The High-Pathogenicity Island Is Absent in Human Pathogens of *Salmonella enterica* Subspecies I but Present in Isolates of Subspecies III and VI

T. A. Oelschlaeger,1 D. Zhang,1 S. Schubert,2 E. Carniel,3 W. Rabsch,4 H. Karch,5 and J. Hacker1*

Institut für Molekulare Infektionsbiologie, University of Würzburg, Würzburg, Max-von-Pettenkofer-Institut für Hygiene und Mikrobiologie, Munich, Robert-Koch-Institut, Wernigerode, and Institut für Hygiene, Universität Münster, Münster, Germany, and Institut Pasteur, Paris, France

Received 10 July 2002/Accepted 13 November 2002

In this study we tested 74 *Salmonella* strains of all eight *Salmonella* groups and were able to demonstrate the presence of two high-pathogenicity island types in strains of *Salmonella* groups IIIa, IIIb, and VI. Most high-pathogenicity island-positive isolates produced yersiniabactin under iron-limited conditions and were positive for the high-molecular-weight proteins HMW1 and HMW2.

The presence of pathogenicity islands in the genomes of bacterial pathogens is one of the main features that differentiate them from closely related non-pathogenic strains or species. Furthermore, particular pathogenicity islands are specific for certain pathotypes. The high-pathogenicity island (HPI), however, has been found in many species of the family *Enterobacteriaceae* (27). The term high-pathogenicity island was coined because this island was shown to be present in all strains of *Yersinia* that are highly virulent for mice or humans but absent from strains with lower virulence (10, 11). This island consists of 12 genes and encodes a highly efficient iron acquisition system driven by the siderophore yersiniabactin. A P4-like integrase (*int*) gene is located at the 5′ end of HPI. The genes ybtS (*irp9*), *irp1*, *irp2*, *ybtT* (*irp4*), and *ybtE* (*irp5*) form the yersiniabactin synthesis gene cluster. Two genes, *ybtQ* (*irp7*) and *ybtP* (*irp6*), encode proteins of the ABC transporter family. The receptor protein for yersiniabactin and pesticin in the outer membrane is encoded by *fyuA*, located at the 3′ end of the HPI core. The expression of HPI-specific genes is induced by iron-limiting conditions and is regulated by the *ybtA* product (13, 22). The gene termed *ybtX* (*irp8*) is of unknown function.

Besides the three pathogenic *Yersinia* species, the HPI was identified in various pathotypes of *Escherichia coli* and in *E. coli* isolates from stool samples of healthy individuals. In addition, HPI was detected in some strains of *Citrobacter* spp., *Enterobacter cloacae*, and *Klebsiella* spp. (3, 28). However, not all HPI-positive strains of the different species of the family *Enterobacteriaceae* produce the iron chelator yersiniabactin or the yersiniabactin receptor (27, 28). Despite several previous attempts, the HPI could not be detected in any *Salmonella* strains belonging to *Salmonella enterica* serovars Typhimurium, Enteritidis, or Typhi, which are pathogenic for humans (3, 27).

In the present work, we investigated various *Salmonella* strains for the presence of the HPI not only of *S. enterica* group I, comprising the majority of *salmonellae* pathogenic for humans and other warm-blooded animals, but also of the other groups of *S. enterica* as well as strains of the species *Salmonella bongori*.

**HPI is present in particular subspecies of *S. enterica***. The genus *Salmonella* consists of two species, *S. enterica*, with seven subspecies, also termed taxonomic groups (I to IV, VI, and VII), and *S. bongori* (group V). The presence of HPI was investigated for 74 *Salmonella* strains, including all strains in *Salmonella* reference collection C (SARC) (5). This was performed by PCRs with primers specific for HPI genes (Fig. 1). The arrangement of the HPI genes was determined by additional PCRs with primer pairs of which each primer was specific for one of two adjacent genes (Fig. 1). All primers were designed according to the HPI sequence of *Yersinia pestis* strain KIM6+ (14). The sequences of the primers employed and the conditions used are listed in a table which can be found on our home page (http://www.uni-wuerzburg.de/infektionsbiologie/PCR-HPI.htm).

Of the 74 strains tested, one strain (no. 1649) of nine of group IIIa, all 11 strains of group IIIb, and two of eight strains of group VI were positive, i.e., resulted in PCR products of the expected sizes in the initial screening (Table 1). For all these strains, the genes covering the HPI from *ybtS* to *fyuA* mapped at the same positions as the corresponding genes of *Y. pestis* strain KIM6+ (14). The only exception was strain 1380 of group IIIb. Its HPI lacked 80% of *irp2* and all of the genes *irp1*, *irp3* (*ybtU*), *irp4* (*ybtT*), *irp5* (*ybtE*), and *fyuA*. The presence of only one HPI per genome of the HPI-positive strains was demonstrated by Southern blot analysis (29) of *Bst*III-digested chromosomal DNA probed either with a PCR fragment representing part of the *fyuA* sequence (PCR 30, Fig. 1A; Fig. 2) or with a PCR fragment obtained with the *ybtS*-specific primers *ybtS4* and *ybtS8* (PCR 31, Fig. 1A) (blot not shown).

**Characterization of Salmonella HPI core region**. Typically, the HPI identified in *Yersinia* spp. and *E. coli* contains at its 5′ end an *int* gene highly homologous to the P4-like integrase gene of *E. coli* (4, 6, 7, 17). This integrase gene located downstream of *ybtS* was only present in the HPI-positive strains...
SARC13 and 1443 of group VI (Fig. 1A). The nucleotide sequence of \textit{int} was identical to that of \textit{Y. pestis} strain KIM6+/H11001.

Nucleotide sequences were compared to those in the data bases with the programs BlastN and BlastX (1).

The HPI of strains SARC13 and 1443 was termed type 1. In contrast, the \textit{int} gene was missing in all HPI-positive strains of groups IIIa and IIIb. Other differences between the HPIs of groups IIIa and IIIb compared to those of group VI and \textit{Yersinia} strains were determined by complete (\textit{int} and ERIC element) or partial sequencing (\textit{ybtSXQPA}, \textit{irp1} to \textit{irp5}, and \textit{fyuA}) for all HPI-positive strains of groups IIIa and IIIb. These differences were (i) a \textit{ybtS} gene with 12 additional nucleotides at its 3' end, (ii) an additional 206 bp (i.e., an ERIC element) between \textit{ybtP} (\textit{irp6}) and \textit{ybtA}, and (iii) a mutation in \textit{irp1} replacing the adjacent amino acids Asp-Ala with Gly-Tyr (Fig. 1). The HPI of \textit{S. enterica} strains of groups IIIa and IIIb was termed type 2 HPI in order to differentiate this type of HPI from that identified in \textit{Yersinia} spp. and other enterobacteria, including \textit{S. enterica} group VI.

**Expression of HPI genes.** The functionality of yersiniabactin biosynthesis was evaluated by detecting the presence of yersiniabactin in the cell-free supernatant of iron-starved bacteria by using a translational reporter gene fusion system with luciferase (20, 28). Synthesis and release of yersiniabactin were observed in all HPI-positive strains tested except 1380, which carries a truncated island, and 50/90 and HPI-negative control strains (Table 1). In addition, expression of the high-molecular-weight proteins HMWP1 and HMWP2, representing multidomain enzymes involved in synthesis of the iron chelator yersiniabactin (9), was detected by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of total proteins extracted from five iron-starved but not iron-replete HPI-positive \textit{Salmonella} strains (IIIa, 1649; IIIb, 845/96, SARC 7, and SARC 8; and VI, 1443) and from \textit{Yersinia enterocolitica} strain Ye8081 (positive control) (data not shown).

**Determination of integration site of \textit{Salmonella} HPIs by analysis of flanking region on 5' side.** The HPI element in most enterobacteria is inserted into an asparagine-specific
tRNA gene, preferentially asnT (6, 21, 27). In order to analyze the HPI site of insertion in the Salmonella strains, PCRs specific for the asn tRNA, int, or ybtS gene were performed. PCR products of the expected sizes were observed for the two HPI-positive strains of group VI (1443 and SARC13) only (Fig. 1A). In addition, analysis of the DNA sequence upstream of the HPI-containing asn tRNA gene was performed after inverse PCR with primer pair int7 and int8 (PCR 21, Fig. 1A) and revealed 86 to 97% identity with the genomic DNA sequence of S. enterica serovar Typhimurium strain LT2 (bases 2081500 to 2083248). Within this region, a sequence 973 bp upstream of the asn tRNA gene was identified that was 100% identical with the serU tRNA gene of E. coli (Fig. 1).

In order to determine the site of insertion of the type 2 HPI in the strains not belonging to S. enterica group VI, inverse PCR starting inside ybtS (PCR 19, Fig. 1B) was performed. Cloning and sequencing of the resulting amplification products confirmed the extension of the ybtS gene and the absence of an int gene at the 3' end of the HPI in these strains (Fig. 1B). Furthermore, 113 nucleotides after the stop codon of ybtS were identical between all HPI-positive strains of Salmonella group IIIa and IIIb and showed no homology to any sequences in the databases. Starting 114 nucleotides downstream of the ybtS gene, the sequence was >90% identical to the genome sequence of S. enterica serovar Typhimurium strain LT2 in all these strains. Identity began with the sequence 5'-TTATG.

The HPI of the HPI-positive strains of Salmonella groups IIIa and IIIb was inserted at a different location of the chromosome and not into any of the asn tRNA genes.

**Determination of region flanking 3' end of HPI in Salmonella spp.** A primer pair (Fig. 1A, PCR 22) for inverse PCR in fyuA yielded an amplification product of the same size as in Yersinia pestis for strains SARCl3 and 1443 of group VI only. Sequence analysis of the amplified fragment (about 1 kb) showed 96% identity with the corresponding region downstream of fyuA in Yersinia pseudotuberculosis strain PBI (24) (Fig. 1A). The 3'-flanking region of the HPI-positive strain 845/96 of group IIIb was determined from a cosmide clone carrying the right half (3' part) of the HPI and about 30 kb of DNA downstream of fyuA. Sequence analysis of 5,237 bp start-

---

**TABLE 1. Determination of the presence of HPI in salmonellae**

<table>
<thead>
<tr>
<th>Salmonella group</th>
<th>Subspecies tested</th>
<th>No. of strains tested</th>
<th>Strain no.</th>
<th>HPI present</th>
<th>Yersiniabactin production</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>S. enterica serovar Typhimurium</td>
<td>5</td>
<td>Various</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica serovar Enteritidis</td>
<td>1</td>
<td>SZ327/94</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica serovar Indiana</td>
<td>3</td>
<td>Various</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica serovar Litchfield</td>
<td>10</td>
<td>Various</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. enterica</td>
<td>1</td>
<td>SARC1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. enterica</td>
<td>1</td>
<td>SARC2</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>II</td>
<td>S. enterica subsp. salamae</td>
<td>6</td>
<td>Various</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. arizonae</td>
<td>1</td>
<td>SARC3</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. arizonae</td>
<td>1</td>
<td>SARC4</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>IIIa</td>
<td>S. enterica subsp. arizonae</td>
<td>6</td>
<td>Various</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. diarizonae</td>
<td>1</td>
<td>1649</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. diarizonae</td>
<td>1</td>
<td>SARC5</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. diarizonae</td>
<td>1</td>
<td>SARC6</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. diarizonae</td>
<td>1</td>
<td>50/90</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>IV</td>
<td>S. enterica subsp. houtenae</td>
<td>7</td>
<td>Various</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. houtenae</td>
<td>1</td>
<td>SARC9</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. houtenae</td>
<td>1</td>
<td>SARC10</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>V</td>
<td>S. bongori</td>
<td>5</td>
<td>Various</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. bongori</td>
<td>1</td>
<td>SARC11</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. bongori</td>
<td>1</td>
<td>SARC12</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>VI</td>
<td>S. enterica subsp. indica</td>
<td>5</td>
<td>Various</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. indica</td>
<td>1</td>
<td>1443</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. indica</td>
<td>1</td>
<td>SARC13</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>S. enterica subsp. indica</td>
<td>1</td>
<td>SARC14</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>VII</td>
<td>S. enterica</td>
<td>1</td>
<td>SARC15</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>S. enterica</td>
<td>1</td>
<td>SARC16</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

* The strain number of each strain that was positive in the HPI gene detection assays is presented. Salmonella strains were tested for the presence of HPI by PCRs specific for the genes ybtS-ybtX, ypt1-ypt2, and ypt4-fyuA. Yersiniabactin production was determined with an indicator strain expressing luciferase under the control of the fyuA promoter (20). Strain 1380 of group IIIb was only positive in PCRs specific for ybtS-ybtX and for part of ypt2. ND, not determined. With the exception of the SARC strains, all but one Salmonella strain (S. enterica serovar Typhimurium strain C17) were from the strain collection of the Robert Koch Institute, Wernigerode Branch, National Reference Center for Salmonellae and other Enterics.

---

**FIG. 2.** Southern blot with chromosomal DNA from HPI-positive Salmonella strains 1649 (lane 1) (group IIIa), group IIIb strains 50/90 (lane 2), 461/95 (lane 3), 658/95 (lane 4), 845/96 (lane 5), 1376/95 (lane 6), 1380 (lane 7), SARC8 (lane 8), 1474 (lane 9), 1757 (lane 10), SARC7 (lane 11), and 1494 (lane 12), and group VI strains 1443 (lane 13) and SARC14 (lane 14). Strain SARC14 (lane 15) was the negative control, and Y. pestis strain KIM6+ (lane 16) served as the positive control. Chromosomal DNA was digested with BssHII and probed with the amplification product of PCR 30 (see table at http://www.uni-wuerzburg.de/infektionsbiologie/PCR-HPI.htm), representing part of the fyuA sequence.
ing with the first nucleotide after the stop codon of fyuA revealed no homology to any nucleotide sequence in the genome databases. However, the deduced amino acid sequence of the last 557 nucleotides of the 5,237-bp region downstream from fyuA exhibited 72% identity with a putative sulfatase of Yersinia pestis encoded by gene YPO3047 (Fig. 1B).

The sequence obtained from the cosmID containing the region downstream from fyuA was used to design primers for PCR amplification of DNA of other HPI-positive Salmonella strains. The 5,237-bp fragments that were obtained for all HPI-positive group IIb strains (50/90, 461/95, 658/95, 845/96, 1376/95, 1494, 1474, 1757, SARC7, and SARC8) and for the HPI-positive group IIIa strain 1649 had a sequence identical to that of the 3'-flanking region of strain 845/96 (Fig. 1B). The only exception was strain 1380, which contains a 3'-truncated HPI and for which no amplified fragment was detected. These results show that the 3'-flanking regions of all HPI-positive strains of groups IIIa and IIb are identical (Fig. 1B).

The yersiniabactin gene cluster is viewed as an entity which can be transferred horizontally between bacteria of the family Enterobacteriaceae. Its presence in Yersinia spp. is correlated with high virulence for mice and humans (9, 18, 19, 25). Therefore, it was termed the high-pathogenicity island. With different methods, it was shown that S. enterica group I strains did not contain the HPI. Furthermore, in other pathogenic enterobacteria such as Shigella boydii, Shigella dysenteriae, Shigella flexneri, and Shigella sonnei, HPI has not been detected (27). However, in other enterobacteria such as E. coli, HPI was also found in substantial numbers in isolates from healthy individuals and in strains from ECOR groups A and B1, which are considered nonpathogenic (12, 28). It can be concluded that HPI might not be a virulence factor in every strain background but might have other functions in commensal and environmental bacteria (16, 21).

Consequently, this study was designed to probe for the presence of HPI in Salmonella group I strains as well as in Salmonella strains of other groups. First, we confirmed the previously published data that the HPI is not present in S. enterica group I. Second, of 53 strains of Salmonella groups II to VII, 14 strains (26.4%) were identified as containing the HPI. All but one strain harbored all the HPI genes from ybtS to fyuA. In contrast to the HPI in other members of the Enterobacteriaceae family, the int gene encoding a P4-like integrase was only present in two HPI-positive Salmonella strains of group VI. This was surprising because this gene is responsible for the integration of HPI into the host chromosome (23).

Further analysis of the Salmonella HPIs revealed that two types of HPI can be distinguished in salmonellae. Type 1, found in strains of Salmonella group VI, is identical to the HPI of Yersinia pestis strain KIM6+ and other members of the enterobacteria. The other HPI type (type 2) of strains of Salmonella groups IIIa and IIb lacks the int gene but encodes a 3'-extended ybtS gene containing an insertion of a 206-bp ERIC-like element and a mutation in irp1. How was the second Salmonella HPI type inserted into the chromosome? Either it was inserted by the product of the int gene, which was lost later on, or its insertion relied on another gene product. This might be an integrase encoded somewhere else in the chromosome or an integrase provided by a phage during the integration process. The former hypothesis is unlikely, because the type 2 HPI of Salmonella spp. differs not just in the lack of the int gene from type 1 HPI but also in the other alterations mentioned. Furthermore, the integration site typical of HPI in Yersinia species as well as in E. coli is an Asn-specific tRNA gene, preferentially asrT. However, only the type 1 Salmonella HPI is integrated into an asr tRNA gene, most likely asrT (26). The Salmonella HPI of type 2 was inserted not adjacent to any tRNA but at a position 41 nucleotides downstream of the ychF gene in the genome of S. enterica serovar Typhimurium, because there the sequence becomes almost identical to the Salmonella genome.

Nevertheless, HPI might have been acquired only once by Salmonella spp. This hypothesis is based on the fact that HPI was found to be present in the three closely related Salmonella subgroups IIIa, IIb, and VI. In turn, this might be the result of HPI acquisition by a Salmonella ancestor which was the predecessor of these three subgroups. The iron-sequestration function encoded by HPI is not controversial (8), and most HPI-positive Salmonella strains produced yersiniabactin. However, this function might be of importance not only in warm-blooded animals but also in other animals or the environment as well. The technique used to identify yersiniabactin relied on the fact that yersiniabactin acts as a regulator in a complex with YbtA for the expression of its receptor, FyuA (13, 20). This regulatory function might work not only on other genes of the HPI (e.g., trp genes) but on non-HPI genes in the same bacterial cell and/or on gene expression of other bacteria expressing the FyuA receptor.

In conclusion, HPI was found in Salmonella strains of groups IIIa, IIb, and VI. In these strains, it might not only function as an iron-sequestring system in host animals and the environment but might also influence the activity of other genes as well.

This work was supported by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 479, and the Fond der Chemischen Industrie.

We thank W.-D. Hardt (Munich) for the kind gift of the SARC strains and C. Albert (Würzburg) for excellent technical assistance.

REFERENCES