

Pathogenicity Islands: Bacterial Evolution in Quantum Leaps

Minireview

Eduardo A. Groisman* and Howard Ochman†

*Department of Molecular Microbiology
Washington University School of Medicine
St. Louis, Missouri 63110

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

"I am the beneficiary of a lucky break in the genetic sweepstakes"

Isaac Asimov

The Molecular Bases of Virulence

For most bacterial pathogens virulence is a multifactorial process requiring two general classes of determinants. The first encompasses genes that participate in physiological processes necessary for survival in host and non-host environments, and these genes are generally found in both pathogenic and non-pathogenic organisms. In *Salmonella*, this set includes regulatory genes, such as *phoP/phoQ*, that control expression of more than 20 loci in response to low magnesium environments; biosynthetic genes, such as *aroA*, that encodes the ability to synthesize aromatic amino acids not found in host tissues; and several additional genes required for cell maintenance and DNA repair (Groisman and Ochman, 1994).

The second class of virulence genes specifies traits that are unique to pathogens, and, not surprisingly, these genes are rarely detected in non-pathogenic organisms. Based on the initial characterizations of plasmids from *Yersinia* and *Shigella*, such sequences were originally thought to be confined to extrachromosomal elements; but recently, several virulence cassettes have been mapped to the chromosome of pathogenic organisms. These segments of the chromosome, termed pathogenicity islands (Hacker et al., 1990; Lee, 1996; see Figure 1), can be of any length, but, most often, they accommodate large clusters of genes contributing a particular virulence phenotype. Pathogenicity islands have been recognized by their association with transmissible DNA, their sporadic distribution in a set of related pathogenic and non-pathogenic organisms, or the unusual features of their encoded genes, and their recent characterization has provided clues regarding their origins, distribution, mode of transfer, and genetic stability.

Pathogenicity Islands Dictate Disease Condition

Incorporation of a pathogenicity island can, in a single step, transform a normally benign organism into a pathogen. For example, enteropathogenic *E. coli* contain a chromosomal segment that mediates attaching and effacing lesions on intestinal epithelial cells (McDaniel et al., 1995; Figure 1). This 35 kb region—termed LEE for locus of enterocyte effacement—is absent from laboratory strains and codes for an outer membrane protein that promotes contact with host cells and a type III secretion system that exports proteins known to modify host cell functions. Type III secretion systems are prevalent among animal and plant pathogens—homologous export systems are responsible

for the secretion of virulence proteins by *Yersinia*, *Erwinia*, *Pseudomonas*, *Xanthomonas*, *Shigella*, and *Salmonella* (Van Gijsegem et al., 1993)—and these systems differ from the *sec*-dependent pathway in that their effector molecules lack a typical signal sequence.

At the same location where LEE has inserted into enteropathogenic strains (McDaniel et al., 1995) a distinct pathogenicity island, PAI-1, has been identified in a uropathogenic strain of *E. coli* (Blum et al., 1994; Figure 1). Both islands are targeted to the same basepair downstream of the *selC* gene, which codes for the selenocysteine-specific tRNA. LEE and PAI-1 contain different sets of genes: PAI-1 is 70 kb in length and encodes a hemolysin. Thus, the specific cassette incorporated at the *selC* locus determines whether strains of *E. coli* are converted into entero- or uropathogens.

Aside from PAI-1 and LEE, three other pathogenicity islands have been identified in *E. coli* (Figure 1); and, in fact, certain uropathogenic strains contain more than one pathogenicity island. In addition to PAI-1, *E. coli* strain 536 also harbors a 190-kb pathogenicity island, PAI-2, that has integrated at the *tRNA^{leuX}* locus and contains genes for an alternate hemolysin and Prf-type fimbriae (Blum et al., 1994). The other pathogenicity islands detected in *E. coli* also contain hemolysin genes but are associated with

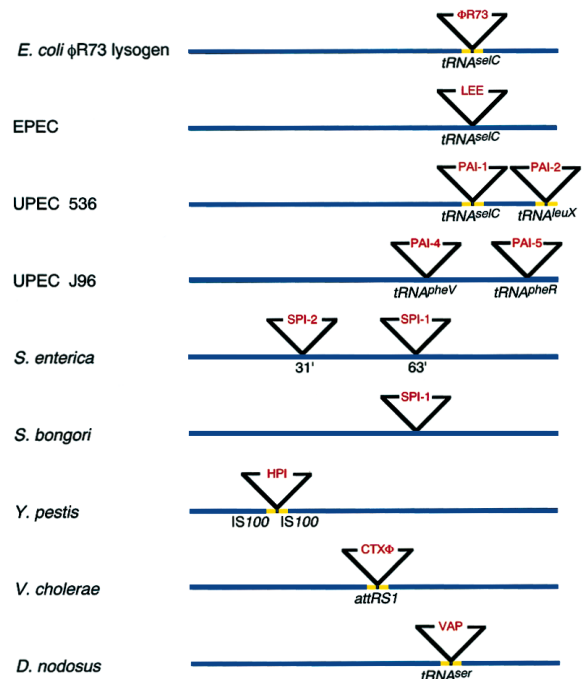


Figure 1. Location of Selected Pathogenicity Islands and Phages of Gram-Negative Bacteria

Chromosomes and pathogenicity islands are depicted as blue lines and black triangles, respectively, and are not drawn to scale. The presence of repeated sequences at the site of insertion, which is shown below the island, is indicated by short yellow lines.

different tRNA loci: the island at the *pheV* locus contains genes for P-fimbriae, which allow the pathogen to adhere to cells lining the urinary tract, whereas the island at the *pheR* locus codes for Prs fimbriae, which bind Gal Nac-1-3-Gal Nac-containing receptors rather the Gal- β -1-4-Gal residues recognized by the P fimbrial adhesins (Blum et al., 1995). In addition, the island at the *pheR* locus codes for a type 1 cytotoxic necrotizing factor, which causes multinucleation and enlargement of eukaryotic cells and can trigger phagocytosis by epithelial cells (Falzano et al., 1993). Sequences similar to those present in the pathogenicity islands of uropathogenic strains have been detected in plasmids of hemolytic strains of *E. coli*, highlighting the mobile nature of these sequences.

Pathogenicity Islands Mediate Specific Phases of the Infection Process

Two pathogenicity islands have been identified in *Salmonella*, each contributing to a specific step in the course of infection. The island at 63', designated SPI-1 (Figure 1), governs the ability of *Salmonella* to invade epithelial cells (Mills et al., 1995; Galán, 1996), and a second island at 31', designated SPI-2, mediates survival within macrophages (Ochman et al., 1996; Shea et al., 1996). Both islands are 40 kb in length and appear to have been acquired by horizontal gene transfer because their G+C contents are lower than those typical of *Salmonella* (Groisman et al., 1993).

The nucleotide sequence of the SPI-1 island is nearly complete, and this region includes at least 25 genes, the majority of which encode components of a Type III secretion system and its effector proteins (Galán, 1996). The assemblage of genes specifying this export apparatus in *Salmonella* is similar in size, sequence, and organization to a set of genes in the *Shigella* virulence plasmid (Groisman and Ochman, 1993), and several of the proteins encoded by these gene clusters are capable of cross-species complementation. However, certain genes within the *Salmonella* invasion island, such as *orgA* and *hil*, have no counterparts in the *Shigella* virulence plasmid and may be responsible for the differential regulation of invasion in *Shigella* and *Salmonella*: this property is controlled by temperature in *Shigella* and by oxygen tension, osmolarity, and pH in *Salmonella*. The vast differences in G+C contents and the phylogenetic distribution of the invasion gene clusters argue that *Salmonella* and *Shigella* gained these sequences independently from an as yet unidentified organism.

The SPI-2 island of *Salmonella* harbors at least 15 genes that code for a distinct Type III secretion system and for a two-component regulatory system (Ochman et al., 1996; Shea et al., 1996). Unlike the invasion island, the SPI-2 region is not needed for invasion of epithelial cells but is essential for intramacrophage survival (Ochman et al., 1996). Moreover, the SPI-2 island is necessary to cause a lethal infection in mice when inoculated either orally or intraperitoneally (Ochman et al., 1996; Shea et al., 1996), whereas mutants defective in SPI-1 are attenuated only when inoculated orally (Galán, 1996), indicating that genes within each of the pathogenicity islands direct distinct functions during the progression of disease.

The phylogenetic distribution of the SPI-1 and SPI-2 pathogenicity islands reveals that their acquisition was critical in the development of *Salmonella* as an intracellular

pathogen. The SPI-1 invasion island is present in representative strains from all subspecific groups of *Salmonella*, although sporadic cases of secondary loss have been reported for individual isolates of *S. enterica* sv. Seftenberg and Litchfield. But, in contrast to SPI-1, SPI-2-hybridizing sequences have not been detected in *S. bongori* (subspecies V), the most divergent lineage of *Salmonella*. The incorporation of the SPI-1 and SPI-2 pathogenicity islands enabled *Salmonella* to invade host cells, evade host defense systems, and cause systemic infections in mammals, while its close relative *E. coli* evolved as a commensal and opportunistic pathogen. On the other hand, the four nominal species of *Shigella*, all of which emerged relatively recently from *E. coli*, gained the ability to enter host cells following the acquisition of the large virulence plasmid, thereby becoming invasive pathogens of primates.

Gain and Loss of Pathogenicity Islands

The occurrence of pathogenicity islands raises several questions about the evolution of virulence in bacteria. First, what advantage does the incorporation of virulence genes into the chromosome provide over retaining them on episomes? Although many virulence determinants occur on extrachromosomal DNAs, integration into the chromosome circumvents the need for autonomously replicating elements. This also lowers the probability that a given gene will be eliminated because episomes must be under continuous selective pressure to be maintained whereas the spontaneous elimination of chromosomal genes is not as high. However, plasmid integration can result in the inactivation of chromosomal loci or in the alteration of the expression of plasmid-encoded genes: for example, insertion of the *Shigella flexneri* virulence plasmid into the *metB* locus leads to methionine auxotrophy and decreased expression of the invasion proteins (Zagaglia et al., 1991).

Second, how are pathogenicity islands incorporated into the bacterial chromosome? Analysis of the junctions of pathogenicity islands can reveal the mechanism used to integrate these sequences. As noted above, the genes encoded by PAI-1 and LEE of *E. coli* are clearly different, but the site of insertion within the *seI/C*-tRNA gene is conserved (McDaniel et al., 1995). This suggests that both islands were acquired by a similar mechanism which was likely to be phage-mediated since *seI/C* also serves as the site of insertion for the retronephage Φ R73 (Inouye et al., 1991).

Bacteriophages have been implicated in the transfer and acquisition of several other pathogenicity islands. The lysogenic conversion of *Corynebacterium diphtheria* and *E. coli* strains to pathogenic variants involves the incorporation of phages harboring toxin genes. Recently, the cholera toxin genes were shown to be encoded by a filamentous bacteriophage that can integrate into the chromosome of *Vibrio cholerae* or replicate autonomously as a plasmid (Waldor and Mekalanos, 1996). Furthermore, the virulence-associated region (*vap*) of *Dichelobacter nodosus* occurs within a tRNA locus and contains an open reading frame with high levels of sequence similarity to the integrases of *E. coli* phages (Cheetham and Katz, 1995).

Features of tRNA genes promote their use as integration sites for phages and other episomes: these sequences

are conserved across species and occur in multiple copies within a genome. The PAI-1 and PAI-2 islands of *E. coli* are flanked by direct repeats of 16 and 18 nucleotides, respectively (Blum et al., 1994), and although these islands integrate at different tRNA genes, the repeats contain a motif, TTCGA, that is within the conserved loop of tRNA genes. PAI-1 and PAI-2 are spontaneously deleted from the *E. coli* chromosome at frequencies of 10^{-4} to 10^{-5} , leaving a single copy of the repeat at the excision site and rendering the organism avirulent (Blum et al., 1994). Deletion of these pathogenicity islands can also modify metabolic properties of the cell by disrupting the tRNA gene at the site of insertion; for example, inactivation of *sefC* will impede the synthesis of enzymes containing selenocysteine and prevent growth under anaerobic conditions.

Although several islands have resulted from the phage-mediated transfer of virulence loci into tRNA genes (Cheetham and Katz, 1995), pathogenicity islands can arise by any mechanism that promotes horizontal gene transfer. In *Yersinia pestis*, incorporation of a 102 kb island encoding determinants required for growth in iron-deprived environments were probably mediated by an insertion sequence (Fetherston et al., 1992). This island is flanked by direct repeats of IS100, an element that is present in some 30 copies per genome and can integrate plasmids into the *Y. pestis* chromosome. Recombination between the directly-repeated IS100 is apparently responsible for the spontaneous loss of the *Yersinia* island, which occurs at a frequency of 10^{-5} .

Pathogenicity Islands and the Evolution of Virulence

Is the acquisition of a pathogenicity island sufficient to transform an organism into a pathogen? Three factors determine the virulence role of pathogenicity islands: the genes within the island, the status of the recipient microorganism, and features of the host that promote the progression of disease. The incorporation of certain genes, such as those encoding cholera toxin, would be sufficient to convert any organism into a pathogen because administration of the toxin itself is sufficient to produce the symptoms of the disease. But for most other cases, the utility of sequences obtained through gene transfer varies with the organism. *E. coli* is a benign constituent of the mammalian intestinal flora and is apparently predisposed to become a pathogen by the introduction of several types of virulence genes. The LEE island is likely to encode all of the genes necessary to produce attachment and effacement lesions, and the acquisition of this island could probably render any strain of *E. coli* pathogenic. And, not surprisingly, pathogenicity islands causing these lesions have been detected in genetically diversified strains of *E. coli*.

The acquisition of pathogenicity islands offers a rapid method of evolving novel functions; however, the introgressed sequences, even when they encode their specific regulators as part of complete functional units, must interact with rest of the genome. Expression of six invasion genes within the SPI-1 island of *Salmonella* is governed by the PhoP/PhoQ regulatory system (Galán, 1996), which is not encoded within the SPI-1 island and is present in both pathogenic and non-pathogenic microorganisms. And even sequences that are maintained

on extrachromosomal elements can be regulated by chromosomal loci: the invasion genes on the *Shigella* virulence plasmid are controlled by the chromosomally-encoded histone-like protein H1.

The association of pathogenicity islands with mobile DNA elements suggests that these sequences could be incorporated into a wide variety of bacterial species. However, presence of a pathogenicity island does not assure the transformation of an organism into a pathogen. A 10 kb region that harbors most known virulence genes in the pathogenic species of *Listeria* (*L. monocytogenes* and *L. ivanovii*) is also present in the chromosome of the nonpathogenic *L. seeligeri* (Gouin et al., 1994). While this gene cluster also encodes a pleiotropic activator, the attenuated phenotype of *L. seeligeri* has been ascribed to down regulation of most virulence genes. Furthermore, due to the coevolution between microbes and their hosts, the virulence potential of pathogenicity islands will only manifest in suitable animals and plants.

Finally, it has been suggested that the deletion of pathogenicity islands may represent a regulatory mechanism to control expression of virulence genes (Blum et al., 1994). This presupposes that the elimination of pathogenicity islands will circumvent the immune response to the encoded determinants and allow rapid microbial replication by reducing chromosome size. While this is a plausible scenario, both pathogenic and non-pathogenic strains of *E. coli* contain surface antigens that can promote an immune response, and the growth rates of natural strains of *E. coli* under lab conditions are not associated with genome size. Therefore, the deletion of pathogenicity islands is probably not a mechanism employed by bacteria to modulate virulence, but simply reflects the intrinsic genetic instability of these sequences.

Selected Reading

- Blum, G., Falbo, V., Caprioli, A., and Hacker, J. (1995). *FEMS Microbiol. Lett.* 126, 189–196.
- Blum, G., Ott, M., Lischewski, A., Ritter, A., Imrich, H., Tschäpe, H., and Hacker, J. (1994). *Infect. Immun.* 62, 606–614.
- Cheetham, B.F., and Katz, M.E. (1995). *Mol. Microbiol.* 18, 201–208.
- Falzano, L., Fiorentini, C., Donelli, G.E. M., Kocks, C.P. C., Cabanie, L., Oswald, E., and Boquet, P. (1993). *Mol. Microbiol.* 9, 1247–1254.
- Fetherston, J.D., Schuetz, P., and Perry, R.D. (1992). *Mol. Microbiol.* 6, 2693–2704.
- Galán, J.E. (1996). *Mol. Microbiol.* 20, 263–272.
- Gouin, E., Mengaud, J., and Cossart, P. (1994). *Infect. Immun.* 62, 3550–3553.
- Groisman, E.A., and Ochman, H. (1993). *EMBO J.* 12, 3779–3787.
- Groisman, E.A., and Ochman, H. (1994). *Trends Microbiol.* 2, 289–294.
- Groisman, E.A., Sturmoski, M.A., Solomon, F.R., Lin, R., and Ochman, H. (1993). *Proc. Natl. Acad. Sci. USA* 90, 1033–1037.
- Hacker, J., Bender, L., Ott, M., Wingender, J., Lund, B., Marre, R., and Goebel, W. (1990). *Microb. Pathogen.* 8, 213–225.
- Inouye, S., Sunshine, M.G., Six, E.W., and Inouye, M. (1991). *Science* 252, 969–971.
- Lee, C.A. (1996). *Infect. Agents Dis.* 5, 1–7.
- McDaniel, T.K., Jarvis, K.G., Donnenberg, M.S., and Kaper, J.B. (1995). *Proc. Natl. Acad. Sci. USA* 92, 1664–1668.

- Mills, D.M., Bajaj, V., and Lee, C.A. (1995). *Mol. Microbiol.* 15, 749–759.
- Ochman, H., Soncini, F.C., Solomon, F., and Groisman, E.A. (1996). *Proc. Natl. Acad. Sci. USA* 93, 7800–7804.
- Shea, J.E., Hensel, M., Gleeson, C., and Holden, D.W. (1996). *Proc. Natl. Acad. Sci. USA* 93, 2593–2597.
- Van Gijsegem, F., Genin, S., and Boucher, C. (1993). *Trends Microbiol.* 1, 175–180.
- Waldor, M.K., and Mekalanos, J.J. (1996). *Science* 272, 1910–1914.
- Zagaglia, C., Casalino, M., Colonna, B., Conti, C., Calconi, A., and Nicoletti, M. (1991). *Infect. Immun.* 59, 792–799.