Pulsed-field gel electrophoresis genomic fingerprinting of hospital Escherichia coli bacteraemia isolates


Department of Pediatrics, Division of Infectious Diseases and *Clinical Microbiology, University of Michigan, Ann Arbor, Michigan 48109-0244, USA

Pulsed-field gel electrophoresis (PFGE), because of the increased sensitivity it affords over other methods of bacterial genotyping, represents a potentially powerful tool for the characterisation of isolates from hospital infections. Genomic fingerprinting by PFGE was applied to all clinical isolates of Escherichia coli obtained from blood during a 6-month period (78 isolates, 58 patients) at the University of Michigan Medical Center. The rare-restriction patterns of these isolates, in contrast to those of isolates from the E. coli reference collection (ECOR), were not randomly distributed through the E. coli species. Four related clusters, which represented c. 21% of the blood isolates, were identified. Two of these genotypic clusters were also clustered temporally, their members all being isolated within the same 2-week period, while the other two clusters spanned the study period. These observations indicate in-hospital endemic vectors or the occurrence of specialised E. coli lineages that are capable of invading the bloodstream and exploiting in-hospital vectors, or both.

Introduction

Escherichia coli is a common commensal of the human gastrointestinal tract. However, under certain conditions strains of E. coli can cause disease, either local or systemic. Strains causing gastrointestinal disease have been categorised as enterotoxigenic, enteroinvasive or enteropathogenic; additional strains have been identified as pyelonephritogenic and neonatal sepsis-associated [1]. Although considerable understanding of some of the virulence determinants associated with recognised subspecific lineages (i.e., colonisation factors, heat-stable and heat-labile enterotoxins, anti-phagocytic capsules and adherence factors) has been accumulated, very little is known about the distribution of E. coli lineages and their virulence determinants within natural and disease-associated populations.

Several techniques have been employed to correlate genetic variation with pathogenicity for E. coli [2, 3]. As multi- locus enzyme electrophoretotyping (MLEE) allows correlation of simple biochemical phenotypes with individual allelic substitutions, it has been a valuable tool in characterising E. coli population genetics [4, 5]. Indeed, studies employing MLEE have indicated not only the occurrence of temporally and geographically stable E. coli lineages in nature, but also the tendency for a subset of these natural lineages to be over-represented among E. coli disease isolates. However, the relative stability of the allelic differences detected by MLEE also precludes use of this method for the tracking of more mutable features of bacterial genomes that might change, for instance, in the course of a single hospital epidemic.

Pulsed-field gel electrophoresis (PFGE) of genomic DNA, following digestion with rare-cutting restriction endonucleases, has allowed highly discriminant genotyping of subspecies varieties of E. coli, Enterococcus faecalis and Klebsiella pneumoniae, among other clinically relevant bacteria [6–8]. However, the pace of the rapidly accumulating genotypic changes, even in the course of spread within a single human host, detected by PFGE has suggested that this technique may be limited in the tracking of bacterial pathogens to temporal and geographic (point-source) outbreaks [9]. To address the potential of this method for characterising hospital-based (yet otherwise temporally and geographically dispersed) infection patterns, all the E. coli blood isolates obtained during a 6-month period at the University of Michigan Medical Center, were analysed by rare-restriction PFGE genotyping.
Materials and methods

Clinical isolates obtained from the clinical microbiology laboratory at the University of Michigan Medical Center between 1 July and 31 Dec. 1994 were stored at −80°C in LB/glycerol 20%. All strains were subsequently grown in Luria-Bertani media with aeration at 37°C.

Minimally sheared DNA from each strain, suitable for restriction with rare-cuts endonucleases, was prepared as previously described [10]. DNAs embedded in agarose were digested with NotI (NEB; Beverly, MA, USA) endonuclease according to the manufacturer’s specifications. Agarose-embedded minimally sheared DNA resistant to digestion with NotI was digested with BlnI (Panvera; Madison, WI, USA).

PFGE was performed on a CHEF (DRIII; BioRad; Hercules, CA, USA) apparatus in FastLane 1.2% agarose gels (FMC; Portland, ME, USA) with a field strength of 6 V/cm and switch times ramped from 18 to 21 s for 11 h and from 45 to 50 s for 10 h. The electrophoresis buffer was 0.5× TBE buffer.

Whereas identical strains were defined by complete identity between bands in pairwise comparisons, pairs of isolates having 75% of their bands in common were defined as related. Related strains were defined as having 75% of the isolate’s bands in common with another isolate’s pattern, in contrast to the proposed guidelines of Tenover et al. [9]. It was felt that the proposed guidelines were too stringent for this study. Use of this degree of genomic structural similarity to indicate genetic relatedness was supported by the following: (i) some pairs of strains isolated from identical sources (and believed to be of common origin) had only 85−90% of their bands in common; (ii) the same criteria applied to the ECOR collection and to a series of E. coli UTI isolates identified, as expected, no related strains; and (iii) the groupings that were formed by pairwise comparisons accepting as the definition of relatedness only 50% band identities, were identical to those formed with the 75% definition.

Results

All blood isolates of E. coli were obtained from the clinical microbiology laboratory during a 6-month period from 1 July to 31 Dec. 1994. Seventy-seven E. coli isolates were obtained from 58 different patient admissions. One patient had two separate admissions for E. coli bacteraemia and was counted twice. Any patient having multiple positive blood cultures during a single admission (16 patients, 37 isolates) was counted as a single instance of E. coli bacteraemia. However, each E. coli isolate (77) was analysed individually by PFGE.

Of the 16 patients with more than one positive blood culture, 12 had identical E. coli isolates from all positive cultures as determined by PFGE (30 isolates). In two of the remaining four patients both isolates were related. The third patient had three positive cultures, two of which were identical and the third of which was unrelated. In the fourth patient it was not possible to analyse one of the two isolates because of what appeared to be endogenous nucleases. These two strains were determined to be unrelated by API biotyping (data not shown).

Inspection of the E. coli genotypes from the 58 different instances of bacteraemia showed that 12 of these formed four related patterns by PFGE (Table 1). Three of these, consisting of one cluster of four strains and two strain pairs (Fig. 1), were defined by related NotI digestion patterns. Seven isolates were resistant to digestion with the NotI restriction endonuclease. Further analysis of these seven isolates demonstrated that four were related by BlnI digestion (Fig. 2). The remaining 45 isolates were unrelated to each other or to any member of the clusters. Thus 20.6% of all isolates in the study belonged to one of the four clusters.

To determine the significance of this genomic structure clustering MLEE data on the ECOR collection were compared with PFGE. Previous reports, examining 35 MLEE loci from the 72-strain ECOR collection (from 66 individuals), found seven clusters (17 isolates) accounting for 24% of the strains [4]. The ECOR collection was examined by PFGE and this technique was found to be more sensitive. Of the 66 isolates from different sources only one pair of isolates was found to be related by NotI restriction and PFGE. In a similar study, of 20 E. coli UTI isolates analysed by NotI restriction and PFGE, none was found to be related [11].

In an attempt to understand the significance of the clustering observed in the hospital’s E. coli bacteraemia isolates both the geographic (i.e., the patient’s

<table>
<thead>
<tr>
<th>Isolate no.</th>
<th>Ward*</th>
<th>Enzyme</th>
<th>Interval between first and last isolation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>6B</td>
<td>NotI</td>
<td>3 months</td>
</tr>
<tr>
<td>2006</td>
<td>NICU</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>6C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2043</td>
<td>NICU</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>8A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>7B</td>
<td>BlnI</td>
<td>5 months</td>
</tr>
<tr>
<td>2025</td>
<td>6B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2070</td>
<td>PCTU</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2059</td>
<td>7B</td>
<td>NotI</td>
<td>&lt; 2 weeks</td>
</tr>
<tr>
<td>2055</td>
<td>OBG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2063</td>
<td>BMU</td>
<td>NotI</td>
<td>&lt; 2 weeks</td>
</tr>
<tr>
<td>2068</td>
<td>BMU</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*NICU, neonatal intensive care unit; PCTU, paediatric cardiothoracic unit; BMU, bone marrow unit; OBG, obstetrics and gynaecology.
Fig. 1. Macrorestriction and PFGE patterns of E. coli blood isolates. Total genomic DNA, embedded in agarose was prepared and digested enzymically as described in Materials and methods. Electrophoresis parameters were as follows: agarose 1.2% gel in 0.5X TBE, field strength of 6 V/cm, re-orientation angle of 120°, switched times ramped from 16 to 19 s over 12 h and 50 to 55 s over 10 h, at 14°C. Lane 1, isolate 2000; 2, 2008; 3, 2025; 4, 2070; 5, 2003; 6, 2006; 7, 2007; 8, 2043; 9, 2055; 10, 2059; 11, 2063; 12, 2068.

Discussion

Different strains of E. coli carry strain-specific genes predisposing them to different types of human diseases [9, 12]. For instance, neonatal sepsis-associated strains carry genes for K1 capsule, allowing evasion of neonatal host humoral immunity [13], while pyelonephritogenic strains carry genes for pyelonephritis-associated pili (PAP), allowing adherence to urinary tract epithelial cells [14]. Although point-source outbreaks comprise one potential cause of relatedness among hospital bloodstream isolates, it is also possible that such relatedness is permitted by a subset of E. coli strains that, by virtue of specialised traits, are capable of bloodstream invasion or of exploiting the hospital environment, or both. Whereas nosocomial spread per se might yield transient temporal clustering of many different types of strains, preferential spread of a select subset of strains might be expected to result in consistent over-representation of that subset in E. coli blood isolates from hospitalised patients.

Four clonally related genotypes of E. coli bloodstream isolates were detected by PFGE fingerprinting, accounting for 20.6% of all hospital isolates. When compared with studies either of UTI strains from the same locale or of the ECOR collection, a much higher rate of PFGE genotypic relatedness was observed among hospital bloodstream isolates than expected. There are two possible explanations for the increase observed: abnormally slow accumulation of PFGE changes among hospital bloodstream isolates or abnormally rapid (point source-like) dissemination of hospital bloodstream isolates, or both. Although the temporal relationship between some of the clusters is suggestive of relatively frequent in-hospital point-source outbreaks, this could also represent spread within the referral community in conjunction with relatively stable PFGE pattern. On the other hand, the four bacteraemias associated with one of the genotypic
clusters were temporally as well as geographically dispersed, and thus could not be explained by a single in-hospital point-source outbreak. This cluster suggests that these results cannot be entirely accounted for simply by the possibly high frequency of point-source outbreaks within the hospital. Moreover, it suggests the ineffectiveness of both temporal and physical layout barriers within the hospital complex. The nature of the clustering observed by PFGE is still unclear and may be multi-factorial.

Typing of bacterial isolates by antibiotic resistances, serotyping [1], phage typing [15], MLEE [6], PFGE [6, 16] and other methods has been used in the epidemiology of bacterial diseases. While both PFGE and MLEE may be used to survey various loci spaced more or less evenly about the chromosome, only MLEE analysis is designed to detect the preservation of ‘clonal frames’ (i.e., common origins of overall chromosomal background) over substantial temporal and geographic distances [4]. However, the particular features detected by MLEE – alleles distinguished by base-pair substitutions in conserved genes – cannot allow tracking of genotypic variations occurring over shorter spacial and temporal distances that may be useful for tracking hospital epidemics. By contrast, PFGE allows detection of many different types of sequence changes that would be invisible by MLEE.

For this reason PFGE analysis may also have failed to detect potential relatedness between isolates in this study that would have been apparent by MLEE (as, in fact, demonstrated by PFGE analyses of the ECOR collection isolates that are related by MLEE). Indeed, by the criteria employed in the present study, many clonal isolates would be scored as unrelated by PFGE after as few as four genome structure-altering events. Even between two isolates from this study (strains 2002 and 2009), which were cultured from the blood of a single patient within a 2-week period, PFGE distinctions were observed (loss of a single NotI band and appearance of two new NotI bands in one of them). It is worth noting that these distinctions were unlikely to have occurred by random base-pair substitution; this is because less than one in $10^4$ substitutions in random sequence would be expected to abolish or create a NotI site.

Base-pair substitutions leading to restriction site divergence between isolates is, of course, not the only possible cause of macrorestriction pattern degeneracy. Events other than base-pair substitution can lead to changes in macrorestriction patterns. Insertion sequence migration, by IS30 for instance (associated with the transposable element Tn5 and known to contain a NotI site [17]), could account for changes. Furthermore, insertions or deletions of pathogenicity
islands in uropathogenic strains of *E. coli* [18] or prophage (lambda phage) [9], gain or loss of plasmids (unpublished observations), or re-arrangements, could all account for macrorestriction pattern changes. Importantly, some for these events could be associated with adaptations contributing to infection.

Although the present study clearly shows that some of the clonal relationships that are traditionally detectable among hospital isolates by MLEE are likewise detectable by PFGE, it also points out a distinct advantage in the use of PFGE versus MLEE for short-term hospital epidemiology. While MLEE tracks all account for macrorestriction pattern changes. (unpublished observations), or re-arrangements, could advantage in the use of PFGE detectable by PFGE, it also points out a distinct clonal relationships that are traditionally detectable among hospital isolates by MLEE are likewise the clonal relationships that are traditionally detectable by PFGE. Conversely, PFGE is generally unlikely to reveal short-term genotypic alterations from re-arrangements, additions or deletions, and neutral or selected base-pair substitutions that can be monitored by PFGE. Conversely, PFGE is generally unlikely to detect genotypic relatedness between strains in epidemiological surveys occurring over extended geographic or temporal distances.

**References**


