MICROBIOLOGICAL REVIEWS, Dec. 1990, p. 502–539
0146-0749/90/040502-38$02.00/0

Vol. 54, No. 4

Organization of the Bacterial Chromosome

STEVEN KRAWIEC[1] AND MONICA RILEY[2]*

Department of Biology and Center for Molecular Bioscience and Biotechnology, Lehigh University, Bethlehem, Pennsylvania 18015,[1] and Marine Biological Laboratory, Woods Hole, Massachusetts 02543[2]

* Corresponding author.

INTRODUCTION

The chromosome is the essential hereditary structure in a bacterium; as such, it serves as the principal repository of genetic information, a site of gene expression, and the vehicle of inheritance. The arrangement of structural genes and regulatory elements on the chromosome constitutes the environment in which genes are expressed and inherited. Studies have been directed toward elucidating some of the fundamental physical and chemical features of this environment and finding whether the arrangement of genetic information in the bacterial chromosome is important to its function. We are beginning to learn about some of the organizational features of the bacterial chromosome, and these features are the focus of this review. Such subjects as the physical and chemical organization, amounts of genetic information present, repeated sequences, recurrent rearrangements, origins, extents and limits of variation, and forbidden alterations of structure are emphasized.

Not all work pertinent to the structure and function of the bacterial chromosome has been covered here by any means, since the breadth of subjects that could have been included prohibited comprehensiveness. Rather than attempt to give full coverage to each topic, we have tried to cite useful review articles and key research papers that should make useful entries to the literature for those who wish to acquire a more detailed understanding of any area.

STRUCTURE AND COMPOSITION OF CHROMOSOMES

Physical and Chemical Attributes of Chromosomes

Nucleoid structure and HU proteins. Chromosomal DNA is compacted into a nucleoid in the bacterial cell. If the DNA of one nucleoid of *Escherichia coli* were laid out as a linear double-stranded molecule alongside an intact cell, it would be about 1,000 times as long as an average cell. Two to four of these outsized molecules are organized in each cell in such a way that the relatively enormous, highly asymmetric DNA molecule is amenable to such activities as replication, recombination, transcription, and complex regulatory processes. The image of a nucleoid as seen by electron microscopy is not an image of the compacted form of the nucleoid as it exists in the cell before release by lysis. Electron micrographs of the released, flattened, and spread nucleoid show a round or oblong shape with numerous supercoiled loops emanating from a center, having a cross-sectional area 4 to 10 times the size of the intact cell (65, 77, 179, 180). Clearly, the nucleoid expands upon extraction from the cell and preparation for electron microscopy. Further expansion occurs when supercoiling is released by strand breakage (179).

The level of negative supercoiling in the bacterial cell is dictated by a balance between DNA gyrase and DNA topoisomerase I, one winding and one unwinding the supercoil turns (72, 74, 364). Levels of DNA gyrase (234), topoisomerase I (352), and ligase (209) are under homeostatic control in that their synthesis responds to current levels of DNA supercoiling. The *E. coli* DNA gyrase binds to REP, a short often-repeated sequence in the chromosomal DNA (see Recurrent Nucleotide Sequences, below), and the binding is stimulated by interaction with HU (381, 382). The gyrase introduces cuts in the DNA at specific sites called toposites that are clustered and about 75 kb apart (323). It is tempting to think that these findings reflect the role of DNA gyrase in establishing and maintaining the architecture of the nucleoid, but the details of the relationship have not yet been established.

The extent of negative supercoiling of DNA in the nucleoid has been determined both in vitro and in vivo. In vitro, the degree of negative supercoiling has been measured as -0.05, in other words one untwist for every 20 turns of the helix. In vivo, the negative superhelicity is about half that, presumably a result of interaction of the DNA with nuclear proteins (75). In the nucleoid, each petal of the daisy, each domain, is separately supercoiled and can be relaxed independently by a nick in one strand. The numbers of domains of supercoiled DNA, estimated as the number of nicks required to relax all domains, is larger in the isolated nucleoid (about 100) (221) than in vivo (about 50) (142).

In addition to interactions with enzymes, the bacterial chromosome exists in situ in association with a variety of structural proteins. In *E. coli*, the most abundant of these is a small basic protein designated HU (73, 75, 268). Although the total amino acid composition of HU is similar to that of histones, there is nonetheless little sequence homology between the two types of protein. HU occurs most commonly as a heterodimer. Ninety percent of the molecules are composed of an α subunit (HU-α), the product of the *hupA* locus (which maps at 90 min) (175), and a β subunit (HU-β), the product of the *hupB* locus (which maps at 10 min) (176, 339). Amino acid sequences of HU-α and HU-β show a 69% identity (75).

The abundance of HU proteins is autogenously regulated (187). Calculations based on the availability of HU and the number of nucleotide base pairs in the *E. coli* chromosome indicate that at any one time approximately one-sixth of the DNA could be organized into structures equivalent in dimension to a eucaryotic nucleosome. Electron micrographs of the structure produced by mixing DNA and HU protein revealed "beads" less regular than eucaryotic nucleosomes (40, 293). Structures with a similar appearance have been observed in the remnants of osmotically ruptured *E. coli* cells (115). Such beaded structures are unstable. The low availability of HU, the lack of uniform appearance, and the instability could mean that nucleosomelike structures in procaryotes are not static but are in a dynamic equilibrium termed transitional coiling by Pettijohn and Hodges-Garcia (268). Indeed, the difference in the number of supercoiled domains seen in vivo and in vitro, along with the fragility of

the HU-DNA complex, further supports the concept of a dynamic structure.

Nonuniform localization of key proteins in the cell also suggests that the DNA in the nucleoid is in a mobile state with transcriptionally active segments at the interface between the surface of the nucleoid body and the cytosol. In the *E. coli* cell, the chromosomal DNA occupies largely ribosome-free space at a concentration of about 20 to 50 mg ml^{-1}, surrounded by the ribosome-filled cytosol (180). Electron micrographs of thin sections of the *E. coli* cell containing immunolabeled RNA polymerase, HU, and topoisomerase located these proteins at the border between the lobular bulk DNA and the cytoplasm, not dispersed throughout the area occupied by DNA (77, 180).

Recent results with *hupA* and *hupB* mutants reveal that disruption or loss of either one of the loci does not have a striking effect on viability (164, 175, 339). By contrast, double mutants are severely enfeebled under some growth conditions (especially at temperatures at the extremes of the permissible range). Double mutants contain less ordered nucleoids and form both filamentous cells and anucleate cells (160, 164, 361). The value of the double mutants is that the roles of the HU proteins can be assessed in vivo rather than solely in model systems. In vivo analysis with *hupA-hupB* mutants showed that HU is required for inversions mediated by several proteins. These include Hin from *Salmonella typhimurium* (172), Pin from the cryptic phage e14; and Gin from phage Mu (164, 362). By contrast, HU has no effect on inversions associated with Cin, a phage P1 protein homologous in structure and function to Gin. Furthermore, phage Mu fails to grow, lysogenize, or transpose in the absence of HU. Finally, HU is necessary for inversions in the R64 "shufflon" (362), a plasmid containing four invertible segments each of which is flanked by seven 19-bp repeats (189).

In addition to effects on gene arrangement, HU has many other effects on vital processes involving protein-DNA interactions, including transcription from bacteriophage templates, interaction of *lac* repressor and operators to form loops (88, 193), and formation of prepriming complexes during initiation of replication from *oriC* (69, 75, 100).

The viability of strains bearing mutations in both *hupA* and *hupB* implies that other DNA-binding proteins which are less abundant in *E. coli* can substitute for HU. Experimental observation substantiates that integration host factor can compensate to a large extent for the absence of HU (174). Additional proteins which may play a role include protein H (similar to histone H_2A), protein H1 (a neutral protein which, nonetheless, has striking avidity for DNA), and the product of *firA* (a gene whose protein product is presently poorly understood) (75).

Although HU has been studied most extensively in *E. coli*, HU-like proteins have also been identified in *Bacillus stearothermophilus*, *Pseudomonas aeruginosa*, *Rhizobium meliloti*, *Clostridium pasteurianum*, *Anabaena* sp., and *Thermoplasma acidophilum*, as well as in *Bacillus subtilis* phage SPO1 (75).

Localized configurations of DNA. Exposure of the cell to environmental shifts can bring about changes in the tertiary configuration of the DNA. The degree of supercoiling is affected by changes in osmolarity (142, 143), shifts between aerobiosis and anaerobiosis (70, 74), and changes of cell growth phase (12). The degree of supercoiling may influence physiological responses by altering the differential expression of groups of genes.

The level of supercoiling both affects and is affected by the process of transcription. The efficiency of transcription can be enhanced by negative supercoiling. The torsional stress of the supercoiled DNA constitutes in effect a pressure to unwind the double helix, so that the helix can open up at A+T-rich promoter regions, facilitating initiation of transcription. The active process of transcription in turn can generate both positive and negative supercoiling as follows. Transcription can be thought of as producing a linear RNA transcript from a revolving DNA template (85). Transcription of a circular DNA molecule lacking a swivel generates positive supercoils in front of the transcribing RNA polymerase and negative supercoils behind (275, 276, 378). The excess positive and negative supercoils are relaxed in *E. coli* by DNA gyrase and topoisomerase I, respectively.

The existence of negative supercoiling in bacterial DNA gives it a dynamic quality, facilitating reversible formation of localized alternate structures of DNA (281, 368–370). It has become clear that the secondary structure of the bacterial DNA is not uniformly a right-handed B-DNA configuration (368). Alternate structures exist over short distances, such as Z-DNA in a left-handed double helix (286, 321, 369), cruciforms (212, 342), triple-stranded DNA (125, 369, 370), and slipped structures (208). The nucleoid in vivo is undoubtedly a complex structure in which the energy of the supercoil is used to drive the formation of localized structural variants that can affect the biological activities of the genes of that region.

Tools have been developed that can detect some of these alternate DNA configurations in vivo. An assay based on the resistance of Z-DNA to methylation has revealed the presence of left-handed Z-DNA in plasmids in vivo (171, 369). Cruciform structures can be detected in vivo by an assay in which the specificity of phage T7 endonuclease is used to seek out sensitive bonds at the central cross of the structure (261). Two kinds of cruciform structures can be extruded from regions of DNA with inverted repeat sequences (211, 212, 342). Other kinds of double-stranded DNA structures can be adopted by sequences with extreme purine or pyrimidine strand bias. Some of these can be detected as regions in plasmid DNA hypersensitive to cleavage with S1 nuclease (370). Also, plasmids with synthetic oligonucleotide inserts can form triple-stranded structures with the pyrimidine-rich oligonucleotide strand folded back into the major groove of the duplex. Such structures are favored by the presence of negative supercoiling.

Other localized DNA structures that may be important to organization and function of the bacterial chromosome are bends and loops. Neither bends nor loops use the energy of supercoiling for formation. Bends are determined by the nucleotide sequence (6). It has been proposed that bent DNA segments might serve to delineate the ends of supercoiled domains (200). Large loops in the DNA on the order of several hundreds or even thousands of base pairs may be formed in part by bends that bring distant loci closer together, but also by the action of specific proteins that bind to each other and to distant DNA loci. Proteins bound to distant DNA sites can interact with each other, thereby forming a protein-nucleic acid complex that creates the loop of DNA. The loop model provides a rationale for how sequences distant from a set of genes can affect their regulation and expression (3, 114, 305, 365). Supercoiling can affect the formation and stability of such loops not through energetics per se but by affecting the relative positions of two distant protein-binding sequences on the faces of the helix (200) and by affecting the degree of condensation of the DNA, thus facilitating or interfering with the interaction of the proteins bound to the two distant loci.

We have a great deal to learn about the dynamics of the activities of the bacterial chromosome within the cell and how the activities are coordinated and regulated. Not only the primary sequence but also the secondary and tertiary structure of the DNA affects gene expression. Some of the coordinated physiological responses of the cell are probably achieved by the adoption of localized specific secondary and tertiary structures at key loci.

Chromosome sizes and geometries. A succession of techniques has been used to estimate the sizes of bacterial chromosomes. The general temporal order of the principal techniques has been : (i) colorimetry (coupled with estimates of numbers of nucleoids), (ii) kinetics of renaturation (data for at least 605 strains of bacteria have been compiled [139]), (iii) two-dimensional gel electrophoresis of restriction fragments, (iv) summation of the sizes of restriction fragments which produce a unique and complete physical map of the chromosome, and (v) pulsed-field gel electrophoresis (PFGE) of macrofragments produced by rare cutting restriction enzymes.

Digestion of chromosomes with conventional restriction endonucleases often yields hundreds or even 1,000 or more fragments; determining that no fragments are overrepresented, that none has been lost, and that the sum of the sizes of the fragments is a complete measure of the chromosome is an imposing task. Short of building a complete restriction map of a chromosome, two-dimensional gel electrophoresis provides spatial separation of most of the fragments (and helps establish the overlap of fragments generated by two restriction endonucleases). Fragments generated by digestion with a restriction endonuclease are separated in one dimensional by gel electrophoresis, digested with a second enzyme, and electrophoresed at right angles to the original axis of migration. Early applications of the technique provided estimates of genome size which were 10 to 20% lower than those obtained from PFGE (see below). However, improvements in two-dimensional gel electrophoresis have recently been reported. Some examples of early estimates of genome sizes are as follows: *Azotobacter chroococcum*, 1,940 kb (291); *Desulfovibrio gigas*, 1,630 kb (274); *Desulfovibrio vulgaris*, 1,720 kb (274); *E. coli*, 3,520 kb (274); and *Myxococcus xanthus*, 5,690 kb (384). Recent values (which still must be regarded as approximations) in closer agreement with results of PFGE are as follows: *Mycoplasma capricolum*, 742 kb; *Acholeplasma laidlawii*, 1,646 kb; *Haemophilus influenzae*, 1,833 kb; and *E. coli*, 4,399 kb (272). In comparison, the currently accepted value for the size of the *E. coli* chromosome is 4.7 kb (56, 324), and for *M. xanthus* the value is 9.45 kb (51).

PFGE (the so-called top-down approach) is conceptually clear-cut, is often straightforward in execution, and is regarded as the most definitive of the techniques used to estimate sizes of chromosomes (322). In practice, chromosomes are released from cells in agar blocks (to prevent mechanical shear) and then digested with a restriction endonuclease which cuts the genome infrequently. Often the products of this procedure are a dozen or several dozen large DNA fragments containing 50 kb or more. These linear fragments (which are too large to be sieved by conventional agarose gel electrophoresis) are successfully distributed by some form of alternating-field electrophoresis. The sizes of fragments are estimated, and the sum of the unique fragments is calculated. Size standards may be yeast chromosomes whose sizes have been independently estimated or concatemers of bacteriophage chromosomes. Recently, some investigators have reported that mobilities may be affected by the G+C content (272, 278); as a consequence, some current estimates may have to be refined.

When observation is confined to PFGE, the sizes of chromosomes per se are seen to range from slightly less than 600 kb in *Mycoplasma genitalium* to 9,454 kb in *Myxococcus xanthus* (Table 1). The smallest chromosomes, less than 1 Mb, are in bacteria lacking cell walls (viz., the *Mycoplasma* and *Ureaplasma* species) or having flexible walls (as in the spirochetes). At the other extreme are the chromosomes of bacteria which exhibit morphological differentiations (e.g., fruiting bacteria).

Size comparisons become much more complicated when observations originate from different techniques. Estimates of the size of the *Myxococcus* chromosome based on autoradiography and colorimetry range from 5,700 to 12,700 kb; a recent PFGE measurement places the size at 9,450 kb (51). Likewise, measurements of the size of the genome (chromosome plus accessory elements) of *Streptomyces coelicolor* by renaturation kinetics and the chromosome size by PFGE differ. The former places the size of the genome at approximately 10,750 kb (161), whereas the latter technique places the size of the chromosome at 7,000 to 8,000 kb (H. M. Kieser, T. Kieser, and D. A. Hopwood, personal communication). Renaturation kinetics of DNA isolated from cyanobacteria indicates that the largest genomes occur in *Calothrix* strains and contain as much as 12,800 kb (140); PFGE data are not yet available to substantiate whether (among currently measured procaryotic genomes) these are the largest.

In addition to determining the size of the bacterial chromosome, restriction fragments produced by rare cutters or more conventional restriction endonucleases can be used to determine the shape or geometry of the chromosome. An endless set of ordered fragments reflects a circular chromosome; indeed, for some bacteria, the large fragments produced by rare cutters have been ordered into circles (Table 1). For a few others, full chromosomal restriction maps are being constructed and perfected. In this so-called bottom-up approach, libraries of many small restriction fragments are generated by two or more restriction enzymes. The fragments can be ordered into contiguous sets (contigs) by any of several methods, among which are hybridization with probes that span adjacent fragments or generation of sufficiently dissimilar fingerprints of regional restriction maps to align overlapping fragments (63, 323). Although the rationale is clear, the technique can be laborious when the number of fragments to be ordered is large. Furthermore, some segments of the genome appear refractory to cloning (63). For example, there was some difficulty in cloning the *oriC* region of *E. coli* (30, 184, 186). Also, some gaps were difficult to resolve; only recently have the persisting gaps among *E. coli* contigs been closed (184). Ordered catalogs of conventional restriction fragments exist for *Mycoplasma pneumoniae* (371) and the W3110 (185, 186), BHB2600 (30), and 803 (184) strains of *E. coli* K-12. Because of the smaller number of fragments produced by rare cutters, the task of ordering the fragments is correspondingly simpler (Table 1). These measurements of shape based on ordering fragments supplement previous determinations of circularity based on one or more of the following observations: a single endless linkage group of chromosomal loci; overlapping patterns of marker transfer by high-frequency donors; and, for *E. coli*, direct imaging of intact chromosomal DNA by autoradiography (43).

Two possible exceptions to a single circular chromosome merit attention. One is *Rhodobacter sphaeroides*, which may have two distinct circular chromosomes (343, 344) (see

TABLE 1. Sizes of bacterial chromosomes measured by PFGE

Organism	Size (kb)[a]	Geometry	Reference
Anabaena sp. strain PCC 7120	6,400	Circle[b]	13
Bacillus cereus	5,700	Circle[b,c]	188
Bacillus subtilis	Incomplete	Circle[b]	359
Borrelia burgdorferi B31	950	Linear	83
Brucella abortus 544^T	2,600	ND[d]	5
Brucella melitensis 16 M^T	2,600	ND	5
Campylobacter coli	1,714	ND	50
Campylobacter jejuni UA580	1,721	Circle[c]	50
Campylobacter laridis UA487	1,451	ND	50
Caulobacter crescentus	4,000	Circle[b]	81
Chlamydia psittaci AB7, 1H, 1B	1,450	ND	95
Chlamydia trachomatis L2	1,450	ND	95
Clostridium perfringens	3,600	Circle[b,c]	46
Enterococcus faecalis	2,600	ND	35
Escherichia coli K-12	4,700[e]	Circle[b,c]	
Haemophilus influenzae V23	1,980	Circle[b]	178
Haemophilus influenzae Rd	1,834	Circle[b]	202
Haemohilus parainfluenzae	2,340	Circle[b]	177
Lactococcus cremoris	2,600	ND	35
Lactococcus lactis	2,500	ND	35
Mycoplasma gallisepticum PG 31	1,050	ND	278
Mycoplasma genitalium	585	ND	340
Mycoplasma hypopneumoniae strain J	1,140	ND	278
Mycoplasma iowae	1,280	ND	278
Mycoplasma mobile	780	Circle[c]	17
Mycoplasma pneumoniae	785[f]	Circle[b,c]	194
Mycoplasma synoviae WVU 1853	900	ND	278
Mycoplasma mycoides			
subsp. *mycoides* Y	1,240	ND	278
subsp. *mycoides* GC 1176-2	1,330	ND	278
Myxococcus xanthus	9,454	ND	51
Porochlamydia buthi	1,550	ND	95
Porochlamydia chironomi	2,650	ND	95
Pseudomonas aeruginosa	5,900	Circle[b,c]	292
Rhodobacter sphaeroides 2.4.1			
Chromosome I	3,046	Circle[b]	343, 344
Chromosome II[g]	914	Circle[b]	343, 344
Rickettsiella grylli	2,100	ND	95
Rickettsiella melolanthae	1,720	Circle[c]	95
Staphylococcus aureus	2,860	ND	366
Staphylococcus aureus	2,748	ND	263
Streptococcus sanguis	2,300	ND	35
Streptococcus thermophilus	1,700	ND	35
Sulfolobus acidocaldarius	ca. 3,100	Circle[b]	380
Thermococcus celer	1,890	Circle[b]	250
Ureaplasma urealyticum 960^T	900	Circle[b,c]	55

[a] Sizes may be the average achieved with several rare-cutting restriction endonucleases.

[b] Geometry was determined by hybridization to electrophoretically distributed macrofragments, using as probes macrofragments produced by a different rare-cutting restriction enzyme, small linking fragments containing the rare-cutting site, fragments produced by random priming of macrofragments, specific gene probes or probes of transposons previously inserted into specific genes, or probes complementary to one end of a macrofragment hybridized to partially digested fragments.

[c] Geometry was determined by constructing restriction maps by full or partial digestion with one or more rare-cutting enzymes in either single or double digests and determination of fragment sizes by either one- or two-dimensional gel electrophoresis.

[d] ND, Not determined.

[e] The sum of *Not*I fragments from *E. coli* K-12 has been estimated to be 4,550 bp (324). More refined analyses with complete sets of contigs produced by six base cutters has yielded an estimate of 4,700 kb (184, 294).

[f] By using a complete set of *Eco*RI contigs or a complete set of (unordered) *Xho*I fragments, the size of the *M. pneumoniae* chromosome has been calculated to be 835 or 849 kb, respectively (371).

[g] Discussed in the text (When Does a Plasmid Become a Chromosome?).

When Does a Plasmid Become a Chromosome?); the other is the spirochete *Borrelia burgdorferi*, which appears to have a linear chromosome in addition to plasmids which are linear and which have covalently closed ends (15). (Linear plasmids have been identified in *B. burgdorferi* [15], *B. hermsii* [183, 270], and several *Streptomyces* spp. [182].). When the chromosomal DNA of *B. burgdorferi* is released into agar blocks, digested with both *Mlu*I and *Sma*I, and subjected to PFGE, a restriction pattern compatible with a linear chromosomal configuration with unique ends (83) is produced. However, hybridization experiments have not yet been reported which would establish whether both enzymes generate fragments characteristic of the ends of linear chromosomes.

In summary, chromosome sizes estimated by PFGE and shape determined from ordered libraries of restriction fragments indicate that bacterial chromosomes are commonly circular and contain 1 to 9 Mb (Table 1).

Physical and genetic maps. Data on the genetic and physical maps and nucleotide sequences of *E. coli* genes are accumulating rapidly. An updated genetic map was completed recently (11). As noted above, detailed restriction site maps exist for *E. coli* K-12 strains W3110, BHB2600, and 803. All nucleotide sequences of segments of the *E. coli* chromosome in the GenBank and EMBL data bases were assembled recently by Kröger et al. (195). At least three laboratories (those of G. Church, F. Blattner, and K. Isono) are currently engaged in determining total DNA sequences of three *E. coli* K-12 strains. Alignment and reconciliation of the existing data have begun. Rudd et al. (294) aligned restriction maps with the 1983 *E. coli* genetic map and found an overall colinear relationship with minor departure from an ideal line. Many of the nucleotide sequences in the major data bases have been positioned on the chromosomal restriction map, with occasional disparities in the number and location of restriction enzyme sites.

The Kohara restriction map of *E. coli* K-12 was constructed for strain W3110 (185, 186). There is a well-known inversion in this strain relative to many other *E. coli* K-12 strains (145). With that inversion reversed, the W3110 sequence is serving at present as the de facto standard for the *E. coli* K-12 physical map. We need to be aware that differences exist between strain W3110 and other *E. coli* K-12 strains. More genetic variation is present among strains or isolates of *E. coli* K-12 than one might expect. There are single-base-pair differences throughout the genomes of different strains of *E. coli* K-12 that cause differences in numbers and locations of restriction sites. Detailed comparisons are beginning to be made between strain W3110 and other K-12 strains such as BHB2600 (30), 803 (184), and MG1655 (62). In one map region, as many as 6 of 22 restriction sites differed between two K-12 strains (62). In another set of comparisons, more variation was reported in the form of rearrangements, insertions, and deletions than in nucleotide differences (233). Differences have even developed in cultures of strain W3110 that were maintained for years in separate facilities (62, 251). Careful strain histories and identifications are becoming important in establishing a standard data set for the *E. coli* K-12 chromosome at the physical map and sequence level.

How far advanced are the physical and genetic maps of bacteria other than *E. coli*? A full restriction map has been assembled for *Mycoplasma pneumoniae* (194, 371). Detailed genetic information has been assembled for *Salmonella typhimurium* (297). Significant strides are being made for other bacteria as well. A detailed genetic map has been made

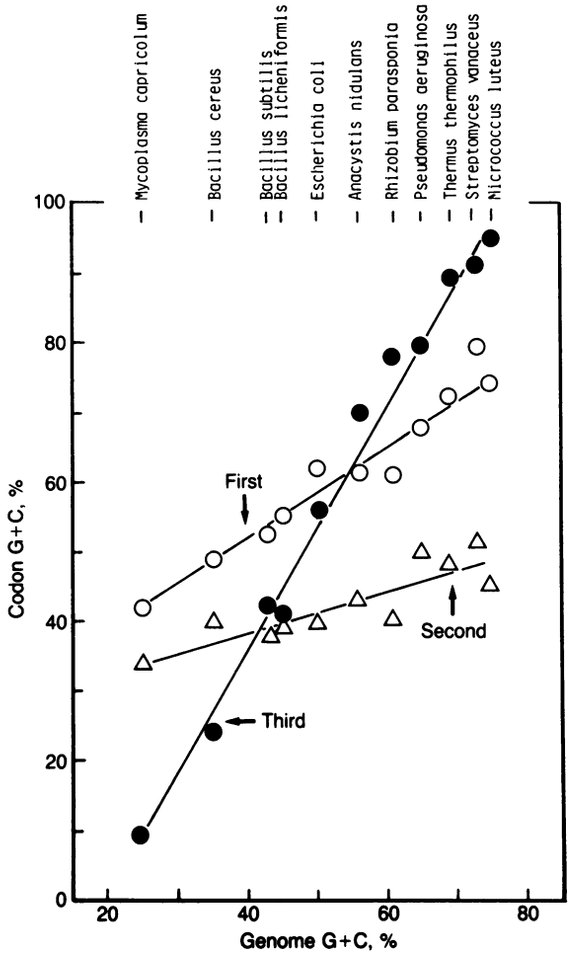


FIG. 1. Correlation of the G+C content between total genomic DNA and the first, second, and third codon positions of sequenced genes. Sources of data and details are in the original article. Reprinted from reference 245 with permission from the author.

for *Bacillus subtilis* (269). The current genetic maps of two *Pseudomonas* species and *Streptomyces coelicolor* are presented and discussed below (see Nonrandom Patterns of Gene Linkages). Genetic maps at earlier stages of development have been compiled for other *Bacillus* spp. (252, 358), *Rhizobium* spp. (192, 252), *Erwinia chrysanthemi* (159), a marine *Vibrio* species (162), *Rhodobacter sphaeroides* (343, 344), *Clostridium perfringens* (46), and others (for citations to earlier work, see reference 287).

Nucleotide composition: mutational and context bias. The nucleotide composition of procaryotic genomes varies greatly, ranging from as low as 25% G+C in *Mycoplasma capricolum* to as high as 75% G+C in *Micrococcus luteus* (23, 245). Comparison of nucleotide sequences in bacteria reveals that the codon composition correlates with overall genomic composition and that differences in codon compositions occur to the greatest extent in the third base position (Fig. 1), a latitude that is an expression of the redundancy of the genetic code. Such biases reflect the preferential use in a species of one or two codons among as many as six specifying a common amino acid.

Comparison of the nucleotide composition of homologous genes clearly establishes that differences in nucleotide composition cannot be attributed solely to the differences in gene contents among bacteria. For example, among 1,188 homologous codon sites of genes encoding ribosomal proteins in *Mycoplasma capricolum* and *E. coli* (245), 199 codons (17%) were found to be identical, 360 (30%) were synonymous, and 169 (14%) specified conservative amino acid substitutions. Approximately 81% of the synonymous codons had differences only in position 3, with *E. coli* favoring the presence of a guanine or cytosine residue and *M. capricolum* favoring an adenine or thymine. Furthermore, among the conservative amino acid substitutions, 67% of the differences were the presence of an adenine or thymine in position 1 or 3 of codons in *M. capricolum*. Among the remaining 459 codons in *M. capricolum* which specified amino acids entirely distinct from amino acids in corresponding positions in *E. coli* proteins, 256 were richer in adenine or thymine when compared with *E. coli* (245). Evidently, *M. capricolum* favors A and T whenever circumstances permit.

Codon bias seems to originate through two principal mechanisms: (i) an intrinsic mutational bias (341) and (ii) a context bias (122, 383). The former is viewed as an inherent feature of the replication process of a particular bacterium. An important aspect of mutational bias is that it is manifest in the absence of selection. Two genes in *E. coli*, *mutT* (4) and *mutY* (254), affect frequencies of transversions. Specifically, MutT mutants have a 1,000-fold increase in the rate at which A · T-to-C · G transversions occur; a product of the *mutY* locus is specific for G · C-to-T · A transversions. The equilibrium between these two forms of change will affect the final composition of DNA replicas (341).

Notably, the favored codons correlate with the abundance of cognate tRNAs in the organism (163). This correlation contributes to the notion that there are optimal codons which "are most accurately and/or efficiently recognized by the most abundant tRNAs," a circumstance which is measured, for example, by a codon adaptation index (316). Further analysis has demonstrated that in some bacteria, e.g., some enteric bacteria (316) but not *P. aeruginosa* (372), optimal codons are used with a higher incidence in highly expressed proteins than among proteins that occur rarely (113). When homologous genes are compared, such constraints on codon usage are seen to be preserved in related organisms (315, 316).

The intrinsic mutational bias may not be uniformly expressed between the origin and terminus of replication. When 67 pairs of homologous genes in *E. coli* and *Salmonella typhimurium* were compared, a group of six genes were recognized which had a much lower occurrence of synonymous codons than would have been expected from the extent of optimal codon usage (i.e., the codon adaptation index). These six genes were all within 5 min of the origin (317).

Context bias reflects relationships of adjacent codons (122, 383). The number of possible pairs of adjacent sense codons is large, viz., 61^2 or 3,721. Accordingly, recognition of a context bias depends on the availability of a large number of established sequences such as is available for *E. coli*. Examination of 237 genes, each of which has more than 100 bp and which collectively make up nearly 236,000 bp, has allowed analysis of context bias. For such a large data set, the expected occurrence of a pair of codons can be compared with the observed occurrence of each pair. Statistical analysis (χ^2) shows that the distribution of codon pairs is markedly different from that of a set of random codon pairs having the same overall composition (122). Pair-by-pair analysis shows that many codon pairs are strikingly overrepresented while others are underrepresented

(even after corrections for the occurrence of amino acid pairs and low frequency of amino acid usage). The observation indicates that the composition of one codon is not unrelated to the compositions of adjacent codons (122, 383). This effect is greatly diminished for pairs of codons separated by an intervening codon and is essentially absent for codon pairs separated by two or more intervening codons. These data suggest that context bias may reflect bindings at the A and P sites of ribosomes. Other analysis and translations of synthetic mRNAs suggest that context bias may, in addition, reflect (i) association of ternary complexes containing a charged tRNA, polypeptide elongation factor Tu, and GTP, or (ii) interactions among tRNAs (198, 325). (Recently, Andersson and Kurland critically reviewed the assumptions and experimental evidence relating codon usage and translation rates [10].)

In general, the translation machinery in any one species may have evolved some differentiating features; a manifestation of such differences would be reflected in the G+C composition of codons. Cumulatively, this preference is seen in the overall G+C composition of the chromosome.

Recurrent nucleotide sequences. Small repeats, some of which occur outside open reading frames (ORFs), are a recurrent motif in procaryotic chromosomes. They serve as sites of protein-DNA interaction; as a modified site which identifies the template strand during semiconservative replication; as signals for recombination, transposition, inversion, and excision; and possibly as sites of slippage for restoring the structure of mutated genes. Other postulated functions are demarcation of chromosomal regions which are folded or condensed and partitioning of replicas subsequent to chromosomal duplication. Thus, these repeats may be regarded as inherent features of the chromosome per se rather than as aspects of genes residing in the chromosome.

Repeated extragenic palindromes (REPs) were first recognized as an intriguing consensus sequence recurring in operons outside coding sequences (335) and often existing near sites where crossing over had occurred (141). In the early characterization of REPs, sequences of ca. 70 bases which could exhibit a variety of dyadic secondary structures were recognized; subsequently, REP was equated with a consensus sequence with 38 bp (or fewer) (106, 385). Although nearly all REPs are extragenic, recently a 30-bp sequence which is highly similar to a REP has been identified within the 3′ end of the *E. coli nrdA* gene (235). When gaps were introduced into the 30-bp sequence so that it could be aligned with a 35-bp consensus of a REP, the sequence was seen to match the consensus at 26 sites. Clusters of two, three, or four REPs (designated a REP element) often occur; each REP sequence in an element is separated from an adjacent sequence by fewer than 25 bp. Within the REP elements, the REP sequences alternate in left-to-right orientation. The dyad nature of the sequences and the alternation in orientation offer opportunities for multiple base-paired DNA structures. The REP elements are estimated to have 100 or 200 occurrences in the *E. coli* and *S. typhimurium* chromosomes, representing about 0.5% of the total (385).

REP sequences bind both HU and DNA gyrase, suggesting a role of REP in the folding of the nucleoid (385). Gel retardation studies with pure gyrase and DNA fragments of a few hundred nucleotides containing two or four REPs indicated that the enzyme has affinity for REPs; the affinity increased in REP elements containing more REPs. The binding appeared to be mediated by the A subunit of the topoisomerase. Additional analyses suggested that gyrase may cleave DNA immediately adjacent to a REP. DNA footprinting indicated that a sequence of approximately 200 bp including REPs was protected by gyrase binding (385). Similarly, REP sequences (also called PU sites) have been found to bind DNA polymerase I as measured by protection against exonuclease III, gel retardation, and DNase I footprinting (105). Involvement of DNA polymerase suggests a role of the REP sequence in replication or repair.

REPs occur in both *E. coli* and *S. typhimurium*. The REP of *S. typhimurium* appears to have an extra guanine residue compared with the homologous structure in *E. coli* (106). Also notable is that the REPs occupy very different positions in the two species in relation to common loci; i.e., genetic location is not conserved. A search in *Bacillus subtilis* for a repetitive stem-loop structure with no more than 39 bases and no fewer than seven consecutive base pairs in the stem failed to detect such an organization among 29,417 bp in 33 sequences (106).

An unanticipated consequence of nucleotide sequencing in *E. coli* in the vicinity of the *iap* locus (which is located at 59.1 min and which codes for an alkaline phosphatase isozyme) was detection of a 29-bp sequence that recurs 14 times (246). Thirteen of the repeats differ from the consensus sequence by only 1 or 2 bp. Eleven of the repeats are separated from one another by intervals of 32 bp; the remaining two repeats are separated by 33 bp. Another reiterated structure with similar features exists at 59.6 min. In this instance, a similar 29-bp sequence separated by 32 or 33 bp recurs seven times. Southern blotting with a 29-bp synthetic probe revealed hybridization with chromosomal DNAs from *E. coli* C600, *E. coli* Ymel, *S. typhimurium* TA1535, and *Shigella dysenteriae* Sh, but not with chromosomal DNA from *Klebsiella pneumoniae* or *P. aeruginosa* (246). These data indicate that components of the striking 815-bp element in *E. coli* appear to be shared among some close relatives.

Hybridization experiments with the cyanobacterium *Calothrix* strain PCC 7601 suggested that highly repeated sequences were present in the chromosome (230). Sequence analysis confirmed the presence of three distinct short tandemly repeated repetitive sequences. Each consists of a heptanucleotide repeated 3 to 11 times at as many as 100 chromosomal locations, and all of them occurred outside ORFs. Their function remains unknown.

Chi is another repeated sequence found in the *E. coli* chromosome and some of its accessory genetic elements. Chi sites are 8-bp sequences, 5′-GCTGGTGG-3′, which stimulate homologous recombination mediated by the RecBCD enzyme in conjunction with the RecA and other proteins (326–328). Biochemical analyses indicate that Chi sequences are the recognition site for cleavage of DNA by the RecBCD endonuclease and generation of a single-stranded length of DNA, a precondition for synapsis. Genetic analyses (in bacteriophage lambda) indicate that the effect of Chi is polar, stimulating recombination at the 5′ side of Chi but not far to the 3′ side (82). Homologous recombination is enhanced in the immediate vicinity of a Chi site (227), although the effect diminishes exponentially with distance by a factor of approximately 2 for each 2.5 kb.

Chi sites occur in *E. coli* and a number of related organisms (306). The ability of RecBCD-like enzymes from numerous enteric bacteria to cut at Chi and to promote Chi-stimulated recombination when cloned into *E. coli* deleted for *recBCD* suggests that Chi is a recombination-enhancing site throughout the members of the family *Enterobacteriaceae*. By contrast, the *Pseudomonas* enzyme does not activate Chi, but may recognize a different sequence for enhancement of recombination (306). The frequency of Chi

sites in *E. coli* was estimated by cloning 20 chromosomal segments of known size into lambda and assaying for homologous recombination (227). The absence of homologous recombination was equated with the null class in the Poisson distribution. Correction for polarity was included when the effect of inverting Chi sites was recognized. These considerations allowed an estimate of one Chi site per 5 kb. Using a current measurement of the size of the *E. coli* chromosome and assuming a completely random occurrence of bases (i.e., disregarding the constraints on DNA composition that exist in parts of the chromosome), one can calculate that approximately 940 Chi sites might be present in the chromosome. Recently, approximately 1.2 Mb of DNA sequences from *E. coli* was scanned for the occurrence of Chi; 220 Chi sites were found (S. Ganesan and G. S. Smith, personal communication). These data indicate that a Chi site occurs on average, in line with expectations, once every 5.5 kb. However, a cluster of 22 Chi sites was observed in a 0.5-min segment (ca. 23 kb) near *oriC*. Thus, in this region, Chi sites occurred five times more frequently than on average. Although the abundance of Chi sites has been estimated, their location relative to ORFs is not yet established.

DNA adenine methylation sites (Dam sites) are important regulatory signals composed of the 4-base sequence, 5′-GATC-3′ (232, 336). One function is to distinguish the conserved template strand from the newly synthesized strand during semiconservative replication, gene conversion, and repair. The distinction is straightforward: addition of a methyl group to the N-6 position of adenine has occurred in the former structure and has yet to occur in the latter. Knowledge of this function does not contribute to an understanding of the distribution of Dam sites. If their occurrence were solely random, they would occur, on average, once every 250 bp. However, in the 245-bp sequence which defines *oriC* in *E. coli*, there are eight Dam sites (232). In 350 bp flanking *oriC*, there are an additional 12 Dam sites (232). Furthermore, many of these sequences are conserved among origins of replication of other enteric bacteria. In addition to their occurrence at *ori*, Dam sites (as suggested from Dam^- mutations) are associated with genes involved in the SOS response, transposon function, a number of structural genes, and bacteriophage infection (336).

Some repeats, e.g., $(CAAT)_n$ recently observed in *Haemophilus influenzae* (367), may regulate alternate gene expressions by shifting reading frames. Hence, these recurrent nucleotide sequences are aptly thought of as contributing to alternate epitopes rather than alternate chromosomal structures (367).

Recently, the chromosome of *Mycoplasma pneumoniae* has been shown to contain several types of repetitive elements each of considerable length. One, *SDC1*, contains 424 bp, may occur as many as eight times, and has no known function (55).

As nucleotide sequences of chromosomal segments accumulate and are more fully analyzed, the relation of small and moderate-sized repeats to chromosome structure in general and coding sequences in particular will emerge.

Summary. The bacterial nucleoid contains hundreds to thousands of genes arrayed in species-specific stable linkage groups. Through association with DNA-binding proteins, the large, typically covalently closed circular molecule exists in situ as a highly compacted structure. Nonetheless, the structure is dynamic; the extents of higher-order structure (secondary, tertiary, and quaternary) vary with environmental and physiological conditions. A variety of nucleotide sequences ranging from 4 to approximately 40 bases exist as repeated sequences; the extents of repetition vary from dozens through hundreds; some of these sequences occur largely or exclusively outside ORFs; some of the repeated sequences may contribute to chromosomal conformation. The G+C content varies among bacteria, and preferred codon usage varies with G+C composition. The variation in the G+C content reflects inherent mutational pressure and probably the structure of translational machinery.

Effective Abundance of Gene Copies

The number of copies of a chromosomal gene may differ because of the position of the locus along a gradient from the origin to the terminus, because additional copies of the gene occur on an accessory genetic element or because there are many replicas of the chromosome in the cell. Furthermore, the effective abundance of gene copies may be affected by organizational states that influence the expression of genes, e.g., sequestration of a chromosome or the occurrence of archived DNA within a chromosome.

Temporal sequence of replication. Replication of the circular chromosome of procaryotes often occurs bidirectionally from the origin and progresses more or less uniformly around the two arms of the chromosome to the terminus. Thus, during the time that the chromosome is being replicated, some positions will already have been replicated whereas others will not have been. By this circumstance alone, some loci will be more abundant than others. The extent of difference is more pronounced when the length of time required for cell duplication is much shorter than the time for chromosomal replication. According to the Helmstetter-Cooper I + C + D model relating chromosome replication and cell division (138), rounds of chromosome replication may be initiated several generations before distribution of the replicated product. Models of this process reveal that there may be four times as many copies of a locus near the origin than at a locus near the terminus.

Presence of accessory elements. The term meroploidy refers to a condition of partial ploidy that affects only one or a few genes as opposed to multiples of entire chromosomes. Meroploidy results when genes are present on an accessory element such as a plasmid as well as on a chromosome. Natural examples of this condition are large plasmids of few copies such as F′ or R′ plasmids, or small multicopy elements such as ColE1 plasmids. Artifactual forms occur when replicons are constructed with gene cloning techniques. The numbers of copies of a gene, the degree of meroploidy, is determined by the numbers of copies of the accessory elements in the cell.

Chromosome reiteration or redundancy. In some species the number of copies of a gene may be very large simply because there are many copies of chromosomes present. Estimates from *Azotobacter vinelandii* indicate that there may be as many as 40 to 80 identical copies of the chromosome in vegetative cells (277, 295); encysted cells have approximately one-quarter as much DNA. The estimates are based on a number of observations. One technique probed for the presence of specific genes in two types of digested samples. One sample was the *A. vinelandii* chromosome; the other was mixtures constituted from various amounts of the isolated gene added to calf thymus DNA. Between 40 and 80 copies of the *leuB* gene or the *nif* gene had to be added to the reconstituted mixture to achieve binding of probes to an extent judged to be equivalent to that seen with native chromosomes. A second technique was to clone a β-lactamase gene into the *A. vinelandii* chromosome and probe for

its abundance after many rounds of replication and segregation. This analysis plus electrophoretic examination of the size of fragments containing the *bla* gene indicate that the 80 copies of the gene are more likely to exist on many chromosomes rather than to exist tandemly reiterated on a single large chromosome (277). Furthermore, renaturation kinetics indicated that the complexity of *A. vinelandii* DNA is similar to that of *E. coli*, whereas results of a diphenylamine assay indicated that *A. vinelandii* has 40 times as much DNA. Finally, the sizes of isolated folded chromosomes appear identical or nearly so (295). Altogether, the results indicate that the ploidy of the *A. vinelandii* genome is at least 40 and may be substantially higher.

Consistent with the results with *A. vinelandii*, vegetative cells of *A. chroococcum* may contain 20 to 25 copies of the genome (291). These estimates were obtained by comparing the size of the genome with the DNA content of cells. The size of the genome was equated with the sum of the sizes of unique restriction fragments generated by two-dimensional electrophoresis.

Two-dimensional electrophoresis has also been applied to *Desulfovibrio gigas* and *D. vulgaris* to estimate the chromosome number (274). For these sulfate reducers, the analyses were performed with organisms grown in chemostats and batch cultures. In the chemostat, *D. gigas* experiencing nitrogen limitation had, on average, 9.1 chromosomes per cell, whereas in batch cultures, the organism registered 17.2 chromosomes per cell. In continuous culture, *D. vulgaris* contained, on average, 4.1 chromosomes per cell.

The conclusions regarding multiplicity of chromosomes are dependent upon successful separation of chromosomal and plasmid DNA, accurate estimates of abundance of DNA per cell, and realistic estimates of the sizes of chromosomes. Error in these measurements may have some effect on the estimated number of chromosomes per cell; nonetheless, the conclusion that some bacterial cells have many copies of a chromosome seems sound.

Chromosome inactivation. Inactivation of entire chromosomes in diploid cells of eucaryotic organisms is relatively common and applies, for instance, to one of each pair of X chromosomes of female mammals (condensed into inactive Barr bodies) (222). The DNA has been changed, or marked, in the inactive chromosome. The marking is a form of modification of the DNA (imprinting) that prevents expression but leaves intact the nucleotide sequence, permits replication, and provides for perpetuation of the imprint (for discussion, see reference 349).

Chromosome inactivation seems to occur in the bacterial world as well. The phenomenon has been described for *B. subtilis*, and studies in early phases indicate that a similar phenomenon also exists in *A. vinelandii*. Ten years ago, Hotchkiss and Gabor (157) reported that one parental set of genes in diploid strains of *B. subtilis* was not expressed. Heterozygous diploid strains of *B. subtilis* can be formed by fusion of protoplasts of genetically different parental strains. Such diploid populations were found to contain some members in which auxotrophic markers of one chromosome were not complemented by the wild-type allele of the other chromosome; these are called noncomplementing diploids. However, the unexpressed genes were clearly present in the noncomplementing diploid because the hidden parental genotype segregated readily. The haploid segregants exhibited normal, expressed phenotypes of the previously unexpressed chromosome.

The chromosome inactivation in *B. subtilis* involved suppression of expression of constituent prophage genes as well. No transcript of the repressor gene of prophage ϕ105 was detected, and the prophage in the inactive chromosome did not express superinfection immunity. Also, functions required for prophage induction were not operational as there was no induction of the prophage even though repressor was absent (120). No free phage was found in Ncd cultures, only uninduced prophage. Hybridization signals for the prophage DNA and a prophage-carried chloramphenicol acetyltransferase gene were orders of magnitude less than signals from the normal chromosomal DNA of a *B. subtilis* lysogen (L. Hirschbein, personal communication). Evidently, the Ncds contain DNA that is physically as well as functionally abnormal.

B. subtilis diploid populations contain two types of noncomplementing diploids, those that were unstable with high rates of segregation of active haploids (10^{-2} to 10^{-4}), and those that were stable with low rates of segregation ($<10^{-7}$). DNA of the former seems to be only loosely inactivated, and that of the latter was more firmly altered (118). DNA extracted from the unstable strains was poor in transformation, but could be reactivated by simple purification of the DNA (31); DNA extracted from the stable strains could not be reactivated this way (118), but regained genetic activity if passed through an unrelated bacterium such as *E. coli* (217).

With respect to physical properties, two classes of nucleoids were found in *B. subtilis* diploids, a class of nucleoids with normal sedimentation properties and a class of faster-sedimenting nucleoids. When assayed genetically, the faster-sedimenting nucleoids had poorer transforming activity than the normal nucleoids, and what little genetic capacity was present reflected the expressed chromosome, not the inactivated chromosome. It seems that the process of chromosome inactivation condensed the nucleoid in a way that made the DNA inoperative in the transformation process (119). The biochemical mechanism of the chromosome inactivation, or imprinting, has not yet been worked out. The *spoA* gene, one known to have pleiotropic effects and to play a role in condensation of chromosomes in sporulation, seems to be important in establishing and maintaining chromosome inactivation (349).

A similar phenomenon of chromosome inactivation may also operate in *A. vinelandii*. Judging by DNA content, *A. vinelandii* is highly polyploid (see Chromosome Reiteration or Redundancy, above). In a sense, a high level of ploidy fits with the large cell size and high level of metabolic activity of *Azotobacter* organisms (181). However, not all of the DNA in the cell seems to be active. In contrast to an apparent chromosome number of 40 to 80, functional measures tended to place the ploidy at a much smaller number (57). Recessive mutations were expressed without undergoing the theoretical number of segregating divisions corresponding to 40 to 80 chromosomes, as if not all of the chromosomes present were functionally active.

Another example of reversibly inactivated genetic material was reported for a P22 prophage in *Salmonella typhimurium* (71). After its purine metabolism was stressed, washed cells of *S. typhimurium* behaved as if the bacterial chromosome harbored a P22 prophage in an unexpressed, archived form. The strain of *S. typhimurium* was not believed to be lysogenic and was not immune to P22 superinfection; nevertheless, under a particular regimen of medium shifts, occasional mature P22 phages were released, revealing the presence of a cryptic prophage in the genome. DNA hybridization behavior of the prophage in the bacterial DNA was not normal. This apparent archiving of the P22 prophage genome in its host chromosome has proven difficult to

analyze, but it may turn out to share features with the inactivation of bacterial chromosomal DNAs.

Summary. Several factors affect the number of functional alleles of a gene per cell. These include the speed of cell growth, which affects the gene dosage factor attributable to multiple replication forks; the copy number of any accessory genetic element(s) bearing the gene; the ploidy of the chromosome itself; and the fraction of the chromosomes present that is active.

Plasmid-Chromosome Relationships

Plasmid-chromosome interactions. Bacteria may contain extrachromosomal (283) or accessory (45) genetic elements; these structures include plasmids, bacteriophages, transposons, and insertion sequences. Two distinguishing features of accessory genetic elements are that they are autonomous and expendable (45). Autonomy means that the number of such elements can be uncoupled from the number of chromosomes; expendability means that the genetic content of the accessory genetic element is not absolutely essential to growth of the organism in its usual environments. A plasmid which ceases to be both autonomous and expendable may be considered an additional chromosome.

The phenotypes specified by genes on plasmids may confer an advantage to organisms which bear the plasmids. At the same time, maintenance of the plasmid and the expression of the genes (some of which may interfere with organismic processes) are costs associated with the presence of plasmids (109, 137). A balance of these effects has been measured in *E. coli* with small, nonconjugative plasmids (viz., pACYC184 and other ColE1 derivatives) in both serial (34) and continuous (137) cultures. As expected, plasmids conferring a specific advantage were maintained (i.e., did not segregate) in the presence of selection; by contrast, in the absence of selection, bacteria with plasmids were initially burdened by their presence and grew more slowly than plasmidless bacteria did. After extended growth, however, the presence of a plasmid appeared to be advantageous in either the presence or absence of selection. The heritable change in the relation between organism and plasmid could be attributed to genetic changes in the chromosome more so than to changes in the plasmid (34, 137). In one instance, *E. coli* B which had been cultivated with a pACYC184 for about 500 generations had a growth advantage over the parental strain of the organism, which had not been cultured with the plasmid but which had been subsequently transformed with the same plasmid (34). Furthermore, the evolved strain cured of the plasmid and transformed with the base-line plasmid outcompeted the evolved strain which had spontaneously lost the plasmid (34). These results may be interpreted in terms of periodic selection with an additional premise, viz., that the fitter mutations are more likely to occur on the chromosome, which, in the present example, is approximately 1,000 times larger than the plasmid.

Maintenance of a 52-kb plasmid (R46) in *E. coli* in carbon-limited or phosphate-limited continuous culture showed that some components of plasmids were preserved while others were lost (109). In the absence of selection, plasmid R46, which confers resistance to ampicillin, streptomycin, sulfonamide, and tetracycline, was maintained for extended periods; however, examination of the antibiotic sensitivities revealed that some drug resistance determinants were lost (109). In continuous cultures inoculated with a mixture of organisms, strains containing the modified plasmids displaced the parental strains containing the intact plasmid.

Another study compared the fate of plasmids pBR322 and pBR325 in two different *E. coli* hosts in nonselective medium in continuous culture limited either by glucose or by ammonium chloride at two different dilution rates, for up to 100 generations (249). Stability of the plasmid depended on the plasmid, the host bacterium, the substance limiting growth, and the growth rate. Whereas plasmid pBR322 was maintained under all growth conditions tested, the related plasmid pBR325 was unstable to some degree under all conditions except for a low dilution rate with limitation of ammonium chloride. Kinetics of plasmid segregation varied with host and culture conditions. Of the three drug resistance markers on pBR325, two followed the plasmid, but one, tetracycline resistance, was lost from the plasmid preferentially under some cultural conditions (249).

The relation of plasmids to chromosomes in continuous culture has been reviewed (79). The relationships appear to be complex and are highly sensitive to growth conditions, identity of the host, and identity of the plasmid.

A complicating circumstance in the relationships between plasmids and chromosomes is that some plasmids can integrate into and excise from chromosomes. Examples include the well-known integration and excision of the F plasmid in *E. coli* and integration of plasmid RP4 into many bacterial genomes such as *Myxococcus* spp. (169), pMEA100 in *Nocardia mediterranei* (223), pSAM in *Streptomyces ambofaciens* (267), and pSE101 in *Saccharopolyspora erythraea* and *Streptomyces lividans* (39). An especially striking example of this condition is manifest in interspecific crosses between *Streptomyces* spp. In such crosses, stable genetic elements in the donor species become plasmids in the recipient strains (153, 258). For example, a cross in which *S. coelicolor* A3(2) served as the donor and *S. lividans* 66 functioned as the recipient revealed the presence of a 17-kb plasmidogenic locus in the former (258). Upon transfer, the locus could exist in the recipient in three forms: transiently as a 17-kb plasmid containing all the genetic content present in the original locus, integrated at a specific site in the *S. lividans* chromosome (203, 259, 260), or as one of several smaller plasmids, 11 to 14.5 kb in size (153). In the smaller plasmids, genetic information for integration and maintenance was deleted. This example illustrates the complexity of plasmid-chromosome interactions and the genetic variety brought to bacterial genomes by accessory elements.

When does a plasmid become a chromosome? Some bacterial species, e.g., *Pseudomonas* and *Rhizobium* species, contain extremely large plasmids. These plasmids often specify traits which are species-specific characteristics.

Members of the genus *Pseudomonas* have a prodigious capacity to degrade organic compounds including aromatic ones. Some of the catabolic capacities are specified by large degradative plasmids, which, in some instances, exceed 500 kb (94). Catabolism of toluene and xylene, for example, is mediated by enzymes of the so-called *meta* pathway. The *xyl* genes specifying these enzymes are found most often on TOL plasmids; TOL plasmids vary in their genetic composition and may be as large as 117 kb (94). Recently, a 39-kb cluster of *xyl* genes and flanking segments has been observed to function as a 70-kb transposon (353, 354). *xyl* genes not only transpose among plasmids, but also sometimes integrate into the chromosome. For example, a 56-kb segment of the TOL plasmid, pWWO, carrying the *xyl* genes has been found integrated into the chromosome of one strain of *P. putida* MW1000 (320). Furthermore, some *Pseudomonas* spp. have the capacity to degrade chlorinated cyclic compounds, in some instances through cometabolic utilization of

enzymes that catalyze reactions with nonchlorinated analogs. Once again, genes conferring such activity (*bph*) have been located both on chromosomes (102) and on plasmids (101). These observations support the notion that genes specifying phenotypically important features of a genus need not be exclusively chromosomal or plasmid borne.

Symbiotic *Rhizobium* species harbor plasmids (pSym) which contain the genetic information for nodulation (*nod* genes) and nitrogen fixation (*nif* and *fix* genes). The plasmids range in size from 180 kb in *Rhizobium trifolii* through 1,600 kb in *R. meliloti* (42, 165). For the latter organism, the megaplasmid is an invariant species characteristic (14). All the *nod*, *nif*, and *fix* genes in *R. trifolii* exist in a 32-kb cluster (165). Repeats of some *nif* and *fix* genes exist at other locations in some pSym plasmids, but for the most part, the genetic function of the bulk of pSym sequences is unknown. Notably, the order of *nif* loci is largely the same in *Rhizobium* and *Klebsiella* species despite the presence of these genes on a plasmid in *Rhizobium* species and on the chromosome in *Klebsiella* species. Furthermore, the genes are regarded as homologous since *Klebsiella* probes are characteristically used to locate *nif* genes on pSym fragments generated with restriction endonucleases (165).

The occurrence of phenetically important determinants on large plasmids raises the formal possibility that these structures are either incipient, established, or transient chromosomes. A critical issue can be identified: are there instances in which essential genes are distributed on two separate genetic elements? If the answer is ever affirmative, neither of the genetic elements can be regarded as merely accessory. Furthermore, neither could be regarded as completely autonomous: replication of all essential genetic elements would have to be coordinated. Both of the genetic elements would have the status of chromosomes.

Suwanto and Kaplan (343, 344) have raised the possibility that *Rhodobacter sphaeroides* contains two chromosomes as well as plasmids. Standard procedures for isolation of plasmids yielded four plasmids containing approximately 100 kb and a fifth containing 42 kb. Digestion of the remaining chromosomal DNA with several rare-cutting enzymes and ordering the large restriction fragments resulted in disposition of the fragments in two independent circles, one of 3,046 kb and the other of 914 kb. Hybridization results indicated that the latter structure has two *rrn* loci while the former has one. Furthermore, the smaller structure contains genes specifying metabolically essential enzymes; however, duplicates or counterparts of these genes also exist on the larger structure. By using aged cells as a source of genomic DNA, two bands were evident in PFGE gels: a diffuse one and a sharp one. The diffuse band was interpreted as being derived from the 3,046-kb chromosome; the sharp one was interpreted to be a linearized form of the 914-kb structure. Southern hybridization of several probes to the DNA in these bands indicated that the bands contained the loci expected for the two structures on the basis of restriction analysis (343, 344).

Thus, collectively, restriction mapping, probing, and physical isolation techniques allowed the identification of two large circular structures in *R. sphaeroides*. Although each structure contains essential genes, only the larger structure has unique copies of essential information. The smaller structure has genes such as *rrn* loci that are also represented on the larger chromosome. Whether the structures should both be considered chromosomes depends on whether they are both required for normal growth. Further investigation of these uniquely interesting bacteria will provide more information on this open question.

Plasmids can be engineered to contain essential genes. If an essential gene is removed from the chromosome and concurrently introduced into a plasmid, the plasmid becomes indispensable. Such a construction has been achieved with *E. coli* AB1157 (273). The *ssb* gene (specifying a single-stranded DNA-binding protein) was eliminated from the chromosome and replaced with a gene conferring kanamycin resistance. The strain harbored pACYC into which a copy of *ssb* had been introduced. When cells with this distribution of the *ssb* gene were grown in continuous culture, plasmidless cells never accumulated (even though regions flanking the *ssb* gene in the plasmid were wholly homologous to segments of the chromosome containing the kanamycin marker). By contrast, continuous cultures of AB1157 containing *ssb* both on the chromosome and in pACYC were rapidly overtaken by plasmidless progeny. In a sense, the *ssb*-carrying pACYC had become an artificial chromosome. Natural counterparts of these engineered conditions have been observed in strains in which F′ plasmids initially arose. Upon formation of the F′ plasmid in the primary strain by excision of the F factor together with adjacent bacterial genes, a plasmid was created that carried unique copies of genes no longer present in the chromosome (300). In the primary strains, F′ plasmids carrying essential genes are not free to segregate. In secondary strains (merodiploid for genes carried by the F′ plasmid) the plasmid is not essential and does segregate.

Summary. The relationships between plasmids and chromosomes are complex. Two enduring characteristics of plasmids are those that are incorporated into the definition of these accessory genetic elements, namely, that they are autonomous and nonessential. Exceptions to both these conditions are neither well established nor commonplace. Accordingly, the existence of more than one naturally occurring, segregationally stable chromosome among procaryotes has yet to be unequivocally demonstrated.

CHROMOSOMAL DYNAMICS

Some Large Recurrent Chromosomal Elements

Repeated elements can serve as sites of intrachromosomal homologous recombination; consequently, duplications, deletions, and inversions arise through these structures. Notably, the occurrence, abundance, organization, and chromosomal locations of some large elements vary widely among bacteria. In some instances, differences are evident when major taxonomic groups are compared; for example, IS elements are common in such families of bacteria as the *Enterobacteriaceae* and have not been observed in such other families as the *Bacillaceae*. In other instances, differences are evident when different strains within a species are compared; for example, *rhs* loci (rearrangement hot spots) occur in many strains of *E. coli* and are absent in others. In yet another instance, *viz*, *rrn* loci, the genetic element is common to all bacteria. Despite the expectation that the organization of a chromosomal component essential for viability might be highly conserved, striking variation in extent of repetition, arrangement, and linkage occurs among procaryotic *rrn* loci. Apparently, bacterial species have acquired and used large repeated genetic elements to differing extents in establishing species-specific or even strain-specific chromosomal organizations.

***rrn* loci.** Ribosomes are abundant, occur universally, and

TABLE 2. Number and linkage in *rrn*loci of eubacteria

Organism	Number of *rrn* loci with 16S-23S-5S linkages	Number of *rrn* loci with other linkages or unknown linkages	Reference
Acholeplasma laidlawii		2	7
Anabaena sp. strain PCC 7120		2	13
Bacillus subtilis	10		219
Clostridium perfringens		9	46
Escherichia coli	7		11
Haemophilus influenzae		6	202
Leptospira interrogans		2 23S 2 16S 1 5S	98
Mycobacterium bovis	≤2		347
Mycobacterium intracellulare	1[a]		22
Mycobacterium leprae	1		311
Mycobacterium lepraemurium	1		345
Mycobacterium phlei	2[a]		22
Mycobacterium smegmatis	2		22
Mycobacterium tuberculosis	1[a]		22
Mycoplasma capricolum		1	7
Mycoplasma capricolum subsp. *mycoides*		2	7
Mycoplasma gallisepticum	1	1 23S-5S 1 16S	52
Mycoplasma hypopneumoniae		1 16S-23S 1 5S	348
Pirellula marina		2 23S-5S 2 16S	210
Pseudomonas aeruginosa	4		130
Salmonella typhimurium	7		288
Streptomyces ambofaciens	4		266
Streptomyces coelicolor A3 (2)	6 (2)[b]		18
Streptomyces lividans TK21	6		346
Thermus thermophilus		2 23S-5S 2 16S	131 129
Ureaplasma urealyticum 960[T]		2	54
Vibrio harveyi		23S-16S-5S[c]	199

[a] Preliminary determination.
[b] Determined for two loci.
[c] Number of loci not determined.

contain (among other components) transcripts of *rrn* genes. Thus, rRNAs are readily available to be used as probes to determine the abundance, arrangement, composition, and location of *rrn* loci.

rrn loci are commonly repeated (Table 2). Both *E. coli* and *S. typhimurium*, members of the *Enterobacteriaceae*, have seven *rrn* loci which occur in equivalent positions in their chromosomes (11, 297). Duplications and deletions arising from crossovers between directly repeated *rrn* loci and inversions from crossovers between inverted *rrn* loci have been observed in these organisms (146). Although the seven *rrn* loci are at equivalent chromosomal positions in *E. coli* and *S. typhimurium*, the *rrn* loci of the latter organism have a striking distinctive feature: many of the sequences coding for the 23S RNA in strain e23566 contain a sequence of approximately 90 bp which is enzymatically cleaved from the mature transcript (41). (Under nondenaturing conditions, the rRNA fragments stay associated through the secondary and tertiary structure.) Intervening sequences also occur in some of the *rrn* loci of several other *Salmonella* spp. (41).

B. subtilis (219) and *Clostridium perfringens* (46), members of a common phylogenetic group, also have similar large numbers of *rrn* loci. However, the extents of repetition of *rrn* loci need not be the same in one family or genus, e.g., compare *Streptomyces* spp.

Repetition of *rrn* loci, while commonplace, is not universal. For example, *Mycobacterium* species can be classified as slow-growing or fast-growing species; the former have one *rrn* locus and the latter have two (22). Some *Mycoplasma* species have one *rrn* locus, whereas others have two identical loci or several loci which differ from one another (7, 52, 348). Likewise, some archaebacteria have a single *rrn* locus, whereas others have as many as four (68). Collectively, these data indicate that some bacteria, including species having genomes estimated to have as many as 6 Mb, have a single locus for each type of rRNA, a condition which precludes *rrn*-mediated rearrangements. By the same consideration, the possibility of rearrangements is greater when there are more repeats of the *rrn* loci, as occurs in some rapidly growing bacteria.

Among eubacteria, there are many different organizations in *rrn* loci. Often, the DNA sequences specifying rRNAs are linked in a single operon, with the order being 16S-23S-5S. However, commonplace order and close linkage are not universal. Information on the organization of *rrn* loci in eubacteria is given in Table 2 and can be summarized as follows. The order of sequences specifying three rRNAs in an operon may vary (Table 2, *Vibrio harveyi*). Not all *rrn* loci code for all three rRNAs (Table 2, *Pirellula marina* and *Thermus thermophilus*). A single type of rRNA may be specified by *rrn* loci containing no other, one other, or two other sequences specifying additional rRNAs. Members of the same genus may have *rrn* loci arranged differently (Table 2, *Mycoplasma* spp.). Finally, there may be no close linkage at all among sequences specifying the different rRNAs (Table 2, *Leptospira interrogans*). Among archaebacteria (members of which are not represented in Table 2), *rrn* templates commonly occur in the order 16S-23S-5S. However, the spacing that separates these components varies greatly (247). Thus, when a spectrum of bacteria including marine bacteria, wall-less forms, spirochetes, and species incertae sedis is examined, a wide variety of arrangements of *rrn* loci is seen. Notably, variation is seen even among phenotypically similar species. Seemingly, *rrn*-mediated chromosomal rearrangements would originate only in instances when recombination can occur between repeated *rrn* loci having equivalent organizations.

The nucleotide sequences of rRNAs are considered to be highly conserved. The compositions of *rrn* loci, rRNA oligonucleotide catalogs, rRNA sequences, and rRNA signatures are very similar for closely related organisms (376). In species having multiple *rrn* loci, the sequence of nucleotides in these loci are often identical or nearly so. Nonetheless, differences in composition occur. In *E. coli*, the 16S rRNA templates of *rrnB* and *rrnG* differ by six bases among the first 674 nucleotides (see reference 146 and references therein). *Halobacterium marismortui*, which has two *rrn* loci, exhibits a high degree of heterogeneity at the 5′ ends of the two 16S RNAs. In particular, 12 of the first 131 nucleotides differ from one another (236).

rrn loci also differ with regard to the presence or absence and identity of tRNA genes in the spacer regions between segments coding for rRNAs. All *rrn* loci in *E. coli* ordinarily have one or two tRNA templates between segments encoding 16S and 23S rRNA and zero, one, or two tRNA templates beyond the segments encoding 5S rRNA; no *rrn* locus codes for more than three tRNAs (91). Inter- and intraspecific comparisons allow identification of differences in *rrn* loci among highly similar species (146). For example, the spacer region separating the 16S and 23S coding region of four loci in both *E. coli* and *S. typhimurium* contain

$tRNA_2^{Glu}$, and three contain $tRNA^{Ile}$ and $tRNA_{1B}^{Ala}$. Although the abundances are the same, the organization of operons in equivalent positions in the genome differs: a reciprocal exchange of spacer regions between *rrnD* and *rrnB* seems to have occurred. *rrnD* in *E. coli* contains $tRNA_{1B}^{Ala}$ whereas *rrnD* in *S. typhimurium* contains $tRNA_2^{Glu}$. The complement of this pattern exists in *rrnB*; *E. coli* contains $tRNA_2^{Glu}$ whereas *S. typhimurium* contains $tRNA_{1B}^{Ala}$ (204). Even within species, recombinational exchange results in differences in spacer regions. Exchange between *rrnB* and either *rrnC* or *rrnE* results in loss of the spacer sequence in *rrnB* of a substrain of *E. coli* K-12 (133). Recently, a series of mutants has been constructed in which the number of spacer regions specifying $tRNA_2^{Glu}$ or $tRNA^{Ile}$ and $tRNA_{1B}^{Ala}$ has been reduced to one. Growth of such mutants was impaired. Typically, the mutants were overgrown by revertants which had higher copy numbers of the spacer encoding the deficient tRNA and which sometimes also had inversions of portions of the chromosome between altered *rrn* loci (132).

Alanine tRNAs are located in the spacer sequence between the 16S and 23S rRNAs of some halobacteria and methanogenic bacteria (68). In addition, the halobacteria have a cysteine tRNA distal to the gene for 5S RNA (68). In *B. subtilis*, the spacer region in two of the *rrn* loci is large (ca. 344 bp) and contains tRNAs; in the remaining eight *rrn* loci, the spacer regions are 164 bp and contain no tRNA genes (219). No tRNA sequences have been detected in *Streptomyces coelicolor* (18), *Streptomyces lividans* (346), or *Streptomyces ambofaciens* (266). Likewise, no distal tRNAs have been detected in the 23S-5S rRNA loci of *Pirellula marina* (210).

The distribution of *rrn* loci in the chromosome also varies among bacteria. In *E. coli*, most of the *rrn* loci are located in the half of the genetic map centered on *oriC* (214). The *rrn* genes are oriented between *ori* and *ter* in such a manner that transcription occurs in the same direction as does replication of the loci (36, 37). In *Clostridium perfringens* (46) and *B. subtilis* (202) the *rrn* genes cluster in about one-third of the genetic maps.

In addition to rRNAs, ribosomes contain proteins. In *E. coli*, the operons specifying the 53 ribosomal proteins are often large and contain structural genes for proteins other than those incorporated into ribosomes. Approximately half the genes for ribosomal proteins occur at 73 min on the standard *E. coli* map while the remainder are distributed with no evident pattern through the remainder of the chromosome (214).

***rhs* loci.** *rhs* sequences were first recognized in *E. coli* by their ability to generate duplications through unequal crossing over between repeated segments (213). Subsequent to the recognition of *rhsA* at 81 min and *rhsB* at 77 min, additional repeats were identified by probing genomic digests. Two of these additional repeats have been cloned and mapped; their identities and locations are *rhsC* at 16 min and *rhsD* at 12 min (296). All four *rhs* loci have a common orientation on the *E. coli* chromosome. The *rhs* loci extend 8 to 9 kb and are composed of 3,714-bp cores plus flanking segments. Sequence analysis indicates that the first nucleotide of the core could correspond to the beginning of an ORF specifying a highly hydrophilic 141-kDa protein composed of 1,238 amino acids (84). The core sequences are highly conserved among three of the *rhs* loci as judged by restriction maps and partial nucleotide sequences. Specifically, the *rhsA* diverges from *rhsB* in 4 of the first 300 nucleotides; *rhsC* differs from *rhsA* and *rhsB* in 7 of the first 300 nucleotides. None of these nucleotide changes alters the amino acid sequence inferred from the nucleotide sequence. In addition, *rhsA* and *rhsC* share a high degree of upstream homology. *rhsD* contrasts with the others by diverging in 49 to 54 of the first 300 nucleotides of the core. Beyond the conserved core are nucleotide sequences which add 139, 173, 159, and 177 3′ codons to *rhsA*, *rhsB*, *rhsC*, and *rhsD* core ORFs, respectively; a second ORF exists in *rhsA* beyond the core extension. Although the inferred amino acid sequences of core proteins are highly conserved, considerable divergence is exhibited among the carboxy termini of these ORFs. Finally, maxicell experiments suggest that two proteins of approximately the appropriate size are produced from *rhsA*, *rhsB*, and *rhsC* (84).

The *rhs* loci occur in *E. coli* K-12 but not in a strain from nature, *E. coli* ECOR 55 (nor strains of *S. typhimurium*) (84). Nucleotide sequences in the ECOR 55 strain are identical to those in the regions which flank *rhsA* in *E. coli* K-12. The segments with identical composition in ECOR 55 and *E. coli* K-12 are interrupted by a sequence of 32 bp which occurs uniquely in ECOR 55 or alternatively by a sequence of 8,249 bp in *E. coli* K-12. The 8,249 bp comprises the *rhsA* locus which contains the *rhsA* core, the 139 codon extension, regions partially repeated, and additional flanking regions. At the *rhsC* locus in ECOR 55 and *E. coli* K-12, regions of precise homology are interrupted by a 10-bp pyrimidine-rich segment which occurs uniquely in ECOR 55, or alternatively a 9.6-kb segment comprising the *rhsC* locus, which occurs uniquely in *E. coli* K-12 (84). Thus, inserts of quite different sizes have taken place in *rhsA* and *rhsC* through independent events in these two *E. coli* lineages.

A similar example of alternate (but not repeated) alleles in *E. coli* is the occurrence of the *gat* gene (galactitol catabolism) in *E. coli* K-12 at the same chromosomal position as the *rtl atl* genes (ribitol and arabitol catabolism) in *E. coli* C (215, 377). Flanking sequences are homologous, but different alleles have become established at the one locus in the two strains.

The occurrence of *rhs* loci has many puzzling features. On the one hand, they occur irregularly among strains and related genera; hence, they are not essential for viability or growth. On the other hand, when they do occur, their structure is highly conserved as if there were no latitude in the composition of the locus and its possible proteinaceous product(s) or as if they arose recently by duplication. As with some accessory elements such as prophages and IS elements, the numbers of copies per cell may vary, but nucleotide sequences of multiple copies are closely similar. However, *rhs* loci are not bounded by short repeated sequences as is the case for prophages and transposons, so in this respect they are not definable as accessory elements of this type (84).

IS elements. The genomes of bacteria often contain many IS elements. In extreme examples, as many as 150 copies of a specific IS element (or related structure) may be present, such as has been seen with IS*1*-related structures in *Shigella dysenteriae* (257). Recently, the genetic locations of several IS elements in strains of *E. coli* have been determined. In particular, the locations of IS*1*, IS*2*, IS*3*, and IS*5* in *E. coli* W3110 and JE5519 (355, 356), IS*5* in W3110 (242), and IS*1*, IS*2*, IS*3*, IS*4*, IS*5*, IS*30*, and IS*150* in BHB2600 (29) have been independently identified. In the latter investigation, additional restriction fragments that contained IS elements in strains W3110 and HB101 were also identified, but the genetic locations were not determined. Comparison of the results shows that the numbers and locations of IS*5* elements were virtually identical in strains BHB2600 (29) and JE5519

(356), whereas strain W3110 contained about twice as many copies (242, 356). Isolates of strain W3110 maintained in separate laboratories gave similar but not identical results for the locations of IS*1*, IS*2*, and IS*3* (29, 355) and IS*5* (29, 356). These mapping studies demonstrate that the distribution of IS elements in the chromosome is not uniform; the chromosomes of W3110 and BHB2600 contain large regions devoid of IS elements and small regions populated densely with clusters of IS elements. For example, BHB2600 contains a 16-min segment in the *oriC* region with no IS elements (29). Both BHB2600 and W3110 have no IS elements from approximately 14 to 22 min and from approximately 52 to 62 min. By contrast, in BHB2600 a 90-kb fragment centered at approximately 9 min contains a cluster of 10 IS elements (29). To some extent, the distribution of the IS elements complements the distribution of *rrn* and *rhs* loci. All the segments lacking IS elements contain one or several *rrn* or *rhs* loci. Furthermore, the regions designated as nondivisible zones (NDZs; see Inversions, below), contain IS elements but are completely devoid of *rrn* and *rhs* loci.

IS elements appear to generate chromosomal rearrangements. A set of duplications of about 14 kb of chromosomal DNA resides at two locations in the W3110 chromosome, one at about position 3140 to 3190 on the Kohara physical map, the other at about position 3300. Each duplicated 14 kb chromosomal segment is flanked by copies of IS*5* (242, 356). Other rearrangements such as inversions and integration-excision of extrachromosomal elements are believed to have arisen by recombination between flanking IS elements (see the discussion and references in references 29, 355, and 356).

tDNA. The most extensive characterizations of procaryotic tRNA organizations are those describing tRNA loci in *E. coli* (91) and *B. subtilis* (360); characterization in *Photobacterium phosphoreum* has begun. Notably, the organizations are quite different in each of these bacteria. In *E. coli*, tRNA loci may exist as a locus specifying only one tRNA, as an operon encoding several tRNAs, or as an operon specifying both tRNAs and a variety of proteins (91). Recently, all loci for the 46 tRNA species presently known to exist in *E. coli* were mapped. The 79 loci occur in 41 operons that are broadly distributed through the chromosome (192). In *B. subtilis*, tDNAs occur within *rrn* genes or in clusters in three regions of the chromosome. The largest cluster encodes the sequence of 21 tRNAs; another site encodes 16 tRNAs; and the final region encodes 6 tRNAs (360).

An especially striking organization has been found in *P. phosphoreum*, with many repeated occurrences of $tRNA^{Pro}$ genes, $tRNA^{His}$ genes, and related pseudogenes (107). Two chromosomal fragments from *P. phosphoreum* were identified by hybridization with a cluster of *E. coli* tDNA sequences and were cloned in plasmids pPPS70 and pPPH12. Nucleotide sequences of the cloned fragments showed a pattern of repeats of tRNA and tRNA-like sequences (Fig. 2). The smaller cluster (in pPPS70) seems to be a duplicate of part of the larger one (in pPPH12): pPPS70 contains 304 bp that are 99% similar to a part of pPPH12. In the larger cluster, there are six copies of the $tRNA^{Pro}$ gene, two copies of the $tRNA^{His}$ gene, and a set of $tRNA^{Pro}$ pseudogenes. The pseudogenes resemble the $tRNA^{Pro}$ sequence, except that there are inserts of unrelated sequences in the anticodon and T loops. Upstream of five of the six $tRNA^{Pro}$ genes is a sequence resembling the 5′ end of the $tRNA^{Pro}$ pseudogene sequence. This elaborate set of repeated tRNA genes is shown in Fig. 2 as a schematic drawing of the possible secondary structure of hypothetical RNAs transcribed from the DNA sequences.

P. phosphoreum is a phosphorescent, purple marine bacterium in the family *Vibrionaceae*. It is not yet clear why it should harbor such extensive duplications or tRNA genes and pseudogenes as compared with either *E. coli* or *B. subtilis*.

Summary. Observations with *rrn* loci, *rhs* loci, IS elements, and tRNA genes show that the occurrence and organization of these structures varies even among closely related bacteria. At the same time, experimental analysis has established that many of these large repeated structures have contributed to rearrangements of chromosomes within experimental strains and may have contributed to chromosomal differences which distinguish natural isolates. In evolutionary terms, these commonplace and fundamental structures can be regarded as having contributed to the establishment and maintenance of species-specific differences in procaryotic chromosomes.

Rearrangements

Bacterial chromosomes become rearranged either by homologous recombination between repeated sequences or by site-specific recombination. Some of these rearrangements are quite stable with reversion rates similar to point mutation rates, others are readily reversible at frequencies in the range 10^{-3} to 10^{-1}.

Amplifications. Tandem duplications of genes have been observed over the years in many bacteria, particularly in antibiotic-resistant mutants. Recently, some amplified mutants have proved to contain massive numbers of relatively large duplicated genetic segments. Under specific conditions of protoplast fusion in the presence of polyethylene glycol followed by regeneration in the presence of tetracycline, some *B. subtilis* mutants arose by a *recE*-dependent process that had 80 to 100 copies of genetic segments including Tc^r markers. The amplified segments varied in separate isolates from 11 to 30 kb, had different endpoints, did not appear to involve flanking repeated sequences, and were all located near the origin of replication (168). In *Streptomyces* spp., mutants with amplifications are even more spectacular. Many *Streptomyces* species mutate at high rate to amplify certain chromosomal sequences into very large arrays of reiterated sequences that make up as much as 30% of the chromosomal DNA. Often accompanying the amplification are deletions of adjacent DNA, sometimes large tracts of DNA that contain known genes and confer mutant phenotypes (for reviews, see references 27, 61, and 159). The genetic instability that is characteristic of most *Streptomyces* species gives rise to spontaneous mutants at rates as high as 1%, rising to 10% and more after such treatments as exposure to UV light or growth with ethidium bromide (159, 201). The generation of a tract of highly amplified sequences is proposed to be equivalent to attaining a hypervariable state, one that generates further mutants at unusually high frequencies, such as deletions of sequences adjacent to the amplified tract (201).

The wild-type parental bacteria that generate these amplified mutants typically contain only a single chromosomal copy of an amplifiable unit of DNA (AUD), usually in the range of 5 to 25 kb in size, flanked by repeated sequences of around 1 to 2 kb. In an amplified mutant, amplified DNA sequences (ADS) are readily detected as strong bands in an agarose gel of restriction enzyme digests of chromosomal DNA. In the mutant, the ADS can be large enough to form a prominent part of the chromosome structure of the mutant. In one example, the parent strain, *Streptomyces fradiae* wild

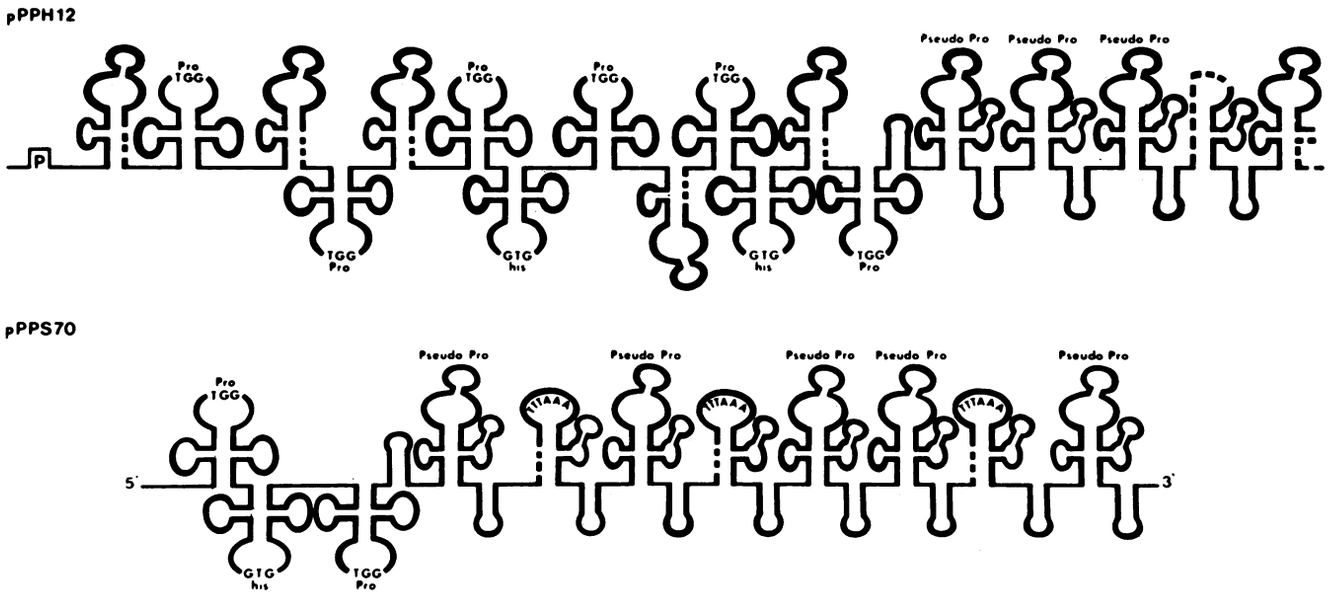


FIG. 2. Schematic drawing of the possible secondary structures of the sequenced fragment of *Photobacterium phosphoreum* DNA including $tRNA^{ProP}$ genes. For clarity, the inserts in the anticodon and T loops are indicated as hairpin structures; in reality these have no base-paired regions and merely increase the sizes of the loops. The dashed lines indicate that no similarity is found between this region and the 3′ half of $tDNA^{Pro}$. Reprinted from reference 107 with permission.

type, contained an AUD organized as 8.3 kb of sequence flanked by two copies of a directly repeated 2.2-kb sequence. The amplifiable unit of 8.3- and 2.2-kb sequences was found in only one or a few chromosomal sites and was often present in the wild type as a tandem repeat (86, 87). For the amplified mutant of *S. fradiae*, kinetics of hybridization of DNA indicated that more than 500 copies of a 10.5-kb AUD are present (86, 87). Were there to be 500 copies of a 10-kb ADS, 5,000 kb of this repeated sequence would be added to the chromosome, an amount greater than the entire *E. coli* chromosome, and a very large fraction of the *Streptomyces coelicolor* chromosome, estimated to be between 7,000 and 8,000 kb (R. Kieser, T. Kieser, and D. A. Hopwood, personal communication).

Similar motifs are present in other *Streptomyces* spp. Four ADS ranging in size from 5.8 to 24.8 kb were identified in a spontaneous pigmentless mutant of *S. ambofaciens*, three of them related to one parental AUD sequence and one to another (67). A similar phenomenon has been documented in *S. achromogenes* subsp. *rubradis* (154, 155). A parental AUD sequence of 8.0 kb (including a weakly expressed determinant for spectinomycin resistance and one copy of the flanking 0.8-kb repeated sequence) was amplified in a spectinomycin-resistant mutant to between 200 and 300 copies of ADS. Accompanying the amplification was deletion of at least 10 kb of DNA adjacent to the ADS array (154, 155). In *S. glaucescens*, amplifications originated from a 100-kb region of the parental genome, capable of generating a heterogeneous collection of ADS-containing mutants with amplifications of up to 500-fold (27, 28, 134). In one amplified mutant, a tandem reiteration of more than 1,500 kb was identified in a restriction fragment by PFGE. Deletions accompanied the amplification, removing two known genes: *strC* (streptomycin resistance) and *melC* (melanin formation). Progressive hybridization with overlapping restriction fragments, DNA walking, established that the genes subject to deletion were physically linked to the AUD locus in the parental chromosome. In the amplified mutants, very large deletions from 270 to more than 800 kb extended away from the ADS array and removed the *strC* and *melC* genes.

How stable are the amplified mutants? When an AUD contains a genetic determinant for resistance to a drug, such as resistance to spectinomycin in *S. achromogenes* subsp. *rubradis*, amplification confers and strengthens drug resistance. When the drug is withdrawn, surprisingly the ADS persists even in the absence of selective pressure during vegetative propagation of the organism. However, formation of spores or protoplasts rapidly rids the organism of the ADS. Seemingly, a life-cycle-related mechanism excises the supernumerary copies of the amplifiable DNA (156).

The extraordinary behavior of *Streptomyces* chromosomes demonstrates that a high degree of genetic variability is not fatal but on the contrary is well tolerated by the species and is compatible with a genetically stable chromosome map (see Physical and Genetic Maps, above). We have also learned from this behavior that a high degree of distortion of chromosomal composition is accommodated with apparent ease by this organism; augmentation at one locus may be between one-third and two-thirds as much DNA as is present in the entire unaltered chromosome. Clearly, the *Streptomyces* chromosome is able to absorb gigantic changes in size and spatial relations of its component genes without severe disadvantage (see Consequences of Moving Genes, below).

Deletions. Small deletions can change the gene dosage, alter a gene product, or alter patterns of gene expression. Deletion mutations arise in a *Streptococcus* strain that cause size variation in the surface protein and virulence factor, the M protein (147). Sizes of the M protein ranged from 41,000 to 80,000 Da, and frequencies of mutants were as high as 1 in 2,000. Nucleotide sequence of the gene for the M protein reveals many internal repeats within the coding region, and these repeated sequences lie at the points of recombination that are believed to generate the various size classes of the M protein (147).

Some strains of *B. subtilis* have 9 rRNA genes, rather than

the more common number of 10. Restriction map analysis of one such *B. subtilis* laboratory strain showed that homologous recombination between closely linked multiple *rrn* loci caused deletion of one of the loci and the loss of genetic material lying between the recombining *rrn* loci (375).

A sequence of programmed deletions has been described for *Anabaena* heterocyst formation. In the absence of nitrate or ammonia, approximately every 10th cell differentiates into a nonreplicating heterocyst. In heterocysts, oxygen-sensitive nitrogenase is protected and the heterocyst provides fixed nitrogen to neighboring cells by using N_2 from the environment and an organic reductant from adjacent cells to produce glutamine (111).

The genetic information for differentiation of a vegetative cell into a heterocyst is both formed and regulated through two programmed excisions of chromosomal segments. In one instance, an 11-kb excision has the effect of lengthening an ORF and also brings a distal gene under the influence of a proximal promoter (111). Three genes in the DNA of vegetative cells, *nifH*, *nifD*, and *nifK*, are involved in the rearrangement. The 11-kb segment lies near the C-terminal end of the *nifD* gene. Excision of the segment changes the terminal section of the ORF, adding 17 amino acids. Transcription of both the *nifH* and *nifD* genes in vegetative cells is coupled, proceeding from the *nifH* promoter. After excision of the 11-kb segment, *nifK* is brought into the transcription unit adjacent to *nifD*. Thus, excision achieves operon fusion and expression of all three structural genes. The excised segment exists in the heterocysts as a circle of DNA.

Another less fully understood excision, of 55 kb, occurs by recombination between direct repeats (110). The 55-kb excision moves the *nifS* gene (which is incompletely characterized) closer to the *rbcL* and *rbcS* genes (specifying the large and small subunits of ribulose biphosphate carboxylase, respectively). To the extent that deleted DNA remains intact as a circle in the cytosol, available for reintegration into the chromosome, the deletion events could be transient and reversible.

A similar mechanism functions in *B. subtilis* to form *sigK*, the structural gene for an RNA polymerase sigma factor which is expressed in the parent cell but not the forespore (197). The proximal and distal sections of the gene, *spoIVCB* and *spoIIIC*, respectively, are separated by a 42-kb segment (called *skin* for *sigK* intervening segment) which is excised as a circle. *skin* is flanked by 5-bp direct repeats which function as the position of site-specific recombination. In both *Anabaena* heterocysts and *B. subtilis* parent cells, the cells with programmed deletions are terminally differentiated and do not function as progenitor cells.

How do duplications and deletions arise? Two main mechanisms are often cited: slipped-strand mispairing during DNA replication (reviewed in reference 208) or unequal crossover between two daughter sets of duplicated sequences, generating one chromosome with another duplication and a second chromosome with a deletion (see, for instance, reference 9). We are only just beginning to collect examples of mutations that occur in one gene and set into motion genetic changes in nearby DNA. An example is the *bglY* gene of *E. coli*. Point mutations in the *bglY* gene at 27 min caused 10- to 100-fold increases in deletions both *cis* to *bglY* in the chromosome and *trans* in a coexisting plasmid (205). Small deletions, then, like small duplications, can lead to other genetic changes, perhaps by encouraging mispairing.

Phase variation and antigenic diversity. Small rearrangements involving inversion of a genetic segment cause on-off switching in gene expression, often called phase variations; small translocations can change coding sequences, creating antigenic diversity. These site-specific, reversible changes affect the expression of one or a few genes for cell surface components.

The range of bacterial genetic systems that use controlled reversible rearrangements to allow for rapid phenotypic switching has been summarized by Borst and Greaves (32). At present there are four types of systems that change the expression of genes affecting several different types of surface antigens.

(i) Inversions of the Hin type. The well-established invertible switch controlling phase variation in *S. typhimurium* uses the site-specific Hin recombinase to invert a 980-bp genetic segment, thereby controlling the expression of two alternate flagellar genes by positioning regulatory sequences in relation to coding sequences (319). Closely similar genetic systems encoding recombinases of like specificity have been found as *gin* in phage Mu, *cin* in phage P1, and *pin* in the inactive prophage e14 residing in the *E. coli* K-12 genome. These all share the attribute of controlling gene expression by a specific inversion that either conjoins a promoter and regulatory sequences to the relevant coding sequence or disjoins them. This group of site-specific inversion systems has been reviewed from different points of view (see, for example, references 60, 108, and 270). Similar systems have been found for pilin genes in *Moraxella bovis* (99, 228) and may also be the mechanism of flagellar phase variation in *Campylobacter coli* (117).

(ii) Other site-specific inversions. An inversion system with different characteristics governs phase variation of type 1 fimbriae (pili) in *E. coli*. The inverted DNA segment that affects the orientation of the *fimA* gene promoter (1) is smaller (314 bp) than the inversion segments of the group related to the *Salmonella hin* inversion (1 to 3 kb), and the nucleotide sequences at the recombination sites of *fim* differ from the sequences of the *hin* family recombination sites. In fact, as it turns out, the *E. coli fim* inversion requires the same *E. coli* protein (integration host factor) that is required for lambda phage integration and excision (70). Integration host factor-binding sequences are located close to *fimA* (85). Thus, *E. coli* fimbrial phase variation is related to the site-specific recombination of the lambda phage *Int-Xis* system, not to the *Salmonella* phase variation system.

(iii) Cassette switching. Another type of programmed rearrangement process is the cassette system, first characterized in *Saccharomyces cerevisiae*. In cassette switching, variants of a protein result from transposition of one of a group of silent genes into an expression site, where the version of the gene at the expression site becomes active. This mechanism is used for expression of *Neisseria gonorrhoeae* pilus genes, leading to antigenic variation in pilus protein composition (330). Presumably the capacity to vary the specificity of surface antigens contributes to the ability to evade host immune defenses. A 5′ portion of the coding sequence of the pilin gene resides at the expression site and is constant in composition. The rest of the gene is variable and is switched into the expression site from copies at multiple sites (123, 309). The recombination event required for the switching is triggered by uptake of transforming DNA (310). A similar mechanism seems to account for antigenic variation in *Borrelia hermsii* (183, 270) and *Campylobacter coli* (116).

(iv) Variation at translation. *N. gonorrhoeae* has another major surface antigen, outer membrane protein P.II, originally noticed as the surface component responsible for an opaque colony appearance. The P.II protein exhibits both

phase variation (on and off switching) and antigenic variation (differences in the sequences of different versions of the protein). These changes are exerted by another mechanism that in this case affects translation. There are many genes for the P.II protein. The nucleotide sequences for the many leader peptides contain from 7 to 28 tandem copies of a pentanucleotide. Addition and/or deletion of variable numbers of these repeated sequences causes changes in the reading frame of codons of the P.II gene. Depending on whether the number of base pairs affected is a multiple of three, translation of the transcript can be either in or out of correct phase, thus controlling output of the gene product (315, 316). Recent studies have focused on the roles of ribosomal frameshifting and slipped-strand mispairing (20).

Summary. As we learn more about large-scale rearrangements in bacterial DNA, we appreciate the importance not just of nucleotide sequences of individual genes and properties of individual gene products, but also of the arrangement of the genes in the chromosome. As discussed in this section, major rearrangements occur in some bacteria at high frequency and play important roles in developmental processes such as heterocyst formation and sporulation. Other rearrangements play important roles in engineering escape from host defenses. Still others set into motion further changes in nearby DNA. We can ask whether there are limits on the plasticity of gene arrangement in the bacterial chromosome. How do bacteria make use of rearrangements without losing essential chromosomal identity? These questions are addressed in the next section on gene position effects.

Gene Position Effects

Since the map locations of homologous genes tend to be conserved in related bacteria, one can ask what kinds of restraints discourage the movement of genes. Are there attributes of chromosomal structure or processes that cannot be disrupted? Are some genetic events made difficult because of chromosome structure? One way of posing these questions is to ask whether there are position effects for bacterial genes.

Consequences of moving genes. Among the ways that genetic locations of genes can be changed, transposition seems to be benign and does not adversely affect gene expression, but some inversions are deleterious in their effects.

(i) Transpositions. The effect has been measured of translocating the *hisD* gene to 16 other chromosomal sites (304) and of integrating F′ plasmids carrying the *lac* operon at 11 (19) and 17 (229) other sites. In these experiments, the closer the transposed genes were to the origin of replication, the higher was the level of gene expression, up to a factor of about 3. The magnitude of the gradient of gene dosage effect for the *hisD* gene increased at higher growth rates (304), in accordance with the differences in gene dosage expected from the model of multifork replication as a function of generation times (58). No map locations were found that prevented gene expression. Apparently no ill effects result from transposing genes to new locations, beyond the relatively small gene dosage effect attributable to the geometry of replication.

(ii) Inversions. Inversions occur much less frequently than other kinds of chromosomal rearrangements such as duplications and deletions (9, 303). Inversions of segments of the chromosome seem either to be more difficult to execute or to have more drastic consequences than other kinds of genetic events.

To throw some light on the relationship of chromosomal structure and gene expression, investigators in two laboratories have systematically engineered and characterized sets of inversions of segments of the chromosomes in *E. coli* and *S. typhimurium*. In the Roth laboratory, inversions were isolated in *S. typhimurium* either by selecting for a recombination event to join a promoter to promoterless genes (303) or by systematic engineering, using elements of Tn*10* and Mu d*lac* phage derivatives to direct locations of recombination and endpoints of inversions (224, 225, 307, 308). Selection for restoration of function made it possible to detect rare events. In the Louarn laboratory, sets of inversions were selected in *E. coli* by using a λ phage derivative carrying marked Tn*10* loci as selectable and portable regions of homology (93, 218, 284). In both of these experimental systems, some recombinant strains with inversions arose with no apparent ill effect or only a small growth disadvantage relative to the wild type; other strains with inversions were severely affected and grew poorly; and yet other inversions were never seen to occur.

The terms permissive and nonpermissive have been used to describe the two main categories of inversions, those with little or no growth disadvantage and those with major growth disadvantage or that have not been observed to occur (308). Examination of the locations and endpoints of permissive and nonpermissive inversions gives clues to the reasons for the rarity of inversions. Absence of inversions of a particular class could reflect the requirement for integrity of a particular genetic arrangement needed for normal function (93, 284), or it might reflect an inability to form that class of inversions for mechanistic reasons (225, 307, 308). As discussed by Segall et al. (307), examples of functional barriers that could affect cellular function are as follows: (i) asymmetry in length of replication arms as a result of the relative positions of *oriC* and *ter*, with consequent imbalances in ratios of copies of genes and gene dosages; (ii) misorientation of the direction of transcription relative to replication (in *E. coli* K-12 [36, 37] and *B. subtilis* [388], the most highly transcribed genes are oriented on genetic maps so that transcription is in the same direction as replication); and (iii) interference with the formation and/or maintenance of chromosomal structure(s) essential to some cellular function.

In contrast to these functional barriers, mechanistic barriers might prevent the formation of a class of inversions because of the nature of the recombination event(s) and structural features of the chromosome (307, 308). The most striking evidence for the existence of a mechanistic barrier is the construction in the Roth laboratory of a so-called nonpermissive inversion of the genetic segment between *his* and *trp* by a genetic route other than inversion, namely, by transduction with two donor transducing elements (308). Although inversion of this genetic segment of about 7 map units was never seen, a recombinant containing the *his-trp* inversion was in fact formed by transduction, was functional, and showed no impairment; therefore, recombination at these particular segment ends was not prohibited, and existence of the nonpermissive inversion is not per se deleterious to the cell. Apparently it is the process of formation of this inversion that is not permitted, not the nature of the end product.

Roth and colleagues propose that the physical juxtaposition of intrachromosomal inverted repeats required for generation of inversions by a single intrachromosomal crossover (308) (Fig. 3) is prevented by a chromosomal structure in the *his-trp* portion of the chromosome (near the terminus region). Other types of recombination not requiring juxtaposi-

tion of loci within a single chromosome would not be affected by the chromosomal structure, such as transduction or interchromosomal recombination. Although interchromosomal recombination can generate deletions and duplications by single crossover events, generation of inversions by interchromosomal recombination requires two independent crossovers and thus would be a rare event (Fig. 3). Inversions not observed could have failed to form because the intrachromosomal endpoints are not accessible to each other, and the alternate route, interchromosomal recombination, is able to produce inversions only at low frequency.

Louarn and colleagues, in analyzing their collection of inversion strains of *E. coli*, have suggested that there are localized structural constraints within a substantial fraction of the bacterial chromosome flanking the terminus of replication (93, 284). By noting effects on growth, Rebollo et al. defined four classes (I to IV) of invertible segments (284). Inversion of class I segments has little, if any, effect on growth. Inversion of class II segments confers a sensitivity to rich medium; inversion of class III segments confers sensitivity to all media. Class IV is a formal group defined by the absence of observed inversions of these segments. The extents of the segments are shown in Fig. 4. Those in panels I, II, and III correspond to inverted segments which have increasingly adverse effects on growth. Those in panel IV correspond to segments which cannot be inverted or which have a lethal effect upon inversion. In addition to the data of Rebollo et al. (284) (solid segments), we have added information on other segments that have been characterized in other laboratories (hatched segments).

Looking first at classes III and IV, Louarn and colleagues pointed out that many members of these classes can be understood in terms of two sensitive regions in the terminus region of the chromosome. The inversions that either do not arise or are highly deleterious share the property that either one or both endpoints of the segment lie in two sensitive parts of the chromosome called NDZs. The entire terminus region of the chromosome extends over almost one-third of the whole, from about 17 to 44 min (Fig. 4). The central section of approximately 5 map units lies between the inside edges of the two terminators, T1 (*terA*, *terD*) and T2 (*terB*, *terC*) (for a discussion of the terminator map locations and subdivision of T1 and T2 into components, see references 196 and 264 and references therein). To either side of this central region are two flanking regions of about 11 map units each, the NDZs. The zones are considered nondivisible because disruption of a portion of the zone by inversion ordinarily either cannot occur or produces ill effects (Fig. 4, segments 21 to 37 and 40 to 47). In agreement, an earlier attempt to invert segment 54 was not successful (191).

By contrast, as long as the NDZs are not disrupted internally, the entire NDZ can be inverted without ill effect (class I, segments 14 to 16). The genetic regions defined as NDZs are almost the same as the two polarized replication zones flanking the terminators, two constrained zones that are replication regions, oriented and polar, containing successive replication pause sites (284). It is not difficult to imagine that interruption and rearrangement of these polar regions, vital to orderly termination of replication, could be deleterious. However, experiments have shown that the act of pausing at the inverted pause sites is not the cause of reduced fitness (92).

Another factor affecting the viability and vigor of inversion mutants is symmetry of the *oriC* and *ter* regions. Class II inversions, those that exhibit rich medium sensitivity (Rms), often have undergone rearrangements grossly asymmetrical with respect to *oriC* and the *ter* region, generating replication arms of markedly unequal lengths (Fig. 4, segments 18 to 20). In addition to the set isolated by Louarn and colleagues recently (93, 284), a very large inversion exhibiting the Rms phenotype, IN (29–78), had been isolated and characterized earlier (218) (Fig. 4, segment 53). Hill and Gray isolated and characterized a large, asymmetrical inversion between *rrnG* and *rrnE* that also exhibits rich medium sensitivity (144) (Fig. 4, segment 52). Unequal replication arms would affect the synchrony of initiation and termination and also the ratios of gene dosages among genes. Gene dosage gradients generated by multiple replication forks in rich medium would be steeper at the higher growth rates in rich medium. In minimal medium at low growth rates, the distortion would not be as great.

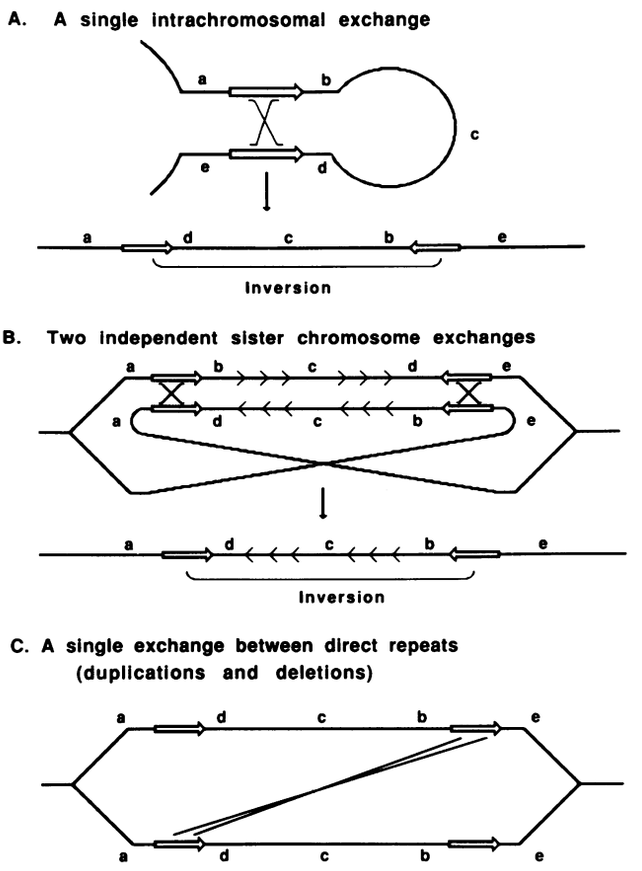


FIG. 3. Recombination events which can give rise to chromosomal rearrangements. Reprinted from reference 308 with permission.

Class I contains the inverted segments that had relatively little effect on the health of *E. coli*. In addition to the segments delineated by the Louarn group, Fig. 4 shows spontaneous inversions discovered residing in commonly used *E. coli* K-12 strains, segments 48 to 50 (145, 379) and the well-known *E. coli-S. typhimurium* inversion, segment 51 (48). The harmless character of almost all these invertible segments can be rationalized in terms of the foregoing discussion. Some of the segments are outside critical zones (segments 1 to 6, 49, and 50), some segments are symmetrical around *ter* or *ori* (segments 10, 48, 49, and 50), some involve the entire termination region without subdivision (segments 10 to 13 and 48), and some flip an intact NDZ without subdivision (segments 14 to 16). The evidently

innocuous inversion found naturally in *E. coli* or *S. typhimurium*, shown in Fig. 4 as segment 51 (53, 289), is symmetrical with respect to *oriC* and *ter*, and so in that respect it is expected to be innocuous. However, both ends lie in the NDZs, similar to segment 46 in class IV, and so from that standpoint it would be expected to be deleterious. Perhaps this ancient rearrangement took place before the full organization of the chromosome in the termination regions of these enteric bacteria had occurred.

Some of the results of the Roth group (307) can be looked at from the point of view of NDZs. In their *lac* inversion series, both the *his-trp* and *trp-pyrC* segments would be expected to interrupt an NDZ, and neither interval has been seen to invert. Likewise, in their *his* series, neither the *his-cobA* nor (again) the *his-trp* interval were seen to invert, consistent with avoiding interruption of an NDZ. However, some of the other observations in these series, such as the set of nonpermissive inversions running clockwise of *his*, have yet to be fully explained.

To summarize, the main findings from these experiments on inversions point to unusual properties of the chromosomes of *E. coli* and *S. typhimurium* in a large region around the terminus. A feature of the organization of the chromosome there seems to facilitate interchromosomal recombination and interferes with intrachromosomal recombination (308). There appear to be recombinationally nondivisible regions related in extent to the two terminal polarized replication zones and pause sites where replication can proceed only toward, not away from, the terminator (66, 284). Combining these observations, a supergenic chromosomal architecture seems to exist in the termination region whose organization interferes with production of inversions. A physical structure of the chromosome flanking the terminators, perhaps membrane bound, could be part of a chromosome segregation mechanism (284).

More information is needed before this view of terminus structure can be extended to other bacteria. Mutants of *B. subtilis* with large-scale rearrangements of the chromosome have been isolated and characterized by Anagnostopolous (8). Some of these preserved symmetry of *ori* and *ter* (inversion of segments Ia and A), others did not (inversion of segments B and C). Unequal replication arms seem to be tolerated in *B. subtilis* more easily than in *E. coli*. There may be a clue to some constraints in *B. subtilis* in that the region immediately flanking *ter* was not seen to vary. The *ter* locus lies at the junction of sections D and II, and this junction was not disturbed in any of the stably rearranged strains (8).

Nonrandom patterns of gene linkages. Any nonrandom arrangement of genes on bacterial genetic maps hints that somehow the positions of genes can affect their function or their regulation at a level of organization higher than the operon or regulon. Below, we summarize interesting observations on apparently nonrandom location of functionally related genes in the genetic maps of certain *Streptomyces* and *Pseudomonas* species with the thought that these nonrandom gene distributions may reflect higher levels of genetic coordination in bacteria which are yet to be perceived and their significance understood.

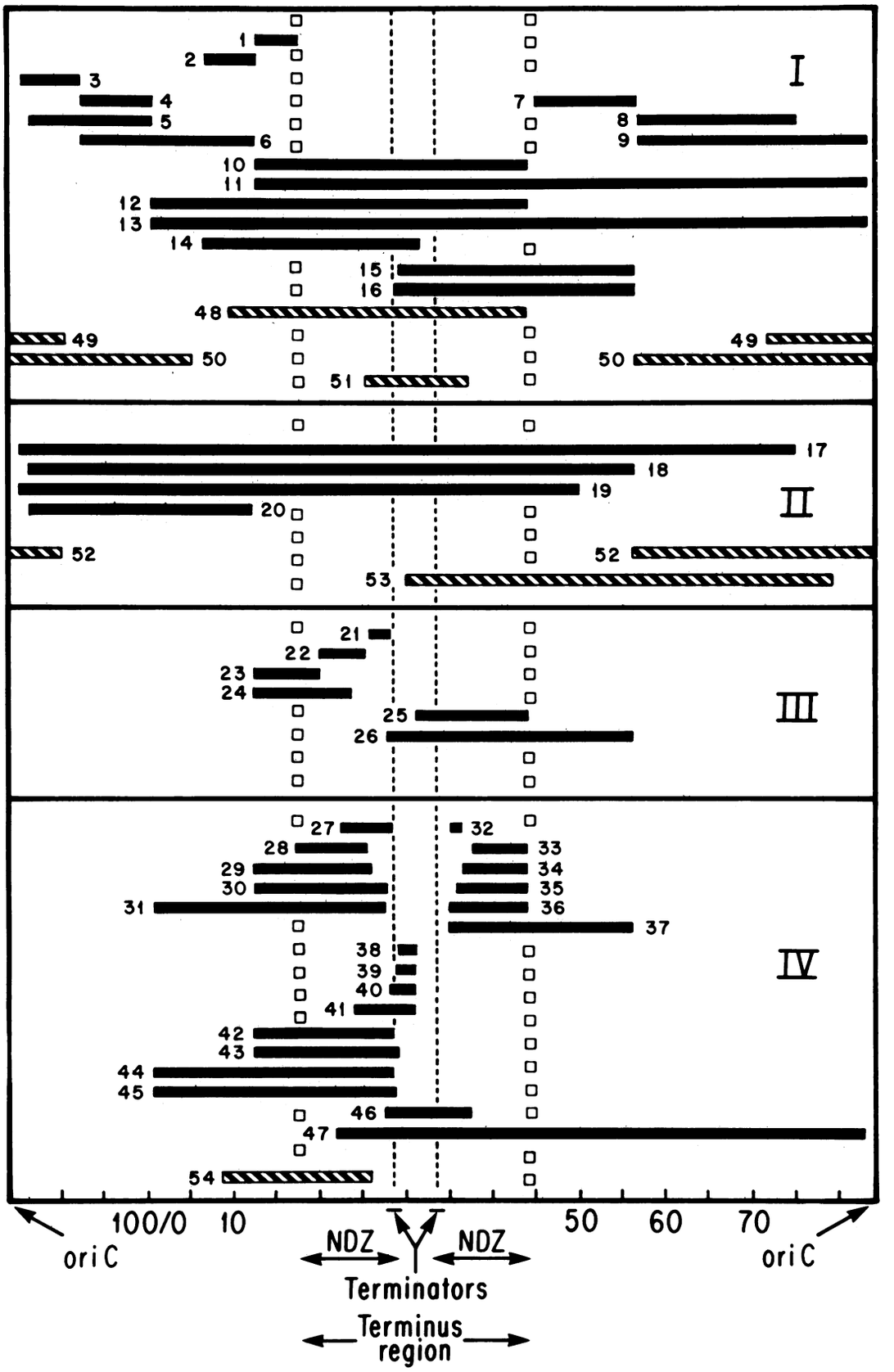
(i) ***Pseudomonas* spp.** Of the *Pseudomonas* species, the most detailed genetic map has been assembled for *P. aeruginosa* (148–150). A remarkable feature of the current map is that many of the genes for anabolic functions and central metabolic conversions (housekeeping functions) tend to be located in one half of the map, not in the other. Of the 52 mapped auxotrophic markers, 46 reside in 59% of the map. Although fewer genes have been mapped in *P. putida*, the same tendency is seen for that organism as well. Of the 43 mapped auxotrophic markers, 39 reside in 47% of the map. In Fig. 5, the thickened lines designate the chromosomal region in which most of the genes for anabolic and housekeeping functions of these two bacteria are located (see reference 149 for a full list of genes and their locations).

There is presently no experimental information bearing on the question of why genes in this broad grouping are confined to half of the map. The grouping could reflect the history of construction of the chromosome from separate parts. One can picture that the genes for essential core metabolism and biosynthetic functions could have been present in a smaller ancestral genome which was expanded to approximately the present size by a fusion with another genetic element, a very large DNA molecule of different function. Incorporation of large plasmids into the chromosome is not unknown in pseudomonads (see Plasmid-Chromosome Interactions above).

Whatever the origins of nonrandom gene position in the chromosome, most positions have been retained over the period since the divergence of *P. aeruginosa* and *P. putida*. Looking at the genes that are mapped in both organisms, Holloway et al. (149) found that the gene locations for anabolic functions are largely comparable and congruent, although exceptions exist. Figure 5 shows specific map positions for the set of auxotrophic markers that have been mapped in both species. Loci of *P. putida* PPN are shown on the outer circle, and the comparable genes in *P. aeruginosa* PAO are shown on the inner circle (adapted from reference 149). Two observations can be made. One is that many of the genes map in similar positions in the two organisms, but where there are differences, one group of differences can be rationalized as having occurred by small inversion events (genes 8 through 13) whereas the other group appears to reflect transpositions over relatively large distances (genes 1 and 2, 3 to 6, 20, 21, and 27). Even though transpositions have changed the gross locations of some of the genes, with few exceptions the genes stayed in the anabolism-rich half of the chromosome.

Catabolic gene clusters, in contrast to the nonrandom locations of anabolic genes, are found in many widely separated genetic locations around the *Pseudomonas* chromosomes, and these locations differ for *P. aeruginosa* and *P. putida* (149). The independent and dissimilar gene locations may reflect the origins of some catabolic genes from extrachromosomal elements. Catabolic functions are often located on plasmids in pseudomonads, and plasmid-borne metabolic genes can on occasion transfer from the plasmid to the chromosome (see When Does a Plasmid Become a Chromosome?; for a review, see reference 150). This may be

FIG. 4. Segment behavior in the inversion test. Segments tested for inversion behavior are arrayed on a simplified *E. coli* chromosome map. The map is linearized, open at *oriC*, with vertical dashed lines representing approximate locations of terminator loci and inside edges of NDZs and open boxes indicating approximate locations of outside edges of NDZs. The segments assayed by Rebollo et al. (284) are indicated by heavy lines and numbered 1 to 47; other segments assembled from the literature are indicated by hatched lines. References to hatched segments are as follows: segment 48, reference 379; segments 49 and 50, reference 145; segment 51, reference 48; segment 52, reference 144; segment 53, reference 218; and segment 54, reference 191. The segments were distributed in four phenotypical classes, I to IV, further described in the text. Modified from a figure in reference 284 and printed with permission from the author.


I
II
III
IV
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
100/0
10
50
60
70
oriC
oriC
NDZ
NDZ
Terminators
Terminus region

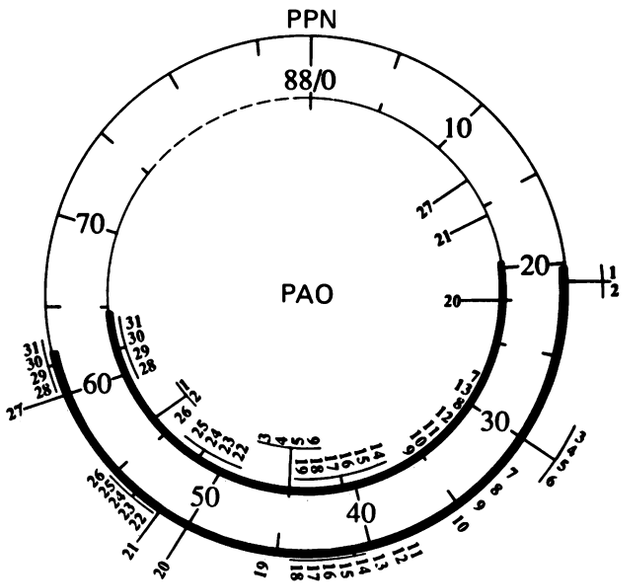


FIG. 5. Genetic maps of *Pseudomonas aeruginosa* PAO and *P. putida* PPN showing auxotrophic loci. A comparison of the *P. aeruginosa* (inner circle) and *P. putida* PPN (outer circle) chromosome maps, with the *P. aeruginosa* map rotated 22 min clockwise to align the centers of the auxotroph-rich regions of each map. These are denoted by thickened lines and have been taken to extend from *val-2* (73 min) to *pyrF* (42 min) for *P. aeruginosa* PAO and from *argG* (21 min) to *met-810* (62 min) for *P. putida* PPN. Markers immediately adjacent to the circles are those whose map position is in close agreement in the two species after *P. aeruginosa* PAO map rotation. The nine markers farther from the circles are markers which may have been transposed from one position to another since divergence from a common ancestor (see text). Redrawn from reference 150. A description of loci is given in the original reference.

the route by which pseudomonads have acquired many of their versatile degradative functions, by independent events of interaction of their chromosomes with visiting plasmids, incorporating genes for new functions by illegitimate recombination at different loci in different bacteria.

Regulation of functionally related genes in *Pseudomonas* species is not exercised, as it is in *E. coli*, through sets of immediately adjacent genes in operons with polycistronic mRNAs; coordination of gene expression is accomplished in *P. aeruginosa* in unlinked or loosely linked groupings that may be the equivalent of *E. coli* regulons. In *P. aeruginosa* and *P. putida*, genes for closely related metabolic functions tend to group near one another, in noncontiguous but loosely linked, jointly regulated clusters (126, 150).

(ii) *Streptomyces* spp. *Streptomyces* species are gram-positive bacteria with complex developmental stages and the capacity for coenocytic mycelial growth. Hopwood has developed and continues to improve genetic systems for *Streptomyces coelicolor*; species; more than 120 genes had been mapped at the last count (154). The genetic map of *S. coelicolor* shows nonrandom distribution of the mapped genes, with a pronounced tendency of the genes to occupy two genetically dense regions on opposite sides of the chromosome. Flanking the dense regions are two regions of much lower genetic density, one region in fact almost empty of known genetic functions (Fig. 6). As in the case of *P. aeruginosa* and *P. putida*, one wonders whether the relatively empty portions of the chromosome reflect recent acquisition of two large tracts of DNA, or whether there is any relationship between the skewed distribution of genetic loci and some high level of functional and structural organization of the chromosome.

Hopwood has noted that genes for biochemically closely related functions are often located on opposite sides of the chromosome (152). A similar relationship exists for some *E. coli* genes (290). There is no information at present on whether these spatial relationships reflect ancient genome aggrandizement events, an operational aspect of chromosomal organization and gene expression, or simply chance positions of genes.

(iii) *Bacillus subtilis*. There is some clustering of genes of related function in the *B. subtilis* genetic map. The *rrn* genes tend to cluster, with 7 or 10 located between map positions 0 and 15° (out of a total of 360°) (170, 269). The ribosomal protein genes are clustered within the same region between 11 and 12°. The chemotaxis genes (*che*) are tightly clustered at 155°, and the sporulation genes (*spo*) are loosely clustered between 210 and 240°. At present, no relationship has been established between clustered map locations and coordination of expression of the functionally related genes.

(iv) Conservation of gene arrangement at origins and termini of replication. The cluster of genes around the origin of replication is similar for many bacteria (97, 256). Although there are differences in details, the broad features of the organization of the chromosome near the origin are conserved more highly than is the arrangement of the bulk of the chromosomal genes. Relative gene position seems to be an integral part of the orchestrated function of the genes of replication initiation. Similarities also exist at the terminus regions of bacteria as distantly related as *E. coli* and *B. subtilis* (196, 363), suggesting that initiation and termination of chromosome replication are coordinated by similar mechanisms in many bacteria.

Summary. We are beginning to learn about some of the complexities of the organization of the bacterial chromosome. It appears that in *E. coli* there are mild constraints on gene location and orientation exerted by gene dosage effects and by the factor of compatibility of directions of transcription and replication. From studies on permissive and nonpermissive inversion intervals in *E. coli* and *S. typhimurium*, we conclude that substantial constraints exist in the one-third of the chromosome centered around the terminators. There seems to be some supergenic organization in the terminal region of the chromosome that cannot be disrupted by inversions and is physically constrained in such a way as to make intrachromosomal recombination unlikely.

Patterns of gene locations in the admittedly incomplete maps of *Streptomyces* and *Pseudomonas* spp. carry clues to past events, hinting at acquisition of catabolic genes from extrachromosomal elements independently in related species. Seemingly nonrandom locations of functionally related genes in one half of a map or another may be quite accidental, reflecting aggrandizement and fusion events in the history of the chromosome, or the nonrandom locations may reflect a high level of organization and regulation of functionally related genes that are not yet properly perceived or understood.

Genetic Variation in Bacteria

Variable mutation rates. One of the classic experiments of bacterial genetics, the Luria-Delbruck fluctuation test (220), showed that bacterial mutations are not directed by selective conditions, but are neutral, and that they occur at random

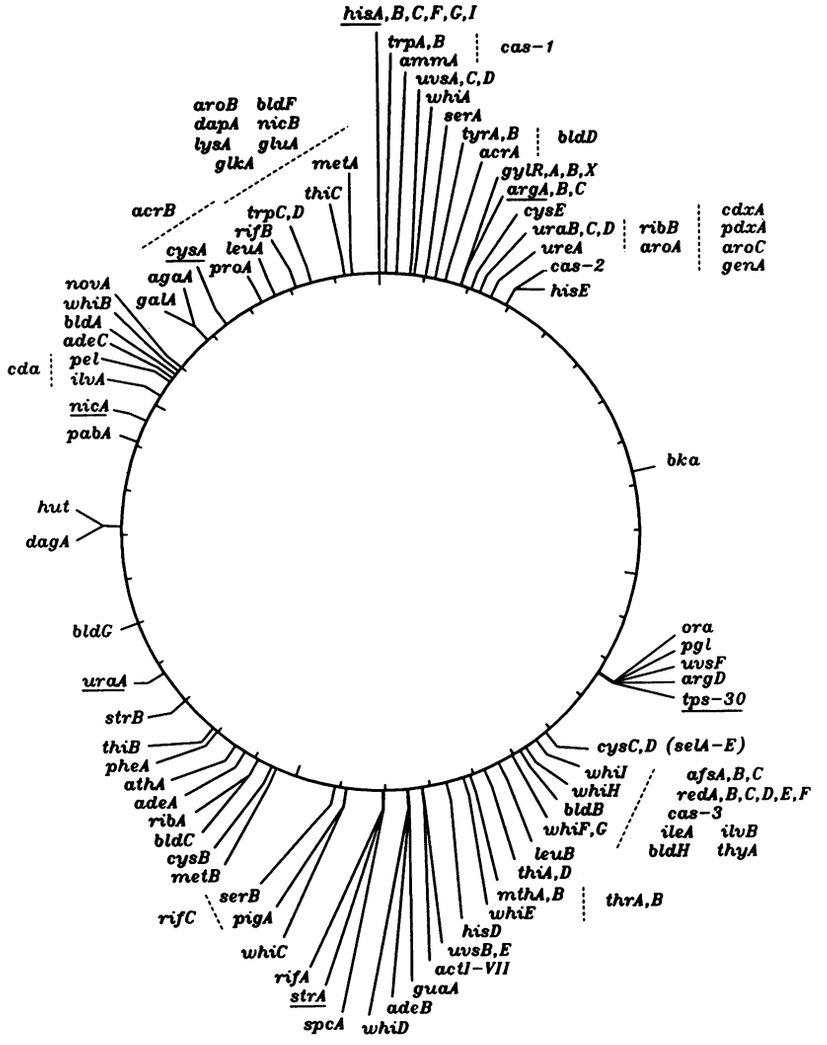


FIG. 6. Genetic map of *Streptomyces coelicolor* A3(2). Genetic loci placed outside of dashed lines have not been ordered relative to each other or to the mapped loci. The map is divided into 10° intervals with *hisA* set at 0°/360°. Its genetic length, as measured by recombination in matings, is some 260 centimorgans. Redrawn from reference 154.

with respect to genetic locus. The experiment did not address whether there also might be another type of mutation that was directly responsive to the needs of the cell under stressful environmental conditions. This possibility has been examined recently.

A fresh look at the factors that influence kinds of bacterial mutations and factors that increase the mutation rate has followed on a report by Cairns et al. (44) about unusually high mutation rates seen under stressful physiological conditions in genes capable of relieving the stress after mutation. In one of the experiments reported, under conditions of carbon starvation and in the presence of lactose, a *lacZ* amber mutant reverted to lac^+ at an unusually high rate, whereas the rate of mutation to valine resistance, a phenotypically unrelated mutation, was not unusually high. Cairns et al. (44) suggested that *E. coli* undergoes two types of mutations, one that they called spontaneous mutations, i.e., those that are random with respect to locus and are not responsive to environmental conditions as documented in the Luria-Delbruck experiments, and another type that they called directed mutations, i.e., those that occur at a higher rate when they are advantageous to the cell.

Other investigators have reported that carbon starvation and provision of an unusable catabolite in the medium have led to high-frequency excision of a mobile genetic element from an *ara*-Mu-*lac* fusion (314) and a *bgl* gene carrying an IS*103* insertion (124), thereby restoring in each case an appropriate and relevant catabolic gene to normal function. Other genes unrelated to the starvation condition did not experience the elevated mutation rate. Are these changes involving deletion of mobile genetic elements also directed mutations in the sense that the cell has a mechanism for accelerating mutations whose outcome will be beneficial to the cell? Hall refers to this type of mutation as Cairnsian mutation (124a).

Several mechanisms have been suggested to explain high-frequency, Cairnsian mutations. One is that variant mRNAs might accumulate in starved cells: these might undergo reverse transcription, and when a relevant genetic function was restored, the bacteria so endowed would survive and persist as successful mutants (44). Another mechanism invokes stalled transcription: genes induced to express by environmental conditions, but unable to be successfully transcribed because of a fault, could be stalled in transcrip-

tion with the DNA in a partly unwound condition, vulnerable to a high rate of mutational change (64). Another mechanism involves fixation of transient mutational changes: since energy and metabolism are required to fix transient mutational changes by DNA replication, perhaps fixation occurs more frequently in cells that, as a result of the mutation, had successfully recovered their physiological health (332). Evidently, there is no shortage of plausible molecular mechanisms (or statistical expectations [337]) for the occurrence and fixation of Cairnsian mutations in starved, stressed bacterial populations.

Recently Hall (124a) examined the production of *trp*$^+$ point revertants from *trp* auxotrophs under conditions of tryptophan deprivation. Again, the phenomenon of apparently specific response was seen. The production of starvation-relieving *trp* prototrophs occurred at high frequency compared with the rate of production of other, selectively irrelevant mutations, valine resistance and *lac* constitutivity. In a control experiment, under conditions of cysteine deprivation, production of irrelevant *trp*$^+$ mutants was not increased. Hall proposes a model of random mutations occurring in a small fraction of the bacteria that have entered a hypervariable state. He pictures that some of the bacteria in the hypervariable state will accumulate irrelevant and/or lethal mutations and will not survive, but bacteria in the hypervariable state that mutate so as to be able to survive under the selective conditions are the visible mutants that are scored. Thus, survivors contain a high proportion of relevant mutations that cope with the selective conditions. Hall predicts, then, that the *trp*$^+$ survivors, having been in the hypervariable state, should have sustained mutations in genes other than *trp* at higher than usual frequencies, and although the numbers to date are small, as far as they go, the data bear out the expectation. Of 110 *trp*$^+$ prototrophs, 2 had become auxotrophic in other markers, well above the expectations for the frequency of independent mutational events in normal cells.

The Hall model proposes that a fraction of a stressed population becomes hypervariable to mutation in all genes. If the model is correct, the essentially random nature of the mutation process is retained, yet an appropriate global cellular response is invoked that would increase the chances of survival of a population faced with deleterious, potentially lethal environmental conditions. Future tests will establish whether this attractive model holds true for other genetic loci and other bacteria.

Variability within species. What do we know about genetic variation in bacterial populations, and what does this tell us about the dynamics of genetic change in bacterial chromosomes? Some bacteria exhibit greater variability than others. Two that exhibit very little variation are *Mycobacterium* and *Rickettsia* species; others that are highly variable include *Rhizobium*, *Halobacterium*, and *Streptomyces* species. Between these extremes lie many eubacteria.

Strains of *Mycobacterium leprae* isolated in India, West Africa, and Louisiana from a human, monkey, and armadillo, respectively, have essentially identical genomes. No polymorphisms were detected among restriction endonuclease digests of these genomes when they were hybridized with a variety of probes. From the fraction of conserved restriction endonuclease sites, it is possible to estimate nucleotide sequence divergence among strains. By using these parameters, the extent of divergence among the isolates from three continents was calculated to be 0.02 to 0.26% (53). By the same measure, *E. coli* isolates from widely separated geographical areas differed in four separate chromosomal locations by about 3% (range, 0.8 to 6.6%) (127). In a more detailed study, *E. coli* isolates differed over a range of 0 to 4% in 15 1.5-kb regions within a 40-kb chromosomal segment (238a).

For *Rickettsia bellii*, an intracellular parasite, no appreciable divergence was observed with more than 53 isolates collected during an 8-year period in a 500-km^2 region of western Ohio. When the sample was expanded to include the entirety of North America, the observed diversity was estimated to be 1/10 the amount seen in *E. coli* (96).

By contrast, other bacteria, such as *Streptomyces* and *Rhizobium* species, are highly variable. These bacteria harbor variable plasmids, insertion elements, and repeated sequences, probably agents of facile change. *Streptomyces* strains are highly variable in culture. One commonly observes variants for colony morphology, pigmentation, and the like. As described in Amplification (above), a particular form of genetic instability is present in many *Streptomyces* spp.: a high degree of amplification of sequences in the size range from 5 to 25 kb occurs; these amplifications may be as great as one-third to two-thirds again the size of the chromosome and may be accompanied by deletions of hundreds of kilobases.

Rhizobium strains are also highly variable in culture. Typical of the variability observed were the patterns of chromosomal Southern hybridizations given by DNA from single-colony isolates from cultures of *Rhizobium phaseoli* in response to some hybridization probes (90). Besides carrying genetically active plasmids that promote rearrangements (24), many *Rhizobium* species harbor multiple copies of IS elements and other repeated sequences (89, 173).

In their degree of genetic variability, many eubacteria lie between the extremes typified by the conserved *Mycobacterium* species and the variable *Streptomyces* species. We are not really sure why some bacteria are characterized by conservatism and stability, whereas others undergo rapid change and variability. One can ask whether there is any pattern to the genetic variation in bacteria. Can the approaches of population genetics be used fruitfully to learn about bacterial population structure and the dynamics of bacterial chromosome structure in genetic terms? Analysis of some of the genetic variation among *E. coli* strains has provided information on the population structure of that organism. In the sections that follow, we summarize information from several approaches. Allelic variation in the electrophoretic mobility of protein gene products (allozyme variation) has provided information on the population structure of *E. coli*; comparative nucleotide sequence data on individual genes have given information on genetic activity within genes; restriction fragment length polymorphism analysis on a longer span of the chromosome encompassing several genes has given information on dynamics of genetical change within the chromosome. These studies will be summarized below.

Clonal structure of bacterial populations. The genetic characteristics of populations of bacteria have been studied by assaying the amount of allelic diversity at several gene loci by multilocus protein electrophoresis. Many genetic loci in bacteria are polymorphic, often producing electrophoretically distinguishable enzymes or outer membrane proteins. Pioneering studies with bacteria were carried out with 829 isolates of *E. coli* from various human and animal hosts, examining five enzyme loci for the presence of allozymes that are detectable as electromorph variants (237, 238). The five loci were found to exhibit a high level of polymorphism.

Later, in a much larger study, the electrophoretic types (ETs) of 1,691 strains of *E. coli* and closely related *Shigella* bacteria were determined by assaying the electrophoretic properties of proteins at 12 loci (254, 312, 374). Bacterial strains can be characterized in terms of their ETs, a summation of information on all loci surveyed. For the 1,691 isolates, although there was a high level of genetic variation at individual loci, only 302 different ETs were observed. There was a pronounced tendency for some combinations of alleles to be associated. A principal component analysis of the allele profiles of these 302 ETs showed that they fell into three overlapping clusters (Fig. 7) (374). The closely related *Shigella* isolates fell within one of those clusters. Thus, there were many fewer types than would be expected if the alleles at the 12 loci had reassorted independently of one another.

The pronounced tendency for some combinations of alleles to be associated and for only three main types of allele combinations to dominate the entire *E. coli* population suggests a clonal subspecies structure of the *E. coli* population in which the conserved unit is basically the entire bacterial chromosome (254, 312) and recombination frequencies are low (206).

Similar studies of 142 isolates of *E. coli* K1 (a pathogenic strain) gave similar results (2, 253). Allozyme variation of 12 enzymes and four electrophoretic properties of outer membrane proteins classified the 142 strains into only two major groups and one smaller group. (Interestingly, serotypes were not well correlated with these groupings, and were found to be poor indicators of clone membership or ET type [2, 128, 253].) The *E. coli* population with relatively few clusters of ETs can be viewed as a population containing relatively few worldwide clones in which the association of all alleles in the chromosome tends to be maintained.

Bacteria other than *E. coli* have also been examined for allozyme patterns by protein electrophoresis. The list includes *Neisseria meningitidis*, *Haemophilus influenzae*, *Legionella* spp., *Serratia* spp., *Salmonella* spp., and *Haemophilus pleuropneumoniae* (see references 21, 49, 103, 243, 244, 313, and references in reference 312). In all cases, the amount of diversity found among independent isolates was much less than expected if all alleles assorted randomly and independently of each other. These and other similar studies have established that chromosome-wide linkage tends to be maintained in many bacteria.

Linkage disequilibrium, or a paucity of genotypes, in bacterial populations would be expected to arise (i) if recombination rates between *E. coli* strains are low, thus maintaining linkage throughout the chromosome; (ii) if, when favorable alleles are acquired, other unselected but linked genes are carried along (hitchhiking effect); and (iii) if occasionally there are extinctions of clones in the population, replaced by other, fitter clones (periodic selection). Each of these factors plays a role; for discussion and literature citations, see references 128, 238a, 255, 312).

Intragenic recombination among *E. coli* strains. The existence of relatively few ETs in *E. coli* strains was an important finding that revealed the essential stability of the chromosomal genotype in a few widely distributed clones. The stability of the chromosomal genotype implies a low rate of recombination in *E. coli* populations. However, although the rate of recombination may be low, this does not signify that recombination is not an important process. The recombination events that do occur can have important genetic and evolutionary consequences.

Some of the recombination that takes place is within genes. Examining the genotypes of *E. coli* strains at a higher level of magnification, i.e., looking at detail within genes, has given information on the genetic activity taking place at the level of nucleotides. Determination of nucleotide sequences of homologous genes among laboratory and wild isolates of *E. coli* has provided information on occurrence of intragenic recombination.

The gene for alkaline phosphatase, *phoA*, was cloned from eight naturally occurring strains of *E. coli*, sequenced, and compared with the sequence of the *phoA* gene from a laboratory strain of *E. coli* K-12 (76). The spatial distribution of the variable nucleotides was found to be highly nonrandom. The sites were clustered and were assorted in combinations consistent with a history of intragenic recombination among the nine *phoA* alleles.

Sequences of the *gnd* gene (for 6-phosphogluconate dehydrogenase) in seven naturally occurring isolates of *E. coli* have also been determined (78, 299). Sequences of the *gnd* genes were much more variable than were the *phoA* sequences, but the *gnd* sequences were similar to *phoA* in that they also contained highly clustered polymorphisms in combinations that suggested intragenic recombination. Statistical analysis of data for both genes provided evidence for multiple intergenic conversion events and evidence that the participating segment sizes in the *gnd* locus were in the range 70 to 200 bp in length (298). The calculated rate of the gene conversion events was on the order of the neutral mutation rate. Evidence for recombination in the regulatory region of *gnd* genes of different *E. coli* strains was found in nucleotide sequences. The leader region of the *E. coli* B/r *gnd* gene appears to be the product of recombination between two allele types found in natural populations (16).

Thus, by examining the structure of *E. coli* genes at a high level of detail, evidence has been found for an active process of intragenic recombinational events such as gene conversion. The intragenic activity can be visualized as a gene conversion activity superimposed on the background of the stable chromosomal structure of the *E. coli* population or as representing the edges of clonal segments that exist within the chromosome (as explained in the following section).

Clonality of segments of the bacterial chromosome within a stable frame. In addition to the comparative nucleotide sequences of the *phoA* and *gnd* genes referred to in the previous section, extensive data have been gathered on sequences and restriction fragment polymorphisms of *trp* and neighboring genes in several *E. coli* isolates (239, 240, 338). Comparative sequences of genes of the *trp* operon among *E. coli* isolates was carried out earlier (239). Later, this study was extended beyond the *trp* operon, through six neighboring ORFs up to the P14 and *tonB* genes, almost 4 kb of DNA (240, 338). More recently, the region examined has been extended to include 40 kb of DNA (238a). Genetic relationships among 14 *E. coli* strains (a laboratory K-12 strain and a roster of isolates from the wild in the ECOR collection) have been examined. Comparative analysis of restriction fragment length polymorphisms has revealed an overlay of point mutations (often neutral third-position codon changes) and some large rearrangements such as apparent acquisitions of substantial segments of DNA on the order of 20 to 50 kb in size (238a). The overall picture is one of extensive regions of identity punctuated by clustered differences.

The analysis of the 4-kb region near *trp* showed that major rearrangement events had occurred in and around four of the six ORFs near *trp*. Relative to a laboratory strain of *E. coli*

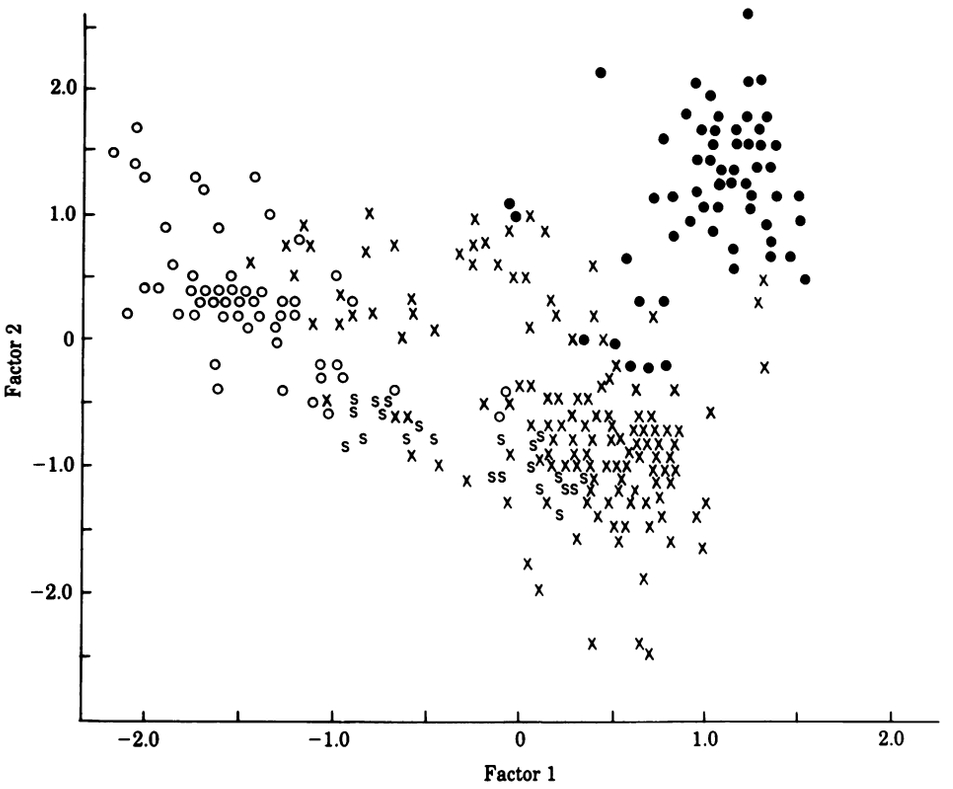


FIG. 7. Factor scores of electrophoretic types of *E. coli* (n = 279) and *Shigella* species (n = 23). Reprinted from reference 374 with permission from the author. Details are given in the original reference.

K-12, insertions or deletions of genetic segments of unknown origin took place in each strain. A deletion affected ORF I and the beginning of ORF II in one strain; an insertion of IS*1* and other genetic material occurred in ORF III in another strain. Nonidentical insertions, sometimes with duplication of part of ORF IV, occurred in four strains. The clustered polymorphisms appeared in a pattern of combinations in the strains analyzed, suggesting that recombination occurred between the strains in the ORF region both between and within the inserted or deleted segments (240, 338).

The clustering and sharing of polymorphic sites among *E. coli* strains suggest a dynamic picture of genetic activity generating complexity and variety in the *E. coli* chromosome. Among the *E. coli* strains examined, the sequences were largely similar, but were sprinkled with clusters of polymorphic sequence differences. Analysis of the composite makeup strongly implies, as with the *phoA* and *gnd* genes, localized recombination events against a background of shared, stable sequences (240, 338).

A picture of the dynamics of change in the chromosome as a whole emerged from the comparative study of the 40-kb region (238a). Milkman and colleagues refer to the background chromosome as a whole as the frame within which individual segments can have independent phylogenetic histories. The comparative data are consistent with population dynamics that entail rapid ascension of an *E. coli* clone or sequence type to worldwide prominence as a result of acquisition of a broadly favorable allele. The chromosomal frame of this fit variant is widely shared. With time, recombination and mutation introduce differences and the chromosomes begin to collect variant regions. The consequences to the makeup of the bacterial chromosome are portrayed schematically in Fig. 8. Sections 1, 2, and 3 portray a progression of clonal segment replacements with time, alternatives a and b showing different consequences according to the relative sizes of the segments exchanged. In section 1, as an arbitrary point of departure, the genotype of the *E. coli* chromosome is represented as being all white at some arbitrary time in the past. In section 2, changes were introduced by acquisition of clonal segments containing favorable alleles, represented as introductions of patterned segments replacing part of the white frame. The acquired segments would be expected to carry unselected, hitchhiking neighbor loci in addition to the favorable allele. The acquired genetic segments are pictured as large relative to the whole chromosome in option a and relatively small in option b. In section 3, repeated acquisition by recombinational replacement of other clonal segments carrying other favorable alleles would in time lead to acquisition of genetic material of different heritage at other loci, represented by appearance of new patterns. If the recombination frequencies were high and recombining segments were large (option a), in time new acquisitions would blur the outlines of the segments acquired in Section 2 and ultimately would obscure the original (white) frame. On the other hand, if recombination frequencies were lower and the segments were small (option b), replacement would proceed relatively slowly and the original frame would still be evident.

On the basis of the comparative restriction fragment length polymorphism data, the rate of replacement was calculated to be approximately 2.4×10^{-12} recombinational replacements per nucleotide per generation (238a). This is the equivalent of 1 replacement per 10^5 genomes per generation. In a million generations, if replacements were 10^5 bp

in length, 20% of the chromosome will have been replaced. If the replacements were only 10^3 bp in length, only 0.2% of the clonal frame would have been replaced in that time. These approximations are beginning to give us informative limits on the kind and frequency of genetic activity that is shaping the *E. coli* genome and determining population structure.

In an independent study of another genus, comparison of nucleotide sequences and hybridization patterns revealed a mosaic organization of the chromosome near the *toxA* gene in 10 strains of *Pseudomonas aeruginosa* (274a).

Summary. The picture of the *E. coli* chromosome that emerges from these studies is one of basic stability with a series of overlaid recombinational and mutational events. By recombination, preexisting alleles are replaced by segments containing favorable alleles together with sets of hitchhiking genes. This activity in turn is overlaid with small intragenic gene conversion events on the order of 70 to 200 bp. At the same time, mutational events occur, ranging from single-nucleotide substitutions to larger rearrangements such as inversions, transpositions, and duplications. Amid this dynamic genetic activity, these multiple sources of genetic variation, the essence of the bacterium *E. coli* is preserved and the basic stability of the chromosomal frame is maintained. Evidently recombination and mutation rates are sufficiently low that only a few stable clonal frames dominate the worldwide *E. coli* population.

The extent to which recombination between sequences obscures the chromosomal frame is a function of recombination frequencies and sizes of recombining segments. The magnitudes of some of the parameters of this process are not yet known, although reasonable estimates have been made (238a). If recombination frequencies are high, more segments would replace homologous sections of the frame, and if the average size of the segments were large, newer segments would encroach on older ones, both old and new segments progressively replacing the original chromosomal frame (Fig. 8, sections 2a and 3a). If recombination frequencies are low (128, 206, 312), and if the average size of segments acquired by recombination is small, the picture is one of a clonal population with small patches of variation throughout the chromosome, with the frame left essentially undisturbed (Fig. 8, sections 2b and 3b). The rate at which broadly favorable alleles arise by mutation will affect the rate of replacement of the relatively few electromorphic types in the entire *E. coli* population (255, 312, 374).

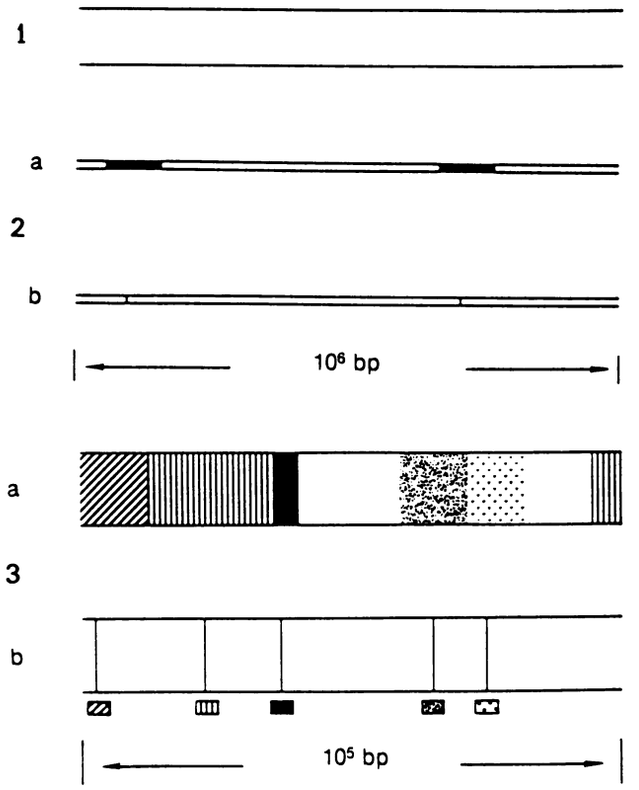


FIG. 8. Stretches of chromosome undergoing replacements with clonal segments over time. In section 1, the chromosome is represented as a single clone, all white, at some arbitrary time in the past (although in reality it will have multiple ancestries from even earlier times). Section 2 shows a 10^6-bp stretch of chromosome one million generations after the (arbitrary) origin of the clone: (a) the average replacement segment size is 10^5 bp; (b) the average replacement size is 10^3 bp. Section 3 shows a 10^5-bp stretch after 10 million generations: (a) the average replacement size is 10^5 bp; (b) the average replacement size is 10^3 bp. See the text for a discussion. Adapted from Milkman and Bridges (238a) with permission.

Integrity of Bacterial Species versus Lateral Transfer

Besides genetic exchange within a species, another source of genetic variation in bacteria is lateral transfer from distantly related organisms. Although lateral transfer may occur infrequently, the acquisition of fresh genetic information is probably important to the vitality of a species. Yet there are limits. The genetic advantage of tapping into novel sources of genetic information must be weighed against the requirement to maintain the integrity of the species, which puts a limit on the amount of interspecific chromosomal genetic exchange.

There is a growing understanding of some of the factors that account for the normally low efficiency of interspecific recombination. Standard modes of genetic transfer require that compatible cellular structures and physiological capabilities be in place. Without appropriate phage receptors, transducing phages cannot infect. Without the capability to mobilize and transfer donor DNA and to build a physiologically feasible mode of intercellular DNA transfer through direct cell contact or a conjugation tube, bacterial conjugation cannot proceed.

Once transferred into an unrelated organism, newly acquired DNA has to be able to survive the protective restriction-modification systems of the host. In gene transfer in *E. coli*, although linear single-stranded DNA is transferred, it is not a stable intermediate in the recipient cell, but becomes double stranded during association with the resident chromosome. Complex recombinational intermediates form that undergo repairlike DNA synthesis (279). If the newly transferred DNA strand and its newly synthesized complement were unmodified, restriction enzyme(s) of the recipient organism could fragment the recombining DNA before stable integration into the chromosome was achieved (for reviews on restriction endonucleases, see references 26 and 386).

All the factors listed above militate against transfer of genetic material between species and incorporation of the transferred DNA into the resident chromosome. Another factor acts to reduce intergeneric recombination: mismatch repair mechanisms.

Mismatch repair limits genetic exchange between dissimilar DNAs. Another barrier to recombination between imper-

fectly matched DNAs is the action of mismatch repair systems (reviewed in references 241 and 280). The methyl-directed system in *E. coli* is a strand-specific repair that is directed by the state of adenine methylation within Dam sites. A major role of the complex is to improve the fidelity of replication by repairing any mismatch in the unmethylated newly synthesized strand so as to pair correctly with the old, methylated template strand. Mutants with mutations in *mutH*, *mutL*, *mutS*, or *mutU* lack parts of the system and have the phenotype of mutator strains, which are poor at correcting mistakes made during DNA replication.

Recent work implicates the mismatch repair system in another role as well: that of restraining interspecific recombination (282). In wild-type bacteria, efficient homologous recombination occurs only when the base-pair match between recombining DNAs is high (318). Recombination rates are severely reduced when there is about 10 to 20% sequence divergence, approximately the same amount of divergence that exists between many *E. coli* and *S. typhimurium* genes (305, 318). Recent experiments have shown that the failure of recombination when the pairing is imperfect is a consequence of methyl-directed mismatch repair systems (282). Degradation of a mismatched strand prevents the formation of interspecific recombinational intermediates, thus aborting the recombination event. Successful recombination by either conjugation or transduction between *E. coli* and *S. typhimurium* was increased when the recipient had mutations in genes for mismatch repair: recombination frequencies were increased 30- to more than 1,000-fold in *mutH*, *mutS*, *mutL*, and *mutU* mutants, with the strongest effects seen for *mutL* and *mutS* mutants (282).

These observations suggest strongly that wild-type mismatch repair systems are an important part of the genetic system that establishes reproductive isolation between diverged species. Such a monitoring system that detects and either corrects or aborts improperly paired DNA could also play a role within a species in regulating the frequency of recombination between polymorphic variants within a bacterial population. Also, monitoring for mismatches might reduce recombination within a chromosome between imperfectly matched copies of repeated genetic elements such as IS elements or tRNA genes. Reducing the frequency of intrachromosomal recombination would minimize generation of duplications or rearrangements of chromosomal gene order.

Fidelity-monitoring systems are obviously an important part of maintaining genetic integrity both within and between bacterial species. However, the level of control must be critical because if the stringency of prevention of recombination between partially mismatched DNAs is too great, opportunities for introduction of genetic diversity will be lost, whether that variation is produced by intrachromosomal, intraspecies, or interspecies recombination.

Plasmids as ferries for bacterial genes. Plasmids are known to carry genetic functions that are useful to bacterial hosts; their variety and role in bacterial populations have been reviewed (see, for example, references 45, 207, and 283) and is discussed above in Plasmid-Chromosome Interactions). Following transfer to a recipient organism, plasmid genes may ultimately be acquired by the recipient chromosome. Entire plasmids, individual genes, or clusters of genes may integrate into the chromosome of the recipient.

A given gene or group of genes may be plasmid borne in some bacterial strains and chromosomally located in others. An example described above (in When Does a Plasmid Become a Chromosome?) is the cluster of *xyl* genes for degradation of toluene and xylene, carried on pWWO, a TOL plasmid, but also found chromosomally located in a strain of *Pseudomonas putida* MW1000 (318). DNA hybridization shows that another cluster of genes for benzoate utilization, *ben*, is related to the *xyl* cluster. Both the *P. aeruginosa* PAO and *P. putida* MW1000 chromosomes carry the *ben* cluster of genes. Evidently the *xyl* degradation genes of plasmid pWWO have duplicated, diverged in function, and dispersed to at least two bacterial species by lateral transfer (149).

How do genes carried by plasmids bypass cellular mechanisms designed to discourage interspecific chromosomal recombination? Some plasmid-borne bacterial genes lie within a transposon in the plasmid and transfer to recipient genome by transposition. For instance, the capacity for citrate utilization is present in some but not all enteric bacteria. The genes for citrate utilization are present in widely distributed plasmids (166). In some cases the genes reside in a transposon, Tn*3411*, flanked by two copies of IS*3411* (167). Similarly, *lac* genes are carried by many plasmids, converting essentially *lac*-less bacterial species to *lac*$^+$ (59, 121, 285); in some cases the *lac* genes reside in a transposon structure, as in Tn*951* (59).

Plasmid transfer across taxonomic boundaries. Physiologically unlikely matings do occur (331). *Agrobacterium tumefaciens* Ti plasmids promote conjugal transfer of plasmid DNA from bacterial to plant cells and integration of plasmid genes into the plant DNA (387). Plasmid replication systems exist that are able to function in many distant bacteria, and even in yeast cells: plasmids of *Staphylococcus aureus* have such broad capabilities (112). Recently, chimeric plasmids have been constructed that contain origins of transfer, mobilization functions, and transfer functions from a mixture of sources and are capable of transferring between distantly related bacterial donors and recipients and of replicating and expressing a transferred gene in the recipient. Conjugal transfer has been engineered between *E. coli* donors and gram-positive recipients such as *Enterococcus* spp., *Streptococcus* spp., *Bacillus* spp., *Staphylococcus* spp. (351), gram-positive corynebacteria (301), and actinomycetes such as *Streptomyces* spp. (231). In the other direction, plasmids have been constructed that pass from *Enterococcus faecalis* and other gram-positive donors to *E. coli* recipients (38, 350). Typically, the chimeric plasmids constructed for this purpose contain origins of replication for both donor and recipient bacteria, an origin of transfer (*oriT*) from a plasmid, and transfer functions (*tra*) including a mobilization function (*mob*) sometimes inserted into the plasmid to be transferred and sometimes supplied in *trans* by another plasmid in the donor cell. To select transconjugants, a selectable marker is included in the chimeric plasmid, often an antibiotic resistance determinant. Frequencies of transconjugant formation by such hybrid conjugal systems were found to range from 5×10^{-9} to 1×10^{-2} (231, 301, 350, 351).

Reaching across the gulf between a procaryote and a eucaryote, a transfer system permitting conjugation between *E. coli* and *Saccharomyces cerevisiae* has been devised (136). A chimeric plasmid was constructed including the necessary functions of an origin of bacterial replication, an origin of plasmid transfer, a selectable yeast *leu* gene, and the capability of replicating in *S. cerevisiae*, derived from the 2 μm circle of *S. cerevisiae*. Mobilization functions of ColE1 were supplied in the donor bacteria in another plasmid. In matings on solid media, frequencies of *E. coli* × *S. cerevisiae* transconjugant formation as high as 5×10^{-5} were

observed (136). The physiology of cell contact and DNA transfer have not yet been worked out in this extraordinary transkingdom genetic transfer.

These experimental systems involving engineered chimeric plasmids show that lateral gene transfers across wide taxonomic gulfs are possible and emphasize the importance of the cellular monitoring mechanisms that regulate frequencies of unorthodox recombinations.

Indications of a history of lateral transfer of genes. We do not know what fraction of any contemporary bacterial chromosome was acquired from external sources by lateral transfer. There is circumstantial evidence indicating that for some genes lateral transfer and acquisition have in fact occurred. A high level of sequence similarity between genes in distant organisms suggests past transfer and acquisition. For instance, two staphylococcal plasmids carry detectably different determinants for chloramphenicol acetyltransferase. The nuclear *cat* genes of some streptococcal strains are similar as judged by DNA hybridization to one staphylococcal *cat* determinant, and genes of other streptococcal strains are similar to the other staphylococcal *cat* determinant, suggesting that there has been genetic exchange between staphylococci and streptococci at some point in the past. Of 21 plasmidless chloramphenicol-resistant *Streptococcus* strains, in eight strains the chromosomal genes were hybridizable to a *cat* determinant in one *Staphylococcus* plasmid, in three they were hybridizable to a different *cat* determinant in another *Staphylococcus* plasmid, and in seven they were not hybridizable to either (265). Intergeneric transfer and chromosomal acquisition probably occurred in the past in a significant fraction of the strains examined.

If nucleotide or amino acid sequences of a gene or gene product are very similar in otherwise distantly related organisms, one can suspect that lateral transfer has occurred, but this possibility has to be distinguished from the possibility that there are stringent functional requirements that do not tolerate mutational change. Examples of highly similar sequences in procaryotes and eucaryotes are phosphotransferase system enzyme III for cellobiose in gram-negative *E. coli* and phosphotransferase system enzyme III for lactose in gram-positive *Staphylococcus aureus* (262); 10-formyl tetrahydrofolate synthetase in *Clostridium acidiurici* and *Saccharomyces cerevisiae* (373); glutamine synthetase II in *Bradyrhizobium japonicum* and in plants (47); hemolysin transport protein in *E. coli* and P glycoprotein in animals (104); certain sugar transport proteins in *E. coli* and in mammals (135, 226); α-amylase in *Streptomyces* spp., invertebrates, and mammals (216); and adenylate cyclase in *Rhizobium meliloti* and mammals (25).

Reports of unusual sequence similarities must remain anecdotal until there is information on the sequences of genes for the same function in many organisms and/or information on the flexibility of the gene product to undergo change while retaining function. This will help distinguish between sequence similarities reflecting conservation of function and those reflecting recent common ancestry through lateral transmission.

Anomalous nucleotide composition relative to the rest of the chromosome also suggests recent external origin. An erythromycin resistance determinant, *ermBC*, was found in a clinical isolate of *E. coli* with a nucleotide sequence closely similar to that of an erythromycin resistance determinant in a transposon of *Enterococcus faecalis* and also to that of a determinant from a plasmid found commonly in *Streptococcus* spp. The G+C content of the *ermBC* determinant in the *E. coli* strain was 33%, far from the typical 50% of *E. coli* genes and characteristic of the chromosomal genes of *E. faecalis* (38).

Another example is the *argF* gene of *E. coli* K-12. Two genes for ornithine transcarbamylase coexist in the chromosome: *argI*, a gene held in common with other *E. coli* strains, and *argF*, peculiar to K-12 and located on a 10-kb segment that is flanked by two IS*1* sequences (158, 385). Even though the nucleotide and amino acid sequences of *argI* and *argF* have evident similarities and could conceivably have arisen by internal gene duplication, transposition, and divergence, the location of *argF* on a segment between flanking IS sequences suggests acquisition by transposition from another genetic source. The relationship between the sequences of the *argF* and *argI* genes recently has been examined further (357). The *argF* gene is anomalous among *E. coli* sequences with respect to its overall G+C content (59% rather than the more typical 52 to 53%) and is anomalous in the G+C content of its codon third positions (75.8% compared with the average for *E. coli* of 56%). Van Vliet et al. (357) suggest that the *argF* gene was not generated by duplication of the *E. coli argI* gene, but was acquired laterally from a bacterial species which has an overall G+C content higher than that of *E. coli*, for instance, *Klebsiella pneumoniae* or *Serratia marcescens*.

Summary. In summary, there can be no doubt that some genes have been acquired by the bacterial chromosome through lateral movement and that acquisition of foreign DNA is an important source of genetic variation. Examples are known of plasmids that can move across wide taxonomic gulfs, bringing genes of different origins into the same cell. Once a foreign gene is in a new environment, there are cellular mechanisms that exercise some control over whether the new gene is incorporated into the resident chromosome. Mismatch repair systems abort recombination between dissimilar DNAs. As discussed in Gene Position Effects (above), there may be some feature of gene arrangement in the bacterial chromosome that cannot be altered. Thus, there is an interplay between the benefits of acquiring new genetic information and the need to retain genetic integrity and species identity. We are making progress in identifying the components of the elements in this balance.

PERSPECTIVES

There are many perspectives from which to view the bacterial chromosome, i.e., as a physical entity, a seat of DNA metabolism and gene expression, a genetic entity, and an important determinant of population structure and dynamics.

One view of the bacterial chromosome is the physical one, the view of the compact intracellular body of concentrated DNA. The nucleoid structure is a dynamic protein-polynucleotide complex, subject to localized associations and disassociations during gene expression. Independent supercoiled domains are capable of responding to environmental change; specific genes and groups of genes can be turned on and off, sometimes through changes in the secondary structure of the DNA and sometimes at a distance through interaction of distant DNA loci mediated by specific proteins. Much more has yet to be learned about the structure of the nucleoid and about localized structures of the DNA and of protein-DNA complexes, all of which affect the expression of the network of interrelated genes of the chromosome. Probably studies in these areas will be teaching us more about supra-operonic regulatory phenomena capable of responding to environmental conditions.

Another view of the bacterial chromosome is as a mosaic of genetic information derived from many sources. The DNA is subject to many avenues of genetic change: through mutation by base substitutions, small insertions and deletions, and major rearrangements; through recombination and genetic exchange within the species; and through lateral transmission of segments from extrageneric sources. These several processes are superimposed on one another, providing rich possibilities for genetic variation.

A balance is struck between the opportunities for change and the need for an essential genetic stability and identity. Some of the conservative factors that limit the range of genetic activity are now being identified. Nucleotide replacements are constrained by factors such as base composition of the genome, codon usage, codon context, and conservation of amino acid function in the gene product. Some rearrangements are apparently deleterious to the cell, such as inversions that rearrange by interrupting either of two large genetic regions that flank the termini of replication. Other rearrangements do not arise because of a barrier to their formation (intrachromosomal recombination in the terminus region apparently does not occur). Other factors such as preservation of the lengths of replicating arms and directions of transcription, also discourage gene shuffling. In these respects, we can now understand some of the constraints that shape the genetics of the bacterial chromosome. However, we still do not understand what factors are responsible for apparently nonrandom arrangements of functionally related genes in *Pseudomonas* spp. and *Streptomyces coelicolor*. The patterns of gene location may be remnants of their history of acquisition of genetic segments by the chromosome, or there may be in place today systems of gene expression or regulation that require maintenance of specific gene locations for optimal cell regulation.

We are also beginning to be able to paint a picture of the dynamics of the bacterial chromosome in terms of bacterial populations. Through determination of multilocus electromorphs and through comparative restriction fragment length polymorphism and sequence analysis of related strains, *E. coli* chromosomes are currently viewed as relatively stable genetic frames that seldom undergo recombination, but do acquire from time to time by recombination a replacement segment of chromosome carrying a favorable allele. The genotype of the fitter strain then rises to worldwide prominence. Models describing the genetic dynamics of populations of bacteria in terms of chromosome structure are currently emerging from a melding of both molecular biology and population genetics, giving a picture of the dynamics of formation, maintenance, and continuing change of the mosaic of the bacterial chromosome. Through these models it is now possible to make reasonable estimates of rates of mutation, recombination, and fixation in natural populations of bacteria. Continued rapid progress in the molecular population biology of bacteria seems likely.

The more we learn, the better we can formulate questions about the areas in which knowledge is incomplete. Clearly, there is much to learn about the functional significance of gene location and gene order, management of the nucleoid and its functions, and the dynamics of population biology. There seems to be no danger of running out of interesting questions to be studied.

ACKNOWLEDGMENTS

M.R. and S.K. acknowledge support by grant DMB8996209 from the National Science Foundation and contract DE-AC22-89PC89903 from the Department of Energy (DE-AC22-89PC89903), respectively.

LITERATURE CITED

1. **Abraham, J. M., C. S. Freitag, J. R. Clements, and B. I. Eisenstein.** 1985. An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. Biochemistry **82**:5724–5727.
2. **Achtman, M., A. Mercer, B. Kusecek, A. Pohl, M. Heuzenroeder, W. Aaronson, A. Sutton, and R. P. Silver.** 1983. Six widespread bacterial clones among *Escherichia coli* K1 isolates. Infect. Immun. **39**:315–335.
3. **Adhya, S.** 1989. Multipartite genetic control elements: communication by DNA loop. Annu. Rev. Genet. **23**:227–250.
4. **Akiyama, M., H. Maki, M. Sekiguchi, and T. Horiuchi.** 1989. A specific role of MutT protein: to prevent dG · dA mispairing in DNA replication. Proc. Natl. Acad. Sci. USA **86**:3949–3952.
5. **Allardet-Servent, A., G. Bourg, M. Ramuz, M. Pages, M. Bellis, and G. Roizes.** 1988. DNA polymorphism in strains of the genus *Brucella*. J. Bacteriol. **170**:4603–4607.
6. **Allewell, N.** 1988. Why does DNA bend? Trends Biochem. Sci. **13**:193–195.
7. **Amikam, D., S. Razin, and G. Glaser.** 1982. Ribosomal RNA genes in *Mycoplasma*. Nucleic Acids Res. **10**:4215–4222.
8. **Anagnostopoulos, C.** 1990. Genetic rearrangements in *Bacillus subtilis*, p. 361–371. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
9. **Anderson, R. P., and J. R. Roth.** 1977. Tandem genetic duplications in phage and fungi. Annu. Rev. Microbiol. **31**:473–505.
10. **Andersson, S. G. V., and C. G. Kurland.** 1990. Codon preferences in free-living microorganisms. Microbiol. Rev. **54**:198–210.
11. **Bachmann, B. J.** 1990. Linkage map of *Escherichia coli* K-12, edition 8. Microbiol. Rev. **54**:130–197.
12. **Balke, V. L., and J. D. Gralla.** 1987. Changes in linking number of supercoiled DNA accompany growth transitions in *Escherichia coli*. J. Bacteriol. **169**:4499–4506.
13. **Bancroft, I., C. P. Wolk, and E. V. Oren.** 1989. Physical and genetic maps of the genome of the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **171**: 5940–5948.
14. **Banfalvi, Z., E. Kondorosi, and A. Kondorosi.** 1985. *Rhizobium meliloti* carries two megaplasmids. Plasmid **13**:129–138.
15. **Barbour, A. G., and C. F. Garon.** 1987. Linear plasmids of the bacterium *Borrelia burgdorferi* have covalently closed ends. Science **237**:409–411.
16. **Barcak, G. J., and R. E. Wolf, Jr.** 1988. Comparative nucleotide sequence analysis of growth-rate-regulated *gnd* alleles from natural isolates of *Escherichia coli* and from *Salmonella typhimurium* LT-2. J. Bacteriol. **170**:372–379.
17. **Bautsch, W.** 1988. Rapid physical mapping of the *Mycoplasma mobile* genome by two-dimensional field inversion gel electrophoresis techniques. Nucleic Acids Res. **16**:11461–11467.
18. **Baylis, H. A., and M. J. Bibb.** 1988. Organization of the ribosomal RNA genes in *Streptomyces coelicolor* A3 (2). Mol. Gen. Genet. **211**:191–196.
19. **Beckwith, J. R., E. R. Signer, and W. Epstein.** 1966. Transposition of the lac region of *E. coli*. Cold Spring Harbor Symp. Quant. Biol. **31**:393–401.
20. **Belland, R. J., S. G. Morrison, P. van der Ley, and J. Swanson.** 1989. Expression and phase variation of gonococcal P.II genes in *Escherichia coli* involves ribosomal frameshifting and slipped-strand mispairing. Mol. Microbiol. **4**:777–786.
21. **Beltran, P., J. M. Musser, R. Helmuth, J. J. Farmer III, W. M. Frerichs, I. K. Wachsmuth, K. Ferris, A. Cravioto, and R. K. Selander.** 1988. Toward a population genetic analysis of *Salmonella*: genetic diversity and relationships among strains of serotypes *S. cholerae-suis*, *S. derby*, *S. dublin*, *S. enteritidis*, *S. heidelberg*, *S. infantis*, *S. newport*, and *S. typhimurium*. Proc. Natl. Acad. Sci. USA **85**:7753–7757.
22. **Bercovier, H., O. Kafri, and S. Sela.** 1986. *Mycobacteria*

possess a surprisingly small number of ribosomal RNA genes in relation to the size of their genome. Biochem. Biophys. Res. Commun. **136**:1136–1141.
23. **Bernardi, G., and G. Bernardi.** 1985. Codon usage and genome composition. J. Mol. Evol. **22**:363–365.
24. **Berry, J. O., and A. G. Atherly.** 1984. Induced plasmid-genome rearrangements in *Rhizobium japonicum*. J. Bacteriol. **157**: 218–224.
25. **Beuve, A., B. Boesten, M. Crasnier, A. Danchin, and F. O'Gara.** 1990. *Rhizobium meliliti* adenylate cyclase is related to eucaryotic adenylate and guanylate cyclases. J. Bacteriol. **172**:2614–2621.
26. **Bickle, T. A.** 1982. The ATP-dependent restriction endonucleases, p. 85–108. *In* S. M. Linn and R. J. Roberts (ed.), Nucleases. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **Birch, A., A. Häusler, and R. Hütter.** 1990. Genome rearrangement and genetic instability in *Streptomyces* spp. J. Bacteriol. **172**:4138–4142.
28. **Birch, A., A. Häusler, M. Vogtli, W. Krek, and R. Hütter.** 1989. Extremely large chromosomal deletions are intimately involved in genetic instability and genomic rearrangements in *Streptomyces glaucescens*. Mol. Gen. Genet. **217**:447–458.
29. **Birkenbihl, R. P., and W. Vielmetter.** 1989. Complete maps of IS1, IS2, IS3, IS4, IS5, IS30 and IS150 locations in *Escherichia coli* K12. Mol. Gen. Genet. **220**:147–153.
30. **Birkenbihl, R. P., and W. Vielmetter.** 1989. Cosmid-derived map of *E. coli* strain BHB2600 in comparison to the map of strain W3110. Nucleic Acids Res. **17**:5057–5069.
31. **Bohin, J. P., K. Ben Khalifa, N. Guillen, P. Schaeffer, and L. Hirschbein.** 1982. Phenotypic expression *in vivo* and transforming activity *in vitro*: two related functions of folded bacterial chromosomes. Mol. Gen. Genet. **185**:65–75.
32. **Borst, P., and D. R. Greaves.** 1987. Programmed gene rearrangements altering gene expression. Science **235**:658–667.
33. **Bouché, J. P.** 1982. Physical map of a 470×10^3 basepair region flanking the terminus of DNA replication in the *Escherichia coli* K12 genome. J. Mol. Biol. **154**:1–20.
34. **Bouma, J. E., and R. E. Lenski.** 1988. Evolution of a bacteria/plasmid association. Nature (London) **335**:351–352.
35. **Bourgeois, P. L., M. Mata, and P. Ritzenthaler.** 1989. Genome comparison of *Lactococcus* strains by pulsed-field gel electrophoresis. FEMS Microbiol. Lett. **59**:65–70.
36. **Brewer, B. J.** 1988. When polymerases collide: replication and the transcriptional organization of the *Escherichia coli* chromosome. Cell **53**:679–686.
37. **Brewer, B. J.** 1990. Replication and the transcriptional organization of the *Escherichia coli* chromosome, p. 61–83. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
38. **Brisson-Noel, A., M. Arthur, and P. Courvalin.** 1988. Evidence for natural gene transfer from gram-positive cocci to *Escherichia coli*. J. Bacteriol. **170**:1739–1745.
39. **Brown, D. P., S.-J. D. Chiang, J. S. Tuan, and L. Katz.** 1988. Site-specific integration in *Saccharopolyspora erythraea* and multisite integration in *Streptomyces lividans* of actinomycete plasmid pSE101. J. Bacteriol. **170**:2287–2295.
40. **Broyles, S., and D. E. Pettijohn.** 1986. Interaction of the *E. coli* HU protein with DNA: evidence for formation of nucleosome-like structures with altered DNA helical pitch. J. Mol. Biol. **187**:47–60.
41. **Burgin, A. B., K. Parodos, D. J. Lane, and N. R. Pace.** 1990. The excision of intervening sequences from *Salmonella* 23S ribosomal RNA. Cell **60**:405–414.
42. **Burkardt, B., D. Schillik, and A. Puhler.** 1987. Physical characterization of *Rhizobium meliloti* megaplasmids. Plasmid **17**: 13–25.
43. **Cairns, J.** 1963. The bacterial chromosome and its manner of replication as seen by autoradiography. J. Mol. Biol. **6**:208–213.
44. **Cairns, J., J. Overbaugh, and S. Miller.** 1988. The origin of mutants. Nature (London) **335**:142–145.
45. **Campbell, A.** 1981. Evolutionary significance of accessory DNA elements in bacteria. Annu. Rev. Microbiol. **35**:55–83.
46. **Canard, B., and S. T. Cole.** 1989. Genome organization of the anaerobic pathogen *Clostridium perfringens*. Proc. Natl. Acad. Sci. USA **86**:6676–6680.
47. **Carlson, T. A., and B. K. Chelm.** 1986. Apparent eukaryotic origin of glutamine synthetase II from the bacterium *Bradyrhizobium japonicum*. Nature (London) **322**:568–570.
48. **Casse, F., M.-C. Pascal, and M. Chippaux.** 1973. Comparison between the chromosomal maps of *Escherichia coli* and *Salmonella typhimurium*—length of the inverted segment in the trp region. Mol. Gen. Genet. **124**:253–257.
49. **Caugant, D. A., L. F. Mocca, C. E. Frasch, O. Froeholm, W. D. Zollinger, and R. K. Selander.** 1987. Genetic structure of *Neisseria meningitidis* populations in relation to serogroup, serotype, and outer membrane protein pattern. J. Bacteriol. **169**:2781–2792.
50. **Chang, N., and D. E. Taylor.** 1990. Use of pulsed-field agarose gel electrophoresis to size genomes of *Campylobacter* species and to construct a *Sal*I map of *Campylobacter jejuni* UA580. J. Bacteriol. **172**:5211–5217.
51. **Chen, H., I. M. Keseler, and L. J. Shimkets.** 1990. The genome size of *Myxococcus xanthus* determined by pulsed field gel electrophoresis. J. Bacteriol. **172**:4206–4213.
52. **Chen, S., and L. R. Finch.** 1989. Novel arrangement of rRNA genes in *Mycoplasma gallisepticum*: separation of the 16S gene of one set from the 23S and 5S genes. J. Bacteriol. **171**:2876–2878.
53. **Clark-Curtiss, J. E., and G. P. Walsh.** 1989. Conservation of genomic sequences among isolates of *Mycobacterium leprae*. J. Bacteriol. **171**:4844–4851.
54. **Cocks, B. G., L. E. Pyle, and L. R. Finch.** 1989. A physical map of the genome of *Ureoplasma urealyticum* 960^T with ribosomal RNA loci. Nucleic Acids Res. **17**:6713–6719.
55. **Colman, S. D., P.-C. Hu, and K. F. Bott.** 1990. Prevalence of novel repeat sequences in and around the *Pl* operon in the genome of *Mycoplasma pneumoniae*. Gene **87**:91–96.
56. **Condemine, G., and C. L. Smith.** 1990. Genetic mapping using large-DNA technology: alignment of *Sfi*I and *Arv*II sites with the *Not*I genomic restriction map of *Escherichia coli* K-12, p. 53–60. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
57. **Contreras, A., and J. Casadesus.** 1987. Tn*10* mutagenesis in *Azotobacter vinelandii*. Mol. Gen. Genet. **209**:276–282.
58. **Cooper, S., and C. E. Helmstetter.** 1968. Chromosome replication and the division cycle of *Escherichia coli* B/r. J. Mol. Biol. **31**:519–540.
59. **Cornelis, G., D. Ghosal, and H. Saedler.** 1978. Tn951: a new transposon carrying a lactose operon. Mol. Gen. Genet. **160**: 215–224.
60. **Craig, N. L., and N. Kleckner.** 1987. Transposition and site-specific recombination, p. 1054–1070. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
61. **Cullum, J., J. Altenbuchner, F. Flett, W. Piendl, and J. Platt.** 1986. DNA amplification and genetic instability in *Streptomyces*. Biotechnol. Genet. Eng. Rev. **4**:59–78.
62. **Daniels, D.** 1990. The complete AvrII restriction map of the *Escherichia coli* genome and comparisons of several laboratory strains. Nucleic Acids Res. **18**:2649–2651.
63. **Daniels, D. L.** 1990. Constructing encyclopedias of genomes, p. 43–51. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
64. **Davis, B. D.** 1989. Transcriptional bias: a non-Lamarckian mechanism for substrate-induced mutations. Proc. Natl. Acad. Sci. USA **86**:5005–5009.
65. **Delius, H., and A. Worcel.** 1974. Electron microscopic visualization of the folded chromosome of *Escherichia coli*. J. Mol. Biol. **82**:107–109.
66. **DeMassy, B., S. Béjar, J. Louarn, J. M. Louarn, and J. P. Bouché.** 1987. Inhibition of replication forks exiting the terminus region of the *Escherichia coli* chromosome occurs at two

loci separated by 5 min. Proc. Natl. Acad. Sci. USA **84**:1759–1763.
67. **Demuyter, P., P. Leblond, B. Decaris, and J.-M. Simonet.** 1988. Characterization of two families of spontaneously amplifiable units of DNA in *Streptomyces ambofaciens*. J. Gen. Microbiol. **134**:2001–2007.
68. **Dennis, P. P.** 1986. Molecular biology of archaebacteria. J. Bacteriol. **168**:471–478.
69. **Dixon, N., and A. Kornberg.** 1984. Protein HU in the enzymatic replication of the chromosomal origin of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **81**:424–428.
70. **Dorman, C. J., and C. F. Higgins.** 1987. Fimbrial phase variation in *Escherichia coli*: dependence on integration host factor and homologies with other site-specific recombinases. J. Bacteriol. **169**:3840–3843.
71. **Downs, D. M., and J. R. Roth.** 1987. A novel P22 prophage in *Salmonella typhimurium*. Genetics **117**:367–380.
72. **Drlica, K.** 1984. Biology of bacterial deoxyribonucleic acid topoisomerases. Microbiol. Rev. **48**:273–289.
73. **Drlica, K.** 1987. The nucleoid, p. 91–103. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (eds.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
74. **Drlica, K., G. J. Pruss, R. M. Burger, R. J. Franco, L.-S. Hsieh, and B. A. Berger.** 1990. Roles of DNA topoisomerases in bacterial chromosome structure and function, p. 195–204. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
75. **Drlica, K., and J. Rouvière-Yaniv.** 1987. Histonelike proteins of bacteria. Microbiol. Rev. **51**:301–319.
76. **DuBose, R. F., D. E. Dykhuizen, and D. L. Hartl.** 1988. Genetic exchange among natural isolates of bacteria: recombination within the *phoA* gene of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **85**:7036–7040.
77. **Dürrenberger, M., M.-A. Bjornsti, T. Uetz, J. A. Hobot, and E. Kellenberger.** 1988. Intracellular location of the histonelike protein HU in *Escherichia coli*. J. Bacteriol. **170**:4757–4768.
78. **Dykhuizen, D. E., and L. Green.** 1986. DNA sequence variation, DNA phylogeny, and recombination in *E. coli*. Genetics **113**:71.
79. **Dykhuizen, D. E., and D. L. Hartl.** 1983. Selection in chemostats. Microbiol. Rev. **47**:150–168.
80. **Eisenstein, B. I., D. S. Sweet, V. Vaughn, and D. I. Friedman.** 1987. Integration host factor is required for the DNA inversion that controls phase variation in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **84**:6506–6510.
81. **Ely, B., T. W. Ely, C. J. Gerardot, and A. Dingwall.** 1990. Circularity of the *Caulobacter crescentus* chromosome determined by pulsed-field gel electrophoresis. J. Bacteriol. **172**: 1262–1266.
82. **Faulds, D., N. Dower, M. M. Stahl, and F. W. Stahl.** 1979. Orientation dependent recombination hotspot activity of bacteriophage λ. J. Mol. Biol. **131**:681–695.
83. **Ferdows, M. S., and A. G. Barbour.** 1989. Megabase-sized linear DNA in the bacterium *Borrelia burgdorferi*, the Lyme disease agent. Proc. Natl. Acad. Sci. USA **86**:5969–5973.
84. **Feulner, G., J. A. Gray, J. A. Kirschman, A. F. Lehner, A. B. Sadosky, D. A. Vlazny, J. Zhang, S. Zhao, and C. W. Hill.** 1990. Structure of the *rhsA* locus from *Escherichia coli* K-12 and comparison of *rhsA* with other members of the *rhs* multigene family. J. Bacteriol. **172**:446–456.
85. **Figueroa, N., and L. Bossi.** 1988. Transcription induces gyration of the DNA template in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **85**:9416–9420.
86. **Fishman, S. E., and C. L. Hershberger.** 1983. Amplified DNA in *Streptomyces fradiae*. J. Bacteriol. **155**:459–466.
87. **Fishman, S. E., P. R. Rosteck, Jr., and C. L. Hershberger.** 1985. A 2.2-kilobase repeated DNA segment is associated with DNA amplification in *Streptomyces fradiae*. J. Bacteriol. **161**: 199–206.
88. **Flashner, Y., and J. Gralla.** 1988. DNA dynamic flexibility and protein recognition: differential stimulation by bacterial histone-like protein HU. Cell **54**:713–721.
89. **Flores, M., V. González, S. Brom, E. Martínez, D. Piñero, D. Romero, G. Dávila, and R. Palacios.** 1987. Reiterated DNA sequences in *Rhizobium* and *Agrobacterium* spp. J. Bacteriol. **169**:5782–5788.
90. **Flores, M., V. González, M. A. Pardo, A. Leija, E. Martínez, D. Romero, D. Piñero, G. Dávila, and R. Palacios.** 1988. Genomic instability in *Rhizobium phaseoli*. J. Bacteriol. **170**:1191–1196.
91. **Fournier, M. J., and H. Ozeki.** 1985. Structure and organization of the transfer ribonucleic acid genes of *Escherichia coli* K-12. Microbiol. Rev. **49**:379–397.
92. **François, V., H. Louarn, J. Pattee, J.-E. Rebollo, and J.-M. Louarn.** 1990. Constraints in chromosomal inversions in *Escherichia coli* are not explained by replication pausing at inverted terminator-like sequences. Mol. Microbiol. **4**:537–542.
93. **François, V., J. Louarn, J. E. Rebollo, and J. M. Louarn.** 1990. Replication termination, nondivisible zones, and structure of the *Escherichia coli* chromosome, p. 351–359. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
94. **Frantz, B., and A. M. Chakrabarty.** 1986. Degradative plasmids in *Pseudomonas*, p. 295–323. *In* J. R. Sokatch (ed.), The bacteria, vol. X. The biology of the *Pseudomonas*. Academic Press, Inc., New York.
95. **Frutos, R., M. Pages, M. Bellis, G. Roizes, and M. Bergoin.** 1989. Pulsed-field gel electrophoresis determination of the genome size of obligate intracellular bacteria belonging to the genera *Chlamydia*, *Rickettsiella*, and *Porochlamydia*. J. Bacteriol. **171**:4511–4513.
96. **Fuerst, P. A., K. P. Poetter, C. Pretzmann, and P. S. Perlman.** 1990. Molecular genetics of populations of intracellular bacteria: the spotted fever group Rickettsiae. Ann. N.Y. Acad. Sci. **20**:785–896.
97. **Fujita, M. Q., H. Yoshikawa, and N. Ogasawara.** 1989. Structure of the *dnaA* region of *Pseudomonas putida*: conservation among three bacteria, *Bacillus subtilis*, *Escherichia coli*, and *P. putida*. Mol. Gen. Genet. **215**:381–387.
98. **Fukunaga, M., and I. Mifuchi.** 1989. Unique organization of *Leptospira interrogans* rRNA genes. J. Bacteriol. **171**:5763–5767.
99. **Fulks, K. A., C. F. Marrs, S. P. Stevens, and M. R. Green.** 1990. Sequence analysis of the inversion region containing the pilin genes of *Moraxella bovis*. J. Bacteriol. **172**:310–316.
100. **Funnell, B., T. Baker, and A. Kornberg.** 1987. In vitro assembly of a prepriming complex at the origin of the *Escherichia coli* chromosome. J. Biol. Chem. **262**:10327–10334.
101. **Furukawa, K., and A. M. Chakrabarty.** 1982. Involvement of plasmids in total degradation of chlorinated biphenyls. Appl. Environ. Microbiol. **44**:619–626.
102. **Furukawa, K., N. Hayase, K. Taira, and N. Tomizuka.** 1989. Molecular relationship of chromosomal genes encoding biphenyl/polychlorinated biphenyl catabolism: some soil bacteria possess a highly conserved *bph* operon. J. Bacteriol. **171**:5467–5472.
103. **Gargallo-Viola, D.** 1989. Enzyme polymorphism, prodigiosin production, and plasmid fingerprints in clinical and naturally occurring isolates of *Serratia marcescens*. J. Clin. Microbiol. **27**:860–868.
104. **Gerlach, J. H., J. A. Endicott, P. F. Juranka, G. Henderson, F. Sarangi, K. L. Deuchars, and V. Ling.** 1986. Homology between P-glycoprotein and a bacterial haemolysin transport protein suggests a model for multidrug resistance. Nature (London) **324**:485–489.
105. **Gilson, E., D. Perrin, and M. Hofnung.** 1990. DNA polymerase I and a protein complex bind specifically to *E. coli* palindromic unit highly repetitive DNA: implications for bacterial chromosome organization. Nucleic Acids Res. **18**:3941–3952.
106. **Gilson, E., D. Perrin, W. Saurin, and M. Hofnung.** 1987. Species specificity of bacterial palindromic units. J. Mol. Evol. **25**:371–373.
107. **Giroux, S., J. Beaudet, and R. Cedergren.** 1988. Highly repetitive $tRNA^{Pro}$-$tRNA^{His}$ gene cluster from *Photobacterium phosphoreum*. J. Bacteriol. **170**:5601–5606.

108. **Glasgow, A. C., K. T. Hughes, M. I. Simon, D. E. Berg, and M. M. Howe.** 1989. Bacterial DNA inversion systems, p. 637–660. *In* D. E. Berg, and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
109. **Godwin, D., and J. H. Slater.** 1979. The influence of the growth environment on the stability of a drug resistance plasmid in *Escherichia coli* K12. J. Gen. Microbiol. **111**:201–210.
110. **Golden, J. W., C. D. Carrasco, M. E. Mulligan, G. J. Schneider, and R. Haselkorn.** 1988. Deletion of a 55-kilobase-pair DNA element from the chromosome during heterocyst differentiation of *Anabaena* sp. strain PCC 7120. J. Bacteriol. **170**:5034–5041.
111. **Golden, J. W., S. J. Robinson, and R. Haselkorn.** 1985. Rearrangement of nitrogen fixation genes during heterocyst differentiation in the cyanobacterium *Anabaena*. Nature (London) **314**:419–423.
112. **Goursot, R., A. Goze, B. Niaudet, and S. D. Ehrlich.** 1982. Plasmids from *Staphylococcus aureus* replicate in yeast *Saccharomyces cerevisiae*. Nature (London) **298**:488–490.
113. **Gouy, M., and C. Gautier.** 1982. Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res. **10**: 7055–7074.
114. **Gralla, J. D.** 1989. Bacterial gene regulation from distant DNA sites. Cell **57**:193–195.
115. **Griffith, J. D.** 1976. Visualization of procaryotic DNA in a regulatory condensed chromatin-like fiber. Proc. Natl. Acad. Sci. USA **73**:563–567.
116. **Guerry, P., S. M. Logan, S. Thornton, and T. J. Trust.** 1990. Genomic organization and expression of *Campylobacter* flagellin genes. J. Bacteriol. **172**:1853–1860.
117. **Guerry, P., S. M. Logan, and T. J. Trust.** 1988. Genomic rearrangements associated with antigenic variation in *Campylobacter coli*. J. Bacteriol. **170**:316–319.
118. **Guillén, N., M. Amar, and L. Hirschbein.** 1985. Stabilized non-complementing diploids (Ncd) from fused protoplast products of *B. subtilis*. EMBO J. **4**:1333–1338.
119. **Guillén, N., M. H. Gabor, R. D. Hotchkiss, and L. Hirschbein.** 1982. Isolation and characterization of the nucleoid of non-complementing diploids from protoplast fusion in *Bacillus subtilis*. Mol. Gen. Genet. **185**:69–74.
120. **Guillén, N., S. Sanchez-Rivas, and L. Hirschbein.** 1983. Absence of functional RNA encoded by a silent chromosome in non-complementing diploids obtained from protoplast fusion in *Bacillus subtilis*. Mol. Gen. Genet. **191**:81–85.
121. **Guiso, N., and A. Ullmann.** 1976. Expression and regulation of lactose genes carried by plasmids. J. Bacteriol. **127**:691–697.
122. **Gutman, G. A., and G. W. Hatfield.** 1989. Nonrandom utilization of codon pairs in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **86**:3699–3703.
123. **Haas, R., and T. F. Meyer.** 1986. The repertoire of silent pilus genes in *Neisseria gonorrhoeae*: evidence for gene conversion. Cell **44**:107–115.
124. **Hall, B. G.** 1988. Adaptive evolution that requires multiple spontaneous mutations. I. Mutations involving an insertion sequence. Genetics **120**:887–897.

124a. **Hall, B. G.** 1990. Spontaneous point mutations that occur more often when advantageous than when neutral. Genetics **126**:5–16.

125. **Hanvey, J. C., M. Shimizu, and R. D. Wells.** 1988. Intramolecular DNA triplexes in supercoiled plasmids. Proc. Natl. Acad. Sci. USA **85**:6292–6296.
126. **Harayama, S., and K. N. Timmis.** 1989. Catabolism of aromatic hydrocarbons by *Pseudomonas*, p. 151–174. *In* D. A. Hopwood, and K. E. Chater (ed.), Genetics of bacterial diversity. Academic Press, Inc. (London), Ltd., London.
127. **Harshman, L., and M. Riley.** 1980. Conservation and variation of nucleotide sequences in *Escherichia coli* strains isolated from nature. J. Bacteriol. **144**:560–568.
128. **Hartl, D. L., and D. E. Dykhuizen.** 1984. The population genetics of *Escherichia coli*. Annu. Rev. Genet. **18**:31–68.
129. **Hartmann, R. K., and V. A. Erdmann.** 1989. *Thermus thermophilus* 16S rRNA is transcribed from an isolated transcription unit. J. Bacteriol. **171**:2933–2941.
130. **Hartmann, R. K., H. Y. Toschka, N. Ulbrich, and V. A. Erdmann.** 1986. Genomic organization of rDNA in *Pseudomonas aeruginosa*. FEBS Lett. **195**:187–193.
131. **Hartmann, R. K., N. Ulbrich, and V. A. Erdmann.** 1987. An unusual rRNA operon constellation: in *Thermus thermophilus* HB8 the 23S/5S rRNA operon is a separate entity from the 16S rRNA operon. Biochimie **69**:1097–1104.
132. **Harvey, S., and C. W. Hill.** 1990. Exchange of spacer regions between rRNA operons in *Escherichia coli*. Genetics **125**:683–690.
133. **Harvey, S., C. W. Hill, C. Squires, and C. C. Squires.** 1988. Loss of the spacer loop sequence from the *rrnB* operon in the *Escherichia coli* K-12 subline that bears the *relA1* mutation. J. Bacteriol. **170**:1235–1238.
134. **Häusler, A., A. Birch, W. Krek, J. Piret, and R. Hütter.** 1989. Heterogeneous genomic amplification in *Streptomyces glaucescens*: structure, location and junction sequence analysis. Mol. Gen. Genet. **217**:437–446.
135. **Hediger, M. A., E. Turk, and E. M. Wright.** 1989. Homology of the human intestinal Na(+)/glucose and *Escherichia coli* Na(+)/proline cotransporters. Proc. Natl. Acad. Sci. USA **86**:5748–5752.
136. **Heinemann, J. A., and G. V. Sprague, Jr.** 1989. Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. Nature (London) **340**:205–209.
137. **Helling, R. B., T. Kinney, and J. Adams.** 1981. The maintenance of plasmid-containing organisms in populations of *Escherichia coli*. J. Gen. Microbiol. **123**:129–141.
138. **Helmstetter, C. E., O. Pierucci, M. Weinberger, M. Holmes, and M.-S. Tang.** 1979. Control of cell division in *Escherichia coli*, p. 517–597. *In* J. R. Sokatch and L. N. Ornston (ed.), The bacteria, vol. VII. Mechanisms of adaptation. Academic Press, Inc., New York.
139. **Herdman, M.** 1985. The evolution of bacterial genomes, p. 37–68. *In* T. Cavalier-Smith (ed.). The evolution of genome size. John Wiley & Sons, Inc., New York.
140. **Herdman, M., M. Janvier, R. Rippka, and R. Y. Stanier.** 1979. Genome size of cyanobacteria. J. Gen. Microbiol. **111**:73–85.
141. **Higgins, C. F., G. F.-L. Ames, W. M. Barnes, J.-M. Clement, and M. Hufnung.** 1982. A novel intercistronic regulatory element in procaryotic operons. Nature (London) **298**:760–762.
142. **Higgins, C. F., C. J. Dorman, and N. Ni Bhriain.** 1990. Environmental influences on DNA supercoiling: a novel mechanism for the regulation of gene expression, p. 421–432. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
143. **Higgins, C. F., C. J. Dorman, D. A. Stirling, L. Waddell, I. R. Booth, G. May, and E. Bremer.** 1988. A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *S. typhimurium* and *E. coli*. Cell **52**:569–584.
144. **Hill, C. W., and J. A. Gray.** 1988. Effects of chromosomal inversion on cell fitness in *Escherichia coli* K-12. Genetics **119**:771–778.
145. **Hill, C. W., and B. W. Harnish.** 1981. Inversions between ribosomal RNA genes of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **78**:7069–7072.
146. **Hill, C. W., S. Harvey, and J. A. Gray.** 1990. Recombination between rRNA genes in *Escherichia coli* and *Salmonella typhimurium*, p. 335–340. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
147. **Hollingshead, S. K., F. A. Fischetti, and J. R. Scott.** 1987. Size variation in group A streptococcal M protein is generated by homologous recombination between intragenic repeats. Mol. Gen. Genet. **207**:196–203.
148. **Holloway, B. W., S. Dharmsthiti, C. Johnson, A. Kearney, V. Krishnapillai, A. F. Morgan, E. Ratnaningsih, R. Saffery, M. Sinclair, D. Strom, and C. Zhang.** 1989. Chromosome organization in *Pseudomonas aeruginosa* and *Pseudomonas putida*, p. 9–13. *In Pseudomonas*: Biotransformations, pathogenesis, and evolving biotechnology. American Society for Microbiology, Washington, D.C.
149. **Holloway, B. W., S. Dharmsthiti, V. Krishnapillai, A. Morgan,**

V. Obeyesekere, E. Ratnaningsih, M. Sinclair, D. Strom, and C. Zhang. 1990. Patterns of gene linkages in *Pseudomonas* species, p. 97–105. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.

150. **Holloway, B. W., and A. F. Morgan.** 1986. Genome organization in *Pseudomonas*. Annu. Rev. Microbiol. **40**:79–105.
151. **Hooykaas, P. J. J.** 1989. Tumorigenicity of *Agrobacterium* on plants, p. 373–391. *In* D. A. Hopwood and K. E. Chater (ed.), Genetics of bacterial diversity. Academic Press, Inc., San Diego, Calif.
152. **Hopwood, D. A.** 1967. A possible circular symmetry of the linkage map of *Streptomyces coelicolor*. J. Cell. Physiol. **70**(Suppl. 1):7–10.
153. **Hopwood, D. A., G. Hintermann, T. Kieser, and H. M. Wright.** 1984. Integrated DNA sequences in three streptomycetes form related autonomous plasmids after transfer to *Streptomyces lividans*. Plasmid **11**:1–16.
154. **Hopwood, D. A., and T. Kieser.** 1990. The *Streptomyces* genome, p. 147–162. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
155. **Hornemann, U., C. J. Otto, G. G. Hoffman, and A. C. Bertinuson.** 1987. Spectinomycin resistance and associated DNA amplification in *Streptomyces achromogenes* subsp. *rubradiris*. J. Bacteriol. **169**:2360–2366.
156. **Hornemann, U., C. J. Otto, and X. Y. Zhang.** 1989. DNA amplification in *Streptomyces achromogenes* subsp. *rhubradiris* is accompanied by a deletion, and the amplified sequences are conditionally stable and can be eliminated by two pathways. J. Bacteriol. **171**:5817–5822.
157. **Hotchkiss, R. D., and M. H. Gabor.** 1980. Biparental products of bacterial protoplast fusion showing unequal parental chromosome expression. Proc. Natl. Acad. Sci. USA **77**:3553–3557.
158. **Hu, M., and R. C. Deonier.** 1981. Mapping of IS1 elements flanking the *arg*F region of the *Escherichia coli* K-12 chromosome. Mol. Gen. Genet. **181**:222–229.
159. **Hugouvieux-Cotte-Pattat, N., S. Reverchon, and J. Robert-Baudouy.** 1989. Expanded linkage map of *Erwinia chrysanthemi* strain 3937. Mol. Microbiol. **3**:573–581.
160. **Huisman, O., M. Faelen, D. Girard, A. Jaffé, A. Toussaint, and J. Rouvière-Yaniv.** 1989. Multiple defects in *Escherichia coli* mutants lacking HU protein. J. Bacteriol. **171**:3704–3712.
161. **Hütter, R., and T. Eckhardt.** 1988. Genetic manipulation, p. 89–184. *In* M. Goodfellow, S. T. Williams, and M. Mordarski (ed.), Actinomycetes in biotechnology. Academic Press, Inc. (London), Ltd., London.
162. **Ichige, A., S. Matsutani, K. Oishi, and S. Mizushima.** 1989. Establishment of gene transfer systems for and construction of the genetic map of a marine *Vibrio* strain. J. Bacteriol. **171**: 1825–1834.
163. **Ikemura, T.** 1981. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. J. Mol. Biol. **151**:389–409.
164. **Imamoto, F., and Y. Kano.** 1990. Physiological characterization of deletion mutants of the *hupA* and *hupB* genes in *Escherichia coli*, p. 259–266. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
165. **Innes, R. W., M. A. Hirose, and P. L. Kuempel.** 1988. Induction of nitrogen-fixing nodules on clover requires only 32 kilobase pairs of DNA from the *Rhizobium trifolii* symbiosis plasmid. J. Bacteriol. **170**:3793–3802.
166. **Ishiguro, N., K. Hirose, M. Asagi, and G. Sato.** 1981. Incompatibility of citrate utilization plasmids isolated from *Escherichia coli*. J. Gen. Microbiol. **123**:193–196.
167. **Ishiguro, N., and G. Sato.** 1988. Nucleotide sequence of insertion sequence IS*3411*, which flanks the citrate utilization determinant of transposon Tn*3411*. J. Bacteriol. **170**:1902–1906.
168. **Ives, C., and K. Bott.** 1990. Characterization of chromosomal DNA amplifications with associated tetracycline resistance in *Bacillus subtilis*. J. Bacteriol. **172**:4936–4944.
169. **Jaoua, S., J. F. Guespin-Michel, and A. Breton.** 1987. Mode of insertion of broad-host-range plasmid RP4 and its derivatives into the chromosome of *Myxococcus xanthus*. Plasmid **18**:111–119.
170. **Jarvis, E. D., R. L. Widom, G. LaFauci, Y. Setoguchi, I. R. Richter, and R. Rudner.** 1988. Chromosomal organization of rRNA operons in *Bacillus subtilis*. Genetics **120**:625–635.
171. **Jaworski, A., W.-T. Hsieh, J. A. Blaho, J. E. Larson, and R. D. Wells.** 1987. Left-handed DNA *in vivo*. Science **238**:773–777.
172. **Johnson, R. C., M. F. Bruist, and M. I. Simon.** 1986. Host protein requirements for *in vitro* site-specific DNA inversion. Cell **46**:531–539.
173. **Kaluza, K., M. Hahn, and H. Hennecke.** 1985. Repeated sequences similar to insertion elements clustered around the *nif* region of the *Rhizobium japonicum* genome. J. Bacteriol. **162**:535–542.
174. **Kano, Y., and F. Imamoto.** 1990. Requirement of integration host factor (IHF) for growth of *Escherichia coli* deficient in HU protein. Gene **89**:133–137.
175. **Kano, Y., K. Osato, M. Wada, and F. Imamoto.** 1987. Cloning and sequencing of the *HU-2* gene of *Escherichia coli*. Mol. Gen. Genet. **209**:408–410.
176. **Kano, Y., M. Wada, T. Nagase, and F. Imamoto.** 1986. Genetic characterization of the gene *hupB* encoding the HU-1 protein of *Escherichia coli*. Gene **45**:37–44
177. **Kauc, L., and S. H. Goodgal.** 1989. The size and a physical map of the chromosome of *Haemophilus parainfluenzae*. Gene **83**:377–380.
178. **Kauc, L., M. Mitchell, and S. H. Goodgal.** 1989. Size and physical map of the chromosome of *Haemophilus influenzae*. J. Bacteriol. **171**:2474–2479.
179. **Kavenoff, R., and B. Bowen.** 1976. Electron microscopy of membrane-free folded chromosomes from *E. coli*. Chromosoma **59**:89–101.
180. **Kellenberger, E.** 1990. Intracellular organization of the bacterial genome, p. 173–186. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
181. **Kennedy, C., and A. Toukdarian.** 1987. Genetics of azotobacters: applications to nitrogen fixation and related aspects of nitrogen metabolism. Annu. Rev. Microbiol. **41**:227–258.
182. **Kinashi, H., M. Shimaji, and A. Sakai.** 1987. Giant linear plasmids in *Streptomyces* which code for antibiotic biosynthesis genes. Nature (London) **238**:454–456.
183. **Kitten, T., and A. Barbour.** 1990. Juxtaposition of expressed variable antigen genes with a conserved telomere in the bacterium *Borrelia hermsii*. Proc. Natl. Acad. Sci. USA **87**:6077–6081.
184. **Knott, V., D. J. Blake, and G. G. Brownlee.** 1989. Completion of the detailed restriction map of the *E. coli* genome by the isolation of overlapping cosmid clones. Nucleic Acids Res. **17**:5901–5912.
185. **Kohara, Y.** 1990. Correlation between the physical and genetic maps of the *Escherichia coli* K-12 chromosome, p. 29–42. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
186. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50**:495–508.
187. **Kohno, K., M. Wada, Y. Kano, and F. Imamoto.** 1990. Promoters and autogenous control of the *Escherichia coli hupA* and *hupB* genes. J. Mol. Biol. **213**:27–36.
188. **Kolstø, A.-B., A. Grønstad, and H. Oppegaard.** 1990. Physical map of the *Bacillus cereus* chromosome. J. Bacteriol. **172**: 3821–3825.
189. **Komano, T., A. Kubo, and T. Nisioka.** 1987. Shufflon: multi-inversion of four contiguous DNA segments of plasmid R64 creates seven different open reading frames. Nucleic Acids Res. **15**:1165–1172.

190. **Kondorosi, A., E. Vincze, A. W. B. Johnston, and J. E. Beringer.** 1980. A comparison of three *Rhizobium* linkage maps. Mol. Gen. Genet. **178**:403–408.
191. **Konrad, E. B.** 1977. Method for the isolation of *Escherichia coli* mutants with enhanced recombination between chromosomal duplications. J. Bacteriol. **130**:167–172.
192. **Kosmine, Y., T. Adachi, H. Inokuchi, and H. Ozeki.** 1990. Genomic organization and physical mapping of the transfer RNA genes in *Escherichia coli* K12. J. Mol. Biol. **212**:579–598.
193. **Krämer, H., M. Amouyal, A. Nordheim, and B. Müller-Hill.** 1988. DNA supercoiling changes in the spacing requirement of two *lac* operators for DNA loop formation with *lac* repressor. EMBO J. **7**:547–556.
194. **Krause, D. C., and C. B. Mawn.** 1990. Physical analysis and mapping of the *Mycoplasma pneumoniae* chromosome. J. Bacteriol. **172**:4790–4797.
195. **Kröger, M., R. Wahl, and P. Rice.** 1990. Compilation of sequences of *Escherichia coli* (update 1990). Nucleic Acids Res. **18**:2549–2587.
196. **Kuempel, P. L., A. J. Pelletier, and T. M. Hill.** 1989. Tus and the terminators: the arrest of replication in prokaryotes. Cell **59**:581–583.
197. **Kunkel, B., R. Losick, and P. Stragier.** 1990. The *Bacillus subtilis* gene for the developmental transcription factor σ^K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. Genes Dev. **4**:525–535.
198. **Kurland, C. G.** 1987. Strategies for efficiency and accuracy in gene expression. Trends Biochem. Sci. **12**:126–128.
199. **Lamfrom, H., A. Sarabhai, and J. Abelson.** 1978. Cloning *Beneckea* genes in *Escherichia coli*. J. Bacteriol. **133**:354–363.
200. **Laundon, C. H., and J. D. Griffith.** 1988. Curved helix segments can uniquely orient the topology of supertwisted DNA. Cell **52**:545–549.
201. **Leblond, P., P. Demuyter, L. Moutier, M. Laakel, B. Decaris, and J.-M. Simonet.** 1989. Hypervariability, a new phenomenon of genetic instability, related to DNA amplification in *Streptomyces ambofaciens*. J. Bacteriol. **171**:419–423.
202. **Lee, J. J., H. O. Smith, and R. J. Redfield.** 1989. Organization of the *Haemophilus influenzae* Rd genome. J. Bacteriol. **171**: 3016–3024.
203. **Lee, S. C., C. A. Omer, M. A. Brasch, and S. N. Cohen.** 1988. Analysis of recombination occurring at SLP1 *att* sites. J. Bacteriol. **170**:5806–5813.
204. **Lehner, K. A., S. Harvey, and C. W. Hill.** 1984. Mapping spacer identification of rRNA operons of *Salmonella typhimurium*. J. Bacteriol. **160**:682–686.
205. **Lejeune, P., and A. Danchin.** 1990. Mutations in the *bglY* gene increase the frequency of spontaneous deletions in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA **87**:360–363.
206. **Levin, B. R.** 1981. Periodic selection, infectious gene exchange and the genetic structure of *E. coli* populations. Genetics **99**:1–23.
207. **Levin, B. R., and F. M. Stewart.** 1979. The population biology of bacterial plasmids: *a priori* conditions for the existence of mobilizable nonconjugative factors. Genetics **94**:425–443.
208. **Levinson, G., and G. A. Gutman.** 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol. Biol. Evol. **4**:203–221.
209. **Liebart, J. C., L. Paolozzi, M. G. Camera, A. M. Pedrini, and P. Ghelardini.** 1989. The expression of the DNA ligase gene of *Escherichia coli* is stimulated by relaxation of chromosomal supercoiling. Mol. Microbiol. **3**:269–273.
210. **Liesack, W., and E. Stackebrandt.** 1989. Evidence for unlinked *rrn* operons in the planctomycete *Pirellula marina*. J. Bacteriol. **171**:5025–5030.
211. **Lilley, D. M.** 1988. DNA opens up—supercoiling and heavy breathing. Trends Genet. **4**:111–114.
212. **Lilley, D. M., and B. Kemper.** 1984. Cruciform-resolvase interactions in supercoiled DNA. Cell **36**:413–422.
213. **Lin, R.-J., M. Capage, and C. W. Hill.** 1989. A repetitive DNA sequence, *rhs*, responsible for duplications within the *Escherichia coli* K-12 chromosome. J. Mol. Biol. **177**:1–18.
214. **Lindahl, L., and J. M. Zengel.** 1986. Ribosomal genes in *Escherichia coli*. Annu. Rev. Genet. **20**:297–326.
215. **Link, C. D., and A. M. Reiner.** 1983. Genotypic exclusion: a novel relationship between the ribitol-arabitol and galactitol genes of *E. coli*. Mol. Gen. Genet. **189**:337–339.
216. **Long, C. M., M.-J. Virolle, S.-Y. Chang, S. Chang, and M. J. Bibb.** 1987. α-Amylase gene of *Streptomyces limosus*: nucleotide sequence, expression motifs, and amino acid sequence homology to mammalian and invertebrate alpha-amylases. J. Bacteriol. **169**:5745–5754.
217. **Lopez, F., N. Guillén, and L. Hirschbein.** 1986. Further characterization of the inactive chromosome from *Bacillus subtilis* stabilized non-complementing diploids, p. 73–86. *In* A. T. Ganesan and J. A. Hoch (ed.), Bacillus molecular genetics and biotechnology applications. Academic Press, Inc., New York.
218. **Louarn, J. M., J. P. Bouché, F. Legendre, J. Louarn, and J. Patte.** 1985. Characterization and properties of very large inversions of the *E. coli* chromosome along the origin-to-terminus axis. Mol. Gen. Genet. **201**:467–476.
219. **Loughney, K., E. Lund, and J. E. Dahlberg.** 1982. tRNA genes are found between the 16S and 23S rRNA genes in *Bacillus subtilis*. Nucleic Acids Res. **10**:1607–1624.
220. **Luria, S. E., and M. Delbruck.** 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28**:491–511.
221. **Lyderson, B. K., and D. E. Pettijohn.** 1977. Interactions stabilizing DNA tertiary structure in the *Escherichia coli* chromosome investigated with ionizing radiation. Chromosoma **62**:199–215.
222. **Lyon, M. F.** 1988. The William Allan memorial award address. X-chromosome inactivation and the location and expression of X-linked genes. Am. J. Hum. Genet. **42**:8–16.
223. **Madon, J., P. Moretti, and R. Hütter.** 1987. Site-specific integration and excision of pMEA100 in *Nocardia mediterranei*. Mol. Gen. Genet. **209**:257–264.
224. **Mahan, M. J., and J. R. Roth.** 1988. Reciprocality of recombination events that rearrange the chromosome. Genetics **120**:23–35.
225. **Mahan, M. J., A. M. Segall, and J. R. Roth.** 1990. Recombination events that rearrange the chromosome: barriers to inversion, p. 341–349. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
226. **Maiden, M. C. J., E. O. Davis, S. A. Baldwin, D. C. M. Moore, and P. J. F. Henderson.** 1987. Mammalian and bacterial sugar transport proteins are homologous. Nature (London) **325**:641–643.
227. **Malone, R. E., D. K. Chattoraj, D. H. Faulds, M. M. Stahl, and F. W. Stahl.** 1978. Hotspots for generalized recombination in the *Escherichia coli* chromosome. J. Mol. Biol. **121**:473–491.
228. **Marrs, C. F., W. W. Ruehl, G. K. Schoolnik, and S. Falkow.** 1988. Pilin gene phase variation of *Moraxella bovis* is caused by an inversion of the pilin genes. J. Bacteriol. **170**:3032–3039.
229. **Masters, M., P. D. Moir, R. Spiegelberg, J. H. Pringle, and C. W. Vermeulen.** 1985. Is the chromosome of *E. coli* differentiated along its length with respect to gene density or accessibility to transcription? p. 335–343. *In* M. Schaechter, F. Neidhardt, J. Ingraham, and N. Kjelgaard (ed.), The molecular biology of bacterial growth. Jones and Bartlett, Boston.
230. **Mazel, D., J. Houmard, A. M. Castets, and N. Tandeu De Marsac.** 1990. Highly repetitive DNA sequences in cyanobacterial genomes. J. Bacteriol. **172**:2755–2761.
231. **Mazodier, P., R. Petter, and C. Thompson.** 1989. Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species. J. Bacteriol. **171**:3583–3585.
232. **McMacken, R., L. Silver, and C. Georgopoulos.** 1987. DNA replication, p. 564–612. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
233. **Medique, C., J. P. Bouché, A. Henant, and A. Danchin.** 1990. Mapping of sequenced genes (700 bp) in the restriction map of the *E. coli* chromosome. Mol. Microbiol. **4**:169–187.
234. **Menzel, R., and M. Gellert.** 1983. Regulation of the genes for *E.*

coli DNA gyrase: homeostatic control of DNA supercoiling. Cell **34**:105–113.

235. **Merino, E., and F. Bolivar.** 1989. The ribonucleoside diphosphate reductase gene (nrdA) of *Escherichia coli* carries a repetitive extragenic palindromic (REP) sequence in its 3′ structural terminus. Mol. Microbiol. **3**:839–841.
236. **Mevarech, M., S. Hirsch-Twizer, S. Goldman, E. Yakobson, H. Eienberg, and P. P. Dennis.** 1989. Isolation and characterization of the rRNA gene clusters of *Halobacterium marismortui*. J. Bacteriol. **171**:3479–3485.
237. **Milkman, R.** 1973. Electrophoretic variation in *E. coli* from natural sources. Science **182**:1024–1026.
238. **Milkman, R.** 1975. Allozyme variation in *E. coli* of diverse natural origin, p. 273–285. *In* C. L. Markert (ed.), Isozymes, vol. 4. Academic Press, Inc., New York.

238a. **Milkman, R., and M. M. Bridges.** 1990. Molecular evolution of the *Escherichia coli* chromosome. III. Clonal frames. Genetics **126**:505–517.

239. **Milkman, R., and I. P. Crawford.** 1983. Clustered third-base substitutions among wild strains of *E. coli*. Science **221**:378–380.
240. **Milkman, R., and A. Stoltzfus.** 1988. Molecular evolution of the *Escherichia coli* chromosome. II. Clonal segments. Genetics **120**:359–366.
241. **Modrich, P.** 1989. Methyl-directed DNA mismatch correction. J. Biol. Chem. **264**:6597–6600.
242. **Muramatsu, S., M. Kato, Y. Kohara, and T. Mizuno.** 1988. Insertion sequence IS5 contains a sharply curved DNA structure at its terminus. Mol. Gen. Genet. **214**:433–438.
243. **Musser, J. M., J. S. Kroll, E. R. Moxon, and R. K. Selander.** 1988. Clonal population structure of encapsulated *Haemophilus influenzae*. Infect. Immun. **56**:1837–1845.
244. **Musser, J. M., V. J. Rapp, and R. K. Selander.** 1987. Clonal diversity in *Haemophilus pleuropneumoniae*. Infect. Immun. **55**:1207–1215.
245. **Muto, A., and S. Osawa.** 1987. The guanine and cytosine content of genomic DNA and bacterial evolution. Proc. Natl. Acad. Sci. USA **84**:166–169.
246. **Nakata, A., M. Amemura, and K. Makino.** 1989. Unusual nucleotide arrangement with repeated sequences in the *Escherichia coli* K-12 chromosome. J. Bacteriol. **171**:3553–3556.
247. **Neumann, H., A. Gierl, J. Tu, J. Leibrock, D. Staiger, and W. Zillig.** 1983. Organization of the genes for ribosomal RNA in archaebacteria. Mol. Gen. Genet. **192**:66–72.
248. **Nghiem, Y., M. Cabrera, C. G. Cupples, and J. H. Miller.** 1988. The *mutY* gene: a mutator locus in *Escherichia coli* that generates G · C→T · A transversions. Proc. Natl. Acad. Sci. USA **85**:2709–2713.
249. **Noack, D., M. Roth, R. Geuther, G. Müller, K. Undisz, C. Hoffmeier, and S. Gaspar.** 1981. Maintenance and genetic stability of vector plasmids pBR322 and pBR325 in *Escherichia coli* K12 strains grown in a chemostat. Mol. Gen. Genet. **184**:121–124.
250. **Noll, K. M.** 1989. Chromosome map of the thermophilic archaebacterium *Thermococcus celer*. J. Bacteriol. **171**:6720–6725.
251. **Nyman, K., K. Nakamura, H. Ohtsubo, and E. Ohtsubo.** 1981. Distribution of the insertion sequence IS1 in gram-negative bacteria. Nature (London) **289**:609–612.
252. **O'Brien, S. J. (ed.).** 1990. Genetic maps. Book 2. Bacteria, algae, and protozoa, 5th ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
253. **Ochman, H., and R. K. Selander.** 1984. Evidence for clonal population structure in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **81**:198–201.
254. **Ochman, H., T. S. Whittam, D. A. Caugant, and R. K. Selander.** 1983. Enzyme polymorphism and genetic population structure in *Escherichia coli* and *Shigella*. J. Gen. Microbiol. **129**:2715–2726.
255. **Ochman, H., and A. C. Wilson.** 1987. Evolutionary history of enteric bacteria, p. 1649–1654. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
256. **Ogasawara, N., M. Q. Fujita, S. Moriya, T. Fukuoka, M. Hirano, and H. Yoshikawa.** 1990. Comparative anatomy of *oriC* of eubacteria, p. 287–295. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
257. **Ohtsubo, H., K. Nyman, W. Doroszkiewicz, and E. Ohtsubo.** 1981. Multiple copies of iso-insertion sequences of IS*1* in *Shigella dysenteriae* chromosome. Nature (London) **292**:640–643.
258. **Omer, C. A., and S. N. Cohen.** 1984. Plasmid formation in *Streptomyces*: excision and integration of the SLP1 replicon at a specific chromosomal site. Mol. Gen. Genet. **196**:429–438.
259. **Omer, C. A., and S. N. Cohen.** 1986. Structural analysis of plasmid and chromosomal loci involved in site-specific excision and integration of the SLP1 element of *Streptomyces coelicolor*. J. Bacteriol. **166**:999–1006.
260. **Omer, C. A., D. Stein, and S. N. Cohen.** 1988. Site-specific insertion of biologically functional adventitious genes into the *Streptomyces lividans* chromosome. J. Bacteriol. **170**:2174–2184.
261. **Panayotatos, N., and A. Fontaine.** 1987. A native cruciform DNA structure probed in bacteria by recombinant T7 endonuclease. J. Biol. Chem. **262**:11363–11368.
262. **Parker, L. L., and B. G. Hall.** 1990. Characterization and nucleotide sequence of the cryptic *cel* operon of *Escherichia coli* K-12. Genetics **124**:455–471.
263. **Patel, A. H., T. J. Foster, and P. A. Pattee.** 1989. Physical and genetic mapping of the protein A gene in the chromosome of *Staphylococcus aureus* 8325-4. J. Gen. Microbiol. **135**:1799–1807.
264. **Pelletier, A. J., T. M. Hill, and P. L. Kuempel.** 1988. Location of sites that inhibit progression of replication forks in the terminus region of *Escherichia coli*. J. Bacteriol. **170**:4293–4298.
265. **Pepper, K., G. DeCespedes, and T. Horaud.** 1988. Heterogeneity of chromosomal genes encoding chloramphenicol resistance in streptococci. Plasmid **19**:71–74.
266. **Pernodet, J.-L., F. Boccard, M.-T. Alegre, J. Gagnat, and M. Guérineau.** 1989. Organization and nucleotide sequence analysis of a ribosomal RNA gene cluster from *Streptomyces ambofaciens*. Gene **79**:33–46.
267. **Pernodet, J. L., J. M. Simonet, and M. Guérineau.** 1984. Plasmids in different strains of *Streptomyces ambofaciens*: free and integrated forms of the plasmid pSAM2. Mol. Gen. Genet. **198**:35–41.
268. **Pettijohn, D. E., and Y. Hodges-Garcia.** 1990. Role of HU protein in transitional DNA coiling, p. 241–245. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
269. **Piggot, P. J.** 1990. Genetic map of *Bacillus subtilis* 168, p. 107–145. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
270. **Plasterk, R. H. A., M. I. Simon, and A. G. Barbour.** 1985. Transposition of structural genes to an expression sequence on a linear plasmid causes antigenic variation in the bacterium *Borrelia hermsii*. Nature (London) **318**:257–263.
271. **Plasterk, R. H. A., and P. Van de Putte.** 1984. Genetic switches by DNA inversions in prokaryotes. Biochim. Biophys. Acta **782**:111–119.
272. **Poddar, S. K., and J. Maniloff.** 1989. Determination of microbial genome sizes by two-dimensional denaturing gradient gel electrophoresis. Nucleic Acids Res. **17**:2889–2895.
273. **Porter, R. D., S. Black, S.Pannuri, and A. Carlson.** 1990. Use of the *Escherichia coli ssb* gene to prevent bioreactor takeover by plasmidless cells. Biotechnology **8**:47–51.
274. **Postgate, J. R., H. M. Kent, R. L. Robson, and J. A. Chesshyre.** 1984. The genomes of *Desulfovibrio gigas* and *D. vulgaris*. J. Gen. Microbiol. **130**:1597–1601.

274a. **Pritchard, A. E., and M. L. Vasil.** 1990. Possible insertion sequences in a mosaic genome organization upstream of the

exotoxin A gene in *Pseudomonas aeruginosa*. J. Bacteriol. **172**:2020–2028.
275. **Pruss, G. J.** 1985. DNA topoisomerase I mutants: increased heterogeneity in linking number and other replicon-dependent changes in DNA supercoiling. J. Mol. Biol. **185**:51–63.
276. **Pruss, G. J., and K. Drlica.** 1986. Topoisomerase I mutants: the gene on pBR322 that encodes resistance to tetracycline affects plasmid DNA supercoiling. Proc. Natl. Acad. Sci. USA **83**:8952–8956.
277. **Punita, N., S. Jafri, M. A. Reddy, and H. K. Das.** 1989. Multiple chromosomes of *Azotobacter vinelandii*. J. Bacteriol. **171**:3133–3138.
278. **Pyle, L. E., L. N. Corcoran, B. G. Cocks, A. D. Bergemann, J. C. Whitley, and L. R. Finch.** 1988. Pulsed-field electrophoresis indicates larger-than-expected sizes for Mycoplasma genomes. Nucleic Acids Res. **16**:6015–6025.
279. **Radding, C. M., J. Flory, A. Wu, R. Kahn, G. DasGupta, D. Gonda, M. Bianchi, and S. S. Tsang.** 1982. Three phases in homologous pairing: Polymerization of recA protein on single-stranded DNA, synapsis, and polar strand exchange. Cold Spring Harbor Symp. Quant. Biol. **47**:821–828.
280. **Radman, M., and R. Wagner.** 1986. Mismatch repair in *Escherichia coli*. Annu. Rev. Genet. **20**:523–538.
281. **Rahmouni, A. R., and R. D. Wells.** 1989. Stabilization of Z DNA in vivo by localized supercoiling. Science **246**:358–363.
282. **Rayssiguier, C., D. S. Thaler, and M. Radman.** 1989. The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. Nature (London) **342**:396–401.
283. **Reanney, D.** 1976. Extrachromosomal elements as possible agents of adaptation and development. Bacteriol. Rev. **40**:552–590.
284. **Rebollo, J. E., V. François, and J. M. Louarn.** 1988. Detection and possible role of two large nondivisible zones on the *Escherichia coli* chromosome. Proc. Natl. Acad. Sci. USA **85**:9391–9395.
285. **Reeve, E., and J. A. Braithwaite.** 1974. The lactose system in *Klebsiella aerogenes* V9A. A comparison of the lac operons of *Klebsiella* and *Escherichia coli*. Genet Res. **24**:323–331.
286. **Rich, A., A. Nordheim, and A. H.-J. Wang.** 1984. The chemistry and biology of left-handed Z-DNA. Annu. Rev. Biochem. **53**:791–846.
287. **Riley, M., and A. Anilionis.** 1978. Evolution of the bacterial genome. Annu. Rev. Microbiol. **32**:519–560.
288. **Riley, M., and S. Krawiec.** 1987. Genome organization, p. 967–981. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
289. **Riley, M., and K. E. Sanderson.** 1990. Comparative genetics of *Escherichia coli* and *Salmonella typhimurium*, p. 85–95. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
290. **Riley, M., L. Solomon, and D. Zipkas.** 1978. Relationship between gene function and gene location in *Escherichia coli*. J. Mol. Evol. **11**:47–56.
291. **Robson, R. L., H. A. Chesshyre, C. Wheeler, R. Jones, P. R. Woodley, and J. R. Postgate.** 1984. Genome size and complexity in *Azotobacter chroococcum*. J. Gen. Microbiol. **130**:1603–1612.
292. **Römling, U., D. Grothues, W. Bautsch, and B. Tümmler.** 1989. A physical genome map of *Pseudomonas aeruginosa* PAQ. EMBO J. **8**:4081–4089.
293. **Rouvière-Yaniv, J., M. Yaniv, and J.-E. Germond.** 1979. *E. coli* DNA binding protein HU forms nucleosome-like structure with circular double-stranded DNA. Cell **17**:265–274.
294. **Rudd, K. E., W. Miller, J. Ostell, and D. A. Benson.** 1990. Alignment of *Escherichia coli* K-12 sequences to a genomic restriction map. Nucleic Acids Res. **18**:313–321.
295. **Sadoff, H. L., B. Shimel, and S. Ellis.** 1979. Characterization of *Azotobacter vinelandii* deoxyribonucleic acid and folded chromosomes. J. Bacteriol. **138**:871–877.
296. **Sadosky, A. B., A. Davidson, R.-J. Lin, and C. W. Hill.** 1989. *rhs* gene family of *Escherichia coli* K-12. J. Bacteriol. **171**:636–642.
297. **Sanderson, K. E., and J. A. Hurley.** 1987. Linkage map of *Salmonella typhimurium*, p. 877–918. *In* F. C. Neidhardt, J. L. Ingraham, B. Magasanik, K. B. Low, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
298. **Sawyer, S.** 1989. Statistical tests for detecting gene conversion. Mol. Biol. Evol. **6**:526–538.
299. **Sawyer, S. A., D. E. Dykhuizen, and D. L. Hartl.** 1987. Confidence interval for the number of selectively neutral amino acid polymorphisms. Proc. Natl. Acad. Sci. USA **84**:6225–6228.
300. **Scaife, J.** 1966. F-prime factor formation in *E. coli* K-12. Genet. Res. **8**:189–196.
301. **Schafer, A., J. Kalinowski, R. Simon, A.-H. Seep-Feldhaus, and A. H. Puhler.** 1990. High-frequency conjugal plasmid transfer from gram-negative *Escherichia coli* to various gram-positive coryneform bacteria. J. Bacteriol. **172**:1663–1666.
302. **Schleif, R.** 1988. DNA looping. Science **240**:127–128.
303. **Schmid, M. B., and J. R. Roth.** 1983. Selection and end point distribution of bacterial inversion mutations. Genetics **105**: 539–557.
304. **Schmid, M. B., and J. R. Roth.** 1987. Gene location affects expression level in *Salmonella typhimurium*. J. Bacteriol. **169**:2872–2875.
305. **Schneider, W. P., B. P. Nichols, and C. Yanofsky.** 1981. Procedure for production of hybrid genes and proteins and its use in assessing significance of amino acid differences in homologous tryptophan synthetase alpha polypeptides. Proc. Natl. Acad. Sci. USA **78**:2169–2173.
306. **Schultz, D. W., and G. R. Smith.** 1986. Conservation of Chi cutting activity in terrestrial and marine enteric bacteria. J. Mol. Biol. **189**:585–595.
307. **Segall, A., M. J. Mahan, and J. R. Roth.** 1988. Rearrangement of the bacterial chromosome: forbidden inversions. Science **241**:1314–1318.
308. **Segall, A. M., and J. R. Roth.** 1989. Recombination between homologies in direct and inverse orientation in the chromosome of *Salmonella*: intervals which are nonpermissive for inversion formation. Genetics **122**:737–747.
309. **Segall, E., P. Hagblom, H. S. Seifert, and M. So.** 1986. Antigenic variation of gonococcal pilus involves assembly of separated silent gene segments. Proc. Natl. Acad. Sci. USA **83**:2177–2181.
310. **Seifert, H. S., R. S. Ajioka, C. Marchal, F. P. Sparling, and M. So.** 1988. DNA transformation leads to pilin antigenic variation in *Neisseria gonorrhoeae*. Nature (London) **336**:392–395.
311. **Sela, S., J. E. Clark-Curtiss, and H. Bercovier.** 1989. Characterization and taxonomic implications of the rRNA genes of *Mycobacterium leprae*. J. Bacteriol. **171**:70–73.
312. **Selander, R. K., D. A. Caugant, and T. S. Whittam.** 1987. Genetic structure and variation in natural populations of *Escherichia coli*, p. 1625–1648. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, D.C.
313. **Selander, R. K., J. M. Musser, D. A. Caugant, M. N. Gilmour, and T. S. Whittam.** 1987. Population genetics of pathogenic bacteria. Microb. Pathog. **3**:1–7.
314. **Shapiro, J. A.** 1984. Observations on the formation of clones containing *ara*B-*lacZ* cistron fusions. Mol. Gen. Genet. **194**: 79–90.
315. **Sharp, P. M.** 1990. Processes of genome evolution reflected by base frequency differences among *Serratia marcescens* genes. Mol. Microbiol. **4**:119–122.
316. **Sharp, P. M., and W.-H. Li.** 1986. An evolutionary perspective on synonymous codon usage in unicellular organisms. J. Mol. Evol. **24**:28–38.
317. **Sharp, P. M., D. C. Shields, K. H. Wolfe, and W.-H. Li.** 1989.

Evolutionary rates in enterobacterial genes vary with chromosomal location. Science **246**:808–810.
318. **Shen, P., and H. V. Huang.** 1986. Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. Genetics **112**:441–457.
319. **Simon, M. I., and M. Silverman.** 1983. Recombinational regulation of gene expression in bacteria, p. 211–227. *In* J. Beckman, J. Davies, and J. A. Gallant (ed.), Gene function in prokaryotes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
320. **Sinclair, M. I., P. C. Maxwll, B. R. Lyon, and B. W. Holloway.** 1986. Chromosomal location of TOL plasmid DNA in *Pseudomonas putida*. J. Bacteriol. **168**:1302–1308.
321. **Singleton, C. K., J. Klysik, S. M. Stirdivant, and R. D. Wells.** 1982. Left-handed Z-DNA is induced by supercoiling in physiological ionic conditions. Nature (London) **299**:312–316.
322. **Smith, C. L., and G. Condemine.** 1990. New approaches for physical mapping of small genomes. J. Bacteriol. **172**:1167–1172.
323. **Smith, C. L., G. Condemine, and S. Ringquist.** 1990. Electrophoretic analysis of large DNA: application to the structure and dynamics of the *Escherichia coli* chromosome, p. 205–210. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
324. **Smith, C. L., J. G. Econome, A. Schutt, S. Klco, and C. R. Cantor.** 1987. A physical map of the *E. coli* K-12 genome. Science **236**:1448–1453.
325. **Smith, D., and M. Yarus.** 1989. tRNA-tRNA interactions within cellular ribosomes. Proc. Natl. Acad. Sci. USA **86**: 4397–4401.
326. **Smith, G. R.** 1989. Homologous recombination in *E. coli*: multiple pathways for multiple reasons. Cell **58**:807–809.
327. **Smith, G. R., S. K. Amundsen, A. M. Chaudhury, K. C. Cheng, A. S. Ponticelli, C. M. Roberts, D. W. Schultz, and A. F. Taylor.** 1984. Roles of RecBC enzymes and Chi sites in homologous recombination. Cold Spring Harbor Symp. Quant. Biol. **59**:485–495.
328. **Smith, G. R., S. M. Kunes, D. W. Schultz, A. Taylor, and K. L. Triman.** 1981. Structure of Chi hotspots of generalized recombination. Cell **24**:429–436.
329. **Snyder, U. K., J. F. Thompson, and A. Landy.** 1989. Phasing of protein-induced DNA bends in a recombination complex. Nature (London) **341**:255–257.
330. **Sparling, F. P., J. G. Cannon, and M. So.** 1986. Phase and antigenic variation of pili and outer membrane protein II of *Neisseria gonorrhoeae*. J. Infect. Dis. **153**:196–201.
331. **Stachel, E. E., and P. C. Zambryski.** 1989. Generic transkingdom sex? Nature (London) **340**:190–191.
332. **Stahl, F.** 1988. A unicorn in the garden. Nature (London) **335**:112–113.
333. **Stern, A., P. Nickel, and T. F. Meyer.** 1986. Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. Cell **47**:61–71.
334. **Stern, A., P. Nickel, T. F. Meyer, and M. So.** 1984. Opacity determinants of *Neisseria gonorrhoeae*: gene expression and chromosomal linkage to the gonococcal pilus gene. Cell **37**: 447–456.
335. **Stern, M. J., G. F.-L. Ames, N. A. Smith, E. C. Robinson, and C. F. Higgins.** 1984. Repetitive extrapalindromic sequences: a major component of the bacterial genome. Cell **37**:1015–1026.
336. **Sternberg, N.** 1985. Evidence that adenine methylation influences DNA-protein interactions in *Escherichia coli*. J. Bacteriol. **164**:490–493.
337. **Stewart, F. M., D. M. Gordon, and B. R. Levin.** 1990. Fluctuation analysis: the probability distribution of the number of mutants under different conditions. Genetics **124**:175–185.
338. **Stoltzfus, A., J. F. Leslie, and R. Milkman.** 1988. Molecular evolution of the *Escherichia coli* chromosome. I. Analysis of structure and natural variation in a previously uncharacterized region between *trp* and *tonB*. Genetics **120**:345–358.
339. **Storts, D. R., and A. Markovitz.** 1988. Construction and characterization of mutations in *hupB*, the gene encoding HU-β (HU-1) in *Escherichia coli* K-12. J. Bacteriol. **170**:1541–1547.
340. **Su, C. J., and J. B. Baseman.** 1990. Genome size of *Mycoplasma geneitalium*. J. Bacteriol. **172**:4705–4707.
341. **Sueoka, N.** 1988. Directional mutation pressure and neutral molecular evolution. Proc. Natl. Acad. Sci. USA **85**:2653–2657.
342. **Sullivan, K. M., and D. M. J. Lilley.** 1986. A dominant influence of flanking sequences on a local structural transition in DNA. Cell **47**:817–827.
343. **Suwanto, A., and S. Kaplan.** 1989. Physical and genetic mapping of the *Rhodobacter sphaeroides* 2.4.1 genome: genome size, fragment identification, and gene localization. J. Bacteriol. **171**:5840–5849.
344. **Suwanto, A., and S. Kaplan.** 1989. Physical and genetic mapping of the *Rhodobacter sphaeroides* 2.4.1 genome: presence of two unique circular chromosomes. J. Bacteriol. **171**:5850–5859.
345. **Suzuki, Y., T. Mori, Y. Miyata, and T. Yamada.** 1987. The number of ribosomal RNA genes in *Mycobacterium lepraemurium*. FEMS Microbiol. Lett. **44**:73–76.
346. **Suzuki, Y., Y. Ono, A. Nagata, and T. Yamada.** 1988. Molecular cloning and characterization of an rRNA operon in *Streptomyces lividans* TK21. J. Bacteriol. **170**:1631–1636.
347. **Suzuki, Y., K. Yoshinaga, Y. Ono, A. Nagata, and T. Yamada.** 1987. Organization of rRNA Genes in *Mycobacterium bovis* BCG. J. Bacteriol. **169**:839–843.
348. **Taschke, C., M.-Q. Klinkert, J. Wolters, and R. Herrmann.** 1986. Organization of the ribosomal RNA genes in *Mycoplasma hypopneumoniae*: the 5S rRNA gene is separated from the 16S and 23S rRNA genes. Mol. Gen. Genet. **205**:428–433.
349. **Thaler, D. S., J. R. Roth, and L. Hirschbein.** 1990. Imprinting as a mechanism for control of gene expression, p. 445–456. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
350. **Trieu-Cuot, P., C. Carlier, and P. Courvalin.** 1988. Conjugative plasmid transfer from *Enterococcus faecalis* to *Escherichia coli*. J. Bacteriol. **170**:4388–4391.
351. **Trieu-Cuot, P., C. Carlier, P. Martin, and P. Courvalin.** 1978. Plasmid transfer by conjugation from *Escherichia coli* to grampositive bacteria. FEMS Microbiol. Lett. **48**:289–294.
352. **Tse-Dinh, Y.** 1985. Regulation of the *Escherichia coli* DNA topoisomerase I gene by DNA supercoiling. Nucleic Acids Res. **13**:4751–4763.
353. **Tsuda, M., and T. Iino.** 1987. Genetic analysis of a transposon carrying toluene degrading genes on a TOL plasmid pWWO. Mol. Gen. Genet. **210**:270–276.
354. **Tsuda, M., and T. Iino.** 1988. Identification and characterization of Tn4653, a transposon covering the toluene transposon Tn4651 on TOL plasmid pWWO. Mol. Gen. Genet. **213**:72–77.
355. **Umeda, M., and E. Ohtsubu.** 1989. Mapping of insertion elements IS*1*, IS*2* and IS*3* on the *Escherichia coli* K-12 chromosome: role of the insertion elements in formation of Hfrs and F′ factors and in rearrangement of bacterial chromosomes. J. Mol. Biol. **208**:601–614.
356. **Umeda, M., and E. Ohtsubu.** 1990. Mapping of insertion element IS*5* in the *Escherichia coli* K-12 chromosome: chromosomal rearrangements mediated by IS*5*. J. Mol. Biol. **213**: 229–237.
357. **Van Vliet, F., A. Boyen, and N. Glansdorff.** 1988. On interspecies gene transfer: the case of the *argF* gene of *Escherichia coli*. Ann. Inst. Pasteur Microbiol. **139**:493–496.
358. **Vary, P. S., and Y.-P. Tao.** 1988. Development of genetic methods in *Bacillus megaterium*, p. 403–407. *In* J. A. Hoch and A. T. Ganesan (eds.), Genetics and biotechnology of bacilli. II. Academic Press, Inc., New York.
359. **Ventra, L., and A. S. Weiss.** 1989. Transposon-medited restriction mapping of the *Bacillus subtilis* chromosome. Gene **78**:29–36.
360. **Vold, B. S.** 1985. Structure and organization of genes for transfer ribonucleic acid in *Bacillus subtilis*. Microbiol. Rev. **49**:71–80.
361. **Wada, M., Y. Kano, T. Ogawa, T. Okazaki, and F. Imamoto.** 1988. Construction and characterization of the deletion mutant

of *hupA* and *hupB* genes in *Escherichia coli*. J. Mol. Biol. **204**:581–591.
362. **Wada, M., K. Kutsukake, T. Komano, F. Imamoto, and Y. Kano.** 1989. Participation of the *hup* gene product in site-specific DNA inversion in *Escherichia coli*. Gene **76**:345–352.
363. **Wake, R. G.** 1990. Termination of chromosome replication in *Bacillus subtilis*, p. 307–312. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
364. **Wang, J. C.** 1985. DNA topoisomerases. Annu. Rev. Biochem. **54**:665–697.
365. **Wang, J. C., and G. N. Giaever.** 1988. Action at a distance along a DNA. Science **240**:300–304.
366. **Weil, M. D., and M. McClelland.** 1989. Enzymatic cleavage of a bacterial genome at a 10-base-pair recognition site. Proc. Natl. Acad. Sci. USA **86**:51–55.
367. **Weiser, J. N., D. J. Maskell, P. D. Butler, A. A. Lindberg, and E. R. Moxon.** 1990. Characterization of repetitive sequences controlling phase variation of *Haemophilus influenzae* lipopolysaccharide. J. Bacteriol. **172**:3304–3309.
368. **Wells, R. D.** 1988. Unusual DNA structures. J. Biol. Chem. **263**:1095–1098.
369. **Wells, R. D., S. Amirhaeri, J. A. Blaho, D. A. Collier, A. Dohrman, J. A. Griffin, J. C. Hanvey, W.-T. Hsieh, A. Jaworski, J. E. Larson, M. J. McLean, A. Rahmouni, M. Rajagopalan, M. Shimizu, F. Wohlrab, and W. Zacharias.** 1990. Biology and chemistry of Z DNA and triplexes, p. 187–194. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
370. **Wells, R. D., D. A. Collier, J. C. Hanvey, M. Shimizu, and F. Wohlrab.** 1988. The chemistry and biology of unusual DNA structures adopted by oligopurine · oligopyrimidine sequences. FASEB J. **2**:2939–2948.
371. **Wenzel, R., and R. Herrmann.** 1989. Cloning of the complete *Mycoplasma pneumoniae* genome. Nucleic Acids Res. **17**: 7029–7043.
372. **West, S. E. H., and B. H. Iglewski.** 1988. Codon usage in *Pseudomonas aeruginosa*. Nucleic Acids Res. **16**:9323–9335.
373. **Whitehead, T. R., and J. C. Rabinowitz.** 1988. Nucleotide sequence of the *Clostridium acidiurici* ("*Clostridium acidiurici*") gene for 10-formyltetrahydrofolate synthetase shows extensive amino acid homology with the trifunctional enzyme C(1)-tetrahydrofolate synthase from *Saccharomyces cerevisiae*. J. Bacteriol. **170**:3255–3261.
374. **Whittam, T. S., H. Ochman, and R. K. Selander.** 1983. Multilocus genetic structure in natural populations of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **80**:1751–1755.
375. **Widom, R. L., E. D. Jarvis, G. LaFauci, and R. Rudner.** 1988. Instability of rRNA operons in *Bacillus subtilis*. J. Bacteriol. **170**:605–610. (Erratum, **170**:2003.)
376. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51**: 221–271.
377. **Woodward, M. J., and H. P. Charles.** 1983. Polymorphism in *Escherichia coli rtl-atl* and *gat* regions behave as chromosomal alternatives. J. Gen. Microbiol. **129**:75–84.
378. **Wu, H.-Y., S. Shyy, J. C. Wang, and L. F. Liu.** 1988. Transcription generates positively and negatively supercoiled domains in the template. Cell **53**:433–440.
379. **Xia, X.-M., and M. Enomoto.** 1986. A naturally occurring large chromosomal inversion in *Escherichia coli* K-12. Mol. Gen. Genet. **205**:376–379.
380. **Yamagishi, A., and T. Oshima.** 1990. Circular chromosomal DNA in the sulfur-dependent archaebacterium *Sulfolobus acidocaldarius*. Nucleic Acids Res. **18**:1133–1135.
381. **Yang, Y., and G. Ames.** 1990. The family of repetitive extragenic palindromic sequences: interaction with DNA gyrase and histone-like protein HU, p. 211–225. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
382. **Yang, Y., and G. F.-L. Ames.** 1989. DNA gyrase binds to the family of prokaryotic repetitive extragenic palindromic sequences. Proc. Natl. Acad. Sci. USA **85**:8850–8854.
383. **Yarus, M., and L. S. Folley.** 1985. Sense codons are found in specific contexts. J. Mol. Biol. **182**:529–540.
384. **Yee, T., and M. Inouye.** 1982. Two-dimensional DNA electrophoresis applied to the study of DNA methylation and the analysis of genome size in *Myxococcus xanthus*. J. Mol. Biol. **154**:181–196.
385. **York, M. K., and M. Stodolsky.** 1981. Characterization of P1*arg*F derivatives from *E. coli* K-12 transduction. I. IS1 elements flank the *argF* segment. Mol. Gen. Genet. **181**:230–240.
386. **Yuan, R.** 1981. Structure and mechanism of multi-functional restriction endonucleases. Annu. Rev. Biochem. **50**:285–315.
387. **Zambryski, P. C.** 1988. Basic processes underlying *Agrobacterium*-mediated DNA transfer to plant cells. Annu. Rev. Genet. **22**:1–30.
388. **Ziegler, D., and D. Dean.** 1990. Orientation of genes in the *Bacillus subtilis* chromosome. Genetics **125**:703–708.