NOTE TO USERS

DISSERTATION

IDENTIFICATION AND ANALYSIS OF NOVEL RADIATION RESPONSE GENES USING A GENE TRAPPED LIBRARY OF HUMAN MAMMARY EPITHELIAL CELLS

Submitted by

Jennifer Malone

Cell and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2004

COLORADO STATE UNIVERSITY

August 10, 2004

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY JENNIFER MALONE ENTITLED IDENTIFICATION AND ANALYSIS OF NOVEL GENES AFFECTED BY GAMMA IRRADIATION USING A GENE-TRAPPED LIBRARY OF HUMAN MAMMARY EPITHELIAL CELLS BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Michael Fox

William Hanneman

Sue Lana

Advisor Robert Ullrich

Department Head Michael Fox

ABSTRACT OF DISSERTATION

IDENTIFICATION AND ANALYSIS OF NOVEL GENES AFFECTED BY GAMMA IRRADIATION USING A GENE-TRAPPED LIBRARY OF HUMAN MAMMARY EPITHELIAL CELLS

Breast cancer is one of the most common cancers among women and is the second leading cause of cancer-associated death in women in the United States. Ionizing radiation is a known carcinogen of the female breast. The female breast is one of the most sensitive tissues to radiation induced cancer and ionizing radiation is one of the few agents identified as a cause of breast cancer in women. Enhanced mutation rates at all levels of radiation exposure were found to cause an increase in transformation.

It is believed that the expression of several unknown genes is directly affected by gamma radiation. Radiation exposure results in altered expression of a wide variety of genes and gene pathways. Many radiation response genes have only recently been identified and more remain to be identified. Abnormal expression of these genes may be involved in the early steps in breast carcinogenesis induced by radiation. An assay has been developed to screen breast cells for a single gene mutation using a technique called gene trapping, which allows changes to be detected in the expression of a specific gene upon treatment with different doses of radiation. These radiation-response genes can then be identified through the rapid amplification of cDNA ends (RACE) procedure and sequenced. Clones affected by radiation exposure were isolated and further analyzed to see if could lead to malignant transformation of the normal breast epithelial cell into a neoplastic cell and effect survival.

Five candidate radiation response genes were identified and their relative mRNA expression levels were analyzed in the MCF10A parental cells both with and without ionizing

ionizing radiation treatment. The five genes are as follows: human androgen receptor, human creatine kinase, human DORA reverse strand protein 1, human eukaryotic elongation translation factor 1 beta 2, and human ribosomal protein L27. They were found to be significantly downregulated relative to the parental at doses of radiation ranging from 0.10 Gy to 4.0 Gy. A radiation dose response was clearly seen for all these genes and that was not cell-type specific. The importance of these genes in radiation-induced cell killing, mutagenesis, clastogenesis and cancer remains to be determined.

Jennifer Lynn Malone
Cell and Molecular Biology
Colorado State University
Fort Collins, Colorado
Fall 2004

TABLE OF CONTENTS

CHAPTER 1-INTRODUCTION

Ionizing radiation is one of the primary modalities used in both the diagnosis and treatment of cancer. While the use of medical radiation has undoubtedly prolonged and saved the lives of many, it is not without side effects. Radiation is one agent that has clearly been identified as a cause of breast cancer in humans. Enhanced mutation rates, cell killing, increased amounts of transformation and cancer have all been observed at all levels of radiation exposure. One of the most sensitive tissues with respect to radiation carcinogenesis is the female breast. A radiation dose-related increase in the incidence of breast cancer has been seen in a wide variety of epidemiological studies. The risk of a woman developing breast cancer following radiation treatment is directly proportional to dose even at very low doses. There is little evidence that altering the dose rate can reduce the risk for a women developing breast cancer once exposed to a radiation dose. An important modifying factor in breast cancer development is age. In epidemiological studies it has been observed that when a woman receives significant radiation prior to the age of 20, she becomes more likely to develop breast cancer while exposure at ages of 45 or greater shows less risk (22). Thus there is considerable interest in understanding the cellular response to radiation. One approach to this problem is to understand the molecular mechanisms underlying the radiation responses of normal mammary epithelial cells. The purpose of this work is to elucidate novel radiation response genes in human mammary epithelial cells that potentially may effect initiation and progression into carcinogenesis.

Breast Cancer Statistics

Breast cancer is the most common form of cancer in women in the United States, other than skin cancer. An estimated 211,300 new cases of invasive breast cancer are expected to occur among women in the United States during 2003 (5). Breast cancer incidence rates continue to increase since 1980, although the rates have slowed down since the early 1990s in comparison to rate increase in the 1980s. In recent times, breast cancer rates have only increased in postmenopausal women. About 1,300 new cases of breast cancer are expected in men in 2003. In addition to invasive breast cancer, 55,700 new cases of in situ breast cancer are expected to occur in women in 2003 (5). Approximately 85% of these will be ductal carcinoma in situ (DCIS). The increase in DCIS cases is believed to be a direct result of the increased utilization of mammography screening which helps to detect invasive breast cancers before they are palpable (5).

An estimated 40,200 deaths (39,800 women and 400 men) are anticipated to occur from breast cancer in 2003 (5). Breast cancer is the second leading cause of cancer-associated death for all women (after lung cancer), and the leading overall cause of death in women between the ages of 40 and 55. Mortality rates have been found to decline by about 1.4% per year during 1989-1995 and by 3.2% afterwards, with the largest decreases seen in younger women from white and African American backgrounds (5). These decreases are most likely due to earlier detection and improved treatment. The earliest sign of breast cancer is typically an abnormality detected by mammography before it is felt by the patient. Physical signs and symptoms of breast cancer include: a breast lump, thickening, swelling, distortion, tenderness, skin irritation or dimpling, nipple pain, scaliness, ulceration, or retraction (5). There are both genetic and environmental components associated with the risk for breast cancer development. Up to 10% of

breast cancer cases in western countries may be attributable to a genetic predisposition (20). Breast cancer incidence increases strongly with age, with higher risks seen in postmenopausal women. Approximately 40% of women with breast cancer will die from the disease (20, 138).

Breast Development

The breast undergoes dramatic changes in size, shape, and function in association with growth, reproduction, and post-menopausal regression. These changes impact a women's lifetime breast cancer risk. The process of mammary gland differentiation is the result of complex interactions of hormones, activation of specific genes, and expression of extracellular matrix proteins in the normal breast. An early first full-term pregnancy exerts a protective effect, emphasizing the need for understanding the role of reproductive influences on breast development and on cancer initiation and progression, and providing a paradigm for developing preventative strategies based on physiological principles. Even though the cause of breast cancer and the ultimate mechanisms through which an early pregnancy protects from cancer development remain largely unknown, a likely explanation for this protection has been provided by in vivo and in vitro models. These studies have led to the conclusions that cancer initiation requires the interaction of a carcinogen with an undifferentiated and highly proliferating mammary epithelium, whereas differentiation of lobular structures of the mammary gland inhibits carcinogenic initiation (153, 193).

Breast cancer is a hormone- and sex-dependent malignancy whose development is influenced by a variety of hormones and growth factors. Epidemiological and clinical evidence indicate that breast cancer risk is associated with prolonged exposure to female ovarian hormones. Among all the complex hormonal influences, estrogens are considered to play a major role in promoting cell proliferation of both the normal and neoplastic breast epithelium.

Estrogens have been demonstrated to be of essential importance in this disease because it is observed in postmenopausal hyperestrogenism resulting from the use of estrogenic hormone replacement therapy and obesity. Estrogens, which are necessary for the normal development of both reproductive and nonreproductive organs, exert their physiological effects by binding to their specific receptors, the estrogen receptors ERα or ERβ. Estrogens can also act through alternative nonreceptor mediated pathways (152).

Breast Cancer Risk Factors

It is not known when breast cancer is initiated in the lifetime of a woman. Recent animal and human studies of breast cancer risk suggest that a population of cells may exist with the potential to become cancerous for many years prior to the onset of the disease. Late menarche and a full-term pregnancy completed before age 24, or early full-term pregnancy, reduce the risk of breast cancer development, whereas early menarche, nulliparity, or late parity is associated with a higher breast cancer incidence. These facts indicate that the period between menarche and the first full-term pregnancy represents a window of high susceptibility for the initiation of breast cancer. Besides an early menarche and late menopause, other risk factors for breast cancer include age over 50, first childbirth after age 30, nulliparity, family history of breast cancer, biopsy-confirmed atypical hyperplasia, increased breast density, recent use of oral contraceptives or postmenopausal estrogens and progestin, obesity, possibly dietary factors such as a high-fat diet, consumption of one or more alcoholic beverages per day, and previous radiation exposure to the chest before age 35. A majority of women who develop breast cancer do not have a known risk factor, which indicates that the causes may be multiple and are not yet clearly understood. Significant increases in risk are associated with a history of breast cancer in a young sister or mother or with a family history of bilateral disease. The risk of developing breast

cancer by age 65 approaches 50% for women with a strong family history of breast cancer (20). Only a few of the genes known to be involved with an increased risk for developing breast cancer have been identified.

Clinical data have indicated that exposure to ionizing radiations at ages younger than 19 increases the incidence of breast cancer, while exposure at older ages or after pregnancy does not increase cancer risk (201). Both animal studies and epidemiological research indicate that the periods during childhood and adolescence might be particularly sensitive for breast cancer initiation (201). Evidence has also suggested that peripubertal exposure to estrogens reduces breast cancer risk. Since estrogens are thought to typically promote breast cancer development, the mechanisms leading to their protective effect in peripubertal girls are likely due to mammary gland morphology. If the mammary gland is exposed to estrogen before puberty the normal pattern of mammary epithelial proliferation and differentiation may be altered (201).

Adolescence is a time of rapid growth and maturation of the breasts. Little is known about the role of estrogens during childhood and its effect on later breast cancer. It has been shown that girls between the ages of 7 through 15 with a high body mass index exhibit a significantly reduced breast cancer risk. Height in childhood has also been found to be directly linked to breast cancer risk; that is, if a girl is tall then her risk of developing breast cancer is significantly increased. Childhood height is inversely linked to circulating estrogen levels. It has been demonstrated that a high fat intake at puberty does not increase breast cancer risk but rather reduces it (203).

Known Breast Cancer Gene Mutations

Although activation of oncogenes, such as Her-2-neu, have clear relevance in selected cases of breast cancer progression and in treatment, studies of hereditary forms of human breast

cancer suggest mutations in one or more tumor suppressor genes are more likely to be assessed with increased cancer risk as early initiating events. As a class, these genes function to maintain genomic integrity and help prevent the propagation of damaged DNA. Aberrations in many tumor suppressor genes directly affect cellular susceptibility to DNA damage and the cellular capacity for DNA damage repair. Others recognize damaged DNA and promote cell cycle arrest before DNA synthesis and mitosis occur. Finally, tumor suppressor gene products may also inhibit propagation of damaged DNA by inducing apoptotic cellular death (29).

Mutations in tumor suppressor genes can either directly or indirectly increase the risk of carcinogenesis. Mutations in certain classes of genes, such as those that regulate the cell cycle or those that promote apoptosis in cells with DNA damage, can directly lead to a malignant cellular change. These functions are the final checkpoints in the maintenance of DNA integrity. A mutation in this gene class leads to a generalized genomic instability. This instability increases the probability that DNA damage may escape the surveillance function of other checkpoint genes such as p53. Mutations in tumor suppressor genes can be classified as germline or somatic. Germline mutations are inherited, and therefore are present in every cell, whereas somatic mutations result from exposure to an external mutagen (such as radiation) or an error in DNA replication. Germline mutations most commonly affect one of the two copies of the inherited gene. This is true for such genes as p53, BRCA1, and BRCA2. Such individuals are heterozygous for the genetic mutation. In these individuals, the remaining copy of the gene often successfully performs the required tumor-suppressor functions. Therefore, these heterozygous mutations may not lead to obvious phenotypic changes. However, the probability of developing a somatic mutation in this remaining copy of the gene, with subsequent loss of the gene function, is more than 1,000 times greater than the two somatic mutations required for an

individual with a normal germline (73). As a result, the corresponding risk of cancer development in individuals with specific germline mutations is significantly greater than in the general population (73, 82).

Approximately 10% of breast cancer cases can be classified as familial. This is a subcategory of breast cancer patients in whom the disease occurs with a very high (over 50%) penetrance in female family members. It is likely that familial breast cancer results from autosomal inherited genetic mutations and certain germline mutations have been discovered in some of these families. The inherited gene mutations that have been clearly demonstrated to increase breast cancer risk involve the p53, ATM, BRCA1, and BRCA2 genes. It is estimated, however, that only 6% to 8% of the total United States breast cancer population have a germline mutation in one of these genes. However, these genes are not found in all cases of familial breast cancer, and efforts to isolate further germline genetic etiologies continue (20, 29). Mutations of the high-risk breast cancer-susceptibility genes BRCA1 and BRCA2 are thought to only account for half of the families with early onset breast cancer. This would mean that there could be numerous other breast cancer susceptibility genes yet undiscovered.

Even within the subgroup of patients with a family history, the inherited genetic factors that lead to their 2.25-fold relative risk for developing breast cancer are unknown except for selected cases with familial breast cancer. Additional work is needed to further characterize the molecular basis for the increased risk seen in individuals with positive family history. In these individuals, it is unlikely that a single genetic mutation is relevant to all cases (138). As advances in molecular genetics continue, a host of mutations will likely be found that contribute to breast cancer risk, but which have variable penetrance and may be insufficient to directly cause malignancy without other contributing factors (138). More than 70% of breast cancer

patients have no clear family history of breast cancer, but data suggests that 20-30% of these cases may have a genetic component.

Tumor suppressor genes and breast cancer

BRCA1, BRCA2, p53 and ATM are all tumor suppressor genes that are involved in maintaining genomic integrity, are key elements in DNA damage response pathways and are all impaired in the pathogenesis of breast cancer. Mutations in BRCA1, BRCA2, and ATM all affect cellular radiosensitivity and the success of double-strand break repair following ionizing radiation. Both BRCA1 and BRCA2 colocalize with Rad51 following radiation-induced double-strand injuries. Mutations in the BRCA1 tumor suppressor gene are found in about 70% of all of the families with inherited breast and ovarian cancers and about 20% of the families with only breast cancer. Considerable evidence points to the fact that both BRCA1 and BRCA2 proteins have multiple, complex roles in the cellular responses to DNA damage. They participate in the processes that implement cell cycle checkpoints in response to double strand breaks, help coordinate the repair of those breaks, facilitate transcription coupled repair of oxidative base damage, and also act as transcriptional modifiers. Although BRCA1 and BRCA2 interact with one another, they have distinct interacting partner proteins and different functions (173).

BRCA1

BRCA1 is a large protein with several functional domains. It can act as a transcriptional activator and is part of the RNA polymerase II holoenzyme complex. p53 and BRCA1 have been demonstrated to form a complex both in vivo and in vitro, with BRCA1 stimulating p53-dependent expression of genes including WAF1. BRCA1 can stimulate WAF1 expression in a p53-independent manner. Several studies have suggested that BRCA1 participates in the G2-M cell cycle checkpoint (173).

Normal function of BRCA1 is required for transcription-coupled repair following damage from ionizing radiation. BRCA1 localization to the sites of double strand breaks produced by ionizing radiation precedes the appearance of Rad50 and Rad51. The finding that BRCA1 strongly binds DNA, with a preference for branched structures, suggests that BRCA1 could act as a damage sensor that helps mediate repair. Phosphorylation of BRCA1 in response to ionizing radiation damage is performed primarily by ATM and Chk2, and secondarily by ATR (173).

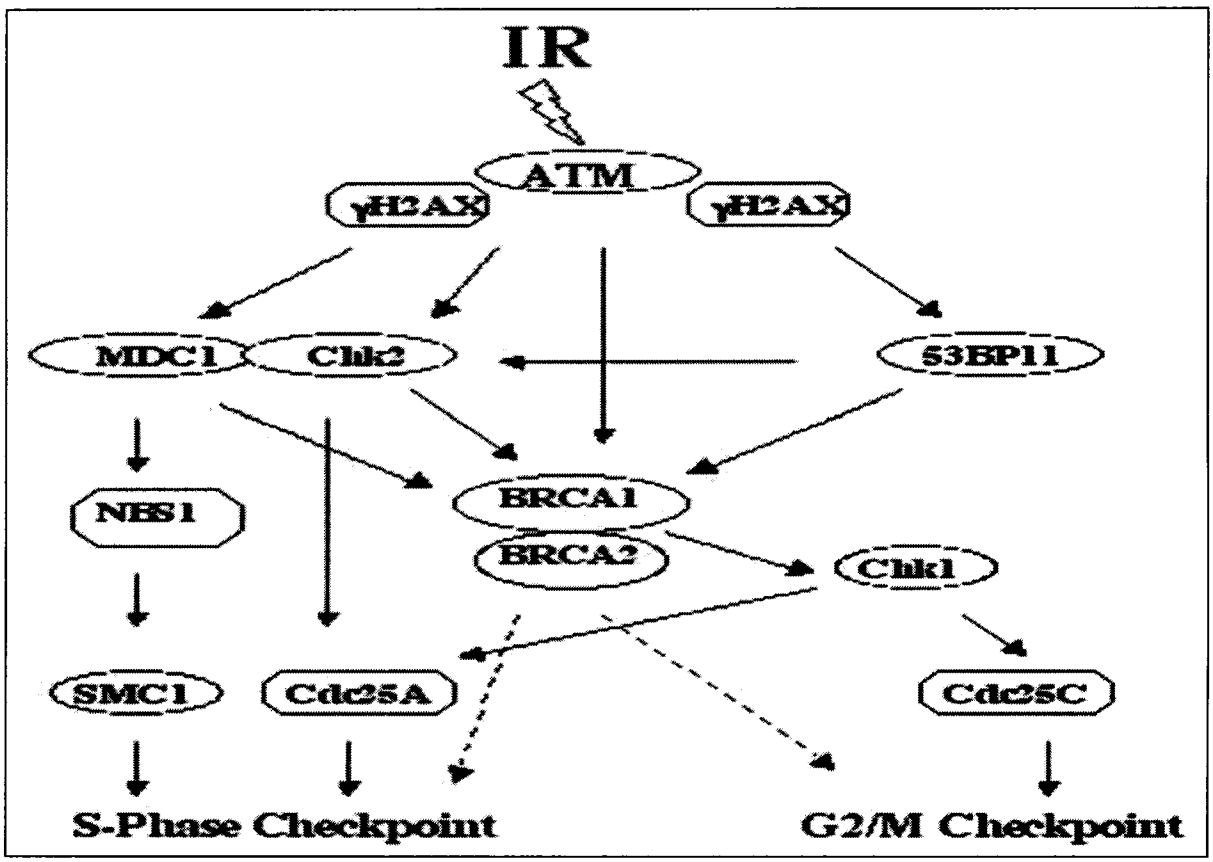


Figure 1.1. Schematic representation of BRCA1/2-dependent checkpoint pathways. Source: Lou, Z and J. Chen. *BRCA proteins and DNA damage checkpoints.* Frontiers in Bioscience, 2003. 8:s718-21.

Clues about BRCA1's specific biochemical functions stem from its interactions with numerous, diverse proteins and its presence in several multimeric complexes that are dynamically altered after DNA replication blockage. The interaction of BRCA1 with the Rad50-Mre11-NBS1 complex could function in directing the repair of double strand breaks through the homologous recombinational repair pathway by mediating the resectioning of broken termini to generate the long single-stranded tails required for assembly of Rad51 nucleoprotein filaments. During this end processing, BRCA2 may participate by regulating the assembly of Rad51 (173). BRCA1 has also been shown to associate with hRad50-hMre11-p95 in directing a cellular DNA damage response following ionizing radiation (29). Evidence that BRCA1 participates in repair of DNA damage includes its phosphorylation and translocation to the nucleus at specific parts of the cell cycle in response to DNA damage. The BRCA1 protein has domains that interact with the RAD51 protein, RAD50 protein, Rb protein, and p53 protein, among others. In the nucleus, BRCA1 associates with RAD51, which is involved in homologous recombination and double-strand break repair. The BRCA1 protein is interacting with proteins involved in DNA repair and recombination, or cell cycle checkpoints. BRCA1 has also been implicated in the non-homologous end joining repair pathway. It is involved in transcription-coupled repair of oxidative damage induced by ionizing radiation or hydrogen peroxide. BRCA1 is found to be associated with the centrosome during mitosis and the γ-tubulin has been found to associate preferentially with a hypophosphorylated form of BRCA1. These observations imply that BRCA1 plays an important role in centrosome regulation and chromosome segregation (173).

It has been demonstrated that BRCA1 exists as a large complex called BASC, or BRCA1-associated genome surveillance complex, which contains ATM, the Nbs1-MRE11-RAD50 complex, mismatch repair proteins MSH2/6 and MLH2, and Bloom's helicase. All the

proteins of this complex have the ability to recognize aberrant DNA structures and could be involved in transmitting the presence of DNA damage to proteins that contribute to the maintenance of genomic integrity during DNA replication and repair. BRCA1 is notable for an N-terminal ring finger domain, a possible site for dimerisation with other proteins, and recently shown to function as an E3 ubiquitin ligase, giving specificity to ubiquitin conjugating E2 activities. The N-terminal RING finger domain has been shown to interact with BARD1 to form a RING-finger dimeric complex. Within this RING domain of BRCA1, are several cancer-predisposing mutations that alter BRCA1's self association or its ability to interact with ubiquitin-conjugating enzymes have been identified. Some breast cancer-associated BRCA1 mutant proteins that lack ubiquitin ligase activity result in sensitivity to ionizing radiation. Through its role with BARD1, BRCA1 could also have a role in down regulating mRNA 3' processing in response to DNA damage (173). The interaction with FANCD2 and the mono-ubiquitination of that protein following DNA damage may reflect the central role of BRCA1 in DNA repair. Recently, BRCA1 has been shown to associate with the Fanconi anemia protein D2, the BACH1 helicase and c-Abl. The association of BRCA1 with c-Abl is disrupted following treatment with ionizing radiation in an ATM-dependent manner. The disruption of the association between BRCA1 and the BRCA1 inhibitor, CtIP, following radiation treatment is mediated by an ATM phosphorylation of CtIP. There appears to be a great amount of cross-talk among the signaling pathways in which these proteins participate. ATM is known to phosphorylate CHK2, a protein that leads to the cell cycle arrest in cells due in part to the phosphorylation of p53. CHK2, like ATM, can also phosphorylate BRCA1 (114). Targeted deletions and disruptions in BRCA1 suggest BRCA1 may also play a role in the G2/M

checkpoint after DNA damage. The transcriptional activator function of BRCA1 may play a role in DNA repair, as BRCA1 activates GADD45 expression (126).

BRCA2

Germline mutations in one allele of BRCA2 are associated with a very high risk (up to 85%) of breast cancer in women and about a 15% risk of ovarian cancer. BRCA2 mutations are the most common inherited genetic alteration so far identified in familial pancreatic cancer. The very large, 3418 amino acid, BRCA2 protein was first linked to homologous recombination by showing direct interactions with Rad51. A role of BRCA2 protein in repair of double-strand breaks has been suggested which includes its direct association with the BRCA1 protein and the RAD51 protein. Consistent with BRCA2 having a role in homologous recombinational repair, cells carrying mutations in both alleles of BRCA2 show moderate sensitivity to killing by ionizing radiation, which is complemented upon transfection of a complementing cDNA. Overexpression of a BRC repeat element in normal cells also confers radiation sensitivity through a dominant negative mechanism. This is explained by the fact that several of the BRC repeats prevent the self-association of Rad51 and inhibit the formation of Rad51 nucleoprotein filaments on single-stranded DNA. The N-terminus of BRCA2 contains two regions that may involve transcriptional activation. The transcriptional co-activator protein, P/CAF (p300/CBP associated factor), possesses histone acetyltransferase activity. Thus, the BRCA2 interaction with P/CAF may regulate transcription through the recruitment of histone-modifying activity of the P/CAF co-activator. As mentioned, Rad51 interacts with BRCA2 through the BRC repeats in exon 11 as well as at the C-terminus next to the nuclear localization signal. The BRC3 and BRC4 polypeptides interfere with Rad51 nucleoprotein filament formation, and cancer-associated mutations in these repeats lack the ability to interfere with Rad51's function. BRCA2

also has a role in controlling the nuclear transport of Rad51 since Rad51 is predominantly in the cytoplasm. Thus, BRCA2 may exert both negative and positive regulatory influences on Rad51. Other proteins that are known to interact with BRCA2 are BCCIPα, a tumor suppressor for breast and brain cancer, and BRAF35, a structure-specific DNA binding protein that binds to chromatin during mitotic prophase and influences cell cycle progression (173).

The complete loss of BRCA2 is incompatible with cell viability in dividing populations. Many BRCA2-defective human tumors have internal deletions or truncations ranging from termination points within exon 11, which contains the BRC repeats, to more benign ones in exons 26 or 27 at the very C-terminus, which contains the nuclear localization signal and a Rad51 interaction region. Very high levels of chromosome breaks and exchanges are seen in BRCA2 mutant cells. BRCA2 mutants illustrate about a 1.5 – 2-fold increase in sensitivity to killing by ionizing radiation and higher levels of sensitivity to crosslinking agents such as mitomyocin C and cisplatin. The Fanconi anemia groups B and D1 have been identified as having causative BRCA2 mutations (173). Deregulation of mutated BRCA2 protein may lead to deficient DNA repair and genomic instability. FANCG/XRCC9 has been found to bind to the highly conserved C-terminal site in BRCA2. BRCA2 is thought to be involved in homologous recombination along with RAD52, RAD54, RAD51 and its five paralogs: XRCC2, XRCC3, RAD51B, RAD51C and RAD51D. Recently an interaction between the FANCD2 protein and BRCA2 has been established. The interaction occurs in the highly conserved BLAT domain in BRCA2, which also binds FANCG/XRCC9 (76).

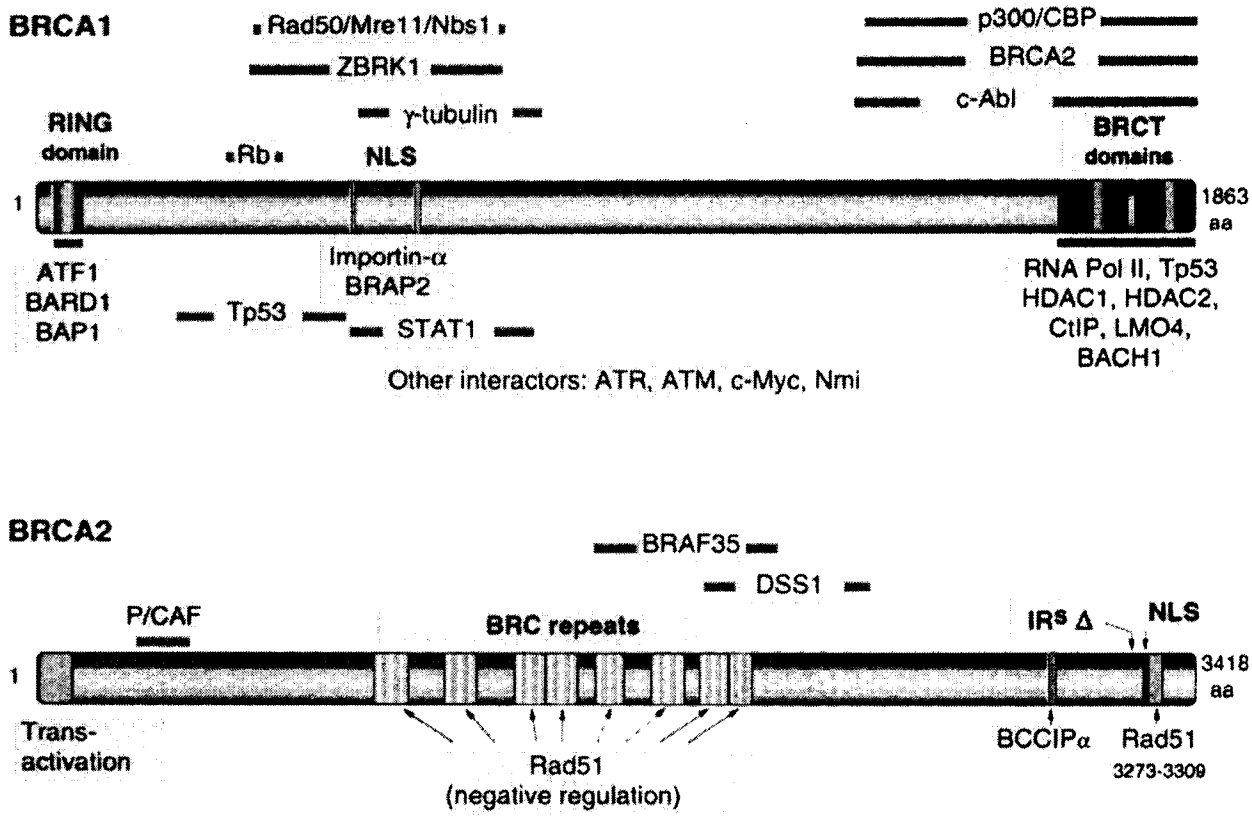


Figure1.2. Domains and interaction regions of the BRCA1 and BRCA2 proteins. Source: L.H. Thompson, Mutation Research, 509 (2002) 49-78.

ATM

Ataxia telangiectasia is an autosomal recessive disorder characterized by progressive cerebellar degeneration, oculocutaneous telangiectasia, and an increased risk for the development of malignancy of the lymphoid system (98). The increased risk of cancer is estimated to be 60 – 180 fold higher than the general population, with ~70% of malignancies being lymphomas and T cell leukemias, a wide variety of solid tumors make up the remainder. ATM patients display marked radiosensitivity upon treatment with low doses of ionizing radiation. The cellular phenotype of AT is characterized by chromosomal breakage, telomere instability, radiosensitivity, radioresistant DNA synthesis, defective cell cycle checkpoints, dysfunctional apoptosis, and a reduced p53 response. AT cell types are typically killed at radiation doses 3 –5 fold lower than those that kill normal cells, and similar sensitivity is manifest as chromosomal aberrations. Levels of spontaneous mutations at the GPA and hprt loci are elevated in AT cells, supporting a causal link between susceptibility to somatic mutation and cancer. Interestingly, AT cells have been shown to have higher levels of residual unrepaired double strand breaks after irradiation (173).

ATM kinase is one of the first molecules to respond to DNA damage and regulates many of the subsequent steps in the radiation response. ATM also plays a critical role in the successful repair of DNA strand breaks following radiation by phosphorylating proteins that lead to cell cycle arrest and to activation of DNA repair pathways (88). In irradiated cells, ATM phosphorylates multiple substrates, which are involved in checkpoint control, repair, apoptosis and other aspects of the response to DNA damage. An important substrate of ATM appears to be BRCA1, a tumor suppressor protein associated with susceptibility to breast and ovarian cancer. Interestingly, ATM heterozygosity has been associated with an increased risk in breast cancer.

Several epidemiological studies indicate that AT families have demonstrated a 4-fold increased risk of developing breast cancer among female heterozygotes. On the basis of the heterozygote frequency and increased breast cancer risk seen in these individuals, AT heterozygosity may account for up to 5% of all breast cancer cases in the general population. There have been several studies on the frequency of AT heterozygotes in nonfamilial breast cancer cases with different conclusions (204). This has lead to a great deal of controversy. This role may in part be explained by the finding that BRCA1 protein function is dependent on phosphorylation by the ATM protein. It is clear that homozygous mutations in any of these three genes (BRCA1, BRCA2, or ATM) result in a radiosensitive phenotype (29). It has been documented that 2 missense or synonymous ATM mutations in 3 of 3 breast cancer patients showed subcutaneous reactions following radiotherapy (209). However, many other such studies have failed to correlate the ATM mutation status with the occurrence of damage of normal tissue (204).

In breast cancer, the risk to close relatives of an affected woman, averaged across all ages, is about twofold. Most of the familial risk is probably genetic in origin. The risk is about the same for the mother, sisters or daughters of a case, suggesting dominant rather than recessive effects. Large population-based studies indicate that only 15-20% of overall familial risk is attributable to mutations in BRCA1 and BRCA2. The possibilities for the remaining 80% are some combination of a small number of moderately strong genes, and a larger number (possibly a hundred or more) of weaker genes. If moderately strong genes exist, it should in principle be possible to identify them by linkage in families. The weaker genes will not, on the whole, result in multiple case families and so must be sought by a different approach. Some of the possible unidentified familial genes include genes encoding steroid hormone receptors and paracrine growth factors, and genes involved in metabolism of exogenous chemicals or in DNA repair.

The variant alleles are associated with risks of around 1.5-fold and are predicted to account for only a few percent of breast cancer incidence. Collectively they account for only a very small fraction of the familial risk. Besides, one-third of all breast cancers that appear to run in families do not illustrate linkage to either BRCA1 or BRCA2, indicating the involvement of one or more unknown genes. Almost certainly there are many more genes to be identified, which together will account for a much higher fraction of cancer incidence than the genes in the inherited cancer syndromes (29, 131).

The identification of these genes will be greatly accelerated by the data from the Human Genome Project. The search relies on cataloguing the DNA sequence variation within the population, and showing that particular variants are significantly associated either with disease susceptibility or with some other aspect of disease phenotype such as treatment response or survival.

Ionizing Radiation and Breast Cancer

It is well known that exposure to ionizing radiation increases the risk for cancer development. One of the most-studied cancer risks after exposure to ionizing radiation is that for female breast cancer. Epidemiological studies of the Japanese atomic bomb survivors, ex-tuberculosis patients who underwent repeated fluoroscopy, Hodgkin's disease patients treated with radiation therapy, and women treated for mastitis or benign breast disease have clearly demonstrated a carcinogenic risk posed by radiation exposure. In these populations, the dose was delivered as a single exposure (atomic bomb survivors) or in multiple fractions ranging from a few cGy (fluoroscopy data) to several Gy (mastitis, Hodgkin's disease). In all these studies the carcinogenic effect was found to be directly proportional to dose with dose rate having little influence on breast cancer risk. What is also apparent from these studies is that radiation

exposure at any age increases breast cancer risks but that risks tend to decrease with increasing age at exposure so that exposures after age 50 carry a lower risk than exposures earlier in life. Regardless of the age of exposure, risks continue to be elevated throughout the remainder of a woman's life, with the largest excess rates occurring late in life, at ages similar to those at which breast cancers are seen in the absence of exposure (141).

Ionizing radiation is one of the primary modalities used in the management of cancer. While the use of medical radiation in cancer therapy has undoubtedly prolonged and saved the lives of many, it is not without side effects, including the development of cancer. Radiation treatments for breast cancer have been linked to increased risks of secondary breast cancers in the contralateral breast but only among women exposed prior to the age of 45 years. No risk was apparent when exposures occurred after the menopausal years (81). There has been increased interest in the possibility that women who are heterozygous for the ATM gene are at an increased risk of breast cancer and at an enhanced risk of radiation-induced breast cancer (88). Ongoing studies of women who develop contralateral breast cancer after radiotherapy will be able to evaluate the hypothesis that radiation is inducing secondary breast cancers (20, 22, 81).

Early detection of mammary tumors, however, reduces mortality considerably. At present, mammography is the most effective method for the early detection of malignant mammary lesions, resulting in a reduction of mortality of 20-40%. Consequently, the mammary glands of many women will be exposed regularly to low doses of radiation with typical average glandular does of 1-4 mGy per mammogram, yielding a total dose of approximately 0.1 Gy over a 20-year screening period. As the effects of such small doses are unknown, estimates of the risk of breast cancer due to mammography are extrapolated from higher doses, assuming that risk decreases linearly with decreasing total dose (15, 16). Randomized controlled trials of

mammographic screening for breast cancer prove that the intervention of offering screening reduces mortality from the disease by 20–30% (206).

Radiation Effects

Since a single dose of ionizing radiation is sufficient to increase cancer risks, ionizing radiation is classified as a complete carcinogen, able to both initiate and promote neoplastic progression. It is also a known carcinogen of the human breast and rodent mammary glands (13). While biological effects of exposure to IR can be measured and have been reported, the immediate effects of IR on the cellular functions that may be responsible for transformation of normal cells to malignant cells are not known. Depending on the dose, kind of radiation, and observed endpoint, the biological effects of radiation can differ widely. The biological effects of radiation can be divided into two general categories, stochastic and deterministic, or nonstochastic. As the name implies, stochastic effects are those that occur in a probabilistic manner, which are all or none effects that do not display an increase in severity. Cancer is one example. If a large population is exposed to a significant amount of radiation, then an elevated incidence of cancer can be expected. Although the expected incidence of cancer increases with dose, the severity of the disease in a stricken individual is not a function of dose. In contrast, deterministic effects are those that show a clear causal relationship between dose and effect in a given individual. There is an increase in incidence and severity with increasing dose. Usually there is a threshold below which no effect is observed, and the severity increases with dose. Skin reddening is an example of a deterministic effect of radiation (148, 159).

The risk of getting cancer from radiation depends on many factors, such as the dose and how it is administered over time, the site, and a person's age, sex, and genetic background. Additional factors, such as exposure to other carcinogens and promoters, are also important.

Cancer causes almost 20% of all deaths in the United States (81). The relatively small contribution made by low levels of radiation to this large total is not statistically evident in epidemiological studies. Radiogenic cancers are not distinguishable from others. Thus, cancer risk at low doses can only be estimated by extrapolating from human data at high doses, where excess incidence of cancer is evident (148). There are currently two theories on how radiation damages a cell, the "one hit" theory and the single cell mutation theory. In the one hit model only one cell is believed to be damaged by radiation and it sends out damage signals to surrounding cells. The single cell mutation theory of radiation-induced cancer suggests that the tumor arises through stages of initiation, promotion and progression. The initial mutation caused by radiation is a rare and stochastic event. In contrast, a genomic instability theory would suggest a common event cause, with gene mutation occurring frequently, but being controlled through cell–cell communication-mediated processes involving bystander signals and responses. Only if the epigenetic control breaks down does progression to cancer occur. This model also allows reversion after radiation exposure, not available in the initiation, promotion and progression model. Bystander effects and genomic instability are both induced at very low doses (207).

Radiation-induced Damage

Stress induced by ionizing radiation has been shown to induce different patterns of DNA damage in mammalian cells, such as single-strand breaks, double-strand breaks and base damage (53). The cellular responses to ionizing radiation include DNA repair, cell cycle control, signal transduction, apoptosis, mutagenesis and perturbed growth (3, 73). Ionizing radiation, such as gamma rays, cause cell cycle delays including G1-phase arrest, G2-phase delay, and G2-phase arrest. An example of a direct effect is a strand break in the DNA caused by an ionization in the

molecule itself. An example of an indirect effect is a strand break that results when an OH radical attacks a DNA sugar at a later time (34).

One of the main forms of damage inflicted by ionizing radiation is a DNA double strand break in the cell. DNA double strand breaks are generated when the two complementary strands of the DNA double helix are broken simultaneously at sites that are sufficiently close to one another that base-pairing and chromatin structure are insufficient to keep the two DNA ends juxtaposed. As a consequence, the two DNA ends generated by a double strand break are liable to become physically dissociated from one another, making ensuing repair difficult to perform and providing the opportunity for inappropriate recombination with other sites in the genome. Another roadblock in the rapid and error-free repair of double strand breaks is the fact that the DNA termini have often sustained base damage causing the double strand break to not be able to re-ligate until processing by DNA polymerase or nucleases has occurred (79).

Double strand breaks are induced deliberately in the V(D)J recombination pathway to provide the basis for the antigen-binding diversity of the immunoglobulin and T-cell receptor proteins. In this pathway, DNA double strand breaks are generated at specific loci by a site-specific nuclease composed of the RAG1 and RAG2 proteins, and the double strand breaks are subsequently repaired by proteins that also function in the repair of double strand breaks that have been generated by mutagenic agents. Double strand breaks are potent inducers of mutations and cell death. There is experimental evidence of a causal link between the generation of double strand breaks and the induction of mutations and chromosomal translocations with tumorigenic potential. Many cancers of lymphoid origin bear oncogenic chromosomal rearrangements that have arisen due to a defective double strand break repair of V(D)J recombination intermediates. A classical example of this is the B-cell malignancy, Burkitt's

lymphoma, where the *c-MYC* gene is often juxtaposed by genome rearrangement to the immunoglobulin heavy-chain genes. The loss or amplification of chromosomal material is characteristic of many cancer cells and is associated with the inactivation of tumor suppressor loci and activation of protooncogenes. Mutations in the factors involved in double strand break signaling and repair lead to increased predisposition to cancer. Defects in cellular responses to double strand breaks may be a frequent initiating event of carcinogenesis (79).

The cellular response to double strand breaks is to activate and induce the levels of DNA repair proteins, which are then recruited to the DNA lesion for repair. Dividing cells respond to DNA double strand breaks by slowing down their progression through the cell cycle. This will provide the cell enough time to allow DNA repair to occur before the lesions are replicated by DNA polymerase. DNA double strand breaks present in G_2-phase prevent entry into mitosis, thereby preventing the mis-segregation of chromosomal fragments during cytokinesis (79).

There are two main pathways for DNA double strand break repair: homologous recombination (HR) and non-homologous end-joining (NHEJ). During HR, the damaged chromosome enters into synapsis with, and retrieves genetic information from a homologous undamaged DNA molecule. During NHEJ, ligation of two DNA double strand breaks occurs without the requirement of homologous sequences between the two (79).

Reports have shown the increased frequency of giant cells, apoptosis, and micronuclei in the progeny of surviving cells (210), indicating that one of the causes of radiation-induced genomic instability is delayed cell death. Delayed chromosomal instability, as evidenced by the induction of nonclonal chromosome aberrations, has also been reported to be transmitted through many generations after irradiation. Observed chromosome aberrations include chromosome breakage-type aberrations and dicentric chromosomes resulting from chromosome fusions

arising from telomeric or interstitial telomeric sequences. Delayed mutagenesis has been examined in the HPRT locus, and increased mutation frequency was detected in the progeny of surviving cells.

Radiation–induced cancer in humans has long latent periods; 10 years for leukemia and over 30 years for solid tumors (208). Thus, there is a long period between radiation exposure and the appearance of tumors. This implies that radiation-induced mutations, due to gene mutations and/or chromosomal damage that can be detected within 24 hrs of radiation exposure are not directly responsible for initiating carcinogenesis in normal human cells. However, such mutations induce genetic instability that make cells more sensitive to accumulation of additional genetic abnormalities caused by exposure to additional radiation doses, chemical mutagens and carcinogens, tumor promoters, oncogenic viruses, or their combinations. Cells may continue to carry genetic abnormalities for a long time until the expression of genes regulating differentiation is altered. This could lead to a cell immortalization phase, the first step in carcinogenesis (208). This is why genes that are unknown to respond to radiation are important to identify.

Known Radiation Response Genes

Rapid alterations in gene expression can occur after cellular exposure to ionizing radiation. Several of such radiation responsive genes are known proto-oncogenes or tumor suppressor genes. Thus identification of new genes participating in the DNA damage response can be particularly important in the unraveling of the carcinogenic process (121). The discovery of radiation-responsive genes provides a starting point for elucidating the molecular processes associated with radiation-induced carcinogenesis. Both constitutive and inducible genes are involved in these responses. *c-fos* is an early response gene that has been shown to be regulated,

either up or down, by ionizing radiation in a pattern that is dependent on the cell and tissue type (3). Genes involved in determining the biological outcome after radiation exposure often act in a tissue- and cell-specific manner. Recent studies have identified hundreds of genes, involved in many different cellular processes, induced in mammalian cells by ionizing radiation. While some genes, such as those with roles in DNA repair or virus activation, may represent responses specific for genotoxic stress, many appear to be more general responses to cell and tissue injury (7). The majority of radiation-responsive genes that have been identified are involved in general metabolic pathways. Less than 10% participate in DNA metabolism or repair. In several studies, subtractive hybridization and differential display of mRNA have been employed to identify new genes activated by ionizing radiation. A recently identified gene, *Csa-19*, has been found by unbiased two-gel cDNA library screening to be down-regulated in cultured human breast carcinoma cells by administration of ionizing radiation (11). Other examples of radiation-responsive genes include *p53*, *c-jun*, *Egr-1*, *TNFα*, *IL-1*, *protein kinase C*, *KIN*, *gadd45*, *WAF-1Waf1/Cip1*, *BAX*, *ATF3* and *mdm2* (53, 83). Tumor suppressor *p53* may play a vital role in the cellular response to radiation (96).

ATM is a known radiation-responsive gene. The ataxia telangiectasia mutated (ATM) gene encodes a protein (ATM) with serine/threonine kinase activity. DNA-double strand breaks are known to increase its kinase activity. AT heterozygous individuals have been reported to have an increased risk for breast cancer. Individuals with AT syndrome are affected by Purkinje cell degeneration, sensitivity to γ radiation, immune disorders, telangiectasias and cancer predisposition, particularly lymphomas. ATM is activated in cells treated with gamma radiation. It plays a major role in checkpoint activation following DNA damage by gamma radiation and oxidative stress, it seems to be less important for checkpoint responses after UV irradiation (71).

ATM was found to be down-regulated in response to radiation in XRS-5 cells deficient in the DNA end-binding protein Ku80 (211). This gene expression assay for radiation responsive genes was performed utilizing microarray technology. The microarray membranes contained 600 cDNAs. Compared to other techniques, such as gene trapping, ATM might not be found to be radiation responsive due to the fact that the retroviral vector is integrating randomly into the genome. This random integration could potentially lead to particular genes being more accessible to trapping and more likely to be trapped than others. Another possibility is that ATM could be located in a region of chromatin that is tightly packed and not as accessible to retroviral integration.

Goal of study

The purpose of this study is to identify novel genes affected by gamma irradiation that could potentially have a role in the initiation and progression of breast cancer. It is imperative to try to identify novel genes affected by radiation since numerous downstream pathways could be effected once the radiation response gene becomes mutated. This could potentially cause detrimental effects in the cell leading to cell death or transformation. Radiation is routinely utilized in cancer treatment and management and its effects on the cell need to be characterized in order to understand it properly. The main form of damage produced by ionizing radiation is a double strand break in the DNA. This cellular damage can eventually lead to downstream genomic instability and carcinogenesis. A disadvantage of looking at changes in gene expression following ionizing radiation treatment is that the effects of the radiation could be occurring many generations later due to genomic instability and the changes in gene expression picked up by this system may not be relevant in affecting cell survival.

The gene trapping system used in this work allows for the identification of novel radiation response genes regardless of their expression level. This technique is ideal for identifying novel genes that are yet uncharacterized, as compared with microarray technology that utilizes known gene sequences. The gene trapping system will identify genes that are transcriptionally affected by radiation treatment. It will not, however, identify genes such as ATM that are modified post-translationally following radiation treatment. This system allows for the identification of genes with changes in their gene expression that are being modified shortly after radiation treatment. I chose to focus on the modification of gene transcription via radiation since the double strand breaks and chromosomal translocations radiation produces are known to leads to genomic instability, which could potentially cause carcinogenesis further down the road.

CHAPTER 2-METHODS BACKGROUND

Gene Trapping

During the early 1990s, alternative strategies for identifying gene function were explored. The possibility of whether the directed screening of libraries of random genetic mutations could be used to identify classes of genes that were relevant to particular biological systems was investigated. This approach was termed "gene trapping". To identify and study novel genes, an excellent experimental strategy is the gene trap. Gene trapping was designed to identify novel, developmentally regulated genes. The principle behind gene trapping is essentially the random insertion of a DNA vector that is designed to signal its presence via the activation of a reporter gene, which both mimics the expression of the endogenous gene and potentially mutates it as well (97). Gene trapping allows the simultaneous identification, sequencing, in vivo expression analysis, and phenotyping of the genes of interest. Gene trapping is an insertional mutagenesis technique in which an exogenous DNA vector integrates randomly into the genome. The vectors used in this approach typically include a splice acceptor site immediately upstream of a promoterless reporter gene (GFP) and the selectable *neo* gene driven by an autonomous promoter. Reporter gene activity accurately reflects the activity of the endogenous gene into which the vector has integrated. An important consideration of a gene trap approach is that the gene trap insertion may be mutagenic. The endogenous locus is usually (but not always) inactivated by vector integration, leading to a loss of function (162, 178).

The most notable limitation of these early vectors is that they target only those genes that are actively transcribed. This restricts the set of genes that can be mutated in any particular cell type. Thus, when transcription in cells is poor or absent drug selection might not select all clones that contain trapped genes (97).

There are two main methods used for the introduction of gene-traps into cells, retroviral vectors and electroporation of naked DNA. Electroporation, which is the method of choice for DNA transformation for homologous recombination, is frequently used to insert gene-trap constructs, but is relatively inefficient. Electroporation facilitates the entry of foreign DNA into living cells by temporarily permeabilizing cell membranes. Although this technique has been used extensively, its efficiency is quite low. Plasmid-based vectors introduced by electroporation most often show random genomic integration. In contrast to electroporated DNA, retroviral vectors provide a much higher efficiency of gene transfer and integrate as an intact single copy. (47). Retroviruses have a propensity for inserting into the 5' portion of a gene, including the 5' untranslated region and the first intron. Vector insertion immediately downstream of the initiation codon generally produces null mutations, which indicates that using retroviruses to introduce gene-trap vectors might give a higher percentage of null mutations, whereas electroporation might be better for generating allelic series. The main advantage of retroviral infection is that it ensures the integration of a single copy of the entire vector, although electroporation strategies can be optimized such that multiple insertions occur in less than 20% of cell lines. Tandem insertions into the same locus are problematic because they can cause ectopic reporter expression and aberrant splicing, which can interfere with the cloning of the trapped fusion transcript. Furthermore, electroporated plasmid DNA is often digested by exonuclease, making the cloning of insertion sites by reverse transcription PCR problematic.

Electroporation can often produce multiple integration sites of the vector and the insertion can cause deletions and rearrangements of the DNA that could affect genes other than the one trapped. In contrast, proviral DNA always retains the long-terminal repeat sequences, which allows the high-throughput cloning of retroviral insertion sites. Retroviral vectors do have three advantages, though: their packaging size is limiting, they can induce retroviral-mediated gene silencing and ectopic reporter gene expression can occur (164).

There are four main types of entrapment vectors: enhancer-trap vectors, gene-trap vectors, promoter trap vectors, and polyA trap vectors (194). Vectors developed for use in mammalian cells, referred to as "gene traps", insert a reporter gene into mostly random chromosomal sites, including transcriptionally active regions. By selecting for gene expression, recombinants are obtained in which the reporter gene is fused to the regulatory elements of endogenous genes. Transcripts generated by these fusions faithfully reflect the activity of the tagged cellular gene and thus provide an effective means to study the expression of genes in their normal chromosomal location. Moreover, gene traps are highly and provide molecular tags to clone any gene whose function is linked to an observable phenotype.

Enhancer Trap Vectors

In contrast to the rational "one gene at a time" approach, enhancer-trapping methods evolved to probe the entire genome simultaneously. Enhancer-trap vectors must be integrated near a cellular enhancer to activate the reporter gene that is fused to a minimal promoter that requires the vector to insert near to a cis-acting enhancer element to produce expression of a lacZ reporter gene. The trapping principle is based on random insertional mutagenesis of a vector ideally containing a promoterless reporter gene. Chromosomal integration of a transposable element tagged the integration site and often mutated the gene into which it inserted. Mutants

resulting from this approach could be selected using genetic screens. Because the insertion site was tagged, the mutated genes could be readily identified, which made this approach more efficient than traditional chemical mutagenesis. In addition, enhancer traps are not mutagenic in eukaryotes (52).

Gene Trap Vectors

The gene trap approach is based on a class of reporter vectors that have been designed containing a splice acceptor site upstream of the β-galactosidase (lacZ) gene and the neomycin resistance gene (neo). Integration of these vectors into a genomic locus downstream of a functional promoter results in the generation of a fusion transcript between the endogenous gene and the lacZ gene. When this vector integrates into an intron, a spliced fusion transcript between the reporter and endogenous gene is generated under the control of endogenous transcriptional regulatory elements of the trapped gene. If the coding region of the reporter remains in-frame with the 5' coding sequence of the trapped gene, translation of the fusion transcript creates a fusion protein between the N-terminus of the gene product and the reporter (66). Fusion transcripts from insertion of these vectors mimic endogenous gene expression at the insertion locus. This expression can be monitored by visualizing the lacZ activity. The tagged genes can then be identified through the use of an anchored PCR procedure. Usually the gene trap vectors will also act as insertional mutagens, disrupting the endogenous gene function. The gene trap method has been demonstrated to effectively knock out genes and that a 5' RACE protocol can be used to obtain sequence information about the disrupted genes (43).

The rationale behind the use of reporter constructs is to tag and detect cis-regulatory sequences by locating the reporter gene within an endogenous gene. In certain integrations, the reporter construct can also disrupt endogenous gene function and can therefore act as a mutagen.

In some cases the full-length gene may still be expressed as a result of differential or inefficient splicing, or an active fragment could be expressed when the trapping element inserts in the 3' region of a gene (43).

In eukaryotic cells, nuclear pre-mRNA introns are excised by a large ribonucleoprotein complex known as the spliceosome, which recognizes sites at the 5' and 3' ends of the intron (the donor and acceptor splice sites, respectively). A minimal splice acceptor consists of a branch point sequence followed by a polypyrimidine tract and the actual splice site, which is recognized by the spliceosome. Splice acceptor vectors must integrate within an intron, and consequently favor genes that contain longer introns because they represent a larger target for insertion. Proper integration and splicing will generate a fusion protein with the trapped locus. The fusion transcript will ideally be prematurely terminated at the poly(A) site of the selection-reporter cassette. In order to identify genes independent of reading frames, an internal ribosome entry site (IRES) from the encephalomyocarditis virus is utilized at the 3' end of the splice acceptor (50). With its dependence on splice acceptors, it would seem logical to develop a gene-trap system that functions independently of the splicing machinery, which removes introns and produces mRNA containing only exons (212). Since the reporter gene of this vector is only expressed when it is integrated into an exon, resistance to neomycin will select only for those cells in which retroviral integration has occurred into an expressed gene (97).

Promoter Trap Vectors

In general, promoter-trap vectors consist of a promoterless reporter gene and selectable marker that are often the same. Reporter expression and mutagenesis occur when the vector inserts into an exon to generate a fusion transcript that comprises upstream endogenous exonic sequence and the reporter gene. Because transcription of the reporter requires that the vector

inserts into an exon, the mutagenicity rate of promoter-trap vectors should be very high, as indicated by there having been very few reports of hypomorphic mutations from promoter traps. Also, since the insertion site is in transcribed DNA, cloning the insertion site will identify the disrupted gene. However, the frequency with which promoter-trap vectors insert into exons is exceedingly low, at least 200-fold lower than that of an enhancer trap. Therefore, promoter-trap vectors generally contain a selectable marker, such as the neomycin resistance gene or the β-galactosidase-Neo (β-geo) fusion marker, as a reporter, so that only clones that contain the vector insertions can be selected. This approach, however, means that only insertions into those genes that are transcriptionally active in the clones will be selected for (164). To identify the trapped gene, DNA flanking the retrovirus integration site is subcloned and sequenced using genomic libraries, PCR methods or a shuttle vector. One drawback of promoter traps is that the sequence tag may or may not contain exonic sequence thus compromising the ability to identify ESTs or cDNAs that match the trapped gene. Promoter traps have been used extensively in ES cells and have proven to be a rapid method for making and identifying mutations with sequence tags (197).

Poly A Trap Vectors

A limitation of the commonly used splice acceptor vector type is that mostly expressed genes can be identified. In contrast, the poly(A) trap vector type allows the identification of all intron-containing genes independent of their expression status. The simplest poly(A) trap vector contains a promoter-driven selector gene without a poly(A) sequence. Thus, drug resistance is only obtained after successfully capturing a poly(A) signal in the genome. Furthermore, poly(A) trap vectors have been developed that contain an additional splice donor sequence as well. The latest poly(A) trap vectors contain a reporter gene with a splice acceptor in front, followed by a

selector gene driven by a strong promoter fused to a splice donor. The only known genes that are not polyadenylated in metazoan organisms are the major histone genes. Therefore, poly(A) trap vectors will basically allow the entrapment of most genes (50).

A poly(A) trap vector design provides an internal constitutive promoter that drives the expression of a selectable marker sequence with a 3' splice donor in place of a poly(A) signal. This vector only confers resistance in the clone when the vector traps an endogenous splice acceptor and a polyadenylation signal. The higher rate of "novel" sequence recovery reported may suggest that these vectors tag genes that are underrepresented in the EST databases or that tagging occurs in genes that are not normally transcribed, such as pseudogenes, line elements, and interspersed repeats. Because these poly(A) trap vectors do not require integration into an actively transcribed gene for selection, they can be useful for identifying genes that are inactive at the time (43).

Poly(A) trap vectors require a spliced poly(A) signal from an endogenous gene in order to generate stable neo mRNA, and in turn, neomycin-resistant clones are produced. Only insertions in genes should generate neomycin-resistant clones, and background intergenic insertions should be lost. Several termination codons follow the selectable marker gene in order to prevent the translation of the 3' trapped exons. Because the stability of many transcripts is linked to their translation and their endogenous 3' untranslated regions, poly(A) trap vectors work best when inserted into the 3' end of the gene. This strategy might bias selection towards 3' insertion events, resulting in fewer null mutations. With the known data on poly(A) trap vectors, it appears as if this bias is unfounded.

A widely used reporter gene for high-throughput screening is the bacterial β-galactosidase (β-gal or lacZ), which has the advantage of being well characterized, stable,

nontoxic, and inexpensive. A limitation of β-gal is some endogenous activity in the mouse. An alternative reporter to β-gal is the bacterial β-lactamase, for which no endogenous activity has been described in the mouse. Human placental alkaline phosphatase (AP) has also been successfully used in gene-trap vector systems as a reporter. The jellyfish green fluorescent protein (GFP), which is available in multiple forms, has the advantage that no substrate is needed and no endogenous activity is observed in murine cells. In addition, antibodies for GFP detection are commercially available. Recently, evidence of GFP toxicity in mouse cells has been reported when highly expressed (34).

Resistance markers in gene-trap vectors are traditionally neomycin phosphotransferase or lacZ/neo fusions (βgeo), but other resistance markers like phleomycin have been used successfully. LacZ/hygromycin cassettes have been utilized in order to trap genes in transfected cells and cells carrying a targeted mutation. Poly(A) trap vectors are available with a neomycin or puromycin resistance cassette (50). The expression of the neomycin-selected marker gene confers the resistance to the selective drug G418, and can therefore be used to facilitate the identification of clones in which gene trapping has occurred. In vectors in which β-geo is the reporter gene, the combined β-galactosidase and neomycin-resistance genes ensure that clones that are selected by resistance to G418 will also express the marker gene activity. β-galactosidase activity can be detected using the chromogenic substrate X-gal, which yields a blue color. Selection on the basis of G418 has been shown to be more sensitive than X-gal staining. Consequently, at least some of the clones that are negative for X-gal may contain weakly expressed genes leading to false negatives. Thus, the combined β-geo reporter gene increases the efficiency of gene trapping and ensures that genes that are expressed at low levels can also be trapped (97).

Gene Trapping Advantages and Disadvantages

Each gene-trap vector has its own inherent advantages and disadvantages. A limitation of both the splice acceptor-type as well as the exon trap vectors is mainly the dependency on integration in active genes. IRES vectors without a splice acceptor that have integrated into intronic regions sometimes fail to be mutagenic due to splicing around the integration site for unknown reasons. In contrast to the splice acceptor-type vectors, it has been postulated that the integration of ubiquitously active transcriptional units could interfere with neighboring genes and thereby complicate the interpretation of the observed phenotypes. It is advised to remove the promoter element of the gene-trap vector by utilizing the Cre or Flp recombinase systems.

One of the major issues concerning the poly(A) trap approach is whether or not the sequences being discovered are actually coding for proteins. The integration in the vicinity of canonical and noncanonical poly(A) sites, which are not related to genes, appear to be a hypothetical disadvantage of the poly(A) trap approach. This integration may lead to a high number of false positive gene-trap clones. The consensus sequence for a canonical poly(A) site is AATAAA, but variations of GU or U-rich sequences are productive and may also lead to proper polyadenylation. There is more to proper polyadenylation than just the poly(A) site, though. The poly(A) site needs to be followed by a more complex signal. RNA polymerase also needs to be released from the DNA template and the transcript needs to be cleaved in order to polyadenylated. This limitation of poly(A) trapping has recently been addressed by the incorporation of mRNA instability elements 3' of the splice donor site which will lead to the degradation of the polyadenylated transcripts that have not been spliced (50).

Two parts of the gene-trap vector mainly influence its mutagenicity with respect to interrupting the endogenous transcript. One, the relative strength of the splice acceptor is critical

for mutagenicity and two, the effective termination of the fusion transcript at the introduced poly(A) signal is also required for mutagenicity. All vectors have been shown to some extent to "splice-around" the inserted vector. Thus, wild-type gene product can be made, and splicing-around leads to variable hypomorphic mutations as compared to null mutants (50).

Gene-trapping methods are proving to be essential in the functional analysis of the genome. Versatility is one of the many strong points of this technology, which is helpful in analyzing complex genomes. Gene-trapping is being constantly developed to form such new traps as "secretory-trap" and "induction-trap". These new traps will help to reveal proteins important for communication between cells. Techniques that can specifically trap and monitor genes expressed at low levels will be valuable for "genome closure" because these genes are currently underrepresented in the compiled EST databases. Reporter cell lines generated by some trapping methods will be useful as tools for investigating the signal transduction pathways that regulate expression of the trapped genes. This will ultimately provide another layer of functional genomic information (43).

Gene trapping is an exceptional tool for gene discovery because genes are trapped regardless of their transcriptional activity. However, as also seen in homologous recombination experiments, there are "hot" and "cold" genomic spots for gene-trap vector insertions (164). The number of genes that can be mutated with gene traps is not known. Conventional gene trapping is good at producing null (or "loss of function") mutations in trapped genes. One possible shortcoming is that this is a sledgehammer approach, thus, it might be more useful in some cases to be able to produce less severe mutations (97). A shortcoming of gene-trap technology is its relative inability to induce subtle or gain-of-function mutations. Because of these limitations, gene-trap technology does not intend to supercede targeted mutations and phenotype-driven

screens. Rather, it is intended to integrate rapid gene discovery, mapping and mutagenesis by minimizing the effort required for cloning. A further problem with gene trapping is that some genes may be more likely to be trapped than others. The use of one type of vector in a study could limit the number of genes that are likely to be targeted; therefore, the use of several vectors is recommended (162, 194).

Real-time PCR analysis of trapped gene(s) expression

Real-time PCR

The development of fluorogenic probes made it possible to eliminate post-PCR processing for the analysis of probe degradation (217). The probe is an oligonucleotide with both a reporter fluorescent dye and a quencher dye attached. While the probe is intact, the proximity of the quencher greatly reduces the fluorescence emitted by the reporter dye by Förster resonance energy transfer (FRET). The reporter dye is located at the 5' end of the probe and the quencher dye is located at the 3' end of the probe. Quenchers accept energy from a fluorophore and dissipate it by one of two mechanisms. First, quenching can occur proximally due to the close contact between fluorophore and quencher and the energy is dissipated as heat (collisional quenching). Second, fluorophore and quencher may be separated by up to 40 bp and quenching is achieved by fluorescence resonance energy transfer (FRET) where the fluorophore donor transfers its energy to the quencher acceptor, which releases the energy as light of a higher wavelength. However, because FRET is distance-dependent, it is important to design dual label probes to be as short as possible (30).

The advantage of fluorogenic probes over DNA binding dyes is that specific hybridization between the probe and target is required to generate fluorescent signal. Thus, with fluorogenic probes, non-specific amplification due to mis-priming or a primer-dimer artifact does

not generate a signal. Another advantage of fluorogenic probes is that they can be labeled with different, distinguishable reporter dyes. The most commonly used fluorescent dyes are: FAM (absorption maximum 518 nm), TET (538 nm) and Cy5 (667 nm) (30).

The RT-PCR reaction exploits the 5′ nuclease activity of the AmpliTaq Gold enzyme to cleave a TaqMan probe during PCR. The TaqMan probe contains a reporter dye at the 5′ end of the probe and a quencher dye at the 3′ end of the probe. During the reaction, cleavage of the probe separates the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter. The accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye. Figure 2.1 shows the forklike-structure-dependent, polymerization-associated, 5′ to 3′ nuclease activity of AmpliTaq Gold DNA Polymerase during PCR.

Reverse transcription combined with the polymerase chain reaction (RT-PCR) has proven to be a powerful method to quantify gene expression. Real-time PCR technology has been adapted to perform quantitative RT-PCR. Real-time polymerase chain reaction is the ability to monitor the progress of the PCR as it occurs, as in "real-time". Data are collected throughout the PCR process, rather than at the end of the PCR reaction. This completely changes the way the DNA and RNA can be quantitated. In real- time PCR, reactions are characterized by the point in time during cycling when amplification of a target is first detected rather than the amount of target accumulated after a fixed number of cycles. The higher starting copy number of the nucleic acid target, the sooner a significant increase in fluorescence is observed. In contrast, an endpoint assay measures the amount of accumulated PCR product at the end of the PCR cycle (54).

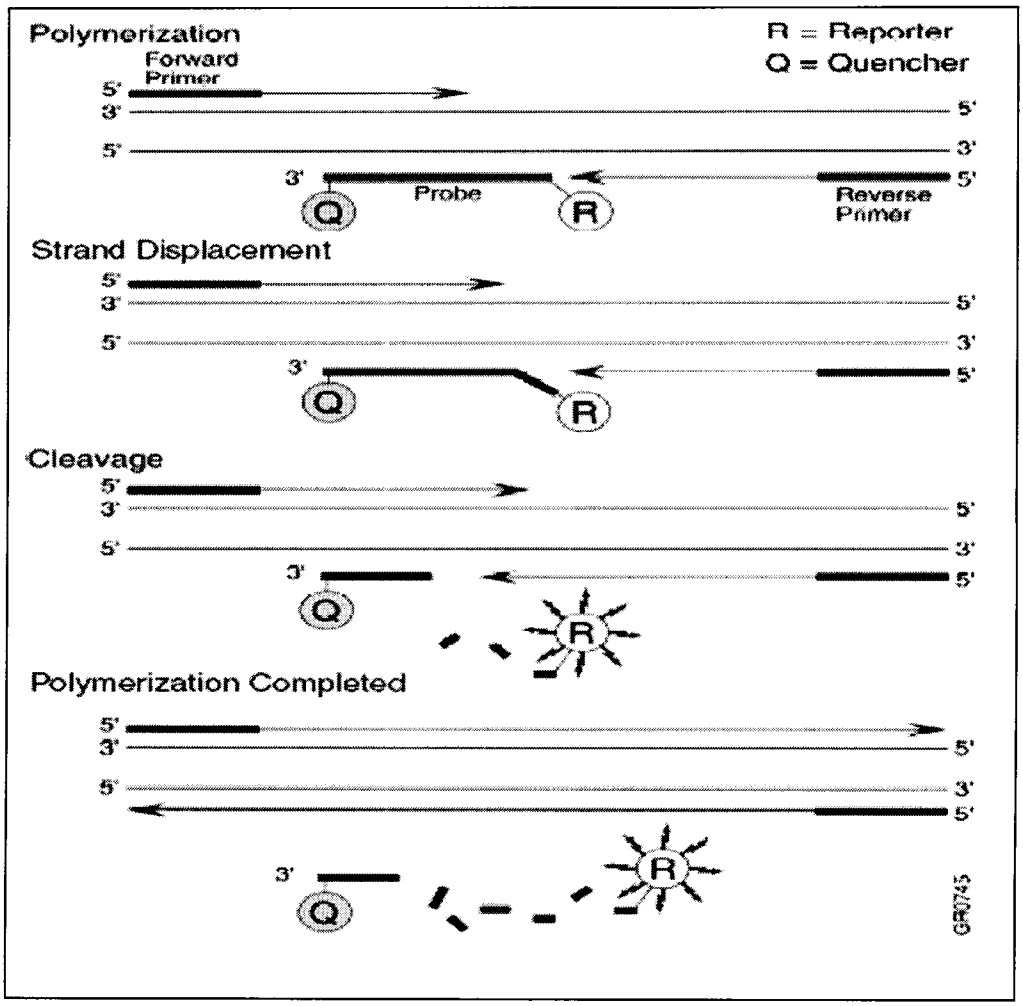


Figure 2.1. Basis of the 5' Nuclease Assay in a real-time PCR experiment.
Source: Applied Biosystems TaqMan Gold RT-PCR Kit Protocol.

The RT step is critical for sensitive and accurate quantification and the amount of cDNA produced by the reverse transcriptase must accurately represent RNA input amounts. The initial step in RT-PCR is the production of a single-strand complementary DNA copy (cDNA) of the RNA through the action of the retroviral enzyme, reverse transcriptase. An oligonucleotide primer is required to initiate cDNA synthesis. The primer anneals to the RNA, and the cDNA is extended toward the 5' end of the mRNA through the RNA-dependent DNA polymerase activity of reverse transcriptase. Primers can either be gene-specific or nonspecific; both have advantages and disadvantages. Random hexamer primers contain all possible nucleotide combinations of a 6-base oligonucleotides and bind to all RNAs present. Similarly, oligonucleotides consisting solely of deoxythymidine residues [oligo (dT)] anneal to the polyadenylated 3' tail found on most mRNAs. Alternatively, a gene-specific primer can be used for the RT reaction. For some genes, especially rare messages, the use of sequence-specific primers increases specificity and decreases background associated with other types of primers. These gene-specific RT primers work well in conjunction with elevated RT reaction temperatures to eliminate random transcripts (54).

Following the RT reaction, the cDNA is amplified by PCR. PCR is usually carried out using an aliquot of the RT reaction or by adding the necessary PCR components directly to the RT reaction. PCR is generally a three-step process, with denaturation, annealing and elongation steps, with temperatures that vary and are subject to a number of considerations that should be determined empirically. The number of cycles depends on the amount of target present and the efficiency of the reaction.

Primer design and selection is of great importance in the PCR step (54). Primer design is typically done with software such as Primer Express oligo design software from Applied

Biosystems or Primer3, a free program from MIT. It is important to screen the genome database with the amplicon sequence to ensure no other sequences besides the target sequence are being amplified. The amplicon for the PCR product should be as small as reasonably possible, 70 – 150 bp for designs with hybridization probes. Typically three primer pairs are constructed to verify which will work best in the experiment. It is important to choose the primers after a probe has been designed. The primers should be designed as close as possible to the probe without overlapping it. The G-C content of the primers should be targeted to the 20 – 80% range. Runs of identical nucleotides, especially guanine, should be avoided in primer design. The melting temperature of the primers should fall within the range of 58 to 60°C. Finally, all five nucleotides at the 3' end of each primer should have no more than two G and/or C bases. Probe design is typically done in the same manner. The probe design follows the same guidelines as primer design except for the following: the probe should not contain a guanine nucleotide at the 5' end, selection should be based on the strand that gives the probe more Cs than Gs and the melting temperature of the probe should be targeted to the range of 68 to 70°C.

The final step in QRT-PCR is the detection and quantification of amplification products. Reaction products must be separated so that target and standard can be discriminated and quantified. The two broad classes of amplification product detection techniques are the traditional "end-point" measurements of product and "real-time" monitoring of product formation. Endpoint determinations analyze the reaction after it is completed, and real-time determinations monitor the reaction in the thermal cycler as it progresses. End-point product measurement can be accomplished through the use of fluorescent intercalating dyes or through the measurement of incorporated radioactivity by autoradiography or phosphor imaging. Hybridization-based protocols, such as Southern blots or fluorescence detection are also used. A

third type of end-point product measurement uses a solid-state approach in which a bound enzyme produces fluorescence or luminescence. The measurement of amplification products can also be done by HPLC or capillary electrophoresis.

During the exponential phase in real-time PCR experiments a fluorescence signal threshold is determined at which point all samples can be compared. This threshold is calculated as a function of the amount of background fluorescence and is plotted at a point in which the signal generated from a sample is significantly greater than the background fluorescence. Therefore, the fractional number of PCR cycles required to generate enough fluorescent signals to reach this threshold is defined as the cycle threshold, or C_T (59). The endpoint of real-time PCR analysis is the threshold cycle, or C_T. The C_T is determined from a log-linear plot of the PCR signal versus the cycle number. Thus, C_T is an exponential and not a linear term (106). These C_T values are directly proportionate to the amount of starting template and are the basis for calculating mRNA expression levels or DNA copy number measurements.

Quantifying real-time PCR results

Two different methods of analyzing data from real-time, quantitative PCR experiments exist: absolute quantification and relative quantification. When calculating the results of quantification assays, you can use either absolute or relative quantitation. Absolute quantification determines the input copy number of the transcript of interest, usually by relating the PCR signal to a standard curve. The absolute quantitation assay is used to quantitate unknown samples by interpolating their quantity from a standard curve. The standard curve method for absolute quantification is similar to the standard curve method for relative quantification, except the absolute quantities of the standards must first by known by some

independent means. Absolute quantification should be performed in situations where it is necessary to determine the absolute transcript copy number (106).

Relative quantification describes the change in expression of the target gene relative to some reference group such as an untreated control or a sample at time zero in a time course study. A relative quantification assay is used to analyze changes in gene expression in a given sample relative to another reference sample. For all experimental samples, the target quantity is determined from a standard curve and divided by the target quantity of the calibrator. Thus, the calibrator becomes the 1x sample, and all other quantities are expressed as an n-fold difference relative to the calibrator. Amplification of an endogenous control may be performed to standardize the amount of sample RNA or DNA added to a reaction. For the quantification of gene expression, researchers have used β-actin, glyceraldehydes-3-phosphate dehydrogenase (GAPDH), ribosomal RNA (rRNA), or other RNAs as an endogenous control. Because the sample quantity is divided by the calibrator quantity, the unit from the standard curve drops out. Thus, all that is required of the standards is that their relative dilutions are known. For relative quantification, this means that any stock RNA or DNA containing the appropriate target can be used to prepare standards.

Quantifying the relative changes in gene expression using real-time PCR requires certain equations, assumptions, and the testing of these assumptions to properly analyze the data. QRT-PCR is inherently an indirect method of measurement. For reliable quantitative information to be extracted from this variable system, consideration of mathematical models is necessary. In this regard, the RT and PCRs are entirely different types of enzymatic reactions and thus should be considered separately. The reverse transcription step of RT-PCR is very basic. There is no amplification, and the sole variable is the percentage of mRNA converted into cDNA. This can

be expressed as: [cDNA] = [RNA] x Efficiency. Efficiency is measured as the percentage of RNA transcribed into cDNA (54).

The comparative C_T method is similar to that of a standard curve method except that is uses the arithmetic formula $2^{-\Delta\Delta C}{}_T$ to achieve the same result for relative quantification. Using the $2^{-\Delta\Delta C}{}_T$ method, the data are presented as the fold change in gene expression normalized to an endogenous reference gene and relative to the untreated control. For the untreated control sample, $\Delta\Delta C_T$ equals zero and 2^0 equals one, so that the fold change in gene expression relative to the untreated control equals one. For the treated samples, evaluation of $2^{-\Delta\Delta C}{}_T$ indicates the fold change in gene expression relative to the untreated control. Similar analysis could be applied to study the time course of gene expression where the calibrator sample represents the amount of transcript that is expressed at time zero (106). For the comparative C_T method to be valid, the efficiency of the target amplification, the gene of interest, and the efficiency of the reference amplification, the endogenous control, must be approximately equal. A sensitive method for assessing if two amplicons have the same efficiency is to look at how ΔC_T varies with template dilution. If the efficiencies of the two amplicons are not equal, then the analysis may need to be performed via the absolute quantification method using standard curves. Alternatively, new primers may be designed or optimized to achieve a similar efficiency for the target and reference amplicons (106).

The purpose of the internal control gene is to normalize the PCRs for the amount of RNA added to the reverse transcription reactions. A standard housekeeping gene usually will suffice for use as an internal control gene. Each internal control gene should be properly validated prior to the experiment to ensure that the gene is unaffected by the experimental treatment being utilized. The simplest choice for a calibrator in a real-time PCR experiment is the untreated

control. Data from the real-time PCR experiments were analyzed using the following equation: $\Delta\Delta C_T = (C_{T,\ Target} - C_{T,\ GAPDH})_{Irradiated} - (C_{T,\ Target} - C_{T,\ GAPDH})_{Untreated}$. The mean and standard deviation are then determined from the triplicate samples at each time point and treatment. Using this analysis, the value of the mean fold change at time zero, or untreated, should be very close to one (106).

RNA Interference

RNAi

RNAi is an important biological mechanism in the regulation of gene expression in plants, fungi, and animals. It is a highly conserved mechanism throughout taxonomical groups (44). RNAi is closely linked to the posttranscriptional gene-silencing (PTGS) mechanism of cosuppression in plants and quelling in fungi. RNAi is critical in the viral defense and in transposon silencing, heterochromatin formation, RNA-dependent DNA and histone methylation in plants, and programmed DNA deletion in *Tetrahymena*. In addition to an antiviral function, RNAi is also thought to suppress the expression of potentially harmful segments of the genome, such as transposons, which might otherwise destabilize the genome by acting as insertional mutagens. The natural function of RNAi and cosuppression appears to be in the protection of the genome against invasion by mobile genetic elements such as transposons and viruses, which produce aberrant RNA or dsRNA in the host cell when activated. mRNA degradation of the transposons and viruses prevent this replication in the cell, although some viruses are able to overcome and prevent this process by expressing proteins that suppress the posttranscriptional gene-silencing mechanism in the cell.

Although the mechanisms of RNAi are not completely understood, it represents the end result of a multistep process comprising initiation and effector stages. On entering the cell, long

dsRNAs are first processed by the RNAse III enzyme Dicer. This functional dimer contains helicase, dsRNA binding, and PAZ (named after piwi, argonaute, and zwille proteins) domains. While the previous two domains are essential for dsRNA unwinding and mediation of protein-RNA interactions, the function of the PAZ domain, which is highly conserved among species, is not fully understood (44).

Dicer has been found to produce dsRNA fragments of 21 – 23 nucleotides in length that contain two 3'nucleotide end overhangs which are known as siRNAs. It has been reported recently that Dicer has functions other than just in the dsRNA cleavage that is required for siRNA-mediated RNAi in mammalian cells. RNAi is affected by a ribonuclease known as RISC (the RNA-induced silencing complex) which, guided by siRNA, recognizes mRNA containing a sequence homologous to the siRNA and cleaves it at a site approximately in the middle of the homologous region. Gene expression then becomes inactivated at the posttranscriptional level (44).

DsRNA triggers degradation of the homologous RNAs specifically in the regions of identity within the dsRNA. The dsRNA is processed to 21 – 23 nucleotide RNA fragments. RNA molecules of similar size have been found to accumulate in plant tissues that express posttranscriptional gene-silencing mechanisms. It has been suggested that the 21 – 23 nucleotide fragments of RNA are guides for recognition in the cell (46). RNA regulatory molecules either reduce or eliminate target gene expression by binding mRNA and targeting it for degradation by cellular enzymes. Genetic and biochemical data suggest that double-stranded siRNAs, 19 – 25 nucleotides in length, are produced from larger RNAs by the Dicer enzyme. Upon binding through base-pairing to target mRNA, siRNAs recruit Rnases to a protein complex called the RNA-induced silencing complex (RISC) that degrades the targeted sequence.

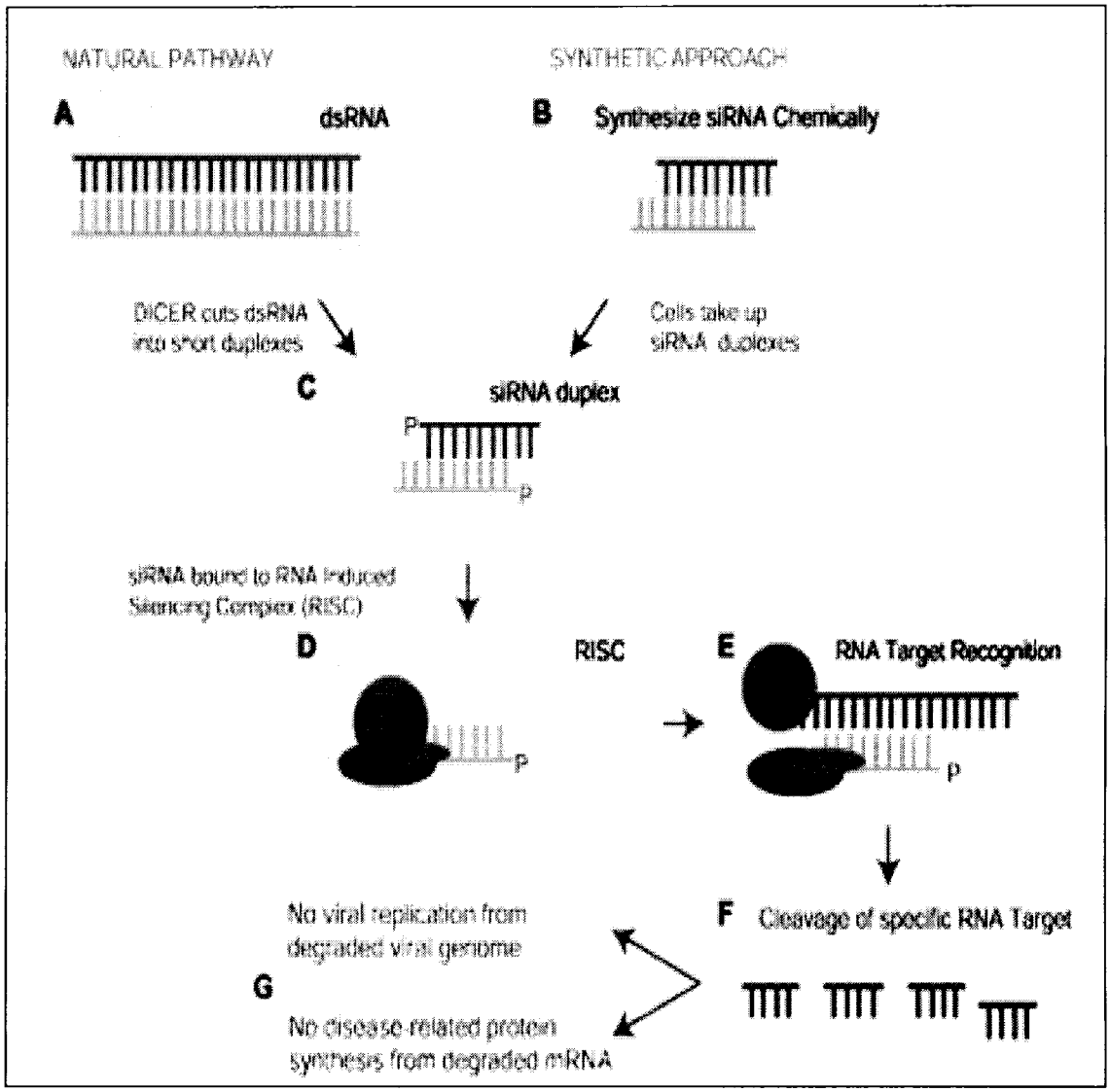


Figure 2.2. Mechanism of RNAi. RNA interference (RNAi) is a cellular mechanism that regulates the expression of genes and the replication of viruses. This mechanism is mediated by double-stranded small interfering RNA molecules (siRNA). Messenger RNA (mRNA) provides the means of implementing the set of instructions contained within the genetic material to produce the cell's machinery. Therefore, by altering the function, the mRNA can be used to modulate the cell's machinery. Source: Sirna Technologies website.

In mammalian cells, RNAi is mediated by 21- to 23-nucleotide siRNAs containing symmetrical two nucleotide overhangs. In order to guide target RNA degradation, 5'-phosphorylation of the target-complementary siRNA strand is required. As long as the 5'-hydroxyl of a siRNA is not blocked by methylation or a 5'-phosphodiester linkage, a cellular kinase rapidly 5'-phosphorylates the siRNA duplex. Duplex siRNAs undergo an ATP-dependent unwinding step prior to formation of RISC, which contains a single-stranded siRNA. Some single-stranded 5'-phosphorylated antisense siRNAs are able to bypass this step and can enter RISC directly, although typically a higher concentration of single-stranded siRNA is required relative to duplex siRNA. The target mRNA cleavage is ATP independent and occurs precisely 10 nucleotides upstream from the complementary residue to the guide siRNA 5'-end. RISC has been found to contain members of the Argonaute family, Ago2 in D. melanogaster and eIF2C1 or eIF2C2 in human. RISC has not yet been reconstituted and the endonuclease activity remains unknown (65).

RNAi has emerged as a powerful genetic tool for the analysis of gene function in mammalian cells. Although RNAi experiments rely on high target specificity, nonspecific effects can occur. An example of this is in mammalian cells treated with 21-bp siRNAs where the interferon system has been found to be activated. It was found that a commonly used siRNA construct can induce an interferon response (25). It is recommended that a test should be done for interferon induction before attributing a particular response to the gene targeted. One precaution to limit the risk of inducing an interferon response is to use the lowest effective dose of siRNA. It was also noted that tumor cells have a defective interferon response to begin with and this phenomenon has not been observed in those cells for that reason (25). This suggests that

when introduced into cells, siRNAs have broad and complex effects beyond the selective silencing of target genes (165).

Suppression by RNAi of gene expression is typically a transient phenomenon. Gene expression usually recovers after 96 to 120 hours after, or 3 – 5 cell divisions post-transfection. This is due primarily to dilution rather than degradation of the siRNAs. The introduction of plasmids which express the siRNA and a selection gene can lead to stable RNAi that can be sustained for up to 2 months posttransfection. Adenoviral and retroviral vectors have been found to produce siRNAs in vivo and stable RNAi has been reported using these approaches. Virus-mediated RNAi may be able to circumvent some of the problems associated with cells that are generally unable to be affected by RNAi, such as transformed primary cells such as fibroblasts and dendritic cells. Retroviral long terminal repeats (LTRs), which are strong promoter elements, may be mostly responsible for driving retroviral hairpin expression rather than the promoter of the expression cassettes.

Looking only at the siRNA sequence it is currently not possible to predict the degree of gene suppression that a particular siRNA will produce. The main variable in constructing an effective siRNA sequence is the gene target site. It is usually recommended that the target site by located at least 100 – 200 nucleotides from the AUG initiation codon to be chosen. Targets within 50 – 100 nucleotides of the termination codon should also be avoided when designing a siRNA sequence. The 5' and 3' untranslated region (UTR) should also be avoided as associated regulatory proteins could potentially compromise the effectiveness of the RNAi. It has been found that siRNAs targeted to the 3' UTR have still been able to induce RNAi. It is imperative to ensure that the siRNA sequence designed targets the target gene by performing a BLAST search.

Equally as important as the choice of the target gene sequence, is the structural characteristics of the siRNA itself. Twenty-one nucleotide siRNAs with 3'-d(TT) or (UU) overhangs are the most effective at silencing. Although nucleotide-protein steric interactions contribute to the relationship between siRNA length and activity, this relationship is not fully understood. For optimal secondary structure, the GC ratio should ideally be between 45 and 55%. Low internal stability of the siRNA at the 5' antisense end is a prerequisite for effective silencing and probably important for duplex unwinding and efficient antisense entry into the RISC. The thermodynamic flexibility of the duplex 5'-end did not appear to correlate with silencing potency, whereas that of the 3'-end correlated well with efficient silencing (147). Multiple identical nucleotides in series, particularly poly(C) and poly(G), should be avoided to determine any requirements for modification, for example fluorophore labeling which can permit siRNA tracking and quantification of transfection efficiency.

In order to induce RNAi, siRNA must be transfected into the cells of interest. There are numerous transfection reagents that exist; the most commonly used are liposomal or amine-based. Cell electroporation may be used, but in some cases cell toxicity can be high. Transfection efficiency is optimized by titrating cell density, transfection time, and the ratio of siRNA-to-transfection reagent. The cell passage number at time of transfection and use of antibiotics can also affect the efficiency of the transfection experiment.

The effect of RNAi should be quantitated at both the mRNA and the protein level. Northern blot analysis is considered to be the standard for quantifying mRNA levels. Real-time reverse transcriptase polymerase chain reaction, incorporating internal controls to quantify housekeeping gene transcript levels can also be used to quantify mRNA levels. Protein knockdown is confirmed by Western blot analysis, immunofluorescence, flow cytometry, and

phenotypic or functional assays. Although RNAi generally occurs within 24 hours of transfection, both the onset and duration of RNAi depend on the turnover rate of the protein of interest, as well as the rate of dilution and longevity of the siRNAs. The duration of gene silencing can also be modified by such factors as the concentration of serum in the culture medium, which can affect the cell cycle rate. It is imperative to determine the time course of silencing in the RNAi experiment (44).

The single most important criteria for a successful RNAi experiment is the time it takes to reduce protein expression below the threshold level that is critical to sustain normal protein function. This is determined by the efficacy of the siRNA to the target mRNA of choice. Protein stability is very important. The time required to reduce protein expression below the critical level is primarily determined by the half-life of that protein. Silencing expression of stable proteins may require very long incubation periods with siRNA that can only be accomplished by stable expression of the siRNA.

CHAPTER 3-EXPERIMENTAL METHODS

Introduction

The aim of this dissertation is to identify radiation responsive genes through the use of gene trapping. Gene trapping tags the radiation-responsive genes and allows for the expression levels to be monitored through the use of the reporter protein, GFP. At the same time, a single allele gene disruption into the radiation-responsive gene through the integration of the poly(A) trap vector causes a decrease in expression. Through the use of the gene-trapping technique unknown genes that are potentially contributing to either familial or non-familial breast cancer can be identified. The radiation-responsive genes identified might then serve as markers for screening and hopefully aid in early detection.

Research Designs & Methods

I. Establishment of a gene-trapped library of clones from human mammary epithelial cells (MCF10A).

To potentially identify genes that are involved in breast epithelial cell transformation by disruption of gene expression, a gene-trapped library of MCF10A clones was developed through the use of the poly A trap retrovirus (RET) designed by Ishida and Leder (77). Virus was generated that contained the pRET vector by transfecting Phoenix amphotropic retrovirus-producing mouse cell line (from Dr. Gary Nolan, Stanford University) with the pRET vector. This viral supernatant was then utilized to infect the non-tumorigenic immortalized human mammary epithelial cells (MCF10A). Clones were selected for neomycin resistance with

G418 and pooled together. Since the infection rate is one virus per cell, this library represents cells in which one functional gene is disrupted by the integration of the vector (127).

MCF10A Cell Line

MCF10A cells arose spontaneously, without viral or chemical treatment, from mortal human diploid mammary epithelial cells of extended life span (>1.2 years). MCF10A cells have had minimal rearrangement and are diploid or near-diploid but not karyotypically normal. The diploid mortal cells (MCF10M) senesce when transferred serially in 1.05 mM Ca^{2+} whereas the immortal line, MCF10A, are attached cells, proliferated for more than 4 years in a medium with either a Ca^{2+} concentration of 1.05 mM or a low Ca^{2+} concentration of 0.04 mM. The non-tumorigenic immortalized human breast epithelial cell line, MCF-10A, was isolated from cells derived from the breast tissue obtained from a mastectomy performed on a 36-year old, parous premenopausal woman with no history of breast malignancy (163). The breast histopathological diagnosis was extensive fibrocystic disease, consisting of increased mammary fibrous stroma containing numerous dilated mammary ducts, benign apocrine metaplasia, and small focal areas of intraductal hyperplasia with no evidence of atypia. The patient was free of disease. MCF10A cells have a near diploid human female karyotype, can organize into monolayers on plastic and ductular structures on collagen gels, form dome formations in post-confluent monolayers, are controlled by hormone and growth factors and are not tumorigenic in nude mice. MCF10A cells express sialomucins and keratins, and ultrastructurally are low cuboidal cells with numerous hemidesmosomes, desmosomes, and short microvilli, all typical attributes of normal human breast epithelium. Collectively, these characteristics make this cell line closest to a normal human breast epithelial cell line, suitable for studying adhesion, spreading and migration (163).

MCF10A cells arose after 849 days in culture. At days 754 to 984, after the MCF10A cells became immortal, a reciprocal translocation, t(3:9)(3p13:9q22) occurred which is believed to have led to the cells spontaneously becoming immortalized. Other nonclonal changes were also observed. MCF10A cells in the range of 1009 to 1527 days in culture were found to be monosomic for chromosomes 3 and 9, although other sporadic losses were seen. All but one of the remaining chromosomes were normal and paired. Derivatives were identified for the missing chromosomes: marker M1 was identical to the der(3) resulting from the t(3:9) translocation described above. Marker M2 was interpreted as a derivative arising from the reciprocal 9p+ translocation that had undergone further rearrangement and translocation. The new translocation at the 9q22 breakpoint was found to have originated from the partial duplication of 9q and probably a small portion of 5q as well. A third marker, M3, was found to contain an unbalanced rearrangement of chromosome 6: t(6:19)(p25:q12). It was first observed after MCF10A cells had been in culture for 1229 days. The cells illustrated some fluctuation in the chromosome number, with a subline showing 48 chromosomes and identical to the stemline except in the acquisition of extra copies of chromosome 8 and 16. Roughly 5% of the MCF10A cells at this point of passage were also found to be tetraploid with duplication of the markers previously described (163).

MCF10 cells have been shown to lack the presence of SV40 and cellular proto-oncogene rearrangement, amplification, and mutational activation were not detected. MCF10A cells display absence of anchorage-independent growth. The cells are estrogen receptor negative and do not produce progressively growing tumors in athymic mice. The MCF10A cell line remains non-tumorigenic when injected at late passage into athymic mice. Each inoculum initially forms a mass; the largest observed was 4.4 mm in diameter. All the masses gradually decreased in size,

so that no mass was palpable or visible by the fifth week. This all supports the fact that MCF10A is not a malignant cell line (170).

MCF10A cells do not require serum, but instead rely on the mitogenic signals provided by EGF for G1 progression and entry into S phase (163). Cortisol is required for growth even when 5% serum is present in the cell culture medium. Upon increased time in cell culture, MCF10A displays the following: cholera entertoxin inhibits cell growth, the doubling time of the cells decreases from 48 hours to 20 hours, and finally the cell number increases 30-fold per T_{25} flask (163).

MCF10A cells are cultured in a 1:1 mixture of Dulbecco's modified Eagle's medium and Ham's F12 medium (DMEM/F12) supplemented with 5% horse serum, hydrocortisone (0.5 μg/ml), insulin (10 μg/ml), epidermal growth factor (20 ng/ml), and penicillin-streptomycin (100 μg/ml each).

MCF7 Cell Line

MCF7 is a human breast carcinoma cell line. The MCF7 cell line was developed by Soule and coworkers in 1974 (172). It was utilized in experiments to represent the expression patterns seen in breast cancer patients following radiation treatment. It possesses two characteristics particularly important for modeling human breast cancer. First, the cells express estrogen receptors in vitro. These receptors are functional in that estrogen treatment results in the up-regulation of the progesterone receptor. Second, inoculation of MCF7 cells into immuno-deficient mice resulted in development of progressively growing tumors, but only if estrogen is administered concurrently (26).

The MCF7 cell line was derived from pleural effusion from a female breast cancer patient with metastatic disease. The cells are cultured in Eagles' minimal essential medium

supplemented with 20 μg insulin and 250 units of penicillin and 250 μg of streptomycin per ml and 10% fetal calf serum (172).

Gene-Trapping

The polyA trap vector was utilized in my experimental system. PolyA addition traps provide a means to both trap non-expressed genes and obtain coding sequence tags. PolyA trap vectors contain a promoter that is active in the target cells that directs the expression of a selectable marker gene. The selectable marker gene does not contain a polyA addition signal and is therefore not expressed unless it lands within an intron of a gene and traps downstream exons that include a polyadenylation signal. This design should allow the trapping of non-expressed genes as the trapping vector carries its own promoter. The gene-trap fusion construct contains a selectable marker sequence fused with exons from the trapped gene on the 3' end. This type of fusion allows the efficient process of 3' RACE to be used to identify sequence tags and these 3' sequences are more likely to contain coding sequences (50, 78, 121).

pRET retroviral vector

The versatile retrovirus vector (RET) was employed to carry out my experiments. It employs an enhanced poly A-trap strategy for the stringent selection of intragenic provirus integration. A very strong splice acceptor and an efficient poly(A) signal are included for the complete disruption of trapped genes, the expression patterns of which can be easily followed in living cells by analyzing green fluorescence protein (GFP) expression. It is then possible to confirm that the mutant phenotype was created by the provirus integration by removing the integrated provirus using Cre/loxP-mediated homologous recombination and restoring the integrity of the interrupted gene. As a further provision, the gene terminator cassette of the RET vector contains multiple stop codons in all three reading frames in the sequences 3' to the bcl-2

splice acceptor. It thereafter encodes an internal ribosome entry site (IRES) sequence, a GFP cDNA, and a BGH poly(A) signal. The GFP serves as a visible marker of transcriptional activity of the trapped gene. RNA splicing between an exon of the trapped gene and the gene terminator cassette generates a fusion transcript consisting of the 5' part of the trapped gene and the GFP cDNA. Ribosomes translating the fusion transcript would be expected to dissociate from it at one of the multiple stop codons and to re-associate at the IRES portion, resulting in bi-cistronic expression of the first half of the trapped gene and GFP. It should also be noted that neither of the two LTRs of the integrated provirus have transcriptional enhancers, thereby minimizing the possibility of insertional activation of trapped or adjacent genes (78).

In the RET system, it is not necessary to isolate all the cells with randomly integrated gene-trap vectors and analyze further for the induced expression of a marker gene included in the vector. The RET vector first selects for intragenic integration events, thereby minimizing the number of clones to be handled in detailed gene induction analyses in later stages. Moreover, the RET system provides for rapid isolation of a cDNA fragment corresponding to an inducible gene by using 3' RACE on chimeric NEO mRNA which is constitutively synthesized from an RNA polII promoter (78).

For pRET, the SupF sequence, a 208 base pair BamHI-BamHI fragment, was replaced with a loxP signal in the U3 portion of the 3' LTR of a gag^+ self-inactivating retrovirus vector pGen-. The gene-terminator cassette of pRET contains four components. The first is the human bcl-2 gene intron 2/exon 3 splice acceptor which is a 234 base pair portion of the intron and a 199 base pair portion of the exon sequence which contains multiple stop codons in all three reading frames. The second is an IRES sequence identical to nucleotides 260 – 845 of the 5' untranslated region of the encephalomyocarditis virus (EMCV-Rueckert) polyprotein RNA. The

third component of the gene terminator cassette is the cDNA coding region for EGFP. Lastly, a poly(A) signal from the bovine growth hormone gene is included. A neomycin cassette is the other major component of the pRET retroviral vector. It consists of two parts, a NEO coding region, gene promoter and an mRNA instability signal. A short form, roughly 224 base pairs, of the RNA polymerase II gene promoter is included to drive expression of the neomycin selection gene. The NEO coding region and the splice donor region of pRET come from the mouse hprt gene exon 8/intron 8, which are 17 and 157 base pairs respectively, and were included in pRET as a 1.2 kb HincII – AatII fragment from the NEO expression vector pKT3NP4. An mRNA instability signal, which is 53 base pairs in length, is derived from the 3' untranslated region of the human GM-CSF cDNA (78).

The sequence of the trapped gene can be identified using techniques that are based on the polymerase chain reaction (PCR). This can lead to the isolation of novel genes regardless of their level of expression in vivo. The sequence of the trapped gene can be identified using RACE (rapid amplification of cDNA ends by PCR). It is a process that produces amplified DNA using only one sequence-specific primer. It is important to note that the level of expression of the reporter gene remains dependent on the rate of transcription of mRNA of the trapped gene.

When a mutant phenotype is observed after a gene-trap experiment, it is desirable to be able to confirm that the resulting phenotype was really brought about by provirus insertion. This is most rigorously done by removing the integrated provirus from the host genome and scoring for a phenotypic reversion. The integrated RET virus acquires an additional loxP site in its 5' LTR in the same orientation as the loxP signal in the 3' LTR. The use of Cre recombinase allows this integration site to be excised. Cre is a 38 kDa recombinase protein from

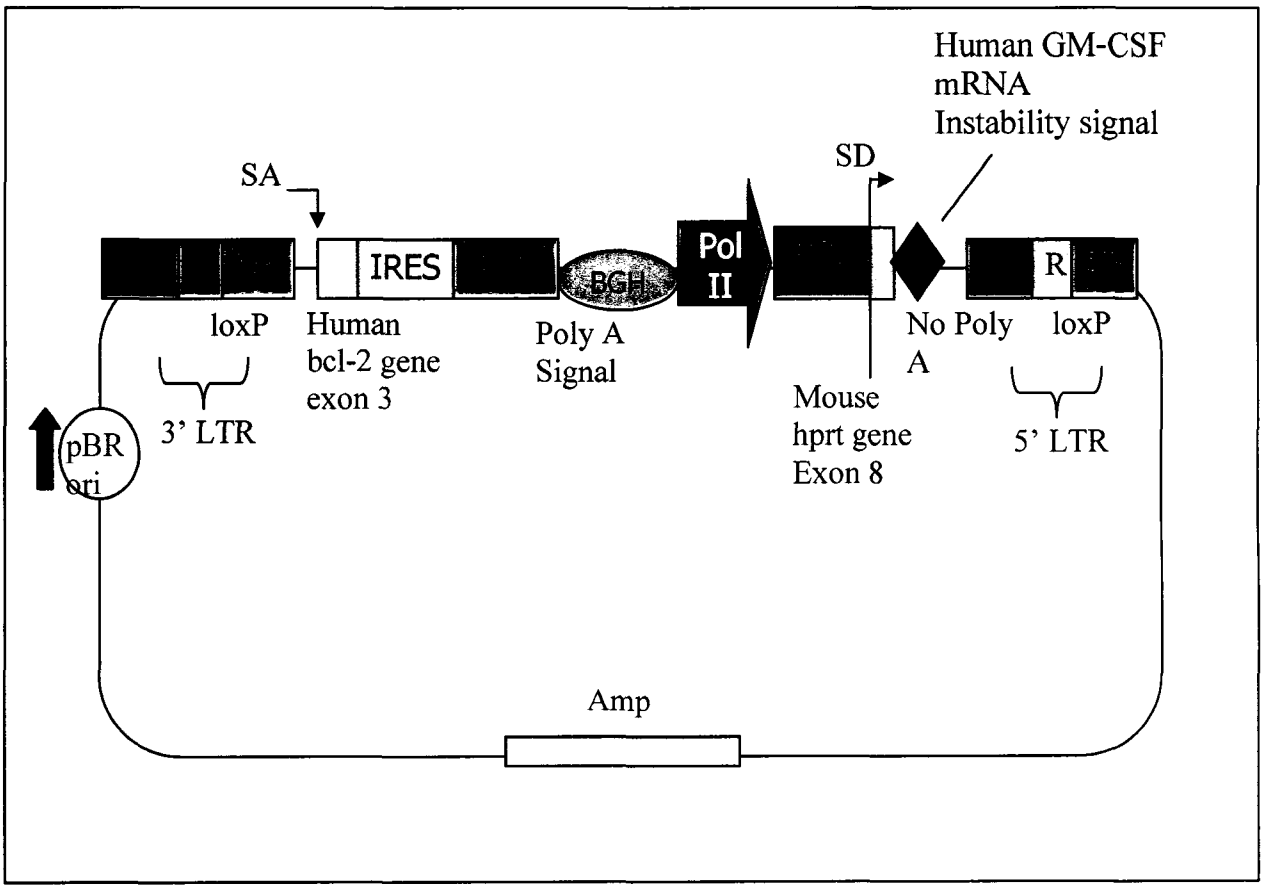


Figure 3.1. The RET gene-trap vector. Source: Ishida, Y. and P. Leder, *RET: a poly A-trap retrovirus vector for reversible disruption and expression monitoring of genes in living cells.* Nucleic Acids Res, 1999.

bacteriophage P1 that catalyzes reciprocal site-specific recombination between 34-base pair loxP sites. In bacteria and in vitro Cre mediates both intramolecular (excision or inversion) and intermolecular (insertion or translocation) recombination at loxP sites. Surprisingly, Cre also promotes both excisive and integrative recombination in eukaryotic cells (55). Cre recombinase is a tool for providing post-integrational modification of targeted genes and transgenes within the cell. It has opened up many possibilities for cell type-specific gene/transgene manipulations and for creation of any imaginable genome alterations in the mouse from subtle changes to large chromosomal aberrations (66).

Apparently, the integration of a trap vector is not entirely random. Often, integrations into the 5' end of genes are observed. 5' integrations result in relatively small fusion proteins, whereas 3' insertions can lead to large fusion complexes between the resistance marker and the trapped gene. Therefore, in the case of 3' integrations, because of adequate protein folding, the resistance marker may become inactive. Accordingly, splice trap vector integrations could be more or less random, but 3' insertions might simply not be functional. Given that this hypothesis is correct, one would expect 3' integrations to be functional in splice donor vectors. In contrast however, for splice donor-type vectors, it has been reported that 88% of integrations occurred either 5' or in the middle of the encoded cDNA. It has not been observed that preferred 5' integration occurs when using a splice acceptor-type vector. Instead, integrations occurred unbiased. Misfolding of the fusion complexes is not applicable to IRES containing vectors since translation of the selector gene takes place independently of the trapped locus from a bicistronic transcript (50).

Viral packaging cell line

It is very important in assembling retroviral vectors that the production of recombinant particles be very high titer. The origin of the cell on which the packaging system is based has great influence on this. It is important that cells chosen to package viral particles have security, absence of pathogens, traceability of the parental cell, and a low rate of homologous recombination. The 293T cell line has the advantage of being efficiently transfected.

For virus production, the Phoenix amphotropic packaging cell line was utilized. The Phoenix packaging cell line is a second-generation retrovirus producer line for the generation of helper free ecotropic and amphotropic retroviruses. The lines are based on the 293T cell line, which is a human embryonic kidney line transformed with adenovirus E1a and carrying a temperature sensitive T antigen co-selected with neomycin. A unique feature of this cell line is that it is highly transfectable by either calcium phosphate mediated transfection or lipid-based transfection protocols, with up to 50% or higher of the cells being transiently transfected. The ecotropic and amphotropic cell lines were created by placing constructs capable of producing gag-pol, an envelope protein, into 293T cells (132).

The lines offered advantages over previous stable packaging systems in that viruses could be produced in just a few days. The ability to monitor gag-pol production on a cell-to-cell basic was accomplished by the introduction of an IRES-CD8 surface marker downstream of the reading frame of the gag-pol construct. Thus, CD8 expression is directly related to the amount of intracellular gag-pol. The stability of the producer cell population's ability to produce gag-pol can then be monitored through the use of flow cytometry. In both the gag-pol construct and the envelope construct, non-moloney promoters were utilized to minimize recombinational potential.

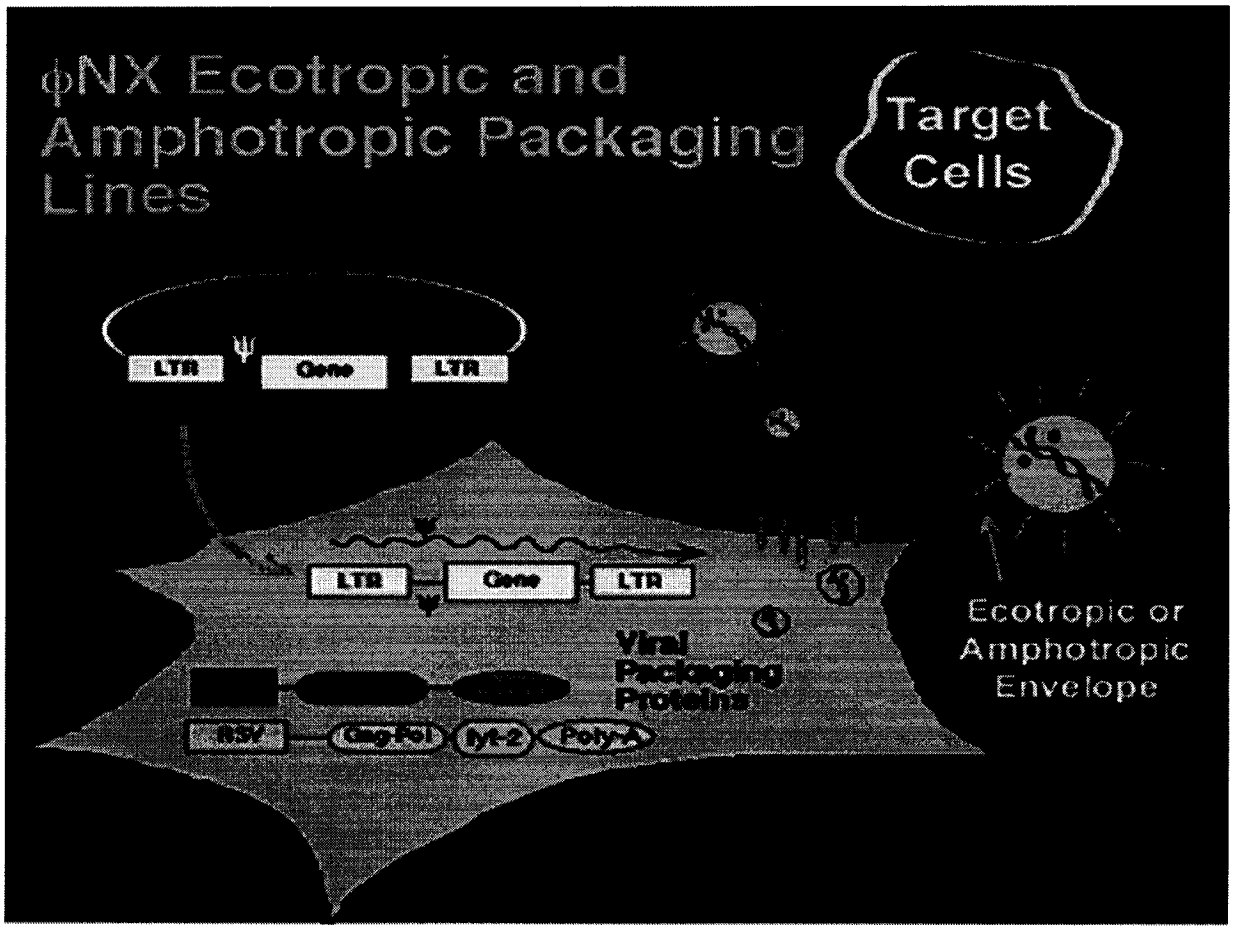


Figure 3.2. Viral packaging of retroviruses in the Ecotropic and Amphotropic packaging cell lines.

Source: http://www.stanford.edu/group/nolan/index.html

II. Assay to detect changes in gene expression in human mammary epithelial cells using gene-trapped MCF10A clones

A transformed library of radiation-induced human breast epithelial cells based on the spontaneously immortalized MCF-10A cells that have been infected with the retrovirus pRET was established. Changes in the promoter activity of the trapped gene were analyzed via changes in the level of expression of the GFP reporter protein through an ELISA sandwich assay. Since the GFP protein contained in the RET vector is promoterless, the expression level of the trapped gene was determined based on GFP expression.

Cell sorting

Cell sorting was performed in order to separate the MCF10A gene trapped pooled population into GFP+ and GFP- populations. This was important in order to determine if the initial GFP expression level changed or had an effect on the clone following radiation treatment. Flow cytometry measurements were made with an EPICS V cell sorter (Coulter, Miami, FL) interfaced to a Cicero data acquisition and display system (Cytomation, Inc., Ft. Collins, CO) in Dr. Mike Fox's laboratory. Cells were illuminated with 500 mW of a 488-nm laser light from an argon ion laser. Clumped cells were mostly eliminated by gating on peak versus integral fluorescence of the PI signal. Cells were sorted based on their varying GFP expression levels. EGFP, which is present in the pRET vector, can be easily detected by flow cytometry. EGFP has a long half life of >24 hours.

Plating of MCF-10A gene-trapped clones in 96-well plates and production of replica plates

Prior to plating the gene trapped pooled clones in 96 well plates, cell sorting by flow cytometry was done in order to sort the pooled clones into GFP+ and GFP- expression levels. The pooled clones of gene trapped MCF10A cells were plated by a limiting dilution of 1 cell/well in 96 well plates and allowed to expand in the presence of G418. After the cells have

grown, media was mixed by pipetting and 25-50 microliters of media and cells were transferred to several 96 well plates and allowed to expand. Replica plates were made from the single-cell assay plates of MCF10A-1023 pooled clones previously sorted into GFP+ and GFP- pools by flow cytometry. These replica plates were used for the GFP quantitation assays and the original plate will serve as the source of the clones for further assays.

Determination of the effect of gamma irradiation on expression of reporter protein GFP

To determine which gene trapped single cell clones were responding to ionizing radiation, an ELISA assay was performed to quantitate the GFP levels of the clones following exposure to a 2.0 Gy dose of ^{137}Cs gamma rays. Since GFP is a marker for the endogenous gene's expression level, measuring the GFP change following IR treatment allows the determination of the gene trapped clones illiciting a response. Those clones showing a 2-fold or higher change in response to the radiation treatment were expanded and 3' RACE was performed in order to determine the sequence of the trapped gene.

Gamma irradiation (0-2 Gy) of 96-well plates

Prior to irradiation, near-confluent cultures of MCF-10A cells were trypsinized, counted and resuspended in DMEM/F12 media at a concentration of 1.5×10^6 cells/ml. Replica plates of MCF-10A clones were then subjected to varying low doses (0-2 Gy) of gamma irradiation using a J. L. Shepard Mark-I ^{137}Cs sealed source gamma irradiator. Replica plates of MCF-10A clones that have not been irradiated were kept as negative controls.

Mark I Irradiator

The Mark I is a self-contained, self-shielded ^{137}Cs sealed source. The source is permanently contained within the unit. Samples are irradiated with 0.622 MeV γ-rays. The Mark I has been designed to irradiate biological samples, including cells in culture, biological

fluids and tissues such as blood, serum and media. Exposure times can vary from a few seconds to hours. A turntable is present in the chamber to insure an even dose distribution. Dose-rates within the chamber can vary from about 40 Gy per minute to 4 cGy per minute, depending upon the position and shielding.

Analysis of GFP expression

The effect of gamma irradiation on the level of GFP expression in the gene trapped clones was determined using an ELISA assay. The 96-well plates were scanned at 450 nm for detection of the GFP antigen through the use of TMB substrate. Fluorescence was quantitated using the available microplate reader software. It allows for fast measurements of low level detection limits.

ELISA analysis of GFP expression

In order to perform the ELISA assay cell lysates of the gene trapped single cell clones must be obtained to use as the antigen. The cell lysates were prepared by first washing the adherent cells with ice-cold PBS. The cells were spun at 1000 rpm for 5 minutes in order to pellet the cells. Ice-cold modified Radioimmunoprecipitation (RIPA) buffer is then added to the cells in order to lyse them. The cells were scraped off the plate with a cooled plastic cell scraper. The cell suspension was then transferred to a centrifuge tube and rocked on an orbital shaker in the cold room for 15 minutes in order to lyse the cells. The cell lysates were centrifuged at 14,000 x g in a precooled centrifuge for 15 minutes. The supernatant was then transferred to a new microcentrifuge tube. The cell lysates must be diluted 1:10 before determining the protein concentration due to the interference of the detergents in the lysis buffer. A microplate reader then quantitates the level of GFP expression in the gene trapped clones both before after gamma irradiation treatment using an ELISA sandwich assay.

Enzyme-linked immunosorbent assay combines the specificity of antibodies with the sensitivity of simple enzyme assays. ELISAs can provide a useful measurement of the antigen or antibody concentration in a sample. One of the most useful of the immunoassays is the two-antibody "sandwich" ELISA. This assay is used to determine the antigen concentration in unknown samples. This ELISA is fast and accurate, and if purified antigen standard is available, the assay can determine the absolute amounts of antigen in unknown samples. The sandwich ELISA requires two antibodies that bind to epitopes that do not overlap on the antigen. This can be accomplished with either two monoclonal antibodies that recognize discrete sites or one batch of affinity-purified polyclonal antibodies.

To utilize this assay, one antibody (the capture antibody) is purified and bound to a solid phase. Antigen is then added and allowed to complex with the bound antibody. Unbound products are then removed with a wash, and a labeled second antibody (the detection antibody) is allowed to bind to the antigen, thus completing the sandwich. This assay is then quantitated by measuring the amount of labeled second antibody bound to the matrix, through the use of a colorimetric substrate. A major advantage of this technique is that the antigen does not need to be purified prior to use, also that these assays are very specific. However, one disadvantage is that not all antibodies can be used. Monoclonal antibody combinations must be qualified at match pairs, meaning that they can recognize separate epitopes on the antigen.

Unlike Western blots, which use precipitating substrates, ELISA procedures utilize substrates that produce soluble products. Ideally the enzyme substrates should be stable, safe and inexpensive. Popular enzymes are those that convert a colorless substrate to a colored product. Substrates used with peroxidase include 3,3'5,5'-tetramethylbenzidine base (TMB), which yields blue color, respectively. The sensitivity of the sandwich ELISA is dependent on

four factors: the number of molecules of the first antibody that are bound to the solid phase, the avidity of the first antibody for the antigen, the avidity of the second antibody for the antigen, and the specific activity of the second antibody. The amount of capture antibody that is bound to the solid phase can be adjusted easily by dilution or concentration of the antibody solution. The avidity of the antibodies for the antigen can only be altered by substitution with other antibodies. The specific activity of the second antibody is determined by the number and type of labeled moieties it contains.

The general protocol for determining the amount of GFP present in the gene trap clones was performed as follows. The unlabeled antibody, monoclonal anti-Green Fluorescent Protein (mouse) was bound to the bottom of each well by adding approximately 50 μl of antibody solution to each well (20 mg/ml in PBS). The plate was incubated at 4˚C overnight to allow for complete antibody binding. The next day the plate was washed twice with PBS. The remaining sites for protein binding on the microtiter plate were then saturated by incubating the plate with blocking buffer composed of 3% BSA/PBS. The blocking buffer did not contain sodium azide since it is a known inhibitor of horseradish peroxidase which is conjugated to the secondary antibody. The wells of the plate were then filled to the top and incubated overnight in a humid atmosphere to ensure efficient blocking. The next day the wells of the plate were washed twice with PBS. A total of 50 μl of the antigen solution, the gene trapped clone cell lysates, was added next. The plate was incubated for at least 2 hours at room temperature in a humid atmosphere. The microtiter plate was then washed four times with PBS. The labeled secondary antibody solution, peroxidase conjugated affinity purified rabbit anti-mouse IgG (heavy and light strands) was added in excess at 100 μl. All dilutions of the antibodies were made in the blocking buffer.

The plate was then incubated for at least 2 more hours and then washed with several changes of PBS.

At this point the TMB substrate was added. TMB is 3,3',5,5 '-tetramethylbenzidine. TMB is a highly sensitive substrate. Due to its rapid reaction rate, it is ideally suited for on-line kinetic analysis. It produces a blue color measurable at a wavelength of 650 nm. TMB can also be utilized in endpoint assays by a stopping the reaction with 2M hydrosulfuric acid. A yellow reaction product forms upon addition of hydrosulfuric acid causing acidification of the reaction which then can then be read at 450 nm. The TMB utilized in this experiment was dissolved in DMSO to a concentration of 6 mg/ml. It is noncarcinogenic, a benefit compared to DAB. It produces a deep blue color during the enzymatic degradation of hydrogen peroxide by horseradish peroxidase and yields a clear yellow color upon addition of the stop solution, 2M hydrosulfuric acid. A total of 100μl TMB substrate was added per well of the 96-well plate. The development time for the substrate was typically 40 minutes and then the stop solution was added. The plate was read in a microplate reader and endpoint analysis is performed.

III. Identification of the "trapped" genes affected by gamma irradiation

Once specific clones were identified as responsive to radiation, the next step was for the trapped genes to be identified. RACE has been used for amplification and cloning of rare mRNAs and may be applied to existing cDNA libraries. Additionally, RACE products can be directly sequenced without any intermediate cloning steps, or the products may be used to prepare probes. Products generated by the 3' RACE procedures may be combined to generate full-length cDNAs. Lastly, the RACE procedures may be utilized in conjunction with exon-trapping methods to enable amplification and subsequent characterization of unknown coding sequences.

3'RACE to amplify trapped gene

Rapid Amplification of cDNA ends (RACE) is a procedure for amplification of nucleic acid sequences from a messenger RNA template between a defined internal site and either the 3' or the 5' end of the mRNA. This methodology of amplification with single-sided specificity has been described as "one-sided" PCR or "anchored" PCR. PCR requires two sequence-specific primers that flank the sequence to be amplified. However, to amplify and characterize regions of unknown sequences, this requirement imposes a limitation.

3'RACE takes advantage of the natural poly (A) tail found in mRNA as a generic priming site for PCR. In this procedure, mRNAs are converted into cDNA using reverse transcriptase (RT) and an oligo-dT adapter primer. Specific cDNA is then amplified by PCR using a gene-specific primer (GSP) that anneals to a region of known exon sequences and an adapter primer that targets the poly (A) tail region. This permits the capture of unknown 3' mRNA sequences that lie between the exon and the poly (A) tail. The 3' RACE procedure is summarized in Figure 3.3.

3'RACE protocol

First strand cDNA synthesis is initiated at the poly (A) tail of mRNA using the adapter primer (AP). After first strand cDNA synthesis, the original mRNA template is destroyed with RNase H, which is specific for RNA:DNA heteroduplex molecules. Amplification is performed, without intermediate organic extractions or ethanol precipitations, using two primers: one is a user-designed GSP that anneals to a site located within the cDNA molecule; the other is a universal amplification primer that targets the mRNA of the cDNA complementary to the 3' end of the mRNA.

The first strand cDNA synthesis reaction is catalyzed by SuperScript II RNase H⁻RT. This enzyme has been engineered to eliminate the RNase H activity, found in other reverse transcriptases that can degrade mRNA during the first strand reaction. SuperScript II RT is not inhibited significantly by ribosomal and transfer RNA and may be used to synthesize first strand cDNA from a total RNA preparation. The RNA template is removed from the cDNA:RNA hybrid molecule by digestion with RNase H after cDNA synthesis to increase the sensitivity of PCR. First strand cDNA was synthesized in a 20 μl reaction from 5μg of total RNA with AD-poly(T) primer (5'-CGTAGCTCTAGACTCCGTGTCCAACT_{20}-3') and SuperScript II reverse transcriptase (Gibco BRL).

Amplification of a target cDNA requires priming with two oligonucleotides and *Taq* DNA polymerase. The sense amplification primer is the NEO 1.5 primer. The antisense amplification primer is the AD primer. Two rounds of 3' RACE are performed in a 50 μl reaction using the Advantage-GC cDNA polymerase mix (Clontech), first with 1 μl of the above cDNA mix (1/20 of the total reaction), NEO 1.5 primer (5'-GCGAA-TGGGCTGACCGCTTCCTCGTGC-3') and AD primer (5'-CGTAGCTCTAGACTCCGTGTCCAA-3'), and second with 1 μl of the first PCR mix (1/50 of the total reaction), NEO 2.0 primer (5'-TACGGTA-TCGCCGCTCCCGATTCGCAG-3') and AD-plus primer (5'-CGTAGCTCTAGACTCCGTGTC-CAACTTTT-3'). The PCR conditions were as follows: (i) 94°C, 3 min (1^{st} round) or 1 min (2^{nd} round), (ii) 8 x (94°C, 40s + 72°C, 4 min), (iii) 32 x (94°C, 40s + 66°C, 2 min + 72°C, 2 min) and (iv) 72°C, 4 min.

After confirming the successful amplification of the DNA fragments by agarose gel electrophoresis, the PCR primers are removed from 30 μl of the final 3'RACE reaction mix by the QIAquick PCR purification kit (Qiagen), and the nucleotide sequence of the purified PCR

fragment(s) is determined by direct sequencing with the NEO.SEQ primer (5'-TGACGAGTTCTTCTGAGGGGATCC-3') containing a *Bam*HI recognition sequence at the end, which is unique to the NEO cassette of the RET vector. This allows for the gene-trapped insert to be removed from the vector. Ultimately, the 3' RACE procedure should produce a single, prominent band on an agarose gel.

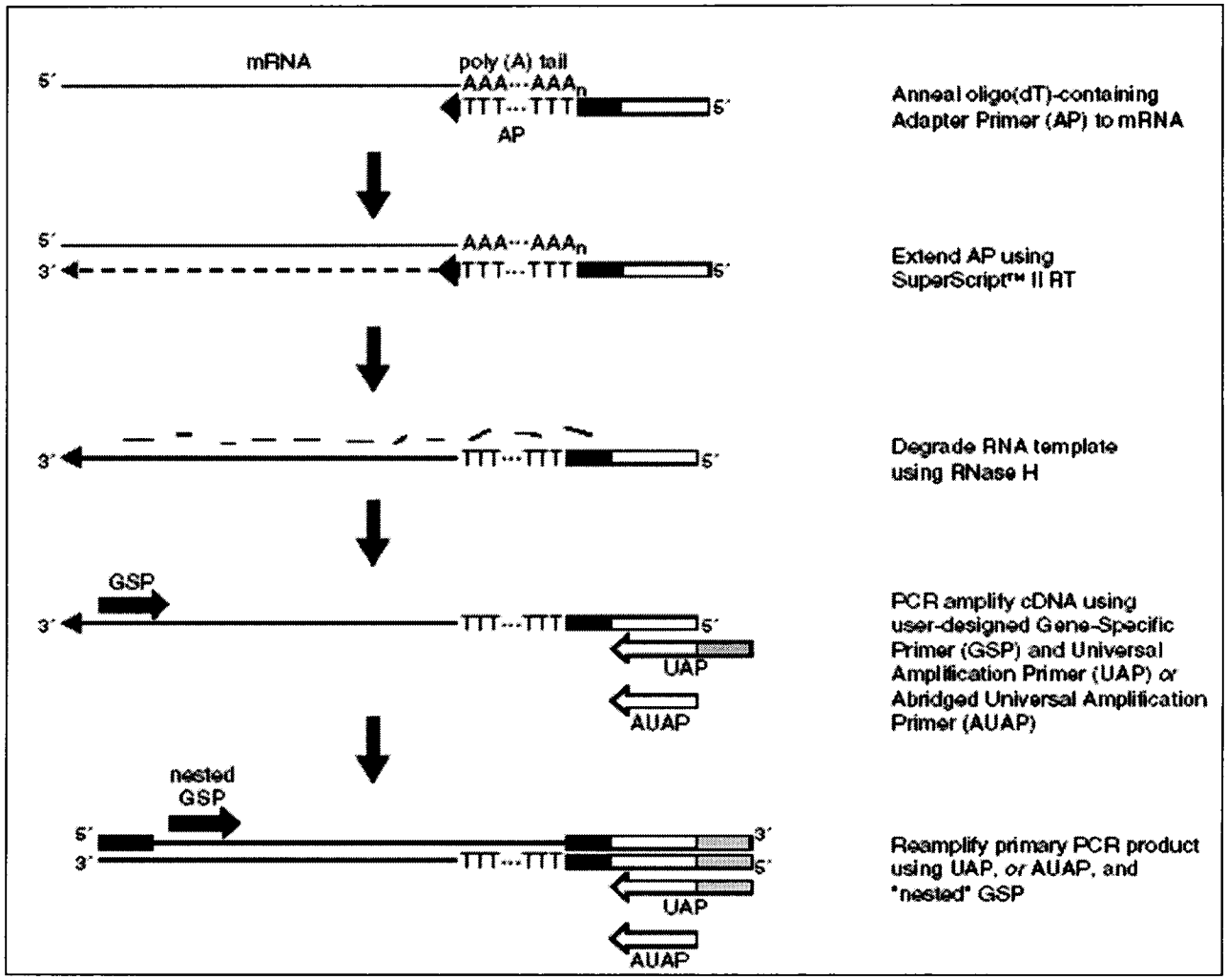


Figure 3.3. 3'RACE procedure is summarized to the right. Source: Gibco technical manual for 3' RACE System for Rapid Amplification of cDNA Ends.

Sequencing of insert

The PCR products were cloned into a sequence vector using TOPO TA Cloning kit for sequencing from Invitrogen. This was performed in order to be able to sequence the clones illustrating a proper band size upon analysis by gel electrophoresis. The sequence of the gene trapped clones sent off for sequencing were then plugged into BLAST in order to search for homologous genes.

TOPO TA cloning

The plasmid vector (pCR®4-TOPO®) is supplied linearized with single 3′ thymidine (T) overhangs for TA Cloning® and the topoisomerase is covalently bound to the vector. Taq polymerase has a nontemplate-dependent terminal transferase activity, which adds a single deoxyadenosine (A) to the 3′ ends of PCR products. The linearized vector supplied in this kit has single, overhanging 3′ deoxythymidine (T) residues. This allows PCR inserts to ligate efficiently with the vector. Topoisomerase I is from the Vaccinia virus and binds to duplex DNA at specific sites and cleaves the phosphodiester backbone after 5′-CCCTT in one strand (215). The energy from the broken phosphodiester backbone is conserved by the formation of a covalent bond between the 3′ phosphate of the cleaved strand and a tyrosyl residue (Tyr-274) of the topoisomerase I. The phospho-tyrosyl bond between the DNA and enzyme can then be subsequently be attacked by the 5′ hydroxyl of the original cleaved strand, reversing the reaction and releasing topoisomerase (215). TOPO® Cloning exploits this reaction to efficiently clone PCR products.

Several positive colonies were selected from each plate and grown overnight. Plasmid DNA was isolated using a QIAprep spin miniprep kit (Qiagen). The nucleotide sequence of the purified PCR fragments was determined by sequencing using the M13 universal primers: M13

universal forward (-20) 5' GTAAAACGACGGCCAGT 3' and M13 universal reverse 5' CACACAGGAAACAGCTATGACCAT 3' or NEO.SEQ primers. The NEO.SEQ primer contains a *Bam*HI recognition sequence at the end, which is unique to the NEO cassette of the RET vector, making it easy to cut out the insert for later work. Sequencing was performed by Davis Sequencing.

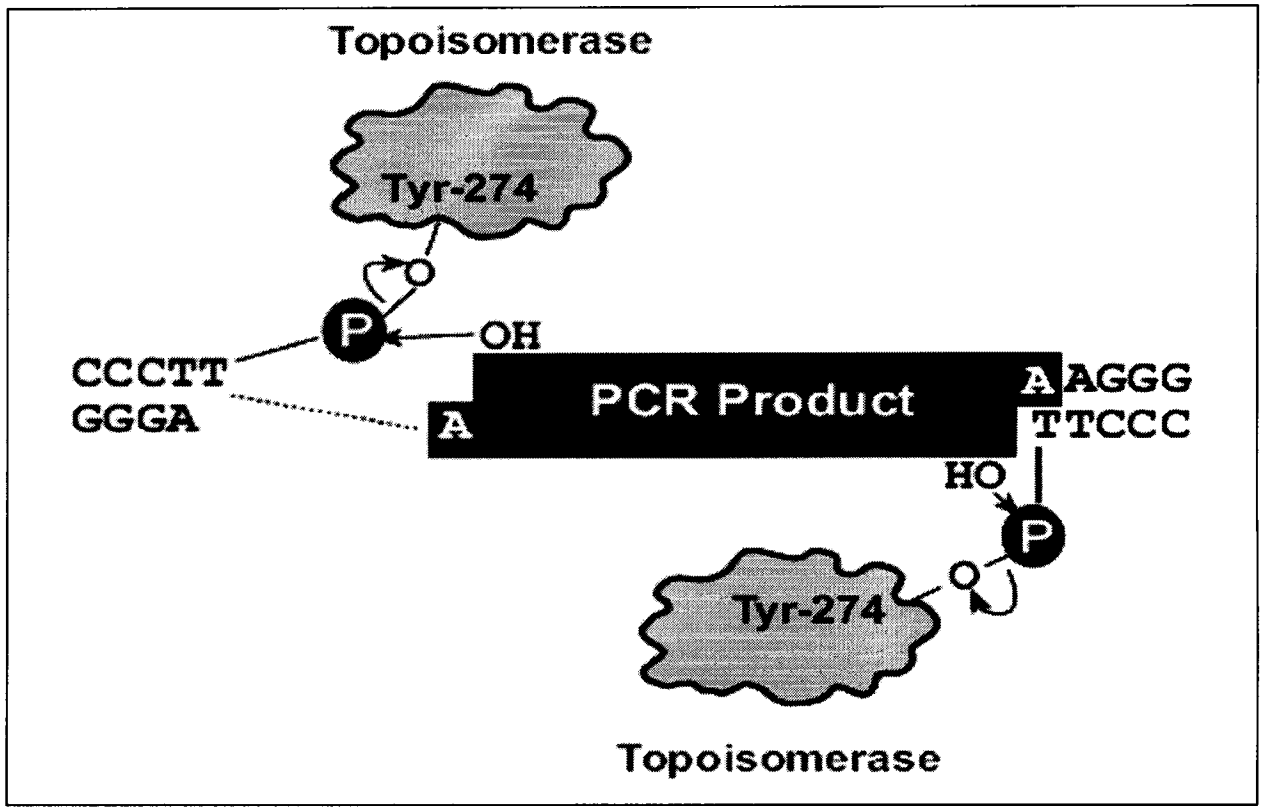


Figure 3.4. TOPO TA Cloning kit for sequencing topoisomerase ligation reaction. Source: TOPO TA cloning kit for sequencing manual.

BLAST search to identify gene from GENBANK

The total number of genes in the human genome is estimated at 30,000 to 35,000. The functions are unknown for over 50% of discovered genes. Over 30 genes have been pinpointed and associated with breast cancer, muscle disease, deafness, and blindness. These genes provide focused targets for the development of effective new therapies. Deriving meaningful knowledge from the DNA sequence will define research through the coming decades to inform our understanding of biological systems. In the past, researchers studied one or a few genes at a time. With whole-genome sequences and new high-throughput technologies, it is possible to study all the genes in a genome.

BLAST analysis of sequences

Sequences obtained were used as templates for a database search using BLASTN and BLASTX in GenBank to look for homologous gene sequences in the human genome. This provided a complete cDNA sequence as well as information on the genomic structure and chromosomal location of homologous genes to our cDNA sequence found in the human genome.

BLAST® (Basic Local Alignment Search Tool) is a set of similarity search programs designed to explore all of the available sequence databases regardless of whether the query is protein or DNA. The BLAST programs have been designed for speed, with a minimal sacrifice of sensitivity to distant sequence relationships. The scores assigned in a BLAST search have a well-defined statistical interpretation, making real matches easier to distinguish from random background hits. BLAST uses a heuristic algorithm that seeks local as opposed to global alignments and is therefore able to detect relationships among sequences which share only isolated regions of similarity (4).

Cell cycle analysis

Cell cycle analysis was performed in order to determine if there was a difference in the proportion of cells in the phases of the cell cycle before and after radiation treatment. Cell cycle analysis allowed for the verification that the relative mRNA expression levels being seen in the real-time PCR experiments were not due to cell cycle factors. This should not be a factor due to the fact that the MCF10A cells utilized in the experiment were not synchronized and thus the majority of cells were expected to be in the G1 phase of the cell cycle.

Cell cycle analysis was performed by flow cytometry. In order to execute the experiment, the cells needed to be fixed with 70% alcohol first. The ethanol fixation began by first rinsing the cells twice with 1X PBS. Trypsin was added and the cells were incubated at 37°C for 15 minutes to remove the cells adhering to the flask. The trypsin reaction was stopped by adding DMEM/F12 complete media and the cells were transferred to a 15 ml centrifuge tube. The tubes were centrifuged for 3 minutes at 1500 rpm. The supernatant was aspirated and the cells were resuspended in 2 ml cold PBS. Cold 100% ethanol was then added dropwise while vigorously vortexing the cells. The cell solution was brought up to 70% ethanol by the addition of more 100% ethanol. Prior to flow cytometry analysis, the cells fixed in ethanol were centrifuged for 3 minutes at 1500 rpm to remove of the ethanol fixative. The supernatant was decanted and the cells were resuspended in 1 ml of propidium iodide containing 40 KU/ml of RNase A. The cells were now able to be analyzed with flow cytometry analysis for changes in the cell cycle.

Real-time PCR protocol

Introduction

Real-time PCR analysis of the mRNA expression of the identified trapped genes was performed in the MCF10A parental cells in order to confirm that a radiation response was indeed being seen. The expression of the mRNA transcript in both the gene trap clones and in the parental MCF10A cell line was performed both with and without radiation treatment was analyzed and compared with real-time PCR. Total RNA from both cell lines was isolated with RNAEasy Kit (Qiagen). Real-time PCR analysis was performed using either the BioRad Icycler or the Applied Biosystems 7000 sequence detection system. The results of the real-time PCR analysis were analyzed using the Applied Biosystems software. This was utilized in order to determine the expression of the radiation response genes following varying doses of ionizing radiation in the parental cell line in order to verify that the genes were truly responding to radiation.

The TaqMan Gold RT-PCR kit from Applied Biosystems was used for real-time PCR amplification of mRNA transcripts. The primers for the reverse transcription reaction are oligo $d(T)_{16}$ primers which are provided with the TaqMan Gold RT-PCR kit. These primers are used to reverse transcribe only eukaryotic mRNAs and retroviral sequences with poly(A) tails. The RT-PCR reaction exploits the 5' nuclease activity of AmpliTaq Gold enzyme to cleave a TaqMan probe during PCR. The TaqMan probe contains a reporter dye, FAM, at the 5' end of the probe and a quencher dye, TAMRA, at the 3' end of the probe. During the reaction, cleavage of the probe separates the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter only if the probe hybridizes to the target. Accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the

reporter fluorescence primarily by Förster-type energy transfer. During PCR, if the target of interest is present, the probe specifically anneals between the forward and reverse primer sites.

MultiScribe Reverse Transcriptase is a recombinant Moloney Murine Leukemia Virus (MuLV) Reverse Transcriptase. AmpliTaq Gold is a thermal stable DNA polymerase. The enzyme has a 5' to 3' nuclease activity, but lacks a 3' to 5' exonuclease activity. AmpErase UNG is a pure nuclease-free, 26-kDa recombinant enzyme encoded by the *Escherichia coli* uracil-N-glycosylase gene. This gene has been inserted into an *Escherichia coli* host to direct expression of the native form of the enzyme. UNG acts on a single- and double-stranded dU-containing DNA. It acts by hydrolyzing uracil-glycosidic bonds at dU-containing DNA sites. The enzyme causes the release of uracil, thereby creating an alkali-sensitive apyrimidinic site in the DNA. The enzyme has no activity on RNA or dT-containing DNA. When two-step RT-PCR is performed with the TaqMan Gold RT-PCR kit, UNG treatment can prevent the reamplification of carryover PCR products. The glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probe is supplied with the TaqMan Gold RT-PCR kit. It utilizes a JOE reporter dye to detect the amplification of the GAPDH control RNA template, a constitutively expressed housekeeping gene.

The reverse transcription reaction contains RNase-free water, 10X TaqMan RT Buffer, 25 mM magnesium chloride, deoxyNTP's mixture, oligo d(T) primers, RNase inhibitor, and MultiScribe Reverse Transcriptase for a total reaction volume of 10 µl per well in a 96-well plate. If more cDNA is needed, a 100 µl reaction volume is utilized. The RT reaction conditions are incubation at 25°C for 10 minutes, RT at 48°C for 30 minutes, and reverse transcription inactivation at 95°C for 5 minutes. The PCR reaction mix contains RNase-free water, 10X TaqMan buffer A, 25 mM magnesium chloride, 10 mM deoxyATP, 10 mM deoxyCTP, 10 mM

deoxyGTP, 20 mM deoxyUTP, TaqMan probe, TaqMan forward primer, TaqMan reverse primer, AmpErase UNG, and AmpliTaq Gold DNA polymerase for a total reaction volume of 50 µl per well in a 96-well plate. The PCR reaction conditions are UNG incubation at 50°C for 2 minutes, AmpliTaq Gold activation at 95°C for 10 minutes, and PCR for 40 cycles with denaturing at 95°C for 15 seconds and anneal/extension at 60°C for 1 minute.

Primer and probes were designed for each of the five radiation response genes. The human androgen receptor primers: F1: CCCTGGCGGCATGGT, F2: ACCCTGGCGGCATGGT, F3: TACCCTGGCGGCATGGT, R1: CCCATTTCGCTTTTGACACA, R2: CCCATTTCGCTTTTGACACAA, and R3: GCCCATTTCGCTTTTGACA. The probe sequence for the androgen receptor is: 6FAM-AGCAGAGTGCCCTATCCCAGTCCCA-TAMRA. The human DORA reverse strand protein 1 (DREV1) primers: F1: GAGGCAGGGTCATCCTTGC, F2: GAGCCAACTAGAGGCAGGGTC, F3: GCCAACTAGAGGCAGGGTCA, R1: CCCACTTGCCACCTACGTTT, R2: TCCCACTTGCCACCTACGTT, and R3: CTCCCACTTGCCACCTACGT. The probe sequence for the DREV1 gene is: 6FAM-CTTGTCCTCCCCTTTCATCCCTATGTGT-TAMRA. The primer sequences for the human creatine kinase gene are: F1: TGCTACCATGGGCACCAGT, F2: TTGCTACCATGGGCACCAGT, F3: TTGCTACCATGGGCACCAG, R1: GCACACACTTTCTGCCGGT, R2: GCACACACTTTCTGCCGGTT, and R3: GGCACTCGGCCATGCA. The probe sequence for the human creatine kinase gene is: 6FAM-TCCTGACCACCGGGTACCTGCTG-TAMRA. The primer sequences for the human eukaryotic translation elongation factor 1 beta 2 are: F1: CACAATTTGCGCGCTCTCT, F2: CCACAATTTGCGCGCTCT, F3: CCACAATTTGCGCGCTC, R1: ACCCATGGTGTCGGCTGTA, R2:

ACCCATGGTGTCGGCTGT, and R3: AACCCATGGTGTCGGCTGTA. The probe sequence for the human eukaryotic translation elongation factor 1 beta 2 is: 6FAM-TCTGCTGCTCCCCAGCTCTCGG-TAMRA. The primer sequences for the human ribosomal protein L27 are: F1: GCCCCTACAGCCATGCTCT, F2: ATCGCCCCTACAGCCATG, F3: TCAGATCGCCCCTACAGCC, R1: CATGGCAGCTGTCACTTTGC, R2: CCCATGGCAGCTGTCACTT, and R3: TCTTGGCGATCTTCTTCTTGC. The probe sequence for the human ribosomal protein L27 is: 6FAM-TGGCTGGAATTGACCGCTACCCC-TAMRA.

RNAi protocol

Introduction

RNA interference was utilized to knockdown the expression of the DREV1 gene in order to assess whether it would affect cell survival. The DREV1 gene was chosen for knockdown due to the fact that it is a novel, uncharacterized gene with little known about it. RNAi was accomplished through the use of siRNAs targeted to regions of the DREV1 gene in order to facilitate knockdown of gene expression.

A mock transfection of the MCF10A cells was first performed. Mock transfection is done without siRNA to ensure success in the RNAi experiment. It is imperative that the siRNA be delivered to the cell efficiently with minimal disruption to the cell. Studies show that transfection disturbs gene expression. This could be due to components in the transfection reagent or the transfection process itself. A mock transfection allows you to determine the degree to which the transfection step affects the expression levels of your target gene. This will allow you to determine whether the transfection is having an effect on gene expression compared to the untreated cells. One limitation of the mock transfection control is that it lacks a siRNA

component. Since lipid-based transfection reagents are cationic, their biological properties, and thus their effects on the cell, are likely to change when complexed with a negatively charged siRNA. A non-functional siRNA control should be included in the siRNA experiment to circumvent this effect. RISC-free siRNAs are a good choice and have been chemically modified to impair their uptake by the RISC. Since they are poor substrates for RISC, they are expected to have few, if any, RNAi-related effects, specific or non-specific. Their effects in the cell, if any, may be through non-RNAi pathways. Thus, these controls can be utilized to check for general side-effects associated with siRNA delivery.

As a further control for potential off-target effects caused by siRNA transfection, siRNA-mediated RNAi experiments should include functional, non-targeting siRNA as a control. Changes in protein or mRNA levels in cells treated with non-targeting siRNA reflect a non-specific baseline cellular response to which the levels in the cells treated with the target-specific siRNA should be compared. Studies have shown that the design of non-targeting control siRNAs are best focused on careful bioinformatic analyses to minimize off-target effects, and that the GC content is a less critical attribute for a non-targeting control siRNA. Non-targeting siRNAs are taken up and processed by RISC, but lack sufficient complementarity to direct target mRNA cleavage. siRNAs must have near-perfect complementarity with mRNA targets to promote RISC-mediated mRNA cleavage. In this experiment siCONTROL non-targeting siRNAs were utilized which are specifically designed to have $\geq$4 mismatches with known human and mouse genes from Dharmacon. Non-targeting siRNAs can be utilized to confirm sequence specificity of the silencing effect observed with experimental, target-specific siRNA. siCONTROL non-targeting siRNAs are strongly recommended over the use of "scrambled" siRNAs, in which the siRNA-sequence has been inverted or altered. Without subsequent testing

and microarray validation this approach is prone to unintended off-target silencing and miRNA effects.

It is also imperative to quantify how efficient transfection occurs in an experiment. A successful transfection is key to accurately identifying the biological functions associated with siRNA-based gene silencing. Poor transfections, for instance those that introduce moderate or high levels of siRNA in only a fraction of the cell population, can generate erroneous or partial phenotypes that can lead to inaccurate and costly assumptions concerning the gene of interest. For this reason, controls that enable quick and accurate quantification of the fraction of transfected cells in a population are critical. The siTOX transfection control by Dharmacon was utilized in this experiment. The siTOX transfection control is a cytotoxic, RNA-based reagent that can be used to optimize nucleic acid transfection efficiencies in different mammalian cell lines. The siTOX transfection control allows for determination of transfection efficiency through the measurement of the level of cell death. Upon entering the cell, the reagent triggers a strong, adverse response, leading to apoptosis and cell death, usually within 24-72 hours. The degree of cell death correlates with the transfection efficiency. Thus in this experiment, roughly 15% of the cells remaining were viable following transfection with the siTOX transfection control, suggesting that 85% of the cells were successfully transfected.

The positive silencing control, cyclophilin B siRNA, was utilized to confirm that siRNA-mediated RNAi was functioning in my experimental system. This control targets both the human and mouse forms of cyclophilin B, a housekeeping gene that is abundantly expressed in most human and mouse cell lines. Since this gene is considered to be non-essential, cell viability is not adversely affected by the addition of this siRNA. The cyclophilin B siRNA is highly efficient as a silencer in a variety of cell lines. The cyclophilin B siRNA typically reduces target

gene expression by 85% or more when silencing is measured at the mRNA level 24 hours after transfection using 100 nM siRNA.

RNAi is accomplished by transfecting siRNA into the parental MCF10A cells targeted to the DREV1 gene. The other four radiation response genes were not targeted in this case due to limited resources and the fact that DREV1 was the only novel, unidentified gene found. RNAi begins by first plating the MCF10A cells at 50% confluency one day prior to transfection with the siRNAs. The oligofectamine reagent and Opti-MEM I reduced serum medium is utilized to increase efficiency of the reaction. The siRNA oligonucleotides concentration was titrated over a dose range of: 100 nM, 20 nM, 10 nM, 5 nM, 1nM, and 0.1 nM. This would allow for the lowest concentration that was the most efficient at silencing to be selected.

The experiment was performed with four siRNAs that were designed to knockdown transcription of the DREV1 gene. The four siRNAs were purchased from Dharmacon as a SMARTpool reagent. SMARTpool reagents are expected to silence target gene expression at the mRNA level by at least 75%. Dharmacon has found that over 85% of the SMARTpool reagents reduce target mRNAs levels by 95% or more. Dharmacon SMARTpool siRNAs are designed with the following features shown to be optimal for effective silencing: twenty-one nucleotide RNA oligonucleotides forming a 19-base-pair duplex core, symmetrical two nucleotide 3'-UU overhangs, and 5'-phosphorylated antisense strand. The sequences are as follows: SMARTpool duplex #1, sense: 5'-GGAAGAGGCUCAAUGUUUGUU-3', antisense: 5'-PO_4-CAACAUUGAGCCUCUUCCUU-3'; duplex #2, sense: 5'-ACACAGAUCUUCUUAAACAUU-3', antisense: 5'-PO_4-UGUUUAAGAAGAUCUGUGUUU-3'; duplex #3, sense: 5'-GGAUGACGCUGUCUUUGUUUU-3', antisense: 5'-PO_4-

AACAAAGACAGCGUCAUCCUU-3'; duplex #4, sense: 5'-GAACUGGGAAGAACAAGUGUU-3', antisense: 5'-PO_4-CACUUGUUCUUCCCAGUUCUU-3'.

Numerous controls were added to the experiment to ensure proper silencing was occurring from the siRNA and not another element. Untreated MCF10A cells, mock transfected MCF10A cells, non-functional RISC complex siRNAs, non-targeting siRNAs, a positive control of an siRNA known to target cyclophilin B, and a transfection efficiency siRNA. Each of these controls, along with the four DREV1 targeted siRNAs, was sampled at 24, 30, 36, and 48 hours post transfection. All of the samples also were measured at 100 nM, 20 nM, 10 nM, 5 nM, 1 nM, and 0.1 nM concentrations. The oligofectamine reagent was complexed with the siRNA and incubated for 20 minutes at room temperature. The complex was then added to the MCF10A cells and incubated for 6 hours. After the 6 hours had passed, DMEM/F12 complete medium was added to the cells. An assay for gene activity was done at 24, 30, 36, and 48 hours post transfection.

Expression of the trapped gene that is believed to be radiation-responsive can be decreased through the use of RNA interference (RNAi). This will allow for the evaluation of the importance the gene has to the survival of the cell. RNAi involves the introduction of a small antisense RNA that will cause decreased expression of the trapped gene. This will be performed in the parental MCF-10A cell line to verify if in fact the transformation observed in the gene-trapped MCF-10A clone is due to the effects of radiation.

Metaphase spreads

Chromatid gaps and breaks and chromosome-type aberrations were analyzed by microscope in MCF10A parental cells with and without radiation treatment. The MCF10A cells

were first transfected with siRNA to knockdown expression of the DREV1 gene and then irradiated with either 2.0 or 4.0 Gy of ionizing radiation. Unirradiated and untransfected MCF10A parental cells were utilized as control to obtain background levels from. The MCF10A cells were synchronized in metaphase by the addition of 10 μl/ml of colcemid. The cells were incubated for four hours and then harvested. The pelleted cells were fixed with a 3:1 mixture of methanol and acetic acid and dropped onto slides. The metaphase spreads were stained with 10% giemsa for counting.

CHAPTER 4-RESULTS

Introduction

The gene-trapped library of human mammary epithelial cells was screened for G418 resistant clones. This library represented 3.0 x 10^6 independent infected clones in which one allele of a functional gene was disrupted by the retrovirus gene-searching vector. After selection by G418 for two weeks, resistance clones were pooled together. The pooled gene trapped clones were then sorted by GFP expression into GFP+ and GFP- sorted pools. These pools were then plated in 96-well plates to obtain single cell clones in order to determine which gene was affected by the gamma irradiation treatment (199).

Results

The gene-trapped MCF-10A clones were then sorted further by flow cytometry into GFP+ and GFP- expression levels, as shown in Figure 4.1.

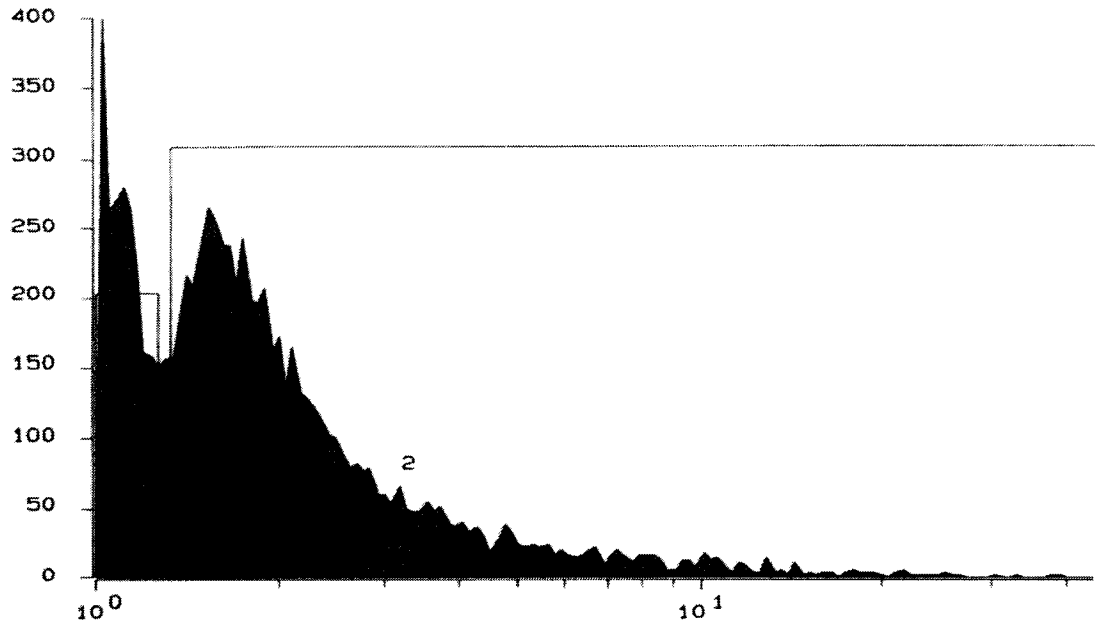
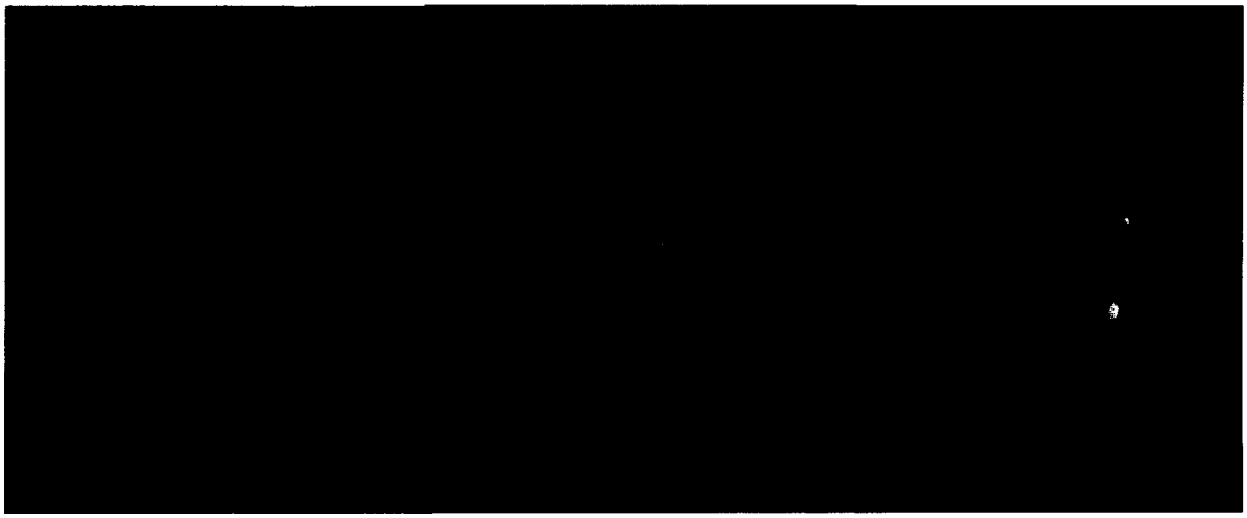


Figure 4.1. Initial sort of gene-trapped MCF-10A clones into GFP+ and GFP- pools. Sort gate 1 is the GFP- pool and sort gate 2 is the GFP+ pool.

After the gene-trapped MCF-10A pooled clones were sorted by flow cytometry they were plated at 1 cell per well in a 96-well tissue culture plate for single cell clone analysis. Basal GFP levels were determined by harvesting cell lysates and performing an ELISA sandwich assay. The cells were then irradiated with a total dose of 2 Gy from a ^{137}Cs gamma source. The gene-trapped MCF-10A clones were then immediately harvested for cell lysates in order to determine the GFP levels following irradiation treatment.

Analysis of the endogenous trapped gene promoter expression level by quantification of GFP by fluorescent microscopy

Gene-trapped MCF10A-1023 cells were plated in 96-well plates to be replica plated in order to obtain single cell clones, not pooled clones. The cells were grown to confluency and expanded into 24-well plates and finally into T25 flasks. The GFP expression was visualized under a fluorescence microscope to verify the expression levels from the endogenous promoter of the trapped gene as either high, medium or low. As shown in Figure 4.2, high, medium and low GFP expression was observed.



A. **B.** **C.**

Figure 4.2. GFP expression of replica plated gene-trapped MCF-10A clones. A. High GFP expressing gene-trapped clone from replica plate. **B.** Medium GFP expressing gene-trapped clone from replica plate. **C.** Low GFP expressing clone from replica plate.

GFP negative gene-trap clone microplate results at 2.0 Gy

The single cell gene trapped MCF10A clones were analyzed for their GFP expression level both before and after a 2.0 Gy dose of ionizing radiation. This was done in order to ascertain which clones were illustrating a radiation response of at least 2-fold in comparison to basal levels. GFP levels were quantitated via a sandwich ELISA assay utilizing the gene trapped clones cell lysate and a monoclonal anti GFP antibody. One 96-well plate containing single cell gene trapped clones was analyzed for both the GFP negative and GFP positive sort plate populations in order to determine whether the initial GFP levels had an effect on the GFP expression following IR treatment.

GFP negative gene-trap clone microplate results at 2.0 Gy

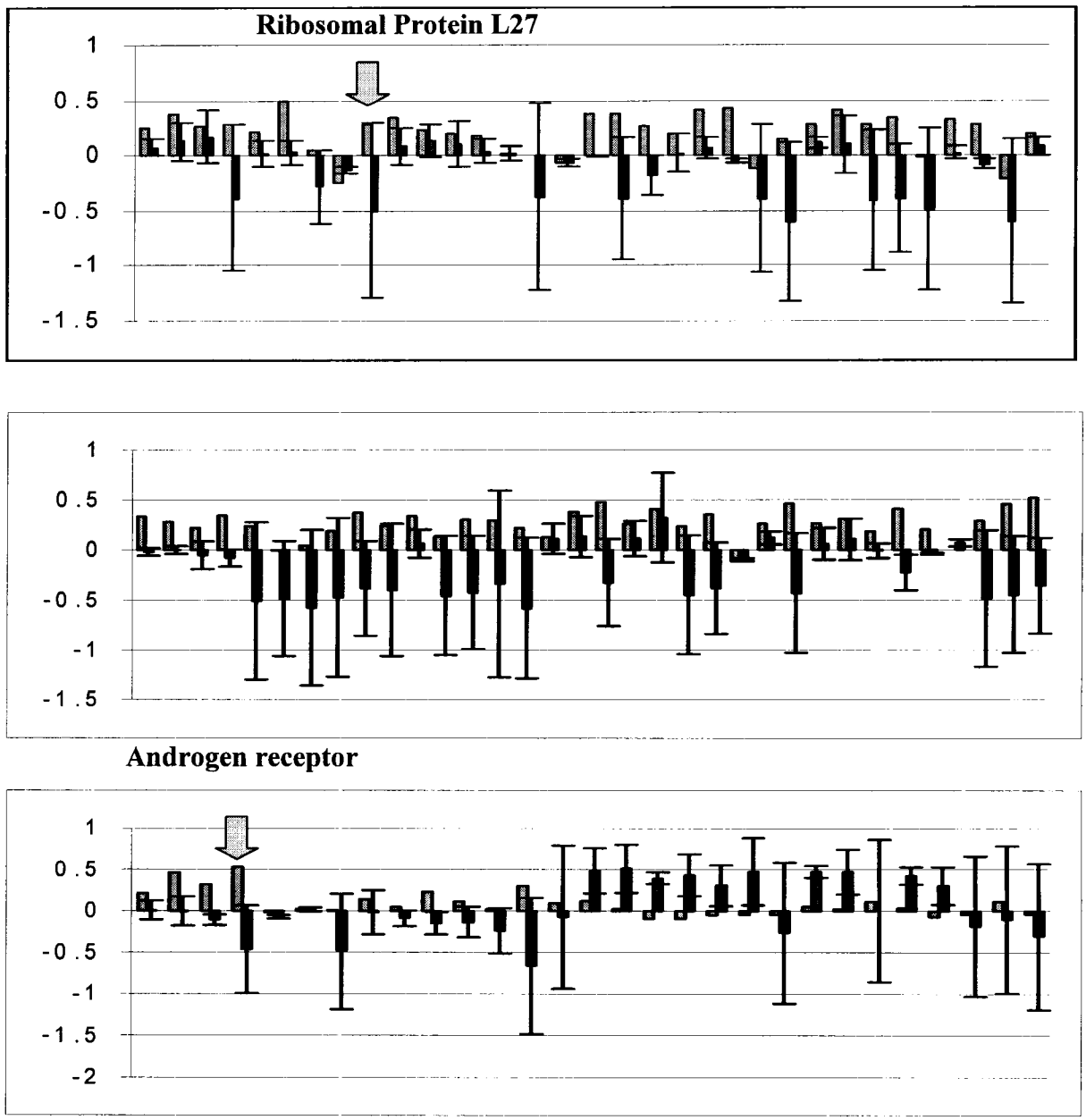


Figure 4.3. GFP negative sort of MCF10A gene trapped single cell clone GFP quantitation by ELISA assay. Each bar on the graph represents a basal level of GFP fluorescence (gray bar) and a GFP fluorescence level following a 2.0 Gy dose of gamma irradiation for a single cell gene trapped clone (black bar). There were a total of 96 single cell gene trapped clones analyzed for the GFP+ sort plate. This is an average of two experiments. The error bars correspond to the standard deviation of the two independent experiments. The identified radiation response genes are identified with an arrow.

GFP positive gene-trap microplate results at 2.0 Gy

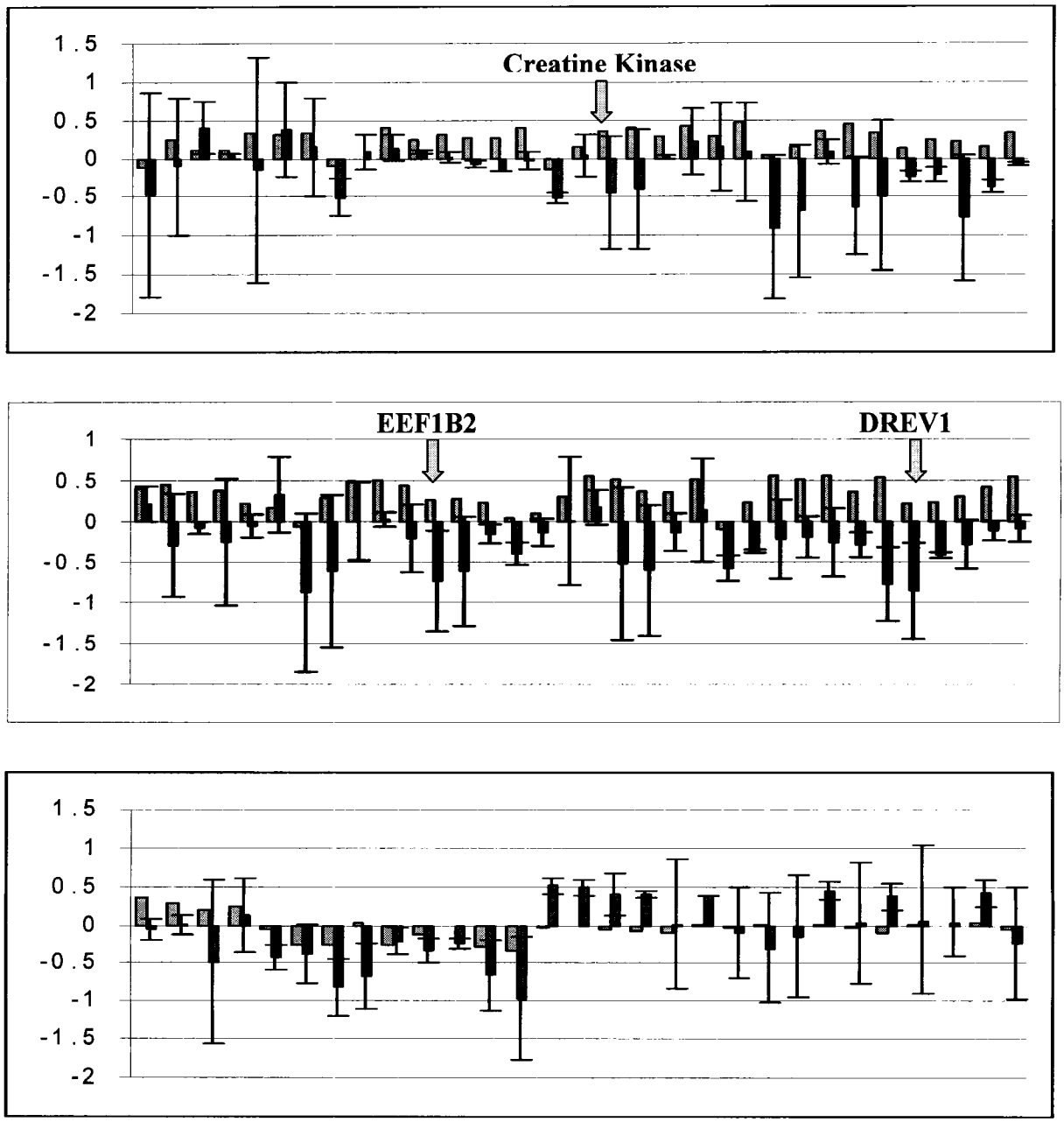


Figure 4.4. GFP positive sort of MCF10A gene trapped single cell clone GFP quantitation by ELISA assay. Each bar on the graph represents a basal level of GFP fluorescence (gray bar) and a GFP fluorescence level following a 2.0 Gy dose of gamma irradiation for a single cell gene trapped clone (black bar). There were a total of 96 single cell gene trapped clones analyzed for the GFP+ sort plate. This is an average of two experiments. The error bars correspond to the standard deviation of two independent experiments. The identified radiation response genes are labeled with an arrow.

GFP expression analysis was performed on both GFP+ and GFP- single cell clones in order to determine if the initial GFP expression level played a role following radiation treatment. It was important to investigate whether a clone with GFP- expression could change to having a GFP+ expression following irradiation, or if the opposite could also be true as well. Since the GFP reporter was a marker of the endogenous gene's expression level, it was possible to determine whether or not the trapped gene was being down or up-regulated in response to the gamma irradiation.

It appeared as though the majority (182 clones, 94.79%) of the single cell gene trapped MCF10A clones were being down-regulated upon radiation treatment independent of their prior GFP expression status. Both clones that were characterized as GFP+ and GFP- prior to radiation treatment were found to be down-regulated upon treatment. It does not appear that the GFP basal expression status changed to the opposite status following radiation treatment for the majority of GFP- clones, though, those clones that were initially GFP+ were down-regulated and appear to have changed to GFP- expressing clones. It was found that only about 10 clones, or 5.21% were being up-regulated in response to the 2.0 Gy radiation dose.

The detection of GFP from a GFP- sort plate and a GFP+ sort plate of single cell gene trapped clones of MCF10A was repeated at both a 1.0 Gy and 2.0 Gy dose of gamma irradiation. This was done to confirm that the same clones from both the GFP- and GFP+ sort plates of single cell gene trapped clones frozen down were still showing a 2-fold or more change upon gamma irradiation treatment. An ELISA assay was utilized once again to quantitate the levels of GFP both basally and after gamma irradiation treatment in the gene trapped MCF10A clones. Upon further analysis it was found that the GFP expression varied a great deal after the experiment was repeated for the 2.0 Gy dose. This variation in the GFP expression following

radiation treatment from one experiment to another could be due to the ELISA assay not being an optimal choice to quantitate the GFP levels. Potentially antibodies are being washed off which could account for some of this variation. Experimental error may also be a cause for the variation seen from one experiment to another. Upon completion of the ELISA experiment I believe it would not be an optimal method for GFP quantification in the future. GFP quantification with antibodies is well utilized in immunohistochemistry, but is not as well defined for ELISA analysis. The variation in GFP expression seen amongst the two experiments contributes to little significance being placed in the results. In the gene trapped clones radiation treatment values ranged roughly +/- 0.20 between the two experiments for the same clone. If this experiment was to be repeated again, I would consider alternate methods for quantitating the GFP expression in the gene trapped clones in order to achieve less variation amongst experiments.

One 96-well plate was analyzed for both GFP+ and GFP- for a total of 192 clones. Out of those, 92 were found to be at least 2-fold or more up- or down-regulated following gamma irradiation treatment. This meant that 47.9% of the gene trapped MCF10A clones were being up- or down-regulated at least 2-fold in response to the ionizing radiation treatment. It does not seem to matter if the gene trapped clone was GFP+ or GFP- prior to radiation treatment. This could be due to the fact that the populations were not sorted correctly and the GFP- population still contained cells that were GFP+. If the sort gate for the flow cytometry analysis was moved higher up then potentially this problem could be eliminated. The endogenous gene into which the pRET retroviral integrated into needed to be determined. In order to accomplish this, 3' RACE was performed.

GFP negative microplate readings at 1.0 Gy

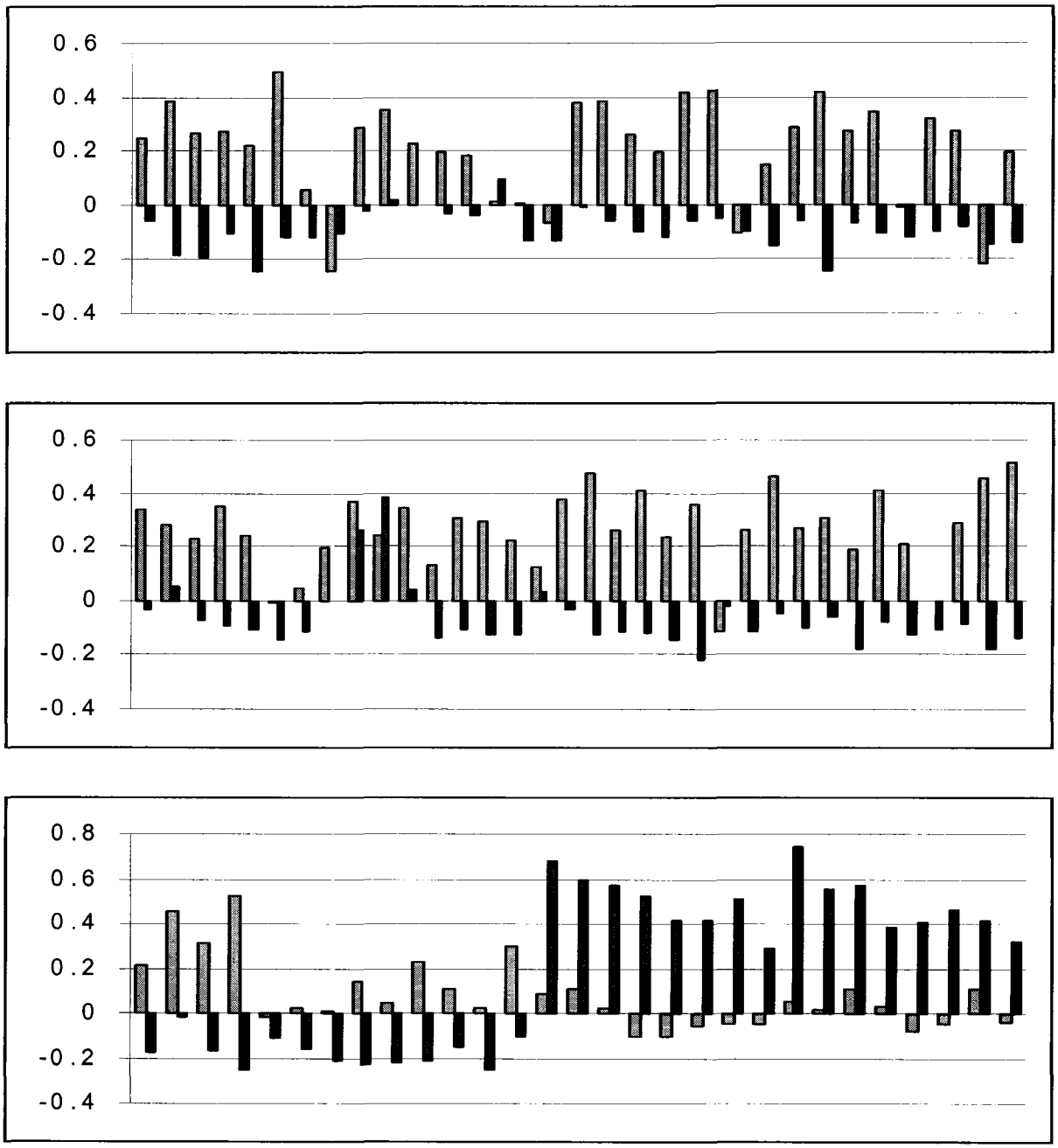


Figure 4.5. GFP negative sort of MCF10A gene trapped single cell clone GFP quantitation by ELISA assay. Each bar on the graph represents a basal level of GFP fluorescence (gray bar) and a GFP fluorescence level following a 1.0 Gy dose of gamma irradiation for a single cell gene trapped clone (black bar). There were a total of 96 single cell gene trapped clones analyzed for the GFP+ sort plate. This data is representative of a single experiment.

GFP positive microplate readings at 1.0 Gy

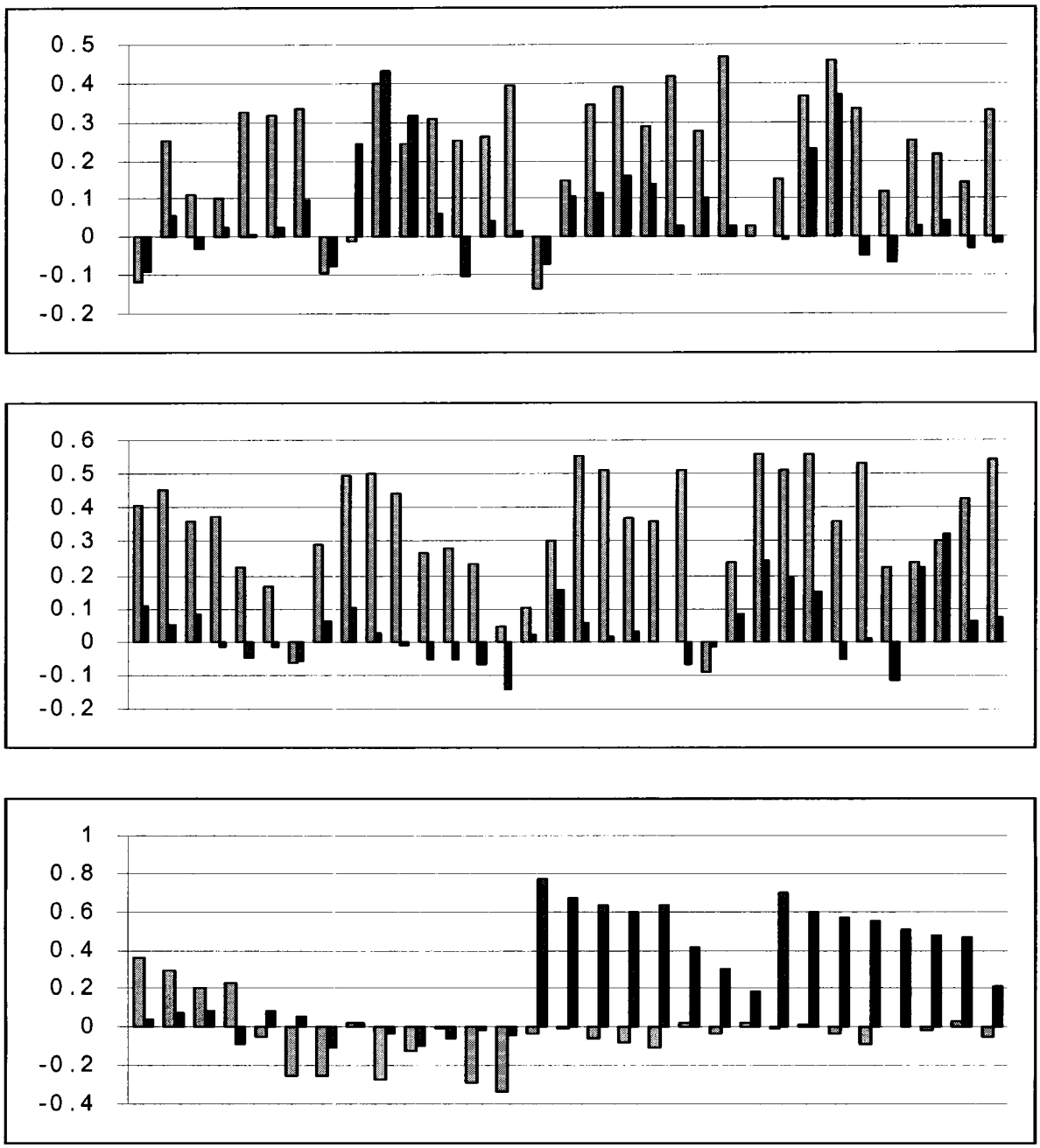


Figure 4.6. GFP positive sort of MCF10A gene trapped single cell clone GFP quantitation by ELISA assay. Each bar on the graph represents a basal level of GFP fluorescence (gray bar) and a GFP fluorescence level following a 1.0 Gy dose of gamma irradiation for a single cell gene trapped clone (black bar). There were a total of 96 single cell gene trapped clones analyzed for the GFP+ sort plate. This data is representative of a single experiment.

3' RACE

The first step in a 3' RACE experiment is the conversion of RNA into cDNA by reverse transcriptase. Two rounds of 3' RACE nested PCR were performed on the MCF10A gene trap clones. The first round amplifies the region upon where the vector integrated through the use of the NEO 1.5 and AD primers and the advantage-GC cDNA polymerase mix. The second round of PCR utilizes the NEO 2.0 and AD-plus primers to amplify the targeted region of the trapped gene. After completion of the second round of 3' RACE, the MCF10A gene trap clone samples were purified from the PCR reaction using the Qiagen QIAquick PCR purification kit. Cloning of the gene trapped clone PCR product into *E.coli* cells is the next step of the experimental procedure. The TOPO TA cloning kit for sequencing is utilized to ligate the PCR product into the pCR4TOPO vector. The ligated gene trapped clones are then transformed into One Shot competent *E. coli* cells and DNA is collected via the Qiagen mini prep kit protocol.

Sequencing Results

Sequencing of the MCF10A gene trapped clones was performed by Davis Sequencing at the University of California at Davis campus in California. The trapped sequences can be found in the appendices. Upon searching by BLAST, six of the 31 clones sent off for sequencing illustrated homology to known genes in the database. The other 25 clones that were sequenced did not yield homology to any known genes. This could have been due to problems in the TOPO TA cloning step where the cloning vector ligated with itself and the trapped gene DNA was not present. When these clones were analyzed by BLAST, homology to cloning vector sequence was found. The following homologies were obtained: clone B5+2 illustrated homology to the human creatine kinase gene, exon 1 located on chromosome 5; clone G10+5 showed homology to the human DORA reverse strand protein 1 (DREV1), clone E8+4 illustrated homology to the

human eukaryotic translation elongation factor 1 beta 2, clone G11-2 showed homology to the human androgen receptor, clone A4-5 illustrated homology to the human ribosomal protein L27, and finally, clone H4+5 showed homology to the human DNA clone RP11-290F20 located on chromosome 20. The last clone was not followed up on due to the fact that it was not a known gene sequence and was felt not worth pursuing.

After identifying the homologous genes the retroviral vector had trapped into in the MCF10A cells, a comparison of the ELISA GFP quantification data was made to compare how much the gene had been down-regulated following gamma irradiation treatment. Clone G11-2, identified homologous to the human androgen receptor gene, was found to have its GFP level down-regulated from .525 to -.652 upon gamma irradiation treatment. Clone G10+5, homologous to the human DORA reverse strand protein 1 (DREV1) gene, was found to have its GFP level down-regulated from .224 to –1.07 upon gamma irradiation treatment. Clone E8+4, homologous to the human eukaryotic translation elongation factor 1 beta 2, was found to have its GFP level down-regulated from .266 to –.951 upon gamma irradiation treatment. Clone B5+2, homologous to the human creatine kinase gene, was found to have its GFP level down-regulated from .345 to -.712 upon gamma irradiation treatment. Lastly, clone A4-5, homologous to the human ribosomal protein L27, was found to have its GFP level down-regulated from .285 to –.780 upon treatment with gamma irradiation. This data is represented in Figure 4.7.

MCF10A gene trapped clones	*Basal GFP level*	*Average GFP level post 2.0 Gy IR*
Clone G11-2 (Androgen receptor)	.525	-.652 +/- .186
Clone G10+5 (DREV1)	.224	-1.07 +/- .20
Clone E8+4 (EEF1B2)	.266	-.951 +/- .22
Clone B5+2 (Creatine Kinase)	.345	-.712 +/- .258
Clone A4-5 (Ribosomal Protein L27)	.285	-.780 +/- .279

Figure 4.7. Analysis of GFP expression changes in MCF10A gene trapped single cell clones following IR treatment. Average GFP expression was taken from two independent experiments.

As a verification for the success of this experimental system, one of the identified radiation response genes had been previously found to be affected by radiation treatment. Creatine kinase is known to be effected by free radical damage induced by gamma irradiation. This supports the hypothesis that the experimental system was indeed identifying genes responding to radiation treatment.

Cell Cycle Analysis

The cell cycle was analyzed on each gene trapped clone, both with and without gamma irradiation treatment, to verify that cell cycle effects were not playing a role in the results being seen. The parental cell line, MCF10A, was also analyzed by flow cytometry to obtain a cell cycle distribution. It was not expected that there would be any differences in the cell cycle distribution of the gene trapped and parental MCF10A cells, but cell cycle analysis via flow cytometry was done in order to rule out the possibility of cell cycle factors contributing to the mRNA expression level results seen. Since the cells used in the experimental conditions were not synchronized, no differences due to cell cycle factors would be expected. The flow cytometry cell cycle analysis for each phase of the cell cycle for each cell type is shown in Figure 4.11.

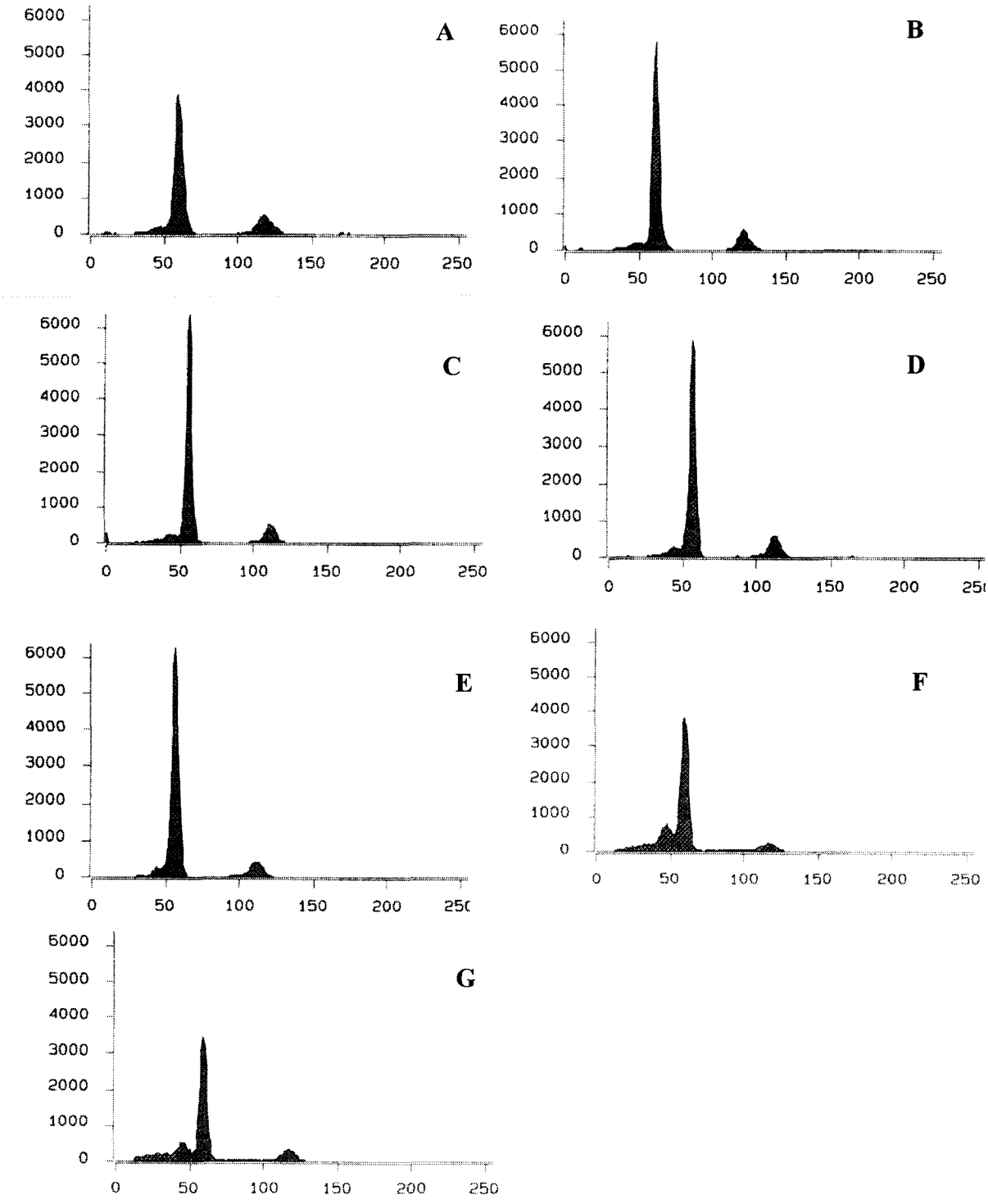


Figure 4.8. Histogram representation of the cell cycle analysis of MCF10A hours post a 2.0 Gy dose of IR. The histograms correspond to **A)** MCF10A parental, **B)** MCF10A cells 2 hours, **C)** 4 hours, **D)** 8 hours, **E)** 12 hours, **F)** 24 hours, and **G)** 30 hours post a 2.0 Gy dose of ionizing radiation.

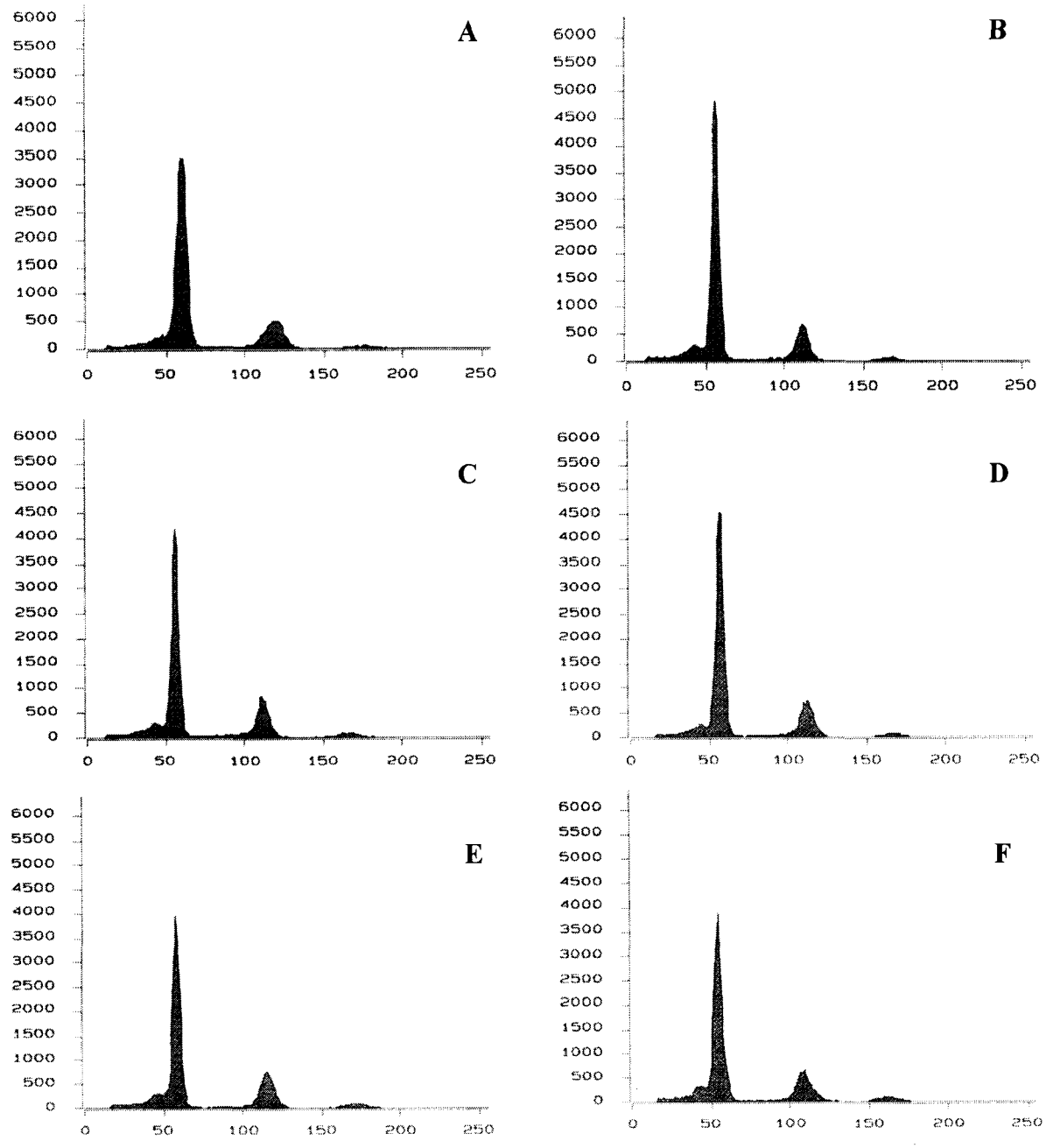


Figure 4.9. Histogram representation of cell cycle analysis of MCF10A gene trapped clones both with and without a 2.0 Gy dose of ionizing radiation. The MCF10A gene trapped clones are as follows from left to right and from top to bottom: **A)** A4-, **B)** A4-IR, **C)** B5+, **D)** B5+IR, **E)** E8+, and **F)** E8+IR. Clone A4- corresponds to human ribosomal protein L27, clone B5+ corresponds to human creatine kinase and clone E8+ corresponds to the human eukaryotic translation elongation factor 1 beta 2.

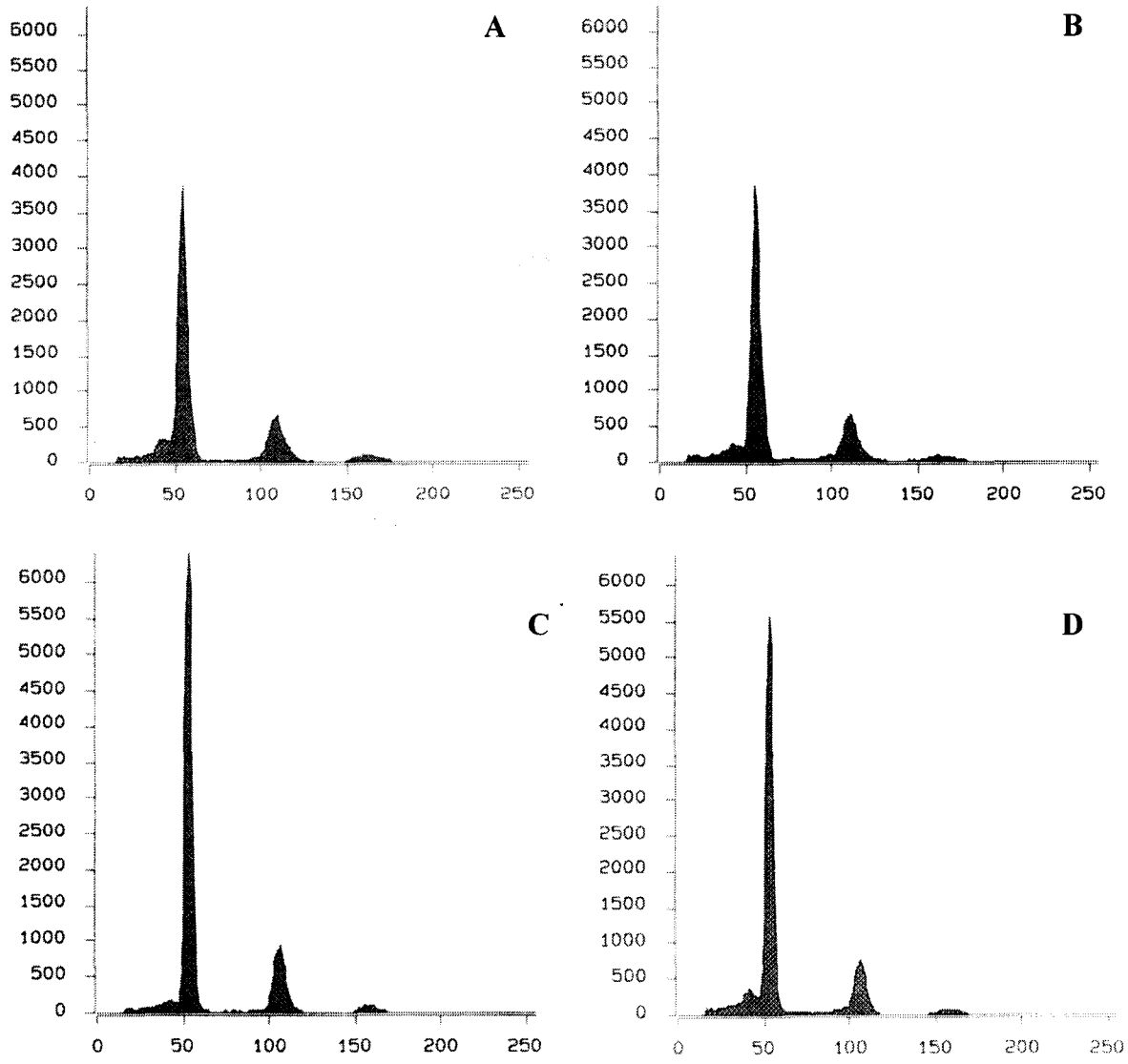


Figure 4.10. Histogram representation of cell cycle analysis of MCF10A gene trapped clones both with and without a 2.0 Gy dose of ionizing radiation. The MCF10A gene trapped clones are as follows from left to right and top to bottom: **A)** G10+, **B)** G10+IR, **C)** G11- and **D)** G11-IR. Clone G10+ corresponds to the DREV1 gene and clone G11- corresponds to the human androgen receptor.

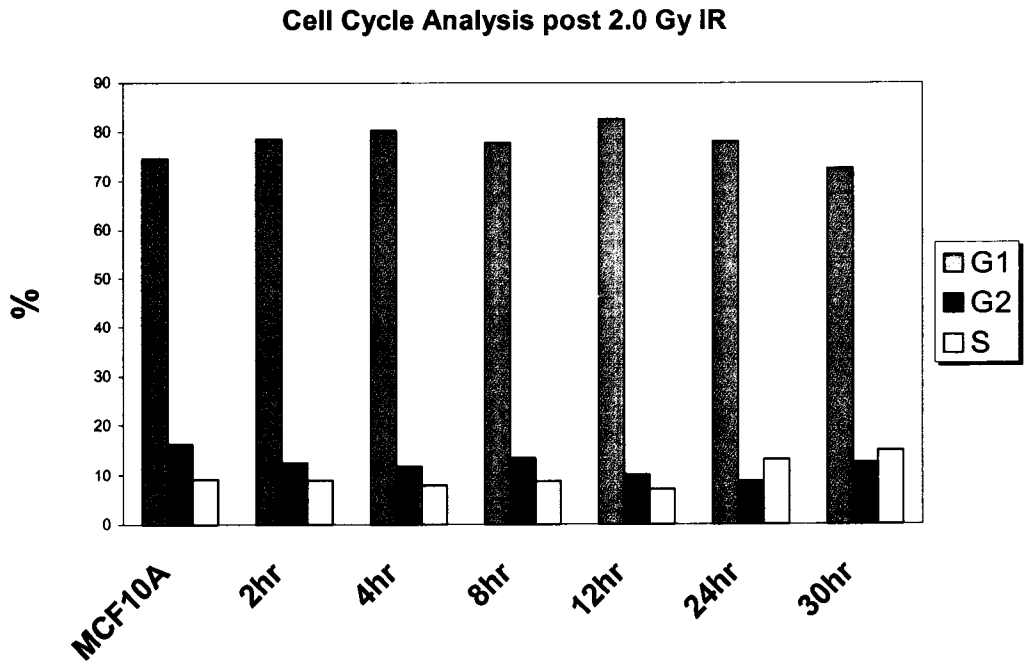


Figure 4.11. Cell cycle analysis of MCF10A parental cells hours post a 2.0 Gy dose of IR. The MCF10A parental cells were analyzed by flow cytometry for the percentage of cells in G1, G2, and S phase at 2, 4, 8, 12, 24 and 30 hours post a 2.0 Gy dose of IR.

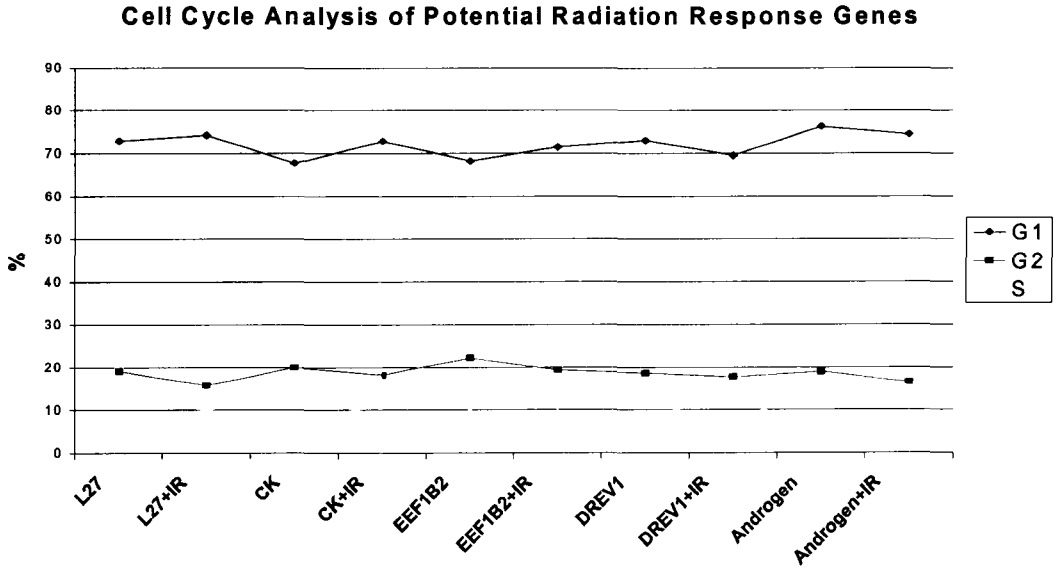


Figure 4.12. Cell cycle analysis of MCF10A gene trapped clones both with a 2.0 Gy dose of IR and without. MCF10A gene trapped clones were analyzed by flow cytometry for the percentage of cells in G1, G2, and S phase after a 2.0 Gy dose of IR. This was compared to the gene trapped clones without IR. The analysis was done on cells harvested within 30 minutes post IR.

Real-time PCR Results

Real-time PCR is a method of quantitating the amount of mRNA present in a sample. The PCR amplification plots and real-time PCR analysis of the five radiation response genes treated with varying doses of ionizing radiation were analyzed relative to MC10A. It is critical to examine whether the mRNA expression level of the five genes of interest vary upon radiation treatment in the MCF10A parental cells. If a radiation response is seen in the parental cells, this would allow the conclusion to be made that the five genes are truly experiencing radiation effects.

The mRNA levels of the human androgen receptor were analyzed in the gene-trapped androgen GFP- clone, from which it was first identified through sequencing analysis. The levels of relative mRNA expression were compared to the parental MCF10A cell line and the breast cancer cell line MCF7 as seen in Figure 4.13. The gene trap androgen GFP+ clone, the opposite clone originating from the GFP + sort population, was analyzed to see if the amount of GFP expression affected the relative mRNA expression levels. As illustrated in Figure 4.13, the relative levels of human androgen receptor mRNA are down-regulated. Upon statistical analysis, it was determined that the results are extremely significantly different from the parental levels with $P = < .0001$. MCF7 cells were utilized to represent what might be seen in breast cancer cells. Androgen receptor mRNA was expressed in the breast cancer cell line, MCF7, but not in the parental human mammary epithelial cell line, MCF10A.

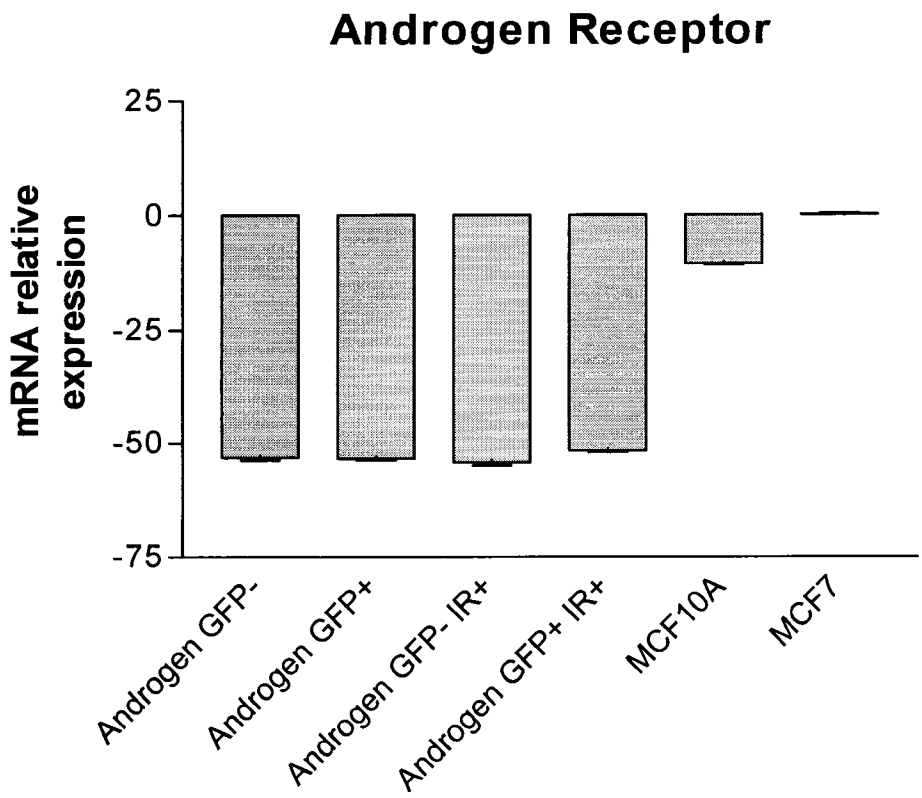


Figure 4.13. Real-time PCR analysis of the androgen receptor mRNA expression levels in the following cell types: Gene trapped androgen clone GFP+ and GFP- with & without 2.0 Gy IR, MCF10A, and MCF7. The mRNA expression was determined relative to the MCF10A parental cell line. The error bars are representative of the standard deviation of triplicate samples for a single experiment.

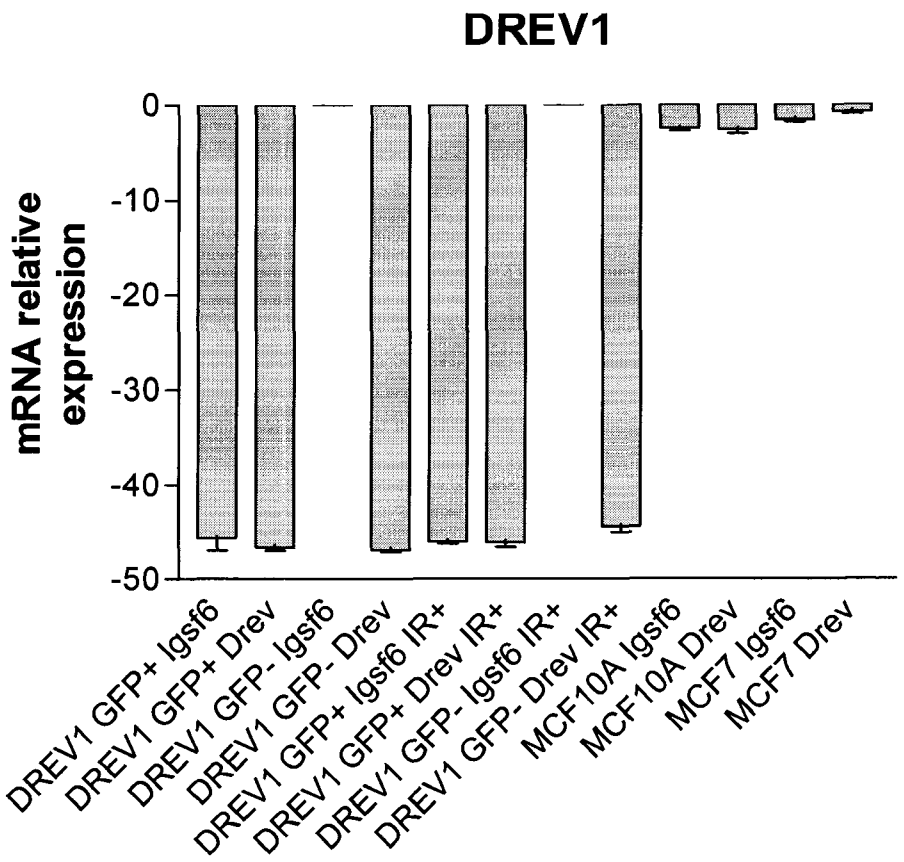


Figure 4.14. Real-time PCR analysis of DREV1 mRNA expression level in the gene trapped DREV1 clone GFP +/- with and without 2.0 Gy IR. The relative mRNA expression was measured relative to the MCF10A parental cell line. The error bars are representative of the standard deviation of triplicate samples from one experiment. The Drev and Igsf6 primers were utilized in the RT step.

The relative mRNA expression levels were quantitated for the DORA reverse strand protein 1 (DREV1) gene in the gene-trapped DREV1 clone GFP +, from which it was initially identified. As illustrated in Figure 4.14, the relative levels of mRNA of the DREV1 gene are down-regulated in comparison to parental levels. Upon statistical analysis, the DREV1 gene trapped clones both with and without ionizing radiation treatment were found to be extremely significantly down-regulated relative to parental MCF10A cells with $P = .0001$. The Igsf6 and Drev in the figure correspond to the gene specific primers used in the reverse transcription reaction upon conversion of the mRNA into cDNA. The Igsf6 primer amplifies the small gene DORA located on the complement strand of DREV1 within intron 4.

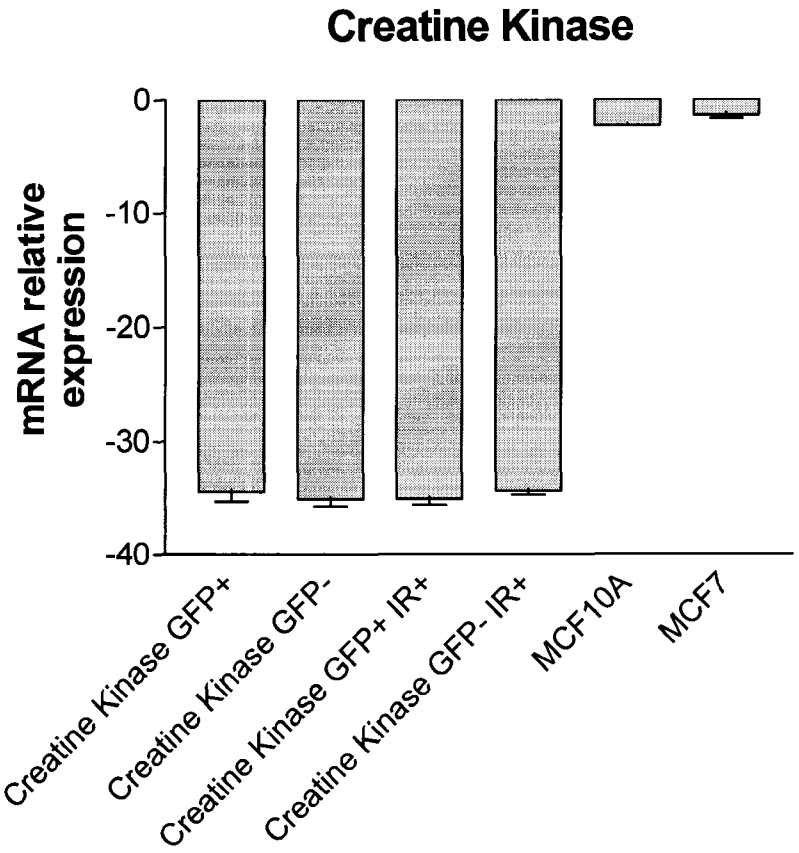


Figure 4.15. Real-time PCR analysis of the creatine kinase mRNA levels in the gene trap clone B5+/-, MCF10A and MCF7. The relative mRNA levels of creatine kinase were quantitated in the gene trap creatine kinase clone GFP +, from which it was originally identified in. The relative mRNA expression was measured relative to the MCF10A parental cell line. Ionizing radiation treatment was performed at 2.0 Gy. The error bars are representative of the standard deviation of triplicate samples in one experiment.

The relative mRNA expression levels were quantitated for the human creatine kinase gene in the gene trapped creatine kinase clone GFP+, from which was initially identified from. As illustrated in Figure 4.15, the relative levels of mRNA of the human creatine kinase gene were found to be down-regulated in comparison to parental levels. The mRNA levels were found to be extremely statistically significant in comparison to MCF10A parental levels upon treatment with 2.0 Gy of IR as well without treatment. Upon statistical analysis the mRNA relative values were found to be extremely statistically significant with P = .0001 to P = .0003. Therefore, in this instance, the mRNA expression of the MCF10A gene trapped clones are being affected by the radiation treatment.

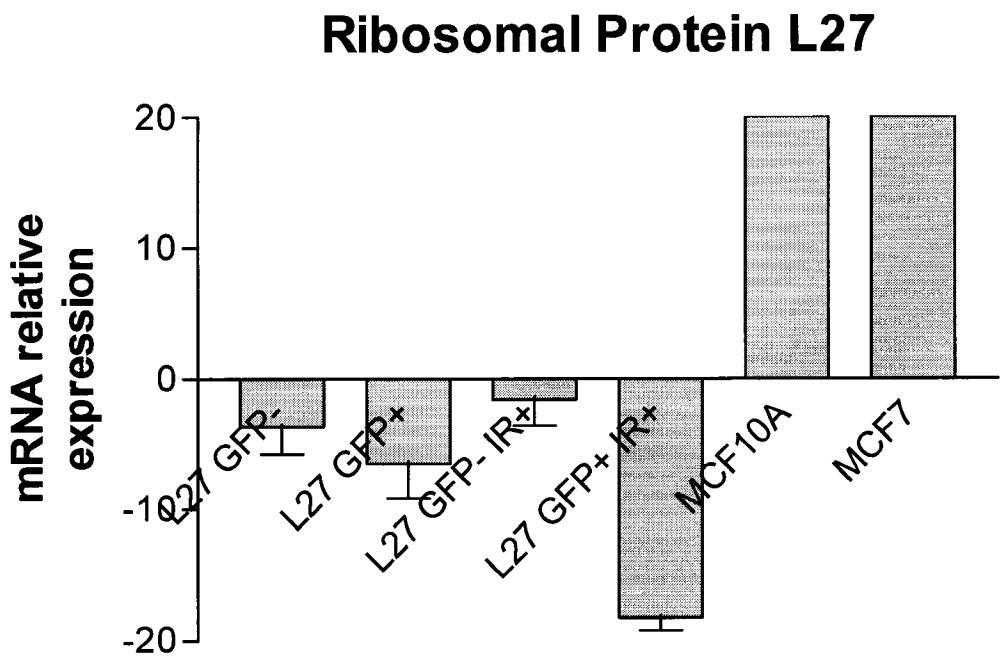


Figure 4.16. Real-time PCR analysis of ribosomal protein L27 mRNA expression levels in gene trap ribosomal protein L27 clone GFP -/+, MCF10A and MCF7. The relative mRNA expression was measured relative to the MCF10A parental cell line. The error bars are representative of the standard deviation of triplicate samples in one experiment. The ionizing radiation dose analyzed in this experiment was 2.0 Gy.

The relative mRNA expression level of human ribosomal protein L27 was quantitated in the gene trap ribosomal protein L27 clone GFP -, which it was first identified in through sequencing. The relative level of mRNA expression in gene trap ribosomal protein L27 clone GFP+/- after a 2.0 Gy dose of gamma irradiation was found to be extremely statistically significant. Down-regulation in comparison to MCF10A parental levels was also observed in the gene trap ribosomal protein L27 clone GFP +/- without radiation treatment. Upon statistical analysis the relative mRNA values were found to be extremely statistically significant with P = .0016 to P = .0074. The relative mRNA expression levels of the MCF10A gene trapped clones with and without radiation treatment are extremely significant from the parental cells.

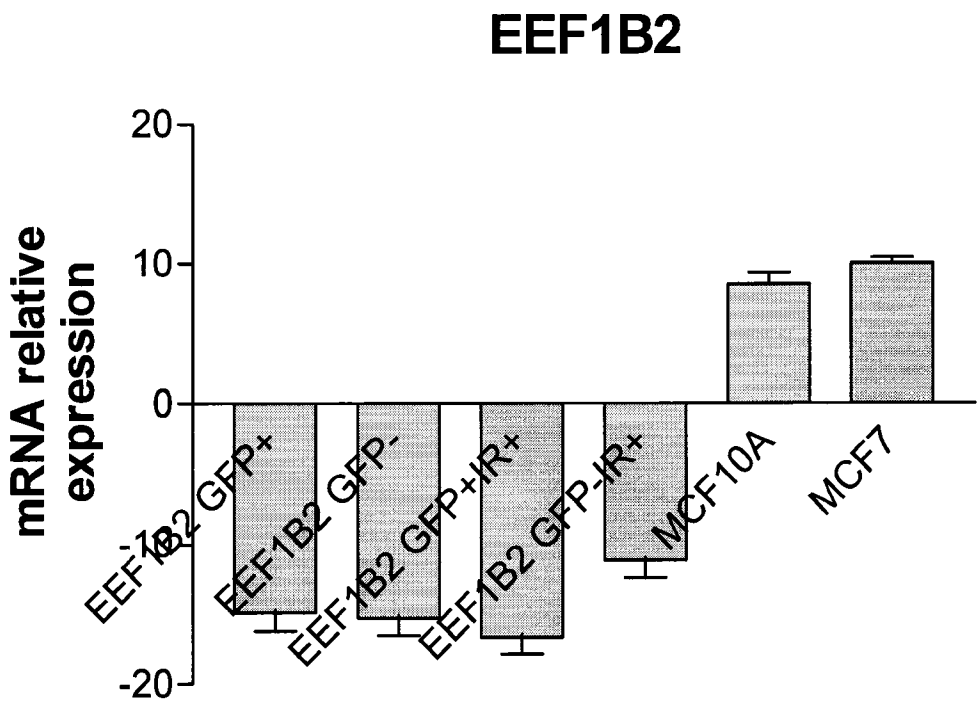


Figure 4.17. Real-time PCR analysis of EEF1B2 mRNA expression levels in gene trap EEF1B2 clone GFP +/-, MCF10A, and MCF7. Relative mRNA expression levels were quantitated in the gene trap EEF1B2 clone GFP +, in which EEF1B2 was first identified as being trapped, MCF10A, and in MCF7. The relative mRNA expression levels were measured relative to the parental MCF10A cell line. The error bars are representative of the standard deviation of triplicate samples in one experiment. A 2.0 Gy dose of ionizing radiation was utilized in this experiment.

The relative mRNA expression level of EEF1B2 following a 2.0 Gy dose of IR was significantly down-regulated in comparison to the parental MCF10A. The relative mRNA expression of both the MCF10A gene trap EEF1B2 clone GFP+ and GFP- following a 2.0 Gy dose of ionizing radiation was found to be down-regulated relative to the parental. The clones were also found to be down-regulated relative to the parental cells without ionizing radiation treatment. Upon statistical analysis, the gene trap clones were found to be very statistically significantly different from parental values with $P = .0017$ to $P = .0030$.

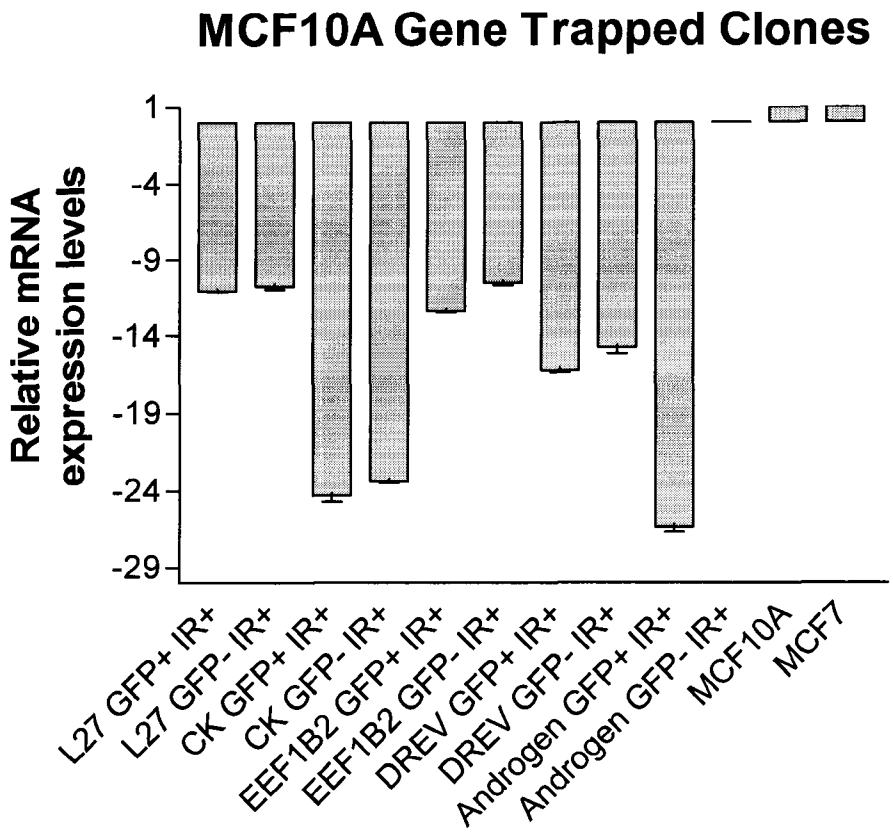


Figure 4.18. Relative mRNA expression of the gene trapped clones following a 2.0 Gy dose of IR. All five radiation response genes were down-regulated relative to the parental in the gene trapped clones. The mRNA expression levels were measured relative to the parental MCF10A cells. The error bars are representative of the standard deviation of triplicate samples in one experiment. A 2.0 Gy dose of ionizing radiation was utilized in this experiment.

In Figure 4.18, the relative mRNA expression levels of the MCF10A gene trapped clones were analyzed relative to the parental cells. Both the GFP+ and GFP- MCF10A gene trapped clones were analyzed relative to the parental following a 2.0 Gy dose of ionizing radiation. For each clone, the basal GFP expression level does not appear to factor into the radiation response being seen. The gene trapped clones at both the GFP+ and GFP- expressing levels were found to be down-regulated relative to the parental. Upon statistical analysis it was found that the relative mRNA expression levels were very statistically significant relative to the parental cells with P = .0030 to P = .0042.

Upon analysis of the MCF10A gene trapped clones a radiation response does appear to be happening in response to the 2.0 Gy dose of ionizing radiation. In order to determine if this radiation response is indeed real, and not due to the gene trapping itself, real-time PCR analysis was performed in order to determine the mRNA expression levels of the five radiation response genes in the MCF10A parental cells. This response was also analyzed at 0.10, 0.25, 0.50, 1.0 and 4.0 Gy doses to verify that a radiation response was occurring at doses other than 2.0 Gy as well. The radiation response was also investigated in various other cell types in order to conclude one way or the other if the radiation response being seen was cell type specific.

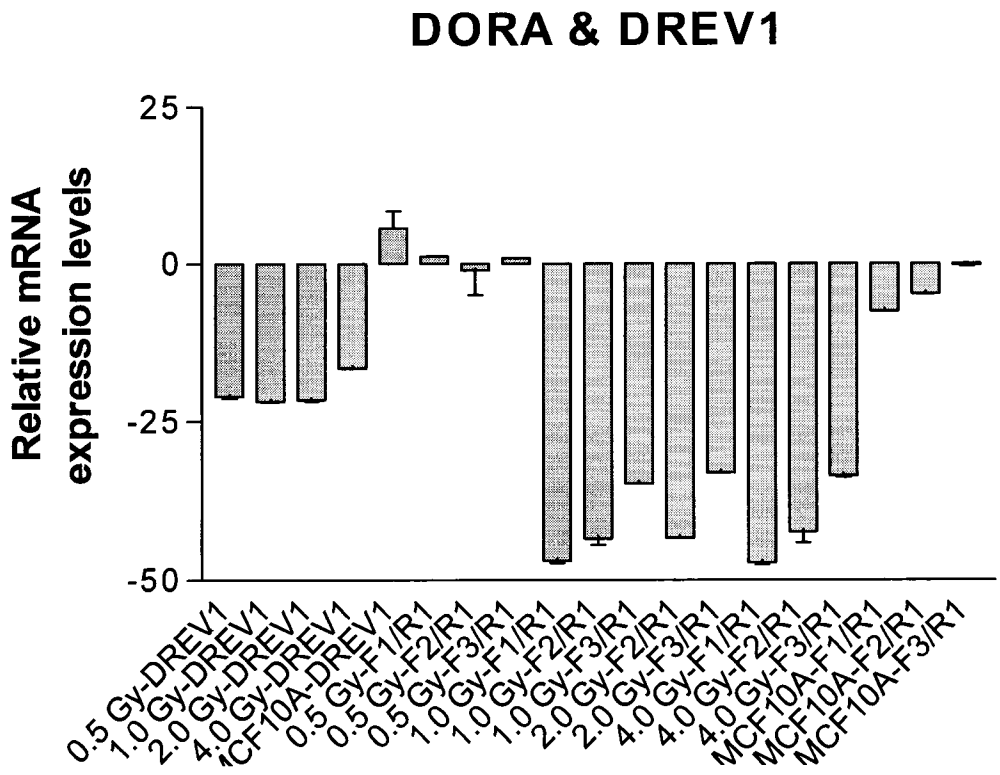


Figure 4.19. Real-time PCR analysis of the mRNA expression levels of the DREV1 gene and the small gene DORA located within DREV1's intron 4 in MCF10A cells following a 0.5, 1.0, 2.0, and 4.0 Gy dose of IR relative to the parental. The F/R corresponds to the primer sets utilized to amplify the DORA gene during real-time PCR. The mRNA expression levels were quantitated relative to the parental MC10A cells. Error bars correspond to the standard deviation of triplicate samples in one experiment.

An interesting characteristic of the DREV1 gene is that a smaller, more characterized gene is located on its complement strand within intron 4. This gene is known as DORA, or IGSF6, and is a novel member of the immunoglobulin superfamily. DORA stands for down-regulated by activation. The relative mRNA expression levels of the two genes were analyzed after the following doses of ionizing radiation: 0.5, 1.0, 2.0 and 4.0 Gy. This was done in order to verify if the mRNA levels of the two genes were related. The mRNA levels were analyzed relative to the parental MCF10A cells.

Upon analysis, there appears to be no relationship between the expressions of the two mRNAs since they were both down-regulated in response to ionizing radiation. If a difference were seen, the opposite effect following radiation treatment would be observed. This is what would be expected based on reports in the literature (1, 71, 98, 180, 182). The expression appears to not take place in the same cell type, consistent with previous reports of this type of gene organization. In the literature, DREV1 and DORA mRNA expression levels had not yet been determined in human mammary epithelial cells (19). In order to prove that the two mRNAs expression levels were not related, real-time PCR analysis was performed. Upon statistical analysis, it was found that the relative mRNA expression levels of DREV1 gene in MCF10A irradiated samples were very statistically significant relative to MCF10A parental cells with $P = .0048$ to $P = .0073$. The mRNA relative expression levels of the DORA gene in the irradiated MCF10A cells were found to be extremely statistically significant relative to parental cells with $P = < .0001$.

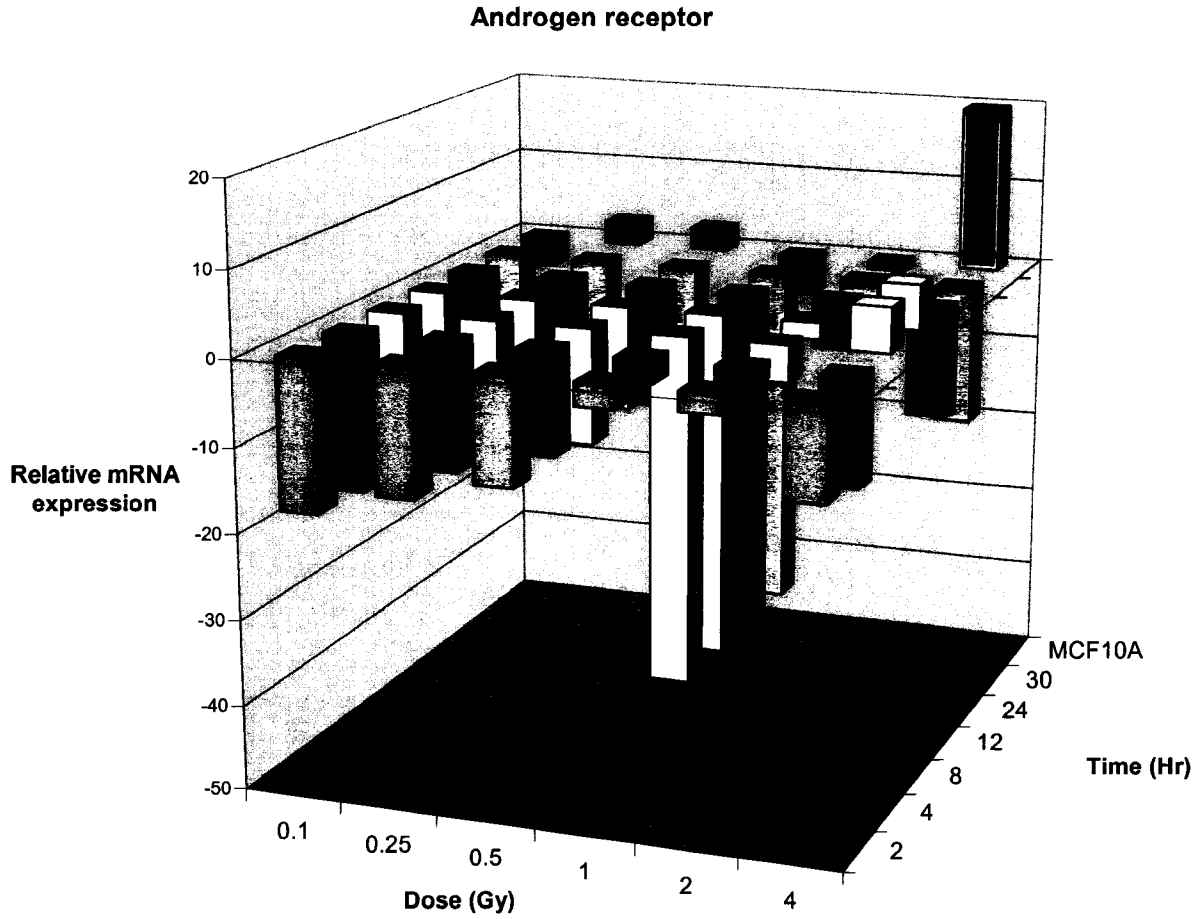


Figure 4.20. Real-time PCR analysis of the mRNA expression levels of the human androgen receptor gene in MCF10A cells at various times following varying doses of IR relative to the parental. The mRNA relative expression was quantitated relative to the parental MCF10A cells. The standard deviation for each sample point is given in the appendices.

The androgen receptor mRNA relative expression was analyzed by real-time PCR in the parental MCF10A cells at 0.10, 0.25, 0.50, 1.0, 2.0 and 4.0 Gy. At each dose of ionizing radiation, samples were taken at 2, 4, 8, 12, 24 and 30 hours post radiation treatment in order to verify at what point the mRNA expression was the most effected. As illustrated in Figure 4.20, the relative mRNA expression of the androgen receptor is down-regulated in MCF10A cells following radiation treatment. Upon statistical analysis, the samples taken following a dose of 0.10 Gy were found to be very statistically significant relative to the parental with P = .0033 to P = .0049. At a dose of 0.25 Gy, the relative mRNA expression of the androgen receptor was determined to be

very to extremely statistically significant relative to the parental with P = .0063 to P = <.0001. At 12 hours post ionizing radiation treatment, the sample was found to not be statistically significant due to greater standard deviation among the triplicate samples run. At a dose of 0.50 Gy, the androgen receptor mRNA expression level relative to the parental was found to be very statistically significant relative to the parental with P = .0017 to P = .0054. Following statistical analysis, the relative mRNA expression of the androgen receptor post a 1.0 Gy dose of ionizing radiation was determined to be statistically to extremely statistically significant relative to the parental with P = .0317 to P = <.0001. At a dose of 2.0 Gy, the androgen receptor mRNA expression relative to the parental was not found. The relative mRNA expression had been determined previously at less than one hour following radiation treatment, but in this instance, expression could not be determined due to experimental error. At a dose of 4.0 Gy, the androgen receptor mRNA expression relative to the parental was found to be extremely statistically significant relative to the parental with P = <.0001 in all instances.

In Figure 4.20, the relative mRNA expression of the androgen receptor is more highly down-regulated at lower doses of ionizing radiation than at the higher doses. At a dose of 1.0 Gy a high amount of down-regulation is occurring that does not follow the pattern seen at the other doses. This could be due to a radiation effect specific for 1.0 Gy that causes increased down-regulated of the androgen receptor in the MCF10A cells.

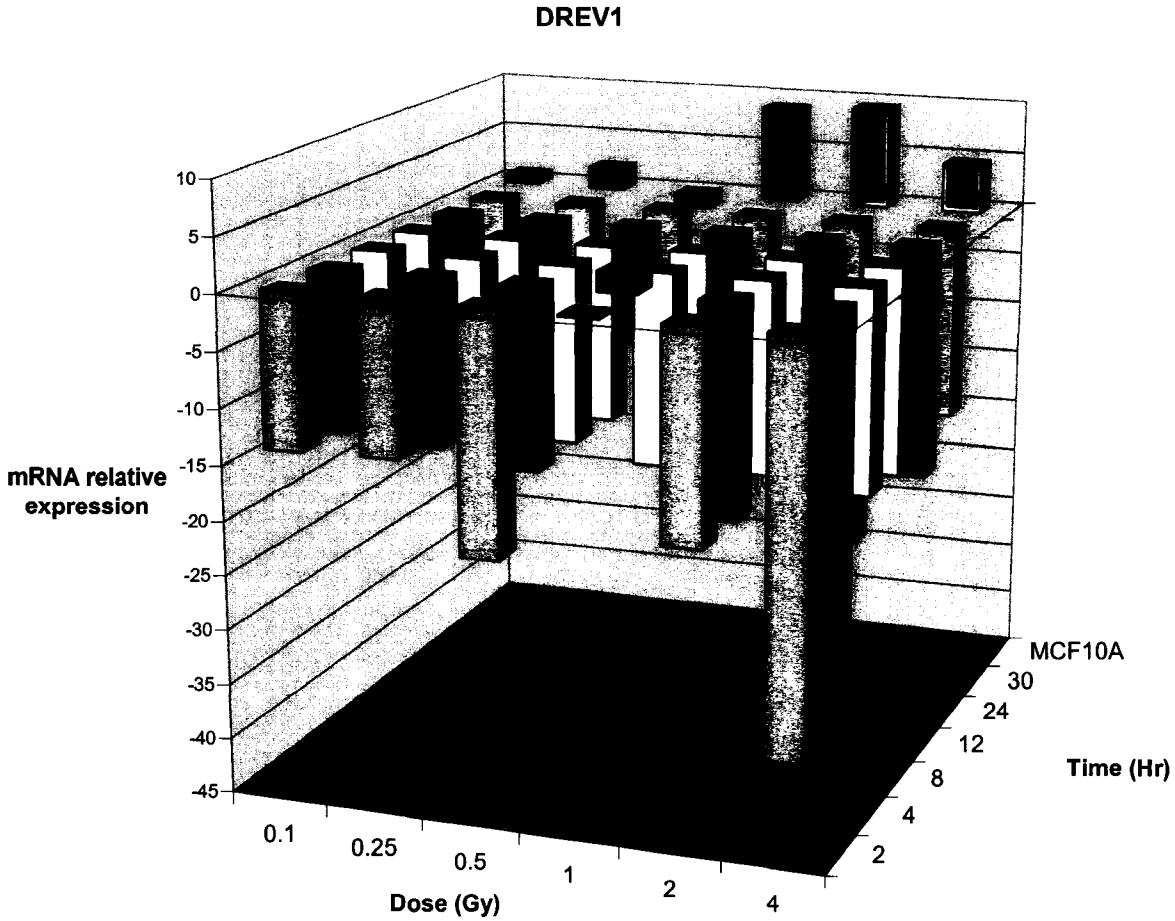


Figure 4.21. Real-time PCR analysis of mRNA expression of the DORA reverse strand protein 1 (DREV1) gene at various hours post varying doses of IR in MCF10A cells. The relative mRNA expression levels were measured relative to the parental MCF10A cells. The standard deviation of triplicate samples of one experiment are given in the appendices.

The DORA reverse strand protein 1 mRNA relative expression was analyzed by real-time PCR in the parental MCF10A cells at 0.10, 0.25, 0.50, 1.0, 2.0 and 4.0 Gy. At each dose of ionizing radiation, samples were taken at 2, 4, 8, 12, 24 and 30 hours post radiation treatment in order to verify at what point the mRNA expression was the most effected. As illustrated in Figure 4.21, the relative mRNA expression of DREV1 is down-regulated in MCF10A cells following radiation treatment. Upon statistical analysis, the samples taken following a dose of 0.10 Gy were found to be very statistically significant relative to the parental with P = .0021 to P = .0057. At a dose of 0.25 Gy, the relative mRNA expression was determined to be statistically to extremely statistically significant with P = .0272 to P = <.0001. Upon statistical analysis, the relative mRNA expression of DREV1 at a 0.50 Gy dose was found to be not quite statistically significant with P = .0814 to .1792. This was due to the high standard deviation values in the triplicate samples run. At a dose of 1.0 Gy the DREV1 relative mRNA expression was determined to be very to extremely statistically significant with P = .0048 to P = <.0001. Upon a dose of 2.0 Gy, the relative mRNA expression of DREV1 relative to the parental was found to be very to extremely statistically significant with P = .0011 to P = .0009. Following a dose of 4.0 Gy, DREV1 relative mRNA expression was determined to be very to extremely statistically significant relative to the parental with P = .0014 to P = .0004.

In Figure 4.21, there is more of an effect seen at the higher doses of ionizing radiation than at the lower doses. At 1.0 Gy, the relative mRNA expression is not as highly down-regulated relative to the parental as at the other doses. This could be an effect specific to the 1.0 Gy dose.

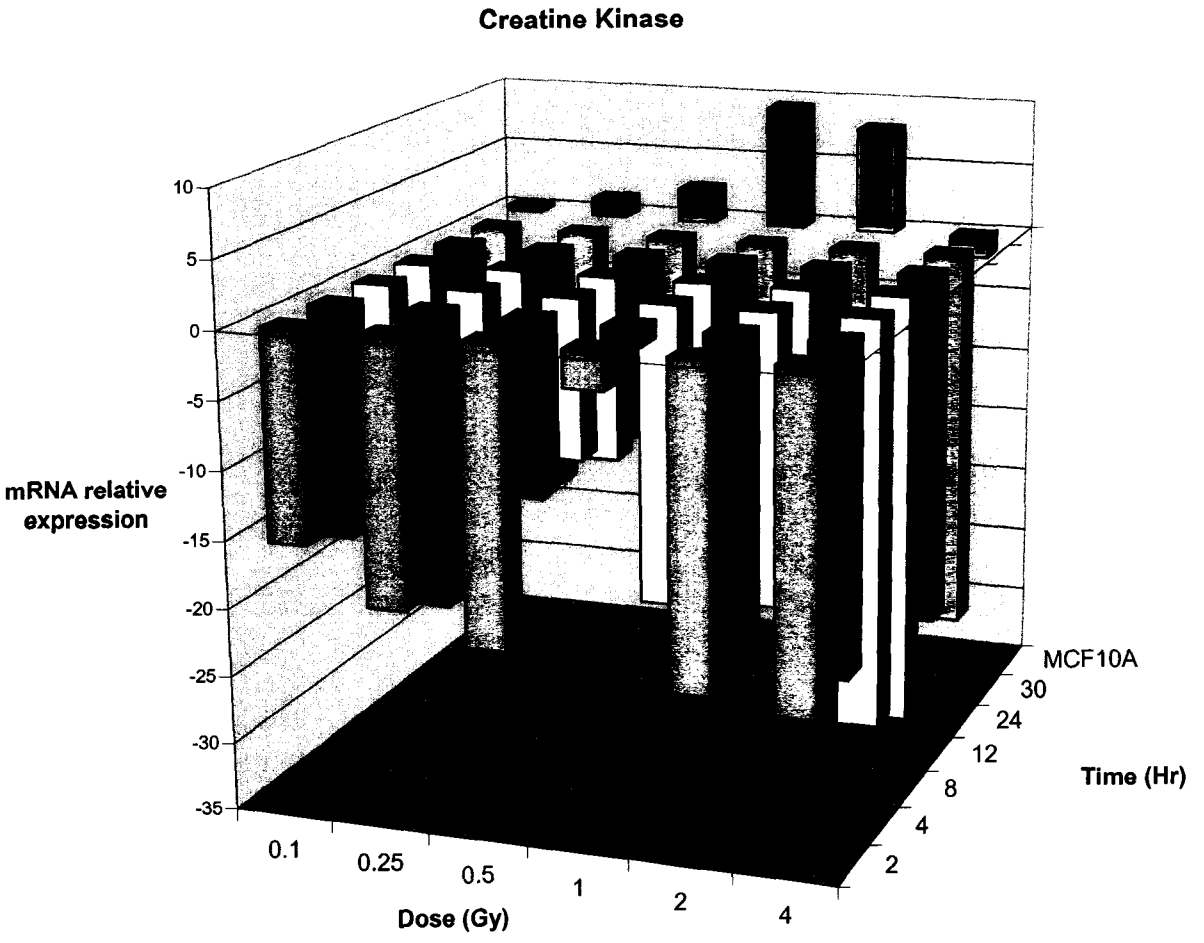


Figure 4.22. Real-time PCR analysis of mRNA expression of the human creatine kinase gene at various hours post varying doses of IR in MCF10A cells. The mRNA expression is relative to the parental MCF10A cells. The standard deviation of triplicate samples in one experiment is given in the appendices.

The creatine kinase mRNA relative expression was analyzed by real-time PCR in the parental MCF10A cells at 0.10, 0.25, 0.50, 1.0, 2.0 and 4.0 Gy. At each dose of ionizing radiation, samples were taken at 2, 4, 8, 12, 24 and 30 hours post radiation treatment in order to verify at what point the mRNA expression was the most effected. As illustrated in Figure 4.22, the relative mRNA expression of creatine kinase is down-regulated in MCF10A cells following radiation treatment. Upon statistical analysis, the samples taken following a dose of 0.10 Gy were found to be very statistically significant relative to the parental with P = .0020 to P = .0035. Following a dose of 0.25 Gy, the relative mRNA expression of the creatine kinase gene was found to be statistically to very statistically significant relative to the parental with P = .0117 to P= .0002. It was determined that at 4 hours post a 0.25 Gy dose of ionizing radiation the sample was not statistically significant due to high standard deviation of the triplicate samples in the real-time PCR run. Upon statistical analysis of a dose of 0.50 Gy, the relative mRNA expression level of the creatine kinase gene was found to be very to extremely statistically significant relative to the parental with P = .0016 to P = .0002. At 2 hours post the 0.50 Gy dose of ionizing radiation the relative mRNA expression of the creatine kinase gene was found not to be statistically significant due to a high standard deviation amongst the triplicate samples in the real-time PCR run. Following a 1.0 Gy dose of ionizing radiation, the relative mRNA expression of the creatine kinase gene was determined to be very to extremely statistically significant relative to the parental with P = .0033 to P = .0001. Upon a 2.0 Gy dose of ionizing radiation, the relative mRNA expression of creatine kinase was found to be extremely statistically significant relative to the parental with P = .0001 in all instances. At a dose of 4.0 Gy, the relative mRNA expression was determined to be very to extremely statistically significant relative to the parental with P = .0045 to P = <.0001.

The dose-rate effect seen in Figure 4.22 is that higher down-regulation of the relative mRNA levels is occurring at the lower doses of ionizing radiation. A dose effect is once again seen at 1.0 Gy, where the down-regulation of the creatine kinase is not to the same degree as that seen at the other doses. This once again could be some effect of the dose itself in which lesser effects are occurring due to potential pathways the cell has developed to deal with a dose of that magnitude.

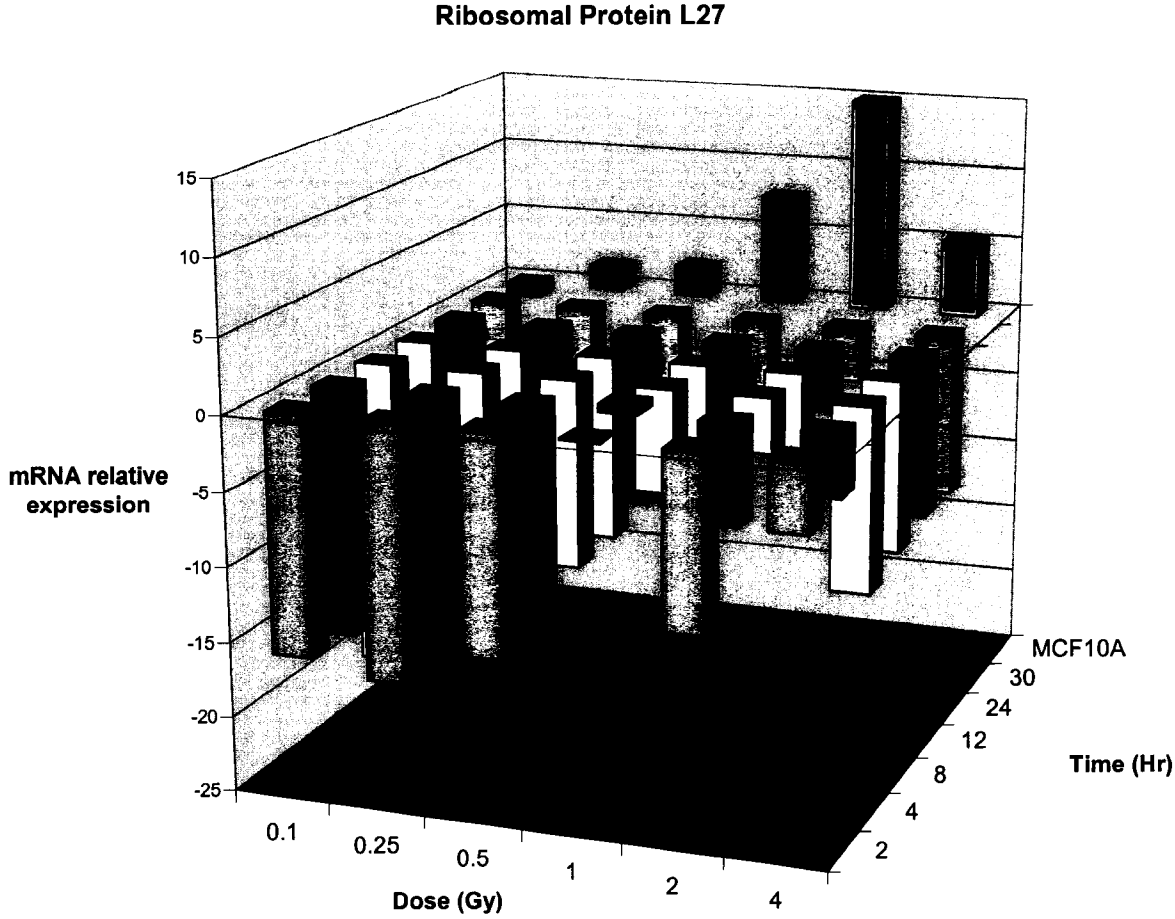


Figure 4.23. Real-time PCR analysis of the mRNA expression levels of the five radiation response genes in MCF10A cells at various times following varying doses of IR relative to the parental. The mRNA expression was measured relative to the parental MCF10A cells. The error bars are representative of the standard deviation of triplicate samples in one experiment.

The ribosomal protein L27 mRNA relative expression was analyzed by real-time PCR in the parental MCF10A cells at 0.10, 0.25, 0.50, 1.0, 2.0 and 4.0 Gy. At each dose of ionizing radiation, samples were taken at 2, 4, 8, 12, 24 and 30 hours post radiation treatment in order to verify at what point the mRNA expression was the most effected. As illustrated in Figure 4.23, the relative mRNA expression of the ribosomal protein L27 is down-regulated in MCF10A cells following radiation treatment. Upon statistical analysis, the samples taken following a dose of 0.10 Gy were found to be very statistically significant relative to the parental with P = .0023 to P = .0042. At a dose of 0.25 Gy, the relative mRNA expression was found to be very to extremely statistically significant with P = .0055 to P = .0003. Upon a 0.50 Gy dose of ionizing radiation, the relative mRNA expression was determined to statistically to very statistically significant relative to the parental with P = .0169 to P = .0061. At a 1.0 Gy dose of ionizing radiation, ribosomal protein L27 mRNA expression was found to be very to extremely statistically significant relative to the parental MCF10A cells with P = .0012 to P = .0006. At a 2.0 Gy dose, the relative mRNA expression was determined to be very to extremely statistically significant relative to the parental with P = .0029 to P = < .0001. Upon statistical analysis of a 4.0 Gy dose of ionizing radiation, the relative mRNA expression level relative to the parental was determined to be very to extremely statistically significant with P = .0012 to P = < .0001.

As seen in Figure 4.23, there is more of a dose effect seen at the lower doses of ionizing radiation than at the higher doses. The relative mRNA expression of the ribosomal protein L27 is more highly down-regulated at the lower doses than at the higher doses. At the 1.0 Gy dose of ionizing radiation, there is once again a radiation dose effect occurring if which the down-regulation of the mRNA expression is not to the extent seen at the other doses. The cause of this is unknown and could potentially be due to the effect the 1.0 Gy dose is having on the cell.

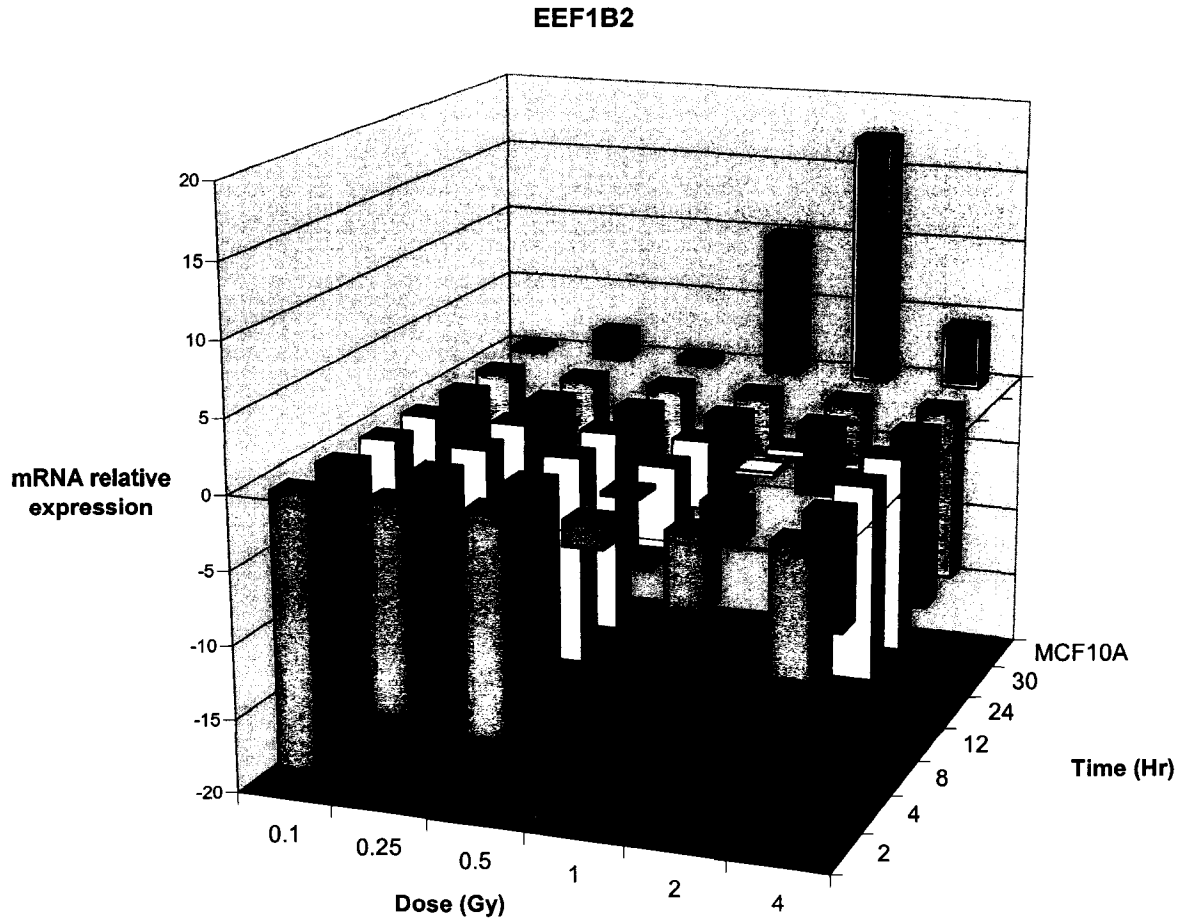


Figure 4.24. Real-time PCR analysis of the mRNA expression levels of the human eukaryotic translation elongation factor 1 beta 2 gene in MCF10A cells at various times following varying doses of IR relative to the parental. The mRNA expression was relative to the parental MCF10A cells. The standard deviation of triplicate samples in one experiment is given in the appendices.

The eukaryotic translation elongation factor 1 beta 2 mRNA relative expression was analyzed by real-time PCR in the parental MCF10A cells at 0.10, 0.25, 0.50, 1.0, 2.0 and 4.0 Gy. At each dose of ionizing radiation, samples were taken at 2, 4, 8, 12, 24 and 30 hours post radiation treatment in order to verify at what point the mRNA expression was the most effected. As illustrated in Figure 4.24, the relative mRNA expression of EEF1B2 is down-regulated in MCF10A cells following radiation treatment. Upon statistical analysis, the samples taken following a dose of 0.10 Gy were found to be very statistically significant relative to the parental with P = .0020 to P = .0038. At a dose of 0.25 Gy, the relative mRNA expression of EEF1B2 was determined to be very to extremely statistically significant relative to the parental with P = .0030 to P = .0007. At a dose of 0.50 Gy, the relative mRNA expression was found to be very to extremely statistically significant relative to the parental with P = .0054 to P = .0001. Following a 1.0 Gy dose of ionizing radiation, the mRNA expression of EEF1B2 was found to be very to extremely statistically significant relative to the parental with P = .0014 to P = .0005. After a 2.0 Gy dose, the relative mRNA levels were determined to be extremely statistically significant with P = <.0001 to P = .0006. At a 4.0 Gy dose, the relative mRNA expression was found to be extremely statistically significant relative to the parental with P = .0001.

Once again, a dose-rate effect is observed with lower doses producing a greater down-regulation in the MCF10A cells than the higher doses. There is a radiation effect occurring at 1.0 Gy that was also observed in the other genes as well. This effect is also being seen at 2.0 Gy for the relative mRNA expression of the EEF1B2 gene as well. Perhaps this effect is specific for only these doses and the cell has developed pathways to better deal with the damage caused at these doses.

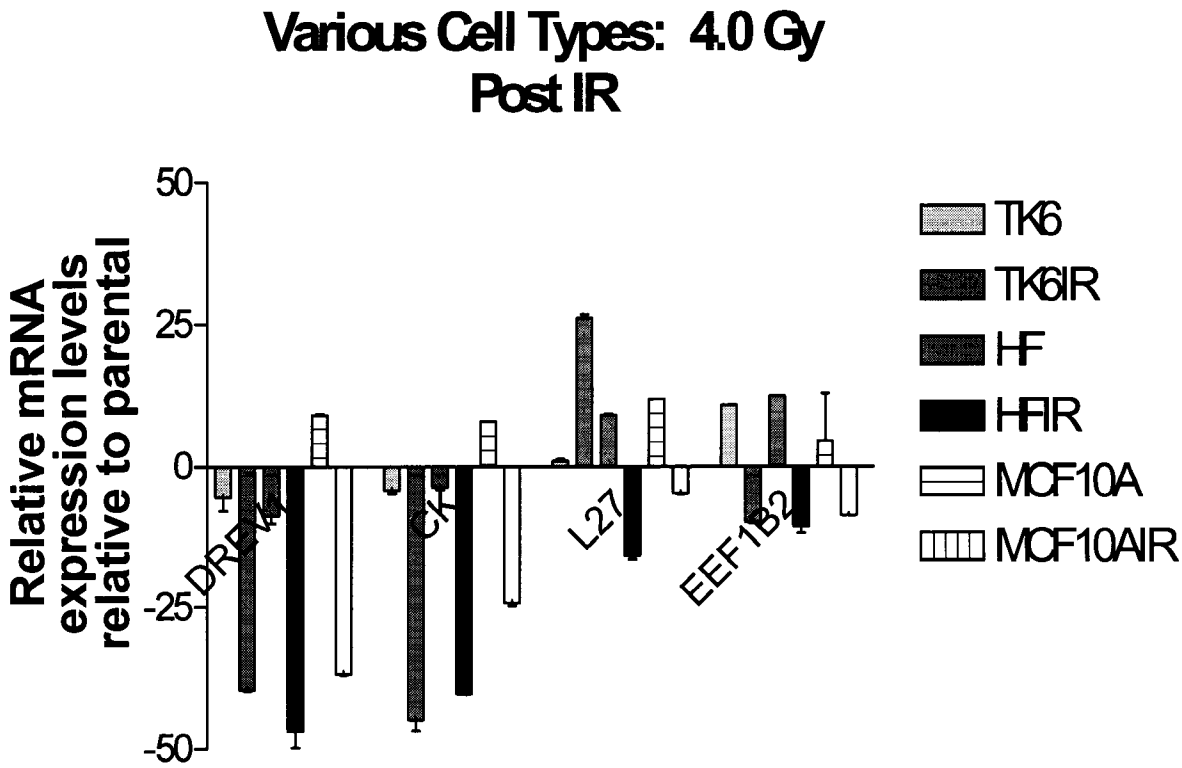


Figure 4.25. Real-time PCR analysis of the mRNA expression levels of the radiation response genes in various cell types following a 4.0 Gy dose of IR relative to the untreated control. The mRNA expression is relative to each cell type's untreated parental cell.

The five radiation response genes were analyzed in two other cell types in order to determine whether the radiation response was cell-type specific. The five radiation response genes were analyzed by real-time PCR for their mRNA expression levels relative to the untreated control in both human fibroblast cells and in the TK6 lymphoblastoid cell line. The mRNA expression levels were analyzed both with a 4.0 Gy dose of ionizing radiation and without. It was found in previous experiments that the five identified genes illustrated a radiation response in epithelial cells. From what is seen in Figure 4.25, the radiation response seen in the five radiation response genes does not appear to be cell-type specific, in fact a radiation response can also be seen in the human fibroblast and TK6 lymphoblastoid cell lines. This can be concluded due to the fact that upon radiation treatment in each cell type, each radiation response gene is being down-regulated. In Figure 4.25 it can be seen how each gene tested is significantly down-regulated upon radiation treatment. There was only one gene found to be up-regulated, ribosomal protein L27, in the TK6 cell line following radiation treatment. This may be due to mutations in the cell line itself. Upon statistical analysis it was determined that the mRNA relative expression levels were very to extremely statistically significant relative to parental cells with $P = < .0001$ to $P = .0037$. All values were found to be statistically significant.

RNAi analysis of DREV1 mRNA expression

The use of RNAi was accomplished by targeting siRNAs to the DREV1 gene in order to ascertain whether or not the gene was essential in MCF10A cells upon gamma irradiation treatment. A dose range of 1 nM to 100 nM siRNAs was utilized in order to determine the lowest optimal concentration of siRNAs needed for efficient gene silencing. It is important to use the lowest siRNA concentration as possible in order to minimize the interferon response. The MCF10A cells were harvested and analyzed at 24, 30, 36 and 48 hours post siRNA

transfection in order to determine at what time point silencing was the most efficient. This can be visualized in the Figures that follow.

Real-time PCR analysis of siRNA knockdown of the mRNA expression of the DREV1 gene

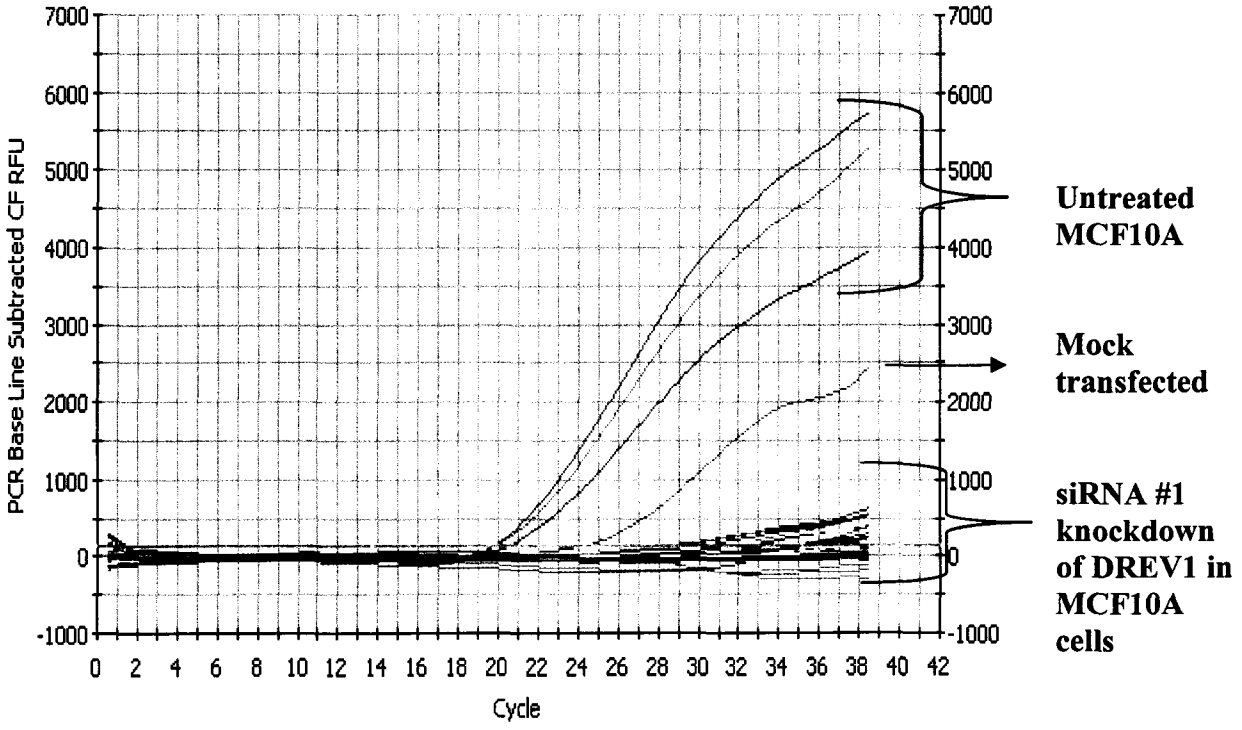


Figure 4.26. Real-time PCR analysis of siRNA knockdown of DREV1 in MCF10A cells. #1 siRNA targeted to the DREV1 gene. The samples were analyzed at 24, 30, 36 and 48 hours post siRNA transfection. Varying concentrations of 100 nM to .1 nM were utilized.

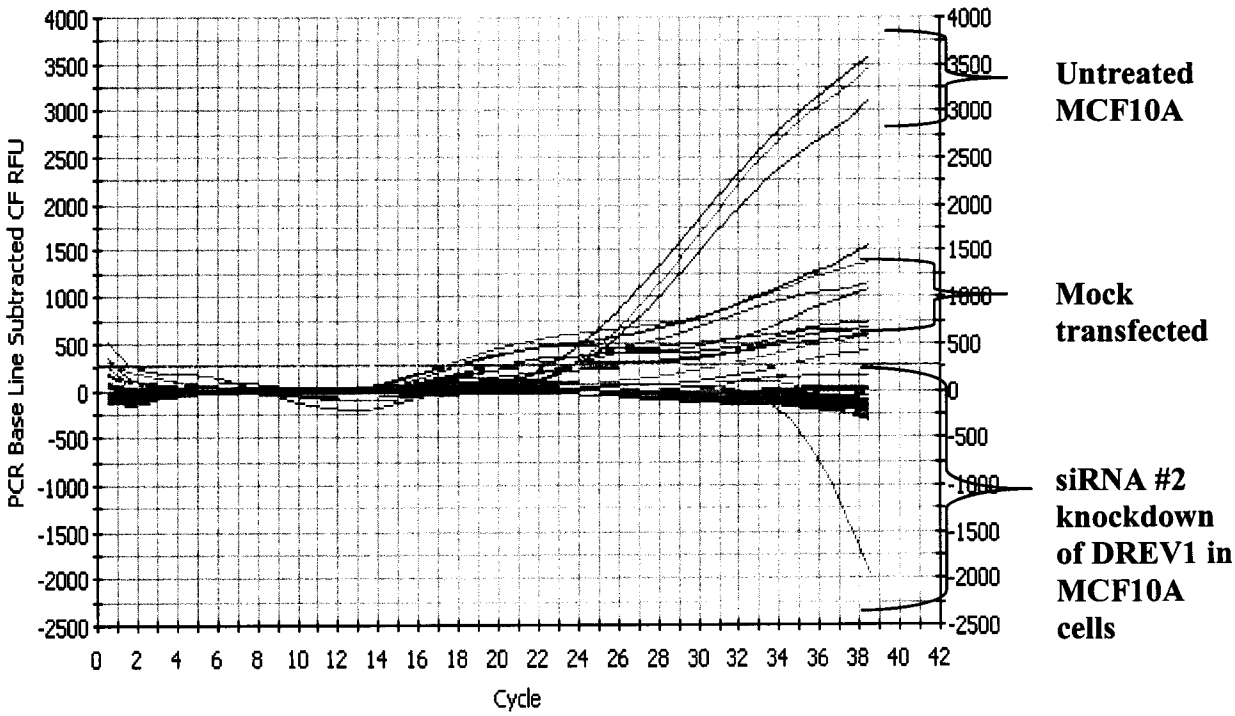


Figure 4.27. Real-time PCR analysis of siRNA knockdown of the DREV1 gene in MCF10A parental cells. #2 siRNA targeted to the DREV1 gene. The samples were analyzed at 24, 30, 36 and 48 hours post siRNA transfection. Varying concentrations of 100 nM to .1 nM were utilized.

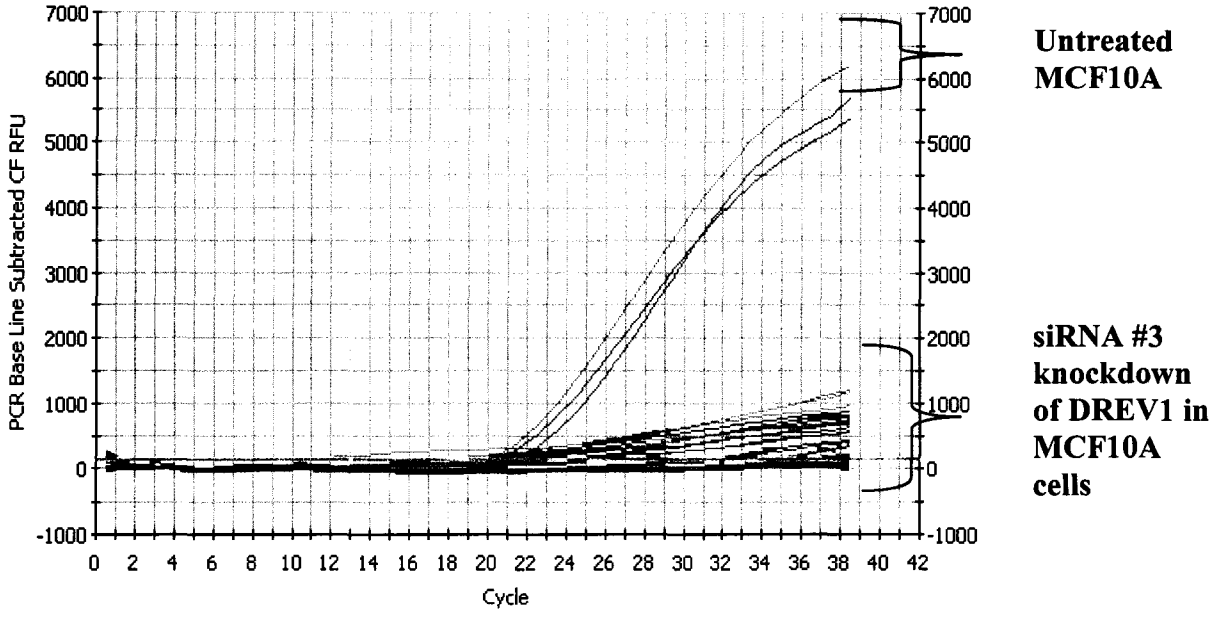


Figure 4.28. Real-time PCR analysis of siRNA knockdown of DREV1 in MCF10A parental cells. #3 siRNA targeted to the DREV1 gene. The samples were analyzed at 24, 30, 36 and 48 hours post siRNA transfection. Varying concentrations of 100 nM to .1 nM were utilized.

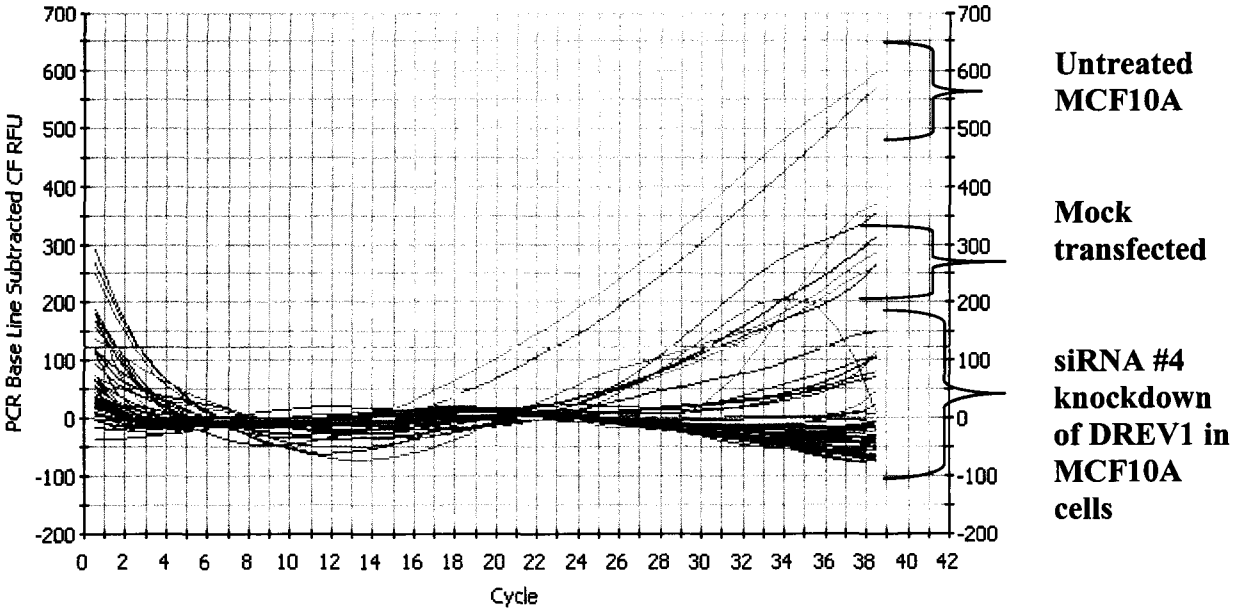


Figure 4.29. Real-time PCR analysis of siRNA knockdown of DREV1 in MCF10A parental cells. #4 siRNA targeted to the DREV1 gene. The samples were analyzed at 24, 30, 36 and 48 hours post siRNA transfection. Varying concentrations of 100 nM to .1 nM were utilized.

RNA interference was performed on the MCF10A parental cells with four siRNAs targeted to the DREV1 gene as shown in Figures 4.26 – 4.29. Only the DREV1 gene was targeted with siRNA out of the five radiation response genes due to the fact that it is a novel protein with very little known about it. It has known homology to DNA methyltransferases and its role in the radiation response in the cell is of great interest. If the cell is unable to methylate particular genes in order to silence them and suppress expression, these genes could then potentially remain unmethylated and actively express protein when they are not supposed to. This could lead to various detrimental effects in the cell, including killing the cell. This question will be answered upon chromosomal and chromatid aberration analysis upon knockdown of the DREV1 gene. As seen in Figures 4.26-4.29, the siRNA targeted to the DREV1 gene efficiently produced knockdown in the MCF10A cells. The small amount of mRNA amplification being seen is from the untreated parental MCF10A cells and the mock transfected MCF10A cells. These samples would not be expected to produce a knockdown in the DREV1 gene due to the fact that no siRNA is present that can target to the DREV1 gene. It appears that the concentration of 20 nM siRNA was the most effective in this experiment and will be utilized in the chromosomal and chromatid aberration study. The samples were analyzed at 24, 30, 36, and 48 hours post siRNA transfection.

Real-time PCR analysis of siRNA controls for the DREV1 knockdown experiment

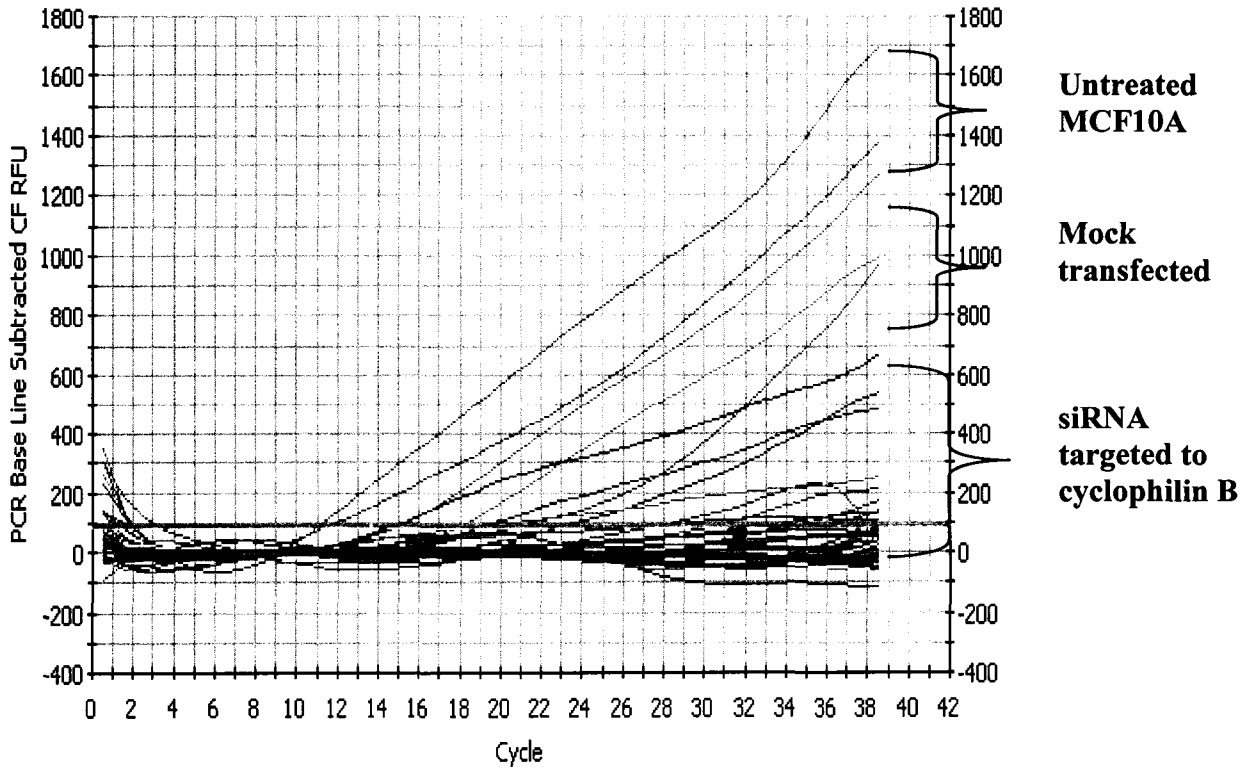


Figure 4.30. Real-time PCR analysis of siRNA knockdown of DREV1 in MCF10A parental cells. Cyclophilin B siRNA positive siRNA targeting control.

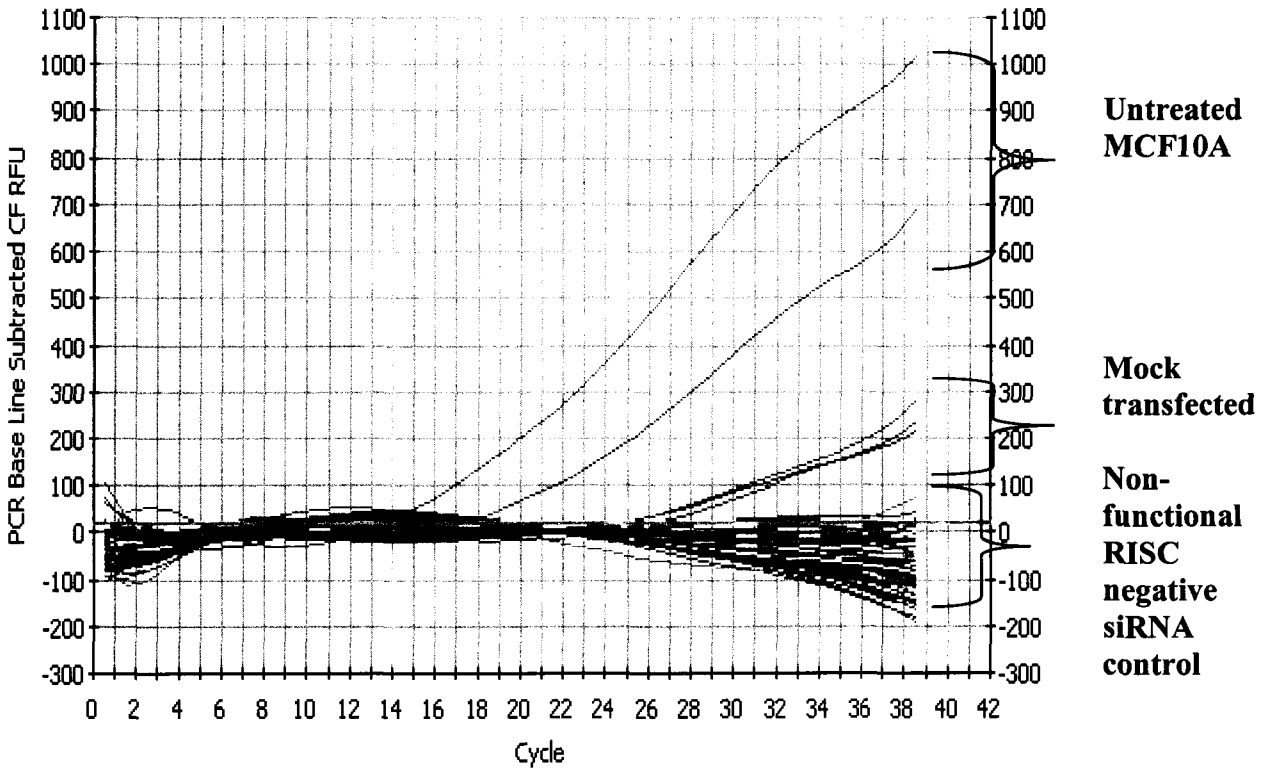


Figure 4.31. Real-time PCR analysis of siRNA knockdown of DREV1 in MCF10A parental cells. Non-functional RISC negative siRNA control.

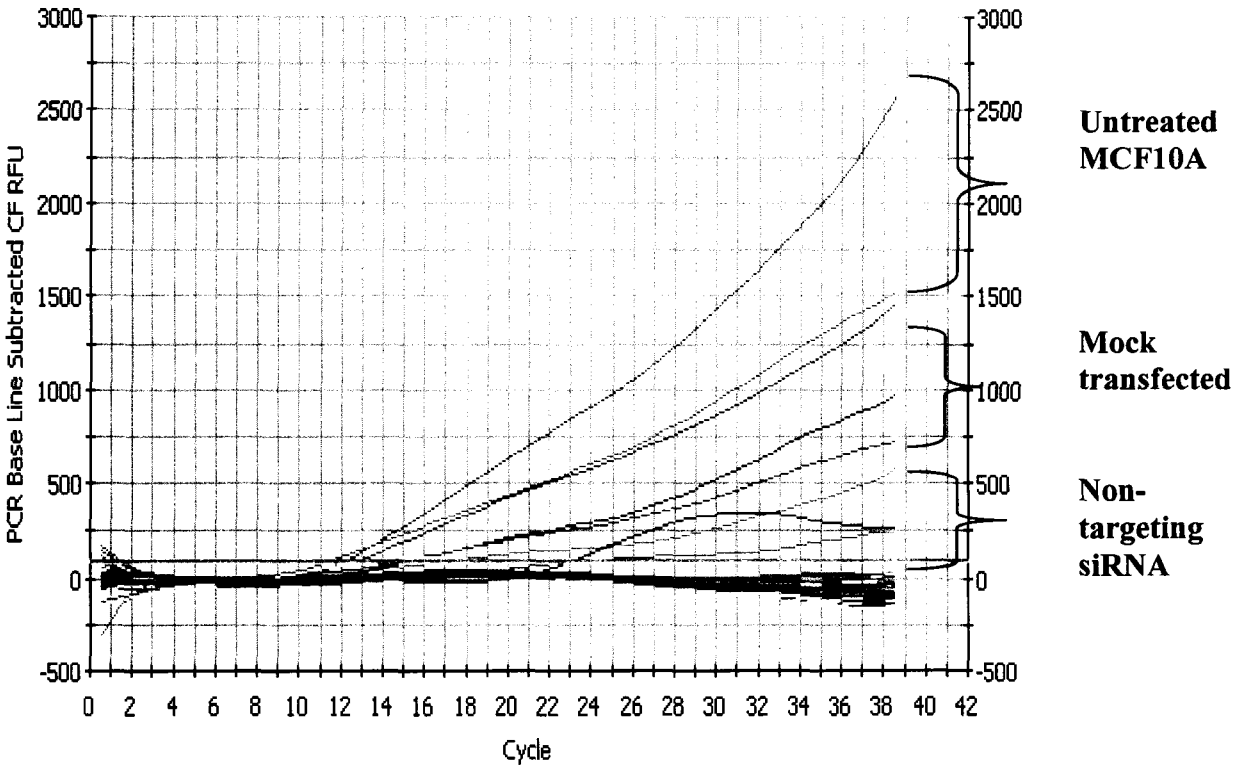


Figure 4.32. Real-time PCR analysis of siRNA knockdown of DREV1 in MCF10A parental cells. Non-targeting siRNA negative control.

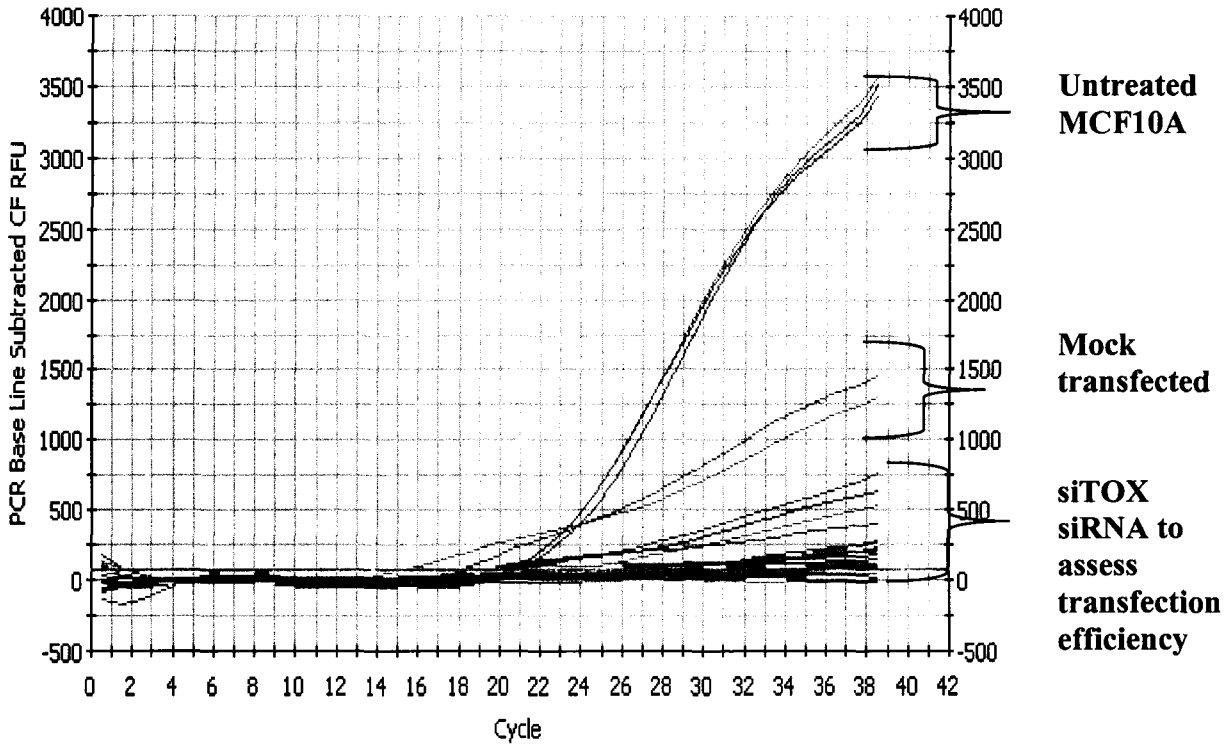


Figure 4.33. Real-time PCR analysis of siRNA knockdown of DREV1 in MCF10A parental cells. siTOX transfection efficiency siRNA control.

siRNA controls were analyzed by real-time PCR for relative mRNA expression levels at 24, 30, 36, and 48 hours post transfection as seen in Figures 4.30-4.33. The following concentrations were utilized in order to determine the optimal concentration of siRNA to be used in the survival analysis experiment: 100 nM, 20 nM, 10 nM, 5 nM, 1 nM, and 0.1 nM. In Figure 4.30, cyclophilin B relative mRNA was analyzed by real-time PCR. Cyclophilin B was utilized as a positive control that was targeted to the cyclophilin B gene, which all cells contain. It serves as a positive control in order to verify that the siRNA system and transfection are functioning correctly. The amplification seen in Figure 4.30 is due to the parental MCF10A untreated control and the mock transfected control, which are not down-regulated since they contain no siRNA. In Figure 4.31, a non-functional RISC complex was transfected into parental MCF10A cells. Once again, the amplification being seen is from the mock and untreated controls. It can be seen that knockdown is still occurring efficiently as expected. In Figure 4.32, a non-targeting siRNA serves as a negative control for the experiment. You would expect that knockdown should not occur. There is knockdown occurring, but the vast majority of the samples run are being amplified as anticipated. In Figure 4.33, the transfection efficiency of the experiment is verified through the use of siTOX transfection control. It should produce knockdown in the cells transfected since it is toxic to the cell. As seen in the Figure 4.33, the majority of samples are being knocked down, with the mock and untreated samples being amplified. The slight amplification seen towards the end of the amplification cycle would imply the fraction of the cell sample not transfected. The siTOX also allows visualization of the cells that were transfected since it glows and is visible under a fluorescent microscope. It appeared as though about 80% of the cells were transfected efficiently.

Analysis of survival potential by scoring for chromosomal and chromatid aberrations

To determine whether the gene disrupted by gene trapping affects the survival of cells after gamma irradiation, chromosomal and chromatid aberration analysis were performed. It is imperative to know whether or not the identified radiation response gene can affect cell survival. This assay allows for easy visualization of any gross chromosomal and chromatid aberrations in the cell following radiation treatment in the parental MCF10A cell line both with and without siRNA knockdown of the radiation response gene.

Chromosomal and chromatid aberration analysis of DREV1 siRNA knockdown cells

Chromosomal and chromatid aberration analysis was performed on the siRNA transfected MCF10A cells in order to determine if knockdown of the DREV1 gene was affecting the survival of the cells. This was done prior to a survival curve in order to first determine if the DREV1 gene was causing enough damage in the cell to decrease cell survival. A total of 50 metaphases for each sample were scored. It appears that the number of chromatid breaks seen in the 2.0 and 4.0 Gy siRNA-knockdown MCF10A cells were equivalent to the background levels seen in the unirradiated MCF10A parental cells. Knocking down the mRNA expression of the DREV1 gene appears to have no effect on the aberration frequency in non-irradiated MCF10A cells or on the frequency of radiation-induced aberrations. The numbers of interstitial deletions, terminal deletions and dicentrics in the 2.0 and 4.0 Gy siRNA-knockdown MCF10A cells are believed to be due to the effects of the ionizing radiation and not from the siRNA-knockdown of the DREV1 gene. The amount and types of breaks seen are characteristic of radiation damage and are found to double from the 2.0 Gy siRNA sample to the 4.0 Gy siRNA sample. This is shown below in Figure 4.34. This once again illustrates the fact that the chromosomal damage seen is due to the radiation effects and not the knockdown of the

DREV1 gene. It appears that the DREV1 gene does not effect the survival of the cells due to the fact that it does not cause any observable chromosomal or chromatid aberrations that would be detrimental to the cell.

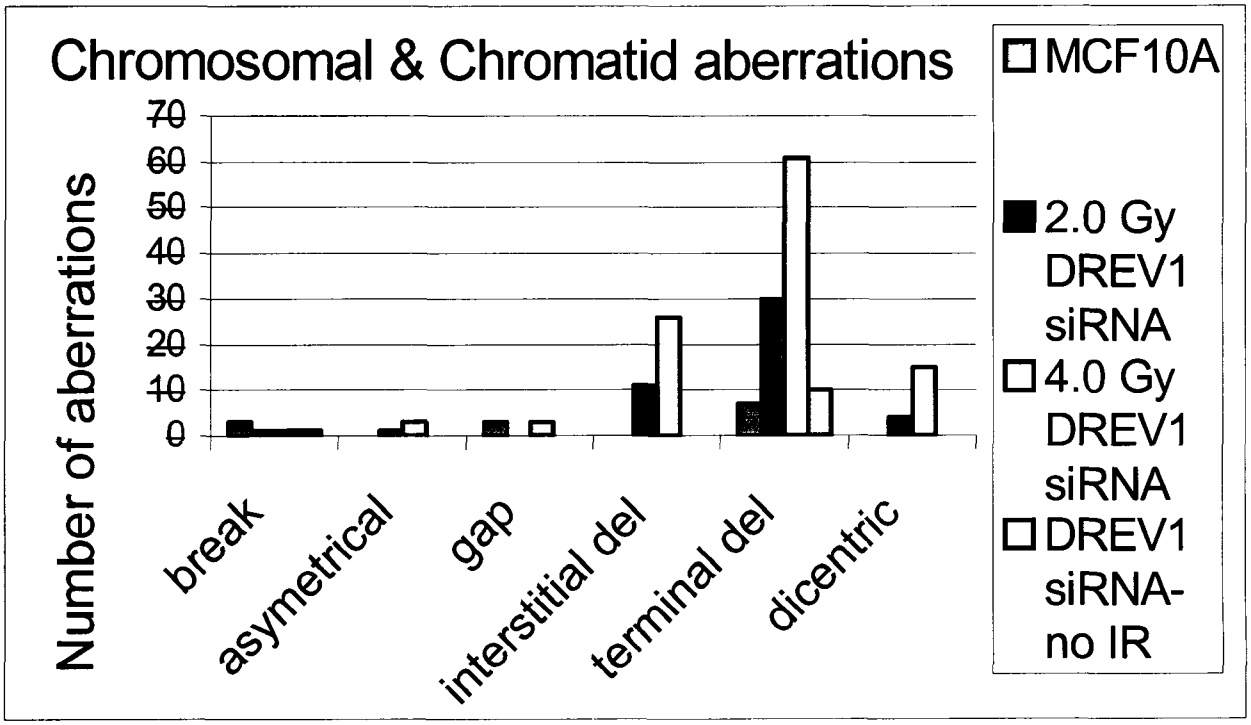


Figure 4.34. Analysis of chromosomal and chromatid aberrations in MCF10A cells irradiated and with a siRNA-knockdown of the DREV1 gene. The metaphase spreads were stained with 10% giemsa stain and scored for chromosomal and chromatid aberrations. A total of 50 metaphase spreads were analyzed for each sample.

CHAPTER 5- DISCUSSION

Introduction

Upon sequencing of the gene trapped MCF10A clones, five genes were identified as responding to the radiation treatment. The five radiation response genes identified were: human androgen receptor, human creatine kinase, human ribosomal protein L27, human eukaryotic translation elongation factor 1 beta 2, and DORA reverse strand protein 1. These five genes were found to be significantly down-regulated relative to the parental cell line at doses of ionizing radiation ranging from 0.10 to 4.0 Gy. This radiation response was also found to occur across various cell types, indicating that this was not a cell-type specific phenomenon. It is important now to consider the molecular pathways in which each of these genes is involved.

Androgen Receptor

It is known that the androgen receptor is essential for prostate development and homeostasis. Androgen, acting through the androgen receptor, regulates PSA, kallikrein gene family members, C (3) protein, sex-limited protein (Slp), and probasin. Androgen response elements (ARE) have been identified in the promoters of all of the aforementioned genes. Androgen receptor is encoded by a 90 kb gene on Xq11-Xq12. It encodes an 11 kb mRNA transcript composed of 8 exons. Mutations in the androgen receptor in the DNA-binding domain results in an inability to bind androgens, which has been seen in brothers with breast cancer. It is well established that androgens play a role in tumorigenesis. Androgens have also been previously identified to be affected by gamma irradiation in sertoli cells.

There are two sources of androgen in women. The main receptor for testosterone is the androgen receptor. Testosterone is produced by the ovary and by conversion of the adrenal androgens dehydroepiandrosterone and dehydroepiandrosterone-sulfate into androstenedione, and further to testosterone in peripheral tissues. In premenopausal women, there is 25% circulating testosterone secreted from the adrenal gland, 25% from the ovary, and 50% from the peripheral conversion of androstenedione. In postmenopausal women, there are lower testosterone levels found. The ovary still continues to produce androstenedione and testosterone. An increase in breast cancer risk in postmenopausal women with increased levels of endogenous testosterone is seen. This is not seen in premenopausal women (103). Treatment with testosterone results in tumor regression and reduction in the estrogen receptor expression. Testosterone may exhert an indirect effect on breast cancer proliferation by sequestering sex hormone binding globulin, leaving more estradiol in the non-protein-bound state and able to act on breast tissue.

In breast cancer there are numerous factors that influence the growth and proliferation of the cells. Steroid hormones such as estrogen receptor, progesterone receptor and androgen receptor all help to regulate the growth and proliferation of the breast cancer cells. Other growth factors, cytokines, and lymphokines, such as epidermal growth factor (EGF), fibroblast growth factor (FGF), and transforming growth factor (TGF-α/β) are also know to play a role in the regulation of breast cancer growth. Growth factors involved in cell proliferation and tumorigenic activity, such as EGF, keratinocyte growth factor (KGF), and insulin-like growth factor (IGF-I) activate the transcription transactivation functions of androgen receptor. In the absence of

androgens, IGF-I, KGF, EGF, luteinizing hormone-releasing hormones (LHRH), neuropeptides, HER-2/neu, and IL-6 increase transcription of the androgen receptor (80).

It is well established that the risk of breast cancer is elevated in postmenopausal women with high estrogen and androgen levels. Androgen receptor is found to be elevated in elderly and postmenopausal women due to compensation for a decline in circulating levels of sex steroids. One might expect that estrogen receptor is the most prevalent hormone expressed in breast cancer, but actually the androgen receptor is expressed more frequently in breast cancer tissues than either the estrogen or progesterone receptors. Androgen receptor positive tumors (60-90%) are more frequent than estrogen receptor alpha positive (60-80%) or progesterone positive (50-70%) tumors. Frequently, all three are co-expressed. This indicates that a subset of breast tumors may express androgen receptor only, which may in turn contribute to breast cancer cell growth and responsiveness to hormonal therapy. An association in breast cancer risk with androgen has been seen epidemiological studies, but the mechanisms are unknown. It has been demonstrated that after aromatization to estrogenic steroids, androgens may enhance tumor progression indirectly (27). It has been identified that in half of the estrogen and progesterone receptor negative primary tumors, that they are in fact androgen receptor positive.

The identification of the human androgen receptor gene as responding to radiation in the gene-trapping experiment is particularly of interest due to the fact that it is known that the androgen receptor is present in 70-90% of primary breast tumors. Since radiation is a known carcinogen, the role a potential radiation response gene has in the development of breast cancer is of great interest. Androgen is currently used clinically in the treatment of breast cancer, with comparable efficiencies to other hormonal therapies, such as tamoxifen. Medroxyprogesterone

acetate (MPA) is other a second line hormone therapy for breast cancer following tamoxifen failure. The antiproliferative effect of MPA is mediated by the androgen receptor. It inhibits proliferation of estrogen receptor alpha negative and progesterone receptor negative cells via the androgen receptor (21). The identification of androgen-regulated genes that mediate growth inhibitory effects of androgen may enable a more precise prediction of the response of the breast cancer to hormonal therapies and define new targets in the treatment of breast cancer.

Androgen acting via its receptor, not only controls the proliferation of female breast tissue, but it is also associated with the development and progression of breast cancer by a complex interplay with the estrogen receptor. An active androgen receptor is essential in the control of breast cancer growth (161). Androgen affects female breast growth and cancer proliferation via two routes: acting with the androgen receptor and transforming to estrogen. Loss of the androgen receptor may lead to a loss of balance between androgen and estrogen effects, subsequently affecting the biological behavior of the breast cancer cells. The androgen receptor is also very important in metastatic disease of the breast, where androgen receptor is often the sole steroid receptor expressed in as much as 25% of the cases (196). Androgen receptor expression in primary breast cancer has been well established, but the data showing the relevant percentages of cases with androgen receptor positive breast cancers is more varied. In 1992, the androgen receptor was identified in roughly 60% of breast cancer samples (218). In 2000, 86% of breast cancer samples were found to be positive for androgen receptor (161). In 1993 it was found that 79% of breast cancer samples were androgen receptor positive (219). And lastly, in 1996, a total of 100% of the breast cancer samples tested were found positive for the androgen receptor (220). Androgen receptor negative status in breast cancer is correlated with the aggressive features of female breast cancer. Given its importance in female breast

growth, androgen receptor may be potentially related to the pathogenesis and biological behavior of female breast cancer.

Allelic loss of the androgen receptor has been found in 9 out of 11 androgen receptor negative breast cancer tumors (161). Loss of an inactive allele represents genetic instability of the CAG repeats in the androgen receptor. Loss of an active allele was found to represent an effect at the level of androgen receptor expression. This is what is believed to cause androgen receptor expression loss in female breast cancer due to methylation of the 5' CpG island or a somatic mutation in the androgen receptor gene. An increase in the androgen receptor CAG repeat length leads to a decrease in the androgen receptor transactivation activity. The normal CAG repeat length of the androgen receptor is between 6 and 39. When the CAG repeat length reaches 40 to 66 repeats, spinal and bulbar muscular atrophy is seen. Long CAG repeat length in postmenopausal women and BRCA1 mutant carriers is associated with an increase in breast cancer risk in a study of women in Quebec. Women diagnosed with breast cancer prior to 45 years of age had an increased number of CAG repeat length in their androgen receptors. Decreased androgenic activity in breast cells expressing a very long androgen receptor CAG repeat allele may result in increased breast cell proliferation. This increased proliferation may in turn affect the penetrance of breast cancer in BRCA1 mutant carriers. Androgen receptor is believed to modulate BRCA1 risk via an endocrine mechanism by altering the levels of circulating hormones, such as estradiol. A paracrine mechanism involves androgens acting on mammary stroma to indirectly affect the growth of the mammary epithelium.

The androgen receptor is known to interact with a great number of genes. The kallikrein gene family is a known target for interaction with the androgen receptor gene. The kallikrein gene family is composed of 15 serine protease genes on chromosome 19q13.4. Numerous

kallikrein gene family members have been found to be expressed in endocrine-related malignancies. KLK10 is a novel tumor suppressor that has been found to be down-regulated as breast cancer progresses. KLK6 has been found to be overexpressed in both breast and ovarian cancers. KLK8 is found to be up-regulated in ovarian cancer. KLK12, KLK15, and KLK13 have all been found to be down-regulated in breast cancer patients. KLK15 is also up-regulated in prostate cancers and has been found to be a favorable prognostic marker in breast cancer patients. KLK3, which encodes the prostate specific antigen (PSA), is known to be expressed in breast tissue. It has been determined to be a favorable prognostic marker in breast cancer patients. All of the mentioned kallikrein gene members are up-regulated by androgens through the androgen receptor pathway and play a role in the cellular proliferation of the breast (195).

Androgen receptor is known to interact with the normal epithelial cell-specific 1 gene (NES1). NES1 was found to be a tumor suppressor, which is down-regulated in progressive breast cancers. An interesting fact about NES1, is that it is expressed in breast cancer at a high rate, but down-regulated at metastatic sites (195). Other genes known to be altered in breast cancer include: RB1, CCND1, MYC, and ERBB2. In poorly differentiated tumors androgen receptor is found to be down-regulated and is correlated to RB1 by down-regulated as well. There is direct regulation of androgen receptor transcription by the c-myc transcription factor via a myc consensus site in an androgen receptor exonic region and downstream regulation of MYC mRNA association with androgen-induced suppression of a transformed phenotype (21).

Androgen receptor can be recognized and activated through androgens, estrogens, gestagens, and is also involved in the steroid-independent MAP kinase signal transduction pathway. The action of androgens is mediated through the androgen receptor that belongs to a steroid hormone subfamily of nuclear receptors. These ligand-regulated transcription factors are

composed of single polypeptides harboring three separate functional domains: the well-conserved C-terminal ligand-binding domain, the highly-conserved DNA-binding domain, and the poorly-conserved N-terminal domain. In the absence of ligand, the androgen receptor is complexed in the cytoplasm to chaperone proteins that keep the receptor in a transcriptionally inactive form. Upon binding the hormone, the androgen receptor dissociates from the chaperones and translocates to the nucleus where it binds to AREs. A ligand-induced conformational change enables the receptor to recruit coactivators and/or transcriptional proteins to target gene promoters. Some examples of coactivators are: steroid receptor coactivator (SRC) family members, such as SRC-1, SRC-2, glucocorticoid receptor-interacting protein 1 (GRIP1), transcription intermediary factor 2 (TIF2) and SRC-3, activator of thyroid and retinoic receptor, amplified in breast cancer protein (AIB1), CREB-binding protein (CBP/p300), and CBP/p300-associated factor (PCAF) (85).

Recruitment of complexes that affect acetylation of chromatin domains has been shown to be important for transcriptional regulation by steroid receptors. Estrogen receptor and coactivators assemble onto estrogen responsive promoter elements in a cyclic fashion and in a specific order, indicating that promoter remodeling by histone acetylation is a dynamic and stepwise process. Core histones, CBP/p300 and PCAF are known to acetylate both the androgen and estrogen receptors. Androgen receptor is also known to interact with TFIIH and TFIIF and a large number of other nuclear coregulatory proteins.

The androgen receptor up-regulates the expression and enzymatic activities of CDK2 and CDK4, both of which promote cell proliferation. Androgen inhibits p16 and up-regulates p21, both of which are essential in cell cycle, DNA repair, and in antiapoptotic activities. p21 is

up-regulated during cell cycle arrest following DNA damage in the cell and protects cancerous cells against p53-mediated apoptosis. p21 is known to be induced by p53, TGF-β, STAT1, vitamin D, and NGF. p21 activation by androgen occurs through a canonical ARE in its proximal promoter region. Sp1 binding sites are involved in the induction of p21 by androgen receptor. It has been found that Sp1 and androgen also interact with one another. It has also been found that androgen induces some target genes via a Sp1 site when the ARE is absent in the gene promoters (108).

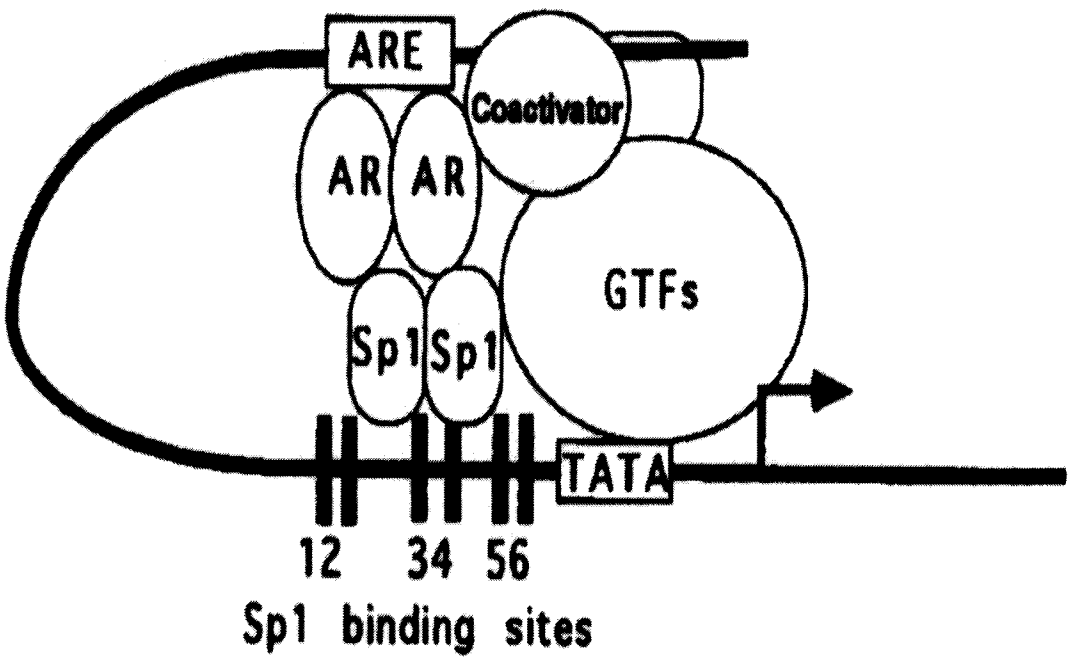


Figure 5.1. The proposed model of ARE and Sp1 site in induction of the p21 gene by androgen. Conformational change of AR upon androgen binding enables the receptor to bind to ARE and interact with transcription factor Sp1. This interaction between AR and transcription factor Sp1 facilitates assembly of transcriptional coactivators and general transcription factors (GFTs) into a transcription preinitiation complex, resulting in enhanced expression of the p21 gene. Source: Lu, S. et al. Androgen induction of cyclin-dependent kinase inhibitor p21 gene: role of androgen receptor and transcription factor Sp1 complex. Molecular Endocrinology, 2000. 14(5): p. 753-760.

Androgens control cell numbers through three different mechanisms: increasing cell proliferation, inhibition of cell death, and inhibition of cell proliferation. Proliferation of androgen is controlled by sex hormones through a two-step mechanism: sex steroids increase proliferation of their target cells by canceling the inhibition exerted by a specific plasma born protein and sex steroids trigger expression of unknown endogenous inhibitors of proliferation (168).

Androgen receptor is known to take part in the regulation of stress response genes. A stress response induced by androgen reflects a protective response to cope with adverse cellular effects of the androgens itself. Linkage of androgen to cellular stress responses may provide a molecular explanation for some of the carcinogenic effects of androgen. A reduction or loss in the expression of such stress response genes in cancer cells may support a pro-carcinogenic effect of androgen (160).

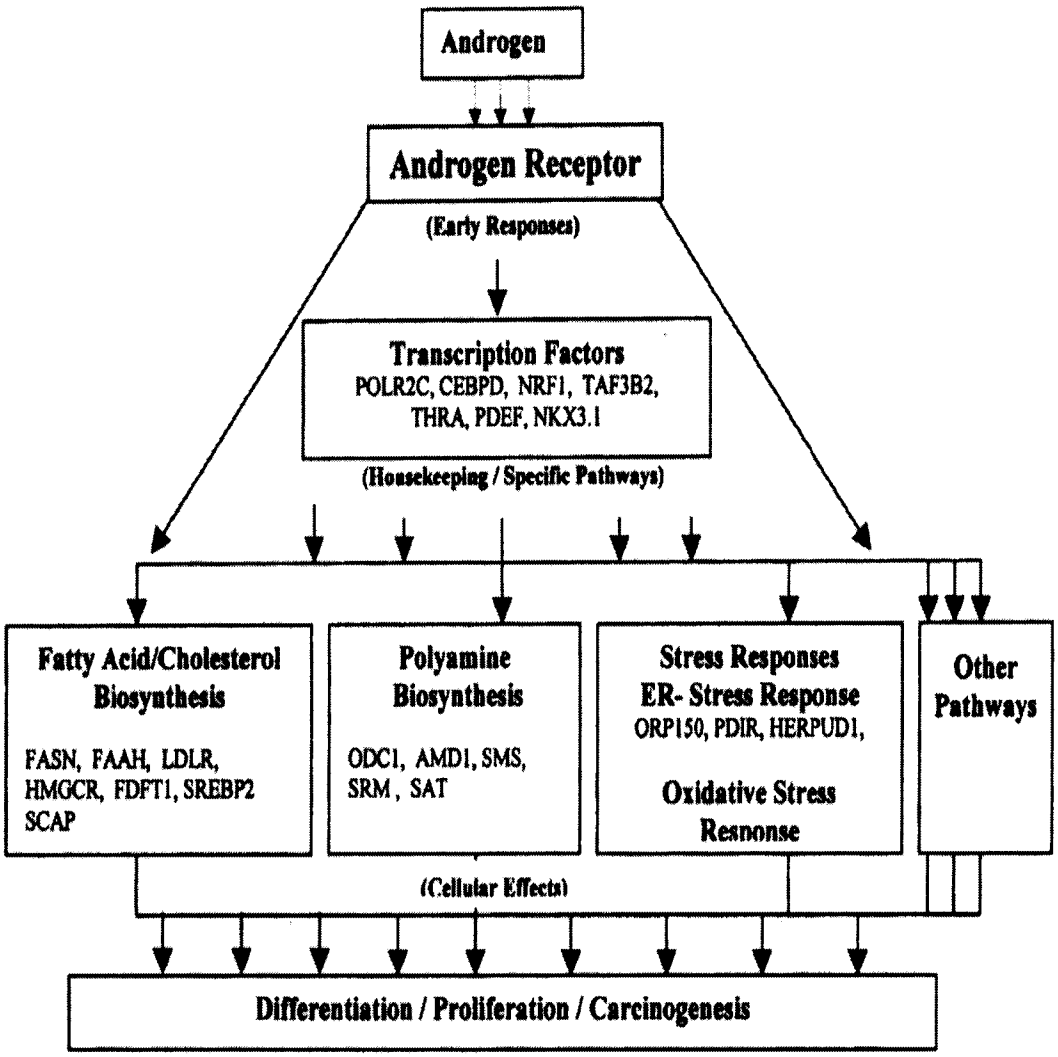


Figure 5.2. Proposed model of the AR transcriptional net work.
Source: Segawa, T. et al. Androgen-induced expression of endoplasmic reticulum (ER) stress response genes in prostate cancer cells. Oncogene, 2002. 21, p. 8749-8758.

Creatine Kinase

Creatine kinase catalyzes the reversible transfer of a γ-phosphate group of ATP to the guanidine group of creatine to produce ADP and phosphorylcreatine. Creatine kinase directly couples oxidative phosphorylation to energy utilization supporting the high and sustained energy requirements needed by the organs and tissues in the body (37). Creatine kinase is expressed in a wide range of tissues.

There have been four isozymes of creatine kinase identified in mammals: muscle and brain, which are both located in the cytosol, and ubiquitous and sarcomeric mitochondrial, which are both located in the mitochondrial intermembrane space. The human muscle creatine kinase gene is 17.5 kb consisting of 8 exons and 7 introns. It encodes a 381 amino acid polypeptide. There is 77-91% in the muscle creatine kinase gene coding sequence homology between human, fish, birds, and other mammals (142).

The human brain isozyme of creatine kinase is located on chromosome 14 and consists of 8 exons and 7 introns. It is expressed in the brain, heart, smooth muscle, uterus, placenta, colon, various tumors, and in cell lines. During myogenesis, cytosolic creatine kinase isozymes undergo an isotype switch from embryonic dimeric BB form to the adult MM form. Steroid hormones modulate the creatine kinase brain isozyme's expression level. Creatine kinase brain isozyme is induced after estrogen receptor or progesterone receptor administration in the uterus of female rats. In male rats, testosterone has the same effect of induced creatine kinase brain isozyme in the bones. In HeLa cells, creatine kinase brain isozyme expression is repressed by p53 and activates creatine kinase muscle isozyme in CV-1 monkey kidney cells.

The ubiquitous mitochondrial creatine kinase isozyme is located on chromosome 15. It is 5.5 kb in length and contains 9 exons and 8 introns. A Sp1 binding site and CpG island have

been confirmed in exon 1 of the gene. This is characteristic of creatine kinase ubiquitous mitochondrial isozymes being a housekeeping gene. It is expressed in the intestine, brain, kidney, placenta, and during pregnancy in the uterus. A consensus nuclear hormone receptor-binding site is also found within the gene. The expression level of this isozyme varies greatly by tissue, during development, and during the estrous cycle and pregnancy. The DNA sequences and protein factors regulating the ubiquitous mitochondrial creatine kinase isozyme gene expression probably include members of the steroid superfamily of transcription factors (142).

The sarcomeric mitochondrial creatine kinase gene isozyme is located on chromosome 5. It is 37 kb in length and contains 11 exons and 3 introns. This isozyme is expressed in cardiomyocytes, slow-oxidative and fast-oxidative-glycolytic skeletal muscles of the heart and skeletal muscle. It is not found in the brain, lung, stomach, intestine, liver, kidney, bladder, testis, or uterus, as some of the other isozymes are. There is a functional interplay between the mitochondrial and cytosolic isoforms of the creatine kinase gene, which is thought to be important for the regulation of cellular energy homeostasis. Cytosolic isozymes re-phosphorylate ADP and upregulate creatine kinase levels in the cell. The cytosolic isozyme activity level is associated with ATPases in subcellular compartments of the sarcoplasmic reticulum or the myofibrillar M-band, where creatine kinase is functionally coupled to the calcium pump and the acto-myosin ATPase. Mitochondrial isozymes catalyze the reaction of creatine kinase to phosphorylated creatine kinase at the expense of ATP. Cytosolic isoforms of creatine kinase form dimeric structures and mitochondrial isoforms form an octamer from four dimers. The mitochondrial isoform octamer-dimer transition could play a crucial role in the regulation of mitochondrial energy metabolism as well as in cellular apoptosis (94).

Mitochondrial energy metabolism is the major source of intracellular reactive oxygen species such as the superoxide anion, hydroxyl radical, hydrogen peroxide, peroxynitrite, nitric oxide, and thiyl peroxyl radical. Defects in the mitochondrial energy production may lead to an increase level of reactive oxygen species. Creatine kinase is a target of reactive oxygen species such as hydrogen peroxide and peroxynitrite. Creatine kinase reacts very sensitively in the presence of various free radicals. Mitochondrial-creatine kinase is a prime target for peroxynitrite inactivation. Creatine kinase inactivation is observed before enzymes of oxidative phosphorylation were affected. Inactivation leads to a deterioration of cellular energetics, eventually followed by elimination of cells with a chronically lowered energy status via apoptosis (93).

Ionizing radiation is a known inducer of free radical damage in cells. The dose of ionizing radiation needed to decay the creatine kinase protein was found to increase with increasing protein concentration. This indicates that the protein damage was due to the indirect radiation effects preferentially initiated by free radical damage (95). An inverse dose-rate effect of damage has been found to be induced in the mitochondrial creatine kinase isoform. Unsaturated lipid molecules undergo free-radical induced lipid peroxidation. Since lipids show an inverse dose-rate effect, the amount of degradation products increases with decreasing dose rate. As a consequence, decay of creatine kinase enzyme activity is accelerated at lower dose rates in the presence of lipids. Lipid hydroperoxidation activates free radicals and is known to produce mammary tumors. Tumors believed to arise from diffusion of long-lived free radical intermediates initiate further radical reactions at remote sites. This diffusion process is presumably aided by the circulatory systems.

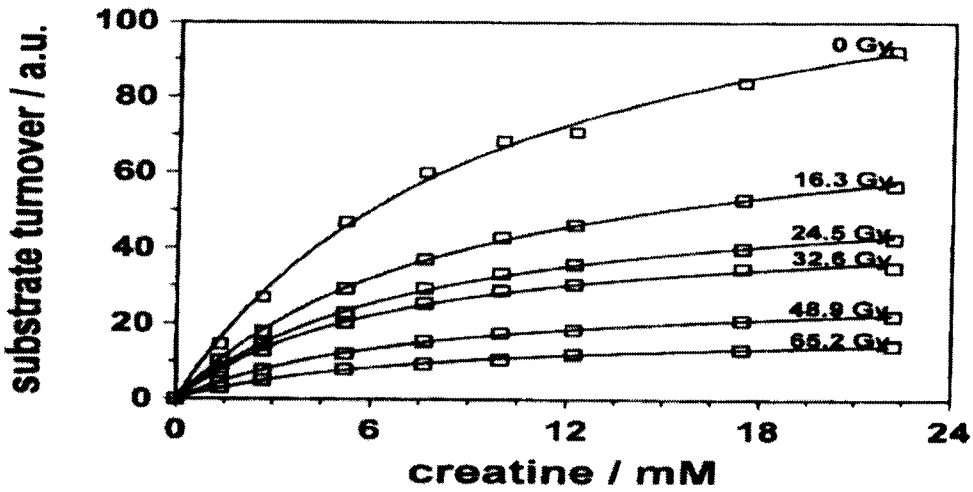


Figure 5.3. Creatine kinase turnover catalyzed by cytosolic muscle creatine kinase as a function of radiation dose. Substrate turnover as a function of creatine concentration and radiation dose. Source: *Koufen, P. and Günther Stark. Free radical induced inactivation of creatine kinase: sites of interaction, protection, and recovery. Biochimica et Biophysica Acta, 2000. 1501: p. 44-50.*

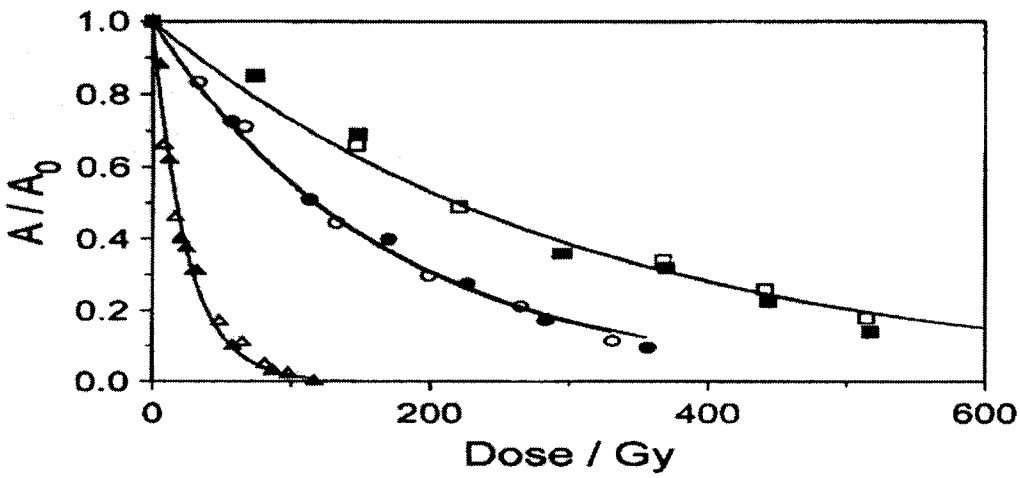


Figure 5.4. The relative activity, A/A_0, of creatine kinase (cytosolic muscle creatine kinase) as a function of dose at two different dose rates. 0.05 mg/ml (▲, Δ) 0.5 mg/ml (●,○) and 1 mg/ml (■, □) were exposed to X-rays at dose rates 98 Gy/min (□,○,Δ) and 5.8 Gy/min (■,●,▲). Source: *Koufen, P. et al. Inverse dose-rate effects at the level of proteins observed in the presence of lipids. International Journal of Radiation Biology, 2000. 76(5): p. 625-631.*

Creatine kinase isozymes contain a single active sulfhydryl group (Cys-278 in Mi-CK and Cys-283 in cytosolic-CK). Modification of the sulfhydryl residue by reactive oxygen species and nitric oxide results in decreased creatine kinase activity. This indicates that the active sulfhydryl group in creatine kinase plays an important role for catalytic activity.

Inactivation of creatine kinase by reactive oxygen species might be involved in energy metabolic regulation in cancer. It has been documented that the enzyme activity from the cytosolic fraction of creatine kinase is higher in normal tissue than in cancer tissue. HeLa cells have been found to contain an increased creatine kinase activity and expression level. Regulation of creatine kinase by reactive oxygen species might be involved in energy metabolism in cancer cells. It is believed that oxidation of creatine kinase by reactive oxygen species is irreversible in tumor cells due to inactivation and activity cannot be recovered. Increased creatine kinase levels are observed in mesotheliomas, brain tumors, and small cell lung cancer. Decreased creatine kinase levels are observed in kidney cancer, lymphomas, and non-small cell lung cancer. The tumors that show low creatine kinase activity could not form colonies in soft agar. This suggests that the enzyme might be required for tumor establishment (37).

Creatine kinase is highly inducible by estrogens in the female rat reproductive tract and in mammary tumors, as well as in human breast tumors and tissues. In breast cancer cells, the creatine kinase gene promoter has multiple binding sites for nuclear transcription factors and at least two regions of the promoter have been characterized as estrogen-responsive. There is not a binding site for the estrogen receptor, but both CCAAT and TA-rich motifs appear important for the estrogen receptor action. The proximal region (-195 to –37) and the upstream region (-627 to –404) contain a nonconsensus estrogen responsive element flanked by two GC-rich Sp1-binding

sites. These sites bind both the estrogen receptor and other transcription factors such as Sp1. Studies in both breast and non-breast cancer cell lines (HeLa and MCF7) show that the creatine kinase promoter is highly inducible by estrogen. The characteristics of these estrogen-responsive GC-rich elements are similar to those observed in other genes regulated by estrogen receptor alpha/Sp1 through GC-rich motifs including DNA polymerase alpha, c-fos, E2F1, retinoic acid, receptor alpha 1, thymidylate synthase, bcl-2, adenosine deaminase, and insulin growth factor binding protein 4. Estrogen receptor alpha/Sp1 action through GC-rich sites has been reported for several other genes in breast cancer cells and recent studies have characterized similar mechanisms for hormone activation of low density lipoprotein receptor, telomerase, progesterone receptor, EGF, receptor for advanced glycation end products in other cell lines.

Functional Sp1 interacts with other members of the nuclear receptor superfamily such as retionic acid/X receptors, progesterone receptor, chick ovalbumin upstream promoter, transcription factors, and androgen receptor. This suggests an expanding role for the nuclear receptor/Sp1 complexes in regulating gene expression. The cell context is important in estrogen receptor alpha/Sp1 action since GC-rich sites in the creatine kinase promoter were estrogen receptor responsive in breast but not in HeLa cells.

Insulin-like growth factors are known to promote skeletal muscle differentiation. This proliferation is mediated by the Raf/MEK/ERK pathway. Differentiation of the skeletal muscle is regulated by the PI3K pathway. PI3K activity is required for muscle creatine kinase isozyme gene expression. Atk/PKB, a serine and threonine kinase, is also regulated by PI3K. Akt1 is known to inhibit muscle creatine kinase isozyme transcription by suppressing myogenin. Akt2 increases the isozyme's transcription as well as myogenin transcription (166).

Ribosomal Protein L27

The ribosome is a cellular organelle responsible for protein synthesis in all cells. It requires four RNA species and 80 proteins. The human cell contains 4 x 10^6 ribosomes. Roughly 80% of cellular RNA and 5-10% of cellular protein levels are ribosomes. The ribosomal protein of interest here, ribosomal protein L27, is located on chromosome 17. Ribosomal proteins are significantly conserved during evolution. The amino acid sequences of ribosomal proteins are nearly identical in mammals. Ribosomal protein genes are dispersed throughout the genome. Ribosomal protein genes are expressed not only ubiquitously but also abundantly during the course of development. Trans-acting regulatory mechanisms, both transcriptional and translational, have been argued to play a substantial role in coordinating production of ribosomal components in mammals through feedback mechanisms. This makes it challenging to investigate the mechanism of ribosomal protein genes in different human disorders (86).

The eukaryotic ribosome contains over 70 different proteins, many of which likely represent recent evolutionary acquisitions by the ribosome to make protein synthesis more efficient. It seems possible that some of these proteins had other cellular roles prior to their recruitment into ribosomes and could conceivably still retain a nonribosomal function. Evidence has accumulated indicating that a number of ribosomal proteins have a second function apart from their role in the ribosome and in protein synthesis. For example, some of the acidic ribosomal P-proteins appear to exchange on and off the ribosome in vivo, supporting the notion that they may not have an absolute role in protein synthesis. In fact, human ribosomal protein PO, whose encoded product is recognized by antibodies present in patients suffering from systemic lupus erythematosus, has recently been shown to be induced by antitumor agents. The

expression of this gene is also elevated in human tumor cell lines defective for a pathway to correct alkylation damage to DNA, suggesting that PO may have a role in DNA repair. Indeed, the *Drosophila* PO gene was cloned using an antibody originally prepared against a human DNA repair protein. The PO gene is also one of six ribosomal genes found to be overexpressed in colorectal cancers (189).

Ribosomal protein S3 is a protein located on the external surface of the 40S ribosomal subunit, which assembles onto preribosomal particles in the nucleus. It forms part of the domain on the ribosome where the initiation of translation occurs and appears to be directly involved in ribosome-mRNA-aminoacyl tRNA interactions during translation. It has been proposed that S3 has functions other than protein synthesis (89). An example of this is ribosomal protein S3 that functions as both a ribosomal protein and an endonuclease. Prior studies have shown that S3 can be cross-linked to eukaryotic initiation factors eIF-2 and eIF-3, suggesting a role in the initiation of protein translation. The protein is also associated with the nuclear matrix, a result perhaps consistent with the presence of a nuclear localization signal for S3. In view of *E. coli* ribosomal protein S3 acting as a DNA binding protein, this provides further support that S3 may indeed be acting in some capacity involved in DNA metabolism.

The activity of S3 was found to be inhibited by tRNA, a characteristic shared by *E.coli* endonuclease III. The DNA repair protein, endonuclease III, is also a β-elimination catalyst. Certain polyamines and peptides are capable of carrying out a β-elimination reaction at AP sites in DNA. S3 proteins have also been found to contain important DNA N-glycosylase activity. Mutants in the gene encoding S3 in *Drosophila* have a recessive lethal and a dominant Minute phenotype, which suggests a possible other role for the protein (89). S3 has also been found to be cutting oligonucleotides adjacent to 8-oxoguanine residues (189). Ribosomal protein S3

cleaves phosphodiester bonds within a cyclobutane pyrimidine dimer. This activity could relax DNA distortions brought about by the dimer and modulate DNA repair in an overstressed damage situation.

Another example is ribosomal protein S6 in *Drosophila*. It functions as a tumor suppressor in the hematopoietic system and is encoded by the air8, aberrant immune response locus. It is unusual among ribosomal proteins in that it shows developmentally regulated phosphorylation of a cluster of serine residues at the C terminus that is often associated with cell growth and tumorigenesis. In quiescent mammals, this phosphorylation is stimulated by treatments that increase protein synthesis and cell proliferation such as serum growth factors, insulin, tumor promoting agents, transforming viruses, mitogens, and chemical carcinogens. The S6 ribosomal protein is rapidly dephosphorylated when growth is arrested by serum deprivation, contact inhibition, or heat shock. It has therefore been suggested that S6 and its specific kinases may have important regulatory roles in controlling cell growth through the selective translation of unknown particular classes of mRNA. S6 and other ribosomal proteins are known to be overexpressed in human colon carcinomas and liver metastases. In mutants animals loss of expression of the S6 ribosomal protein gene caused growth inhibition in some tissues, and mitotic recombination experiments illustrated that the S6 gene product is required for egg development in the ovary and for cell survival in developing imaginal discs in *Drosophila*. Surprisingly, loss of the same gene product in the hematopoietic system does not cause cell lethality but instead leads to lymph gland hyperplasia, precocious differentiation of plasmatocytes into lamellocytes, and melanotic tumor formation. This pleiotropic syndrome eventually leads to death (185).

Ribosomal protein S2 is known to function in oogenesis. Ribosomal protein L5 forms a complex with Mdm2 and p53 in humans. Lastly, ribosomal protein L22 forms an RNP complex with the Epstein-Barr virus encoded RNA. With this evidence in hand, the role of ribosomal protein other than in protein synthesis needs to be investigated further.

Ribosomal protein L27, the radiation response gene identified through the use of gene-trapping, is relatively unstudied. It is 476 nucleotides in length with a coding region of 155 amino acids. It has 100% identity with the ribosomal protein L27 of rat and chicken. Ribosomal protein L27 is developmentally regulated in the kidney and has been found to be down-regulated 20-fold in comparison with the fetal kidney (56). Ribosomal protein L27 was identified in a proliferation gene cluster through the use of microarray technology in a breast cancer cell line (221). DNA microarrays were used to survey the variation in approximately 8,000 distinct human transcripts in 60 cancer cell lines (150).

Although the ribosome is essential for cell growth and development, the effects of ribosomal mutations and their role in human disease have been largely ignored. One might predict that genetic defects in ribosomal components would cause serious problems with the translational apparatus and result in embryonic death. Quantitative deficiencies in human ribosomal proteins will result in reduced translation capacity and thereby yield abnormal phenotypes (176).

There have been numerous findings of mutations in various ribosomal proteins. Deficiencies in ribosomal protein genes may underlie certain congenital disorders in humans. Ribosomal protein SR (RPS4) is encoded by the X and Y-chromosomes and has been found to be important in Turner syndrome. Turner syndrome is where there is a monosomy of the X chromosome. Deficiency of RPS4 may be responsible for lymphedema and neck webbing

observed in 45 X monosomy individuals and in 45 X monosomy fetuses. Many Turner patients with structurally abnormal X chromosomes have two or more active copies of RPS4 on the X chromosome. Quantitative deficiencies of individual ribosomal proteins could result from gross chromosomal deletions or point mutations in the ribosomal protein genes.

Ribosomal protein 19 has been found to be mutated in patients with Diamond-Blackfan anemia. The mutations found in ribosomal protein 19 in these patients are nonsense, missense, splice sites, and frameshift mutations. Diamond-Blackfan anemia is a rare, chronic anemia characterized by an absent or a decreased level of erythroid precursors in the bone marrow.

The ribosomal protein RPS16, RPL30, and RPL32 lacks a canonical TATA motif an initiate transcription at a C residue that is embedded in a pure pyrimidine stretch flanked by GC rich sequences. RPL30 and RPL32 gene have localized regulatory elements and binding sites for trans-acting factors to regions that span about 80 to 130 base pair of 5' flanking sequence and about 80 base pair of internal sequence. RPS16 promoter has been found to contain an element that specifically binds one or more nuclear factors which points to a transcriptional function. A part of the ribosomal protein S16 was found to bind the Sp1 transcription factor (69).

Ribosomal protein L6 (RPL6) has been mapped within the interval containing the Noonan syndrome locus. Noonan syndrome is an autosomal dominant disorder characterized by a congenital heart defect, typical facial dysmorphy, and short stature. It has an estimated incidence of 1:1000 to 1:2500. RPL6 has been mapped to chromosome 12 at locus 12q22-qter. Ribosomal protein S9 (RPS9) is found in the critical region for retinitis pigmentosa 11. This is an autosomal dominant form of inherited retinal degeneration. The region this protein maps to is on chromosome 19 at 19q13.4.

The L5 protein has been shown to bind specifically to 5S RNA during ribosomal biogenesis, first in the nucleus and then in the nucleolus, as well as on the large ribosomal subunit of the ribosomes. It has been reported that the ribosomal protein L5 complexes with the mdm-2 protein. The L5 protein appears to be a subunit of one of or more isoforms of the mdm-2 protein. mdm-2-L5 protein complexes and mdm-2-L5-p53 complexes have been shown to associate with 5S RNA as well. It is believed that the mdm-2-L5-5S RNA and p53-5.8S RNA complexes are held together in a ribonuclear protein complex by protein-RNA and RNA-RNA interactions at their C-terminal ends. The mdm-2-L5-5S RNA complexes have been detected in cells that overproduce mdm-2 proteins. While the majority of p53 and mdm-2 proteins in a cell reside in the nucleus, some may be associated with the ribosomes. The importance of this interaction could be in the recognition of DNA damage and preventing the cell from translating selected mRNAs. This interaction could potentially block transition from the G_1 to S phase by selectively inhibiting the translation of cyclins or cyclin-dependent kinase mRNAs. What role L5 has in complexing with p53 and mdm-2 still remain unknown (115).

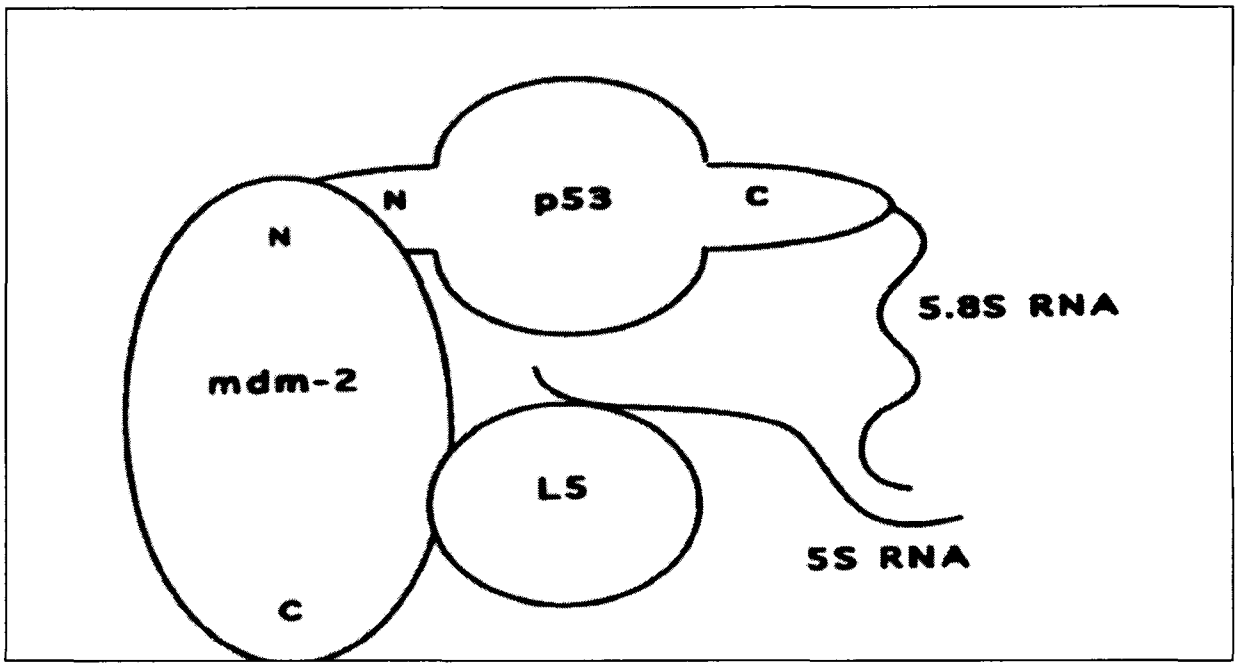


Figure 5.5. Schematic diagram of the putative p53-mdm-2-L5-5S RNA-5.8S RNA ribonucleoprotein complex. p53 and mdm-2 interact through their N-termini, and the 5.8S RNA is bound to the penultimate serine of p53. The L5 protein forms a tight complex with mdm-2 and is bound to the 5S RNA and 5.8S RNA species in a ternary complex. Source: *Marechal, V. et al. The ribosomal L5 protein is associated with mdm-2 and mdm-2-p53 complexes. Molecular and Cellular Biology, Nov 1994, 14(11) p. 7414-7420.*

Eukaryotic Translation elongation factor 1 beta 2

Eukaryotic protein elongation factors are involved in a wide-range of cellular functions including cell division, differentiation, ageing, and transformation (137). The pentameric elongation factor 1 (EF-1) is composed of two alpha subunits, one beta subunit, one gamma subunit, and one delta subunit. The four EF-1 subunits are very abundant and together may constitute up to 5% of the total cellular protein in actively proliferating cells. Tumor and cultured cells have EF-1 subunit levels up to 20-fold higher than those in normal tissue (156).

EF-1 ensures the rapid transport and accurate delivery of aminoacyl tRNA to the ribosome. Translation elongation factor 1 β (EF-1β) is a guanine nucleotide exchange factor (GEF). GEF activity resides in the highly conserved C-terminal domain. EF-1β is responsible for the regeneration of the active form of EF-1α (133). EF-1 has been found to be expressed in higher levels in pancreatic tumors. It has significant homology with EF-1γ. EF-1γ and EF-1α have been found to be highly expressed in certain tumors. EF-1β levels are known to be elevated in tumor cells relative to the other two. There are increased amounts of EF-1γ in pancreatic and colon tumors. Increased levels of EF-1α are seen in colon, breast, lung, and gastric tumors relative to the normal tissue. Increased levels of EF-1 are known to increase the susceptibility of cells to undergo oncogenic transformation (155). EF-1 is overexpressed in both islet and ductal-appearing tumors; this suggests that it does not correspond to a duct cell-specific gene and that its overexpression is a tumor-associated phenomenon. A mutation in EF-1 in yeast can result in translational errors. The potential role of translation in tumorigenesis has largely been unexplored (100).

Translation elongation is the process of adding amino acylated tRNA to the growing polypeptide chain. Translation elongation factor 1 alpha transfers an aminoacylated tRNAs to

the A site of the ribosome. This is a GTP-dependent process catalyzed by eEF1βα (35). This is a brain and muscle-specific isoform of the translation elongation factor 1-β, now called eEF1βα. The human gene encoding this isoform is eEF1B. This gene is encoded by three different genes: EEF1B1, EEF1B2, and EEF1B3.

EEF1B1 is the gene I identified through gene trapping to be radiation responsive EEF1B1. EEF1B1 is mapped to chromosome 15; EEF1B3 is mapped to chromosome 5 at locus 5q12-14. It lacks introns and is truncated relative to the other two genes.

EEF1B2 maps to chromosome 2 and it contains 6 exons. EEF1B2 is located in the endoplasmic reticulum. It maps to locus 2q33-34. A rare, recessive, juvenile-onset motor neuron disease/amyotrophic lateral sclerosis gene (ALS2), which occurs in a single inbred Tunisian family has been mapped to locus 2q33 (35). Amyotrophic lateral sclerosis (ALS) is a heterogeneous group of progressive neurodegenerative disorders associated with the loss of motor neurons in the cerebral cortex, brain stem, and spinal cord. Most cases of the disease are sporadic, with only 5-10% of the cases due to a familial ALS (FALS) heredity. Four types of FALS have been mapped to defined regions in the human genome. The autosomal dominant form with late onset (ALS1), maps to chromosome 21q22.1-q22.2 with is associated with mutations in the Cu-Zn superoxide dismutase-1 gene. The dominant form of juvenile ALS (ALS4), maps to region 9q34. There are two types of autosomal recessive forms of juvenile ALS, ALS2 and ALS5. ALS2 maps to region 2q33-q35 and ALS5 maps to region 15q12-q21. ALS2, type 3 autosomal recessive ALS (RFALS), is found in a large inbred Tunisian kindred. It is characterized by loss of upper motor neurons and spasticity of limb and facial muscles accompanying distal amyotrophy of the hands and feet (61). No coding regions of EEF1B2 have been found to be mutated in ALS2 patients, but noncoding regulatory regions may be mutated.

EEF1B2 has been found to be expressed in a wide variety of tissue types, as expected of a single gene encoding a protein predicted to by essential.

Deletion of EEF1A2 is known to give rise to the wasted mutation, which is a mouse model for motor neuron disease. The autosomal recessive wasted mutation (wst) arose spontaneously in inbred mouse colony HRS/J at Jackson Laboratories in 1972. Homozygous wasted mice (wst/wst) have neurological defects, waste away, and immune system abnormalities, such as a defective response to DNA damage. The mice are normal until weaning at about 21 days old. Then they develop tumors and ataxia. There is extensive neuronal vacuolar degeneration of anterior horn cells of the spinal cord with less severe abnormalities of motor nuclei in the brain stem. The mice lose weight through muscle wasting, have progressive paralysis and die by about 28 days old. The mice also experience progressive atrophy of the spleen and thymus. There are also decreased levels of circulating lymphocytes in the blood. The wasted mice show a defective response to radiation-induced DNA damage, which is reminiscent of a checkpoint defect. It has been found that EF-1α decreases 95% within the first month of postnatal life. EF1α2 is found to increase during this same period. The wasted mutation has been identified as a deletion of 15.8 kb that removes the promoter region and first noncoding exon of EEF1A2, which abolishes transcription of the gene. EEF1A2 is switched on in the skeletal muscle and in the heart postnatal. EEF1A is down-regulated in these tissues until detectable at roughly 25 days old when it becomes essential. EEF1A2 is necessary for normal function once synaptogenesis has been completed (36).

DORA reverse strand protein 1

Dora reverse strand protein 1 (DREV1) is a novel protein found to have a weak homology to ribosomal RNA adenine dimethylases and the NNMT/PNMT/TEMT family of

methyl transferases. It is believed that this protein may have a housekeeping function in the cell. This protein has a function in the recognition of nucleotide methyl groups and possibly a role in methyl group transfer (19). DREV1 contains 5 exons. The 5' UTR region and the upstream region of DREV1 have a high G+C content and the putative promoter region contains a CpG island, which is typical of a housekeeping gene. DREV1 is believed to be a member of an ancient gene family. The DREV1 protein has a motif that is conserved in methylated nucleotide binding.

DREV1 has known homology to DNA methyltransferases. DNA methyltransferases methylate the cytosine residues in CpG dinucleotide islands at specific and selected sites. It is a process associated with transcriptional silencing of genes during mammalian development and differentiation. In malignant cells, methylation patterns are significantly altered.

The DNA methyltransferase 1 (DNMT1) has been found to contain p53-binding sites. p53 was found to increase in its binding to DNMT1 following doses of ionizing radiation of 10 Gy. p53 is then released from the DNMT1 promoter following the DNA damage caused by ionizing radiation, which then allows for an increase in the levels of DNMT1. Elevated levels of DNA methyltransferases may result in the permanent loss of p53 function and may be sufficient to drive the altered patterns of methylation seen in human cancers (134).

Another example of a DNA methyltransferase in the DNA damage response is 0^6-methylguanine-DNA methyltransferase (MGMT). It plays an important role in the protection of cells against the mutagenic and carcinogenic effects of environmental DNA damaging agents. Observed increases in the level of MGMT protein and repair activity upon genotoxic treatment have been found to be due to the stimulation of MGMT gene expression. Experimental evidence supports the idea that the primary cellular signal for the induction of MGMT expression is DNA

damage, notably DNA strand breaks. Ionizing radiation is known to induce the expression of MGMT. p53 is required for MGMT induction following ionizing radiation treatment. The MGMT promoter is known to contain several putative Sp1 binding sites as well (60).

Analysis of the DREV1 gene revealed that the gene conforms to the GT-AG intron/exon splicing rule. The rodent and human DREV genes retain the same intron/exon structure, but C. elegans and D. melanogaster have a different predicted gene structure. This reflection of the different evolutionary lineages could be used to analyze lower vertebrates such as Xenopus species or the cartilaginous fishes which have an adaptive immune system and in which the DORA-encoding gene would appear to be present and could be used as a tool for the estimation of evolutionary distances. DREV1 has been found to be expressed in most cell types and tissues. DREV1 mRNA appears to be expressed by an uncharacterized cell population present in most organs.

An interesting characteristic of this gene is that a smaller, more characterized gene is located on the complement strand within the DREV1 intron 4. This gene is known as DORA, or IGSF6, and is a novel member of the immunoglobulin superfamily. DORA stands for down-regulated by activation. It is 12kb in length and contains 6 exons. It is completely encoded within intron 4 of the DREV1 gene. This organization is conserved between human and mouse. DORA is localized to chromosome 16p13, in a region with few other identified genes. The closest neighbor to this gene is μ-crystallin and no other gene or gene family of this type has been localized to this region. CD19, an immunoglobulin superfamily member expressed on B cells and follicular dendritic cells, has also been found to be located on chromosome 16p11 and is the only other member of the immunoglobulin superfamily located in this region (19).

DORA is a type I membrane protein with a single IgV_1J-type loop. The protein is composed of an extracellular single V loop region and a J chain domain, which is closest in homology to the variable loop regions of the T-cell receptors and the CD8/CD4 receptors. A number of human immune system-derived molecules containing single V domains have been described including CD7, CD8, PD-1, CTLA-4, CD79b, Thy-1, CD83, CD28, and CMRF-35. These molecules perform a variety of functions, and it is important to remember that one cannot define a protein function simply by the expression of an Ig-like domain. However, a number of common features between the above proteins and DORA exist. It is known that these molecules interact primarily with other members of the immunoglobulin superfamily as adhesion or recognition molecules. The J chain region is generally believed to be the region where dimerisation acting via the diglycine bulge and is present on such immunoglobulin family members as: CD8β, CD7, MRC, TCRs, CTLA-4, CD4, CD28 and OX-2. T-cell receptors and CD8β are closest in homology to DORA. It is believed that the J chain region contained in these genes gave rise to the dimerisation of DORA. DORA does not have the cysteine residue present that is necessary to form an intermolecular disulphide bridge. Interestingly, both the human and rat homologues of DORA contain a single tyrosine residue in their transmembrane domains that may aid in the dimerisation through this domain. DORA contains an intracellular domain that is 65 amino acids long and highly charged. This type of domain is accessible to interact with other molecules located in the cytoplasm. The motif just described is believed to be a novel site of phosphorylation. This novel motif may indicate a new mechanism of signal transduction in dendritic cells, including the presence of other novel kinases and phosphatases (18).

Its expression is limited to cells of the lymphohematopoeitic system. It is expressed principally in immature dendritic cells of the myeloid lineage and also in monocytes and

macrophages. The expression of DORA in dendritic cells and cells of the myeloid lineage is restricted to resting cells. DORA was detected in monocytes, skin-derived Langerhans cells, dendritic cells, granulocytes, neutrophils, T cells and in macrophages. It is believed that DORA encodes two separate mRNAs at the same levels in the cell. This is believed to be due to the presence of alternative transcription stop sites on the DORA gene. However, this does not seem to indicate an alternative splicing of this gene. A retrotransposon –like sequence is present in the human genomic DNA between the 3' end of the DORA gene and exon 4 of the DREV1 gene (18).

The protein also contains an intracellular domain containing tyrosine residues. DORA contains a novel motif that may indicate a new mechanism of signal transduction in dendritic cells. Dendritic cells are professional antigen presenting cells, required for the initiation of immune responses. They present antigen in all cell types and can be isolated in various stages of maturation. The conservation in the cytoplasmic domain in exon 5 of DORA suggests a role in signaling via either the conserved tyrosine residue or via the equally well-conserved serine and threonine residues. DORA possibly has a role in antigen uptake, perhaps as a co-receptor in an antigen-receptor complex, or as a molecule involved in homing and recirculation of dendritic cells. DORA has been found to be associated with the development of the vertebrate immune system (19).

The activation of CD40L has also been found to down-regulate DORA. CD40 is expressed on monocytes, dendritic cells, B cells, basophils, eosinophils, T cells, and non-hematopoietic cells such as fibroblasts, endothelial and epithelial cells. The interaction of CD40 with its ligand causes activation of these cells and can lead to differentiation and secretion of cytokines. Ligation of CD40 upregulates co-stimulatory molecules on dendritic cells and

enhances their survival. DORA is one of the first genes identified whose expression is specifically down-regulated by this pathway. The induction and down-regulation of genes by CD40 ligation probably represents a switch in the function of dendritic cells from cells involved in antigen uptake to cells primed for antigen presentation (18).

It is localized to chromosome 16 at the16p11-p12 locus that is postulated to be involved in inflammatory bowel disease (IBD). Other genes localized to this region include: CD19, interleukin-4 receptor, and the CD11 gene complex. Inflammatory bowel disease can involve either or both of the small and large bowels. Inflammatory bowel disease is the general name for diseases that cause inflammation in the small intestine and colon. Crohn's disease and ulcerative colitis are clinical subtypes of idiopathic inflammatory bowel disease. Active IBD is characterized by acute inflammation. Chronic IBD is characterized by architectural changes of crypt distortion and scarring. Crypt abscesses, active IBD consisting of neutrophils in crypt lumens, can occur in many forms of IBD, not just in ulcerative colitis. Ulcerative colitis is a disease that causes inflammation and ulcers in the lining of the large intestine. It involves the colon as a diffuse mucosal disease with distal predominance. The inflammation usually occurs in the rectum and lower part of the colon. Ulcerative colitis rarely affects the small intestine except in the terminal ileum region. The inflammation causes the colon to frequently empty leading to diarrhea. Ulcers form in the inflammation regions where cells have been killed in the lining of the colon. Ulcerative colitis is most common in persons of Caucasian background, particularly in women ages 20 – 25 years old. Patients with prolonged ulcerative colitis are at a greater risk for developing colon cancer and liver diseases including sclerosing cholangitis and bile duct carcinoma. It is not known what causes ulcerative colitis, but it is believed that the

immune system may have abnormalities resulting from a viral or bacterial infection in the inflamed intestinal wall (200).

Crohn's disease and ulcerative colitis are the two most severe digestive afflictions. Tens of thousands of Americans are affected. In contrast to ulcerative colitis, Crohn's disease may affect the small intestine as well as the large intestine. Crohn's disease differs from ulcerative colitis because it causes inflammation deeper within the intestinal wall. Crohn's disease can involve any part of the gastrointestinal tract, but most frequently involves the distal small bowel and colon. It can also occur in the mouth, esophagus, stomach, duodenum, large intestine, appendix and anus. Inflammation is typically transmural and can produce anything from a small ulcer over a lymphoid follicle, an aphthoid ulcer, to a deep fissuring ulcer to transmural scarring and chronic inflammation. One third of the cases of Crohn's disease have granulomas, and extracolonic sites such as lymph nodes, liver, and joints may also have granulomas. The transmural inflammation leads to the development of fistulas between loops of bowel and other structures. Inflammation is typically segmental with uninvolved bowel separating areas of involved bowel. The etiology is unknown, though infectious and immunologic mechanisms have been proposed. There is a bimodal incidence and an increased risk in women and members of the Caucasian race. It is believed that diet is a likely contributor to the incidence of the disease. Genetic factors play a significant role in the pathogenesis of inflammatory bowel disease. It is believed that DORA is unlikely to be involved in the susceptibility to either of the clinical phenotypes of the IBD (200).

There appears to be no relationship between the mRNA expressions of the DORA and the DREV1 genes. The expression appears to not take place in the same cell type, consistent with previous reports of this type of gene organization. Human chromosome 16 has undergone

extensive duplication and rearrangement. Since few other members of the immunoglobulin superfamily are present on this chromosome, DORA is likely to have arisen by duplication, followed by a gene shuffling and transposition to this locus. Pseudogenes are known to reside within introns. That this locus represents the correct gene for DORA is supported by the fact that the intron/exon boundaries are conserved and functionally correct, which is rarely the case for pseudogenes. Most pseudogenes are retrotranscribed copies of functional genes and have no introns.

Little is known about the function of introns and very few examples are available of large human introns harboring other genes (120). The function of the DNA strand opposite the transcribed strand is to serve as a template for double-stranded, semiconservative DNA replication (1). The presence of genes within an intron and transcribed from the opposite strand of another gene has been reported for other genes such as: oligodendrocyte-myelin glycoproteins and the uncharacterized EVI2A and EVI2B which are embedded in the neurofibromatosis type I gene, F8B is embedded in intron 22 of the human factor VIII gene, the src-like adapter protein in the 64 kb intron of the thyroglobulin gene, and the TIMP3 gene and the HCF2 gene on chromosome 22. These events are rare in mammalian cells, although, genes that overlap partially or wholly have been described. It is believed that this type of organization adds to the complexity of the mammalian genome and probably has an importance in the regulation of gene expression or function in the cell.

Neurofibromatosis type 1 (NF1) is a heritable condition that affects 1 in 4,000 people. There are two genes, EV12A and EV12B, which are known to be embedded in introns of the NF1 gene. Both are transcribed from the complementary strand of NF1. EV12A and EV12B both encode putative transmembrane peptides. They both have single introns separating 5'

noncoding exons from single 3' coding exons. Both genes map between two NF1 translocation breakpoints. So far, their functions appear to be unknown. A third gene, OMgp, has also been found to be lying within the same intron as EV12A and EV12B in the NF1 gene. OMgp is an extracellular peptide linked to the outer cell membrane through a glycosylphosphatidylinositol lipid. OMgp may function as a cell adhesion molecule in the myelin central nervous system (181).

Another example of a gene located within an intron of another gene is in the gene encoding factor VIII. Hemophilia is an X-linked bleeding disorder resulting from a defect in the coagulation factor VIII. The gene encoding coagulation factor VIII has 26 exons and spans roughly 200 kb. Factor VIII associated gene (F8A) has been found to be located in the largest intron of the coagulation factor VIII gene. Another gene, factor VIII associated gene (F8B) has also been localized to intron 22 of the factor VIII gene and is transcribed in the same direction as the factor VIII gene. These two genes are believed to function as housekeeping genes since they are associated with a CpG-rich region (99).

Summary & Conclusions

Upon sequencing the gene-trapped MCF10A clones, five candidate radiation response genes were identified through a BLAST homology search. The five genes were: human androgen receptor, human creatine kinase, human ribosomal protein L27, human translation elongation factor 1 beta 2, and DORA reverse strand protein 1. Upon treatment with 0.10, 0.25, 0.50, 1.0, 2.0, and 4.0 Gy the five radiation response genes were found to be down-regulated in comparison to the parental MCF10A cells at all doses. The mRNA expression of the five radiation response genes was evaluated by real-time PCR analysis in various other cell types, such as human fibroblast cells and the TK6 lymphoblastoid cell line, in order to see if the

radiation response being seen was cell-type specific. Upon analysis, the radiation response was not found to be cell-type specific. The relative mRNA expression levels of all five radiation response genes were found to be down-regulated in the other cell types tested compared to the parental MCF10A cells. The fact that the radiation response being seen was not cell-type specific is of importance. This could mean that the five genes being affected by the gamma irradiation treatment must be present in numerous cell types and have important cellular functions with respect to radiation damage response pathways. Clearly these genes are being affected by radiation and the mechanisms by which this is happening need to be elucidated. This requires the determination of specific gene involved so that the molecular pathways can be elucidated.

Of the five radiation response genes identified, DREV1 was the only novel uncharacterized gene discovered. DREV1 is known to have homology to DNA methyltransferases and this is intriguing due to the fact that it has been found to respond to radiation treatment. The radiation response gene DREV1 has been identified to contain the small gene, DORA, on the complement strand within its intron 4. In previous studies, this type of gene organization in mammalian cells does not appear to affect the mRNA expression levels of the two genes. It has been documented in numerous cases that the mRNA expression levels of the two genes are not related. This was verified through the use of real-time PCR in order to analyze the mRNA expression levels of both DREV1 and DORA following ionizing radiation treatment. The MCF10A cells were treated with 0.5, 1.0, 2.0 and 4.0 Gy doses of gamma irradiation and samples were taken at 2, 4, 8, 12, 24, and 30 hours post IR treatment. As documented in the literature, it was found that the mRNA expression levels of the two genes were not in fact related. This type of organization in mammalian cells is rare, so little is known why the mRNA

expression levels of the two genes would not be related. One theory is that the genes are transcribed at different times and by different polymerases. It is also believed that the gene within a gene is transcribed in the opposite direction, although there have been instances where the genes have been transcribed in the same direction.

If a gene such as DREV1 is being affected by radiation, all sorts of downstream effects can possibly occur. Typically DNA methyltransferases methylate genes in order to silence and turn off their expression. If the DREV1 DNA methyltransferase is down-regulated in response to radiation treatment it can no longer properly silence genes, such as tumor suppressors. This could potentially lead to many downstream effects and possibly carcinogenesis if a gene like a tumor suppressor is not silenced in the proper fashion. This type of potential downstream effect produced by a DNA methyltransferase becoming down-regulated was imperative to analyze. This was accomplished through the use of RNA interference to knockdown the expression of the DREV1 gene. Knockdown of the DREV1 gene through siRNA transfection would enable for the detection of cell survival following downregulation of the DREV1 DNA methyltransferase. Upon knockdown and radiation treatment, there did not appear to be any significant change in the number of chromosomal and chromatid aberrations seen in the MCF10A treated cells. This would imply that the DREV1 gene downregulation was not important enough in cellular function to affect survival. The DREV1 gene could still be definitely having an affect on the cell, just not to that high of a degree where survival becomes compromised. Possibly the targets of the DREV1 gene for silencing are not tumor suppressor genes and thus would not lead to an immediate decrease in the survival of the cell. Upon radiation treatment of the DREV1 gene mutagenesis could occur, leading to aberrant methylation patterns in the cell. This could potentially pose another problem in that upstream gene targets of DREV1 may be able to interact

with a gene that they were not supposed to. If a gene upstream in the DREV1 molecular pathway is able to bind and interact with a gene that was supposed to be transcriptionally silenced through DREV1 methylation aberrant events may possibly occur in the cell.

It was demonstrated in an earlier experiment that the DREV1 mRNA expression was down-regulated in comparison with parental up to 30 hours following doses ranging from 0.10 to 4.0 Gy. Downstream effects of the DREV1 target genes may potentially take longer than 30 hours to demonstrate an effect on the cell survival.

The next experimental step needed to take would be to analyze the DREV1 target genes. It is imperative to known what DREV1 is binding to and what downstream cascades it may be involved in. It is also essential to know whether the gene targets DREV1 is binding to in order to ascertain what other genes they interact with that could be upstream in the molecular pathway. This study would benefit from the development of an antibody to the DREV1 gene in order to be able to analyze gene targets. Currently there are no DREV1 antibodies commercially available. Antibodies would allow for the gene targets of DREV1 to be elucidated through such experiments as gel shift and binding assays. If antibodies were prepared, Western blots for protein expression could be performed in order to look at the expression of the DREV1 gene in various cell types at various hours following ionizing radiation treatment. Up to this point only the relative mRNA expression levels of the gene could be quantitated due to this fact. The DREV1 protein could be extremely stable and knockdown might not be enough to produce an effect on cell survival. This could be the case since only the mRNA levels were being knocked down and they are independent of protein expression. The characterization of the DREV1 gene is very important since it is uncharacterized and should be looked into further.

Of the other four genes, creatine kinase had been previously found to be responsive to radiation. The creatine kinase protein has been found to be down-regulated in response to radiation. It is known to be affected by free radical damage induced by ionizing radiation. The identification of the creatine kinase gene through the use of gene trapping has demonstrated that the gene trapping experiment utilized here does have the potential to identify genes responding to radiation treatment.

The functions and characteristics of the five identified radiation response genes have been described in this chapter. Further studies are required to determine the function and significance of the genes or their pathways in the radiation response. It is essential to be able to determine what upstream and downstream genes in the DREV1 molecular pathway are being affected following radiation treatment. Since DREV1 was found to be down-regulated in response to radiation treatment, these upstream and downstream gene targets may potentially be able to interact and bind to genes they may otherwise not due to methylation. Besides DREV1, the gene targets of the other four identified radiation response genes should be analyzed as well. Unlike DREV1, the molecular pathways for androgen receptor and creatine kinase have been fairly well documented. The fact that the androgen receptor was found to respond to radiation in human mammary epithelial cells causes speculation in its role in radiation induced mammary carcinogenesis.

The human ribosomal protein L27 and human eukaryotic translation elongation factor 1 beta 2 have been less characterized. The gene targets of human ribosomal protein L27 are not known at all. It has been documented that numerous other ribosomal proteins have functions in the cell besides protein synthesis. This could potentially be the case for ribosomal protein L27. It is very important to investigate what pathways this gene may be involved in and what other

genes it is binding to following radiation treatment. The human eukaryotic translation elongation factor 1 beta 2 is also relatively uncharacterized. Little is known what genes it is binding to, and the fact that it maps to the region of the ALS2 disease locus is of particular interest. It is very critical that binding assays and gel shift assays be performed for all of these genes since their responses in radiation have not been investigated thoroughly.

To determine the function and significance of the changes in gene expression the elucidation of the genes in the DREV1 molecular pathway are required. This can be accomplished through the use of gel shift and binding assays once DREV1 antibodies have been developed. This next experimental step is essential to perform since DREV1 is a novel gene and none of the gene targets it binds and interacts with have been identified. Currently the radiation responses of the DREV1 gene and the other four identified genes have not been determined. The pathways of this response need to be analyzed in order to determine what the cell is undergoing upon radiation treatment. There are commercially available antibodies for the androgen receptor and creatine kinase. Eukaryotic translation elongation factor 1 beta 2 and the ribosomal protein L27 currently do not have commercially available antibodies, so these would need to be developed as well in order to elucidate what upstream and downstream gene targets are involved in their radiation response pathways. Once antibodies are available for the identified genes the appropriate gel shift and binding assays can be performed in order to elucidate the pathways in the radiation response mechanism. Once the genes involved in the upstream and downstream pathways for each of the five identified genes have been found their response following radiation treatment can be analyzed by determining the change in gene expression via real-time PCR. This would allow for the visualization of how the genes mRNA expression levels are changing dependent upon one another following radiation treatment. This experiment could also be done

by knocking out the expression of one of the genes involved in the molecular pathway of the radiation response gene and analyzing by real-time PCR to see how the upstream and downstream genes in the pathway are responding.

These experiments would be essential to perform so that the mechanism of the radiation response seen can be elucidated and further explored. It is essential to understand the underlying molecular mechanisms in which these identified radiation response genes are involved. Their overall function in the cell is important to know so that upstream and downstream effects occurring following radiation treatment can be elucidated. It is uncertain what role these genes have in the cell, since survival was found to be unaffected upon DREV1 mRNA knockdown. The other genes have yet to be examined for their role in cell survival; an experiment deemed relevant. Since the DREV1gene appears to not be relevant in cell survival, it then becomes necessary to attempt to postulate what role it has in the cell. That is why it becomes imperative to investigate what gene partners it interacts with and whether DREV1 knockdown has an effect in these upstream and downstream molecular partners. This could potentially lead to an effect in cell survival, the knockdown of DREV1 mRNA just may not be an immediate and sudden detrimental effect to the cell since other factors in the molecular pathways in which it is involved in may exert their effects at a later time. This could potentially be a domino effect in which the knockdown of the DREV1 gene causes various downstream and upstream effects in its gene targets, which ultimately lead to cell death.

Prior to measurement of GFP expression levels in the gene-trapped clones, the pooled gene-trapped clone population was sorted into GFP+ and GFP- pooled populations. This was done in order to determine if the initial GFP expression level in the gene-trapped clone would make a difference following radiation treatment. The clones were sorted by flow cytometry

analysis and ELISA assays were performed for the GFP+ and GFP- populations. As seen in Figure 4.1, the pooled gene-trap population of MCF10A cells was sorted into GFP+ and GFP- populations. One potential problem with this sort is that the sort gate for the GFP+ population was not set high enough. This could lead to GFP+ cells being contained in the GFP- pooled population. As seen in Figures 4.3 – 4.6, the basal GFP readings from the ELISA assay for the GFP- population contain numerous GFP+ samples and the opposite is also true for the GFP+ population. Since the sort was most likely done incorrectly and the initial GFP expression level did not appear to have an effect following radiation in the GFP- population, the flow cytometry analysis could have been eliminated.

One of the goals of this work was to be able to quantitate the GFP levels in the MCF10A gene-trapped clones both basally and following a 2.0 Gy dose of ionizing radiation. This step was essential in determining what clone's GFP levels changed at least 2-fold following the radiation treatment. Since the GFP reporter protein is a measure of the endogenous trapped gene's expression level, we could then sequence and analyze the trapped gene further to elucidate its response to radiation. This was accomplished through the use of a sandwich ELISA assay. This experimental procedure was repeated twice, with widely varied results. As seen in Figure 4.7, the GFP expression level in the gene trapped clones identified to be homologous to known genes varied up to +/- 0.20 between the two independent ELISA experiments conducted. In comparison with the real-time PCR analysis of GFP expression, the ELISA assay did not generate reproducible results from the same gene-trapped clones upon repeat examinations. This would lead one to conclude that the ELISA assay is not very reliable for the quantitation of GFP levels in this experimental case. Perhaps, the GFP levels in the clones are not stable for the long periods of time required to set up and run the sandwich ELISA assay. Another possibility could

be the numerous protease inhibitors and reagents utilized in the lysis buffer needed to prepare the cell lysates for the sandwich ELISA assay. Potentially one of these reagents could interfere with GFP expression or cause degradation of the protein before it could be measured.

Real-time PCR and the ELISA assay are quantitatively different in the fact that real-time PCR measures mRNA levels and the ELISA assay is a measure of protein expression. This difference could also contribute to the variation seen in the ELISA. Real-time PCR appears more sensitive possibly because it is measuring the mRNA levels which may not all be translated into protein. This could be one factor contributing to the variability seen among the results. In contrast to the ELISA method of quantifying GFP expression, real-time PCR is very efficient and generates reproducible results consistently. Real-time PCR takes under 3 hours to set-up and complete, whereas the ELISA assay typically takes roughly 2 days to generate results. This time difference could most likely cause the significant variability in the results produced with the ELISA assay. Real-time PCR is also extremely sensitive and specific only for the targeted gene region. The ELISA assay utilized anti-GFP monoclonal antibodies, but the antibodies could still potentially bind to non-GFP antigens causing misleading results. In this instance the ELISA assay did detect gene-trapped clones responding to radiation treatment which did eventually end up to be identified as genes demonstrating a radiation response. It has been clearly shown, though, that this procedure is most likely not the best choice for a future experiment. If more sensitive instruments had been available for use, the ELISA assay would not have even been utilized for this experiment.

As demonstrated through the use of real-time PCR, the five identified genes were all found to respond to radiation at various doses and time points. Statistical analysis proved that the majority of mRNA expression levels analyzed were statistically significant relative to the

parental MCF10A cells. This would allow for the conclusion to be made that the genes are responding to radiation and the responses are significantly different from what is seen in the parental cells.

LITERATURE CITED

1. Adelman, J. et al. *Two mammalian genes transcribed from opposite strands of the same DNA locus*. Science, 1987. 235(4795): p. 1514-1517.
2. Akiyama, K. et al. *A novel method for constructing gene-targeting vectors*. Nucleic Acids Research, 2000. 28(16): p. e77
3. Al-Assar, O., et al., *Regulation of FOS by different compartmental stresses induced by low levels of ionizing radiation*. Radiation Research, 2000. 154: p. 503-514.
4. Altschul, S. F., et al., *Basic local alignment search tool*. J Mol Biol, 1990. 215(3): p. 403-410.
5. American Cancer Society. Breast Cancer Facts and Figures for 2003.
6. Amundson, S. et al. *Differential responses of stress genes to low dose-rate γ irradiation*. Molecular Cancer Research, 2003. 1: p. 445-452.
7. Amundson, S. et al., *Physiological function as regulation of large transcriptional programs: the cellular response to genotoxic stress*. Comparative Biochemistry and Physiology Part B, 2001. 129: p. 703-710.
8. Andres, J.L., et al., *Regulation of BRCA1 and BRCA2 expression in human breast cancer cells by DNA-damaging agents*. Oncogene, 1998. 16(17): p. 2229-41.
9. Andreú, T. et al. *Gene trapping identifies inhibitors of oncogenic transformation*. The Journal of Biological Chemistry, 1998. 273(22): p. 13848-13854.
10. Baker, R. et al. *In vitro preselection of gene-trapped embryonic stem cell clones for characterizing novel developmentally regulated genes in the mouse*. Developmental Biology, 1997. 185: p. 201-214.
11. Balcer-Kubiczek, E., et al., *Csa-19, a radiation-responsive human gene, identified by an unbiased two-gel cDNA library screening method in human cancer cells*. Oncogene, 1997. 14: p. 3051-3057.
12. Band, V., *Preneoplastic transformation of human mammary epithelial cells*. Seminars in Cancer Biology, 1995. 6: p. 185-192.
13. Barcellos-Hoff, M. and S. Ravani, *Irradiated mammary gland stroma promotes the expression of tumorigenic potential by unirradiated epithelial cells*. Cancer Research, 2000. 60: p. 1254-1260.
14. Barnes, D., *p53, apoptosis and breast cancer*. J Mamm Gland Biol Neoplasis, 1996. 1(2): p. 163-75.
15. Bartstra, R., et al, *The effects of fractionated gamma irradiation on induction of mammary carcinoma in normal and estrogen-treated rats*. Radiation Research, 2000. 153: p. 557-569.
16. Bartstra, R., et al., *Induction of mammary tumors in rats by single-dose gamma irradiation at different ages*. Radiation Research, 1998. 150: p. 442-450.
17. Basolo, F., *Transformation of human breast epithelial cells by c-Ha-ras oncogene*. Molecular Carcinogenesis, 1991. 4: p. 25-35.

18. Bates, E.E.M. et al. *CD40L activation of dendritic cells down-regulates DORA, a novel member of the immunoglobulin superfamily.* Molecular Immunology, 1998. 35: p. 513-524.

19. Bates, E.E.M. et al. *The mouse and human IGSF6 (DORA) genes map to the inflammatory bowel disease 1 locus and are embedded in an intron of a gene of unknown function.* Immunogenetics, 2000. 52: p. 112-120.
20. Bennett, L., *Breast Cancer: Genetic predisposition and exposure to radiation.* Molecular Carcinogenesis, 1999. 26(3): p. 143-9.
21. Bièche, I. et al. *Quantitation of androgen receptor gene expression in sporadic breast tumors by real-time RT-PCR: evidence that MYC is an AR-regulated gene.* Carcinogenesis, 2001. 22(9): p. 1521-1526.
22. Boice, Jr., J. D., *Radiation and breast carcinogenesis.* Medical and Pediatric Oncology, 2001. 36: p. 508-513.
23. Brennan, J. and W.C. Skarnes, *Gene trapping in mouse embryonic stem cells.* Methods Mol Biol, 1999. 97: p. 123-38.
24. Brezniceanu, M-L. et al. *HMGB1 inhibits cell death in yeast and mammalian cells and is abundantly expressed in human breast carcinoma.* The FASEB Journal, 2003. 17: p. 1295-1297.
25. Bridge, A. et al. *Induction of an interferon response by RNAi vectors in mammalian cells.* Nature Genetics, 2003. 34(3): p. 263-264.
26. Brooks, S., *Estrogen receptor in a human cell line (MCF-7) from breast carcinoma.* Journal of Biological Chemistry, 1973. 245: p. 6251-6253.
27. Bryś, M. et al. *Androgen receptor status in female breast cancer: RT-PCR and western blot studies.* J Cancer Res Clin Oncol, 2002. 128: p. 85-90.
28. Buchholz, T. and Xifeng Wu, *Radiation-induced chromatid breaks as a predictor of breast cancer risk.* Int J Radiation Oncology Biol Phys, 2001. 49(2): p. 533-537.
29. Bucholz, T., et al., *Tumor suppressor genes and breast cancer.* Radiation Oncology Investigations, 1999. 7: p. 55-65.
30. Bustin, S.A. *Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems.* Journal of Molecular Endocrinology, 2002. 29: p. 23-39.
31. Calaf, G. and Tom Hei, *Establishment of a radiation- and estrogen-induced breast cancer model.* Carcinogenesis, 2000. 21(4): p. 769-776.
32. Cathers, L.E. and M.N. Gould, *Human mammary cell survival following ionizing radiation.* Int J Radiat Biol Relat Stud Phys Chem Med, 1983. 44(1): p. 1-16.
33. Cecconi, F. and Barbara Meyer. *Gene trap: a way to identify novel genes and unravel their biological function.* FEBS Letters, 2000. 480: p. 63-71.
34. Chalfie, M., et al., *Green fluorescent protein as a marker for gene expression.* Science, 1994. 263(5148): p. 802-805.
35. Chambers, D. et al. *Comparative genomic analysis of genes encoding translation elongation factor 1Bα in human and mouse shows EEF1B1 to be a recent retrotransposition event.* Genomics, 2001. 77(3): p. 145-148.
36. Chambers, D. et al. *The lethal mutation of the mouse wasted (wst) is a deletion that abolishes expression of a tissue-specific isoform of translation elongation factor 1α, encoded by the Eef1a2 gene.* Proc. Natl. Acad. Sci. USA, 1998. 95: p. 4463-4468.

37. Choi, H. et al. *Creatine Kinase B is a target molecule of reactive oxygen species in cervical cancer.* Mol. Cells, 2001. 12(3): p. 412-417. Church, D. et al. *Isolation of genes from complex sources of mammalian genomic DNA using exon amplification.* Nature Genetics, 1994. 6: p. 98-104.
38. Cole, L. and Peter Nowell. *Radiation carcinogenesis: The sequence of events.* Science, 1965. 150(3705): p. 1782-1786.
39. Crompton, N., *Programmed cellular response in radiation oncology.* Acta Oncologica, 1998. 11 Suppl: p. 1-49.
40. Davis, R. et al. *Functional androgen receptor confers sensitization of androgen-independent prostate cancer cells to anticancer therapy via caspase activation.* Biochemical and Biophysical Research Communications, 2003. 309: p. 937-945.
41. Dunham, I. et al. *The DNA sequence of human chromosome 22.* Nature, 1999. 402: p. 489-495.
42. Durick, K. et al. *Hunting with Traps: Genome-wide strategies for gene discovery and functional analysis.* Genome Research, 1999. 9: p. 1019-1025.
43. Duxbury, M. et al. *RNA interference: a practical approach.* Journal of Surgical Research, 2004. 117: p. 339-344.
44. Elbashir, S. et al. *Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells.* Nature, 2001. 411: p. 494-498.
45. Elbashir, S. et al. *RNA interference is mediated by 21- and 22-nucleotide RNAs.* Genes & Development, 2001. 15: p. 188-200.
46. Eldridge, S.R., M.N. Gould, and B.E. Butterworth, *Genotoxicity of environmental agents in human mammary epithelial cells: a trigger for human breast cancer.* Prog Clin Biol Res, 1997. 396: p. 125-31.
47. Ethier, S. P. *Human breast cancer cell lines as models of growth regulation and disease progression.* Journal of Mammary Gland Biology and Neoplasia, 1996. 1(1): p. 111-121.
48. Evans, M. *Gene trapping – a preface.* Developmental Dynamics, 1998. 212: p. 167-169.
49. Evans, M. J., M.B. Carlton, and A.P. Russ, *Gene trapping and functional genomics.* Trends Genet, 1997. 13(9): p. 370-4.
50. Fabre-Suer, C. and Stephen Hauschka. *A novel site in the muscle creatine kinase enhancer is required for expression in skeletal but not cardiac muscle.* The Journal of Biological Chemistry, 1996. 271(9): p. 4646-4652.
51. Floss, T. and Wolfgang Wurst. *Functional genomics by gene-trapping in embryonic stem cells.* Methods in Molecular Biology, 2000. 185: p. 347-379.
52. Fornace, Jr., A. J., et al., *The complexity of radiation stress responses: analysis by informatics and functional genomics approaches.* Gene Expression, 1999. 7: p. 387-400.
53. Freeman, W.M. *Quantitative RT-PCR: pitfalls and potential.* Biotechniques, 1999. 26: p. 112-115.
54. Fukushige, S. and Brian Sauer. *Genomic targeting with a positive-selection lox integration vector allows highly reproducible gene expression in mammalian cells.* Proc. Natl. Acad. Sci. USA, 1999. 89: p. 7905-7909.

55. Gallagher, R. and Afshan Malik. *The gene encoding human ribosomal protein L27 is developmentally regulated in human kidney*. Biochemical Society Transactions, 1994. 22: p. 245S.
56. Gentile, M. et al. *Cell cycle arrest and apoptosis provoked by UV radiation-induced DNA damage are transcriptionally highly divergent responses*. Nucleic Acids Research, 2003. 31(16): p. 4779-4790.
57. Gewirtz, D., *Growth arrest & cell death in the breast tumor cell in response to ionizing radiation and chemotherapeutic agents that induce DNA damage*. Breast Cancer Res Treat, 2000. 62(3): p. 223-35.
58. Ginzinger, D. *Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream*. Experimental Hematology, 2002. 30: p. 503-512.
59. Grombacher, T. et al. *p53 is involved in regulation of the DNA repair gene O^6-methylguanine-DNA methyltransferase (MGMT) by DNA damaging agents*. Oncogene, 1998. 17: p. 845-851.
60. Hadano, S. et al. *A yeast artificial chromosome-based physical map of the juvenile Amyotrophic lateral sclerosis (ALS2) critical region on human chromosome 2q33-q34*. Genomics, 1999. 55: 106-112.
61. Hadjantonakis, A-K. and Andras Nagy. *The color of mice: in the light of GFP-variant reporters*. Histochem. Cell. Biol., 2001. 115: p. 49-58.
62. Halaby, D.M. and J.P.E. Mornon. *The immunoglobulin superfamily: an insight on its tissular, species, and functional diversity*. Journal of Molecular Evolution, 1998. 46: p. 389-400.
63. Harborth, J. et al. *Identification of essential genes in cultured mammalian cells using small interfering RNAs*. Journal of Cell Science, 2001. 114: p. 4557-4565.
64. Harborth, J. et al. *Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing*. Antisense and Nucleic Drug Development, 2003. 13: p. 83-105.
65. Hardouin, N. and Andras Nagy. *Gene-trap-based target site for cre-mediated transgenic insertion*. Genesis, 2000. 26: p. 245-252.
66. Hardouin, N. and Andras Nagy. *Mouse models for human disease*. Clinical Genetics, 2000. 57(4): p. 237-244.
67. Hardwick, M. et al., *Peripheral-type benzodiazepine receptor levels correlate with the ability of human breast cancer MDA-MB-231 cell line to grow in SCID mice*. Int J Cancer, 2001. 94: p. 322-327.
68. Hariharan, N. and R.P. Perry. *A characterization of the elements comprising the promoter of the mouse ribosomal protein gene RPS16*. Nucleic Acids Research, 1989. 17(13): p. 5323-5337.
69. Harshman, K. et al. *Comparison of the positional cloning methods used to isolate the BRCA1 gene*. Human Molecular Genetics, 1995. 4(8): p. 1259-1266.
70. Heinloth, A. et al. *ATM-dependent and –independent gene expression changes in response to oxidative stress, gamma irradiation, and UV irradiation*. Radiation Research, 2003. 160: p. 273-290.
71. Henikoff, S. et al. *Gene within a gene: nested Drosophila genes encode unrelated proteins on opposite DNA strands*. Cell, 1986. 44: p. 33-42.
72. Hilakivi-Clarke, L., *Estrogens, BRCA1, and Breast Cancer*. Cancer Research, 2000. 60: p. 4993-5001.

73. Holzschu, D. et al. *A molecular strategy designed for the rapid screening of gene traps based on sequence identity and gene expression pattern in adult mice.* Transgenic Research, 1997. 6: p. 97-106.
74. Hood, E. *RNAi: What's all the noise about gene silencing*? Environmental Health Perspectives, 2004. 112(4): p. A224-A229.
75. Hussain, S. et al. *Direct interaction of FANCD2 with BRCA2 in DNA damage response pathways.* HMG, 2004.
76. Hutchinson, F., *Molecular biology of mutagenesis of mammalian cells by ionizing radiation.* Seminars in Cancer Biology, 1993. 4: p. 85-92.
77. Ishida, Y. and P. Leder, *RET: a poly A-trap retrovirus vector for reversible disruption and expression monitoring of genes in living cells.* Nucleic Acids Res, 1999. 27(24): p. e35.
78. Jackson, S.P. *Sensing and repairing DNA double-strand breaks.* Carcinogenesis, 2002. 23(5): p. 687-696.
79. Jenster, G. *Ligand-independent activation of the androgen receptor in prostate cancer by growth factors and cytokines.* Journal of Pathology, 2000. 191: p. 227-228.
80. Johansson, S., H. Svensson, and J. Denekamp, *Timescale of evolution of late radiation injury after postoperative radiotherapy of breast cancer patients.* Int J Radiat Oncol Biol Phys, 2000. 48(3): p. 745-50.
81. Jongmans, W., *Cellular responses to radiation and the risk of breast cancer.* Eur J Cancer, 1999. 35(4): p. 540-8.
82. Jung, M. and A. Dritschilo, *NF-κB signaling pathway as a target for human tumor radiosensitization.* Seminars in Radiation Oncology, 2001. 11(4): p. 346-351.
83. Jung, R. *Quantitative PCR.* Clin. Chem. Lab. Med., 2000. 38(9): p. 833-836.
84. Kang, Z. et al. *Involvement of proteasome in the dynamic assembly of the androgen receptor transcription complex.* The Journal of Biological Chemistry, 2002. 277(50): p. 48366-48371.
85. Kenmochi, N. et al. *A map of 75 human ribosomal protein genes.* Genome Research, 1998. 8: p. 509-523.
86. Keyse, S.M., *The induction of gene expression in mammalian cells by radiation.* Seminars in Cancer Biology, 1993. 4: p. 119-128.
87. Khanna, K., *Cancer risk and the ATM gene: a continuing debate.* J Natl Cancer Instit, 2000. 92(10): p. 795-802.
88. Kim, J. et al. *Implication of mammalian ribosomal protein S3 in the processing of DNA damage.* The Journal of Biological Chemistry, 1995. 270(23): p. 13620-13629.
89. King, K. et al. *Genetic variation in the IGSF6 gene and lack of association with inflammatory bowel disease.* European Journal of Immunogenetics, 2003. 30: p. 187-190.
90. Kiosses, W. et al. *Characterization of morphological and cytoskeletal changes in MCF10A breast epithelial cells plated on laminin-5: comparison with breast cancer cell line MCF7.* Cell Communication and Adhesion, 2001. 8(1): p. 29-44.
91. Klein, S. et al. *Regulatory element analysis and structural characterization of the human sarcomeric mitochondrial creatine kinase gene.* The Journal of Biological Chemistry, 1991. 266(27): p. 18058-18065.

92. Koufen, P. and Günther Stark. *Free radical induced inactivation of creatine kinase: sites of interaction, protection, and recovery.* Biochimica et Biophysica Acta, 2000. 1501: p. 44-50.
93. Koufen, P. et al. *Free radical-induced inactivation of creatine kinase: influence of the octameric and dimeric states of the mitochondrial enzyme (Mi_b-CK).* Biochem. J., 1999. 344: p. 413-417.
94. Koufen, P. et al. *Inverse dose-rate effects at the level of proteins observed in the presence of lipids.* International Journal of Radiation Biology, 2000. 76(5): p. 625-631.
95. Kuerbitz, S., et al., *Wild-type p53 is a cell cycle checkpoint determinant following irradiation.* Proc Natl Acad Sci, 1992. 89: p. 7491-7495.
96. Lako, M. and Nicholas Hole. *Searching the unknown with gene trapping.* Expert Reviews in Molecular Medicine, 2000.
97. Leong, T., et al., *Mutation analysis of BRCA1 and BRCA2 cancer predisposition genes in radiation hypersensitive cancer patients.* Int J Radiation Oncology Biol Phys, 2000. 48(4): p. 956-965.
98. Levinson, B. et al. *Evidence for a third transcript from the human factor VIII gene.* Genomics, 1992. 14: p. 585-589.
99. Lew, Y. et al. *Expression of elongation factor-1 gamma-related sequence in human pancreatic cancer.* Pancreas, 1992. 7(2): p. 144-152.
100. Li, S. et al. *Functional link of BRCA1 and ataxia telangiectasia gene product in DNA damage response.* Nature, 2000. 406: p. 210-215.
101. Liberto, M. et al. *Rho regulates $p21^{CIP1}$, cyclin D1, and checkpoint control in mammary epithelial cells.* Oncogene, 2002. 21: p. 1590-1599.
102. Lillie, E. et al. *The role of androgens and polymorphisms in the androgen receptor in the epidemiology of breast cancer.* Breast Cancer Research, 2003. 5(3): p. 164-173.
103. Lipskaya, T. Yu. *Mitochondrial creatine kinase: properties and function.* Biochemistry (Moscow), 2001. 66(10): p. 1098-1111.
104. Liu, W. and D.A. Saint. *Validation of a quantitative method for real time PCR kinetics.* Biochemical and Biophysical Research Communications, 2002. 294: p. 347-353.
105. Livak, K. and Thomas Schmittgen. *Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C}{}_T$ method.* Methods, 2001. 25: p. 402-408.
106. Lomen-Hoerth, C. *Characterization of Amyotrophic lateral sclerosis and frontotemporal dementia.* Dementia and Geriatric Cognitive Disorders, 2004. 17: p. 337-341.
107. Lu, S. et al. *Androgen induction of cyclin-dependent kinase inhibitor p21 gene: role of androgen receptor and transcription factor Sp1 complex.* Molecular Endocrinology, 2000. 14(5): p. 753-760.
108. Lu, S. et al. *Androgen regulation of the cyclin-dependent kinase inhibitor p21 gene through an androgen response element in the proximal promoter.* Molecular Endocrinology, 1999. 13(3): p. 376-384.
109. Lu, Y., *Disruption of the ATM gene in breast cancer.* Cancer Genet Cytogent, 2001. 126(2): p. 97-101.
110. Lubahn, D. et al. *Cloning of human androgen receptor complementary DNA and localization to the X chromosome.* Science, 1988. 240(4850): p. 327-330.
111. Maiti et al. *Irradiation selectively inhibits expression from androgen-dependent pem homeobox gene promoter in sertoli cells.* Endocrinology, 2001. 142(4): p. 1567-1577.

112. Malucelli, A. et al. *Quantification of androgen receptor mRNA in tissues by competitive co-amplification of a template in reverse transcription-polymerase chain reaction.* J. Steroid Biochem. Molec. Biol., 1996. 58(5/6): p. 563-568.
113. Mamon, H. et al. *Differing effects of breast cancer 1, early onset (BRCA1) and ataxia-telangiectasia mutated (ATM) mutations on cellular responses to ionizing radiation.* International Journal of Radiation Biology, 2003. 79(0): p. 1-13.
114. Marechal, V. et al. *The ribosomal L5 protein is associated with mdm-2 and mdm-2-p53 complexes.* Molecular and Cellular Biology, 1994. 14(11): p. 7414-7420.
115. Matsuda, E. et al. *Expression profiling with arrays of randomly disrupted genes in mouse embryonic stem cells leads to in vivo functional analysis.* PNAS, 2004. 101(12): p. 4170-4174.
116. Medema, R. Optimizing RNA interference for application in mammalian cells. Biochemical Journal, 2004.
117. Medico, E. et al. *A gene trap vector system for identifying transcriptionally responsive genes.* Nature Biotechnology, 2001. 19: p. 579-582.
118. Mei, J. et al. *Transformation of non-cancerous human breast epithelial cell line MCF10A by the tobacco-specific carcinogen NNK.* Breast Cancer Research and Treatment, 2003. 79: p. 95-105.
119. Meijerink, P. et al. *The gene for the human Src-like adapter protein (hSLAP) is located within the 64-kb intron of the thyroglobulin gene.* Eur. J. Biochem., 1998. 254: p. 297-303.
120. Menichini, P., et al., *A gene-trap approach to isolate mammalian genes involved in the cellular response to genotoxic stress.* Nucleic Acids Research, 1997. 25(23): p. 4803-4807.
121. Midgley, C., et al., *Coupling between gamma irradiation, p53 induction and the apoptotic response depends upon cell type in vivo.* Journal of Cell Science, 1995. 108: 1843-1848.
122. Miki, Y. et al. *A strong candidate for the breast and ovarian susceptibility gene BRCA1.* Science, 1994. 266(5182): p. 66-71.
123. Mitchell, K. et al. *Functional analysis of secreted and transmembrane proteins critical to mouse development.* Nature Genetics, 2001. 28: p. 241-249.
124. Moore, M., *The role of chemoattraction in cancer metastases.* BioEssays, 2001. 23: p. 674-676.
125. Moses, R. *DNA damage processing defects and disease.* Annu. Rev. Genomics Hum. Genet., 2001. 2: p. 41-68.
126. Muller, P. *Processing of gene expression data generated by quantitative real-time RT-PCR.* BioTechniques, 2002. 32(6): p. 2-7.
127. Nisole, S. and Ali Saïb. *Early steps of retrovirus replicative cycle.* Retrovirology, 2004. 1(9): p. 1-20.
128. Pagès, J.C. and Thierry Bru. *Toolbox for retrovectorologists.* The Journal of Gene Medicine, 2004. 6: S67-S82.
129. Paine, T. M., et al., *Characterization of epithelial phenotypes in mortal and immortal human breast cells.* Int J Cancer, 1992. 50: p. 463-473.
130. Park, S., et al., *Hsp25-induced radioresistance is associated with reduction of death by apoptosis: involvement of Bcl2 and the cell cycle.* Radiation Research, 2000. 154: p. 421-428.

131. Pear, W. et al. *Production of high-titer helper-free retroviruses by transient transfection.* Proc. Natl. Acad. Sci. USA, 1993. 90: p. 8392-8396.
132. Pérez, J. et al. *The solution structure of the guanine nucleotide exchange domain of human elongation factor 1β reveals a striking resemblance to that of EF-Ts from Escherichia coli.* Structure, 1999. 7: p. 217-226.
133. Peterson, E. et al. *p53-mediated repression of DNA methyltransferase 1 expression by specific DNA binding.* Cancer Research, 2003. 63: 6579-6582.
134. Petkau, A. *Radiation carcinogenesis from a membrane perspective.* Acta. Physiol. Scand. Suppl., 1980. 492: p. 81-90.
135. Pfaffl, M. *A new mathematical model for relative quantification in real-time RT-PCR.* Nucleic Acids Research, 2001. 29(9): p. 2002-2007.
136. Pizzuti, A. et al. *Human elongation factor EF-1β: cloning and characterization of the EF1β5a gene and assignment of EF-1β isoforms to chromosomes 2, 5, 15 and X.* Biochemical and Biophysical Research Communications, 1993. 197(1): p. 154-162.
137. Ponder, B., *Cancer genetics.* Nature, 2001. 411: p. 336-341.
138. Ponnaiya, B, M. Cornforth, and R. L. Ullrich, *Radiation-induced chromosomal instability in BALB/c and C57BL/6 mice: the difference is as clear as black and white.* Radiation Research, 1997. 147: p. 121-125.
139. Ponnaiya, B. and R.L. Ullrich, *Induction of chromosomal instability in human mammary cells by neutrons and γ-rays.* Radiation Research, 1997. 147: p. 288-294.
140. Preston, D. et al. *Radiation effects on breast cancer risk: a pooled analysis of eight cohorts.* Radiation Research, 2002. 158: p. 220-235.
141. Qin, W. et al. *Molecular characterization of the creatine kinases and some historical perspectives.* Molecular and Cellular Biochemistry, 1998. 184: p. 153-167.
142. Raeymaekers, L. *Basic principles of quantitative PCR.* Molecular Biotechnology, 2000. 15: p. 115-122.
143. Ramakers, C. et al. *Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data.* Neuroscience Letters, 2003. 339: p. 62-66.
144. Rebbeck, T. et al. *Modification of BRCA1-associated breast cancer risk by the polymorphic androgen-receptor CAG repeat.* Am. J. Hum. Genet., 1999. 64: p. 1371-1377.
145. Ree, A. H. et al. *Repression of mRNA for the PLK cell cycle gene after DNA damage requires BRCA1.* Oncogene, 2003. 22: p. 8952-8955.
146. Reynolds, A. et al. *Rational siRNA design for RNA interference.* Nature Biotechnology, 2004. p. 1-5.
147. Ron, E., *Ionizing radiation and cancer risk: evidence from epidemiology.* Radiation Research, 1998. 150 (Suppl.): p. S30-S41.
148. Roshon, M. et al. *Gene trap mutagenesis of hnRNP A2/B1: a cryptic 3' splice site in the neomycin resistance gene allows continued expression of the disrupted cellular gene.* BMC Genomics, 2003. 4(2): p. 1-11.
149. Ross, D. et al. *Systematic variation in gene expression patterns in human cancer cell lines.* Nature Genetics, 2000. 24: p. 227-235.
150. Rothfuss, A. et al. *Induced micronucleus frequencies in peripheral lymphocytes as a screening test for carriers of a BRCA1 mutation in breast cancer families.* Cancer Research, 2000. 60: p. 390-394.

151. Russo, J. et al. *Neoplastic transformation of human breast epithelial cells by estrogens and chemical carcinogens.* Environmental and Molecular Mutagenesis, 2002. 39: p. 254-263.
152. Russo, J., et al., *Cancer risk related to mammary gland structure and development.* Microscopy Research and Technique, 2001. 52: p. 204-223.
153. Sanders, J. et al. *Elongation factor-1 messenger-RNA levels in cultured cells are high compared to tissue and are not drastically affected further by oncogenic transformation.* Nucleic Acids Research, 1992. 20(22): p. 5907-5910.
154. Sanders, J. et al. *Nucleotide sequence of human elongation factor-1β cDNA.* Nucleic Acids Research, 1991. 19(16): p. 4551.
155. Sanders, J. et al. *Immunofluorescence studies of human fibroblasts demonstrate the presence of the complex of elongation factor-1βγδ in the endoplasmic reticulum.* Journal of Cell Science, 1996. 109: p. 1113-1117.
156. Santner, S. et al. *Malignant MCF10CA1 cell lines derived from premalignant human breast epithelial MCF10AT cells.* Breast Cancer Research and Treatment, 2001. 65: p. 101-110.
157. Schmittgen, T. *Real-time quantitative PCR.* Methods, 2001. 25: p. 383-385.
158. Schwartz, J., et al., *Dose-dependent changes in the spectrum of mutations induced by ionizing radiation.* Radiation Research, 2000. 153: p. 312-317.
159. Segawa, T. et al. *Androgen-induced expression of endoplasmic reticulum (ER) stress response genes in prostate cancer cells.* Oncogene, 2002. 21: p. 8749-8758.
160. Shan, L. et al. *Active allele loss of the androgen receptor gene contributes to loss of androgen receptor expression in female breast cancers.* Biochemical and Biophysical Research Communications, 2000. 275: p. 488-492.
161. Skarnes, W. C., *The identification of new genes: gene trapping in transgenic mice.* Curr Opin Biotechnol, 1993. 4(6): p. 684-9.
162. Soule, H. D., et al., *Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10.* Cancer Research, 1990. 50: p. 6075-6086.
163. Stanford, W. et al. *Gene-trap mutagenesis: past, present and beyond.* Nature Reviews Genetics, 2001. 2: p. 756-768.
164. Stribinskis, V. and Kenneth Ramos. *Decoding the riddle: The dawn of RNAi for the study of gene-gene and gene-environment interactions.* Environmental Health Perspectives, 2004. 112(4): p. A210-A211.
165. Sumitani, S. et al. *Akt1 and Akt2 differently regulate muscle creatine kinase and myogenin gene transcription in insulin-induced differentiation of C2C12 myoblasts.* Endocrinology, 2002. 143(3): p. 820-828.
166. Suy, S., et al., *Association of Grb2 with Sos and Ras with Raf-1 upon gamma irradiation of breast cancer cells.* Oncogene, 1997. 15(1): p. 53-61.
167. Szelei, J. et al. *Androgen-induced inhibition of proliferation in human breast cancer MCF7 cells transfected with androgen receptor.* Endocrinology, 1997. 138(4): 1406-1412.
168. Tabotta, W. et al. *Genetic reshuffling reconstitutes functional expression cassettes in retroviral vectors.* The Journal of Gene Medicine, 2001. 3: p. 418-426.
169. Tait, L., *Ultrastructural and immunocytochemical characterization of an immortalized human breast epithelial cell line, MCF-10.* Cancer Research, 1990. 50: p. 6087-6094.

170. Thomas, T. et al. *A new gene trap construct enriching for insertion events near the 5' end of genes*. Transgenic Research, 2000. 9: p. 395-404.
171. Thompson, E., et al., *The invasive and metastatic properties of hormone-independent but hormone-responsive variants of MCF-7 human breast cancer cells*. Clin Exp Metastasis, 1993. 11: p. 15-26.
172. Thompson, L.H. and David Schild. *Recombinational DNA repair and human disease*. Mutation Research, 2002. 509: p. 49-78.
173. Tonin, P. N., *Genes implicated in hereditary breast cancer syndromes*. Seminars in Surgical Oncology, 2000. 18: p. 281-286.
174. Tsuda, H., *p53 as a prognostic factor of breast cancer*. Nippon Rinsho, 2000. 58 Suppl: p. 413-7.
175. Uechi, T. et al. *A complete map of the human ribosomal protein genes: assignment of 80 genes to the cytogenetic map and implications for human disorders*. Genomics, 2001. 72: p. 223-230.
176. Ullrich, R. L., *Cellular and molecular changes in mammary epithelial cells following irradiation*. Radiation Research, 1991. 128: p. S136-S140.
177. Vallis, K.A., et al., *Identification of radiation-responsive genes in vitro using a gene-trap strategy predicts for modulation of expression by radiation in vivo*. Radiation Research, 2002. 157: p. 8-18.
178. van Bekkum, D. W. and J. J. Broerse, *Induction of mammary tumors by ionizing radiation*. Radiat Environ Biophys, 1991. 30: p. 217-220.
179. van Dijk, J. et al. *A novel, essential control for clonality analysis with human androgen receptor gene polymerase chain reaction*. American Journal of Pathology, 2002. 161(3): p. 807-812.
180. Viskochil, D. et al. *The gene encoding the oligodendrocyte-myelin glycoproteins is embedded within the neurofibromatosis type 1 gene*. Molecular and Cellular Biology, 1991. 11(2): p. 906-912.
181. Voss, A. et al. *Efficiency assessment of the gene trap approach*. Developmental Dynamics, 1998. 212: 171-180.
182. Vuoristo, J.T. et al. *C18orf2, a novel, highly conserved intronless gene within intron 5 of the GNAL gene on chromosome 18p11*. Cytogenetics and Cell Genetics, 2001. 93: p. 19-22.
183. Wang, F. et al. *Transcriptional activation of rat creatine kinase B by 17β-estradiol in MCF-7 cells involves an estrogen responsive element and CG-rich sites*. Journal of Cellular Biochemistry, 2002. 84: p. 156-172.
184. Watson, K. et al. *Drosophila homolog of the human S6 ribosomal protein is required for tumor suppression in the hematopoietic system*. Proc. Natl. Acad. Sci. USA, 1992. 89: p. 11302-11306.
185. Weichselbaum, R., et al., *Biological consequences of gene regulation after ionizing radiation exposure*. Journal of the National Cancer Institute, 1991. 83(7): p. 480-483.
186. Weil, M., et al., *Radiation induces genomic instability and mammary ductal dysplasia in Atm heterozygous mice*. Oncogene, 2001. 20: p. 4409-4411.
187. Wempe, F. et al. *Gene trapping identifies transiently induced survival genes during programmed cell death*. Genome Biology, 2001. 2(7): p. 1-10.

188. Wilson III, D.M. et al. *Drosophila ribosomal protein S3 contains an activity that cleaves DNA at apurinic/apyrimidinic sites*. The Journal of Biological Chemistry, 1994. 269(14): p. 25359-25364.
189. Xu, B. et al. *Phosphorylation of serine 1387 in Brca1 is specifically required for the Atm-mediated S-phase checkpoint after ionizing radiation*. Cancer Research, 2002. 62: p. 4588-4591.
190. Xu, X. et al. *Genetic interactions between tumor suppressors Brca1 and p53 in apoptosis, cell cycle and tumorigenesis*. Nature Genetics, 2001. 28: p. 266-271.
191. Xu, Y. et al. *Functional comparison of single- and double-stranded siRNAs in mammalian cells*. Biochemical and Biophysical Research Communications, 2004. 316: p. 680-687.
192. Yamanouchi, H. et al., *Relationship between stages of mammary development and sensitivity to gamma-ray irradiation in mammary tumorigenesis in rats*. Int J Cancer, 1995. 60: p. 230-234.
193. Yoshida, M., et al., *A new strategy of gene trapping in ES cells using 3'RACE*. Transgenic Res, 1995. 4(4): p. 277-87.
194. Yousef, GM et al. *Identification and characterization of KLK-L4, a new kallikrein-like gene that appears to be down-regulated in breast cancer tissues*. The Journal of Biological Chemistry, 2000. 275(16): p. 11891-11898.
195. Yousef, GM et al. *The androgen-regulated gene human kallikrein 15 (KLK15) is an independent and favorable prognostic marker for breast cancer*. British Journal of Cancer, 2002. 87: p. 1294-1300.
196. Zambrowicz, B. and G. A. Friedrich, *Comprehensive mammalian genetics: history and future prospects of gene trapping in the mouse*. Int J Dev Biol, 1998. 42: p. 1025-1036.
197. Zhang, L. et al. *Effects of the Bowman-Birk inhibitor on clonogenic survival and cisplatin- or radiation-induced cytotoxicity in human breast, cervical, and head and neck cancer cells*. Nutrition and Cancer, 1999. 33(2): p. 165-173.
198. Zou, W. et al. *Caveolin-1 haploinsufficiency leads to partial transformation of human breast epithelial cells*. Anticancer Research, 2003. 23(6C): p. 4581-4586.
199. Zouali, H. et al. *Etiology of inflammatory bowel disease*. Int. J. Colorectal. Dis, 1999. 14(1): p. 2-9.
200. Cabanes, A. et al. *Prepubertal estradiol and genistein exposure up-regulate BRCA1 mRNA and reduce mammary tumorigenesis*. Carcinogenesis, 2004. 25(5): p. 741-748.
201. Dorgan, J. et al. *Diet and sex hormones in girls: Findings from a randomized controlled clinical trial*. Journal of the National Cancer Institute, 2003. 95(2): p. 132-141.
202. Hilakivi-Clarke, L. et al. I *Maternal and prepubertal diet, mammary development and breast cancer risk*. Journal of Nutrition, 2001. 131: p. 154S-157S.
203. Angèle, S. et al. *ATM haplotypes and cellular response to DNA damage: Association with breast cancer risk and clinical radiosensitivity*. Cancer Research, 2003. 63: p. 8717-8725.
204. Thorstenson, Y. et al. *Contributions of ATM mutations to familial breast and ovarian cancer*. Cancer Research, 2003. 63: p. 3325-3333.
205. Duffy, S. W. et al. *The relative contributions of screen-detected in situ and invasive breast carcinomas in reducing mortality from the disease*. European Journal of Cancer, 2003. 39: p. 1755-1760.

206. Mothersill, C. and Colin Seymour. *Radiation-induced bystander effects, carcinogenesis and models.* Oncogene, 2003. 22: p. 7028-7033.
207. Prasad, K. et al. *Health risks of low doses of ionizing radiation in humans: a review.* Experimental Biology and Medicine, 2004. 229(5): p. 378-382.
208. Iannuzzi, C..M. et al. *ATM mutations in female breast cancer patients predict for an increase in radiation-induced late effects.* Int. J. Radiat. Oncol. Biol. Phys., 2002. 52(3): p. 606-13.
209. Trott, K. R. and A. Teibe. *Lack of specificity of chromosome breaks resulting from radiation-induced genomic instability in Chinese hamster cells.* Radiat. Environ. Biophys., 1998. 37(3): p. 173-6.
210. Chan, J. Y. et al. *Differential gene expression in a DNA double-strand-break repair mutant XRS-5 defective in Ku80: analysis by cDNA microarray.* J. Radiation Research (Tokyo), 2001. 42(4): p. 371-385.
211. Zhang, Y. and M. A. Frohman. *Using rapid amplification of cDNA ends (RACE) to obtain full-length cDNAs.* Methods in Molecular Biology, 1997. 69: p. 61-87.
212. Hicks, G. G. et al. *Functional genomics in mice by tagged sequence mutagenesis.* Nature Genetics, 1997. 16(4): p. 338-44.
213. Chowdhury, M. et al. *GA/GC-rich sequence confers Tat responsiveness to human neurotropic virus promoter, JCVL, in cells derived from central nervous system.* Oncogene, 1993. 8(4): p. 887-92.
214. Skarnes, W. C. et al. *Capturing genes encoding membrane and secreted proteins important for mouse development.* Proc. Natl. Acad. Sci. USA, 1995. 92(14): p. 6592-6.
215. Shuman, S. *Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase.* Journal of Biological Chemistry, 1994. 269(51): p. 32678-84.
216. Lee, L. G. et al. *Allelic discrimination by nick-translation PCR with fluorogenic probes.* Nucleic Acids Research, 1993. 21(16): p. 3761-6.
217. Kuenen-Boumeester, V. et al. *Immunohistochemical determination of androgen receptors in relation to oestrogen and progesterone receptors in female breast cancer.* Int. Journal Cancer, 1992. 52(4): p. 581-4.
218. Isola, J. J. *Immunohistochemical demonstration of androgen receptor in breast cancer and its relationship to other prognostic factors.* Journal of Pathology, 1993. 170(1): p. 31-5.
219. Hall, R. E. et al. *Expression of the androgen receptor and an androgen-responsive protein, apolipoprotein D, in human breast cancer.* British Journal of Cancer, 1996. 74(8): p. 1175-80.
220. Yang, G. P. et al. *Combining SSH and cDNA microarrays for rapid identification of differentially expressed genes.* Nucleic Acids Research, 1999. 27(6): p. 1517-23.

APPENDICES

The standard deviations for the real-time PCR experiments analyzing the dose-rate effect of the five identified genes at various time points follow.

	Dose (Gy)	Time (Hr)	Average mRNA expression relative to the parental	Standard deviation
Androgen receptor	0.10	2	-17.910000	1.778000
		4	-18.490000	0.296000
		8	-19.450000	0.000000
		12	-19.190000	1.660000
		24	-15.680000	0.098000
		30	-14.860000	0.155000
		MCF10A	-3.550000	1.868000
	0.25	2	-15.050000	0.197000
		4	-15.050000	0.403000
		8	-11.830000	0.049000
		12	-19.090000	8.444000
		24	-18.520000	2.390000
		30	-13.140000	0.247000
		MCF10A	2.770000	0.898000
	0.50	2	-12.550000	0.148000
		4	-12.140000	1.061000
		8	-13.620000	1.102000
		12	-12.230000	0.597000
		24	-14.700000	0.056000
		30	-16.550000	0.381000
		MCF10A	2.730000	1.142000
	1.0	2	-2.760000	0.393000
		4	-2.716670	0.404000
		8	-40.920000	0.563000
		12	-40.970000	0.000000
		24	-44.260000	0.000000
		30	-41.180000	2.080000
		MCF10A	-4.850000	0.374000
	2.0	2	-37.458330	1.100730
		4	-36.773330	4.821290
		8	-36.525000	5.575537
		12	-36.321670	0.5609714
		24	-39.450000	0.208000
		30	-40.500000	0.135000
		MCF10A	2.340000	0.342000
	4.0	2	-10.240000	0.000000
		4	-12.406700	2.210000
		8	-5.210000	0.192000
		12	-5.340000	0.021200
		24	-13.050000	0.806000
		30	-16.490000	0.106000
		MCF10A	19.770000	1.185000

	Dose (Gy)	Time (Hr)	Average mRNA expression relative to the parental	Standard deviation
DREV1	0.10	2	-13.920	0.149
		4	-14.630	0.077
		8	-14.200	1.250
		12	-14.970	0.438
		24	-13.400	0.346
		30	-10.770	6.323
		MCF10A	0.810	0.000
	0.25	2	-13.750	0.629
		4	-15.090	4.077
		8	-13.350	0.311
		12	-13.550	0.206
		24	-12.420	0.155
		30	-13.320	0.014
		MCF10A	2.030	0.098
	0.50	2	-21.810	7.060
		4	-16.350	0.770
		8	-16.090	0.360
		12	-16.280	0.763
		24	-11.210	0.163
		30	-10.540	0.184
		MCF10A	-1.290	6.430
	1.0	2	-0.11667	0.120
		4	-1.840	0.696
		8	-17.430	0.538
		12	-6.450	0.176
		24	-6.680	0.381
		30	-6.970	0.560
		MCF10A	8.980	0.107
	2.0	2	-18.980	0.226
		4	-19.100	0.982
		8	-17.400	1.244
		12	-20.820	0.905
		24	-34.230	0.875
		30	-31.550	0.500
		MCF10A	9.440	0.282
	4.0	2	-36.740	0.183
		4	-40.140	0.134
		8	-18.420	0.007
		12	-18.960	0.516
		24	-21.500	2.404
		30	-17.950	0.438
		MCF10A	4.560	0.000
	Dose (Gy)	**Time (Hr)**	**Average mRNA expression relative to the parental**	**Standard deviation**
Creatine Kinase	0.10	2	-15.53	0.198
		4	-17.10	0.756
		8	-14.68	0.113
		12	-19.61	0.742
		24	-15.22	0.148
		30	-16.40	0.166

		MCF10A	0.41	1.717
	0.25	2	-19.57	4.38
		4	-21.57	7.565
		8	-13.67	0.982
		12	-12.26	2.07
		24	-18.41	0.19
		30	-16.13	1.459
		MCF10A	1.36	0.346
	0.50	2	-21.76	13.79
		4	-12.86	0.445
		8	-12.08	0.367
		12	-13.99	0.0707
		24	-14.26	0.396
		30	-24.68	0.307
		MCF10A	2.58	0.461
	1.0	2	-2.73833	0.247
		4	-1.38667	0.26
		8	-22.28	0.22
		12	-7.76	0.33
		24	-7.68	1.33
		30	-6.15	0.318
		MCF10A	9.37	0.388
	2.0	2	-23.38	0.384
		4	-21.87	0.197
		8	-21.66	0.494
		12	-20.79	0.113
		24	-26.29	0.035
		30	-25.18	0.097
		MCF10A	8.02	0.345
	4.0	2	-24.23	0.028
		4	-23.90	0.033
		8	-29.92	0.622
		12	-31.97	0.813
		24	-26.69	0.242
		30	-28.98	0.1697
		MCF10A	-1.52	0.106
	Dose (Gy)	**Time (Hr)**	**Average mRNA expression relative to the parental**	**Standard deviation**
Ribosomal Protein L27	0.10	2	-16.14	0.00
		4	-16.78	3.401
		8	-20.69	0.035
		12	-18.51	3.48
		24	-16.57	0.19
		30	-17.18	1.44
		MCF10A	-0.97	3.696
	0.25	2	-17.11	0.551
		4	-17.70	2.036
		8	-12.96	0.254
		12	-14.30	0.452
		24	-12.76	1.367
		30	-16.29	5.599
		MCF10A	1.71	0.212
	0.50	2	-14.79	0.049
		4	-17.29	4.133

		8	-12.73	0.749
		12	-12.57	0.332
		24	-11.95	0.066
		30	-13.87	0.096
		MCF10A	2.12	1.88
	1.0	2	-0.10333	0.061
		4	-0.496667	0.126
		8	-6.99	0.557
		12	-9.43	1.05
		24	-8.05	0.374
		30	-6.86	0.155
		MCF10A	7.75	0.473
	2.0	2	-11.67	1.98
		4	-6.73	0.784
		8	-3.67	0.307
		12	-5.83	0.24
		24	-3.85	0.466
		30	-3.68	0.077
		MCF10A	14.98	0.13
	4.0	2	-4.67	0.084
		4	-4.20	0.09
		8	-12.45	0.304
		12	-11.64	0.205
		24	-11.02	0.177
		30	-11.14	0.15
		MCF10A	5.30	0.121
	Dose (Gy)	**Time (Hr)**	**Average mRNA expression relative to the parental**	**Standard deviation**
EEF1B2	0.10	2	-18.6800	4.2230
		4	-18.7500	3.6510
		8	-16.5600	2.9560
		12	-15.3400	0.0000
		24	-15.0000	1.2500
		30	-14.7000	1.0300
		MCF10A	0.2900	1.0300
	0.25	2	-14.0900	0.3110
		4	-19.9100	0.3250
		8	-14.1900	1.6780
		12	-16.6900	2.6970
		24	-14.0800	1.0300
		30	-14.2200	0.4590
		MCF10A	2.0200	0.1480
	0.50	2	-14.8600	0.5860
		4	-14.8800	1.5900
		8	-13.9800	0.0490
		12	-13.8100	0.0000
		24	-11.7300	0.0670
		30	-11.2800	0.0990
		MCF10A	0.5600	0.2400
	1.0	2	-1.6800	0.0580
		4	-0.28333	0.2420
		8	-5.5000	0.8310
		12	-4.6100	0.1620
		24	-5.5800	0.1270

		30	-5.4800	0.4240
		MCF10A	10.0900	0.5020
	2.0	2	-4.9400	0.2750
		4	-2.6100	0.0210
		8	0.3000	0.2590
		12	-0.4800	0.2190
		24	-4.9400	0.5230
		30	-4.8700	0.4140
		MCF10A	17.4300	0.5930
	4.0	2	-8.5900	0.2900
		4	-7.6600	0.1690
		8	-12.9500	0.1060
		12	-13.0700	0.4380
		24	-12.2100	0.1697
		30	-12.0500	0.1500
		MCF10A	4.3100	0.0490