Rearrangements of the Bacterial Chromosome: Formation and Applications

JOHN R. ROTH, NICHOLAS BENSON, TIMOTHY GALITSKI, KENNETH HAACK, JEFFREY G. LAWRENCE, AND LYNN MIESEL

120

INTRODUCTION

One of the adjuncts to DNA replication and repair is the formation of occasional chromosomal rearrangements—deletions, tandem duplications, and inversions. Several earlier reviews have discussed aspects of this subject (6, 61, 83). Rearrangements differ greatly from the point mutations that are more often considered in thinking about evolution and analytical genetics. Deletions can remove multiple functions and are irreversible. Duplications can amplify a coding region and are so highly reversible that they might be considered a temporary "regulatory state" rather than a mutation. Inversions change the orientation of a sequence in the chromosome but disrupt that sequence only at the two endpoints. Inversions are not highly reversible. All of these mutations have consequences for the structure of the chromosome and may lead, on an evolutionary timescale, to changes in the genetic map. In this review, we will describe the formal characteristics of these rearrangements and the ways in which they may be formed. Duplications and deletions, despite their distinct properties and consequences, can be generated by similar sister strand exchanges and will be discussed together. Inversions are fundamentally more complicated and will be discussed separately, with some of their implications for chromosome structure.

Chromosome rearrangements are valuable additions to the arsenal of techniques available for bacterial genetics. Historically, deletions have been used extensively in fine-structure genetic mapping, but duplications also have properties that are particularly valuable as analytical tools, especially in complementation testing. Rearrangements provide a system for assaying recombination that closely approximates the natural role of the recombination system in DNA damage repair. The genetic exchanges assayed by sexual exchanges are probably only a subset of the normal activities of the recombination system. Since natural populations of enteric bacteria are largely clonal, mating is very rare (101). We suggest that bacterial recombination systems have evolved largely because of their role in repair and formation of adaptive rearrangements of the chromosome. We will describe experimental methods for construction of various rearrangements and the ways in which these rearrangements can be applied to the study of recombination.

The chromosome maps of *Salmonella typhimurium* (official designation, *Salmonella enterica* serovar Typhimurium) and *Escherichia coli* have remained well conserved over the estimated 100 million years since the divergence of these groups (82). This conservation is remarkable in that some chromosomal rearrangements occur frequently (in the laboratory) and some types can result in stable alterations of the genetic map. Duplications form and segregate frequently, and this ability appears to be selectively valuable to bacteria in allowing a population to adapt to stressful conditions. Thus, we are faced with a paradox, i.e., that the chromosome seems stable on a long timescale but unstable on a short timescale. We suggest that the selective forces that act to conserve the genetic map have secondarily shaped the recombination system of bacteria, leading to evolution of a system that acts preferentially on direct-order repeated sequences (useful in repair and formation of deletions and duplications) and avoids making the exchanges between inverse-order sequences that can lead to inversions.

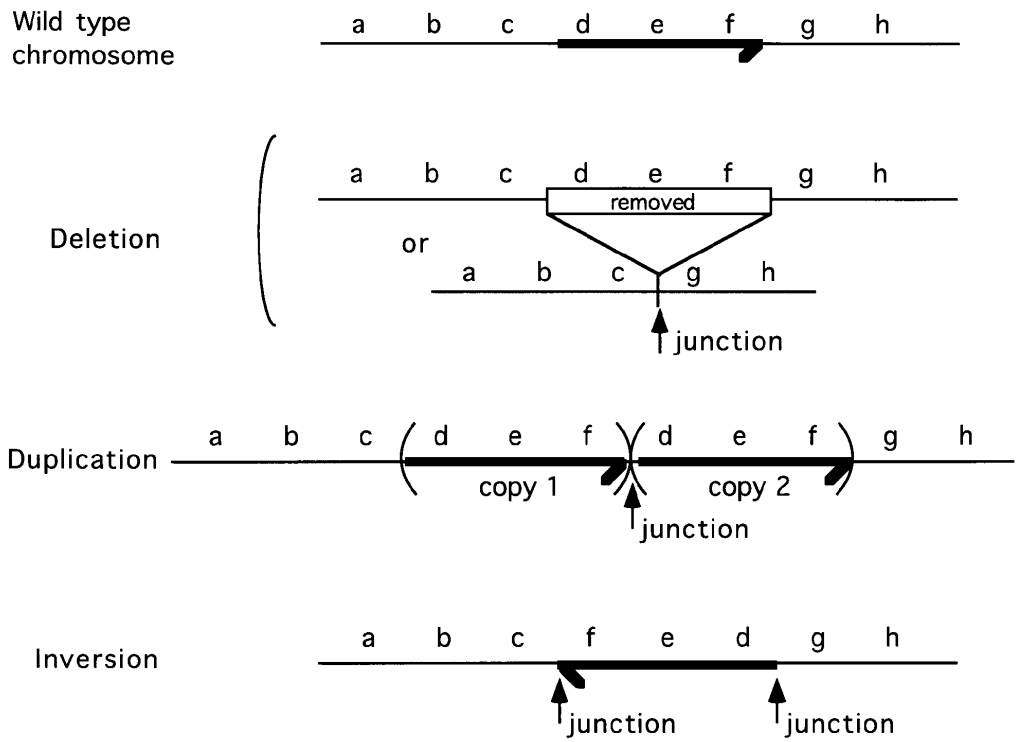

FIGURE 1 Properties of four basic rearrangement types. In the top line, the heavy line designates the portion of the wild-type chromosome affected by each of the rearrangements that follow. The next line presents a deletion; the missing material is designated by an open box. Then alternative diagram of the same deletion emphasizes that this mutuation generates a novel sequence at the deletion junction point as well as causing a loss of material. The following line presents a duplication; note that this rearrangement, like a deletion, has only one novel sequence element, present at the junction point. The bottom line describes an inversion. Note that an inversion causes no loss or gain of material but disrupts the wild-type sequence and creates two new sequence elements at the two junction points

PROPERTIES AND FORMATION OF REARRANGEMENTS

Defining Rearrangements

Essential Features. The rearrangements considered here are diagrammed in Fig. 1. In this figure, the lowercase letters denote sequence elements and the heavy line indicates the portion of the sequence affected by each mutation. Note the following basic features of these rearrangement types. (i) Both deletions and duplications confer a change in the amount of information. Deletions reduce the copy number of a segment by one (to zero), and duplications increase it by one (to two). Both rearrangements leave a chromosome with a single novel sequence element at the junction point. (ii) The phenotype of a deletion can be due to loss of information or to functional consequences of the new sequence element formed at the junction point. (iii) A tandem duplication does not cause the strain to lose expression of any genetic information as long as the duplication is not entirely included within a single functional element (gene or operon). (iv) The phenotype of a duplication can be due to the increase in gene dosage or to functional consequences of the new sequence element formed at the junction point. (v) Inversions involve no loss or gain of chromosomal material but alter the wild-type sequence at the two junction points. (vi) The

phenotype of an inversion could be due to disruption of wild-type sequence at one or both junction points, to new properties of the sequences created at these points, or to consequences that the reversed orientation of the segment might have for gene expression. (vii) All classes of rearrangements could have effects on folding of the chromosome and could therefore have wide-ranging indirect effects on phenotype. All the rearrangements in Fig. 1 can be imagined to occur by recombinational exchanges between repeated direct- or inverse-order chromosomal sequences. (In fact, as we shall see, some may occur by other means.)

Formal Requirements for Forming Deletions and Tandem Duplications. One method of generating deletions and duplications is by recombination between repeated sequence elements. Figure 2 describes the similarity in the behavior of duplications and deletions. In considering the events diagrammed in Figure 2, the following points should be noted. (i) On the left, deletions and duplications can both be formed by a sister strand exchange between short direct-order repeats (the open boxes at the ends of the affected segments). A full exchange (rejoining both flanking sequences) between such repeats generates both a deletion and a duplication in a single event. However, a half-exchange (one pair of flanks rejoined, the other pair left as broken ends) is sufficient to make either of the two rearrangement types singly. (ii) A deletion can also be formed (in principle) by an intrachromosomal half exchange between two repeats in a single chromosome (top right). If this were a full exchange, both a deletion and a free circular fragment would be made. A half exchange can form a deletion and a linear fragment. (iii) Once a tandem duplication is present, it is made unstable by the same sort of event that generates a deletion. That is, a sister strand exchange between different copies of the repeated segments can return the chromosome to a haploid state (or generate further amplification). Alternatively, the duplication can be lost by an exchange between the repeated segments in a single chromosome (lower right). Loss of a duplication can result from an exchange at any point in the entire duplicated segment. Since this is typically a long region (>1 min), most duplications are extremely unstable. Their rate of loss can be influenced by the size of the duplicated segment (and by other factors), but rates on the order of 1% per generation are common. (iv) The recombination events that generate duplications and deletions are formally similar to the events that lead to loss (segregation) of a duplication and further amplification of the copy number. For all these events, either a full or a half exchange will suffice. Events can be intrachromosomal or occur between sister chromosomes.

One might imagine that all exchanges between sister chromosomes (including the unequal exchanges that can cause duplications and deletions) would have consequences for chromosome partitioning. A single exchange between sister chromosomes (or an odd number of such exchanges) results in formation of a chromosome dimer. An additional exchange resolves the dimer and allows partitioning. It has been hypothesized that mechanisms in the terminus region of the chromosome provide this function. In particular, the site-specific recombinase XerCD catalyzes full exchanges at the Dif site in the terminus region (16, 27, 67, 89). These Dif-Dif exchanges can in principle reverse the deleterious topological consequences of sister strand exchanges involved in repair and rearrangement of a circular chromosome.

Formal Requirements for Inversion Formation. Generation of an inversion is a more demanding event than those described above. Two ways of visualizing this process are depicted in Fig. 3. The simplest possibility is an intrachromosomal full exchange occurring between inverse-order repeated sequences; in this exchange, both pairs of flanking sequences must be rejoined (left side of Fig. 3). However, one can also imagine an inversion forming by a series of half exchanges occurring either simultaneously or sequentially (right side of Fig. 3). It is important to remember these alternative possibilities when using inversions to study recombination, because it is not yet clear whether bacteria are able to perform a full exchange. They may make only half exchanges and use the generated double-strand breaks to stimulate highly efficient repair of the second pair of flanking sequences. Most of the evidence that has been interpreted as supporting full exchanges, such as circle cointegration and formation of inversions, can also be explained by a series of half exchanges such as those diagrammed at the right of Fig. 3. For a discussion of this point, see the appendix to reference 99.

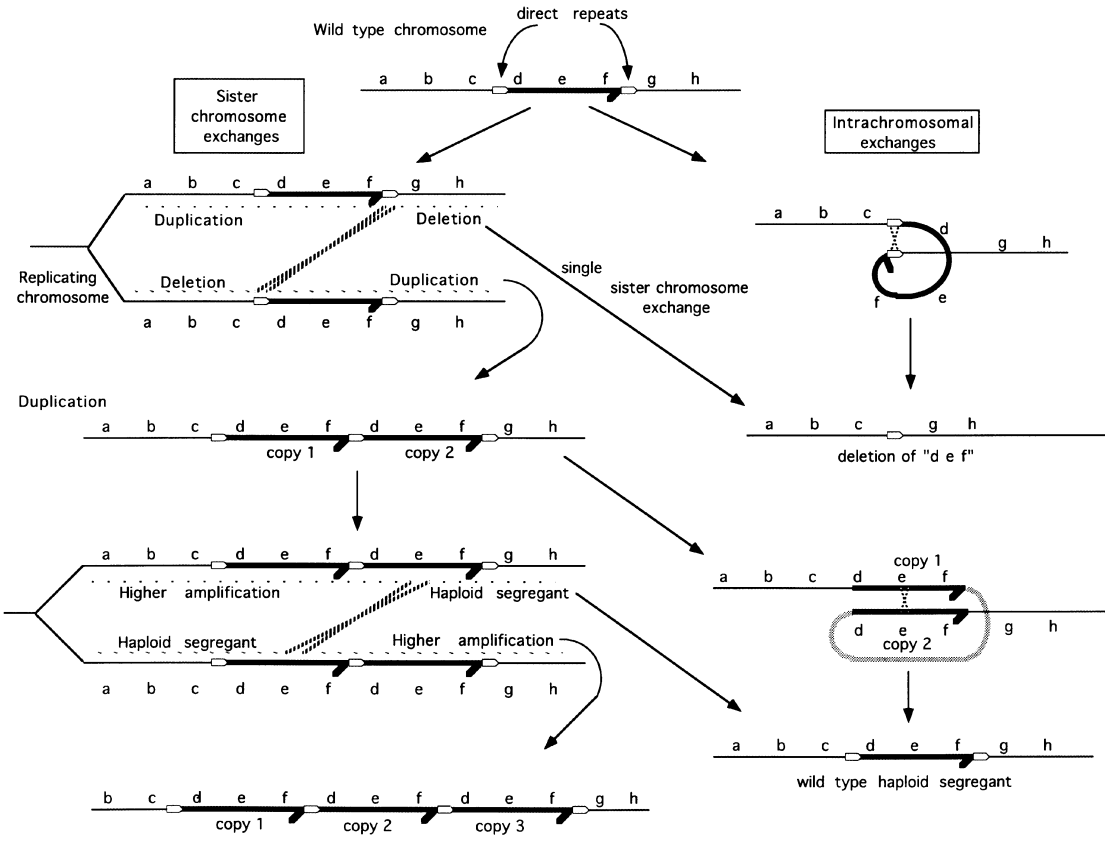

FIGURE 2 Unequal exchanges involved in formation of duplications and deletions. The top line presents a wild-type chromosome with small naturally repeated sequences. Events diagrammed at the left of the figure are sister chromosome exchanges that form a deletion or duplication of material between repeats. Even diagrammed at the right are intrachromosomal exchanges that can form a detection or segregate a duplication.

Sequences Used in Rearrangement Formation. Any type of repeated sequence could provide regions of homology at which recombination could occur to form a rearrangement. The seven *rrn* loci are prominent regions of repeated sequence (9, 46). The role of these long repeated sequences in chromosome rearrangement has been reviewed by Petes and Hill (83). In *S. typhimurium,* insertion sequences and REP elements can be used as passive sites for recombinational exchanges that generate duplications (7, 102; K. Haack and J. R. Roth, unpublished results). In *E. coli,* the Rhs sequences are used as homologies in formation of rearrangements (66). A variety of other long sequence repeats probably serve at frequencies that reflect the length of the repeat and the degree of the sequence similarity. Rearrangments can also be formed by transpositional activities of tranposable elements and by site-specific recombination systems, but these events are outside the scope of this review.

Frequency of Rearrangements. It is difficult to compare the relative frequencies of the three rearrangement types discussed here, because the detection systems used are not equivalent. Deletions are extremely frequent when they can occur between large direct sequence repeats (25). However, this situation is very seldom seen unless these repeats are experimentally placed close together in the chromosome without any intervening essential gene. In a wild-type chromosome, deletions are generally limited (by their potential lethality) to small regions of the chromosome devoid of essential genes. This means that the size of most deletions is less than 1% of the chromosome. Because regions of this size are unlikely to contain natural repeats of a large sequence element, deletions must form by using minimal sequence repeats in a small region or by a variety of mechanisms that are far less efficient than the homologous recombination between long repeats. This puts a strong limitation on the frequency of detectable deletions. Typically, deletions represent about 10% of the spontaneous null mutations at a locus; therefore, they are expected at a frequency of roughly 10^{-6} in an unselected cell population. The actual frequency of deletions for any particular gene will depend on the extent of the local region that can be deleted without lethal effects and on local sequence features.

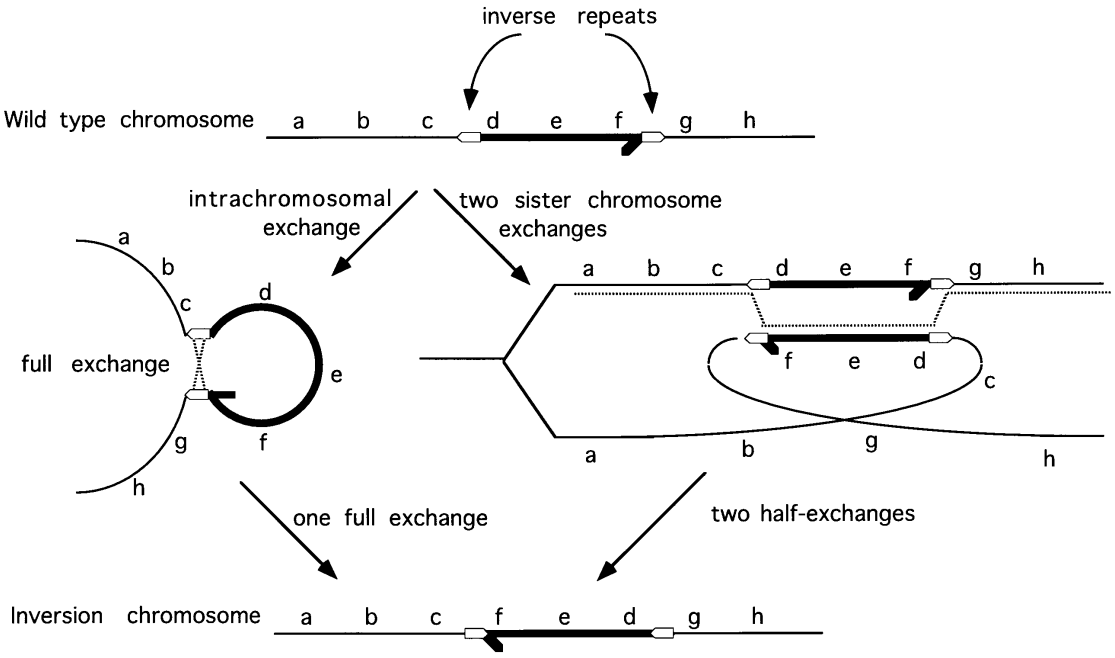

FIGURE 3 Exchanges involved in formation of inversions. The top line describes a wild-type chromosome with a segment (heavy line) flanked by natural inverse-order repeats. In principle, inversions can form by either intra- or interchromosomal exchanges. The left side of the figure diagrams inversion formation in a single circular chromosome, which requires a full exchange (both pairs of flanking sequences are rejoined). The right side of the figure describes formation of an inversion by interaction of sister chromosomes. In this case, two exchanges are required by half exchanges are sufficient (one pair of flanks rejoined).

Duplications are almost unrestricted in size and location in the chromosome. A large set of spontaneous duplications of various loci of the *Salmonella* chromosome included many that encompassed nearly half the chromosome (9). We have been able to construct duplications of all regions of the chromosome except the replication terminus region (D. Hillyard and J. R. Roth, unpublished results). The only

recovered duplications found to cause a severe growth defect were those that included both the main ribosomal protein operon (map min 73 in *E. coli*) and an additional large block of chromosomal material (Hillyard and Roth, unpublished). The failure to detect duplications of the terminus region may reflect a lethal effect of having two copies of this region or may be due to severe instability of such duplications. This instability could be caused by action of the machinery for segregating chromosomes (e.g., Dif/XerCD); this possibility was suggested by David Sherratt (personal communication).

The high measured frequency of large duplications (and the strong RecA dependence) is almost certainly due to widely separated long repeated sequences (*rrn* loci and IS sequences) being used as efficient recombining sites to form the rearrangement. When one isolates duplications that include any specified point in the chromosome, one usually encounters these large duplications that form frequently by exchanges between long repeats (5, 48; Hillyard and Roth, unpublished). In an unselected *Salmonella* culture, the frequency of duplications of a particular locus varied from 3% (for genes between the closest *rrn* loci) to about 10^{-5} for genes near the terminus (*trp*); a typical locus is duplicated in about 10 of the cells in an unselected culture. We estimate that nearly 10% of the cells in an unselected culture carry a duplication of some region of the chromosome. When considered together with the instability of duplications, these estimates suggest that the bacterial chromosome is in a constant state of flux, with duplications being continually acquired and lost.

Inversion frequency is very difficult to measure, because the main detection methods rely on observing phenotypes generated at the two inversion junction points; this is difficult to do on a large scale in a wild-type chromosome. One might expect that gene disruption at inversion junction points might have caused some of the many spontaneous mutations isolated in the course of characterizing bacterial operons. To our knowledge, no inversion mutants were reported in any of these collections, while many deletion mutants were reported. Grabnar et al. tested *Salmonella his* mutations for inversions and found none (41). Since operons are unlikely to include extensive sequences repeated elsewhere in the chromosome, both deletions and inversions must presumably form by using very short, repeated sequences. Thus, the observation of deletions and the failure to observe inversions among spontaneous null mutations suggests that inversion formation may be mechanistically difficult. If inversions can occur only by recombination between extensive inverse-order sequence repeats, the failure to detect inversion breakpoints among mutants with a nutritional phenotype may simply reflect the fact that the operons studied do not include such repetitive sequences.

Most work on inversions has involved placing particular characterized sequences in inverse order at known positions in the chromosome and observing their recombination. As discussed below, sequences placed in inverse order at some chromosomal sites recombine to yield inversions while those placed at other sites (nonpermissive sites) do not support inversion formation (98). When the same recombining sequences are placed in direct order at these nonpermissive sites, recombination occurs and supports frequent duplication formation (102). The failure to observe inversion of some intervals does not seem to be due to lethal effects (78). This suggests that chromosome structure or sequences flanking the recombining sequences may somehow restrict the formation of particular inversions. It should be kept in mind that most experiments on inversions involve the use of sequences added to the chromosome. We do not have a good estimate of the frequency of inversions that would form in a wild-type chromosome by using natural sequence repeats. Aspects of this problem are discussed further below.

Formation of Deletions and Duplications

Because of the formal considerations outlined above, it seems reasonable that formation of spontaneous deletions and duplications might depend on the RecA protein, which promotes homologous pairing and strand exchange and is essential for most homologous recombination events in *E. coli* and *Salmonella* species. Tests of this expectation revealed a more complex situation; the dependence on RecA function varied depending on the nature of the test. When large homologous repeats were provided or when the

locations of possible endpoints were unconstrained so that natural large direct-order repeats were available, duplications and deletions were frequent and their occurrence was heavily dependent on RecA (25, 83). Conversely, when the test system demanded exchanges in particular specified small regions (unlikely to contain extensive direct-order repeats), duplications and deletions occurred a lower frequency and showed only weak dependence on RecA function (37, 56). Thus, it appeared that a second mechanism existed for formation of duplications and deletions. In fact, it seems likely that multiple RecA-independent mechanisms are at work. Assays involving long direct repeats (kilobases in length) and those which depend on much smaller sequence elements are discussed separately.

Deletions between Long Sequence Repeats. Removal of material between large repeated sequence elements is highly dependent on RecA. This is seen when Tn*10* elements (providing 9.3 kb of repeated sequence) were placed on both sides of the *his* operon (25). Deletions of the operon were found at a frequency of 0.5% (15% if the region was on an F′ plasmid), and the frequency dropped 100-fold in the presence of a *recA* mutation. Deletion events between long repeats can also be scored by observing the frequency with which tandem duplications segregate and return to the haploid state. We have recently developed a recombination assay by constructing duplications in *Salmonella* species with a functional *lac* operon fusion at the duplication junction point (37a). Every loss of the duplication (segregation) causes loss of the Lac^+ phenotype and can be scored as white sectors in blue colonies formed on a medium with X-Gal or can be quantitated using fluctuation tests done with unselected liquid cultures. The segregant frequency dropped at least 10^4-fold in a *recA* mutant strain but was essentially unaffected by a *recB* or *recF* mutation, tested individually. The rate of segregation dropped more than 10-fold in strains with both a *recB* and a *recF* mutation, suggesting that these functions contribute to alternative recombinational pathways of deletion formation. (Uses of this system in the study of recombination are discussed below.)

Deletion Formation without Long Sequence Repeats. To pursue the role of shorter repeats in deletion formation, a variety of plasmid assay systems were developed in which two copies of a DNA segment were placed in direct order such that formation of the deletion between the repeats generated an observable phenotype. These assays used sequences varying in length from several hundred bases to over 1 kb. Many of these assays show a relatively small *recA* dependence (5- to 40-fold), suggesting strongly that some deletions are formed without RecA. It should be remembered that in these assay systems, although the dependence on RecA function is reduced, the majority of deletions are still being formed in a RecA-dependent manner (26, 28, 30, 59, 75, 76). In general, in these assays, dependence on RecA decreases with the length of the recombining sequence and the frequency of deletion formation decreases as the repeats are moved farther apart.

Formation of deletions in the bacterial chromosome, without any provided sequence repeats, is reduced very little by a *recA* mutation (37, 56). Under these conditions, endpoints must fall in small regions unlikely to contain extensive natural sequence repeats. This result suggests a mode of deletion formation that is independent of both extensive homologies and RecA function. When such deletion mutations were sequenced, it was found that many (but not all) occur between short, direct-order sequence repeats of about 10 bp (2, 34). Some deletion hot spots appear to be associated with such short repeats. This is consistent with several models for deletion formation described below.

It is not clear how short sequence repeats are used to generate deletions without involving recombination functions, but one attractive possibility is related to the "strand slippage" model for frameshift mutagenesis (106). This is outlined in Fig. 4. The basic proposal is that during replication or repair, a single-strand end becomes unpaired from its complement and pairs with a similar sequence in the immediate vicinity on the template. If synthesis is resumed from the mispaired end, either a duplication or a deletion can be formed depending on the direction of the slippage. Evidence for this model has been reviewed by Ehrlich (32). Supporting evidence for this model is the observation that deletion between direct repeats is stimulated by including an inverse repeat between the direct repeats. (The inverse repeats

could stabilize the unpaired loop by forming a hairpin structure.)

The strand slippage model was further tested by Trinh and Sinden, who constructed a sequence that included three direct repeats; two adjacent copies of this repeat could be sequestered by an inverse repeat (109, 110). When this sequence appears in a single-stranded region of the template, the inverse repeats can form a stem-loop, leaving a single available copy of the direct repeat on one side of the paired stem-loop. If the growing complementary strand with one repeat at its end can slip-pair across the stem-loop in the template and find a free complement on the other side, deletion formation should be favored. Deletions were most frequent when the constructed sequence element was placed in an orientation that permitted the lagging-strand template to form a hairpin and provide the available direct repeat for further replication (109, 110). The results strongly suggest that looping of the sort diagrammed in Fig. 4 (for the leading strand) is most likely to occur on the template for lagging-strand replication. This is consistent with the fact that replication of this template is interrupted by Okazaki fragments and might be expected to be able to provide the ends needed for slipped pairing. These results provide support for the idea that some deletions form in the course of replication.

Lovett et al. have developed a different model for deletion formation that involves recombination but is nevertheless independent of RecA and most other known recombination functions. The model is based on a plasmid assay of the type in which deletions are formed between 787-base repeats (68). Deletion formation is frequent and shows a smaller dependence on RecA (two- to fourfold) than in the plasmid systems described above. When the assay system was placed in the chromosome, it showed a deletion frequency and RecA dependence similar to those observed on the plasmid. When this system is used to select deletions in a *recA* mutant strain, the act of deletion is frequently associated with formation of a plasmid dimer. (Recombination is not predicted by a strand slippage mechanism.) These results suggest that the deletions are formed by a RecA-independent mechanism that makes a full exchange between different copies of the repeated sequence. The dimer-associated deletions do not appear to form by exchanges between two complete plasmids but, rather, by sister chromosome exchanges. A very compelling model is suggested whereby RecA-independent sister strand exchange is stimulated in the single-stranded regions generated by stalled replication fork. Such single strands can pair to initiate an exchange without a need for the RecA strand exchange protein. A variation of this assay system was used to show that the frequency of RecA-independent exchanges is reduced by increasing the distance between recombining sequences (68a). These assay systems, despite the large size of the repeated sequences used, behave in a manner that closely resembles chromosomal deletion formation.

Some deletions form without involvement of direct repeats. These cannot be easily explained by either strand slippage or sister strand exchange. Ripley and coworkers have pointed out the importance of palindromic structures in the formation of many deletions and frameshifts (38, 88). They observed that many deletions without direct repeats at their ends occur in regions that include palindromic sequences capable of forming hairpin structures; these suggest several plausible mechanisms for deletion formation. The inverse sequences could stimulate the strand slippage mechanism if the hairpin structures stabilized the unpaired loops of the type diagrammed in Fig. 4. Alternatively, such a hairpin in the template strand can cause replication to leave out all or part of the paired sequence by template switching while copying the paired region. The latter possibility is similar to the mechanism suggested earlier to explain internal deletions (imprecise excisions) of Tn*10* elements; in these events, material is removed between the inverse flanking copies of IS*10* as described by Foster et al., leaving the outside ends of the inverse repeats at the site (36). There are multiple examples of palindromic sequences that stimulate deletion formation either with or without associated flanking direct repeats (2, 58, 103, 112).

Several additional ideas have been suggested to explain how deletions might form without involvement of the RecA strand exchange protein. A simple possibility is that ends generated by double-strand chromosome breaks are resected, leaving 3′ single-stranded extensions which could be rejoined with the short repeats for initial pairing followed by replication. This might be RecA independent if the pairing partners were complementary single-stranded overhangs generated by exonucleases. Some deletions could

also be generated by errors made by gyrase or topoisomerase in the course of introducing or relaxing supercoils (20, 55, 74, 81). Thoughts on all of these processes have been reviewed (3, 11, 54).

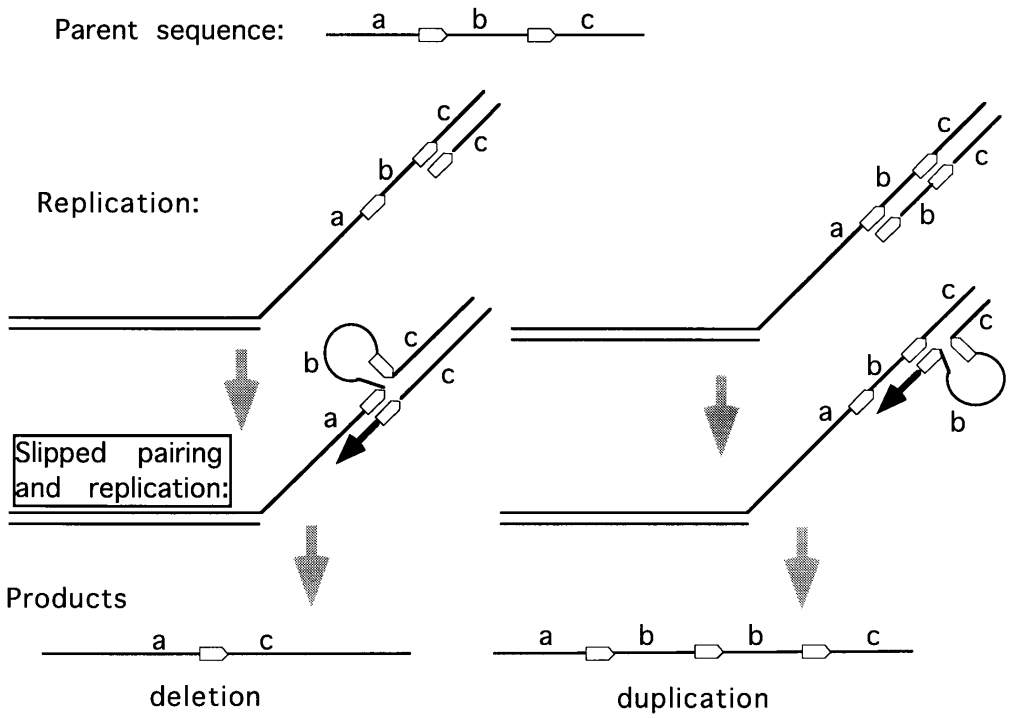

FIGURE 4 Formation of delections and duplications by strand slippage. One model for rearrangement formation without use of RecA protein involves slipped pairing of replication intermediates. The parental chromosome (above) has a natural repeat flanking central element b. If one copy of the repeat is replicated but comes unpaired from the template, it can anneal with the second copy ahead in the template. Resumption of replication leads to deletion formation. Slippage in the opposite direction can lead to duplication formation. Although this is diagrammed here for leading-strand replication, there is evidence that this sort of event is more likely to occur during lagging-strand replication, in which new synthesis is discontinuous, leaving Okazaki fragments that may be prone to slippage before they are joined.

Thus, it appears that deletion formation without RecA is likely to be a complex, multipathway phenomenon. When deletion frequency is measured under conditions that restrict the length of the deletable segment or provide particular recombining sequences, the results are likely to be heavily influenced by the particular sequence used; for many such systems, the results may be difficult to generalize. The possible mechanisms that seem attractive for explaining the formation of short deletions (the types that have been analyzed in most detail) are less attractive as explanations of very long deletions. The strand slippage model and repair of palindromic structures seem less applicable to large chromosomal deletions. Direct ligation of free ends following a double-strand break would require large tracts of degradation to make distant pairing partners single stranded. It is possible that very large deletions will show a higher RecA dependence, but to our knowledge this has not been directly tested. (See, however, the discussion of duplication frequency below.) It should be kept in mind that most of the systems used to analyze deletion formation are on plasmids, which have their own idiosyncrasies such as high copy number and distinct

replication machinery. It will be important to test any model for deletion formation at a chromosomal site and to demonstrate that it applies in multiple sequence contexts.

Mutations That Affect Deletion Frequency. One approach to determining the mechanisms of deletion formation is to look for mutations that increase or reduce the frequency of deletion formation. This approach has the disadvantage that the effects noted might be due to abnormal processing of DNA lesions so as to bias the outcome of repair and thereby reflect an artificial alteration in deletion frequency. Like the mechanistic studies of deletions described above, the effects of mutations on deletion frequency may be peculiar to the assay system used. The litany that follows describes a variety of assay systems and the effects that were noted. It seems important to test these effects in a variety of chromosomal contexts.

Whoriskey et al., using a deletion detection scheme on an F′ plasmid, found a chromosomal mutation which increases the frequency of two particular deletion types on the F′ plasmid but does not increase the frequency of precise excision of Tn*9* (114). The gene affected by this mutation has since been shown to encode a topoisomerase (95). Deletions of this gene may also be responsible for enhanced deletion formation on a different plasmid (115).

Allgood and Silhavy, using a plasmid deletion selection system in which no particular repeats were given, found that mutations in the *sbcB* (*xon*) gene increased the frequency of deletions (4). The affected gene encodes the 3′ single-stranded exonuclease, ExoI.

Imprecise excision of Tn*10,* which involves formation of deletions with one endpoint in each inverse repeat, is stimulated by particular alleles of the *recBC* genes (69). A variety of other repair defects appear to affect deletion in this assay (11, 12, 43, 65, 69).

Formation of Duplications with Long Recombining Sequences. Duplication formation is generally found to be highly dependent on RecA function (29, 48, 84). In typical assays, one is selecting for any event that will provide two copies of a particular locus. This allows detection of large duplications, which can form by recombination between natural large sequence repeats. The known repeated sequences that are major contributors to duplication formation are rRNA (*rrn*) operons (9, 48), the Rhs elements of *E. coli* (66), and the IS sequences of *Salmonella* species (Haack and Roth, unpublished).

Formation of duplications has also been observed by direct selection for exchanges between particular chromosomal sequences. Two mutant copies of the *lac* operon were placed in direct order at separated positions in the chromosome (100). Since each copy of the *lac* operon has a different *lacZ* insertion mutation, the resulting strain is phenotypically Lac^- but can become Lac^+ by recombination between the marked *lac* operons (Fig. 5, left). This strain generates Lac^+ recombinants at a frequency of 10^{-4} (1% of these are duplications; the rest are due to conversion or double exchanges). The formation of all Lac^+ recombinant types is reduced (>1,000-fold) by a *recA* mutation; a *recB* mutation reduces the frequency of total Lac^+ recombinants about 10-fold but does not reduce the frequency of duplications. Thus, taken together, it appears that the formation of both duplications and deletions is highly dependent on RecA function when the assays are set up to use long recombining sequences.

Formation of Duplications without Long Repeats. The possible recombining sites available for duplication formation can be restricted by selecting for juxtaposition of particular sequences at a duplication junction point. This has been done with a deletion mutation that removes the promoter of the *his* operon. Selection for restored expression of a structurally intact gene in the operon (*hisD*) requires formation of any rearrangement that will bring a new promoter near the silent *hisD* gene (Fig. 5, right). In this assay, restoration of *hisD* function requires that the region immediately upstream of the *hisD* gene be joined to a distant sequence that can provide a promoter. The short repeated sequence (REP; 40 bp) lies immediately upstream of the *hisD* gene and is frequently used as the recombining sequence in the events detected by this assay. Another REP sequence, located within the transcript of the *argA* gene, can provide the second recombining site (102). Duplications formed by recombination between these short repeats

places the *hisD* gene under the regulatory control of the *argA* promoter (C. Conner and J. R. Roth, unpublished results).

The frequency of these duplications is low because of the restrictions placed on the possible location of the duplication junction point. With this restriction, the dependence on RecA function is also greatly reduced because some duplication events that express the *hisD* gene can occur in a RecA-independent manner. In a RecA+ strain, all duplications tested had one endpoint within the REP element adjacent to the hisD gene; in a Rec– strain, the frequency of duplications was reduced sixfold and the duplications recovered were not limited to this REP element (7). We conclude that exchanges between the 40-bp REP sequences used are dependent on RecA but occur at a frequency that is close to that of a variety of other, RecA-independent events. Once RecA is eliminated, the second class of events is detectable. It is clear that duplications, like deletions, can form even when no extensive recombining sequences are provided, but under these conditions RecA function is not required. The nature of the RecA-independent exchanges that form between widely separated sites is unclear. Most of the models for RecA-independent deletion formation seem unlikely to apply (strand slippage, sister strand exchanges near stalled forks, palindromes).

Amplification beyond Duplication. When cells are subjected to a selection for increased dosage of a particular gene product, variants in the population which have more copies of a particular gene are selected. These variants have been shown to carry long tandem arrays of gene copies. In several cases, the selected clones have a surprisingly large number of tandem copies of the gene in question. In two carefully analyzed cases, copy numbers in the range of 10 to 100 have been reported (31, 108). In one case, the selection was for increased expression of a *lacZ* gene fused to the *lacI* gene by a deletion. Expression of the *lacZ* function was low because of the weakness of the *lacI* promoter and was further reduced by a *lacI* amber mutation which made translation of *lacZ* sequences dependent on an internal reinitiation site. In another case, selection was for resistance to increasing levels of ampicillin provided by amplification of a chromosomal gene normally present in single copy. In both cases, the amplified segment is small (<50 kb) compared with the duplications usually isolated. Such high amplification seems to restrict the permissible size of the duplicated segment, presumably because the cell could not support such high amplification of large chromosomal regions. This size restriction makes it unlikely that a large repeated sequence element is included, and formation of these amplified arrays occurs with rather short sequence signals in the same size class as those used in deletion formation (31, 113).

It is generally assumed that amplified arrays are generated by a two-step process in which an initial duplication is generated using short repeated sequence elements (113). One would expect this initial step to depend only weakly on RecA function (since it uses short repeats like the formation of spontaneous deletions). However, once a duplication is generated (involving the order of 10 kb of chromosomal sequence), long sequence repeats are present which can support standard RecA-dependent recombination events that increase the copy number of the repeated element. These secondary standard recombination events in long duplicated sequence elements (>10 kb) would be expected to be highly dependent on recombination functions (108). This sequence of events is diagrammed on the left side of Fig. 2. The second step in this process seems difficult, since it would require multiple sister chromosome events, probably occurring over many cell generations.

A more radical possibility for how these arrays might be formed is suggested by work with *Bacillus* species by Petit et al. (85). The proposal suggests that a double-stranded end, generated by a chromosome break, engages in a RecA-mediated homology hunt in the course of attempted repair. During this hunt, the end could loop around and find a short sequence match on its own chromosome, where a replication fork could be initiated. If this fork is directed toward the broken end, it leads to a rolling-circle structure, which could generate a tandem array of copies of the material within the initial loop (Fig. 6). Ultimately, this tandem array would have to recombine back into the chromosome. The experimental feasibility of this sequence of events has been demonstrated (85) and provides a possible mechanism for formation of amplified arrays without the necessity for slow amplification by sister strand exchanges over many

generations. We suggest that this sort of mechanism might be called a “do-loop” by analogy with the computer programmer’s device of repeatedly using the same information.

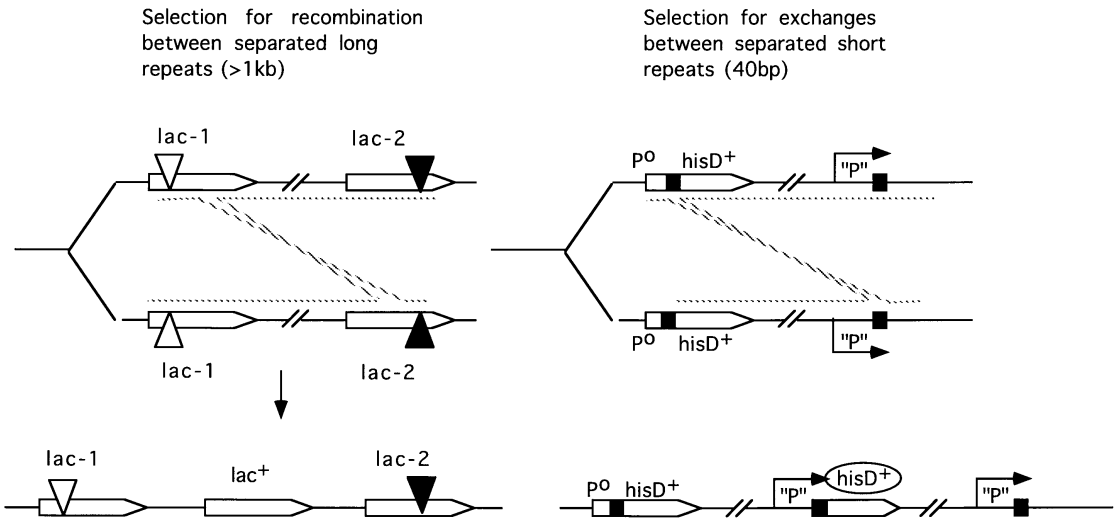

FIGURE 5 Use of novel junction point structures to select duplications. The left side of the figure diagrams a method that has been used to select duplications that form in the process of generating a Lac^+ recombinant by sister chromosome exchange. The chromosome carries two different mutant *lac*Z alleles at separated sites. Unequal recombination between copies on different sister chromosomes can generate a functional $lacZ^+$ allele at the junction point of a duplication; this exchange can occur between the extensive shared sequences between the insertion sites of the mutations used. The right side of the figure shows a selection for duplications that demands exchanges in regions unlikely to include extensive repeats. The parental strain carries an intact *his*D gene that lacks a promoter. Selection for $HisD^+$ recombinants demands that a promoter be brought close to the silent *his*D gene. This can occur by a sister strand exchange between the region upstream of the *his*D gene and the region immediately downstream of an appropriately oriented promoter. This exchange generates a duplication with an expressed *his*D gene at the junction point. Since the exchange must occur between two restricted sequences, it is unlikely that extensive repeats will be shared. In the case illustrated, a small REP sequence is frequently used as the region of homology.

Inversions

As described above (Fig. 3), formation of an inversion appears to be a mechanistically more difficult process than formation of a deletion or duplication. We suggest that this is because inversion requires rejoining of both pairs of sequences flanking the recombining site, while a half exchange (one pair of flanks rejoined) is sufficient for duplication and deletion formation. As outlined in Fig. 3, inversion could be achieved by a single full exchange or by two sequential half exchanges. It is not clear which of these routes is used. Since inversions alter the parental sequence only at the two junction points, surveying the frequency of this rearrangement in bacteria is more difficult than surveying duplications and deletions which add or remove blocks of material. While a few naturally occurring inversions have been identified, characterized, and reviewed (83), most studies of the inversion process have involved adding inverse-order sequences at known sites in the chromosome. These sequences can be designed to provide a selection for

the exchange event involved in inversion formation. As with duplications and deletions, these studies have involved both long-homology and short-homology assays.

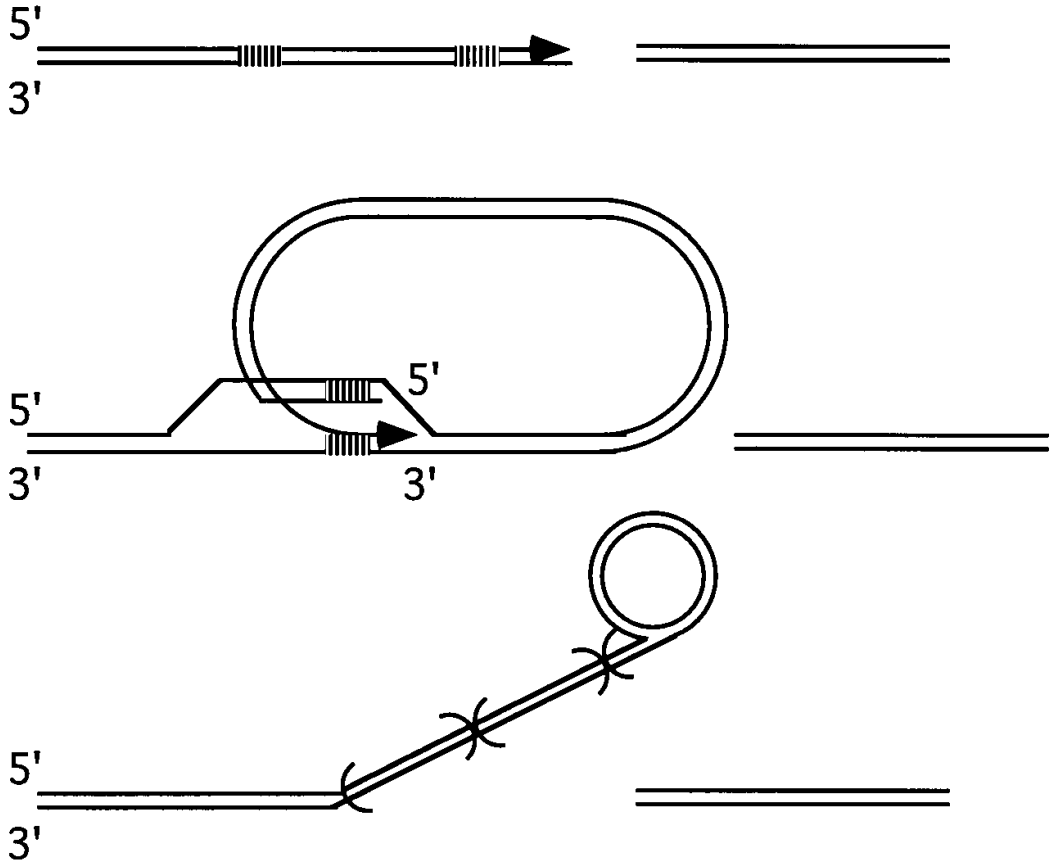

FIGURE 6 Proposal that a "do-loop" might generate long amplified tandem arrays. In the course of trying to repair a double-strand break, an end might engage in a RecA-mediated homology search nearby in the same chromosome. Short sequence repeats (like those used to support delection formation) might allow a replication fork to form. If the fork is oriented toward the end, this forms a rollingcircle structure that would generate a long tandem array of copies. Stable inheritance of this structure would require repair of the end or formation of a circular array of repeats that could be integrated.

Inversion Formation with Long Inverse Repeats. Inversions were first sought by placing different mutant alleles of a gene at separated sites in the same bacterial chromosome and selecting recombinants that restore a functional gene (60, 70, 72, 98, 100, 116). The selected recombinants could then be scored for the presence of an inversion. Two basic schemes for selecting inversions are described in Fig. 7, with *lac* sequences as an example. In the first assay, Lac^+ recombinants can arise by a single full exchange (exchange 1, which causes an inversion) or by double exchanges (exchanges 1 and 2, exchanges 2 and 3) or gene conversion events (which do not lead to inversion). In the second assay (bottom of Fig. 7), double

exchanges and conversions are prevented (by use of a deletion mutation) and the selection is satisfied only by inversion formation.

Nonpermissive Inversion Intervals. Initially, such procedures yielded no inversions when tested with a standard *lac* mutation (at the normal *lac* locus) and a second mutant *lac* allele placed at a distant site in the chromosome (60, 116). Later, this general approach was tried with more sequences at a wider range of chromosomal sites. This led to the surprising observation that in both *E. coli* and *Salmonella* species, the ability to form an inversion depends on the chromosomal position of the recombining sequences (86, 98). That is, recombining sequences at some pairs of sites (permissive intervals) recombine to produce many inversions among their recombinants, while the same sequences placed at other chromosomal sites (nonpermissive intervals) recombine by conversion or double exchange but do not produce inversions among their recombinants.

Two general explanations of nonpermissive intervals have been offered. One is that inversion of these regions is lethal because it disrupts aspects of chromosomal structure important to the cell division process (86). An alternative is that the inversions are not lethal per se but their formation is prevented by some limitation placed on recombination by the chromosomal position of the recombining sequences. While neither alternative has been eliminated, inversions of two nonpermissive intervals have been constructed by two-fragment transduction crosses (78). Strains with an inversion of either of these two different nonpermissive regions are viable, despite the fact that these inversions could not be formed by intrachromosomal exchanges (98, 100). This suggests that (at least for these two intervals) the nonpermissivity is not due to a lethal effect of the inversion. However, it is not clear how chromosome position might influence the recombination process to prevent inversion formation. Recently, mutant strains in which nonpermissive intervals can be inverted by intrachromosomal recombination have been isolated; analysis of these mutants may provide some clues to the nature of "forbidden intervals" (L. Miesel and J. R. Roth, unpublished results).

Effect of *rec* Mutations on Inversion Formation. In the long-repeat assays described above (using a permissive interval), inversion formation is heavily dependent on RecA function, but the situation is more complicated for *recB* mutations. For various intervals, the effect of a *recB* mutation was a 3- to 1,000-fold reduction in the frequency of inversion (71, 99; Miesel and Roth, unpublished). The most modest effects of a *recB* mutation were seen when recombining sequences were not interrupted by active insertion elements, suggesting that these insertions might cause breaks that could be used as substrates for the RecBCD enzyme. This effect of transposons has been noted previously (44). However, when the same sets of recombining sequences were tested with active and inactive inserted elements, no systematic differences were seen. In experiments with recombining sequences with no transposon insertion, *recB* and *recF* mutations alone had modest effects but together caused a 50-fold reduction in inversion frequency (Miesel and Roth, unpublished). Thus, the deletion (duplication segregation) assay and the inversion assays seem to show similar dependence on these *rec* functions.

It is difficult to resolve the question whether bacteria can perform a full exchange (rejoining both flanking sequences), since products that appear to result from a full exchange (circle integration, inversion formation) can also be achieved by a sequence of half exchanges as described above and diagrammed in Fig. 3. The assay diagrammed at the top of Fig. 3 can generate Lac^{+r}ecombinants with or without an inversion. For a permissive interval, roughly 50% of the Lac+ recombinants carry an inversion. Since inversion was not selected, this result suggests to us that the inversion (where permitted) can form without a sequence of rare events. This could reflect the ability of cells to complete a full exchange or could indicate that the ends generated by the first half exchange are extremely active in recombination and have a high probability of stimulating the second exchange.

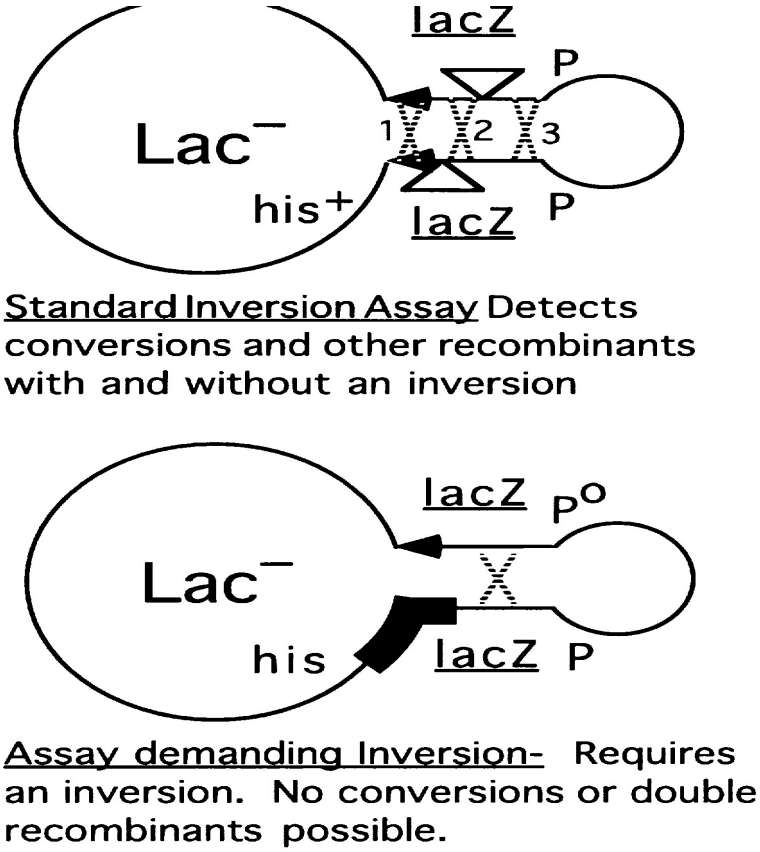

FIGURE 7 Assays for inversion formation by exchanges between provided repeated sequences. The strain carries two different mutant lacZ alleles placed in inverse order. The upper figure is a diagram of the situation if the mutations are sites within the *lacZ* gene. Exchange 2 would generate a Lac^+ recombinant carring an inversion of the chromosomal region between the two copies of the lacZ gene. Paired exchanges, either (exchange 1, exchange 2) or (exchange 2, exchange 3), generate a Lac+ recombinant without at inversion. Gene conversion events that lead to loss of either of the *lac*Z mutations can also generate a Lac^+ recombinant without an inversion. The lower figure shows a modified assay in which one allele is intact but lacks a promoter and the second allele has a promoter but has a distal *lac*Z deletion that extends out of the region of shared homology. This structure does not support formation of Lac^+ recombinants by either double exchanges or conversion events; all Lac^+ recombinants must carry an inversion.

Inversion with Short Sequence Repeats. To examine small inversions occurring by exchanges between specific, short, inverse-order sequences, Schofield et al. (96) constructed an 800-bp inversion extending between the *lacI* and *lacZ* genes with either 12 or 23 bp of inverse-order repeated sequences at the junction points. This Lac^- inversion was placed on an F′ *lac* plasmid, and recombination events that restored the normal orientation (reinversions) were detected as Lac^+ recombinants. The demanded inversions were detected at a frequency of about 10^{-7} in a wild-type strain; this frequency was reduced 1,000-fold by either a *recA* or a *recC* mutation. A *recJ* mutation had no effect.

The above results suggest that the standard recombination system (RecABC) can generate exchanges in extremely short homologies. However it should be kept in mind that these events may depend on the F′ plasmid location of the recombining sequences, since recombination seems to be frequent on this plasmid, as seen by Syvanen et al. (107) and Chumley and Roth (25), and discussed by Whoriskey et al. (114). Since the RecBCD enzyme is thought to act at double-strand breaks (80a), the RecBC dependence of these events and that seen for some inversions formed by using longer inverse repeats raises the question of the source of the double-strand breaks. For the F′ plasmid, the replication origin is a likely suspect. Both single- and

double-strand ends could be generated by mating (between F′ cells) on the selection plate. We have found that F′ transfer functions play an important role in the Rec-dependent reversion of some *lac* mutations when those mutations are present on the F′ plasmid (37a).

MANIPULATION AND APPLICATION OF REARRANGEMENTS

Selection and Construction of Chromosome Rearrangements

Selection of Spontaneous Deletions. Isolation of spontaneous deletion mutants is an art form that usually exploits features of the particular region of interest or requires extensive screening of spontaneous mutants. However, the Bochner selection for tetracycline-sensitive derivatives of Tn*10* insertions has provided an extremely powerful method of isolating deletions of any region for which a Tn*10* insertion is available (18). This method has been upgraded several times to make it more effective (53, 73). When the method is used with a complete Tn*10* element, the bulk of the survivors are "imprecise excisions" of the Tn*10* element, which leave a bit of sequence at the insertion site. Many others are IS*10*-generated deletions that extend from one internal end of IS*10,* removing the resistance determinant and extending into the chromosomal region on the opposite side of the parent element; this leaves one copy of IS*10* at the deletion junction point (36).

The Bochner selection (for tetracycline sensitivity) can also be imposed on strains with the defective element Tn*10*dTc. In this case, the frequency of Tc^s survivors is 10- to 100-fold lower than that seen for a complete Tn*10* element and the deletions recovered include standard spontaneous deletions, arising independently of any function of the insertion element and using non-Tn*10* sequence features in the adjacent regions to dictate deletion endpoints. The fraction of deletions among the bulk of simple Tc^s point mutations is heavily dependent on the particular locus used, suggesting that local sequence features have a big effect on deletion frequency. This method is generally effective and provides a selection for spontaneous deletions that can be applied to virtually any region of the chromosome.

Recently a derivative of phage Mu has been constructed that includes a kanamycin resistance (Kan^r) determinant and a *sacB* (levansucrase) gene from *Bacillus subtilis* (63). When this element is inserted in the chromosome, the *sacB* gene renders the strain sensitive to sucrose and makes this element a deletion-selection device, using the same principles that applied to the defective Tn*10* element. The Kan^r determinant allows isolation of new insertions, as is done with standard Mud transposons, or it can be used to select recombination of the *sacB* gene into available Mud elements having a different antibiotic resistance determinant.

Selection of Spontaneous Duplications. As with deletions, many selections for duplications are based on characteristics of a particular locus that allow selection for increased gene dosage or for some novel feature of the duplication junction point (7, 8). These methods fall into two classes—those that demand two copies of a particular region, and those that demand formation of a novel sequence at the duplication junction point.

One general method of duplication identification is based on transductional crosses in which one selects for simultaneous inheritance of two selectable alleles of the same locus (one from the donor and one from the recipient). When selection is made for only the donor allele, haploid recipient cells normally inherit that allele and lose the recipient alternative; however, rare recipient cells with a preexisting duplication of the locus can acquire the donor allele by recombination into one copy and retain the recipient allele in the other copy. This selection can be made conveniently with any auxotrophic insertion mutation that provides an antibiotic resistance; this procedure is diagrammed in Fig. 8. One selectable allele is antibiotic resistance provided by the donor insertion mutation; the second is the functional allele of the locus, present in the prototrophic recipient. When antibiotic resistance is transduced into the prototrophic recipient, both alleles must be retained for growth on minimal medium containing the antibiotic. Selection

for antibiotic-resistant, prototrophic transductants provides a way of identifying recipient cells with a preexisting duplication of the region in question. The selected recombinants carry a duplication which can be held by maintaining selection for the two dominant alleles (9).

This general procedure also provides a method of mapping the extent of a chromosomal duplication isolated by any means (9). By using the duplication strain as a recipient, one transduces in a series of insertion mutations which confer antibiotic resistance and (in a haploid strain) cause auxotrophy. Antibiotic resistance is selected for, and the auxotrophy of the resulting antibiotic-resistant transductants is scored. If the recipient duplication includes the gene corresponding to the donor insertion mutation, transductants will remain prototrophic because of the second copy of the recipient gene; if the gene lies outside the duplication, the bulk of the transductants will acquire an auxotrophic requirement. Methods of mapping duplications by pulsed-field gel electophoresis have been described (47); the relative resolution of the two methods depends on the number of suitable restriction enzymes compared with the number of available antibiotic-resistant, auxotrophic insertions.

Properties of the duplication junction point can also provide a means of identifying duplications. One method is to start with an intact gene lacking a promoter. By selecting for expression of this gene, any mutation or rearrangment event that can either create a new promoter or bring a preexisting promoter near the gene under selection is detected (7). A variation of this method is to place direct repeats of a gene at separated sites in the chromosome. If these gene copies are mutant at different sites, selection for restored function demands a recombination event that generates a functional allele at the junction point of a duplication. These methods were described above and are diagrammed in Fig. 5.

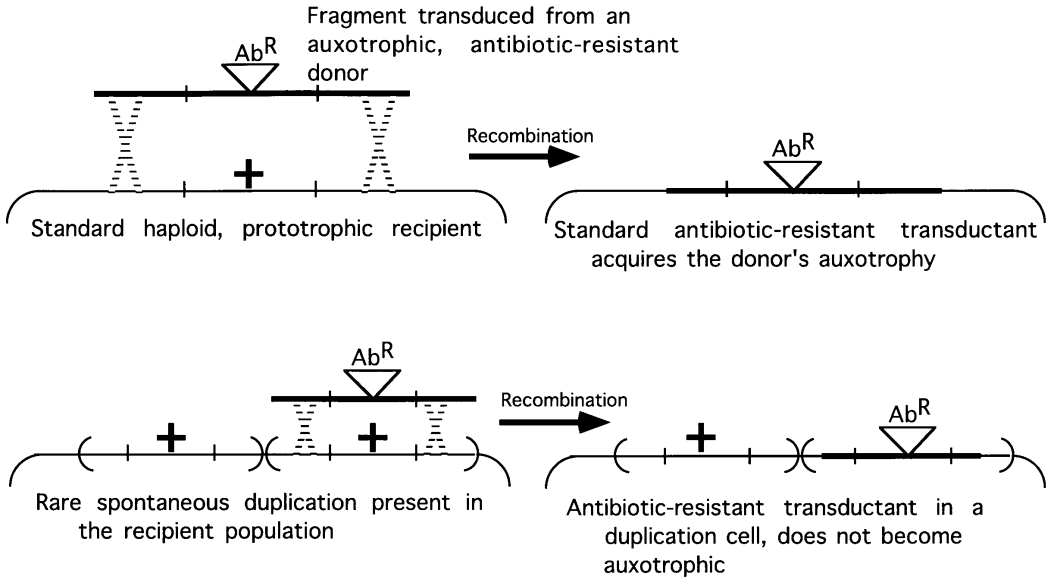

FIGURE 8 Selective "trapping" of spontaneous duplications. The top of the figure diagrams a standard transduction cross in which an auxotrophic, antibiotic-resistant (Ab^R) insertion mutation is inherited by a haploid recipient; the selected resistant transductant acquires the donor auxotrophy. The bottom diagrams the same cross as it would affect rare cells in the recipient population that carry a duplication of the relevant locus. The resistance determinant is inherited by one copy of the duplicated region, but the transductant remains prototrophic because it retains a functional allele of the locus in the other copy. This depends on the fact that the donor insertion confers two pheno-types-a dominant resistance phenotype and a recessive auxotrophic defect.

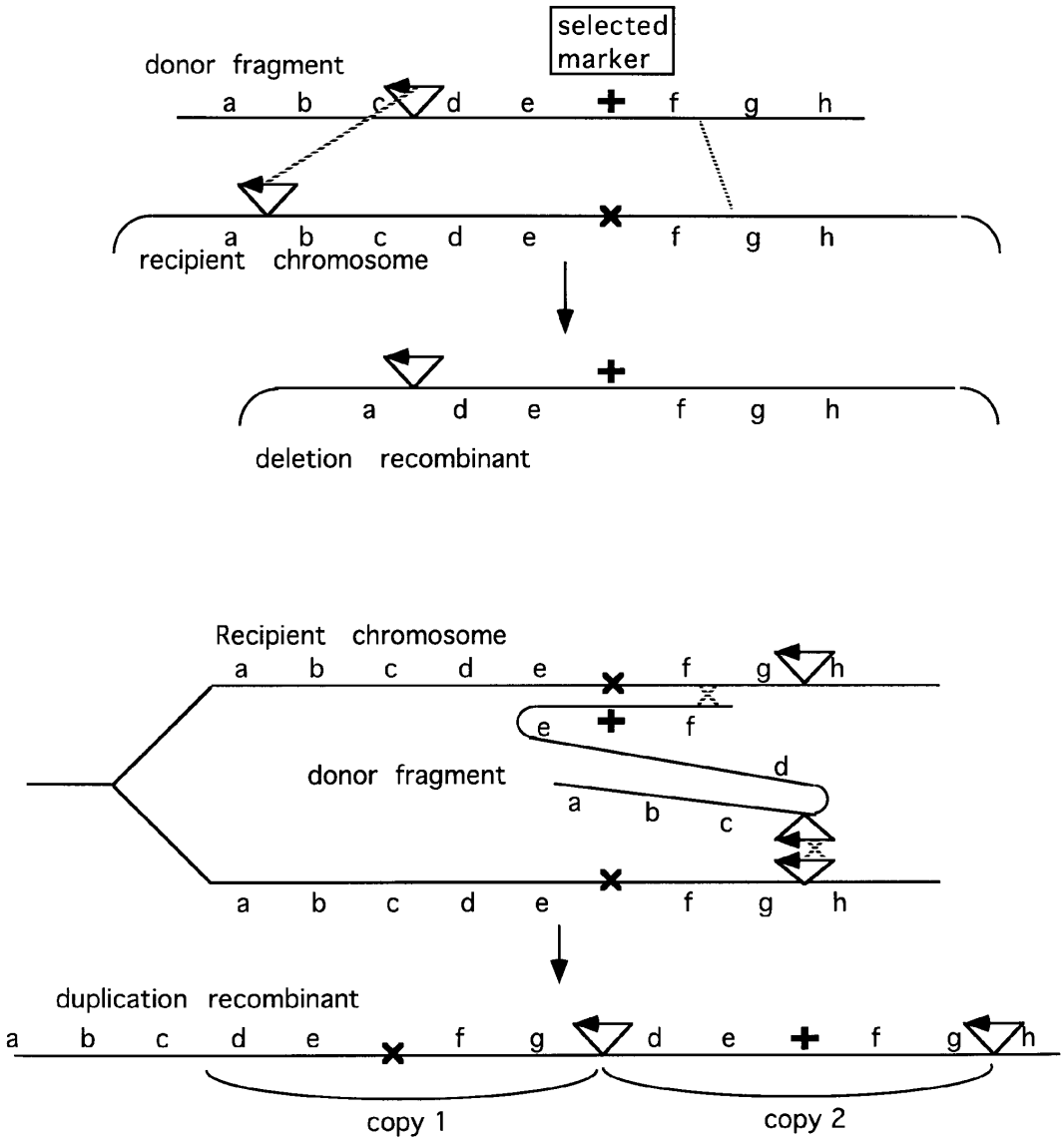

FIGURE 9 Formation of deletions and duplications by using insertions as a portable region of homology. The upper figure shows a transduction cross in which the selectable donor marker has a contransducible insertion. The recipient has ain insertion of the same element at a greater distance from the selectable market. An exchange at the right, combined with an exchange between the two inserted elements, generates a deletion of the chromosomal material between the insertion sites. The lower figure shows a similar cross that generate a duplication. The donor is the same as that use above, with an insertion near the selectable marker. The recipient has in insertion on the opposite side of the selectable marker. If one exchange is in the chromosomal region at the right of the selectable marker and the second is between the donor inserted element and the element in a sister chromosome of the recipient, a duplication is formed that includes the region between the insertion sites.

Construction of Directed Deletions and Duplications by Transduction.

Use of Tn*10* insertions and unequal recombination. Transductional crosses can also be used to construct duplications and deletions with predetermined endpoints. In doing this, transposable elements are used as regions of homology to support unequal recombination; two elements are placed in the same orientation at different points in the donor and recipient chromosomes. The procedure is diagrammed in Fig. 9. At the top is a cross in which the donor carries a Tn*10* element near the selected gene. The recipient has an insertion of the same element (in direct order) on the same side of the selected marker but farther away. Since the length of the transposon is a considerable fraction of the entire transduced fragment, there is a high probability that one of the exchanges that incorporate the selected marker will occur between the two insertions, generating a deletion. It should be noted that the recipient element can be at a substantial distance from the selected marker and could be farther away than the length of a transduced fragment.

The bottom of Fig. 9 describes a cross involving the same donor strain, with a recipient arrangement designed to construct a duplication. In this case, the recipient insertion is on the opposite site of the selected marker from the donor insertion. The cross can be visualized as involving two sister chromosomes. One end of the transduced fragment recombines by chromosomal homology (to the right of the selected marker). The second exchange occurs between the transposon at the left of the selected marker (in the donor fragment) and the recipient transposon at the left of the selected marker (in the recipient chromosome). The donor and recipient insertions must be in the same orientation but can be located very far from the selected marker. The region between the insertion sites is duplicated.

Use of Mud elements in two-fragment transduction crosses. Two incoming transduced fragments from different donors are used to form the junction point of a deletion or duplication (52). The two donor Mud insertions are in direct orientation at widely separated sites in the chromosome. Each introduced fragment carries part of one of the two chromosomal Mud elements. When two such transduced fragments recombine with each other, they can create the junction fragment of a rearrangement (deletion or duplication); this hybrid fragment can recombine into the chromosome. The procedure relies on the fact that the Mud1 transposon and its transposition-defective derivative, MudA, are so large (39 kb) that a single P22-transduced fragment (44 kb) rarely includes the entire element. The element is therefore usually transduced (by using P22 at high phage multiplicity) by two cooperating transducing particles which must recombine with each other to effect the transduction of even a simple element (52). This is diagrammed at the top of Fig. 10.

The middle of Fig. 10 shows a cross that uses one fragment derived from the outside half of each of two different insertions. Coinheritance of these fragments generates a deletion of the intervening material. The bottom of Fig. 10 shows a cross that uses the two inside fragments from different insertions. These are coinherited to generate a duplication of the intervening region. In all crosses, two overlapping fragments must be coinherited and recombine with each other. Thus, if the size and concentration of each fragment type are the same, all three types of transductants will occur equally frequently. In the cross, one selects for the antibiotic resistance encoded by the MudA element (Ap^r). If the two donor insertions are both auxotrophic, a simple inheritance of one Mud causes one of the auxotrophic requirements. Formation of a deletion generates both auxotrophic requirements plus any additional phenotype of the deletion. Formation of a duplication gives an ampicillin-resistant transductant that has neither auxotrophic requirement. Thus, once a collection of insertion mutations is available, duplication construction is operationally simple.

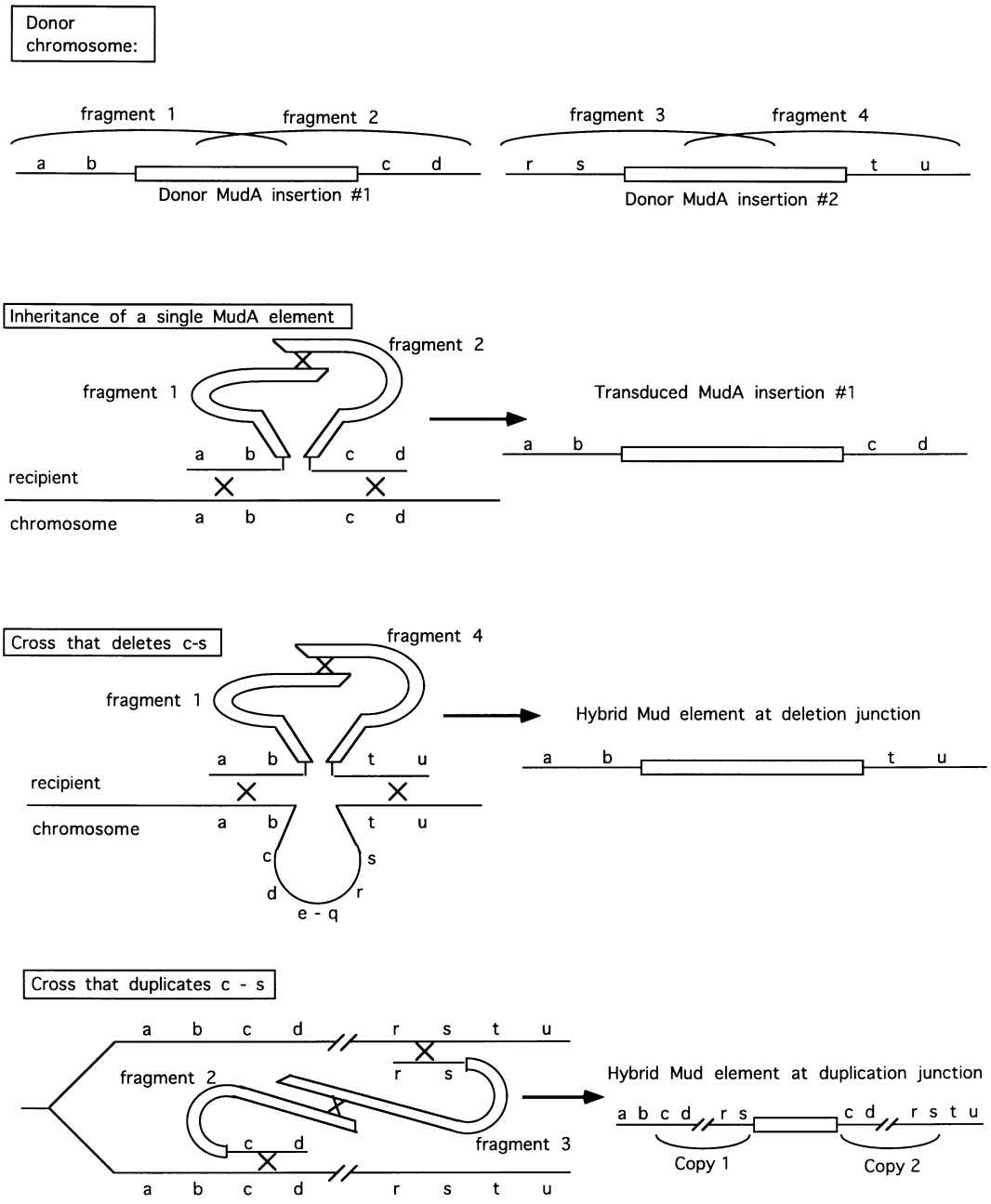

FIGURE 10 Directed formation of deletions and duplications by two-fragment transductional crosses. The top line of the figure shows two different insertions of the MudA element (open boxes) in the bacterial chromosome. The second line shows inheritance of one insertion by a two-fragment transduction; two fragments are needed because the MudA element is large (39 kb) compated with the size of a P22 transduced fragment (44 kb). The third line shows formation of a deletion by use of two fragments, one from each donor MudA element; these can be provided by different transducing lysates. The fourth line shows formation of a duplication by a two-fragment transduction involving two sister chromosomes of the recipient.

Transduction of Large Duplications into a New Recipient. Events very similar to those involved in duplication construction make it possible to transduce a large duplication from one strain to another. This is possible because the duplication junction sequence is the only element unique to this rearrangement. If the junction fragment can be incorporated into a transducing virion, its inheritance in the recipient chromosome reconstructs the duplication. In doing this, only the junction fragment is inherited; the bulk of the sequences of the inherited duplication are derived from recipient sister chromosomes. This cross is diagrammed in Fig. 11. Duplication-bearing recombinants arise at frequencies very close to those expected for inheritance of the same selective marker with standard flanking sequences (6). Therefore, the cross can also be used to measure transductional linkage between a junction marker and any mutation present in the recipient chromosome.

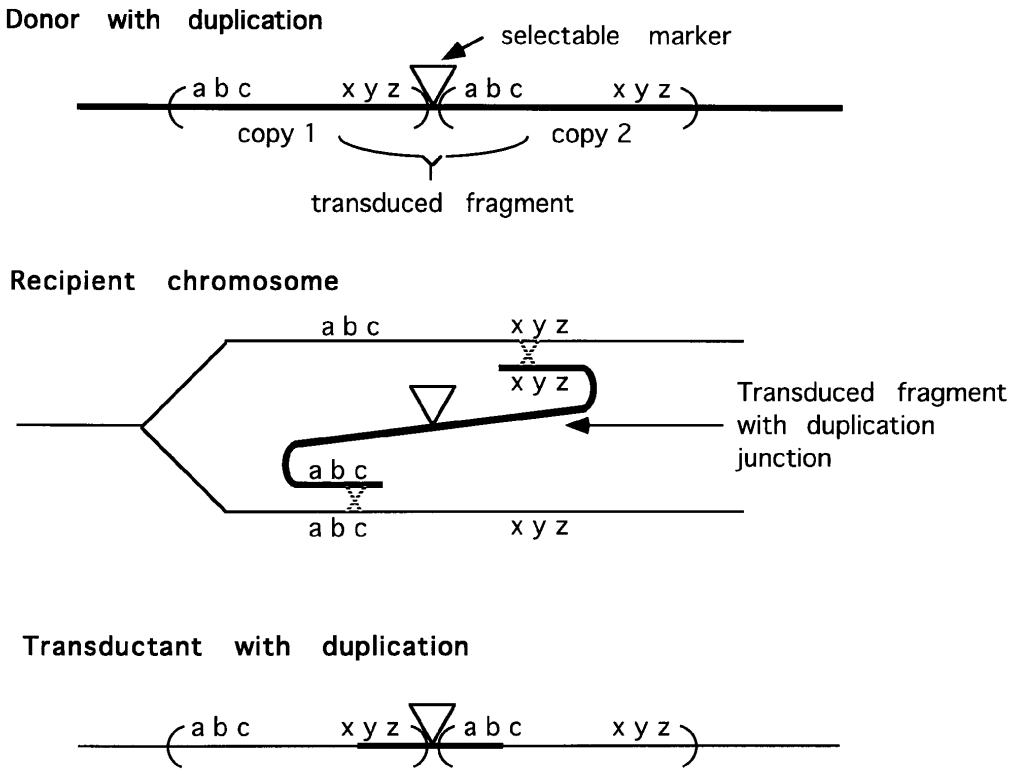

FIGURE 11 Transduction of a large duplication into a new recipient. At the top is a donor strain carrying a duplication with a selectable marker at the junction point. The transduced fragment includes that selectable marker. In the middle is a diagram of how this junction fragment could be inherited between two recipient sister chromosomes. This process reconstructs the duplication in the transductant. Note that the donor provides material only at the junction point (designated by a heavy line). The resulting duplication can be much larger than the transduced fragment, since most of the material is derived from the two sister chromosomes of the recipient.

Selection of Spontaneous Inversions. Inversions have been selected by a modification of the operon fusion method described above for duplications. Selection is made for expression of an intact gene lacking a promoter. This requires that a new promoter be either constructed or moved near the silent gene. Several

inversions have been made by this method (93, 94). Selected clones were screened for possession of an inversion. Some inversions were identified by the fact that the revertant phenotype could not be transduced into a new recipient (because of linkage disruption at the relevant inversion junction point). One inversion was identified because of a null phenotype generated by gene disruption at the second junction point. In this case, a silent *hisD* gene was expressed by an inversion that fused it to the *purB* gene promoter; this inversion disrupted the *purB* gene and caused purine auxotrophy.

Construction and Repair of Inversions by Two-Fragment Transduction. Construction and repair of inversions by transduction is possible but is more complicated than the procedure outlined above for duplications. Since an inversion has two sites at which the chromosomal sequence is abnormal, creation or repair of an inversion requires the simultaneous involvement of two transduced fragments, one for each junction point. This two-fragment transduction is possible with the high-transducing mutants of phage P22. The procedure for this is diagrammed in Fig. 12. For more details of these crosses, see references 78 and 93.

Use of Duplications in Genetic Analysis

Complementation and Dominance Tests. Because of the haploid nature of bacteria, complementation tests are the most difficult aspect of formal genetic analysis to perform on a routine basis in these organisms. Classically, this has been done by using plasmids to create merodiploids that can allow the tests that sort mutations into individual complementation groups. The disadvantage of classifying large numbers of mutations in this way is the labor involved in constructing all the needed plasmids. Other difficulties arise in verifying the presence of both mutations in the diploid used for the test. The most difficult problem is that of copy number. Many complementation tests are strongly affected by the relative copy numbers of the two alleles being tested. Tandem duplications can provide a solution to these problems.

A tandem duplication provides the two alleles in normal low copy number and allows the diploid to contain one copy of each allele. Once the duplication is constructed and its phenotype is tested, the genetic structure can be verified by looking at the phenotype of haploid segregants, which will include both of the haplotypes present in the duplication.

A recent example of this application of duplications can be seen in the *cis*/*trans* tests of regulatory mutations for the *cob*/*pdu* regulon (1, 17). In these tests, the duplications carried a *lac* reporter insertion in a Mud element at the junction point of a duplication constructed by both of the two methods outlined above. The duplication state was maintained by selecting for ampicillin resistance (Ap^r) or kanamycin resistance (Km^r) determinants associated with the Mud element at the junction point. Mutations to be tested could be placed *cis* (upstream) of the fusion or in the other copy of the duplicated region, located far downstream of the insertion. This test gave clear, consistent results for a system in which the available plasmids gave highly irregular results, probably because the system is extremely sensitive to small differences in the level of its positive regulatory protein. Constructing the diploids was simplified by using insertion mutations which can be selectively transduced into the duplication. The position of the insertion (*cis* or *trans* to the fusion) was assessed by measuring the relative frequency of the two haploid segregants. A large duplication was used such that mutations of interest located *cis* to the fusion are close to the junction point and are most frequently lost when the duplication segregates. Mutations in *trans* to the fusion are at the most distant end of the other copy of the duplication and are relatively rarely lost when the duplication segregates.

Another recent use of duplications was in dominance testing of alleles of the *sbcB* gene, a suppressor of *recBC* mutations. It has long been argued that this gene has multiple functions, only some of which are eliminated by the standard alleles of the gene, or that these mutations might have some novel dominant phenotype. Distinguishing subtle dominance effects can be difficult if either allele is expressed from a high-copy-number plasmid. In the duplication test, it could be shown that the phenotypes of insertion mutations of the *sbcB* gene (presumed to be null alleles) are completely recessive to the wild type for all

phenotypes. The more effective point mutations are recessive to the wild type for suppression of the recombination defect of *recBC* mutations but are codominant with the wild-type allele when tested for correction of the UV or mitomycin sensitivity phenotype. This provides evidence that *sbcB* point mutant alleles do in fact produce a protein that contributes to the cell phenotype (13).

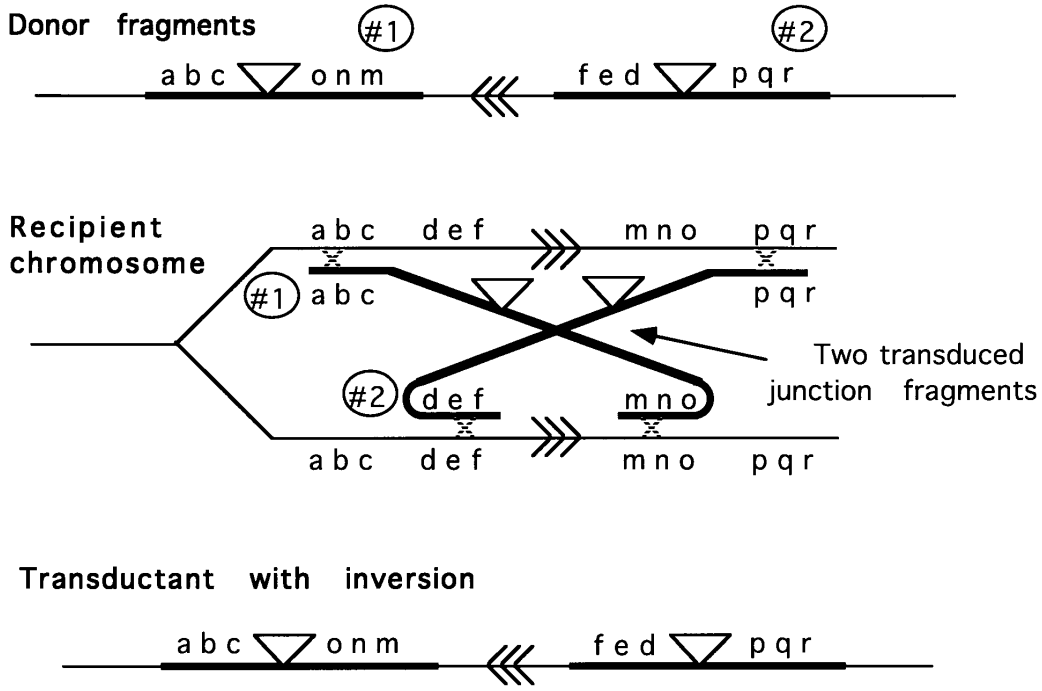

FIGURE 12 Formation or repair of an inversion by a two-fragment transduction. The top part of the figure presents the chromosome of a strain with an inversion. Triangles indicate selectable markers at the inversion junction points; letters designated base sequences of flanking regions. In the middle is a diagram of events that would occur if selection were made for inheritance of either of the donor's junction markers. Inheritance of either marker requires cooperation of a second fragment carrying other markers, to restore an intact chromosome. (Use of either fragment alone causes a chromosome break.) The result is a transductant that acquires both junction markers and an inversion of the intervening region. Similar events can repair an inversion if the donor is wild type and the recipient carries an inversion.

Determining the Direction of Transcription of a New Gene. The two-fragment transduction crosses that form duplications (described above) provide a reliable way to determine the direction of transcription of a new Mud-*lac* insertion mutation. Since formation of a duplication requires that both participating Mud insertions be in the same orientation, one can test the unknown Mud insertion for the ability to form duplications in combination with a series of Mud insertions (at other loci) whose orientation is known. This allows one to assign an orientation by a single set of transduction crosses (52). It is even possible to assign orientations to transcripts by this method when the orientation of neither of the Mud insertions is known;

this requires knowing the relative positions of the two insertions in the chromosome. This has been described in detail elsewhere (40).

Experimental Situations in which Duplications Cause Problems

The high frequency with which duplications occur in bacteria can cause problems in some experimental situations. Whenever a selection condition is imposed that can be circumvented by amplification of a normal gene, duplications are likely to be the major "revertant" type that appears. This is fine if duplications are desired, but it can be a serious problem if rare point mutations with an altered phenotype are being sought. Increasing the severity of the selection may minimize the problem by preventing detection of simple two-copy duplications, but, as described above, the Lac^+ selection of Tlsty et al. (108) yielded amplification revertants with more that 20 tandem copies of the selected allele. These appeared at a frequency of about 1 per 10^7 plated cells but may have been common in part because the parental Lac^- mutation is leaky and cells grow slowly on the selection plate. Use of a *recA* mutation may reduce the frequency of duplications and (on the basis of the Lac example above) would be expected to prevent the high level of amplification.

Whenever a high-resolution transductional cross is done in which the mutations being crossed can complement, duplications will be recovered among the recombinants. This was exploited as an early method of duplication detection (5), but it can interfere with analysis if it is not expected. Crosses were made between very closely linked, complementary *his* mutations. Standard haploid His^+ recombinants were expected to be rare because of the proximity of the donor and recipient mutations and the necessity of a recombination event between them. Rare cells in the recipient population having a duplication of the *his* region can be transduced to His^+ by unrestricted (high-frequency) exchanges that replace the recipient mutation in one copy of the duplication with the closely linked donor mutation. The result is a prototrophic recombinant that has complementary donor and recipient mutations in different copies of a spontaneous duplication of the *his* region. These recombinants arise without an exchange in the small region between the two mutations. Their frequency is limited by the frequency of duplications in the recipient population.

The above situation arises frequently in the use of two insertion mutations with different associated antibiotic resistances. In trying to construct a doubly mutant strain, one may cross one insertion into a recipient with a nearby insertion of different resistance. The "true" recombinants (with both closely linked insertions) may be rare. However, recipient cells with a preexisting duplication can inherit the donor antibiotic resistance without restriction. The donated antibiotic resistance replaces the recipient antibiotic resistance in one copy of the duplicated region. This leaves a transductant with both resistances, which looks as if it occurred by an exchange between the two elements; in fact, it carries a duplication with one element in each copy of the duplicated region. This can be a serious problem when two insertions are used for fine-structure mapping. An example of this problem and how it was resolved is described in the appendix to reference 17.

Rearrangements as an Assay System for Recombination

Most genetic analysis of the process of recombination has been done with sexual crosses. In these crosses, a linear chromosome fragment is introduced from a donor strain (by conjugation, transduction, or transformation) into a recipient cell. The process of recombination is monitored by selecting for incorporation of some donor marker by replacement of the corresponding region of the recipient chromosome. While this procedure has been successful in revealing many genes that affect the recombination process, the phenotypes of these mutations suggest that some of them may normally serve a very different purpose from that required for integration of a linear fragment. Only *recA* and *recBC* mutations were identified in the initial hunts. Most other classes of *rec* mutations have little or no phenotype alone and were revealed by the defect they cause in strains with a *recBC* mutation and one or

more suppressor mutations (*sbcABCD*). It is sometimes argued that these extra *rec* genes contribute to subsidiary pathways of recombination that are activated or amplified by the suppressors. It seems strange that most recombination functions are silent in a wild-type strain, and it has been suggested that they normally act on substrates other than the recombinogenic ends provided in the standard sexual recombination assays. We suggest that these assays, by providing a double-stranded end, may have biased our view of recombination in favor of events that are initiated by such substrates. We may have underestimated the importance of functions that act at other initiating structures (e.g., nicks, gaps, and loops). Perhaps the *sbc* mutations serve to allow these alternative pathways to act at the unfamiliar double-stranded ends provided in the sexual recombination assay.

It is our contention that some of the problems inherent in sexual recombination assays might be circumvented by use of chromosomal rearrangements to score recombination events. These recombination assays involve exchanges between two chromosomal sequences and provide no initial structures to stimulate exchange initiation. In these assays, a variety of initiating structures are probably generated spontaneously and are acted upon by functions that contribute to multiple normal paths of recombination. Thus, these assays may allow visualization of new aspects of recombination. It should be noted that recombination in eukaryotes is studied under conditions that also require any initiating structures to be generated internally. Initial results of this approach are promising. While *recBC* mutations reduce sexual recombination about 100-fold, they have very little effect on the rate of duplication segregation (70, 71, 97, 100). In a *recBC* mutant, an added *recF* mutation causes a reduction (>10-fold) in duplication segregation. This suggests that the RecF function is contributing strongly to duplication segregation in a *recBC* strain without requiring a suppressor (*sbc*) mutation to activate the alternative pathway. The RecBC and RecF pathways appear to be acting simultaneously to cause segregation, perhaps by using different initiating structures (37a). The duplication assay has been used to identify new *rec* mutations, and initial results suggest that a novel set of genes have been identified; these may contribute to the formation of initiating structures (37a).

The use of inversions for a recombination assay may provide a feature that is distinct from the duplication assays. Inversion formation may occur most frequently by a full exchange in which both pairs of flanking sequences are exchanged. Duplication formation and segregation formally requires only a half exchange (joining one pair of flanking sequences). It is possible that these distinctions are reflected by differential effects of *rec* mutations on the two process. Thus far, only *recA, recB, recC, recF,* and *recJ* mutations have been tested in both assays, and they show similar effects on both duplications and inversion (37a; Miesel and Roth, unpublished).

IMPLICATIONS OF REARRANGEMENTS FOR POPULATIONS AND EVOLUTION

The previous sections of this review have described the fundamental properties of rearrangements and how they can be manipulated in the laboratory. Presumably, all of the events observed in the laboratory can also occur in natural populations. Below, we discuss some more global ramifications of chromosome rearrangements.

Selective Value of Duplications

Amplification as a Regulatory Mechanism. Duplications are common bacterial responses to selective pressure under laboratory conditions (31, 105). It is difficult to escape the possibility that the same response will prove selectively valuable under stressful conditions encountered in nature. The basic idea is that standard gene regulatory mechanisms allow cells to adjust their metabolism to the range of conditions encountered most frequently. When extreme conditions cannot be accommodated by gene regulatory mechanisms, selection is imposed for increasing the copy number of a gene or set of linked genes that can improve growth. Duplications arising in this way mimic a regulatory response in that they are reversible.

The cell is not committed to its new genotype, because duplications are held at high frequency in the population only as long as they provide a selectable phenotype. When growth conditions change and the duplication is no longer selected, haploid revertants again take over the population. Thus, duplications might be viewed as a primitive regulatory mechanism whose specificity is determined by natural selection.

While this notion is difficult to test directly for natural populations, it seems to work well under laboratory conditions that seem to be reasonable approximations of the natural situation (35, 47, 50, 90, 105). If duplications are selectively valuable in the real world, some general predictions are made that might be subject to testing. A gene whose amplification is frequently advantageous might be found with local flanking sequences that would provide homology for increasing the frequency of that duplication. Sets of genes that act synergistically when amplified might be found grouped with direct repeats flanking the entire group. These predictions are fulfilled in a case which is probably the best example of this phenomenon, the amplification of the cholera toxin gene during the course of a mouse infection (77). In this case, the amplification is selected by conditions in the host (a real-world situation). The duplication is formed by RecA-dependent exchanges between repeated sequences that are found to flank the toxin gene that is subject to amplification (39).

In addition to the variation in gene copy number provided by duplications, the actual expression level of a gene and the nature of its control mechanism can be changed by a duplication that fuses a gene copy (at the junction point of the duplication) to a novel promoter on the other side of that junction. This possibility has been discussed above in the section on formation of duplications and methods of duplication detection (Fig. 5). This sort of regulatory change could be selectively valuable in a natural population, since a reversible rearrangment could temporarily provide a gene with increased expression or altered regulatory specificity. An example of this event has been seen in *Salmonella* species; exchanges between REP elements generated a duplication whose junction point fused the *hisD* gene (and the rest of the *his* operon) to the transcribed region of the *argA* locus (7, 102). This placed the *his* operon under the control of the *argR* repressor (Conner and Roth, unpublished). Rearrangements of this sort (like the amplification described above) are reversible, since segregation events cause loss of the duplication junction site and restore the normal haploid chromosome and transcription control to its normal specificity. So far, this mechanism has been seen only in a laboratory situation.

Transposable Elements as Regions of Homology. Insertion sequences are among the direct repeats that contribute heavily to formation of large spontaneous duplications by providing sites for unequal recombination. In *S. typhimurium,* species, four (of the six) copies of *IS200* are almost evenly spaced in the half-chromosome segment counterclockwise of the replication origin. All of the IS*200* elements in this region are in the same orientation, and roughly half of the duplications of sites in this region are formed by unequal recombination between flanking pairs of IS*200* copies (Haack and Roth, unpublished). In *S. typhimurium,* the IS*200* elements appear to be closely conserved, since they recombine readily. If duplications generated by recombination between elements are frequently of selective value, the movement of IS elements within the genome might provide a way of tuning the repertoire of most frequently duplicated regions. This would impart a positive selective value to IS elements and to their transposition ability.

We think it likely that duplications generated by recombination between IS elements may be responsible for the apparently unstable genomes reported in some bacteria (14, 15, 35, 64, 91, 92). In these experiments, independent clones derived from a common parent culture show a different pattern of genome fragments in Southern hybridizations probed with a repetitive element. The genome appears to be rearranging rapidly, or the element is transposing very frequently. We submit that the new bands observed may represent the novel fragment generated at the junction point of a duplication formed by recombination between repeated elements. The lost bands may represent duplications in the parent strain that were lost in the new subclone. We know that duplications are continuously forming and segregating in unselected populations of *Salmonella* species. If the same process occurs with higher frequency in other organisms, it

would predict the observations reported. This could occur without transposition and with a completely conserved basic genetic map.

Evolutionary Stability of Genetic Maps

A striking feature of the chromosomes of *Salmonella* species and *E. coli,* the organisms to which this volume is dedicated, is the degree to which their genetic maps have been conserved. This conservation is particularly impressive when one considers that the rearrangements discussed here occur in the laboratory at easily detectable frequencies and that duplications, in particular, are extremely common. This suggests that natural selection has operated to conserve the genetic map and to select against variants with rearrangements of this gene order. What selective forces have directed formation of the map we see today and might be acting to conserve it?

There is a tendency for genes to be oriented such that their transcription and replication proceed in the same direction (19, 21). Brewer (19) has suggested that this tendency might have been driven by the fact that a replication fork can be stalled by an opposing transcription complex. Although large inversions reverse the orientation of many genes with respect to the direction of replication without lethal consequences, the possibility remains that these inversions cause subtle reductions in growth rate that could represent a potent selective force under natural conditions. Hill and Gray have shown that an inversion of the origin region in *E. coli* causes a slight decrease in growth rate (49). Many of the *E. coli* inversions described by Rebollo et al. (86) impair growth. In *Salmonella* species only 1 of about 30 different inversions tested showed a major growth defect, but these inversions were not tested in sensitive competition experiments (98).

Another force for map conservation might reflect the fact that the pattern of bacterial DNA replication increases the dosage of genes located near the origin of replication. These genes are replicated early and are present at a higher dosage (averaged over the cell cycle) than are genes near the terminus. The strengths of all promoters may have been evolutionarily "optimized" in the light of this dosage difference. An inversion would change the position of many genes with respect to the origin of replication and thus would change the average dosage of many genes. While these changes may be small, they would affect a large number of genes and might be a significant selective force for conserving chromosome order. This idea would suggest that the order of genes is physiologically unimportant but that once it is established and promoters are tuned, any alteration of that order would be deleterious.

A striking feature of the chromosome is the observation that the Chi sequences, which activate the RecBCD recombinase, are present at a frequency that is in 60-fold excess to the expected frequency on the basis of base composition. More striking, perhaps, is the fact that 80% of these sequences are oriented such that they would activate a RecBCD complex moving toward the origin of replication from either side (21). The degree to which the Chi sequences are preferentially oriented reflects approximately the degree to which they are overrepresented in the genome. If the position of these sequences is important for the orderly repair of chromosome damage, rearrangements that reverse this bias might be counterselected.

The chromosome may include additional strategically placed sequence features that help in the orderly folding, replication timing, and partitioning of the chromosome. If this is so, large rearrangements might alter the relative position of these sequence features to the detriment of the ability of the cell to grow. Such sequences could act as a conservational force on chromosome structure without regard to the position or orientation of protein coding regions.

In view of the impressive conservation of the maps of *S. typhimurium* and *E. coli,* it is surprising to note the recent finding that the genetic map of *S. typhi* is substantially rearranged compared with that of *S. typhimurium* (66a). The rearrangement is not a complete shuffling of genes but appears to be due only to recombination between *rrn* loci. As molecular methods (e.g., pulsed-field gel mapping) are used to assemble maps of other enteric bacteria, it will be possible to see how general the map conservation is and may suggest new interpretations of the map conservation seen for *S. typhimurium* and *E. coli*.

Implications of Rearrangements for Evolution of the Recombination System

Recent work on the population structure of bacteria suggests that many bacteria, including *S. typhimurium* and *E. coli,* grow as substantially clonal populations with infrequent sexual episodes in which small blocks of information are transferred (101, 104). The frequency of transfer of a particular region may approximate the frequency of mutations (42, 79). The low frequency of sexual exchanges (and the low viability of *rec* mutants) suggests that the recombination system is most heavily used for DNA repair and perhaps adaptive duplication. Thus, it may be selected mainly for its action on internal substrates and more weakly for its role in sexual exchange.

The paragraphs above suggest that the recombination system might be under conflicting selection. In the short term, recombination is needed as an important DNA repair function and to form selectively valuable, reversible duplications. Over the long term, selection seems to have operated (regardless of the functional basis) to maintain gene order. This would put a secondary selection on the recombination system to minimize chromosome rearrangement. We propose that inversions are the main contributor to map instability. Inversions are not known to be advantageous and are poorly reversible. (Reversal of an inversion is less likely than occurrence of a secondary inversion, using different inverse repeats of the same size.) Thus, minimizing inversion formation might contribute to map stability.

The conflicting pressures on the recombination system could be resolved if the system could distinguish between direct- and inverse-order substrates. For the purposes of repair and duplication, it must favor homologous sister strand exchanges (needed for repair), including exchanges between separated direct repeats (formation and resolution of duplications). To minimize inversion formation, the recombination system might be under selection to avoid full exchanges between inverse repeats. We suggest that the rarity of inversions and the finding that many chromosomal regions are mechanistically nonpermissive for inversion provide evidence that the recombination system can make some distinction between direct and inverse repeats. We do not know how this distinction is made.

Selective Forces That Drive Formation of Gene Clusters

While the chromosome gene orders of *S. typhimurium* and *E. coli* seem evolutionarily stable, it is clear that on a larger timescale, chromosome rearrangements have played a large role in the evolution of these chromosomes. The prominent clustering of functionally related genes in the bacterial chromosome is evidence that extensive rearrangements have occurred and formed the gene clusters that we observe today. While one can suggest selective values for the final groupings, it is more difficult to envision how selection might operate to drive the series of intermediate approximations that are likely to have preceded the present arrangement.

Genes in a particular cluster usually encode proteins that catalyze mechanistically diverse reactions which together serve a common function or contribute to a phenotype. An example is the set of genes whose proteins catalyze the reactions required for synthesis of an amino acid from a common metabolic substrate. It was once thought that the proteins in a gene cluster might have evolved from a common precursor (51). Very few cases support this idea, but some do exist (33). In general, sequence comparisons suggest that proteins in a single pathway are unrelated. Instead, proteins with a common detailed mechanism of action show similarity regardless of the more general function to which each contributes (111). Thus, the clustering of genes observed in bacteria appears to be a derived state. That is, each gene in the cluster initially arose from a copy of a gene encoding a mechanistically or structurally similar protein acting in a different context. We suspect that new genes in a pathway were initially formed at widely scattered sites in the chromosome and were later brought together by rearrangements to form the operons that are characteristic of the present-day organisms. The ability to synthesize the end product of a pathway is not likely to depend on the map positions of the genes encoding biosynthetic enzymes. Therefore, one must find some other way of accounting for the selective forces that drive assembly of gene clusters.

It has been suggested that economies achieved by regulation of gene clusters (the operon) might provide a selective value (57). The problem with this idea (and other ideas of the physiological value of clustering) is that the benefit provided by the clustering could not be realized until the relevant genes are finally and precisely placed. This would require that an astronomically rare act of suddenly juxtaposing two genes and fusing their transcripts would have to provide an extremely high selective value such that the rare event would be quickly driven to fixation. We find it much more likely that genes are brought together and eventually fused into operons by a slow multistep process. This requires a different sort of selective force—one that can operate on all the intermediate series of approximations that ultimately lead to operon formation.

We suggest two models, either or both of which may provide the selective force for progressive gene clustering (J. Lawrence and J. R. Roth, unpublished results). For both models, the advantage of clustering increases as the genes are brought closer together. The first involves the use of duplications as a primitive regulatory mechanism (see above). If a block of functionally related genes are frequently amplified in response to selection, this will provide a selective force for clustering, since the closer genes are together, the more likely it is that they will be amplified together. This might provide a selective force for clustering even in a strictly clonal population.

A more potent influence on gene clustering may involve the transfer of chromosome segments between cells of the same or different species. Most bacterial phenotypes are only occasionally placed under selection. During the long periods of relaxed selection, the genes contributing to a function can accumulate mutations and clones of mutant cells might become very large. As soon as multiple mutations occur, the function is effectively lost and can be reacquired only by mating with a cell which can provide a new version of the information. When the mutations affect different genes, repair is more likely if the donated genes are close together and can be simultaneously introduced by a single transferred fragment. The more closely clustered the genes are, the more likely is such multigene repair (45; Lawrence and Roth, unpublished).

Another view of this scenario is one we call the "selfish operon." Imagine a selectable phenotype that depends on three genes. The immediate fitness (growth ability) of an organism might be the same regardless of whether these three genes are clustered or widely separated. Transfer of this phenotype to a naive cell requires transfer of all three genes. (While individual genes might be transferred singly, they would not confer a selectable phenotype and are unlikely to persevere in the population.) The clustered version of these genes is more likely to be transferred to organisms that never possessed the function or have lost the phenotype by mutational loss of multiple functions. Clustering facilitates cotransfer because prominent mechanisms of bacterial mating involve transfer of small genome fragments. Since the clustered versions are more portable, they are more likely to spread horizontally among bacteria or to replace scattered mutant versions of the gene set and would thus be detected in modern-day isolates (Lawrence and Roth, unpublished).

This notion is difficult to test without a better knowledge of the conditions faced by bacteria over their entire range of habitats. However, one prediction might be that rarely selected phenotypes are most subject to mutational loss (and subsequent reacquisition by mating); they would therefore be under the strongest selection to cluster. Conversely, groups of genes that are essential or are under nearly constant selection are less likely to be clustered, since they would be present in all organisms and new mutants would be eliminated from the population before there was an opportunity for reacquisition. Aside from the ribosomal protein operons, which seem to contradict this prediction, the model fits a variety of available information (e.g., genes encoding pathways for utilization of relatively rare carbon sources are virtually all clustered in the genomes of *S. typhimurium* and *E. coli*, while genes for synthesis of many essential amino acids, purines, and pyrimidines are scattered). A better evaluation could be made if more were known about the relative abundance of the various nutrients in natural habitats.

Duplication Model for "Adaptive Mutation"

Gene amplification in response to stress provides a plausible mechanism whereby bacteria might appear to be able to direct mutability to base pairs whose alteration improves growth (24). The idea is based on the observations (described above) that bacterial duplications frequently provide improved ability to grow under stressful conditions. The model, outlined below, is described as it would apply to the experimental setup designed to investigate the possibility of adaptive mutation (22, 23). (i) Most revertible point mutations are detectably leaky (10, 62, 80). (ii) Duplications are common (6, 83) and are frequently selected when a particular gene product limits growth (50, 87, 105). (iii) Higher amplification of a mutant gene may occur if additional copies permit improved growth. Tlsty et al. selected multicopy arrays of the *lac* operon by demanding reversion of a leaky Lac^- strain (108). Edlund and Normark found 50-fold amplification of a chromosomal gene by selecting for ampicillin resistance (31). (iv) The large number of copies per cell (as a result of amplification) and the growth of a small clone of cells with this amplification will ultimately give rise to a sufficient number of copies of the mutant gene to allow the reversion event to be realized. (v) Once a base change corrects the mutation of one of the amplified copies, rapid growth starts and selection for the amplified array is relaxed. Haploid segregants with the revertant allele then grow up from these clones. Since only clones with the revertant are counted and the extra copies needed to realize the mutation are not assessed, enhanced mutability appears to be directed to the few base pairs that contribute to enhance growth. In fact, the mutation may have arisen from an unexpectedly large pool of gene copies.

Events outlined above provide a scenario that would appear to represent selection-directed mutability. The revertant frequency increases because there are more copies of the mutant gene present on the selection plate than one would estimate from the number of cells plated. The extra copies could be present in foci on the plate that ultimately generate revertant haploid colonies and thus might be missed by assays for background cell growth. The increased mutability appears to be directed to the exact base pairs that limit growth, because the final revertant clone carries a revertant allele that is selected from a large array of copies that are lost. It is not yet clear whether this scenario is appropriately scaled to account for the phenomenon of adaptive mutation, but it must, in principle, work under some conditions of bacterial growth.

Stahl has suggested that this sequence of events would be a way to account for the evolution of new genes under continuously selective conditions (F. Stahl, personal communication). A preexisting gene would acquire a very slight ability to perform a second, selectively valuable function. The fitness provided by this new activity would cause an increase in the frequency of the bifunctional allele in the population. Duplications could provide further advantage by amplifying the number of copies of the allele with the new minor activity, and the new gene copies could be driven to fixation in the population. The duplication would provide the dispensable gene copy, which permits the new activity of one gene copy to be enhanced by mutations which eliminate the original function of that copy. Thus, a gene with a new function is formed under continuous selection.

SUMMARY AND PROSPECTS

The events responsible for chromosomal rearrangements fall into two general classes, depending on their requirement for the RecA strand exchange protein. The most frequent exchanges are those that require RecA function and involve extensive sequence repeats (multiple kilobases). As rearrangements are restricted to smaller repeated sequences, the frequency of exchanges drops and a larger fraction of the observed exchanges occur by mechanisms independent of RecA protein. The recombination mechanisms responsible for exchanges between long sequence repeats appear to be able to distinguish between sequences in direct-order (including sister chromosomes) and inverse-order repeats. Pairs of sequences in inverse order recombine freely (double exchanges and gene conversion) regardless of their position or orientation, but

inverse-order sequences at particular chromosomal sites fail to undergo the full exchanges needed for inversion formation. Direct-order sequences (and sister chromosomes) do not seem to be subject to these restrictions. We suggest that the existence of "forbidden inversion intervals" results from the recombination system having been under selection to minimize the occurrence of inversions which stably alter the genetic map.

The RecA-independent mechanisms (which operate on short sequence repeats) are capable of forming deletions and duplications. Most studies of this phenomenon have focused on deletions of small regions of plasmids. This work has suggested multiple RecA-independent pathways of deletion formation. Many of these mechanisms are difficult to apply to formation of long deletions and duplications with widely separated endpoints. It seems likely that deletions and duplications (regardless of size) will ultimately be proved to form by similar mechanisms. Since the mechanisms that have been proposed for deletions depend heavily on small size and local sequence features, it is important to test them in a wide variety of contexts, for widely separated sequences, and particularly in the bacterial chromosome.

We suggest that RecA-dependent exchanges between long sequence repeats in the bacterial chromosome provide useful recombination assays that do not limit us to the study of events initiated by the double-stranded ends (which are necessarily provided by sexual assays). Rearrangement assays may test the notion that the variety of identified *rec* mutations normally act on a variety of initiating structures (not limited to double-strand breaks). Chromosomal recombination assays may also be able to resolve the question whether bacteria can perform full exchanges (in which both flanking sequences are rejoined in a single concerted event).

LITERATURE CITED

1. **Ailion, M. E., T. A. Bobik, and J. R. Roth.** 1993. Two global regulatory systems (*crp*/*cya* and *arcA*/*arcB*) control the propanediol/cobalamin (*pdu*/*cob*) regulon of *Salmonella typhimurium*. *J. Bacteriol*. **175:**7200–7208.
2. **Albertini, A. M., M. Hofer, M. P. Calos, and J. H. Miller.** 1982. On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. *Cell* **29:**319–328.
3. **Allgood, N. D., and T. J. Silhavy.** 1988. Illegitimate recombination in bacteria, p. 309–330. *In* R. Kucherlapati and G. R. Smith (ed.), *Genetic Recombination*. American Society for Microbiology, Washington, D.C.
4. **Allgood, N. D., and T. J. Silhavy.** 1991. Escherichia coli *xonA* (*sbcB*) mutants enhance illegitimate recombination. *Genetics* **127:**671–680.
5. **Anderson, R. P., C. G. Miller, and J. R. Roth.** 1976. Tandem duplications of the histidine operon observed following generalized transduction in *Salmonella typhimurium*. *J. Mol. Biol*. **105:**201–218.
6. **Anderson, R. P., and J. R. Roth.** 1977. Tandem genetic duplications in phage and bacteria. *Annu. Rev. Microbiol*. **31:**473–505.
7. **Anderson, R. P., and J. R. Roth.** 1978. Tandem chromosomal duplications in *Salmonella typhimurium:* fusion of histidine genes to novel promoters. *J. Mol. Biol*. **119:**146–166.
8. **Anderson, R. P., and J. R. Roth.** 1978. Tandem genetic duplications in *Salmonella typhimurium* amplification of the histidine operon. *J. Mol. Biol*. **126:**53–72.
9. **Anderson, R. P., and J. Roth.** 1981. Spontaneous tandem genetic duplications in *Salmonella typhimurium* arise by unequal recombination between ribosomal RNA (*rrn*) cistrons. *Proc. Natl. Acad. Sci. USA* **78:**3113–3117.
10. **Atkins, J. F., D. Elseviers, and L. Gorini.** 1972. Low activity of β-galactosidase in frameshift mutants of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **69:**1192–1195.
11. **Balbinder, E.** 1993. Multiple pathways of deletion formation in Escherichia coli. *Mutat. Res*. **299:**193–209.

12. **Balbinder, E., B. Coll, J. Hutchinson, A. S. Bianchi, T. Groman, K. A. Wheeler, and M. Meyer.** 1993. Participation of the SOS system in producing deletions in E. coli plasmids. *Mutat. Res.* **286:**253–265.
13. **Benson, N., and J. R. Roth.** 1994. Suppressors of *recB* mutations in *Salmonella typhimurium*. *Genetics* **138:**11–29.
14. **Birch, A., A. Häusler, and R. Hütter.** 1990. Genome rearrangement and genetic instability in *Streptomyces* spp. *J. Bacteriol.* **172:**4138–4142.
15. **Birch, A., A. Häusler, C. Ruttener, and R. Hütter.** 1991. Chromosomal deletion and rearrangement in *Streptomyces glaucescens*. *J. Bacteriol.* **173:**3531–3538.
16. **Blakely, G., G. May, R. McCulloch, L. K. Arciszewska, M. Burke, S. T. Lovett, and D. J. Sherratt.** 1993. Two related recombinases are required for site-specific recombination at *dif* and *cer* in E. coli K12. *Cell* **75:**351–361.
17. **Bobik, T. A., M. Ailion, and J. R. Roth.** 1992. A single regulatory gene integrates control of vitamin B_{12} synthesis and propanediol degradation. *J. Bacteriol.* **174:**2253–2266.
18. **Bochner, B., G. Huang, G. Schieven, and B. N. Ames.** 1980. Positive selection for loss of tetracycline resistance. *J. Bacteriol.* **143:**926–933.
19. **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.
20. **Bullock, P., J. J. Champoux, and M. Botchan.** 1985. Association of crossover points with topoisomerase I cleavange sites: a model for nonhomologous recombination. *Science* **230:**954–958.
21. **Burland, V., F. Plunkett III, D. Daniels, and F. R. Blattner.** 1993. DNA sequence and analysis of 136 kilobases of Escherichia coli genome: organizational symmetry around the origin of replication. *Genomics* **16:**551–561.
22. **Cairns, J., and P. L. Foster.** 1991. Adaptive reversion of a frameshift mutation in *Escherichia coli*. *Genetics* **128:**695–701.
23. **Cairns, J., and P. L. Foster.** 1992. Mechanisms of directed mutation. *Genetics* **131:**783–789.
24. **Cairns, J., J. Overbaugh, and S. Miller.** 1988. The origin of mutants. *Nature* (London) **335:**142–145.
25. **Chumley, F. G., and J. R. Roth.** 1980. Rearrangement of the bacterial chromosome using Tn*10* as a region of homology. *Genetics* **94:**1–14.
26. **Cohen, A., and A. Laban.** 1983. Plasmidic recombination in *Escherichia coli* K12: the role of the *recF* gene function. *Mol. Gen. Genet.* **189:**471–474.
27. **Cornet, F., I. Mortier, J. Patte, and J. M. Louarn.** 1994. Plasmid pSC101 harbors a recombination site, *psi,* which is able to resolve plasmid multimers and to substitute for the analogous chromosomal *Escherichia coli* site, *dif*. *J. Bacteriol.* **176:**3188–3195.
28. **Dianov, G. L., A. V. Kuzminov, A. V. Mazin, and R. I. Salganik.** 1991. Molecular mechanisms of deletion formation in Escherichia coli plasmids. I. Deletion formation mediated by long direct repeats. *Mol. Gen. Genet.* **228:**153–159.
29. **Dimpfl, J., and H. Echols.** 1989. Duplication mutation as an SOS response in *Escherichia coli:* enhanced duplication formation by a constitutively activated RecA. *Genetics* **123:**255–260.
30. **Doherty, M. J., P. T. Morrison, and R. Kolodner.** 1983. Genetic recombination of bacterial plasmid DNA: physical and genetic analysis of the products of plasmid recombination in *Escherichia coli*. *J. Mol. Biol.* **167:**539–560.
31. **Edlund, T., and S. Normark.** 1981. Recombination between short DNA homologies causes tandem duplication. *Nature* (London) **292:**269–271.
32. **Ehrlich, S. D.** 1989. Illegitimate recombination in bacteria, p. 799–832. *In* D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
33. **Fani, R., P. Liò, I. Chiarelli, and M. Bazzicalupo.** 1994. The evolution of the histidine biosynthetic

genes in prokaryotes: a common ancestor for the *hisA* and *hisF* genes. *J. Mol. Evol.* **38:**489–495.
34. **Farabaugh, P. J., U. Schmeissner, M. Hofer, and J. H. Miller.** 1978. Genetic studies of the *lac* repressor VII. On the molecular nature of spontaneous hotspots in the *lacI* gene of *Escherichia coli*. *J. Mol. Biol.* **126:**847–863.
35. **Flores, M., S. Brom, T. Stepkowski, M. L. Girard, G. Davila, D. Romero, and R. Palacios.** 1993. Gene amplification in Rhizobium: identification and in vivo cloning of discrete amplifiable DNA regions (amplicons) from Rhizobium leguminosarum biovar phaseoli. *Proc. Natl. Acad. Sci. USA* **90:**4932–4936.
36. **Foster, T. J., V. Lundblad, S. Hanley-Way, S. M. Halling, and N. Kleckner.** 1981. Three Tn*10*-associated excision events: relationship to transposition and role of direct and inverted repeats. *Cell* **23:**215–227.
37. **Franklin, N.** 1967. Extraordinary recombinational events in Escherichia coli. Their independence of the rec^+ function. *Genetics* **55:**699–707.
37a. **Galitski, T., and J. R. Roth.** 1995. Evidence that F plasmid transfer replication underlies apparent adaptive mutations. *Science* **268:**421–423.
38. **Glickman, B. W., and L. S. Ripley.** 1984. Structural intermediates of deletion mutagenesis: a role for palindromic DNA. *Proc. Natl. Acad. Sci. USA.* **81:**512–516.
39. **Goldberg, I., and J. J. Mekalanos.** 1986. Effect of a *recA* mutation on cholera toxin gene amplification and deletion events. *J. Bacteriol.* **165:**723–731.
40. **Goldman, B. S., and J. R. Roth.** 1993. Genetic structure and regulation of the *cysG* gene in *Salmonella typhimurium*. *J. Bacteriol.* **175:**1457–1466.
41. **Grabnar, M., R. C. Stahl, Z. Hartman, and P. E. Hartman.** 1964. Genes and proteins involved in histidine biosynthesis in *Salmonella typhimurium*. *Brookhaven Symp. Biol.* **17:**15–52.
42. **Guttman, D. S., and D. E. Dykhuizen.** 1994. Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. *Science* **266:**1380–1383.
43. **Hafner, L. M., and D. G. MacPhee.** 1991. Precise excision of Tn*10* in *Salmonella typhimurium:* effects of mutations in the *polA, dam, mutH* and *mutB* genes and of methionine or ethionine in the plating medium. *Mutat. Res.* **263:**179–184.
44. **Hageman, A. T., and N. L. Craig.** 1993. Tn*7* transposition creates a hotspot for homologous recombination at the transposon donor site. *Genetics* **133:**9–16.
45. **Hartman, P. E., J. C. Loper, and D. Serman.** 1960. Fine structure mapping by complete transduction between histidine-requiring *Salmonella* mutants. *J. Gen. Microbiol.* **22:**323–353.
46. **Harvey, S., and C. W. Hill.** 1990. Exchange of spacer regions between rRNA operons in Escherichia coli. *Genetics* **125:**683–690.
47. **Heath, J. D., and G. M. Weinstock.** 1991. Tandem duplications of the lac region of the Escherichia coli chromosome. *Biochimie* **73:**343–352.
48. **Hill, C. W., R. H. Grafstrom, B. W. Harnish, and B. W. Hillman.** 1977. Tandem duplications resulting from recombination between ribosomal RNA genes in *Escherichia coli*. *J. Mol. Biol.* **116:**407–428.
49. **Hill, C. W., and J. A. Gray.** 1988. Effects of chromosomal inversion on cell fitness in *Escherichia coli* K-12. *Genetics* **119:**771–778.
50. **Horiuchi, R., S. Horiuchi, and A. Novick.** 1963. The genetic basis of hypersynthesis of β-galactosidase. *Genetics* **48:**157–169.
51. **Horowitz, N. H.** 1965. The evolution of biochemical synthesis: retrospect and prospect, p. 15–23. *In* V. Bryson and H. J. Vogel (ed.), *Evolving Genes and Proteins*. Academic Press, Inc., New York.
52. **Hughes, K., and J. R. Roth.** 1985. Directed formation of deletions and duplications using Mud(Ap, lac). *Genetics* **109:**263–282.
53. **Hughes, K. T., J. R. Roth, and B. M. Olivera.** 1991. A genetic characterization of the *nadC* gene of *Salmonella typhimurium*. *Genetics* **127:**657–670.

54. **Hutchinson, F.** 1993. Induction of large DNA deletions by persistent nicks: a new hypothesis. *Mutat. Res.* **299:**211–218.
55. **Ikeda, H., A. Kazuko, and A. Naito.** 1982. Illegitimate recombination mediated in vitro by DNA gyrase of *Escherichia coli:* structure of recombinant molecules. *Proc. Natl. Acad. Sci. USA* **79:**3724–3728.
56. **Inselburg, J.** 1967. Formation of deletion mutations in recombination-deficient mutants of *Escherichi coli. J. Bacteriol.* **94:**1266–1267.
57. **Jacob, F., and J. Monod.** 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3:**318–356.
58. **Kazic, T., and D. E. Berg.** 1990. Context effects in the formation of deletions in *Escherichia coli. Genetics* **126:**17–24.
59. **Kolodner, R. D., R. A. Fishel, and M. Howard.** 1985. Genetic recombination of bacterial plasmid DNA: effect of RecF pathway mutations on plasmid recombination in *Escherichia coli. J. Bacteriol.* **163:**1060–1066.
60. **Konrad, E. B.** 1977. Method for the isolation of *Escherichia coli* mutants with enhanced recombination between chromosomal duplications. *J. Bacteriol.* **130:**167–172.
61. **Krawiec, S., and M. Riley.** 1990. Organization of the bacterial chromosome. *Microbiol. Rev.* **54:**502–539.
62. **Langridge, J., and J. H. Campbell.** 1969. Classification and intragenic position of mutations in the beta-galactosidase gene of *Escherichia coli. Mol. Gen. Genet.* **103:**339–347.
63. **Lawes, M., and S. Maloy.** 1995. Mu*d*SacI, a transposon with strong selectable and counterselectable markers: use for rapid mapping of chromosomal mutations in *Salmonella typhimurium. J. Bacteriol.* **177:**1383–1387.
64. **Leblond, P., P. DeMuyter, J.-M. Simonet, and B. Decaris.** 1991. Genetic instability and associated genome plasticity in *Streptomyces ambofaciens:* pulsed-field gel electrophoresis evidence for large DNA alterations in a limited genomic region. *J. Bacteriol.* **173:**4229–4233.
65. **Levy, M. S., P. Pomposiello, and R. Nagel.** 1991. RecA-dependent increased precise excision of Tn*10* in *Salmonella typhimurium. Mutat. Res.* **255:**95–100.
66. **Lin, R.-J., M. Capage, and C. W. Hill.** 1984. A repetitive DNA sequence, Rhs, responsible for duplication within the *Escherichia coli* K12 chromosome. *J. Mol. Biol.* **177:**1–18.
66a. **Liu, S. L., and K. E. Sanderson.** 1995. Rearrangements in the genome of the bacterium *Salmonella typhi. Proc. Natl. Acad. Sci. USA* **92:**1018–1022.
67. **Lobner-Olesen, A., and P. L. Kuempel.** 1992. Chromosome partitioning in *Escherichia* coli. *J. Bacteriol.* **174:**7883–7889.
68. **Lovett, S. T., P. T. Drapkin, V. A. Sutera, Jr., and T. J. Gluckman-Peskind.** 1993. A sister-strand exchange mechanism for *recA*-independent deletion of repeated DNA sequences in *Escherichia coli. Genetics* **135:**631–642.
68a. **Lovett, S. T., T. J. Gluckman, P. J. Simon, V. A. Sutera, and P. T. Drapin.** 1994. Recombination between repeats in *Escherichia coli* by a *recA*-independent, proximity-sensitive mechanism. *Mol. Gen. Genet.* **245:**294–300.
69. **Lundblad, V., A. F. Taylor, G. R. Smith, and N. Kleckner.** 1984. Unusual alleles of recB and recC stimulate excision of inverted repeat transposon Tn*10* and Tn*5*. *Proc. Natl. Acad. Sci. USA* **81:**824–828.
70. **Mahan, M. J., and J. R. Roth.** 1988. Reciprocality of recombination events that rearrange the chromosome. *Genetics* **120:**23–36.
71. **Mahan, M. J., and J. R. Roth.** 1989. Role of RecBC function in formation of chromosomal rearrangements a two-step model for recombination. *Genetics* **121:**433–444.
72. **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.

73. **Maloy, S. R., and W. D. Nunn.** 1981. Selection of loss of tetracycline resistance by *Escherichia coli*. *J. Bacteriol*. **145:**1110–1112.
74. **Masurekar, M., K. N. Kreuzer, and L. S. Ripley.** 1991. The specificity of topoisomerase-mediated DNA cleavage defines acridine-induced frameshift specificity within a hotspot in bacteriophage T4. *Genetics* **128:**193–202.
75. **Matfield, M., R. Badawi, and W. J. Brammar.** 1985. Rec-dependent and Rec-independent recombination of plasmid-borne duplications in Escherichia coli K12. *Mol. Gen. Genet*. **199:**518–523.
76. **Mazin, A. V., A. V. Kuzminov, G. L. Dianov, and R. I. Salganik.** 1991. Mechanisms of deletion formation in Escherichia coli plasmids. II. Deletions mediated by short direct repeats. *Mol. Gen. Genet*. **228:**209–214.
77. **Mekalanos, J. J.** 1983. Duplication and amplification of toxin genes in *Vibrio cholera*. *Cell* **35:**253–263.
78. **Miesel, L., A. Segall, and J. R. Roth.** 1994. Construction of chromosomal rearrangements in Salmonella by transduction: inversions of non-permissive segments are not lethal. *Genetics* **164:**919–932.
79. **Milkman, R., and M. M. Bridges.** 1990. Molecular evolution of the *Escherichia coli* chromosome. III. Clonal frames. *Genetics* **126:**505–517.
80. **Miller, J. H., and A. M. Albertini.** 1983. Effects of surrounding sequence on the suppresion of nonsense codons. *J. Mol. Biol*. **164:**59–71.
80a. **Myers, R. S., and F. W. Stahl.** 1994. Chi and the RecBCD enzyme of *Escherichia coli*. *Annu. Rev. Genet*. **28:**49–70.
81. **Naito, A., S. Naito, and H. Ikeda.** 1984. Homology is not required for recombination mediated by DNA gyrase of *Escherichia coli*. *Mol. Gen. Genet*. **193:**238–243.
82. **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,* vol. 2. American Society for Microbiology, Washington, D.C.
83. **Petes, T. D., and C. W. Hill.** 1988. Recombination between repeated genes in microorganisms. *Annu. Rev. Genet*. **22:**147–168.
84. **Petit, M. A., J. Dimpfl, M. Radman, and H. Echols.** 1991. Control of large chromosomal duplications in Escherichia coli by the mismatch repair system. *Genetics* **129:**327–332.
85. **Petit, M.-A., J. M. Mesas, P. Noirot, F. Morel-Deville, and S. D. Ehrlich.** 1992. Induction of DNA amplification in the *Bacillus subtilis* chromosome. *EMBO J*. **11:**1317–1326.
86. **Rebollo, J.-E., V. Francois, 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.
87. **Rigby, P. W. J., B. D. Burleigh, and B. S. Hartley.** 1974. Gene duplication in experimental enzyme evolution. *Nature* (London) **251:**200–204.
88. **Ripley, L. S.** 1990. Frameshift mutation: determinants of specificity. *Annu. Rev. Genet*. **24:**189–213.
89. **Roecklein, B. A., and P. L. Kuempel.** 1992. In vivo characterization of *tus* gene expression in *Escherichia coli*. *Mol. Microbiol*. **6:**1655–1661.
90. **Rownd, R. H.** 1982. R plasmid drug resistance gene amplifiation in bacteria, p. 273–282. *In* R. T. Schimke (ed.), *Gene Amplification*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
91. **Sapienza, C., and W. F. Doolittle.** 1982. Unusual physical organization of the *Halobacterium* genome. *Nature* (London) **295:**384–389.
92. **Sapienza, C., M. R. Rose, and W. F. Doolittle.** 1982. High-frequency genomic rearrangements involving archaebacterial repeat sequence elements. *Nature* (London) **299:**182–185.
93. **Schmid, M. B., and J. R. Roth.** 1983. Genetic methods for analysis and manipulation of inversion mutations in bacteria. *Genetics* **105:**517–538.
94. **Schmid, M. B., and J. R. Roth.** 1983. Selection and end point distribution of bacterial inversion

mutations. *Genetics* **105:**539–558.
95. **Schofield, M. A., R. Agbunag, M. L. Michaels, and J. H. Miller.** 1992. Cloning and sequencing of *Escherichia coli mutR* shows its identity to *topB,* encoding topoisomerase III. *J. Bacteriol.* **174:**5168–5170.
96. **Schofield, M. A., R. Agbunag, and J. H. Miller.** 1992. DNA inversions between short inverted repeats in Escherichia coli. *Genetics* **132:**295–302.
97. **Sclafani, R. A., and J. A. Wechsler.** 1981. High frequency of genetic duplications in the *dnaB* region of the *Escherichia coli* K12 chromosome. *Genetics* **98:**677–690.
98. **Segall, A., M. J. Mahan, and J. R. Roth.** 1988. Rearrangement of the bacterial chromosome: forbidden inversions. *Science* **241:**1314–1318.
99. **Segall, A., and J. R. Roth.** 1994. Approaches to half-tetrad analysis in bacteria: recombination between repeated inverse-order chromosomal sequences. *Genetics* **136:**27–39.
100. **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:**737–748.
101. **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.
102. **Shyamala, V., E. Schneider, and G. F. Ames.** 1990. Tandem chromosomal duplications: role of REP sequences in the recombination event at the join-point. *EMBO J.* **9:**939–946.
103. **Sinden, R. R., G. X. Zheng, and R. G. Brankamp.** 1991. On the deletion of inverted repeated DNA in Escherichia coli: effects of length, thermal stability, and cruciform formation in vivo. *Genetics* **129:**991–1005.
104. **Smith, J. M., N. H. Smith, M. O'Rourke, and B. G. Spratt.** 1993. How clonal are bacteria? *Proc. Natl. Acad. Sci. USA* **90:**4384–4388.
105. **Sonti, R. V., and J. R. Roth.** 1989. Role of gene duplications in the adaptation of *Salmonella typhimurium* to growth on limiting carbon sources. *Genetics* **123:**19–28.
106. **Streissinger, G.** 1985. Mechanisms of spontaneous and induced frameshift mutations. *Genetics* **109:**633–659.
107. **Syvanen, M., J. D. Hopkins, T. J. Griffin IV, T.-Y. Liang, K. Ippen-Ihler, and R. Kolodner.** 1986. Stimulation of precise excision and recombination by conjugal proficient F′ plasmids. *Mol. Gen. Genet.* **203:**1–7.
108. **Tlsty, D. T., A. M. Albertini, and J. H. Miller.** 1984. Gene amplification in the *lac* region of *E. coli. Cell* **37:**217–224.
109. **Trinh, T. Q., and R. R. Sinden.** 1991. Preferential DNA secondary structure mutagenesis in the lagging strand of replication in E. coli. *Nature* (London) **352:**544–547.
110. **Trinh, T. Q., and R. R. Sinden.** 1993. The influence of primary and secondary DNA structure in deletion and duplication between direct repeats in Escherichia coli. *Genetics* **134:**409–422.
111. **Tsou, A. Y., S. C. Ransom, J. A. Gerlt, D. D. Buechter, P. C. Babbitt, and G. L. Kenyon.** 1990. Mandelate pathway of *Pseudomonas putida:* sequence relationships involving mandelate racemase, (S)-mandelate dehydrogenase and benzoylformate decarboxylase and expression of benzoylformate decarboxylase in *Escherichia coli. Biochemistry* **29:**9856–9862.
112. **Weston-Hafer, K., and D. E. Berg.** 1989. Palindromy and the location of deletion endpoints in *Escherichia coli. Genetics* **121:**651–658.
113. **Whoriskey, S. K., V. H. Nghiem, P.-M. Leong, J.-M. Masson, and J. H. Miller.** 1987. Genetic rearrangements and gene amplification in *Escherichia coli:* DNA sequences at the junctures of amplified gene fusions. *Genes Dev.* **1:**227–237.
114. **Whoriskey, S. K., M. A. Schofield, and J. H. Miller.** 1991. Isolation and characterization of

Escherichia coli mutants with altered rates of deletion formation. *Genetics* **127:**21–30.
115. **Yi, T., D. Stearns, and B. Demple.** 1988. Illegitimate recombination in an *Escherichia coli* plasmid: modulation by DNA damage and a new bacterial gene. *Mol. Gen. Genet.* **170:**2898–2903.
116. **Zieg, J., and S. R. Kushner.** 1977. Analysis of genetic recombination between two partially deleted lactose operons of *Escherichia coli*. *J. Bacteriol.* **131:**123–132.