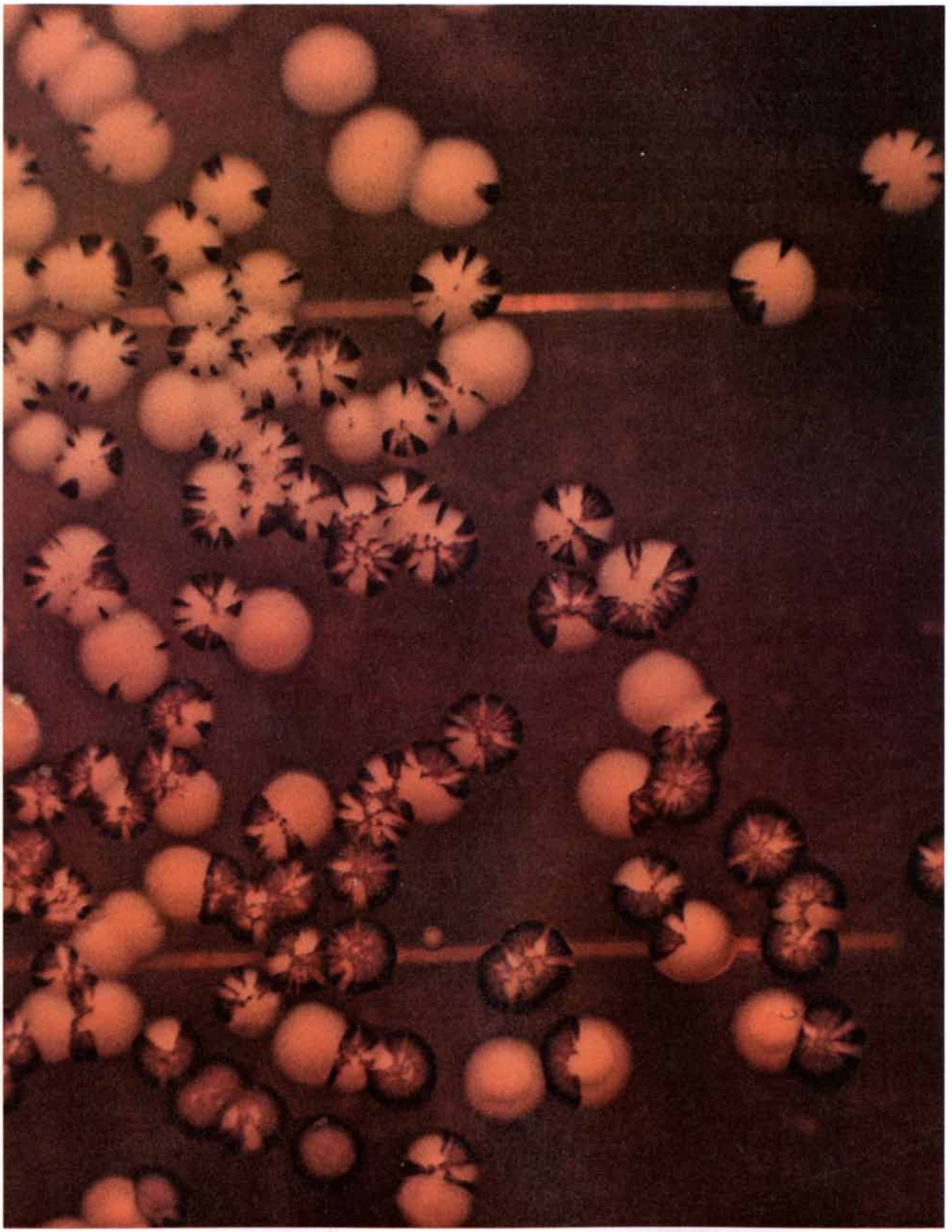


PRESENCE OF PLASMIDS in bacteria and their autonomous behavior are revealed in a photograph, made by the author, of colonies of *Staphylococcus aureus* growing on a temperature gradient: from 30 degrees Celsius at the bottom to 40 degrees at the top. Most of the cells originally plated on the agar contain a plasmid with a gene coding for penicillinase. Cells that have the plasmid and produce penicillinase stain purple; those without the plasmid, lacking penicillinase, stain orange. The plasmid's replication (but not the host cell's) is temperature-sensitive, so that the higher the temperature is, the more often cells lacking the plasmid and penicillinase are generated. When a plasmid is lost from a cell, the cell's progeny stain orange rather than purple, giving rise to colonies with radial purple and orange sectors.

Plasmids

These accessory genetic elements in bacteria, best known as carriers of resistance to antibiotics and as vehicles for genetic engineering, are actually subcellular organisms poised on the threshold of life

by Richard P. Novick

The living cell is the only natural environment within which the intricate chemistry of life can proceed. As such it is the fundamental unit of biological organization for multicellular organisms as well as for unicellular ones. It is also an ecological niche occupied by a variety of subcellular, submicroscopic organisms—organisms whose structures and reproductive dynamics are so profoundly different from those of cellular forms that they should probably be assigned to a new taxonomic kingdom.

The best-known of these curious life forms are the viruses: elegantly symmetrical particulate structures composed essentially of a molecule of nucleic acid encapsulated in a protein coat. They can exist outside the cell, but there they are inert. Once inserted into a cell, however, their nucleic acid (RNA or DNA) reprograms the cell's metabolic apparatus to the service of the virus; the nucleic acid replicates and is encapsulated, in the process destroying the cell and releasing a new crop of virus particles.

There is a less familiar class of subcellular organisms: the plasmids, found for the most part in bacteria. Plasmids are simpler in organization than viruses. They have no protein coat and no extracellular phase; they are no more than circular molecules of double-strand DNA that multiply independently within host cells and are inherited in a regular manner as those cells proliferate. For a time after their discovery about 30 years ago plasmids were thought to be something of an oddity. Now they are known to be extremely common. Plasmids are found in virtually all bacterial species; they are probably present in most individual bacterial cells. They are not essential to the everyday metabolism of the host cell, and so it is almost always possible to isolate a variant of any plasmid-carrying strain that no longer has any plasmids. The loss is permanent: the cell cannot regenerate a new plasmid but can only acquire it from another bacterium. The information encoded in a plasmid (or in any other DNA molecule) can be developed only through the prolonged and tortuous process of organic evolution.

Plasmids account for only a small part of a cell's genome: typically between a fraction of 1 percent and 2 or 3 percent. That small fraction of the cell's hereditary information, however, codes for important accessory genetic traits that are not ordinarily encoded by the bacterial chromosome. Plasmids alone carry the information for conjugation, or "mating," between bacteria. They are uniquely responsible for several diseases of plants and animals. They enable their bacterial hosts to utilize many complex substances as nutrients. And they confer on their host cells the ability to resist a wide variety of toxic agents, including antibiotics.

It was the importance of these genetic traits, and in particular the clinical importance of antibiotic resistance, that first drew the close attention of biologists to plasmids and led to some understanding of their role in bacterial genetics. In the past five years plasmids have been intensively investigated, dissected and manipulated for another reason: they were found to be ideal vehicles for introducing nonbacterial genes into bacteria by the techniques of molecular cloning and thereby purifying and greatly amplifying the foreign genes. The molecular cloning of DNA has already revolutionized genetic analysis and, because the bacteria can sometimes be induced to manufacture nonbacterial proteins, is the basis of a new "applied genetics" with great promise for medicine and industry.

Plasmids as Organisms

In a broader perspective plasmids are significant because they force biologists to reappraise the essential nature of living things. Plasmids are poised at the very threshold of life, between the inanimate and the animate. Are they merely bits of DNA that are integral parts of the cell's genome, contributing a little extra genetic information? Or are they independently evolving subcellular organisms more or less on a par with the viruses? The second view is strongly supported by a number of significant observations.

In the first place, each plasmid autonomously controls its own "copy number": the number of replicas of the plasmid present in a host cell. Moreover, conjugation between bacterial cells can lead to the exchange of plasmids between different species, and even genera, that are quite unable to exchange chromosomal genes. Finally and most significant, this exchange can result in the transfer from one species to a competing species of a plasmid carrying genes that enable the recipient to survive at the expense of the donor. In other words, the plasmids have evolved the ability to survive regardless of the fate of their host species—something that would be inconceivable, within the framework of evolution through natural selection, for an element that was merely a component of a particular organism's genome.

These observations imply that plasmids are indeed independent organisms, members of a hierarchy of subcellular life. Within that hierarchy animal viruses such as those that cause poliomyelitis or yellow fever and bacteriophages, or bacterial viruses, such as T4 can be regarded as predators because they always kill the cells they infect. "Temperate" phages and certain animal viruses, which sometimes kill the host cell but also sometimes coexist stably within the cell without apparently harming or benefiting it, can be considered subcellular parasites. Plasmids, then, can be regarded as endosymbionts: they exist stably within the host cell and often perform a service by supplying the host with adaptive genetic functions, the quid pro quo of the symbiotic state.

The notion that naked molecules of DNA are organisms cannot be squared with the conventional cell-based concept of life. What is common to all living things is that they respond independently to evolutionary forces and duplicate their genetic material autonomously. Any nucleic acid system that controls its own replication should therefore be regarded as an organism. An examination

of the details of the molecular processes that determine the life cycle of plasmids should advance understanding of the very essence of the living state and hence illuminate the borderline between the animate and the inanimate.

The Discovery of Plasmids

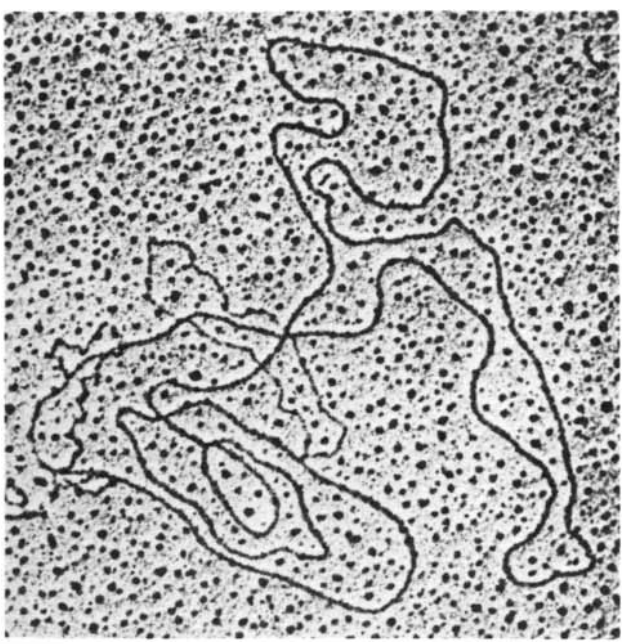
In the early 1950's, soon after the discovery of conjugation in the bacterium *Escherichia coli* by Joshua Lederberg, it became clear that there were two genetically determined "mating" types and that genetic information was physically transferred from a "male," or donor, type to a "female," or recipient. The genetic character of "maleness"—the F (for fertility) factor—was transferred far more often than any other trait. Moreover, the F factor was transferred independently of any other known donor genes: it was not "linked" to any of them. In the case of a eukaryotic cell (a higher cell with a nucleus and with its DNA organized in discrete chromosomes visible under a light microscope) such a finding would have meant that the trait was encoded by an extrachromosomal gene. Lederberg saw that the F factor was somehow similar to extrachromosomal genetic elements in the cytoplasm of higher organisms, and in 1952 he coined the word plasmid to refer to all such extrachromosomal genetic systems. The structural organization of the genome in bacteria was not well understood at that time, and so the precise significance of the absence of linkage was not entirely clear.

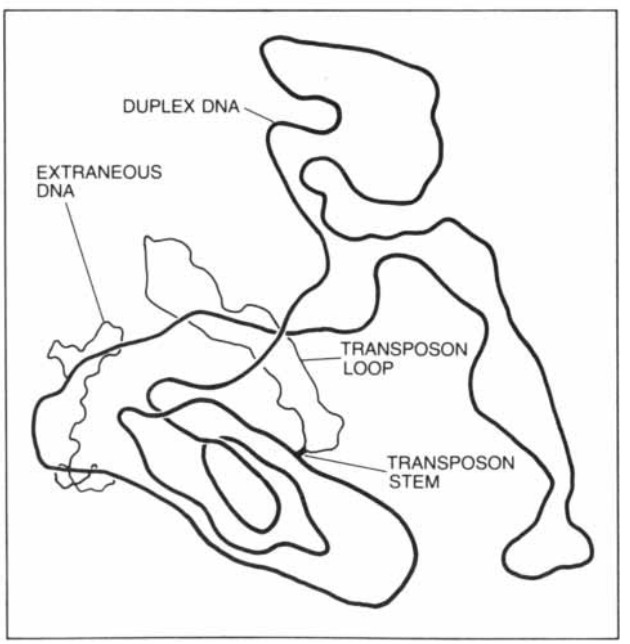
It was soon clarified, however, by further investigation of the F factor. In 1953 William Hayes, who was then working at Hammersmith Hospital in London, discovered that under certain circumstances the F factor could in fact become linked to standard markers and could thereupon bring about their sequential transfer during conjugation. It developed that most of the genes of *E. coli* are arrayed along a single continuous structure that behaves in formal genetic terms like a closed loop—the bacterial analogue of the eukaryotic chromosome. The F factor is physically attached to this structure at a specific site; the loop opens at the site of F attachment during conjugation and passes as a linear structure to the recipient cell. It was also observed that the attached F factor could occasionally become detached, sometimes entraining a segment of the loop near its attachment site. Given a bacterial chromosome, it was now possible to consider the F factor to be a specifically extrachromosomal element that sometimes became integrated into the chromosome.

At the Pasteur Institute in Paris François Jacob and Elie L. Wollman noted similarities between the behavior of the F factor and that of a temperate bacterial virus, phage lambda, and another plasmid, ColE1 (which encodes a colicin, a protein that kills *E. coli*). They coined a new term, episome, for a genetic element that could be replicated in either of two states: attached to a host-cell chromosome or free in the cytoplasm.

For a time phage lambda, the F factor, ColE1 and one or two other elements remained the only known episomes. Then in about 1959 a remarkable discovery was made in Japan in the course of a study of patients with bacterial dysentery who did not respond to treatment with ordinarily efficacious antibiotics. The causative bacteria, *Shigella dysenteriae,* contained genes that made them resistant to several antibiotics at once, and the resistance genes were transferred to other intestinal bacteria in much the same way as the F factor [see "Infectious Drug Resistance," by Tsutomu Watanabe; SCIENTIFIC AMERICAN, December, 1967]. These R (for resistance) factors, as they came to be called, had several properties in common with the F factor, including the ability to promote their own intercell transfer by conjugation; some have since been shown to have partial DNA sequences in common with F. Whereas the F factor had been considered a rather special entity, genetically very interesting but not necessarily widespread, the discovery of R factors made it suddenly clear that extrachromosomal genetic elements were important components of the bacterial microcosm, with clinical significance as well as scientific.

Soon after the discovery of R factors in the enteric bacteria I found similar factors in staphylococci. They carried a gene for an extremely potent enzyme, penicillin beta-lactamase (penicillinase), that broke down penicillin, and they were responsible for the resistance to penicillin that was compromising treatment of "hospital staph" in many parts of the world in the early 1960's. The staphylococcal R factors are apparent-





PLASMID HETERODUPLEX is seen in an electron micrograph made by Marjorie M. Schwesinger in the author's laboratory. Plasmids pI258 and pI6187, identical except that the former carries a transposon specifying resistance to erythromycin, were denatured and the separated strands were allowed to reanneal. The heteroduplex has one strand from each of the plasmids. The two strands match perfectly except for the sequence constituting the transposon. As the drawing indicates, the two strands therefore form a double-strand molecule with a single-strand loop: the transposon, which is attached by a short double-strand stem formed by its "inverted repeat" termini.

ly unable to promote their own transfer by conjugation. They are transferred passively by transduction: a process in which a bit of bacterial DNA is encapsulated in a phage particle (in place of the phage's own DNA) and then introduced into another cell infected by the phage. A substantial proportion of the R factors of enteric bacteria are similarly unable to promote their own conjugative transfer and rely instead on transduction or on mobilization by another, conjugative plasmid that is present in the same cell.

By the mid-1960's it was apparent that most of the R factors in staphylococci and the enteric bacteria (and ColE1 too) were different from the F factor and lambda in that they remained extrachromosomal instead of undergoing reversible integration into the chromosome, and so they did not fit the restrictive definition of an episome. It seemed to me the requirement of reversible integration made it impossible to group together a wide range of clearly related extrachromosomal genetic systems, and in 1963 I suggested that Lederberg's more general term was more appropriate; other workers in the field agreed, and gradually over the next 15 years these mobile extrachromosomal genetic elements in bacteria came to be generally known as plasmids rather than episomes.

Gradually too it became evident that plasmids are responsible for a remarkably wide range of biological activities in bacteria. Appreciation of the role of plasmids (and of some other extrachromosomal genetic systems that have come to light even more recently) has produced a rather dramatic shift in biologists' thinking about genetics. The traditional view was that the genetic makeup of a species was about the same from one cell to another and was constant over long periods of time. Now a significant proportion of genetic traits, not only in bacteria but also in higher organisms, are known to be variable (present in some individual cells or strains, absent in others), labile (subject to frequent loss or gain) and mobile (transferable between individual cells or transposable from one site to another in a cell)—all because those traits are associated with plasmids and other atypical genetic systems.

The Variety of Plasmids

In 1964 Mark Richmond and Eric Johnston of the National Institute for Medical Research in England showed that the staphylococci responsible for an outbreak of postsurgical "suture line" infections carried a plasmid enabling them not only to destroy penicillin but also to grow in the mercury-based disinfectant used to sterilize the sutures. In my laboratory at the Public Health Research Institute of the City of New York, Christine Roth and I found that our staph plasmids carried genes for resistance to penicillin and mercury compounds and also to a variety of other heavy-metal compounds lethal to staphylococci: salts of cadmium, bismuth, lead and antimony and the ions arsenate and arsenite. Various heavy-metal resistance genes have since been found on some *E. coli* R plasmids.

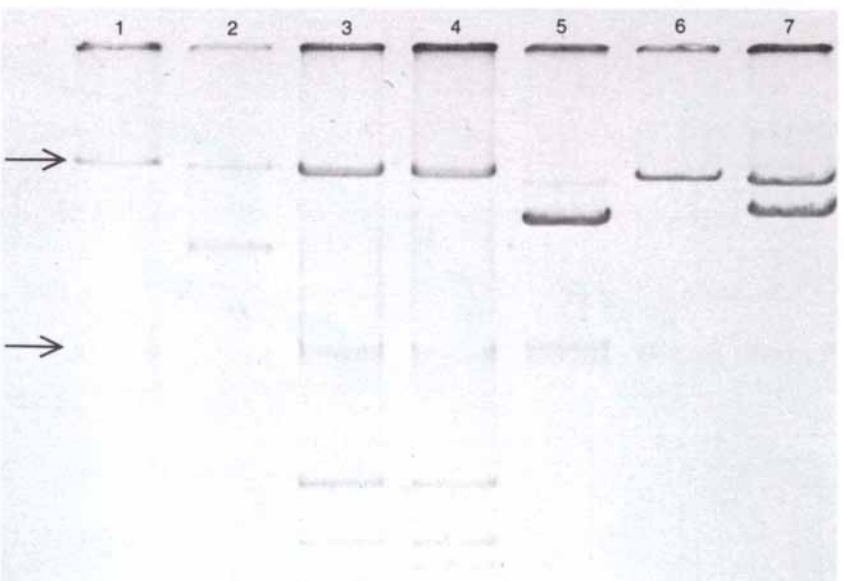
Many other kinds of specialty genes are carried by plasmids rather than by the bacterial chromosome, including the genes for some human diseases such as traveler's diarrhea and staphylococcal impetigo, for the fermentation of milk by lactic acid bacteria to produce cheese and for much of the metabolic versatility of the pseudomonads, soil bacteria that can degrade a variety of complex hydrocarbons (and may have applications such as cleaning up oil spills and converting petroleum into edible proteins). Plasmids give rise to a curious gas vacuole in salt-loving bacteria; they are thought to control the synthesis by *Bacillus thuringiensis* of an insecticide widely applied to control the gypsy moth and tent worms.

Apart from the rather special nature of many of their functions, plasmid genes are not basically different from other genes. For the most part they encode proteins that are active enzymes, and these proteins are synthesized by the same metabolic machinery—supplied by the cell—as other proteins.

The most intensively studied of the plasmid-derived behaviors has been resistance to antibiotics and other substances toxic to bacteria. It was the first widespread plasmid-carried property to be encountered; it presents a grave medical problem; it is a convenient selective trait for experiments in bacterial genetics and in molecular cloning, and analyzing it sheds light on the origin of plasmids and on how they accumulate a variety of genes.

Resistance to Antibiotics

At least three different strategies have evolved by which plasmid-associated genes resist toxic substances, including those we know as antibiotics. The simplest is outright destruction of the antibiotic. This is the case for resistance to penicillin, chloramphenicol, aminoglycoside antibiotics such as streptomycin, and mercury compounds. Another strategy involves blocking access of the toxic substance to its particular target in the cell. The macrolide antibiotics, such as erythromycin, act by binding to the cellular organelles called ribosomes and thereby blocking the synthesis of proteins; Bernard Weisblum and his colleagues at the University of Wisconsin Medical School have shown that the plasmid-borne resistance genes attach a methyl group to the ribosomal RNA and so prevent the binding of the antibiotic. Tetracycline is apparently kept out



GEL ELECTROPHORESIS reveals the presence of plasmids in bacteria. Daniel A. Portnoy and Stanley Falkow of the University of Washington at Seattle prepared cell-free extracts of seven strains of the bacterium *Escherichia coli,* eliminated most of the chromosomal DNA and placed samples in slots at one end of an agarose gel slab. Application of an electric current caused the negatively charged DNA molecules to migrate toward the positive electrode at a rate inversely proportional to the logarithm of their size. The DNA, labeled with a fluorescent dye, is visualized as a set of linear bands. The top band represents material still in the origin wells. The next band (*top arrow*), representing a plasmid responsible for traveler's diarrhea, is common to all samples except No. 5—a strain that does not cause the disease. The other distinct bands correspond to various plasmids, some of which specify antibiotic resistance. The faint, more diffuse bands (*bottom arrow*) correspond to segments of residual chromosomal DNA.

Our widely accepted *UNIX*™ operating system is simplifying the use of computers. Designed to handle a variety of applications, it is being used to manage and maintain the telecommunications network, control experiments, develop software, process text, prepare documents, and teach computer science.

The *UNIX* system allows users to take small programs and assemble them like building blocks to perform complex tasks. In text processing, for example, the command "Spell Bell Labs Ad" tells a computer to proofread this ad against a dictionary filed in its memory. The program that performs the task was created by simply combining several smaller *UNIX* programs.

Another useful feature of our *UNIX* operating system is its ease in typesetting mathematical expressions. To typeset an equation like

$$\sum_{i=1}^{\infty} \frac{1}{x_i} = \pi$$

someone need only type: "Sum from i = 1 to infinity 1 over x sub i = pi." The computer does everything else.

Since 1969 the Bell System has installed more than 1100 *UNIX* systems. Along with other Bell Labs innovations in computing technology, these systems are enhancing the way the nation's telecommunications network is planned, designed, built, and operated. Through licensing agreements with Western Electric, universities have installed over 800 *UNIX* systems, and government and industrial facilities are using over 400.

The *UNIX* operating system can be used with computers of different manufacturers because it is small, cleanly designed, and written in a general-purpose programming language. Such portability in a computer operating system saves time and money.

Building on the past

The *UNIX* system is just one of many Bell Labs advances in computing science and technology over the years. Among our innovations:

- Application of telephone switching concepts and technology to early computers
- First demonstration of remote computer terminal and data link
- Conception of electronic analog computer
- First design of AND and OR gates for diode circuitry
- Formulation of Information Theory
- Invention of error-detecting and error-correcting codes
- Demonstration of first general-purpose transistorized digital computer
- Development of computer operating systems
- Design of computer languages, including ALTRAN, SNOBOL, L6, and C
- Creation of computer graphics techniques for storing, manipulating, and presenting information
- Development of Fast Fourier Transform
- Design of central processors for switching systems having virtually no downtime

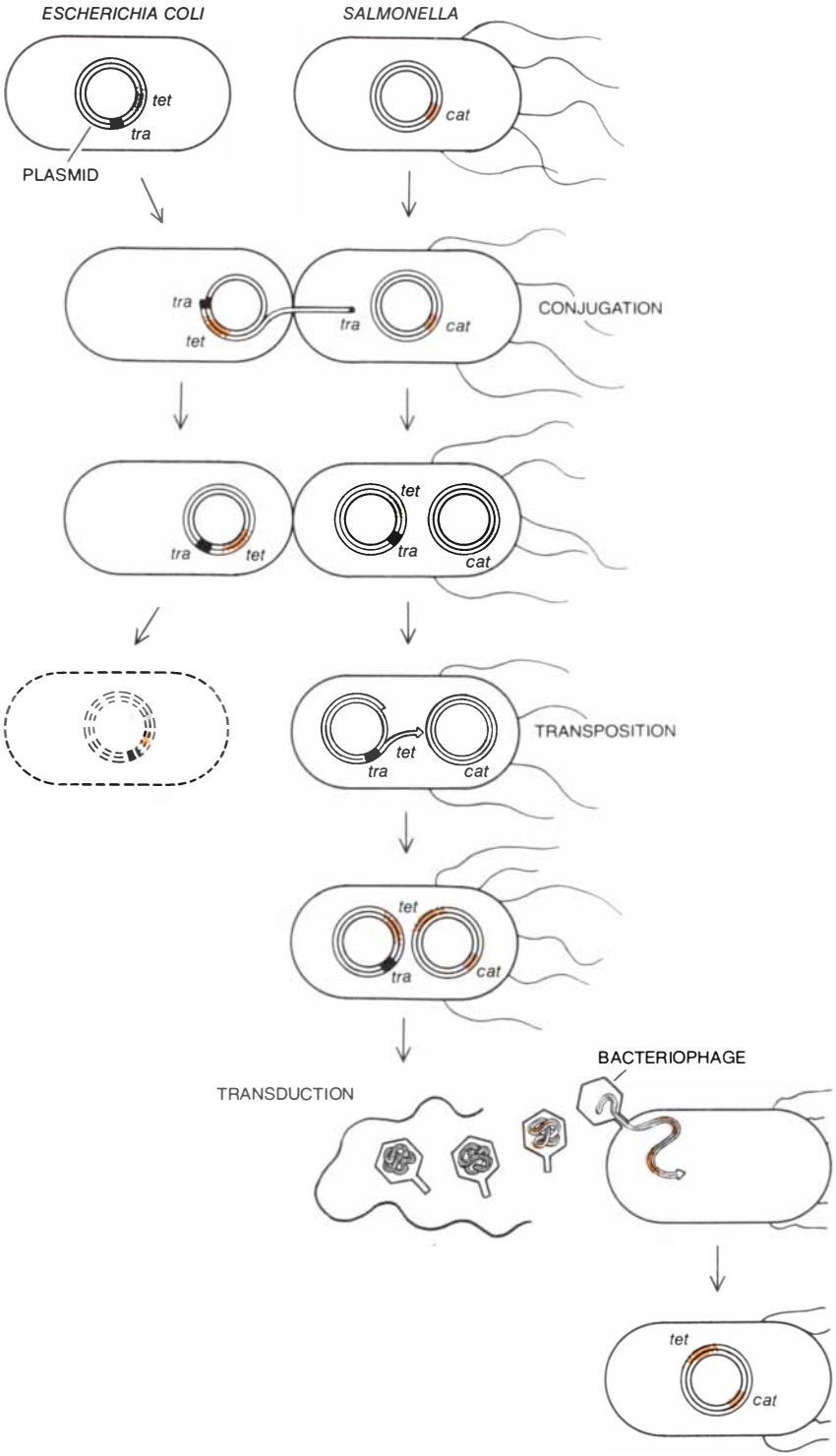
Looking ahead

Computing technology is having a major impact on the telecommunications business. It's increasing the Bell System's ability to provide new services and handle existing ones more efficiently. Today, for example, the nationwide telecommunications network links thousands of software-controlled electronic systems, making it by far the world's largest distributed processing network. We and our partners—Western Electric and the telephone companies of the Bell System—are putting technology to work so that the network will continue to evolve and expand its information-handling capabilities. The object, of course, is to meet the fast-growing and changing needs of Bell System customers.

For information about employment, write:
Bell Laboratories, Room 3C-303, 600 Mountain Avenue, Murray Hill, N.J. 07974
Please include a resume.
An equal opportunity employer.

Bell Laboratories

From Science: Service



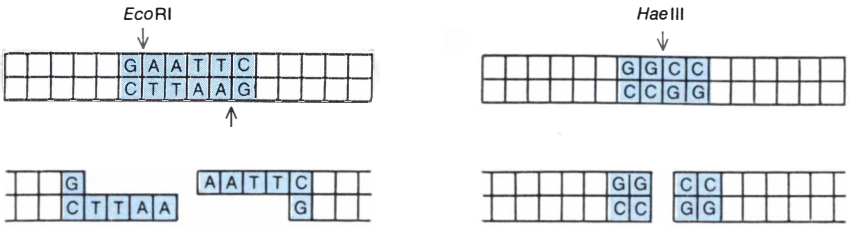
GENES FOR RESISTANCE to antibiotics are collected by plasmids and are transferred from one bacterial cell to another by various mechanisms, sometimes enabling a plasmid to survive the death of its host. Here two bacteria are depicted (*top*): an *Escherichia coli* cell containing a plasmid with genes for transmission by conjugation (*tra*) and for tetracycline resistance (*tet*) and a *Salmonella* cell with a plasmid carrying a gene for resistance to chloramphenicol (*cat*). The two cells conjugate and the *tet*-carrying plasmid is transferred to the salmonella, rendering it resistant to tetracycline as well as to chloramphenicol. In an environment containing both antibiotics the *E. coli* die, but their plasmid survives in the successful host. The *tet* gene is on a transposon that subsequently moves from one plasmid to the other, which then carries genes for resistance to both antibiotics. Finally the double-resistance plasmid may be transferred again, by transduction. A bacterial virus infects the salmonella and proliferates, killing the cell; one phage particle incorporates the plasmid instead of viral DNA and transfers it to new cell.

of the cell by plasmid-encoded proteins that modify some specific transport system in the cell envelope. Simon D. Silver of Washington University has found that staphylococcal resistance to the arsenate ion is the result of an active pumping process somehow mediated by plasmid genes; the ions enter the cell without difficulty but are pumped out again as fast as they get in.

The third strategy, substitution of a bypassing process for a process inhibited by the antibacterial agent, provides resistance to the sulfa drugs and trimethoprim, which act by inhibiting bacterial enzymes required for the synthesis of the essential vitamin folic acid. (Human beings, unlike bacteria, are unable to synthesize the vitamin; they depend on dietary sources, which provide too little folic acid to sustain infecting bacteria.) Plasmid-borne resistance to these drugs involves the substitution for the drug-inhibited enzyme of another enzyme that has the same function but is insensitive to the inhibition.

Clearly such genetic strategies as these could not have developed overnight; they must have arisen in the slow course of organic evolution, and their existence must predate the clinical application of antibiotics by millions of years. As a matter of fact an examination of bacteria isolated and stored long before the antibiotic era, or isolated recently in remote places unlikely to have been exposed to antibiotics, often reveals the presence of fully developed R plasmids.

I should make it clear that plasmids are not the only source of bacterial antibiotic-resistance traits. Mutations in chromosomal genes can also give rise to such traits, and for several years chromosomal mutations were mistakenly assumed to be responsible for clinical antibiotic resistance. The confusion arose from laboratory experiments in which pure strains of bacteria were challenged by an antibiotic. If a spontaneous mutation arose on a bacterial chromosome and fortuitously enabled a cell to resist the drug, that cell proliferated as the other cells died. Such a mechanism could not, however, explain the naturally occurring resistance to several drugs that appeared simultaneously in bacteria infecting a person being treated with a single antibiotic. Moreover, the resistant mutants isolated in laboratory experiments were shown to grow poorly in the absence of the particular antibiotic that had selected for them. Chromosomal resistance mutations modify the antibiotic-sensitive cellular target—say a ribosome or the cell wall—in a way that reduces its general efficiency; the resistant mutants are evolutionary cripples, and under natural conditions they rapidly die out. Plasmid-determined resistance traits, in contrast, have evolved over time to be exquisitely specific: they



RESTRICTION ENDONUCLEASES are enzymes that cleave DNA at sites determined by the sequence of the DNA's component nucleotide bases *A, G, T* and *C*. Each enzyme (two of which are illustrated here) recognizes a particular palindromic sequence of base pairs and cuts each DNA strand between specific bases of that sequence. If the sites are offset from each other as in the case of *Eco*RI (*left*), the cleavage yields fragments with complementary, overlapping single-strand ends. Such "sticky end" fragments join readily with other fragments generated by same enzyme and are particularly suitable for making new combinations of fragments.

protect the antibiotic's target in the cell without significantly reducing the cell's general adaptivity. That is the key to their success.

Evolution of Plasmids

How and why did the resistance genes of plasmids evolve? Julian E. Davies of the University of Wisconsin, noting that antibiotics are manufactured by bacteria and fungi, has suggested that the same microorganisms evolved resistance genes to protect themselves from their own products. He and his co-workers have shown that *Bacillus circulans,* which produces the aminoglycoside antibiotic butirosin, synthesizes an enzyme that inactivates butirosin and several other aminoglycosides, including neomycin. The gene specifying the enzyme can be transferred to *E. coli* by recombinant-DNA techniques, rendering the *E. coli* resistant to neomycin.

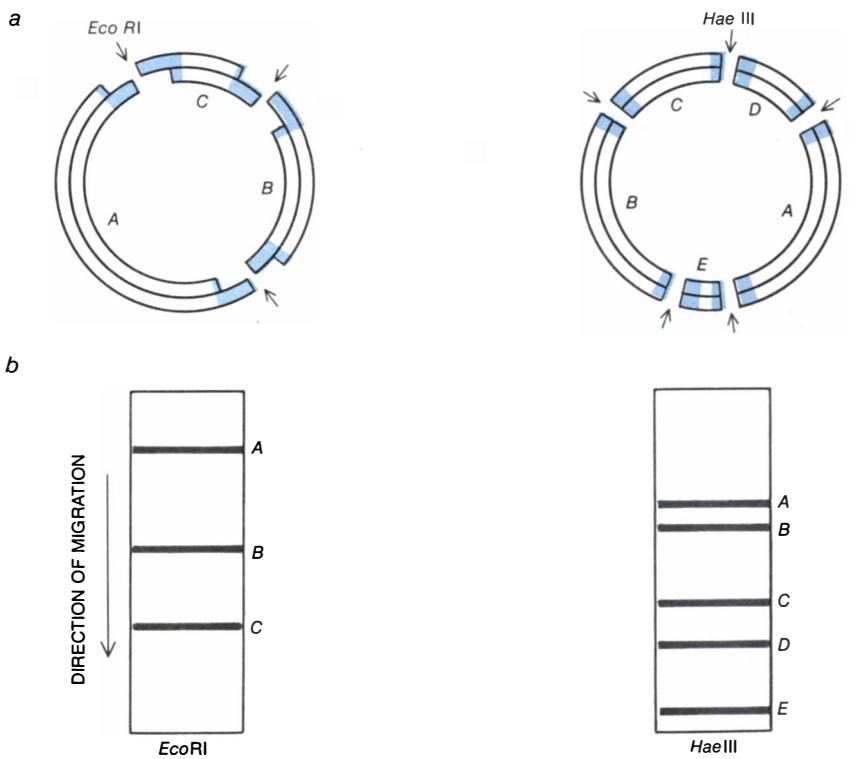
In my own view there is an additional possibility. Antibiotics are manufactured by organisms in the soil, which also contains traces of toxic heavy metals. The presence of these compounds in the soil (even at low concentrations) could have evoked the evolution of resistance genes among sensitive soil bacteria, as primary plasmid genes or as chromosomal genes that were later transferred to plasmids. In either case the genes would have made their way to nonsoil bacteria that cause infections in animals.

This view is supported by the fact that although penicillin is manufactured by fungi (which are eukaryotes, not bacteria), many nonpathogenic soil bacteria synthesize a penicillinase similar to the enzyme responsible for penicillin resistance in pathogenic animal bacteria. Plasmids specifying penicillinase have been found in staphylococcus strains isolated before the discovery of penicillin. And plasmids specifying resistance to tetracycline have been found in several species of soil bacteria. In my laboratory June Polak has shown that one such plasmid can be transferred to *Staphylococcus aureus,* which thereupon becomes resistant to tetracycline; the plasmid is almost identical with a plasmid that is native to staphylococci and specifies resistance to kanamycin. The implication of this soil-bacteria hypothesis is that the large-scale application of antibiotics in medicine and agriculture has disrupted a very old ecosystem in which a balance had been achieved between producers of antibiotics and target organisms with resistance genes.

When R plasmids were first discovered, they commonly carried one or two resistance traits. Today it is not uncommon to find a plasmid carrying genes for resistance to as many as 10 antibiotics. The rather frightening clinical implication of this accumulation of resistance genes is that treating a patient with a single drug can promote the selection of an organism resistant to everything in sight. Analysis of this phenomenon has revealed another astonishing property of plasmids and related genetic systems: genetic transposition.

At first it was assumed, by analogy with the F factor's behavior, that chromosomal mutations specifying resistance had been "picked up" by R plasmids temporarily integrated into the chromosome, which later emerged entraining a segment of chromosomal DNA. It was impossible, however, to find DNA sequences on plasmids that matched sequences on their host-cell chromosomes, as would have been expected if the plasmids had picked up significant amounts of the chromosomal DNA. Moreover, if the plasmids could routinely acquire chromosomal genes, how had they apparently been able to choose only resistance genes? Even if some selective force was responsible, surely some sequences on each side of the resistance genes would have been picked up. Actually, as it has turned out, many R plasmids do not insert themselves into the chromosome in the manner of F factors.

The real answer to multiple antibiotic resistance came only in 1974, first from Alan E. Jacob and Robert W. Hedges



PLASMID IS CLEAVED by different restriction enzymes at sites peculiar to each enzyme, yielding a set of fragments (*a*). The length of those fragments (and thus the cleavage sites for the enzyme) can be determined by gel electrophoresis. After migration through the gel the DNA is dyed, revealing a pattern of bands (*b*), each one composed of fragments of the same size. The distance of each band from the origin gives the size of the fragments making up that band.

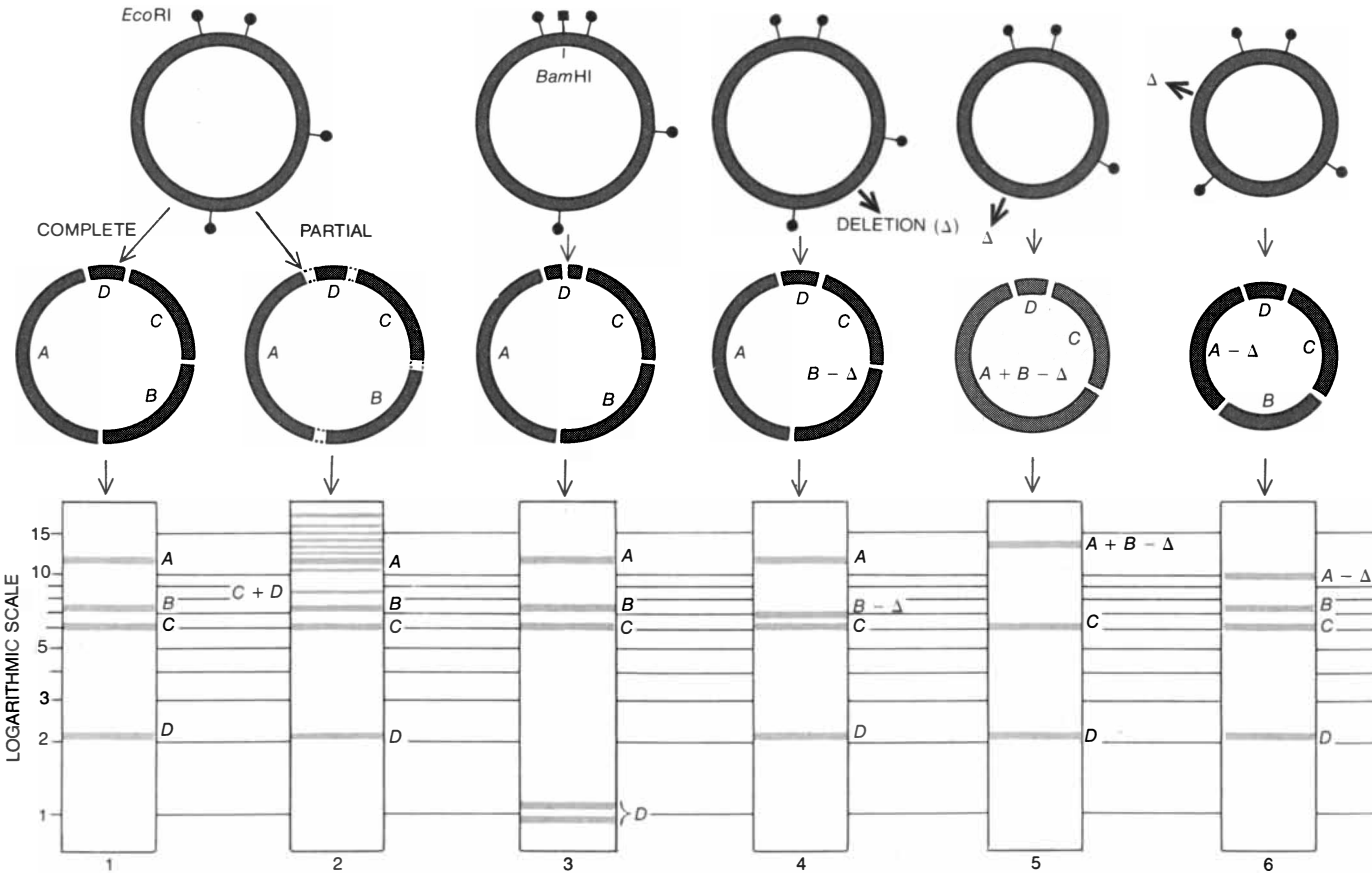
of the University of London Royal Postgraduate Medical School and then quickly from at least five other investigative groups. The resistance genes in question, it developed, were on novel genetic elements that can transfer themselves bodily from one genetic location to another in a cell; they do so without benefit of the enzymes that mediate more typical recombination mechanisms and without any need for the matching DNA base sequences required by those mechanisms. These new genetic elements, known as transposons, are responsible not only for the accumulation by plasmids of multiple resistance genes but also for gross genetic rearrangements in bacteria and apparently in eukaryotes too [see "Transposable Genetic Elements," by Stanley N. Cohen and James A. Shapiro; SCIENTIFIC AMERICAN, February]. Probably the prevalence of resistance genes on the transposons discovered so far will turn out to have the same basis as the prevalence of R plasmids themselves: the radical change in the processes of evolution brought about by the wholesale dissemination of antibiotics.

The Structure of Plasmids

Plasmid genes seem to be arranged so as to ensure hereditary stability and at the same time allow great structural flexibility. Unlike viruses, plasmids can acquire new genes and grossly rearrange old ones to maintain a store of genetic information consistent with the needs of their current host organism—all without compromising their own replicative efficiency. To appreciate the plasmid's genetic strategy one needs to know a little about the experimental analysis that has revealed it.

Plasmids do not readily yield their organizational secrets to the mapping techniques that have revealed the structure and function of genetic systems in animals, plants and even viruses. Those techniques are based on tracking gene mutations in genetic crossing experiments. In plasmids it is hard to isolate mutations affecting replication functions, and the analysis of genetic recombination is discouragingly complex. Plasmids have, however, turned out to be well suited to some recently developed techniques of molecular genetics, in particular the application of gel electrophoresis to the analysis of DNA and the dissection and cloning of DNA sequences with enzymes, known as restriction endonucleases, that are synthesized by bacteria.

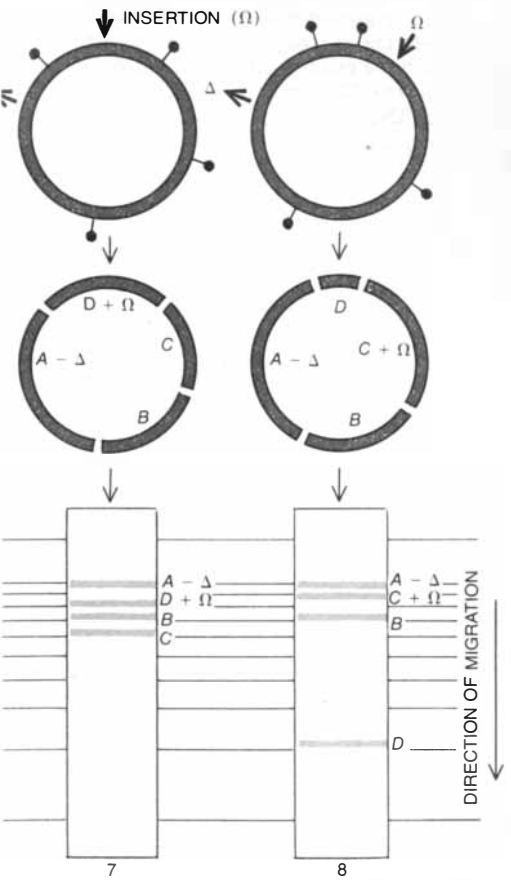
A plasmid, like most other DNA molecules, is a double helix each strand of which is a linear array of the four nucleotide bases adenine (*A*), guanine (*G*), thymine (*T*) and cytosine (*C*). The sequence of bases on one chain is complementary to the sequence on the other chain: *A* always pairs with *T* and *G* pairs with *C*. Each restriction enzyme recognizes a specific short sequence of bases and cleaves DNA at a particular site within that sequence wherever it may be in any DNA molecule. Given a large



DATA FOR MAPPING plasmid pI258 are obtained by cleaving the plasmid and variants of it with restriction enzymes and analyzing the fragments by gel electrophoresis. The plasmids are shown in the top row with cleavage sites indicated; the cleavage products are shown in the next row. Below each set of fragments is a schematic representation of the track on a gel that establishes the size of each fragment. Track No. 1 shows that enzyme *Eco*RI cuts the plasmid into four fragments, *A*, *B*, *C* and *D*, that are respectively about 13, seven, six and two kilobases (thousands of base pairs) long. The sequence of the fragments around the loop is determined by "partial digestion," in which not all the plasmid molecules are completely cleaved. The presence in track No. 2 of uncleaved fragments comprising *C* and *D* (but not any combining *B* and *D*) establishes the sequence: *A*, *B*, *C*, *D*. To make smaller fragments for detailed mapping the mixture of fragments can be treated with another enzyme. Track No. 3, for example, shows that enzyme *Bam*HI cuts *D* into two pieces, establishing the location of the single site of *Bam*HI cleavage on this plasmid. Track No. 4 is for a plasmid in which a deletion mutation has disrupted a sequence encoding penicillinase; the penicillinase gene must lie at least partly in *B*, because the *B* band is now smaller (it travels farther). In track No. 5 a different deletion (eliminating resistance to arsenate and arsenite ions) has joined fragments *A* and *B*, confirming the adjacency of *A*

collection of identical DNA molecules, such as a pure sample of plasmid DNA, each molecule will be cut by a particular endonuclease at precisely the same sites, yielding as many subsets of identical molecular fragments as there are cleavage sites on the original molecule. These subsets of fragments can be conveniently separated from one another according to size (or molecular length) by gel electrophoresis. The cleaved DNA is placed in a rectangular slot near one end of a slab of gel; an electric current is applied and the DNA molecules migrate through the gel toward the positive electrode at a rate inversely proportional to the logarithm of their molecular length. Each subset of identical fragments forms a narrow horizontal band whose position can be visualized by soaking the gel in a fluorescent dye such as ethidium bromide (which binds to DNA) and photographing it under ultraviolet radiation.

By dissecting a plasmid with several restriction enzymes, separately and in



and *B* and locating the resistance genes at or near the *A-B* junction. A transposon carrying erythromycin resistance has been deleted in the plasmid of track No. 6; it must have been in *A* because the *A* band now travels farther. Two secondary insertions of the transposon, one in fragment *D* (track No. 7) and the other in fragment *C* (track No. 8), both abolish resistance to mercury disinfectants. The gene responsible for that resistance must therefore be partly in fragment *D* and partly in *C*.

combination, one can generate various sets of fragments that can be pieced together to form a fragmentation map. The location of the plasmid's genes can be determined in several ways, including the analysis of plasmids with structurally identifiable mutations such as deletions and insertions. Some idea of the procedure can be gained by considering the first steps in the analysis of the staphylococcal R plasmid pI258, which has an overall length of some 28,000 base pairs and carries genes for resistance to erythromycin and penicillin and to a number of toxic metal compounds.

The plasmid is first cleaved, by the restriction enzyme *Eco*RI, into four fragments, which track No. 1 of the gel shows to be respectively about 13, seven, six and two kilobase pairs (thousands of base pairs) long [*see illustration on preceding two pages*]. The actual sequence of those fragments around the circle is determined by a "partial digestion," shown in track No. 2. The sites at which other enzymes cut the plasmid can be determined by a large number of experiments such as the one shown in track No. 3. The remaining tracks show the fragments obtained when derivative plasmids are dissected from which segments bearing known traits have been deleted or into which a transposon has been inserted. Examination of a plasmid with a deletion affecting a particular gene shows which segment has been shortened by the deletion, thereby locating the deleted gene. The insertion of a transposon lengthens a fragment and disrupts the function of any gene into which it is inserted, thus pinpointing the site of the inactivated gene.

By examining only these few physical rearrangements of the plasmid DNA one can get a great deal of information about the structure and genetic organization of the plasmid. By combining the results of a larger number of electrophoresis experiments with the direct determination of a plasmid's entire base sequence (which has been accomplished only for much smaller plasmids) one can hope to obtain the detailed information needed for a complete understanding of the plasmid as a genetic entity.

Plasmid Autonomy

Since plasmids represent the ultimate stage of symbiosis, being dependent on the host cell for all life-support systems except the "autonomy functions" (those that are most intimately linked with its claim to being an organism and that serve to define its individuality), the identification and analysis of those functions reveal the minimal essential components of the living state. Put another way, every plasmid must contain in its genome a solution to the problem of how a nonessential genetic system can be stably inherited and how it can control its own replication and the distribution of its copies so that a constant relationship with its host is maintained. What are those solutions?

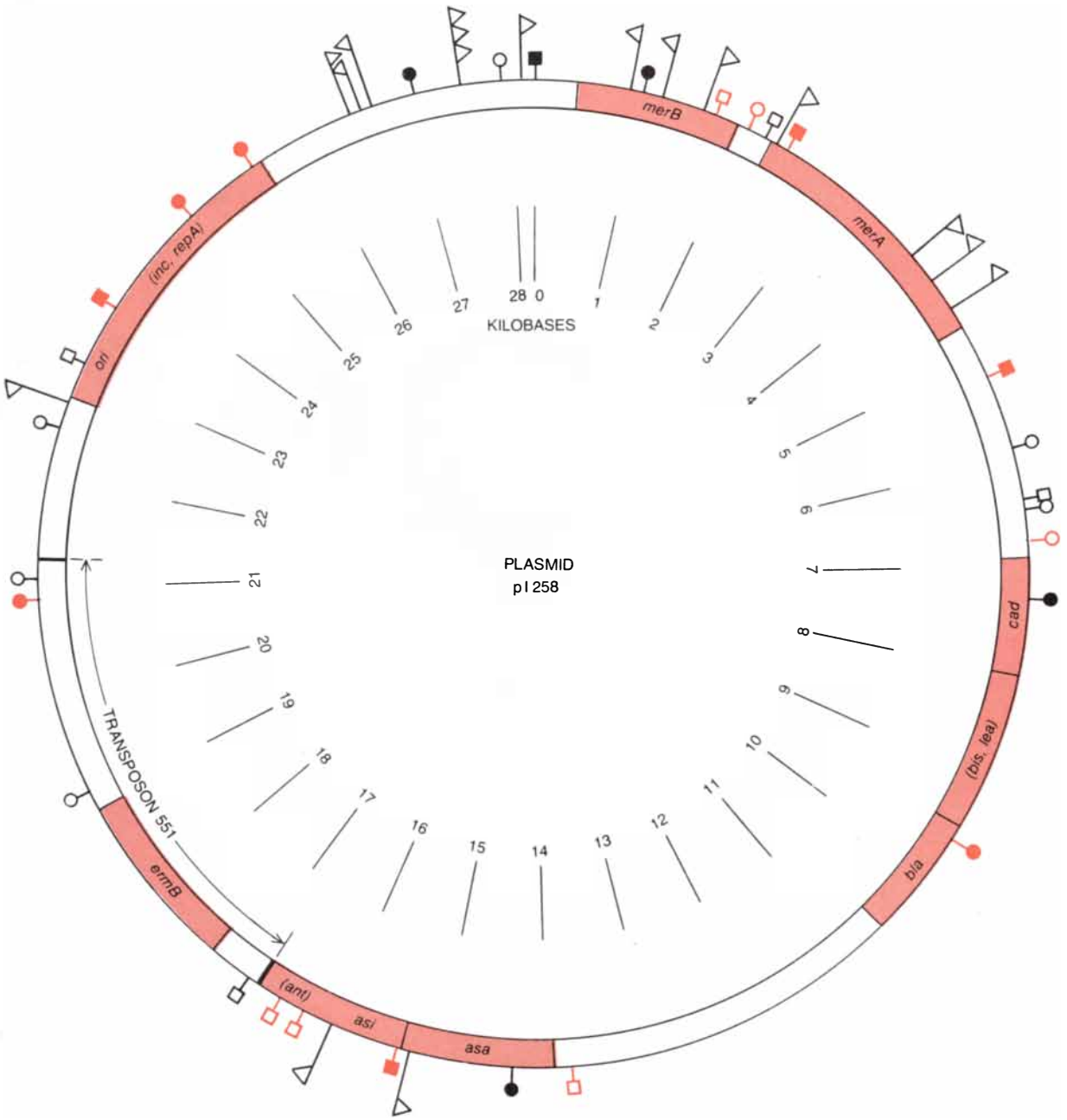
The basic theoretical foundation for considering this question was developed in 1963 by Jacob, Sydney Brenner and François Cuzin. They proposed that for any given "replicon," or replication unit, whether it is a bacterial chromosome, a segment of a eukaryotic chromosome, a virus, a plasmid or some other form, replication begins at a specific point (the origin) and proceeds in a sequential and linear manner to another specific point (the terminus), where it stops; an "initiator" substance, encoded by the replicon itself, acts at the origin to trigger replication. They proposed that in bacteria all autonomous replicons are permanently attached to the inner surface of the cell membrane at a site near the cell's equatorial plane; the attachment is essential for replication and for partitioning, or copy distribution. At a particular stage of the cell cycle, then, a new attachment site would be formed for each replicon on the side of the equatorial plane opposite the existing site, triggering a replication cycle during which the newly synthesized DNA molecule would become attached to the new site. The growth of a division septum along the equatorial plane would allot one of the two DNA molecules to each daughter cell. This brilliant theoretical synthesis, years ahead of its time and formulated on the basis of scanty evidence, has served to guide virtually all research on the control of bacterial replication and on the genetics of plasmid autonomy.

The theory predicted that very few genetic functions—only a small part of any plasmid—would be required for autonomy. This was confirmed by the isolation of derivative plasmids from which large segments of DNA had been deleted but that were still viable. In a study of deletions affecting the staphylococcal R plasmid pI258 I found that as much as two-thirds of the plasmid could be deleted without affecting the plasmid's autonomy functions. No other type of genome can survive deletions nearly as large as that. More important, there was a small "forbidden" region of the viable plasmids that was never affected by deletions and that turned out to carry all the autonomy functions; deletion of any part of this region destroyed the plasmid's viability.

Similar results have been obtained for many other plasmids, and it is now well established that plasmid autonomy genes are clustered in what Ken-ichi Matsubara of Osaka University has termed a "replication drive unit." In 1966, when my experiments were completed, molecular techniques had not yet advanced to the stage allowing precise determination of the size of the replication drive unit. More recently it has been found that the unit cannot be larger

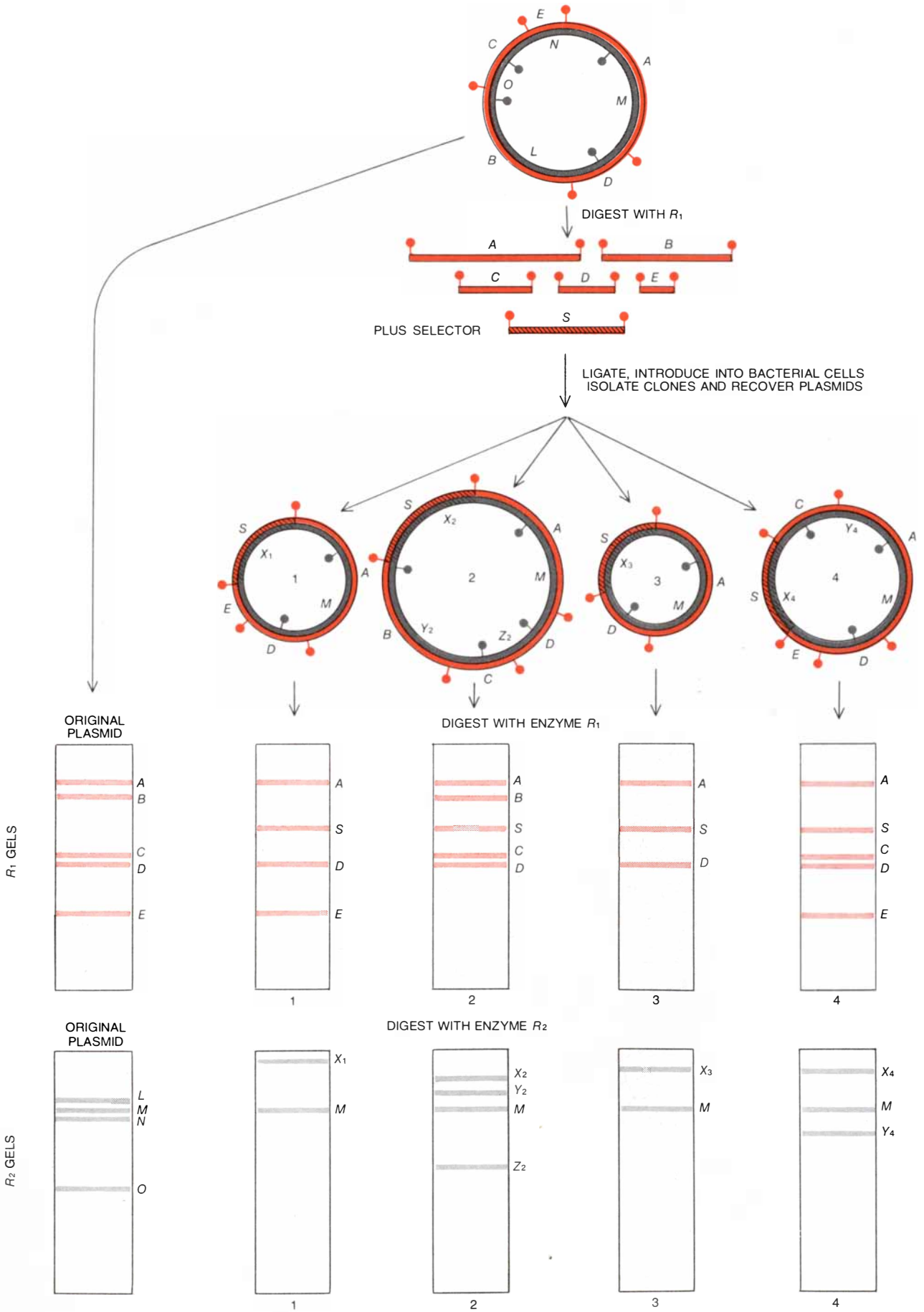
than about 3,000 base pairs, or just over a tenth of the pI258 plasmid.

Recent molecular-cloning experiments have defined more precisely the replication drive units of several *E. coli* plasmids. In these experiments a plasmid is digested with a restriction enzyme and the collection of fragments is mixed with a nonreplicating "selector" segment of DNA generated by the same endonuclease and containing a known marker gene. Since all the fragments are cleaved from DNA by the same enzyme, they have matching and symmetrical unpaired bases at their ends ("sticky ends") and can therefore combine with one another, in any number, order and relative orientation, by base pairing. When the mixture is treated with the enzyme DNA ligase, the base-paired fragments form new, scrambled plasmids: random rearrangements of the original plasmid. When these are introduced into recipient cells, only those that contain the region or regions required for autonomous replication are able to replicate and only those that also contain the selector can be recovered. Treatment with the original endonuclease shows which of the original frag-



EcoRI
BamHI
Bgl II
Hpa I
Xba I
Pst I
Sma I
Sal I

PHYSICAL-GENETIC MAP of plasmid pI258 reflects the results of many experiments of the kind shown in the preceding illustration. Locations are given by the kilobase scale, which indicates the distance from the single site at which the restriction enzyme *Bam*HI cleaves the circular molecule; the symbols (*see key*) mark the recognition sites for different restriction enzymes. Pennants represent various alternative sites at which transposon 551 has been mapped and indicate the orientation of the transposon. Plasmid genes that have been mapped include those for resistance to penicillin (*bla*), arsenate (*asa*), arsenite (*asi*), antimony (*ant*), erythromycin (*ermB*), inorganic mercury (*merA*), organic mercury (*merB*), cadmium (*cad*), bismuth (*bis*) and lead (*lea*) as well as for the replication origin (*ori*), the initiator protein (*repA*) and incompatibility (*inc*). Genes whose precise location is still not certain are indicated in parentheses.



ments are present and analysis with a second endonuclease establishes their sequence and relative orientations.

If all replication-function genes are contained within a single original fragment, then that fragment is always found to be present in viable new plasmids. Kenneth N. Timmis, Felipe Cabello and Stanley N. Cohen of the Stanford University School of Medicine found that to be the case for two large *E. coli* plasmids, F and R6. In both cases a single fragment about a tenth as long as the original plasmid was found to contain the entire replication drive unit and thus to ensure autonomous replication in the same way as the full original plasmid. Further reduction in size of the derivative plasmids indicated that a DNA segment of no more than 2,000 base pairs encompasses the entire replication drive unit of plasmids more than 50 times as long.

If two fragments each contain independently necessary replication functions, both are always found to be present in the new plasmids, without constraint on their relative location. If a restriction-enzyme site falls within the confines of a required gene, again two of the original fragments are always recovered from the new plasmids, but in this case they are always adjacent and in the same orientation as in the original.

Replication Drive Unit

Four elements of a typical replication drive unit have so far been defined experimentally: the origin (the site of initiation of replication), the initiator substance, the genetic functions controlling copy number and those that ensure partitioning, or the distribution of replicas to daughter cells. The central element of the autonomy system is the replication origin, a unique sequence of bases within which the actual "start site" for each replication cycle is situated. A replication origin is defined functionally as the smallest DNA sequence that can serve as the start of a replication cycle.

An extension of the fragment-scrambling experiment makes it possible to isolate the origin sequence. As before, one ligates a mixture of restriction-enzyme fragments with a selector. Now, however, the new plasmids are introduced into a specially engineered bacterial strain. It contains a second, unrelated plasmid to which has been attached a fragment of the first one that is known to encode whatever diffusible substances, such as the initiator protein, are necessary for its replication. As before, the experiment pinpoints a fragment whose presence is required if a scrambled plasmid is to be viable. In this case, however, the required fragment need not include the entire replication drive unit because the diffusible substances encoded by that unit will have been supplied by the complementary plasmid already present in the recipient cells. Therefore any fragment that is always present in viable new plasmids must specifically contain the DNA that constitutes the origin sequence. In this way Donald R. Helinski of the University of California at San Diego and his colleagues isolated the origins of the *E. coli* plasmids R6K and RK2, and similar methods have been applied to isolate the origins of other plasmids.

In each case the origin region has been found to be several hundred base pairs long. Within that region the synthesis of a new DNA chain always begins at essentially the same base pair. The rest of the origin region presumably contains specific sequences to which various proteins (and perhaps other large molecules) adhere to influence the initiation process. For example, it now appears that the first step in the replication of plasmid (and other) DNA is the traversal of the origin sequence by the enzyme RNA polymerase. This interaction not only synthesizes an RNA chain that primes the replication of the DNA but also separates the two DNA strands, perhaps allowing intrastrand loops to form; these loops may attract the enzymes and other proteins involved in the actual biosynthesis of new DNA.

The first hint of the existence of plasmid genes required for replication came when Jacob, Brenner and Cuzin isolated F-plasmid mutants that were unstable at high temperatures, apparently because they were defective in replication. We isolated similar mutants of various staphylococcal plasmids that we could show were indeed unable to replicate at an elevated temperature at which the host cells grew perfectly well [*see illustration on page 102*]. These plasmids were enabled to replicate when they were complemented by other copies of the same plasmid carrying the nonmutant allele (variant) of the temperature-sensitive gene. This meant the mutant gene encoded a diffusible protein (generally assumed to be the initiator) that was supplied by the nonmutant plasmid. The mutants were not enabled to replicate by a different but closely related plasmid, establishing the high degree of specificity of this protein-DNA reaction and fulfilling the prediction of a specific initiator for each replicon. Manabu Inuzuka and Helinski have demonstrated that such a protein, encoded by plasmid R6K, is required for replication of the plasmid in cell-free extracts of *E. coli*.

At first it was assumed that there would be found to be as many copies of a particular plasmid in a cell as there were chromosomes, with plasmid replication coupled to and regulated by chromosomal replication. That is not the case. In the plasmids studied to date there appears to be a continuum of copy numbers from one or two per cell to more than 100; in general the larger the plasmid, the smaller its copy number. And plasmid replication is generally regulated independently of chromosome replication. Since different plasmids present in the same host can have very different copy numbers, the copy number must be determined by a plasmid-encoded regulatory system. The first such system was discovered in 1972 by Kurt Nordström, now of Odense University in Denmark, for the *E. coli* plasmid R1, and similar systems have been revealed for other plasmids, in staphylococci as well as in *E. coli*. In R1, at least, copy number appears to be controlled by a protein or proteins inhibiting replication, fulfilling a 1968 prediction by Robert H. Pritchard, Peter T. Barth and John F. Collins of the University of Leicester that replication is controlled negatively, that is, by a repressor.

When in the course of the cell cycle do plasmids replicate, and what determines the sequence in which particular copies of a given plasmid will replicate? There might possibly be a precise moment when all plasmid molecules replicate in synchrony, but there are indications that the replication events are instead random ones, distributed throughout the cell cycle. An elegant experiment reported in 1967 by Robert H. Rownd of Wisconsin showed that all the copies of a particular plasmid act as members of a homogeneous pool, with copies chosen for replication by random selection without respect to their recent replication history: the regulatory system does not discriminate between "new" and "old" DNA. Other workers have shown that plasmids replicate throughout the cell cycle. One simple interpretation of these observations is that at any constant host-cell growth rate there is a corresponding constant concentration of a repressor that fixes the probability of plasmid replication in any time period.

How two elements, a positive initiator

FRAGMENT-SCRAMBLING experiment defines the region responsible for plasmid autonomy functions. A plasmid is diagrammed (*top*) to indicate the sites at which hypothetical restriction enzymes R_1 (*colored symbols*) and R_2 (*gray symbols*) cleave it. The plasmid is first digested with R_1 and the fragments (*A–E*), mixed with a "selector" fragment (*S*) having matching ends, are ligated to make a variety of new plasmids, which are cloned in bacteria. Viable plasmids (those containing the autonomy functions) replicate in the bacteria and can be recovered if they also contain the selector. Digestion of the scrambled plasmids with R_1 reveals that each contains a different set of fragments, but all of them have the selector and also fragments *A* and *D;* the last two must between them contain all autonomy functions. Dissection of the original plasmid with enzyme R_2 yields a different group of fragments, *L, M, N* and *O;* scrambled plasmids yield *M* and variety of new fragments. Because only *M* is always present in viable plasmids it must include autonomy functions; its constant presence shows *A* and *D* are adjacent and in same relative orientation in all viable scrambles.

and a negative repressor, might interact to regulate the rate of replication is not yet clear. It is probably different in different systems. A repressor could control either the rate of synthesis of the initiator protein, the rate at which the initiator interacts with the origin or the frequency with which the origin region is transcribed by RNA polymerase. Any such copy-control mechanism is quite different from the one predicted by Jacob, Brenner and Cuzin, which formally implies the presence of only a positive regulatory element, consisting of a limited number of structural sites to which a replicon must attach to be replicated.

Partition

Jacob, Brenner and Cuzin recognized that if a nonessential genetic element present in a limited number of copies in each cell was to maintain its hereditary stability, there would have to be a mechanism to ensure the equal distribution (partition) of the copies of the element among daughter cells. Such a mechanism has not yet been clearly identified, but the first direct evidence that there is one came from a study in our laboratory of the plasmid mutants that have a temperature-sensitive initiator. As the host cells divide at the elevated temperature the plasmids originally present in them are not destroyed, but neither are they replicated; they are assorted among the dividing cells until no cell has more than one plasmid.

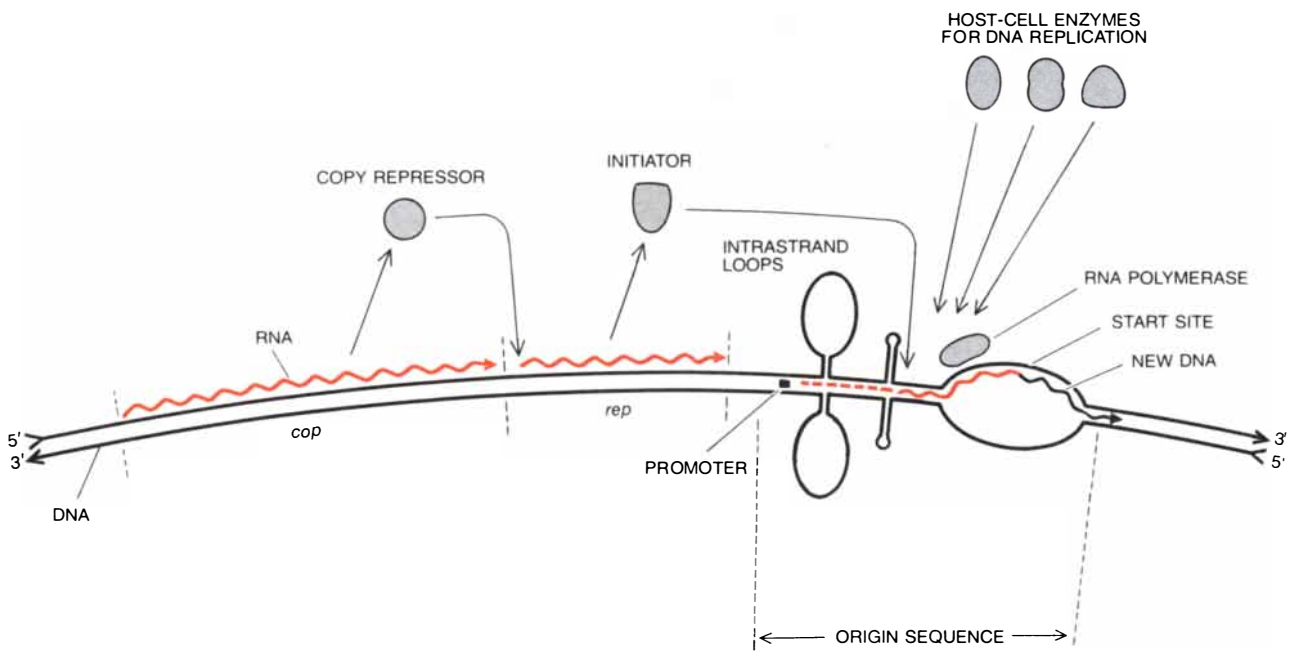
We did a statistical analysis of this assortment in the case of two different plasmids, one with an average copy number between two and three and the other with about 30 copies per cell. In both cases the copies originally present were divided approximately in half at each generation until there was one copy per cell. Thereafter at each cell division one daughter cell had a single plasmid and the other cell had none. We concluded that at least these two plasmids do contain a specific equipartition mechanism. Here again the mechanism is likely to have at least two components, an effector molecule and a recognition site, and the cell envelope is probably involved.

Nordström and Rownd independently have recently found that when a particular segment of an R plasmid is excised by recombinant-DNA techniques, the plasmid becomes unstable and is inherited in an apparently random way; the tentative interpretation is that the deleted segment must be implicated in equipartition. In collaboration with Carmen Sanchez-Rivas of the University of Paris we have studied plasmids in staphylococcal protoplasts: bacteria from which the tough cell wall has been removed, leaving only the thin cell membrane. When staphylococcal protoplasts regenerate a cell wall, many of the regenerated cells fail to inherit plasmids that were originally present. The plasmid DNA is not physically expelled from the protoplasts and it can replicate normally, and so the failure to inherit must be related to a series of cell divisions that take place during cell-wall regeneration. Because the normal location of the plane of division is determined by the cell wall these divisions are abnormal, often resulting in "cells" without a chromosome (which of course cannot survive) or cells without any plasmid.

This result implies that the intact cell wall is intimately involved in the coordination of division and partitioning, in keeping with the theory of Jacob, Brenner and Cuzin. The theory, however, specified that after replication the two copies of a replicon are automatically separated, one to each daughter cell. It is now clear that this cannot be true for plasmids because of a phenomenon, unique to plasmids, known as incompatibility.

Incompatibility

Pairs of closely related plasmids usually cannot be stably maintained together in a single cell line: they are "incompatible." It may be possible to develop a line containing two such plasmids, but it is difficult to keep them both present. As the cells proliferate, subclones with only one or the other of the two plasmids are continually generated; the culture approaches, asymptotically, a state in which it consists of two pure sublines, one for each of the original plasmids. Frank C. Hoppensteadt of the University of Utah and I have shown, by statistical analysis of experiments done in our laboratory and in Nordström's, that the process is clearly stochastic: the results



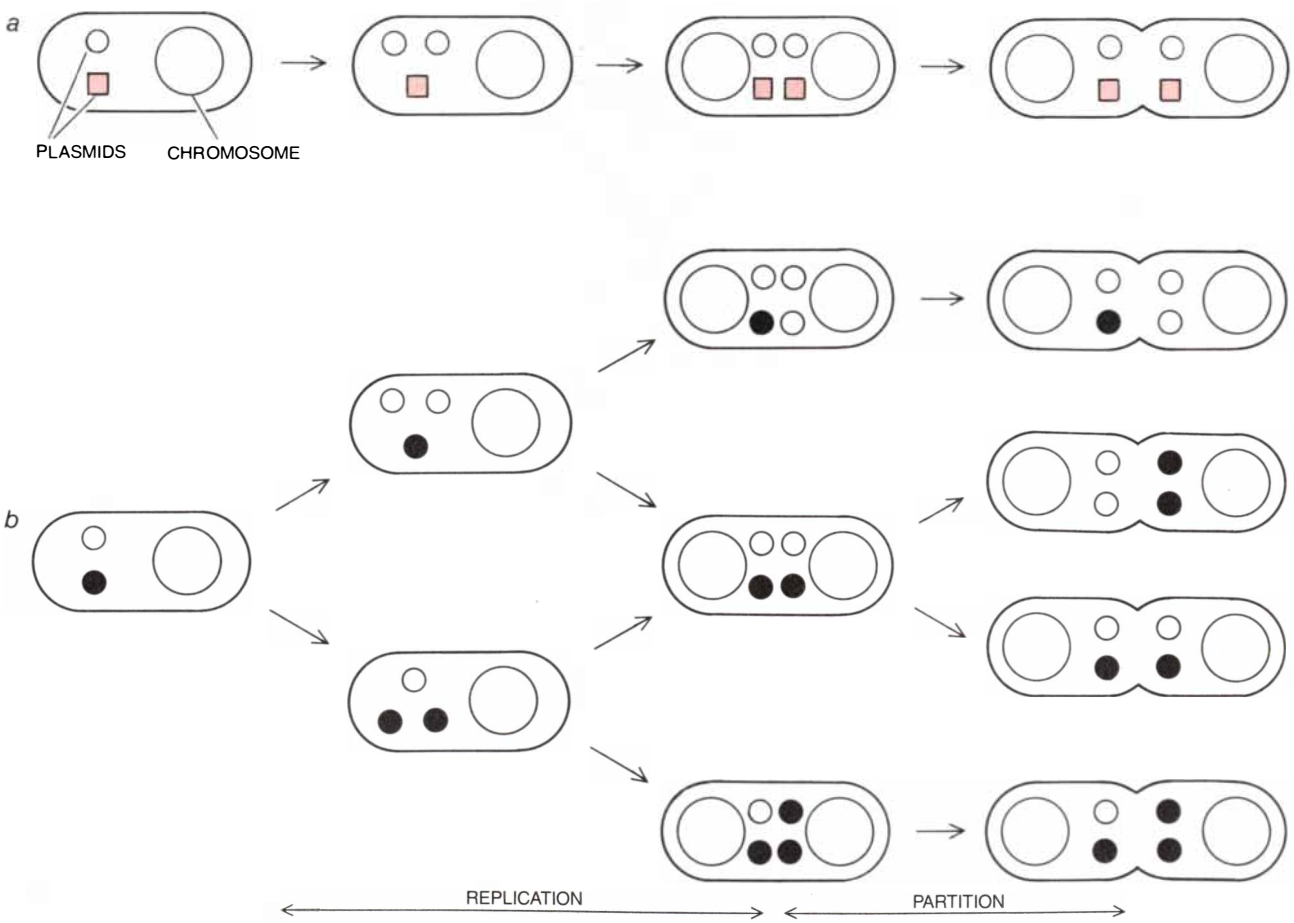
PLASMID REPLICATION might be controlled as is suggested by this model. An initiator protein is encoded by a replication gene (*rep*) whose activity is controlled by a repressor protein that is specified by a copy-control gene (*cop*). The initiator recognizes loops in the two strands of plasmid DNA; the loops form (by intrastrand base pairing) when the strands are dissociated as the enzyme RNA polymerase traverses the origin sequence, beginning at the promoter, to synthesize an RNA transcript that will serve as a primer for DNA replication. The resulting complex of the initiator and the intrastrand loops attracts a number of host-cell enzymes, including DNA polymerase, that mediate the synthesis of DNA; these proceed to attach new nucleotide bases to the RNA primer, starting a new replication cycle.

can be interpreted as the net effect of random replication and random assortment. Two incompatible plasmids *A* and *B* in the same cell form a common pool from which individual copies are chosen at random for replication until the total number has doubled; then they are again chosen at random for partitioning, with half of the total going to each daughter cell. The relative numbers of *A* and *B* will therefore vary from cell to cell in a mathematically predictable way, even though the sum of *A* and *B* remains constant. Since the two plasmids have similar or identical replication drive units, they are not recognized as being different by the plasmid and cellular systems that regulate replication and partitioning. And so the disproportions persist, giving rise (at a predictable rate determined by the overall copy number) to progeny cells having only *A* or only *B*.

In contrast to plasmids with similar or identical drive units, pairs of unrelated plasmids can be maintained together in a cell and are regarded as being compatible; naturally occurring bacteria are quite commonly observed to have as many as seven or eight different plasmids. The contrast between compatibility and incompatibility indicates a high degree of specificity in the interactions between cells and plasmids. There must be specific gene loci on a plasmid that are responsible for the differing behavior of compatible and incompatible plasmids. Perhaps plasmid partition is specific for each incompatibility type, in that any two incompatible plasmids share the same partition apparatus. In that case there must be at least as many different partition systems as there are different groups of mutually incompatible plasmids, and a substantial number of such incompatibility groups have already been identified in several bacterial species. All of this suggests that the specificity of the partition system is determined solely by the plasmid.

Consideration of plasmid partition and incompatibility bears directly on the essence of the symbiotic state. For most of its needs a subcellular endosymbiont presumably exploits a host structure, metabolic pathway or other function for the same purpose as the host itself does. For example, a plasmid presumably depends on a host RNA polymerase to transcribe its DNA and on host DNA polymerases for DNA replication and repair. Alternatively a subcellular organism can subvert a host function to serve a need peculiar to itself. For example, the system whereby a bacterial cell transports the sugar maltose across its membrane is exploited by bacteriophage lambda for a very different purpose: as its site of attachment to the wall of a cell it is infecting. In other cases, such as the control of copy number, there is presumably no appropriate cellular function to be exploited, and the plasmid itself apparently encodes whatever control molecules are required.

What about partition, where a variety of plasmids must each be able independently to develop a specific structural interaction with the host cell? I would suggest that each plasmid may have "learned" to attach itself to a different structure in the cell (perhaps on the interior of the cell membrane or even, as suggested by Bruce C. Kline of the Mayo Clinic, on the chromosome itself) by evolving a protein that binds to a specific site. If this notion is correct, the isolation and study of such systems should help to clarify the special kind of symbiosis at the molecular level that is the hallmark of the plasmid way of life.



REPLICATION AND PARTITION of compatible plasmids (*a*) are contrasted with the same processes in incompatible plasmids (*b*). Unrelated plasmids are "compatible" and can be maintained together in a cell. They may replicate at different stages of the cell cycle, but both types replicate and the copies are assorted equally during cell division. Closely related plasmids are "incompatible." One or the other is chosen at random for replication (*middle columns*) and at cell division they are randomly assorted (*right*), that is, sister plasmid copies are not always separated from each other. The result is that cell lines lacking either one plasmid or the other are inevitably produced.