JOURNAL OF BACTERIOLOGY, Feb. 1997, p. 1105–1111 Vol. 179, No. 4
0021-9193/97/$04.00+0

Analysis of the Boundaries of *Salmonella* Pathogenicity Island 2 and the Corresponding Chromosomal Region of *Escherichia coli* K-12

MICHAEL HENSEL,[1] JACQUELINE E. SHEA,[2] ANDREAS J. BÄUMLER,[3] COLIN GLEESON,[2] FREDERICK BLATTNER,[4] AND DAVID W. HOLDEN[2]*

Department of Infectious Diseases and Bacteriology, Royal Postgraduate Medical School, Hammersmith Hospital, London W12 0NN, United Kingdom[2]; Lehrstuhl für Bakteriologie, Max von Pettenkofer-Institut für Hygiene und Medizinische Mikrobiologie, D-80336 Munich, Germany[1]; Department of Microbiology and Immunology, Oregon Health Science University, Portland, Oregon 97201-3098[3]; and Laboratory of Genetics, University of Wisconsin-Madison, Madison, Wisconsin 53706[4]

Received 1 October 1996/Accepted 21 November 1996

We recently identified a pathogenicity island (SPI2) located at 30.7 centisomes on the *Salmonella typhimurium* chromosome. SPI2 contains genes encoding a type III secretion system whose function is distinct from that of the type III secretion system encoded by a pathogenicity island (SPI1) at 63 centisomes which is involved in epithelial cell entry. An analysis of the boundaries of SPI2 and comparison with the corresponding region of the *Escherichia coli* chromosome revealed that SPI2 inserted adjacent to the $tRNA^{Val}$ gene. The *E. coli* chromosome contains 9 kb of DNA at the region corresponding to the SPI2 insertion point which appears to be absent in *S. typhimurium*. The distribution of SPI1 and SPI2 was examined in various *Salmonella* isolates. In contrast to type III secretion system genes of SPI1, those of SPI2 are not present in *Salmonella bongori*, which diverged at the first branch point in the *Salmonella* lineage. These and other data indicate that SPI2 was acquired by a *Salmonella* strain already harboring SPI1 by horizontal transfer from an unknown source.

Diseases caused by *Salmonella* spp. range from self-limiting gastroenteritis to typhoid fever, which can be fatal (28). A useful tool for investigation of the systemic form of salmonellosis is the murine model of typhoid-like illness caused by *Salmonella typhimurium* (full taxonomical name, *Salmonella enterica* subsp. *enterica* serotype Typhimurium). Many virulence factors required at different stages of *S. typhimurium* infection have been characterized at the molecular level. The genes encoding these virulence factors are distributed on the *Salmonella* chromosome and the 92-kb virulence plasmid (20).

An important stage in *S. typhimurium* pathogenesis is invasion of the gut epithelium. A large number of genes is required for epithelial cell invasion, and it has been shown that these encode the structural components of a type III (contact-dependent) secretion system, the secreted effector proteins, and associated regulatory proteins (17). These genes are clustered at 63 centisomes (cs) on the *S. typhimurium* chromosome, and recent analysis of this locus revealed that it constitutes a pathogenicity island (PAI) (29). PAIs comprise large, sometimes unstable chromosomal regions harboring clusters of virulence genes and are often either flanked by insertion sequence elements (16) or appear to have inserted in or adjacent to tRNA genes (6, 12). The locus for enterocyte effacement of enteropathogenic *Escherichia coli* is a PAI; like the invasion locus of *S. typhimurium*, it contains genes for a type III secretion system. Homologs of these genes are also found on the virulence plasmids of *Yersinia* spp. (for a review, see reference 13) and *Shigella* spp. (for a review, see reference 32). The DNA base composition of PAIs often differs from those of the bacterial chromosomes in which they are located, indicating that they have probably been acquired by horizontal gene transfer (for a review, see reference 11).

During a search for new virulence genes of *S. typhimurium* by signature-tagged mutagenesis (22), we discovered a second *Salmonella*-specific PAI of 40 kb located at 30.7 cs (40). We termed this PAI *Salmonella* pathogenicity island 2 (SPI2) to distinguish it from the PAI at 63 cs, which we suggested be termed *Salmonella* pathogenicity island 1 (SPI1). Nucleotide sequence analysis of regions of SPI2 revealed genes encoding a second type III secretion apparatus. Mutations in these genes result in a profound attenuation of virulence in mice following oral or intraperitoneal inoculation (40). This demonstrates that SPI2 has a crucial role at a stage(s) of pathogenesis subsequent to epithelial cell penetration and that SPI1 and SPI2 are functionally distinct (40). More recently, it was reported that SPI2 genes are required for survival in macrophages (30).

In this paper, we report the structure of the boundaries of SPI2, their relation to the corresponding region of the *E. coli* K-12 genome, and a detailed analysis of the phylogenetic distribution of SPI2 in the salmonellae.

MATERIALS AND METHODS

Bacterial strains. Throughout this paper, the common Latin binominal *S. typhimurium* is used for *S. enterica* subsp. *enterica* serotype Typhimurium. The *Salmonella* Reference Collection B, consisting of 72 isolates and representing 37 serotypes of *S. enterica* subspecies I, has been reported recently (8). The collection of *Salmonella* serotypes representing *Salmonella bongori* and six subspecies of *Salmonella enterica* was described by Reeves et al. (34). Further strains and plasmids used in the study are listed in Table 1.

DNA biochemistry. The selection of a set of clones covering SPI2 from a library of *S. typhimurium* LT2 in λ 1059 (26) was described previously (40). Fragments of phage DNA as well as of DNA prepared from mitomycin-induced lysates of *S. typhimurium* TT15244 (3) were used for subcloning of regions of SPI2 and its boundaries by standard methodology (37) as detailed in Table 1. For the analysis of the corresponding chromosomal region in *E. coli* K-12, DNA fragments of phage 319 of the Kohara collection (23) were subcloned as indicated in Table 1.

* Corresponding author. Phone: 44 0 181 383 3487. Fax: 44 0 181 383 3394. E-mail: dholden@rpms.ac.uk.

TABLE 1. Bacterial strains, plasmids, and phages

Strain, plasmid, or phage	Properties	Reference or source
Strain		
S. typhimurium LT2	Wild type	25; K. Sanderson
S. typhimurium TT15244	See reference	3; K. Sanderson
E. coli K-12	Wild type	Lab stock
E. coli TA One Shot	See reference	Invitrogen
E. coli DH5α	See reference	Gibco BRL
Plasmids		
p7-13	5.5-kb *Pst*I-*Kpn*I fragment of λ7 in pSK$^+$ (Stratagene)	This study
pTT-1	6.8-kb *Eco*RI-*Xba*I fragment of TT15244 lysate in pSK$^+$	This study
Phages		
λ1; λ7	Clones harboring SPI2 fragments selected from genomic library of *S. typhimurium* LT2	40
Kohara 319		23

Bacteria were cultured in Luria-Bertani broth (37) aerobically at 37°C overnight, and isolation of chromosomal DNA was performed as described previously (1). Chromosomal DNA was restricted with *Eco*RI, and the fragments were separated on a 0.5% agarose gel. Southern transfer onto nylon membranes and hybridization at 65°C in solutions without formamide were performed by standard methods (1). Hybridization was followed by two 15-min washes under nonstringent conditions in 2× SSC–0.1% (wt/vol) sodium dodecyl sulfate (SDS) at room temperature (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate.) Hybridization and washing conditions were such that they permitted 35% mismatch at the nucleotide level. Hybrids were detected by means of the digoxigenin (DIG) chemiluminescent detection system as described in the instructions of the manufacturer (Boehringer Mannheim).

DIG-labeled DNA fragments were generated by random-primed incorporation of DIG-dUTP with the DIG DNA labeling system of Boehringer Mannheim. To generate an *invA*-specific probe, an internal fragment of this gene was amplified from *S. typhimurium* as described previously (33) and the PCR product was subsequently labeled as described above.

Nucleotide sequence analysis. The nucleotide sequence was determined for a 5.5-kb fragment of the 31-cs boundary of SPI2 and a 4.6-kb fragment of the 30-cs boundary of SPI2. Sequencing was performed by the dideoxy method (39) with the Pharmacia T7 sequencing system. Sequence assembly was performed with AssemblyLign and MacVector software (Kodak) on a Macintosh PowerPC. Database searches were performed with the BLAST and FASTA programs, and sequence alignments were generated by means of the BESTFIT program of the Genetics Computer Group package (14). The DNA sequence of *E. coli* K-12 strain MG1655 was determined by the Sanger dideoxy method using dye terminator chemistry on ABI 377 sequencing instruments. The Janus m13 vector (10) was employed to isolate shotgun library clones from a mechanically sheared DNA fragment containing the relevant portion of the *E. coli* genome. The fragment, isolated from a pulsed-field gel, was defined by *I-Sce*I sites introduced into the genome on mini-Tn*10* transposons (5, 36). After initial assembly at fivefold coverage of *E. coli* DNA sequence, selected Janus clones were flipped to establish closure and minimal depth of coverage three times at each point, including at least one determination on each strand. Sequence assembly and database searches were performed with DNAstar's SEQMAN and GENEMAN software packages. Open reading frames (ORFs) were identified by the GenePlot program based on the algorithm of Borodovsky and McIninch (7).

Nucleotide sequence accession numbers. The nucleotide sequences described in this paper have been deposited in the EMBL database under accession numbers X99944 (*S. typhimurium* SPI2 31-cs boundary) and X99945 (*S. typhimurium* SPI2 30-cs boundary) and in GenBank under accession number U68703 (*E. coli* chromosomal region *ribC* to *pykF*).

RESULTS

The 31-cs boundary. The DNA sequence of the 5.5-kb *Pst*I-*Kpn*I fragment of plasmid p7-13 (Table 1) covering the end of SPI2 towards 31 cs was determined. We refer to the orientation of the boundaries by their approximate chromosomal positions as the 31-cs boundary and the 30-cs boundary based on the genetic map of *S. typhimurium* edition VIII (38). A series of ORFs was identified in the 5.5-kb fragment (Fig. 1B) as follows. (i) A 5′ portion of an ORF is identical over 15 amino acid residues to the N-terminal region of *E. coli* RibC (SwissProt accession number P29015). (ii) An ORF with a coding capacity for a protein of 456 amino acids is 94% identical over 230 amino acid residues with the N-terminal portion of the hypothetical protein YdhE of *E. coli*. The similarity probably extends beyond this region, but DNA sequence coding for only these N-terminal 230 amino acid residues of *E. coli* YdhE was available (SwissProt accession number P37340). Database searches also revealed similarity to the hypothetical protein YdhE of *Haemophilus influenzae* (SwissProt accession number P45272; 46% identity over 460 amino acid residues). (iii) Downstream of the *ydhE* homolog of *S. typhimurium*, there are 77 bp with 92% identity to the *E. coli* tRNA^{Val} gene. In *E. coli*, the tRNA^{Val} gene at 37.5 cs is present as a tandem repeat with six base changes (tRNA$^{Val\ vw}$) (24). However, there is only one tRNA^{Val} gene at the corresponding position in *S. typhimurium*, showing equivalent sequence similarity to *E. coli* tRNA$^{Val\ v}$ and tRNA$^{Val\ w}$. We refer to the *S. typhimurium* tRNA^{Val} as tRNA$^{Val\ v}$. (iv) The tRNA$^{Val\ v}$ gene in *S. typhimurium* is followed by the first ORF of SPI2, which encodes a protein similar to YscU of *Yersinia pseudotuberculosis* (36% identity over 351 amino acid residues) and which lacks a counterpart in *E. coli* (40). This ORF is followed by ORFs encoding proteins with similarity to YscT, YscS, YscR, and YscQ of *Y. pseudotuberculosis* (4). These ORFs represent genes encoding components of the type III secretion system of SPI2. A detailed analysis of these and other genes in SPI2 will be published elsewhere.

The DNA base composition changes dramatically at the 31-cs boundary (Fig. 1B). The *ribC*, *ydhE*, and tRNA^{Val} genes collectively have a G+C content of 50.6%, but the five ORFs with similarity to YscQRSTU of *Y. pseudotuberculosis* at the 31-cs end of SPI2 have a significantly lower G+C content of 41.4%.

The 30-cs boundary. The 30-cs boundary of SPI2 was analyzed by DNA sequencing of a 4.6-kb *Hin*dIII-*Xba*I fragment from plasmid pTT-1 (Table 1). Database searches with the deduced amino acid sequences of two ORFs, ORF32 and ORF48 (which could encode proteins of 31.6 and 48.5 kDa), revealed similarity to the sequence of proline iminopeptidase (Pip) of *Lactobacillus delbrueckii* (SwissProt accession number P46544; 29% identity over 295 amino acid residues) and that of the hypothetical amino acid permease YeeF of *E. coli* (SwissProt accession number P33016; 32% identity over 449 amino acid residues), respectively. Previous analysis indicated that this region is specific to *Salmonella* and may be just inside the 30-cs boundary of the island, because the region downstream of ORF48 is present in both *S. typhimurium* and *E. coli* K-12 (40). This observation was confirmed by the identification of the 52.9-kDa ORF starting 401 bp downstream of ORF48 encoding a homolog of *E. coli* PykF (31). The deduced amino acid sequence is 94% identical to the sequence of *E. coli* PykF over 463 residues. In contrast to the sharp change in G+C content at the 31-cs boundary of SPI2, no substantial change in G+C content was found when regions present only in *S. typhimurium* and those present in both *E. coli* and *S. typhimurium* were compared (Fig. 1B). Average G+C contents of 49.9 and 48.7% were observed for the two ORFs specific to *Salmonella* and *pykF*, respectively (Fig. 1B).

Comparison of the insertion point of SPI2 with the corresponding region in *E. coli* K-12. We found that SPI2 is located between homologs of the *E. coli* genes *ydhE* and *pykF* at 30 to

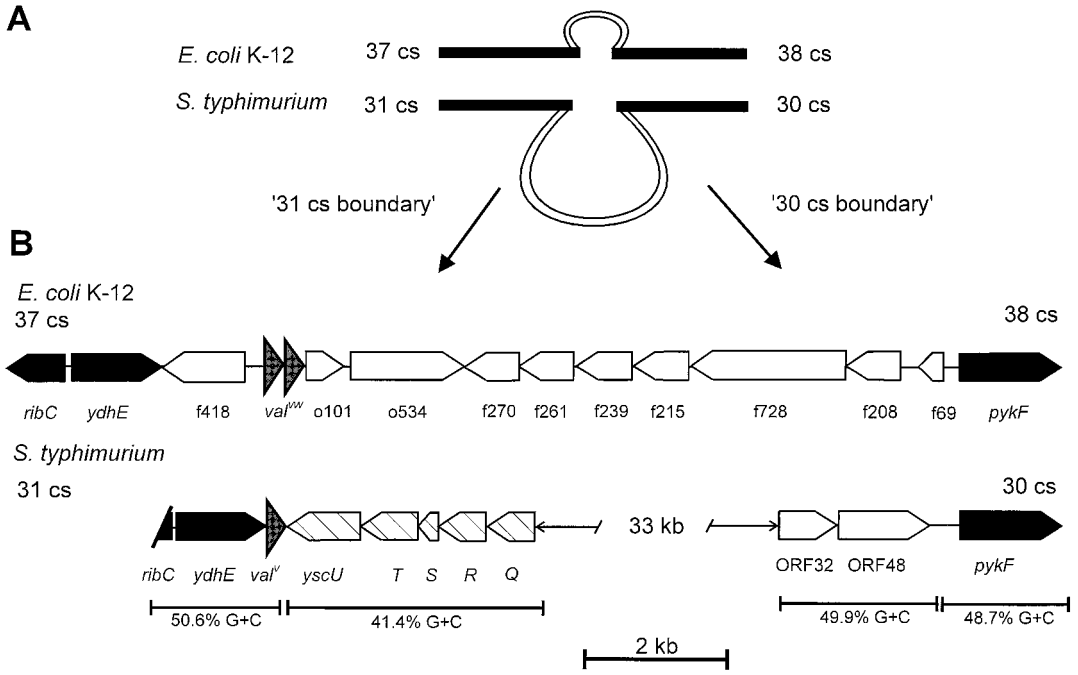

FIG. 1. Organization of the boundaries of SPI2. (A) Schematic representation of the chromosomal region of SPI2 of *S. typhimurium* aligned with the corresponding region in *E. coli* K-12. Filled lines represent regions common to both species. Unfilled lines represent regions not present in the other species. (B) Map of ORFs at the 31- and 30-cs boundaries of SPI2 aligned with the corresponding region in *E. coli* K-12. Filled arrows represent genes present in both species. Open and hatched arrows represent genes present only in *E. coli* or *S. typhimurium*. Hatched arrows represent genes in SPI2 encoding components of the type III secretion system. The G+C content of DNA fragments outside and inside of SPI2 were calculated for a 5.5-kb DNA fragment for the 31-cs boundary and a 4.6-kb fragment for the 30-cs boundary, respectively.

31 cs on the chromosome of *S. typhimurium* (40). To determine the boundaries of SPI2 more precisely and to investigate the organization of the corresponding region of the *E. coli* chromosome, the physical distance between *E. coli ydhE* and *pykF* was determined. DNA fragments derived from *ydhE* and *pykF* of *E. coli* K-12 were used as probes in hybridization experiments with digests of *E. coli* K-12 genomic DNA and with digests of phage 319 of the Kohara ordered library, which covers the corresponding chromosomal region of *E. coli* (23). A distance of approximately 9 kb was determined between *E. coli ydhE* and *pykF* (data not shown). The physical mapping (40) as well as the DNA sequencing of SPI2 indicated a distance of 39 kb between $tRNA^{Val\ v}$ and *pykF* of *S. typhimurium*. A schematic representation comparing the chromosomal region of *S. typhimurium* with that of *E. coli* is given in Fig. 1A. The *E. coli* chromosomal DNA sequence from 37 to 38 cs was determined as part of the *E. coli* genome project and was used for further comparative analysis between both species. There is an ORF (f418) present in *E. coli* between *ydhE* and $tRNA^{Val\ v}$ which is absent from the corresponding region in the *S. typhimurium* genome. The region between the *S. typhimurium* $tRNA^{Val\ v}$ gene and the *yscU* homolog shows no similarity to that of *E. coli* $tRNA^{Val\ w}$ or the region downstream from this (Fig. 2A). An alignment of DNA sequences at the proposed boundary of SPI2 at 30 cs is shown in Fig. 2B.

Database searches were performed with ORFs predicted from the DNA sequence between *E. coli* $tRNA^{Val\ vw}$ and *pykF* (see Fig. 1B). Amino acid sequences deduced from ORFs f270, f215, o534, and f69 did not reveal significant similarity to known or potential protein sequences in the databases. ORF f728 is 31.5% identical to the aldehyde:ferredoxin oxidoreductase of *Pyrococcus furiosus* (EMBL accession number X79777). ORF f208 and f239 both encode an amino acid sequence motif characteristic of electron-transporting proteins (4Fe4S cluster). Overall, the deduced amino acid sequences for ORFs f208 and f239 are, respectively, 31.2 and 52% identical to the electron transporter PhsB of *S. typhimurium* (SwissProt accession number P37601). The product of ORF f261 is 55.8% identical to *S. typhimurium* PhsC (SwissProt accession number P37602), the membrane-anchoring protein for the hydrogen sulfide production (Phs) system. Although the products of these ORFs are similar to PhsB and PhsC, it is unlikely that they are functionally equivalent because most *E. coli* strains do not reduce thiosulfate to hydrogen sulfide (2). None of the *E. coli* ORFs has similarity to ORF32 and ORF48 at the 30-cs boundary of SPI2. When DNA fragments derived from the 9-kb *E. coli*-specific region were used to probe digests of *S. typhimurium* genomic DNA under nonstringent hybridization conditions, no hybridizing fragments were identified (data not shown), indicating that DNA corresponding to the region between $tRNA^{Val\ vw}$ and *pykF* of *E. coli* is not present at another position on the *S. typhimurium* chromosome. The sequence analysis of both boundary regions did not reveal phage attachment sites or sequence repeats of significant length.

Distribution of SPI1 and SPI2 among the salmonellae. Hybridization analysis using a probe derived from SPI2 suggested that this PAI is present throughout the salmonellae but is not present in related bacterial species (40). To determine the distribution of SPI1 and SPI2 in *Salmonella* spp., hybridization studies were performed under nonstringent conditions with a larger number of *Salmonella* isolates whose genetic relationships have been established (34) (Fig. 3). Probes specific for the *invA* gene on SPI1 of *S. typhimurium* hybridized to DNA from all *Salmonella* isolates tested. However, probes corresponding to the homologs of genes of the type III secretion system in SPI2 (Fig. 3, probes 7-12 and 7-5) did not hybridize

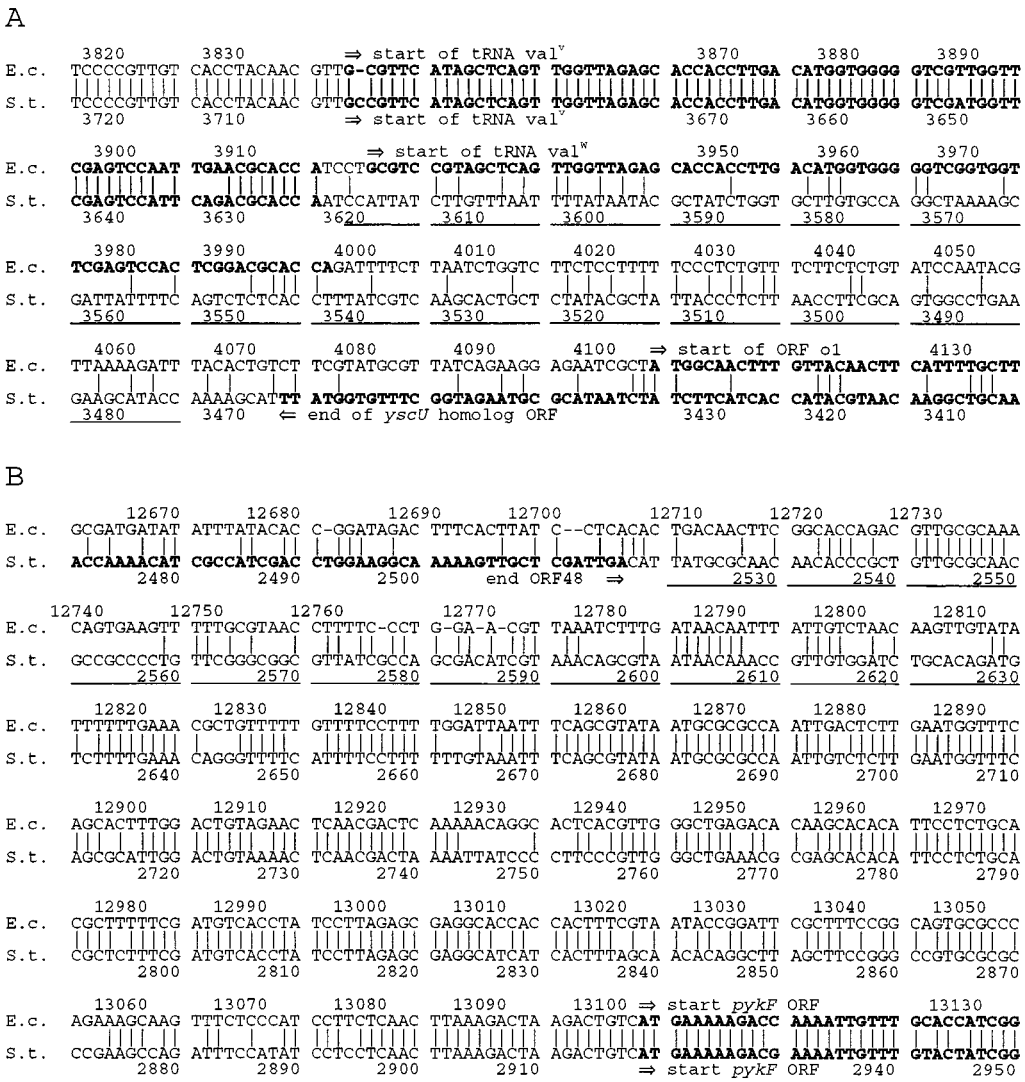

FIG. 2. Nucleotide sequence alignments at the boundaries of SPI2. The DNA sequence of the 31-cs boundary (A) and the 30-cs boundary (B) of SPI2 (S.t.) is aligned with the DNA sequence of the corresponding regions of the *E. coli* chromosome (E.c.). Alignments were performed with the MacVector program. The numbering of nucleotides corresponds to that of GenBank accession number U68703 for the *E. coli* sequence and to GenBank accession numbers X99944 and X99945 for the *S. typhimurium* sequence in panels A and B, respectively. Coding sequences and transcriptional orientation of genes are indicated by bold characters and arrows, respectively. The underlined regions are likely to include the site at which SPI2 inserted into the genome.

to serovars of *S. bongori* but did hybridize to all serovars of *S. enterica*. Probes corresponding to regions outside SPI2 (Fig. 3, probes 7-19 and TT-3) hybridized to all *Salmonella* isolates as well as to *E. coli* K-12. Phylogenetic analysis of *Salmonella* (9, 34) groups the large number of serovars in two species. These are *S. enterica*, comprising a large number of pathogenic serotypes, including *S. typhimurium*, *Salmonella typhi*, and *Salmonella enteritidis*, and the distant relative *S. bongori*. Hybridization experiments were performed with a DNA fragment derived from the region of SPI2 approximately 2.5 kb from the 30-cs boundary (Fig. 3A, probe 1-12). In contrast to the results obtained with probes 7-12 and 7-5 (Fig. 3A), hybridization signals were not obtained with six of the serovars of *S. enterica* but were obtained with two of the *S. bongori* serovars (Fig. 3B).

The PAIs of uropathogenic *E. coli* (35) and SPI1 of some serotypes of *Salmonella* (18) show instability. To investigate whether SPI2 is unstable, 12 *Salmonella* serovar *enteritidis* isolates from infected humans as well as the *Salmonella* Reference Collection B (8) containing 72 serotypes of the *S. enterica* subspecies I were analyzed for the presence of SPI2. By use of probe 1-12 (Fig. 3A), the presence of SPI2 was confirmed for all isolates analyzed (data not shown).

DISCUSSION

The genetic analysis of virulence in various bacterial pathogens has shown that many of the associated genes are closely linked or carried on PAIs. Clusters of virulence genes have been found on plasmids, for example, in *Yersinia* spp. (13) and *Shigella* spp. (32), and chromosomally located PAIs have been identified in uropathogenic (6) and enteropathogenic *E. coli* (27) and *Dichelobacter nodosus* (12). The *S. typhimurium* genome contains at least two PAIs, one at 63 cs (SPI1) containing genes for a type III secretion system involved in epithelial cell entry (29), and a second at 30.7 cs (SPI2), which also contains genes for a type III secretion system, the function of which is

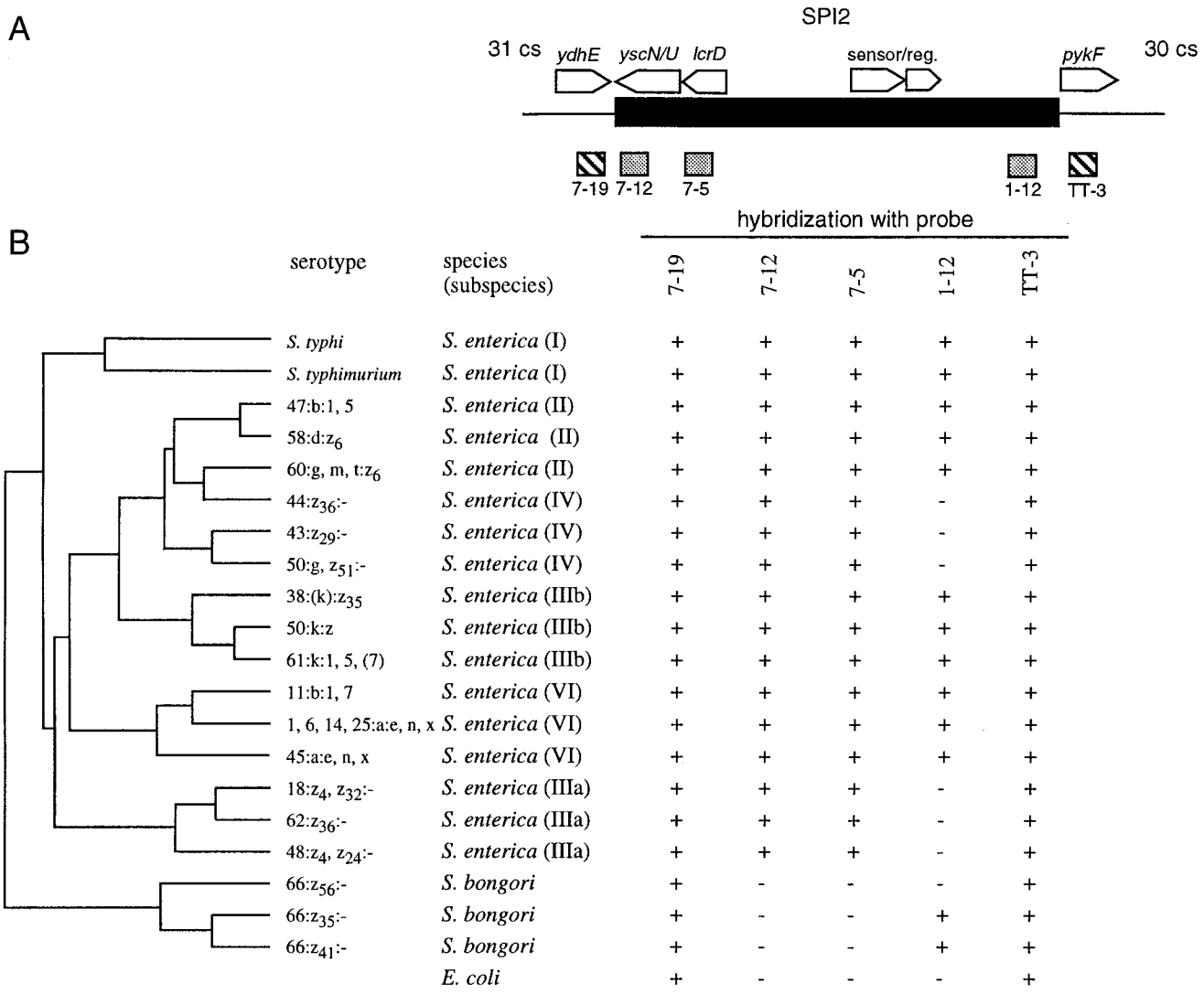

FIG. 3. Distribution of SPI1 and SPI2 among the salmonellae. (A) Schematic representation of SPI2 and approximate positions of genes which have been identified previously (40). The approximate locations of probes used for hybridization studies from within SPI2 (7-13, 7-5, 1-12 [stippled]) and adjacent regions outside SPI2 (7-19, TT-3 [hatched]) are indicated. A probe complementary to SPI1 was generated by the PCR with primers specific for *invA* of *S. typhimurium*. (B) Results of Southern hybridization analysis between probes in panel A and restriction digests of genomic DNA of various *Salmonella* isolates and *E. coli* K-12. The presence or absence of hybridizing fragments is indicated by + and −, respectively. The tree indicating the relationship of the various *Salmonella* isolates was derived by Reeves et al. (34) using multilocus enzyme electrophoresis analysis.

different from that encoded by SPI1 (30, 40). A characteristic of PAIs is that they are large chromosomal loci often inserted in the vicinity of tRNA genes (6, 12) or flanked by insertion sequence elements (16). The PAIs identified to date also have G+C compositions different from those of the rest of the genomes in which they reside. These observations have suggested that PAIs could be transferred between different bacteria by transducing phages (11). Indeed, some bacteriophages carry tRNA genes in their genomes which could facilitate site-specific integration via their conserved sequences into bacterial genomes (11). We found that the *S. typhimurium* $tRNA^{Val\ v}$ gene lies between the *ydhE* (present in both *E. coli* and *S. typhimurium*) and the *Salmonella*-specific gene whose product is similar to *Y. pseudotuberculosis* YscU. The high degree of sequence identity between *E. coli* and *S. typhimurium* in the region immediately to the left of the 3′ end of *E. coli ydhE* (Fig. 1B) and the lack of significant sequence similarity downstream from *E. coli* $tRNA^{Val\ w}$ suggest that SPI2 inserted at this site. The absence of f418 and a second copy of $tRNA^{Val}$ in *S. typhimurium* could be explained by their loss, probably during the SPI2 insertion event. Alternatively, f418 may have been lost from *E. coli* and the $tRNA^{Val}$ may have undergone a duplication after divergence of the two species from their common ancestor. Since the uropathogenic *E. coli* PAI-associated tRNA genes influence PAI gene expression (35), it is conceivable that $tRNA^{Val\ v}$ has a similar role as that of SPI2. Comparison of *E. coli* and *S. typhimurium* nucleotide sequences at the 30-cs region indicates that the boundary might lie in the 113-bp region between the 3′ end of ORF48 and the start of the *pykF* promoter. The nucleotide sequence comparison also showed that in *E. coli* there is 8.9 kb of *E. coli*-specific DNA at the site occupied in *S. typhimurium* by SPI2. This region may have been present originally in *S. enterica* and lost as a result of the SPI2 insertion event or may have been acquired subse-

quently by *E. coli*. An examination of the corresponding region in the *S. bongori* genome could help to resolve this question.

The absence of the type III secretion system genes of SPI2 in *S. bongori* indicates that these genes were probably acquired by *S. enterica* after the divergence of these two species. The finding that probe 1-12 failed to hybridize to DNA from several *S. enterica* isolates but did hybridize to two of the *S. bongori* isolates warrants further investigation. It is of interest that none of the virulence-attenuating transposon insertions that we have obtained in SPI2 map to the 13-kb region between ORF13, which encodes a regulatory subunit of a two-component regulatory system, and the 30-cs boundary (40). These observations suggest that this region might be dispensable for the virulence function of SPI2, and we are currently mutating this region to address this question.

A comparison of SPI2 genes encoding components of its type III secretion system with homologs in *Salmonella* and other bacteria showed that the structure and arrangement of SPI2 genes are no more similar to those of the SPI1 genes than they are to those of homologs in *Yersinia* spp., *Shigella* spp., or plant pathogens (22a, 40). These observations favor the hypothesis that SPI1 and SPI2 were acquired independently rather than by duplication of one of the PAIs. In contrast to SPI2, SPI1 is present in *S. bongori*, which has been shown by multilocus enzyme electrophoresis analysis to be the most distal and perhaps phylogenetically oldest group within the salmonellae (8, 9, 34). Therefore, it seems reasonable to propose that SPI1 was acquired first, leading to the ability of *Salmonella* to invade epithelial cells, thereby escaping competition from other bacterial species in the intestine (15). The subsequent acquisition of SPI2 may then have assisted *S. enterica* to survive in the internal organs of the host organism. Experimental support for this hypothesis might be gained from examination of the host range and pathology associated with *S. bongori.*

The discovery of two complex PAIs in the *S. typhimurium* genome raises the question of the existence of additional PAIs (and possibly type III secretion systems) in *Salmonella*. The horizontal transfer of blocks of virulence genes (19, 21) enabling exploitation of a host tissue or organ as a growth habitat seems to have played an important part in the evolution of *Salmonella* virulence.

ACKNOWLEDGMENTS

This project was supported by an MRC grant to D.H. and an NIH grant to A.B. M.H. was a fellow of the HCM program of the E.U., and work in Munich is supported by the DFG (grant He 1964/2-1). Sequencing of the *E. coli* genome is supported by grant PO1 HGO 1428 from the Human Genome Mapping Project of the National Institutes of Health (Bethesda, Md.). This work also benefited from the use of the HGMP facilities and the ECDC database.

We thank Andries Guilde for technical assistance, Jürgen Heesemann and Frederick Heffron for discussion and support, and Herbert N. Arst for critical review of the manuscript.

ADDENDUM IN PROOF

After acceptance of this paper, the distribution within *Salmonella* spp. of a 5.7-kb *Bam*HI DNA fragment representing the central region of SPI2 was reported (H. Ochman and E. A. Groisman, Infect. Immun. **64:**5410–5412, 1996).

REFERENCES

1. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York, N.Y.
2. **Barrett, E. L., and M. A. Clark.** 1987. Tetrathionate reduction and production of hydrogen sulphide from thiosulfate. Microbiol. Rev. **51:**192–205.
3. **Benson, N. R., and B. S. Goldman.** 1992. Rapid mapping in *Salmonella typhimurium* with Mu*d*-P22 prophages. J. Bacteriol. **174:**1673–1681.
4. **Bergman, T., K. Erickson, E. Galyov, C. Persson, and H. Wolf-Watz.** 1994. The *lcrB* (*yscN*/*U*) gene cluster of *Yersinia pseudotuberculosis* is involved in Yop secretion and shows high homology to the *spa* gene clusters of *Shigella flexneri* and *Salmonella typhimurium*. J. Bacteriol. **176:**2619–2626.
5. **Bloch, C. A., C. K. Rode, V. H. Obreque, and J. Mahillon.** 1996. Purification of *Escherichia coli* chromosomal segments without cloning. Biochem. Biophys. Res. Commun. **223:**104–111.
6. **Blum, G., M. Ott, A. Lischewski, A. Ritter, H. Imrich, H. Tschape, and J. Hacker.** 1994. Excision of large DNA regions termed pathogenicity islands from tRNA-specific loci in the chromosome of an *Escherichia coli* wild-type pathogen. Infect. Immun. **62:**606–614.
7. **Borodovsky, M., and J. McIninch.** 1993. GenMark: parallel gene recognition for both DNA strands. Comput. Chem. **17:**123–133.
8. **Boyd, E. F., F. S. Wang, P. Beltran, S. A. Plock, K. Nelson, and R. K. Selander.** 1993. *Salmonella* reference collection B (SARB): strains of 37 serovars of subspecies I. J. Gen. Microbiol. **139:**1125–1132.
9. **Boyd, E. F., F. S. Wang, T. S. Whittam, and R. K. Selander.** 1996. Molecular genetic relationships of the salmonellae. Appl. Environ. Microbiol. **62:**804–808.
10. **Burland, V. A., D. L. Daniels, G. Plunkett, and F. R. Blattner.** 1993. DNA sequencing on both strands: the Janus strategy. Nucleic Acids Res. **21:**3385–3390.
11. **Cheetham, B. F., and M. E. Katz.** 1995. A role for bacteriophages in the evolution and transfer of bacterial virulence determinants. Mol. Microbiol. **18:**201–208.
12. **Cheetham, B. F., D. B. Tattersall, G. A. Bloomfield, J. I. Rood, and M. E. Katz.** 1995. Identification of a gene encoding a bacteriophage-related integrase in a *vap* region of the *Dichelobacter nodosus* genome. Gene **162:**53–58.
13. **Cornelis, G.** 1994. *Yersinia* pathogenicity factors. Curr. Top. Microbiol. Immunol. **192:**243–263.
14. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
15. **Falkow, S.** 1996. The evolution of pathogenicity in *Escherichia*, *Shigella*, and *Salmonella*, p. 2723–2728. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, Jr., B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.
16. **Fetherston, J. D., P. Schuetze, and R. D. Perry.** 1992. Loss of the pigmentation phenotype in *Yersinia pestis* is due to the spontaneous deletion of 102 kb of chromosomal DNA which is flanked by a repetitive element. Mol. Microbiol. **6:**2693–2704.
17. **Galan, J. E.** 1996. Molecular genetic bases of *Salmonella* entry into host cells. Mol. Microbiol. **20:**263–271.
18. **Galan, J. E., and P. J. Sansonetti.** 1996. Molecular and cellular bases of *Salmonella* and *Shigella* interactions with host cells, p. 2757–2773. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, Jr., B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.
19. **Groisman, E. A., and H. Ochman.** 1993. Cognate gene clusters govern invasion of host epithelial cells by *Salmonella typhimurium* and *Shigella flexneri*. EMBO J. **12:**3779–3787.
20. **Groisman, E. A., and H. Ochman.** 1994. How to become a pathogen. Trends Microbiol. **2:**289–293.
21. **Groisman, E. A., M. H. Saier, Jr., and H. Ochman.** 1992. Horizontal transfer of a phosphatase gene as evidence for mosaic structure of the *Salmonella* genome. EMBO J. **11:**1309–1316.
22. **Hensel, M., J. E. Shea, C. Gleeson, M. D. Jones, E. Dalton, and D. W. Holden.** 1995. Simultaneous identification of bacterial virulence genes by negative selection. Science **269:**400–403.

22a. **Hensel, M.** Unpublished data.

23. **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.
24. **Komine, Y., T. Adachi, H. Inokuchi, and H. Ozeki.** 1990. Genomic organisation and physical mapping of the transfer RNA genes in *Escherichia coli* K12. J. Mol. Biol. **212:**579–598.
25. **Lilleengen, L.** 1948. Typing *Salmonella typhimurium* by means of bacteriophage. Acta Pathol. Microbiol. Scand. Suppl. **77:**11–125.
26. **Maurer, R., B. C. Osmond, E. Shekhtman, A. Wong, and D. Botstein.** 1984. Functional interchangeability of DNA replication genes in *Salmonella typhimurium* and *Escherichia coli* demonstrated by a general complementation procedure. Genetics **108:**1–23.
27. **McDaniel, T. K., K. G. Jarvis, M. S. Donnenberg, and J. B. Kaper.** 1995. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. Proc. Natl. Acad. Sci. USA **92:**1664–1668.
28. **Miller, S. I., E. L. Hohmann, and D. A. Pegues.** 1995. *Salmonella* (including *Salmonella typhi*), p. 2013–2033. *In* G. L. Mandell, J. E. Bennett, and R. Dolin (ed.), Mandell, Douglas and Bennett's principles and practice of infectious diseases. Churchill Livingstone, Inc. New York, N.Y.

29. **Mills, D. M., V. Bajaj, and C. A. Lee.** 1995. A 40 kb chromosomal fragment encoding *Salmonella typhimurium* invasion genes is absent from the corresponding region of the *Escherichia coli* K-12 chromosome. Mol. Microbiol. **15:**749–759.
30. **Ochman, H., F. C. Soncini, F. Solomon, and E. L. Groisman.** 1996. Identification of a pathogenicity island required for *Salmonella* survival in host cells. Proc. Natl. Acad. Sci. USA **93:**7800–7804.
31. **Ohara, O., R. L. Dorit, and W. Gilbert.** 1989. Direct genomic sequencing of bacterial DNA: the pyruvate kinase I gene of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **86:**6883–6887.
32. **Parsot, C.** 1994. *Shigella flexneri*: genetics of entry and intercellular dissemination in epithelial cells. Curr. Top. Microbiol. Immunol. **192:**217–241.
33. **Rahn, K., S. A. De Grandis, R. C. Clarke, S. A. McEwen, J. E. Galan, C. Ginocchio, R. Curtiss, and C. L. Gyles.** 1992. Amplification of an *invA* gene sequence of *Salmonella typhimurium* by polymerase chain reaction as a specific method of detection of *Salmonella*. Mol. Cell Probes **6:**271–279.
34. **Reeves, M. W., G. M. Evins, A. A. Heiba, B. D. Plikaytis, and J. J. Farmer III.** 1989. Clonal nature of *Salmonella typhi* and its genetic relatedness to other salmonellae as shown by multilocus enzyme electrophoresis, and proposal of *Salmonella bongori* comb. nov. J. Clin. Microbiol. **27:**313–320.
35. **Ritter, A., G. Blum, L. Emody, M. Kerenyi, A. Bock, B. Neuhierl, W. Rabsch, F. Scheutz, and J. Hacker.** 1995. tRNA genes and pathogenicity islands: influence on virulence and metabolic properties of uropathogenic *Escherichia coli*. Mol. Microbiol. **17:**109–121.
36. **Rode, C. K., V. H. Obreque, and C. A. Bloch.** 1995. New tools for integrated genetic and physical analysis of the *Escherichia coli* chromosome. Gene **166:**109–121.
37. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
38. **Sanderson, K. E., A. Hessel, and K. E. Rudd.** 1995. Genetic map of *Salmonella typhimurium*, edition VIII. Microbiol. Rev. **59:**241–303.
39. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
40. **Shea, J. E., M. Hensel, C. Gleeson, and D. W. Holden.** 1996. Identification of a virulence locus encoding a second type III secretion system in *Salmonella typhimurium*. Proc. Natl. Acad. Sci. USA **93:**2593–2597.