

Native Insertion Sequence Elements: Locations, Distributions, and Sequence Relationships

RICHARD C. DEONIER

111

INTRODUCTION

Insertion sequences (ISs) are transposable elements whose only genes are related to promotion and regulation of transposition. IS elements typically fall within the 700- to 2,000-bp size range, and they are normal constituents of many bacterial chromosomes and plasmids. IS elements were originally recognized by the phenotypes resulting from their insertion into other genes. An excellent overview of the earlier studies was presented by Starlinger and Saedler (55). The biology, structure, regulation, transposition mechanisms, and other attributes of many IS elements have been comprehensively reviewed by Galas and Chandler (18).

IS elements participate in a range of events. Transpositional processes include insertion, adjacent deletion (a consequence of intramolecular transposition in many cases), and transpositional inversion of adjacent genes. RecA-mediated processes such as inversion, cointegration, or deletion involving pairs of identical IS elements are a natural consequence of the presence of multiple copies of particular IS elements. Precise excision (removal of the element and one of the flanking direct repetitions) is unrelated to transposition. The multiplicity of transpositional and recombinational events associated with IS elements allows them to contribute to plasticity of bacterial chromosomes and plasmids in which they are found.

The magnitude and nature of influences by IS elements on a bacterial chromosome will depend upon the exact types of IS elements that are present. For example, *Salmonella typhimurium* (official designation, *Salmonella enterica* serovar Typhimurium) LT2 appears to contain only IS200, and this particular element transposes infrequently. In contrast, *Escherichia coli* K-12 contains 8 to 12 different types of IS element, and they contribute significantly to the collection of mutations that have been obtained for various genes. For example, in the *lacI* gene (excluding a hot spot for point mutation), two IS1 insertions were detected in a mutant pool containing 24 point mutations (and other deletion or frameshift mutations) (16), and in a collection of 25 *cI* mutations in λ lysogens, 15 were caused by IS insertions (32). Because of their ability to induce mutations and genome rearrangements, the identification of IS elements and determination of their locations on the genome map are necessary for any global view of the genetics of a given bacterial type.

This chapter focuses on the distribution and sequence relationships of chromosomal IS elements in *E. coli* and *Salmonella* spp., particularly in the common laboratory bacteria *E. coli* K-12 and *S. typhimurium* LT2. Nine of the most significant of these elements are listed in Table 1, and representative structures are illustrated in Fig. 1. Similar elements are found in a broad spectrum of other bacteria, including archaeobacteria. A detailed discussion of IS1, IS10, and IS911 structures and transposition mechanisms is provided in chapter 124 of this volume.

MAPPING IS ELEMENTS

The IS elements that have been mapped in the *E. coli* K-12 and *S. typhimurium* LT2 chromosomes are listed in Table 1. Presence and absence of cleavage sites for several restriction endonucleases are indicated to serve as a guide for preliminary identification of these elements when new transpositions in these particular bacterial strains are identified. Isoforms for some of these elements are known, and these sequence variants may differ in restriction patterns from those given.

TABLE 1 Insertion sequences from *E. coli* K-12 and *S. typhimurium* LT2 and their sensitivities to cleavage by selected diagnostic restriction endonucleases^a

Element	Length (bp)	GenBank accession no. ^b	Cleavage by:						
			<i>Ava</i> I	<i>Bam</i> HI	<i>Bgl</i> II	<i>Bst</i> EII	<i>Eco</i> RI	<i>Hind</i> III	<i>Pst</i> I
IS1	768	V00609	-	-	-	+	-	-	+
IS2	1,327	V00610	+	-	-	-	-	+	-
IS3	1,258	X02311	-	-	-	+	-	+	+
IS4	1,426	J01733	+	-	-	-	-	-	+
IS5	1,195	J01734	-	-	+	+	+	-	-
IS30	1,221	X00792	-	-	+	-	-	+	-
IS150	1,443	X07037	-	-	+	+	-	+	-
IS186	1,338	X03123	+	+	-	-	-	-	+
IS200	707	X56834	-	-	-	-	+	+	-

^aNote that not all IS elements of a given type in these bacteria need have identical sequences. For example, there are four different sequence types for IS1 in the *E. coli* K-12 chromosome (61). Consequently, some variants of these elements may have an altered spectrum of sensitivity to the indicated restriction endonucleases.

^bTypical type sequence. Additional accession numbers for individual chromosomal elements of each type may also be available.

IS Elements in the *E. coli* K-12 Chromosome

Because IS elements can transpose, their status as genetic loci is less secure than that of conventional genes. This is emphasized by the absence of some types of elements from particular *E. coli* strains but not from others (49). Nevertheless, the positions of some IS elements of *E. coli* K-12 are sufficiently conserved that they appear in most commonly used contemporary isolates.

Earlier IS mapping experiments were directed at limited regions of the chromosome using heteroduplex or restriction mapping of F' plasmids (20) or specialized restriction mapping approaches (see, for example, reference 29). More recent studies have mapped most of the chromosomal IS elements in the entire chromosome by screening lambda or cosmid libraries representing the entire genome (3, 4, 58–60). Unfortunately, the *E. coli* K-12 strain W3110 used to make the λ libraries (27) contained chromosomal rearrangements, and strain BHB2600 used to make the cosmid library had a pedigree that involved extended laboratory manipulation. Some of the mapped IS elements may be peculiar to one or the other of these strains, given the propensity of IS elements to transpose and to induce genome rearrangements (e.g., adjacent deletion). Strain MG1655 is considered to be a better representative of wild-type *E. coli* K-12 (43).

The set of IS elements that were present in the original *E. coli* K-12 isolate represents the starting point from which the complement of insertion sequences in any contemporary laboratory strain has evolved, and members of this original, "standard" set are likely to be present in them. During the past 50 years of experimentation with *E. coli* K-12, transpositions have occurred in many lineages, particularly during storage on nutrient agar stabs. Since chromosomes lacking transpositional alterations with respect to at least one IS type may not be available, the likely original sets of IS1, IS2, IS3, IS5, and IS30 were estimated from the contemporary sets of nine laboratory strains by using parsimony arguments and the known pedigrees.

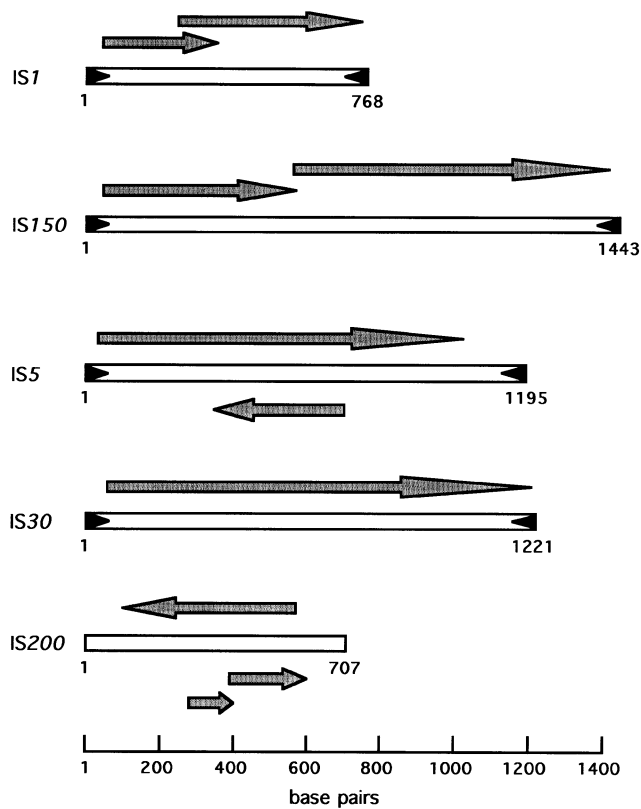


FIGURE 1 Structures of some typical IS elements from *E. coli* and *S. typhimurimum*. Shaded arrows designate the most important ORFs, and filled triangles represent the terminal inverted repetitions. Terminal inverted repetitions may be imperfect, and other shorter ORFs that might be present have been omitted (see reference 18 for complete descriptions). IS510 is the member of the IS3 family, whose members include IS3, IS2, IS150, IS911, and other elements(51). IS5 is the member of the IS4 family, whose members include IS4, IS5, IS186, and other elements (45). Data for IS200, which lacks terminal inverted repetitions, were supplied by J. Casadesus (personal communication).

TABLE 2 Presence or absence of *Eco*RI fragments containing various IS elements in selected *E. coli* K-12 strains

Element	Fragment size (kb)	Presence or absence in strain ^a :									Basic element ^b
		K-12 (CGSC 5073)	W1485F ⁻	W3110	W6F ⁻	χ5 9	BHB2600	χ101F ⁻	HB101		
		Reference 58	Reference 3								
IS1	31	1	0	0	0	0	0	0	0	0	
	26	0	1	1	1	0	1	1	1	1	IS1F
	22	1	0	0	0	0	1	1	0	0	
	20	1	1	1	1	1	1	1	1	1	IS1A
	18	0	0	0	0	0	0	1	0	0	
	15	0	0	0	0	0	0	1	0	0	
	15	?	?	1	1	?	?	1	?	1	IS1B
	14	0	0	0	0	0	0	1	0	0	
	14	0	0	0	0	0	0	1	1	0	
	14	1	0	1	1	0	0	1	0	1	IS1E ^c
	13.1	0	0	0	0	0	1	0	1	0	
	12.7	0	0	0	0	0	1	0	0	0	
	12.0	0	1	0	0	0	0	1	0	0	
	10.5	1	0	0	0	0	0	1	0	0	
	8.0	0	0	0	0	0	0	0	0	1	
	4.8	0	0	?	1	0	0	0	0	0	
	3.8	0	1	1	1	1	1	1	?	1	IS1C
3.7	1	1	1	1	1	1	1	?	1	IS1D	
IS2	39	0	0	1	0	0	0	0	0	0	
	25–30	0	0	0	1	1	0	0	0	0	
	20	1	1	1	1	1	1	1	1	1	IS2A
	17	0	0	0	0	1	0	0	1	0	
	12.5	0	1	1	1	1	1	1	1	1	IS2K
	12.0	0	?	1	1	1	0	0	0	0	
	11.0	1	1	1	1	1	1	1	1	1	IS2I
	9.8	0	0	1	1	0	0	0	0	0	
	9.1	0	0	0	1	0	0	0	0	0	
	8.7	1	1	1	1	1	1	1	1	1	IS2F
	8.2	0	0	0	0	0	0	1	1	1	
	7.9	0	0	1	1	0	0	0	0	0	
	7.0	0	0	0	0	1	0	0	0	0	
	6.3	1	?	1	1	1	?	0	?	0	IS2J
	5.5	1	0	0	1	0	0	1	0	1	
	4.9	0	0	1	0	0	0	0	0	0	
	4.0	0	1	1	1	1	1	1	1	1	IS2E
3.0	1	1	?	1	1	1	1	1	1	IS2D	
2.9	0	0	0	1	0	0	0	0	0		
IS3	18	—	1	1	1	1	0	0	1	1	IS3A
	17	—	0	0	1	0	0	0	0	0	
	13.5	—	1	1	1	1	1	1	1	0	IS3C
	12.5	—	0	0	0	0	0	1	0	0	
	12.3	—	1	1	1	1	1	1	1	0	IS3D
	12.0	—	0	0	0	0	0	1	0	0	
	11.8	—	1	1	1	1	1	1	1	1	IS3B
	10.5	—	0	0	0	0	0	0	0	1	
	8.7	—	0	0	0	0	0	0	0	1	
	8.0	—	0	0	0	1	0	0	0	0	
	5.9	—	1	1	1	0	1	1	1	1	IS3E
IS5	1.9	—	0	0	0	0	1	1	0	0	
	24	—	—	1	1	—	—	1	—	1	IS5T ^d
	—	—	—	1	0	—	—	0	—	1	
	13.5	—	—	1	1	—	—	1	—	0	IS5B
	12.9	—	—	0	1	—	—	1	—	1	“IS5Y”
	11.4	—	—	1	1	—	—	1	—	1	IS5H
	11.0	—	—	1	0	—	—	0	—	1	
	10.5	—	—	1	1	—	—	1	—	1	IS5LO
	8.9	—	—	1	1	—	—	1	—	1	IS5D
	8.0	—	—	1	0	—	—	0	—	0	
	7.7	—	—	0	1	—	—	0	—	0	
7.6	—	—	0	1	—	—	0	—	0		
7.4	—	—	1	1	—	—	1	—	1	IS5K	
5.9	—	—	1	1	—	—	1	—	1	IS5R	
5.9	—	—	1	0	—	—	0	—	0		

TABLE 2 Continued

Element	Fragment size (kb)	Presence or absence in strain ^a :								Basic element ^b	
		K-12 (CGSC 5073)	W1485F ^c	W3110	W6F ^c	χ59	BHB2600	χ101F ^c	HB101		
				Reference 58	Reference 3						
	5.7	—	—	1	1	—	—	1	—	1	IS5A
	5.5	—	—	0	1	—	—	0	—	0	
	5.1	—	—	1	1	—	—	0	—	0	
	5.0	—	—	0	0	—	—	1	—	0	
	4.9	—	—	1	0	—	—	0	—	1	
	4.6	—	—	0	0	—	—	0	—	0	
	4.0	—	—	0	1	—	—	1	—	1	“IS5X”
	3.8	—	—	0	1	—	—	0	—	0	
	3.8	—	—	1	1	—	—	0	—	0	
	3.2	—	—	0	0	—	—	0	—	1	
	2.9	—	—	1	1	—	—	1	—	1	IS5E
	2.7	—	—	1	0	—	—	0	—	1	
	1.8	—	—	0	1	—	—	0	—	1	
	1.6	—	—	0	0	—	—	0	—	1	
	1.3	—	—	0	1	—	—	0	—	0	
	1.2	—	—	1	1	—	—	0	—	0	
	1.1	—	—	1	1	—	—	0	—	1	

^a1, presence; 0, absence; ?, presence uncertain because of hybridization from a vector containing an IS probe; —, data not available. Data for *EcoRI* fragment sizes derived from each strain are taken directly from references 3, 4, 15, 23, and 57 or from maps of cloned fragments (27) for data from references 58–60.

^bBasic elements are those inferred by parsimony to probably have been present in the original isolate of *E. coli* K-12. Names enclosed in quotation marks are provisional.

^cMay or may not have been ancestral; parsimony unable to discriminate between these alternatives.

^dThe 31-kb fragment from W3110 (58) was considered to be equivalent to the 24-kb fragments from the other strains, based on map location and larger experimental error in this size range.

The following procedure was used. Sizes of *EcoRI* fragments containing these elements have been reported (or can be inferred) for a number of *E. coli* K-12 strains (3, 4, 23, 57–60). (Other published data for *E. coli* K-12 restriction fragments containing IS elements were not used because accurate sizes had not been presented.) The data for IS1, IS2, IS3, and IS5 are summarized in Table 2. If fragments hybridizing to an IS had identical sizes (within estimated experimental error) for two or more different strains, the fragments were assumed to be identical. This assumption was not made if fragments of apparently identical size appeared only in strains widely separated by pedigree. For example, if all strains contained a 20-kb *EcoRI* fragment hybridizing to IS2, and if a 20-kb fragment mapping at 8.3 min had been shown to contain IS2A, then all strains were assumed to contain IS2A. Undoubtedly there are particular strains for which this assumption is wrong, but this would require (i) that IS2A be lost and (ii) that another IS2 transpose to generate a fragment of identical size by insertion or adjacent deletion.

Given the assumptions stated above, the presence or absence of IS elements in the collection of nine strains was scored for IS1 to IS3, and four strains were scored for the presence of IS5 (Table 2; data for IS30 not shown). The phylogenetic relationships of these strains are known (Fig. 2), thanks to committed and meticulous detective work by Barbara J. Bachmann (chapter 133; see the legend to Fig. 2 for a discussion regarding nodes of the tree). The phylogenetic tree was used together with the complement of IS elements in the individual strains to infer the most parsimonious set of IS elements present in the original *E. coli* K-12. The scoring at each node in Fig. 2 is shown for a 5.5-kb *EcoRI* fragment containing IS2 (Table 2) and found in all three strains studied by Birkenbihl and Vielmetter (3) and in *E. coli* K-12 (*E. coli* Genetic Stock Center strain CGSC 5073 [23]). The assigned score is 1 if the element is inferred to be present in a strain or at a node, 0 if the element is inferred to be absent, or 1,0 if a decision cannot be made. This 5.5-kb fragment was detected in four of the nine strains, but by parsimony arguments, it was considered not to have been present in the original *E. coli* K-12. Assuming that the excision frequency is not greater than the insertion frequency (see below), the most parsimonious explanation for the observed pattern of strains having this fragment invokes four

transposition events: insertions in one of the W3110 strains, in BHB2600, in HB101, and in CGSC 5073. If there had been an IS2 in an identical 5.5-kb fragment of the original *E. coli* K-12, then five excisions would have been required to generate the observed pattern. In some cases, such as IS2B, one cannot determine whether the element was ancestral from the set of strains in Fig. 2 and the data in Table 2.

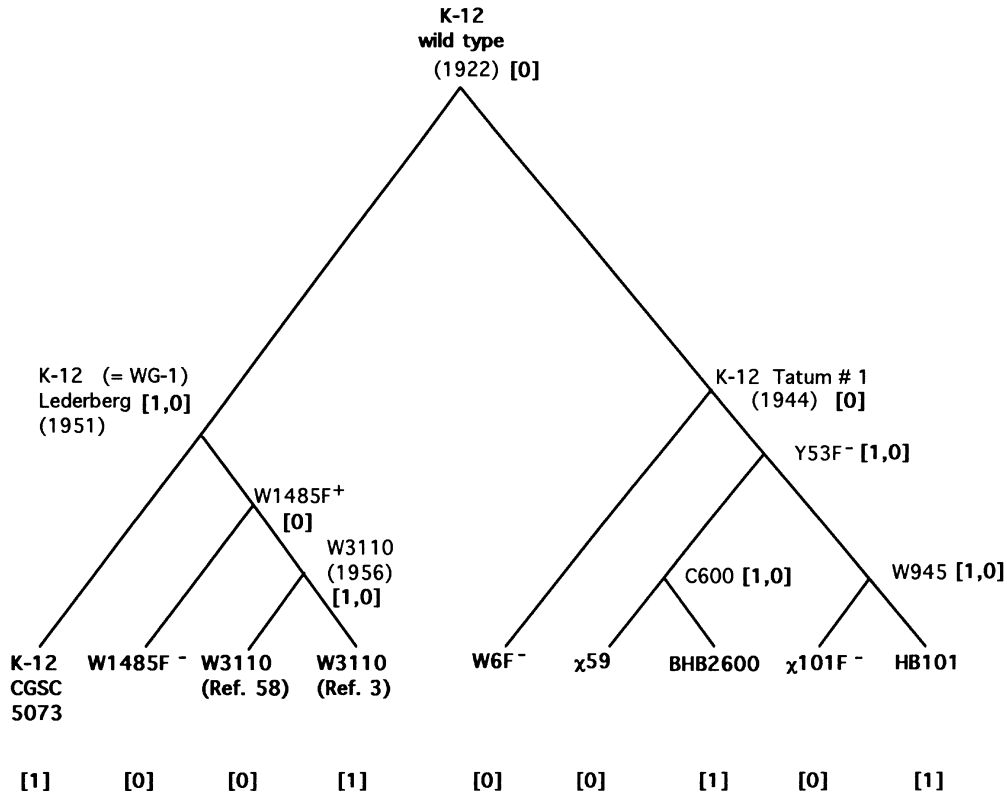


FIGURE 2 Abbreviated phylogenetic tree of some laboratory strains of *E. coli* K-12. Many intermediate steps have been omitted for clarity. Some steps (e.g., K-12 wild type [1922] → WG-1[1951]) involved no intentional mutagenic steps: the strains either were propagated in the laboratory. Some nodes on this tree (e.g for W1485) correspond to times when samples of same strain were taken from storage for further study or dissemination. The complement of IS elements in W3110 was independently studied by two groups (3, 58). The separate W3110 isolates have diverged in their IS complements since W3110 was first isolated in 1956. Strain χ 59 is the laboratory stock of C600 from the laboratory of Roy Curtiss III. HB101 has undergone a complicated construction history (b. Bachmann, personal communication). Its placement on this tree is conjectural, based on recipient markers retained after several conjugations steps. The maximum parsimony argument is illustrated for an IS2 found on a 5.5-kb *EcoRI* fragment in four of the strains by scoring each strain at the bottom (0 if the element is absent and 1 if the element is present). Moving up to higher and higher nodes, scores are 0 if the element is absent in paired branches, 1 if the element is present in paired branches, and 1,0 if its presence is neither supported nor rejected.

Figure 3 shows the map of the elements inferred to have been present in the original *E. coli* K-12; these elements are listed by name and map location in Table 3. In most cases, map locations were taken from EcoMap6 (K. E. Rudd, personal communication). The possible significance of the apparently

uneven distribution is discussed in a later section. Some IS elements mapped in previous studies (3, 58, 59) were not placed on this map because they were judged not to be ancestral. Designation of the basic set for IS5 was the most problematic because data for fewer strains were directly accessible. In addition, the apparent number of IS5 transposition events in three of the four strains analyzed was larger than for any of the other elements. Also, W3110 had been the victim of amplification probably caused by IS5 × IS5 reciprocal recombination (59). Because the starting point was a “restored” version of the W3110 chromosome implicit in EcoMap6, some IS5 elements mapped in W3110 are excluded. Most of the IS5 elements represented as indeterminate in Table 3 are unlikely to be members of the standard set, because other analyses using strains from different branches of the tree are more consistent with 10 to 11 copies per chromosome (50, 56). This number is consistent with data for strains BHB2600 (3), JE5519 (59), and JE5527 (40).

Additional IS Elements in the *E. coli* K-12 Chromosome

E. coli K-12 carries at least four other types of IS elements, some of whose locations have not yet been determined: IS421 (47), IS600 (38), IS629 (36), and IS911 (44). The latter three elements were first isolated from *Shigella* species. Note that by electrophoretic enzyme analysis, DNA sequence similarity, and other criteria, *Shigella* species are closely related to *E. coli* (see reference 52 and chapter 148).

IS421 is a member of the IS4 family (see below), and four copies are present on chromosomes of some *E. coli* K-12 strains. IS600, IS629, and IS911 are all members of the IS3 family. IS600 and IS629 were first detected in *Shigella sonnei*, and they each hybridize weakly to one chromosomal fragment from *E. coli* K-12 strain JM109, suggesting the presence of fragments or isoforms of these elements in *E. coli* K-12 (37). The sequence of IS629 is closely related to that of IS3411 (25), whose presence was not detected in another *E. coli* K-12 strain. This apparent disagreement may reflect different hybridization stringencies.

IS911 was isolated and identified after it had transposed into bacteriophage λ which was integrated in the *Shigella dysenteriae* chromosome (44). Prère et al. found in chromosomal DNA from an *E. coli* K-12 strain C600 derivative four *Bgl*III-*Pst*I restriction fragments that hybridized to an IS911 probe (44). Subsequent mapping has shown that these fragments are derived from two deleted IS911 elements. One contains IS30A inserted at IS911 bp 334 (with IS911 sequences beyond bp 762 deleted); the other contains IS30D and an adjacent 327 bp of IS600 substituting for IS911 sequences between bp 335 and 454 (M. F. Prère, M. Chandler, and O. Fayet, personal communication). In other words, the IS911 sequences in the *E. coli* K-12 chromosome map coincidentally with IS30A and IS30D (the four hybridizing chromosomal fragments arose because the IS30 elements contain a *Bgl*III site). The fragmentary nature of the IS911 sequences explains why no IS911 insertions were detected during the initial use of λ as an IS trap in *E. coli* K-12 (32). The transposition properties of IS911 are discussed in chapter 124.

IS200 in the Chromosome of *S. typhimurium* LT2

The IS map of *S. typhimurium* LT2 is simpler because this bacterium appears to contain only IS200 from among the elements described in this chapter, and because IS200 does not transpose as frequently as the *E. coli* K-12 elements (7). The six IS200 elements in the *S. typhimurium* LT2 chromosome were initially mapped by identifying chromosomal Tn10 insertions that affected mobilities of restriction fragments that contain them (29). Subsequent refinement of this mapping has been achieved by using hybridization of IS200 to chromosomal *Bln*I and *Xba*I fragments separated by pulsed-field gel electrophoresis and P22-mediated cotransduction frequencies (46).

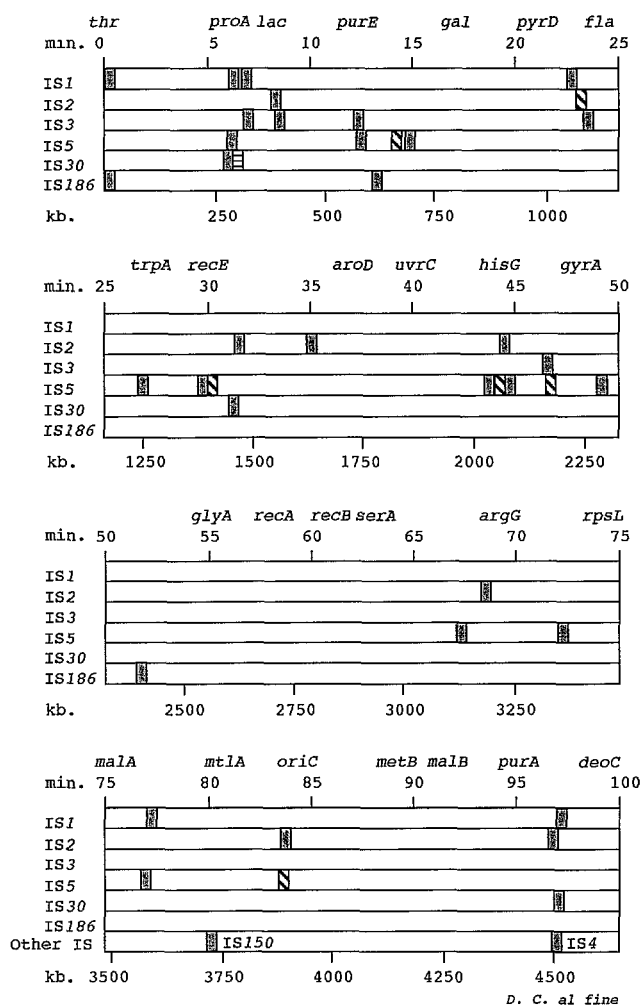


FIGURE 3 Locations of standard IS elements in the chromosome of *E. coli* K-12. Coordinates are physical map minutes (above the position fields) or kilobases (below the position fields). Precise designations, locations, and orientations are presented in Table 3. Shaded boxes represent the standard sets, diagonally hatched boxes represent elements whose status was not resolved by parsimony arguments, and the horizontally hatched box represent a fragment of IS30, IS elements appearing as single copies appeared only in the last quadrant of the map (Other IS).

The locations of the IS200 elements in *S. typhimurium* LT2 are as follows: IS200(I), 65.2 min, clockwise of *cysC*; IS200(II), 74.5 min, clockwise of *crp*; IS200(III), 94.4 min, counterclockwise of *purA*; IS200(IV), 42.5 min, clockwise of *cysB*; IS200(V), 53.6 min, clockwise of *ptsG*; and IS200(VI), 22.2 min, clockwise of *galT*. IS200 elements obviously are relatively widely spaced around the *S. typhimurium* LT2 chromosome.

Variations in IS Copy Numbers for *E. coli* K-12 Strains

Patterns of restriction fragments containing IS elements can be affected by genome rearrangements, by mutations that remove or create restriction sites, and by transposition. *E. coli* K-12 laboratory strains exhibit examples of all of these types of alterations.

Perkins et al. (43) have documented and correlated changes in the *Xba*I, *Bln*I, *Not*I, and *Sfi*I restriction maps of *E. coli* K-12 strains AB1157, EMG2 (one representative of supposedly wild-type *E. coli* K-12), MG1655, W1485, and W3110. Even though MG1655 was derived directly from W1485, the MG1655 chromosome contained deletions of 8, 7, 6, 1, and 1 kb and insertions of 12 and 7 kb compared with W1485. One of the deletions removed IS5G, which was not included as a standard element in the present compilation (Table 3). Exact correlations of differences in restriction maps of MG1655 and W1485 with IS elements would require mapping with additional enzymes and Southern blotting. It can be seen, however, that the insertion/deletion regions would potentially affect only the following IS elements (standard or nonstandard as defined above, locations given in kilobase coordinates and minutes): IS2C (not standard), 1,305 kb, 27.4 min; IS5K (standard), 2,299 kb, 46.6 min; IS2G (not standard), 2,331 kb, 47.5 min; IS2H (not standard), 3,011 kb, 61.5 min; IS5LO (standard), 3,147 kb, 64.4 min; and IS2I (standard), 3,203 kb, 66.5 min.

TABLE 3 Inferred standard sets of selected IS elements^a in the *E. coli* K-12 chromosome

Element	Standard element ^b									
IS1	A	B	C	D	E	F				
	0.4	6.2	6.4	2.6	77.	97.3				
IS2	←	←	←	→	→	←				
	A	D	E	F	I	J	K	(B)		
	8.3	31.	35.	44.	68.	83.8	96.9	23.8		
		6	5	5	6					
IS3	←	→	?	→	←	←	→	→		
	A	B	C	D	E					
	6.9	8.5	12.	23.	46.					
IS5 ^c	→	←	→	←	→					
	A	B	D	E	Y	X	H	K	LO	
	6.1	12.	14.	26.	30.	43.8	44.5	49.2	67.4	
		4	9	8	0					
	←	←	←	←	←	←	←	←	→	
	R	T	(C)	(F)	(I)	(J)	(U)			
72.4	78.	14.	30.	45.2	46.7	83.7				
IS30	←	←	→	→	←	→	←			
	A	B ^d	C	D						
	6.0	6.2	31.	97.						
IS186	→	→	→	←						
	A	B	C							
	0.3	13.	51.							
	2	8								
	→	→	→							

^aIS4 (one copy, 97.0, →) and IS150 (one copy, 80.3, → [54]) are also present. Unmapped elements are IS911, IS600, and IS629 elements or isoforms.

^bMap positions are in physical map minutes and were taken from EcoMap6. Symbols → and ← indicate clockwise and counterclockwise orientations of the elements relative to the genetic map, respectively. The direction of the arrow on each element corresponds to the transcriptional direction of ORFs for the known or putative transposases. Elements listed in boldface were mapped prior to 1987 and have been confirmed by subsequent studies; see reference 13 for primary references, procedures, and rationales. Data for IS5T are taken from reference 54. Parentheses indicate that the status of the element after parsimony analysis is indeterminate.

^cCopy number in C600 was 10 to 11 (50), copy number in JE5527 was approximately 10 (40), and copy number in W1485 was estimated at 11 (56). These results suggest that the upper limit for the number of ancestral (standard) IS5 elements is 11.

^dOnly 181 bp of IS30 are present at this site (60).

Most of the variation in IS element distribution appears to arise from transposition, as witnessed by increases in IS copy numbers. One hundred eighteen clones isolated from W3110 after storage for 30 years in a nutrient agar stab revealed extensive polymorphism in sizes and numbers of restriction fragments containing IS2, IS3, IS5, and IS30 but little variation in fragments containing IS1, IS4, IS150,

and IS186 (39). A phylogenetic tree for the 118 clones was constructed, and it was estimated that at least 174 IS-associated mutations had occurred during storage. The restriction pattern for IS1 was strictly conserved in all strains, but 31 new fragments containing IS5 were detected, indicating that IS5 was more transpositionally active during storage than were other elements.

The results for the laboratory strains described in Table 2 are consistent with these observations. The numbers of ancestral IS copies estimated by Naas et al. (39) are in reasonable agreement with the numbers of standard elements reported in Table 3, differing by at most one for any element type. The total changes in copy numbers for IS1, IS2, IS3, and IS5 from the presumed original copy numbers in *E. coli* K-12 are presented in Table 4 for strains at the bottom of the pedigree chart in Fig. 2. Although the number of generations along each branch differs, the number of events for each element type can be compared along any chosen pathway. Note that the number of generations of active growth and the length of time in storage will differ for each branch of the tree.

Along the pathway from C600 to BHB2600, there were five IS1 transpositions and one IS2 transposition, corresponding to 0.8 IS1 transposition per ancestral IS1 and to 0.14 IS2 transposition per ancestral IS2. In contrast, the lineage from the early W1485 isolate to W3110 (3) experienced one IS1 transposition (0.17 transposition per ancestral IS1) and seven IS2 transpositions (1.0 transposition per ancestral IS2). In each lineage, the number of generations during which IS1 and IS2 could have transposed was the same, yet in the C600 → BHB2600 lineage, IS1 appeared to transpose five times more frequently than IS2, while in the early W1485 → W3110 (3) lineage (again having the same number of generations allowed for each element), IS2 appeared to transpose six times more frequently than IS1.

TABLE 4 Inferred accumulated transposition events for different IS elements in various *E. coli* K-12 sublines

Element	No. of transposition events in strain:							
	W1485F ⁻	W3110	W3110	W6F ⁻	χ59	BHB2600	χ101	HB101
		Reference 58	Reference 3					
IS1	1	1	1	0	3	7	1	2
IS2	1	5	7	5	0	2	2	2
IS3	0	1	1	1	1	3	0	2
IS5	— ^a	9	10	—	—	1	—	8
IS30	0	1 or 2 ^b	2 + 1 ^c	—	—	1 ^c	—	1 ^c

^a —, not tested.

^b Two elements identified but not mapped (60).

^c Hybridization intensity lower than that expected for a single element.

The difference in apparent IS activity may partly be attributed to periods of active growth for these laboratory strains, during which IS1 can contribute significantly to the IS mutational spectrum (32). This is in sharp contrast to the behavior of IS1 during long-term storage of W3110 in stabs (39). Interestingly, the W3110 lineages in Table 4 show IS2 (5 and 7 transposition events) and IS5 (9 and 10 transposition events) to have been among the most transpositionally active IS elements, as was observed by Naas et al. (39). Moreover, IS30 did not appear to be transpositionally active in these strains, as was the case for a large subset of the W3110 clones isolated after long-term storage. Strains BHB2600 and HB101 (Table 5) do not display the same spectrum of IS transpositions as W3110, perhaps reflecting their different histories.

TABLE 5 Copy numbers of selected insertion sequences among *E. coli* and *S. typhimurium* isolates

Element	Standard copy no. ^a		ECOR set			SARA set, fraction of strains positive
	<i>E. coli</i> K-12	<i>S. typhimurium</i> LT2	Fraction of strains positive	Copy no. range	Avg copy no. (of positives) ^b	
IS1	6	0	60/71	0–27	7.5	9/66 ^c
IS2	7	0	43/71	0–17	4.5	0/10 ^d
IS3	5	0 ^d	48/71	0–6	2.5	31/40
IS4	1	0	28/71	0–14	5.1	0 ^d
IS5	11	0 ^d	25/71	0–21	3.4	~18/40 ^d
IS30	3	0 ^d	35/71	0–5	1.9	0/40 ^d
IS150	1	?	30/71	0–6	1.8	?
IS186	3	?	?	?	?	?
IS200	0	6	14/72	0–11	3.3	47/66 ^c

^a*E. coli* K-12 copy numbers for IS1, IS2, IS3, and IS5 determined by parsimony; *S. typhimurium* LT2 numbers from reference 28 except as otherwise noted.

^bExcluding strains with copy number of 0.

^cFor IS1, the number range is 0 to 5, with an average number of 3.1 among those strains having IS1. For IS200, the number range is 0 to 19, with an average of 6.5 for those containing IS200.

^dM. Biserić and H. Ochman, personal communication.

DISTRIBUTIONS OF IS ELEMENTS IN *E. COLI* AND *SALMONELLA* STRAINS

Earlier studies had focused primarily on IS elements in *E. coli* K-12 and derivatives or in *S. typhimurium* LT2 (see reference 13 for review). There were some published reports that had included a broader range of bacteria (41). A more comprehensive view has emerged with studies of the *E. coli* ECOR and the *Salmonella* SARA collections. Chosen to be more representative of the natural diversity of *E. coli*, members of the ECOR set have been characterized by electrophoretic typing at a number of chromosomal loci, and this allowed construction of phylogenies for members of the set (52). The SARA strains (1) are more closely related to each other than are members of the ECOR set and do not represent the diversity of *Salmonella* spp. in the wild.

IS Elements in the ECOR Strains

Among the *E. coli* ECOR strains, presence or absence of IS1 to IS5, IS30, and IS150 has been determined (19, 21, 49). Results are summarized in Table 3. IS2 to IS5 or IS30 is absent in 40 to 60% of the strains. In contrast, nearly 90% of the strains contain IS1. Three of the ECOR strains (strains 47, 48, and 68) contained no chromosomal copies of any of these elements. Fourteen of the ECOR strains contained IS200, originally identified in *S. typhimurium* (6, 28).

By using the observed distributions of copy numbers among the ECOR strains, Sawyer et al. (49) tested various mathematical models which differed in the functional dependence of transposition and cell death on IS element copy number n . More than one model could account for the observed distributions for IS2, IS4, IS5, and IS30. In contrast, IS3 and IS150 distributions were most consistent with models implying strong regulation of transposition (e.g., transposition rates proportional to $1/n$). IS1 distributions could be described by a model in which IS1 transposition was unregulated. The results for the laboratory strains presented in Table 2, though representing a much smaller sample, are consistent with this result. Over that portion of the tree not containing *E. coli* K-12 (CGSC 5073), IS3 has transposed 6 times (1.1 per element), whereas IS1, for example, has transposed 13 times (2.1 per element).

The IS distributions also provide information on how IS elements are acquired. The appearance of unrelated IS elements together on chromosomes more frequently than would have been expected for random distributions suggested that IS elements are transmitted horizontally by bacteriophages or plasmids (22). This is also suggested by the distribution of individual elements. For example, IS200 is distributed sporadically among many branches of the ECOR phylogenetic tree (6), as would be expected

if it were disseminated horizontally. In contrast to IS200, most (21 of 25) of one major phylogenetic division (the A group) contain chromosomal IS150 elements, suggesting that most of the members of this group inherited IS150 vertically from a common ancestor. However, sparse and sporadic appearances of IS150 in sister groups to group A suggest that IS150 was not present at the root of the *E. coli* phylogenetic tree.

IS elements are found on plasmids as well as chromosomes in the ECOR collection. Approximately half of the 72 ECOR strains possessed one or more IS1 to IS5 or IS30 elements on plasmids, whereas over 90% carried one or more of these elements on their chromosomes. IS150 distributions differed, showing more frequent association with plasmids (21). Clearly, the presence of these elements on plasmids, which may be conjugative or mobilizable, is consistent with horizontal transmission of IS elements to chromosomes.

IS Elements in the SARA Strains

Because of the nature of the SARA collection (see above), the proportions of strains containing various elements are unlikely to apply to other populations. Nevertheless, they give examples of the magnitudes of proportions that are likely to be encountered.

IS200, which was first identified in *S. typhimurium* LT2 (28), appears in the majority of the SARA strains (47 of 66 tested strains [5]). IS1 and IS3 were also detected in a 40-member subset of the SARA collection (6), contrary to previous observations for a more restricted collection of strains (28). IS3 was found in 31 of 40 strains, and in 4 of 40 strains, it appeared together with both IS200 and IS1. IS200 was more frequently encountered (24 of 40 strains) than was IS1 (13 of 40 strains); moreover, the IS200 average copy number among strains possessing the element (6.5) was larger than that for IS1 (3.1). Given the infrequent transposition of IS200, this finding suggests a longer, more stable association for IS200 with *Salmonella* chromosomes than for IS1. This point has been addressed more rigorously from studies of the sequences of these elements, as will be described below. The sporadic distribution of IS1 among branches of the SARA phylogenetic tree suggests that these elements have been transmitted horizontally, an observation consistent with the relatively frequent association of IS1 with plasmids. In contrast, IS200 is rarely found on plasmids.

In addition to IS1, IS3, and IS200, some members of the SARA collection also carry IS5 and IS30 (H. Ochman, personal communication). However, IS2 and IS4 have not been found in this group of bacteria.

IS ELEMENT SEQUENCE DIVERGENCE

The various IS elements of *E. coli* and other bacteria show sequence similarities indicating that different families of IS elements exist. For example, the IS3 family includes the *E. coli* elements IS3, IS2, IS150, and IS3411, as well as *Shigella* elements IS600, IS629, and IS911 (8, 44, 51). The latter three elements (or fragments or isoforms of them) are present in the *E. coli* K-12 chromosome (see above). The IS3 family, which includes elements from organisms not included among the *Enterobacteriaceae*, is characterized by sequence similarities among the putative transposases and by the presence of two tandem open reading frames (ORF A and ORF B), with the downstream ORF B in the -1 frame relative to ORF A. Translational frameshifting has been demonstrated for some of these elements, and other members of this family contain sequences capable of promoting frameshifting (see chapter 124 for more details). A phylogenetic tree relating to some of these elements is presented in Fig. 4.

Another grouping of IS elements is the IS4 family, which includes *E. coli* elements IS4, IS5, IS186, and IS421, in addition to members from other genera (45). The ORF B regions of transposases from both the IS3 and IS4 families show sequence similarities to a portion of the integrase genes of retroviruses (17). The significance of this observation for possible origins of prokaryotic IS elements has not yet been fully explored.

Sequence variations among individual members of particular IS types also provide clues about the time and mode of their acquisition by bacteria. For example, predicted protein sequences from IS3 elements of *E. coli*, *Shigella dysenteriae*, *Escherichia fergusonii*, and *Shigella odorifera* differed from each other to approximately the same extent that chromosomal genes from these bacteria differed, and these differences correlated with the phylogenies inferred from chromosomal gene sequences (30). This indicates that IS3 may have been associated with the common ancestor of these bacteria. Since only 23 members of the 72-member ECOR collection do not carry IS3, and since these strains are distributed among separated branches of the phylogenetic tree, mechanisms for removal or “curing” of IS3 appear to have operated in some of the strains since they diverged from the common ancestor.

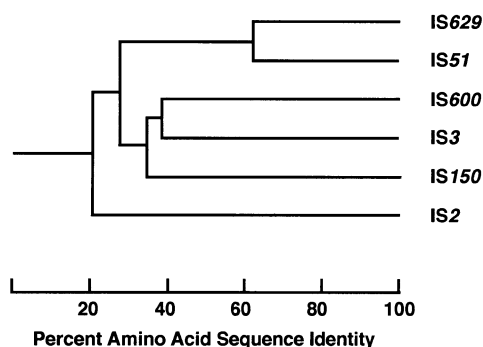


FIGURE 4 Phylogenetic relationships of members of the IS3 family. The organization of principal genetic features of these element (e.g., IS150) is illustrated in Fig. 1. The relationships were inferred from predicted amino acid sequences of the second (ORF B) reading frames (redrawn from reference 51). IS911 is also a member of this group, and its ORF B shows 34% amino acid identity to IS600 ORF B and 28% identity to IS150 ORF B (44).

In contrast to IS3, divergence of IS1 sequences does not parallel the divergence of the chromosomal genes in organisms from which they are isolated (30). This indicates that some of the IS1 elements were evolving independently of the chromosomes in which they are found today, which implies that they have been acquired by horizontal transfer. In the ECOR collection, bacteria have acquired either IS1F or IS1R versions of IS1, but not both, again emphasizing the relative recency and independence of IS1 (30). *E. coli* K-12 is an exception: it contains both IS1F and IS1R versions (61). DNA sequences of IS1 and IS3 elements from various isolates of the ECOR collection usually show little sequence variation, ranging from identity to the *E. coli* K-12 prototype sequences up to one to two substitutions in most cases.

IS200, like IS3, appears to have been present in *S. typhimurium* and *E. coli* since before these species diverged (5). *E. coli* IS200 copies differ from those of *S. typhimurium* by approximately 7%, which is similar to the differences between chromosomal genes from these two organisms. Among ECOR strains tested, sequence divergence among IS200 examples is approximately 3% (5), which is similar to the extent of divergence of chromosomal genes within the ECOR collection.

General recombination also appears to play a role in the evolution of IS elements. A particular IS3 from ECOR strain 63 contained a 107-bp sequence block in ORF A that was only 63% identical to the corresponding IS3 region from *E. coli* K-12, and comparisons of IS3 sequences from *E. coli* K-12, ECOR strain 63, and *Shigella dysenteriae* suggested that the IS3 from ECOR strain 63 was a recombinational composite (30). The same study showed that the middle one-third of IS1 elements isolated from *E. coli*, *E. fergusonii*, *E. hermannii*, and *E. vulneris* shows a lesser percentage of nucleotide sequence identity than do the flanking regions, again indicating that different portions of IS elements may have experienced different evolutionary histories.

INFLUENCES OF IS ELEMENTS

Map Locations

Do the locations and distributions of IS elements on the chromosomes reflect intrinsic properties of various chromosomal regions? Are the apparent clusters and gaps historical accidents, as might occur if an initially random IS insertion were to make more likely interactions of other IS elements with the same regions (e.g., by cointegration with plasmids containing a variety of IS elements)? Might a tendency toward *cis*-acting transposition lead to accumulation within topologically restricted chromosomal domains? A preliminary approach to these issues is to ask whether the relative abundance or paucity of IS elements in particular chromosomal regions in *E. coli* K-12 might have occurred by chance.

The distribution of IS elements (Fig. 2) shows regions with relatively high densities of elements (the interval from 5 to 15 min) and regions devoid of IS elements (the interval from 53 to 67 min). For a random distribution, the expected pattern of elements depends upon the distribution function relating frequency of occurrence to the element density (number of elements per defined interval). Jurka and Savageau (26) showed that regions of high gene density in the *E. coli* K-12 map were a natural consequence of the log normal gene density distribution function. (The normal, or gaussian, distribution represents frequency by using the gene density as the independent variable; however, with the log normal distribution, the frequency is a function of the logarithm of the gene density.) If IS elements were to obey a similar log normal distribution function, then clustering would be expected to occur for them as well.

One approach is to make the simplifying assumption that IS elements are distributed in a Poisson manner and to estimate the expected maximal and minimal spacings between elements. If the expected value for the smallest spacing, $X_{(1)}$, is defined as $E(X_{(1)}) = \mu/n$ (equation 6 in reference 9), where μ is the average spacing (2.86 min) and n is the number of elements (35 elements), then for the present case, $E(X_{(1)}) = 0.08 \text{ min} = 3.8 \text{ kb}$. The expected value for the largest interval, $X_{(n)}$, is $E(X_{(n)}) = \mu[(0.5772 + \lambda v(n))] = 11.8 \text{ min} = 552 \text{ kb}$ (equation 8 in reference 9). The predicted maximum and minimum spacings are not very different from the observed values (0.03 min minimum interval, 14 min maximum interval). Equation 9 from Churchill et al. (9) indicates that the probability of observing by chance a maximum spacing greater than 14 min would be 0.23. Thus, the observed maximum spacing is reasonably probable for a random distribution obtained by using Poisson statistics.

An alternative approach to analyzing the distribution of IS elements is to ask whether the observed distribution differs significantly from a uniform distribution, in which the number of IS elements encountered during progression along the genome increases linearly with distance. The Kolmogorov-Smirnov test associates the maximum deviation of a computed test statistic (equation 2 in reference 9) with the probability that the observed data follow the hypothetical (in this case, uniform) distribution. By using the locations of the 35 unambiguously identified elements in Table 3, one can calculate the probability P that the observed distribution is identical to a uniform distribution. This calculated P value falls between 0.02 and 0.05, indicating that the nonuniformity of the IS distribution along the chromosome is significant but not highly significant.

Clustering of IS elements might result from characteristics of a particular chromosomal region (such as transcriptional activity, superhelix density, or DNA sequence composition), or it might merely reflect historical contingency. For example, if an IS3 were to transpose by chance into a region of the chromosome lacking IS elements, then insertion at that locus of F or other plasmids containing IS3 would be facilitated. These plasmids might carry other IS elements, which then might preferentially transpose into the immediate neighborhood. Alternatively, IS locations might be correlated as a result of acquisition of composite transposons like the one proposed to be responsible for the duplicate gene *argF* in *E. coli* K-12 (24). A third possibility is that IS elements are preferred targets for other types of IS elements (e.g., possible transposition of IS30 into IS911 [Prère et al., personal communication]).

Genome Rearrangements

The homology associated with multiple IS copies provides a pathway for several (presumably *recA*-mediated) processes that lead to chromosome reorganization. IS3A can recombine with IS3B to invert the *lac* region of *E. coli* K-12 (48). Inversions of this type had been noted in Hfr strains (2), and they can be explained by the same type of recombination event. IS5 elements also can cause extensive chromosomal inversions. The *oxa1* mutations (12) and other, similar *oxa* mutations (33) are a result of recombination between IS5Y and a presumably inverted IS5 between *argG* and *xylA*. The correspondence between inversion endpoints with the locations of IS elements in *E. coli* K-12 indicates that recombination between inverted IS elements of the same type is a major inversion pathway in this strain.

The integration of the F plasmid at chromosomal IS2, IS3, or $\gamma\delta$ (*Tn1000*) elements to form Hfr strains (11) has been experimentally important in the development of *E. coli* K-12 genetics (34). At chromosomal sites that have been examined in Hfr strains, there is excellent correlation between Hfr points of origin and chromosomal IS2A and IS3A to C elements (14). Although F can presumably integrate by transpositional cointegration, the *recA* dependence of F integration (10) suggests that most F integration occurs by reciprocal homologous exchange. Similarly, the directly repeated chromosomal IS5A and IS5B elements can function in the excision of F' plasmids from the bacterial chromosome (56). The roles of these and other IS elements in F integration and F' excision have been summarized by Umeda and Ohtsubo (58), and the effects of IS elements and other repeated sequences on genome rearrangements are discussed in chapter 112.

Coordinated Transposition

Transposition of IS elements is particularly evident for *E. coli* strains that have been stored in stab cultures (19, 39). This phenomenon also occurs for *S. typhimurium*, in which copy numbers of IS200 may increase from 6 to 11 after serial propagation in stabs (C. R. Beuzón and J. Casadesús, personal communication). For IS200, the increase appears to occur abruptly. This result differs from observations with IS elements in *E. coli* K-12, which reportedly show linear increases in transposition events with time (32).

In the case of IS30, bursts of transposition may arise after rearrangements that form the tandem structure (IS30)₂ (42). It would be interesting to know how these rearrangements contribute to accumulation of IS-associated events in stab cultures and whether they appear gradually over time or primarily after a threshold period of storage.

Evolutionary Questions

Because IS elements can promote genome rearrangements and create mutations, they potentially could contribute to the evolution of bacterial chromosomes. These effects will be more readily documented as DNA sequences of complete bacterial chromosomes become available. It already appears that *E. coli* K-12 is missing DNA sequences that are presumably present in other *E. coli* strains, since IS1A, IS1B, IS1C, and IS1F are not flanked by direct repetitions (61). The loss of direct repetitions might indicate that these elements promoted adjacent deletions of chromosomal DNA. It would be interesting to check members of the ECOR collection for their chromosomal DNA sequences adjacent to the locations of the IS elements in *E. coli* K-12 to determine what may have been lost from *E. coli* K-12.

The horizontal dissemination of IS elements makes it likely that sequence variants of other classes of IS elements will be found in the *E. coli* K-12 and *S. typhimurium* chromosomes. It may prove fruitful to search the *E. coli* K-12 chromosomal DNA sequence (soon to be available) for limited similarities to sequences of other IS elements and for imperfect inverted repetitions of appropriate-size flanking 700- to 2,000-bp regions that just fail to contain appropriate-size ORFs or that include appropriate-size regions containing only a few stop codons. This approach might reveal derelict IS elements that are no longer functional in *E. coli* K-12. The H-rpt sequence found in some Rhs elements (62) might be an example of an IS that is rarely or not at all functional in *E. coli* K-12 (see chapter 112).

The dynamics of inheritance of IS elements over long periods of time is another issue to be addressed. Does the persistence in the population of elements that transpose by “cut and paste” mechanisms differ from that for elements that transpose replicatively? Why are some elements found preferentially on chromosomes, whereas others are more frequently found on plasmids? Is this bias related to the transposition mechanisms?

Similarly, it will be interesting to know what determines the spectrum of IS elements that transpose into particular chromosomal genes. In some cases, this will be attributable to biochemical idiosyncrasies of different elements and the particular target sequences. For example, over half of the IS2 insertions in bacteriophage P1 fall in a fragment representing <2% of the P1 genome but containing sequence similarities to IS2 (53). A recent study showed that all IS insertions in the *hemB* gene of *E. coli* K-12 were IS2 (31). Is this only a reflection of the sequence properties of *hemB*, or is it partly a consequence of the close proximity (8-kb separation along the contour length) of IS2 as the nearest neighboring IS element?

Some elements like IS3 appear to have been present in *E. coli* for a very long time, yet approximately one-third of the ECOR strains lack IS3 on their chromosomes. Do these elements shuttle back and forth between chromosomes on the one hand and plasmids and phages on the other? How are IS elements removed from chromosomes? Is precise excision sufficient, or are there other mechanisms for rectification of chromosomal DNA (e.g., replacement of an IS-containing segment by the corresponding uninterrupted DNA introduced during conjugation with an Hfr)? Precise excision frequencies range from 10^{-7} to 10^{-9} , depending on sequence context (31, 55). Excision frequencies also are influenced by external factors, such as the *ref* gene of bacteriophage P1. When *ref* is derepressed, it stimulates the efficiency of excision of IS1 from *galT* by a factor of 10^5 (35). Thus, rectification of IS mutations in the bacterial chromosome may depend on particular historical contingencies (e.g., whether they were infected by a particular type of phage or plasmid).

IS Element Chimeras

Some IS elements may be recombinationally composite (30), presumably as a result of general recombination. Other composites formed from pieces of different IS elements are known. The 181-bp IS30 fragment IS30B abuts an end of IS1B, suggesting that it may be a remnant of an adjacent deletion event (60). The origin of the IS911 deletant containing IS30D and an IS600 fragment (see above) is not so readily explained. Are IS elements themselves particularly susceptible to attack by other elements, or are IS elements accumulated on plasmids and phages and then, after transpositional deletion and other events, mobilized to the chromosome?

ACKNOWLEDGMENTS

I thank Barbara J. Bachmann of the *E. coli* Genetic Stock Center for assembling and transmitting additional information on *E. coli* pedigrees, and I thank Kenneth E. Rudd for unpublished mapping information and extensive documentation and discussion relating to EcoMap6. Michael Waterman contributed most helpful information on statistical analysis. Howard Ochman provided unpublished information on IS elements in the SARA collection and helpful insights. J. Casadesús, M.-F. Prère, M. Chandler, and O. Fayet generously provided unpublished data. I thank Elaine Freund for her thoughtful comments on the manuscript.

LITERATURE CITED

1. **Beltran, P., S. A. Plock, N. H. Smith, T. S. Whittam, D. C. Old, and R. H. Selander.** 1991. Reference collection of strains of the *Salmonella typhimurium* complex from natural populations. *J. Gen. Microbiol.* **137**:601–606.
2. **Berg, C. M., and R. C. Curtiss III.** 1967. Transposition derivative of an Hfr strain of *Escherichia coli* K-12. *Genetics* **56**:503–525.
3. **Birkenbihl, R. P., and W. Vielmetter.** 1989. Complete maps of IS1, IS2, IS3, IS4, IS5, IS30, and IS150 locations in *E. coli* K12. *Mol. Gen. Genet.* **220**:147–153.
4. **Birkenbihl, R. P., and W. Vielmetter.** 1991. Completion of the IS map in *E. coli*: IS186 positions on the *E. coli* K12 chromosome. *Mol. Gen. Genet.* **226**:318–320.
5. **Biserić, M., and H. Ochman.** 1993. The ancestry of insertion sequences common to *Escherichia coli* and *Salmonella typhimurium*. *J. Bacteriol.* **175**:7863–7868.
6. **Biserić, M., and H. Ochman.** 1993. Natural populations of *Escherichia coli* and *Salmonella typhimurium* harbor the same classes of insertion sequences. *Genetics* **133**:449–454.
7. **Casadesús, J., C. R. Beuzon, and I. Gilbert.** 1992. IS200, basic and applied. *Genetics (Life Sci. Adv.)* **11**:179–186.
8. **Chandler, M., and O. Fayet.** 1993. Translational frameshifting in the control of transposition in bacteria. *Mol. Microbiol.* **7**:497–503.
9. **Churchill, G. A., D. L. Daniels, and M. S. Waterman.** 1990. The distribution of restriction enzyme sites in *Escherichia coli*. *Nucleic Acids Res.* **18**:589–597.
10. **Cullum, J., and P. Broda.** 1979. Chromosome transfer and Hfr formation by F in *rec*⁺ and *recA* strains of *Escherichia coli* K-12. *Plasmid* **2**:358–365.
11. **Davidson, N., R. C. Deonier, S. Hu, and E. Ohtsubo.** 1975. Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*. X. Deoxyribonucleic acid sequence organization of F and of F-primes, and the sequences involved in Hfr formation, p. 56–65. In D. Schlessinger (ed.), *Microbiology—1974*. American Society for Microbiology, Washington, D.C.
12. **de Massey, B., J. Patte, J.-M. Louarn, and J.-P. Bouché.** 1984. *OriX*: a new origin of replication in *E. coli*. Correction. *Cell* **38**:333.
13. **Deonier, R. C.** 1987. Locations of native insertion sequence elements, p. 982–989. In F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington, D.C.
14. **Deonier, R. C., and R. G. Hadley.** 1980. IS2-IS2 and IS3-IS3 relative recombination frequencies in F integration. *Plasmid* **3**:48–64.
15. **Deonier, R. C., R. G. Hadley, and M. Hu.** 1979. Enumeration and identification of IS3 elements in *Escherichia coli* strains. *J. Bacteriol.* **137**:1421–1424.
16. **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.
17. **Fayet, O., P. Ramond, P. Polard, M.-F. Prère, and M. Chandler.** 1990. Functional similarities between retroviruses and the IS3 family of bacterial insertion sequences. *Mol. Microbiol.* **4**:1771–1777.
18. **Galas, D. J., and M. Chandler.** 1989. Bacterial insertion sequences, p. 109–162. In D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
19. **Green, L., R. D. Miller, D. E. Dykhuizen, and D. L. Hartl.** 1984. Distribution of DNA insertion element IS5 in natural isolates of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **81**:4500–4504.
20. **Hadley, R. G., M. Hu, M. Timmons, K. Yun, and R. C. Deonier.** 1983. A partial restriction map of the *proA-purE* region of the *E. coli* K-12 chromosome. *Gene* **22**:281–287.
21. **Hall, B. G., L. L. Parker, P. W. Betts, R. F. DuBose, S. A. Sawyer, and D. L. Hartl.** 1989. IS103, a new insertion element in *Escherichia coli*: characterization and distribution in natural

- populations. *Genetics* **121**:423–431.
22. **Hartl, D. L., and S. A. Sawyer.** 1988. Why do unrelated insertion sequences occur together in the genome of *Escherichia coli*? *Genetics* **118**:537–541.
 23. **Hu, M., and R. C. Deonier.** 1981. Comparison of IS1, IS2 and IS3 copy number in *Escherichia coli* strains K-12, B, and C. *Gene* **16**:161–170.
 24. **Hu, M., and R. C. Deonier.** 1981. Mapping of IS1 elements flanking the *argF* gene region on the *Escherichia coli* K-12 chromosome. *Mol. Gen. Genet.* **181**:222–229.
 25. **Ishiguro, N., and G. Sato.** 1988. Nucleotide sequence of insertion sequence IS3411, which flanks the citrate utilization determinant of transposon Tn3411. *J. Bacteriol.* **170**:1902–1906.
 26. **Jurka, J., and M. A. Savageau.** 1985. Gene density over the chromosome of *Escherichia coli*: frequency, distribution, spatial clustering, and symmetry. *J. Bacteriol.* **163**:806–811.
 27. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. *Cell* **50**:495–508.
 28. **Lam, S., and J. R. Roth.** 1983. IS200: a *Salmonella*-specific insertion sequence. *Cell* **34**:951–960.
 29. **Lam, S., and J. R. Roth.** 1983. Genetic mapping of IS200 copies in *Salmonella typhimurium* strain LT2. *Genetics* **105**:801–811.
 30. **Lawrence, J. G., H. Ochman, and D. L. Hartl.** 1992. The evolution of insertion sequences within enteric bacteria. *Genetics* **131**:9–20.
 31. **Lewis, L. A., D. Lewis, V. Persaud, S. Gopaul, and B. Turner.** 1994. Transposition of IS2 into the *hemB* gene of *Escherichia coli* K-12. *J. Bacteriol.* **176**:2114–2120.
 32. **Lieb, M.** 1981. A fine structure map of spontaneous and induced mutations in the lambda repressor gene, including insertions of IS elements. *Mol. Gen. Genet.* **184**:364–371.
 33. **Louarn, J. M., J.-P. Bouché, R. Legendre, J. Louarn, and J. Patte.** 1985. Characterization and properties of very large inversions of the *Escherichia coli* chromosome along the origin-to-terminus axis. *Mol. Gen. Genet.* **201**:467–476.
 34. **Low, K. B.** 1972. *Escherichia coli* K-12 F-prime factors, old and new. *Bacteriol. Rev.* **36**:587–607.
 35. **Lu, S. D., D. Lu, and M. Gottesman.** 1989. Stimulation of IS1 excision by bacteriophage P1 *ref* function. *J. Bacteriol.* **171**:3427–3432.
 36. **Matsutani, S., and E. Ohtsubo.** 1990. Complete sequence of IS629. *Nucleic Acids Res.* **18**:1899.
 37. **Matsutani, S., and E. Ohtsubo.** 1993. Distribution of the *Shigella sonnei* insertion elements in Enterobacteriaceae. *Gene* **127**:111–115.
 38. **Matsutani, S., H. Ohtsubo, Y. Maeda, and E. Ohtsubo.** 1987. Isolation and characterization of IS elements repeated in the bacterial chromosome. *J. Mol. Biol.* **196**:445–455.
 39. **Naas, T., M. Blot, W. M. Fitch, and W. Arber.** 1994. Insertion sequence-related genetic variation in resting *Escherichia coli* K-12. *Genetics* **136**:721–730.
 40. **Nakamura, K., and M. Inouye.** 1981. Inactivation of the *Serratia marcescens* gene for the lipoprotein in *Escherichia coli* by insertions sequences, IS1 and IS5; sequence analysis of junction points. *Mol. Gen. Genet.* **183**:107–114.
 41. **Nyman, K., K. Nakamura, H. Ohtsubo, and E. Ohtsubo.** 1981. Distribution of the insertion sequence IS1 in Gram-negative bacteria. *Nature (London)* **289**:609–612.
 42. **Olasz, F., R. Stalder, and W. Arber.** 1993. Formation of the tandem repeat (IS30)₂ and its role in IS30-mediated transpositional DNA rearrangements. *Mol. Gen. Genet.* **239**:177–187.
 43. **Perkins, J. D., J. D. Heath, B. R. Sharma, and G. M. Weinstock.** 1993. *Xba*I and *Bln*I genomic cleavage maps of *Escherichia coli* K-12 strain MG1655 and comparative analysis of other strains. *J. Mol. Biol.* **232**:419–445.
 44. **Prère, M.-F., M. Chandler, and O. Fayet.** 1990. Transposition in *Shigella dysenteriae*: isolation and analysis of IS911, a new member of the IS3 group of insertion sequences. *J. Bacteriol.* **172**:4090–4099.
 45. **Rezsöhazy, R., B. Hallet, J. Delcour, and J. Mahillon.** 1993. The IS4 family of insertion sequences: evidence for a conserved transposase motif. *Mol. Microbiol.* **9**:1283–1295.
 46. **Sanderson, K. E., P. Sciore, S.-L. Liu, and A. Hessel.** 1993. Location of IS200 on the genomic

- cleavage map of *Salmonella typhimurium* LT2. *J. Bacteriol.* **175**:7624–7628.
47. **Sato, S., Y. Nakada, and A. Shiratsuchi.** 1989. IS421, a new insertion sequence in *Escherichia coli*. *FEBS Lett.* **249**:21–26.
 48. **Savić, D. J., S. P. Romac, and S. D. Ehrlich.** 1983. Inversion in the lactose region of *Escherichia coli* K-12: inversion termini map within IS3 elements $\alpha_{3\beta 3}$ and $\beta_{5\alpha 5}$. *J. Bacteriol.* **155**:943–946.
 49. **Sawyer, S. A., D. E. Dykhuizen, R. F. DuBose, L. Green, T. Mutangadura-Mhlanga, D. F. Wolczyk, and D. L. Hartl.** 1987. Distribution and abundance of insertion sequences among natural isolates of *Escherichia coli*. *Genetics* **115**:51–63.
 50. **Schoner, B., and R. G. Schoner.** 1981. Distribution of IS5 in bacteria. *Gene* **16**:347–352.
 51. **Schwartz, E., M. Droger, and B. Rak.** 1988. IS150: distribution, nucleotide sequence, and phylogenetic relationships of a new *E. coli* insertion element. *Nucleic Acids Res.* **16**:6789–6802.
 52. **Selander, R. K., D. A. Caugent, 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*. American Society for Microbiology, Washington, D.C.
 53. **Sengstag, C., J. C. W. Shepherd, and W. Arber.** 1983. The sequence of the bacteriophage P1 genome region serving as a hot target for IS2 insertion. *EMBO J.* **2**:1777–1781.
 54. **Sofia, H. J., V. Burland, D. L. Daniels, G. Plunkett III, and F. R. Blattner.** 1994. Analysis of the *Escherichia coli* genome. V. DNA sequence of the region from 76.0 to 81.5 minutes. *Nucleic Acids Res.* **22**:2576–2586.
 55. **Starlinger, P., and H. Saedler.** 1976. IS-elements in microorganisms. *Curr. Top. Microbiol. Immunol.* **75**:111–152.
 56. **Timmons, M. S., A. M. Bogardus, and R. C. Deonier.** 1983. Mapping of chromosomal IS5 elements that mediate type II F-prime plasmid excision in *Escherichia coli* K-12. *J. Bacteriol.* **153**:395–407.
 57. **Timmons, M. S., K. Spear, and R. C. Deonier.** 1984. IS121 is near *proA* in the chromosomes of *Escherichia coli* K-12 strains. *J. Bacteriol.* **160**:1175–1117.
 58. **Umeda, M., and E. Ohtsubo.** 1989. Mapping of insertion elements IS1, IS2 and IS3 on the *Escherichia coli* K-12 chromosome. Role of the insertion elements in formation of Hfrs and F' factors and in rearrangement of bacterial chromosomes. *J. Mol. Biol.* **208**:601–604.
 59. **Umeda, M., and E. Ohtsubo.** 1990. Mapping of insertion element IS5 in the *Escherichia coli* K12 chromosome. Chromosomal rearrangements mediated by IS5. *J. Mol. Biol.* **213**:229–237.
 60. **Umeda, M., and E. Ohtsubo.** 1990. Mapping of insertion element IS30 in the *Escherichia coli* K12 chromosome. *Mol. Gen. Genet.* **222**:317–322.
 61. **Umeda, M., and E. Ohtsubo.** 1991. Four types of IS1 with differences in nucleotide sequence reside in the *Escherichia coli* K12 chromosome. *Gene* **98**:1–5.
 62. **Zhao, S., C. H. Sandt, G. Feulner, D. A. Vlazny, J. A. Gray, and C. W. Hill.** 1993. *Rhs* elements of *Escherichia coli* K-12: complex composites of shared and unique components that have different evolutionary histories. *J. Bacteriol.* **175**:2799–2808.