Evidence of persistence of *Prevotella* spp. in the cystic fibrosis lung


**Abstract**

**Purpose.** *Prevotella* spp. represent a diverse genus of bacteria, frequently identified by both culture and molecular methods in the lungs of patients with chronic respiratory infection. However, their role in the pathogenesis of chronic lung infection is unclear; therefore, a more complete understanding of their molecular epidemiology is required.

**Methodology.** Pulsed Field Gel Electrophoresis (PFGE) and Random Amplified Polymorphic DNA (RAPD) assays were developed and used to determine the degree of similarity between sequential isolates (*n*=42) from cystic fibrosis (CF) patients during periods of clinical stability and exacerbation.

**Results.** A wide diversity of PFGE and RAPD banding patterns were observed, demonstrating considerable within-genus heterogeneity. In 8/12 (66.7 %) cases, where the same species was identified at sequential time points, pre- and post-antibiotic treatment of an exacerbation, PFGE/RAPD profiles were highly similar or identical. Congruence was observed between PFGE and RAPD (adjusted Rand coefficient, 0.200; adjusted Wallace RAPD-PFGE 0.459, PFGE-RAPD 0.128). Furthermore, some isolates could not be adequately assigned a species name on the basis of 16S rRNA analysis: these isolates had identical PFGE/RAPD profiles to *Prevotella histicola*.

**Conclusion.** The similarity in PFGE and RAPD banding patterns observed in sequential CF *Prevotella* isolates may be indicative of the persistence of this genus in the CF lung. Further work is required to determine the clinical significance of this finding, and to more accurately distinguish differences in pathogenicity between species.

**INTRODUCTION**

Improved culture and culture-independent methodologies have resulted in an increased appreciation of the role of anaerobes, including members of the genus *Prevotella*, in a range of infections. *Prevotella* spp. comprise a heterogeneous group of approximately 40 species of Gram-negative, obligately anaerobic, non-spore forming, non-motile, pleomorphic rods [1]. Members of this genus are physiologically diverse and this is reflected in the capacity of *Prevotella* spp. to colonize a spectrum of ecological niches, ranging from soil to various body sites [2, 3]. Although well-established as oral pathogens, *Prevotella* spp. have also been identified as causal agents in several infection types, including bone and joint infections, and sinusitis [4, 5].

Recent work has focused on the potential role of *Prevotella* spp. in chronic lung disease. Following the identification of anaerobic niches in the lungs of patients with cystic fibrosis (CF) [6], *Prevotella* spp. have been consistently identified in high numbers by culture and/or molecular methods in several chronic lung diseases, including CF [7, 8], chronic obstructive pulmonary disease (COPD) [9–11], asthma [12, 13] and bronchiectasis [14]. However, the role of *Prevotella* in the pathogenesis of chronic lung disease remains unclear and has yet to be established in full detail. *Prevotella* spp. have been isolated in longitudinal samples taken from patients when clinically stable and during acute infective exacerbations. However, no studies have examined the genotypic similarity of these isolates to determine whether this is evidence of persistence. Indeed, metagenomic analyses of the respiratory microbiome, as described in the above studies, are not sufficiently discriminatory to distinguish between *Prevotella* at the species level. Given the heterogeneous nature of this group, it is conceivable that there are...
potentially significant differences in phenotype and virulence within the genus.

To better understand the contribution of *Prevotella* spp. to the pathogenesis of lung infection and inflammation in chronic respiratory diseases, it is important to understand the potential for *Prevotella* to persist within the lung and thereby provide further insight into their role as established lung colonizers. Therefore, in this study, we developed PFGE and random amplified polymorphic DNA (RAPD) typing schemes to determine the degree of similarity between *Prevotella* spp. isolated from patients with CF during periods of clinical stability and exacerbation.

**METHODS**

**Clinical samples**

*Prevotella* spp. were isolated as part of a multicentre study to define the natural history and pathogenicity of anaerobic bacteria in CF using enhanced culture techniques [Office for Research Ethics Committees Northern Ireland reference 10/NIR01/41; Integrated Research Approval System (IRAS) Project no. 41579]. Sputum was collected from CF patients (*n*=10) when either clinically stable (*n*=1) or during an exacerbation (*n*=9) (Table 1). Stable isolates were cultured initially (S1) and 3 months later, when no exacerbation was recorded (S2 NE). Exacerbation isolates were cultured from sputum collected prior to antibiotic treatment of an acute infective exacerbation (E1), post-antibiotic treatment of exacerbation (E2) and at least 4 weeks later, once the patients had returned to clinical stability (S2 PE).

**Bacterial isolates**

Isolates of differing colony morphotype were isolated from CF sputum following incubation on either anaerobic blood agar or kanamycin-vancomycin laked blood agar, in strict anaerobic conditions (10% H2/10% CO2/80% N2) in an anaerobic workstation (Don Whitley Scientific). Following confirmation that these isolates were strictly anaerobic Gram-negative rods (i.e. no growth observed following incubation at 37°C in air for 48 h), their identity was confirmed by comparison of 16S rRNA sequences, as described previously [7, 15]. Briefly, forward and reverse primers (27F 5'-AGAGTTTGATCMTGGCTCAG-3' and 1492R 5'-TACGGYTACCTTGTTACGACTT-3') [16] were used to generate a full length 16S rRNA product, and an additional primer (926R 5'-CCGTCAATTCCTTTRAGTTT-3') added to the subsequent sequencing reaction [17]. The full length 16S rRNA gene was then assembled from the three generated amplicons and compared with those in both the Reference and Representative Genomes (refseq_representative_genomes) database using the BLAST algorithm, and also

<table>
<thead>
<tr>
<th>Patient</th>
<th>S1</th>
<th>S2 NE</th>
<th>E1</th>
<th>E2</th>
<th>S2 PE</th>
</tr>
</thead>
<tbody>
<tr>
<td>B003</td>
<td>P. melaninogenica</td>
<td>P. nigrescens*</td>
<td>P. nigrescens*</td>
<td>P. veroralis/P. histicola</td>
<td></td>
</tr>
<tr>
<td>(n=4)</td>
<td>P. nigrescens*</td>
<td>P. denticola</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B005</td>
<td>P. histicola (n=2)</td>
<td>P. melaninogenica</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=4)</td>
<td>P. melaninogenica</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B008</td>
<td>P. histicola*</td>
<td>P. salivae*</td>
<td>P. histicola*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=4)</td>
<td>P. salivae*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B009</td>
<td>P. histicola†</td>
<td>P. melaninogenica/P. histicola</td>
<td>P. histicola†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=4)</td>
<td>P. veroralis/P. histicola</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B011</td>
<td>P. histicola†</td>
<td>P. histicola†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B014</td>
<td>P. histicola†</td>
<td>P. veroralis/P. histicola†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=4)</td>
<td>P. denticola†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B021</td>
<td>P. denticola†</td>
<td>P. nigrescens</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B023</td>
<td>P. veroralis/P. histicola†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B032</td>
<td>P. salivae†</td>
<td>P. histicola†</td>
<td>P. salivae†</td>
<td>P. nigrescens* (n=2)</td>
<td></td>
</tr>
<tr>
<td>(n=11)</td>
<td>P. histicola†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B040</td>
<td>P. histicola†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Prevotella* spp. from respective patients were not found in the same PFGE or RAPD group (i.e. banding patterns had <80% homology or <50%, respectively).
†Prevotella spp. were found in the same PFGE or RAPD group (i.e. banding patterns had >80% or >50% homology, respectively).
‡Prevotella spp. from respective patients were found in the same RAPD but not PFGE group.

---

Table 1. PFGE and RAPD comparison of *Prevotella* spp. isolated from patients with CF at various clinical time points

Isolates were sampled while the patient was stable (S1), within 4 months of the S1 sample with no exacerbation (S2 NE), pre- and post-antibiotic treatment of an acute exacerbation (E1 and E2, respectively), and at least 4 weeks post-E2 sample when the patient had returned to clinical stability (S2 PE). Where the same species was identified at more than one time point from one patient, this is denoted in bold and timpoints where no *Prevotella* spp. from respective patients were not found in the same PFGE or RAPD group (i.e. banding patterns had <80% homology or <50%, respectively).
to sequences in the Ribosomal Database Project using the Classifier algorithm.

**Genotypic analysis**

**PFGE**
Whole genome DNA extraction was carried out as described elsewhere [18, 19], with the following modification: 100 μM thiourea was added as a reducing agent to the electrophoresis running buffer and agarose gel prior to electrophoresis. Briefly, agarose plugs containing genomic DNA were digested using restriction enzyme XbaI (Life Technologies) and products separated by electrophoresis (switch time 5.3–49.9 s; run time 20 h, 6 V cm⁻², included angle 120°). On completion, the gel was stained using 0.5 mg ethidium bromide ml⁻¹ and banding patterns visualized under ultra-violet illumination. Banding patterns were normalized and compared using GelCompar II software (v 5.10; Applied Maths). Dendrograms were constructed using the unweighted pair group method with arithmetic mean (UPGMA) and the DICE coefficient.

**RAPD**
RAPD was carried out in accordance with previously published methods [20–22] using primer 272 (5’-AGCGGGC-CAA-3’). A dendrogram was constructed from the resulting banding patterns using UPGMA analysis with the DICE coefficient (GelCompar II; Applied Maths). The degree of congruence between partitions assigned by RAPD and PFGE was compared by the adjusted Rand’s and Wallace’s coefficients using an online tool (www.comparingpartitions.info) [23, 24].

**RESULTS**
Forty-two isolates from ten patients were obtained (Table 1) and confirmed as *Prevotella* spp. by comparison of full length 16S rRNA as described earlier, which determined that the *Prevotella* spp. isolated belonged to eight species. In some cases, the isolate could not be accurately determined based on the high similarity obtained between multiple database hits; therefore, the isolate was recorded with both species names.

**Genotypic analysis**

**PFGE**
To determine the overall degree of similarity between all isolates (*n*=42), a dendrogram was constructed and isolates were assigned to groups on the basis of similarity of banding patterns (>80 %) (Fig. 1). This cut-off level was a pragmatic choice, based both on previous literature [18] and inspection of the clustering observed, and generated seven PFGE groups. Of these PFGE groups, 5/7 contained multiple *Prevotella* spp. Analysis of the isolates from individual patients showed that *Prevotella* of the same species were isolated at more than one time point in 9/10 patients (Table 1). In patient B003, *Prevotella nigrescens* isolates were cultured during clinical stability. In patients B008 and B011, *Prevotella salivae* and *Prevotella histicola*, respectively, were isolated when stable (S1) and prior to exacerbation (E1). There were a further nine examples of where the same *Prevotella* spp. was isolated both pre-exacerbation (S1, E1 or both) and post-exacerbation (E2, S2 PE or both). Furthermore, where the same *Prevotella* species was identified at different time points, these isolates were found to belong to the same PFGE group in only 6/12 cases.

Seven *Prevotella* spp. could not be accurately identified to the species level by 16S rRNA sequencing, due to the high similarity obtained between multiple database hits: the isolates were, therefore, recorded with both species names [Prevotella veroralis/P. histicola (*n*=6) and *Prevotella melaninogenica*/P. histicola (*n*=1)]. All *P. veroralis*/P. histicola isolates were found to cluster in the same PFGE group (i.e. share >80 % PFGE banding pattern homology) as the *P. histicola* isolated from their respective patients.

Further analysis of the dendrogram in Fig. 1 identified ten groups of isolates that had identical (i.e. 100 % homology; identified with vertical lines) PFGE banding patterns from eight patients. A total of 3/10 PFGE homologous groups comprised isolates from one patient only, with the remaining 7/10 included isolates from different patients (Fig. 2). In 4/8 patients, PFGE homologues were identified at different clinical time points in that patient; and in 3 patients, this spanned a period of clinical exacerbation.

**RAPD**
To determine the overall degree of similarity between all isolates (*n*=42), a dendrogram was constructed and isolates assigned to groups on the basis of banding pattern homology (>50 %) as described for PFGE analysis (Fig. 3). On this basis, isolates clustered into 14 distinct RAPD groups. As with PFGE, RAPD groups of >1 isolate frequently contained more than one *Prevotella* spp. (*n*=4/6 RAPD groups of >1 isolate). Four sets of isolates were found to have identical RAPD banding patterns; in all cases these RAPD-identical isolates were from the same patient.

As described for PFGE above and shown in Table 1, *Prevotella* of the same species were isolated at more than one time point in 9/10 patients. Isolates from all but one of these patients spanned a period of exacerbation. There was good agreement between PFGE and RAPD banding patterns obtained from these isolates, with one exception: in patient B032, *P. nigrescens* isolated at E1 and S2 PE had RAPD banding patterns >50 % homology and hence these isolates were assigned to the same RAPD group, whereas they had differing PFGE groups. As with PFGE analysis, isolates that could not be accurately identified to the species level by 16S rRNA sequencing, due to the high similarity obtained between multiple database hits (*n*=1 *P. melaninogenica*/P. histicola; *n*=6 *P. veroralis*/P. histicola) were found in each case to have the same RAPD banding profile as *P. histicola*.

**Congruence between PFGE and RAPD**
The levels of congruence observed between PFGE and RAPD were as follows: adjusted Rand coefficient, 0.200; 95 % CI, 0.083–0.325; and the probability that isolates were in the same RAPD and PFGE group was approximately...
45.9% [adjusted Wallace \(RAPD \rightarrow PFGE\) 0.459 (95% CI 0.204–0.714)]. However, the probability that isolates in the same PFGE group were found in the same RAPD group was considerably lower, at 12.8% [adjusted Wallace \(PFGE \rightarrow RAPD\) 0.128 (95% CI 0.020–0.235)].

**DISCUSSION**

The identification of anaerobic niches in the lung, together with increasing data generated from metagenomic analysis of the lung microbiome, has focused attention on the potential pathogenicity of anaerobes, particularly *Prevotella* spp. [6, 7]. Although recognized as endogenous pathogens in various body sites, their role in the development of chronic lung infection has not been clearly defined. There is, therefore, a need for tools to accurately identify and discriminate between *Prevotella* isolates in order to more fully investigate the molecular epidemiology of the genus. This study used PFGE and RAPD to identify genotypically similar *Prevotella* isolates in...
CF patients when clinically stable and during acute infective exacerbations. While this may be indicative of repeated colonization from a common source, a more likely scenario is persistence and establishment of colonization by *Prevotella* from the oral cavity. Other authors have suggested that inadequate clearance of *Prevotella* spp. from the oral cavity is the likely precursor to persistence [25]. Further work is required to investigate how environmental factors, such as iron availability, pH and antibiotic selective pressure, may drive *Prevotella* persistence in the CF lung.

The likely clinical impact of *Prevotella* persistence is not yet clear. This may be due in part to a lack of tools to more fully investigate their molecular epidemiology, complicated further by their fastidious nature and difficulties involved in culture. Some classical indicators that are suggestive of *Prevotella* as the causative agent of infection have previously been described in the literature, and include detection of increased titres of anti-*Prevotella* antibodies during exacerbation in chronic bronchitis [26, 27] and the association between the presence of supraglottic taxa, including *Prevotella* spp. [28]. Furthermore, in a mouse model of chronic lung infection, high numbers of *Prevotella intermedia* were found to induce inflammation [29]. These findings are in contrast to other work that suggests *Prevotella* spp. are associated with less inflammation and improved lung function compared to classical pathogens such as *Pseudomonas aeruginosa* [30].

Comparison of the inflammatory response between established airway pathogens and *Prevotella* spp. suggested that colonization with the latter was only weakly inflammatory, likely to be tolerated by the immune system and suggestive of a potentially disease protective role [31, 32]. They may also exert an indirect effect on the lung microbiome: *Prevotella* spp. have been shown to produce β-lactamases, and contribute to antibiotic resistance within the lung microbiome [33, 34]. In this current study, persistence did not appear to be affected by the antibiotic therapy given either at the time of exacerbation or as part of normal CF maintenance therapy. However, it is not routine practice in CF to administer antibiotic therapy specifically targeting anaerobes in the treatment of exacerbations. The difficulty in interpreting these conflicting results and, hence, assigning a pathogenic role to *Prevotella* spp., may lie in the relatively blunt tools used to identify the phenotypically diverse members of the genus. Metagenomic non-culture-based analyses investigating the lung microbiome are able to identify *Prevotella* spp. at family and genus level only, and differences in pathogenicity between *Prevotella* species may, therefore, be missed. Many laboratories lack the capacity to reliably culture anaerobes, and this, coupled with the relatively recent realization that these bacteria may be involved in infection, has led to a lower representation of *Prevotella* spp. in comparison databases.

Reasonable congruence was observed between PFGE and RAPD typing methods, with both techniques yielding highly similar banding patterns from longitudinal *Prevotella* isolates. PFGE is a highly discriminatory method for bacterial typing, considered by many as the ‘gold standard’ of molecular epidemiology [35, 36]. Although pulsed-field patterns are generally stable and reproducible, the technique has several limitations, primarily the length of time taken to generate results and cost. In addition, it cannot be definitely determined if bands of the same size represent the same DNA, and for some bacterial strains PFGE can be poorly discriminatory [37]. Finally, PFGE requires highly harmonized laboratory procedures in order to allow comparison between laboratories. In this study, comparison of the PFGE banding patterns obtained from all *Prevotella* spp. analysed yielded a diverse range of genotypes which showed only limited congruence with 16S species designation. However, at a patient level, highly similar isolates of the same species and genotype were observed at different sampling points in each patient.

**Fig. 2.** *Prevotella* isolates with identical PFGE banding patterns isolated at different clinical time points from eight patients. Clinical time points examined were when the patient was clinically stable (S1), stable with no recorded exacerbation event (S2 NE), prior to (E1) and post- (E2) antibiotic treatment of an acute infective exacerbation, and a post-exacerbation sample (S2 PE), taken at least 4 weeks following antibiotic treatment of exacerbation. Isolate homology was determined by UPGMA analysis with the DICE coefficient, and isolates with 100% homology were assigned to the same PFGE homology (PH) group.
clusters generated. More within-group homogeneity was observed with RAPD compared to PFGE, with respect to 16S rRNA species designation. Where identical RAPD profiles were generated, in the majority of cases these isolates were of the same species designation. As with PFGE, RAPD analysis also showed highly similar banding profiles in isolates from the same patient at different time points, during both periods of exacerbation and clinical stability.

There are a number of limitations with the present study. Prevotella spp. isolates were not cultured at each patient sampling point and it cannot be determined from this study whether that was due to the limitations of our culture technique or whether Prevotella spp. were below the limits of culture detection. Isolates were collected on the basis of differing colony morphology, and this phenotypic heterogeneity was reflected in the diverse range of species and PFGE/
RAPD banding patterns obtained. However, the extent of intra-species variation occurring among the banding patterns observed is not known. Isolates were assigned to the same PFGE or RAPD group if their banding patterns showed >80 % or >50 % homology, respectively; these were arbitrary cut-off points based on inspection of the patterns of clustering obtained. This meant that some RAPD groups contained only one isolate, and analysis using both RAPD and PFGE produced different numbers of groups and cut-off values. This is, however, consistent with other comparisons of typing methods [23, 24, 38–41] and while this may make statistical analysis of the techniques more complex, it is nonetheless a reflection of the diversity apparent within the Prevotella genus. Use of the adjusted Rand’s and Wallace’s coefficients, which have been widely used to compare partitions generated by different typing methods [23, 24], suggests that there was reasonable congruence between the two typing methods. Of interest was that in some instances, species identity could not be accurately identified due to high similarity between multiple hits when using both the BLAST algorithm and rdp classifier tool; however, the PFGE/RAPD profiles obtained from these isolates showed very high similarity with other isolates and, hence, added clarity to their identification. Therefore, 16S rRNA may not always be able to adequately discriminate between isolates of this phenotypically diverse genus. From the limited numbers of isolates and patients investigated in this current study, it is not possible to determine whether there is the possibility for shared clones between patients. Decreasing costs and turnaround time associated with whole genome sequencing are likely to make this a viable option for future molecular epidemiology studies. Indeed, recent studies have shown that it appears to generate superior resolution to PFGE [42, 43].

In summary, this study has shown that genotypically similar Prevotella spp. can be isolated longitudinally from the sputum of CF patients, during periods of clinical stability and pre- and post-antibiotic treatment of an exacerbation, which is strongly suggestive of persistence within the lungs of people with CF. Furthermore, it has demonstrated the capacity of PFGE and RAPD for genotypic comparisons of Prevotella spp. A greater understanding of the molecular epidemiology of this diverse genus is required in order to determine how this relates to patient outcome in chronic lung infection.

Funding information
This work was funded by a US–Ireland Partnership grant from the Health and Social Care Research and Development Office, Public Health Agency Northern Ireland. K. N. was funded by the Department for Education and Learning for Northern Ireland.

Conflicts of interest
No conflict of interest is declared by any of the authors.

Ethical statement
This work involved human participants and ethical approval was granted by the Office for Research Ethics Committees Northern Ireland reference no. 10/NIR01/41; IRAS project no. 41579. Informed consent was obtained from all study participants.

References


---

Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.