CAPACITY OF ESTABLSIHED BREAST CELL LINES IN CONTRAST TO LATIMER-GENERATED EXPLANTS AND CELL LINES USED FOR MOLECULAR ANALYSIS

4.1 INTRODUCTION

4.1.1 Immortalization

The generation of cell lines begins when cells that were once part of a tissue *in vivo* are removed and placed into a culture system where they are temporarily maintained. These cultures are called explants. Primary explants are defined as the first culture generated after removal from the tissue of origin. Depending upon the cell type and the quality of the primary culture, these cells can be passaged by multiple techniques, including mechanical disaggregation, a method often used with embryo and extraembyonic culture, or enzymatic disaggregation such as is done with trypsin. These passaged cells represent secondary explants, and if passaging is continued, extended explants.

Typically, cells will complete only a set number of generations *in vitro* before they stop growing and eventually die (367). The *in vitro* process whereby cells permanently stop dividing is called senescence. The combination of senescence with associated changes in morphology, including flattening of cell and increased vacuolization (368), is referred to as “crisis”.

Molecularly, this is characterized *in vitro* for human skin fibroblasts by the progressive shortening of telomeres and the activation of the TP53 and retinoblastoma tumor suppressor pathways (367). It is also characterized by apoptosis and widespread chromosomal instability secondary to the shortening of telomeres. As DNA replicates, telomeres progressively shorten, ultimately resulting in unprotected DNA ends that can cause end-to-end fusions. Explants are therefore of use for only a limited time while they are still capable of growth.

Subpopulations of cells, however, can acquire the ability to survive this crisis period and continue growth, a process known as immortalization. Once past crisis, the population of cells is referred to as a cell line. The mechanism of immortalization past crisis are currently under investigation, as described below.

A qualification to this usage of the term "cell line" must also be made with respect to normal breast reduction cultures. Martha Stampfer has suggested that primary culture should only refer to the cells the first time they are placed in culture (*i.e.*, tissue outgrowths as presented in Chapter 2) (369). Cells which have been subcultured are no longer primary cultures. She is careful to note that "while in technical cell culture parlance, they could be called cell lines once subcultured, but I find this usage confusing and only use cell line to refer to cells with indefinite growth potential." Rather than cell lines, she calls high passage cultures of normal cells "strains with long-term growth potential." It is unknown how this quality can be predicted *a priori*.

Spontaneous immortalization is a rare event, and thus, many techniques for inducing immortalization have been introduced (367). Methods that are still used in common practice include transfection with oncogenic or vial DNA such as SV-40 large T-antigen, the HPV E6 protein, and the Epstein-Barr virus. Indeed, these methods go beyond simple immortalization and into cellular transformation, a distinction that is discussed further below. The intent of these

methods is to block the TP53 and retinoblastoma tumor suppressor pathways required for senescence. More recent advances have been made with the use of transfection with the catalytic subunit of telomerase to prevent the shortening of telomeres with successive generations.

4.1.2 Immortalization of Human Mammary Epithelial Cells

The paradigm for investigation of senescence and the molecular events surrounding it in human cells has largely been based on observations of human skin fibroblasts. This includes the proposed mechanisms described briefly above regarding the TP53 and retinoblastoma tumor suppressor pathways and telomere length. Epithelial cells, however may not follow the same patterns.

One model proposed by Tlsty *et al.* is that human mammary epithelial cells go through two periods of crisis instead of one (368). Observation of 5 different breast epithelial cell cultures showed that after an initial period of growth, the growth rate of human mammary epithelial cells reached a plateau. Some cells "escaped" the plateau and continued growth for a time, only to reach a second growth plateau. The first growth plateau is called stasis and cannot be reversed by the ectopic expression of telomerase, distinguishing it from the senescence characterized in human fibroblasts (369).

Escape from the first plateau is proposed to be due to reduced p16INK4A expression. p16INK4A is a necessary factor in the retinoblastoma tumor suppressor pathway. Compromise of this pathway is one path to immortalization. Cells are proposed to decrease expression of p16INK4A through epigenetic silencing. Once this has occurred, cells can grow to the second

plateau, called agonescence, caused by progressive telomere shortening (369). In contrast to the first crisis, or growth plateau, this can be overcome by ectopic telomerase expression.

Freshney, however, has postulated that this survival through spontaneous immortalization in epithelial cell populations is due to the presence of stem cells (370). While he offers no data to support this hypothesis, it is an interesting one. Multiple studies are currently ongoing to identify stem cells in normal mammary epithelium, with multiple laboratory groups using multiple methods to attempt to isolate them (272,273). Isolation after flow sorting with multiple markers has yielded cells capable of continued growth in culture. The contribution of these putative stem cells to the continued growth exhibited by some breast reduction epithelium cell cultures has not yet been evaluated. The work of multiple other groups, including Wicha, Bissell, and Gudjonsson, all support this argument (272,273,371).

4.1.3 Effects of Common Immortalizing Agents on DNA Repair

Multiple agents commonly used to transform cells for immortal growth in culture are specifically known to interfere with DNA repair. These agents include the Human Papilloma Virus E6 protein (292), SV40 large T antigen (290,291), Epstein-Barr virus (372) and telomerase expression (293). Expression of the Human Papilloma Virus E6 protein specifically inhibits the ability of cells to remediate thymidine dimers after UV-exposure and therefore NER (373). Fibroblasts immortalized with it are UV-sensitive and display decreased global genomic repair (295). Similarly, cells transformed with SV-40 are unable to repair cyclobutane pyrimidine dimers (362). Epstein-Barr virus has not been evaluated for its specific effects on NER.

Telomerase constitutive expression has been shown to produce cell lines with phenotypes relevant to their tissue of origin without the induction of a transformed phenotype (291,374). Furthermore, recent publications have shown that there is no deficit in the ability of telomerase-transformed cells in their ability to deal with cyclobutane pyrimidine dimers (375). They are also efficient in global genomic repair of UV-induced damage (376). This would seem to suggest that this may be a promising method of cell immortalization for the study of NER, but it is important to remember that this is still an exogenous transforming agent. While the full functional capacity may still be present as determined in these two studies, other effects may have taken place that are as yet not described. Overexpression of the catalytic subunit of telomerase, hTERT, in human FF has been shown to abrogate the cellular response to double strand breaks, and radiosensitize these cells, indicating that telomerase overexpression is not benign with respect to all types of DNA repair (377).

When SV-40 large T-antigen, the telomerase catalytic subunit, and an H-ras oncoprotein are introduced into primary mammary epithelial cells they become transformed and tumorigenic in mice (378). While the combined activity of three immortalizing agents was necessary to create this transformed phenotype, it is important to remember that all three were needed and contributed to transformation. Though the specific effects of telomerase on DNA repair do not seem to be substantial at first glance, they are known to make contributions to transformation. The use of 3 agents of transformation to immortalize cells in which you want to study the process of malignant transformation as it occurs in the body seems extreme. The resulting cell lines are not always stable and can have limited growth, ie, it does not always work.

Mutagenesis with benzo[a]pyrene has also been utilized to induce cells into an immortalized state (cell lines 184A1 and 184B1) (294), although this is a much more nonspecific

method than the ones discussed immediately above. Since NER remediates lesions of benzo[*a*]pyrene, this method is specifically unadvisable for creating cell lines used in the study of DNA repair. Because it is an agent known to cause malignant transformation, it is also unadvisable to use these cell lines to study benzo[*a*]pyrene-independent modes of cancer etiology.

4.1.4 Transformation versus Immortalization

Immortalization in cell culture does not necessarily constitute *in vitro* malignant transformation. *In vitro* malignant transformation is a multistage process requiring multiple steps. Freshney describes it as "the event by which a cell line undergoes a critical change that endows it with an infinite lifespan and increased growth rate (367)." The first step in Freshney's scheme, which was later reiterated in a work by Madsen *et al.*, is immortalization (379). The next is a change in growth control to generate more rapidly dividing cells that have a higher saturation density and possibly, anchorage-independent growth. Thus, immortalization and malignant transformation are two different phenomenon, with immortalization existing as a component of *in vitro* malignant transformation. Oncoproteins like the SV-40 large T-antigen and HPV E6 protein are involved in the process of immortalization, and not necessarily transformation. Implicit in this idea is that such transformed cells will manifest genetic instability and abnormal karyotypes.

Tumorigenicity, on the other hand, is a separate consideration. Immortalized cell lines are not necessarily tumorigenic, nor are transformed cell lines. This phenotype, the ability to form tumors in immunocompromised mice, is an independent descriptor of a cell line and the gold standard for tumorigenicity. For example, the MCF10A cell line, derived from a breast

reduction mammoplasty, demonstrates continuous growth in culture, but is not tumorigenic in mice (380), *i.e.*, it is immortal but not transformed. MCF7/Ly2, a subpopulation of a cancer cell line derived from a malignant stage IV pleural effusion secondary to invasive ductal carcinoma of the breast, however, is not capable of forming tumors in nude mice (381), while its parental line can do so when supplemented with estrogen. Both of these cell lines are discussed further below.

4.1.5 Establishment of Breast Tumor Cell Lines

Studies of breast cancer typically utilize *in vitro* models such as the well-established breast cancer-derived cell lines: MCF-7, MCF-7/LY2, CaMa-1 MDA-MB-231, BT-20, and SK-BR-3 (314). These and other cell lines provide a renewable resource for proteins, nucleic acids, and cells, and therefore the means to accomplish many molecular and biochemical studies in which large cell numbers are required. While these experimental systems have provided important advances in breast cancer research, it is important to be aware of their limitations.

Breast cancer cell lines are difficult to establish in culture (382), with a historical success rate of only approximately 15% for tumors (300). These tumors may, in fact, represent a small subset of the original tumors that were particularly permissive for *in vitro* growth (*i.e.,* 15% of all breast tumors). Because of the difficulty in establishing cultures, the majority of human breast cancer cell lines have been initiated from tumor metastases, and in particular pleural effusions (383). Twenty percent of the 50 breast cell lines available from ATCC have been cultured from malignant pleural effusions. Pleural effusions were considered advantageous by those who developed cell lines because of the large number of cells that could be collected, the lack of

contaminating fibroblasts, their ability to demonstrate anchorage independent growth *in vivo*, and the possibility of sequential samples from the same patient (384). However, these metastatic cancer cells may not always be representative of their tumor of origin.

4.1.6 Description of Commonly Used Breast Cancer Cell Lines

Perhaps the single most studied and best-characterized breast cancer cell line is MCF7. This cell line maintains its expression of the estrogen receptor due to an over 100 fold amplification of the ER gene. The mitogenic effects of estrogen and the inhibitory effects of antiestrogens were first defined in these cells by basic scientists and the pharmacology industry (385). The MCF7/Ly2 derivative variant has been selected for resistance to the antiestrogen LY1170718. The line remains partially responsive to estrogen, making it an important model of the progression of many cases of human breast cancer (381). It is also considered the most stable of the antiestrogen-resistant variants. CaMa-1, while not as widely used as MCF7, is another estrogen-dependent breast cancer cell line (386). The expression of ER in these cells in nowhere near the level of overexpression seen in MCF7. This cell line has been used for its known expression of the androgen receptor.

MDA-MB231 is among the most widely used ER- breast cancer cell lines and is often used in labs as a negative control for the endocrine regulation of breast cancer cell growth (383). BT-20 was the first human breast carcinoma cell line, derived in 1958 (298). In stark contrast to many other cell lines in wide use, BT-20 was derived directly from an intraductal carcinoma and not from a metastatic growth, using physical disaggregation methods similar to those of Latimer et al (described in the methods of Chapter 2). While it expresses ER, the receptor has a deletion

in exon 5, rendering it nonfunctional. SK-BR-3 is a popular cell line because of its overexpression of HER2/c-erb-2 at 128 fold over normal fibroblasts, and the secretion of truncated c-erb into its culture medium (387). The ultrastructure of these SKBR3 cells has been well characterized using transmission electron microscopy (388).

Table 4.1 Summary of features of commonly used breast cell lines.

Cell Line	Tissue	Patient	Morph-ology	Karyo-type	Receptors Expressed	Tumorigen-icity
MDA MB231 (383)	Breast adeno-carcinoma metastasis: pleural effusion; treated with steroids, CAF, CMF, oophor-ectomy	51-y.o. white female	Epithelial, adherent	Hypo-triploid	TGFαR+, ER-, PR-	highly tumorigenic, metastasizes to lung in nude mice
MCF7 (385)	Breast adeno-carcinoma metastasis: pleural effusion; treated with hormones & radiation	69-y.o. white female	Epithelial, adherent	Hyper-triploid / hypo-tetraploid	ER+++, PR+, androgen, LHRH, gluco-corticoids, insulin, RARα and γ, prolactin, IGFs, TGF-α, EGF, FGF, PDGF	Estrogen-dependent for growth; low metastatic potential-considered as an "early" breast cancer phenotype
MCF7/ Ly2 (381)	Stable variant of MCF-7 resistant to antiestrogen LY117018, still partially responsive to estrogen	69-y.o. white female	Epithelial, adherent		ER+, though less than the parent line, PR-	blunted response to estrogen, cannot form tumors even when given estrogen in mice
BT-20 (298)	Infiltrating ductal carcinoma of the breast	74-y.o. white female	Epithelial, adherent	X may be absent, normal 3, 4, 9, 13, 14	ER-: has a deletion in exon 5, PR-, TGFαR+, gluco-corticoid, amplified EGFR	tumorigenic, but non-metastatic in athymic nude mice

Table 4.1 (continued)

Cell Line	Tissue	Patient	Morph-ology	Karyo-type	Receptors Expressed	Tumorigen-icity
CaMa-1 (386)	Breast adeno-carcinoma metastasis: pleural effusion	51-y.o. white female	Epithelial, adherent in multilayer colonies	Hyper-triploid	ER+, TGFβR-	
SK-BR-3 (387)	Breast adeno-carcinoma metastasis: pleural effusion	43-y.o. white female	Epithelial, adherent	Hyper-triploid	ER-, over-expresses HER2/c-erb-2 128-fold over normal fibroblasts, EGFR	
MCF-10A (297)	Fibro-adenoma of the breast from a woman with fibrocystic disease	36-y.o. white female	Epithelial, adherent	Normal diploid female except for a stable t(3:9)(3p 13:9p22	ER-	called either a "normal human breast epithelial cell line" or a "non-tumorigenic cell line"

Some known limitations of human breast tumor cell lines include the relative lack of ER+ lines. In order to be able to develop and screen new therapeutic agents and to determine the mechanism of action of current endocrine therapies *in vitro*, models of breast cancer that exhibit an endocrine response are necessary. While most breast cancers that are diagnosed are ER+, only three tumor cell lines that are commonly studied, MCF-7, T-47D, and ZR-75-1, are estrogen responsive. Since the loss of ER expression in a common phenomenon in the culture of breast and breast tumors it has been assumed that the cellular loss of "*in vivo* context" is related to this phenomenon (389).

Tumors that are ER- are generally faster growing (390), more aggressive, and associated with a poorer prognosis (391,392). These characteristics may make ER- tumors more permissive for *in vitro* growth, *i.e.*, less breast specific juxtacrine and paracrine signal are required, and explain their greater representation in the tumor cell line population. Alternatively, since we do not actually know what the ER status of the original tumors were (385,393), it is also possible that current culture systems are not consistent with the maintenance of the estrogen receptor and the estrogen-dependent phenotype.

Human breast tumors are known to be made up of heterogeneous populations of cells with different phenotypic characteristics (394). While tumors may arise from a single mutated cell, as that cell divides into daughters cells these progeny acquire new traits that favor tumor growth in different ways. These traits may include more complex chromosomal abnormalities, decreased dependence on growth factors, the ability to grow in new environments, the ability to recruit vasculature, and the secretion of proteins that can affect the host's metabolism. These new characteristics are acquired over time and possibly independently of one another.

We have assumed that cell lines may be more homogeneous populations and treated them as such. While sometimes homogeneity is intentionally created through single cell cloning, it may also be achieved through passaging cultures and clonal evolution over a long period of time. The cells most suited to *in vitro* culture conditions will eventually out-compete other, less permissive clones. While these cell lines may not fully reflect the complexity of a human tumor, they do allow the thorough investigation of a tumor subpopulation (314) through the use of critical masses of DNA, RNA, and protein.

Many breast tumor cell lines were derived from women who had received either radiotherapy, chemotherapy, or hormone therapy prior to surgery. Thus, the cell lines are not therapeutically naïve and probably represent a subpopulation of cells that have evolved *in vivo* to be resistant to these treatments.

While these cell lines have been created and subsequently used to generate reproducible sources of cells, DNA, RNA, and protein, their phenotypes are not always necessarily stable. The intrinsic genomic instability demonstrated by tumor cells in general may also be present in these cell lines, causing changes over time. Investigators have shown that certain biologic behaviors, such as response to 17β-estradiol, is different in four different stocks of MCF-7 cells from different stock sources (395). Furthermore, another recent study showed that two substrains of MCF7 had 10-fold differences in mutation frequencies (396). Thus, further changes have occurred in the MCF7 cell line over the years in culture. This challenges whether the results from different laboratories using MCF7 cells can be always directly compared since it is unknown just how alike their two stocks of MCF7 are.

4.1.7 Establishment of Nondiseased Breast Cell Lines

Few immortalized non-malignant human mammary epithelial cell lines exist to answer questions about normal breast epithelial cell biology and provide controls for analysis of these established breast cancer cell lines. Existing lines include HBL-100, HMT-3522, and MCF-10A and F (314). HBL-100 cells were originally derived from a presumably normal source of mammary tissue, the milk of a 27-year old apparently healthy woman (397). However, later analysis showed that this line contains an integrated SV40 virus genome and is tumorigenic in nude mice (398). HMT-3522 were generated from fibrocystic breast tissue in a chemically defined cell culture medium (399). HMT-3522 cells were nontumorigenic in immunocompromised mice, expressed epithelial cytokeratins as its intermediate filaments, and showed monolayers of polarized epithelium by electron microscopy. They were also characterized by a lack of expression of the ER and a diploid karyotype with marker chromosomes.

The MCF-10A cell line was created from cells isolated from a woman with fibrocystic disease after she underwent a mastectomy (297). After 849 days in culture these cells spontaneously immortalized, giving rise to the MCF-10A (attached) and MCF-10F (floating) cell lines. The cells resemble luminal epithelium in their ultrastructure, specifically the presence of desmosomes and microvilli, and by their expression of several keratins detected by immunocytochemistry (297). Their original karyotype was a normal female diploid except for a stable t(3:9)(3p13:9p22) translocation. We, however, have show that the karyotype of MCF10A cells in our laboratory is 48 XX, del(3)(p13), der(6) t(6;12)(p26;q22), +8, der(9) t(7;9)(q11.2;p24), +16 (Figure 4.3), even though these cells were directly purchased from the Michigan Cancer Foundation. The original MCF-10A cells are nontumorigenic in nude mice

and do not exhibit anchorage-independent growth. Even after being transfected with erbB-2 and Ha-ras oncogenes, this line is still not tumorigenic in mice (380). MCF-10A is widely used as a "normal" nontumorigenic control. While this cell line has been derived from a benign origin, it has been hypothesized that the immortalization of this line constitutes the initial step in malignant transformation (370,379).

Despite the caveats associated with these cell lines, they are important experimental models, and once their limitations are acknowledged, they are still powerful tools. They have been used to elucidate the cellular and molecular events associated with endocrine responsiveness, malignant progression, invasiveness, and metastatic potential (314). They can also provide continuous sources of cells for experimental manipulation, DNA, RNA, and protein. For this reason, we evaluated whether these cells lines could be used as viable surrogates for the breast reduction (MCF10A) and breast tumor primary cultures that were evaluated in Chapter 3. If these cell lines were able to reiterate the repair capacities of the primary cultures, we would be able to use them for further molecular characterizations of the phenotype of decreased DNA repair capacity in early stage breast tumors.

Although there have been 50-70 breast cancer cell lines described in the literature, only approximately 20 have been adequately characterized and are commonly used. Since use of all these 50-70 cell lines is beyond the scope of this project, we instead focused on a subset of these 20 widely used cell lines and a variant of one of them. These lines were chosen because they represent some of the major classifications of breast cancer cell lines. Some are ER+ (MCF-7 and CaMa-1) while the others are ER-. Some have been derived from a solid tumor in the breast (BT-20), while the others are from malignant pleural effusions. Others exhibit clinically relevant breast cancer phenotypes such as the antiestrogen independence of MCF7/Ly2, and the

overexpression of HER2/c-erb-2 (SK-BR-3). And finally, MCF10A was included as an example of nondiseased breast epithelial line used as a control.

4.2 MATERIALS AND METHODS

Tissue Procurement and Establishment of Primary Explant Cultures were performed as described in the methods of **Chapter 2.** Primary culture of breast tumor samples are discussed in the methods of **Chapter 3** (283).

4.2.1 Culture of Established Cell Lines

MDA MB231, MCF7, MCF7/Ly2, BT-20, CaMa-1, and SK-BR-3 cell lines were obtained from the American Type Culture Collection (ATCC) (Rockville, MD). MCF10A was obtained from the Michigan Cancer Foundation Cell Lines Facility (Detroit, MI). MDA MB231, MCF7, and MCF7/Ly2 were cultured in DMEM (Gibco) supplemented with 10% heat inactivated FCS, while BT-20, CaMa-1, SK-BR-3 and MCF10A were grown in RPMI (Mediatech) supplemented with 10% heat inactivated FCS. All cell lines were grown on uncoated chamber slides and were maintained in a humidified incubator at 37°C in 5% CO_2. Cell lines were plated subconfluently 1-2 days prior to the UDS assay.

4.2.2 Unscheduled DNA Synthesis Assay

NER, specifically global genomic repair was measured utilizing the autoradiographic UDS assay as described in **Chapter 3** (140,191,400). Four slides of each of the cell lines were run in 3 independent experiments. At least two slides of each of the primary cultures derived from breast reduction epithelium and two slides of each of the primary cultures derived from intraductal carcinomas were examined. Each of these primary cultures was run in one UDS assay experiment. One hundred nuclei were counted on both the irradiated and unirradiated sides of two slides for each sample by two independent counters (400 nuclei per counter, per sample, per experiment for a total of 2,400 nuclei per cell line).

4.2.2.1 Statistical Analysis

Final grain counts from slides of the same cell type within the same experiment were pooled together and expressed as a percentage of the concurrently analyzed FF. NER values for the cell lines were pooled or averaged over all experiments. Comparisons between different cell lines were performed using two tailed t-tests with $\alpha<0.05$, using statistical analysis package included in the Microsoft Excel spreadsheet program. Comparisons between populations (*e.g.,* breast reduction versus cancer cell lines) were also carried out using two tailed t-tests with $\alpha < 0.05$, in Microsoft Excel.

4.2.2.2 Z-test Comparisons

Comparisons between a population and a single point (*e.g.,* breast reduction versus the MCF10A cell line) were carried out using a two-tailed z-test with $\alpha<0.05$, also in Microsoft Excel. The z-test allows for the assessment of how likely a single point is to be a member of a particular population (assuming that population has a normal distribution), using the mean of that population and its standard deviation to represent its distribution. The possible effects of S phase index and age were evaluated using linear regression at the same level of significance using MiniTAB.

4.2.3 Generating and Maintaining Latimer Cell Lines

Extended explant cultures were generated from primary explants of both normal and tumor cells. Primary cultures were plated on matrigel-coated 12.5-cm tissue culture dishes (Falcon) andwere allowed to grow to confluence. Once this was achieved, cells were split at a 1:2 ratio with 0.25% trypsin (Sigma) into two 12.5-cm flasks. These cells were allowed to attach to the substratum for 24 hours and then fed fresh MWRI medium every 48 hours. Passaging was carried out at a 1:2 ratio each time as confluence was achieved. Before 13 passages, samples were considered to be extended explants. After 13 passages these explants were designated cell lines. No exogenous immortalizing agent was used at any time on any of these samples.

4.2.4 Cytokeratin and Estrogen Receptor Staining

Cytokeratin 18 and Estrogen Receptor Staining patterns were determined on confluent cell line cultures grown on chamber slides as described for the Unscheduled DNA synthesis assay. Both breast reduction and breast tumor lines were stained in accordance with the protocols described in **Chapter 2**.

4.2.5 Karyotyping

Karyotyping was carried out in collaboration with either Sharon Wenger at the Clinical Cytogenetics Laboratory of West Virginia University or Susanne Gollin at the University of Pittsburgh Cancer Institute Cytogenetics Facility.

Confluent cell cultures were passaged and cell growth was monitored. When enough mitoses (rounded up cells) were observed, usually after 1-2 days, 20-45 μl of colcemid (stock concentration of 10 μg/ml) was added and the culture was left in the incubator overnight (the amount and the length of time of colcemid addition varied depending on the mitotic index of the culture). The culture was then visualized under the microscope for mitotic cells. If enough mitotic cells were found, the culture was harvested.

Cell were harvested by trypsinization and resuspended in 0.1% sodium citrate hypotonic solution and incubated for 10 min at 37°C. Cells were collected by centrifugation and fixed in 3:1 solution of methanol to acetic acid at 4°C for 1 hour. This fixation was repeated for a total of three times with harvest of cells by centrifugation between each fix. The cells were then dropped using a 5.75 inch pasteur pipette on a preheated super frosted microscope slide held at a 45°

angle from a distance of about 10 inches. The slides were reheated in a 65°C water bath. Slides were air dried and incubated at 60°C overnight. Cells were banded by exposure to trypsin for 20 to 30 seconds followed by 2 saline rinses. Slides were stained with 5% Giemsa for 6.5 minutes, rinsed with tap water and air-dried. Finally, the slides were examined under the microscope for metaphases. A minimum of 20 metaphases were randomly chosen and used to determine the karyotype of the each culture.

4.2.6 Microarray Analysis

Microarray expression analysis was carried out at the Clinical Genomics Facility of the University of Pittsburgh Cancer Institute under the guidance of Dr. Federico Monzon using the Affymetrix U133 2.0 Plus chip. Before samples were used, RNA integrity was evaluated using the Agilent Bioanalyzer technology in the Microarray facility. The output of this technique is an RNA integrity number (RIN), a tool designed by Agilent to estimate the integrity of the RNA samples. RIN ranges from 1-10 with 10 representing a perfect RNA without any degradation products. Our samples all fell within the 6-10 range.

The Affymetrix U133 2.0 Plus chip set comprises 54,000 probe sets to analyze the expression of 47,000 transcripts and variants, including 38,500 well-characterized human genes. Sequences used in the design of the array were selected from GenBank®, dbEST, and RefSeq. Biotin-labeled cRNA synthesis, hybridization, washing, staining, and scanning were done following the manufacturer's protocols (Affymetrix). Total mRNA was labeled with fluorescent tags by a polymerase chain reaction and then hybridized to the chip set. Results were captured using a GeneChip Scanner 3000 with GeneChip Operating Software (GCOS).

Cell lines were analyzed using unsupervised hierarchical cluster analysis to observe relationships between them with Avadis software.

4.3 RESULTS

4.3.1 Nucleotide Excision Repair Capacity of Commonly Used Breast Cancer Cell Lines

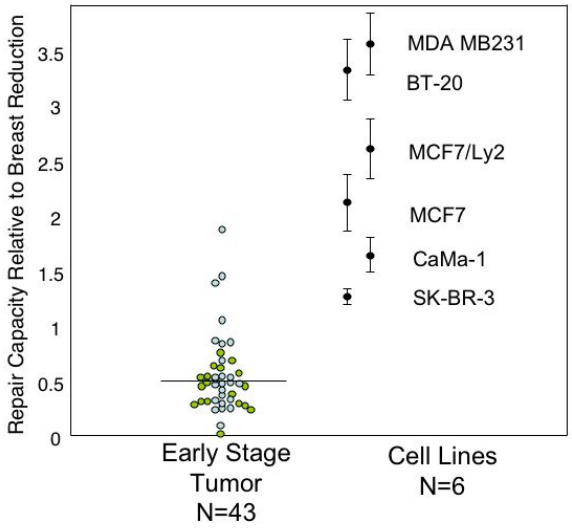
In order to determine whether established breast tumor cell lines reflect the baseline NER capacity of the breast tumor primary cultures, we compared primary cultures of breast tumor tissue with 6 breast tumor cell lines and the nontumorigenic cell line MCF-10A using the UDS assay (Figure 4.1A). Similarly, we compared the 6 tumor cell lines with primary cultures of late stage breast tumors and breast reduction epithelium (Figure 4.1B). With this information, researchers will be able to understand an important caveat in the utilization of breast cancer cell lines in the investigation of the role of DNA repair in breast cancer etiology.

When compared to the populations of early stage tumors, the cell lines as a group were statistically significantly increased in their repair by 4.35 times the average of 43 early stage tumors ($p = 0.00426$) (Figure 4.1A). Using a two-tailed z-test, these cell lines each showed that they were individually significantly different from the mean of the early stage tumor population (all p values were <0.001). Thus these cell lines cannot be considered as surrogates for this population.

In comparison to the early stage tumors, when these cell lines were comparable to the average of the late stage tumors they demonstrated a 1.66 times greater repair capacity as a

population, but this was not a significant difference ($p = 0.0689$) (Figure 4.1B). When each cell line was considered as an individual point, a z test showed that only the SK-BR-3 and CaMa-1 cell lines were likely to be members of the population of late stage tumors ($p = 0.576$ and $p = 0.558$, respectively). All other cell lines had p values less than our α of 0.05 (range from 0.0423 for MCF7 to 4.04×10^{-10} for MDA MB231). The range of these commercial cell lines compared to the range seen in the late stage tumors, seems to suggest that these cell lines are more related to one another than the late stage primary cultures are related to one another. This could possibly be due to the fact that our primary cultures reflect more of the heterogeneity of progressive breast cancer.

A



B

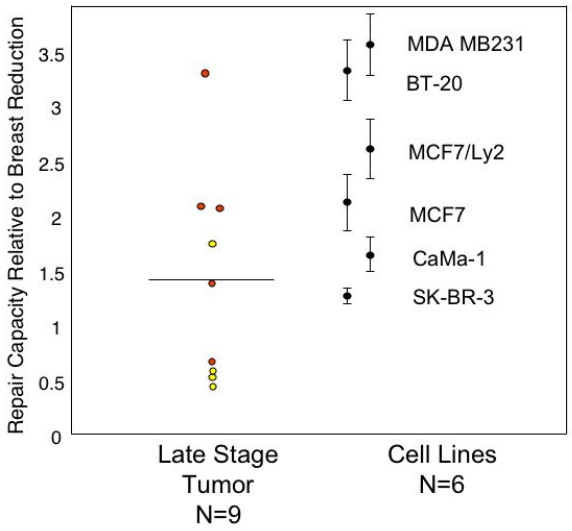


Figure 4.1 Repair capacity of commercially available cell lines compared to early and late stage primary cultures. A. Established breast cancer cell lines are not consistent with the repair capacity of early stage breast cancer primary cultures. Comparisons of the functional NER capacities of 43 primary explant cultures of stage I and stage II ductal carcinomas with 6 commonly used breast tumor cell lines are shown. 19 stage I tumors (green) and 24 stage II tumors (blue) are shown. Standard error over three independent experiments is indicated for the cell lines. The horizontal bar represents the average of the breast tumor data.

Comparisons made between the cell lines themselves show that the two ER+ cell lines, MCF-7 and CaMa-1 are each significantly lower in repair than the ER- cell lines BT-20 and MDA MB231 ($p < 0.05$) (Table 4.2). In contrast, they are both more efficient at repair than the other ER- cell line SK-BR-3, but this relationship did not reach significance for either. MCF7 and CaMa-1 are not significantly different from one another. When compared to one another MCF-7 is not significantly different from its subline MCF7/Ly2, however MCF7/Ly2 is significantly different from the other tumor cell lines tested, with the exception of BT-20.

When the breast tumor cell line BT-20, which was derived directly from an intraductal breast carcinoma, is compared to the other cell lines, which were each derived from malignant pleural effusions, it was found to be significantly higher than MCF-7, CaMa-1, and SK-BR-3, but not significantly lower than MDA MB231.

The cell line SK-BR-3, which overexpresses the HER2/c-erb-2 oncoprotein, is significantly lower in repair capacity the MCF7 subline MCF7/Ly2, BT-20, and MDA MB231.

Unfortunately the treatment protocols carried out on the patients from whom many of these tumors were derived were not published. We, do, however, have information on the MDA MB231 cell line, treated with the genotoxic combination chemotherapies of CMF (cyclophosphamide, methotrexate, and 5-flurouracil) and CAF (cyclophosphamide adriamycin (doxorubicin), and 5-flurouracil), steroid, and oophorectomy , and the MCF7 cell line, treated with hormones and external bean radiation. Both women were treated with endocrine therapy, although through different methods (oophorectomy versus hormonal therapy) and genotoxic insults (genotoxic chemotherapy and radiation). The MDA MB231 patient was also treated with steroids for unknown reasons.

MDA MB231 and MCF7 have significantly different NER capacities from one another, with the MDA MB231 cell line demonstrating the higher repair capacity. Although this reflects the results of only two samples, it is interesting that the tumor sample that was treated with genotoxic agents that form bulky adducts remediated by the NER pathway, specifically cyclophosphamide, has an increased repair capacity relative to the tumor sample treated with radiation. Perhaps the high rate of repair for the MDA MB231 cells reflects a population of cells that were able to resist the cyclophosphamide treatment due to increased NER capacity.

The NER capacity of each of these breast cancer cell lines ranged from 1.22 times the average of breast reduction epithelium for SK-BR-3 to 3.44 times the repair capacity of normal breast reduction epithelium for MDA MB231 (Figure 4.2). When the commercially available cell lines are considered together as a population against the breast reductions they are statistically significantly different from this group ($p = 0.0104$) with the cell lines demonstrating an average of 2.41 times the repair capacity of the breast reduction cultures. When compared individually to the breast reduction cultures using a two-tailed z-test, each of these individual cancer cell lines were significantly different than the mean of the breast reduction population (p value for the z scores ranged from <0.001 to 0.0131).

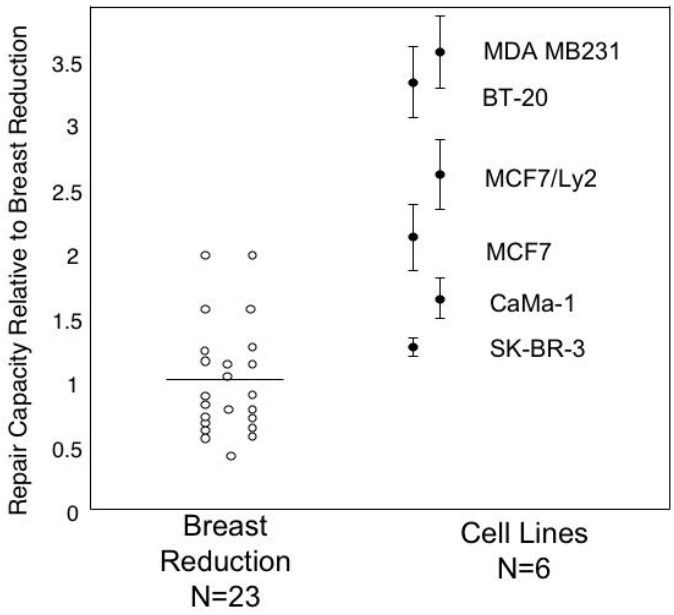


Figure 4.2 Primary cultures of breast reductions are lower in NER capacity than commercially available breast cancer cell lines. Comparisons of the functional NER capacities of 23 primary explant cultures of breast reduction mammoplasty tissue with 6 commonly used breast tumor cell lines are shown. Standard error over three independent experiments is indicated for the cell lines. The horizontal bar represents the average of the breast reduction data.

4.3.2 Nucleotide Excision Repair Capacity of the MCF10A Cell Line

When we karyotyped the nontumorigenic line MCF-10A it was determined to be 48 XX, del(3)(p13), der(6) t(6;12)(p26;q22), +8, der(9) t(7;9)(q11.2;p24), +16 (Figure 4.3). This is in contrast to the published karyotype of female diploid except for a stable t(3:9)(3p13:9p22). This stable translocation was not seen in the strain of MCF-10A that we received. These cells were purchased directly from the Michigan Cancer Foundation in 1995.

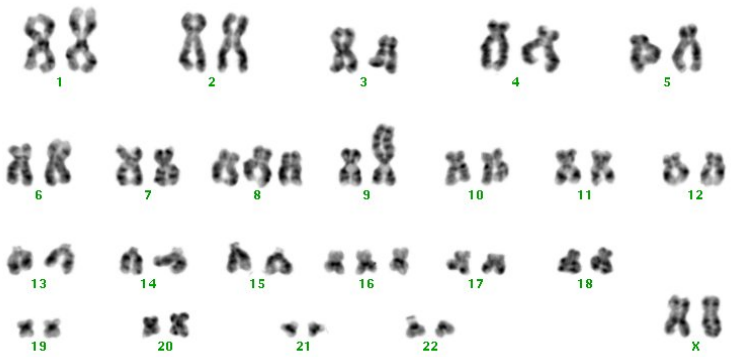


Figure 4.3 The karyotype of MCF10A is consistent with that of a transformed cell. This cell line was purchased by the Latimer lab directly from the Michigan Cancer Foundation.

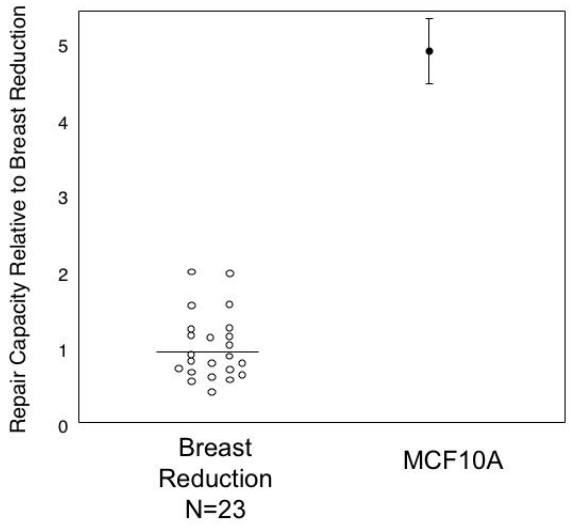
Description:

48 XX, del(3)(p13), der(6) t(6;12)(p26;q22), +8, der(9) t(7;9)(q11.2;p24), +16

When MCF-10A was compared to primary cultures of breast epithelium, early stage tumors, and late stage tumors by two tailed z-test the p value was calculated at <0.001, indicating that this data point is not a member of any of these populations. MCF10A exhibits more than a 5 fold increase in repair capacity relative to normal breast reduction cultures, a 9.27 fold increase over the early stage tumors, and a 3.55 fold increase over the late stage tumors (Figure 4.4). It also manifests an extremely slow growth rate even when cultures in MWRI, a medium containing 20% serum.

When MCF10A was compared to the cancer cell lines assayed here, it was significantly higher than each line except for MDA MB231, for which the trend of increased repair capacity failed to reach significance (p= 0.0593).

A



B

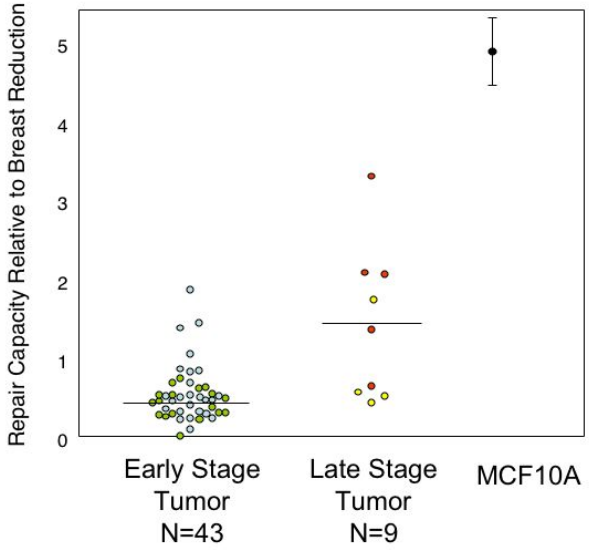


Figure 4.4 MCF10A has significantly higher repair capacity than breast reduction mammoplasties and early stage tumors.A. Comparisons of the functional NER capacities of 23 primary explant cultures of breast reduction mammoplasty tissue with the MCF10A cell line. Standard error over three independent experiments is indicated for MCF10A. The horizontal bar represents the average of the breast reduction data.

B. Comparisons of the functional NER capacities of 43 primary explant cultures of stage I and stage II ductal carcinomas, 9 primary explant cultures of chemotherapy-naïve stage III and stage IV ductal carcinomas, and the MCF10A cell line. 19 stage I tumors (green), 24 stage II tumors (blue), 5 stage III tumors (red), and 4 stage IV tumors (yellow) are shown. Standard error over three independent experiments is indicated for MCF10A. The horizontal bars represent the averages of the breast tumor data.

Table 4.2 Two-tailed t-test p-values comparing the repair capacities of six breast cancer cell lines and the nonmalignant cell line MCF10A. Significant associations ($\alpha < 0.05$) are shown in red.

	MCF7/Ly2	**CaMa-1**	**SK-BR-3**	**BT-20**	**MDA MB231**	**MCF10-A**
MCF7	0.257	0.198	0.0665	0.0324	0.00502	0.00897
MCF7/Ly2	X	0.0490	0.0307	0.135	0.0428	0.0153
CaMa-1	X	X	0.121	0.0115	5.26 x 10-5	0.00947
SK-BR-3	X	X	X	0.0131	2.56 x 10-6	0.0112
BT-20	X	X	X	X	0.574	0.0451
MDA MB231	X	X	X	X	X	0.0593

4.3.3 Effects of Proliferative Index on Repair in Breast Cell Lines

When the effect of S-phase index was determined by a linear regression against NER capacity for these tumor cell lines, it was determined that the relationship was not significant (p= 0.600) (Figure 4.5). However, when data from the fibrocystic line MCF-10A was used the p value decreased to a significance of p= 0.0228. Yet this data point is 63% responsible for the significance of the trend observed.

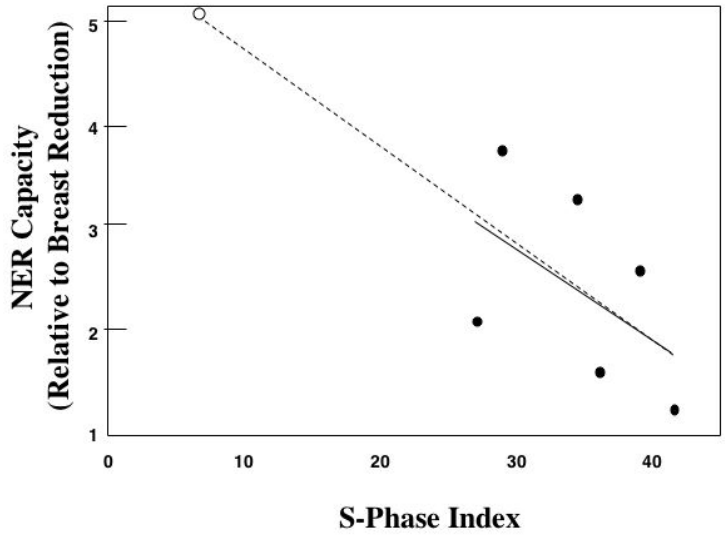


Figure 4.5 NER capacity is reduced with increasing S-phase in established breast cancer cell lines. S-phase index linear regression graph of NER capacity of 6 established breast tumor cell lines and the fibrocystic-derived cell line MCF-10A is shown. The solid line represents the regression based only on the tumor cell lines, while the dashed line represents the regression when the data from MCF-10A is included.

4.3.4 Characteristics of an Ideal Breast Cell Line

As these commercially available cell lines were not reasonable surrogates for the DNA repair capacity exhibited in early stage breast cultures, another source of cells, DNA, RNA, and protein was needed. Because of the known effects of currently used methods of immortalization on DNA repair, and because of the many effects yet to be enumerated, it was necessary to find a method of generating cell lines that did not involve the use of an exogenous immortalizing agent. However, since the rate of spontaneous immortalization in current systems is rare, it was clear that a new model for breast cell culture was needed for both breast reduction tissue and tumors. Ideally, the cell lines created should maintain a phenotype as close to the human breast as possible. The generation of a physiologically relevant system that can provide large numbers of cells and a renewable source of DNA, RNA, and protein is critical not only to this work with DNA repair capacity, but to the understanding of the molecular biology of both breast development and breast carcinogenesis.

These cells should be robust enough to withstand experimental manipulation, and normal and tumor cells should be culturable by the same system to decrease variability in experiments. In the case of normal breast epithelium, cultures should be capable of generating architecture, maintaining cytokeratin staining specificity, preserving a normal karyotype, and expressing the ER as was described for the primary culture system in Chapter 2.

With respect to cell lines generated from breast tumors, it is also hoped that they would maintain as many of the features of the tumor of origin as possible including a lack of architectural organization, the production of appropriate epithelial cytokeratins, an abnormal

karyotype and the possible expression of the ER. The system should also be capable of generating cell lines derived from early stage breast tumors, including ductal carcinoma *in situ*, something which is completely lacking in the current research climate. Cell lines currently available as discussed above have been made from either stage IV or metastatic tumors cells, making them more suitable for the study of tumor progression than tumor etiology.

Table 4.3 Desired characteristics of organotypic breast cell lines.

Characteristic	**Breast Reduction Cell Line**	**Breast Tumor Cell Line**
Organization / architecture	Should reiterate the architecture of the breast- ducts and lobules lined by epithelial cells	Lack of architecture; perhaps representing the epithelial to mesenchymal transition believed to be involved in the acquisition of an invasive phenotype
Karyotype	Diploid, XX	Abnormal
Cytokeratin Profile	Consistent with luminal epithelial cells- CK19 and CK18 positive; myoepithelial cells- CK14 positive	Varied; different breast cancer subtypes express different combinations of cytokeratins
Epithelial Membrane Antigen	positive	Varied; different breast cancer subtypes display different phenotypes
Estrogen Receptor	Positive for 11% of epithelial cells	Varied; lost in the pathogenesis of 25% of breast tumors (see Introduction) Ideally should be concordant with tumor of origin

4.3.5 Generation of Extended Explants and Cell Lines in the Latimer Lab

Primary cell cultures of breast reduction epithelium and breast tumors were conservatively passaged by trypsinization to passage 13 and beyond (Figure 4.6). No exogenous immortalizing agents are used, yet the cultures remain viable to passage 30 and beyond. Explants generated by passaging primary cultures are considered extended explants until passage 13, and then cell lines one they continue past passage 13. This is consistent with the presentation of terms by Stampfer introduced above (369) in the discussion of immortalization. Primary cultures are generated directly from breast tissue, and then passaged. Where she calls these subcultures "strains with long-term growth potential," we simply continue to call them explants in accordance with the cell culture literature (401). Once these cultures pass their crisis and exhibit what has been for us an as-yet unlimited growth potential, we call them cell lines.

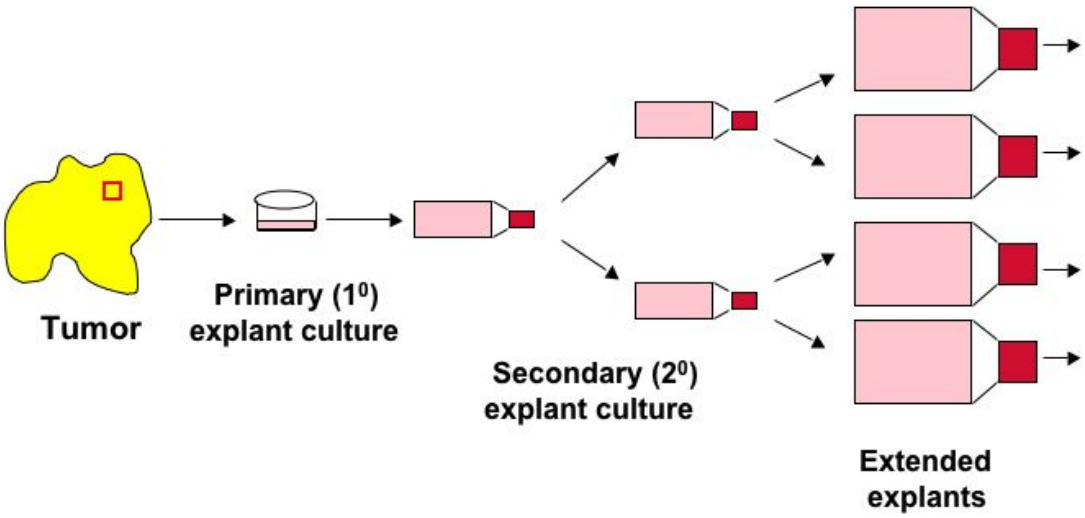


Figure 4.6 Schematic of the generation of explants and cell lines in the Latimer lab.

To date the Latimer lab has generated over 119 cell lines from multiple tissue and tumor types, including those listed in Table 4.4. Characterization of many of these lines has begun and is described below with respect to the desired traits of an organotypic cell line.

Table 4.4 Summary of cell lines generated in the Latimer lab.

Tissue of Origin	**Number of Cell Lines**
Breast Reduction Mammoplasty	27
Ductal Carcinoma *in situ*	5
Tissue adjacent to Ductal Carcinoma *in situ*	1
Invasive breast carcinoma	46
Tissue adjacent to invasive breast carcinoma	11
Contralateral breast of a woman who had invasive breast carcinoma	2
Prophylactic breast mastectomy of healthy BRCA1 or BRCA2 mutation carrier	4
Tumor of a BRCA1 or BRCA2 mutation carrier	4
Tissue adjacent to Tumor of a BRCA1 or BRCA2 mutation carrier	3
Contralateral breast of a BRCA1 or BRCA2 mutation carrier who had breast cancer	2
Breast fibroadenoma	4
Tissue adjacent to breast fibroadenoma	3
Normal ovarian epithelial scrapings	1
Ovarian tumors	6

4.3.6 Formation of Architecture

Like the primary explant cultures (Chapter 2), some breast reduction cell lines (JL BRLs) are able to maintain the ability to form structures in cultures at least up to passage 14 (Figure 4.7A). In contrast to this, many breast tumor lines (JL BTLs) neither show this ability in primary culture nor manifest it as cell lines (Figure 4.7B).

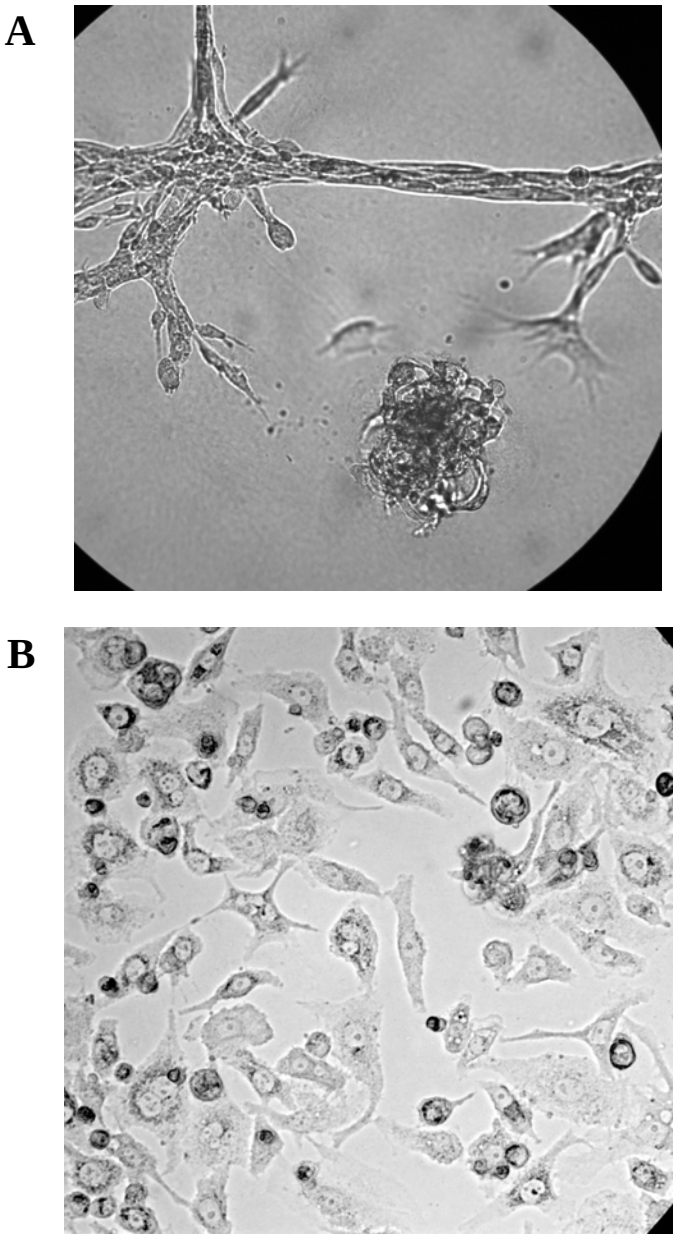


Figure 4.7 DIC images of cell lines generated from nondiseased and tumor tissue. A. DIC image of JL BRL-14, a cell line derived from breast reduction mammoplasty tissue at passage 8, still forming ductal architecture. B. DIC image of JL BTL-12, a cell line derived from a stage III ductal carcinoma at passage 22.

4.3.7 Cytokeratin Staining Patterns

Two of our tumor cell lines, one derived from a stage III invasive ductal carcinoma (Figure 4.8), and one derived from a ductal carcinoma *in situ* were both shown to express cytokeratin 18. Characterization of breast reduction lines with cytokeratin 18 is ongoing. Expression of cytokeratin 18 is consistent with breast ductal epithelial cells.

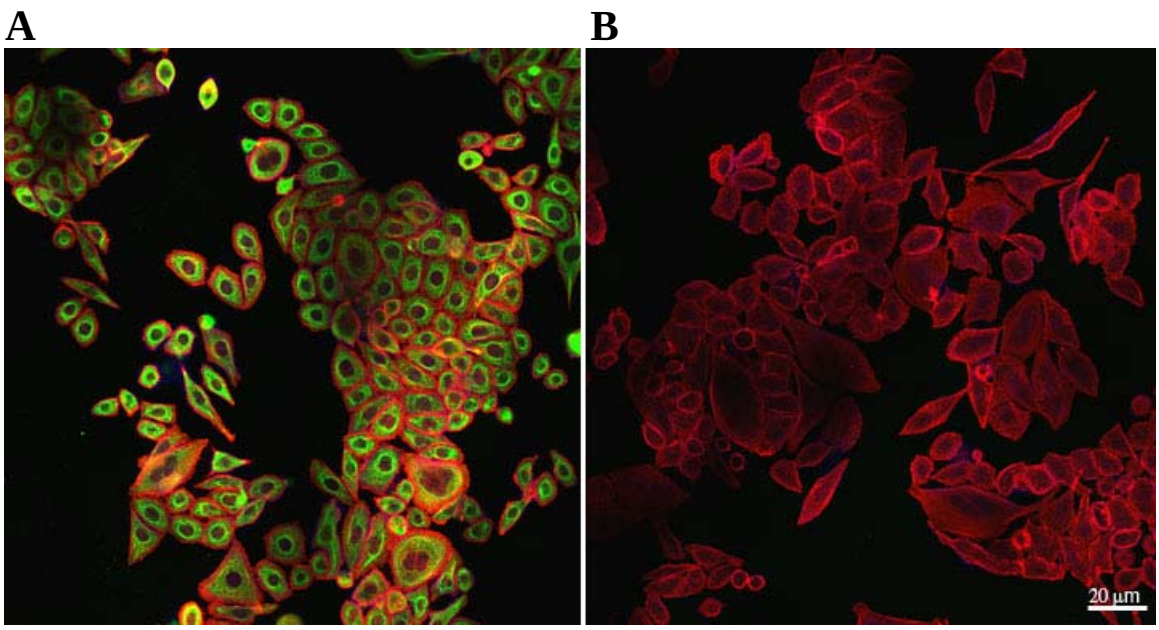


Figure 4.8 Cytokeratin 18 staining of the epithelial cell line JL BTL-12, derived from a stage III chemotherapy-naïve invasive ductal carcinoma. Positive staining for cytokeratin 18 (green) is shown on the left. The nonimmune control is shown on the right. Counterstaining was done by a rhodamine-conjugated phalloidin, recognizing filamentous actin.

4.3.8 Epithelial Membrane Antigen Staining

Both normal and tumor cell lines generated using the Latimer method express epithelial membrane antigen staining, also consistent with their identity as epithelial cells. Nonspecific staining of the matrigel substratum can be seen surrounding the JL BRL-14 cells in Figure 9A in the nonimmune control (Figure 9B). The JL DCIS-3B line, derived from a ductal carinoma *in situ*, also shows positive staining with EMA, consistent with its identity as an epithelial cell (Figure 9 C and D).

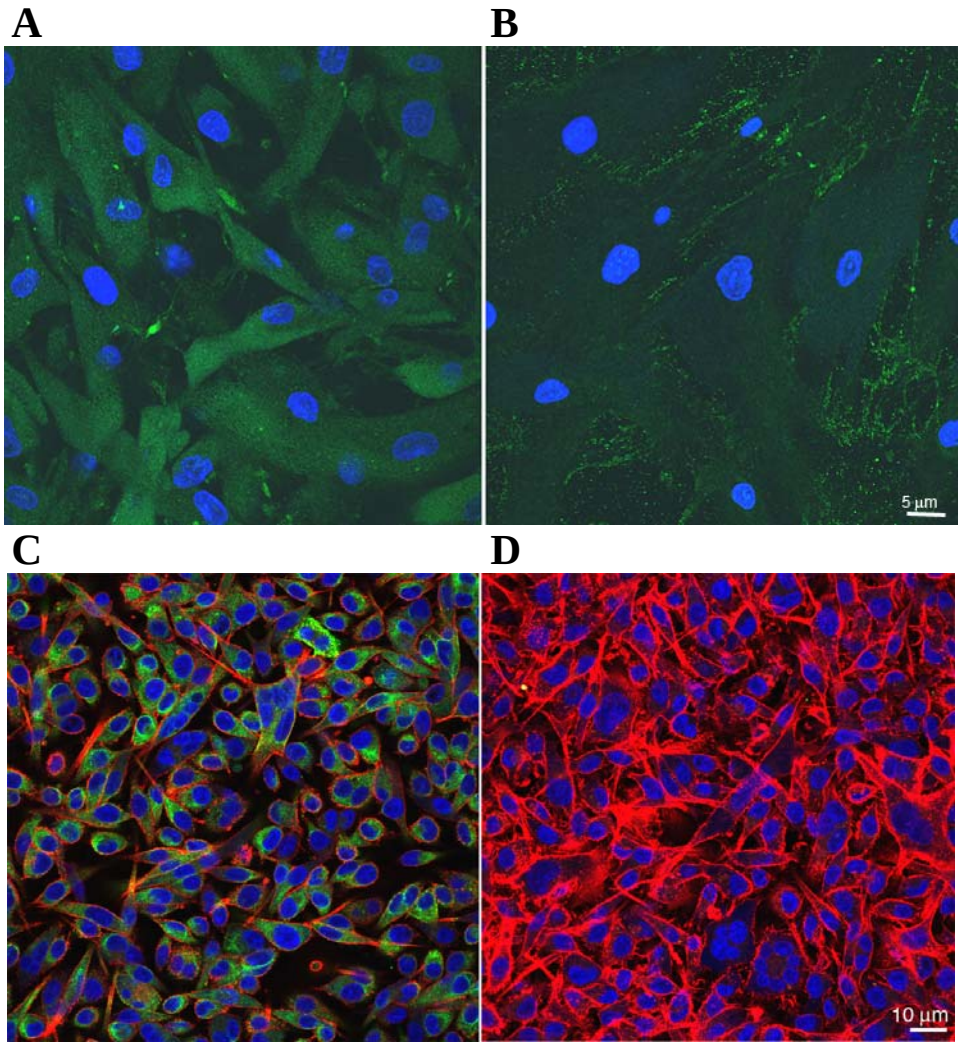


Figure 4.9 Epithelial Membrane Antigen staining of cell lines generated from nondiseased and tumor tissue. A. Epithelial membrane antigen staining of JL BRL-14, a cell line derived from breast reduction epithelium and B. nonimmune control. C. Epithelial membrane antigen staining of JL DCIS-3B, a cell line derived from a ductal carcinoma *in situ* and D nonimmune control. Epithelial membrane antigen is represented by the green stain in both A and B. All panels include contain nuclei counterstained with the chemical DRAQ 5. The counterstain of rhodamine-conjugated phalloidin is also used in panel B demonstrating the presence of filamentous actin.

4.3.9 Estrogen Receptor Staining Patterns

Lines generated in the Latimer lab from both breast reduction tissue (Figure 4.10A and B) and tumor (Figure 4.10 B and C) also show the presence of ER in a subset of nuclei.

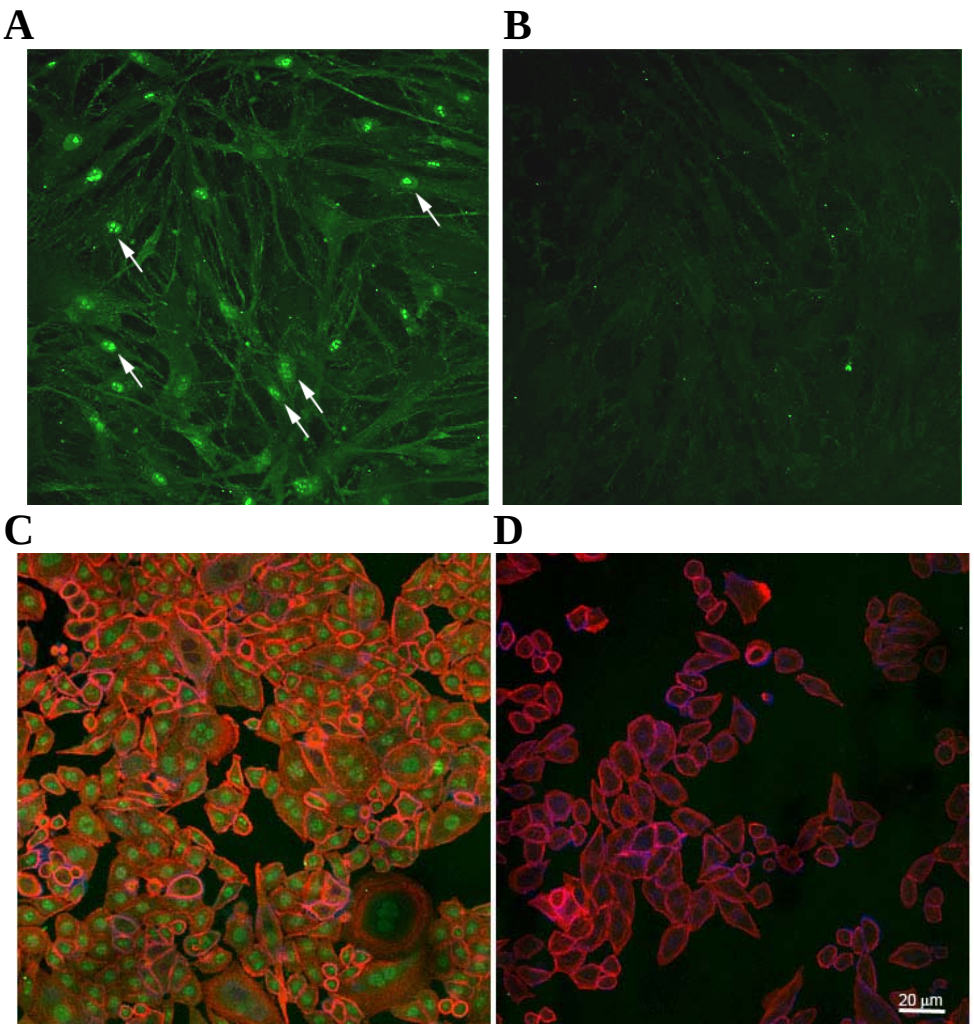


Figure 4.10 Estrogen Receptor α staining of cell lines generated from nondiseased and tumor tissue. A. Estrogen receptor α expression of JL BRL-14, a cell line from normal breast reduction epithelium, and B its nonimmune control. Arrows indicate some of the ER+ nuclei in this field. C. Estrogen receptor staining of JL BTL-12, a cell line established from an estrogen-receptor positive stage III ductal carcinoma, and its D respective nonimmune control. Nuclei are stained with ERα (green). C and D include the counterstain of rhodamine-conjugated phalloidin to demonstrate the presence of filamentous actin.

4.3.10 Karyotypes of Breast Reduction-Derived Cell Lines

Normal female diploid karyotypes have been obtained for 4 breast reduction epithelial cell lines generated in the lab called JL BRL-11, JL BRL-14, JL BRL-16, and JL BRL-23. These cells not only show a normal autosomal karyotype, but also maintain both of their X chromosomes in culture (Figure 4.11).

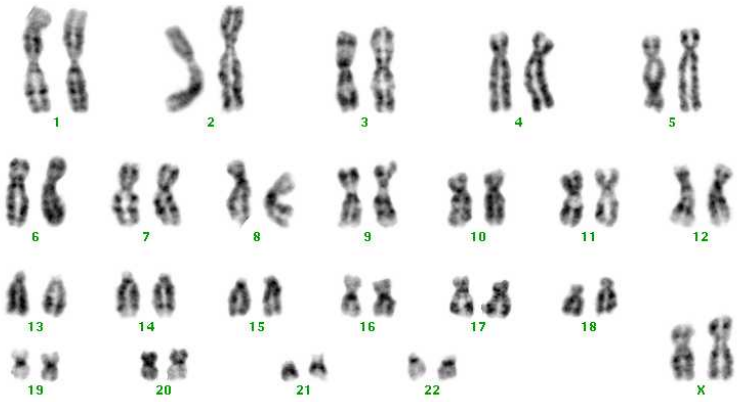


Figure 4.11 JL BRLs have normal karyotypes. The Geimsa banding pattern karyotype of JL BRL-14, a cell line generated in the Latimer lab from a breast reduction mammoplasty is shown. Other BRLs with the sample karyotype are JL BRL-11, JL BTL-16, JL BTL-23, and JL BRL-24. The karyotype for all of these is 46XX.

4.3.11 Karyotypes of Breast Tumor-Derived Cell Lines

In contrast to this, we have seen abnormal karyotypes for multiple tumor cell lines generated in the same way. JL DCIS-3A and JL DCIS-3B, two cell lines generated from a patient with a 7 centimeter ductal carcinoma *in situ* mass, were both show to have a grossly abnormal hypertriploid / hypotetratploid karyotype with multiple marker chromosomes. JL BTL-12, derived from a stage III invasive ductal carcinoma also demonstrated an abnormal karyotype as seen in Figure 4.12. This karyotype is also grossly abnormal, with an unknown marker chromosome and significant aneuploidy. This cell line was used for subsequent analysis of mRNA and protein expression in Chapter 5.

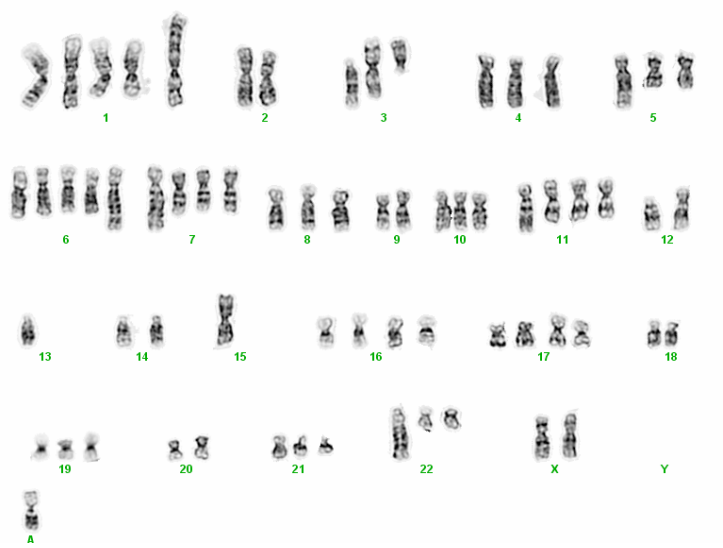


Figure 4.12 Breast tumor lines have abnormal karyotypes. The geimsa-banding pattern karyotype of JL BTL-12, derived from a stage III invasive ductal carcinoma of the breast. The karyotype is: 49-75,X,-X[17],del(X)(q23)[12],+1[12],+1[3], del(1)(q31)[16], del(1)(q31)x2[4],der(1)t(1;3;7)(1qter→1p10::3q10→3q24::7q11.2→7qter)[18],der(1;3)(p10;q10)[15],-2[15],del(2)(q12q22)[20],del(2)(q12q22)x2[15],+3[7],del(3)(q10)[20], der(3)del(3)(p10)t(3;?)(q26.1;?)[16],-4[7],del(4)(p13p15)[10],-5[7],+5[7], del(5)(q13.2q31.2)[19], del(5)(q13.2q31.2)x2[5],-6[5],+6[9],add(6)(q24)[11], add(6)(p22.3)[8],+7[11],+7[5],-7[4],der(7)t(5;7)(q14;q31.3)[16],-8[5],del(8)(p22)[12], -9[17],del(9)(q33)[15],-10[9],del(10)(p12.3)[13],+11[9],add(11)(q23)[4], der(11;15)(q10;p10)[17],-12[10],del(12)(p10)[17],-13[19],-13[11],add(13)(p10)[5], del(13)(q13;q21.3)[5],-14[10],+14[5],-15[12],i(15)(q10)[14],i(15)(q10)x2[2],+16[8], add(16)(p12)[12],add(16)(p12)x2[4],-17[11],add(17)(p10)[11],add(17)(p10)x2[4], -18[15], -18[6],-19[7],+19[5],-20[9],+20[4],del(20)(q13.2)[9],-21[8],-22[14], add(22)(q11.2)[18],+mar1[14],+mar2[6],+mar3[2][cp20]

4.3.12 Microarray Analysis

Total RNA from multiple cell lines was analyzed for its expression patterns using the Affymetrix U133 2.0 Plus microarray chip. Included in the analysis are the commercial cell lines and Latimer-generated cell lines listed in Table 4.5. A selection of these tumor cell lines were used for further characterization of the molecular mechanism behind the observed loss and gain of functional repair capacity: JL BRL-1, JL BRL-2, JL BRL-6, JL BTL-8, JL BTL-3 (not yet included in microarray analysis), JL BTL-29, JL BTL-12, MCF7, and MDA MB231. Details of the tumor cell lines used for further analysis are listed in Chapter 5.

Table 4.5 Established and Latimer-generated cell lines used in microarray analysis.

Cell Line	**Tissue of Origin**
MCF10A	Breast reduction mammoplasty
MDA MB231	Stage IV pleural effusion
MCF7	Stage IV pleural effusion
SK-BR-3	Stage IV pleural effusion
Ca-Ma-1	Stage IV pleural effusion
BT-20	Stage IV invasive ductal carcinoma
FF (Foreskin Fibroblasts)	Explant culture of neonatal foreskin tissue
Hep3B	Hepatocellular Carcinoma
HeLa 229	Cervical Carcinoma
JL BRL-1	Latimer-generated; Breast Reduction
JL BRL-2	Latimer-generated; Breast Reduction
JL BRL-6	Latimer-generated; Breast Reduction
JL BRL-14	Latimer-generated; Breast Reduction
JL BRL-16	Latimer-generated; Breast Reduction
JL DCIS-3A	Latimer-generated Ductal Carcinoma *in situ*
JL DCIS-3B	Latimer-generated Ductal Carcinoma *in situ*- same patient as JL DCIS-3A, but a more advanced area of the tumor
JL BTL-8	Latimer-generated Stage I invasive ductal carcinoma
JL BTL-45	Latimer-generated Stage I invasive ductal carcinoma
JL NTAL-45	Latimer-generated Histologically normal tissue located 1 cm away from the Stage I invasive ductal carcinoma used to generate JL BTL-45
JL BTL-2	Latimer-generated Stage II invasive ductal carcinoma
JL BTL-29	Latimer-generated Stage II invasive ductal carcinoma
JL BTL-12	Latimer-generated Stage III invasive ductal carcinoma Chemotherapy-naive
JL BTL-20	Latimer-generated Stage IV invasive ductal carcinoma Treated with genotoxic chemotherapy

The microarray chip contains all the known genes of the human genome as well as all estimated genes. Overall gene expression patterns were compared to one another using an unsupervised analysis using the Avadis software program. This analysis relates cell lines to one another based upon their total gene expression profile with equal weight given to all genes tested. Data from this analysis is represented as a dendrogram in Figure 4.13. Horizontal lines can be drawn across the dendrogram to create different clusters of cell lines. The higher up the dendrogram the horizontal line, the less related those cell lines in the cluster are. The horizontal line shown in Figure 4.13 creates 3 general clusters of cells. (1): MCF10A, MDA MB231, and JL DCIS3B. Interestingly, the nondiseased MCF10A cell line clusters here with an ER- pleural effusion and one of our tumor cell lines rather than with our breast reduction lines. This correlates with our repair data showing how different this cell line is now from breast reduction epithelium. (2) The second cluster is comprised of the bulk of the Latimer-derived tumor and breast reduction cell lines and the foreskin fibroblast explants. Within this cluster some additional trends can be seen. For example, the breast reduction lines JL BRL-1, JL BRL-2, and JL BRL-6 are more related to one another than they are to the stage I cell line JL BTL-8 or the stage II line JL BTL-29. (3) The third and final cluster represents the majority of the commercially available cell lines and the stage II Latimer cell line JL BTL-12, emphasizing how abnormal this cell line truly is in terms of its expression patterns.

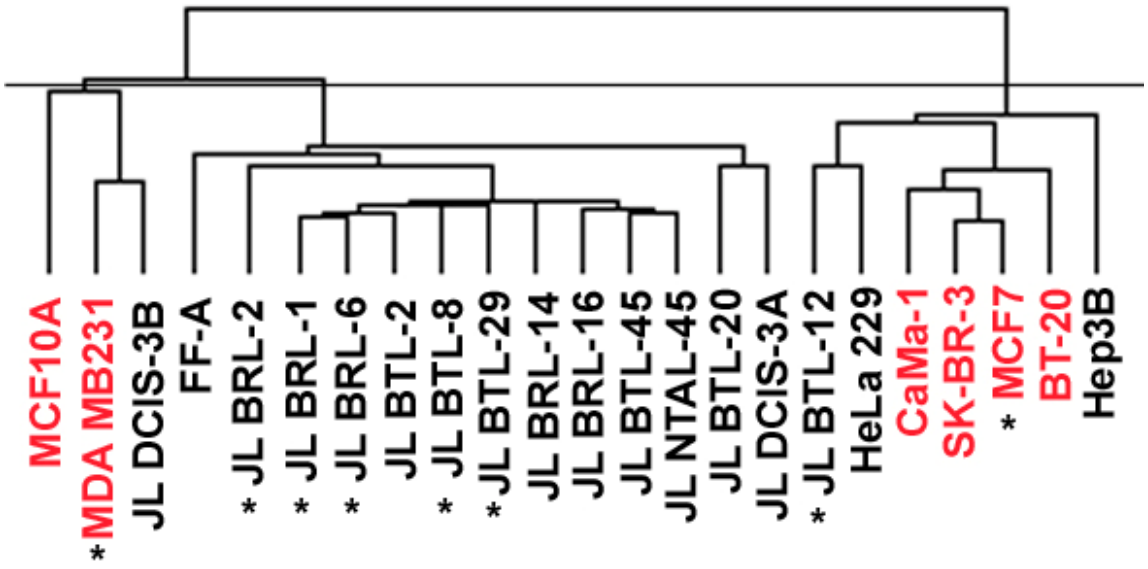


Figure 4.13 Dendrogram representing the unsupervised analysis of expression data from multiple cell lines. Five of the commercially-available breast tumor cell lines evaluated above for their repair capacity, MCF10A, a foreskin fibroblast explant, the cell lines HeLa 229 and Hep3B, eight Latimer-generated tumor cell lines, one Latimer-generated line from the histologically normal tissue adjacent to a stage I tumor, and five breast-reduction generated cell lines were all analyzed for their expression patterns using the Affymetrix U133 2.0 Plus microarray chip. Starred cell lines were chosen for use in subsequent experiments (Chapter 5). Commercial cell lines analyzed for DNA repair as shown in red. The horizontal line across the dendrogram breaks the cell lines up into three clusters of related lines.

4.4 DISCUSSION

4.4.1 High DNA Repair Capacities in Established Cell Lines Relative to Early Stage Tumors and Breast Reductions

While established cell lines are popular *in vitro* models of breast cancer, we show here that their NER capacity is not necessarily indicative of *in vivo* tumors. When these breast cancer cell lines are considered against the early stage breast tumor data that we have presented, there is a statistically significant difference. The repair phenotype seen in early stage tumors is lost in this group. Only SK-BR-3 and CaMA-1 are even within the range demonstrated by the early stage breast tumors and by z-test, they are statistically different than the mean of the early stage tumor group. This supports the idea from Chapter 3 that late stage tumors are not good models for the study of cancer etiology, in that there are different phenotypes associated with late and early stage tumors. When considered in the context of their stage, these breast cancer cell lines, like our stage III and stage IV chemotherapy naïve data, demonstrate that late stage tumors do in fact show an increased repair capacity relative to breast reduction. Thus, they may serve as adequate models for progression of breast cancer.

When compared to primary cultures derived from women who were undergoing breast reduction mammoplasty surgery these established cell lines were between 122% and 344% of the average breast reduction repair capacity. SK-BR-3, CaMa-1, and MCF7 were all within the range of the breast reduction primaries, however MCF7/Ly2, BT-20, and MDA MB231 were all well above the highest repairing breast reduction cultures. Thus, a study of repair capacity with these cell lines relative to the tissue specific control of breast reduction epithelium would yield

the result that breast cancer has competent, and even increased NER capacity. However, repair fidelity in tumors versus nondiseased cells was not evaluated, only total genomic repair.

When the growth rate of the cell lines was considered, there was no association between increased proliferative index and decreased DNA repair. This is again inconsistent with the mitogen-mutagen hypothesis, *i.e.,* the idea that as the cell cycles more rapidly, DNA repair is compromised and the cells are at greater risk of DNA damage, and consistent with our data as represented in Chapter 3.

4.4.2 MCF10A Usage as "Normal"

Although this cell line is used as a normal, nontumorigenic control, its repair is drastically different than the repair of breast reduction primaries. It is unknown whether the difference in repair can be attributed to passaging, increased immortalization steps in the MCF10A cells compared to the entirely normal breast reductions, clonal evolution, or a selection after passaging. The change is karyotype from the original is evidence of the evolution of this culture over time. This repair capacity data calls into question the utilization of this cell line as a normal control.

4.4.3 Tumor Cell Lines Versus Primary Cultures

4.4.3.1 Time in Culture

Differences between the cell lines and the primary tumor cultures could be attributable to many factors. These cell lines are also a clonal outgrowth of one or a few cells that were present in the original tumor that were more permissive for *in vitro* growth, therefore only representing only a small portion of the heterogeneous cells that are included in a tumor *in vivo*, while the primary cultures are a conglomeration of many cell types. Perhaps only tumors with normal or enhanced NER were able to grow in old culture system(s). Even if the initial culture had a low NER capacity, growth in culture may select for higher NER. Conditions within the culture systems used to generate these cell lines are quite different from those seen in the human body.

4.4.3.2 Staging

Also of consideration is the stage of the tumors used to make the cell lines, as compared to the tumors used to generate the primary cultures considered in this study. The cell lines have mostly been generated from stage IV pleural effusions, with one line, BT-20, having been generated directly from a stage IV tumor in the breast. However, even BT-20 is significantly higher in its repair capacity than the primary cultures, and it is not significantly different than the other tumor cell lines used in this paper with the exception of the three lowest ones. Later stage tumors arise from the most aggressive subpopulations of tumor; may be related to normal or even enhanced DNA repair. However, since the NER capacity in the cell lines is so much higher, this phenomenon could also be the result of gene amplification of the genes involved in excision repair.

4.4.3.3 Drug Treatment

Our breast tumor primary cultures have also been harvested from women who have not yet received any treatment for their breast cancer. However, several of the established tumor lines were derived from women who received some form of treatment, be it hormonal therapy, chemotherapy, or radiotherapy. It is possible that the cell lines derived from these women have undergone a selection for therapy-resistant clones, leading to a higher repair capacity in the resultant outgrowths.

4.4.3.4 Tissue Specificity

Another consideration is that since repair capacity is tissue specific, the breast reduction primary cultures demonstrate a tissue specific repression of repair capacity (Chapter 3). Breast reduction explants had a lower repair capacity than foreskin fibroblasts, which manifest a phenotype consistent with their direct exposure to UV during lifetime, *i.e.,* greatest necessity for protection from UV-induced DNA lesions. Such tissue specificity of gene expression is controlled by epigenetic activation and inactivation based on DNA methylation (or lack thereof). DNA methylation is not faithfully reiterated in transformed cells, so established cell lines may be showing eventual de-repression to FF levels of expression.

4.4.4 Comparison of Established Breast Cancer Cell Lines

The small number of cell lines used here prevent a full analysis of the different characteristics of the tumor cell lines used. Preliminary evaluation of ER status based upon the ER+ cell lines yielded no definitive correlation with regard to repair relative to ER- cell lines (MCF7 and

CaMa-1 are lower than MDA MB231 and BT-20 but higher than SK-BR-3). The MCF7/Ly2 subline is also not statistically significantly different that its parental line, suggesting that resistance to antiestrogens does not effect repair capacity, although this is only one example. Whether the cell line is generated from the tumor of origin or from a malignant pleural effusion does not significantly effect DNA repair, in that BT-20 falls within the range generated by the pleural effusion samples. Again, we are only considering a single sample. This is difficult to ascertain, however, as the tumors of origin for the pleural effusion cell lies are not available for analysis.

4.4.5 Representation of the Spectrum of Breast Cancer

One of the variables of human breast cancer that is not accounted for in commonly used breast cancer cell lines was race. All of these cell lines were derived from Caucasian women (Table 4.1). Studies have shown that there is a marked difference in African American breast cancer as compared to Caucasian breast cancer this includes age at diagnosis, stage at diagnosis, ER status, correlation of Her-2/Neu status with outcome, and overall survival (see Chapter 1). Thus, the biology of disease in black women is not represented in the studies which exclusively use these cell lines. The paucity of cell lines generated from African American subjects can likely be traced to a number of factors including psychosocial and ethical issues of medicine in the African American community.

Two cell lines from African American women with breast cancer exist in the literature. The first was created in 1979 from members of the same group that created MDA MB231. This was created from the pleural effusion of a chemotherapy-treated 51-year old African American

woman (402). A cell line was more recently derived in 1997 from a 27-year old African American breast cancer patient using transformation with SV-40 (403). In addition to these two minority patients, a cell line has been characterized that represents the tumor of a Native American woman with breast cancer (404). While it is important that such cell lines are being generated, the limitations of cell lines from pleural effusions or with viral immortalization dampens enthusiasm.

4.4.6 The Latimer Culture Method and Representation of Breast Cancer Heterogeneity

As discussed in Chapter 2, the inability of current culture systems to efficiently grow cells from both normal and diseased breast prohibits the field from being able to evaluate *in vitro* the characteristics of more than a few breast tumors. Breast cancer is a heterogeneous disease, with great diversity in tumors. Some of the variability of tumors has been described by microarray analysis. Sorlie *et al.* has used microarray analysis on tumor tissue to show that breast cancer heterogeneity can be expressed as five different subtypes based upon expression profiles (147,148). These cell lines fail to take into account the diversity of breast cancer represented in these studies. Our own microarray analysis has shown that there is also heterogeneity in the cell lines being generated with the Latimer culture method.

Some cell lines, like JL DCIS-3B show more of a relationship with the ER- MDA MB231 than with the other cell lines generated using our culture method. Others, like the stage III cell line JL BTL-12 resemble the cervical carcinoma line HeLa229 even more than they resemble other lines from the breast. The majority, however, are comparable to one another, but are not perfectly related. There are layers of inter-related mRNA expression profiles that will be analyzed in the

future with supervised analyses including the use of the intrinsic gene set used to define the 5 breast cancer subtypes by Sorlie *et al.* (147,148), and previous studies of breast cancer heterogeneity performed with microarray (405).

Not only is the Latimer culture system capable of generating both nondiseased and tumor primary cell culture, but it is also able to create novel cell lines derived from passaging these primary cultures. The success rates seen with the Latimer system in generating both primary culture and cell lines opens up the possibility of studies of larger numbers of breast cancer patients. Cell lines from multiple patients, representing more of the heterogeneity of breast cancer can be created with this system. Thus, the culture system is not only more physiologically relevant for the characteristics discussed in Chapter 2, but more relevant to the field of breast cancer for its creation of a way to study a much broader subset of breast cancers.

4.4.7 Latimer Cell Lines are Consistent with their Tissue of Origin

As discussed in Chapter 2, problems in the field of mammary biology have included the culture diploid cells from karyotypically abnormal breast tumor tissue, suggesting an overgrowth of normal, non-breast epithelial cells in these cultures. We have shown here, however, that not only do the cell lines we have generated from tumor tissue have a grossly abnormal karyotype, they are still able to express cytokeratin 18 as an intermediate filament. Both of these pieces of evidence support their identity as malignant luminal epithelial cells. Furthermore, nondiseased epithelium maintains a normal karyotype in cell culture including both of its X chromosomes. Although the karyotypes are consistent, this may not be true with all internal markers.

5.0 MOLECULAR ANALYSIS OF NUCLEOTIDE EXCISION REPAIR

5.1 INTRODUCTION

5.1.1 Molecular Mechanisms of Nucleotide Excision Repair Capacity in Breast Cancer

In order to understand the molecular mechanism(s) behind the functional loss of DNA repair capacity that was seen in early stage breast tumors, we began by analyzing the steady state mRNA expression levels of genes in the NER pathway. To evaluate the role that expression of these genes played in cancer etiology, the steady state mRNA levels of 20 genes in the NER pathway were evaluated in three breast reduction cell lines relative to three early stage tumor-derved cell lines. To evaluate the role that expression of these genes could play in cancer or tumor progression, these same genes were evaluated for their repair capacity in a Latimer-derived stage III breast cancer cell line and the established stage IV cell lines MCF7 and MDA MB231.

The multiplex RNase Protection Assay offered the ability to detect the expression of 10 different genes in the NER pathway in each of two kits (20 genes total) using only 2 μg of RNA from each cell line. In contrast, previous efforts in the Latimer laboratory to quantify ERCC1 mRNA levels in these cell lines by Northern analysis required 30 μg of total RNA. It has been reported that the XPA gene is expressed at such a low level that Northern blot analysis of total

RNA cannot be completed (406). Thus, Rnase protection allows for the quantification of the expression of multiple genes with less than a tenth of the RNA needed for Northern analysis, provided that Northern analysis would even be feasible.

In order to evaluate other genes that are involved in the NER pathway and NER-related genes, we used microarray analysis with the Affymetrix U133 2.0 Plus chips. Not only were we able to look for loss of expression of these additional genes, but we were also able to observe the overall expression profiles of the cell lines, as shown in Chapter 4. Furthermore, microarray results provided an independent analysis of the mRNA expression levels of the genes we had analyzed by RNase protection.

Protein expression for some genes in the pathway was also quantified by Western analysis. Since it is protein that is most closely related to function, it was important to evaluate this level of expression. This is also important for context within the field of DNA repair, which is primarily driven by investigations at the protein level. DNA repair mechanisms have been clarified through studies of protein-protein interactions in cell-free *in vitro* systems.

Since the expression data we observed was consistent with an epigenetic mechanism of regulation, we altered the methylation status of early stage tumor-derived cell lines via treatment with the demethylation agent 5-aza-2'deoxycytidine. We also used concurrent treatment with the histone deacetylase inhibitor Trichostatin A to alter chromatin structure. Alterations in mRNA steady state expression were measured using both RNase protection and microarray analysis. For genes that were not controlled by methylation, we used transcription factor analysis of the promoter regions of these genes to determine possible mechanisms of regulation.

5.1.2 Studies of mRNA Expression in Cancer Using Cancer Cells

Some reports in the literature have studied the possible concurrent regulation of NER genes. Quantitative RT-PCR analysis of expression in hepatocellular carcinoma cells reported that mRNA expression of the NER genes ERCC1, XPA, and XPC gene were increased in hepatocellular carcinoma compared to normal liver tissue (407). The same study showed that there was no change in four other genes in the pathway, CSA, CSB, hHR23B, and XPB. While this study did specify that patients were chemotherapy-naïve, they did not discuss the staging or grading of these patients.

In one study of ovarian cancer, RT-PCR was used to determine the expression levels of ERCC1 and XPB relative to different histologic subtypes. These researchers found that these two genes were expressed two-fold higher in clear cell ovarian tumors than in the other subtypes tested (serous, mucinous, poorly differentiated, and endometroid) (408). These researchers, however, did not include samples of normal ovarian epithelium, and thus whether this result is correlated with a gain of repair in tumors is unknown.

One study looked at the protein expression levels rather than the mRNA expression levels to screen for decreased NER genes in cancer. This group, however, compared protein levels between cancer cell lines from cancers that are typically sensitive to platinum-based therapies and cell lines from cancers that are not. Platinum-based therapies generate inter- and intrastrand crosslinks that are remediated by the NER pathway. This study demonstrated that ERCC1, XPF, and XPA proteins were lower in protein expression in sensitive testis tumor cell lines than in cell lines from prostate, bladder, breast, lung, oral, cervical, and ovarian cancers that are less sensitive to treatment with platinum-based agents (406). When mRNA expression of these same

cell lines was assayed to see if there was also an association with loss of the steady-state transcripts, no relationship was seen. This study also does not compare any of these tumors with normal organotypic tissue.

The mRNA expression levels of ERCC1 in breast cancer biopsies was examined by relative quantification reverse transcriptase PCR. However, results using this technique have not been published as of yet, only the technique itself was reported (409).

A summary of these studies of mRNA expression of NER genes in cancer is presented in Table 5.1.

5.1.3 Studies of mRNA Expression in Cancer Using Peripheral Blood Lymphocytes

While some previous studies have used PBLs for the analysis of repair capacity (70,323,351), results from this laboratory have shown that repair capacity is tissue-specific and that PBLs may not, in fact, faithfully represent the relative repair capacity of the tissue of interest (140). With this caveat in mind with regards to functional analysis, there are several studies which have reported the use of PBLs as a surrogate in analyzing the mRNA expression of NER genes in cancer patients. Another caveat that must be considered in the evaluation of this literature is that tumor stage-specificity was never considered in any of these studies. The researchers did not look at what the stage of cancer in the patient who donated the lymphocytes was. They therefore mix together and confuse the separate issues of etiology and progression.

A large prospective study of lung cancer in Denmark resulted in finding 265 individuals in its cohort developed lung cancer. The mRNA expression levels of ERCC1 and XPD in lymphocytes from these patients plus another 272 matched controls were measured using

quantitative PCR and no association was found between expression levels of these two genes and the disease (410) (Table 5.1). A similar study was conducted on PBLs from patients with squamous cell cancers of the head and neck compared to cancer-free controls. Multiplexed RT-PCR was used to assess the expression of 5 different NER genes and significant decreases in expression in cases versus controls were found for the ERCC1, XPB, XPG, and CSB genes, but not for XPC (411). This same group also analyzed lymphocytes in lung cancer cases relative to controls and found a decrease only in XPG and CSB (256).

Table 5.1 Summary of mRNA expression studies of NER genes in tumor versus nondiseased tissue. Stage of the cancer of interest was not considered or analyzed in any of these studies. All studies were conducted using quantitative reverse transcriptase PCR.

Gene	mRNA Expression in Cancer Relative to Control	Cancer	Tissue Used for Analysis of mRNA Levels
XPA	Up (407)	Hepatocellular carcinoma	Carcinoma
XPB	Same (407) Same (256) Down (411)	Hepatocellular carcinoma Lung Squamous cell cancer of the head and neck	Carcinoma PBLs PBLs
XPC	Same (411) Same (256) Up (407)	Squamous cell cancer of the head and neck Lung Hepatocellular carcinoma	PBLs PBLs Carcinoma
XPD	Same (410)	Lung	PBLs
XPG	Down (411) Down (256)	Squamous cell cancer of the head and neck Lung	PBLs PBLs
ERCC1	Same (256) Same (410) Up (407) Down (411)	Lung Lung Hepatocellular carcinoma Squamous cell cancer of the head and neck	PBLs PBLs Carcinoma PBLs
CSA	Same (407)	Hepatocellular carcinoma	Carcinoma
CSB	Same (407) Down (256) Down (411)	Hepatocellular Carcinoma Lung Squamous cell cancer of the head and neck	Carcinoma PBLs PBLs
hHR23B	Same (407)	Squamous cell cancer of the head and neck	PBLs

5.1.4 Methylation as Epigenetic Regulation in Breast Cancer

Methylation of tumor suppressor genes is a known phenomenon in many types of cancer including breast cancer. Multiple genes have been shown to be methylated in breast cancer, including ER and PR, E-cadherin, 14-3-3 σ, and several genes involved in DNA repair: BRCA1 (216) and the mismatch repair genes hMLH1 and HMSH2 (412). Disruption of each of these tumor suppressor genes plays a role in the transformation process. Loss of responsiveness to estrogen curtails the effectiveness of endocrine therapy and allows the cells to find new modes of estrogen-independent growth. Loss of E-cadherin expression is implicated in the epithelial to mesenchymal transition, believed to be part of the mechanism by which cancer becomes invasive. Loss of DNA repair genes prevents the cells from being able to deal with genotoxic insults either from exogenous agents or endogenous processes and increases genetic instability.

DNA methylation has been also implicated in the expression of three NER genes in the IL-6 responsive multiple myeloma cell line KAS-6/1 (258). Researchers hypothesized that pro-inflammatory cytokines may be mediating an effect on transcriptional regulation in these cancer cells through an epigenetic mechanism. They therefore chose to look at the effect of IL-6 methylation. Because there are a number of genes involved in DNA repair that are affected by methylation in cancer, they chose to used RNase protection as a way to survey multiple genes involved in DNA repair. Using the Pharmingen hNER1 RNase protection assay kit, they showed that the expression of hHR23B, RPAp32, and DDB1 genes were all increased by the treatment of these cells with the methylation inhibitor zebularine. Zebularine, like the 5-aza-2'deoxycytidine used in our assays, is another nucleoside analogue that is able covalently bind the DNMT1 enzyme after incorporation into DNA. Furthermore, methylation of the hHR23B gene was

restored, and expression decreased, when the drug was removed and IL-6 was added to the culture. However, these researchers did not discuss whether there had been a decrease in expression of these genes relative to a nondiseased line, nor did they discuss any functional assays of DNA repair activity performed on these cells.

5.1.5 Modeling Transcription Factors

Software programs exist to search DNA sequences for the presence of defined sequences and sequence elements such as known transcription factor binding sites. Analysis can be as simple as the identification of how many potential binding sites exist within a single DNA sequence to complex models linking multiple genes through multiple transcription factors. The specificity of a regulatory unit (set of *cis*-elements) is usually determined by the spatial organization of binding sites. Therefore, programs now exist that determine not only the number of binding sites, but also their context (*i.e.,* which sites and their location in the sequence): PROMOTER SCAN, FunsiteP, GeneLang, GeneParser, GRAIL, and Frameworker (193).

5.2 MATERIALS AND METHODS

5.2.1 Lipofection

Cells were plated to achieve 90-95% percent confluence in a 6-well cluster dish (Corning). For each well a ratio of 25:1 experimental pGL3 plasmid (Promgea) to control pRL-TK plasmid (Promega) was used. 5 µg total plasmid DNA was diluted in 250 µl DMEM per well. Three wells of each cell line were transfected with DNA in the presence of Lipofectamine 2000, and three wells were mock transfected with plasmid DNA only and no transfection reagent. 15 µl Lipofectamine 2000 was mixed with 250 µl DMEM per well and allowed to incubate at room temperature. The Lipofectamine DMEM was never allowed to sit for more than 5 minutes at room temperature to prevent inactivation of the Lipofectamine agent. After this incubation, the diluted Lipofectamine was added to the diluted DNA to yield 500 µl of DMEM with a combination of Lipofectamine, experimental plasmid and control plasmid per well. For the mock transfected wells, the diluted DNA mixture was combined with DMEM only. These mixtures were incubated at room temperature for 20 minutes to allow DNA-Lipofectamine complexes to form. When the incubation was complete, medium was removed from each of the cell wells and replaced with 500 µl of the desired mixture: DNA-Lipofectamine for the experimental wells and DNA in DMEM for the controls. These cells were incubated at 37^{o}C in a 5% CO_2 incubator for 1 hour. After the 1 hour incubation, 1.5 mL media with serum was added to each well.

At 44 hours post transfection, cells were harvested according to the instructions in Promega's Technical Manual No. 040, Dual Luciferase® Reporter Assay System. Briefly,

medium was removed from each of the wells and the cells washed with phosphate buffered saline (PBS, CellGro). 500 µl 1x passive lysis buffer (PLB) was added to each well and incubated for 20 minutes with agitation. PLB is a proprietary creation of Promega generated for use with the Dual Luciferase Assay Kit. It contains detergents and reducing and chelating agents. The formula is not available from Promega. Cells were scraped from the bottom of the wells with a rubber policeman halfway through the incubation. Lysates were collected into microcentrifuge tubes and cleared by centrifugation to provide cell lysates free of unlysed cells, genomic DNA, membrane lipids, *etc*. Protein concentrations were quantified by Pierce's BCA Protein Assay Reagent Kit against a standard dilution set of BSA in accordance with Pierce's BCA Protein Assay Reagent Kit manual.

Luciferase activity was assayed using Promega's Dual Luciferase® Reporter Assay System. Luciferase Assay Substrate and Stop & Glo® Substrate were each added by automatic injection by the Zylux Femtomaster FB15 luminometer (Maryville, TN). Measurement time was set to 10 seconds for each reaction. Three seconds were allowed for maximal activity of the firefly luciferase (experimental plasmid) and Fourteen seconds were allowed for quenching of this reaction and maximal activity of the sea pansy luciferase (control plasmid). Background luminescence was measured against triplicate samples of 1x PBL lysis buffer. 26 ng of protein from each well (three experimental and three mock transfected) was quantified in triplicate. Output was represented in Relative Light Units (RLUs).

Average background luminescence was subtracted from each reading. Triplicate readings from each well were averaged and then the averages across the three replicate wells were calculated. Finally, the mock transfected well average was subtracted from the experimental wells.

$$\text{RLU per well} = \frac{\text{pGL3 luminescence} - \text{background 1x PBL luminescence}}{\text{pRL-TK luminescence} - \text{background 1x PBL luminescence}}$$

RLU for a Cell Line = Average RLU per well – Average RLU for mock transfected

5.2.2 RNA Isolation

Total RNA was isolated from cells using a standard guanidium isothiocyanate cesium chloride preparation method. Confluent monolayers of cells were washed with PBS to remove dead cells and residual growth medium. Cells were then lysed in the presence of a solution containing the RNase inhibitor guanidium isothiocyanate (Fluka), and the detergent sarkosyl (Sigma). Cells were dounced to ensure the lysis of the nuclear membrane. This cell lysate was then run through a 5.7 M cesium chloride (Sigma) gradient at 35,000 rpm for 16 hours to yield pelleted RNA. Finally, RNA was harvested using ethanol precipitation.

Quantification of the RNA concentration was determined via a spectrophotometer by absorbance at 260 nm. The integrity of the total RNA was checked for each sample through the use of a 1.5% agarose 17% formaldehyde gel run in 1M sodium phosphate pH 6.5. The 28S and 18S ribosomal RNA bands were visualized using ethidium bromide on a transilluminator. Images were captured on Polaroid Positive Negative film and the bands on the negative film were quantified using scanning densitometry using a control RNA that was fully intact as a standard of comparison. Samples were snap-frozen stored at -80°C.

5.2.3 RT-PCR

Total RNA was reverse transcribed with the Ambion RETROScript kit using random decamer primers (Ambion Inc., Austin, TX) at 42°C for 45 minutes. Controls for possible contaminating genomic DNA were performed by the omission of the MMLV-RT enzyme in the RT reaction. The resultant cDNAs were subjected to PCR with Ambion SuperTaq enzyme for 35 cycles on the Eppendorf Mastercycler epgradient S thermocycler. Specific conditions for each primer are listed below.

ERα RT-PCR primers

The amplimer size for ER was 650 bp.

ERα Primer Sequence (413)

FORWARD: 5' TACTGCATCAGATCCAAGGG 3'

REVERSE: 5' ATCAATGGTGCACTGGTTGG 3'

β casein RT-PCR Primers

The amplimer size for β casein was 320 bp.

The β Casein Primer sequence was generated in the Latimer Lab:

FORWARD: 5' GGT GGC TCT TGC TCT TGC AAG G 3'

REVERSE: 5' GGC ATC ACT CTG CCC TTA GTG 3'

5.2.3.1 RT-PCR for 14-3-3 σ

RNA samples that were used for RT-PCR with primers for 14-3-3 σ were also subjected to DNase I treatment because this was an intronless genes. This prohibited the design of primers

that would span multiple exons, differentiating the genomic DNA template from spliced mRNA. Total RNA harvested with the above procedure was bound to a silica gel membrane from the Qiagen RNeasy micro kit (Qiagen, Valencia, CA). Bound RNA was then incubated at 37°C for 30 minutes with a 27 U of RNase-free DNase I (Qiagen). The RNA was eluted from the column and bound to a second column for a subsequent incubation with DNase I for 15 min at 37°C. The RNA samples were eluted with RNase-free water and used immediately for reverse transcription.

14-3-3 σ RT-PCR primers

The amplimer size for 14-3-3 σ was 279 bp.

14-3-3 σ Primer Sequence (360):

FORWARD: 5' GTG TGT CCC CAG AGC CAT GG 3'

REVERSE: 5' ACC TTC TCC CGG TAC TCA CG 3'

5.2.3.2 Controls for the presence of cDNA

The presence of cDNA (and RNA) for all genes analyzed by RT-PCR was confirmed by the use of control primers targeting the rig/S15 gene (414). The same cDNA samples were subjected to PCR using these primers.

Rig/S15 RT-PCR Primers

The amplimer size was 361 bp.

Rig/S15 Primer Sequence:

FORWARD: 5' TTCCGCAAGTTCACCTACC 3'

REVERSE: 5' CGGGCCGGCCATGCTTTACG 3'

5.2.3.3 Gel Electrophoresis for RT-PCR Products

The amplified products for all RT-PCR reactions were analyzed on a 2.0% agarose gel (SeaKem) with ethidium bromide (Sigma) staining for visualization on a UV transilluminator. Images were captured on Polaroid film and a BioRad chemidoc system using Quantity One software.

5.2.4 RNase Protection Analysis

The multi-probe RPA kits hNER1 and hNER2 (BD Riboquant, Pharmingen, San Jose, CA) were utilized to analyze the mRNA expression of 20 genes in the NER pathway. The hNER1 contained probes designed for the genes XPC, hHR23B, XPA, RPAp70, RPAp32, RPAp14, XPG, XPF, DDB1, and DDB2. The hNER2 kit contained XPB, XPD, TFIIHp34, TFIIHp44, TFIIHp52, cdk7, CycH, ERCC1, CSA, and CSB.

Radiolabeled (^{32}P-UTP, NEN Perkin Elmer Life Sciences, Wellesley, MA) anti-sense RNA probes were synthesized using the Pharmingen In Vitro Transcription Kit. pPMG vectors containing sequences of the sense strand of each gene downstream of a T7 promoter were transcribed using the DNA-dependent T7 RNA polymerase. This reaction was carried out in the presence of unlabelled ribonucleoside triphosphates, radiolabeled ^{32}P-UTP, and an RNase inhibitor. Probe strength was determined as CPM both by Cherenkov counting and scintillation counting.

RNA samples were dried to dust in a vacuum evaporator and resuspended in a hybridization buffer (80% formamide, 1 mM EDTA, 400 mM NaCl, 40 mM PIPES, pH 6.7) provided by Pharmingen. Control samples in each experiment included two replicates each of

RNAs isolated from the cells used as controls in our UDS analyses: FF, MCF-7, and MDAMB 231. Two replicates of yeast tRNA was used as a negative control. 2 μg of each sample was used. In addition each experiment contained a dilution series of control RNA. Desitometric quantification was performed against a serial dilution of the provided RNA standard, comprised of 2 μg, 0.67 μg, 0.22 μg, and 0.07 μg of human control RNA-2 (a control RNA provided by Pharmingen generated from pooling total RNA from three cell lines, NC-37, Daudi, and THP-1).

The labeled antisense probes were allowed to hybridize to the total RNA samples at 56 °C for 16 hours. RNase A was then incubated with each sample for 45 minutes at 30°C to digest unhybridized probe and other single-stranded "unprotected" RNAs. After Proteinase K treatment for 15 min at 37°C to remove protein and a phenol:chloroform extraction to remove protein and destroy any of the remaining enzyme activities, protected probes were collected by ethanol precipitation and resolved by denaturing gel electrophoresis (40 cm, 0.4-mm thick 4.75% acrylamide gel at 40 Watts for 2.5 hours). Unhybridized probe was run as a size standard on each gel. These acrylamide gels were transferred to Whatman 3MM chromatography paper, dried under a vacuum at 80°C for 1.5 hours, and allowed to expose Kodak AR Scientific Imaging Film 35cm x 43cm. Exposure times varied for each gene-specific band based on the band intensities.

Band identities were confirmed using the unhybridized probe size standard. Scanning densitometry was used to quantify results from the RNase Protection Assays using the Harmony program. Each band was individually quantified. Local background under each band was subtracted from the band itself. Normalization of band intensities was performed using the densitometric analysis of the GAPH housekeeping gene. Using an exposure showing a linear relationship for the GAPDH gene in our dilution series, a "normalization factor" was derived for

each lane relative to the human control RNA-2 2 µg standard. This normalization factors was generated by dividing the intensity of the GAPDH band for each sample by the intensity of the human control RNA-2 2 µg sample. Because of the range of intensities that were seen for the bands corresponding to different genes, only the exposure that demonstrated a linear relationship between the human control RNA-2 dilution series was also chosen for quantification of that gene's bands. Once the intensities for the bands corresponding to all genes were quantified from appropriate exposures, these values were normalized for loading with the normalization factor determined from the GAPDH loading control. Next, the normalized intensity of the bands for each gene were expressed relative to the average of the FF control. Normalization with the GAPDH control accounted for loading differences in the samples and normalization against the FF control allowed for comparison across experiments. Finally, gene expression was listed relative to the average of the three breast reduction cell lines analyzed.

5.2.4.1 Data Analysis

At least three independent experiments were run for each sample with each of the two kits used. In general, the total results reflect approximately 10 experiments (gels) per cell line. Three breast reduction cell lines, three early stage tumor-derived cell lines, one stage III cell line, MCF7, and MDA MB231 were used. At least two, and sometimes three independent harvests of RNA were used to generate this data. The average of the normalized intensities relative to FF were created first within a cell line sample. Standard error was also calculated within the individual cell lines representing experimental variation. Averages of the sample groups were then also generated: (1) three breast reduction cell lines; (2) three early stage tumor-derived cell lines; and (3) three late stage tumor-derived cell lines. Standard error between the three samples

present in each of these groups was also calculated to represent inter-individual variation. Comparisons between groups were done using 2-tailed t-tests assuming unequal variance with α set to < 0.05 for significance. Comparisons between a population of cell lines (*e.g.*, breast reduction lines or early stage tumor-derived lines) and a single cell line were carried out using a 2-tailed z-test with $\alpha < 0.05$.

5.2.5 Microarray Analysis

Microarray expression analysis was carried out at the Clinical Genomics Facility of the University of Pittsburgh Cancer Institute by Dr. Federico Monzon using the Affymetrix U133 2.0 Plus chip. This chip set comprises 54,000 probe sets to analyze the expression of 47,000 transcripts and variants, including 38,500 well-characterized human genes. Sequences used in the design of the array were selected from GenBank® http://www.ncbi.nlm.nih.gov/Genbank/, dbEST http://www.ncbi.nlm.nih.gov/dbEST/, and RefSeq http://www.ncbi.nlm.nih.gov/RefSeq/. Total mRNA was labeled and hybridized to the chip set. Results were captured using a GeneChip Scanner 3000 with GeneChip Operating Software (GCOS).

Cell Lines were first analyzed using unsupervised hierarchical cluster analysis to observe relationships between them with Avadis software. In the 2-dimensional cluster analysis, gene clustering and cell line clustering was performed independently using an agglomerative hierarchical clustering algorithm. For gene clustering, pairwise similarity measurements among genes were calculated on the basis of expression ratios across all significant genes.

Analysis of the relationships of genes between specific sets of cell lines were carried out by both using a t-tests ($\alpha<0.05$) and SAM tests ($\alpha<0.05$) by Dr. Uma Chaundran.

Specific genes were also interrogated using microarray analysis data. All probe set IDs corresponding to the gene of interest were averaged. Normalization relative to the average of the two probe set IDs corresponding to GAPDH was then performed, similar to the way that RNase protection was normalized for sample loading. Samples were initially expressed relative to FF, as they had been in both UDS and RNase protection assays. Finally, gene expression was reported as relative to the average of the three breast reduction cell lines that were tested.

5.2.6 Total Protein Isolation

Total protein was harvested from snap-frozen cell pellets using whole cell lysis in the presence of Triton. Cells were incubated for 30 minutes on ice in an equivalent volume of extraction buffer containing 1x protease inhibitors (Sigma), and 0.1% Triton X-100, in a base of 0.05 M Tris pH 7.5, 0.25 M NaCl, and 0.001 M EDTA. The lysate was cleared of unlysed cells, lysed membranes, and genomic DNA by centrifugation at 13,000 rpm for 20 min at 4°C. The supernatant was collected, snap frozen, and stored at -80°C.

Quantification of protein concentration was done by a Bradford Protein Assay (BioRad, Hercules, CA). 1 µl of each sample was mixed with 800 µl water and 200 µl BioRad Bradford Assay reagent. Bovine Serum Albumin (Sigma) controls of 1, 3, 6, 9, and 12 µg/ml were also prepared to generate a standard curve. After 5 minutes of incubation at room temperature, absorbance at 595 nm was read for all samples. The BSA standards were used to generate a standard curve and the concentrations of each of the samples were determine by interpolation on this graph. Proteins that were more concentrated than 10 µg/µl were diluted.

5.2.7 Western Blot

Western analyses was carried out for the CSB, CPA, DDB1, and DDB2 genes. These genes were chosen because antibodies that recognized these proteins were available, whereas antibodies for many of the NER pathway genes are either not commercially available or of poor quality. A commercially available antibody was present for the CSB protein, one of our most dramatically lost genes in terms of mRNA expression. The XPA antibody was the generous gift of Dr. Richard Wood who also provided information about the optimization of the use of this antibody. Antibodies for the DDB1 and DDB2 proteins were used in collaboration with the laboratory of Dr. Vesna Rapic-Otrin.

Total protein was resolved using an 8% SDS glycine polyacrylamide gel. A 4% SDS glycine stacking gel was added to the top of each resolving gel. Controls on each gel included lysate from FF, similar to the RNase protection experimental design, and, for experiments with DDB1 and DDB2 probing, the normal human fibroblast line GM5757 (Corielle Cell Repositories), a standard practice in the laboratory of Vesna Rapic-Otrin where this work was performed. A high-range rainbow molecular weight marker (Amersham) was used as a size standard on each gel. Proteins were denatured at 100°C for 10 minutes in the presence of a loading buffer: 50 mM Tris pH 6.8, 1% SDS, 4 M Urea, 0.03% pyronine dye and 0.03% B-mercaptoethanol. The pyronine dye front was used to track the migration of proteins through the gel. Electrophoresis was carried out at 120 V in Tris Glycine SDS (25 mM Tris, 192 mM glycine, pH 8.3).

After gel resolution, the proteins were transferred in the presence of a tris glycine methanol buffer (0.5x TG (12.5 mM Tris, 96 mM glycine, 0.05% SDS, pH 8.3), 20% methanol)

to a PVDF membrane (Immobilon-P). This membrane was washed to remove the methanol in either Phosphate-Buffered Saline (PBS) or Tris-Buffered Saline (TBS), and blocked (Table 5.2). Blocks for DDB1, DDB2 and tubulin were carried out for 1 hour at room temperature while blocking for CSB and XPA were carried out overnight at 4 degrees C. Primary and secondary antibodies were prepared in the appropriate block. Primary antibodies were applied overnight at 4 degrees C for DDB1, DDB2 and tubulin and 1 hour at room temperature for CSB and XPA. All secondary antibody probes were carried out at room temperature for 30 minutes.

In order to visualize both the DDB2 and α-tubulin or the CSB and the DDB1 on the same membrane, the PVDF was stripped using a glycine stripping buffer (1.5% glycine, 0.1% SDS, and 1% Tween-20, pH 2.2) for 1 hour at room temperature. The membrane was then blocked and probed as described above.

Table 5.2 Conditions used for western blot antibodies.

Protein	**Block**	**Primary Antibody**	**Primary Antibody Dilution**	**Secondary Antibody**	**Secondary Antibody Dilution**	**Size of Protein on 8% tris-glycine gel**
CSB	5 % milk TBST	Santa Cruz sc-25370	1:1000	Goat anti-Rabbit- AP Biorad	1:20,000	168 kDa
DDB1	2% I-block	Santa Cruz sc-16292	1:5000	Goat anti-Rabbit- AP Biorad	1:20,000	127 kDa
DDB2	2% I-block	Santa Cruz sc-16295	1:1000	Goat anti-Rabbit- AP Biorad	1:20,000	48 kDa
XPA	20% milk TBST	XPA bleed5, the generous gift of Dr. Richard D. Wood	1:10,000	Goat Anti-Rabbit-HRP Biorad	1:100,000	36 kDa
Tubulin (loading control)	2% I-block	Santa Cruz sc-5546	1:800	Goat anti-Rabbit- AP Biorad	1:20,000	45 kDa

Visualization for alkaline phosphatase conjugated secondary antibodies was performed using the Tropix CPD Star (Tropix, Bedford, MA) assay agents. Visualization for the horse radish peroxidase conjugated secondary antibody was carried out using the SuperSignal West Femto kit from Pierce (Rockford, IL). Images were captured both on Kodak Biomax XR film and a BioRad Chemidoc system using the Quantity One software package.

5.2.8 CpG Island Mapping

Sequences of chosen NER genes from NCBI's nucleotide database were analyzed using the CpG Island Searcher program for the presence of CpG islands in the 5' regions of these genes. The original definition of a CpG island as proposed in 1987 by Gardiner-Garden and Frommer was a 200-bp stretch of DNA with a C+G content of 50% and an observed CpG /expected CpG in excess of 0.6 (209). Since then, stricter definitions of a CpG island have been adopted. The algorithm used to determine this definition is also available at the CpG Island Searcher site [http://cpgislands.usc.edu] (208). For our analyses we searched using the parameters as discussed in (211). Lower limits were set at: *greater than 55% GC*, and the *Ratio of Observed to Expected CpG dinucleotides greater than 0.65*.

Histograms showing the frequency of CpG dinucleotides were created with the MacVector software program (Accelrys Software Inc., San Diego, CA). Transcriptional start sites as denoted in the NCBI nucleotide database and CpG islands as determined by CpG island searcher were manually added to this graphical output.

5.2.9 Azacytidine and Trichostatin A Treatment

Cell Lines were plated at a density to yield 60% confluence 24 hours after plating. After 24 hours the cells were then switched to a growth medium (MWRI) without cytosine nucleotide supplementation to clear their nucleoside pools of cytosine and facilitate the uptake of the 5-aza-2'deoxycytidine and β-mercaptoethanol to allow for growth assessment using the MTT assay. After 24 hours, 5-aza-2'deoxycytidine (Sigma) was added to the experimental cells at final

concentrations of 1 μM and 5 μM. Cells were allowed to grow through two doubling periods (5 days) to allow for incorporation of the cytosine analogue during replication. Images of the treated and untreated cells were captured daily. Trichostain A was either added alone or in concert with azacytidine treatment. In one experiment, 100 ng/ml (0.331 μM) Trichostatin was added during the last 15 hours of treatment to a set of cells exposed to 5 μM azacytdine or for a total of 36 hours as a single agent.

When treatment was complete, cells were harvested for RNA. An aliquot of the harvested RNA was treated with DNase I and used to determine the expression level of 14-3-3 σ by RT-PCR.

5.2.10 Promoter Analysis

The first issue that programs analyzing promoters must address is the identification of the sequence to which a transcription factor can attach. As described above in the introduction, transcription factor binding sites are 10-30 nucleotides long, yet only a small core is required for binding. Variability exists in these sequences such that transcription factors binding sites cannot be listed by a simple sequence for recognition. Therefore, the field has now advanced to programs like MatInspector, which uses weight matrices to account for variability. These weight matrices represent the possible binding sequences of a transcription factor with their estimated binding affinity (415). A threshold score is determined for binding to occur and applied to the sequence of DNA being analyzed. If the DNA sequence being analyzed has a score above the binding affinity threshold, it is considered a potential binding site. Most families of transcription factors correspond to only one matrix, but some transcription factors possess different binding

domains, each with its own binding site. There also may be different weight matrices describing the same site. This comes as an artifact of the different training sets used to define them.

The second issue that these programs must address is determining the context in which these transcription factor binding sites exist. Order, orientation, and distance between transcription factors have a profound effect on the structure of the transcription initiation complex (415). The combination of transcription factor binding sites is called a promoter module, defined as "two or more individual elements that act in a coordinated way with those elements arranged in a defined distance and sequential order" (415). It is the context of the sequence within the entire promoter region together with its binding affinity that must ultimately determine the likelihood that it functions as a transcription factor binding site.

In order to account for both binding variability and sequence context when analyzing promoters, Genomatix has created a novel program called FrameWorker. This program determines the presence of frameworks within a promoter region in order to determine if there is coregulation in coexpressed genes. It has been designed to interact with expression array analysis to subdivide groups of genes that are seen to be coexpressed. It is hoped that the identification of frameworks in their promoters will yield functionally related genes through their coregulation. Once a framework has been identified, it can also then be applied across the human genome to find other genes with the same pattern. Specificity for models found in this manner can be determined by applying the framework to a random set of "background promoter sequence set" of 5000 human promoters. The output is a p-value score of the probability of finding the same number of sequences with a model match in a randomly drawn sample of the same size as the input sequence set. The lower the probability is, the higher the specificity of the model.

Initial promoter analysis was carried out using Genomatrix program MatInspector available at: http://www.genomatix.de/. Co-regulated candidate genes included in the Pharmingen hNER-1 and hNER-2 kits were included in our analyses. Genomic sequences for these genes including portions of their 5' untranslated regions (including their promoter regions) were used as our inputs. The 5' regulatory regions of the genes were determined using Genomatix's Gene2Promoter software. This software allows for access to their Genomatix Promoter database including promoters from the human genome that have been confirmed with 5' oligo capping experiments as determined by the company's literature searches. For genes where no such information from Genomatix was available, we searched independently in the literature for confirmation of the 5' sequence in experimental literature or information regarding the generation of transgenic mice. When no such information was available, we used Genomatix's Gene2Promoter software to determine the most likely length of the 5' promoter region by sequence analysis.

We then analyzed the 5' regions of coregulated genes for commonalities. These sequences were analyzed first for the presence of common transcription factors using the MatInspector program. MatInspector uses weight matrices to determine the presence of the a transcription factor binding site using a 2-D table of weighted consensus sequences (415). We then continued analysis by looking for frameworks involving any subset of these coregulated genes using the FrameWorker program. The program was set to look for frameworks involving as few as two sequences out of all input sequences and involving as few as two elements in the model. Elements were allowed to be as few as 5 bp in distance from one another, and a maximum of 50 bp apart. The distances for the consensus sequences are determined from an anchor position. The anchor position for the MatInspector program is defined as the central

nucleotide of the sequence to allow for maximum compatibility between sequences of uneven length. No mandatory elements were specified, meaning that we did not require our model to contain any specific transcription factor. These permissive parameters were chosen to generate the maximum number of possible models. The specificity of these models was determined within the FrameWorker program by applying the model to a "background promoter sequence set" of 5000 human promoters. The output is a p-value score of the probability of finding the same number of sequences with a model match in a randomly drawn sample of the same size as the input sequence set. The lower the probability is, the higher the specificity of the model.

Once models were generated, we used the ModelInspector program to apply them to the entire human genome. This program takes the desired model and find how many other genes within the human genome also contain them. We allowed for matches in any regions of any genes with our models. Included in the output for ModelInspector were the matrix similarity and model similarity. A perfect score for either value is 1.00. For the matrix similarity this means that each sequence position corresponds to the highest conserved nucleotide at that position in the matrix. A "good match" is defined as a similarity of > 0.80 (193). Model similarity if the percentage of the individual elements in a model that are present. A model similarity of 1.00 means that all elements of the model were found.

5.3 RESULTS

5.3.1 Latimer Cell Lines are Transfectable using Lipofection

We evaluated the transfectablity of Latimer-generated cell lines for possible applications including transfection-based assays of NER and complementation. Cell lines generated in the Latimer laboratory were seen to be transfectable using a lipofection technique (Lipfectamine2000 from Invitrogen). Cells showed transfection efficiencies similar to the established cell lines MCF 7 and MDA MB231 (Figure 5.1). The breast reduction lines JL BRL-11 and JL BRL-16, the stage I tumor line JL BTL-45, the stage III tumor line JL BTL-12 have all been shown to be transfectable by lipofection.

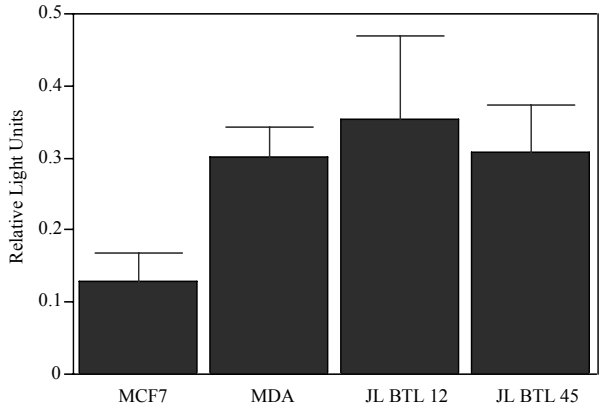


Figure 5.1 Established cell lines as well as tumor explants generated in the Latimer lab are transfectable by lipofection. Results represent the relative light units generated from firefly luciferase normalized to sea pansy luciferase, present on a cotransfected plasmid to control for transfection efficiency, measured with Promega's Dual Luciferase Assay System. Two independent experiments are averaged for each cell line.

5.3.2 Latimer Cell Lines Express Estrogen Receptor α and β Casein mRNA

RT-PCR was used to determine whether cell lines generated in the Latimer lab were able to express ERα (Figure 5.2A). The MCF7 cell line, which overexpresses Erα, was used as a positive control and the ER- cell line MDA MB231 and FF were used as a negative controls. Data presented in Chapter 3 regarding the ER and PR status of these tissues was done by collaborating pathologists in the Tissue Procurement Program at Magee-Womens Hospital. The

method they used was immunoperoxidase staining of the tissue, an antibody-based technique that detects protein. These RT-PCR analyses presented here were carried out on cell lines and detected the presence of mRNA. Three breast reduction lines expressed ERα, consistent with their tissue of origin. Tumor lines, however, varied in their concordance with the ERα protein expression determined by pathologic analysis of their tumor of origin. The tumor cell lines JL BTL-45, JL BTL-8, and JL BTL-12 all show concordance with their tumors of origin. JL BTL-2, derived from a stage II tumor, and JL BTL-35 derived from a stage IV tumor, were both determined to ER- tumors by immunoperoxidase staining of the tumor tissue during pathologic analysis. Here, however, both of these cell lines are shown to express ERα mRNA. RT-PCR is not quantitative, however. This discordance may be explained by (1) a very low level of expression of ER mRNA with a concurrent low expression of ER protein below the level of detection by immunoperoxidse; (2) a lack of correlation between ER mRNA and functional protein expression. Ultimately, 78% of the Latimer-generated breast cancer cell lines, and 100% of the breast reduction cell lines were concordant with their tissue of origin. Besides being useful for the validation of our cell lines, ERα results were important for planning our azacytidine experiments. The ER gene can be used as a re-activation control, as shown by Ferguson *et al.* (240).

RT-PCR was also used to evaluate the potential expression of β casein mRNA (Figure 5.2B). β casein is a milk protein expressed in the human breast. The human liver tissue specific control was shown to be negative by RT-PCR for mRNA expression. MCF7 and MDA MB231 cell lines were used as positive controls, consistent with the literature demonstrating expression of β casein in both cell lines by RT-PCR methodology (416). Cell lines generated from breast reduction tissue all showed expression of this milk protein mRNA. This result is consistent with

their cell type of origin- luminal breast epithelial cells, and therefore provides further validation of the integrity of these nondiseased cell lines and the methodology used to establish them. With respect to breast tumor tissue, only 17% of breast tumors have been reported to express β casein protein (289). Our results show that β casein is expressed in 50% of the breast tumor cell lines tested. While this is higher than the expected frequency of β casein expression in tumors from the general population, the expression within the particular tumors from which these lines were derived is unknown. In that it is only a subset of the tumors that express β casein, our results are consistent with previous findings.

Remarkably, we do not see the tissue-specific extinction of either ERα or β casein genes in our cell culture system. This is in direct contrast to the phenomenon seen in breast cancer cell lines where the majority of cell lines generated are ER-. As discussed in Chapter 4, the majority of breast cancer cell lines are ER- while the majority of breast cancers diagnosed in the US are ER+. Two possibilities explaining this doscordance is that either ER- tumors are more permissive in their growth and therefore more easily cultured with previously established culture systems or the expression of the ER is lost from these tumor cells when they are placed in culture. Unfortunately, the ER status of the tumor of origin has not been published for the vast majority of these cell lines, so we cannot distinguish between these two possibilities. However, we can report that we are able to maintain the expression of these genes concordant with their tissues of origin, eliminating this problem from the Latimer culture system.

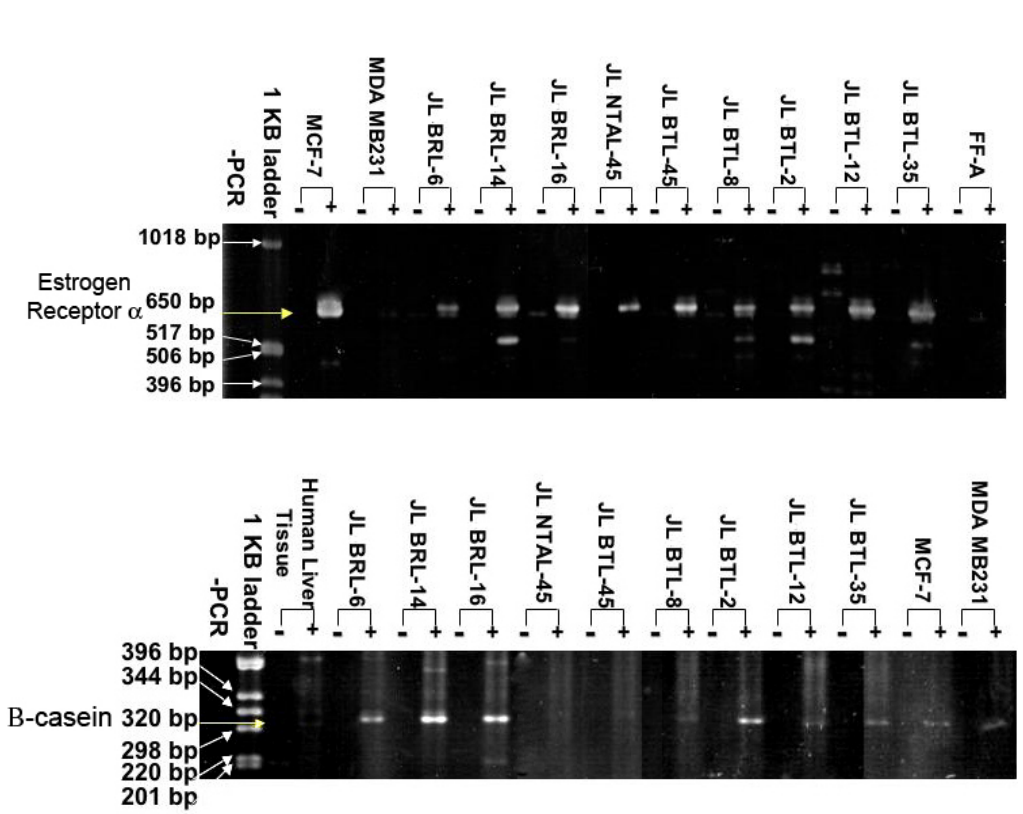


Figure 5.2 Estrogen receptor and β-casein expression in Latimer-generated cell lines by RT-PCR. A. Representative RT-PCR analysis of ERα showing expression in breast reduction and breast tumor lines. The amplimer size for the primers used was 650 bp. The MCF-7 (ER+) and MDA MB231 (ER-) cell lines and a FF explant (also ER-) were used as controls. B. Representative RT-PCR analysis of β-casein showing expression in nondiseased breast cell lines and tumor cell lines derived in the Latimer lab. The primers yield a 320 bp amplimer. The MCF-7 and MDA MB231 cell lines (positive) as well as human liver tissue (negative) were used as controls. Controls for contaminating genomic DNA were performed by the omission of the MMLV-RT enzyme in the RT reaction, shown as either + or – samples. Products were analyzed on a 2.0% agarose gel.

5.3.3 Cell Lines Chosen for Further Analysis of mRNA and Protein Expression Levels

Three breast reduction cell lines representing two pre-menopausal samples and one post-menopausal sample were chosen for analysis of expression of NER genes by RNase protection. For the breast reductions, one high repair sample, one sample with repair near the average of the population, and one with repair lower than the average of the population were chosen (Table 5.3, Figure 5.3). Similarly, early stage breast cancer was represented by three early stage tumor-derived cell lines (one stage I and two stage II) that represented the loss of repair seen in this population (Table 5.3, Figure 5.3). Advanced stage breast cancer was represented by one late stage cell line (stage III), and the established cell lines MCF7 and MDA MB231 (Table 5.3, Figure 5.3). For the late stage tumors, we chose a Latimer cell line that represented the increase in repair seen in stage III tumors. MCF7 and MDA MB231, besides representing late stage tumors, gave our studies a context in a field where these two cell lines are the most commonly used models.

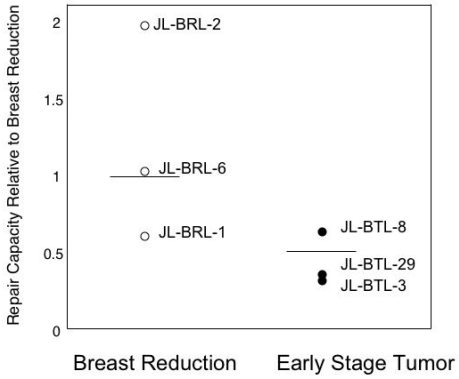
Demographic information is summarized in Table 5.3 for each of the samples with the exception of JL BRL-2. The woman from whom this sample was derived declined to give any of the information requested.

Table 5.3 Cell lines used in the analysis of mRNA and protein expression of NER genes. Repair is listed as measured by the UDS assay relative to the 23 nondiseased breast samples described in Chapter 3.

Cell Line	**Age**	**Gravida**	**Para**	**Height**	**Weight**	**Menopausal Status / Hormone Use**	**Repair**
JL BRL-1	34	2	0	5'4"	170 lbs	premenopausal	0.61
JL BRL-2	19	-	-	-	-	-	1.97
JL BRL-6	62	4	4	5'6"	158 lbs	Hysterectomy at age 30; premarin use	1.03

Cell Line	**Age**	**Stage**	**Grade**	**Size**	**ER**	**PR**	**Lymph Nodes**	**Recurrence**	**Repair**
JL BTL-8	52	I	2	2	+	+	0/20	3-14-02 negative	0.64
JL BTL-3	54	II	3	2.3	+	++	5/19	9-14-01 negative; 9-12-02 alive	0.32
JL BTL-29	39	II	2	2.2	-	-	0/22	5-29-01 negative; 8-14-02 alive	0.36
JL BTL-12	37	III	3	3	+	+	8/16	Recurred with bone and liver metastases 8-15-97 Died 11-26-99	2.08
MCF7	69	IV pleural effusion			+	+			2.07
MDA MB231	51	IV pleural effusion			-	-			3.44

A



B

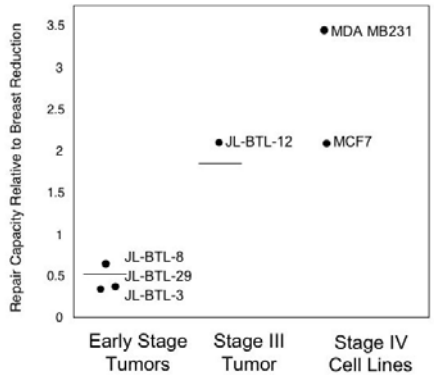


Figure 5.3 Repair capacities of the primary cultures from which the cell lines chosen for further analysis were derived. A. The repair capacity of primary cultures derived from both breast reduction and early stage breast tumors that were subsequently used to make the cell lines chosen for molecular analysis. The horizontal bar represents the average repair capacity of the population from which each type of sample was chosen. B. The repair capacity of primary cultures derived from early stage breast tumors, a stage II chemotherapy naïve breast tumor that were subsequently used to make the cell lines chosen for molecular analysis and the established lines MCF7 and MDA MB231. The horizontal bar represents the average repair capacity of the population from which they were chosen.

5.3.4 mRNA Expression of 20 Nucleotide Excision Repair Genes by RNase Protection Assay in Early Stage Tumor-Derived Cell Lines

mRNA expression levels for 20 genes in the NER pathway were determined by RNase Protection Analysis using the In Vitro Transcription Kit and RNase Protection Kits from Pharmingen. Pharmingen has developed two different probe sets each containing cDNAs representing sequences from 10 of the genes known to be involved in NER (Table 5.4). When we applied this technique to three breast reduction lines relative to three low-repair early stage tumor-derived lines, we saw significant trends in multiple genes (Table 5.4). Specifically, we saw a statistically significant loss of mRNA expression in 6 genes: XPA, CSB, XPB, TFIIHp44, Cdk7, RPAp14.

Since the intensity of the RPAp14 band is very weak it was not further analyzed. TFIIHp52 approached statistical significance with the three chosen early stage tumor-derived ($p = 0.0622$). When data from an additional stage I cell line showing low repair was added to the data from these three early stage breast tumor cell lines, the trend became significant at $p = 0.0169$. However, data from this cell line only represented one experiment. Work is ongoing in the laboratory to grow enough cells from this cell line to create more RNA for further analysis. Furthermore, when a z-test was used to compare each of the early stage tumor-derived cell lines to the population of breast reduction cell lines for TFIIHp52, each one individually had a significant loss of expression: JL BTL-8 $p = 2.64 \times 10^{-12}$, JL BTL-3 $p = 2.82 \times 10^{-5}$, JL BTL-29 $p = 4.85 \times 10^{-6}$. Because of the loss of each individual early stage tumor-derived cell line and because of the loss of the population of early stage tumor-derived with the addition of another point, TFIIp52 is considered a candidate gene for loss.

Thus, our candidate gene list as defined by loss of mRNA expression in early stage tumor-derived relative to the expression of breast reduction cell lines is: XPA, CSB, XPB, TFIIHp52, TFIIHp44, TFIIp34, and Cdk7 (Figure 5.4). Notably, all the candidate genes except for CSB and XPA are known members of the TFIIH complex.

Table 5.4 The expression of mRNA in three early stage breast tumor lines with low repair relative to three breast reduction lines as measured by RNase protection analysis. The p values calculated by a two-tailed t-test are also shown. Candidate genes are identified in red.

Gene	Function in the NER Pathway	Expression in Early Stage Tumors Lines Relative to Breast Reduction Cell Lines	P value by Two-Tailed T-Test
CSA	Transcription-coupled repair	1.079	0.806
CSB	Transcription-coupled repair	0.356	0.0187
DDB1	Damage recognition	0.743	0.0622
DDB2	Damage recognition	0.472	0.120
XPC	Damage recognition	0.673	0.266
hHR23B	Damage recognition	0.818	0.431
XPA	Lesion Verification	0.717	0.0314
RPAp70	Lesion Verification	0.646	0.0995
RPAp32	Lesion Verification	0.755	0.405
RPAp14	Lesion Verification	0.616	0.0049
XPF	Incision / Excision	0.508	0.225
ERCC1	Incision / Excision	0.860	0.441
XPG	Incision / Excision	0.678	0.0692
XPB	Unwinding- TFIIH member	0.676	0.00743
XPD	Unwinding- TFIIH member	0.585	0.113
TFIIHp52	TFIIH	0.699	0.0622
TFIIHp44	TFIIH	0.665	0.0174
TFIIHp34	TFIIH	0.583	0.0448
Cdk7	TFIIH- CAK subcomplex	0.741	0.0299
CycH	TFIIH- CAK subcomplex	0.790	0.435

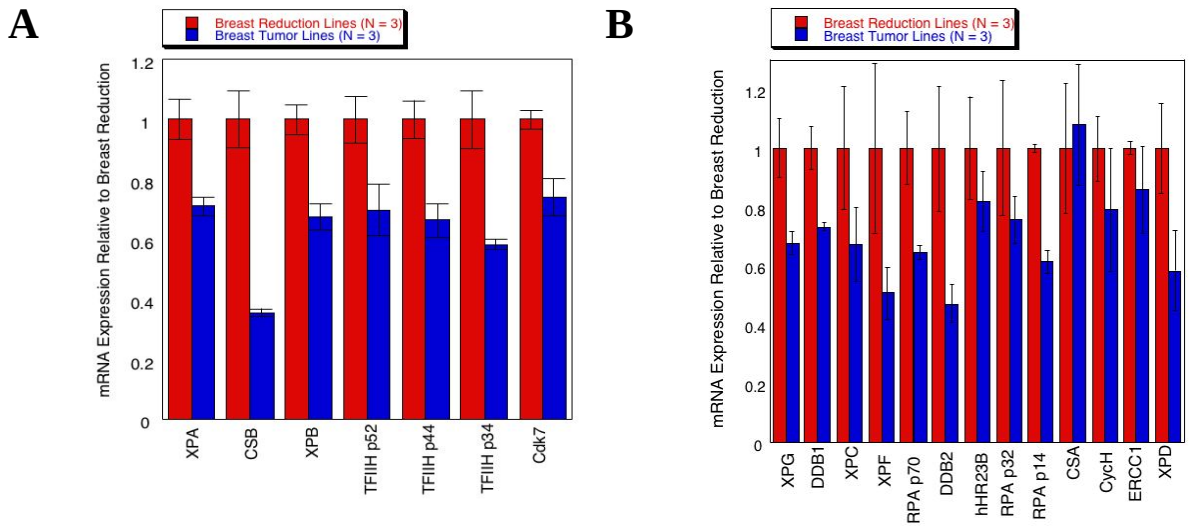


Figure 5.4 Expression of mRNA in early stage cell lines relative to breast reduction lines as measured by RNase protection analysis. A. Seven Candidate genes that demonstrated a statistically significant loss of mRNA expression when comparing three BRLs with three early stage BTLs, measured by RNase protection analysis.

B. Noncandidate genes that did not show a significant loss of RNA expression using the same six cell lines. Error bars represent the standard error across three cell lines (interindividual variation). mRNA expression is expressed relative to the average of the three BRLs.

Microarray analyses were run on multiple cell lines generated in the Latimer laboratory, including the three BRLs used for RNase protection analysis and two of the early BTLs, JL BTL-8 and JL BTL-29. RNA is currently being generated in the laboratory to complete analysis for the third early stage cancer cell lines used, JL BTL-3. Analysis was both carried out to examine the same genes as were detected by the RNase protection analysis, as well as 14

additional genes that were either directly involved in steps of the NER pathway or putatively related to it that had not been included in the RPA kits. Each gene was represented as the average expression of all probe set ids correlating to the same human sequence number. These averages were then normalized by the expression of GAPDH, as was done for the RNase Protection Analysis and expressed relative to FF as was done in the RNase protection analysis. Finally, the expressions were made relative to the tissue specific control of the three breast reduction lines used.

Analysis showed that while significance could not be achieved for any of the genes also analyzed by RNase protection the trends observed with RNase Protection Analysis were preserved for candidate (Figure 5.6) and noncandidate (Figure 5.7) genes. This suggests that additional microarray analysis could generate support for our RNase protection results. The lack of significance is most likely due to the large amount error seen in the samples, due to the fact that only 1 run of each cell line has been shown.

In addition to the genes that had been analyzed for mRNA expression with RNase protection, 14 more genes that are either involved in the NER pathway or putatively related to it were analyzed for loss with microarray. These genes include ones which are directly involved in the NER pathway through their associations with TFIIH and those involved in repair synthesis (Table 5.5). These genes were not included in the kits offered by Pharmingen for RNase Protection analysis.

BRCA1 has been implicated in the control of expression of two NER genes, XPC and DDB2. When BRCA1 is overexpressed in osteosarcoma cell lines, the mRNA expression of these two genes is also increased (319). Though it is unknown whether this regulation also takes

place in the breast, or if it occurs at physiologic expression levels of BRCA1, investigation of this gene seemed warranted.

Four genes were found to have a statistically significant loss in the early stage tumor-derived cell lines group as compared to the breast reduction lines (Table 5.5, Figure 5.5). These genes included the MNAT1 gene, a member of the TFIIH complex, the PCNA and POLD2 genes involved in DNA repair synthesis, and BRCA1, the putative regulator of XPC and DDB2 expression.

Table 5.5 The expression of mRNA in three early stage breast tumor lines with low repair relative to two BRLs as measured by microarray analysis. The p values calculated by a two-tailed t-test are also shown. Candidate genes are identified in red.

Gene	Function	Expression in Early Stage Tumors Lines Relative to Breast Reduction Cell Lines	P value by Two-Tailed T-Test
hHR23A	Damage Recognition	0.6628	0.1989
TFIIHp62	TFIIH	0.6865	0.3094
TTDA	TFIIH	0.7865	0.2414
MNAT1	TFIIH- CAK subcomplex	0.6429	0.0163
RFC1	Repair Synthesis	0.7431	0.1635
RFC2	Repair Synthesis	0.5612	0.1568
RFC3	Repair Synthesis	0.5599	0.1462
RFC4	Repair Synthesis	0.4377	0.0671
RFC5	Repair Synthesis	0.6050	0.1918
PCNA	Repair Synthesis	0.4440	0.0147
POLD1	Repair Synthesis	0.7362	0.3545
POLD2	Repair Synthesis	0.6271	0.0275
LIG1	Repair Synthesis	0.4908	0.2221
BRCA1	Putative regulation of DDB2 and XPC	0.4419	0.0434

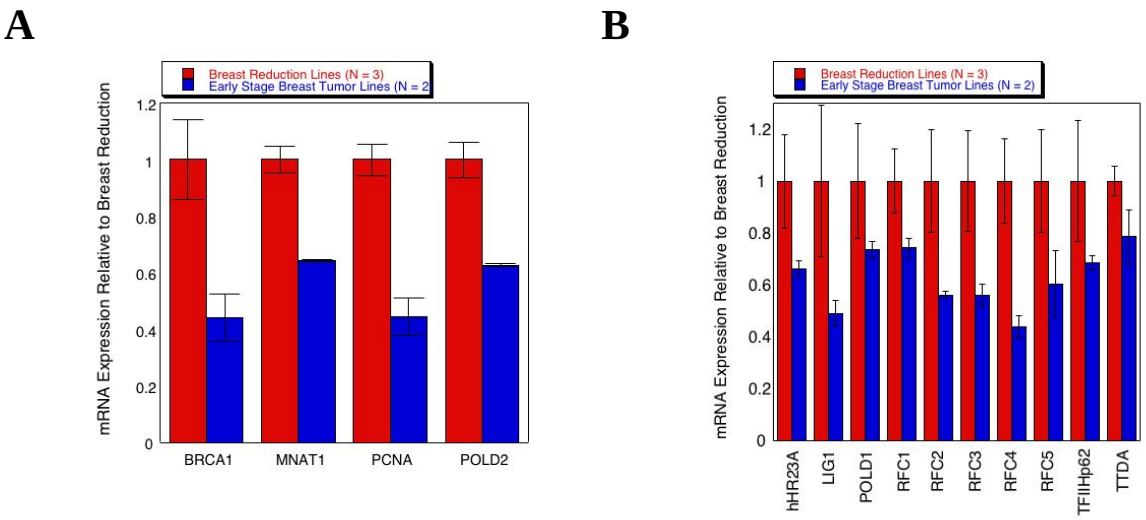


Figure 5.5 Expression of mRNA in early stage cell lines relative to breast reduction lines as measured by microarray analysis. A. Four genes associated with the NER pathway that were not included in the RNase protection panel were analyzed by micorarray analysis and also showed significant loss in mRNA expression in two early stage BTLs relative to BRLs. Data represents the average of three BRLs and two early stage BTLs (JL BTL-8 and JL BTL-29). All cell lines, with the exception of JL BTL-29, have been run once. JL BTL-29 has been analyzed twice. B. Ten more genes that are also associated with the NER pathway and not included in the RNAse protection panel were not significantly decreased in early stage tumors relative to breast reductions. The same cell lines were used as in panel A.

5.3.5 mRNA and Protein Expression Agree

Total protein was harvested from our breast reduction and early stage tumor-derived lines to examine whether the protein expression levels of two candidate and two non-candidate genes would correlate with our mRNA expression levels. In order to be relevant functionally, the loss of mRNA expression also has to be present at the protein level, since it is in protein that the actually activity of the gene lies.

The two candidate genes, CSB and XPA both showed a loss of protein expression that was consistent with the loss of mRNA expression in early stage tumor-derived cell lines (Figure 5.6). However, only the loss of XPA was statistically significant ($p= 7.34 \times 10^{-6}$) (CSB $p= 0.287$). The fact that there is also a loss of protein with XPA validates the relationship of transcriptional regulation with function

The trends shown by two of our noncandidate genes, DDB1 and DDB2 are also consistent between mRNA and protein levels (Figure 5.7). Just as with the mRNA expression, neither of these genes exhibit a statistically significant loss of protein expression ($p = 0.323$ and 0.0978, respectively). Retrospective power analysis of our results for DDB2 show that it may be possible to drive the relationship to significance if 7 additional cell lines were added to the analysis.

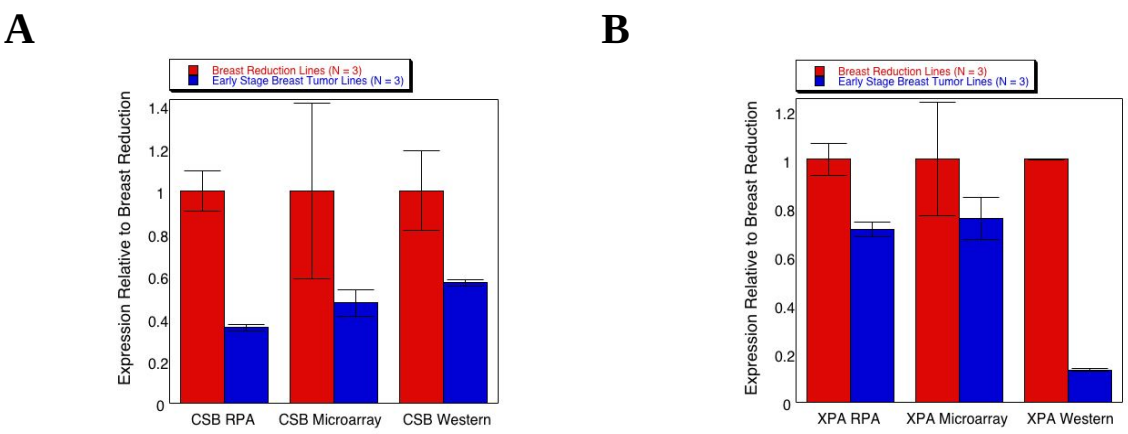
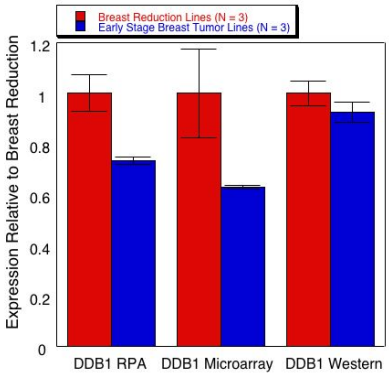


Figure 5.6 mRNA and protein expression of XPA and CSB in early stage tumor cell lines. A. CSB shows consistent trends of mRNA expression using two independent methods of mRNA analysis and similar trends with protein expression. Loss of expression is seen with all three methods. RNase protection and microarray analysis were each used to determine mRNA expression levels. Western blotting was performed with the sc-25370 antibody from Santa Cruz. Analysis was performed on three early stage tumor-derived cell lines relative to breast reduction cell lines. B. XPA shows consistent trends of mRNA expression using two independent methods of mRNA analysis and similar trends with protein expression. Analyses were performed on cell lines as in panel A. The antibody XPA bleed5 from Dr. Richard Wood was used for Western analysis.

A



B

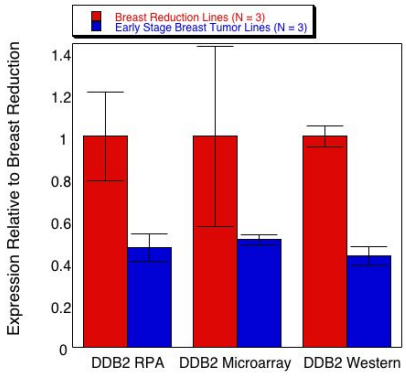


Figure 5.7 mRNA and protein expression of DDB1 and DDB2 in early stage tumor cell lines. A. DDB1 shows consistent trends of mRNA expression using two independent methods of mRNA analysis and similar trends with protein expression. No loss of expression is seen with any method. RNase protection and microarray analysis were each used to determine mRNA expression levels. Western blotting was performed with the sc-16292 antibody from Santa Cruz. Analysis was performed on three early stage tumor-derived cell lines relative to breast reduction cell lines. B. DDB2 shows consistent trends of mRNA expression using two independent methods of mRNA analysis and similar trends with protein expression. Significant loss of protein, but not mRNA expression was seen. Methods were performed on cell lines as in panel A. The antibody sc-16295 from Santa Cruz was used for Western analysis.

5.3.6 Gain of Expression Levels in Late Stage Tumors-Derived Cell Lines

In order to futher investigate the gain of DNA repair capacity observed by UDS analysis in Chapter 3 in late stage tumors, mRNA expression levels were also determined for three late stage tumor-derived cell lines: including JL BTL-12, generated in the Latimer lab from a stage III chemotherapy naïve patient, and the established cell lines MCF7 and MDA MB231 (both pleural effusion-derived). mRNA expression levels were determined for these three cell lines using the standard RNase protection analysis (Table 5.6, Figure 5.8). When compared to early stage tumor-derived cell lines, 13 out of 20 genes, including 6 of our 7 candidate genes, showed a significant increase in repair capacity as reported by a z-test with α set to < 0.05. Expression was significantly increased in 19 of the 20 genes assayed for both MCF and MDA MB231. Alternatively, these three late stage tumors can be compared as a population against the early stage breast tumor cell lines (Table 5.7, Figure 5.8). When this is done, however, far fewer genes show a significance in gain of expression due to the high variability in the three late stage tumor-derived cell lines.

Table 5.6 The expression of mRNA in three early stage BTLs with low NER capacity relative to: a stage III chemotherapy naïve cell line with high NER capacity, and the established lines MCF7 and MDA MB231 by RNase protection analysis. Both the change in expression and the p value as calculated by a two-tailed z-test are shown. Candidate genes are identified in red.

Gene	BTL-12 fold change	BTL-12 p value	MCF7 fold change	MCF7 p value	MDA MB231 fold change	MDA MB231 p value
XPG	0.541	1.82×10^{-14}	1.39	4.20×10^{-11}	1.28	2.11×10^{-6}
DDB1	0.971	0.106	2.15	0	2.14	0
XPC	0.905	0.613	2.21	9.44×10^{-11}	2.21	1.12×10^{-10}
XPF	1.85	6.05×10^{-7}	3.91	0	3.09	0
RPAp70	1.55	0	3.09	0	3.32	0
DDB2	2.84	0	4.78	0	3.43	0
hHR23B	1.03	0.796	1.90	3.13×10^{-13}	1.95	2.46×10^{-14}
XPA	2.08	0	4.50	0	4.66	0
RPAp32	2.48	0	2.61	0	1.99	0
RPAp14	10.94	0	11.5	0	7.49	0
CSB	2.01	0	2.21	0	2.59	0
XPB	3.51	0	2.46	0	2.70	0
TFIIHp52	3.79	0	3.58	0	4.85	0
TFIIHp44	0.949	0.546	3.06	0	2.21	0
CSA	0.688	0.0998	1.85	7.04×10^{-6}	1.60	0.00179
Cdk7	1.16	0.0430	2.19	0	2.33	0
CycH	1.24	0.359	1.59	0.0245	1.82	0.0179
TFIIHp34	1.92	0	2.32	0	3.86	0
ERCC1	1.45	0.00859	0.881	0.484	1.48	0.0531
XPD	1.32	0.169	1.80	6.24×10^{-4}	2.84	3.33×10^{-15}

Table 5.7 The mRNA expression of early stage BTLs versus late stage BTLs when all three late stage tumors are considered together as a population. Significant gains of repair are shown in red.

Gene	Function	Expression in Late Stage Tumors Relative to Early Stage Tumors	P-value by Two-tailed T-test
CSA	transcription-coupled repair	1.378	0.413
CSB	transcription-coupled repair	2.271	0.0136
DDB1	damage recognition	1.753	0.194
DDB2	damage recognition	3.684	0.037
XPC	damage recognition	1.775	0.210
hHR23B	damage recognition	1.627	0.158
XPA	lesion verification	3.746	0.0810
RPAp70	lesion verification	2.650	0.097
RPAp32	lesion verification	2.362	0.00707
XPF	incision / excision	2.950	0.0727
ERCC1	incision / excision	1.268	0.359
XPG	incision / excision	1.072	0.812
XPB	unwinding- TFIIH member	2.891	0.0236
XPD	Unwinding- TFIIH member	1.985	0.146
TFIIHp52	TFIIH	4.075	0.0105
TFIIHp44	TFIIH	2.073	0.220
Cdk7	TFIIH- CAK subcomplex	1.892	0.130
CycH	TFIIH- CAK subcomplex	1.548	0.165

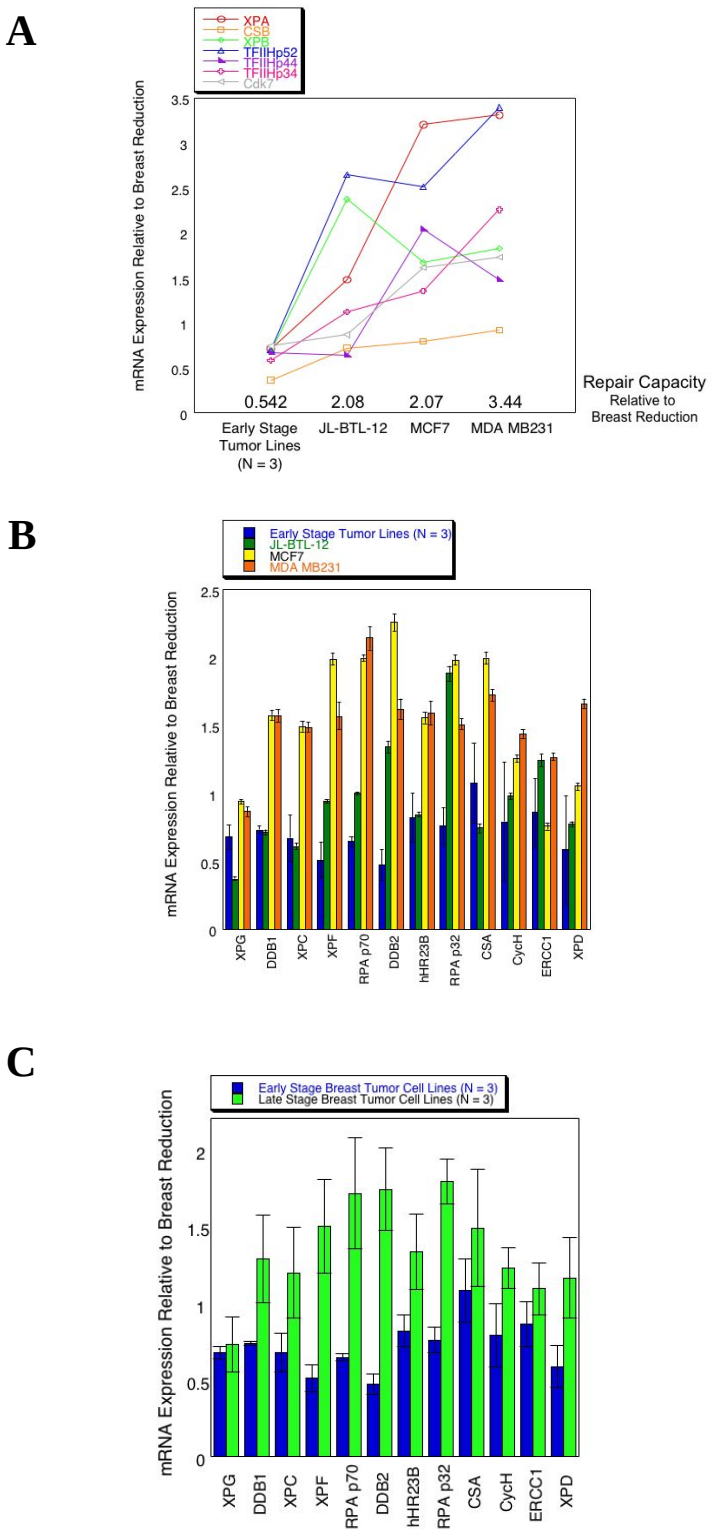


Figure 5.8 Expression of mRNA in late stage cell lines relative to early stage cell lines as measured by RNase protection analysis.

A. Gain of mRNA expression of candidate genes in late stage BTLs measured by RNase protection. One stage III tumor-derived cell line generated in the Latimer laboratory, JL BTL-12, as well as the stablished breast cancer cell lines MCF7 and MDA MB231 relative to three early stage tumor-derived Latimer cell lines is reported relative to the average of the three breast reduction lines.

B. Gain of mRNA expression of non-candidate genes measured by RNase protection. The same cell lines are used as in panel A. Data for RPAp14 could not be reliably reproduced and is not shown. Error between three cell lines is shown for the early stage tumor-derived cell lines, and experimental error is reported for JL BTL-12, MCF7, and MDA MB2331.

C. Gain of mRNA expression in noncandidate genes by RNase protection. This data represents the same data shown in panel B, with the data from the three late stage tumor-derived cell lines: JL BTL-12, MCF7 and MDA MB231 averaged. Standard error is given as an estimate of variability between these three cells lines.

mRNA expression was also evaluated by microarray analysis of all three late stage tumor-derived cell lines for the 20 NER genes included in the RNase protection assay and 14 additional NER genes. Trends for the increases in expression matched between the microarray and RNase protection with one exception, the candidate gene CSB. In contrast to the other genes tested, the mRNA expression of this gene was higher in the early stage tumor-derived cell lines than in each of the three late stage tumors when tested by microarray (Figure 5.10).

Trends for the 14 additional genes that were analyzed by microarray showed that the 4 candidate genes that had previously showed a loss of expression in early stage tumors, also now showed a gain of expression in late stage tumors (Figure 5.9A). The 10 non-candidate genes (those that had not shown a loss of mRNA expression in early stage tumors) also showed an increase in expression in the late stage tumor-derived cell lines with the exceptions of TFIIHp62 and TTDA (Figure 5.9B). Because of the low number of analyses (only one for each of these cell lines) statistical significance is not estimated for microarray comparisons. The fold difference seen in expression relative to early stage tumors is seen in Table 5.8. As was the case with the genes assayed by RNase protection, the majority of genes demonstrated an increase in expression in each of these three late stage tumor-derived cell lines: JL BTL-12, MCF7, and MDA MB231 (Figure 5.9). The notable exceptions to this were the genes TFIIHp62 and TTDA. These two genes decreased in expression in JL BTL-12 and MCF7, but did show an increase in MDA MB231.

Table 5.8 The expression of mRNA in three early stage breast tumor-derived cell lines with demonstrated low NER relative to: a stage III chemotherapy naïve cell line with high repair JL BTL-12, and the estsablished lines MCF7 and MDA MB231 by microarray analysis. The fold change in expression is shown. Due to the lack of repetition of samples and the low number of samples, p-values could not yet be calculated. Candidate genes are identified in red.

Gene	JL BTL-12 Fold Change relative to early stage tumors	MCF7 Fold Change relative to early stage tumors	MDA MB231 Fold Change relative to early stage tumors
hHR23A	1.3478	0.9176	1.7023
TFIIHp62	0.4357	0.6831	2.0674
TTDA	0.6841	0.6311	1.3718
MNAT1	1.0939	1.5512	2.0899
RFC1	1.6555	1.5290	2.8188
RFC2	7.2761	2.6567	4.7494
RFC3	5.4307	5.2153	5.6202
RFC4	8.4873	2.4373	7.5441
RFC5	7.4052	4.3907	5.8758
PCNA	3.8426	2.5265	5.3181
POLD1	4.8020	1.9316	3.9056
POLD2	2.8462	1.0648	2.0736
LIG1	2.6815	4.8391	6.2825
BRCA1	5.2947	3.8245	5.0378

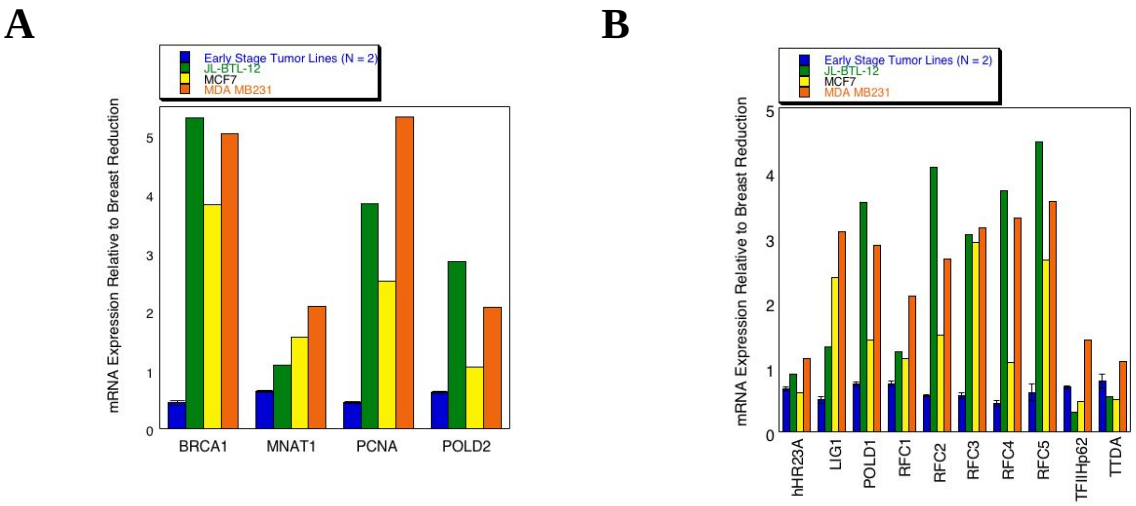


Figure 5.9 Expression of mRNA in late stage cell lines relative to early stage cell lines as measured by microarray analysis. A. Gain of mRNA expression in late stage tumors of 4 candidate genes identified by their loss in early stage tumors by microarray analysis. Data includes one stage III tumor-derived cell line, JL BTL-12, generated in the Latimer laboratory as well as MCF7 and MDA MB231, expressed relative to two early stage tumor-derived Latimer cell lines measured by microarray. mRNA expression is reported relative to the average of three breast reduction cell lines. Error is expressed across samples for the early stage tumor group. B. Gain of mRNA expression for 10 non-candidate genes by microarray analysis. The same cell lines are used as in panel A.

5.3.7 Agreement of Protein with mRNA Expression for Gain

Using Western analysis, we determined that there was also a coordinate expression of protein with mRNA expression as determined by RNase protection. Two candidate genes, CSB and XPA, and two noncandidate genes, DDB1 and DDB2, were examined (Figure 5.10). Gain of protein expression between the early stage tumor-derived cell lines and JL BTL-12 were significant only for the two candidate genes, CSB and XPA ($p < 0.0001$ and $p = 0.000491$, respectively). Gain of protein expression was significant between the early stage tumors and MCF7 for all four of these genes.

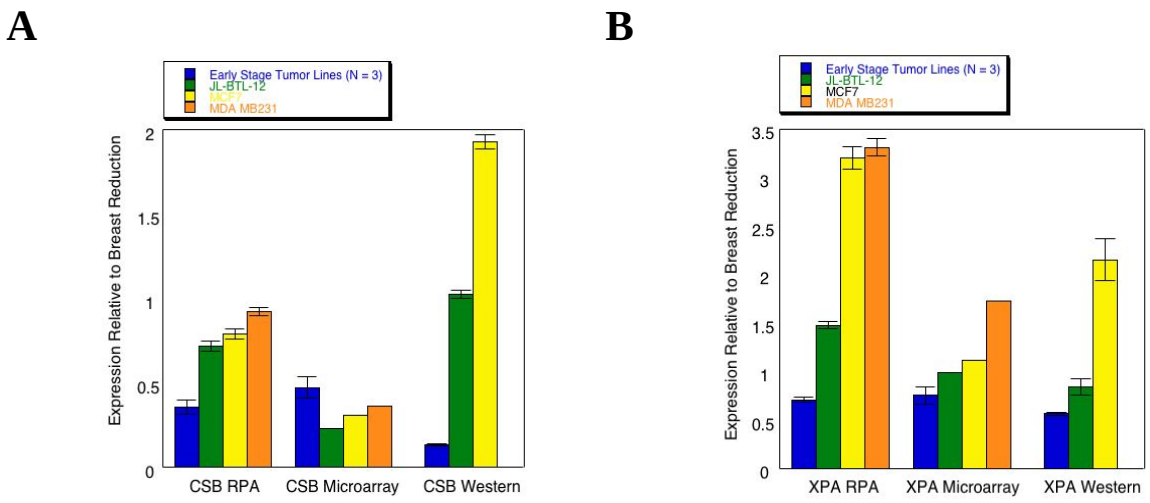


Figure 5.10 mRNA and protein expression of XPA and CSB in late stage tumor cell lines. A. CSB shows different trends of mRNA expression using two independent methods of mRNA analysis but a similar trend between protein expression and mRNA as measured by RNase protection. RNase protection and microarray were both used to determine mRNA expression levels. Western blotting was performed with the sc-23570 antibody from Santa Cruz. The average of three separate early stage tumor-derived cell lines is shown with the error between them expressed. The error expressed for the stage III chemotherapy naïve stage III cell line JL BTL-12 and the stage IV MCF7 and MDA MB231 cell lines represents experimental error. MDA MB231 has not yet been analyzed by Western analysis. B. XPA shows consistent trends of mRNA expression using two independent methods of mRNA analysis and a similar trend with protein expression. RNase protection and microarray were both used to determine mRNA expression levels. Western blotting was performed with the XPA bleed 5 antibody from Dr. Richard Wood. Cell lines and error are the same as in panel A.

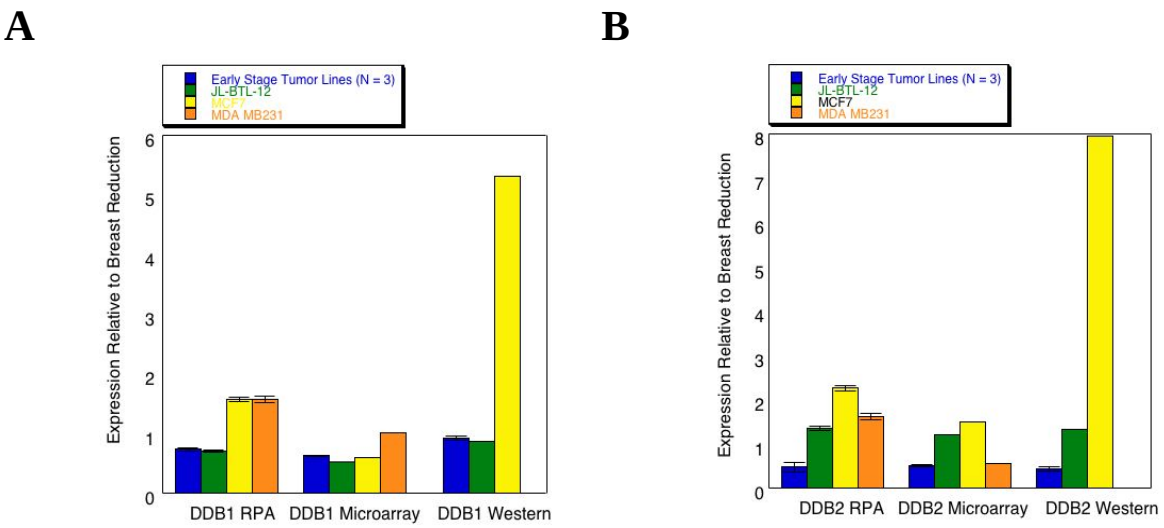


Figure 5.11 mRNA and protein expression of DDB1 and DDB2 in late stage tumor cell lines. A. DDB1 shows consistent trends of mRNA expression using two independent methods of mRNA analysis and a similar trend with protein. RNase protection and microarray were both used to determine mRNA expression levels. Western blotting was performed with the sc-16292 antibody from Santa Cruz. The average of three separate early stage tumor-derived cell lines is shown with the error between them expressed. The error expressed for the stage III chemotherapy naïve stage III cell line JL BTL-12 and the stage IV MCF7 and MDA MB231 cell lines represents experimental error. MDA MB231 has not yet been analyzed by Western analysis. B. DDB2 shows consistent trends of mRNA expression using two independent methods of mRNA analysis and a similar trend with protein expression. RNase protection and microarray were both used to determine mRNA expression levels. Western blotting was performed with the sc-16295 antibody from Santa Cruz. Cell lines and error are the same as in panel A.

Possible mechanisms to explain the patterns of NER gene regulation that we have seen here suggested both genetic and epigenetic regulation. With respect to genetic regulation, loss could be achieved through gene mutation and gain through reversion. Alternatively, a transcription factor that controls these genes could be regulating their transcription level. We have observed, however, that multiple genes in the NER pathway are lost in early stage tumors and that these same candidate genes are gained in late stage tumors. Furthermore, we have seen that this regulation of expression is present in multiple breast tumor lines derived from multiple women relative to multiple breast reduction cell lines. This makes the possibility of a genetic mechanism such as loss and reversion highly unlikely. There are also currently no common transcription factors that have been reported to regulate our set of candidate genes.

Epigenetic regulation mechanisms include the methylation of the promoter regions of these genes, effectively silencing them. Methylation of genes is a reversible phenomenon that would be a much more likely explanation of the effect that we have seen in multiple genes in multiple cell lines derived from multiple people. Since methylation of genes occurs in CpG rich regions of the genome known as CpG islands, we evaluated the 5' ends of our genes to determine whether epigenetic regulation was even possible based upon sequence.

5.3.8 CpG Island Analysis

In order to evaluate the control of gene expression in our candidate genes, we looked for the presence of CpG dinucleotides, the site of methylation in the mammalian genome in each gene evaluated by RNase protection. Gene sequences as listed in the NCBI database including a

region 2000 nucleotides upstream of the promoter region were analyzed. This includes the most liberal definition of the region where a CpG island can lie.

Table 5.9 Criterion for the definition of Type I, Type II, and Type IIa CpG Islands.

Type	**% GC**	**Observed CG / Expected CG**	**Length, nucleotides**
Type I (209)	50 %	0.60	200
Type II (210)	55 %	0.65	500
Type IIa (211)	55 %	0.65	200 ***located within 100 nucleotides upstream or downstream of the transcription start site

Our analysis revealed that 19 out of the 20 NER genes analyzed by RNase protection had type I CpG islands, and all 20 genes had Type II and IIa CpG islands (Table 5.10). These islands were located in regions immediately surrounding the transcriptional start site. Using the less stringent Type I definition, the majority of genes were seen to have multiple islands, including ones that were not in the region of their transcriptional start site: XPB, XPC, hHR23B, XPD, DDB1, DDB2, XPF, ECRR1, XPG, TFIIHp52, TFIIHp44, TFIIHp34, Cdk7, CSA, RPAp70, RPAp32 and RPAp14. Only two of these islands that were outside the region of the transcription start site also met the criterion required for a Type II CpG island: RPAp32 and TFIIHp52. In the case of TFIIHp52, this may be due to the presence of the putative gene LOC646563 that is presumed to be present antisense to TFIIHp52 by computer analysis of the region's sequence.

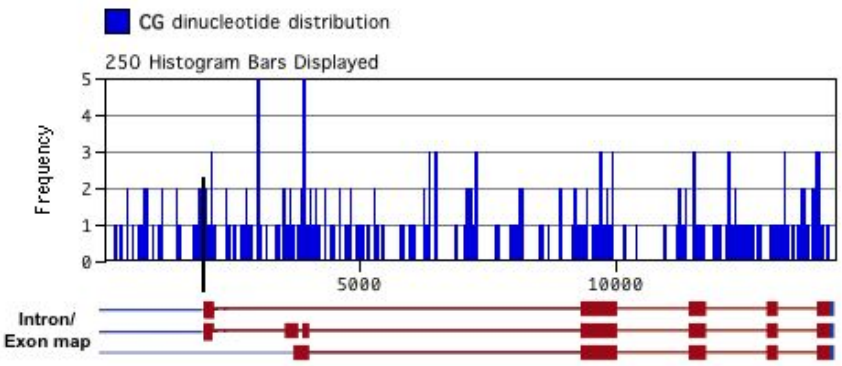
The control genes GAPDH, a housekeeping gene used for normalization of the RNase protection analysis, and α-1-antitrypsin, a tissue specific gene, were also included. This showed that there were no CpG islands in the tissue specific gene, but were present in the housekeeping gene.

The CpG island analysis was also represented graphically by overlaying the islands over the CpG dinucleotide distribution as mapped by MacVector (Figure 5.12). For perspective, the intron/exon sequences of the genes and the transcription start sites were also included. Shown here are the housekeeping gene GAPDH, the tissue specific gene α-1-antitrypsin, the noncandidate gene DDB1, and the seven candidate genes: XPA, CSB, TFIIHp52, TFIIHp44, TFIIHp34, and Cdk7.

Table 5.10 Presence of CpG islands and overall CG richness of the 20 genes evaluated using RNase protection.

Gene	CpG Dinucleotide Frequency	Observed / Expected Ratio of CpG Dinucleotides	Type I	Type II	Type IIa
XPA	0.011	0.27	1	1	1
XPB	0.014	0.26	4	1	1
XPC	0.014	0.29	3	1	1
hHR23B	0.011	0.31	1	1	1
XPD	0.028	0.36	6	2	1
DDB1	0.016	0.30	2	2	1
DDB2	0.018	0.33	3	2	1
XPF	0.013	0.32	3	1	1
ERCC1	0.021	0.31	3	1	1
XPG	0.014	0.35	2	1	1
TFIIHp52	0.022	0.35	2	2	1
TFIIHp44	0.012	0.32	5	1	1
TFIIHp34	0.017	0.36	7	1	1
RPAp70	0.020	0.40	18	1	1
RPAp32	0.018	0.36	7	2	1
RPAp14	0.009	0.25	6	2	0
Cdk7	0.012	0.28	4	1	1
CycH	0.010	0.29	1	1	1
CSB	0.010	0.24	4	1	1
CSA	0.008	0.23	5	1	1
GAPDH	0.038	0.42	1	1	1
α-1-antitrypsin	0.014	0.21	0	0	0

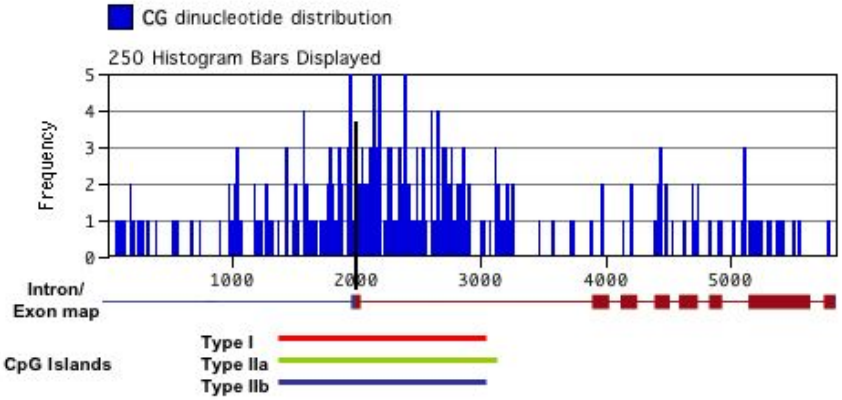
A



No cpg islands

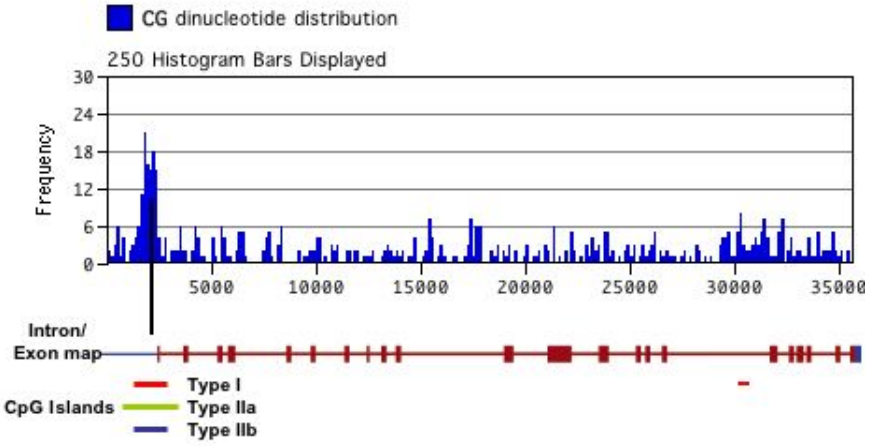
Alpha-1-antitrypsin

B



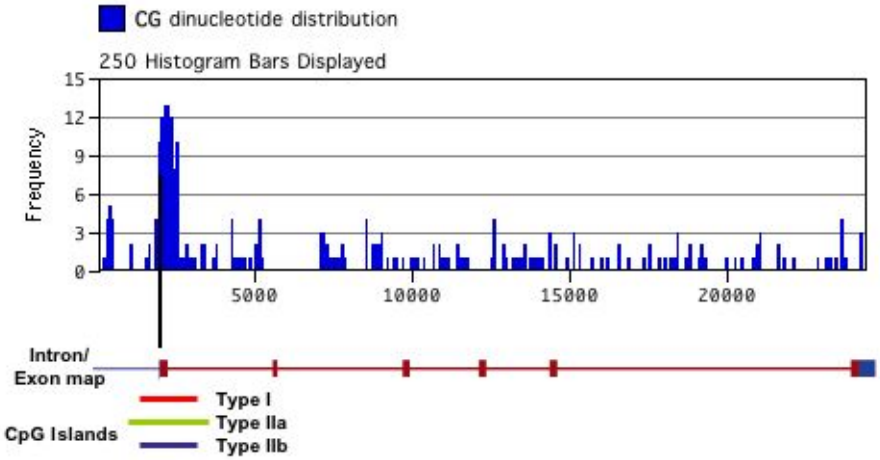
GAPDH

C



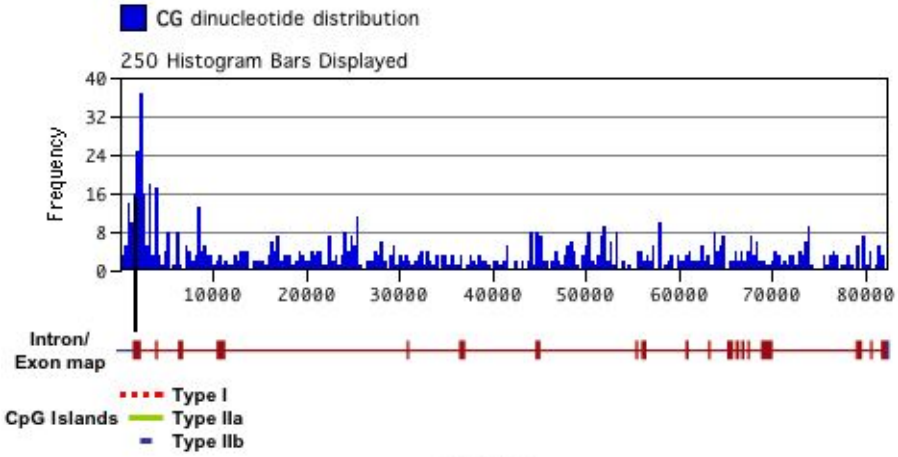
DDB1

D



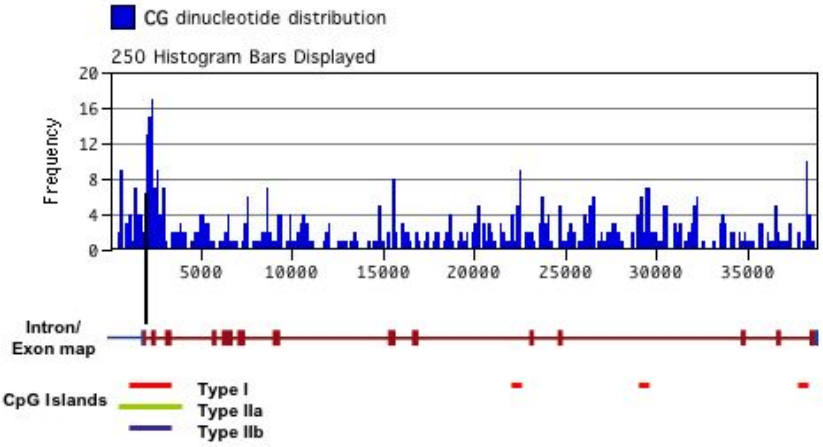
XPA

E



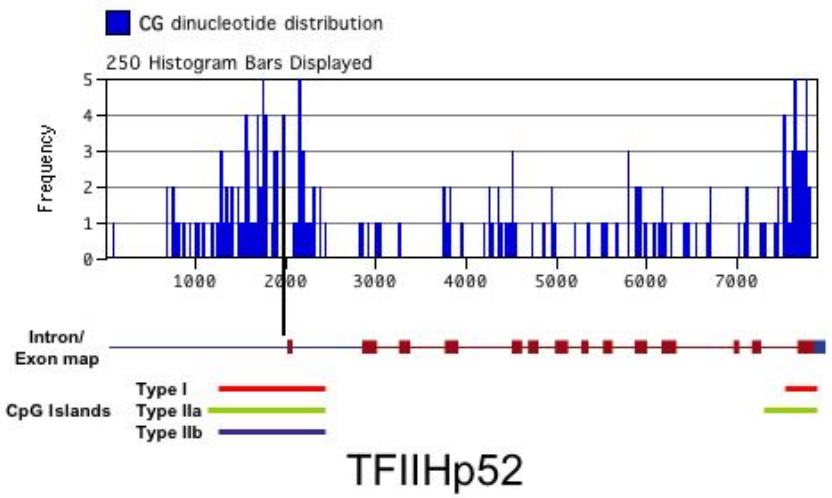
CSB

F



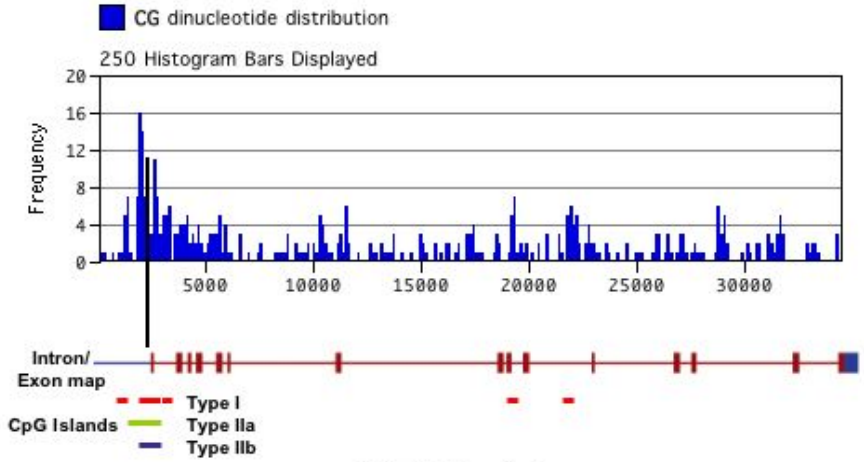
XPB

G



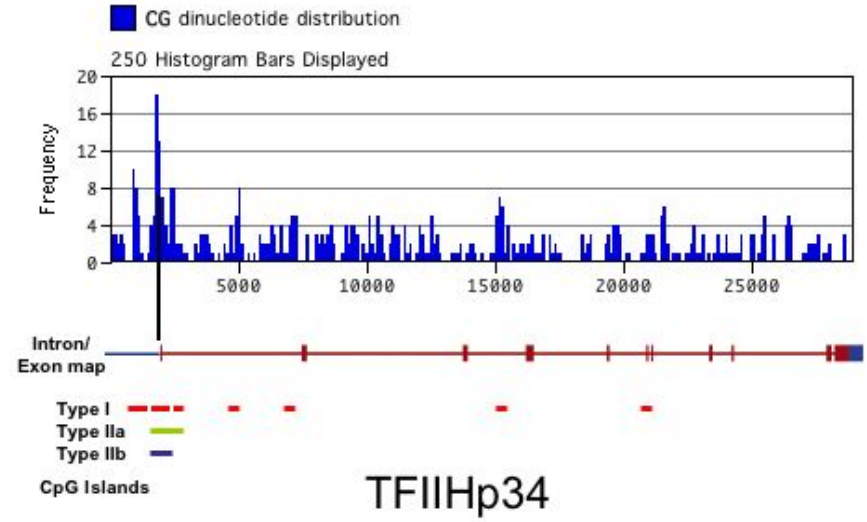
TFIIHp52

H



TFIIHp44

I



TFIIHp34

J

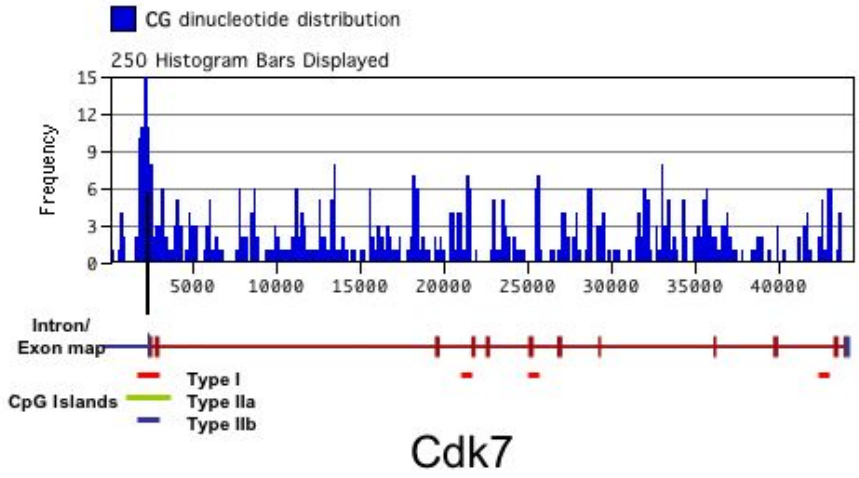


Figure 5.12 The CpG distribution and presence of CpG islands in nine selected genes. A-J. The CpG dinucleotide distribution through 2 control genes, 1 noncandidate NER gene, and all 7 NER candidate genes identified by RNase Protection Analysis are shown. The histogram bars represent the frequency of CpG dinucleotides per 250 bases. The transcriptional start site is indicated on each histogram by the black vertical line. The presence of CpG islands of each of the types discussed are shown on each histogram corresponding to their location (Type 1 = red; Type II = green; Type Iia = blue). In addition, the intron/exon map of each of the genes in shown. The region upstream of the transcription start site, including the promoter, is shown in blue. The coding region of the genes are shown in brown. Introns are represented by lines, and exons are represented by boxes. If multiple splice variants or multiple transcription start sites were known for the gene, all were represented.

5.3.9 Utilization of 5-aza-2'Deoxycytidine to Restore Early Stage Nucleotide Excision Repair Gene Expression to Nondiseased Levels

Once our CpG analysis revealed that our candidate genes had sequences that supported the idea of methylation as a regulation mechanism, we wanted to experimentally determine if methylation was responsible for the loss of gene expression. To do this, we treated early stage tumor-derived cell lines with the demethylation agent 5-aza-2'deoxycytidine in an effort to restore gene expression. Dose was determined by exposing the early stage tumor-derived cell lines JL BTL-29 and JL BTL-3 to a range of treatments used in the literature in an investigation of methylation of the ER in MDA MB231 cells: 0.1 μM, 0.5 μM, 1 μM, 5 μM, and 10 μM (240). Growth rates of these cells were substantially slowed, and cells appeared vacuolated in the 1 μM and 5 μM doses of drug. Substantial changes in morphology and cell death were observed at the 10 μM dose. Treated cells were able to double in approximately 4 days for each of these cell lines in each of these doses. Therefore, the doses of 1 μM and 5 μM were originally chosen for analysis with the JL BTL-29 cell line. Subsequent analysis of the JL BTL-8 cell line was carried out with only the 5 μM dose over a 5 day period. Five days included the minimum time we believed needed for a replication to have occurred and time for expression to take place. This is a 10-fold higher dose than was required for minimal reactivation of the ER used with the MDA MB231 cell line, but the same amount of time.

A similar report also published by the research group who showed reactivation of ER with azacytidine noted that the addition of 100 ng/ml (0.331 μM) of the histone deacetylase inhibitor trichostatin A could potentate the action of azacytidine. Addition of trichostatin A had to occur in a particular temporal pattern: after the use of azacytidine was begun, and concurrent

with azacytidine for its activity. In other words, cells spend the first 87.5% of treatment time in azacytidine alone, and the remaining 12.5% in azacytidine and trichostatin A. Therefore, in our experiments with JL BTL-8, we included a 100 ng/ml dose of trichostatin A to evaluate its effects. An evaluation of the dose range for Trichostatin A was not performed. As the authors of their publication noted, any additional Trichostatin A caused widespread death of the cultures. We used the same dose as was published in this report, with the same relative time of exposure of azacytidine alone to azacytidine with trichostatin A before harvest.

5.3.10 14-3-3 σ Can Be Reactivated with Azacytidine in Early Tumor-Derived Stage Cell Lines

We used the gene 14-3-3 σ as a test for whether or not our demethylation activity had been successful. 14-3-3 σ is reportedly absent in expression in 94% of breast tumors, an effect that was due to promoter methylation in 91% of these tumors (360). This effect was reversed with treatment by 5-aza-2'deoxycytidine in a cell line that was originally negative for 14-3-3 σ expression (360). Expression of 14-3-3 σ was absent from our early stage tumor-derived cell lines, but present in the stage III tumor-derived cell line JL BTL-12 (Figure 5.13). MCF7 was used as a positive control for expression, as it was already known to express this mRNA as measured by Northern blot (360). Treatment with azacytidine was able to restore expression, as determined by RT-PCR in both the stage I cell line JL BTL-8 and the stage II cell line JL BTL-29 (Figure 5.14).

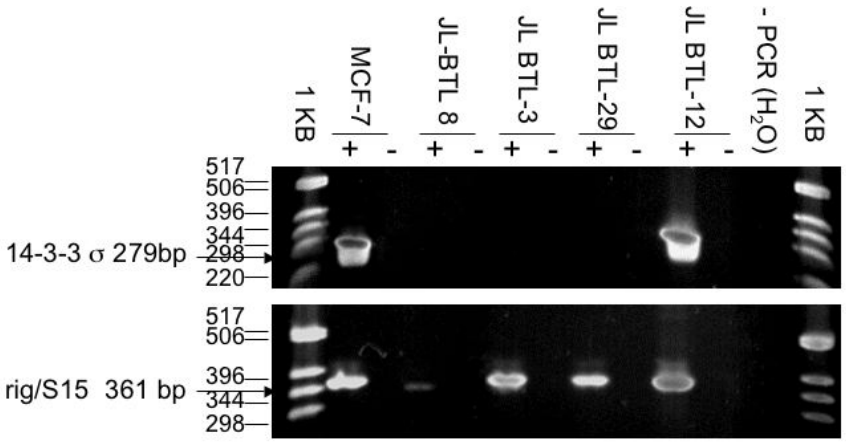


Figure 5.13 14-3-3 σ is not expressed in the early stage tumor-derived JL BTL-8, JL BTL-3, and JL BTL-29. The presence of 14-3-3 σ mRNA was assayed by RT-PCR in 5 cell lines. MCF7 is a known positive for expression of this gene and water was used as a negative control for the PCR reaction. The presence of genomic DNA contamination was assessed by the use of a negative RT control, done with out the MMLV-RT reverse transcriptase enzyme in the RT reaction. The amplimer size expected for the 14-3-3 σ primers was 179 base pairs. The presence of DNA in the RT reaction was confirmed by the use of RT-PCR with primers designed for the rig/S15 gene. This control reaction was carried out on the same cDNA synthesized by reverse transcription as was used for the 14-3-3 σ analysis. Amplimer size for rig/S15 was 361 base pairs.

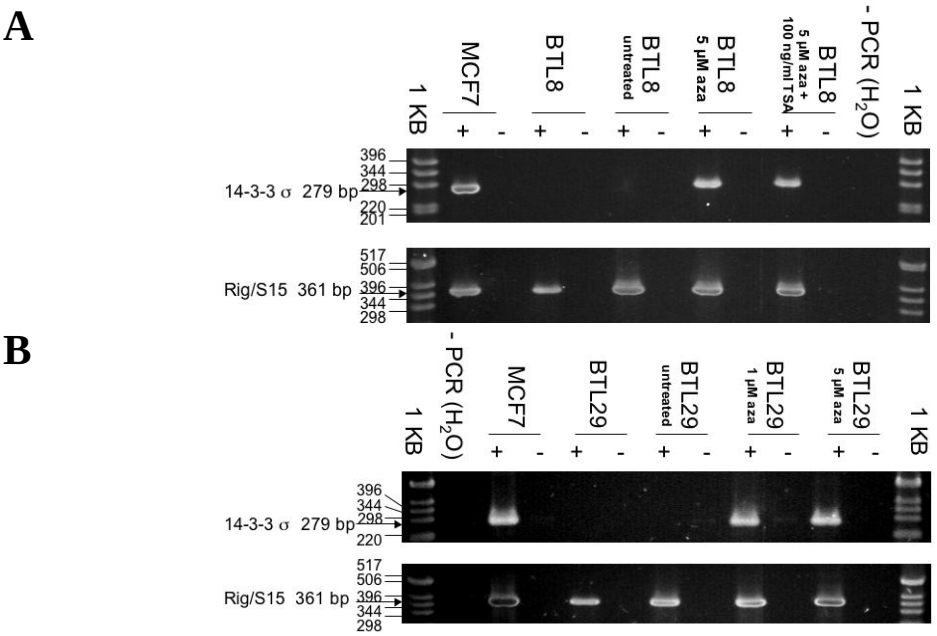


Figure 5.14 Reactivation of the 14-3-3 σ gene in two early stage tumor-derived lines. Results of RT-PCR analyses of 14-3-3 σ in the stage I cell line JL BTL-8 (A) and the stage II cell line JL BTL-29 (B) after treatment with 5-aza-2'deoxycytidine (aza) are shown. The MCF-7 cell line was included as positive control for the presence of 14-3-3 σ. Controls for possible contaminating genomic DNA (because the mammalian σ gene does not contain introns) were performed by the omission of the MMLV-RT enzyme in the RT reaction. The resultant cDNAs were subjected to PCR with Ambion SuperTaq enzyme. The amplimer size for 14-3-3 σ was 279 bp. The presence of cDNA (and RNA) was confirmed by the use of control primers targeting the rig/S15 gene (lower panels of A and B). The same cDNA samples were subjected to PCR using these primers and the same conditions as were used for the σ 14-3-3 analysis. Amplimer size for rig/S15 was 361 bp.

5.3.11 mRNA Expression of Three Candidate Genes is Restored by Azacytidine Treatment

mRNA expression was assessed in total RNA harvested from the early stage tumor-derived cell lines JL BTL-8 and JL BTL-29 after azacytidine or azacytidine with trichostatin A treatment. Four genes: RPAp32, TFIIHp52, TFIIHp34, and Cdk7 all showed increases in their expression of RNA after treatment with azacytidine in both of these cell lines (Table 5.8). TFIIHp52, TFIIHp34, and Cdk7 had all been identified as candidate genes. RPAp32 mRNA expression, however, had not been statistically significantly lost in early stage tumor-derived cell lines relative to breast reduction cell lines. Data for the cell line JL BTL-8 comes from two repetitions of the azacytdine treatment on passage 16 and passage 19 JL BTL-8 cells.

RNase protection was carried out in at least in triplicate with the hNER2 template set for both experimental replicates. RNase protection was carried out twice for the passage 16 experiment and once for the passage 19 experiment. Therefore, significance can be assessed using these experimental replicates.

Expression of all 20 genes was not significantly different in the untreated controls of either experimental replicate and thus were pooled with the previous examinations of JL BTL-8 expression. Likewise, there was no statistically significant difference in the expression in the two different passages of azacytidine treated JL BTL-8s and these groups were also pooled. The same was found for the two passages of azacytidine with trichostatin A. These pooled data were used to determine significance.

Work is ongoing in the laboratory for more repetitions of the RNase protection analysis of both of these experimental replicates of the JL BTL-8 cell line with the hNER1 template set.

Data for the JL BTL-29 cell line was generated from 1 treatment of the cells with 1 μM or 5 μM azacytidine. This RNA was subjected to triplicate analysis on RNase protection with the hNER2 set. Unfortunately, the yield of RNA prevented further analysis by RNase protection. Untreated JL BTL-29 cells were not significantly different than the previous results of analysis of JL BTL-29 mRNA expression by RNase protection, and thus these data were pooled for comparisons. Further experimental replicates of the JL BTL-29 cell line are currently underway.

mRNA expression was also assessed by microarray analysis. Total RNA from each experimental replicate of the JL BTL-8 and the single experiment of JL BTL-29 were subjected to microarray expression analysis. Due to the low number of repetitions (only 1 for each sample) the relationships between the experimental replicates could not be assessed and therefore data from the two passages of JL BTL-8 and between the untreated JL BTL-29 and previous data for this cell line were not pooled. Trends for expression agreed between the microarray analysis and the RNase Protection Analysis. Genes where there had been a statistically significant increase in expression by RNase protection also showed an increase in expression by microarray (Table 5.11, Figure 5.15). Neither the other candidate genes nor the other noncandidate genes for which there was no increase in expression by RNase protection showed an increase in mRNA expression by RNase protection or microarray (Figure 5.16).

These experiments demonstrated that our cells are able to tolerate treatment with a toxic drug long enough to show effects of demethylation. Multiple genes showed that the epigenetic mechanism of methylation was indeed playing a role in their regulation. Sodium bisulfite sequencing of these genes is planned to confirm whether this is through a direct effect of methylation in the CpG islands of the restored genes. However, not all of the NER genes that were significantly lost were gained again after treatment. There are two possibilities to explain

this. (1) The first possibility is that chromatin configuration was also a part of the regulation in addition to methylation. Because the dose and time course of the trichostatin A drug were not rigorously worked out for these cell lines, it is possible that the drug was not effective. A suitable positive control for the effect of this drug must be identified and the conditions worked out in the future. (2) The second possibility is that there is another mechanism that is also controlling the expression of some of these genes.

Table 5.11 mRNA expression of JL BTL-8 by RNase protection analysis after azacytidine treatment. The addition of trichostatin A (TSA) was never significantly different than the azacytidine (aza) only groups with fold differences between the treatments ranging from 0.978 to 1.06. Genes with significant increases are shown in red. Candidate genes are bolded.

Gene	**Azacytidine** Treatment Fold Difference	P Value for Gain with **Azacytidine** Only	**azacytidine** treatment + **TSA** Fold Difference	P Value for Gain with **Azacytidine** and **TSA**
CSA	1.03	0.288	1.02	0.523
CSB*	1.03	0.798	1.03	0.940
DDB1	1.00	0.867	1.02	0.428
DDB2	1.01	0.101	1.06	0.114
XPC	1.01	0.897	1.02	0.715
hHR23B	1.10	0.128	1.11	0.150
XPA*	1.03	0.204	1.04	0.119
RPAp70	1.02	0.390	1.04	0.426
RPAp32	1.32	3.14E-06	1.35	2.80E-03
RPAp14	1.07	0.282	1.07	0.0735
XPF	1.01	0.996	1.03	0.528
ERCC1	0.977	0.408	0.991	0.793
XPG	1.02	0.717	1.00	0.740
XPB*	0.987	0.749	1.01	0.323
XPD	0.978	0.622	1.07	0.286
TFIIHp52*	1.37	1.85E-05	1.41	1.17E-06
TFIIHp44*	1.00	0.763	1.00	0.807
TFIIHp34*	1.77	2.02E-08	1.88	2.61E-11
Cdk7*	1.94	9.73E-09	1.99	3.33E-08
CycH	1.03	0.206	1.00	0.799

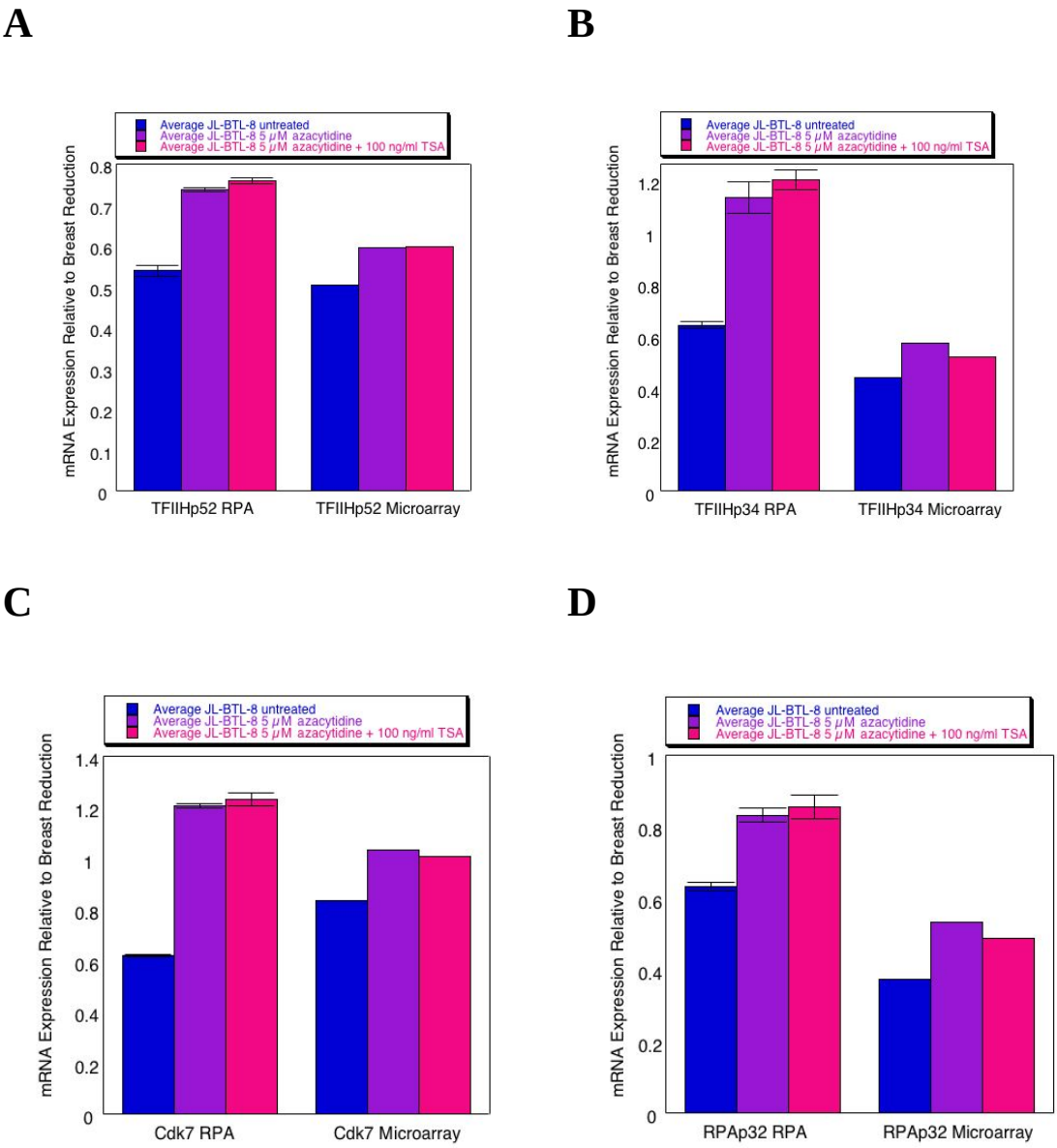


Figure 5.15 Four genes show epigenetic regulation after treatment of the stage I cell line JL BTL-8 with 5-aza-2'deoxycytidine. RNase protection results represent the pooling of two experimental replicates of JL BTL-8. Microarray analysis is represented by a single analysis of passage 16 JL BTL-8.

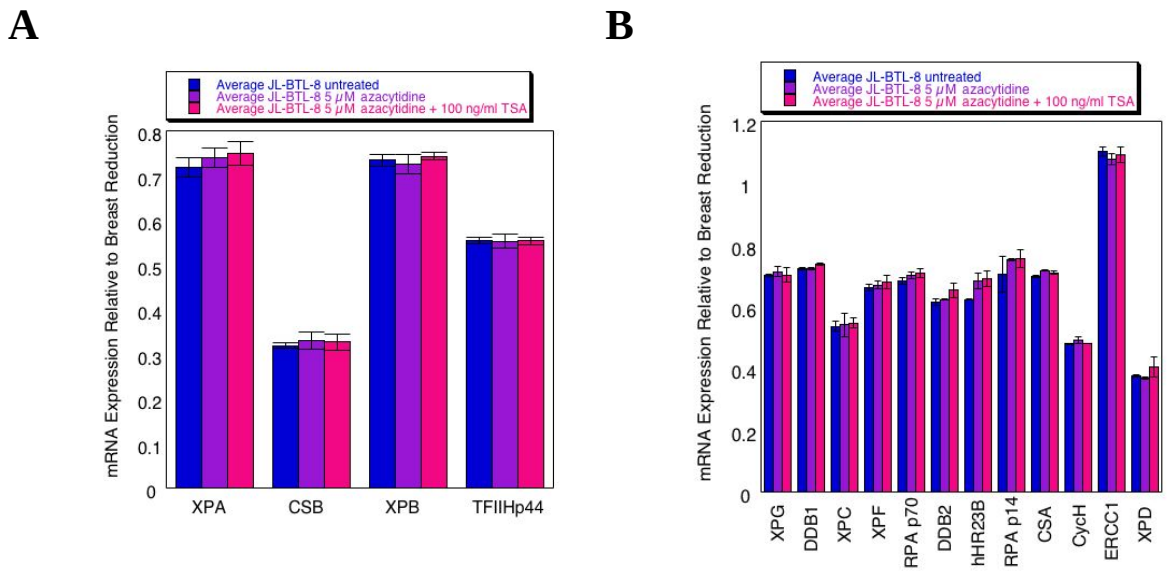


Figure 5.16 Sixteen genes do not show epigenetic regulation after treatment of the stage I cell line JL BTL-8 with 5-aza 2'deoxycytidine. A. The remaining 4 candidate genes that were identified by RNase protection were not seen to be regulated by methylation. Data represents the average of multiple (≥ 3) RNase protection analyses on two independent experimental replicates of JL BTL-8 treatment. Error between the replicates is shown.

B. 12 of the non-candidate genes were also not shown to be regulated by methylation. Treatments represented are the same as in panel A.

5.3.12 Restoration of expression of Candidate Genes Identified by Microarray Analysis

The use of microarray analysis not only allowed us to begin to find support for the results we had shown by RNase protection as discussed above, but also to assess whether any of the additional 14 genes we identified as being involved or related to the NER pathway were also significantly increased in their mRNA expression relative to untreated controls. We were also able to look at a quantification of the level of increase of the 14-3-3 σ gene that was seen with the nonquantitative RT-PCR method (Figure 5.17A). Of the candidate genes, only the expression of POLD2 was increased in the both the azacytidine and azacytidine with TSA groups, although the magnitude of increased expression was less than that seen with the 14-3-3 σ gene (Figure 5.15). Three genes that had not shown significant loss also seemed to show a gain of expression after azacytidine treatment: RFC2, TFIIHp62, and TTD-A. Two of these genes, TFIIHp62 and TTDA are again members of the TFIIH complex.

A **B**

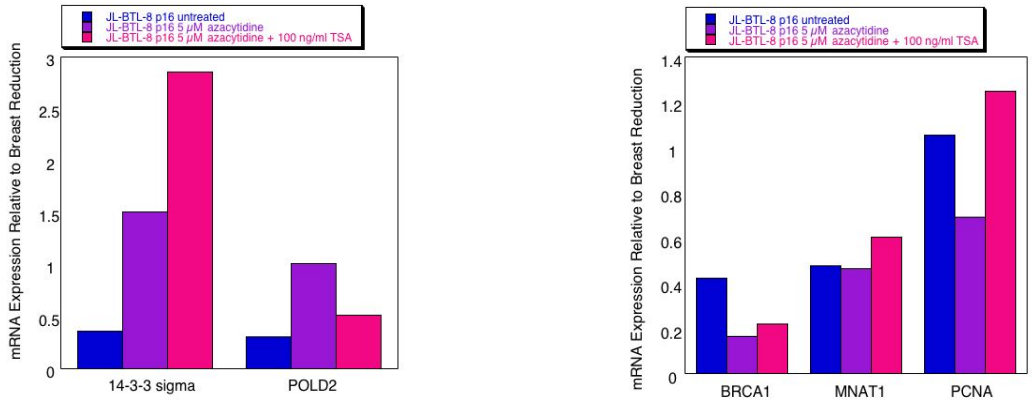


Figure 5.17 Five genes analyzed by microarray analysis show epigenetic regulation after treatment of the stage I cell line JL BTL-8 with 5-aza-2'deoxycytidine. A. 14-3-3 σ and POLD2 are regulated by methylation. The genes 14-3-3 σ, used as a control, and POLD2, which was identified as a candidate gene for loss based upon microarray analysis, both demonstrated increased mRNA levels after treatment with 5-aza-2'deoxycitidine relative to untreated controls. Results from the JL BTL-8 cell line are shown. B. Three other candidate genes identified by their loss in early stage tumor-derived cell lines were not regulated by methylation. BRCA1, MNAT1, and PCNA mRNA expression levels were not significantly increased after treatment with 5-aza-2'deoxycitidine relative to untreated controls. The same cell line is shown as in panel A.

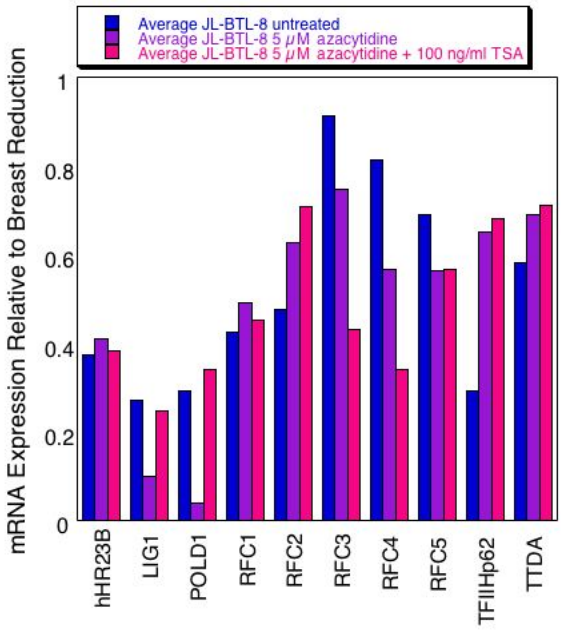


Figure 5.18 Ten genes that were not considered candidates for loss of expression in early stage tumor-derived cell lines were also evaluated for gain in expression after treatment with 5-aza-2'deoxycitidine by microarray analysis. Two genes, TFIIHp62 and TTDA showed gains in mRNA expression after treatment, even though they had not shown a significant loss in the comparison of early stage tumor-derived cell lines with those derived from breast epithelium. The results of treatment of the stage I cell line JL BTL-8 are shown.

5.3.13 Analysis of Transcription Factors in the Promoters of Candidate Genes

Because not all of our candidate genes were regulated by methylation, we also looked at the possibility of an as yet unknown set of transcription factors as the mechanism for co-regulation. Sequence analysis was carried out on the promoter regions of the candidate genes shown to have a decreased expression in early stage tumors to determine whether there was a coordinated set of transcription factors (a transcription factor model) that could explain their coordinate regulation. Promoters were identified from the Genomatix Promoter Database as described. When the information in the database was not listed as "gold" status, meaning that it had not been confirmed experimentally, we performed additional literature searchers to determine whether there had been a more recent update regarding the promoter status of the gene. Of the 20 genes from the RNase protection kits that we analyzed for expression, only ten had been experimentally verified. Of the remaining ten, we could find additional references for the promoter region of two genes: TFIIHp44 and TFIIHp34. Four genes were listed as "silver," or sequences that had been determined by the use of the Genomatix program Promoter Inspector: ERCC1, XPG, RPAp70, and DDB1. An additional four genes were listed as only having promoter sequences listed as annotated sequences: XPA, RPAp32, RPAp14, and XPF.

We carried out transcription factor analysis on the original set of candidate genes as defined by loss of repair capacity in early stage tumors relative to breast reduction using the Genomatix program Frameworker. There **were no promoter modules that were able to link all seven of these genes**. However, there were promoter modules that could explains subsets of these genes. The most specific of these were the four seven-element models that linked XPA and CSB (Figure 5.19). These models included the transcription factors: RXR, HNF1, HOXF,

PAX2/MYT1, TBPF, ERE/RORA, and HOMF (Table 1.22). The weight matrices assigned to the PAX2 and MYT1 as applied to these sequences made either of these two transcription factor families possible for binding at this location. Likewise, the ERE and RORA transcription factor families were also both candidates for the sixth binding position found. Interestingly, there was no subset that linked the three candidate genes whose expression was restored by azacytidine treatment.

The seven element model that was generated for XPA and CSB was applied across a database of human promoters compiled by Genomatix to see how many other promoters could be regulated by s a similar model. The specificity for this model was calculated at 7.99×10^{-7}, meaning that there was a 1 in 7.99×10^{7} chance of another human promoter sharing this model. *After 59,215 sequences and 3,119,345,110 base pairs were checked, only 5 listings had promoter sequences that also fit this model.* These corresponded to the genomic contiguous sequences of the XPA and CSB genes, the separate genomic sequences of these genes, and one clone that contained the CSB gene. Thus, *no other genes in the human genome to date share this pattern.*

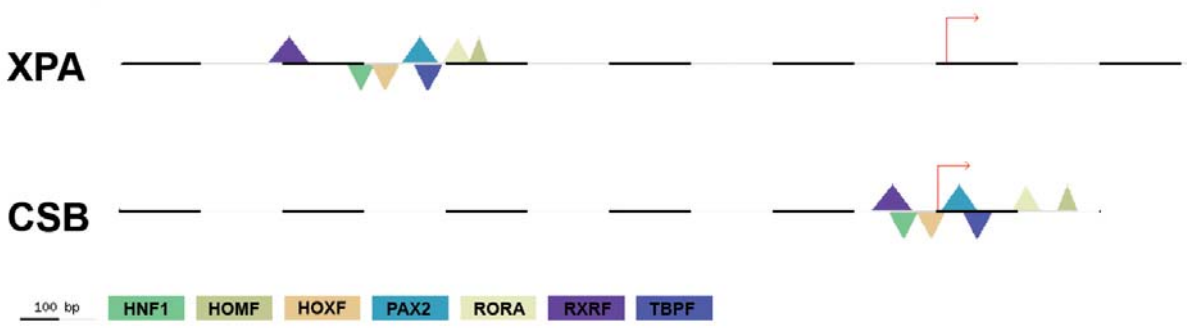


Figure 5.19 A seven-element model of transcription factors was found common to CSB and XPA. The position labeled as PAX2 could also be bound by MYT1, and the position bound by RORA could also be bound by ERE.

5.4 DISCUSSION

5.4.1 Loss of mRNA and Protein Expression in Early Stage Tumors

Analysis of mRNA expression levels with RNase protection and with microarray analysis showed a total of eleven genes had a significantly decrease in expression in three early stage tumor-derived NER deficient cell lines (derived from three different breast cancer patients) relative to three breast reduction-derived cell lines (derived from an additional three patients). Seven of these genes were identified through RNase protection and then supported by consistent results confirmed by microarray analysis, and an additional four genes appeared to decreased in mRNA expression by microarray analysis: MNAT1, PCNA, POLD2, and BRCA1. The genes that showed significant decreased expression varied in their functions in the repair pathway

(Figure 5.20). CSB is involved in transcription-coupled repair. The hypothesized function is to bind to the site of a RNA pol II complex that has during transcriptional elongation due to a DNA damage lesion.

As discussed in Chapter 1, XPA is involved in "lesion verification" in both global genomic repair and transcription-coupled repair. After damage has been recognized by either the transcription-coupled or global genomic DNA repair pathways, XPA is recruited to the site of the lesion as a dock on which other repair factors can assemble. PCNA and PolD2 are both involved in repair synthesis. PCNA binds to the RFC complex at the site where the lesion-containing segment of DNA has been excised. It helps to orient the DNA polymerase, polymerase δ so that it can synthesize the missing segment. POLD2 is the regulatory subunit of polymerase δ.

BRCA1 is included in this analysis because it is known to affect the regulation of two other genes in the NER pathway: XPC and DDB2. While both of these show a trend towards decrease of expression in early stage tumor-derived cell lines, neither is statistically significant, whereas the loss of BRCA1 mRNA expression is.

The remaining genes that have shown a significant decrease are all members of the TFIIH complex: TFIIHp52, p44, p34, XPB (p89), Cdk7, and MNAT1. The three others members of the complex: XPD (p80), Cyclin H, and TTD-A showed decreases that did not reach significance. The TFIIH complex is involved in unwinding the DNA helix around the site of a lesion in order to make it accessible for incision, excision, and replication. XPB and XPD are the helicases, and TFIIHp52, p44, and p34 all help to orient the TFIIH complex on the DNA and target these helicases. Cdk7, CycH, and MNAT1 are the CAK subcomplex. Their specific role in NER is unknown.

While it was only these 11 genes out of the total 34 (about 30%) that were analyzed that had a significant loss of repair, the majority of the genes tested showed some level of decreased expression. Therefore it is possible that there are contributions to this loss of function phenotype from many genes of the pathway. The inclusion of more samples may help to bring some of these additional genes to significance. This general trend points to the possibility that there is a level of control being exerted over the entire nucleotide excision repair pathway due to a global phenomenon in breast cancer.

Breast reduction-derived cell lines that were chosen for this analysis represented the range of DNA repair capacities seen in the primary tumors: a high repair, an average repair, and a lower than average repair. Likewise, the early stage tumor-derived cell lines were included one cell line that had greater than average repair for the early stage tumors and two that had less than average repair. The effect of this was that we chose samples that yielded higher interindividual variability. Therefore, when significance is achieved when comparing these three early stage tumors to these breast reduction samples, it has a greater chance of being able to explain the repair phenotype represented in the breast reduction and early stage tumor populations.

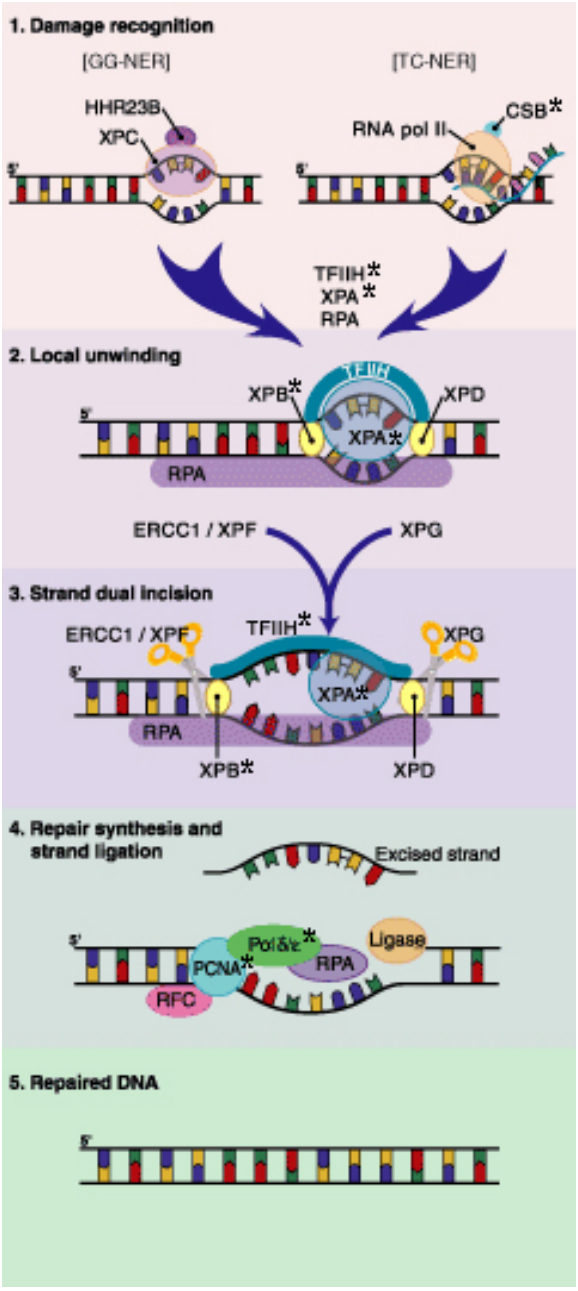


Figure 5.20 Schematic of NER with candidate genes identified by their loss of mRNA expression in early stage tumors marked with an asterisk (*). They are located both early and late in the pathway. First Printed in R&D Systems 2003 Catalog; reprinted with permission.

5.4.2 Systemic Polymorphisms in NER Genes Causing Cancer

Since we have seen a decrease in expression of our candidate genes, the idea that the women who are being diagnosed with breast cancer have single nucleotide polymorphisms (SNPs) that are responsible for their loss of repair capacity is less likely. Such a polymorphism would have to be associated with functional differences in gene expression or enzyme activity. While this is possible, *i.e.,* the polymorphism could be present in the promoter region of the gene effecting the binding of a needed transcription factor, it makes this idea less plausible.

5.4.3 Correlation of mRNA and Protein Expression

Few studies have been performed that relate the mRNA expression levels of NER genes with their protein expression levels. Some of these show that there is a correlation between the expression of mRNA and protein for ERCC1 (417), while others show that there is not (406). The researchers who showed that they did not find a correlation between ERCC1 mRNA and protein also did not find one for XPF and XPA.

In the field of DNA repair, protein biochemists have focused on the interactions of the repair proteins themselves in *in vitro* and *in vivo* situations. Should a difference in protein levels be found between samples, focus is given to post-translational modification rather than transcriptional regulation.

Molecular epidemiologists seek to use the mRNA expression levels to explain phenotypes across large populations. Those who study DNA repair capacity with one of the few tools available to do so, look at the order of recruitment of different proteins to the site of

damage, the effect of differentiation on repair capacity, and the interplay between transcription-coupled and global genomic repair. These fields stay unfortunately and artificially separated, shedding little light on how protein and mRNA levels agree.

In our study, however, we show that the loss and the gain of mRNA expression levels in two of our candidate genes and the lack of this pattern in two of our noncandidate genes is also observed at the protein level. This suggests that the expression of the XPA and CSB is controlled at the level of transcription rather than with post-translational modification or stability.

5.4.4 Gain of mRNA and Protein Expression in Late Stage Tumors

In contrast to loss of repair gene expression in early stage tumors relative to breast reduction, gain of repair gene expression in late stage tumors relative to early stage tumors seems to be a more global phenomenon. Of the 34 genes tested, 25 were increased in expression in the stage III tumor-derived cell line JL BTL-12, including six of our seven candidate genes. The large number of genes showing an increase in expression may in part be due to the use of only one cell line for the investigation of chemotherapy-naïve late stage tumors. More samples may help to show that like in the case of early stage tumors, where many genes were seen with a trend towards decreased expression, significance is only reached in some of them.

When early stage tumors are compared to the MCF7 cell line 31 out of 34 genes showed an increase in expression. Only ERCC1, TFIIHp62, and TTDA did not. In the case of MDA MB231, only ERCC1 failed to show a significant increase in mRNA expression. The high expression seen in the MCF7 and MDA MB231 cell lines is consistent with the very high level

of repair also seen in these cell lines. It may also reflect that fact that these cell lines are not therapy-naïve, and have been through some form of treatment. These cells may be representing only a subpopulation of the woman's tumor that survived treatment. The survival of the subpopulation may have been due to a more transformed phenotype, including greater deregulation of gene expression. These two cell lines also represent malignant pleural effusions rather than tumors that were confined to the breast and regional lymph nodes. This high expression may be a result of further genetic and epigenetic alterations required to move to a metastatic site.

The differences in expression seen with the MCF7 and MDA MB231 cell lines could also possibly result from the different culture conditions in which they are placed as compared to the cell lines derived in the Latimer lab and maintained using our own culture system. MCF-7 and MDA MB231 cell are maintained in DMEM with 10% fetal calf serum. In contrast, MWRI contains 20% serum in DMEM in addition to other nutrients. The effect that the rich medium MWRI and the use of recombinant basement membrane (matrigel) may have on expression is unknown and may be contributing to some of the differences seen in these stage IV cell lines relative to our cell lines.

When both early loss and then progressive gain of NER gene expression are considered, using all cell lines tested and all methods used to test them, our candidate genes are: CSB, XPA, XPB, TFIIHp52, TFIIHp44, TFIIHp34, Cdk7, MNAT1, PCNA, and POLD2.

5.4.5 Possible Mechanisms for a Loss and Gain of mRNA Expression

Two separate types of mechanisms can be considered for the regulation of DNA repair genes observed in early and late stage breast cancer cell lines. These include post-transcriptional controls, genetic, and epigenetic mechanisms.

The first type of mechanism that could be used is the genetic type. This could either involve the direct mutation of these genes such that their expression levels were compromised, or the mutation of a protein responsible for regulating that transcription. If we had seen only a single gene in the pathway that was affected, we could consider the mutation rate in human genome of 1 in 10^{12} (1 in 10^{6} per allele) as the likelihood of having this happen. Reversion (a 1 in 10^{8} phenomenon) of this mutation, such that the repair gene expression could be reversibly restored in a late stage tumor would then be 1 in 10^{20}.

The possibility of this occurring in all 10 of the candidate genes where gene expression was significantly lost and gained makes this theory even less feasible. The possibility that it is a transcription factor that is mutated and therefore driving both the decrease and increase of expression follows the same pattern. Loss would require the 1 in 10^{12} mutation of the transcription factor in the early stage tumors, and the 1 in 10^{20} chance that this would be reversed in the late stage. This theory could, however explain why multiple genes are effected, since it would theoretically take the mutation of only one gene to cause multiple downstream regulation effects. Unfortunately, there are no known transcription factors that effect this constellation of genes. Mutation in this transcription factor would also have to have happened in multiple patients since the same effect was seen in multiple cell lines, again making this a less feasible option for explaining this regulation phenomenon.

The second type of mechanism we have considered is epigenetic regulation, literally, control above the level of the gene's sequence. Epigenetic regulation includes both effects on chromatin structure and methylation of the promoter region. Both of these affect the ability of transcription factors to bind to the promoter region of the gene and effect transcription. Epigenetic regulation is also reversible. Both methylation and regulation of chromatin structure involve reversible covalent modifications of either the DNA itself (methylation) or the histones to which it is bound (chromatin structure).

In breast cancer, both DNA hypomethylation and hypermethylation are known to concurrently exist. Furthermore, such hypermethylation is a stage-specific phenomenon, showing that changes within the methylation state of a particular gene are possible not only during cancer etiology, but also during the process of tumor progression (*i.e.*, methylation of a particular gene is reversible). The fact that loss of DNA repair gene expression is seen in multiple early stage breast cancer cell lines derived from multiple patients relative to multiple breast reduction cell lines also derived from multiple patients and that this loss of repair also seems to be a reversible phenomenon, with later stage tumors, in general, expressing a higher level of these same repair genes makes epigenetic regulation the most plausible explanation of the three mechanisms.

If mRNA stability was altered, this would result in both a decrease of the transcript and the ultimate gene product being measured, as we have observed here. This idea could be used to explain the loss of repair that is seen in early versus late stage tumors. However, the mechanism behind the regulation of mRNA stability and its dysregulation in cancer is not fully understood. It is possible that this may be a stage specific and reversible phenomenon, but more study must be done in order to discover this in breast cancer.

5.4.6 The 5' regions of Most Nucleotide Excision Repair Genes are CpG Rich

Regulation of genes by methylation is associated with the methylation of CpG rich regions called CpG islands in the 5' regulatory regions of the genes. We therefore analyzed the sequences of our DNA repair genes to determine whether methylation was possible based on the presence of the CpG islands in their 5' ends. Analysis with three different definitions of a CpG island showed that 19 out of 20 of the genes included in the hNER1 and 2 kits from Pharmingen contained CpG islands in their 5' regions. Further analysis of the 14 genes of the NER pathway analyzed by microarray for their expression showed that 13 of these genes also contained Type II CpG islands.

Overall 94% of NER genes analyzed contained CpG islands, compared to the average of only 60% of genes in the human genome. The presence of these CpG islands makes regulation by CpG methylation plausible.

5.4.7 Determination of a Mechanism for the Functional Loss of NER in Breast Cancer Etiology

The addition of methyl groups onto cytosines, is a reversible process. The expression of ER (240), PR (418), BRCA1 (419), 14-3-3 σ (360,420), E-cadherin (421), and hMHL1 (422) have each been reactivated *in vitro* through the use of the cytosine analogue 5-aza-2'deoxycytidine. We therefore sought to use this drug treatment to evaluate the role of methylation in the regulation of our candidate genes.

When cells were treated with 5-aza-2'deoxycytidine, and their mRNA expression levels assayed by either RNase protection or microarray analysis, 4 of the 11 genes that had shown a

loss of DNA repair in early stage tumors showed a gain of expression: TFIIHp52, TFIIp34, Cdk7, and POLD2. This indicates that methylation can explain the loss of expression of these genes and may partly explain the loss of NER function. The other 7 genes, while they all possessed CpG islands in their regulatory regions, did not show a significant increase after treatment.

Furthermore the addition of the HDAC inhibitor Trichostain A did not significantly increase the expression of any of these genes. This can indicates that the chromatin was already in an open configuration for these genes.

Methylation has been further implicated in the phenomenon of DNA repair capacity in Chinese Hamster Ovary (CHO) cells. A hypermethylated derivative of a CHO line was established by passaging the parental line over months in increasing concentrations of 5-azacytidine. These cells then stably showed a more than 60% decrease in their total methylation. When transcription coupled repair of the dihydrofolate reductase reporter gene was examined, little or no effect was seen, except for repair in a small portion of the transcribed region of the gene a the 3' end. Researchers did not see a difference in either the non-transcribed strand or the 5' end of the transcribed strand. Thus, they reached the conclusion that methylation played an indirect role in the regulation of DNA repair, perhaps through affecting the chromatin structure. While this study is interesting, it has been carried out in a rodent repair system, which for reasons discussed in Chapter 1 is not a good comparison to the human system. Therefore, it would be of great interest to test the effect of global demethylation on DNA repair in a human cell line.

Within the context of what has been published in the literature comparing tumor versus normal (either in tissue or in the surrogate PBLs) for expression of mRNA, our results agree with

the decreased expression of CSB mRNA seen for both lung and squamous cell cancers of the head and neck. Our results with XPB mRNA agree with the loss of mRNA expression seen for squamous cell cancers of the head and neck, and also agree with the lack of loss of CSA and hHR23B mRNA in hepatocellular carcinoma. In the case of ERCC1, where previous studies have shown either increase (hepatocellular carcinoma), decrease (squamous cell cancer of the head and neck), or no change (lung) in expression depending upon the cancer type, our results agree with those who show no change. In contrast to the literature, we have seen a decrease, rather than an increase in XPA expression, and XPC shows a trend of decreased expression rather than increase. XPD also shown a trend towards decrease while the literature has shown no loss in the lymphocytes of lung cancer patients.

The lack of agreement in studies regarding whether mRNA expression of NER genes is increased or decreased in cancer may come from multiple sources. All of these cancers may have different etiologies. In hepatocellular carinoma, it is entirely possible that an increase in repair capacity occurs as an early event in carcinogenesis and is present in all stages. It is also just as possible that a loss of repair capacity is an early phenotype of lung and squamous cell carcinomas, and that different genes are responsible for the same phenotypes in different tissues.

In hepatocellular carcinoma, expression of the NER genes were either unchanged or increased. In lung and squamous cell cancer of the head and neck they were either unchanged or decreased in expression. One important difference that exists between the study of hepatocellular carcinoma in comparison to the studies on lung and squamous cell cancer was that only hepatocellular carcinoma study used the actual solid tumor of interest. The other three studies all used lymphocyte data as a surrogate.

Furthermore, none of these studies considered the stage of the tumors or sought to show that expression might be different with stage. If there was a preponderance of a particular stage in one of these patient samples, then the relative expression would have been driven by this group. For example, if the majority of the hepatocellular carcinomas in the study mentioned above were of a high stage, and there is a stage specific increase in expression as we have seen in breast cancer, then this would explain why researchers have observed an increase in mRNA expression of their three genes. It would not be an intrinsic mark of all hepatocellular carcinomas, rather a stage specific phenotype with these cancers.

5.4.8 Global Trends of Loss

An interesting alternative possibility for the genesis of loss of expression of NER genes during carcinogeneis may be the effects of environmental exposures. Rather than looking for a purely genetic reason why these repair genes are expressed at low levels, this loss of expression phenotype may represent the interplay of genetics and environmental exposure. One study showed that there was an inverse relationship between the level of arsenic found in the toenails of individuals who had been exposed to drinking water with high levels of arsenic and the expression of multiple genes involved in NER in their lymphocytes (423). ERCC1, XPF, and XPB were all shown to be decreased in expression in people who had high levels of arsenic in their toenail tissue. XPA and XPG were also tested, but did not show such an effect. This opens up the possibility that effects of the environment, including exposure to toxic agents, may be modulating the repair capacity of exposed individuals. Not only would increased risk of cancer

be due to the exposure of these people to the genotoxic effects of the chemical, but a combination effect with the modification of their ability to deal with the damage.

5.4.9 Transcription Factors Linking Candidate Genes

The constellation of transcription factors that were shown to link the CSB and XPA transcription factors, RXR, HNF1, HOXF, PAX2/MYT1, TBPF, ERE/RORA, HOMF can be considered in terms of other known effects of these transcription factors already implicated in human cancers.

The HOX1 family includes 39 genes that are used in development to determine the identity of different body segments and organs along the body's axis. Any of these family members could bind to the HOXF site identified in our model. Quantitative RT-PCR experiments have also shown that 11 of these HOX genes are dysregulated in human breast cancer. Two are putative oncogenes during carcinogenesis, HOXA1 and HOXB7.

HOXA1 overexpression causes the transcriptional upregulation of the anti-apoptotic factor BCL-2. HOXB7, which is amplified in over 80% of primary breast tumors, causes direct transcriptional upregulation of Fibroblast Growth Factor (FGF). In contrast to these HOXA5 acts as a tumor suppressor gene that is lost in 60% of breast cancers due to promoter hypermethylation. HOX5A expression is associated with upregulation of p53 expression and increased apoptosis in breast tumor cells.

HOX5A also directly induces expression of E-cadherin, a gene whose loss is implicated in the epithelial to mesenchymal transition associated with breast cancer invasion. Since hypermethylation causes the loss of a factor that induces E-cadherin, the loss of this transcription factor may play a role in the conversion of cells from an epithelial to mesenchymal phenotype in

cancer. Since the epithelial to mesenchymal transition is something that we have associated with the behavior of tumors in our breast cancer systems, it is consistent with the idea that this transcription factor may be influencing regulation.

PAX2 is a member of the PAX family of transcription factors, usually associated with the maintenance of pleuripotence of stem cells during development. In the normal breast, PAX2 in required for progesterone-induced cell growth and branching of the normal mammary epithelial cells during pregnancy. PAX2 can also act as an oncogene that is overexpressed in 50% of breast cancers. PAX2 expression in primary breast tumors correlates positively with the expression of the ER and the PR, and PAX2 is itself a target for ER regualtion in endometrial cancers, although this effect has not need shown in breast tumors as yet. Thus, PAX2 may play a role in hormone-mediated carcinogenesis. This transcription factor may also have something to do with the tissue-specific nature of DNA repair capacity of both breast epithelium and ovarian epithelium, as discussed in Chapter 2. If this transcription factor does play a role in the regulation of expression of these two genes in the NER pathway, and is present in normal breast, perhaps it may regulate these two genes in normal breast. Also interesting is the fact that this transcription factor is expressed in the normal breast for activities during pregnancy, and full-term pregnancy is known to be a protective event with respect to breast cancer. Both of these ideas are highly speculative, however.

5.4.10 Increased Expression of Nucleotide Excision Repair Genes in Other Cancers and the Association with Chemoresistance

Because the products of the NER genes are involved in the remediation of cross-links induced by platinum-based agents, researchers have looked for a link between expression of some NER genes and response to chemotherapy. The majority of this research has been done in the area of ovarian cancer therapy, since the platinum-based chemotherapies are still typically used for treatment of this cancer. This research is based upon the idea that if cells are expressing higher amounts of genes in the NER pathway, they are more capable of dealing with the genotoxic effects of the pathway. If they can deal with the genotoxic events, they are more likely to survive the insult, resulting in a decreased response to the chemotherapeutic regimen and a decreased survival. Studies have been published that show all three possibilities: an increase in expression correlates with decreased response and survival, no change in expression, and a decrease in expression leads to a decreased survival (424).

One study showed that patients who are resistant to platinum-based therapies express higher levels of XPB, CSB, XPA, and ERCC1 (425). In colorectal cancer, decreased ERCC1 mRNA expression correlated with better survival after platinum therapy (426). Yet another study that was conducted using the MCF7 cell line and its alkylating-agent resistant subline MCF7-MLNr showed that there was no difference in the mRNA expression levels of ERCC1, XPD, and XPB between the two lines even though there was a difference in the detected level of transcription-coupled repair (417). In contrast, yet another study showed that there was an association between decreased ERCC1 expression and decreased patient survival, suggesting that increased ERCC1 expression is protective in non-small-cell lung cancer (427).

The results of our analyses showing that there is stage specificity with respect to mRNA expression suggest that in breast cancer stage must be considered when looking for a relationship between mRNA levels and chemotherapeutic responsiveness. When put into the context of literature, it again suggests that different cancer types have different etiologies and should be considered separately. The effects that are present in lung cancer may not be true for ovarian cancer. Cisplatin-based therapies are not commonly used in breast cancer patients unless they are being used as "rescue" therapies, those that are reserved for recurrent patients or those who have failed to respond to their initial course of therapy. However, the drug cyclophosphamide creates bulky adducts which are also remediated by the NER pathway. This type of study could potentially be adapted to the breast using our candidate genes and response to chemotherapeutic agents.

6.0 CONCLUSIONS

Protein biochemists traditionally focus on the molecular interactions of various proteins in the NER pathway. Lines of research have included (1) the elucidation of the factors necessary for NER to be re-created in a cell-free in vitro system, (2) the ability of the "recognition" proteins-XPC-hHR23B, DDB1, and DDB2 to bind to different types of DNA damage, (3) the existence of these proteins in pre-formed complexes in the nucleus,

And, more recently, (4) the kinetics of binding of each of these proteins to the sites of DNA damage within the cell. Biochemical analysis takes place in great part in cell-free *in vitro* systems. When these researchers wish to present a biological context for the results observed *in vitro*, their chosen *in vivo* systems frequently include SV-40 transformed skin fibroblasts from patients with XP. These studies have generated innumerable pieces of information for the DNA repair community, including the functions of many of the genes involved in NER, and in large-part the stepwise mechanism of the pathway itself.

Because of the bias of protein biochemists, Western blot analysis has become the standard currency of discussion for DNA repair deficiencies. Attention has been placed on protein expression levels and protein stability to the exclusion of functional assays and potential transcriptional regulation. Other potential reasons for this bias away from functional assays are presented below.

In contrast to the work of protein biochemists, the current trend in the molecular epidemiology literature appears to be towards PCR-based molecular analyses of gene expression and of DNA repair gene polymorphisms. They do, however, agree with the protein biochemists in their lack of interest in functional assays of DNA repair capacity. Lines of research have included both (1) differentiation of transcription levels of genes involved in NER in cancer cases versus controls by multiplex RT-PCR analysis, and (2) the investigation of a wide array of polymorphisms, and SNPs in particular, in genes of the NER pathway as they relate to increased cancer risk.

While molecular epidemiologists are interested in the mRNA expression of repair genes, they seem completely uninterested in the mechanism of regulation. Techniques of 5' end analysis have been available now for several years, but shockingly little has yet to be published on the 5' end analyses of the genes of the NER pathway.

Neither the protein biochemists nor the molecular epidemiologists demonstrate an interest in the use of functional assays in the cell or tissue type of interest, but rather use PBLs from patients with breast cancer and their associated controls. Protein biochemists associate protein expression with function, and molecular epidemiologists associate cancer with function. Neither one of these associations is necessarily correct.

We have presented in this work a marriage between the three lines of investigation of DNA repair: the often-ignored functional repair capacity, mRNA expression, and protein expression. The placement of the context of functional repair capacity on these expression data gives them a relevance sorely lacking from other research, while the inclusion of both mRNA and protein expression data allows this work to be accessible to both the protein biochemists and molecular epidemiologists.

6.1 THE DEFICIT OF FUNCTIONAL ASSAYS TO ASSESS NUCLEOTIDE EXCISION REPAIR

The lack of focus on functional assays may in large part be due to the lack of functional assays that can quantify DNA repair. Three main techniques are used in the repair community by a small subset of researchers. The first of these techniques, which is the one we have used, is the UDS assay. This assay is technically challenging, with a great requirement for attention to details, and ideally, the construction of a light source that can reproducibly generate a specified dose of UVC light. This is prohibitive for many researchers and is thus not a common practice. UDS can detect problems in all steps of the pathway up until the ligation of the newly synthesized fragment of DNA. This is exemplified by the normal repair measured by the UDS assay in the 46BR case, where a mutation in the DNA ligase 1 gene was found.

More recently, a new technique using damage-specific antibodies has been created. Cells are damaged with either UV light or benzo[*a*]pyrene diol-epoxide (BPDE) and damage directed antibodies are allowed to bind. These antibodies recognize cyclobutane pyrimidine dimers and 6-4 photoproducts in case of UV light damage, and distinctive BPDE-DNA adducts in the case of BPDE exposure. The rate of the disappearance of these bound antibodies over time is considered a measure of the number of sites of damage that NER has repaired.

This antibody-based technique, like the UDS assay, measures total genomic repair. Due to its method of detection, it cannot measure the efficiency at stages in the pathway past the level of excision of the DNA containing the damage lesion. Thus, deficiencies of the RFC complex, PCNA, or the DNA polymerase (delta or eta) would not be detected. The damage site would be gone, but the gap would still be unfilled by repair synthesis. Full mutations in any of these genes

would most likely be incompatible with life, as is seen by the lack of a known human mutation in any of the genes involved in this synthesis: PCNA, RFC1-5, and DNA polymerase delta. These genes also have functions outside of repair synthesis, specifically in DNA replication that would make their omission or mutation catastrophic in an organism.

Our data, however, suggests that there may be regulation of these genes at the transcriptional level. Microarray analyses demonstrate a significant loss of POLD2 expression in early stage tumors relative to breast reduction samples. This suggests that there is regulation at least at the transcriptional level of factors involved in DNA repair synthesis. It will be interesting in future studies to determine if there is also a decrease in protein expression of this gene. Furthermore, this could lead to cross-over with another line of investigation in DNA repair: error-prone replication.

Researchers are currently investigation the role of low-fidelity DNA polymerases in DNA damage tolerance. Perhaps if there is a decrease in the relatively high fidelity DNA polymerase δ, there is a co-ordinate increase of expression in other DNA polymerases such at polymerase ζ to compensate. This is truly speculative, however, but an interesting potential direction for studies in the future. Antibodies to polymerase ζ components are still under production, so it will be some time before this speculation can even be addressed.

6.2 FUTURE EXPERIMENTATION

6.2.1 Implications of the Latimer Culture System for the Study of Breast Cancer

The Latimer culture system offers the possibility of marrying epidemiologic studies, where interindividual variation is accepted, and understood. Studies with molecular biology where n = 1 can generate an accepted phenomenon. As has been mentioned before, the Latimer culture system of primary breast reductions and tumor tissue, as well as the generation of cell lines from these primary cultures, presents an opportunity to observe the heterogeneity of breast cancer and the variation in the human population in general. Specific attention can be paid in future experiments to the inclusion of breast cancer patients of different races because of this. Further analysis of current repair data is being completed regarding race and repair at this time, however, more samples will be needed to generate significant trends.

Many elegant molecular biologic characterizations have been made of physiologically irrelevant phenomenon secondary to poor cell culture choices. The advent of this cell culture technology can remediate some of these deficiencies in future experiments.

Further analysis can be completed and is currently being completed on functional repair capacity data presented here. These analyses include multivariate analysis taking in to account the pathologic characteristics of the tumors: ER, PR, grade, and lymph node status. If NER function is independent of these currently used clinical markers, it shows vastly more promise as a clinical marker of its own because of the power that it can add to the analysis.

Data is being collected on the tumors which we have characterized by functional repair capacity assay to determine whether low repair correlates with increased disease-free and overall survival. If this correlation proves to be true, the identification of molecular markers that corresponding to functional repair capacity (such as mRNA expression) could serve as surrogates for the repair capacity assay. This would prevent the possibility of having to run the UDS assay on all breast cancer tumor samples. This would be too prohibitive in terms of cost and time.

Further analysis for the normal ovarian epithelial tissue samples will also be completed in the near future. Demographic information is being collected on these samples to be able to look at the effect on hormonal factors such as hormone replacement therapy or other exogenous sources of estrogen, menopausal status, gravity, and parity. This is of particular interest with regards to the mitogen-mutagen theory of cancer.

6.2.2 Future Analysis of the Molecular Mechanism of DNA Repair Deficiency

If I were to continue this work I would investigate the individual effects that each of these genes exert on overall repair capacity. Does the decrease in expression of one of these genes cause the co-ordinate decrease in expression of others ?

One way to begin to answer these questions would be with the use of siRNA technology. If the mRNA level of a specific gene could be targeted and knocked down, the functional repair capacity measured by UDS could then be assessed and the relative contribution of this gene to the loss of repair phenotype could be determined. In this knocked down breast reduction cell line, the mRNA expression levels of the other genes could then also be assessed in order to determine if there is co-ordinate regulation.

Specifically since the CSB gene is involved in transcription-coupled repair, the observed reduction in CSB expression should result in a functional deficiency in transcription-coupled repair. Since transcription-coupled repair is a component of total genomic repair, this drop could, in part, account for the drop in functional DNA repair. In order to determine the role of CSB and transcription-coupled repair in total genomic repair we propose to perform the Host

Cell Reactivation Assay (HCR) using a damaged luciferase vector transfected into repair deficient cell lines.

Because they are part of a complex, each one of the TFIIH genes will have a measurable effect on functional NER. These genes and CSB may also interact with each other via feedback mechanisms.

One could now perform siRNA knock-down experiments for each of the candidate genes and evaluate the effect on the expression of both the gene of interest and additional NER genes via multiplex RNase protection. Comparisons of levels of expression will also be made using isogenic pair of normal and tumor cell lines. Finally, one could perform the functional UDS assay on these treated cell lines, within the 48 hour window for siRNA effect, to quantitatively establish the effect of reduced gene expression on NER capacity.

BIBLIOGRAPHY

1. Hanahan, D. and Weinberg, R. (2000) The hallmarks of cancer. *Cell*, **100**, 57-70.

2. Ernberg, I. (1990) Oncogenes and tumor growth factors in breast cancer. *Acta Oncol*, **29**, 331-334.

3. Tripathy, D. and Benz, C. (1992) Activated oncogenes and tumor suppressor genes involved in human breast cancers. *Breast Cancer Res Treat*, **63**, 13-60.

4. Aaltonen, L., Peltomaki, P., Leach, F., Sistonen, P., Pylkkanen, L., Mecklin, J., Jarvinen, H., Powell, S., Jen, J. and Hamilton, D. (1993) Clues to the pathogenesis of familial colorectal cancer. *Science*, **260**.

5. Friedberg, E., Walker, G., Siede, W., Wood, R., Schultz, R. and Ellenberger, T. (2006), *DNA Repair and Mutageneis, Second Edition*. ASM Press, Washington, D.C.

6. Bjelland, S. and Seeberg, E. (2003) Mutagenicity, toxicity, and repair of DNA base damage induced by oxidation. *Mutat Res*, **531**, 37-80.

7. Demple, B. and Linn, S. (1982) 5,6-Saturated thymide lesions in DNA: production by ultraviolet light and hydrogen peroxide. *Nucleic Acids Res*, **10**, 3781-3789.

8. Friedberg, E., Walker, G., Siede, W., Wood, R., Schultz, R. and Ellenberger, T. (2006), *DNA Repair and Mutageneis, Second Edition*. ASM Press, Washington, D.C.

9. Singer, B. (1986) O-Alkyl pyrimidines in mutagenesis and carcinogenesis: ocurrence and significance. *Cancer Res*, **46**, 4879-4885.

10. Kozekov, I., Nechev, L., Moseley, M., Harris, C., Rizzo, C., Stone, M. and Harris, TM. (2003) DNA interchain cross-links formed by acrolein and crotonaldehyde. *J Amer Chem Soc*, **125**, 50-61.

11. Hearst, J., Isaacs, S., Kanne, D., Rapoport, H. and Straub, K. (1984) The reaction of psoralens with deoxyribonucleic acid. *Q Rev Biophys*, **17**, 1-44.

12. Falany, C. and Wilborn, T. (1990) Biochemistry of cytosolic sulfotransferases involved in bioactivation. *Adv Pharmacol*, **27**, 301-363.

13. Weinstein, I., Jeffrey, A., Jenete, K., Blobstein, S., Harvey, R., Harris, C., Autrup, H., Kasai, H. and Nakanishi, K. (1976) Benzo[a]pyrene diol epoxides as intermediates in nucleic acid binding in vitro and in vivo. *Science*, **193**, 592-594.

14. Smela, M., Currier, S., Bailey, E. and Essigmann, J. (2001) The chemistry and biology of aflatoxin B1: from mutational spectrometry to carcinogenesis. *Carcinogenesis*, **22**, 535-545.

15. Cavalieri, E., Frenkel, K., Liehr, J., Rogan, E. and Roy, D. (2000) Estrogens as endogenous genotoxic agents--DNA adducts and mutations. *J Natl Cancer Inst Monogr*, **27**, 75-93.

16. Povirk, L. (1996) DNA damage and mutageneis by radiometric DNA-cleaving agents: bleomycin, neocarzinostatin and other enediynes. *Mutat Res*, **355**, 71-89.

17. Ferguson, L. and Baguley, B. (1996) Mutagenicity of anti-cancer drugs that inhibit topoisomerase enzymes. *Mutat Res*, **355**, 91-101.

18. (2003) *Physician's Desk Reference Oncology Reference Guide*. Thompson PDR, Montvale, NJ.

19. Povirk, L. and Shuker, D. (1994) DNA damage and mutagenesis induced by nitrogen mustards. *Mutat Res*, **318**, 205-226.

20. Drablos, F., Feyzi, E., Aas, P., Vaagbo, C., Kavli, B., Bratlie, M., Pena-Diaz, J., Otterlei, M., Slupphaug, G. and Krokan, H. (2004) Alkylation damage in DNA and RNA--repair mechanisms and medical significance. *DNA Repair (Amst)*, **3**, 1389-1407.

21. Chen, J. and Stubbe, J. (2004) Bleomycins: new methods will allow reinvestigation of old issues. *Curr Opin Chem Biol*, **8**, 175-181.

22. Binaschi, M., Bigioni, M., Cipollone, A., Rossi, C., Goso, C., Maggi, C., Capranico, G. and Animati, F. (2001) Anthracyclines: selected new developments. *Curr Med Chem Anticancer Agents*, **1**, 113-130.

23. Costi, M., Ferrari, S., Venturelli, A., Calo, S., Tondi, D. and Barlocco, D. (2005) Thymidylate synthase structure, function and implication in drug discovery. *Curr Med Chem*, **12**, 2241-2258.

24. Disinfection by UV Light. *http://www.lenntech.com/will1.htm#uvc*.

25. Mitchell, D. and Nairn, R. (1989) The biology of the 6-4 photoproduct. *Photochem Photobiol*, **49**, 805-819

26. Park, H., Zhang, K., Ren, Y., Nadji, S., Sinha, K., Taylor, J. and Kang, C. (2002) Crystal structure of a DNA decamer containing a cis-syn thymine dimer. *Proc Natl Acad Sci U S A*, **99**, 15965-15970.

27. Tornaletti, S., Rozek, D. and Pfiefer, G. (1993) The distribution of UV photoproducts along the human p53 gene and its relationship to mutations in skin cancer. *Oncogene*, **8**, 2051-2057.

28. Friedberg, E., Walker, G., Siede, W., Wood, R., Schultz, R. and Ellenberger, T. (2006), *DNA Repair and Mutageneis, Second Edition*. ASM Press, Washington, D.C.

29. Friedberg, E., Walker, G., Siede, W., Wood, R., Schultz, R. and Ellenberger, T. (2006), *DNA Repair and Mutageneis, Second Edition*. ASM Press, Washington, D.C.

30. Szymkowski, D., Lawrence, C. and Wood, R. (1993) Repair by human cell extracts of single (6-4) and cyclobutane thymine-thymine photoproducts in DNA. *Proc Natl Acad Sci U S A*, **90**, 9823-9827.

31. Kantor, G. and Setlow, R. (1981) Rate and extent of DNA repair in nondividing human diploid fibroblasts. *Cancer Res*, **41**, 819-825.

32. Hanawalt, P., Ford, J. and Lloyd, D. (2003) Functional characterization of global genomic DNA repair and its implications for cancer. *Mutat Res*, **544**, 107-114.

33. Mullenders, L., Stary, A. and Sarasin, A. (2001) Nucleotide Excision Repair. *Atlas Genet Cytogenet Oncol Haematol*, **URL :** **http://AtlasGeneticsOncology.org/Deep/ExcisRepairID20014.html**.

34. Branum, M., Reardon, J.T. and Sancar, A. (2001) DNA repair excision nuclease attacks undamaged DNA. A potential source of spontaneous mutations. *J Biol Chem*, **276**, 25421-25426.

35. Mullenders, L.H. (1998) Transcription response and nucleotide excision repair. *Mutat Res*, **409**, 59-64.

36. Mellon, I., Bohr, V., Smith, C. and Hanawalt, P. (1986) Preferential DNA repair of an active gene in human cells. *Proc Natl Acad Sci U S A*, **83**, 8878-8882.

37. Bohr, V., Smith, C., Okumoto, D. and Hanawalt, P. (1985) DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall. *Cell*, **40**, 359-369.

38. Seeberg, E. and Fuchs, R. (1990) Acetylaminofluorene bound to different guanines of the sequence -GGCGCC- is excised with different efficiencies by the UvrABC excision nuclease in a pattern not correlated to the potency of mutation induction. *Proc Natl Acad Sci U S A*, **87**, 191-194.

39. van Hoffen, A., Venema, J., Meschini, R., van Zeeland, A. and Mullenders, L. (1995) Transcription-coupled repair removes both cyclobutane pyrimidine dimers and 6-4 photoproducts with equal efficiency and in a sequential way from transcribed DNA in xeroderma pigmentosum group C fibroblasts. *EMBO J*, **14**, 360-367.

40. Citterio, E., Van Den Boom, V., Schnitzler, G., Kanaar, R., Bonte, E., Kingston, R.E., Hoeijmakers, J.H. and Vermeulen, W. (2000) ATP-dependent chromatin remodeling by the Cockayne syndrome B DNA repair-transcription-coupling factor. *Mol Cell Biol*, **20**, 7643-7653.

41. Selby, C.P. and Sancar, A. (1997) Cockayne syndrome group B protein enhances elongation by RNA polymerase II. *Proc Natl Acad Sci U S A*, **94**, 11205-11209.

42. Bradsher, J., Auriol, J., Proietti de Santis, L., Iben, S., Vonesch, J.L., Grummt, I. and Egly, J.M. (2002) CSB is a component of RNA pol I transcription. *Mol Cell*, **10**, 819-829.

43. Kamiuchi, S., Saijo, M., Citterio, E., de Jager, M., Hoeijmakers, J.H. and Tanaka, K. (2002) Translocation of Cockayne syndrome group A protein to the nuclear matrix: possible relevance to transcription-coupled DNA repair. *Proc Natl Acad Sci U S A*, **99**, 201-206.

44. Wakasugi, M. and Sancar, A. (1999) Order of assembly of human DNA repair excision nuclease. *J Biol Chem*, **274**, 18759-18768.

45. Wakasugi, M., Shimizu, M., Morioka, H., Linn, S., Nikaido, O. and Matsunaga, T. (2001) Damaged DNA-binding protein DDB stimulates the excision of cyclobutane pyrimidine dimers in vitro in concert with XPA and replication protein A. *J Biol Chem*, **276**, 15434-15440.

46. Wold, M. and Kelly, T. (1988) Purification and characterization of replication protein A, a cellular protein required for in vitro replication of simian virus 40 DNA. *Proc Natl Acad Sci U S A*, **85**, 2523-2527.

47. de Laat, W., Appeldoorn, E., Sugasawa, K., Weterings, E., Jaspers, N. and Hoeijmakers, J. (1998) DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes Dev*, **12**, 2598-2609.

48. Jawhari, A., Laine, J., Dubaele, S., Lamour, V., Poterszman, A., Coin, F., Moras, D. and Egly, J. (2002) p52 Mediates XPB function within the transcription/repair factor TFIIH. *J Biol Chem*, **277**, 31761-31767.

49. Egly, J. (2001) TFIIH: from transcription to clinic. *FEBS Lett*, **498**, 124-128.

50. Jawhari, A., Boussert, S., Lamour, V., Atkinson, R., Kieffer, B., Poch, O., Potier, N., van Dorsselaer, A., Moras, D. and Poterszman, A. (2004) Domain architecture of the p62 subunit from the human transcription/repair factor TFIIH deduced by limited proteolysis and mass spectrometry analysis. *Biochemistry*, **43**, 14420-14430.

51. Samkurashvili, I. and Luse, D. (1996) Translocation and transcriptional arrest during transcript elongation by RNA polymerase II. *J Biol Chem*, **271**, 23495-23505.

52. Conaway, J., Dvir, A., Moreland, R., Yan, Q., Elmendorf, B., Tan, S. and Conaway, R. (1998) Mechanism of promoter escape by RNA polymerase II. *Cold Spring Harb Symp Quant Biol*, **63**, 357-364.

53. Reines, D., Conaway, R. and Conaway, J. (1999) Mechanism and regulation of transcriptional elongation by RNA polymerase II. *Curr Opin Cell Biol*, **11**, 342-346.

54. Svejstrup, J. (2003) Rescue of arrested RNA polymerase II complexes. *J Cell Sci*, **116**, 447-451.

55. Lodish, H. (2000) In Lodish, H., Berk, A., Zipursky, L., Matsudaira, P., Baltimore, D. and Darnell, J. (eds.), *Mol Cell Biol*. W. H. Freeman, New York.

56. Léveillard, T., Andera, L., Bissonnette, N., Schaeffer, L., Bracco, L., Egly, J. and Wasylyk, B. (1996) Functional interactions between p53 and the TFIIH complex are affected by tumour-associated mutations. *EMBO J*, **15**, 1615-1624.

57. Lu, H., Fisher, R., Bailey, P. and Levine, A. (1997) The CDK7-cycH-p36 complex of transcription factor IIH phosphorylates p53, enhancing its sequence-specific DNA binding activity in vitro. *Mol Cell Biol*, **17**, 5923-5934.

58. Enzlin, J. and Scharer, O. (2002) The active site of the DNA repair endonuclease XPF-ERCC1 forms a highly conserved nuclease motif. *EMBO J*, **21**, 2045-2053.

59. Painter, R. and Cleaver, J. (1969) Repair replication, unscheduled DNA synthesis, and the repair of mammalian DNA. *Radiat Res*, **37**, 451-466.

60. Cleaver, J. and Thomas, G. (1981) In Friedberg, E. and Hanawalt, P. (eds.), *DNA Repair, A Laboratory Manual of Research Procedures, Volume I*. Marcel Dekker, New York, Vol. 277-287.

61. Cleaver, J.E. (1984) DNA repair deficiencies and cellular senescence are unrelated in xeroderma pigmentosum cell lines. *Mech Ageing Dev*, **27**, 189-196.

62. Cleaver, J.E. (1968) Defective repair replication of DNA in xeroderma pigmentosum. *Nature*, **218**, 652-656.

63. Nakagawa, A., Kobayashi, N., Muramatsu, T., Yamashina, Y., Shirai, T., Hashimoto, M., Ikenaga, M. and Mori, T. (1998) Three-dimensional visualization of ultraviolet-induced DNA damage and its repair in human cell nuclei. *J Invest Dermatol*, **110**, 143-148.

64. Imoto, K., Kobayashi, N., Katsumi, S., Nishiwaki, Y., Iwamoto, T., Yamamoto, A., Yamashina, Y., Shirai, T., Miyagawa, S., Dohi, Y. *et al.* (2002) The total amount of DNA damage determines ultraviolet-radiation-induced cytotoxicity after uniformor localized irradiation of human cells. *J Invest Dermatol*, **119**, 1177-1182.

65. Nishiwaki, Y., Kobayashi, N., Imoto, K., Iwamoto, T., Yamamoto, A., Katsumi, S., Shirai, T., Sugiura, S., Nakamura, Y., Sarasin, A. *et al.* (2004) Trichothiodystrophy fibroblasts are deficient in the repair of ultraviolet-induced cyclobutane pyrimidine dimers and (6-4)photoproducts. *J Invest Dermatol*, **122**, 526-532.

66. Kennedy, D., Agrawal, M., Shen, J., Terry, M., Zhang, F., Senie, R., Motykiewicz, G. and Santella, R. (2005) DNA repair capacity of lymphoblastoid cell lines from sisters discordant for breast cancer. *J Natl Cancer Inst*, **97**, 127-132.

67. Matijasevic, Z., Precopio, M., Snyder, J. and Ludlum, D. (2001) Repair of sulfur mustard-induced DNA damage in mammalian cells measured by a host cell reactivation assay. *Carcinogenesis*, **22**, 661-664.

68. Johnson, J.M. and Latimer, J.J. (2005) Analysis of DNA repair using transfection-based host cell reactivation. *Methods Mol Biol*, **291**, 321-335.

69. Protic-Sabljic, M. and Kraemer, H. (1985) One pyrimidine dimer inactivates expression of a transfected gene in xeroderma pigmentosum cells. *Proc Natl Acad Sci U S A*, **82**, 6622-6626.

70. Athas, W., Hedayati, M., Matanoski, G., Farmer, E. and Grossman, L. (1991) Development and field-test validation of an assay for DNA repair in circulating human lymphocytes. *Cancer Res*, **51**, 5786-5793.

71. Berwick, M. and Veneis, P. (2000) Markers of DNA repair and susceptibility to cancer in humans: an epidemiologic review. *J Natl Cancer Inst*, **92**, 847-897.

72. Nouspikel, T. and Hanawalt, P. (2002) DNA repair in terminally differentiated cells. *DNA Repair (Amst)*, **1**, 59-75.

73. Robbins, J., Kraemer, K., Lutzner, M., Festoff, B. and Coon, H. (1974) Xeroderma Pigmentosum. An inherited disease with sun sensitivity, multiple cutaneous neoplasms, and abnormal DNA repair. *Ann Intern Med*, **80**, 221-248.

74. Takebe, H., Miki, Y., Kozuka, T., Furuyama, J. and Tanaka, K. (1977) DNA repair characteristics and skin cancers of xeroderma pigmentosum patients in Japan. *Cancer Res*, **37**, 490-495.

75. Woods, C. (1998) DNA repair disorders. *Arch Dis Child*, **78**, 178-184.

76. Vermeulen, W., Scott, R., Rodgers, S., Muller, H., Cole, J., Arlett, C., Kleijer, W., Bootsma, D., Hoeijmakers, J. and Weeda, G. (1994) Clinical heterogeneity within xeroderma pigmentosum associated with mutations in the DNA repair and transcription gene ERCC3. *Am J Hum Genet*, **54**.

77. Norris, P., Limb, G., Hamblin, A., Lehmann, A., Arlett, C., Cole, J., Waugh, A. and Hawk, J. (1990) Immune function, mutant frequency, and cancer risk in the DNA repair defective genodermatoses xeroderma pigmentosum, Cockayne's syndrome, and trichothiodystrophy. *J Invest Dermatol*, **94**, 94-100.

78. Cole, J., Green, M., Stephens, G., Waugh, A., Beare, D., Steingrimsdottir, H. and Bridges, B. (1990) HPRT somatic mutation data. *Prog Clin Biol Res*, **340C**, 25-35.

79. Swift, M. and Chase, C. (1979) Cancer in families with xeroderma pigmentosum. *J Natl Cancer Inst*, **62**, 1415-1421.

80. Welshimer, J. and Swift, M. (1982) Congenital malformations and developmental disabilities in ataxia telangiectasia, Fanconi anemia, and xeroderma pigmentosum families. *Am J Hum Genet*, **34**, 781-793.

81. Nishigori, C., Moriwaki, S., Takebe, H., Tanaka, T. and Imamura, S. (1994) Gene alterations and clinical characteristics of xeroderma pigmentosum group A patients in Japan. *Arch Dermatol*, **130**, 191-197.

82. Moriwaki, S., Nishigori, C., Teramoto, T., Tanaka, T., Kore-eda, S., Takebe, H. and Imamura, S. (1993) Absence of DNA repair deficiency in the confirmed heterozygotes of xeroderma pigmentosum group A. *J Invest Dermatol*, **101**, 69-72.

83. Rainbow, A. (1980) Reduced capacity to repair irradiated adenovirus in fibroblasts from xeroderma pigmentosum heterozygotes. *Cancer Res*, **40**, 3945-3949.

84. Day 3rd, R. (1974) Studies on repair of adenovirus 2 by human fibroblasts using normal, xeroderma pigmentosum, and xeroderma pigmentosum heterozygous strains. *Cancer Res*, **34**, 1965-1970.

85. Giannelli, F. and Pawsey, S. (1976) DNA repair synthesis in human heterokaryons. III. The rapid and slow complementing varieties of xeroderma pigmentosum. *J Cell Sci*, **20**, 207-213.

86. Cleaver, J.E. (1970) DNA repair and radiation sensitivity in human (xeroderma pigmentosum) cells. *Int J Radiat Biol Relat Stud Phys Chem Med*, **18**, 557-565.

87. Bootsma, D., Mulder, M., Cohen, J. and Pot, F. (1970) Different inherited levels of DNA repair replication in xeroderma pigmentosum cell strains after exposure to ultraviolet irradiation. *Mutat Res*, **9**, 507-516.

88. Regan, J., Setlow, R., Kaback, M., Howell, R., Klein, E. and Burgess, G. (1971) Xeroderma pigmentosum: a rapid sensitive method for prenatal diagnosis. *Science*, **174**, 147-150.

89. Kleijer, W., de Weerd-Kastelein, E., Sluyter, M., Keijzer, W., de Wit, J. and Bootsma, D. (1973) UV-induced DNA repair synthesis in cells of patients with different forms of xeroderma pigmentosum and of heterozygotes. *Mutat Res*, **20**, 417-428.

90. Neumann, A., Sturgis, E. and Wei, Q. (2005) Nucleotide Excision Repair as a Marker for Susceptibility to Tobacco-Related Cancers: A Review of Molecular Epidemiogical Studies. *Mol Carcinog*, **42**, 65-92.

91. Park, J., Park, S., Choi, J., Lee, S., Jeon, H., Cha, S., Kim, C., Park, J., Kam, S., Park, R. *et al.* (2002) Polymorphisms of the DNA repair gene xeroderma pigmentosum group A and risk of primary lung cancer. *Cancer Epidemiol Biomarkers Prev*, **11**, 993-997.

92. Wu, X., Zhao, H., Wei, Q., Amos, C., Zhang, K., Guo, Z., Qiao, Y., Hong, W. and Spitz, M. (2003) XPA polymorphism associated with reduced lung cancer risk and a modulating effect on nucleotide excision repair capacity. *Carcinogenesis*, **24**, 505-509.

93. Sanyal, S., Festa, F., Sakano, S., Zhang, Z., Steineck, G., Norming, U., Wijkstrom, H., Larsson, P., Kumar, R. and Hemminki, K. (2004) Polymorphisms in DNA repair and metabolic genes in bladder cancer. *Carcinogenesis*, **25**, 729-734.

94. Forsti, A., Angelini, S., Festa, F., Sanyal, S., Zhang, Z., Grzybowska, E., Pamula, J., Pekala, W., Zientek, H., Hemminki, K. *et al.* (2004) Single nucleotide polymorphisms in breast cancer. *Oncol Rep*, **11**, 917-922.

95. Zhou, W., Liu, G., Miller, D., Thurston, S., Xu, L., Wain, J., Lynch, T., Su, L. and Christiani, D. (2002) Gene-environment interaction for the ERCC2 polymorphisms and cumulative cigarette smoking exposure in lung cancer. *Cancer Res*, **62**, 1377-1381.

96. Sturgis, E., Zheng, R., Li, L., Castillo, E., Eicher, S., Chen, M., Strom, S., Spitz, M. and Wei, Q. (2000) XPD/ERCC2 polymorphisms and risk of head and neck cancer: a case-control analysis. *Carcinogenesis*, **21**, 2219-2223.

97. Rybicki, B., Conti, D., Moreira, A., Cicek, M., Casey, G. and Witte, J. (2004) DNA repair gene XRCC1 and XPD polymorphisms and risk of prostate cancer. *Cancer Epidemiol Biomarkers Prev*, **13**, 23-29.

98. Xing, D., Tan, W., Wei, Q. and Lin, D. (2002) Polymorphisms of the DNA repair gene XPD and risk of lung cancer in a Chinese population. *Lung Cancer*, **38**, 123-129.

99. Butkiewicz, D., Rusin, M., Enewold, L., Shields, P., Chorazy, M. and Harris, C. (2001) Genetic polymorphisms in DNA repair genes and risk of lung cancer. *Carcinogenesis*, **22**, 593-597.

100. Smith, T., Levine, E., Perrier, N., Miller, M., Freimanis, R., Lohman, K., Case, L., Xu, J., Mohrenweiser, H. and Hu, J. (2003) DNA-repair genetic polymorphisms and breast cancer risk. *Cancer Epidemiol Biomarkers Prev*, **12**, 1200-1204.

101. Jeon, H., Kim, K., Park, S., Lee, S., Choi, J., Lee, G., Kam, S., Park, R., Kim, I., Kim, C. *et al.* (2003) Relationship between XPG codon 1104 polymorphism and risk of primary lung cancer. *Carcinogenesis*, **24**, 1677-1681.

102. Kumar, R., Hoglund, L., Zhao, C., Forsti, A., Snellman, E. and Hemminki, K. (2003) Single nucleotide polymorphisms in the XPG gene: determination of role in DNA repair and breast cancer risk. *Int J Cancer*, **103**, 671-675.

103. Zhou, W., Liu, G., Park, S., Wang, Z., Wain, J., Lynch, T., Su, L. and Christiani, D. (2005) Gene-smoking interaction associations for the ERCC1 polymorphisms in the risk of lung cancer. *Cancer Epidemiol Biomarkers Prev*, **14**, 491-496.

104. Horibata, K., Iwamoto, Y., Kuraoka, I., Jaspers, N.G., Kurimasa, A., Oshimura, M., Ichihashi, M. and Tanaka, K. (2004) Complete absence of Cockayne syndrome group B gene product gives rise to UV-sensitive syndrome but not Cockayne syndrome. *Proc Natl Acad Sci U S A*, **101**, 15410-15415.

105. Spivak, G. (2004) The many faces of Cockayne syndrome. *Proc Natl Acad Sci U S A*, **101**, 15273-15274.

106. Draptchinskaia, N., Gustavsson, P., Andersson, B., Pettersson, M., Willig, T.N., Dianzani, I., Ball, S., Tchernia, G., Klar, J., Matsson, H. *et al.* (1999) The gene encoding ribosomal protein S19 is mutated in Diamond-Blackfan anaemia. *Nat Genet*, **21**, 169-175.

107. Wade, M. and Chu, E. (1979) Effects of DNA damaging agents on cultured fibroblasts derived from patients with Cockayne syndrome. *Mutat Res*, **59**.

108. Norris, P.G., Limb, G.A., Hamblin, A.S., Lehmann, A.R., Arlett, C.F., Cole, J., Waugh, A.P. and Hawk, J.L. (1990) Immune function, mutant frequency, and cancer risk in the DNA repair defective genodermatoses xeroderma pigmentosum, Cockayne's syndrome, and trichothiodystrophy. *J Invest Dermatol*, **94**, 94-100.

109. Broughton, B., Lehmann, A., Harcourt, S., Arlett, C., Sarasin, A., Kleijer, W., Beemer, F., Nairn, R. and Mitchell, D. (1990) Relationship between pyrimidine dimers, 6-4 photoproducts, repair synthesis and cell survival: studies using cells from patients with trichothiodystrophy. *Mutat Res*, **235**, 33-40.

110. Itin, P., Sarasin, A. and Pittelkow, M. (2001) Trichothiodystrophy: update on the sulfur-deficient brittle hair syndromes. *J Am Acad Dermatol*, **44**, 891-820.

111. Broughton, B., Berneburg, M., Fawcett, H., Taylor, E., Arlett, C., Nardo, T., Stefanini, M., Menefee, E., Price, V., Queille, S. *et al.* (2001) Two individuals with features of both xeroderma pigmentosum and trichothiodystrophy highlight the complexity of the clinical outcomes of mutations in the XPD gene. *Hum Mol Genet*, **10**, 2539-2547.

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

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

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

115. Barnes, D., Tomkinson, A., Lehmann, A., Webster, A. and Lindahl, T. (1992) Mutations in the DNA ligase I gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents. *Cell*, **69**, 495-503.

116. *Mouse Mutation Database version 5 http://pathcuric1/swmed.edu/research/research Accessed August 2006.*

117. Hanawalt, P. (2001) Revisiting the Rodent Repairadox. *Environ Mol Mutagen*, **38**, 89-96.

118. van Zeeland, A., Smith, C. and Hanawalt, P. (1981) Sensitive determination of pyrimidine dimers in DNA of UV irradiated mammalian cells: introduction of T4 endonuclease V into frozen and thawed cells. *Mutat Res*, **132**, 129-138.

119. Mellon, I., Spivak, G. and Hanawalt, P. (1987) Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell*, **51**, 241-249.

120. Tang, J., Huang, B., Ford, J., Hanawalt, P. and Chu, G. (2000) Xeroderma pigmentosum p48 gene enhances global genomic repair and supresses UV-induced mutagenesis. *Mol Cell*, **5**, 737-744.

121. Surveillance, Epidemiology, and End Results (SEER) Program. *SEER*Stat Database: Incidence (1973-2003) November 2005 Submission Released April 2006*. National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch.

122. American Cancer Society. *Breast Cancer Facts and Figures 2005*. Atlanta.

123. Vogel, V. and Bevers, T. (2003) *Handbook of Brast Cancer Risk-Assessment*. Jones and Bartlett Publishers, Sudsbury, MA.

124. Dawson, D. and Thompson, G. (1989) Breast Cancer Risk Factors and Screening: United States, 1987. *National Center for Health Statistics. Vital Health Stat 10*, **172**.

125. Lippman, M. and Swain, S. (1992) *Endocrine-responsive cancers in humans*. WB Saunders Company, Philadelphia.

126. Wolfe, J. (1976) Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer*, **37**, 2486-2492.

127. Wolfe, J. (1976) Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer*, **37**, 2486-2492.

128. Gierach, G. and Vogel, V. (2004) In Singletary, S. E., Robb, G. L. and Hortobagyi, G. N. (eds.), *Advanced Therapy of Breast Disease*. B.C. Decker, Inc., Hamilton, Ontario.

129. Bain, C., Speizer, F., Rosner, B., Belanger, C. and Hennekens, C. (1980) Family history of breast cancer as a risk indicator for the disease. *Am J Epidemiol*, **111**, 301-308.

130. Cocquyt, V. and Van Belle, S. (2005) Lobular carcinoma in situ and invasive lobular cancer of the breast. *Curr Opin Obstet Gynecol*, **17**, 55-60.

131. (2002) *American Joint Committee on Cancer. (2002) AJCC Staging Manual, 6th Ed. New York: Springer-Verlag.*

132. Bieche, I. and Lidereau, R. (1995) Genetic alterations in breast cancer. *Genes Chromosomes Cancer*, **14**, 227-251.

133. Bloom, H.G.J. and Richardson, W.W. (1957) Histologic grading and prognosis in breast cancer: a study of 1409 cases of which 359 have been followerf for 15 years. *Br J Cancer*, **11**, 359-377.

134. Robbins, P., Pinder, S., de Klerk, N., Dawkins, H., Harvey, J., Sterrett, G., Ellis, I. and Elston, C. (1995) Histological grading of breast carcinomas: a study of interobserver agreement. *Hum Pathol*, **26**, 873-879.

135. Driscoll, M., Sathya, G., Muyan, M., Klinge, C., Hilf, R. and Bambara, R. (1998) Sequence Requirements for Estrogen Receptor Binding to Estrogen Response Elements. *J Biol Chem*, **273**, 29321-29330.

136. Kardinal, C. In Donegan, W. and Spratt, J. (eds.), *Cancer of the Breast, Fourth edition*. W.B. Saunders, Philadelphia, pp. 534-580.

137. Rosen, P. (1997) *Rosen's Breast Pathology*. Lippincott-Raven, Philadelphia.

138. Henderson, B.E., Ross, R.K., Judd, H.L., Krailo, M.D. and Pike, M.C. (1985) Do regular ovulatory cycles increase breast cancer risk? *Cancer*, **56**, 1206-1208.

139. Brugge, D., de Lemos, J. and Oldmixon, B. (2005) Exposure pathways and health effects associated with chemical and radiological toxicity of natural uranium: a review. *Rev Environ Health*, **20**, 177-193.

140. Latimer, J.J., Nazir, T., Flowers, L.C., Forlenza, M.J., Beaudry-Rodgers, K., Kelly, C.M., Conte, J.A., Shestak, K., Kanbour-Shakir, A. and Grant, S.G. (2003) Unique tissue-specific level of DNA nucleotide excision repair in primary human mammary epithelial cultures. *Exp Cell Res*, **291**, 111-121.

141. Li, C.I., Daling, J.R. and Malone, K.E. (2003) Incidence of invasive breast cancer by hormone receptor status from 1992 to 1998. *J Clin Oncol*, **21**, 28-34.

142. Hudelist, G., Singer, C., Manavi, M., Pischinger, K., Kubista, E. and Czerwenka, K. (2003) Co-expression of ErbB-family members in human breast cancer: Her-2/neu is the preferred dimerization candidate in nodal-positive tumors. *Breast Cancer Res Treat*, **80**, 353-361.

143. Zhou, B. and Hung, M. (2003) Dysregulation of cellular signaling by HER2/neu in breast cancer. *Semin Oncol*, **30**, 38-48.

144. Baldwin, A. (2001) Control of oncogenesis and cancer therapy resistance by the transcription factor NF-kappaB. *J Clin Invest*, **107**, 241-246.

145. Paik, S., Hazan, R., Fisher, E., Sass, R., Fisher, B., Redmond, C., Schlessinger, J., Lippman, M. and King, C. (1990) Pathologic findings from the National Surgical Adjuvant Breast and Bowel Project: prognostic significance of erbB-2 protein overexpression in primary breast cancer. *J Clin Oncol*, **8**, 103-112.

146. Nathanson, S., Slater, R., Debruyn, D., Kapke, A. and Linden, M. (2006) Her-2/neu expression in primary breast cancer with sentinel lymph node metastasis. *Ann Surg Oncol*, **2**, 205-213.

147. Sorlie, T., Perou, C., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M., van de Rijn, M., Jeffrey, S. *et al.* (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*, **98**, 10869-10874.

148. Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S. *et al.* (2003) Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A*, **100**, 8418-8423.

149. Hu, Z., Fan, C., Oh, D., Marron, J., He, X., Qaqish, B., Livasy, C., Carey, L., Reynolds, E., Dressler, L. *et al.* (2006) The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics*, **7**.

150. Field, T., Buist, D., Doubeni, C., Enger, S., Fouayzi, H., Hart, G., Korner, E., Lamerato, L., Bachman, D., Ellis, J. *et al.* (2005) Disparities and survival among breast cancer patients. *J Natl Cancer Inst Monographs*, **35**, 88-95.

151. Stark, A., Claud, S., Kapke, A., Lu, M., Linden, M. and Griggs, J. (2005) Race modifies the association between breast carcinoma pathologic prognostic indicators and the positive status for HER-2/neu. *Cancer*, **104**, 2189-2196.

152. Carey, L., Perou, C., Livasy, C., Dressler, L., Cowan, D., Conway, K., Karaca, G., Troester, M., Tse, C., Edmiston, S. *et al.* (2006) Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA*, **295**, 2492-2502.

153. Martin, M., Meyricke, R., O'Neill, T. and Roberts, S. (2006) Mastectomy or breast conserving surgery? Factors affecting type of surgical treatment for breast cancer--a classification tree approach. *BMC Cancer*, **6**.

154. Herbert, S. (1998) In Coia, L. and Moylan, D. (eds.), *Introduction to Clinical Radiation Oncology, Third edition*. Medical Physics Publishing, Madison, pp. 153-176.

155. Khan, A., Kirk, M., Mehta, P., Seif, N., Griem, K., Bernard, D., Chu, J. and Dickler, A. (2006) A dosimetric comparison of three-dimensional conformal, intensity-modulated radiation therapy, and MammoSite partial-breast irradiation. *Brachytherapy*, **5**, 183-188.

156. Trudeau, M., Charbonneau, F., Gelmon, K., Laing, K., Latreille, J., Mackey, J., McLeod, D., Pritchard, K., Provencher, L. and Verma, S. (2005) Selection of adjuvant chemotherapy for treatment of node-positive breast cancer. *Lancet Oncol*, **6**, 886-889.

157. van Rijk, M., Peterse, J., Nieweg, O., Oldenburg, H., Rutgers, E. and Kroon, B. (2006) Additional axillary metastases and stage migration in breast cancer patients with micrometastases or submicrometastases in sentinel lymph nodes. *Cancer*, **107**, 467-471.

158. Ring, A.E. and Ellis, P.A. (2005) Taxanes in the treatment of early breast cancer. *Cancer Treat Rev*, **8**, 618-627.

159. Baquiran, D. (2001) *Lippincott's Cancer Chemotherapy Handbook, Second Edition*. Lippincott Williams and Wilkins, Philadelphia.

160. Garces, C. and Cance, W. (2004) Neoadjuvant chemotherapy of breast cancer. *Am Surg*, **70**, 565-569.

161. Vogel, V., Costantino, J., Wickerham, D., Cronin, W., Cecchini, R., Atkins, J., Bevers, T., Fehrenbacher, L., Pajon ER, J., Wade 3rd, J. *et al.* (2006) Effects of tamoxifen vs raloxifene on the risk of developing invasive breast cancer and other disease outcomes: the NSABP Study of Tamoxifen and Raloxifene (STAR) P-2 trial. *JAMA*, **295**, 2727-2741.

162. Francucci, C., Romagni, P. and Boscaro, M. (2005) Raloxifene: bone and cardiovascular effects. *Endocrinol Invest*, **28**, 85-89.

163. Howell, A., Cuzick, J., Baum, M., Buzdar, A., Dowsett, M., Forbes, J., Hoctin-Boes, G., Houghton, J., Locker, G., Tobias, J. *et al.* (2005) Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. *Lancet*, **365**, 60-62.

164. Rio, P., Maurizis, J., Peffault de Latour, M., Bignon, Y. and Bernard, -.G.D. (1999) Quantification of BRCA1 protein in sporadic breast carcinoma with or without loss of heterozygosity of the BRCA1 gene. *Int J Cancer*, **80**, 823-826.

165. Taylor, J., Lymboura, M., Pace, P., A'hern, R., Desai, A., Shousha, S., Coombes, R. and Ali, S. (1998) An important role for BRCA1 in breast cancer progression is indicated by its loss in a large proportion of non-familial breast cancers. *Int J Cancer*, **79**, 334-342.

166. Lynch, E., Ostermeyer, E., Lee, M., Arena, J., Ji, H., Dann, J., Swisshelm, K., Suchard, D., MacLeod, P., Kvinnsland, S. *et al.* (1997) Inherited mutations in PTEN that are associated with breast cancer, cowden disease, and juvenile polyposis. *Am J Hum Genet*, **61**, 1254-1260.

167. Swift, M., Reitnauer, P., Morrell, D. and Chase, C. (1987) Breast and other cancers in families with Ataxia-Telangiectasia. *N Engl J Med*, **316**, 1289-1294.

168. Inskip, H., Kinlen, L., Taylor, A., Woods, C. and Arlett, C. (1999) Risk of breast cancer and other cancers in heterozygotes for ataxia telangiectasia. *Br J Cancer*, **79**, 1304-1307.

169. Thompson, D., Duedal, S., Kirner, J., McGuffog, L., Last, J., Reiman, A., Byrd, P., Taylor, M. and Easton, D. (2005) Cancer risks and mortality in heterozygous ATM mutation carriers. *J Natl Cancer Inst*, **97**, 813-822.

170. Osborne, A., Elledge, S. and Zou, L. (2002) Checking on the fork: the DNA-replication stress-response pathway. *Trends Cell Biol*, **12**, 509-516.

171. Boardman, L., Thibodeu, S. and Schaid, D. (1998) Increased risk for cancer in Peutz-Jeghers syndrome. *Ann Intern Med*, **128**, 896-899.

172. Jenne, D., Reimann, H., Nezu, J., Friedel, W., Foff, S. and Muller, O. (1997) Peutz-Jeghers syndrome is caused by mutations in a novel serine theronine kinase. *Nat Genet*, **18**, 38-43.

173. Malkin, D., Li, F., Strong, L., Fraumeni, J.J., Nelson, C., Kim, D., Kassel, J., Gryka, M., Bischoff, F. and Tainsky, M. (1990) Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science*, **250**, 1233-1238.

174. Sidransky, D., Tokino, T., Helzlsouer, K., Zehnbauer, B., Rausch, G., Shelton, B., Prestigiacomo, L., Vogelstein, B. and Davidson, N. (1992) Inherited p53 gene mutations in breast cancer. *Cancer Res*, **52**, 2984-2986.

175. Ford, J. and Hanawalt, P. (1995) Li-Fraumeni syndrome fibroblasts homozygous for p53 mutations are deficient in global DNA repair but exhibit normal transcription-coupled repair and enhanced UV resistance. *Proc Natl Acad Sci U S A*, **92**, 8876-8880.

176. Ford, J. and Hanawalt, P. (1997) Expression of wild-type p53 is required for efficient global genomic nucleotide excision repair in UV-irradiated human fibroblasts. *J Biol Chem*, **272**, 58073-58080.

177. Wani, M., Zhu, Q., El-Mahdy, M., Venkatachalam, S. and Wani, A. (2000) Enhanced sensitivity to anti-benzo(a)pyrene-diol-epoxide DNA damage correlates with decreased global genomic repair attributable to abrogated p53 function in human cells. *Cancer Res*, **60**, 2273-2280.

178. Hwang, B., Ford, J., Hanawalt, P. and Chu, G. (1999) Expression of the p48 xeroderma pigmentosum gene is p53-dependent and is involved in global genomic repair. *Proc Natl Acad Sci U S A*, **96**, 424-428.

179. Adimoolam, S. and Ford, J. (2002) p53 and DNA damage-inducible expression of the xeroderma pigmentosum group C gene. *Proc Natl Acad Sci U S A*, **99**, 12985-12990.

180. Wang, X., Yeh, H., Schaeffer, L., Roy, R., Moncollin, V., Egly, J., Wang, Z., Freidberg, E., Evans, M. and Taffe, B. (1995) p53 modulation of TFIIH-associated nucleotide excision repair activity. *Nat Genet*, **10**, 188-195.

181. Yoshida, K. and Miki, Y. (2004) Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. *Cancer Sci*, **95**, 866-871.

182. Gail, M., Brinton, L., Byar, D., Corle, D., Green, S., Schairer, C. and Mulvihill, J. (1989) Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst*, **81**, 1879-1886.

183. Ford, D., Easton, D., Stratton, M., Narod, S., Goldgar, D., Devilee, P., Bishop, D., Weber, B., Lenoir, G., Chang-Claude, J. *et al.* (1998) Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet*, **62**, 676-689.

184. Rubinstein, W.S. (2004) Hereditary breast cancer in Jews. *Familial Cancer*, **3**, 249-257.

185. Nanda, R., Schumm, L., Cummings, S., Fackenthal, J., Sveen, L., Ademuyiwa, F., Cobleigh, M., Esserman, L., Lindor, N., Neuhausen, S. *et al.* (2005) Genetic testing in an ethnically diverse cohort of high-risk women: a comparative analysis of BRCA1 and BRCA2 mutations in American families of European and African ancestry. *JAMA*, **294**, 1925-1933.

186. (1997) Breast Cancer Linkage Consortium. Pathology of familial breast cancer: differences between breast cancers in carriers of BRCA1 or BRCA2 mutations and sporadic cases. *Lancet*, **349**, 1505-1510.

187. Wang, Y., Cortez, D., Yazdi, P., Neff, N., Elledge, S. and Qin, J. (2000) BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes Dev*, **14**, 927-939.

188. Le Page, F., Randrianarison, V., Marot, D., Cabannes, J., Perricaudet, M., Feunteun, J. and Sarasin, A. (2000) BRCA1 and BRCA2 are necessary for the transcription-coupled repair of the oxidative 8-oxoguanine lesion in human cells. *Cancer Res*, **60**, 5548-5552.

189. Yu, X., Fu, S., Lai, M., Baer, R. and Chen, J. (2006) BRCA1 ubiquitinates its phosphorylation-dependent binding partner CtIP. *Genes Dev*, **20**, 1721-1726.

190. Kutler, D., Singh, B., Satagopan, J., Batish, S., Berwick, M., Giampietro, P., Hanenberg, H. and Auerbach, A. (2003) A 20-year perspective on the International Fanconi Anemia Registry (IFAR). *Blood*, **101**, 1249-1256.

191. Latimer, J.J., Rubinstein, W.S., Johnson, J.M., Kanbour-Shakir, A., Vogel, V.G. and Grant, S.G. (2005) Haploinsufficiency for BRCA1 is associated with normal levels of DNA nucleotide excision repair in breast tissue and blood lymphocytes. *BMC Med Genet*, **6**, 26.

192. Lewin, B. (1997), *Genes VI*. Oxford University Press, New York, pp. 847-884.

193. Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M. and Werner, T. (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics*, **21**, 2933-2942.

194. *Genomatix Software GmbH. 2006. Genomatix Portal Family Information.* *http://www.genomatix.de*.

195. (2006) National Center for Biotechnology Information. Entrez Gene. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene.

196. Chiurazzi, P. and Neri, G. (2003) Reactivation of silenced genes and transcriptional therapy. *Cytogenet Genome Res*, **100**, 56-64.

197. Bates, G. (2001) Huntington's disease. Exploiting expression. *Nature*, **413**, 693-694.

198. Steffan, J., Bodai, L., Pallos, J., Poelman, M., McCampbell, A., Apostol, B., Kazantsev, A., Schmidt, E., Zhu, Y., Greenwald, M. *et al.* (2001) Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in Drosophila. *Nature*, **413**, 739-743.

199. Spotswood, H. and Turner, B. (2002) An increasingly complex code. *J Clin Invest*, **110**, 577-582.

200. Xodo, L., Alunni-Fabbroni, M. and Manzini, G. (1994) Effect of 5-methylcytosine on the structure and stability of DNA. Formation of triple-stranded concatenamers by overlapping oligonucleotides. *J Biomol Struct Dyn*, **11**, 703-720.

201. Kress, C., Thomassin, H. and Grange, T. (2001) Local DNA demethylation in vertebrates: how could it be performed and targeted? *FEBS Lett*, **494**, 135-140.

202. Kapoor, A., Agius, F. and Zhu, J. (2005) Preventing transcriptional gene silencing by active DNA demethylation. *FEBS Lett*, **579**, 5889-5898.

203. Zhu, B., Zheng, Y., Hess, D., Angliker, H., Schwarz, S., Siegmann, M., Thiry, S. and Jost, J. (2000) 5-methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex. *Proc Natl Acad Sci U S A*, **97**, 5135-5139.

204. Zhu, B., Benjamin, D., Zheng, Y., Angliker, H., Thiry, S., Siegmann, M. and Jost, J. (2001) Overexpression of 5-methylcytosine DNA glycosylase in human embryonic kidney cells EcR293 demethylates the promoter of a hormone-regulated reporter gene. *Proc Natl Acad Sci U S A*, **98**, 5031-5036.

205. Yan, L., Yang, X. and Davidson, N. (2001) Role of DNA methylation and histone acetylation in steroid receptor expression in breast cancer. *J Mammary Gland Biol Neoplasia*, **6**, 183-192.

206. Larsen, F., Gundersen, G., Lopez, R. and Prydz, H. (1992) CpG islands as gene markers in the human genome. *Genomics*, **13**, 1095-1107.

207. Suzuki, Y., Tsunoda, T., Sese, J., Taira, H., Mizushima-Sugano, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y. *et al.* (2001) Identification and characterization of the potential promoter regions of 1031 kinds of human genes. *Genome Res*, **11**, 677-684.

208. Takai, D. and Jones, P.A. (2003) The CpG island searcher: a new WWW resource. *In Silico Biol*, **3**, 235-240.

209. Gardiner-Garden, M. and Frommer, M. (1987) CpG islands in vertebrate genomes. *J Mol Biol*, **196**, 261-282.

210. Takai, D. and Jones, P.A. (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci U S A*, **99**, 3740-3745.

211. Vinogradov, A. (2005) Dualism of gene GC content and the CpG pattern in regard to expression in the human genome: magnitude versus breadth. *Trends Genet*, **21**, 639-643.

212. Zhang, L., Kasif, S., Cantor, C.R. and Broude, N.E. (2004) GC/AT-content spikes as genomic punctuation marks. *Proc Natl Acad Sci U S A*, **101**, 16855-16860.

213. Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N.P. and Bickmore, W.A. (2004) Chromatin Architecture of the Human Genome: Gene-Rich Domains Are Enriched in Open Chromatin Fibers. *Cell*, **11**, 555-566.

214. Szyf, M., Pakneshan, P. and Rabbani, S. (2004) DNA methylation and breast cancer. *Biochem Pharmacol*, **68**, 1187-1197.

215. Jones, P. and Baylin, S. (2002) The fundamental role of epigenetic events in cancer. *Nat Rev Genet*, **3**, 415-428.

216. Esteller, M., Silva, J., Dominguez, G., Bonilla, F., Matias-Guiu, X., Lerma, E., Bussaglia, E., Prat, J., Harkes, I., Repasky, E. *et al.* (2000) Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. *J Natl Cancer Inst*, **92**, 564-569.

217. Nuovo, G., Nakagawa, H., Sotamaa, K. and Chapelle Ade, L. (2006) Hypermethylation of the MLH1 promoter with concomitant absence of transcript and protein occurs in small patches of crypt cells in unaffected mucosa from sporadic colorectal carcinoma. *Diagn Mol Pathol*, **15**, 17-23.

218. Widschwendter, M. and Jones, P. (2002) DNA methylation and breast carcinogenesis. *Oncogene*, **21**, 5462-5482.

219. Huang, T., Laux, D., Hamlin, B., Tran, P., Tran, H. and Lubahn, D. (1997) Identification of DNA methylation markers for human breast carcinomas using the methylation-sensitive restriction fingerprinting technique. *Cancer Res*, **57**, 1030-1034.

220. Huang, T., Perry, M. and Laux, D. (1999) Methylation profiling of CpG islands in human breast cancer cells. *Hum Mol Genet*, **8**, 459-470.

221. Chen, C., Chen, H., Hsiau, T., Hsiau, A., Shi, H., Brock, G., Wei, S., Caldwell, C., Yan, P. and Huang, T. (2003) Methylation target array for rapid analysis of CpG island hypermethylation in multiple tissue genomes. *Am J Pathol*, **163**, 37-45.

222. Soares, J., Pinto, A., Cunha, C., Andre, S., Barao, I., Sousa, J. and Cravo, M. (1999) Global DNA hypomethylation in breast carcinoma: correlation with prognostic factors and tumor progression. *Cancer*, **85**, 112-118.

223. Bernardino, J., Roux, C., Almeida, A., Vogt, N., Gibaud, A., Gerbault-Seureau, M., Magdelenat, H., Bourgeois, C., Malfoy, B. and Dutrillaux, B. (1997) DNA hypomethylation in breast cancer: an independent parameter of tumor progression? *Cancer Genet Cytogenet*, **97**, 83-89.

224. Feinberg, A., Gehrke, C., Kuo, K. and Ehrlich, M. (1988) Reduced genomic 5-methylcytosine content in human colonic neoplasia. *Cancer Res*, **48**, 1159-1161.

225. Esteller, M., Fraga, M., Guo, M., Garcia-Foncillas, J., Hedenfalk, I., Godwin, A., Trojan, J., Vaurs-Barriere, C., Bignon, Y., Ramus, S. *et al.* (2001) DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. *Hum Mol Genet*, **10**, 3001-3007.

226. Sadikovic, B., Haines, T., Butcher, D. and Rodenhiser, D. (2004) Chemically induced DNA hypomethylation in breast carcinoma cells detected by the amplification of intermethylated sites. *Breast Cancer Res*, **6**, 329-337.

227. Chen, R., Pettersson, U., Beard, C., Jackson-Grusby, L. and Jaenisch, R. (1998) DNA hypomethylation leads to elevated mutation rates. *Nature*, **395**, 89-93.

228. Nass, S., Ferguson, A., El-Ashry, D., Nelson, W. and Davidson, N. (1999) Expression of DNA methyl-transferase (DMT) and the cell cycle in human breast cancer cells. *Oncogene*, **18**, 7453-7461.

229. Sano, H., Shiomi, N., Imanishi, K., Maie, O. and Shiomi, T. (1989) DNA methylation in xeroderma pigmentosum. *Mutat Res*, **217**, 141-151.

230. Suzuki, H., Gabrielson, E., Chen, W., Anbazhagan, R., van Engeland, M., Weijenberg, M., Herman, J. and Baylin, S. (2002) A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. *Nat Genet*, **31**, 141-149.

231. Glaser, K., Staver, M., Waring, J., Stender, J., Ulrich, R. and Davidsen, S. (2003) Gene expression profiling of multiple histone deacetylase (HDAC) inhibitors: defining a common gene set produced by HDAC inhibition in T24 and MDA carcinoma cell lines. *Mol Cancer Ther*, **2**, 151-163.

232. Yoshida, M., Horinouchi, S. and Beppu, T. (1995) Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. *Bioessays*, **17**, 423-430.

233. Monneret, C. (2005) Histone deacetylase inhibitors. *Eur J Med Chem*, **40**, 1-13.

234. Butler, L., Agus, D., Scher, H., Higgins, B., Rose, A., Cordon-Cardo, C., Thaler, H., Rifkind, R., Marks, P. and Richon, V. (2000) Suberoylanilide hydroxamic acid, an inhibitor of histone deacetylase, suppresses the growth of prostate cancer cells in vitro and in vivo. *Cancer Res*, **60**, 5165-5170.

235. Takai, N., Desmond, J., Kumagai, T., Gui, D., Said, J., Whittaker, S., Miyakawa, I. and Koeffler, H. (2004) Histone deacetylase inhibitors have a profound antigrowth activity in endometrial cancer cells. *Clinical Cancer Res*, **10**, 1141-1149.

236. Gottlicher, M., Minucci, S., Zhu, P., Kramer, O., Schimpf, A., Giavara, S., Sleeman, J., Lo Coco, F., Nervi, C., Pelicci, P. *et al.* (2001) Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells. *EMBO J*, **20**, 6969-6978.

237. Louis, M., Rosato, R., Brault, L., Osbild, S., Battaglia, E., Yang, X., Grant, S. and Bagrel, D. (2004) The histone deacetylase inhibitor sodium butyrate induces breast cancer cell apoptosis through diverse cytotoxic actions including glutathione depletion and oxidative stress. *Int J Oncol*, **25**, 1701-1711.

238. Chavez-Blanco, A., Perez-Plasencia, C., Perez-Cardenas, E., Carrasco-Legleu, C., Rangel-Lopez, E., Segura-Pacheco, B., Taja-Chayeb, L., Trejo-Becerril, C., Gonzalez-Fierro, A., Candelaria, M. *et al.* (2006) Antineoplastic effects of the DNA methylation inhibitor hydralazine and the histone deacetylase inhibitor valproic acid in cancer cell lines. *Cancer Cell Int*, **6**.

239. Ghoshal, K., Datta, J., Majumder, S., Bai, S., Kutay, H., Motiwala, T. and Jacob, S. (2005) 5-Aza-deoxycytidine induces selective degradation of DNA methyltransferase 1 by a proteasomal pathway that requires the KEN box, bromo-adjacent homology domain, and nuclear localization signal. *Mol Cell Biol*, **25**, 4727-4741.

240. Ferguson, A., Lapidus, R., Baylin, S. and Davidson, N. (1995) Demethylation of the estrogen receptor gene in estrogen receptor-negative breast cancer cells can reactivate estrogen receptor gene expression. *Cancer Res*, **55**, 2279-2283.

241. Yang, X., Ferguson, A., Nass, S., Phillips, D., Butash, K., Wang, S., Herman, J. and Davidson, N. (2000) Transcriptional activation of estrogen receptor alpha in human breast cancer cells by histone deacetylase inhibition. *Cancer Res*, **60**, 6890-6894.

242. Yang, X., Phillips, D., Ferguson, A., Nelson, W., Herman, J. and Davidson, N. (2001) Synergistic activation of functional estrogen receptor (ER)-alpha by DNA methyltransferase and histone deacetylase inhibition in human ER-alpha-negative breast cancer cells. *Cancer Res*, **61**, 7025-7029.

243. Marquez, V., Kelley, J., Agbaria, R., Ben-Kasus, T., Cheng, J., Yoo, C. and Jones, P. (2005) Zebularine: a unique molecule for an epigenetically based strategy in cancer chemotherapy. *Ann N Y Acad Sci*, **1058**, 246-254.

244. Chuang, J., Yoo, C., Kwan, J., Li, T., Liang, G., Yang, A. and Jones, P. (2005) Comparison of biological effects of non-nucleoside DNA methylation inhibitors versus 5-aza-2'-deoxycytidine. *Mol Cancer Ther*, **4**, 1515-1520.

245. Deng, C., Lu, Q., Zhang, Z., Rao, T., Attwood, J., Yung, R. and Richardson, B. (2003) Hydralazine may induce autoimmunity by inhibiting extracellular signal-regulated kinase pathway signaling. *Arthritis Rheum*, **48**, 746-756.

246. Lin, X., Asgari, K., Putzi, M., Gage, W., Yu, X., Cornblatt, B., Kumar, A., Piantadosi, S., DeWeese, T., De Marzo, A. *et al.* (2001) Reversal of GSTP1 CpG island hypermethylation and reactivation of pi-class glutathione S-transferase (GSTP1) expression in human prostate cancer cells by treatment with procainamide. *Cancer Res*, **61**, 8611-8616.

247. Latimer, J.J., Hultner, M.L., Cleaver, J.E. and Pedersen, R.A. (1996) Elevated DNA excision repair capacity in the extraembryonic mesoderm of the midgestation mouse embryo. *Exp Cell Res*, **228**, 19-28.

248. Crallan, R., Lord, P., Rees, R. and Southgate, J. (2002) Inter-individual variation in urothelial DNA repair gene expression in vitro. *Toxicol In Vitro*, **16**, 383-387.

249. Li, Q., Gardner, K., Zhang, L., Tsang, B., Bostick-Bruton, F. and Reed, E. (1998) Cisplatin induction of ERCC-1 mRNA expression in A2780/CP70 human ovarian cancer cells. *J Biol Chem*, **273**, 23419-23425.

250. Codegoni, A., Broggini, M., Pitelli, M., Pantarotto, M., Torri, V., Mangioni, C. and D'Incalci, M. (1997) Expression of genes of potential importance in the response to chemotherapy and DNA repair in patients with ovarian cancer. *Gynecol Oncol*, **65**, 130-137.

251. Zeng-Rong, N., Paterson, J., Alpert, L., Tsao, M., Viallet, J. and Alaoui-Jamali, M. (1995) Elevated DNA repair capacity is associated with intrinsic resistance of lung cancer to chemotherapy. *Cancer Res*, **55**, 4760-4764.

252. Dabholkar, M., Bostick-Bruton, F., Weber, C., Egwuagu, C., Bohr, V. and Reed, E. (1993) Expression of excision repair genes in non-malignant bone marrow from cancer patients. *Mutat Res*, **293**, 151-160.

253. Dabholkar, M.D., Berger, M.S., Vionnet, J.A., Overton, L., Thompson, C., Bostick-Bruton, F., Yu, J.J., Silber, J.R. and Reed, E. (1996) Comparative analyses of relative ERCC3 and ERCC6 mRNA levels in gliomas and adjacent non-neoplastic brain. *Mol Carcinog*, **17**, 1-7.

254. Wei, Q., Xu, X., Cheng, L., Legerski, R. and Ali-Osman, F. (1995) Simultaneous amplification of four DNA repair genes and beta-actin in human lymphocytes by multiplex reverse transcriptase-PCR. *Cancer Res*, **55**, 5025-5029.

255. Cheng, L., Guan, Y., Li, L., Legerski, R., Einspahr, J., Bangert, J., Alberts, D. and Wei, Q. (1999) Expression in normal human tissues of five nucleotide excision repair genes measured simultaneously by multiplex reverse transcription-polymerase chain reaction. *Cancer Epidemiol Biomarkers Prev*, **8**, 801-807.

256. Cheng, L., Spitz, M., Hong, W. and Wei, Q. (2000) Reduced expression levels of nucleotide excision repair genes in lung cancer: a case-control analysis. *Carcinogenesis*, **21**, 1527-1560.

257. Merkel, P., Khoury, N., Bertolotto, C. and Perfetti, R. (2003) Insulin and glucose regulate the expression of the DNA repair enzyme XPD. *Mol Cell Endocrinol*, **201**, 75-85.

258. Peng, B., Hodge, D., Thomas, S., Cherry, J., Munroe, D., Pompeia, C., Xiao, W. and Farrar, W. (2005) Epigenetic silencing of the human nucleotide excision repair gene, hHR23B, in interleukin-6-responsive multiple myeloma KAS-6/1 cells. *J Biol Chem*, **280**, 4182-4187.

259. Tobon, H. and Salazar, H. (1974) Ultrastructure of the human mammary gland. I. Development of the fetal gland throughout gestation. *J Clin Endocrinol Metab*, **40**, 443-456.

260. Tobon, H. and Salazar, H. (1974) Ultrastructure of the human mammary gland. II. Postpartum lactogenesis. *J Clin Endocrinol Metab*, **40**, 834-844.

261. Salazar, H., Tobon, H. and Josimovich, J. (1975) Developmental, gestational and postgestational modifications of the human breast. *Clin Obstet Gynecol*, **18**, 113-137.

262. Russo, I.H. and Russo, J. (1996) Mammary gland neoplasia in long-term rodent studies. *Environ Health Perspect*, **104**, 938-967.

263. Dabelow, A. (1957) In Barmann, W. (ed.), *Handbuch der Mikroskopischen Anatomic des Munschen, Vol. 3, Part 3 Haut and Sinnes Organs*. Springer-Verlag, Berlin, pp. 277-485.

264. Vorherr, H. (1974) In Vorherr, H. (ed.), *The Breast: Morphology, Physiology and Lactation*. Academic Press, New York, pp. 1-18.

265. Easty, G., Easty, D., Monaghan, P., Ormerod, M. and Neville, A. (1980) Preparation and identification of human breast epithelial cells in culture. *Int J Cancer*, **26**, 577-584.

266. Janss, D., Hillman, E., Malan-Shibley, L. and Ben, T. (1980) In Harris, C., Trump, B. and Stoner, G. (eds.), *Methods in Cell Biol.*. Academic Press, New York, pp. 107-134.

267. Stampfer, M., Hallowes, R. and Hackett, A. (1980) Growth of normal human mammary cells in culture. *In Vitro*, **16**, 412-425.

268. Yang, J., Balakrishnan, A., Hamamote, S., Elias, J., Rosenau, W., Beattie, C., Gupta, T., Welling, S. and Nandi, S. (1987) Human breast epithelial cells in serum-free collagen gel primary culture: growth, morphological, and immunocytochemical analysis. *J Cell Physiol*, **133**, 228-234.

269. Mork, C., van Deurs, B. and Petersen, O. (1990) Regulation of vimentin expression in cultured human mammary epithelial cells. *Differentiation*, **43**, 146-156.

270. Taylor-Papadimitrou, J. and Stampfer, M. (1992) In Freshney, R. (ed.), *Culture of Epithelial Cells*. Wiley-Liss, New York, pp. 107-133.

271. Ethier, S., Mahacek, M., Gullick, W., Frank, T. and Weber, B. (1993) Differential isolation of normal luminal mammary epithelial cells and breast cancer cells from primary and metastatic sites using selective media. *Cancer Res*, **53**, 627-635.

272. Gudjonsson, T., Villadsen, R., Nielsen, H.L., Ronnov-Jessen, L., Bissell, M.J. and Petersen, O.W. (2002) Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties. *Genes Dev*, **16**, 693-706.

273. Dontu, G., Abdallah, W.M., Foley, J.M., Jackson, K.W., Clarke, M.F., Kawamura, M.J. and Wicha, M.S. (2003) In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev*, **17**, 1253-1270.

274. Barcellos-Hoff, M., Aggeler, J., Ram, T. and Bissell, M. (1989) Functional differentiation and alveolar morphogenesis of primary mammary cultures on reconstituted basement membrane. *Development*, **105**, 223-235.

275. Streuli, C., Bailey, N. and Bissell, M. (1991) Control of mammary epithelial differentiation: basement membrane induces tissue-specific gene expression in the absence of cell-cell interaction and morphological polarity. *J Cell Biol*, **115**, 1383-1395.

276. Lin, C. and Bissell, M. (1993) Multifaceted regulation of cell differentiation of extracellular matrix. *FASEB J*, **7**, 737-743.

277. Niemann, C., Brinkman, V., Spitzer, E., Hartmann, G., Sachs, M., Naundorf, H. and Birchmeier, W. (1998) Reconstitution of mammary gland development in vitro: requirement of c-met and c-erbB2 signaling for branching and alveolar morphogenesis. *J Cell Biol*, **143**, 533-545.

278. Wang, C., Goulet, F., Tremblay, N., Germain, L., Auger, F. and Tetu, B. (2001) Selective culture of epithelial cells from primary breast carcinomas using irradiated 3T3 cells as a feeder layer. *Pathology*, **197**, 175-181.

279. Matouskova, E., Dudorkinova, D., Krasna, L. and Vesely, P. (2000) Temporal in vitro expansion of the luminal lineage of human mammary epithelial cells achieved with the 3T3 feeder layer technique. *Breast Cancer Res Treat*, **60**, 241-249.

280. Robertson, E. (1987) In Robertson, E. (ed.), *Teratocarcinomas and embryonic stem cells a practical approach*. IRL Press Limited, Oxford, pp. 71-112.

281. Blatchford, D., K.A.K. Hendry, K. and Wilde, C. (1998) Autocrine regulation of protein secretion in mouse mammary epithelial cells. *Biochem Biophys Res Commun*, **248**.

282. Love, S. (1995), *Dr. Susan Love's Breast Book, Second Edition*. Addison Wesley Publishing Company, Reading, pp. 223.

283. Latimer, J.J. (2000) United States patent no. 6074874.

284. Blatchford, D., Quarrie, L., Tonner, E., McCarthy, C., Flint, D. and Wilde, C. (1999) Influence of microenvironment on mammary epithelial cell survival in primary culture. *J Cell Physiol*, **181**, 304-311.

285. German, T. and Barash, I. (2002) Characterization of an epithelial cell line from bovine mammary gland. *In Vitro Cell Dev Biol Anim*, **38**, 282-292.

286. Hurley, W., Blatchford, D., Hendry, K. and Wilde, C. (1994) Extracellular matrix and mouse mammary cell function: comparison of substrata in culture. *In Vitro Cell Dev Biol Anim*, **30A**, 529-538.

287. Sharp, J., Cane, K., Mailer, S., Oosthuizen, W., Arnould, J. and Nicholas, K. (2006) Species-specific cell-matrix interactions are essential for differentiation of alveoli like structures and milk gene expression in primary mammary cells of the Cape fur seal (Arctocephalus pusillus pusillus). *Matrix Biol*.

288. Kim, D., Oberley, T. and Clifton, K. (1993) Primary culture of flow cytometry-sorted rat mammary epithelial cell (RMEC) subpopulations in a reconstituted basement membrane, matrigel. *Exp Cell Res*, **209**, 6-20.

289. Monaco, M., Bronzert, D., Tormey, D., Waalkes, P. and Lippman, M. (1977) Casein production by human breast cancer. *Cancer Res*, **37**, 749-754.

290. Bartek, J., Bartkova, J., Kyprianou, N., Lalani, E., Staskova, Z., Shearer, M., Chang, S. and Taylor-Papadimitrou, J. (1991) Efficient immortalization of luminal epithelial cells from human mammary gland by introduction of simian virus 40 large tumor antigen with a recombinant retrovirus. *Proc Natl Acad Sci U S A*, **88**, 3520-3524.

291. Shay, J., Van Der Haegen, B., Ying, Y. and Wright, W. (1993) The frequency of immortalization of human fibroblasts and mammary epithelial cells transfected with SV40 large T antigen. *Exp Cell Res*, **209**.

292. Sun, Q., Tang, S., Pater, M. and Pater, A. (1997) Different HPV16 E6/E7 oncogene expression patterns in epithelia reconstructed from HPV16-immortalized human endocervical cells and genital keratinocytes. *Oncogene*, **15**, 2399-2408.

293. Bodnar, A.G., Ouellette, M., Frolkis, M., Holt, S.E., Chiu, C.P., Morin, G.B., Harley, C.B., Shay, J.W., Lichtsteiner, S. and Wright, W.E. (1998) Extension of life-span by introduction of telomerase into normal human cells. *Science*, **279**, 349-352.

294. Stampfer, M.R. and Bartley, J.C. (1985) Induction of transformation and continuous cell lines from normal human mammary epithelial cells after exposure to benzo[a]pyrene. *Proc Natl Acad Sci U S A*, **82**, 2394-2398.

295. Ford, J., Baron, E. and Hanawalt, P. (1998) Human fibroblasts expressing the human papillomavirus E6 gene are deficient in global genomic nucleotide excision repair and sensitive to ultraviolet irradiation. *Cancer Res*, **58**, 599-503.

296. Jian, Y., Schmidt-Grimminger, D., Chien, W., Wu, X., Broker, T. and Chow, L. (1998) Post-transcriptional induction of p21cip1 protein by human papillomavirus E7 inhibits unscheduled DNA synthesis reactivated in differentiated keratinocytes. *Oncogene*, **17**, 2027-2038.

297. Tait, L., Soule, H. and Russo, J. (1990) Ultrastructural and immunocytochemical characterization of an immortalized human breast epithelial cell line, MCF-10. *Cancer Res*, **50**, 6087-6094.

298. Lasfargues, E.Y. and Ozzello, L. (1958) Cultivation of human breast carcinomas. *J Natl Cancer Inst*, **21**, 1131-1147.

299. Ethier, S. (1996) Human Breast Cancer Cell Lines as Models of Growth Regulation and Disease Progression. *J Mammary Gland Biol Neoplasia*, **1**, 111-121.

300. Gazdar, A.F., Kurvari, V., Virmani, A., Gollahon, L., Sakaguchi, M., Westerfield, M., Kodagoda, D., Stasny, V., Cunningham, H.T., Wistuba, II *et al.* (1998) Characterization of paired tumor and non-tumor cell lines established from patients with breast cancer. *Int J Cancer*, **78**, 766-774.

301. Li, M., Aggeler, J., Farson, D., Hatier, C., Hassell, J. and Bissell, M. (1987) Influence of a reconstituted basement membrane and its components on casein gene expression and secretion in mouse mammary epithelial cells. *Proc Natl Acad Sci U S A*, **84**, 136-140.

302. Kirkland, W., Yang, N., Jorgensen, T., Longley, C. and Furmanski, P. (1979) Growth of normal and malignant human mammary epithelial cells in culture. *J Natl Cancer Inst*, **63**, 29-41.

303. Gomm, J., Coope, R., Browne, P. and Coombes, R. (1997) Separated human breast epithelial and myoepithelial cells have different growth factor requirements in vitro but can reconstitute normal breast lobuloalveolar structure. *J Cell Physiol*, **171**, 11-19.

304. Dontu, G., Jackson, K.W., McNicholas, E., Kawamura, M.J., Abdallah, W.M. and Wicha, M.S. (2004) Role of Notch signaling in cell-fate determination of human mammary stem/progenitor cells. *Breast Cancer Res*, **6**, R605-615.

305. Sympson, C., Bissell, M. and Werb, Z. (1995) Mammary gland tumor formation in transgenic mice overexpressing stromelysin-1. *Semin Cancer Biol*, **6**.

306. Sternlicht, M., Lochter, A., Sympson, C., Huey, B., Rougier, J., Gray, J., Pinkel, D., Bissell, M. and Werb, Z. (1999) The stromal proteinase MMP3/stromelysin-1 promotes mammary carcinogenesis. *Cell*, **98**, 137-146.

307. Wolman, S., Smith, H. and Hackett, A. (1985) Growth of diploid cells from breast cancers. *Cancer Genet Cytogenet*, **16**, 49-64.

308. Taylor-Papadimitrou, J., Stampfer, M., Barter, J., Lewis, A., Boshell, M., Lane, E. and Leith, I. (1989) Keratin expression in human mammary cells cultures from normal and malignant tissue: relation to in vivo phenotypes and influence of medium. *J Cell Sci*, **94**, 403-413.

309. Thiery, J. and Sleeman, J. (2006) Complex networks orchestrate epithelial-mesenchymal transitions. *Nat Rev Mol Cell Biol*, **7**, 131-142.

310. Thompson, L. (1998) In Nickoloff, J. and Hoekstra, M. (eds.), *DNA Damage and Repair, Volume 2: DNA Repair in Higher Eukaryotes*. Humana Press, Totowa, pp. 335-393.

311. Mullenders, L.H. and Berneburg, M. (2001) Photoimmunology and nucleotide excision repair: impact of transcription coupled and global genome excision repair. *J Photochem Photobiol B*, **65**, 97-100.

312. Friedberg, E., Walker, G., Siede, W., Wood, R., Schultz, R. and Ellenberger, T. (2006), *DNA Repair and Mutageneis, Second Edition*. ASM Press, Washington, D.C.

313. Killary, A. and Fournier, R. (1984) A genetic analysis of extinction: trans-dominant loci regulate expression of liver-specific traits in hepatoma hybrid cells. *Cell*, **38**, 523-534.

314. Clarke, R., Leonessa, F., Brunner, W. and Thompson, E. (2000) In Harris, J., Lippman, M., M, M. and Osborne, C. (eds.), *Diseases of the Breast*. Lippincott Williams and Wilkins, Phladelphia, Vol. 347-48.

315. Loeb, L.A. (1991) Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res*, **51**, 3075-3079.

316. Ishikawa, T., Ide, F., Qin, X., Zhang, S., Takahashi, Y. and Sekiguchi, M. (2001) Importance of DNA repair in carcinogenesis: evidence from transgenic and gene targeting studies. *Mutat Res*, **477**, 41-49.

317. Jasin, M. (2002) Homologous repair of DNA damage and tumorigenesis: the BRCA connection. *Oncogene*, **21**, 8981-8993.

318. Rosen, E., Fan, S., Pestell, R. and ID, G. (2003) BRCA1 gene in breast cancer. *J Cell Physiol*, **196**, 19-41.

319. Hartman, A. and Ford, J. (2002) BRCA1 induces DNA damage recognition factors and enhances nucleotide excision repair. *Nat Genet*, **32**, 180-184.

320. Takimoto, R., MacLachlan, T., Dicker, D., Niitsu, Y., Mori, T. and el-Deiry, W. (2002) BRCA1 transcriptionally regulates damaged DNA binding protein (DDB2) in the DNA repair response following UV-irradiation. *Cancer Biol Ther*, **1**, 177-186.

321. Jaloszynski, P., Kujawski, M., Czub-Swierczek, M., Markowska, J. and Szyfter, K. (1997) Bleomycin-induced DNA damage and its removal in lymphocytes of breast cancer patients studied by comet assay. *Mutat Res*, **85**, 223-233.

322. Nascimento, P., da Silva, M., Oliveira, E., Suzuki, M. and Okazaki, K. (2001) Evaluation of radioinduced damage and repair capacity in blood lymphocytes of breast cancer patients. *Braz J Med Biol Res*, **34**, 165-175.

323. Kovacs, E., Stucki, D., Weber, W. and Muller, H. (1986) Impaired DNA-repair synthesis in lymphocytes of breast cancer patients. *Eur J Cancer Clin Oncol*, **22**, 863-869.

324. Xiong, P., Bondy, M., Li, D., Shen, H., Wang, L., Singletary, S., Spitz, M. and Wei, Q. (2001) Sensitivity to benzo[a]pyrene diol-epoxide associated with risk of breast cancer in young women and modulation by glutathione S-transferase polymorphisms: a case-control study. *Cancer Res*, **61**, 8465-8469.

325. Boyum, A. (1968) Separation of leukocytes from blood and bone marrow. *Scand J Clin Lab Invest*, **21**, 7-106.

326. Forlenza, M., Latimer, J. and Baum, A. (2000) The effects of stress on DNA repair capacity. *Psychol Health*, **15**, 881-891.

327. Steier, H. and Cleaver, J. (1969) Exposure chamber for quantitative ultraviolet photobiology. *Lab Practice*, **18**, 1295.

328. Dell'Orco, R. and Anderson, L. (1981) Unscheduled DNA synthesis in human diploid cells of different donor ages. *Cell Biol Int Rep*, **5**, 359-364.

329. Hennis 3rd, H., Braid, H. and Vincent, R. (1981) Unscheduled DNA synthesis in cells of different shape in fibroblast cultures from donors of various ages. *Mech Ageing Dev*, **16**, 355-361.

330. Grossman, L. (1992) The role of DNA damage and its repair in the aging process. *Aging*, **4**, 252-257.

331. Andersson, B., Sadeghi, T., Siciliano, M., Legerski, R. and Murray, D. (1996) Nucleotide excision repair genes as determinants of cellular sensitivity to cyclophosphamide analogs. *Cancer Chemother Pharmacol*, **38**, 406-416.

332. Trainor, K., Wigmore, D., Chrysostomou, A., Dempsey, J., Seshadri, R. and Morley, A. (1984) Mutation frequency in human lymphocytes increases with age. *Mech Ageing Dev*, **27**, 83-86.

333. Akiyama, M., Kyoizumi, S., Hirai, Y., Kusunoki, Y., Iwamoto, K. and Nakamura, N. (1995) Mutation frequency in human blood cells increases with age. *Mutat Res*, **338**, 141-149.

334. Moriwaki, S., Ray, S., Tarone, R., Kraemer, K. and Grossman, L. (1996) The effect of donor age on the processing of UV-damaged DNA by cultured human cells: reduced DNA repair capacity and increased DNA mutability. *Mutat Res*, **364**, 117-123.

335. Grossman, L. (1997) Epidemiology of ultraviolet-DNA repair capacity and human cancer. *Environ Health Perspect*, **105**, 927-930.

336. Wei, Q. (1998) Effect of aging on DNA repair and skin carcinogenesis: a minireview of population-based studies. *J Invest Dermatol Symp Proc*, **3**, 19-22.

337. Breast Cancer Information Core [http://research.nhgri.nih.gov/bic/].

338. Rubinstein, W., Vogel, V., Sumkin, J., Huerbin, M. and Latimer, J. (2006) A study of 0.5 Tesla dedicated magnetic resonance imaging for the detection of breast cancer in young, high risk women. *BMC Womens Health*, **6**, 10.

339. Boyd, N., Lockwood, G., Martin, L., Knight, J., Byng, J., Yaffe, M. and Tritchler, D. (1998) Mammographic densities and breast cancer risk. *Breast Dis*, **10**, 113-126.

340. Kishimoto, S., Tomino, S., Inomata, K., Kotegawa, S., Saito, T., Kuroki, M., Mitsuya, H. and Hisamitsu, S. (1978) Age-related changes in the subsets and functions of human T lymphocytes. *J Immunol*, **121**, 1773-1780.

341. Song, L., Kim, Y., Chopra, R., Proust, J., Nagel, J., Nordin, A. and Adler, W. (1993) Age-related effects in T cell activation and proliferation. *Exp Gerontol*, **28**, 313-321.

342. Goukassian, D., Gad, F., Yaar, M., Eller, M., Nehal, U. and Gilchrest, B. (2000) Mechanism and implications of the age-associated decrease in DNA repair capacity. *FASEB J*, **14**, 1325-1334.

343. Yager, J. and Liehr, J. (1996) Molecular mechanisms of estrogen carcinogenesis. *Annu Rev Pharmacol Toxicol*, **36**, 203-232.

344. Henderson, B. and Feigelson, H. (2000) Hormonal carcinogenesis. *Carcinogenesis*, **41**, 427-433.

345. Cole, J., Green, M., James, S., Henderson, L. and Cole, H. (1988) A further assessment of factors influencing measurements of thioguanine resistant mutant frequency in circulating lymphocytes. *Mutat Res*, **204**, 493-507.

346. Khaidakov, M. and Glickman, B. (1996) Possible factors leading to a misjudgement of mutant frequencies in HPRT assay. *Mutat Res*, **354**, 9-14.

347. Robbins, J. and Kraemer, K. (1972) Prolonged ultraviolet-induced thymidine incorporation into xeroderma pigmentosum lymphocytes: studies on its duration, amount, localization and relationship to hydroxyurea. *Biochem Biophys Acta* **277**, 7-14.

348. Andrews, A., Robbins, J., Kraemer, K. and Buell, D. (1974) Xeroderma pigmentosum long-term lymphoid lines with increased ultraviolet sensitivity. *J Natl Cancer Inst*, **53**, 691-693.

349. Langlois, R., Bigbee, W. and Jensen, R. (1990) In Mendelsohn, M. and Albertini, R. (eds.), *Mutation and the Environment, Part C: Somatic and Heritable Mutation, Adduction and Epidemiology*. Wiley-Liss, New York, pp. 47-57.

350. Martin, G., Ogburn, C., Colgin, L., Gown, A., Edland, S. and Monnat Jr, R. (1996) Somatic mutations are frequent and increase with age in human kidney epithelial cells. *Hum Mol Genet*, **5**, 215-221.

351. Kovacs, E. and Almendral, A. (1987) Reduced DNA repair synthesis in healthy women having first degree relatives with breast cancer. *Eur J Cancer Clin Oncol*, **23**, 1051-1057.

352. Buchholz, T., Wu, X., Hussain, A., Tucker , S., Mills, G., Haffty, B., Bergh, S., Story, M., Geara, F. and Brock, W. (2002) Evidence of haplotype insufficiency in human cells containing a germline mutation in BRCA1 or BRCA2. *Int J Cancer*, **97**, 557-561.

353. Mamon, H., Dahlberg, W. and Little, J. (2003) Hemizygous fibroblast cell strains established from patients with BRCA1 or BRCA2 mutations demonstrate an increased rate of spontaneous mutations and increased radiosensitivity. *Int J Radiat Oncol Biol Phys*, **57**, S356-357.

354. Rothfuss, A., Schutz, P., Bochum, S., Volm, T., Eberhardt, E., Kreienberg, R., Vogel, W. and Speit, G. (2000) Induced micronucleus frequencies in peripheral lymphocytes as a screening test for carriers of a BRCA1 mutation in breast cancer families. *Cancer Res*, **60**, 390-394.

355. Das, R. and Grant, S. (2002) Unusual pattern of age-related somatic mutation in cell lines derived from carriers of mutations in the BRCA1 and 2 genes [abstract]. *Am J Hum Genet*, **71 Suppl**, 251.

356. Nieuwenhuis, B., Van Assen-Bolt, A., Van Waarde-Verhagen, M., Sijmons, R., Van der Hout, A., Bauch, T., Streffer, C. and Kampinga, H. (2002) BRCA1 and BRCA2 heterozygosity and repair of X-ray-induced DNA damage. *Int J Radiat Biol*, **78**, 285-295.

357. Trenz, K., Rothfuss, A., Schutz, P. and Speit, G. (2002) Mutagen sensitivity of peripheral blood from women carrying a BRCA1 or BRCA2 mutation. *Mutat Res*, **500**, 89-96.

358. Kraemer, K., Lee, M. and Scotto, J. (1984) DNA repair protects against cutaneous and internal neoplasia: evidence from xeroderma pigmentosum. *Carcinogenesis*, **5**, 511-514.

359. Giglia, G., Dumaz, N., Drougard, C., Avril, M., Daya-Grosjean, L. and Sarasin, A. (1998) p53 mutations in skin and internal tumors of xeroderma pigmentosum patients belonging to the complementation group C. *Cancer Res*, **58**, 4402-4409.

360. Ferguson, A.T., Evron, E., Umbricht, C.B., Pandita, T.K., Chan, T.A., Hermeking, H., Marks, J.R., Lambers, A.R., Futreal, P.A., Stampfer, M.R. *et al.* (2000) High frequency of hypermethylation at the 14-3-3 sigma locus leads to gene silencing in breast cancer. *Proc Natl Acad Sci U S A*, **97**, 6049-6054.

361. Helzlsouer, K., Harris, E., Parshad, R., Perry, H., Price, F. and Sanford, K. (1996) DNA repair proficiency: a potential susceptibility factor for breast cancer. *J Natl Cancer Inst*, **88**, 754-755.

362. Bowman, K., Sicard, D., Ford, J. and Hanawalt, P. (2000) Reduced global genomic repair of ultraviolet light-induced cyclobutane pyrimidine dimers in simian virus 40-transformed human cells. *Mol Carcinog*, **29**, 17-24.

363. Kaneko, M. and Cerutti, P.A. (1980) Excision of N-acetoxy-2-acetylaminofluorene-induced DNA adducts from chromatin fractions of human fibroblasts. *Cancer Res*, **40**, 4313-4319.

364. Shibutani, S., Reardon, J., Suzuki, N. and Sancar, A. (2000) Excision of tamoxifen-DNA adducts by the human nucleotide excision repair system. *Cancer Res*, **60**, 2607-2610.

365. Koberle, B., Grimaldi, K., Sunters, A., Hartley, J., Kelland, L. and Masters, J. (1997) DNA repair capacity and cisplatin sensitivity of human testis tumour cells. *Int J Cancer*, **70**, 551-555.

366. Koberle, B., Masters, J., Hartley, J. and Wood, R. (1999) Defective repair of cisplatin-induced DNA damage caused by reduced XPA protein in testicular germ cell tumours. *Curr Biol*, **9**, 273-276.

367. Freshney, R. (1994), *Culture of Animal Cells, a Manual of Basic Technique*. Wiley-Liss, New York, pp. 9-16.

368. Romanov, S., Kozakiewicz, B., Holst, C., Stampfer, M., Haupt, L. and Tlsty, T. (2001) Normal human mammary epithelial cells spontaneously escape senescence and acquire genomic changes. *Nature*, **409**, 633-637.

369. Stampfer, M. (2000) HMEC: Human Mammary Epithelial Cell Extended Life Culture; The Review of HMEC Culture System *http://www.lbl.gov/~mrgs/*.

370. Freshney, R. (1992), *Culture of Specialized Cells: Culture of Epithelial Cells*. Wiley-Liss, New York, pp. 1-25.

371. Dontu, G., Al-Hajj, M., Abdallah, W.M., Clarke, M.F. and Wicha, M.S. (2003) Stem cells in normal breast development and breast cancer. *Cell Prolif*, **36 Suppl 1**, 59-72.

372. Yeager, T. and Reddel, R. (1999) Constructing immortalized cell lines. *Curr Opin Biotechnol*, **10**, 465-469.

373. Giampieri, S. and Storey, A. (2004) Repair of UV-induced thymine dimers is compromised in cells expressing the E6 protein from human papillomaviruses types 5 and 18. *Br J Cancer*, **90**, 2203-2208.

374. Jiang, X.R., Jimenez, G., Chang, E., Frolkis, M., Kusler, B., Sage, M., Beeche, M., Bodnar, A.G., Wahl, G.M., Tlsty, T.D. *et al.* (1999) Telomerase expression in human somatic cells does not induce changes associated with a transformed phenotype. *Nat Genet*, **21**, 111-114.

375. Bates, S., Zhou, N., Federico, L., Xia, L. and O'Connor, T. (2005) Repair of cyclobutane pyrimidine dimers or dimethylsulfate damage in DNA is identical in normal or telomerase-immortalized human skin fibroblasts. *Nucleic Acids Res*, **33**, 2475-2495.

376. Porter, P., Clark, D., McDaniel, L., McGregor, W. and States, J. (2006) Telomerase-immortalized human fibroblasts retain UV-induced mutagenesis and p53-mediated DNA damage responses. *DNA Repair (Amst)*, **5**, 61-70.

377. Masutomi, K., Possemato, R., Wong, J., Currier, J., Tothova, Z., Manola, J., Ganesan, S. and Lansdorp. (2005) The telomerase reverse transcriptase regulates chromatin state and DNA damage responses. *Proc Natl Acad Sci U S A*, **102**, 8222-8227.

378. Elenbaas, B., Spirio, L., Koerner, F., Fleming, M., Zimonjic, D., Donaher, J., Popescu, N., Hahn, W. and Weinberg, R. (2001) Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. *Genes Dev*, **15**, 50-65.

379. Madsen, M.W., Lykkesfeldt, A.E., Laursen, I., Nielsen, K.V. and Briand, P. (1992) Altered gene expression of c-myc epidermal growth factor receptor-a, and c-erb-B2 in an immortalized human breast cancer cell line, HMT-3522, in association with decreased growth factor requirements. *Cancer Res*, **52**, 1210-1217.

380. Ciardello, F., Gottardis, M. and Bsolo, F. (1992) Additive effects of c-erbB-2, c-Ha-ras, and transforming growth factor-α genes on in vitro transformation of human mammary epithelial cells. *Mol Carcinog*, **6**, 43.

381. Bronzert, D., Greene, G. and Lippman, M. (1985) Selection and characterization of a breast cancer cell line resistant to the antiestrogen LY117018. *Endocrinology*, **117**, 1409-1417.

382. Smith, H., Wolman, S. and Hackett, A. (1984) The biology of breast cancer at the cellular level. *Biochem Biophys Acta*, **734**, 103-123.

383. Cailleau, R., Olive, M. and Cruciger, Q. (1978) Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. *In Vitro*, **14**, 911-915.

384. Band, V., Zajchowski, D., Swisshelm, K., Trask, Kulesa, V., Cohen, C., Connely, J. and Sager, R. (1990) Tumor progression in four mammary epithelial cell lines derived from the same patient. *Cancer Res*, **50**, 7351-7357.

385. Soule, H.D., Vazquez, J., Long, A., Albert, S. and Brennan, M. (1973) A human cell line from a pleural effusion derived from a breast carcinoma. *J Natl Cancer Inst*, **51**, 1409-1413.

386. Fogh, J., Wright, W. and Loveless, J. (1977) Absence of HeLa cell contamination in 169 Cell lines derived from human tumors. *J Natl Cancer Inst*, **58**, 209-214.

387. Fogh, J. and Trempe, G. (1975) In Fogh, J. (ed.), *Human Tumor Cells In Vitro*. Plenum Press, New York, pp. 115-159.

388. ATCC. (Accessed 2003) ATCC Catalogue HTB-30: SK-BR-3; http://www.atcc.org.

389. Novaro, V., Radisky, D., Ramos Castro, N., Weisz, A. and Bissell, M. (2004) Malignant mammary cells acquire independence from extracellular context for regulation of estrogen receptor alpha. *Clin Cancer Res*, **10**, 402S-409S.

390. Ballare, C., Bravo, A.I. and Laucella, S. (1989) DNA synthesis in estrogen receptor positive human breast cancer takes place preferentially in estrogen-receptor negative cells. *Cancer*, **64**, 842.

391. Shek, L.L. and Godlophin, W. (1989) Survival with breast cancer: the importance of estrogen receptor quantity. *Eur J Cancer Oncol*, **25**, 243.

392. Clark, G.M. and MacGuire, W.L. (1988) Steroid receptors and other prognostic factors in primary breast cancer. *Semin Oncol*, **15**, 20.

393. Keydar, I., Chen, L., Karby, S., Weiss, F.R., Delarea, J., Radu, M., Chaitcik, S. and Brenner, H.J. (1979) Establishment and characterization of a cell line of human breast carcinoma origin. *Europ J Cancer*, **15**, 659-670.

394. Van Netten, J.P., Armstrong, J.B. and Carlyle, S.S. (1988) Estrogen receptor distribution in the peripheral, intermediate, and center regions of breast cancers. *Eur J Cancer Clin Oncol*, **24**, 1885.

395. Villalobos, M., Olea, N., Brotons, J., Olea-Serrano, M., Ruiz de Almodovar and Pedraza, V. (1995) The E-screen assay: a comparison of different MCF7 cell stocks. *Environ Health Perspect*, **103**, 844-850.

396. Wegner, S., Senft, J., Linda, M., Bamezai, R., Bairwa, N. and Grant, S. (2004) Comparison of Established Cell Lines at Different Passages by Karyotype and Comparative Genomic Hybridization. *Biosci Rep*, **24**, 631-639.

397. Gaffney, E., Pigott, D. and Grimaldi, M. (1979) Human serum and the growth of human mammary cells. *J Natl Cancer Inst*, **63**, 913-918.

398. Caron de Fromentel, C., Nardeux, P., Soussi, T., Lavialle, C., Estrade, S., Carloni, G., Chandrasekaran, K. and Cassingena, R. (1985) Epithelial HBL-100 cell line derived from milk of an apparently healthy woman harbours SV40 genetic information. *Exp Cell Res*, **160**, 83-94.

399. Briand, P., Petersen, O. and Van Deurs, B. (1987) A new diploid nontumorigenic human breast epithelial cell line isolated and propagated in chemically defined medium. *In Vitro Cell Dev Biol*, **23**, 181-188.

400. Kelly, C.M. and Latimer, J.J. (2005) Unscheduled DNA synthesis: a functional assay for global genomic nucleotide excision repair. *Methods Mol Biol*, **291**, 303-320.

401. Schaeffer, W. (1990) Terminology Associated with Cell, Tissue and Organ Cultures, Molecular Biology and Molecular Genetics. *In Vitro Cell Dev Biol*, **26**, 97-101.

402. Pathak, S., Siciliano, M., Cailleau, R., Wiseman, C. and Hsu, T. (1979) A human breast adenocarcinoma with chromosome and isoenzyme markers similar to those of the HeLa line. *J Natl Cancer Inst*, **62**, 263-271.

403. Ochieng, J., Warfield, P. and Johnson, K. (1997) Establishment of breast cell cultures and lines from peoples of African origin. *Am J Obstet Gynecol*, **176**, 240-245.

404. Pattillo, R., Hussa, R. and Story, M. (1997) Characteristics of antigens from an HCG alpha secreting cell line (ElCo) derived from human breast carcinoma in a Native American patient. *Am J Obstet Gynecol*, **176**, 246-254.

405. Hedenfalk, I., Duggan, D., Chen, Y., Radmacher, M., Bittner, M., Simon, R., Meltzer, P., Gusterson, B., Esteller, M., Kallioniemi, O. *et al.* (2001) Gene-expression profiles in hereditary breast cancer. *N Engl J Med*, **344**, 539-548.

406. McGurk, C., Cummings, M., Koberle, B., Hartley, J., Oliver, R. and Masters, J. (2006) Regulation of DNA Repair Gene Expression in Human Cancer Cell Lines. *J Cell Biochem*, **97**, 1121-1136.

407. Fautrel, A., Andrieux, L., Musso, O., Boudjema, K., Guillouzo, A. and Langouet, S. (2005) Overexpression of the two nucleotide excision repair genes ERCC1 and XPC in human hepatocellular carcinoma. *J Hepatol*, **43**, 288-293.

408. Reed, E., Yu, J., Davies, A., Gannon, J. and Armentrout, S. (2003) Clear cell tumors have higher mRNA levels of ERCC1 and XPB than other histological types of epithelial ovarian cancer. *Clinical Cancer Res*, **9**, 5299-5305.

409. Juhasz, A., Frankel, P., Cheng, C., Rivera, H., Vishwanath, R., Chiu, A., Margolin, K., Yen, Y., Newman, E., Synold, T. *et al.* (2003) Quantification of a chemotherapeutic target gene mRNA expression in human breast cancer biopsies: comparison of real-time reverse transcription PCR vs. relative quantification reverse transcription PCR utilizing DNA sequences analysis of PCR products. *J Clin Lab Anal*, **17**, 184-194.

410. Vogel, U., Nexo, B., Tjonneland, A., Walling, H., Hertel, O. and Raaschou-Nielsen, O. (2006) ERCC1, XPD, and RAI mRNA levels in lymphocytes are not associated with lung cancer risk in a prospective study of Danes. *Mutat Res*, **593**, 88-96.

411. Cheng, L., Sturgis, E., Eicher, S., Spitz, M. and Wei, Q. (2002) Expression of nucleotide excision repair genes and the risk for squamous cell carcinoma of the head and neck. *Cancer*, **95**, 393-397.

412. Murata, H., Khattar, N., Gu, L. and Li, G. (2005) Roles of mismatch repair proteins hMSH2 and hMLH1 in the development of sporadic breast cancer. *Cancer Lett*, **233**, 143-150.

413. Lau, K., LaSpina, M., Long, J. and Ho, S. (2000) Expression of estrogen receptor (ER)-alpha and ER-beta in Normal and Malignant Prostatic Epithelial Cells: Regulation by Methylation and Involvement in Growth Regulation. *Cancer Res*, **60**, 3175-3182.

414. Inoue, C., Shiga, K., Takasawa, S., Kitagawa, M., Yamamoto, H. and Okamoto, H. (1987) Evolutionary conservation of the insulinoma gene rig and its possible function. *Proc Natl Acad Sci U S A*, **84**, 6659-6662.

415. Werner, T., Fessele, S., Maier, H. and Nelson, P. (2003) Computer modeling of promoter organization as a tool to study transcriptional coregulation. *FEBS J*, **17**, 1228-1237.

416. Engel, L. and Young, N. (1978) Human breast carcinoma cells in continuous culture: a review. *Cancer Res*, **38**, 4327-4339.

417. Yen, L., Woo, A., Christopoulopoulos, G., Batist, G., Panasci, L., Roy, R., Mitra, S. and Alaoui-Jamali, M. (1995) Enhanced host cell reactivation capacity and expression of DNA repair genes in human breast cancer cells resistant to bi-functional alkylating agents. *Mutat Res*, **337**, 179-189.

418. Liu, Z., Zhang, X., Zhang, Y. and Yang, X. (2004) Progesterone receptor gene inactivation and CpG island hypermethylation in human leukemia cancer cells. *FEBS Lett*, **567**, 327-332.

419. Wei, M., Grushko, T., Dignam, J., Hagos, F., Nanda, R., Svenn, L., Xu, J., Fackenthal, J., Tretiakova, M., Das, S. *et al.* (2005) BRCA1 promoter methylation in sporadic breast cancer is associated with reduced BRCA1 copy number and chromosome 17 aneusomy. *Cancer Res*, **65**, 10692-10699.

420. Suzuki, H., Itoh, F., Toyota, M., Kikuchi, T., Kakiuchi, H. and Imai, K. (2000) Inactivation of the 14-3-3 sigma gene is associated with 5'CpG island hypermethylation in human cancers. *Cancer Res*, **60**, 4353-4357.

421. Tsutsumida, A., Hamada, J., Tada, M., Aoyama, T., Furuuchi, K., Kawai, Y., Yamamoto, Y., Sugihara, T. and Moriuchi, T. (2004) Epigenetic silencing of E- and P-cadherin gene expression in human melanoma cell lines. *Int J Oncol*, **25**, 1415-1421.

422. Arnold, C., Goel, A. and Boland, C. (2003) Role of hMLH1 promoter hypermethylation in drug resistance to 5-fluorouracil in colorectal cancer cell lines. *Int J Cancer*, **106**, 66-73.

423. Andrew, A., Karagas, M. and Hamilton, J. (2003) Decreased DNa repair gene expression among individuals exposed to arsenic i nthe United States. *Int J Cancer*, **104**, 263-268.

424. Reed, E. (1998) Platinum-DNA adduct, nucleotide excision repair and platinum based anti-cancer chemotherapy. *Cancer Treat Rev*, **24**, 331-344.

425. Dabholkar, M., Thornton, K., Vionnet, J.A., Bostick-Bruton, F., Yu, J. and Reed, E. (2000) Increased mRNA levels of xeroderma pigmentosum complementation group B (XPB), and Cockayne's Syndrome complementation group B (CSB) without increased mRNA levels of multidrug-resistance (MDR1) or metallothionein-II (MT-II) in platinum-resistant human ovarian cancer tissues. *Biochem Pharmacol*, **60**, 1611-1619.

426. Shirota, Y., Stoehlmacher, J., Brabender, J., Xiong, Y., Uetake, H., Danenberg, K., Groshen, S., Tsao-Wei, D., Danenberg, P. and Lenz, H. (2001) ERCC1 and thymidilate synthase mRNA levels predict survival for colorectal cancer patients receiving combination oxaliplatin and fluorouracil chemotherapy. *J Clin Oncol*, **19**, 4298-4304.

427. Lord, R., Brabender, J., Gandara, D., Alberola, V., Camps, C., Domine, M., Cardenal, F., Sanchez, J., Gumerlock, P., Taron, M. *et al.* (2002) Low ERCC1 expression correlates with prolonged survival after cisplatin plus gemcitabine chemotherapy in non-small cell lung cancer. *Clinical Cancer Res*, **8**, 2286-2291.