Transposable Genetic Elements

They bypass the rules of ordinary genetic recombination and join together segments of DNA that are unrelated, transferring groups of genes among plasmids, viruses and chromosomes in living cells

by Stanley N. Cohen and James A. Shapiro

Natural selection, as Darwin recognized more than a century ago, favors individuals and populations that acquire traits conducive to survival and reproduction. The generation of biological variation, which gives rise to new and potentially advantageous combinations of genetic traits, is therefore a central requirement for the successful evolution of species in diverse and changing environments.

Hereditary information is encoded in the sequence of the building blocks, called nucleotides, that constitute a molecule of DNA, the genetic material. The basic step in the creation of genetic variation is the mutation, or alteration, of the DNA within a gene of a single individual. Mutations involve changes in nucleotide sequence, usually the replacement of one nucleotide by another. This can lead to a change in the chain of amino acids constituting the protein encoded by the gene, and the resulting change in the properties of the protein can influence the organism's biological characteristics. Spontaneous mutations are too rare, however, for genetic variation to depend on new mutations that arise in each generation. Instead variation is generated primarily by the reshuffling of large pools of mutations that have been accumulated within a population in the course of many generations.

In higher organisms this reshuffling is done in the process of sexual reproduction. The genes are arrayed on two sets of chromosomes, one set inherited from the female parent and the other set from the male parent, so that there are two copies of each gene. Sometimes the nucleotides of a genetic sequence differ slightly as a result of earlier mutation, producing alleles, or variant forms of a gene. In the formation of gametes (egg or sperm cells) the breakage of structurally similar pairs of chromosomes can result in the reciprocal exchange of alleles between the two members of a pair of chromosomes. Such genetic recombination requires that the segments of DNA undergoing exchange be homologous, that is, the sequence of nucleotides on one segment of DNA must be very similar to the sequence on the other segment, differing only at the sites where mutations have occurred.

The ability of segments of DNA on different chromosomes to recombine makes it likely that in complex plants or animals the particular collection of genes contained in each egg or sperm cell is different. An individual produces many eggs or sperms, which can potentially interact with sperms or eggs from many other individuals, so that there is a vast opportunity for the generation of genetic diversity within the population. In the absence of intentional and extended inbreeding the possibility that any two plants or animals will have an identical genetic composition is vanishingly small.

Genetic variation is also important in the evolution of lower organisms such as bacteria, and here too it arises from mutations. Bacteria have only one chromosome, however, so that different alleles of a gene are not normally present within a single cell. The reshuffling of bacterial genes therefore ordinarily requires the introduction into a bacterium of DNA carrying an allele that originated in a different cell. One mechanism accomplishing this interbacterial transfer of genes in nature is transduction: certain viruses that can infect bacterial cells pick up fragments of the bacterial DNA and carry the DNA to other cells in the course of a later infection. In another process, known as transformation, DNA released by cell death or other natural processes simply enters a new cell from the environment by penetrating the cell wall and membrane. A third mechanism, conjugation, involves certain of the self-replicating circular segments of DNA called plasmids, which can be transferred between bacterial cells that are in direct physical contact with each other.

Whether the genetic information is introduced into a bacterial cell by transduction, transformation or conjugation, it must be incorporated into the new host's hereditary apparatus if it is to be propagated as part of that apparatus when the cell divides. As in the case of higher organisms, this incorporation is ordinarily accomplished by the exchange of homologous DNA; the entering gene must have an allelic counterpart in the recipient DNA. Because homologous recombination requires overall similarity of the two DNA segments being exchanged, it can take place only between structurally and ancestrally related segments. And so, in bacteria as well as in higher organisms, the generation of genetic variability by this mechanism is limited to what can be attained by exchanges between different alleles of the same genes or between different genes that have stretches of similar nucleotide sequences. This requirement imposes severe constraints on the rate of evolution that can be attained through homologous recombination.

Until recently mutation and homologous recombination nevertheless appeared to be the only important mechanisms for generating biological diversity. They seemed to be able to account for the degree of diversity observed in most species, and the implicit constraints of homologous recombination—which prevent the exchange of genetic information between unrelated organisms lacking extensive DNA-sequence similarity—appeared to be consistent with both a modest rate of biological evolution and the persistence of distinct species that retain their basic identity generation after generation.

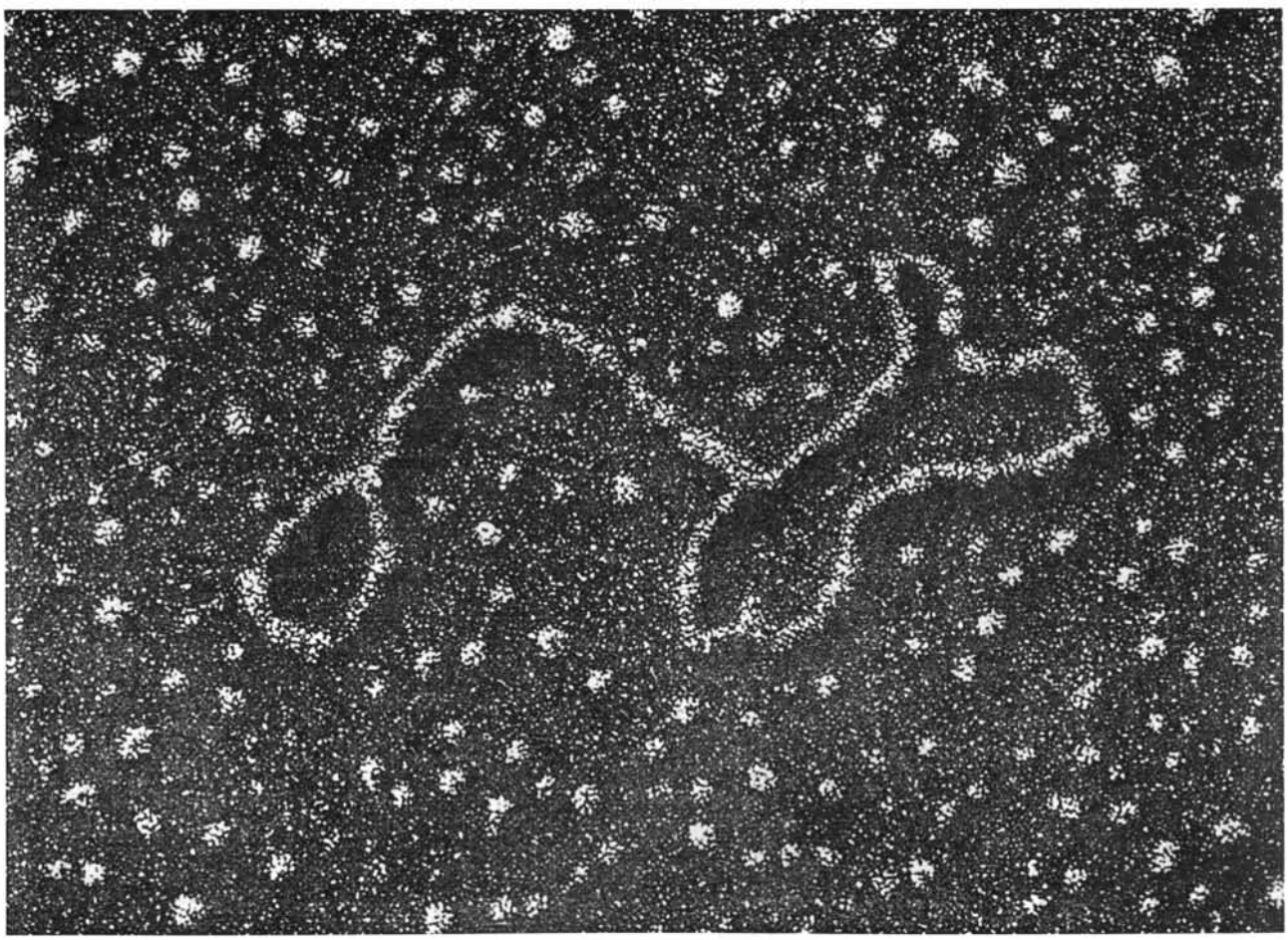
Within the past decade or so, however, it has become increasingly apparent that there are various "illegitimate" recombinational processes, which can join together DNA segments having little or no nucleotide-sequence homology, and that such processes play a significant role in the organization of genetic information and the regulation of its expression. Such recombination is often effected by transposable genetic elements: structurally and genetically discrete segments of DNA that have the

ability to move around among the chromosomes and the extrachromosomal DNA molecules of bacteria and higher organisms. Although transposable elements have been studied largely in bacterial cells, they were originally discovered in plants and are now known to exist in animals as well. Because illegitimate recombination can join together DNA segments that have little, if any, ancestral relationship, it can affect evolution in quantum jumps as well as in small steps.

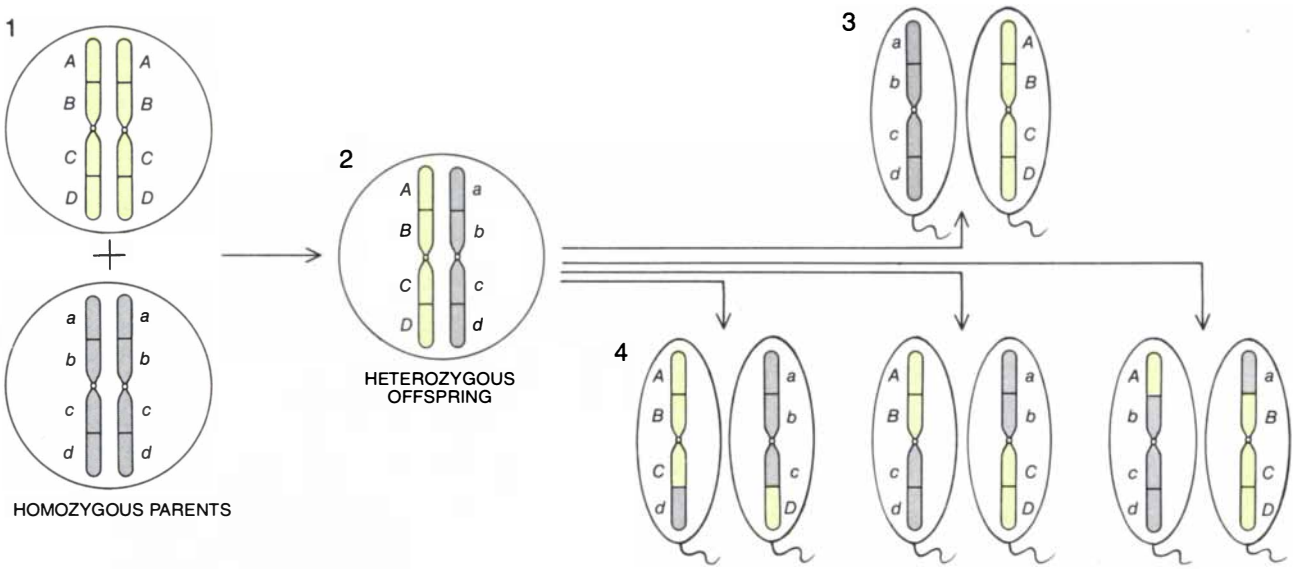
In the late 1940's Barbara McClintock of the Carnegie Institution of Washington's Department of Genetics at Cold Spring Harbor, N.Y., first reported a genetic phenomenon in the common corn plant, *Zea mays,* that would later be found to have parallels in other biological systems. While studying the inheritance of color and the distribution of pigmentation in plants that had undergone repeated cycles of chromosome breakage she found that the activity of particular genes was being turned on or off at abnormal times. Because some of these genes were associated with the development of pigments in kernels as well as in the plant itself, certain kernels were mottled, showing patches of pigmentation against an otherwise colorless background. The patterns of this variegation were reproduced in successive generations and could be analyzed like other heritable traits. After painstaking study of many generations of corn plants McClintock concluded that the variegation she observed was the result of the action of distinct genetic units, which she called controlling elements, that could apparently move from site to site on different maize chromosomes; as they did so they sometimes served as novel biological switches, turning the expression of genes on or off.

McClintock's genetic analysis showed that some patterns of variegation affected three or more genes simultaneously, suggesting that the structure of one of the plant's chromosomes had been rearranged at the site of a controlling element. Direct microscopic examination of maize chromosomes containing controlling elements confirmed that these genetic elements did in fact serve as specific sites for the breakage and resealing of DNA, thereby giving rise to either minute or gross changes in chromosome structure.

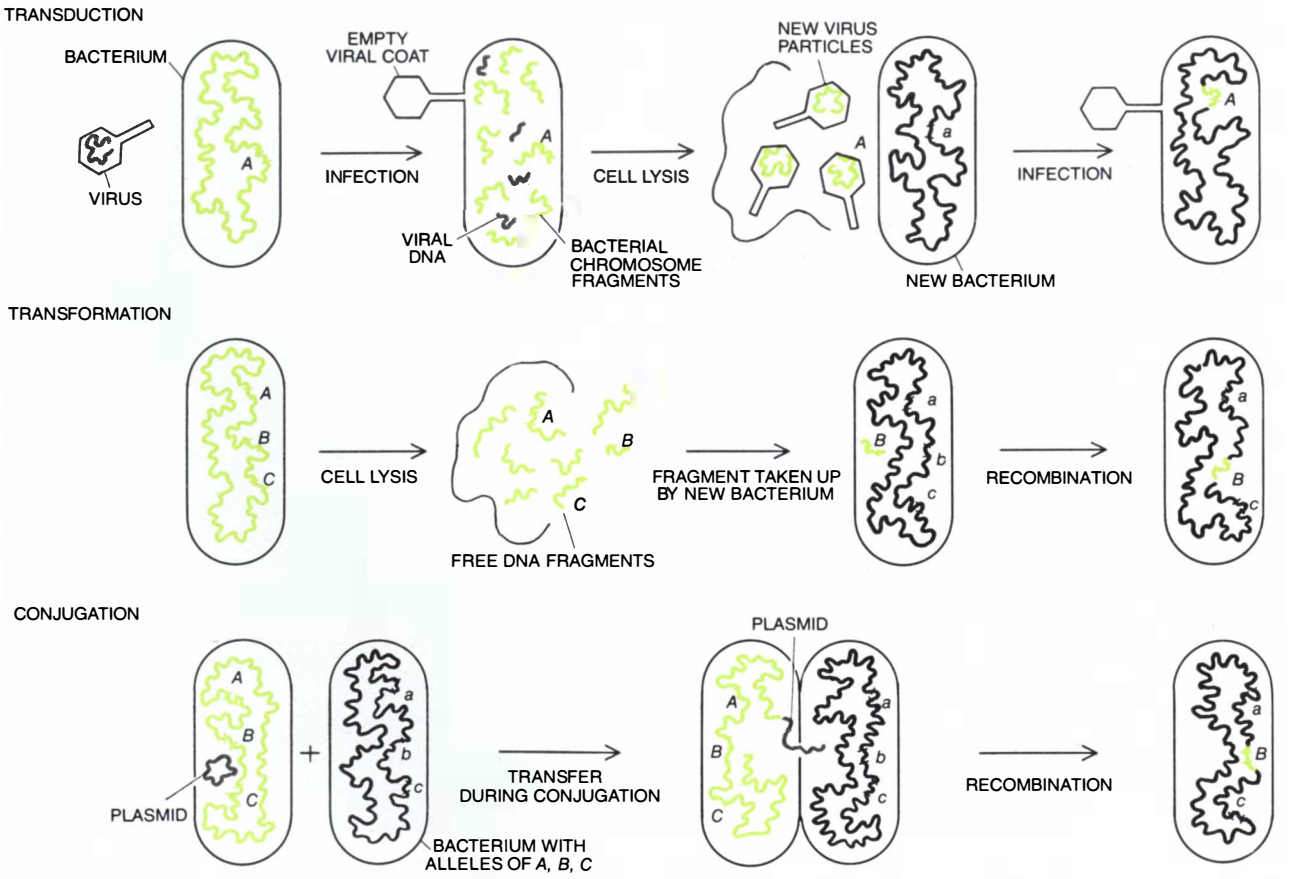
Almost 20 years after McClintock reported her earliest studies on controlling elements in the corn plant Michael Malamy, who is now at the Tufts University School of Medicine, Elke Jordan, Heinz Saedler and Peter Starlinger of the University of Cologne and one of us (Shapiro), who was then at the University of Cambridge, found a new class of mutations in genes of a laboratory strain of the common intestinal bacterium *Escherichia coli.* They were unusual in that their effects were detectable beyond the borders of the mutated genes them-



DNA OF TRANSPOSABLE GENETIC ELEMENT (a transposon) forms a characteristic stem-and-loop structure, which is seen here in an electron micrograph made by one of the authors (Cohen). The structure results from the "inverted repeat" nature of the nucleotide sequences at the two ends of the transposon DNA (*see upper illustration on page 44*). The double-strand DNA of the plasmid *pSC*105, into which the transposon had been inserted, was denatured and complementary nucleotide sequences on each strand were allowed to "reanneal." The joining of the complementary nucleotides constituting the transposon's inverted-repeat termini formed the double-strand stem. The smaller loop was formed by the segment of single-strand transposon DNA between the inverted repeats, a segment that includes a gene conferring resistance to the antibiotic kanamycin. The larger loop represents the single-strand DNA of a miniplasmid derivative of the host plasmid. DNA was spread with formamide and shadowed with platinum-palladium. Enlargement is 230,000 diameters.



HOMOLOGOUS RECOMBINATION is accomplished in higher organisms by the "crossing over" of structurally similar chromosome segments during sexual reproduction. Here the process is shown for a hypothetical animal each of whose somatic (body) cells has a single chromosome pair carrying four genes, each of which may be present in either of two variant forms (alleles). Homozygous parents having the same set of alleles on both paired chromosomes (*1*) give rise to heterozygous offspring (*2*), which in turn can produce gametes (sperms or eggs) containing copies of the original chromosomes (*3*). As a result of crossing over and reciprocal homologous recombination, alleles can be reshuffled in various ways (*4*), producing gametes containing chromosomes that are different from either of original chromosomes.



RECOMBINATION IN BACTERIA requires the introduction into a bacterial cell of an allele obtained from another cell. In transduction an infecting phage, or bacterial virus, picks up a bacterial-DNA segment carrying allele *A* and incorporates it instead of viral DNA into the virus particle. When such a particle infects another cell, the bacterial-DNA segment recombines with a homologous segment, thereby exchanging allele *A* for allele *a*. In transformation a DNA segment bearing allele *B* is taken up from the environment by a cell whose chromosome carries allele *b;* the alleles are exchanged by homologous recombination. In conjugation a plasmid inhabiting one bacterial cell can transfer the bacterium's chromosome, during cell-to-cell contact, to another cell whose chromosome carries alleles of genes on the transferred chromosome; again allele *B* is exchanged for allele *b* by recombination between homologous DNA segments.

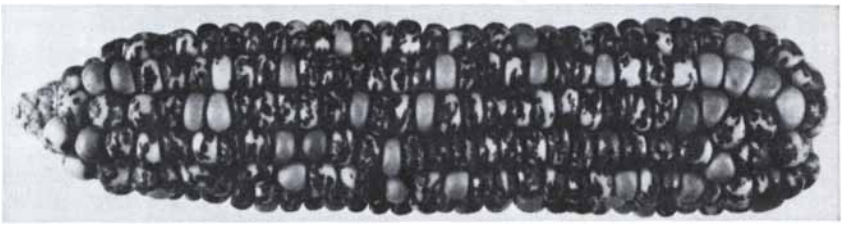
selves; this property could not be explained by any known mutational mechanism.

When the DNA segments carrying these mutations were inserted into particles of a bacterial virus and the density of the virus was compared with that of viruses carrying normal genes, it became clear that the mutated DNA was longer than the normal DNA: the mutations had been caused by the insertion of sizable DNA fragments into the mutated gene. It further developed that a limited number of other kinds of distinguishable DNA segments, which were up to 2,000 nucleotides in length, could also insert themselves within many different genes, interrupting the continuity of the gene and turning off its activity. These elements were named insertion sequences, or IS elements. The observation that a small number of specific DNA segments could be inserted at a large number of different sites in the bacterial chromosome suggested that some type of nonhomologous recombination was taking place; it seemed to be unlikely that an IS element could be homologous with the nucleotide sequences at so many different insertion sites.

At about the same time that IS elements were discovered other microbiologists and geneticists made observations hinting that certain genes known to be responsible for resistance to antibiotics by bacteria were capable of transfer from one molecule of DNA to another. Results obtained by Susumu Mitsuhashi and his colleagues at the University of Tokyo in the mid-1960's suggested that a gene encoding a protein that inactivates the antibiotic chloramphenicol could move from its normal site on a plasmid-DNA molecule to the chromosome of a bacterium or to the DNA of a virus.

Similar instances of the apparent transfer of antibiotic-resistance genes between different DNA molecules in the same cell were reported from the U.S. and Britain. The first direct evidence that such transfer is by a process analogous to the insertion of IS elements was published in 1974. R. W. Hedges and A. E. Jacob of the Hammersmith Hospital in London found that the transfer from one plasmid to another of a gene conferring resistance to antibiotics such as penicillin and ampicillin was always accompanied by an increase in the size of the recipient plasmid; the recipient could donate the resistance trait to still other plasmids, which thereupon showed a similar increase in size.

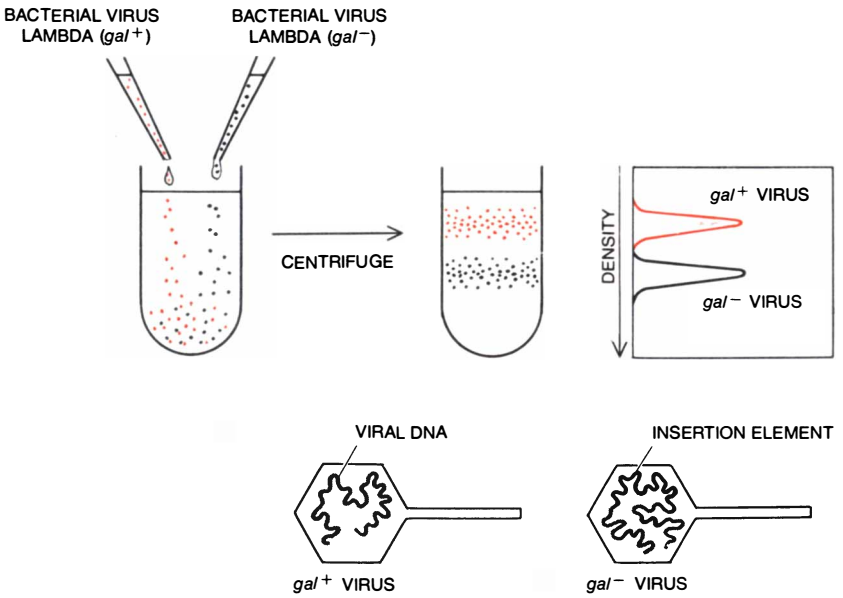
Hedges and Jacob postulated that the gene for ampicillin resistance was carried by a DNA element that could be "transposed," or could move from one molecule to another, and they called such an element a transposon. Their discovery of a transposable element that carries an antibiotic-resistance gene was an important advance. In earlier studies the movement of IS elements had been tracked only indirectly by genetic techniques: by observing the effects of insertions on various genetic properties of the host organism. It now became possible to track a transposable element's intermolecular travels directly by observing the inheritance of the antibiotic-resistance trait.



DIFFERENCES IN PIGMENTATION in kernels of the corn plant *Zea mays* reflect the action of a two-element control system discovered by Barbara McClintock of the Cold Spring Harbor Laboratory. Both elements are transposable. One element is at the locus of a gene whose action it modulates to yield the faintly and homogeneously pigmented kernels. The other element acts on the first one to produce the variegated pattern that is seen in many of the kernels.

While the Hedges and Jacob experiments were being carried out Dennis J. Kopecko and one of us (Cohen), at the Stanford University School of Medicine, were studying the acquisition of a gene for resistance to ampicillin by still other plasmids. It emerged, as Hedges and Jacob had found, that the ampicillin-resistance trait present on one plasmid could be acquired by another plasmid. Surprisingly, however, it also developed that such transfer could take place in mutated bacteria lacking a particular protein, the product of a gene designated *rec*A, known to be necessary for homologous recombination. Examination of the plasmid DNA with the electron microscope revealed that a 4,800-nucleotide segment carrying the ampicillin-resistance trait was being transferred as a characteristic and discrete structural unit. Moreover, the segment could become inserted at many different sites on the recipient plasmid DNA.

Electron microscopy also showed that the two ends of the transposable DNA segment had a unique feature: they consisted of nucleotide sequences that were complementary to each other but in the reverse order. This finding calls for some explanation. The four nitrogenous bases that characterize DNA nucleotides are linked in complementary pairs by hydrogen bonds to form the double helix of DNA: adenine (*A*) is linked to thymine (*T*) and guanine (*G*) to cytosine (*C*). The nucleotide sequence *AGCTT*, for example, is complementary

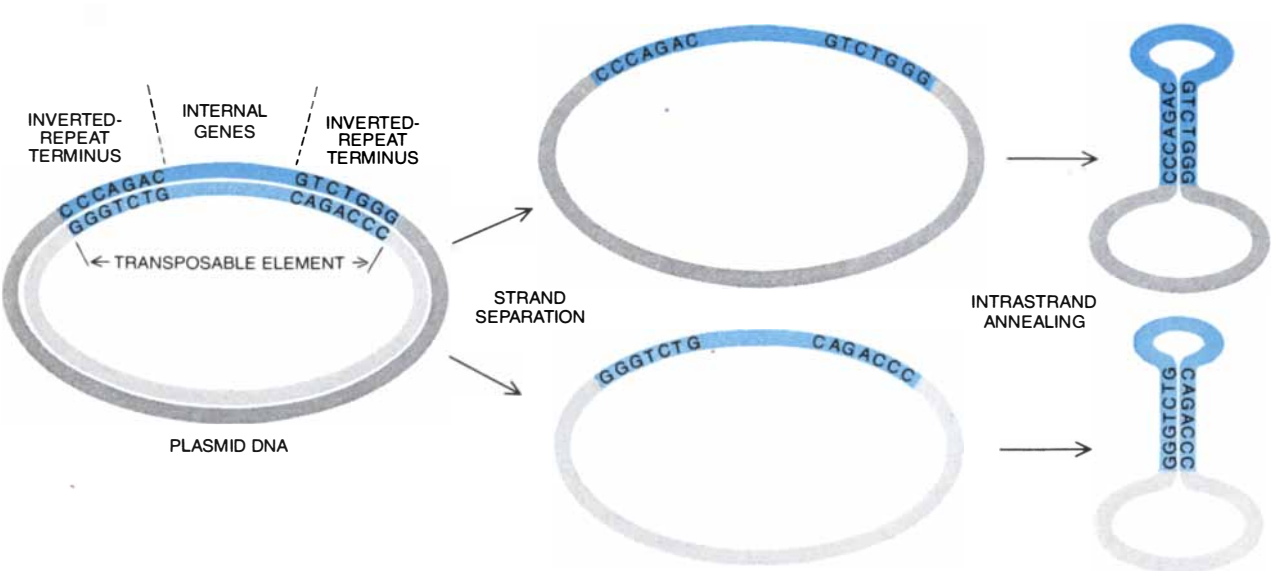


MUTATION BY INSERTION was demonstrated by one of the authors (Shapiro) with phage-lambda particles carrying the bacterial gene for galactose utilization (*gal*$^+$) and particles carrying the mutant gene *gal*$^-$. The viruses were centrifuged in a cesium chloride solution. The *gal*$^-$ particles were found to be the denser. Because the virus particles all have the same volume and their outer shells all have the same mass, increased density of *gal*$^-$ particles showed they must contain a larger DNA molecule: *gal*$^-$ mutation was caused by insertion of DNA.

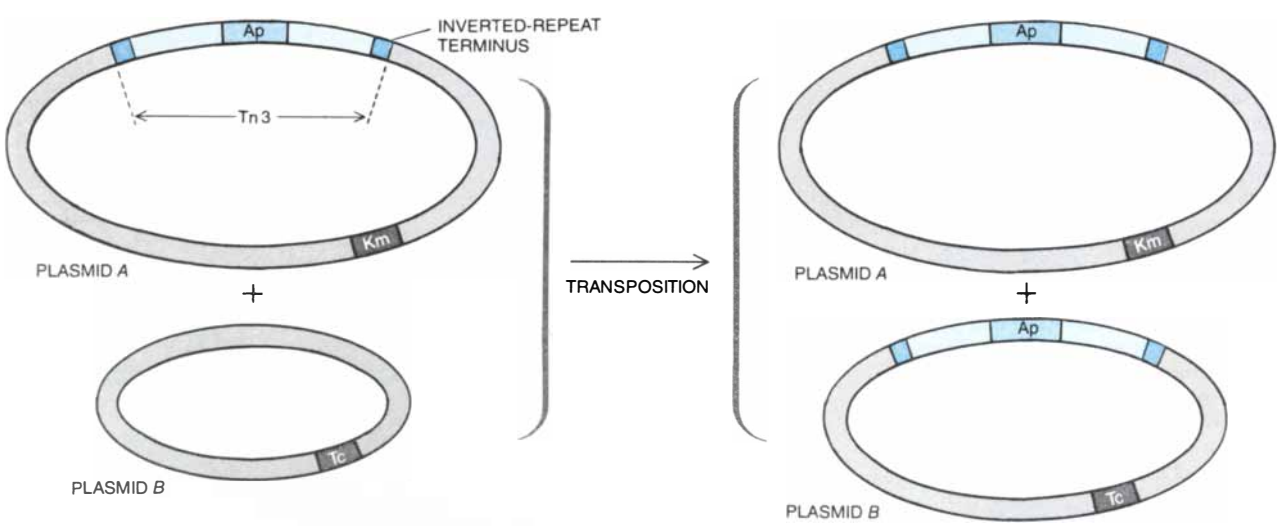
to the sequence *TCGAA*. The nucleotide sequence at one end of the transposable DNA segment was complementary in reverse order to a sequence on the same strand at the other end of the element [*see upper illustration on this page*]. These "inverted repeats" were revealed when the two strands of the double-strand plasmid DNA carrying the transposon were separated in the laboratory and each of the strands was allowed to "reanneal" with itself: a characteristic stem-and-loop structure was formed by the complementary inverted repeats.

The result of the transposition process is that a segment of DNA originally present on one molecule is transferred to a different molecule that has no genetic homology with the transposable element or with the donor DNA. The fact that the process does not require a bacterial gene product known to be necessary for homologous recombination indicates that transposition is accomplished by a mechanism different from the usual recombinational processes.

Subsequent experiments done in numerous laboratories have shown that DNA segments carrying genes encoding a wide variety of antibiotic-resistance traits can be transferred between DNA molecules as discrete units. Moreover, genes encoding other traits, such as resistance to toxic mercury compounds, synthesis of bacterial toxins and the capacity to ferment sugars or metabolize hydrocarbons, have been shown to be capable of transposition. All the transposons studied so far have ends consisting of inverted-repeat sequences, which range in length from only a few nucleotides to as many as 1,400. The ends of at least two transposons actually consist of two copies of the insertion sequence IS1 (which itself has been found to have terminal inverted-repeat sequences). Recent evidence has suggested that the insertion of any gene between two trans-



STEM-AND-LOOP STRUCTURES demonstrate the inverted-repeat nucleotide sequences of the ends of transposable elements. The four bases adenine (*A*), guanine (*G*), thymine (*T*) and cytosine (*C*) of DNA's four nucleotide building blocks are linked to form a helix (shown here schematically as a double strand); *A* always pairs with *T* and *G* pairs with *C*. The termini of a transposable element have sequences (seven nucleotides long here) that are bidirectionally and rotationally symmetrical. When the two strands of a plasmid containing an element are separated and each strand is allowed to self-anneal, the complementary nucleotides at the termini pair with each other, forming a double-strand stem (*right and in electron micrograph on page 41*). The remainder of the DNA is seen as single-strand loops.



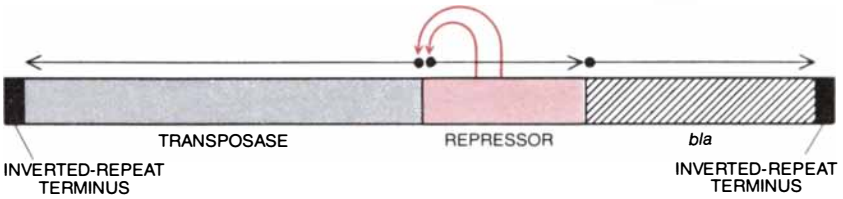
TRANSPOSITION of the transposon Tn3, which carries a gene conferring resistance to the antibiotic ampicillin (*Ap*), is diagrammed. It is shown as originally being part of plasmid *A*, which also includes a gene for resistance to kanamycin (*Km*). A plasmid *B*, which confers resistance to tetracycline (*Tc*), acquires a copy of the transposon. The new plasmid *B* confers resistance to ampicillin and tetracycline.

posable elements makes possible the transfer of the gene to a structurally unrelated DNA molecule by nonhomologous recombination.

Since transposable elements are transferred as discrete and characteristic genetic units, there must be some highly specific enzymatic mechanism capable of recognizing their inverted-repeat ends and cleaving DNA precisely at these locations. The first evidence that genes carried by the transposable elements themselves can encode such enzymes came from a series of experiments carried out by Frederick L. Heffron, Craig Rubens and Stanley Falkow at the University of Washington and continued by Heffron and his colleagues at the University of California at San Francisco. When these investigators introduced mutations that interrupted the continuity of genes at various locations within the ampicillin-resistance transposon designated Tn3, they found alterations in the ability of the element to function as a transposon. Mutation in the inverted-repeat ends or in a particular region of the DNA segment between the ends prevented transposition. On the other hand, mutations within another region of Tn3 actually increased the frequency of movement of Tn3 between the different plasmids, suggesting that this region might contain a gene modulating the ability of Tn3 to undergo transposition.

Recently published work by Joany Chou, Peggy G. Lemaux and Malcolm J. Casadaban in the laboratory of one of us (Cohen) and by Ronald Gill in Falkow's laboratory has shown that the Tn3 transposon does in fact encode both a "transposase"—an enzyme required for transposition—and a repressor substance that regulates both the transcription into RNA of the transposase gene and the repressor's own synthesis. Analogous experiments at the University of Chicago, the University of Wisconsin and Harvard University have shown that other transposable elements also encode proteins needed for their own transposition.

Even though transposons can insert themselves at multiple sites within a recipient DNA molecule, their insertion is not random. It has been recognized for several years that certain regions of DNA are "hot spots" prone to multiple insertions of transposons. Experiments recently reported by David Tu and one of us (Cohen) have shown that Tn3 is inserted preferentially in the vicinity of nucleotide sequences similar to sequences within its inverted-repeat ends, even in a bacterial cell that does not make the *rec*A protein required for ordinary homologous recombination. It therefore appears that recognition of homologous DNA sequences may play some role in determining the frequency and site-specificity of transposon-associated recombination, even though the actual recombinational mechanism differs from the one commonly associated with the exchange of homologous segments of DNA.

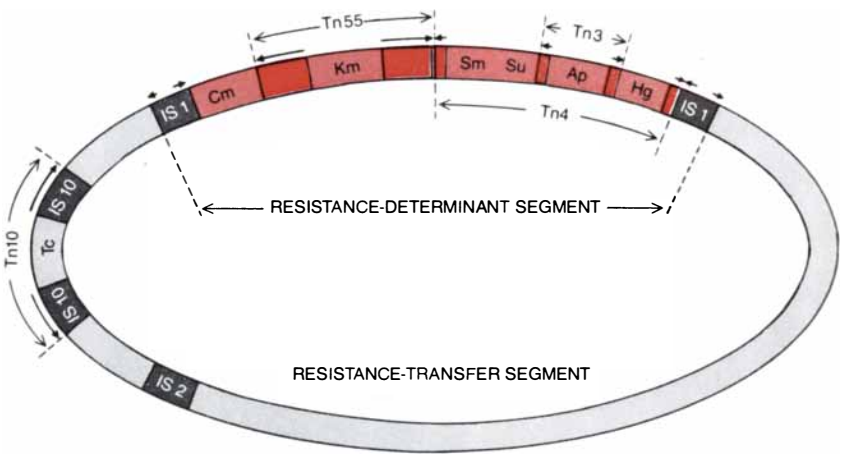


FUNCTIONAL COMPONENTS of the transposon Tn3 are diagrammed (not to scale). Genetic analysis shows there are at least four kinds of regions: the inverted-repeat termini; a gene for the enzyme beta-lactamase (*bla*), which confers resistance to ampicillin and related antibiotics; a gene encoding an enzyme required for transposition (a transposase), and a gene for a repressor protein that controls the transcription of the genes for transposase and for the repressor itself. The arrows indicate the direction in which DNA of various regions is transcribed.

The discovery of the process of transposition explains a puzzling phenomenon in bacterial evolution that has serious implications for public health: the rapid spread of antibiotic resistance among bacteria. Under the selective pressure of extensive administration of antibiotics in human and veterinary medicine and their use as a supplement in animal feeds, bacteria carrying resistance genes have a great natural advantage. For some time it has been known that resistance to several different antibiotics can be transmitted simultaneously to a new bacterial cell by a plasmid, but until transposition was discovered it was not known how a number of genes conferring resistance to different antibiotics were accumulated on a single plasmid-DNA molecule. The explanation seems to be that the resistance-determinant segments of drug-resistance plasmids have evolved as collections of transposons, each carrying a gene that confers resistance to one antibiotic or to several of them.

Work carried out at Stanford and by Phillip A. Sharp and others in the laboratory of Norman R. Davidson at the California Institute of Technology has made it clear that certain bacterial plasmids are constructed in a modular fashion. Plasmids isolated in different parts of the world show extensive sequence homology in certain of their DNA segments, whereas in other segments there is no structural similarity at all. In some instances plasmids can dissociate reversibly at specific sites. Transposable IS elements are found both at these sites and at sites where the plasmid interacts with chromosomal DNA to promote the transfer of chromosomes between different bacterial cells.

Identical transposons are commonly found in bacterial species that exchange genes with one another in nature. In addition antibiotic-resistance transposons



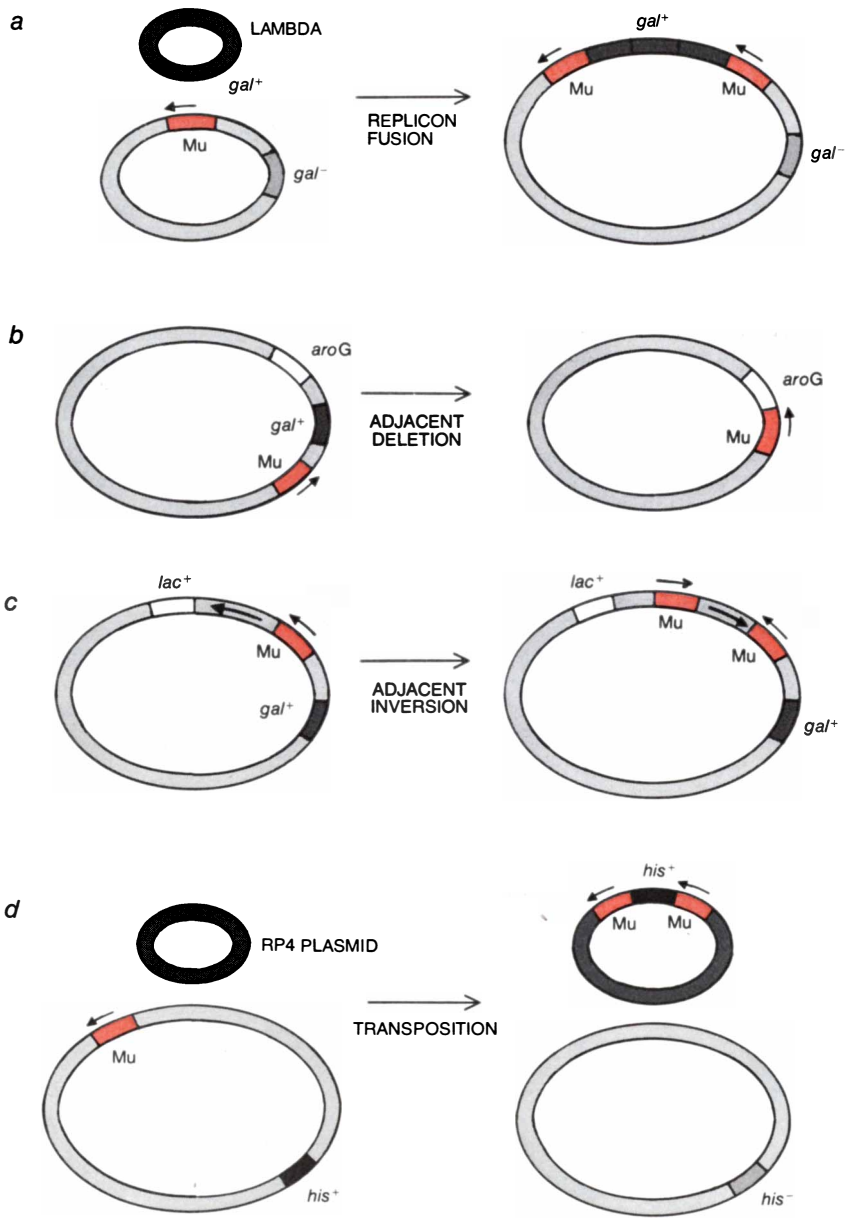
ROLE OF TRANSPOSABLE ELEMENTS in the evolution of antibiotic-resistance plasmids is illustrated by a schematic map of a plasmid carrying many resistance genes. The plasmid appears to have been formed by the joining of a resistance-determinant segment and a resistance-transfer segment; there are insertion elements (*IS1*) at the junctions, where the two segments sometimes dissociate reversibly. Genes encoding resistance to the antibiotics chloramphenicol (*Cm*), kanamycin (*Km*), streptomycin (*Sm*), sulfonamide (*Su*) and ampicillin (*Ap*) and to mercury (*Hg*) are clustered on the resistance-determinant segment, which consists of multiple transposable elements; inverted-repeat termini are designated by arrows pointing outward from the element. A transposon encoding resistance to tetracycline (*Tc*) is on the resistance-transfer segment. Transposon Tn3 is within Tn4. Each transposon can be transferred independently.

appear to be able to move among very different bacterial species that have not previously been known to exchange genes. For example, DNA sequences identical with part of Tn3 have recently been found to be responsible for penicillin resistance in two bacterial species unrelated to those commonly harboring Tn3 and in which such resistance had not previously been observed. Transposable elements seem, in other words, to accomplish in nature gene manipulations akin to the laboratory manipulations that have been called genetic engineering.

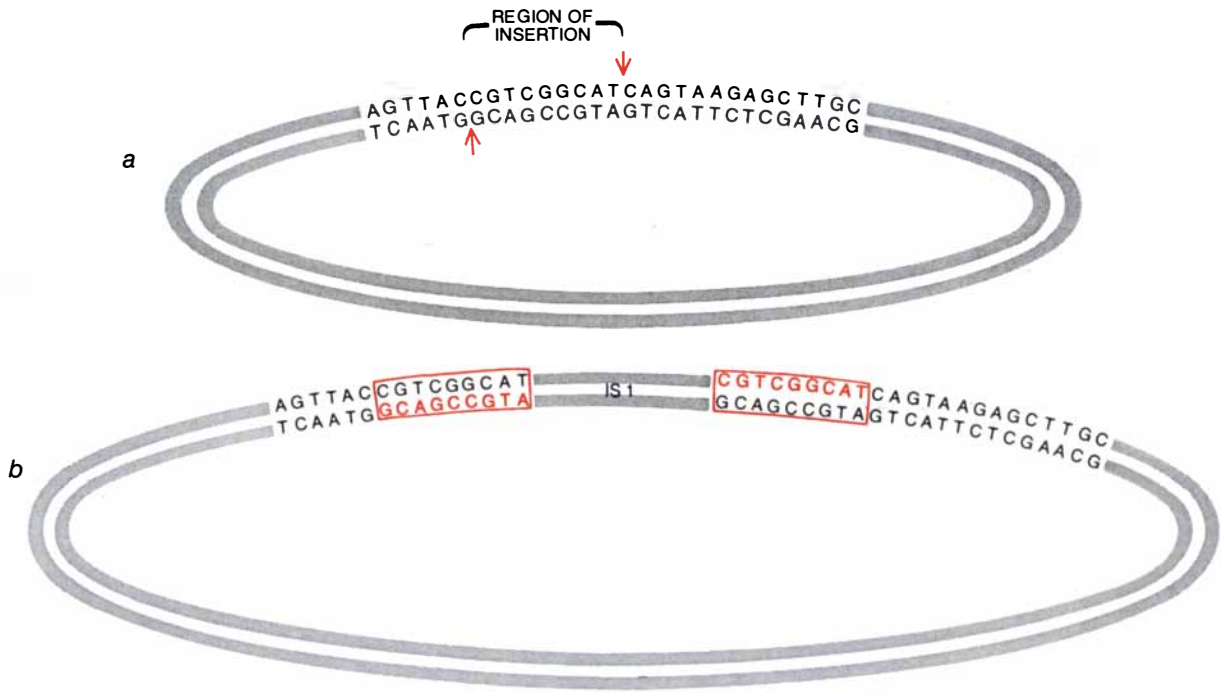
The effects of transposable genetic elements extend beyond their ability to join together unrelated DNA segments and move genes around among such segments. These elements can also promote both the rearrangement of genetic information on chromosomes and the deletion of genetic material. An awareness of these effects has emerged most clearly from studies of a peculiar phage, or bacterial virus, discovered in 1963 by Austin L. Taylor of the University of Colorado. Like other "temperate" bacterial viruses Taylor's phage could insert its DNA into a bacterial chromosome, creating a latent "prophage" that coexists with the bacterial cell and is transmitted to the bacterial progeny when the cell divides. Unlike other temperate phages, however, this one could become inserted at multiple sites within the chromosome, thereby causing many different kinds of mutation in the host bacterium. Because of this property Taylor called his phage Mu, for "mutator."

Further studies have shown that Mu is actually a transposable element that can also exist as an infectious virus. In the virus particle the Mu DNA is sandwiched between two short segments of bacterial DNA it has picked up from a bacterial chromosome. When the Mu virus infects a new cell, it sheds the old bacterial DNA and is transposed to a site in the new host chromosome. Ahmad I. Bukhari and his colleagues at the Cold Spring Harbor Laboratory have shown that Mu's ability to replicate is closely associated with its ability to be transposed; the virus has apparently evolved in such a way that its life span is dependent on transposition events.

While the structure of Mu's DNA and the details of the phage's life cycle were being unraveled Michel Faelen and Arianne Toussaint of the University of Brussels were doing genetic experiments aimed at understanding how the Mu DNA interacts with other DNA in a bacterial cell. The results of experiments carried out over a period of almost 10 years have demonstrated that Mu can catalyze a remarkable series of chromosome rearrangements. These include the fusion of two separate and independently replicating DNA molecules ("replicons"), the transposition of segments of the bacterial chromosome to plasmids, the deletion of DNA and the inversion of segments of the chromosome. Significantly, all these rearrangements seem to involve the nucleotide sequences at the ends of Mu DNA and to require the expression of a Mu gene that had been found earlier to be necessary both for transposition and for virus replication. Experiments done by Hans-Jorg Reif and Saedler at the University of Freiburg and by other groups have shown that many other transposable elements can, like Mu, promote the deletion of DNA; Nancy E. Kleckner and David Botstein and their associates at Harvard and at the Massachusetts Institute of Technology have shown that such elements can also bring about the inversion of DNA sequences. Indeed, there is evidence that some transposable elements may participate in specific re-



CHROMOSOME REARRANGEMENTS mediated by the bacterial virus Mu include replicon fusion, adjacent deletion, adjacent inversion and transposition of chromosome segments to a plasmid. Mu is shown in color; a small arrow gives its orientation. In a cell lysogenic for Mu (having Mu DNA integrated in its chromosome) and containing DNA of a lambda *gal*+ virus (*a*) the viral DNA becomes integrated into the chromosome between two copies of Mu. In a lysogenic cell in which Mu is near integrated *gal*+ genes (*b*) the *gal*+ genes are deleted. In a lysogenic male bacterium (*c*) with Mu near the origin of chromosome transfer (*large arrow*) the origin becomes inverted between two oppositely oriented copies of Mu. In a lysogenic bacterium carrying a plasmid (*d*) a bacterial *his*+ gene is transposed to plasmid between copies of Mu.



DUPLICATION of five, nine or 11 pairs of nucleotides in the recipient DNA is associated with the insertion of a transposable element; the two copies bracket the inserted element. Here the duplication that attends the insertion of IS1 is illustrated in a way that indicates how the duplication may come about. IS1 insertion causes a nine-nucleotide duplication. If the two strands of the recipient DNA are cleaved (*colored arrows*) at staggered sites that are nine nucleotides apart (*a*), then the subsequent filling in of single strands on each side of the newly inserted element (*b*) with the right complementary nucleotides (*color*) could account for the duplicated sequences (*colored boxes*).

arrangements of DNA more frequently than they do in transposition events.

New methods for determining DNA nucleotide sequences rapidly and simply have provided an important tool for elucidating the structure of transposable elements as well as the biochemical mechanisms involved in transposition and in chromosome rearrangements. The sequence of a transposable element (IS1) has been determined in its entirety by Hisako Ohtsubo and Eiichi Ohtsubo of the State University of New York at Stony Brook. An important insight into the mechanism of transposition has resulted from DNA-sequence observations initially made by Nigel Grindley of Yale University and the University of Pittsburgh and by Michele Calos and Lorraine Johnsrud of Harvard, working with Jeffrey Miller of the University of Geneva. Both groups examined the DNA sequences at the sites of several independently occurring insertions of the IS1 element. They found that the insertion of IS1 results in the duplication of a sequence of nine nucleotide pairs in the recipient DNA. The duplicated sequences bracket the insertion element and are immediately adjacent to its inverted-repeat ends. Since the sequence of the recipient DNA was different at each of the various insertion sites studied, different nucleotides were duplicated for each insertion.

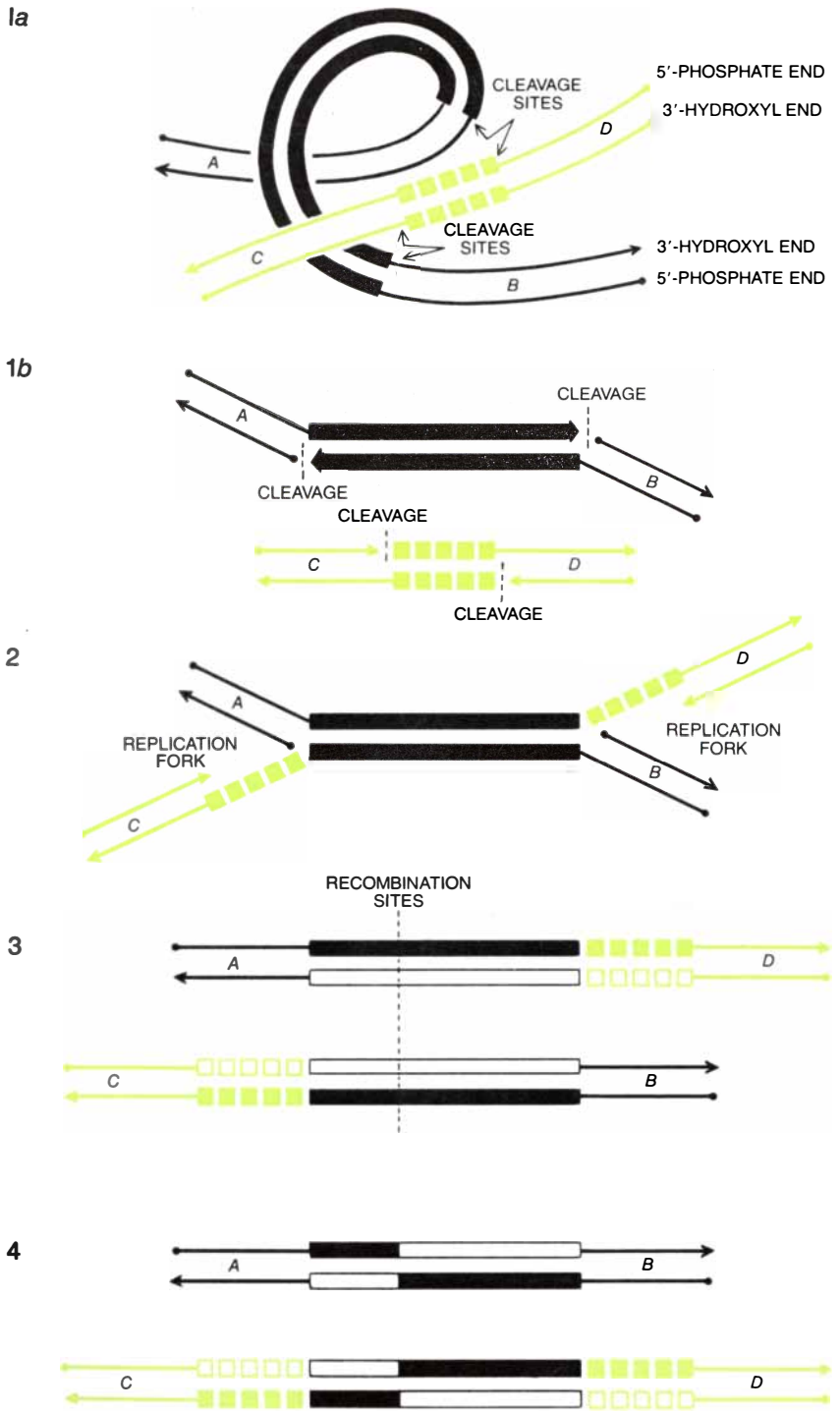
Subsequent reports from many laboratories have shown that similar duplications of a short DNA sequence result from the insertion of other transposable elements. Some elements generate nine-nucleotide duplications and others generate duplications five or 11 nucleotides long. As Calos and her colleagues and Grindley have pointed out, these observations suggest that a step in the insertion process involves staggered cleavage (at positions five, nine or 11 nucleotides apart) of opposite DNA strands at the target site for transposition. The filling in of the single-strand segments following such cleavage would require the synthesis of short single-strand stretches of complementary DNA and would result in the nucleotide-sequence duplication. Faelen and Toussaint had also concluded that DNA synthesis is required in the generation of chromosome rearrangements by Mu: they had noted that the rearranged bacterial chromosome often included two copies of the prophage, the inserted form of Mu.

On the basis of these observations one of us (Shapiro) has proposed a model to explain transposition, chromosome rearrangements and the replication of transposable elements such as phage Mu as variations of a single biochemical pathway. The pathway is such that transposable elements can serve two functions in the structural reorganization of cellular DNA: they specifically duplicate themselves while remaining inserted in the bacterial chromosome, and they bring together unrelated chromosomal-DNA segments to form a variety of structural rearrangements, including fusions, deletions, inversions and transpositions.

If this model is at all close to reality, then the nonhomologous recombination events associated with transposable elements are rather different from other types of illegitimate recombination, such as the integration of the phage-lambda DNA into the bacterial chromosome, that do not involve DNA synthesis. It seems likely that bacterial cells will turn out to have several different systems for carrying out nonhomologous recombination, just as they have multiple pathways for homologous recombination.

The potential for multiple mechanisms of illegitimate recombination is important to bear in mind when comparing phenomena that appear to be similar in bacteria and higher cells. Transposition phenomena that are analogous genetically may not be similar biochemically. There is some genetic evidence indicating that the movement of controlling elements in maize from one chromosomal site to another may be brought about by a mechanism different from that of transposition in bacteria.

Genetic rearrangements can have biological importance on two time scales: on an evolutionary scale, where the effects of the rearrangement are seen



POSSIBLE MOLECULAR PATHWAY is suggested to explain transposition and chromosome rearrangements. The donor DNA, including the transposon (*thick bars*), is in black, the recipient DNA in color. Arrowheads indicate the 3′-hydroxyl ends of DNA chains, dots the 5′-phosphate ends; the letters *A, B, C* and *D* identify segments of the two DNA molecules. The pathway has four steps, beginning with single-strand cleavage (*1a*) at each end of the transposable element and at each end of the "target" nucleotide sequence (*colored squares*) that will be duplicated. The cleavages expose (*1b*) the chemical groups involved in the next step: the joining of DNA strands from donor and recipient molecules in such a way that the double-strand transposable element has a DNA-replication fork at each end (*2*). DNA synthesis (*3*) replicates the transposon (*open bars*) and the target sequence (*open squares*), accounting for the observed duplication. This step forms two new complete double-strand molecules; each copy of the transposable element joins a segment of the donor molecule and a segment of the recipient molecule. (The copies of the element serve as linkers for the recombination of two unrelated DNA molecules.) In the final step (*4*) reciprocal recombination between copies of the transposable element inserts the element at a new genetic site and regenerates the donor molecule. The mechanism of this recombination is not known; it does not require proteins needed for homologous recombination, and at least in Tn3 it is mediated by sequences within element.

after many generations, and on a developmental time scale, where the effects are apparent within a single generation. It is known that transposable genetic elements can serve as biological switches, turning genes on or off as a consequence of their insertion at specific locations. In some instances the insertion of an IS element in one orientation turns off nearby genes, whereas an unexpressed gene can be turned on when the element is inserted in the opposite orientation.

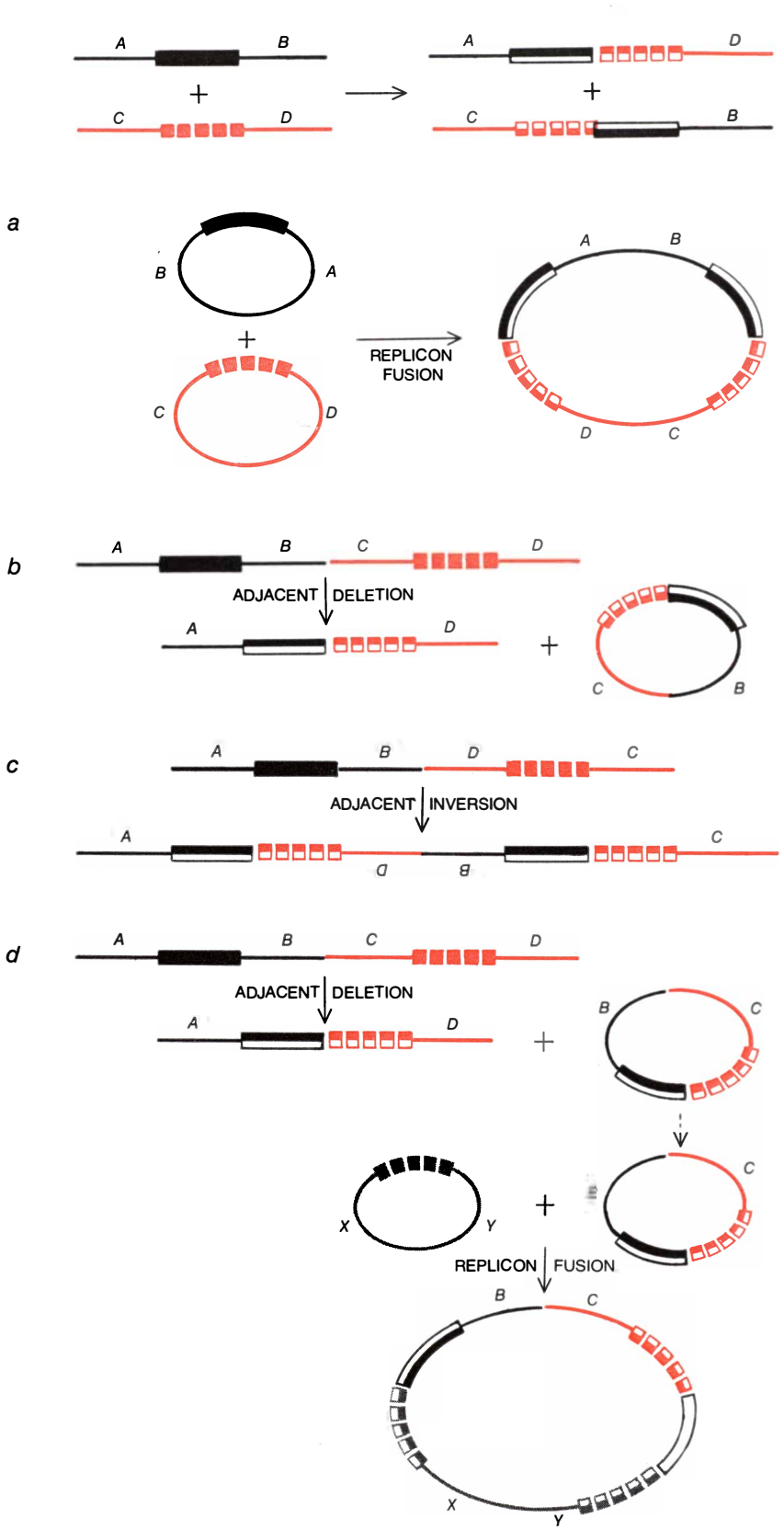
An analogous regulation of gene expression through chromosome rearrangement is "phase variation," which is seen in certain disease-producing bacteria that can invade the gastrointestinal tract. The phase, or immunological specificity, of a hairlike flagellum on these bacteria can change suddenly within a single bacterial generation. Melvin Simon of the University of California at San Diego and his colleagues have recently shown that the choice between the expression of one *Salmonella* flagellum gene and the expression of its counterpart, which specifies a different phase, is controlled by the inversion of a particular segment of the bacterial chromosome. The inversion takes place in the absence of proteins needed for homologous recombination, and so it appears to depend on recombination enzymes that recognize the ends of the invertible segment. Whether the switching mechanism responsible for phase variation operates by a molecular process similar to transposition remains to be determined, but the process clearly falls within the category of recombination events that were considered "illegitimate" a few years ago.

Although molecular studies on transposable elements have so far been carried out primarily in bacteria, there has been extensive genetic evidence for the existence of similar elements in higher organisms for years. The pioneering work of Barbara McClintock not only established the existence of transposable genetic elements in the corn plant but also showed by genetic analysis that the movement of a controlling element from one site to another in the maize chromosome depends on the action of genes on certain of the elements themselves, genes presumably analogous to those encoding the transposases of Tn elements and of phage Mu. McClintock also showed that some controlling elements (called regulators) regulate the expression of distant genes carrying insertions of other controlling elements (called receptors). Groups of genes are expressed synchronously at specific times during plant development, and McClintock suggested that the transposition of receptor elements could provide a mechanism for the rapid evolution of control mechanisms in situations in which several genes must be switched on or off at the same time, as they are in the course of development.

As often happens in science, the significance of McClintock's work was not entirely understood or appreciated until later studies carried out with the much simpler bacterial systems provided actual physical evidence for the existence of insertion sequences and transposons as discrete DNA segments and also established that transposition is brought about by a mechanism different from previously understood recombinational processes. Numerous other examples of transposable elements have now been recognized in higher organisms, such as the fruit fly *Drosophila* and the yeast *Saccharomyces cerevisiae.* The possible role of these elements in the generation of chromosome rearrangements is being actively investigated. Recent work on the control of immunoglobulin synthesis in mice by Susumu Tonegawa and his associates at the Basel Institute for Immunology has shown that the ability of mammalian cells to produce specific antibody molecules in response to injected foreign proteins also involves chromosome rearrangements. There is little doubt that additional instances will soon be found in which illegitimate recombination events play a major role in the expression of genes during cellular differentiation.

Even in bacteria much remains to be learned about the basic molecular mechanisms that accomplish the transposition of genetic elements and the associated rearrangement of DNA molecules. The various biochemical steps in the transposition pathway need to be more fully defined. What is the mechanism for recognition of the inverted-repeat ends of transposable elements? What proteins other than those encoded by the transposon play a role in transposition? What are the additional genetic aspects of the regulation of transposition? In a broader sense, what is the role of illegitimate recombination in the organization and expression of genes, not only in bacteria but also in higher organisms? Although the mechanisms that have been studied in bacteria provide a working model for the mechanisms of similar events in higher organisms, the parallels are probably incomplete.

It is already clear that the joining of structurally and ancestrally dissimilar DNA segments by transposable elements is of great importance for the production of genetic diversity and the evolution of biological systems. The discovery of such a fundamentally different recombinational process at a time when many molecular biologists believed virtually all the important aspects of bacterial genetics were understood in principle—with only the details of particular instances remaining to be learned—leads one to wonder whether still other fundamentally new and significant basic biological processes remain to be discovered.



FIRST THREE STEPS OF PATHWAY are summarized schematically at the top of this illustration. These steps achieve reciprocal recombination between unrelated DNA molecules and explain all rearrangements shown in the illustration on page 46, as follows. If the donor and recipient molecules are circular, the three steps result in replicon fusion (*a*). If the donor and recipient regions are part of a single molecule, the steps generate an adjacent deletion (*b*) or an adjacent inversion (*c*), depending on the positions of regions *A, B, C* and *D*. Two successive events (deletion and then replicon fusion) can result in the transposition to a plasmid of a DNA segment adjacent to the transposable element, along with two copies of the element (*d*).