1 of 1

Energy & Technology

*Unlocking the Mysteries
of DNA Repair*

University of California • Lawrence Livermore National Laboratory

About the Cover

DNA, the double-helix blueprint for life, has the remarkable ability to repair itself through the use of certain genes. For more than a decade, Lawrence H. Thompson and members of the DNA repair group have been studying mutations affecting DNA repair genes in an attempt to understand these genes and how DNA repairs itself. This issue of *Energy and Technology Review* is an attempt to document that effort.

The group's research has focused on repair processes that correct damage caused by chemical mutagens and radiation (both ionizing and ultraviolet). The particular emphasis has been excision repair, a multistep process in human cells

in which damaged pieces of DNA are removed and replaced with intact pieces.

Using classical genetic procedures as well as new gene-transfer and recombinant-DNA techniques, Thompson and the group have been able to identify, isolate, clone, and characterize several human repair genes. They have also shown that a defect in one of these repair genes can be linked to the recessive genetic disorder xeroderma pigmentosum. Their next projects will be to isolate repair proteins using cloned DNA repair genes and to develop special lines of DNA-repair-deficient mice that can serve as animal models of human repair syndromes.

Prepared for DOE under contract
No. W-7405-Eng-48

About the Journal

The Lawrence Livermore National Laboratory, operated by the University of California for the United States Department of Energy, was established in 1952 to do research on nuclear weapons and magnetic fusion energy. Since then, we have added other major programs, including laser fusion and laser isotope separation, biomedical and environmental science, strategic defense, and applied energy technology. These programs, in turn, require research in basic scientific disciplines, including chemistry and materials science, computer science and technology, engineering, and physics. The Laboratory also carries out a variety of projects for other Federal agencies. *Energy and Technology Review* is published monthly to report on unclassified work in all our programs. Please address any correspondence concerning *Energy and Technology Review* to Mail Stop L-3, Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, CA 94551.

■ April 1993

Scientific Editor

William A. Quirk

Editorial Staff

June Canada	Harriet Kroopnick
Lauren de Vore	Lori McElroy
Kevin Gleason	Nona M. Sanford
Robert D. Kirvel	

Art Staff

Treva Carey	Paul M. Harding
Lynn M. Costa	Kathryn House
Janet Crampton-Pipes	George Kitrinos

Compositor

Louisa Cardoza

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California and shall not be used for advertising or product endorsement purposes.

Printed in the United States of America
Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, Virginia 22161
Price codes: printed copy A03, microfiche A01

UCRL-52000-93-4
Distribution Category UC-700
April 1993

Energy & Technology

DNA Repair Research at Lawrence Livermore National Laboratory: A Summary iii

Unlocking the Mysteries of DNA Repair 1

We are studying the interaction of known DNA-damaging agents, repair genes, and their protein products in an attempt to decipher the biochemical pathways that allow DNA to encode its own repair. This article summarizes our decade-long efforts to understand the most important DNA repair process in human cells, called excision repair.

Why We Use Rodent Cells 4

Rodent and human cells are remarkably similar in terms of their DNA repair processes, and several rodent cell lines are used to isolate repair-deficient mutants.

Why We Use Hybrid Cells 8

By using rodent-rodent and rodent-human hybrid cells, we can find out how many genes are involved in repair processes and identify the human chromosomes on which repair genes reside.

The Excision Repair Process 11

In excision repair, damaged or inappropriate pieces of DNA are identified and removed and then replaced with correct pieces.

LLNL Strategy for Studying Human Repair Genes 15

The strategy for studying human repair genes involves classical genetic procedures as well as new gene-transfer and recombinant-DNA techniques.

Recent Findings on DNA Repair Genes 20

Since 1984, researchers at LLNL and elsewhere have identified, isolated, cloned, and characterized approximately 15 genes involved in the repair of human DNA.

Human DNA Repair Syndromes 22

Several relatively rare diseases in humans are directly associated with defects in repairing damaged DNA. Lab researchers have linked the *ERCC2* repair gene to the inherited disease xeroderma pigmentosum.

Future Studies 25

Two exciting areas of future work are the use of cloned DNA repair genes to isolate repair proteins and the development of a new technology involving special lines of mice that are deficient in DNA repair, which will give us valuable animal models of the human repair syndromes.

Glossary 29

Abstract 34

MASTER

THIS DOCUMENT IS UNCLASSIFIED

DNA Repair Research at Lawrence Livermore National Laboratory: A Summary

The broad mission of biological research at Lawrence Livermore National Laboratory (LLNL) is to identify and characterize effects on human health arising from the development and use of energy sources. Research conducted in the Biology and Biotechnology Research Program at the Laboratory is directly linked to the long-term goals of the U.S. Department of Energy (DOE), in particular, the Office of Health and Environmental Research (OHER). One part of our multidisciplinary effort, largely funded by the DOE, emphasizes an understanding of the structure of mammalian chromosomes and chromatin, which packages the DNA. An important aspect of our research program is to advance our knowledge about what environmental agents and what spontaneous biological processes can be responsible for damaging the macromolecule that makes up chromosomes—the DNA.

After being subjected to certain types of insult, the usual double-helix structure of DNA can be changed, and defects or distortions, called “lesions,” can result. We now know that DNA encodes information to enable its own repair—in less technical terms, DNA has the remarkable ability to fix itself by the action of repair proteins. For more than a decade, laboratory researchers have been attempting to determine what genes are involved in DNA repair, how many genes are required for each type of repair mechanism, and the structure and function of proteins associated with DNA repair genes.

The idea that DNA could encode its own repair was demonstrated in bacteria almost three decades ago. However, unraveling the detailed molecular mechanisms by which DNA damage can be repaired in mammals, especially humans, has been a daunting task requiring an explicit scientific strategy. As discussed in the first two sections of this report, our research strategy at LLNL makes use of both nonhuman and human starting materials. Human cells alone can be difficult to manipulate and clone (copy asexually), and the human repair pathways we seek to understand are biochemically complex. We are learning that human and rodent cells are surprisingly similar in terms of their DNA repair processes. Using rodent cells and rodent-human hybrid cells gives us ways to determine how many genes are involved in a repair process and to identify the chromosomes on which the genes reside.

Several different types of DNA damage-containment systems are known today. Some are widely studied, whereas the details for others remain mysterious. Research at LLNL focuses on the most important of these repair systems in human cells—one that corrects DNA damage produced by chemicals and radiation, both ionizing and ultraviolet. The third section of this report describes the present state of our knowledge about this multistep and exquisitely precise process, called excision repair.

The redundancy of information in the duplex structure of DNA inherently lends itself to repair. In essence, a certain length of the DNA backbone on both sides of the damaged site is cut, a segment with the damage is removed, and

Photomicrograph of chromosomes from a mutant line of the Chinese hamster. Each chromosome has been duplicated as two closely aligned sister chromatids in preparation for cell division. One chromatid is stained red and the other yellow (see arrow), but the chromatids of most chromosomes have a checkered appearance. This pattern reflects exchanges or crossovers between the two DNA molecules from a process called sister chromatid exchange. These exchanges arise from various types of insult, including radiation and some chemicals, and are greatly enhanced in certain repair-deficient cells, such as those in individuals with Bloom's syndrome, a cancer-prone disorder. Such exchanges may result in mutations, that is, heritable genetic changes.

new material is synthesized in a predetermined direction to produce a repair patch. The most recent evidence we have suggests that 30 to 50 different proteins play a role in this damage-recognition, cutting, and mending process. We believe that a defect in any one of these proteins may reduce or eliminate the repair activity.

The pace of work on DNA-repair mechanisms is accelerating at various research centers and universities throughout the world. Even though much of this research involves well-known genetic techniques, no two laboratories conduct their investigations in exactly the same way. At LLNL, we continually adopt new approaches, and, in that sense, virtually all of our work pushes the limits of known procedures. In this project and other biological work at LLNL, we serve as a resource to the broader scientific community by providing cell lines and DNA probes and by developing new technologies with which to study biological systems. The fourth section explains how our overall strategy combines classical genetic procedures with new gene-transfer and recombinant-DNA techniques.

What we have learned to date about DNA-repair processes is the subject of the fifth section of this report. In 1988, we published our results on the isolation and biological characterization of the second human DNA repair gene to be cloned anywhere in the world. This gene is now known to be involved in one of the cancer-prone DNA repair disorders in humans called xeroderma pigmentosum. More recently, we reported the isolation of yet another excision-repair gene. One Laboratory researcher has the distinction of being the first person to have cloned *two* human DNA repair genes. In addition to many other discoveries, LLNL researchers were the first to assign repair genes to specific human chromosomes. To date, we have assigned three different human repair genes to chromosome 19 and another to chromosome 16.

The DNA that makes up chromosomes and genes has been described as the genetic blueprint for life. The significance of our work is that DNA repair is fundamental to maintaining life and genetic diversity among species. Today, we are learning that unrepaired or improperly repaired DNA damage, if it accumulates over time, is likely to trigger the onset of cancer and may underlie the biological process of aging. The next section describes some relatively rare inherited diseases in humans that are associated with defects in repairing damaged DNA. Such individuals suffer striking developmental abnormalities and neurological defects, and many are at greatly increased risk of developing cancer and dying at an early age. In terms of specific clinical applications, an important goal in the field is to isolate the genes that could correct the molecular defects in the various human repair syndromes. Whether therapeutic means can be devised to actually correct some of the repair defects and clinical abnormalities in patients who already have DNA repair deficiencies remains an open question.

Although many different repair genes have been identified in human cells, we are still at the early stages of achieving a formidable long-term goal: isolating all the genes involved in repairing human DNA and classifying them into specific repair pathways. Our recent success in cloning repair genes is opening many new doors for future research at LLNL, which is the subject of the final section of this report. A gene is a stretch of DNA responsible for

synthesizing a specific protein. However, characterizing those proteins that play a role in DNA repair is a considerable challenge. Having a supply of cloned repair genes gives us a means to isolate for the first time relatively large quantities of protein products specified by the genes. Sufficient amounts of the repair proteins will allow us to assess what they do and how some proteins might interact with others during repair. Moreover, knowing the amino acid sequence of a protein at the molecular level can give us clues about its interaction with DNA. With respect to the broader issue of human exposure to mutagens in the environment, our emerging techniques for developing mouse models that are deficient in their ability to repair DNA should help clarify some of the risk factors and molecular mechanisms by which mutations arise.

Molecular biologists are experiencing an explosion of new information as the pace of their advances gathers momentum. The language associated with their research is expanding with each new strategy, technology, and discovery. Terms like "transgenics," "intron," "exon," and "nucleotide excision repair" and the names of specific repair genes will be new to the nonexpert, but they have become the familiar language of genetics and molecular biology. In this report, technical terms are defined at their first appearance, and some of the most common techniques and strategies are illustrated. For quick reference, a glossary of frequently used terms is provided on pp. 29-33. As this dedicated issue of *Energy and Technology Review* demonstrates, enriching our vocabulary with a few new words can have a high payoff, namely gaining insight into some of the ways by which the body preserves its unique genetic blueprint.

Unlocking the Mysteries of DNA Repair

DNA, the genetic blueprint, holds the code for repairing itself following many types of damage. We are studying the interplay of DNA-damaging agents, repair genes, and their protein products in an attempt to decipher the biochemical pathways that control this repair. Our long-term goals are to link repair proteins to human genetic disorders involving defects in DNA repair and to develop lines of repair-deficient mice that will serve as models for the human DNA-repair disorders.

SINCE 1963, LLNL researchers in the biological sciences have been working to understand how ionizing radiation produces adverse health effects, such as mutations or cancers. In the early days, our principal charge was to study the dose to man and the potential biological effects of radiation from nuclear accidents or explosions. As our work evolved in response to changing national needs, we expanded

our focus to include the possible damaging effects of various pollutants arising from fossil fuel technology.

Basic to the study of genetic injury is knowledge of the organization and content of deoxyribonucleic acid (DNA), the double-stranded macromolecule that carries all our genetic information. We know that damage to DNA can arise from exposure to both ionizing and

ultraviolet (UV) radiation, from many chemicals found in the environment or our food, or from spontaneous errors in cell metabolism. For more than a decade, we have been studying how damaged DNA is repaired.

To understand what we mean by DNA repair and how repair processes were first identified, we need a basic familiarity with the structure of this fundamental molecule. As shown in

The first-ever image of the helical coils of DNA, taken in 1989 by researchers from LLNL and the Lawrence Berkeley Laboratory using the scanning tunneling microscope that was jointly developed by the two groups.

Figure 1. Structure of chromosomes, DNA, and nucleotides. (a) Inside the nucleus of every human body (somatic) cell, the genetic information is packaged into 46 chromosomes (23 pairs) as a compressed and entwined packet of deoxyribonucleic acid (DNA). (b) One chromosome in each pair comes from each parent. (c) DNA is a double-stranded helix (a duplex) of repeating nucleotide units. Each chromosome contains one very long, highly coiled DNA molecule. (d) Each nucleotide sub-unit consists of a sugar (deoxyribose), a phosphate group, and a nitrogenous base. (e) The four chemical bases—the “rungs” in the DNA ladder shown here in color—are adenine (A), thymine (T), guanine (G), and cytosine (C). These bases form pairs in a fixed manner. A base pair refers to the unvarying partnership of A with T or of C with G in a DNA double helix. The linear sequence of bases along the DNA backbone is, however, slightly variable from one person to another and forms the genetic code. A gene is a structural and functional unit of double-stranded DNA that specifies how to build a protein using up to 20 different amino acids encoded by the linear sequence of DNA bases. Human genes vary in size from fewer than 1000 base pairs to more than 2 million, and they reside along DNA much like beads on an extremely long string.

Figure 1, the human genetic blueprint, DNA, consists of about 100,000 structural and functional units—that is, specific sections of chromosomal DNA—called genes. Genes contain all the information needed to dictate the development of the human embryo and its growth into an adult. Genes contain the biochemical code that specifies how to build proteins out of amino acids; as such, they are responsible for determining all cell functions and thus the body's entire structure.

As the bottom part of Figure 1 shows, DNA consists of nucleotide subunits, each consisting of a sugar (deoxyribose), a phosphate group, and a nitrogenous base. Each sugar-phosphate group is chemically bonded to the adjacent group to form the ribbon-like spine or "backbone" of DNA, which does not vary in structure along the length of the macromolecule. The four chemical bases—or "rungs" in the DNA ladder shown in color—are adenine (A), thymine (T), guanine (G), and cytosine (C). These four bases form pairs in a fixed manner: one base pair refers to the unvarying partnership of A with T or of C with G in the DNA double helix. The linear sequence of bases along the DNA backbone is, however, variable almost without limit across different species. It is the precise sequence of bases along the backbone that forms the genetic code of each individual and specifies how to build proteins.

During the course of evolution, virtually all living cells—from bacteria and yeast to man—have acquired remarkable biochemical machinery for preserving their genetic material. If DNA repair does not take place or if it is done incorrectly, then mutations can occur. Sometimes, mutations occur in genes responsible for DNA repair. If these mutations occur in the sperm and egg (germ) cells, they are passed on to the next generation.

Occasionally, a child will inherit "bad" copies of a repair gene from both parents. The consequences for the affected individual are profound, as is vividly demonstrated by the cancer-prone nature of several hereditary diseases in humans caused by repair deficiencies. These relatively rare diseases produce striking clinical features, including neurological defects and gross developmental abnormalities. Such clinical observations, together with the isolation of repair-deficient microorganisms (for example, bacteria) that have been known for many years, gave us some of the first clues that DNA actively undergoes repair. When DNA repair is disrupted, the consequences are dramatic and often fatal for the cell or organism.

In 1978, we began our first hunts for mutant mammalian cell lines that are deficient in their ability to repair damaged DNA. We reported our successful results in 1980.¹ About five years later, we used the first of several mutants to identify the location of a human gene that corrects the repair deficiency on one of the 24 different human chromosomes. Many of the mutant cell lines were isolated at UC Berkeley and characterized at LLNL in collaborative studies.^{2,3}

Today, we are working to isolate the proteins that several repair genes produce and are studying what those proteins actually do at the molecular level to repair DNA. We are beginning to link faulty DNA repair proteins to specific human diseases. On the horizon, we will also attempt to model human repair disorders by producing special lines of mice that are genetically deficient in the DNA repair process.

Studying the intricate network of multistep DNA repair mechanisms has become a broad area of research involving highly complex molecular biology and chemistry. Only a decade

ago, progress in the field was slow because successful genetic approaches were not available. Now, intense competition in publication is taking place among many international research groups as each new repair gene is sought and discovered. At LLNL, we are dedicating significant resources and have formed an expanding team to unlock the mysteries of DNA repair.

LLNL serves as a resource to the scientific community as we continue to develop many of the new techniques needed to eventually identify all of the human genes involved in DNA repair. Because we are only at the early stages of achieving this formidable long-term goal, we still do not know exactly how many genes may be involved. As each gene is isolated, we then classify it into the pathways—the molecular processes—by which defective DNA is repaired.

In this article, we discuss some of the most common repair processes, but we focus attention on one type called excision repair. Today, we know that this extensively studied pathway is nature's way of mending a remarkably broad range of distortions in the regular structure of DNA, including damage caused by exposure to the UV component of sunlight.

The significance of our work is that DNA repair is fundamental to maintaining life and genetic diversity among species. The accumulation of unrepaired or improperly repaired DNA damage can trigger the onset of cancer and may underlie the biological process of aging. In terms of more specific clinical applications, a widely sought goal in the field is to isolate the genes that will correct the repair deficiency in cells from various human repair syndromes. At LLNL, we are making exciting discoveries that will help us to unravel the biochemical details of how the body preserves its unique genetic blueprint.

Why We Use Rodent Cells

Rodent and human cells are remarkably similar in terms of their DNA repair processes. We are using several rodent cell lines to isolate repair-deficient mutants.

MECHANISMS of DNA repair were discovered in bacteria more than 25 years ago. However, isolating the genes and proteins involved in DNA repair in mammals, especially humans, requires a scientific strategy. An explicit strategy that makes use of both human and nonhuman starting materials, namely repair-deficient mutant cells, is essential for several reasons. For one, it is difficult to do genetic studies with human cells alone because they are often difficult to manipulate and clone efficiently. For another, the human repair pathways we want to understand are biochemically complex, and many repair proteins are likely to be present only in low abundance. We cannot simply grind up human cells and purify the proteins because those involved in DNA repair would be present at very low levels, and we would not know precisely what to look for in the inordinately complex mixture.

Thus, the initial step of our strategy involves the use of classical procedures to create and isolate mutant cells in culture. The original cells come from rodents, usually

hamsters or mice. (The box on p. 6 provides a background discussion of mutants and some of the DNA-damaging agents, called mutagens, that we use to cause genetic changes associated with repair.) The value of using rodent cells as a model for understanding DNA repair in human cells has become rather obvious today, but this was not always the case. Indeed, earlier researchers often regarded rodent cell lines as a poor choice for attempting to isolate mutants. However, as our grasp of DNA repair has developed, we have gained insights into how remarkably similar rodent and human cells are in terms of their repair processes.

During the last 15 years, several rodent cell lines, especially cells from a Chinese hamster ovary (CHO), have been used to isolate repair-deficient mutants. In fact, despite intense interest in human DNA repair genes, very few repair mutants have been isolated in human cell lines. The CHO cells offer advantages because they are fast growing in culture and because some mutants arise from changes in hamster chromosomes that have no partners. (In the context of our work, the term "repair-deficient mutant" refers to a cell or an

organism in which a biochemical deficiency in repair is present because of mutations.)

The body (or somatic) cells of humans and most other mammals contain pairs of homologous chromosomes—one pair of number 1 chromosomes, one pair of number 2 chromosomes, and so forth, except for the sex chromosomes. As shown in Figure 2a, homologous chromosomes are designated by geneticists as $+/+$ to indicate that each member of the pair has a normal gene at a particular locus (at the same genetic location along the chromosome). This duplication of the genes greatly hampers the recovery of mutants. Figure 2b shows that each copy of a specific gene can have alternate forms, called alleles. If one chromosome has a normal allele, and its partner has a mutated allele, geneticists designate the gene status as $+/-$. An individual with identical alleles is called homozygous; one with different alleles is heterozygous.

A third condition is also possible. As geneticists put it, the entire genetic material (or genome) of certain rodent cell lines contains regions of hemizyosity. As shown in Figure 2c, hemizyosity simply means that a

Using cells from rodents, like this hamster, we can create and isolate DNA-repair-deficient mutants and use them to investigate the complex mechanisms of DNA repair. As our understanding of DNA repair has developed, we have gained insights into how remarkably similar rodents and humans are in terms of their repair processes.

repair gene may have no counterpart (no allele) because the other chromosome is partly or completely absent. Geneticists designate this gene status as $+/0$, where the 0 indicates the absence of a second allele. When this is the case, only one copy of the gene has to be inactivated by some DNA-damaging agent to obtain cells ($-/0$) that are deficient in DNA repair. The feature of hemizyosity greatly simplifies our ability to recognize and

recover by conventional screening methods mutant cells for certain genes. (Even for hemizygous genes, using manual methods to recover mutants imposes rather severe experimental limitations because the frequency of detection is very low, about one cell in a thousand.)

The first demonstrations that researchers could greatly elevate mutation frequencies in cultured cells by adding potent mutagens were

made using CHO and V79 Chinese hamster cell lines back in the late 1960s. So far, mutants obtained in the CHO line are the only rodent cells derived in culture to be successfully used for isolating DNA repair genes. Furthermore, many of the human repair genes isolated to date have come from the use of CHO mutants as the vehicle for cloning (purifying unlimited numbers of identical copies).

(a) Homozygosity ($+/+$)

(b) Heterozygosity ($+/-$)

(c) Hemizyosity ($+/0$)

Figure 2. Homozygosity, heterozygosity, and hemizyosity.

(a) Homologous chromosomes are those that carry the same genetic loci. ("Locus" is a term geneticists use to designate the precise location of a DNA sequence on a chromosome, usually referring to a gene.) Each copy of a specific gene may have alternate forms called alleles. When two alleles are identical (normal or mutant), the locus is designated as homozygous. (b) A chromosome or individual with different alleles at a particular locus is called heterozygous for that trait. (c) The genetic material of certain rodent cell lines contains regions of hemizyosity. This means that a rodent gene, for example one that is responsible for a DNA repair process, may have no counterpart (no allele) on the other paired chromosome, if present at all. Because only one allele need be inactivated to obtain a repair-deficient mutant, the feature of hemizyosity is one of the reasons rodent cells are valuable in DNA repair research.

Mutations, Mutagens, and Mutants

All organisms are constantly subjected to events that can cause structural alterations to their genetic material. Some of these events are spontaneous within an organism, and others are environmentally induced. Not all alterations or damage to DNA are harmful to a cell, an organism, or a species. Indeed, genetic diversity can be highly beneficial from an evolutionary point of view because it provides a pool of different

traits from which to select. As Charles Darwin showed more than 100 years ago, those traits that are adaptive to an organism can help it to survive.

Some internal or external events trigger chromosome changes called mutations. First the DNA is damaged. If this damage is not repaired, a mutation can result when the DNA replicates. A *mutation* involves breakage or rearrangement of chromosomes, or a biochemical

change in the order or sequence of the four molecular units (bases) that make up genes. A *mutagen* is any agent that tends to increase the frequency or extent of mutation by interacting with the DNA.

We have long known that agents like toxic chemicals, a wide range of drugs, and radiation can cause some of the biological changes associated with mutations. Ionizing radiation (x rays and gamma rays) and some drugs used in chemotherapy are now known to damage the nitrogenous bases that form the informational part of the DNA molecule. For example, damage may involve the addition of an oxygen molecule or a methyl group (CH_3) to a DNA base. Even though this type of damage may seem relatively "minor" compared to larger scale chromosomal changes, damage to even a single base can lead to base mispairing and later destroy the ability of a gene to properly code for its protein. As shown in the figure, the class of damage that involves one or a few DNA bases but does not distort the DNA helix is often called "x-ray-like."

Sunlight is both a life-giving and a life-threatening agent. Part of the spectrum of sunlight overlaps the absorption spectrum of proteins and nucleic acids, and direct exposure to this ultraviolet (UV) light can produce damaging photoproducts in DNA. The major class of such photoproducts is the cyclobutane pyrimidine dimers (molecules linking adjacent bases). These photoproducts from UV light can cause large structural defects—bulky lesions—that distort the double-stranded DNA structure and block the ability of DNA to replicate or cause replication errors. For this reason, the class of damage involving bulky lesions is often called "UV-like." Many mutagens (e.g., heterocyclic amines) from cooked meat in our diet are bulky molecules that also join to DNA bases and cause mutations.

Mutagens can cause many other types of damage on even larger scales. For example, ionizing radiation, such as that associated with x rays or the Chernobyl accident, produces free radicals (namely $\text{OH}^\bullet$) when it interacts with the water in our body. Free radicals are highly reactive with the genetic material. The result can be breaks in one or sometimes both DNA strands.

A portion of a chromosome may be lost from a cell, an event that can push a cell toward the cancerous state. On even larger scales, whole chromosomes can be lost when errors occur during cell division.

We use a variety of known mutagens in our work on DNA repair. UV light from a germicidal lamp, predominantly at a wavelength of 254 nm, has been a widely used agent for studying repair since it efficiently induces cyclobutane pyrimidine dimers. *Cis*-platinum is a highly effective chemotherapy agent for treating testicular malignancies, but it is also a known mutagen. (The prefix "*cis*" means that organic molecules are bonded on the same side of the platinum atom.) Another DNA-damaging agent is the drug mitomycin C (MMC). The repair of damage caused by both *cis*-platinum and MMC seems to involve not only the genes that participate in UV-like damage repair but other genes as well. Using different mutagenic agents gives us clues about what aspect of DNA repair might be involved following damage to the genetic material. (Several different repair pathways are known, as is discussed in the box on p. 13, and others may be discovered.)

The term *mutant* can refer to a specific gene, a chromosome, a cell, a line of cells that reproduce in culture, or even an organism carrying a given mutation. However, when we say that our DNA repair work begins with the hunt for mutants, we are searching for repair-deficient cells that are derived from an established cell line originally taken from a rodent and placed in culture dishes. Thus, the objective of our mutant searches is to isolate cloned lines of cells in which we have deliberately induced mutations by exposure to some known DNA-damaging agent. In our studies, the mutant cells we seek are said to be mutagen-sensitive or hypersensitive because they can no longer repair their DNA and thereby reproduce in culture after exposure to one or more doses of the same or another DNA-damaging agent. Upon such exposure, they are killed. The UV-sensitive hamster repair mutants we have studied in detail appear to arise mainly from structural gene mutations and not, say, from gene inactivation (a process by which a nonmutated gene becomes metabolically silent and makes no protein).

Why We Use Hybrid Cells

By using rodent-rodent and rodent-human hybrid cells, we can find out how many genes are involved in a repair process and can identify the human chromosomes on which repair genes reside.

ONCE we obtain a series of rodent repair mutants that are altered in various genes, we need to know how many genes are involved in the repair process that is connected with a given expressed defect (phenotype), such as high sensitivity to cell killing by UV light. In genetic terms, we need to classify the mutants into different genetic groups, called complementation groups. This step of our strategy uses a rather efficient and simple test of cell killing versus cell survival to identify the complementation groups and thus the number of genes in the repair process.

Figure 3 shows how we can form hybrid cells in a petri dish by placing two, independently derived, mutant cell lines in culture and allowing some of them to fuse together (or hybridize). About 1 to 10% of the cells hybridize so that they contain complete sets of chromosomes (genomes) from both parent cells. The test for cell survival begins by exposing the entire culture to a DNA-damaging agent.

Two mutant lines are defined as complementing when the repair-deficient phenotype (say, UV sensitivity in each parent line) is converted to a repair-proficient phenotype (survival after exposure to UV light) in a hybrid cell containing both genomes. In other words, the defective repair protein of each parent cell is compensated in the hybrid by the contribution from the other parent. Failure to show complementation (that is, continued cell killing from failure to correct the defect) usually means that the two mutations affect the same protein and are in the same place (at the same genetic locus) in the DNA. The occurrence of complementation (hybrid cell survival, growth, and colony formation) suggests that different loci, and thus different genes, are involved.

We developed this complementation test of cell survival at LLNL in the early 1980s to identify five different groups of rodent (CHO) cell lines that showed high sensitivity to UV light. This early evidence for five different UV-sensitive groups suggested that at least five different

genes—perhaps more—were centrally involved in DNA repair processes after UV damage. (Additional complementation groups in CHO and other rodent cells have been more recently isolated, as is discussed later.)

After we have identified a rodent repair gene by genetic complementation, we test whether that gene is functionally equivalent to a human repair gene. Lab researchers developed one of the first ways to determine such equivalence by, once again, making hybrid cells. This time, however, we made the hybrid cells by mixing mutant rodent cells (usually hamster) with normal human cells, as shown in Figure 4.

We obtain human white blood cells (lymphocytes) and add them to rodent mutant cells in petri dishes under conditions that allow them to fuse together. The hybrid cells, which contain the chromosomes from both "parents" (hamster and human), are then grown in the presence of a DNA-damaging agent, such as repeated irradiation with UV light. When damage-resistant hybrid cells

Christine A. Weber, one of the researchers in the LLNL DNA repair group.

grow and form colonies (clones), many of the human chromosomes are lost. By the time the hybrid clones have multiplied to mass culture (about 10 million cells), only one or a few human chromosomes are left in each cell. We examine cells from many hybrid colonies (usually 20 to 30) to obtain a clear picture of their chromosomal

makeup. If only one human chromosome is consistently present in the damage-resistant rodent-human hybrid, then we know the DNA-repair gene is located on that human chromosome. Thus, we can assign (or map) the repair gene to that particular human chromosome. We base our chromosome assignments on >90%

concordance between the repair trait and the chromosomal marker. (A "marker" is any identifiable place on a chromosome, such as the locus of a gene, that can be assigned, or mapped, to the chromosome.) To our surprise, the first three human repair genes to be mapped were assigned to a single, small human chromosome, number 19.

Figure 3. Use of hybrid cells and the complementation test to determine how many genes may be involved in DNA repair. Rodent cells are exposed to DNA-damaging agents, called mutagens, that induce mutations causing repair deficiency. When pairs of different mutant cells are mixed together in equal numbers on a petri dish in the presence of certain chemicals, the two different types of cells can be made to physically fuse together (or hybridize), as shown at the top of this illustration. On fusing, the chromosomes from the two parents combine into one nucleus, and the hybrid retains complete sets of chromosomes from both parent cells. After exposure to a mutagen to which both parent cell types are sensitive, only complementing hybrids survive and multiply to form colonies (bottom right). Such survival means that defects in different genes cause the sensitivities. Hybrids that fail to complement one another are killed (bottom left), usually implying that the two mutations are in the same genetic locus, that is, the same location on the chromosome.

Figure 4. LLNL-developed technique to test whether a rodent repair gene is functionally equivalent to a human repair gene. This technique also allows us to assign human DNA repair genes to chromosomes. Human white blood cells (lymphocytes) are added to rodent mutant cells in petri dishes under conditions that allow them to fuse together. The hybrid cells, initially containing chromosomes from both "parents," are grown in the presence of a DNA-damaging agent, such as repeated UV irradiation. When damage-resistant hybrid cells grow and form colonies (clones), human chromosomes are preferentially lost. By the time the hybrid clones multiply to mass culture (~10 million cells), only one or a few human chromosomes are left in each cell. By examining a sufficient number of independent hybrids (20 to 30), we obtain a clear picture of their chromosomal makeup. If only one particular human chromosome is consistently present among a series of damage-resistant rodent-human hybrids, we can assign (or map) the DNA repair gene to that human chromosome.

The Excision Repair Process

In excision repair, damaged or inappropriate pieces of DNA are identified and removed and then replaced with correct pieces.

OUR research focuses on repair processes that correct DNA damage produced by chemicals and radiation, both ionizing and ultraviolet. Although several other types of DNA repair are known (see the box on p. 13), the most important type of repair in human cells is called excision repair. In this process, damaged or inappropriate pieces of DNA are removed and replaced with correct ones. As shown in Figure 5, excision repair can be divided into two types, *nucleotide* or *base* excision repair, depending on what part of the DNA molecule is cut and mended.

Nucleotide excision repair (NER) is a major, ubiquitous, and multistep pathway. Recall from Figure 1 that in DNA, a nucleotide consists of a nitrogenous base, a sugar (deoxyribose), and a phosphate backbone. Although we know that the complex NER process is initiated by damage recognition, we have little idea how a structural distortion in DNA is detected. Nevertheless, as shown in Figure 5a, the DNA backbone is cut on both sides of the

site of the recognized nucleotide distortion (lesion) by a process called incision. This recognition and incision step alone probably requires about 10 or more different proteins. Next, a segment of DNA containing the damage—a string of 29 nucleotides—is removed (excised). The undamaged complementary strand of DNA is then used as a template to synthesize new material in the resulting gap (polymerization). The enzyme called DNA polymerase is required to carry out this repair synthesis (that is, to fill in the gap on a single strand of DNA). Synthesis occurs in a manner similar to that in which DNA is replicated during normal cell growth and division. Finally, in the ligation step, the new repair patch is tied into the contiguous parental DNA strand by an enzyme called DNA ligase, closing the single-stranded nick left over after detachment of DNA polymerase.

The NER pathway recognizes a wide spectrum of damage. It is responsible for removing DNA damage caused by a multitude of chemicals, including many

environmental pollutants that produce bulky adducts (chemical addition products) in DNA and distort the backbone. Historically, UV radiation from germicidal lamps has been a convenient and widely used agent for studying repair by the NER pathway. Today, we know that the NER mechanism replaces a patch of DNA containing the damage by excising a run of 29 nucleotides, give or take one nucleotide. Our best evidence suggests that some 10 to 20 different proteins are involved in this damage-recognition and incision/excision process. A defect in any one of these proteins may reduce or eliminate the repair activity.

The second excision-repair pathway is called base excision repair. As shown in Figure 5b, this mode of repair removes damaged or inappropriate individual bases of DNA and replaces them with correct ones. Along the length of DNA, the four nucleotide bases (see Figure 1) are adenine, thymine, cytosine, and guanine. Although subtle damage or spontaneous change (such as oxidation) in the structure of one DNA

base may not significantly distort the DNA structure, it may still change the coding properties of that base. The altered base may code for a different amino acid, or the coding property may be disrupted altogether. The alterations

associated with base excision repair are often called "x-ray-like" because they are induced by ionizing radiation, but they can be caused by many environmental mutagens and chemotherapy drugs as well.

Like NER, base excision repair also involves recognition, incision, excision, polymerization, and ligation. Whereas the first several steps for the two classes of damage repair are done by different enzymes,

(a) Nucleotide excision repair

(b) Base excision repair

Figure 5. Excision repair processes. Excision repair is believed to be the most versatile DNA-repair pathway in terms of the kinds of DNA damage it can handle. Two general classes of excision repair have been defined by molecular biologists. (a) Nucleotide excision repair (NER) removes large structural defects that distort the double-stranded DNA helix and block DNA replication. NER results in relatively large repair patches of about 29 nucleotides. (b) Base excision repair removes specific modified DNA bases having more subtle damage, resulting in a small repair patch of one or a few nucleotides. After a stretch of defective nucleotides, or one or more bases, is cut out (excised), the undamaged complementary strand of DNA is used as a template to synthesize new material in the resultant gap [the polymerization step in (a) and (b)]. Finally the repair patch is tied to the contiguous parental DNA strand (ligation).

Other Types of DNA Repair

Compared to our knowledge of the excision-repair pathway, other DNA repair mechanisms are less clearly understood. One way to describe other damage-containment systems is in terms of the type of injury or "lesion" to the DNA macromolecule that is repaired. Damage to DNA is defined as any change that introduces a deviation from the usual double-helix structure. Several types of repair systems have been described in various organisms, but the molecular mechanisms and repair genes associated with them have not been fully identified.

Repair of DNA crosslinks. The presence of interstrand crosslinks (those between bases on opposite strands) can inhibit DNA from replicating by preventing the strands of DNA from separating. We know that mutagens like MMC, *cis*-platinum, and nitrogen mustard produce interstrand DNA crosslinks and are extremely toxic to cells. The repair of DNA crosslinks arising from MMC or *cis*-platinum seems to involve not only the genes that participate in UV-damage repair but other genes as well. The molecular mechanisms for this kind of repair are still mysterious in mammalian cells.

Photoreactivation. This well-studied mechanism, sometimes called enzymatic photoreversal, involves the repair of one type of intrastrand crosslink. It directly reverses the links formed between two adjacent thymine bases (called thymine dimers) that are the most common photoproduct of UV light. As shown in the illustration, photoreactivation reverses the links in thymine dimers by the combined action of a light-dependent enzyme (photolyase) and visible light. We know that photoreactivation in the bacterium *E. coli* depends on the protein product of a single gene.

Mismatch repair. This type of repair acts on incorrectly paired bases often arising from unfaithful DNA replication. Because the mispaired partnership of bases was created when one of them was incorrectly inserted during replication, the repair system has no way of knowing which is the correct base (called the wild type) and which is the mutant. In many cases, the direction of the repair process is not random, but is influenced by the local sequence around the mutated base. In mismatch repair, all the bases are normal in structure but mismatched, whereas in base excision repair, a base is modified or subtly damaged.

Recombinational repair. The breakage and subsequent rejoining of DNA strands in a highly programmed way is a normal process in some cells. This process, called recombination, occurs during the formation of egg and sperm cells (during meiosis). Strand breaks are also a major form of damage arising from ionizing radiation. In yeast and bacteria, double strand breaks are known to be repaired by a recombinational pathway. Recombinational repair involves a retrieval system, which requires the target DNA being repaired to use a replica that acts as a sort of scaffold. Recombination repair proteins work by copying information from one duplex of DNA to repair another duplex. For this type of repair to occur, two homologous chromosomes must be present in a cell.

the final steps of polymerization and ligation are common to both pathways and may share the same enzymes.

Base excision repair is initiated by enzymes (glycosylases) that recognize specific modified DNA bases, such as uracil, 3-methyladenine, 7-methylguanine, or urea. These glycosylases cleave the damaged bases from the sugar to which they are attached without cutting the sugar-phosphate backbone of DNA, as shown near the top of Figure 5b. The removal of the abnormal base creates an abasic, sugar-containing site along the DNA phosphate backbone that is then

incised or cut by other enzymes (endonucleases). Ultimately, a small repair patch is made involving one or a few nucleotide bases.

By examining many different repair-deficient mutants and drawing on the many sophisticated research techniques now available to molecular biologists, we can begin to identify the details of each repair pathway. The integrated study of naturally occurring human mutations and those induced in culture provides us with a framework for tackling how these repair processes are interrelated with DNA replication and recombination and for eventually learning how repair deficiencies can cause human disease.

LLNL Strategy for Studying Human Repair Genes

Our strategy for studying human repair genes involves classical genetic procedures as well as new gene-transfer and recombinant-DNA techniques.

OUR strategy for studying human repair genes is summarized in Figure 6.⁴ These approaches have also been adopted by some other research groups. Although eight separate and sequential steps are shown in this figure, and advances for a particular gene generally follow this progression, our work is not always so neatly compartmentalized. We and researchers elsewhere are adopting varied experimental approaches and addressing many different aspects of DNA repair simultaneously.

The first two steps in Figure 6 involve classical genetic procedures for dissecting a complex biochemical process. As already described, these and some later steps consist of isolating mutagen-sensitive rodent cell mutants, classifying the different mutations into complementation groups that usually correspond to different genes, and using human-rodent hybrids to map the DNA repair genes to specific human chromosomes.

A variety of mutagenic agents are used to hunt for mutants (step 1). These DNA-damaging agents include UV radiation, x rays, and cancer-treatment drugs, such as *cis*-platinum and mitomycin C (MMC). When we

try to characterize a new mutant that has been isolated on the basis of its hypersensitivity to a particular agent, the spectrum and degree of sensitivity to other agents give us insights into what aspect of repair might be involved. The UV-sensitive CHO repair mutants we have studied appear to arise mainly from structural gene mutations rather than the "silencing" of a gene by some developmental "switch."

Rodent cell mutants exhibiting various degrees of sensitivity to UV light now fall into 11 different complementation groups (step 2). This finding suggests the involvement of at least 11 different genes in rodents and a high degree of genetic complexity underlying the excision repair pathway. About 10 complementation groups have been described for x-ray sensitivity, and these genes should be involved in base excision repair. Additional complementation groups will likely be found in the future. Identifying all the genes that play a role in mammalian DNA repair is a daunting task that has been pursued for the last 15 years in several laboratories.

In the past, our next step was to assign repair genes to individual human chromosomes (shown as step 4

in Figure 6). As discussed earlier, we have produced rodent-human hybrids to map a gene to a whole chromosome. To help identify exactly which human chromosome ultimately allows a given hybrid to survive low exposures to a DNA-damaging agent (that is, to find the complementing chromosome), Michael Siciliano, one of our collaborators at the University of Texas, Houston, has used various enzyme markers in combination with DNA "probes" that uniquely identify each type of human chromosome.

In the future, we plan to proceed with the cloning step (step 3) before mapping a particular gene to a chromosome (step 4). The reason for this shift in our strategy is that it is now relatively easy to map an already cloned gene to a chromosome or even a specific region along the chromosome—the process only takes a couple of days. Under this approach, the repair gene is tagged with a dye and then allowed to bind (hybridize) with target chromosomes on a microscope slide. The technique, called *in situ* hybridization, is effective because fusion of the DNA probe and target chromosome only occurs when the complementary base sequence is found. A refinement of

Lawrence H. Thompson, leader of the LLNL DNA repair group.

this approach involves using already mapped DNA probes (cloned segments of human DNA with known positions along the DNA strand) to locate the region of human repair genes with respect to other mapped genes. These DNA probes have become available to us through work being conducted at LLNL and elsewhere, especially through the Human Genome Project (see the April/May 1992 edition of *Energy and Technology Review*).

There are many methods and variations for cloning mammalian DNA repair genes. Some repair genes are too large or difficult to clone intact, so they were cloned as overlapping shorter fragments (often 20,000 to 40,000 bases long). Our

recent efforts to clone a specific repair gene took about one year. Figure 7 shows one of the ways we can clone a human DNA-repair gene after it has been introduced in a CHO cell's DNA and selected.

Step 3 in Figure 6, cloning, and the subsequent steps of our broad strategy make use of techniques involving gene transfer and recombinant DNA methodologies that have evolved during the last decade. In general, recombinant DNA technology makes use of molecular biology tools, such as enzymes that function as molecular scissors, to cut pieces of DNA at specific sites. With other enzymes (ligases) used for rejoining, we can combine the DNA from one organism with that of another to form a

"recombinant" molecule. As shown in Figure 7, certain DNA carriers, called cloning vectors, allow fragments of foreign (that is, human) DNA to be reproduced in large quantities in host cells (usually bacterial cells).

One of the cloning vectors that is useful in our work is the cosmid vector. A cosmid is an artificially made, ring-shaped DNA molecule. The molecule contains one particular gene of the lambda virus used for packaging DNA into viruses (the *COS* gene, thus the name cosmid), a gene for drug resistance, and an introduced DNA fragment we want to copy. Cosmids are inserted into the lambda virus protein coat for efficient infection into a bacterial host. As the bacterium divides and multiplies, the cosmid also replicates to produce unlimited, identical clones of the DNA we have inserted.

Another approach to gene cloning with long-term potential is to use complementary DNA (cDNA). The DNA making up a gene has specific protein-coding sequences, called exons, and much longer sequences, called introns, that interrupt the protein-coding exons. (See box on p. 18 for a discussion of interrupted genes and how proteins are made.) The cDNA is a type of DNA in which the introns, including highly repetitive sequences, are removed. The cDNA is essential in our studies because it contains little more than the genetic material necessary to make the protein product we want to isolate. If we put a whole repair gene into bacteria, these microorganisms would not distinguish between the exons (the direct template for a protein) and the introns (the intervening sections). The cDNA provides a "mini-gene" that is easy to handle and that can express protein in convenient organisms.

Thus, in addition to helping us map genes to chromosomes, purified copies of repair genes and cDNAs (the clones in step 3 of Figure 6) give

Figure 6. Strategy for studying human repair genes at LLNL. This strategy, which has also been adopted by other laboratories, uses the methods of somatic cell genetics to study DNA repair. The first several steps involve classical procedures of mutant isolation, complementation analysis, and gene mapping. Some of the subsequent steps use methods of recombinant DNA technology devised at the Lab. Different genes are at different stages of investigation in this sequence.

us a way to produce and isolate the repair proteins, which they specify how to build (step 5). Once we have purified a sufficient amount of a particular repair protein (we need milligram quantities), we can study its normal role in DNA repair (step 6) and begin to study possible links between that protein and human disease (step 7).

Mice that are deficient in DNA repair will provide us with animal models (step 8) for some of the rare human genetic disorders in which DNA repair is defective and associated with elevated cancer risk. These repair-deficient rodents, called transgenic mice, can be made by inducing a mutation in a specific mouse repair gene. Rodent analogs

of human DNA repair syndromes will be of great value because of the inherent limitations of doing studies with humans. Step 7 and step 8 represent work that began here two years ago. These two steps are being done in concert with other recent efforts, such as the isolation of repair proteins and cloning of new genes.

Figure 7. Cloning human repair genes. Cloning is a research tool to produce the quantities of purified and homogeneous DNA we need. Our cloning technique makes use of gene transfer and recombinant DNA methodologies. We start with a human repair gene embedded in hamster DNA. Restriction enzymes that function as molecular scissors cut pieces of DNA at specific sites to yield smaller fragments. With other enzymes used for rejoining, such as DNA ligase, we combine our DNA fragment with DNA of another organism to form a "recombinant" molecule. Certain DNA carriers, called cloning vectors, allow the fragments of foreign (i.e., human) DNA to be reproduced in large quantities in host (usually bacterial) cells. A cloning vector is inserted into a bacterial host, which then grows and multiplies repeatedly on a growth medium to yield unlimited, identical copies of the DNA we have inserted. These copies represent a clone.

From Genes to Proteins

A gene is a sequence of DNA responsible for synthesizing a given protein. The synthesis of proteins is a complex process that can be described somewhat like an assembly line. This assembly line for proteins depends on small factories, the ribosomes. Ribosomes are spherically shaped particles in cells that are the site of protein synthesis. Today, we know that all ribosomes in a cell are the same and that it is another species of nucleic acid, the messenger RNA (mRNA) that is the actual template used to make proteins.

The simple view that a continuous stretch of DNA corresponds directly with its protein product was revised in 1977 with the discovery of interrupted genes in mammals. We now know that the sequences of DNA making up a gene are divided into two categories. *Exons* are the regions used to produce a protein product. *Introns* are other, usually much larger, regions of DNA that do not code for a protein and must be

removed to give the mRNA, which consists of the series of protein-coding exons with control sequences at each end. Thus, a gene embedded in a chromosome is usually far longer than the actual sequence of nucleotides that codes for a protein.

The idea that each protein consists of a fixed series of amino acids dates from Sanger's work on the hormone insulin (a protein) in the 1950s. Twenty different amino acids are commonly found in proteins. Examples of common amino acids are lysine and proline. The individual amino acids of a protein are joined together in a string by peptide bonds to form a polypeptide chain. Both the structure and function of each protein are uniquely determined by the linear sequence of from a hundred or so to more than one thousand of the 20 amino-acid building blocks.

How can a sequence of nucleotides in DNA represent a sequence of amino acids and, eventually,

a protein? As shown in the middle portion of the illustration, the genetic code is read from a fixed starting position as a series of triplets of nucleotides in the DNA sequence. Each triplet sequence that represents one amino acid is called a *codon*. For example, the triplet of bases A-A-G codes for the amino acid lysine, the triplet C-C-C codes for proline, and so forth.

However, before a protein can be made, the coding information in double-stranded DNA of a gene must be converted into single-stranded RNA, as shown at the top of the illustration. In the first step, called *transcription*, a long RNA molecule is synthesized from one strand of the DNA. In a second step, this transcript is processed by enzymatic cutting and

splicing to remove all internal sequences that contain no protein-coding information. The result of these reactions is the mRNA. Finally, during *translation*, the mRNA molecule is used at the ribosomes to produce a specific protein. The ribosomes attach to the mRNA and provide the catalytic environment for the step-by-step attachment of amino acids as successive triplets of mRNA bases are read.

At each step of translation, the growing chain that will become the protein product becomes longer by one amino acid. Thus, through the stages of transcription, processing, and translation, the genetic information carried by the sequence of DNA is converted into amino acid sequences of proteins. This idea is a central concept of molecular biology.

Recent Findings on DNA Repair Genes

Since 1984, researchers at LLNL and elsewhere have identified, isolated, cloned, and characterized approximately 15 genes involved in the repair of human DNA.

THE isolation of the first human DNA repair gene, *ERCC1*, was reported in 1984 by researchers in Rotterdam. Even the name of this gene contains much information about its role in DNA repair. The acronym stands for excision repair cross-complementing rodent group 1 (by convention, gene names are italicized). This name means that the gene is involved in the NER pathway and that the DNA correction occurs across species, from human to hamster. That is, this human gene complements the excision-repair defect of rodent-cell mutants that are hypersensitive to UV radiation and belong to complementation group 1.

In 1988, LLNL researchers reported⁵ the isolation and biological characterization of the second human repair gene to be cloned, *ERCC2*. This human gene corrects the repair defect and UV sensitivity of rodent complementation group 2. We recently reported the isolation at LLNL of yet another excision repair gene, *ERCC4*, and we are now cloning

the cDNA. Indeed, LLNL's Kerry Brookman has the distinction of being the only person in the world to have cloned *two* human DNA repair genes (*ERCC4* and *XRCC1*, discussed below).

Table 1 lists six human genes (*ERCC1* through *ERCC6*) that are known to be involved in excision repair. In addition to these six genes, researchers in Japan and Texas have isolated and cloned two more genes involved in the NER pathway. These genes, *XPAC* and *XPCC*, have been linked to the human repair disorder xeroderma pigmentosum (XP), which is characterized by increased sensitivity to UV light (discussed later). The five other genes listed in Table 1 are involved in the repair of DNA damage induced by x rays. These x-ray repair cross complementing genes are known as *XRCC1*⁶ through *XRCC5*.

Table 1 shows the assignments of all these repair genes to human chromosomes. The first three repair genes mapped to any human chromosome were assigned by LLNL

researchers to chromosome 19. This type of clustering on a single chromosome is not merely a matter of chance. Rather, it can be understood in terms of a principle known as evolutionary conservation. Recall that our work involves hamster cells with hemizygous genes that have no second or backup copy. When we induce a mutation in a hemizygous gene, we see the mutants immediately and thus recover mutants *preferentially* for that gene. Evolutionary conservation among all mammals means that blocks of genes are maintained or conserved among species, in this case from hamster to human. Therefore, some DNA repair genes are linked together along a stretch of chromosomal DNA in humans in the same order as they are in hamsters. *ERCC1* and *ERCC2* are quite close together (no more than 250,000 bases apart) on chromosome 19. In a recent collaborative study with researchers at the University of Texas, we found that *ERCC4* is located on the short arm of chromosome 16.

Kerry W. Brookman, one of the researchers in the LLNL DNA repair group.

Among the many genes needed for excision repair in humans, structural homologs (that is, genes whose encoded proteins are similar in their sequence of amino acids) have been discovered not only in rodents but in other organisms, most notably the baker's yeast *Saccharomyces cerevisiae*. Table 2 lists six cloned human genes showing clear homology with the cloned genes from *S. cerevisiae* that are also required for nucleotide excision repair. For

example, the predicted protein product of the human repair gene *ERCC1* has significant homology with the predicted protein product of the yeast repair gene *RAD10*. In addition, *ERCC2* shows a striking 52% identity of its protein, in terms of deduced amino acid sequences, with that of the yeast *RAD3* protein.

Findings like these indicate that the human and yeast genes are descendants of the same ancestral gene and may well have analogous

functions. We now believe that both the *ERCC2* and *RAD3* proteins help the spiral DNA molecule to unwind during excision repair. Enzymes that unwind the duplex DNA molecule are known as helicases. There is also good evidence that both proteins have an essential role in cell viability as well as in DNA repair. They may be required for DNA replication or gene expression (in which RNA templates for protein are made, as described in the box on pp. 18-19).

Table 1. Human DNA repair genes and their assignments to chromosomes.

Name of human gene	Human chromosome
ERCC1	19
ERCC2	19
ERCC3	2
ERCC4	16
ERCC5	13
ERCC6	10
XPAC	9
XPCC	3
XRCC1	19
XRCC2	7
XRCC3	14
XRCC4	5
XRCC5	2

Table 2. Human excision-repair genes cloned using hamster or human mutants as the vehicle for cloning.

Gene name	Gene size, kilobases	Homologous gene in yeast	Confirmed involvement in human disease
ERCC1	15	RAD10	—
ERCC2	19	RAD3	Xeroderma pigmentosum, group D
ERCC3	35	RAD25	Xeroderma pigmentosum, group B
ERCC4	<40	RAD1?	Xeroderma pigmentosum, group F?
ERCC5	32	RAD2	Xeroderma pigmentosum, group G
ERCC6	~100	?	Cockayne's syndrome, group B
XPAC*	25	RAD14	Xeroderma pigmentosum, group A
XPCC*	?	RAD4	Xeroderma pigmentosum, group C
XRCC1	33	?	—

*Cloned using human mutants.

Human DNA Repair Syndromes

Several relatively rare diseases in humans are directly associated with defects in repairing damaged DNA. We are beginning to link alterations in repair genes to these hereditary diseases.

GIVEN that almost all organisms, from yeast to man, have developed an array of self-defense mechanisms that help to repair their genetic blueprint, what happens if one or more DNA repair processes are altered or seriously disrupted? If the disruption of DNA repair involves enzymes, which, in addition to their role in repair, are essential to other cell processes fundamental to sustaining life, then we would expect that a developing cell or embryo expressing profound defects would not survive. Some types of nonlethal DNA repair alterations would be expected to result in survival, but at the cost of abnormalities and deficiencies in the mature organism. These organisms and the deficiencies they express could give us further insights into how the pathways of DNA processing work.

It turns out that several relatively rare, inherited diseases in humans are directly associated with defects in repairing damaged DNA. The field of DNA repair is an example from biology in which "mutants" have been identified in humans, whereas almost none have been found in other animals. Human diseases involving repair defects are of considerable interest not only to geneticists and molecular biologists but also to the medical community because they

place the afflicted individuals at increased risk of developing cancer.

Xeroderma Pigmentosum

A direct correlation between unrepaired DNA and cancer susceptibility was first established in 1968 for one of these human diseases, xeroderma pigmentosum (XP).⁷ XP is a genetically recessive disease that strikes about one person out of 100,000 worldwide. Its victims have extreme skin sensitivity to the sun, a high incidence of skin tumors, eye defects, and, very often, neurological disorders. The incidence of skin cancer in XP individuals is about 2000 times that of the general population. XP involves defects in repairing DNA damage produced by UV radiation, and we now know that these biochemical defects reside in the NER pathway.

All cell types from one XP individual show the same repair deficiency. However, the situation is not so straightforward when different XP patients are studied. At present, genetic analysis has revealed at least seven different complementation groups in XP, each with its unique characteristics. These groups are designated XP group A through XP group G. Since each complementation group represents a different gene involved in the repair process, the genetic complexity of this disease is

considerable. Indirect evidence has suggested that, whereas the precise defect differs in its details for the different complementation groups, the overall defect in all seven groups is in the early stages of the NER process, before or during incision at damaged sites in DNA.

Cockayne's Syndrome

Cockayne's syndrome (CS) is another inherited disease associated with a deficiency in nucleotide excision repair, and it shares some clinical features with XP. (A few XP patients have also been diagnosed with CS.) CS individuals are acutely sensitive to the sun, but they also exhibit mental and physical retardation and other deformities as well. It is curious that the elevated risk of skin cancer seen in XP is absent in CS patients. Three complementation groups have been identified, but one of these, CS group C, is identical to XP group B. This overlap means that a mutation in the XP group B gene is causing both XP and CS.

Ataxia Telangiectasia

Ataxia telangiectasia (AT) is another inherited, genetically recessive disease that places individuals at increased risk for

Chromosome strands (chromatids) are being pulled apart as a cell prepares to divide.

developing cancer. This disease involves impairments of the immune system and nervous system as well as skin disorders. An individual with AT who does not succumb to infections will develop cancer before the age of 20. Like XP, the incidence of AT is about one per 100,000 live births. Whereas XP individuals are sensitive to UV exposure but not to x rays, the reverse holds true for AT; that is, AT victims are extremely sensitive to x radiation. For this reason, AT was originally thought to be an x-ray analog of XP. Although four complementation groups have been identified for AT, and we know the disease involves DNA-repair abnormalities, the precise nature of the DNA-repair deficiency has yet to be established.

Fanconi's Anemia

Fanconi's anemia (FA), which has also been associated with defective DNA repair, places individuals at higher risk of developing tumors,

especially an acute form of leukemia. Four complementation groups have been described in FA. Cultured cells from some FA individuals are highly sensitive to a class of mutagens, such as the drug MMC, that chemically link the two strands of DNA.

Relation of Mutants and Repair Genes to Human Syndromes

It is possible that all the inherited cancer-prone diseases discussed here may have their primary cause in DNA-repair deficiencies. However, XP represents our best example of a biochemical connection between unrepaired DNA lesions and the initiation of cancer (carcinogenesis) in humans. For this and other reasons, an important goal in the field of DNA repair research is to identify and clone all the genes that will correct the many complementation groups of XP and other human repair syndromes. Having access to the relevant genes associated with these disorders along with sufficient quantities of their

protein products gives us a direct approach to understanding the mechanisms involved in the syndromes.

The first repair gene to be cloned, *ERCC1*, did not correct defects in any of the XP groups. In recent work funded by the DOE, we attempted to determine whether a defect in the *ERCC2* gene was responsible for the repair deficiency in any of the XP groups. (Recall that our *ERCC2* gene was the second mammalian NER gene to be cloned.)

We introduced our *ERCC2* gene (in mini-gene form) into cells obtained from patients that represent the various XP groups. We then exposed the cells to UV light and incubated them for several days. Our reasoning was that if the *ERCC2* gene corrects the repair defect in a particular group, then those cells should have enhanced survival and growth relative to control cells. (The DNA we introduced in controls did not express *ERCC2*.) As shown in Figure 8, we found that *ERCC2* enhanced the survival of XP

Figure 8. Cytotoxicity assay demonstrating correction of the defect in XP group D cells. We introduced our *ERCC2* mini-gene into cells from patients that represent the various XP groups. These cells, along with controls, were then placed in growth chambers, exposed to UV light of varying intensity, and incubated for several days. The darker stain indicates where cells have grown into a solid layer. Introducing the *ERCC2* repair gene into XP group D cells (row 1) enhances cell survival after UV exposure relative to control XP group D cells (row 2). Introducing the *ERCC2* repair gene into XP group F cells (row 3) does not enhance cell survival after UV exposure relative to control XP group F cells (row 4).

group D cells but not that of other XP groups. Our quantitative measurements in other studies have showed that the *ERCC2* gene can enhance survival of XP-D cells to normal levels after UV exposure.

To date, we know of several other DNA repair genes that specifically correct defects in human repair syndromes. As shown in Table 2, *ERCC3*, for example, corrects the defect in XP group B, and *ERCC6* has been linked to CS group B in a collaborative study with the DNA-repair group in Rotterdam who used CHO mutant cells we characterized at LLNL. One gene involved in FA was also recently cloned. An important observation about many studies using repair-deficient human cells is that the transfer of the complementing DNA repair gene, or even a whole

chromosome carrying the gene, into these cells may not fully correct the defect. If a correction is less than 50%, uncertainties arise as to whether the proper gene has actually been found.

Despite several decades of research on human repair syndromes and our recent success in cloning several human repair genes, the biochemical bases of these complicated disorders remain to be clarified. Very recent evidence is suggesting that many repair proteins act in physical association as complexes and that such molecular ensembles are necessary for certain types of repair to take place. As other human repair genes are cloned, it will be important to test their proteins for the ability to correct mutant cell lines using both individual proteins and protein complexes.

Future Studies

Two exciting areas of future work are the use of cloned DNA repair genes to isolate repair proteins and the development of a new technology involving special lines of mice that are deficient in DNA repair, which will give us valuable animal models of the human repair syndromes.

OUR relatively recent access to cloned DNA repair genes is opening many new doors in terms of laboratory research and potential applications. Two exciting areas of work are the use of cloned repair genes to isolate repair proteins and the development of a new technology at LLNL involving transgenic mice that are deficient in their ability to repair DNA. Such animal models of the human repair syndromes give us a valuable tool for studying mutagenesis and carcinogenesis. The techniques for producing both the repair proteins and transgenic mice are now at hand.

Isolation of Repair Proteins

Many DNA repair proteins are virtually impossible to isolate and collect with conventional biochemical methods of purification from normal cells, such as chromatography. The biggest handicap is that we have no way to identify the function of a given protein and thus guide the purification process. However, cloned repair genes are now giving us a new way to produce and isolate

sufficient quantities of repair proteins in bacteria, yeast, and even mammalian cells.

Under a new, Lab-funded project, we plan to produce large quantities of the proteins specified by three repair genes (*XRCC1*, *ERCC2*, and *ERCC4*) in various systems. Our goal is to isolate milligram quantities of protein product for each gene for biochemical studies. To assess what the proteins actually do in the repair pathways, we will measure the activity of each highly purified protein in cell-free assays that detect particular steps in DNA repair. These measurements are often done using small, circular DNA molecules called plasmids (mini-chromosomes found in bacterial cells) that contain known types of DNA damage. We can use radioactive tracers to detect various biochemical reactions as the damaged plasmid is repaired.

Our protein work is under way with the *XRCC1* protein, which we have partially purified from mammalian cells. Our new procedure involves using a polyhistidine tag (six histidine residues attached to the protein) that allows us to identify and

obtain the *XRCC1* protein product without destroying its function. From about 4 liters of cell suspension culture, containing roughly the amount of tissue equivalent to a few grams of human liver, we expect to obtain about a milligram of protein product. Very recent findings with the *XRCC1* gene suggest that it is required at the late step in repair, when a ligase enzyme seals the break in the DNA backbone (refer to Figure 5). Biochemical studies have identified three distinct ligases in mammalian cells. We believe that the *XRCC1* protein is not a ligase, but rather a protein that is physically associated with DNA ligase III and is essential for its activity. We hope to understand the nature of this protein interaction.

We can also produce altered proteins containing specific mutations and analyze their functions in both cells and cell-free systems. The important functional sites in a protein can be identified in this way. Furthermore, at the molecular level, knowing the amino acid sequence of a protein can provide important clues about its interaction with DNA.

A special line of mice that is deficient in its ability to repair DNA is being developed to provide an animal model of human DNA repair syndromes and to give us a valuable tool for studying mutagenesis and carcinogenesis.

Development of Transgenic Mice

A transgenic mouse is one that is genetically altered by experimental manipulation. Two kinds of transgenics are proving to be invaluable scientifically. In the first approach, a foreign gene is introduced into the single-cell mouse embryo (at the pronuclear stage) by a method called pronuclear injection. The "trans" (foreign) gene is inserted into a chromosome of the embryo and becomes a permanent part of its genetic makeup as the embryo develops into a mouse. The second approach, which is more complex and of direct relevance to the field of DNA repair, seeks to alter the expression of a specified mouse gene. The goal of this manipulation can be to destroy the function of a particular gene, an alteration called a gene knockout. Alternatively, more subtle changes allow us to modify the properties of a gene without destroying the protein it encodes—a procedure called gene targeting. We have recently demonstrated

pronuclear injection and are well on our way to acquiring the technical capability and knowledge to perform a DNA-repair gene knockout.

Gene targeting is a gene-replacement process that requires the use of some recent but rather poorly understood approaches. We start with cells, derived from a mouse embryo, that will grow in culture and yet maintain their ability to participate in embryological development when introduced into another embryo. These cells are called embryonic stem (ES) cells. We can introduce foreign DNA into the ES cells and enrich for rare cells in which a gene segment has been replaced with one containing the desired alteration. The resulting mutant cells are then cloned into large, genetically pure populations. We then analyze DNA samples from many clones to test whether the desired mutation is present. Once identified, mutant cells are injected into mouse embryos at an early stage of development called the blastocyst stage. As these embryos develop into mice, some are

chimeras, that is, genetic mosaics recognizable by predictable coat color patterns. Our goal is to obtain chimeras in which the germ tissue (ovaries or testes) contains the mutation. Germ-line chimeras can be bred to "fix" the mutation into all cells of the mice in the second generation.

Once we have obtained high-quality ES cells, we plan to produce a gene knockout for the *XRCC1* DNA repair gene (remember that this gene is required for rejoining breaks in the DNA molecule). This procedure requires using the mouse *XRCC1* gene, which we have cloned and have found to be very similar to the human *XRCC1* gene. Transgenic mice that are defective in repairing DNA strand breaks should be unusually susceptible to mutation by radiation or chemical mutagens, thereby providing a sensitive test system for studying the origin of mutations. Other repair-deficient mice will also provide animal models for the human genetic disorders, including XP, in which DNA repair is defective and associated with an elevated cancer risk.

Conclusions

We have seen that the *Journal of the American Statistical Association* has a long history of publishing research in the field of statistics. The journal has been a leading source of information for statisticians and for those who are interested in the application of statistics to various fields. The journal has been a leading source of information for statisticians and for those who are interested in the application of statistics to various fields. The journal has been a leading source of information for statisticians and for those who are interested in the application of statistics to various fields.

References

1. L. H. Thompson, J. S. Rubin, J. E. Cleaver, G. F. Whitmore, and K. Brookman. "A Screening Method for Isolating DNA Repair-Deficient Mutants of CHO Cells." *Somatic Cell Mol. Genet.* **6**, 391 (1980).
2. L. H. Thompson, D. B. Busch, K. W. Brookman, C. L. Mooney, and D. A. Glaser. "Genetic Diversity of UV-Sensitive DNA Repair Mutants of Chinese Hamster Ovary Cells." *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3634 (1981).
3. D. Busch, C. Greiner, K. Lewis, R. Ford, G. Adair, and L. Thompson. "Summary of Complementation Groups of UV-Sensitive CHO Cell Mutants Isolated by Large-Scale Screening." *Mutagenesis* **4**, 349 (1989).
4. L. H. Thompson. "Somatic Cell Genetics Approach to Dissecting Mammalian DNA Repair." *Environ. Mol. Mutagenesis* **14**, 264 (1989).
5. C. A. Weber, E. P. Salazar, S. A. Stewart, and L. H. Thompson. "Molecular Cloning and Biological Characterization of a Human Gene, *ERCC2*, that Corrects the Nucleotide Excision Repair Defect in CHO UV5 Cells." *Mol. Cell. Biol.* **8**, 1137 (1988).
6. L. H. Thompson, K. W. Brookman, N. J. Jones, S. A. Allen, and A. V. Carrano. "Molecular Cloning of the Human *XPC1* Gene, which Corrects Defective DNA Strand-Break Repair and Sister Chromatid Exchange." *Mol. Cell. Biol.* **10**, 6160 (1990).
7. J. E. Cleaver. "Defective Repair Replication of DNA in Xeroderma Pigmentosum." *Nature* **218**, 652 (1968).

**For further
information
contact
Lawrence H.
Thompson
(510) 422-5658.**

Glossary

Allele	An alternative form of a gene that can result in the expression of a different characteristic, such as blue eyes versus brown eyes. If one allele is dominant and the other is recessive, the organism will have the appearance (phenotype) of the dominant allele.
Adduct	A molecular addition product caused by chemical agents, including many environmental pollutants (such as polycyclic aromatic hydrocarbons) and heterocyclic amines from cooked meat, that can cause lesions in DNA.
Assay	In the context of our DNA-repair work, a laboratory analysis to determine the role of DNA-repair proteins in a repair pathway.
Base pair	The partnership or pairing of the four nitrogenous bases in DNA. Adenine (A) always pairs with thymine (T), and guanine (G) always pairs with cytosine (C) in the DNA double helix.
Chimera	An organism consisting of tissues of diverse genetic constitution. Our chimeric mice are derived from embryonic cells of two different mouse strains. These chimeric mice are genetic mosaics recognizable by predictable coat color patterns.
CHO	Chinese hamster ovary line of cells. Cells taken from this rodent's ovaries many years ago are widely used in genetic research because they are robust and fast growing in culture. After exposure to mutagens, some of the mutants obtained from the CHO line are used for isolating human DNA-repair genes.
Chromosome	Discrete, threadlike particles (46 in humans, or 23 pairs) in the nucleus of a cell in which reside the genes. Each chromosome consists of a very long stretch of supercoiled DNA containing many genes in order, like beads on a string. Chromosomes can be easily visualized in most cells only during cell division.
Cloning	Asexually reproducing the cells, genes, or DNA fragments from a single ancestor. The multiple copies of genetically identical material are called clones.
Cloning vector	DNA molecule from a plasmid (see below), virus, or cell that can contain one or more inserted DNA fragments, sometimes from several sources. Vectors (the DNA carriers) allow foreign DNA to be reproduced in large quantities in host cells. Examples of cloning vectors include plasmid vectors, cosmid vectors, bacterial viruses, and yeast artificial chromosomes.
Codon	Triplets of nucleotides in the DNA sequence that represent one amino acid during the synthesis of a protein. For example, the triplet of three cytosine bases, C-C-C, codes for the amino acid proline.
Cosmid	An artificially made cloning vector that contains the <i>cos</i> gene of the lambda virus and an introduced DNA fragment. Cosmids are placed into the lambda virus protein coat for infection into the bacterial host <i>E. coli</i> . Both the virus and bacterium multiply to produce genetically identical copies (clones) of relatively large DNA inserts (up to 45,000 base pairs).

DNA	Deoxyribonucleic acid. The spiral, double-stranded macromolecules that contain the genetic blueprint.
ERCC	Excision repair cross-complementing. A family of human genes involved in the repair of DNA damage induced by UV light.
Excision	The removal of a single-stranded sequence of DNA containing damaged bases. Excision is one step of the multistep excision repair pathway.
Exon	A protein-coding sequence of DNA in a gene. Compare to intron (see below).
Free radical	A reactive atom or molecule that has one or more unpaired electrons. Some free radicals (such as OH•) arise from ionizing radiation, such as that associated with a nuclear explosion. Free radicals are highly reactive with the genetic material and can result in breaks in one or sometimes both DNA strands.
Gene	The fundamental physical and functional unit of heredity. Structurally, human genes are sections of chromosomal DNA that vary in size from fewer than 10,000 to more than 2 million base pairs. Functionally, each gene carries the code that specifies how to build a protein. Humans are now believed to have up to 100,000 different genes. These genes are composed of regions that code for a protein (exons) and regions that do not so encode (introns).
Genetic locus	A term used to designate a precise location of a DNA sequence on a chromosome, usually referring to the location of a gene.
Genome	The total genetic material, or DNA, in the chromosomes of a single cell. Most chromosomes are duplicated, so the size is often expressed as the total number of base pairs in one complete set of the two copies, or 3 billion base pairs in humans and other mammals.
Genotype	The genetic constitution of an individual. Compare to phenotype.
Germ cell	Sperm or egg (ovum), which contain 23 chromosomes in humans compared to 23 pairs of chromosomes found in most other body cells.
Hemizygous	The condition of having a gene present in a single dose.
Heterozygous	The condition of having one or more pairs of dissimilar alleles.
Homolog/homologous	Correspondence in structure. Homologous chromosomes have the same or allelic genes with homologous loci usually arranged in the same order.
Homozygous	The condition of having identical alleles in the corresponding loci of homologous chromosomes.
Hybridize	To fuse or bind. For example, two different rodent cells, or one human and one rodent cell, can be fused together (hybridized) so that the new hybrid cell contains chromosomes from both parental cells. In mapping gene fragments, a probe (tagged DNA fragment) hybridizes (binds) to a homologous sequence of bases along the chromosome target (see <i>in situ</i> hybridization).

Incision	The process of cutting the DNA backbone at or near the site of a damaged or modified nucleotide (lesion). Incision is one of the early steps in the multistep excision repair pathway, requiring many different repair proteins.
<i>In situ</i> hybridization	A mapping tool that tags DNA fragments, called probes, with one or more chemical labels, often dyes. When the probes bind to (hybridize with) a target piece of DNA with a complementary base sequence, the location can be seen under a microscope or with some other apparatus.
Intron	A DNA sequence that interrupts the protein-coding sequence of a gene. Compare to exon (see above).
Lambda	A bacterial virus (phage) used to copy, or clone, human DNA fragments. The fragments are inserted into the chromosomes of the lambda virus, which infects the bacterium <i>E. coli</i> , the host organism. Upon replication of both the virus and bacterium, large quantities of identical human DNA segments can be obtained for further study.
Ligation	The formation of a bond to link two adjacent bases in one strand of a double helix of DNA. Ligation is the final step of the excision repair pathway, requiring an enzyme called ligase.
Marker	Any identifiable place on a chromosome, such as the locus of a gene, that can be assigned, or mapped, to that chromosome.
Meiosis	Reduction division. The cell-division process that generates the germ (sperm and egg) cells, which contain 23 chromosomes (the haploid number) in humans.
Mitosis	The process by which somatic (body) cells divide and give rise to two identical daughter cells each containing the full complement of 46 chromosomes (the diploid number) in humans.
Mutagen	An agent that produces mutation. The DNA-damaging agents commonly used in DNA repair research include UV light, x rays, and chemicals (such as <i>cis</i> -platinum and mitomycin C).
Mutation	A structural change in the molecular units that make up genes, or a physical change in the relation of one chromosome to another. Point mutations, for example, are substitutions of single base pairs. Some mutations can prevent the gene in which they occur from functioning properly or at all.
NER	Nucleotide excision repair. A major, multistep DNA repair pathway responsible for removing DNA damage caused by nearly all bulky chemicals and the photoproducts of UV radiation. The NER mechanism replaces a patch of DNA containing the damage by excising a run of about 29 nucleotides.
Nucleotide	DNA is a double-stranded helix of repeating nucleotide units. Each nucleotide consists of a five-carbon sugar (deoxyribose), a phosphate group, and a nitrogenous base (see base pair).

Phage	A virus that uses a bacterial cell as its natural host for replication.
Phenotype	The expressed characteristics of an organism or cell, such as eye color or blood type, that arise from the genotype.
Plasmid	An autonomously replicating, circular DNA molecule that can be used to move genes into and out of bacteria and other cells. Plasmids are mini-chromosomes in bacterial cells, which consist of 1000 to 10,000 base pairs. They are separate from the main chromosomes and are not essential for cell survival. Artificially constructed plasmids are used as cloning vectors (see above).
Polymerization	One step of the excision repair pathway that also occurs during normal DNA replication. An undamaged complementary strand of DNA is used as a template to synthesize new material in the gap that results after a damaged, single-stranded sequence of DNA has been excised. The enzyme DNA polymerase is required to carry out this repair synthesis.
Probe	A DNA research tool, such as a cloned segment of human DNA. Probes can be labeled with fluorescent or radioactive tags. Such probes are used to locate human repair genes with respect to other mapped genes or to uniquely identify each type of human chromosome.
Protein	An essential class of macromolecules made of one or more chains of up to 20 different amino acids in a specific sequence. The sequence is determined by the gene coding for that protein. Besides water, proteins form the bulk of an organism's cells and tissues. Examples include hormones, such as insulin; antibodies, which fight infections; structural proteins that provide cell shape; and numerous enzymes that keep cells alive.
Recombinant DNA technology	The use of molecular biology tools to cut pieces of DNA (see restriction enzyme) and then combine the DNA from one organism with that of another.
Restriction enzyme	A class of proteins that functions as molecular scissors. Some restriction enzymes bind to DNA, travel along the double helix until they recognize a specific base sequence (often four to six bases), and cut the DNA to produce fragments, some with overhanging single-stranded ends. Because these "sticky" ends can recombine with any other piece of DNA with a complementary single-stranded overhang, restriction enzymes are used to join, or combine, DNA from one species to another to form a recombinant molecule.
Stem cell	Precursor cells, such as those in an embryo or in the bone marrow of an adult, that are undifferentiated (unspecialized). Over time, stem cells divide, giving rise to cells that are differentiated with respect to structure and function, such as red and white blood cells.
Transcription	The first step of the multistep process by which the genetic code is expressed as a protein product. During transcription, a single-stranded RNA molecule is synthesized from one strand of DNA.

- Transgenics** An experimental manipulation used as a research tool that alters the genetic makeup of an organism, often a mouse. A foreign or "trans" gene can be inserted into a chromosome, the function of a gene can be destroyed by disruption arising from insertion, or a gene can be targeted to modify its properties without destroying the protein it encodes.
- Translation** The process by which the triplets of nucleotides (codons) in messenger RNA are converted into amino acid sequences of proteins.
- XP** Xeroderma pigmentosum. A recessive human disease involving a defect in repairing DNA damage that arises from solar UV radiation and bulky chemicals. XP patients have extreme sun sensitivity, a high incidence of skin tumors, eye defects, and neurological disorders.
- XRCC** X-ray repair cross-complementing. A family of human genes involved in the repair of DNA damage induced by x rays.

Abstract **Unlocking the Mysteries of DNA Repair**

DNA, the genetic blueprint, has the remarkable property of encoding its own repair following diverse types of structural damage induced by external agents or normal metabolism. We are studying the interplay of DNA-damaging agents, repair genes, and their protein products to decipher the complex biochemical pathways that mediate such repair. Our research focuses on repair processes that correct DNA damage produced by chemical mutagens and radiation, both ionizing and ultraviolet. The most important type of DNA repair in human cells is called excision repair. This multistep process removes damaged or inappropriate pieces of DNA—often as a string of 29 nucleotides containing the damage—and replaces them with intact ones. We have isolated, cloned, and mapped several human repair genes associated with the nucleotide excision repair pathway and involved in the repair of DNA damage after exposure to ultraviolet light or mutagens in cooked food. We have shown that a defect in one of these repair genes, *ERCC2*, is responsible for the repair deficiency in one of the groups of patients with the recessive genetic disorder xeroderma pigmentosum (XP group D). We are exploring ways to purify sufficient quantities (milligrams) of the protein products of these and other repair genes so that we can understand their functions. Our long-term goals are to link defective repair proteins to human DNA repair disorders that predispose to cancer, and to produce DNA-repair-deficient mice that can serve as models for the human disorders.

Contact: Lawrence H. Thompson (510) 422-5658.

**DATE
FILMED**

11 / 9 / 93

END