Cold Spring Harbor Symposia on Quantitative Biology

Gene Duplication in Bacteria: Alteration of Gene Dosage by Sister-chromosome Exchanges

R. P. Anderson and J. R. Roth

Cold Spring Harb Symp Quant Biol 1979 43: 1083-1087
Access the most recent version at doi:10.1101/SQB.1979.043.01.120

Gene Duplication in Bacteria: Alteration of Gene Dosage by Sister-chromosome Exchanges

R. P. ANDERSON* AND J. R. ROTH

Department of Biology, University of Utah, Salt Lake City, Utah 84112

Unequal exchanges between sister chromosomes can result in duplications (or deletions) of segments of the chromosome. These events are quite common and have been studied in a variety of systems in phage and bacteria (for review, see Anderson and Roth 1977). We have isolated and mapped a series of duplications of the *his* region of *Salmonella typhimurium*. We infer from the results that most duplications result from legitimate (*rec*-dependent) sister-chromosome exchanges between homologous sequences located at different positions on the map. Duplications also seem to arise by illegitimate (*rec*-independent) exchanges, although less frequently. Here we review the evidence for these two types of exchanges and present preliminary evidence that the resulting increase in gene dosage may provide a significant regulatory mechanism for bacteria.

Predictions

Figure 1 depicts the events which we believe lead to gene duplication. Exchange between points *ab* and *fg* on the sister chromosomes lead to generation of a duplication and a deletion. The position of the exchange determines the endpoints of the duplicated segment. Exchanges of this sort, which are due to legitimate recombination, would be expected to yield duplications having two properties: (1) They would arise only in rec^+ cells. (2) They would repeatedly arise with the same endpoints, since the exchanges could only involve sites that share sequence homology.

Illegitimate exchanges would (by definition) also occur in a rec^- cell. It is not possible to predict the expected distribution of duplication endpoints resulting from such exchanges. If the exchanges are truly random, then a wide scattering of endpoints is expected. If an insertion sequence (IS) mediates the exchanges, one duplication endpoint might be fixed and the other distributed. If the illegitimate exchanges result from site-specific recombination, then both endpoints might be fixed.

*Present address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England.

Selection for Duplication in the *his* Region

Two methods have been used for selecting duplications of the *his* region. The first (dosage selection) demands resistance to an inhibitor of one of the *his* enzymes; under the conditions used, a twofold increase in gene dosage achieves this resistance (R. P. Anderson and J. R. Roth, 1978b). This selection makes no demands on the endpoints of the duplicated segment; it only requires that the *his* operon be included in the duplication. The second selection method (fusion selection) requires duplications whose endpoints are positioned so that one of the *his* genes is placed under control of a foreign promoter (operon fusion). This is diagramed in Figure 2. The parent strain (*his203*) has no *his* promoter and cannot express the *hisD* gene. Duplication events can lead to a $HisD^+$ phenotype if the critical exchange events leave the *hisD* gene at the join point of the duplication and near a functional foreign promoter (Anderson and Roth 1978a). This selection places heavy restrictions on the position of this exchange (i.e., on the permissible endpoints of the duplications): one end must be near the end of the *hisD* gene; the other must be near a functioning foreign promoter oriented in the appropriate direction.

rec Dependence of Duplication Formation

Table 1 presents the frequency of *his* duplications isolated by the two selection methods described above. The dosage selection yields duplications at a high rate; this is expected, since there are few restrictions on the permissible endpoints. It appears that most duplications (>99%) isolated by this method depend on recombination (*recA*) for their formation. No duplications were detected in the rec^- parent strain. Technical problems with the dosage selection preclude conveniently looking for duplications less frequent than 1×10^{-8}.

The fusion selection yields duplications less frequently. Although available sites for duplication-generating exchanges are severely limited, the occurrence of duplications is again highly *rec*-dependent; the

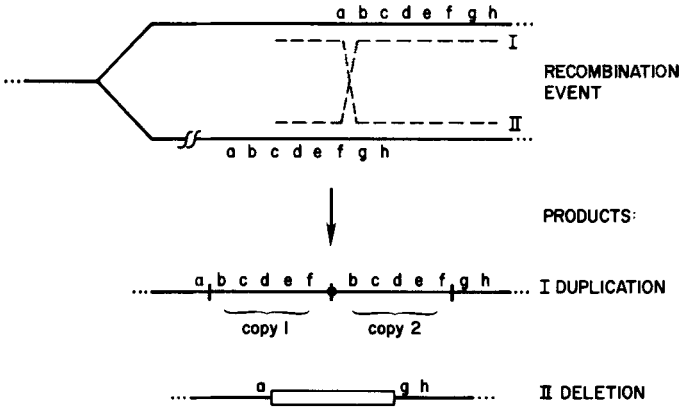

Figure 1. A likely mechanism for duplication formation. Genetic exchanges between different sites on sister chromosomes can generate duplications and deletions. Many of the deletions are probably lethal, but large duplications cause no loss of gene function.

frequency of duplication is sixfold higher in *rec*$^+$ strains than in *rec*$^-$ strains. This selection, unlike the dosage selection, permits detection of duplications in a *rec*$^-$ parent, in spite of their low rate of occurrence (Anderson and Roth 1978a).

Mapping Duplication Endpoints

Auxotrophic insertion mutations made by the transposable drug-resistance element Tn*10* have been used as genetic markers in transductional crosses to map the duplication endpoints (Kleckner et al. 1975, 1977). The duplication-carrying strain is used as a recipient; the auxotrophic Tn*10*-insertion mutant is used as donor. Selection is made for inheritance of Tn*10* (tetracycline resistance [Tc^r]) on minimal medium. Thus, one selects for simultaneous inheritance of the insertion mutation (*tet*^{r}) and the wild-type allele of the gene involved. If the recipient duplication includes the site of the Tn*10* insertion, prototrophic recombinants arise by inheritance of Tn*10* in one of the two copies of the locus. If the duplication does not include the donor Tn*10* site, then all recombinants inheriting Tn*10* will of necessity become auxotrophs and will fail to grow. The resolution of this mapping method is currently limited only by the number of characterized Tn*10* insertions available.

Distribution of Duplication Endpoints

The endpoint distribution for *his* duplications obtained from each of the above selections is presented in Figure 3. The dosage selection yields (from a *rec*$^+$ parent) two duplication types which occur repeatedly. This conclusion is based largely on mapping of the right endpoint; the left endpoint mapping is of very low resolution, owing to the lack of Tn*10* insertions on this side of the *his* operon (Anderson and Roth 1978b).

The fusion selection (from a *rec*$^+$ parent) yielded two predominant types of duplication. They have one endpoint between the *hisG* and *hisD* genes and the other at a particular position on the map far from the *his* operon. The endpoint of the majority type (marked with an asterisk in Fig. 3) has been more accurately determined by its contransducibility with the *argA* gene.

Duplications arising in a *rec*$^-$ parent can also be detected using the fusion selection. Their endpoint distribution (see Fig. 3) is very different from that of the *rec*$^+$ duplication: their endpoints are variable, both within the *his* operon and to the right of *his*. Most importantly, the most frequent class found in *rec*$^+$ cells was not encountered among the duplications arising in *rec*$^-$ cells. These results suggest that even when the range of possible exchange points is strongly restricted, most duplication formation is *rec*-dependent. Duplications that do arise in a *rec*$^-$ strain show an endpoint distribution that gives little evidence for favored endpoints (Anderson and Roth 1978a).

Selecting Duplications at Sites Other Than the *his* Locus

Duplications have been selected for many regions of the chromosome. They were isolated following transductional crosses in which auxotrophic Tn*10* insertions were used as donors and wild-type cells were used as recipients. Selection was made for tetracycline resistance and growth on minimal medium. Since generalized transduction proceeds by replacing the recipient allele of the selected marker with donor material, the

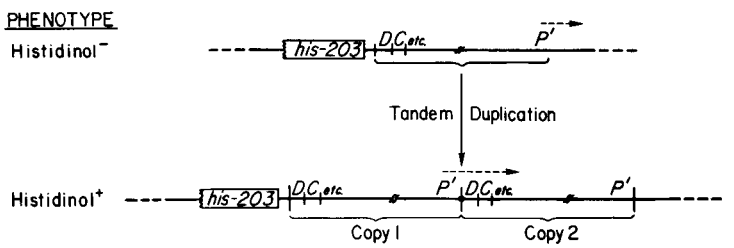

Figure 2. Selecting duplications by operon fusion. Parental mutant *his203* lacks the *his* promoter but retains a structurally normal, unexpressed *hisD* gene. Selection for growth on histidinol demands expression of the *hisD* gene. One of the classes of HisD$^+$ revertants is described above. The duplication event leaves the *hisD*$^+$ gene near a foreign promoter at the duplication join point (Anderson and Roth 1978a).

Table 1. Frequency of *his* Duplications

Selection	rec$^+$	rec$^-$
Dosage	6×10^{-5}	$<10^{-8}$
Fusion	3×10^{-10}	5×10^{-11}

vast majority of the transductants that inherit Tn*10* also inherit the concomitant auxotrophy. However, a few prototrophic transductants arise by transduction into recipient cells carrying preexisting duplications. Recipients carrying a duplication of the Tn*10* site can inherit the Tn*10* element in one copy of the region while the wild-type allele is maintained in the second copy. The frequency of duplications of the Tn*10* site can easily be estimated by comparing the number of Tet^r transductants obtained on minimal medium to the number obtained on medium supplemented with the nutrient required by the Tn*10* insertion mutant. Correction is made for the prototrophic Tet^r transductants arising by means not necessitating duplication.

Duplication Frequencies at Many Sites

The duplication frequencies for many sites in the chomosome are presented in Figure 4. With one major exception, duplication frequencies are between 10^{-4} and 10^{-3}. The exceptional region is located just clockwise from the origin of replication. Here a peak duplication frequency of 3×10^{-2} is observed for broth-grown cells. Nearby points in the chromosome also show an elevated frequency of duplications. The region involved includes four rRNA loci (*rrnA*, *B*, *C*, *E*), which are not contiguous but are all in the same orientation (Nomura et al. 1977; Kenerley et al. 1977; Ikemura and Nomura 1977). Duplications of the *glyT* gene, also located in the region, have been studied extensively by Hill and coworkers. They have demonstrated that duplications of the *glyT* gene arise by unequal recombination between pairs of these several *rrn* loci (Hill and Combriato 1973; Hill et al. 1977).

To determine whether the frequent duplications observed in *Salmonella* involve *rrn* genes, we mapped a large number of duplications affecting this region of the chromosome. Results are presented in Figure 5. The duplications presented were isolated as outlined previously, except that Tn*5* (kanamycin resistance [Km^r]) was used in place of Tn*10*. The duplications were mapped using auxotrophic Tn*10* insertion mutations as described previously. As indicated in Figure 5, the extent of duplications including the *ilv* or *metF* regions is consistent with their being generated by recombination between various *rrn* loci. Of 100 such duplications mapped, only one is inconsistent with this interpretation (dashed line in Fig. 5).

A second point emerged from mapping these duplications. Of 11 duplications of the *thr* locus, eight have their left endpoint in a region including an *rrn* locus and their right endpoint between the *leu* and *pro* regions. Similarly, of eight *leu* duplications mapped, five have endpoints consistent with their having arisen by recombination between *rrn* loci. These data suggest that an *rrn* gene is located between the *leu* and *pro* genes in the same orientation as the *rrnA*, *B*, *C*, and *E* genes. These duplications are presented in Figure 5.

Gene Duplication as a Regulatory Mechanism

One observation strongly suggests that gene duplications may play a significant regulatory role: the frequency of duplication involving the *rrn* loci is influenced by previous growth conditions. Cells grown in rich medium at a fast doubling time show a high rate of duplication; cells grown more slowly on minimal salts medium have a relatively low rate (Fig. 4). These effects are not seen for all loci, although one other

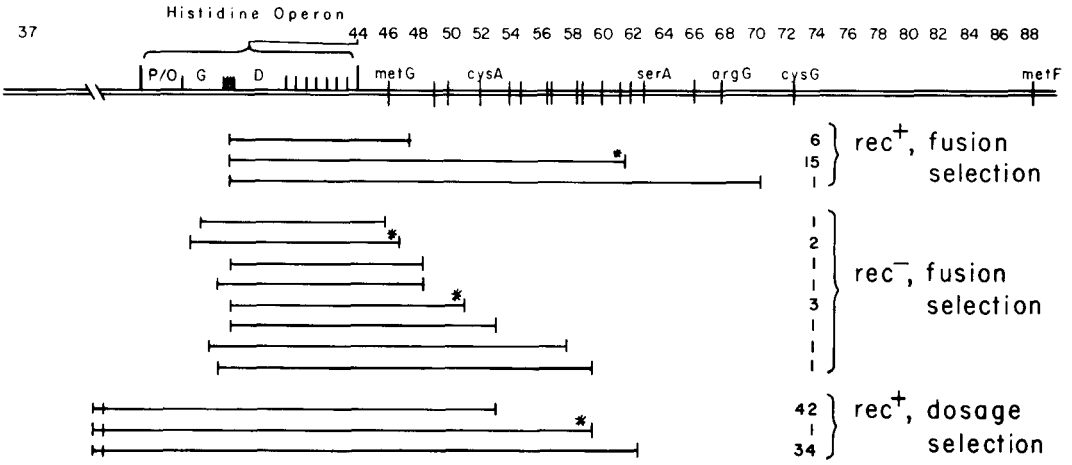

Figure 3. Map distribution of *his* duplications. Duplications were mapped using Tn*10* auxotrophic insertion mutations. Each vertical line crossing the genetic map indicates the position of one of the Tn*10* insertions used; only selected mutations are labeled. Horizontal bars below the genetic map indicate the extent of chromosome duplication in one duplication type. The members at the right of the duplication bars indicate the number of independent isolates of that duplication type. Asterisks denote duplication types whose join point shows P22-mediated cotransduction with the rightmost-mapping mutation included in the duplicated segment (Anderson and Roth 1978a,b).

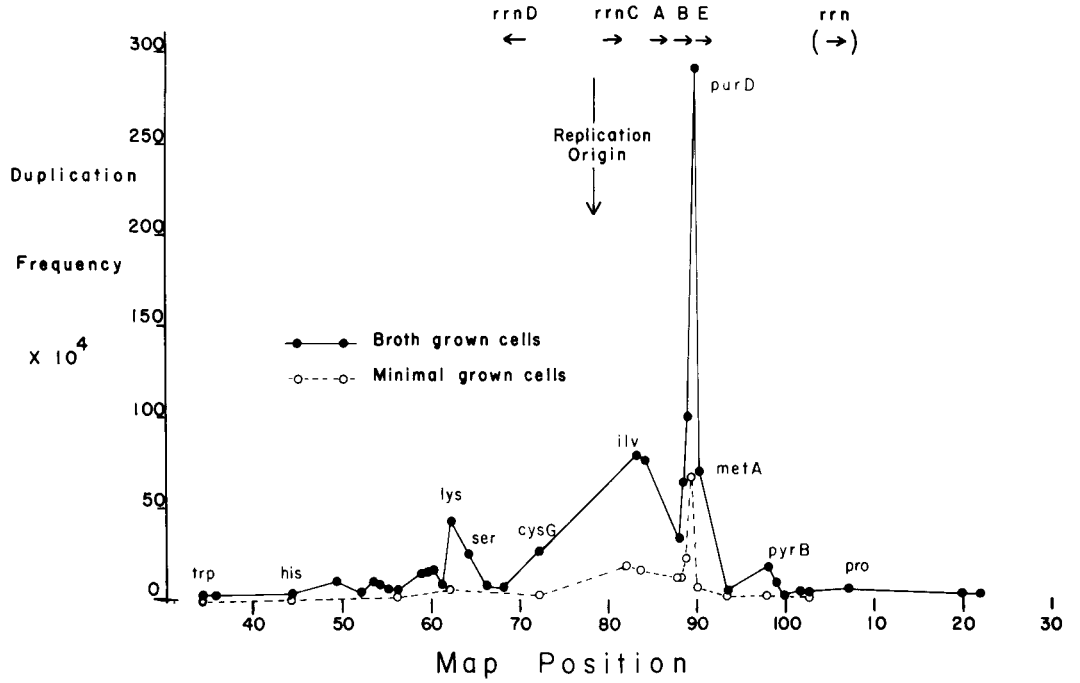

Figure 4. Frequency of duplications at many points on the bacterial chromosome. Duplication frequency was assayed using Tn*10* insertion auxotrophs at each point to be checked. A wild-type recipient was transduced to Tet^r with and without the nutritional supplementation required by the donor mutant. Only recipients carrying a duplication of the relevant gene can give rise to prototrophic Tet^r transductants. By comparing the frequency of Tet^r transductants with and without supplementation, one can estimate the duplication frequency in the recipient population.

region (60–65 min) shows effects similar to those seen for the *rrn* loci.

An increase in dosage of *rrn* genes would seem to be an adaptive response to conditions of fast growth rate, since it is likely that growth rate in rich medium is limited by the supply of ribosomes (for review, see Nierlich 1978). The position and orientation of *rrn* genes is such that their amplification is favored. The four *rrn* loci located clockwise from the origin of replication (*rrnA*, *B*, *C*, *E*) are close together and all in the same orientation; this arrangement permits duplication formation by homologous recombination between different *rrn* loci. It should be noted, however, that even under conditions favoring duplication, only 3% of the individuals tested carried duplications. We presume that this frequency may increase further when cells are grown for extended time periods under these conditions.

The mechanism for the variation of duplication frequency with growth rate is not known, but several possibilities are being entertained: (1) Duplications occur by spontaneous recombination between *rrn* loci; cells carrying such duplications achieve a higher doubling time and increase in frequency in the population. (2) After a shift-up, new chromosome replication forks provide increased dosage of the four *rrn* loci near the origin. This provides enhanced opportunity for duplication, causing the duplication frequency to increase. (3) A specific recombination mechanism may exist which generates duplications in response to some

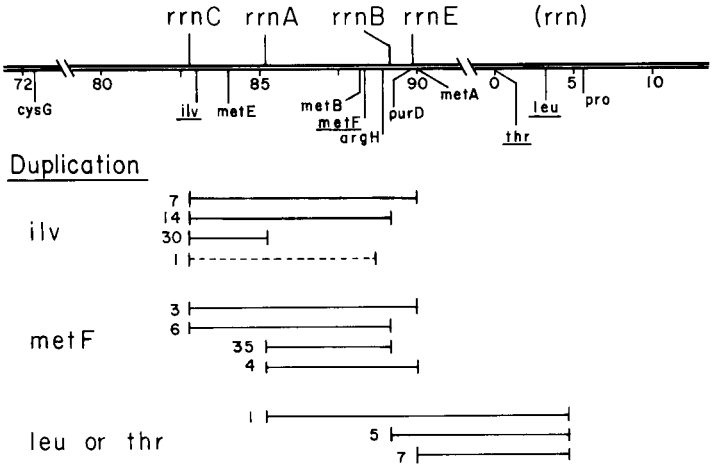

Figure 5. Mapping of duplications in the *rrn* region. Duplications of the *ilv*, *met*, *thr*, and *leu* genes were selected as described in Fig. 4 except that Tn*5* insertion auxotrophs were used. The duplications were mapped using Tn*10* auxotrophs (indicated below the genetic map). The extent of the duplicated segment is indicated by the horizontal bars below the genetic map. Each solid line depicts a duplication whose endpoints lie within a region including a *rrn* locus. Numbers adjacent to duplication bars indicate number of independent occurrences of that duplication type. The *leu* and *thr* duplications do not represent all of the duplications isolated (see text). The *leu* duplications presented are distributed 1:3:1 among the three types found; the *thr* duplications are distributed 0:2:6.

metabolic signal. At present, we are unable to determine which of these mechanisms is operating.

Acknowledgment

This work was supported by grant PCM76-15048 from the National Science Foundation.

REFERENCES

ANDERSON, R. P. and J. R. ROTH. 1977. Tandem genetic duplications in phage and bacteria. *Annu. Rev. Microbiol.* **31:**473.

———. 1978a. Tandem chromosomal duplications in *Salmonella typhimurium*: Fusion of histidine genes to novel promoters. *J. Mol. Biol.* **119:**147.

———. 1978b. Tandem genetic duplications in *Salmonella typhimurium*: Amplification of the histidine operon. *J. Mol. Biol.* **126:**53.

HILL, C. and G. COMBRIATO. 1973. Genetic duplications induced at very high frequency by ultraviolet irradiation of *E. coli*. *Mol. Gen. Genet.* **127:**197.

HILL, C., R. GRAFSTROM, B. HARNISH, and B. HILLMAN. 1977. Tandem duplications resulting from recombination between ribosomal RNA genes in *Escherichia coli*. *J. Mol. Biol.* **116:**407.

IKEMURA, T. and M. NOMURA. 1977. Expression of spacer tRNA genes in ribosomal RNA transcription units carried by hybrid ColE1 plasmids in *E. coli*. *Cell* **11:**779.

KENERLEY, M., E. MORGAN, L. POST, L. LUNDAHL, and M. NOMURA. 1977. Characterization of hybrid plasmids carrying individual ribosomal ribonucleic acid transcription units of *Escherichia coli*. *J. Bacteriol.* **132:**931.

KLECKNER, N., J. ROTH, and D. BOTSTEIN. 1977. Genetic engineering *in vivo* using translocatable drug-resistance elements: New methods in bacterial genetics. *J. Mol. Biol.* **116:**125.

KLECKNER, N., R. CHAN, B.-K. TYE, and D. BOTSTEIN. 1975. Mutagenesis by insertion of a drug-resistance element carrying an inverted repetition. *J. Mol. Biol.* **97:**561.

NIERLICH, D. 1978. Regulation of RNA synthesis. *Annu. Rev. Microbiol.* **32:**393.

NOMURA, M., E. MORGAN, and S. R. JASKUNAS. 1977. Genetics of bacterial ribosomes. *Annu. Rev. Genet.* **11:**297.