INT. J. RADIAT. BIOL 1997, VOL. 71, NO. 6, 667–674

What will molecular biology contribute to our understanding of radiation-induced cell killing and carcinogenesis?

E. J. HALL

Abstract. The vast body of radiobiological data accumulated with mammalian systems *in vitro* and *in vivo* has had an enormous impact on radiotherapy. However, while quantitative, these data are essentially phenomenological, and it is only in the last decade or so that the techniques of molecular biology allow basic mechanisms to be understood. This will be illustrated by two examples, one involving cell killing and the other carcinogenesis. The identification and sequencing of repair and checkpoint control genes in the yeast *S. pombe* allow the mechanism of sensitivity/resistance to radiation to be understood at the molecular level. The development of techniques to identify mutations in mismatch repair genes have made it possible to show that such mutations are associated with a wide range of human cancers and are a likely mechanism of radiation induced malignancies. Tikvah Alper would have been delighted to see the central role that microorganisms have played in these recent developments.

1. Introduction

I knew Tikvah Alper for close to 40 years, from the time that I was a graduate student at Oxford. When I gave my first talks at the British Institute of Radiology, and the Association for Radiation Research, a principal fear was that Tikvah would be instantly on her feet to probe the weakness in the argument with an incisive question, or put her finger on the missing piece of data!

At that time in the 1950s most radiobiological research was conducted with simpler systems, such as bacteria (Tikvah's choice) or *Vicia faba* (my choice). Over the years, there were great pressures to move to more 'relevant' biological systems with the development, first, of *in vitro* techniques to grow mammalian cells in culture, followed by quantitative transplantable tumour systems in animals, and finally normal tissues *in vivo* scoring both early- and late-responding tissues. The move from prokaryotes to mammalian cells and from the petri dish to the animal was prompted by the conviction that data with more complex systems would be more relevant to the problems of radiation therapy. History has shown this to be the case. Forty years has seen the accumulation of a vast body of highly quantitative radiobiology that led to the application of particle therapy, sensitizers and protectors, hyperfractionation and accelerated treatment, all familiar bandwagons of the 1970s and 1980s. There is no denying that mammalian cell radiobiology had an enormous impact on radiation therapy.

However, while quantitative, this body of data is essentially phenomenological since the basic mechanisms involved in the endpoints observed are seldom known or understood. It is only during the past decade or so that the techniques of molecular biology have allowed a few glimpses of the basic mechanisms involved in radiation effects. Our knowledge is extremely limited at the present time. It would have been to Tikvah's delight to note that the key observations that led to our present insights resulted from experiments in simpler systems, such as bacteria and yeast. I propose to illustrate this with two examples from work in our own laboratory, one involving cell killing and the other carcinogenesis.

2. The *rad9* repair gene in yeast

Lieberman and colleagues (1992) produced the gamma ray dose response curves (Figure 1) for the wild type fission yeast *Schizosacchomyes Pombe* as well as for cells containing the mutant allele, *rad9–192*. The mutant is more sensitive both to ionizing radiation and UV light by a factor close to 100. This radiosensitivity can be completely reversed by the transfection of the normal $rad9^+$ gene from wild type cells. Both $rad9^+$ and *rad9–192* genes have been isolated and their DNA sequences determined. The sequencing gel is shown in Figure 2; there is a single nucleotide base-pair difference, cytosine-guanine substituting for thymine-adenine.

The gene contains a total of about 1200 base-pairs, (Figure 3) and yet the alteration of a single base pair leads to this dramatic change of radiosensitivity. The reason is not difficult to understand. This single base change from T to C in the gene is translated into the coding for a different amino acid. The amino acid change is from leucine to proline, which leads to a distortion in the protein produced as shown in Figure 4. The amino acids valine, leucine, and alanine can all join in a linear fashion, but the substitution of proline causes a 90° distortion in the protein which affects its ability to fold and makes it non-functional. The protein is made, but cannot

Center for Radiological Research, College of Physicians & Surgeons of Columbia University, New York, NY. USA.

0955–3002/97 $12.00

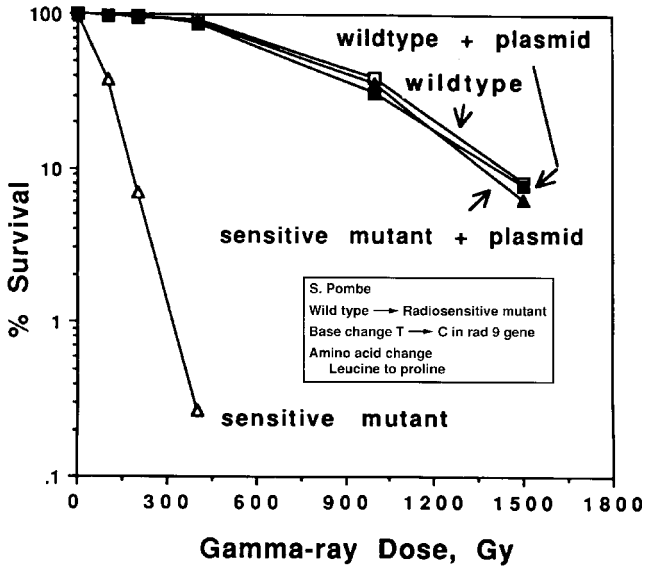

Figure 1. Gamma ray dose-response curves for the yeast *Schizosaccharomyces pombe*. Data for the wild type are shown as open squares, with those for the radiosensitive mutant shown as open triangles. The closed triangles show data for the radiosensitive strain, into which a plasmid had been transformed containing the *rad9* wild-type gene; wild-type resistance is restored by this means; introducing the cosmid into wild-type cells did not affect radiosensitivity (closed squares). (Redrawn from Lieberman *et al.* 1992. Reprinted by kind permission of Springer–Verlag, New York Inc.)

function. Thus, the mechanism of inactivation of the *rad9* gene can be understood at the molecular level.

At first, it was thought that the *rad9* gene in *S. pombe* was involved solely in the repair of DNA damage. In fact it turns out to have two quite separate functions; in addition to being a DNA repair gene, it is also a molecular checkpoint gene. This was demonstrated in some ingenious experiments in which mutants of intermediate radiosensitivity were isolated which showed the ability to repair, without regaining the G_2/M checkpoint control (Lieberman 1994, 1995).

Molecular checkpoint genes, in general, serve the function of ensuring that the initiation of late events are dependent upon the completion of early events. In specific terms, in mammalian cells exposed to radiation, the function of a checkpoint gene is to hold the cells in G_2, to check for the integrity of their chromosomes before the complex task of mitosis has begun. In *S. pombe*, it turns out that the *rad9* gene is a molecular checkpoint gene, as well as a repair gene. Cells that are deficient in this gene, do not stop in G_2 after radiation, but proceed immediately through mitosis. This lack of a checkpoint function is just as important as a lack of repair in making these mutant cells hypersensitive to radiation.

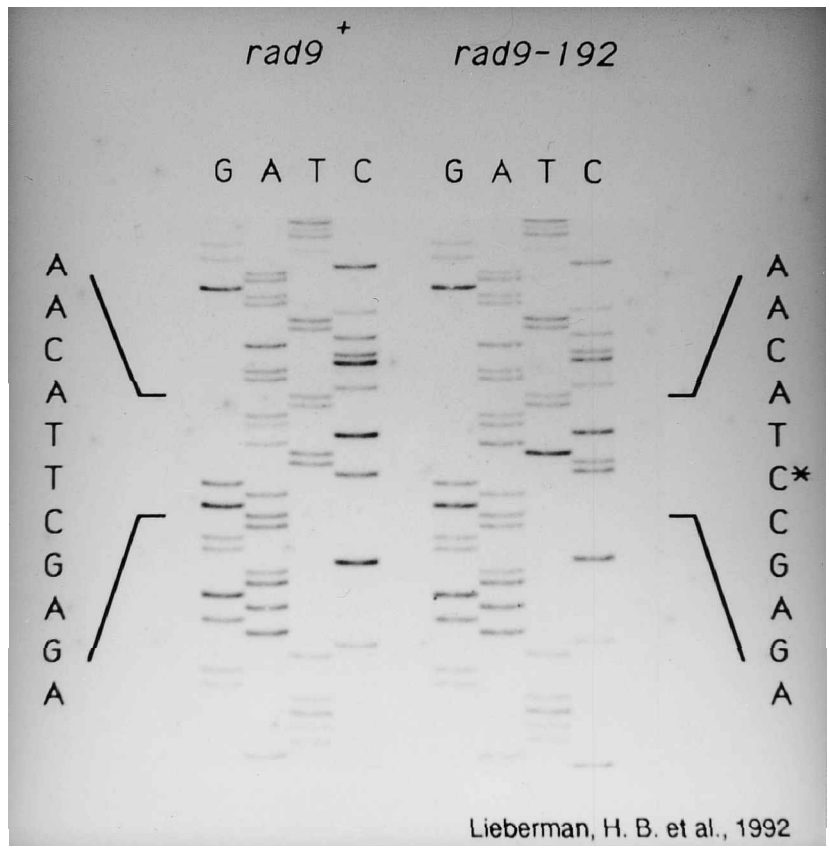

Figure 2. Part of a DNA sequencing gel indicating the single base-pair difference between the *S. pombe rad9*$^+$ and *rad9–192* genes. Derived from Lieberman 1995 and reproduced with kind permission of the publishers ©Genetics Society of America.

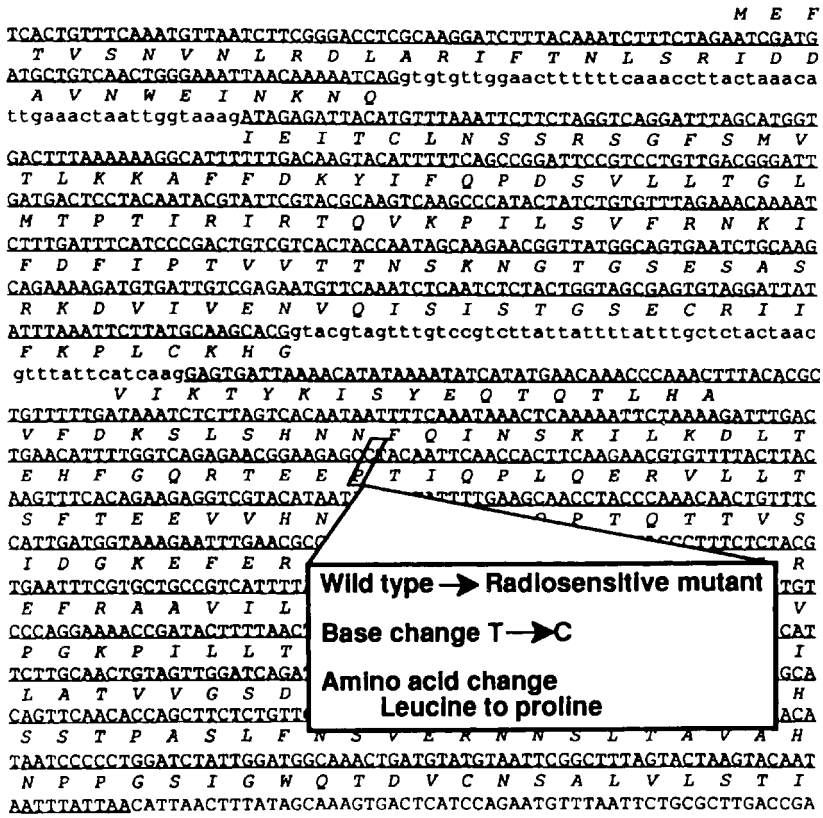

Figure 3. The radioresistant wild-type *S. pombe* contains the *rad9*$^{+}$ gene. The radiosensitive strain contains the mutant allele *rad9–192*. Both genes contain a 1092 bp open-reading frame. A comparison of the DNA sequence of the two genes reveals a single nucleotide base-pair difference, a CG substituting for a TA. This would cause the coding for proline instead of leucine in the mutant gene, promoting a dramatic change in the three-dimensional structure of the protein. Derived from Lieberman 1995 and reproduced with kind permission of the publishers ©Genetics Society of America.

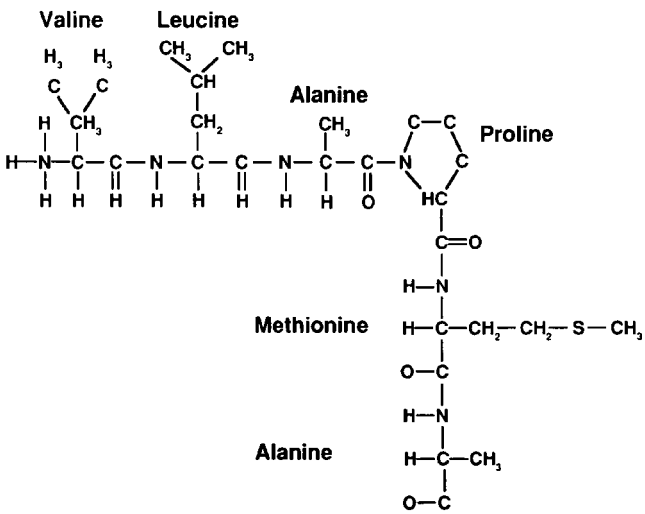

Figure 4. The amino acids valine, leucine, and alanine can join in a linear fashion. Proline, however, causes a 90° distortion in the protein chain produced. This occurs as a result of the mutation in the *rad9–192* gene which inactivates the protein.

It is interesting to note that this *rad9* gene is conserved to varying degrees across several strains of yeast, and a human homologue has recently been identified on human chromosome 11q 13.1 to 13.2 (Lieberman 1996, unpublished data).

3. Mismatch repair

Interest in mismatch repair genes heightened with the discovery that they were responsible for the mutator phenotype associated with a predisposition for hereditary non-polyposis colon cancer (HN PCC) and possibly other familial cancers. The initial clue to this novel molecular mechanism was the discovery of deletions of long monotonic (dA-dT) runs in a sub set of human colon cancers (Peinado *et al.* 1992). Soon afterwards insertions or deletions at mono-, di- and tri nucleotide repeat sequences were discovered in sub sets of colon tumours (Aaltonen *et al.* 1993, Ionov *et al.* 1993, Thibodeau *et al.* 1993), as well as in a majority of colon cancers from individuals with HNPCC (Peltomaki *et al.* 1993). This phenotype has

also been detected in several other types of human malignancies especially those associated with Type 2 Lynch syndrome (Lynch *et al.* 1993). These various investigations culminated in the identification and cloning of the human hMSH 2 gene (Fischel *et al.* 1993, Leach *et al.* 1993), which maps to a locus linked to HNPCC on chromosome 2p21-22 and whose homologues in *Saccharomyces cerevisiae* and *Escherichia coli* are involved in the process of DNA mismatch repair (Cleaver 1994).

3.1. *Mismatch repair in* E. coli

The primary function of mismatch repair genes in *E. coli* appears to be to scan the genome as it replicates and to spot errors of mismatch as the DNA is replicated, i.e. as the new strand is laid down using the stable methylated strand as a template.

Errors of mismatch that result from slippage are particularly likely. When a mismatch error is noted, the repair genes locate a recognition site (the base sequence CTAG) on either side of the error and cut the DNA at these sites. The incorrect piece of DNA is removed and a new strand synthesized using the stable methylated strand as a template. (This is illustrated in Figure 5). This mechanism can correct an error of a single base, or a loop of up to 4 bases. The genes in *E. coli* are well known and have their homologues in human cells.

4. Mismatch repair in human cells

Mismatch repair, then, is a process well understood in bacteria such as *E. coli* and appears to be highly conserved. There is substantial homology in the gene products between *E. coli* and the human, though the precise mechanisms by which the human miss-match repair genes work is by no means as clear as in the case of *E. coli*.

A growing family of human genes has been associated with HN PCC by means of linkage analysis and studies of mutational mapping. (Lindblom *et al.* 1993, Bronner *et al.* 1994, Nicolaides *et al.* 1994, Papadopoulos *et al.* 1994).

Table 1 is the current listing of human mismatch repair genes together with their chromosomal location. The list may not yet be complete. The mismatch repair process in yeast and bacteria involves a large

Table 1. Human mismatch repair genes.

Gene	Chromosomal location	Germline mutations in HNPCC cases (reported family studies)
hMSH2	2p21–22 (2p16?)	60%
hMLH1	3p21	30%
hPMS1	2q31–33	4%
hPMS2	7P21	4%
GTBP	2P16 (2p21–22?)	0%

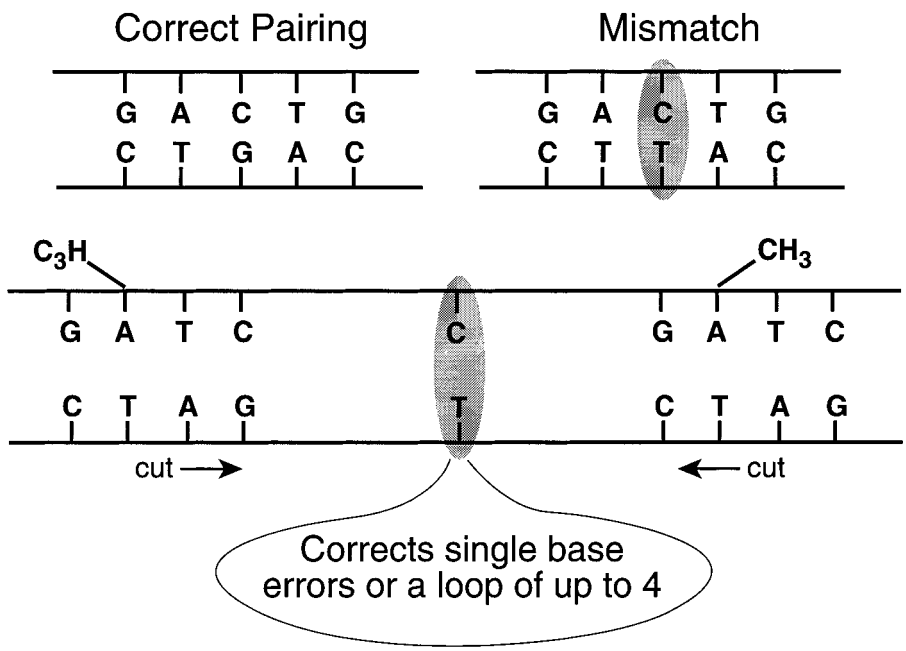

Figure 5. Simplified schema to illustrate the function of mismatch repair genes in *E. coli*. Correct base pairing includes G with C and A with T. When an incorrect pairing occurs due, for example, to slippage during the synthesis of the new strand of DNA, mismatch repair genes find 'recognition sites' (the sequence GATC) on either side of the error and cut the new DNA strand. A matching base is then inserted. The error may be a single base, or sometimes a loop of up to 4 bases.

number of proteins and so it is likely that additional causes of HN PCC remain to be uncovered.

4.1. *Identification of mismatch repair*

Cells with defective or non-functioning mismatch repair genes can be identified by two quite different techniques.

(a) The use of a selectable reporter system that inserts an exogenous long repeat sequence into the cells in question and measures the mutation rate in it; or,
(b) Measuring the mutation rate in one or more of the many endogenous repeat sequences that already exist in every human cell; the so-called micro satellite instability assay.

Both techniques have strengths and weaknesses and both are far from perfect. The limitation of the selectable reporter system is that it can be applied only to cells that are actively growing, and growing well. This is a serious limitation, since establishing a cell line from a human tumour is no small task. The technique cannot be used retrospectively with archival material such as paraffin block sections of tumours. By contrast microsatellite instability requires only a sample of tumour DNA, and does not require growing cells, i.e. the technique can be applied to archival material. Its weakness stems from the fact that there are over 100 000 repeat sequences to choose from in the human genome; which one does one choose, and how many are enough? There are also problems of repeatability and inter-laboratory agreement that have not been reconciled.

5. The selectable reporter system

The construct illustrated in Figure 6 consists of the sequence CA repeated 12 or 13 times, immediately downstream of the ATG start codon of the Hygromycin B resistance gene. There is also a neo resistance gene to simply identify cells that have taken up the construct. The construct is packaged into a retro-viral vector in order to readily infect cells. This ingenious system was devised by Kahn *et al.* (1995). The procedure to determine whether cells have a functional mismatch repair gene is as follows:- Control cells are infected with a construct containing 12 CA repeats, as illustrated in Figure 7. This codes exactly for 8 amino acids, the hygromycin resistance gene is in frame, so that the cells will grow in the presence of the drug. The test cells are infected with a construct containing 13 CA repeats; this puts the hygromycin gene out of frame and so the cells will *not* grow in the presence of the antibiotic. However, if the cells have a defect in mismatch repair, this leads to mutations, such as deletions or insertions, most likely due to slippage of DNA strands during replication. If this leads to a repeat sequence

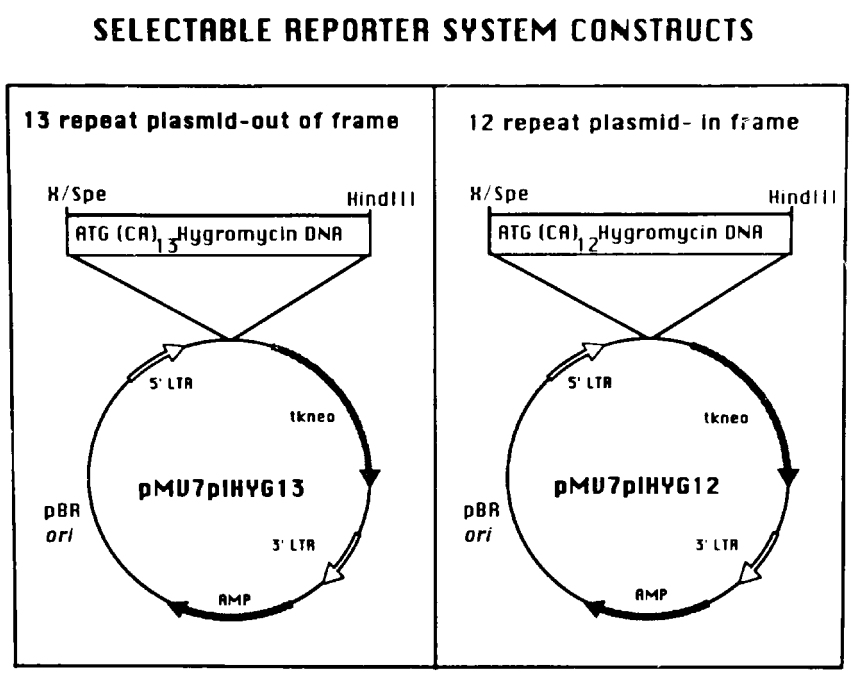

Figure 6. Representation of the constructs having (CA)13 and (CA)12 repeat tracts inserted immediately downstream of the hygromycin B resistance gene ATG start codon. The neomycin resistance gene constitutively driven by the herpes virus thymidine kinase promoter (tk-neo) provides a selectable marker.

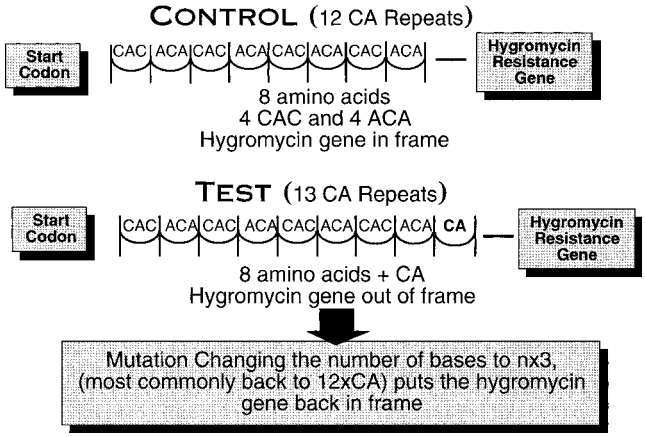

Figure 7. Simplified schema to illustrate the selectable reporter system. Twelve CA repeats between the start codon and the hygromycin resistance gene code for exactly 8 amino acids. The hygromycin gene is in frame and the cells containing the construct will grow in the presence of the antibiotic. The test construct contains 13 CA repeats, which does not translate into an exact number of amino acids and so the hygromycin resistance gene is out of frame. Cells containing this construct will not grow in the presence of the antibiotic. However, if a mutation occurs in the CA repeats which changes the number of bases to any integral number of triples, the hygromycin gene will again be in frame and cells will grow in the presence of hygromycin.

that is an exact multiple of 3 bases, the hygromycin resistance gene will be in frame again and the cells will grow in the presence of the antibiotic. In typical experiments with cell lines derived from colon cancer, cell lines with functional mismatch repair exhibit a mutation rate in the inserted repeat sequence of about 10^{-5}, while cells with defective mismatch repair, have a mutation rate two orders of magnitude higher at about 10^{-3}. This technique is being used to assess which spontaneous human tumours are associated with mismatch repair defects, and in particular to investigate whether this is a common mechanism for radiation induced malignancies.

6. Microsatellite instability

In this technique a region of DNA from the tumour containing a multiple repeat sequence is amplified by PCR and then studied to determine mutation frequency in that repeat sequence. For a given tumour, several different multiple repeat sequences are usually studied. (Aaltonen *et al.* 1993, Thibodeau *et al.* 1993), Table 2 compiled by Dr. Scott Kahn, shows the status of studies with mismatch repair in a variety of human tumours as of late 1995. It appears to be a factor of some importance in a wide variety of human tumours, but the variations in the results highlight the problems inherent in the technique.

Table 2. Microsatellite instability in sporadic tumours.

Multiple loci		Single loci	
Colorectal	11–28%	Brain	2%
Endometrial	15–22%	Liver	3–10%
Gastric	18–39%	Bladder	3–23%
Pancreas	67%	Ovary	10–16%
CML	74%		
Lung (small cell)	45%		
Renal	25%		
Breast	0–20%		
Esophagous	10%		
Germ cell	3–18%		
Lung (non SCC)	2–34%		
Prostate	6–65%		

6.1. *Homology between species*

Mismatch repair has been studied for many years in microorganisms and is well understood. A knowledge of the DNA sequences of the mismatch repair genes in *E. coli* was a vital step in the identification of homologues in human cells. Table 3 shows the homology between *E. coli* and human cells in this regard. The old saying that an elephant is just an *E. coli* only more so, sounds more and more reasonable as time goes by and additional homologues across the species are identified.

6.2. *Mismatch repair, ploidy and radiosensitivity*

It is a consistent observation in tumour cells from a variety of sites that those with defective or non-functional mismatch repair genes are invariably diploid. This is a remarkable observation since the vast majority of tumour cells are anything but diploid. There is no *known* reason for an association between ploidy and mismatch repair, but the strength of the association points to the likelihood that the mismatch repair genes have functions over and above those described earlier and that are associated with base errors. Not only are cells with mismatch repair defects invariably diploid, but they also tend to be radiosensitive compared to similar tumour types

Table 3. Human genes encoding protein homologs of bacterial genes.

hMSH2	*Mut S*
hMLH1, hMPS1, hMPS2	*Mut L*

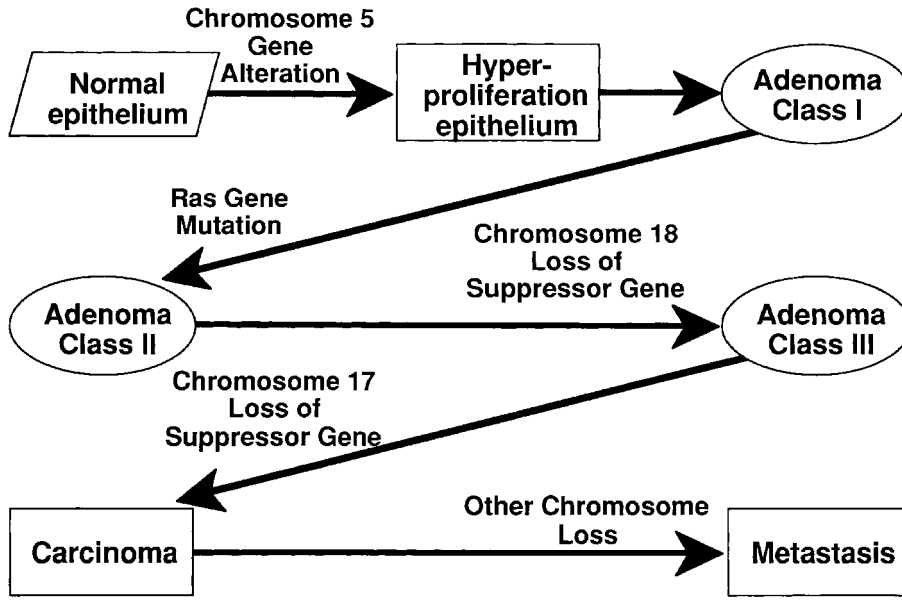

Figure 8. Cancer has long been thought to be a multi-step process, with operational terms such as initiation, promotion and progression. Colon cancer is an excellent example where this multi-step process can be characterized by a series of chromosomal and molecular events that accompany the evolution from hyperproliferation to a metastatic malignant tumour. (Redrawn from the ideas of Fearon *et al.* 1990).

that are aneuploid (Dr. Scott Kahn, personal communication 1996).

6.3. *Multistep nature of cancer*

Perhaps the most pervasive dogma in cancer research is that carcinogenesis is a multistage process. The implication is that there are a number of distinct events that may be separated in time. This idea is almost 69 years old and is exemplified by the skin cancer experiments in mice which introduced the concepts of initiation, promotion and progression as stages in tumour development. The most recent reincarnation of this idea is the model suggested by Vogelstein and his colleagues to describe the carcinogenic process in colorectal tissue (Fearon *et al.* 1990). This is illustrated in Figure 8. The process from normal epithelium to a metastatic tumour appears to include multiple chromosomal changes as well as mutations in several oncogenes and tumour suppressor genes. At least one aspect of the model must cause consternation to the radiation biologist, namely how can a single brief exposure to a low dose of radiation result in six or seven mutations at different loci? This number of mutations may be consistent with the interpretation of Armitage and Doll as the incidence of cancer in adults increases with the sixth power of age, but that so many mutations could be caused by a single modest dose of radiation is, to say the least, improbable. And yet there is hard evidence from the Japanese survivors of the A-bomb that a prompt exposure to a modest dose of radiation can induce a whole spectrum of malignancies including cancers of the digestive tract. A possibility that seems more likely is that the radiation causes a mutation in one of the genes responsible for the stability of the genome and/or the fidelity of replication; leading to what has been called a mutator phenotype.

The discovery of mutations in a mismatch repair gene in the case of hereditary non-polyposis colorectal cancer shows that a single induced mutation can lead to instability and a mutator phenotype. This instability, in turn, leads to a cascade of events with multiple changes possible. This is an attractive hypothesis to account for radiation induced cancer by a single modest dose.

7. Conclusion

In her own research, Dr. Alper chose to concentrate on loss of proliferative capacity as the main effect of ionizing radiation, studying the phenomenon in bacteria and yeast in the belief that the simpler the system, the greater the likelihood that the mechanisms could be understood. In her widely used book 'Cellular Radiobiology' she also concentrated on loss of reproductive integrity pointing out that for radiation to produce a cancer, the affected cell must first survive! Had she lived to see the spectacular advances in molecular genetics that have occurred in the past decade, she would be heartened to see the seminal contribution from studies with micro-organisms.

Acknowledgements

Based on research supported by CA49065 awarded by the National Cancer Institute. The author takes great pleasure in acknowledging the part played by Drs Howard Lieberman and Scott Kahn, who not only read and corrected the manuscript, but developed many of the ideas upon which it is based.

References

Aaltonen, L., Peltomäki, P., Leach, F., Sistonen, P., Pylkkänen, L., Mecklin, J., Järvinen, H., Powel, S., Jen, J., Hamilton, S., Peterson, G., Kinzler, K., Vogelstein, B. and De La Chapelle, A., 1993, Clues to the pathogenesis of familial colorectal cancer. *Science*, **260,** 812–816.

Bronner, C., Baker, S., Morrison, P., Warren, G., Smith, L., Lescoe, M., Kane, M., Earabino, C., Lopford, J., Lindblon, A., Tannergard, P., Bollag, R., Godwin, A., Ward, D., Nordenskjold, M., Fishel, R., Kolodner, R. and Liskay, R., 1994, Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature*, **368,** 258–261.

Cleaver, J., 1994, It was a very good year for DNA repair. *Cell*, **76,** 1–4.

Fearon, E. R., Cho, K. R., Nigro, J. M., Kern, S. E., Simons, J. W., Ruppert, J. M., Hamilton, S. R., Preisinger, A. C., Thomas, G., Kinzler, K. W. and Vogelstein, B., 1990, Identification of a chromosome 189 gene that is altered in colorectal cancers. *Science*, **247,** 49–56.

Fischel, R., Lescoe, M. K., Rao, M., Copeland, N., Jenkins, N., Garber, J., Kane, M. and Kolodner, R., 1993, The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell*, **75,** 1027–1038.

Ionov, Y., Peinado, M., Malkhosyan, S., Shibata, D. and Perucho, M., 1993, Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature*, **263,** 558–561.

Kahn, S. M., Klein, M. G., Jiang, W., Xing, W. Q., Xu, D. B., Perucho, M. and Weinstein, I. B., 1995, Design of a selectable reporter for the detection of mutations in mammalian simple repeat sequences. *Carcinogenesis*, **16,** 1223–1228.

Leach, F., Nicolaides, N., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L., Nystrom-Lahti, M., Guan, X.-Y., Zhang, J., Meltzer, P., Yu, J.-W., Kao, F.-T., Chen, D., Cerosaletti, K., Fourneir, R., Todd, S., Lewis, T., Leach, R., Naylor, S., Weissenbach, J., Mecklin, J., Jarvinen, H., Petersen, G., Hamilton, S., Green, J., Jass, J., Watson, P., Lynch, H., Trent, J., De La Chapelle, A., Kinzler, K. and Vogelstein, B., 1993, Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell*, **75,** 1215–1225.

Lieberman, H. B., 1995, Extragenic suppressors of *Schizosaccharomyces pombe* rad9 mutations uncouple radioresistance and hydroxyurea sensitivity from cell cycle checkpoint control. *Genetics*, **141,** 107–117.

Lieberman, H. B. and Hopkins, K. M., 1994, *Schizosaccharomyces malidevorans* and *Sz. octosporus* homologues of *Sz. pombe* rad9, a gene that mediates radioresistance and cell-cycle progression. *Gene*, **150,** 281–286.

Lieberman, H. B., Hopkins, K. M., Laverty, M. and Chu, H. M., 1992, Molecular cloning and analysis of *Schizosaccharomyces pombe* rad9, a gene involved in DNA repair and mutagenesis. *Molecular Gen. Genet.*, **232,** 367–376.

Lindblom, A., Tannergard, P., Werclius, B. and Nordenskjold, M., 1993, Genetic mapping of a second locus prediposing to hereditary non-polyposis colon cancer. *Nature Genetics*, **5,** 279–282.

Lynch, H., Smyrk, T., Watson, P., Lanspa, S., Lynch, J., Lynch, P., Cavalieri, J. and Boland, C., 1993, Genetics, natural history, tumor spectrum, and pathology of hereditary nonpolyposis colorectal cancer: an updated review. *Gastroenterology*, **104,** 1535–1549.

Nicolaides, N., Papadopoulos, N., Liu, B., Wei, Y.-F., Carter, K., Ruben, S., Rosen, Haseltine, W., Fleichmann, R., Fraser, C., Adams, M., Venter, J., Dunlop, M., Hamilton, S., Petersen, G., De La Chapelle, A., Vogelstein, B. and Kinzler, K., 1994, Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature*, **371,** 75–80.

Papadopoulos, N., Nicolaides, N., Wei, Y.-F., Ruben, S., Carter, K., Rosen, C., Haseltine, W., Fleischmann, R., Fraser, C., Adams, M., Venter, J., Hamilton, S., Petersen, G., Watson, P., Lynch, H., Peltomaki, P., Mecklin, J., De La Chapelle, A., Kinzler, K. and Vogelstein, B., 1994, Mutation of a mutL homolog in hereditary colon cancer. *Science*, **263,** 1625–1629.

Peinado, M., Malkhosyan, S., Velazquez, A. and Perucho, M., 1992, Isolation and characterization of allelic losses and gains in colorectal tumors by arbitrarily primed polymerase chain reaction. *Proceedings of the National Academy of Science* USA, **89,** 10065–10069.

Peltomaki, P., Aaltonen, L., Sistonen, P., Pylkkanen, L., Mecklin, J., Jarvinen, H., Green, J., Jass, J., Weber, J., Leach, F., Petersen, G., Hamilton, S., De La Chapelle, A. and Vogelstein, B., 1993, Genetic mapping of a locus predisposing to human colorectal cancer. *Science*, **260,** 810–812.

Thibodeau, S., Bren, G. and Schaid, D., 1993, Microsatellite instability in cancer of the proximal colon. *Science*, **260,** 816–819.