# Natural Populations of *Escherichia coli* and *Salmonella typhimurium* Harbor the Same Classes of Insertion Sequences

**Marina Biserčić and Howard Ochman**

*Department of Biology, University of Rochester, Rochester, New York 14627*

Manuscript received July 17, 1992
Accepted for publication November 12, 1992

ABSTRACT

Despite their close phylogenetic relationship, *Escherichia coli* and *Salmonella typhimurium* were long considered as having distinct classes of transposable elements maintained by either host-related factors or very restricted gene exchange. In this study, genetically diverse collections of *E. coli* and *S. typhimurium* (subgroup I) were surveyed for the presence of several classes of insertion sequences by Southern blot analysis and the polymerase chain reaction. A majority of salmonellae contained IS*1* or IS*3*, elements originally recovered from *E. coli*, while IS*200*, a Salmonella-specific element, was present in about 20% of the tested strains of *E. coli*. Based on restriction mapping, the extent of sequence divergence between copies of IS*200* from *E. coli* and *S. typhimurium* is on the order of that observed in comparisons of chromosomally encoded genes from these taxa. This suggests that copies of IS*200* have not been recently transferred between *E. coli* and *S. typhimurium* and that the element was present in the common ancestor to both species. IS*200* is polymorphic within *E. coli* but homogeneous among isolates of *S. typhimurium*, providing evidence that these species might differ in their rates of transfer and turnover of insertion sequences.

INSERTION sequences (IS) are segments of DNA, typically 1–2 kilobases (kb) in length, with the ability to transpose within and among bacterial genomes (GALAS and CHANDLER 1989). These mobile genetic elements exist on the host chromosome and on plasmids, and can mediate several types of genetic rearrangements (IIDA, MEYER and ARBER 1983). IS are widely distributed among populations and have been detected in most species of bacteria.

Despite their widespread occurrence and potential for horizontal transfer, the observed distributions of IS elements among strains of *Escherichia coli* and *Salmonella typhimurium* remain somewhat enigmatic. Early studies indicated that these two closely related enteric species maintain substantially different pools of IS elements. Natural populations of *E. coli* harbor at least 20 classes of IS (GALAS and CHANDLER 1989) and are highly polymorphic with respect to their distribution and copy numbers (SAWYER *et al.* 1987). The classes of IS detected within *E. coli* are generally not present in the species typed to Salmonella, although some have been detected in other enteric genera (NYMAN *et al.* 1981; SCHONER and SCHONER 1981; LAWRENCE, OCHMAN and HARTL 1992; but see NISEN, PURUCKER and SHAPIRO 1979). In contrast to the situation in *E. coli*, a single class of element–IS*200*–has been recovered from *S. typhimurium* and this element is widely distributed among salmonellae. Prior to this study, this element had not been detected in *E. coli* populations (LAM and ROTH 1983a,b; Gibert, Barbe and CASADESUS 1990).

The "genus-specificity" of IS*200* has been attributed to two factors that may also account for the absence of other IS elements from Salmonella: (1) *Host determinants*, IS*200* may rely upon cellular components of the Salmonella host for its maintenance and transposition while other factors may act to exclude some classes of insertion sequences from the Salmonella genome. (2) *Ancestry*, The restricted distribution of IS*200* could reflect a relatively recent origin and dissemination of the element within salmonellae. Given that the frequency of interspecific genetic exchange among enteric bacteria is relatively low (MAYNARD-SMITH, DOWSON and SPRATT 1991), there might have been insufficient time for the element to become established in other species.

To investigate the origin and ancestry of the distinct classes of IS in these bacterial species, we surveyed a collection of strains of *S. typhimurium* (subgroup I) (BELTRAN *et al.* 1991) for IS*1* and IS*3*, which are widely distributed among natural populations of *E. coli* (NYMAN *et al.* 1983; SAWYER *et al.* 1987), and for IS*200*, which had already been detected in numerous serovars of Salmonella (LAM and ROTH 1983a,b; GIBERT, BARBE and CASADESUS 1990; STANLEY, JONES and THRELFALL 1991; STANLEY, GOLDSWORTHY and THRELFALL 1992). In addition, we examined the occurrence of IS*200* among natural isolates of *E. coli*. In contrast to previous studies, we determined that natural populations of *E. coli* and *S. typhimurium* harbored three of the same classes of insertion sequences, and demonstrated that reported differences in the distri-

butions of these IS elements were the result of insufficient sampling rather than features of the host genome or the insertion sequences.

## MATERIALS AND METHODS

**Bacterial strains:** Strains of *E. coli* and *S. typhimurium* (subgroup I) were taken from the ECOR (OCHMAN and SELANDER 1984) and SARA (BELTRAN *et al.* 1991) reference collections, respectively. The ECOR collection consists of 72 natural isolates and encompasses the range of genotypic diversity found within *E. coli*. The SARA collection comprises strains representative of the diversity found within subgroup I of the *S. typhimurium* complex and includes strains of the serovars *S. typhimurium*, *S. saintpaul*, *S. heidelberg*, *S. paratyphi B* and *S. muenchen*. We examined 65 (SARA 1-4, 7-13, 15-30, 34-43, 45-72) of the 72 reference strains within the SARA collection. Both the ECOR and SARA collections contain isolates whose genealogical relationships have been inferred by protein electrophoresis (SELANDER, CAUGANT and WHITTAM 1987; BELTRAN *et al.* 1991). Species identification of each strain was confirmed biochemically on ES MicroPlates (Biolog, Inc.).

**Insertion sequence surveys:** Strains were surveyed for the presence of each class of IS elements–IS*1*, IS*3*, IS*200* and IS*630*–by the polymerase chain reaction (PCR) (SAIKI *et al.* 1988). IS elements were amplified either from purified chromosomal DNA or from lysed single colonies. For colony-based PCR, single colonies were suspended in 50 μl of $ddH_20$, and 1.0 μl of this cell suspension was used in the PCR. Lysis of cells typically occurred during the denaturation step of the PCR cycle and reactions proceeded for 30 cycles. All isolates of the *S. typhimurium* complex assayed by colony-based PCR for the presence of IS*1*, IS*3* and IS*200* harbored at least one of the classes of IS, indicating that the inability to amplify a specific element was due to the absence of sequences complementary to the PCR primers rather than poor cell lysis. Primers utilized in the amplification reactions were based on available sequences (IS*1*, OHTSUBO and OHTSUBO 1978; IS*3*, TIMMERMAN and TU 1985; IS*200*, K. HECK and J. R. ROTH, personal communication; IS*630*, MATSUTANI *et al.* 1987) and listed in the 5′ to 3′ direction:

IS*1*: Forward, GATTAGTGTATGATGG; Reverse, GATAGTGTTTTATGTTC

IS*3*: Forward, GGACACGCGGCTAAGTG; Reverse, TGGACACAGGCCTAAGCG

IS*200*: Forward, CAGATGCGCCTATAAGGCT; Reverse: CTAGGCTGGGGGTTCCGGAA

IS*630*: Forward, CCGTGACGAACGACGCCTGAT; Reverse: CTGCACTGATGATTACGCGTT

Primers were designed to anneal internally to the inverted repeats (except for IS*200*, which does not have inverted repeats) and to allow for the amplification of almost complete elements. Utilizing these primer pairs, the expected sizes of PCR products were as follows: IS*1*, 729 base pairs (bp); IS*3*, 1220 bp; IS*200*, 660 bp; IS*630*, 1038 bp. Additional primers whose sequences were complementary to the central portions of IS*1* or IS*3* were employed for some reactions. Primer 5′-ATGGACGAACAGTGGGGC-3′, whose 5′ end anneals at position 375 of IS*1*, was used in conjunction with the IS*1* reverse primer to amplify a 354-bp fragment. A 700-bp fragment within IS*3* was amplified using the forward primer 5′-CTTCTGAACGTGAACTGGA-3′ and the reverse primer 5′-GTGAACGATAACGTTCCGGG-3′ whose 5′-ends anneal at positions 241 and 940, respectively. Chromosomal DNA from ECOR 23 served as a control in the PCR involving IS*1* and IS*3* because it was known to contain both elements (SAWYER *et al.* 1987); and SARA 2, a *S. typhimurium* LT2 derivative (BELTRAN *et al.* 1991) was used as the control in the amplification of IS*200*. Amplification products were electrophoresed through 1% agarose gels, stained with ethidium bromide and visualized with UV-illumination.

**Southern blotting:** Chromosomal and plasmid DNAs from bacterial strains were digested to completion with the restriction enzyme *Hin*dIII. Fractionated DNAs were electrophoresed through 0.9% agarose gels and transferred to nylon membranes (HYBOND-N, Amersham). A hybridization probe for the detection of IS*200* was prepared by PCR amplification of the element from *S. typhimurium* LT2 DNA using the primers listed above. The 660-bp product was precipitated by the addition of ethanol, labeled to high specific activity by the method of FEINBERG and VOGELSTEIN (1983) and hybridized to membranes containing digested bacterial DNAs for 16 hr at 65° in a high phosphate buffer (0.5 M NaCl, 0.1 M $NaH_2P0_4$, 5 mM EDTA) supplemented with 0.2% Sarkosyl. Filters were washed at 20° for a total of 30 min in five changes of 10 mM Tris, 1 mM EDTA (pH 8.0) prior to autoradiography.

**RFLP analysis:** IS*200* elements were mapped with the following restriction endonucleases: *Eco*RI, *Eco*RV, *Hae*III, *Hph*I, *Sau*3A and *Taq*I. All amplication reactions yielded a single PCR product, which was subsequently used in restriction enzyme digests without additional purification. For enzyme digests, 3 μl of the PCR product was diluted in 15 μl of $ddH_2O$, mixed with 2 μl of the appropriate reaction buffer and digested according to the manufacturers' specifications (Boehringer Mannheim; New England Biolabs). Digested DNAs were resolved by electrophoresis on 1.5% agarose gels and their RFLP patterns compared with that predicted from the sequence of IS*200* from *S. typhimurium* LT2.

## RESULTS

**Insertion sequences within Salmonella:** We initially screened 40 strains from the SARA collection for the presence of IS*1*, IS*3* and IS*200*. In this set, every strain tested contained at least one of these elements, and four harbored all three classes of IS (Figure 1).

**IS*200* in Salmonella:** IS*200*, which was originally detected in *S. typhimurium* LT2, was present in 24 of the 40 *S. typhimurium* (subgroup I) strains tested. The element occurred in all strains typed to *S. typhimurium* and *S. heidelberg*, but was absent from the *S. paratyphi* B and *S. muenchen* serovars (Figure 1). Only one strain of *S. saintpaul* contained IS*200* (SARA 22). Due to the diverse relationships of strains typed to the different serovars of Salmonella, we expanded this analysis to include additional strains of *S. paratyphi* (SARA 45–62), *S. heidelberg* (SARA 35, 39 and 40) and *S. muenchen* (SARA 63). In this augmented set of 22 strains, those typed to *S. paratyphi* and *S. heidelberg* were polymorphic for the presence of IS*200*. Of 19 *S. paratyphi* strains, only those most closely related to *S. typhimurium* (SARA 57–60) contained IS*200* (delineated as group **C** in BELTRAN *et al.* 1991). *S. muenchen* (SARA 63) and one strain of *S. heidelberg* (SARA35) did not contain IS*200*. The distribution of IS*200*

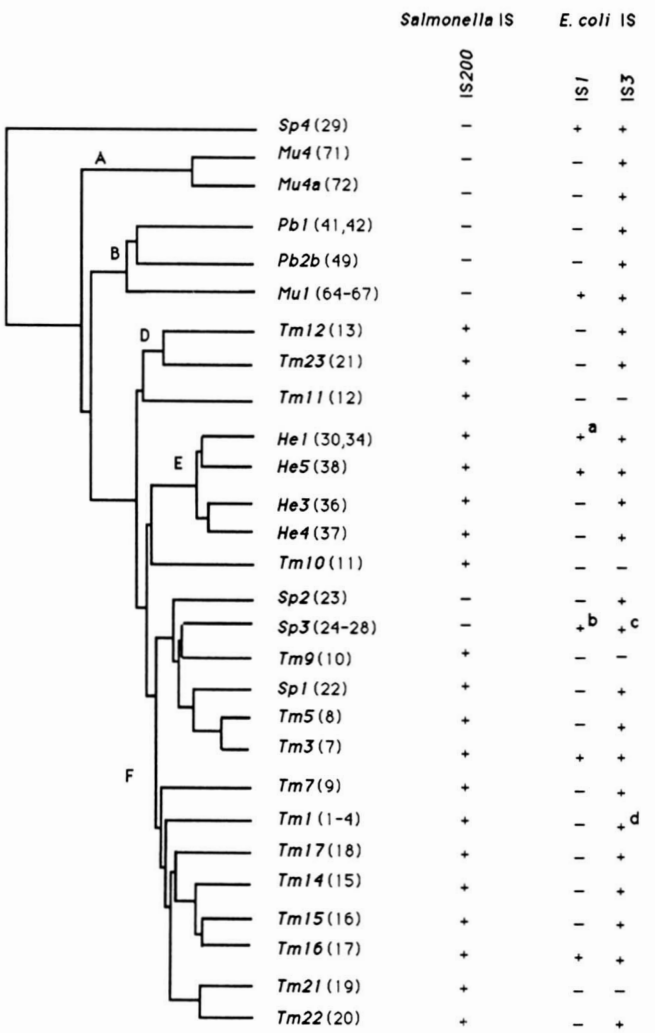


FIGURE 1.—Phylogenetic tree showing the relationships among strains of the *Salmonella typhimurium* complex (SARA collection) containing the insertion sequences IS*1*, IS*3* and IS*200*. The presence or absence of a particular class of elements is indicated by "+" or "−," respectively. The tree topology, strain numbers (in brackets) and serovar abbreviations (Tm, *S. typhimurium*; Sp, *S. saintpaul*; He, *S. heidelberg*; Pb, *S. paratyphi B*; Mu, *S. muenchen*) are adapted from BELTRAN *et al.* (1991). [a] Only SARA 34 contains IS*1*; [b] SARA 24 through 28 contain IS*1*; [c] Only SARA 24 and 25 contain IS*3*; [d] Only SARA 1, 2 and 4 contain IS*3*.

within the SARA collection broadly follows the phylogenetic relationships established by protein electrophoresis (BELTRAN *et al.* 1991). None of the strains included in groups **A** and **B** contain IS*200* while a majority of those in groups **C** through **F** have the element.

**IS*1* and IS*3* in Salmonella:** Serovars of *S. typhimurium* were surveyed for the presence of IS*1* and IS*3*, elements originally isolated from *E. coli*. As shown in Figure 1, IS*1* was detected in 11 of the 40 strains, and in each case the amplification yielded a single product of about 729 bp, corresponding to the size predicted from the nucleotide sequence of IS*1* from *E. coli* K12. To further determine the extent of similarity between the IS*1* detected in *S. typhimurium* and those isolated

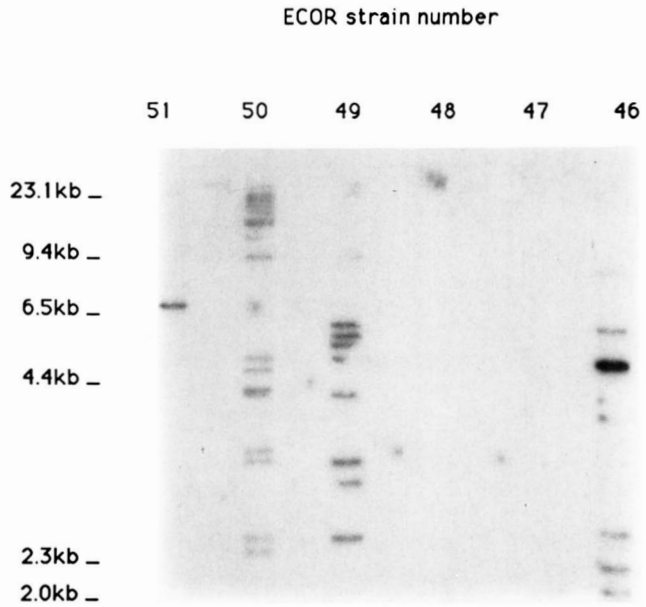


FIGURE 2.—Southern blot of natural isolates of *E. coli* from the ECOR collection (OCHMAN and SELANDER 1984) probed for IS*200*. DNAs were digested to completion with *Hin*dIII, and hybridized with IS*200* prepared by PCR from *S. typhimurium* LT2, as described in the Materials and Methods. ECOR strain numbers are listed above each lane.

from *E. coli*, we performed an additional amplification reaction employing a primer complementary to the center of IS*1* on a subset of the SARA strains (SARA 7, 17, 34 and 64). And in all cases these secondary reactions amplified a fragment of proper length (354 bp).

A total of 31 of the 40 tested strains within *S. typhimurium* complex contain IS*3* (Figure 1). To establish if the IS*3* elements in these salmonellae were homologous to those present in *E. coli*, we attempted to amplify an internal segment of IS*3*, of 700 bp in length, from representative strains SARA 8, 15, 34, 64 and 71. This fragment, however, could only be amplified from the control strain of *E. coli* (ECOR 23) and not from the representatives of the SARA collection.

**Insertion sequences in *E. coli*.** (a) IS*200* and IS*630*: All strains of *E. coli* from the ECOR collection were screened for the presence of IS*200* by Southern blot analysis. A total of 14 strains contained chromosomal sequences homologous to IS*200* (ECOR 1, 8, 16, 46, 49, 50, 51, 53, 58, 64, 65, 66, 70, 72). These strains were polymorphic with respect to the distribution and copy numbers of the element, and no plasmid-borne copies of IS*200* were detected. Representative examples of IS*200* profiles are shown in Figure 2. Multiple copies of IS*200* were detected in ECOR 46 (7 copies), 49 (8 copies), 50 (11 copies), 64 (3 copies) and 66 (8 copies). All other ECOR strains harboring IS*200* had single copies of the element; and ECOR 51 and 53, which are closely related based on electrophoretic

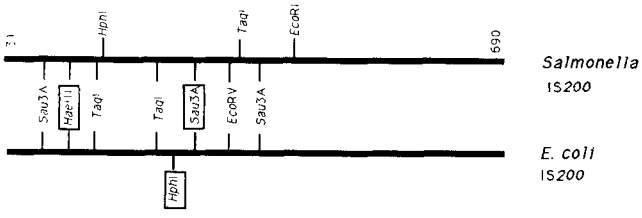


FIGURE 3.—Comparative restriction analysis of IS*200* from *E. coli* (ECOR 8, 46, 49, 50, 51, 53, 58, 64, 65, 66, 72) and the *S. typhimurium* complex (SARA 2, 7, 17, 22, 34, 36, 38, 57, 60). Restriction maps of IS*200* begin at base position 31 and end at 690 of the element. Restriction enzymes listed between the two maps represent the positions of recognition sites conserved in the elements from both species while sites on the outside indicate those that are unique to the IS*200* from Salmonella (above) or *E. coli* (below). IS*200* polymorphisms detected within *E. coli* are boxed: the *Hae*III (boxed) site is not present in ECOR 72 and the *Sau*3A (boxed) site in *Salmonella* occurs in ECOR 49. An additional haplotype of IS*200* present in ECOR 46 lacks the boxed *Hae*III and *Hph*I sites.

profiles, displayed identical patterns of hybridization.

Enzymatic amplification of IS*200* from these strains of *E. coli* yielded products of the same length as those from *S. typhimurium* (660 bp). Using primers complementary to the IS*200* of *S. typhimurium* LT2, this element could be amplified from all ECOR strains shown to harbor the element by Southern blot analysis, except ECOR 1, 16 and 70. *E. coli* strains containing IS*200* were also screened by the PCR for the presence of IS*630*, an IS element originally detected in Shigella and thought to be related to IS*200* (MATSUTANI *et al.* 1987; GIBERT, BARBE and CASADESUS 1990). Only ECOR 51 and ECOR 53 harbor copies of IS*630* in addition to IS*200*.

(b) IS*200* polymorphisms: To determine the degree of similarity between the IS*200* from *E. coli* and *S. typhimurium*, we produced restriction maps of the elements from the 11 strains of *E. coli* and from a representative subset of the *S. typhimurium* complex (SARA 2, 7, 17, 22, 34, 36, 38, 57 and 60). Amplified elements were digested with six restriction enzymes yielding a total of 11 restriction sites. Restriction maps of IS*200* were invariant within salmonellae but were polymorphic among strains of *E. coli*. There were four haplotypes of IS*200* present in *E. coli*, and each haplotype is distinguishable by a single restriction site polymorphism. The IS*200* from ECOR 46, 49 and 72 differed from those of other ECOR strains at sites noted in Figure 3, and each strain contained a unique IS*200* haplotype. The fourth haplotype of IS*200* comprised ECOR 1, 8, 16, 50, 51, 53, 58, 64, 65, 66 and 70.

The tested strains of *S. typhimurium* (subgroup I) all contained the same haplotype of IS*200*, which was distinct from those detected in *E. coli* (Figure 3). (The IS*200* from *S. typhimurium* yielded a restriction pattern differing from that expected on the basis of nucleotide sequence: *Taq*I produced an additional band of approximately 120 bp.) Five restriction sites were polymorphic between the IS*200* of *E. coli* and *S. typhimurium*, and the average level of nucleotide sequence divergence observed between the haplotypes of IS*200* in *E. coli* and the single haplotype detected within the *S. typhimurium* complex was 0.083 ± 0.018 (NEI 1987). Although some strains of *E. coli* were known to harbor multiple copies of IS*200* (Figure 2), the elements within each strain–except for ECOR 46–yielded a uniform RFLP pattern. The *Hae*III and *Hph*I digests of IS*200* amplification products from ECOR 46 yielded RFLP patterns containing both uncut and cleaved fragments. Based on the cleavage patterns of several other enzymes in ECOR 46, as well as the fidelity of *Hae*III and *Hph*I in all other reactions, these IS*200* polymorphisms within ECOR 46 can not be attributed to incomplete enzyme digestions and indicate that this strain harbors two different forms of IS*200*.

## DISCUSSION

Most previous studies considered Salmonella as having a single class of insertion sequences—IS*200*—which was thought to be largely confined to that genus (LAM and ROTH 1983a,b; GIBERT, BARBE and CASADESUS 1990). By surveying collections of natural isolates, we have established that *E. coli* and *S. typhimurium* contain three of the same classes of insertion sequences—IS*1*, IS*3* and IS*200*—and that the previously reported differences were artifacts of limited sampling. The observed distributions of IS*1*, IS*3* and IS*200* could arise if each class of elements were present in the common ancestor of *E. coli* and *S. typhimurium* or could be the result of lateral transfer after these species diverged.

**Distribution and ancestry of IS*200*:** Based on restriction mapping, the IS*200* elements from *E. coli* and *S. typhimurium* are genotypically distinct. The level of nucleotide divergence between IS*200* from *E. coli* and *S. typhimurium* is on the order of that observed for pairs of homologous chromosomal genes from the two species, which suggests that IS*200* is ancestral to both species. [The amount of nucleotide sequence divergence between homologous genes from *E. coli* K12 and *S. typhimurium* ranges from 1% to 27.5%, with an average of 15.6% (SHARP 1991), while the extent of sequence divergence between the haplotypes of IS*200* from these species, as estimated from restriction maps, is about 8%.]

The populations of IS*200* detected in *E. coli* and in *S. typhimurium* differ with respect to their degree of polymorphism. Based on the restriction mapping of elements from several serovars, IS*200* appears to be genetically homogenous within the *S. typhimurium* complex. In contrast, we identified four IS*200* haplotypes within the collection of *E. coli* isolates, none

of which was detected in salmonellae. This variation suggests that the dynamics of IS*200* differ in these closely related bacterial species. Such differences could result from the nucleotide substitutions detected in the elements or from host factors, both of which might affect rates of transposition. These possibilities can be assessed by nucleotide sequence analysis of elements from strains of known phylogenetic relationships and by the experimental analysis of transposition rates (*e.g.*, CASADESUS and ROTH 1990; LAWRENCE, OCHMAN and HARTL 1992).

Unlike other insertion sequences, IS*200* lacks inverted terminal repeats and does not generate duplications of host DNA sequences upon insertion. It has been postulated that transposition of IS*200* may be constrained in some hosts (CASADESUS and ROTH 1989), although strains may harbor as many as 25 copies of the element (LAM and ROTH 1983a,b; GIBERT, BARBE and CASADESUS 1990). The distribution and copy numbers of IS*200* within *E. coli* are similar to those seen for other classes of insertion sequences (SAWYER *et al.* 1987). However, *E. coli*, and most salmonellae (STANLEY, JONES and THRELFALL 1991), lacks plasmid-borne copies of IS*200*, which could account for its low rate of transfer both within and among species. That IS*200* has become established within the genomes of numerous strains of *E. coli*, Shigella and Salmonella (GIBERT, BARBE and CASADESUS 1990) suggests that species-specific factors are not required for its maintenance and transposition.

**Distribution and Ancestry of IS*1* and IS*3*:** The presence of IS*1* and IS*3* in a large number of salmonellae is somewhat unexpected in that only a single class of insertion sequence had previously been detected in representative strains of this genus. IS*1* and IS*3* are prevalent in natural populations of *E. coli* (SAWYER *et al.* 1987), and the present study revealed that IS*3* was more prevalent in *S. typhimurium* strains from the SARA collection than IS*200*. Despite the distribution and abundance of IS*3* among enteric species, the frequency of horizontal transfer of this element between bacteria typed to different species appears to be low (LAWRENCE, OCHMAN and HARTL 1992). For example, the inferred phylogenetic relationships of homologues of IS*3* from several species of enteric bacteria are consistent with those derived from chromosomal genes. It is currently not known whether the IS*3* elements in Salmonella are ancestral, or obtained through horizontal transfer, but our inability to amplify a central portion of this element (using oligonucleotide primers based on the sequence of IS*3* from *E. coli*) suggests that the IS*3* from *S. typhimurium* differs from that detected in *E. coli*.

Aside from *E. coli* and the *S. typhimurium* complex, IS*1* has been detected in four species of Shigella (OHTSUBO *et al.* 1984), *Escherichia fergusonii*, *E. hermanii*, *E. vulneris* (although typed as Escherichieae, these species have diverged substantially; see LAWRENCE, OCHMAN, and HARTL 1992), *Klebsiella arogenes* and *Serratia marcescens* (NISEN, PURUCKER and SHAPIRO 1979; NYMAN *et al.* 1981). The nucleotide sequences of certain forms of IS*1* are well-conserved in many of these species, suggesting that this element, unlike IS*3*, has been transferred among distantly related lineages (LAWRENCE, OCHMAN and HARTL 1992); however, the origins of IS*1* within Salmonella can not be fully established without additional nucleotide sequence information.

Cumulatively, these studies have provided a new perspective on the distribution of insertion sequences among natural isolates of *E. coli* and Salmonella. The full extent of overlap as well as the ancestry of the classes of IS elements common to *E. coli* and Salmonella are presently not known. By comparing the nucleotide sequences, evolutionary relationships and the abundance of homologous elements within these species, it will be possible to determine some of the factors influencing the rates of transfer and accumulation of transposable elements within natural populations of bacteria.

## LITERATURE CITED


BELTRAN, P., S. A. PLOCK, N. H. SMITH, T. S. WHITTAM, D. C. OLD and R. K. SELANDER, 1991 Reference collection of strains of the *Salmonella typhimurium* complex from natural populations. J. Gen. Microbiol. **137:** 601–606.

CASADESUS, J., and J. R. ROTH, 1989 Absence of insertions among spontaneous mutants of *Salmonella typhimurium*. Mol. Gen. Genet. **216:** 210–216.

FEINBERG, A. P., and B. VOGELSTEIN, 1983 A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:** 6–13.

GALAS, D. J., and M. CHANDLER, 1989 Bacterial insertion sequences, pp. 109–162 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society for Microbiology, Washington, D.C.

GIBERT, I., J. BARBE and J. CASADESUS, 1990 Distribution of insertion sequence IS*200* in *Salmonella* and *Shigella*. J. Gen. Microbiol. **136:** 2555–2560.

IIDA, S., J. MEYER and W. ARBER, 1983 Prokaryotic IS elements, pp. 159–221 in *Mobile genetic elements*, edited by J. A. SHAPIRO. Academic Press, New York.

LAM, S., and J. R. ROTH, 1983a Genetic mapping of IS*200* copies in *Salmonella typhimurium* LT2. Genetics **105:** 801–811.

LAM, S., and J. R. ROTH, 1983b IS*200*: A Salmonella-specific insertion sequence. Cell **34:** 951–960.

LAWRENCE, J. G., H. OCHMAN and D. L. HARTL, 1991 Molecular and evolutionary relationships among enteric bacteria. J. Gen. Microbiol **137:** 1911–1921.

LAWRENCE, J. G., H. OCHMAN and D. L. HARTL, 1992 The evolution of insertion sequences within enteric bacteria. Genetics **131:** 9–20.

MATSUTANI, S., H. OHTSUBO, Y. MAEDA and E. E. OHTSUBO, 1987 Isolation and characterisation of IS elements repeated in the bacterial chromosome. J. Mol. Biol. **196:** 445–455.

MAYNARD SMITH, J., C. G. DOWSON and B. G. SPRATT, 1991 Localized sex in bacteria. Nature **349:** 29–31.

NEI, M., 1987 *Molecular Evolutionary Genetics*, Columbia University Press, New York.

NISEN, P., M. PURUCKER and L. SHAPIRO, 1979 Deoxyribonucleic acid sequence homologies among bacterial insertion sequence elements and genomes of various organisms. J. Bacteriol. **140:** 588–596.

NYMAN, K, K. NAKAMURA, H. OHTSUBO and E. OHTSUBO, 1981 Distribution of the insertion sequences IS*1* in Gram-negative bacteria. Nature **289:** 609–612.

NYMAN, K, H. OHTSUBO, D. DAVISON and E. OHTSUBO, 1983 Distribution of insertion element IS*1* in natural isolates of *Escherichia coli*. Mol. Gen. Genet. **189:** 516–518.

OCHMAN, H., and R. K. SELANDER, 1984 Standard reference strains of *Escherichia coli* from natural populations. J. Bacteriol. **157:** 690–693.

OHTSUBO, H., and E. OHTSUBO, 1978 Nucleotide sequence of an insertion element IS*1*. Proc. Natl. Acad. Sci. USA **75:** 615–619.

OHTSUBO, E., H. OHTSUBO, W. DOROSZKIEWICZ, K. NYMAN, D. ALLEN and D. DAVISON, 1984 An evolutionary analysis of *iso*-IS*1* elements from *Escherichia coli* and *Shigella* strains. J. Gen. Appl. Microbiol. **30:** 359–376.

SAIKI, R. K., D. H. GELFAND, S. STOFFEL, S. J. SCHARF, R. HIGUCHI, G. T. HORN, K. B. MULLIS and H. A. ERLICH, 1988 Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239:** 487–491.

SAWYER, S. A., D. DYKHUIZEN, R. DUBOSE, L. GREEN, T. MUTANGADURA-MHLANGA, D. WOLCZYK and D. HARTL, 1987 Distribution and abundance of insertion sequences among natural isolates of *Escherichia coli*. Genetics **115:** 51–63.

SCHONER, B. and R. G. SCHONER, 1981 Distribution of IS*5* in bacteria. Gene **16:** 347–352.

SELANDER, R. K., D. A. CAUGANT and T. S. WHITTAM, 1987 Genetic structure and variation in natural populations of *Escherichia coli*, pp. 1625–1648 in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, edited by F. C. NEIDHARDT. American Society for Microbiology, Washington, D. C.

SHARP, P.M., 1991 Determinants of DNA sequence divergence between *Escherichia coli* and *Salmonella typhimurium*: codon usage, map position, and concerted evolution. J. Mol. Evol. **33:** 23–33.

STANLEY, J., C. S. JONES and E. J. THRELFALL, 1991 Evolutionary lines among *Salmonella enteritidis* phage types are identified by insertion sequence IS*200* distribution. FEMS Microbiol. Lett. **82:** 83–90.

STANLEY, M. GOLDSWORHTY and E. J. THRELFALL, 1992 Molecular phylogenetic typing of pandemic isolates of *Salmonella enteritidis*. FEMS Microbiol. Lett. **90:** 153–160.

TIMMERMAN, K., and C. P. D. TU, 1985 Complete sequence of IS*3*. Nucl. Acids Res. **13:** 2127–2139.

Communicating editor: A. G. CLARK