# DNA Repair Comes Into Its Own

The linkage of mutations in DNA repair genes to hereditary cancers is only one of many advances taking place in this extremely active research area

In the familiar TV advertisement, the Maytag repairman laments never having enough to do because his company's appliances so rarely break down. It would be great if the same were true of DNA. But it's not. The genetic material, which contains all the information for building and operating cells, is in constant danger of malfunctioning, both because of errors introduced during its manufacture and damage caused later by environmental insults such as ultraviolet radiation and chemicals. As a result, the cell's DNA "repairman" has plenty of work to do to prevent breakdowns that could cause harmful effects—such as cancer.

Studying just how that work is done has become one of the most intriguing and rapidly moving fields of contemporary molecular biology. "DNA repair has essentially come into its own in the past few years," says Paul Modrich of Duke University, one of the pioneers in the field. Another DNA repair expert, James Cleaver of the University of California, San Francisco, agrees: "It's one of those times when lots of very disparate disciplines are converging. It's terribly fast-moving at the moment."

Helping propel the work is interest generated by the discovery during the past year that mutations in DNA repair genes can cause one of the most common hereditary cancers, hereditary nonpolyposis colon cancer (HNPCC). And researchers have recently identified a different set of DNA repair genes that when defective cause xeroderma pigmentosum (XP), a rare hereditary condition characterized by high rates of all kinds of skin cancers.

Whether this information will improve therapies for these cancers is uncertain, but at the very least it should lead to better tests for identifying people who carry the defective genes and thus are likely to develop the cancers. This will be especially valuable for HNPCC, which afflicts perhaps one person in 200, accounting for up to 15% of all colon cancers—about 20,000 cases every year in the United States. In contrast, only about one person in 250,000 gets XP.

But as important as the cancer discoveries are, they are just the tip of the large and still growing iceberg of DNA repair research. By dissecting components of the major DNA repair pathways in species ranging from bacteria to humans, researchers are also coming up with some surprises about the biology of normal cells. They have found, for example, that many of the proteins needed for DNA repair participate in other important cellular activities. Among these are gene regulation, DNA replication, and certain kinds of gene shuffling. "It's remarkable how versatile these proteins are. I'm always flabbergasted because these seem to be such different [cellular] activities," says geneticist Satya Prakash of the University of Texas Medical Branch in Galveston.

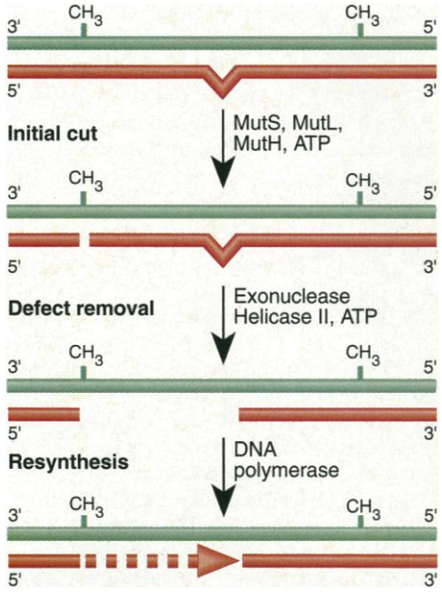



ILLUSTRATION: K. SUTLIFF

**Mismatch removal.** A DNA mismatch (indicated by bend) can be repaired from the 5' side, as shown here, or from the 3' side.

Researchers have made these discoveries while studying the two major types of DNA repair pathways, each of which has its own distinct function. "Mismatch repair," the one that is defective in HNPCC, corrects errors that creep in while DNA is being copied. These include simple mismatches, in which the wrong nucleotide building block is put into the new DNA strand. In contrast, "nucleotide excision repair," which goes awry in XP, snips out and replaces DNA damaged by chemicals and radiation.

### Making a match with cancer

Of these two major pathways, mismatch repair is the simpler. Researchers, including Modrich, Richard Fishel of the University of Vermont, Burlington, and Richard Kolodner of Harvard University's Dana-Farber Cancer Institute, have so far found that only a few—probably no more than a half dozen—proteins are needed for the initial steps of mismatch repair: recognition of the repair site and the first nicking of the damaged strand. Modrich has found, for example, that just three proteins, MutS, MutL, and MutH, are needed to accomplish this task in the bacterium *Escherichia coli* (though other cellular enzymes are needed to cut away, then rebuild, the DNA segment containing the mistake).

The cloning and sequencing of the mismatch repair genes has also revealed that the proteins they encode are very similar in species from bacteria all the way up the evolutionary ladder to human beings. "The mechanism is extremely conserved. It's clearly a very old system," notes Modrich, who adds that this is what would be expected for an activity as important as maintaining the integrity of an organism's genetic information. As it turned out, this evolutionary conservation was extremely helpful in the work leading to the identification of the HNPCC genes.

The first clue that a mismatch repair defect might cause HNPCC came about 18 months ago when several groups, including those of Manuel Perucho of the California Institute of Biological Research in La Jolla, Stephen Thibodeau of the Mayo Clinic, and Bert Vogelstein of the Johns Hopkins University School of Medicine and Albert de la Chapelle of the University of Helsinki, Finland, noted that HNPCC and some other colon cancer cells show an abnormality called microsatellite instability (*Science*, 7 May 1993, pp. 751, 810, 812, and 816).

Microsatellite DNA consists of short base sequences repeated over and over and dispersed throughout the genome. While these sequences vary in length from person to person, in any one individual's cells, they should all be the same. But the microsatellite DNA of HNPCC tumors varied in length from that of normal cells from the same patient, indicating that the sequences had gained or lost bases during tumor formation.

That was intriguing, because microsatellite instability is one of the defects mismatch repair is supposed to prevent. It can come about during replication of repetitive sequences if the two strands of DNA—the strand that is being copied and the new one being synthesized—slip relative to one another. As a result, the new strand will be

either longer or shorter than the old strand, depending on the direction of slippage. This in turn means the new double-stranded DNA molecule will contain a small loop of unpaired DNA. Normally, mismatch repair would target that loop, but apparently in the tumors something was going awry.

Still, finding microsatellite instability in the tumor cells didn't necessarily mean that defective mismatch repair causes colon cancer development. Cancer biologists have known for years that cancer cells show many kinds of chromosomal abnormalities—some of which may well be effects, rather than causes, of the malignant state.

Strong evidence that defective mismatch repair might be causing the colon cancer was not long in coming, however. At the same time the cell studies came out, the de la Chapelle–Vogelstein group reported genetic studies which linked HNPCC in some families to a gene on chromosome 2. This finding implied that microsatellite instability was the cause and not the effect of the cancer, Fishel says: "Here was a gene for colon cancer, and people who inherited it also displayed microsatellite instability."

By December, that implication was amply borne out as two teams, one led by Fishel and Kolodner, the other by Vogelstein and Kenneth Kinzler (also of Johns Hopkins) isolated the HNPCC gene on chromosome 2 and showed that it is the human equivalent of the bacterial mismatch repair gene *MutS*. The linkage studies indicated that the chromosome 2 gene, now known as *hMSH2*, accounts for only 50% to 60% of HNPCC cases. Within 3 months of its identification, however, the same two groups had identified the gene that causes most of the remaining cases: *hMLH1*, the human equivalent of the bacterial mismatch repair gene known as *MutL* (*Science*, 10 December 1993, p. 1645, and 18 March, pp. 1559 and 1625). And just this September, the Vogelstein team reported that it had linked mutations in two additional human *MutL* homologs, designated *hPMS1* and *hPMS2*.

Although so far the conclusion that defective mismatch repair underlies HNPCC is based mainly on the structural similarities between the human repair proteins and their bacterial counterparts, researchers are acquiring biochemical evidence to buttress the case. Modrich, working with the Johns Hopkins group, and more recently a team led by Thomas Kunkel of the National Institute of Environmental Health Sciences (NIEHS) in Research Triangle Park, North Carolina, have found that tumor cells with microsatellite instability are unable to carry out mismatch repair. "So far we have a perfect correlation," says Kunkel.

In addition, Fishel and his colleagues have shown that purified hMSH2 binds with high specificity to DNA sequences containing mismatched nucleotides, but not to those in which the nucleotides pair perfectly. In this regard, hMSH2's activity parallels that of its bacterial counterpart, MutS, whose job is to recognize and bind to appropriate DNA defects, getting mismatch repair under way. What apparently happens in HNPCC is that the DNA repair mutations destabilize the genome, presumably leading to other genetic alterations, such as the activation of oncogenes or the inactivation of tumor suppressor genes, that might cause cancer.

Although the parallels between the human mismatch repair proteins and those of bacteria and yeast have played a large role in identifying the HNPCC genes, recent work, including a report on page 814 from Kunkel's group, suggests that the mismatch repair capabilities of bacterial and mammalian cells may not be identical. Researchers previously noted that bacterial mismatch repair can't handle unpaired loops containing more than three nucleotides. But, as Kunkel points out, bigger loops may well form during replication of human microsatellite DNA, which is not only far more abundant than that of bacteria, but also often consists of repeating sequences up to five or six base pairs long. "The human genome is so full of [microsatellite] DNA, it offers a big opportunity for that kind of slippage loop formation," he says. Those observations led Kunkel and his NIEHS colleagues Asad Umar and Jayne Boyner to ask whether humans have a repair system not possessed by bacteria to repair the lesions.

The new results indicate that the answer is yes. They suggest that human cells have the ability, not possessed by bacteria, to target repair to loops containing five or more nucleotides. It is possible that this is a new activity acquired by hMSH2, although Kunkel cautions that the experiments do not rule out the involvement of other proteins. The evolution of proteins to repair the larger loops "makes a lot of sense from an organism's standpoint," remarks Fishel, who is also investigating the repair of larger loop mismatches in human cells. "Since the bacterial genome has very few of these repeats, the bacterial enzyme doesn't have to recognize these errors."

### Versatile but complex

These results raise the possibility that mismatch repair in higher organisms may be more complicated than researchers previously assumed. Yet that pathway still looks a good deal simpler than the other major pathway for correcting DNA errors: nucleotide excision repair (NER), which uses the products of at least a dozen genes to recognize and clip out damaged DNA segments.

The pathway's additional complexity may reflect the fact that NER has more work to do in the cell. "Nucleotide excision repair defends the body against damage caused both spontaneously and by exogenous agents," says Richard Wood of the Imperial Cancer Research Fund Clare Hall Laboratories in South Mimms, U.K. "It's the most versatile mechanism." Indeed, agrees Aziz Sancar of the University of North Carolina, Chapel Hill, who studies excision repair in bacteria, "in the beginning of my career, I tried all kinds of DNA-damaging agents, and every time I published a paper because I couldn't find anything [the system] didn't repair."

Because NER protects the cell against damage from ultraviolet (UV) radiation and chemicals, defects in this pathway could potentially predispose to any kind of cancer. But so far the only cancers definitively linked to mutations in NER genes are the skin cancers that afflict XP patients. That link had been expected since the late 1960s, when UCSF's Cleaver found that skin cells from XP patients cannot repair the damage induced by ultraviolet light. Cleaver's observation suggested that mutations causing XP would be in the genes needed for DNA repair, leaving the patients extremely susceptible to UV's carcinogenic effects. Indeed, their risk of skin cancer is some 1000 to 2000 times greater than normal.

It soon became clear that there are several of these XP genes—a conclusion reached by fusing cells from two different XP patients.

**SOME OF THE GENES NEEDED FOR NUCLEOTIDE EXCISION REPAIR**

| Human Gene | Yeast Equivalent | Protein Function |
|---|---|---|
| *XPA* | RAD14 | Bind damaged DNA |
| *XPB* | SSL2 (RAD25) | Helicase TFIIH component |
| *XPC* | RAD4 | ? |
| *XPD* | RAD3 | Helicase TFIIH component |
| *XPE* | ? | Bind damaged DNA |
| *XPF* | RAD1 | Works with ERRC1 (RAD10) protein to cut DNA |
| *XPG* | RAD2 | Cuts DNA |

Surprisingly, the hybrid cells were often no longer deficient in DNA repair, indicating that the two parent cells had different defects, and so when they fused, the resulting hybrids had functional copies of all the genes needed for NER. By systematically doing such fusions, researchers learned that XP can be caused by mutations in any of seven genes. But even that didn't exhaust the battery of NER genes, as other work indicated that at least five or six more genes are needed.

Isolating the XP genes proved to be a major challenge, mainly because the DNA transfer strategy commonly used to isolate genes didn't work for XP cells. But beginning in the mid-1980s, researchers developed more effective systems for spotting XP genes. Now, at least six XP genes have been isolated, and researchers are making rapid headway toward understanding the functions of the proteins they encode.

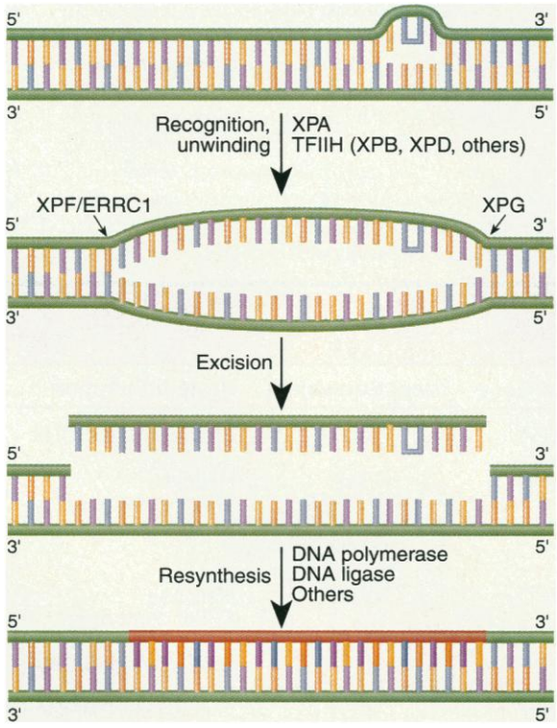
The picture being developed shows that the XPA protein recognizes and binds to damaged DNA, while others, including XPB and XPD, are helicases, enzymes that unwind the DNA double helix so that the damaged segment of DNA can be clipped out. The clipping requires still other proteins. In work reported this fall, Errol Friedberg's group at the University of Texas Southwestern Medical Center in Dallas showed that it takes two proteins, XPF and one called ERRC1, which has not been linked to XP, to make the cut on the right-hand (5') side of the damage site; Wood's group demonstrated that XPG makes the cut on the left-hand (3') side. As had been shown by North Carolina's Sancar, these enzymatic snippers remove a DNA segment about 30 nucleotides long, which must be replaced by the cell's DNA-synthesizing enzymes.

But that's only the opening chapter of the story of what the NER proteins do. Much recent evidence shows that these enzymes also play roles far removed from the arena of DNA repair. Work by Galveston's Satya and Louise Prakash, for example, suggests that the yeast equivalents of ERRC1 and XPF bring about the DNA cutting needed for certain types of genetic rearrangements.

The discovery of another nonrepair function—the involvement of NER proteins in gene expression—may help solve a mystery that has long puzzled XP researchers. Some XP patients also develop either of two other rare hereditary diseases, Cockayne's syndrome (CS) and trichothiodystrophy (TTD), whose symptoms are very different from those of XP: CS and TTD are characterized by neurological and other developmental abnormalities, including severe mental retardation, but not by a big increase in skin cancers.

How could mutations in the same genes produce such wildly varying symptoms? Last year, a team led by Jean-Marc Egly of the INSERM unit of the University of Strasbourg, France, and Jan Hoeijmakers of Erasmus University in Rotterdam provided part of the answer: They found that XPB is one of several proteins in an important "transcription factor," TFIIH, which helps regulate the activity of all protein-encoding genes. Since then, other teams, including one led by Friedberg and Roger Kornberg of Stanford University School of Medicine and another led by North Carolina's Sancar and Danny Reinberg of the University of Medicine and Dentistry of New Jersey, have shown that at least five other NER proteins are also part of TFIIH.

These results suggest that the CS and TTD symptoms are caused, not by defective DNA repair as XP is, but rather by impairment of the proteins' other function: the regulation of gene transcription. "From the human side of things, the findings have opened a whole new horizon in thinking about these diseases," Friedberg says.




**Cutting it out.** In excision repair, the nucleases XPF/ERRC1 and XPG clip out about 30 nucleotides in all.

The discovery that certain NER proteins are components of TFIIH is not the only link that has been found between repair and transcription. In the mid-1980s, Philip Hanawalt's group at Stanford noted that the DNA damage caused by UV light is preferentially repaired in the transcribed strands of active genes. As Hanawalt puts it, the transcribing enzyme RNA polymerase II (pol II) somehow "serves as an antenna to recruit the repair enzymes" to the sites of the lesions it encounters. How that happens is also beginning to be clarified.

Sancar's group has identified the protein produced by a bacterial gene called *Mfd* (for mutation frequency decline) as the coupling factor that recognizes the bacterial RNA polymerase when it has stalled at a lesion, pulls it off, and attracts the repair enzymes. In mammalian cells, Hoeijmakers, Dirk Bootsma, also of Erasmus University, and their colleagues have identified a protein known as ERRC6 as a likely coupling factor. But Hanawalt suggests that mammalian transcription-repair coupling may differ in a significant way from that in bacteria. Rather than pulling pol II off the gene entirely, his group's work suggests that it stays in place, while the RNA it is making is partially degraded to allow access for the repair enzymes. Otherwise, Hanawalt says, pol II might spend hours copying a long mammalian gene, only to have to start all over again if it encounters a lesion near the end.

While it's clear that researchers have learned a great deal about both mismatch and nucleotide excision repair, they have not solved all the problems. One outstanding question concerns whether repair defects are associated with cancers other than HNPCC and skin cancers. There are some indications that they might be. Researchers have detected microsatellite instability in several kinds of cancers. For example, based on work on head and neck, bladder, and lung tumors, David Sidransky and his colleagues at Johns Hopkins have suggested that the abnormalities might be used as diagnostic markers. There are also indications that XP patients might be at increased risk of some internal cancers, although the increase is nowhere near that seen with the skin cancers.

But with the DNA repair work going so well, researchers are optimistic that answers to these—and other—questions will be found. As Satya Prakash puts it, "When I started out, I thought we would have some fun, but I didn't think we would have so much fun." And all because, unlike the work of the Maytag repairman, which rarely begins, the work of DNA repair enzymes hardly ever ends.

**–Jean Marx**

ILLUSTRATION: K. SUTLIFF


Additional Reading

J. E. Cleaver, "It was a very good year for DNA repair," *Cell* **76**, 1 (1994).

P. C. Hanawalt, B. A. Donahue, K. S. Sweder, "Repair and transcription: Collision or collusion?" *Current Biology* **4**, 518 (1994).

K. Tanaka and R. D. Wood, "Xeroderma pigmentosum and nucleotide excision repair of DNA," *Trends in Biochemical Sciences* **19**, 83 (1994).