Mapping of noninvasion TnphoA mutations on the Escherichia coli O18:K1:H7 chromosome

Craig A. Bloch [a],*, Sheng-He Huang [b], Christopher K. Rode [a], Kwang Sik Kim [b]

[a] Department of Pediatrics and Epidemiology, University of Michigan, Ann Arbor, MI 48109-0656, USA
[b] Department of Pediatrics, University of Southern California, Los Angeles, CA 90007, USA

Abstract

The most virulent newborn meningitis-associated *Escherichia coli* are of the serotype O18:K1:H7. We previously isolated a large number of *E. coli* O18:K1:H7 mutants resulting from transposon TnphoA mutagenesis that fail to invade brain microvascular endothelial cells. We have now determined the locations of 45 independent insertions. Twelve were localized to the 98 min region, containing a 120 kb segment that is characteristic of *E. coli* O18:K1:H7. Another, the previously described insertion *ibe*-10::TnphoA, was localized to the 87 min region, containing a 20 kb segment found in this *E. coli*. These noninvasion mutations may define new O18:K1:H7 pathogenicity islands carrying genes for penetration of the blood-brain barrier of newborn mammals.

Keywords: Bacterial genomics; *Escherichia coli* meningitis; I-SceI restriction endonuclease; Pathogenicity island; Pulsed-field gel electrophoresis; Rare-restriction-site mapping

1. Introduction

Escherichia coli serovar O18:K1:H7 predominates among colonic *E. coli* in ~3% of human hosts but is responsible for ~30% of instances of *E. coli* meningitis in newborns (cf. [1,2]). Moreover, epidemiologic studies suggest that nearly all premature infants becoming colonized may also develop disease (cf. [1,2]). Thus, greater understanding of the genomic determinants distinguishing *E. coli* O18:K1:H7 strains is needed to allow improved surveillance and treatment.

Screening of mutations that affect bacterial surface proteins affords an avenue for identification of new virulence genes: such changes may alter the surface properties of pathogenic bacteria and, in turn, the interactions between them and mammalian-host cells. Thus, the transposon TnphoA, which is designed to generate secreted alkaline-*pho*sphatase fusion proteins that can be readily distinguished by their blue colonies (PhoA+) formed on 5-bromo-4-chloro-3-indolyl phosphate indicator medium, has allowed identification of new virulence genes from a variety of species, including *Salmonella enterica*, *Vibrio cholera*, *Bordetella pertussis*, *Yersinia pestis*, and *E. coli* [3]. Indeed, in the *E. coli* O18:K1:H7 prototype strain RS218 [4], screening PhoA+ TnphoA mutants for invasion of brain microvascular endothelial cells (BMEC) has recently allowed isola-

* Corresponding author. Tel.: +1 (313) 763-2440;
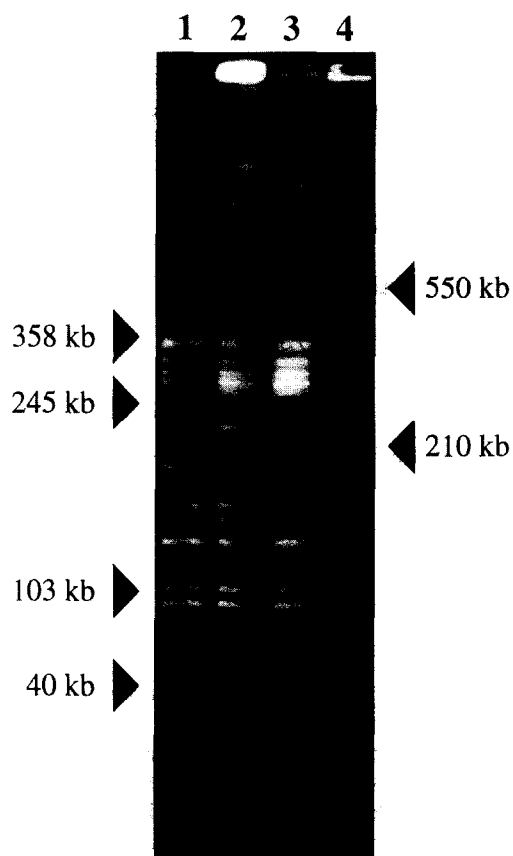Fax: +1 (313) 764-6837; E-mail: CBloch@Umich.edu

1 2 3 4

550 kb

358 kb ▶

245 kb ▶

210 kb

103 kb ▶

40 kb ▶

Fig. 1. NotI restriction mapping of RS218::TnphoA insertions. Lanes 1–3: strains RS218, HK1021 (genotype; RS218 ibe-211::TnphoA), χM3003 (genotype; HK1021 zjh-226::Tn10dSpcRCP2), respectively, digested with Not I. Lane 4: yeast chromosomes. White bars indicate the native NotI fragment B missing due to TnphoA and/or Tn10dRCP2 insertion(s), while black bars indicate the corresponding new subfragments generated. PFGE pulse times were ramped from 11 to 16 s for 12 h and from 50 to 55 s over 9 h.

tion of the mutant RS218ibe-10::TnphoA, which encodes a truncated peptide interrupting both invasion of brain endothelium in vitro and penetration of the blood-brain barrier in vivo [5]. Determination of the sequence of the disrupted gene suggested potential epidemiologic utility as well as a possible new pathogenesis motif, as it appears unrelated to genomic DNA sequenced from a variety of other pathogenic and nonpathogenic bacteria.

In the same study, 49 other independent RS218 ibe::TnphoA mutants were isolated. However, the available data provide few additional characteristics

distinguishing the mutants in this collection. Physical mapping of these would be expected to allow inference of separate ibe loci as well as of any clusters of ibe::TnphoA insertions. Such data would also be expected to reveal which insertions occur within RS218-specific chromosomal segments [6,7] and, thereby, may define RS218 pathogenicity islands.

The transposon TnphoA carries, along with its design features that are well suited for genetic dissection of virulence, a NotI site from the native sequences of its parent transposon Tn5 [8]. This NotI site and recent availability of a NotI restriction map for strain RS218 [6] should allow rapid physical mapping of RS218::TnphoA insertions (cf. [9]). Herein, NotI digestion and pulsed-field gel electrophoresis (PFGE) of genomic DNAs from RS218::TnphoA mutants were used to assign physical-map locations to many of the above-mentioned ibe::TnphoA insertions.

2. Materials and methods

2.1. Bacterial genetics techniques

E. coli strains used in this study were grown in Luria-Bertani (LB) medium [10] with aeration or on solid LB medium. Media were supplemented with kanamycin (Kan; 25 μg/ml), spectinomycin (Spc; 100 μg/ml), and/or chloramphenicol (15 μg/ml) as required. Cultures were incubated at 37°C, or at 32°C for P1 transduction of strain RS218 [11] and derivatives. Cells were stored long term by suspending in LB/glycerol (85:15%, v/v), and cooling to −80°C. Bacteriophage P1 stocks were propagated and used to transduce recipient strains as described by Sternberg and Maurer [12]. Strain MG1655 [13] served as laboratory E. coli K-12 prototype; all Tn10dRCP2 insertions (carrying the Rare Cutting Polylinker 2) used in this study (car-101, zah-107, zba-111, zbi-121, zcg-128, zda-135, zed-154, hisA158, zfa-163, zfi-172, zib-203, zib-204, ilvG211, and zjh-226) were generated in this K-12 strain [14]. Introduction of NotI sites and I-SceI sites into the RS218 background (and of second-site I-SceI sites into the MG1655 background) was carried out by P1 transduction using lysates of these Tn10dRCP2 insertion mutants.
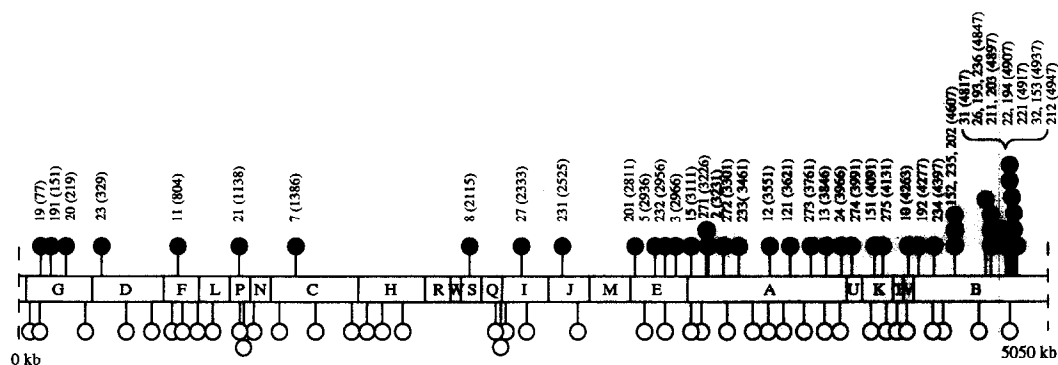
Fig. 2. Locations of PhoA⁺ RS218::Tn*phoA* insertions. A linearly represented *Not*I restriction map, viewed with the circular chromosome of strain RS218 [5] opened at coordinate 0/5050 kb (25 kb counterclockwise of the native *Not*I site nearest the *carAB* locus [5]), is shown with markers at the sites of Tn*phoA* insertions. RS218 *ibe*::Tn*phoA* insertions are indicated by filled markers projecting upward from the map, with the allele number and physical map coordinate (kb) [5] of each mutation provided. Two of these insertions, one previously described (*ibe*-10::Tn*phoA*) and one described in detail herein (see Section 3) (*ibe*-211::Tn*phoA*), are distinguished by solid fill. Tn*phoA* mutations failing to abolish BMEC monolayer invasion are indicated by unfilled markers projecting downward from the map.

2.2. Genomic DNA biophysical techniques

Genomic DNAs were purified and subjected to restriction digestion and electrophoretic separation as described [6,14]. Briefly, after digestion of agarose dot-embedded DNAs with I-*Sce*I (Boehringer-Mannheim; Indianapolis, IN) for 1 h, or *Not*I (New England Biolabs, Beverly, MA) for 4–5 h, dots were melted (70°C) and gently pipetted with plastic 200 μl tips into sample wells in 1.2% agarose (PFGE-approved; FastLane; FMC, Portland, ME) gels for electrophoresis in 0.5×TBE buffer (0.045 M Tris borate/0.045 M boric acid/0.001 M EDTA) in a contour-clamped homogeneous electric field apparatus (DR-III; Bio-Rad, Hercules, CA) at a field strength of 6 V/cm. Pulse-ramping parameters were determined as described elsewhere [15]. Gels were analyzed as described [6].

Mapping of noninvasion Tn*phoA* mutations was carried out by *Not*I digestion of genomic DNAs from the corresponding *ibe*::Tn*phoA* mutants. Double-insertion mutants (constructed by transferring one of a series of previously mapped *Not*I-site-carrying (Tn*10*dRCP2) insertions [14] into each *ibe*::Tn*phoA* mutant) were used to determine which of the two possible locations delimited by the *Not*I pattern from each original *ibe*::Tn*phoA* mutant was correct. Tn*phoA* mutations failing to abolish invasion, in contrast to the noninvasion mutations, were loca-

lized only to the two possible locations delimited by the *Not*I pattern from the original mutant; one of these was randomly selected for display purposes.

Purification of artificial I-*Sce*I fragments was carried out as described [14] by introduction of two Tn*10*dRCP2 insertions (each carrying a different antibiotic resistance) into the same strain background. (Because each insertion carries a unique I-*Sce*I site, which is not found in native *E. coli* sequences [14], pairs of them could be used to cleave the chromosome at unique pairs of sites.)

3. Results and discussion

Genomic DNAs from 50 *ibe*::Tn*phoA* insertion mutants, identified by their inability to *i*nvade *b*rain microvascular *e*ndothelial-cell monolayers (from a total of 704 PhoA⁺ RS218::Tn*phoA* mutants [5]), were subjected to *Not*I digestion and PFGE to determine the locations of their Tn*phoA* insertions. As an example, the changes introduced by one of the insertions (*ibe*-211::Tn*phoA*) can be seen in Fig. 1: The *Not*I fragment B (785 kb) in strain RS218 (lane 1) was missing from the corresponding digest of the RS218 mutant containing the *ibe*-211::Tn*phoA* insertion (lane 2); instead, this digest yielded subfragments of 635 and 150 kb. These data limited the *ibe*-211::Tn*phoA* insertion to one of two positions
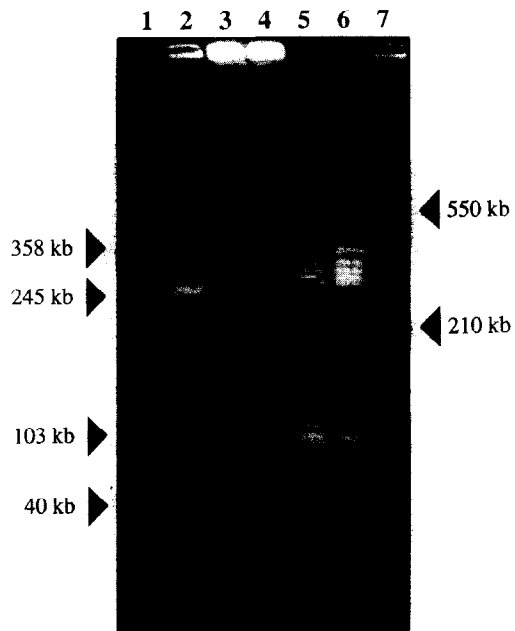
Fig. 3. Purification of a pair of corresponding I-SceI fragments from strains RS218 and K-12 harboring, in the former, the site of a cluster of 12 ibe::TnphoA insertions. Lanes 1,2,5,6: strains MG1655, χM3001 (genotype; MG1655zjh-226::Tn10dSpcRCP2, car-101::Tn10dKanRCP2), χM3002 (genotype; RS218zjh-226::Tn10dSpcRCP2, car-101::Tn10dKanRCP2), and RS218 digested with NotI, respectively. Lanes 3,4: strains χM3001 and χM3002 digested with I-SceI, respectively. White bars indicate the native NotI bands missing, from both the RS218 and K-12 strain backgrounds, due to incorporation of the identical pair of Tn10dRCP2 alleles [12], while black bars indicate the corresponding subfragments generated. PFGE pulse times were ramped from 11 to 16 s over 12 h and from 50 to 55 s over an additional 9 h.

within fragment B (at either ~4441 or ~4926 kb on the RS218 chromosomal NotI map [6]). The correct clockwise/counterclockwise orientation of the two subfragments from fragment B was then distinguished by introduction (by P1 transduction) of the NotI-site-carrying zjh-226::Tn10dSpcRCP2 allele at ~4796 kb, into the same strain. This indicated that (i) the zjh-226::Tn10dSpcRCP2 insertion was contained within the 635 kb subfragment (lane 3), (ii) the 150 kb subfragment was oriented clockwise of ibe-211::TnphoA, and (iii) the ibe-211::TnphoA insertion was located at ~4926 kb. Analogous procedures were repeated for each of the 50 RS218 ibe::TnphoA mutants and revealed chromosomal locations for the insertions carried by 45 mutants, in-

cluding the ibe-10::TnphoA mutant [5]. Three others carried TnphoA insertions within the single 110 kb plasmid [6] of strain RS218; two had no detectable insertion. The locations of 45 RS218 ibe::TnphoA insertions are depicted in Fig. 2 as filled circles projecting above the RS218 chromosomal physical map.

A lack of random distribution of the ibe::TnphoA insertions on the RS218 physical map could signify hotspots for the genes of surface proteins required in blood-brain barrier invasion. Alternatively, the same finding could result (i) from the chromosomal copy-number inequalities introduced during exponential growth or (ii) from nonrandom distribution of the targets for TnphoA insertion or of the genes for secreted proteins. To distinguish between these possibilities, 50 RS218::TnphoA mutants were randomly chosen from the remaining group of 654 that were PhoA$^+$ and yet retained invasiveness on BMEC monolayers. Mapping revealed chromosomal locations for 42 of their insertions. (Six contained plasmid insertions; two no detectable insertions.) These locations are depicted in Fig. 2 as open circles projecting below the RS218 map.

Several features distinguishing the ibe::TnphoA insertions were apparent. First, a cluster at 98 min (~4930 kb), formed by 12 insertions (27%) that were located within a contiguous ~140 kb region (<3%) of the RS218 chromosome (~5050 kb in total size [6]), was present in this group. Interestingly, this cluster was also located at a 120 kb RS218-specific chromosomal segment, which could be purified by PFGE within an artificial, unique I-SceI fragment (435 kb; see Fig. 3). This 120 kb, demonstrated by the size difference between I-SceI restriction fragments generated by the identical pair of Tn10dRCP2 insertions in the K-12 and RS218 backgrounds (Fig. 3; lanes 3 and 4, respectively), had been previously identified as one of 10 > 20 kb RS218-specific chromosomal segments totaling ~600 kb (~12% of the RS218 physical map) [6]. Its location suggested the possibility that it might be inserted at the same site as the pathogenicity island II (at leuX; 98 min) of uropathogenic E. coli strain 536 [16]; however, this was not clear from the available data. An unusual feature of the chromosomal region containing this segment was the number of distinct ibe::TnphoA insertions within it. This suggested a substantial grouping of genes encoding var-

ious secreted and/or surface associated proteins required for invasion of the blood-brain barrier.

Second, the *ibe*-10::Tn*phoA* insertion, whose phenotype and adjacent sequences we previously described [5], was also located at one of the 10 > 20 kb RS218-specific chromosomal segments – that at 87 min containing 20 kb of RS218 DNA [6]. The locations of (i) this insertion, (ii) the above-mentioned cluster of 12 insertions, and (iii) the recognized RS218 virulence gene clusters *kpsA* [17], *rfb* [6], and *sfa* (results not shown) at these sites suggest that the > 20 kb RS218-specific segments may contain a variety of acknowledged as well as yet-uncharacterized O18:K1:H7 virulence genes.

Third, the 45 *ibe*::Tn*phoA* insertions tended to be located in the left half of the RS218 chromosome – that is, in the region surrounding the *oriC* locus, from *rrnG* (57 min) clockwise to *rrnH* (5 min). Indeed, only seven were located outside of this interval. By contrast, 21 insertions in the invasive mutants were located outside the same interval. It is interesting that the region that was most devoid of *ibe*::Tn*phoA* insertions, near the putative terminus of *E. coli* RS218, corresponds to the so-called 'garbage can' region of *E. coli* K-12 which contains an overrepresentation of sequences that either are unexpressed or are unshared with other *E. coli* [18].

We conclude that targets for *ibe*::Tn*phoA* insertions, in contrast to those for PhoA$^+$ Tn*phoA* insertions in general, are nonrandomly distributed on the *E. coli* RS218 chromosome. Moreover, physical mapping of these insertions provides a useful guide for their further investigation: The *ibe*-10 insertion and the 12 clustered insertions (at 87 and 98 min, respectively) are of immediate interest as, together with comparative *Not*I and I-*Sce*I data, they may define two new pathogenicity islands carrying genes required for invasion of the blood-brain barrier of newborn mammals.

Acknowledgments

References

[1] Achtman, M., Mercer, A., Kusecek, B., Pohl, A., Heuzenroeder, M., Aaronson, W., Sutton, A. and Silver, R. (1983) Six widespread bacterial clones among *Escherichia coli* K1 isolates. Infect. Immun. 39, 315–335.

[2] Sarff, L., McCracken, G., Jr., Schiffer, M., Glode, M., Robbins, J., Ørskov, I. and Ørskov, F. (1975) Epidemiology of *Escherichia coli* K1 in healthy and diseased newborns. Lancet 1, 1099–1104.

[3] Kaufman, M. and Taylor, R. (1994) Identification of bacterial cell-surface virulence determinants with Tn*phoA*. Methods Enzymol. 235, 426–48.

[4] Silver, R.P., Aaronson, W., Sutton, A. and Schneerson, R. (1980) Comparative analysis of plasmids and some metabolic characteristics of *Escherichia coli* K1 from diseased and healthy individuals. Infect. Immun. 29, 200–206.

[5] Huang, S.H., Wass, C., Fu, Q., Prasadarao, N.V., Stins, M. and Kim, K.S. (1995) *Escherichia coli* invasion of brain microvascular endothelial cells in vitro and in vivo: molecular cloning and characterization of invasion gene *ibe10*. Infect. Immun. 63, 4470–4475.

[6] Rode, C.K., Obreque, V.H. and Bloch, C.A. (1995) New tools for integrated genetic and physical analyses of the *Escherichia coli* chromosome. Gene 166, 1–9.

[7] Bloch, C.A., Rode, C.K., Obreque, V.H. and Russell, K.Y. (1994) Comparative genome mapping with mobile physical map landmarks. J. Bacteriol. 176, 7121–7125.

[8] Smith, C.L. and Kolodner, R.D. (1988) Mapping of *Escherichia coli* chromosomal Tn*5* and F insertions by pulsed field gel electrophoresis. Genetics 119, 227–236.

[9] Okada, N., Sasakawa, C., Tobe, T., Talukder, K.A., Komatsu, K. and Yoshikawa, M. (1991) Construction of a physical map of the chromosome of *Shigella flexneri* 2a and the direct assignment of nine virulence-associated loci identified by Tn*5* insertions. Mol. Microbiol. 5, 2171–2180.

[10] Miller, J.H. (1992) A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for *Escherichia coli* and Related Bacteria, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

[11] Bloch, C.A., Thorne, G.M. and Ausubel, F.M. (1989) General method for site-directed mutagenesis in *Escherichia coli* O18ac:K1:H7: Deletion of the inducible superoxide dismutase gene, *sodA*, does not diminish bacteremia in neonatal rats. Infect. Immun. 57, 2141–2148.

[12] Sternberg, N.L. and Maurer, R. (1991) Bacteriophage-mediated generalized transduction in *Escherichia coli* and *Salmonella typhimurium*. Methods Enzymol. 204, 18–43.

[13] Bachmann, B.J. (1987) Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12. In: *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology (Neidhardt, F.C. Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M. and Umbarger, H.E., Eds.), pp. 1190–1219, American Society for Microbiolgy, Washington, DC.

[14] Bloch, C.A., Rode, C.K., Obreque, V.H. and Mahillon, J. (1996) Purification of *Escherichia coli* chromosomal segments

without cloning. Biochem. Biophys. Res. Commun. 223, 104–111.

[15] Birren, B. and Lai, E. (1993) Pulsed Field Gel Electrophoresis: A Practical Guide, Academic Press, San Diego, CA.

[16] Blum, G., Ott, M., Lischewski, A., Ritter, A., Imrich, H., Tschape, H. and Hacker, J. (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.

[17] Bloch, C.A. and Rode, C.K. (1996) Pathogenicity island evaluation in Escherichia coli K1 by crossing with strain K-12. Infect. Immun. 64, 3218–3223.

[18] Moir, P.D., Spiegelberg, R., Oliver, I.R., Pringle, J.H. and Masters, M. (1992) Proteins encoded by the Escherichia coli replication terminus region. J. Bacteriol. 174, 2102–2110.