Maternal oral origin of *Fusobacterium nucleatum* in adverse pregnancy outcomes as determined using the 16S–23S rRNA gene intergenic transcribed spacer region

Cecilia Gonzales-Marin,1 David A. Spratt2 and Robert P. Allaker1

1Institute of Dentistry, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, 4 Newark Street, London E1 2AT, UK
2Division of Microbial Diseases, Eastman Dental Institute, University College London, 256 Gray’s Inn Road, London WC1X 8LD, UK

*Fusobacterium nucleatum*, a common Gram-negative anaerobe prevalent in the oral cavity, possesses the ability to colonize the amniotic cavity and the fetus. However, *F. nucleatum* may also be part of the vaginal microbiota from where it could reach the amniotic tissues. Due to the heterogeneity of *F. nucleatum*, consisting of five subspecies, analysis at the subspecies/strain level is desirable to determine its precise origin. The aims of this study were: (i) to evaluate the use of the 16S–23S rRNA gene intergenic transcribed spacer (ITS) region as a tool to differentiate subspecies of *F. nucleatum*, and (ii) to design a simplified technique based on the ITS to determine the origin of *F. nucleatum* strains associated with adverse pregnancy outcomes. Amplified fragments of the 16S–23S rRNA gene ITS region corresponding to the five subspecies of *F. nucleatum* were subjected to cloning and sequencing to characterize the different ribosomal operons of the subspecies. Distinctive length and sequence patterns with potential to be used for identification of the subspecies/strain were identified. These were used to evaluate the origin of *F. nucleatum* identified in neonatal gastric aspirates (swallowed amniotic fluid) by sequence comparisons with the respective oral and vaginal maternal samples. A simplified technique using a strain-specific primer in a more sensitive nested PCR was subsequently developed to analyse ten paired neonatal–maternal samples. Analysing the variable fragment of the ITS region allowed the identification of *F. nucleatum* subsp. *polymorphum* from an oral origin as potentially being involved in neonatal infections. Using a strain-specific primer, the *F. nucleatum* subsp. *polymorphum* strain was detected in both neonatal gastric aspirates and maternal oral samples in cases of preterm birth from mothers presenting with localized periodontal pockets. Interestingly, the same strain was not present in the vaginal sample of any case investigated. The 16S–23S rRNA gene ITS can be a useful tool to determine the origin of *F. nucleatum*. The results of this study strongly indicate that *F. nucleatum* subsp. *polymorphum* of oral origin could be involved with pregnancy complications.

**INTRODUCTION**

Maternal periodontal infection has been investigated for a number of years as a risk factor for adverse pregnancy outcomes, particularly preterm delivery and low birthweight (Offenbacher et al., 1996; Jeffcoat et al., 2011); however, to date there is no conclusive evidence to support this association. The biological mechanisms can possibly be explained by translocation of bacteria from the oral cavity to the amniotic environment followed by a local immune reaction (Han et al., 2004; Liu et al., 2007). Although

**Abbreviations:** ATCC, American Tissue Culture Collection; DSMZ, German Collection of Microorganisms and Cell Cultures; HOMD, Human Oral Microbiome Database; ITS, intergenic transcribed spacer; NCTC, UK National Collection of Type Cultures; RDP, Ribosomal Database Project; TA, annealing temperature.


Three supplementary figures are available with the online version of this paper.
periodontal disease is characterized by a polymicrobial infection, *Fusobacterium nucleatum* presents with the strongest evidence regarding translocation properties and association with various extra-oral infections (Bolstad *et al.*, 1996).

*F. nucleatum* is a Gram-negative anaerobe and the most common bacterial species isolated from amniotic fluid in cases of preterm labour and intact membranes (Hill, 1998). This bacterium can invade epithelium, endothelium and placental tissues (Han *et al.*, 2000, 2004; Saito *et al.*, 2008). It has been detected in relation to high-risk pregnancies, present and past pregnancy complications and, as demonstrated recently, stillbirth (Bearfield *et al.*, 2002; Han *et al.*, 2009, 2010; Tateishi *et al.*, 2012). Furthermore, studies in animals have confirmed that *F. nucleatum* is able to translocate via the bloodstream, specifically invade the amniotic tissues and cause fetal death (Han *et al.*, 2004). Nevertheless, *F. nucleatum* may also have originated from non-oral sources, such as the female genito-urinary tract or the gastrointestinal tract (Strauss *et al.*, 2008; Gonzales-Marin *et al.*, 2011); hence, the importance of determining its true origin.

Using the 16S rRNA gene, we have previously confirmed the presence of *F. nucleatum* in samples of gastric aspirates acquired from neonates that were born from complicated pregnancies (Gonzales-Marin *et al.*, 2011, 2012). In terms of bacterial presence, the aspirates were shown to contain swallowed amniotic fluid, and thus have been proposed as an alternative to investigate fetal infections associated with adverse pregnancy outcomes (Miralles *et al.*, 2005; Gonzales-Marin *et al.*, 2012). In our previous studies, *F. nucleatum* was observed at high levels in the neonatal samples (up to $2.32 \times 10^3$ cells ml$^{-1}$, representing 49.5% of the total bacterial load), which strongly supports its possible role in causing complications. It was also present in the oral cavity of the respective mother and, although in considerably lower numbers and at smaller proportions, the pathogen was likewise observed in the low vaginal swabs acquired from the same woman (Gonzales-Marin *et al.*, 2011); therefore, due to its close physical proximity to the uterus, this site should also be considered as a potential source of the pathogen.

The heterogeneity of *F. nucleatum* on the basis of physiological, biochemical and molecular variability demonstrates that distinct subgroups exist within the species. Five subspecies of *F. nucleatum* have been proposed: *polymorphum*, *nucleatum*, *animalis*, *fusiforme* and *vincentii* (Dzink *et al.*, 1990; Gharbia & Shah, 1992). In the oral cavity, it was determined that *F. nucleatum* subsp. *nucleatum* was more prevalent in diseased sites, whereas isolates from healthy sites were generally identified as *F. nucleatum* subsp. *polymorphum* and *fusiforme* (Gharbia *et al.*, 1990). However, another study did not detect any difference between healthy and diseased sites (Berres *et al.*, 2003). In contrast, the subspecies *animalis* and *nucleatum* have been shown to be present in vaginal samples, whilst DNA from the subspecies *polymorphum*, *fusiforme* and *vincentii* was not amplified from this site (Gonzales-Marin *et al.*, 2011). Although *F. nucleatum* subsp. *animalis* has been reported occasionally in the oral cavity (Dewhirst *et al.*, 2010), it has also been reported as commonly prevalent in the human gut (Strauss *et al.*, 2008), and a transient status due to oral-genital practice could explain the sporadic findings. Indeed, the presence of a possible opportunistic *F. nucleatum* subsp. *animalis* strain has been observed primarily in refractory cases of periodontitis, necrotizing ulcerative gingivitis and human immunodeficiency virus infection (Paster *et al.*, 2001; Gmür *et al.*, 2006). It is therefore important to identify more precisely the subspecies/substrains involved in adverse pregnancy outcomes, as this could provide a more reliable indication of the origin of the infection.

With the advent of culture-independent molecular methods, the 16S rRNA gene has been established as the gold standard for identification and taxonomic classification of bacterial species. Nevertheless, a disadvantage of using the 16S rRNA component is the lack of sufficient resolution to enable a precise differentiation between strains of the same species. The 16S–23S rRNA gene intergenic transcribed spacer (ITS) region is increasingly being used for the analysis of phylogenetic relationships between species and subspecies of the same genus (Tyrrell *et al.*, 1997; Sadeghifard *et al.*, 2006), due to a higher number of variable sites compared with the more conserved 16S rRNA gene (Söller *et al.*, 2000). The heterogeneity of the ITS region in terms of both copy number and length has raised the possibility of including this fragment as part of the methodology for bacterial identification and typing purposes (Gürtler & Stanisich, 1996; Popa *et al.*, 2009). Previously, the ITS has been used to examine inter- and intra-species relationships of *Fusobacterium* spp., demonstrating a high phylogenetic resolution (Conrads *et al.*, 2002). However, there remains a lack of information regarding the ability of this genetic region to investigate evidence of translocation from one site to another.

This study aimed to examine the ITS region as a validated tool to differentiate subspecies/substrains of *F. nucleatum*. This was followed by the application of this methodology to identify precisely the origin of *F. nucleatum* in neonatal gastric aspirates by comparison with its counterparts in the maternal oral cavity and vaginal sequences, and through the construction of a strain-specific primer to identify the source.

**METHODS**

**Reference strains and samples.** The most accessible culture bacterial collections were searched for available strains of *F. nucleatum* that have been fully characterized phenotypically from human oral samples. With the exception of *F. nucleatum* subsp. *nucleatum*, all other subspecies had a single representative strain in the American Tissue Culture Collection (ATCC), UK National Collection of Type Cultures (NCTC), German Collection of Microorganisms and Cell Cultures (DSMZ), Japan Collection of Microorganisms (JCM), Culture Collection, University of Göteborg and Institute Pasteur Collection. All the available strains were isolated originally by Dzink *et al.* (1990) and Gharbia & Shah (1992). Therefore, for the purpose
of this study, one strain from each of the five subspecies of *F. nucleatum* was investigated fully.

The strains *F. nucleatum* subsp. *polymorphum* NCTC 10562, *F. nucleatum* subsp. *nucleatum* ATCC 25586, *F. nucleatum* subsp. *vincentii* DSM 19507, *F. nucleatum* subsp. *fusiforme* DSM 19508 and *F. nucleatum* subsp. *animalis* DSM 19679 were grown on blood agar under anaerobic incubation conditions as detailed previously (Gonzales-Marin et al., 2011). The neonatal gastric aspirate samples and those from the maternal oral cavity and vagina have been investigated previously by our group using broad-range PCR primers directed to the 16S rRNA gene (Gonzales-Marin et al., 2012) and quantitative PCR approaches (Gonzales-Marin et al., 2011); details on sample collection, extraction of genomic DNA and bacterial content are detailed in these studies. Briefly, neonatal samples were obtained through aspiration from newborns who presented with a clinical manifestation of possible neonatal sepsis or were considered at risk of developing an infection (i.e. preterm and low-birthweight infants).

The environmental contamination of samples was determined as possible; however, contamination from oral sources was not likely, as demonstrated previously (Gonzales-Marin et al., 2012). The samples from the maternal oral cavity represented a pool of saliva, tongue scrapings and suprap- and subgingival plaque. The vaginal samples consisted of a self-collected low vaginal swab. Ethical approval was obtained from the Outer North East London Research Ethics Committee, ref. no. 08/H0701/61 (currently National Research Ethics Service, Committee London-East, UK).

### Specific amplification of the 16S–23S rRNA gene ITS region.

Fn1F (5′-GACAGAAGCTTTGCGTCC-3′), a forward primer directed to a fragment in the 16S rRNA gene and specific for the genus *Fusobacterium* (Naganlo et al., 2007), and the 456R universal primer (5′-CCTTTCCTACGCTACT-3′) within the 23S rRNA gene (Gürtler & Stanisich, 1996) were used to amplify a fragment containing the 16S rRNA gene, the complete ITS region and a fragment of the 23S rRNA gene (~2000 bp) from both reference (Gürtler & Stanisich, 1996) were used to amplify a fragment (5′-GACAGAAGCTTTGCGTCC-3′) and 456R, producing sequences of 800–1000 bp. The quality of the sequences was checked using the Chromas software package (http://www.technelysium.com.au/?page_id=13). Sequences were finally aligned and corrected using Multalign, a multiple sequence alignment tool (Corpet, 1988).

Bacterial sequences were identified by comparison with libraries of sequences based on ≥97% similarity with the closest relative, as suggested previously by Patel (2001). The online libraries used are extensive and were comprised of both 16S rRNA gene and other sequence segments; these libraries were the Ribosomal Database Project (RDP) (Cole et al., 2007), the Human Oral Microbiome Database (HOMD) (Chen et al., 2010) and GenBank of the National Center for Biotechnology Information (Benson et al., 2011).

### Strain-specific primer design and nested PCR.

Using the ITS information from *F. nucleatum* subsp. *polymorphum*, a strain-specific forward primer was constructed and optimized in a nested PCR in combination with the 456R primer. The FnpF (5′-CCTTCTAAAGGAGAATAGTC-3′) oligonucleotide contained the highly variable fragment (underlined) of the ITS identified in this study. The set of primers used specifically amplified almost the complete ITS region plus a portion of the 23S rRNA gene and therefore they included all the different sizes of the *rnl* operons. The annealing temperature (Tₐ) used was determined experimentally by running PCR assays at a range of temperatures from 43 to 53 °C. The total reaction volume of 25 μl contained 0.5 μl sterile water, a final concentration of 0.2 μM each primer, 1.1 × ReddyMix PCR Master Mix (1.5 mM MgCl₂, Thermo Fisher Scientific) and 1 μl bacterial DNA extracted from 10⁸ c.f.u. ml⁻¹ of each *F. nucleatum* subspecies. PCR conditions consisted of a denaturation step of 95 °C for 5 min, followed by 35 cycles of 95 °C for 25 s, 35 s at the Tₐ and 40 s of 72 °C, with 3 min final extension at 72 °C. A temperature of 52 °C was determined as the optimal temperature to ensure primer specificity (Fig. 2a). Sensitivity was also tested in this study by running a serial dilution of bacterial DNA in the nested PCR approach and was determined to be <10⁵ c.f.u. ml⁻¹ (Fig. 2b), as a very faint band could not be detected.

### Extended agarose gel analysis.

An extended agarose gel analysis was used to confirm amplification of the ITS region and to discriminate variations of the fragment sizes in the gel. The amplified products of ~2000 bp were resolved on a 0.8 % agarose gel in Tris-acetate/EDTA electrophoresis buffer at 65 V, using HyperLadder I or IV (Bioline) size markers. An extended electrophoresis time of 4 h was determined as the most effective period to resolve the bands (data not shown).

### Sequencing, alignment and sequence comparative analysis.

All sequencing was carried out at the Genome Centre, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, using BigDye 3.1 chemistry with visualization on an ABI 3700 automated DNA sequencer. The primers used for sequencing were 1114F (5′-GAAGACGAGCGCAACCG-3′) and 456R, producing sequences of 800–1000 bp. The quality of the sequences was checked using the Ribosomal Database Project (RDP) (Cole et al., 2007), the Human Oral Microbiome Database (HOMD) (Chen et al., 2010) and GenBank of the National Center for Biotechnology Information (Benson et al., 2011).

Sequencing, alignment and sequence comparative analysis. All sequencing was carried out at the Genome Centre, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, using BigDye 3.1 chemistry with visualization on an ABI 3700 automated DNA sequencer. The primers used for sequencing were 1114F (5′-GAAGACGAGCGCAACCG-3′) and 456R, producing sequences of 800–1000 bp. The quality of the sequences was checked using the Chromas software package (http://www.technelysium.com.au/?page_id=13). Sequences were finally aligned and corrected using Multalign, a multiple sequence alignment tool (Corpet, 1988).

Bacterial sequences were identified by comparison with libraries of sequences based on ≥97% similarity with the closest relative, as suggested previously by Patel (2001). The online libraries used are extensive and were comprised of both 16S rRNA gene and other sequence segments; these libraries were the Ribosomal Database Project (RDP) (Cole et al., 2007), the Human Oral Microbiome Database (HOMD) (Chen et al., 2010) and GenBank of the National Center for Biotechnology Information (Benson et al., 2011).

### Cloning analysis.

The amplicons generated by PCR were cloned into One Shot TOP10 chemically competent *Escherichia coli* using a TOPO TA Cloning kit (Invitrogen), as recommended by the manufacturer. PCR products obtained with primers Fn1F and 456R were then purified using a QIAquick PCR Purification kit (Qiagen) for sequencing.

**Fig. 1.** Schematic representation of the 16S rRNA gene, ITS and 23S rRNA gene regions of the *Fusobacterium* species amplified in this study, including the position of primers used for amplification and sequencing. The expected lengths of the ITS (*) were confirmed by comparison with a previous study (Conrads et al., 2002).
be seen for the $10^2$ c.f.u. ml$^{-1}$. For the nested PCR, the first PCR was performed with primers Fn1F and 456R following the conditions described above. The second PCR was identical to that described in this paragraph. Finally, the products were run in 1 % agarose gels in an extended analysis.

**RESULTS**

**Sequence analysis of the subspecies of F. nucleatum grown in culture**

The ITS region was amplified from the five subspecies of *F. nucleatum*. The products from each subspecies were cloned and 20 randomly selected clones from each subspecies were run in an extended agarose gel. Three different fragment sizes could be differentiated in the gel for *F. nucleatum* subsp. *polymorphum* (Fig. 3), whilst two fragment sizes were observed for the other subspecies. As observed in the gels, fragment lengths ranged between ~1900 and 2300 bp. Six to 12 representative clones for each strain were sequenced. Interestingly, variable results were observed when sequences were identified using the current standard methods for species identification. These consisted of sequence comparisons against three different libraries available: GenBank, RDP and HOMD (Table 1). In general, *F. nucleatum* subsp. *nucleatum* was the only strain correctly identified by the three libraries. *F. nucleatum* subsp. *polymorphum* was accurately identified by RDP, whilst subsp. *fusiforme* and *vincentii* were identified by GenBank. In the case of *F. nucleatum* subsp. *animalis*, clones amplified from this strain were also recognized as subsp. *animalis* by the three libraries; however, it was equally identified as subsp. *nucleatum* in all

---

**Fig. 2.** Strain-specific primer test in a nested PCR. (a) Specificity test based on $T_A$ variation. (b) Nested PCR using the optimal $T_A$ (52 °C) to test serial dilutions of *F. nucleatum* subsp. *polymorphum* and samples obtained from case study 10. $H_N$, HyperLadder IV (Bioline); Fnp, *F. nucleatum* subsp. *polymorphum*; Fna, *F. nucleatum* subsp. *animalis*; Fnn, *F. nucleatum* subsp. *nucleatum*; Fnf, *F. nucleatum* subsp. *fusiforme*; Fnv, *F. nucleatum* subsp. *vincentii*; (–), negative control; GA10, neonatal gastric aspirate from case study 10; OC10, maternal oral sample from case study 10; V10, vaginal sample from case study 10; No Fnp mix, equal mixture of $10^5$ c.f.u. ml$^{-1}$ concentration from all *F. nucleatum* subspecies with the exception of subsp. *polymorphum*. The primers used to amplify the ITS from the sample of GA10 (indicated by *) were the universal primer 1114F (instead of FnF1) and 456R. The arrows indicate a very faint band observed in the gel.
cases. Similar results were observed with the other strains, as detailed in Table 1, and in a few cases, they failed to be identified precisely.

Published data obtained from the Oral Pathogen Sequence Database, Los Alamos National Laboratory (http://oralgen.lanl.gov/), relating to the five different operons of *F. nucleatum* subsp. *nucleatum* ATCC 25586 were used in this study to identify the sizes of the ITS region and to validate the results produced. The analysis indicated that this subspecies presented four short, identical operons and one long operon as a result of size variation of the ITS fragment (155 and 330 bp, respectively). A comparison between the published data and data produced from the pure culture of the ATCC 25586 strain using the protocol outlined in this study demonstrated that the methodology successfully amplified the two different fragments of *F. nucleatum* subsp. *nucleatum* and these were 100% identical to the published data (Fig. S1, available in JMM Online). When the long and short fragments were compared by alignment, it could be seen that the long fragment of the ITS contained an extra 9 bp fragment, followed by one tRNA\textsuperscript{Ala} (77 bp), 3 bp (TGT), one tRNA\textsuperscript{Ile} (77 bp) and another 9 bp fragment in the ITS (Fig. 4).

Alignments of the other *F. nucleatum* strain sequences against *F. nucleatum* ATCC 25586 data were performed similarly to calculate the lengths and sequences of the ITS region for all strains. A total of five different lengths of the ITS were identified in the five strains: three bands for *F. nucleatum* subsp. *polymorphum* and two bands for the other strains. As observed, the short and long fragments were identical in size for *F. nucleatum* subspsp. *animalis*, *nucleatum* and *polymorphum* (short: 156 bp, long: 331 bp), and subsp. *polymorphum* also contained a medium-sized ITS (222 bp). These differed from the corresponding fragment lengths in *F. nucleatum* subspsp. *fusiforme* and *vincteini* (short: 147 bp, long: 322 bp) with no medium-sized fragment observed. It was observed that both long ITS fragments contained tRNA\textsuperscript{Ala} and tRNA\textsuperscript{Ile} sequences. It was evident that the variability in length was due to the presence or absence of certain regions (of different lengths) along the ITS region (Fig. S2). It was also observed that the five strains presented few mismatches along the 16S rRNA gene, ITS and the 23S rRNA gene.

Interestingly, a short variable region contained in the ITS (nt 28–32) could clearly be identified with variations specific for the different strains irrespective of the operon or fragment length (Table 2). Two variable fragments corresponding to *F. nucleatum* subsp. *animalis* were observed. *F. nucleatum* subspsp. *polymorphum* and *nucleatum* also seemed to share an identical variable fragment (AATGT), but this was only present in the medium-sized fragment for *F. nucleatum* subsp. *polymorphum* and in both short and long fragments of subsp. *nucleatum*, which did not have a medium-sized ITS. The *F. nucleatum* subsp. *polymorphum* strain also presented a unique fragment (AATAA) in all three sizes.

**Sequence analysis of *F. nucleatum* from clinical samples**

Paired neonatal–maternal samples from a case study (no. 10) known to contain *F. nucleatum* in the neonatal, oral and vaginal samples (as shown using 16S rRNA gene PCR) was tested with the comprehensive cloning and sequence comparison technique described above. This was a case of a preterm infant, delivered by caesarean section at 35 weeks of gestation, and the woman had suffered a previous miscarriage and presented with generalized chronic gingivitis and localized areas of periodontitis with probing depths of 4–5 mm in four sites. This case study was observed to contain high levels of *Fusobacterium* spp. in the neonatal gastric aspirate sample, representing 37% of the total bacterial load.

As the primer set Fn1F/456R used in this study to amplify the ITS region was not exclusive for the *F. nucleatum* subspecies but was also functional for all species of the genus *Fusobacterium*, other species were also identified. Seven clones from the neonatal gastric aspirate of case 10 were identified as *Sneathia sanguinegens*, whilst the remaining seven clones were confirmed as *F. nucleatum*. Two clones of *Leptotrichia* spp. and two clones of *Leptotrichia buccalis* were also identified in the oral cavity and vagina, respectively. These clones were excluded from the comparisons; consequently, 20 clones from the oral cavity and 14 vaginal clones identified as *F. nucleatum* were included in the alignments for comparison (Fig. S3a–d).

It is likely that, due to the low number of clones analysed, it was not possible to obtain a 100% match between one of the clones from the neonatal sample and a clone from either the oral cavity or the vagina. However, when analysing the variable region of the ITS, it could be seen that the gastric aspirates contained AATAA (*F. nucleatum* subsp. *polymorphum*) and also AATGT in the long fragment (*F. nucleatum* subsp. *polymorphum*), which was only observed in the oral strains and not in the vaginal strains (Table 3).

**F. nucleatum** subsp. *polymorphum*-specific primer and nested PCR

Following the observations above, a strain-specific primer for *F. nucleatum* subsp. *polymorphum* was designed and optimized in a nested PCR approach. The use of a nested
Table 1. Results of the conventional library sequencing method compared against three sequence databases (GenBank, RDP and HOMD) used to identify the clones obtained from the five strains of *F. nucleatum* grown in this study

Fragments were amplified using primers Fn1F and 456R and contained part of the 16S rRNA gene (1290 bp), the complete ITS (147–331 bp) and part of the 23S rRNA gene (460 bp). The sizes of the ITS were determined as long: A1, 331 bp; A2, 322 bp; medium: B, 222 bp; and short: C1, 156 bp; C2, 147 bp.

<table>
<thead>
<tr>
<th>Species</th>
<th>Clone no.</th>
<th>ITS size</th>
<th>GenBank*</th>
<th>Query coverage/maximum identity range (%)</th>
<th>RDP</th>
<th>Similarity score (range)</th>
<th>HOMD*</th>
<th>Identity range (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>F. nucleatum</em> subsp.</td>
<td>14, 15, 17 A1</td>
<td>Subspecies <em>animalis</em></td>
<td>100/100 Subspecies <em>animalis</em> JCM 11025</td>
<td>Subspecies <em>animalis</em> JCM 11025</td>
<td>0.991–0.997</td>
<td>Subspecies <em>animalis</em> 99.0–99.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>animalis</em> DSM 19679</td>
<td></td>
<td>DSM 19679 Subspecies <em>nucleatum</em> ATCC 25586</td>
<td></td>
<td>Subspecies <em>nucleatum</em> ChDC OS50; AF543300</td>
<td>0.991–0.997</td>
<td>Subspecies <em>animalis</em> 99.0–99.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1, 8, 9, 19, 12, 16 C1</td>
<td>Subspecies <em>animalis</em></td>
<td>100/100 Subspecies <em>animalis</em> JCM 11025</td>
<td>Subspecies <em>animalis</em> JCM 11025</td>
<td>0.994–0.997</td>
<td>Subspecies <em>animalis</em> 99.2–99.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DSM 19679 Subspecies <em>nucleatum</em> ATCC 25586</td>
<td></td>
<td>Subspecies <em>nucleatum</em> ChDC OS50; AF543300</td>
<td>0.994–0.997</td>
<td>Subspecies <em>animalis</em> 99.2–99.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> subsp.</td>
<td>7, 16 A1</td>
<td>Subspecies <em>nucleatum</em></td>
<td>98–99/100 Subspecies <em>nucleatum</em> ATCC 25586</td>
<td>Subspecies <em>nucleatum</em> ATCC 25586</td>
<td>0.997</td>
<td>Subspecies <em>nucleatum</em> 99.7–100</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>nucleatum</em> ATCC 25586</td>
<td></td>
<td>ATCC 25586 Subspecies <em>nucleatum</em> ATCC 25586</td>
<td></td>
<td>Subspecies <em>nucleatum</em> ATCC 25586</td>
<td>0.994–0.997</td>
<td>Subspecies <em>nucleatum</em> 99.2–100</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2, 4, 5, 8, 13, 14, 15 C1</td>
<td>Subspecies <em>nucleatum</em></td>
<td>99–100/99–100 Subspecies <em>nucleatum</em> ATCC 25586</td>
<td>Subspecies <em>nucleatum</em> ATCC 25586</td>
<td>0.997</td>
<td>Subspecies <em>nucleatum</em> 99.5–99.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> subsp.</td>
<td>2, 8, 11, 15 A1</td>
<td>Subspecies <em>animalis</em></td>
<td>100/98–100 Subspecies <em>polymorphum</em> JCM 12990</td>
<td>Subspecies <em>polymorphum</em> JCM 12990</td>
<td>0.97</td>
<td>Subspecies <em>polymorphum</em> 99.5–99.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>polymorphum</em> NCTC 10562</td>
<td></td>
<td>DSM 19679 Subspecies <em>animalis</em> ATCC 25586</td>
<td></td>
<td>Subspecies <em>polymorphum</em> JCM 12990</td>
<td>0.989–0.994</td>
<td>Subspecies <em>polymorphum</em> 99.0–99.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1, 3, 4, 7, 9, 12 B</td>
<td>Subspecies <em>nucleatum</em></td>
<td>93/98–100 Subspecies <em>polymorphum</em> JCM 12990</td>
<td>Subspecies <em>polymorphum</em> JCM 12990</td>
<td>0.989–0.994</td>
<td>Subspecies <em>polymorphum</em> 99.0–99.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ATCC 25586 Subspecies <em>nucleatum</em> ATCC 25586</td>
<td></td>
<td>Subspecies <em>polymorphum</em> JCM 12990</td>
<td>0.994–0.997</td>
<td>Subspecies <em>polymorphum</em> 99.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5, 10 C1</td>
<td>Subspecies <em>nucleatum</em></td>
<td>100/98 Subspecies <em>polymorphum</em> JCM 12990</td>
<td>Subspecies <em>polymorphum</em> JCM 12990</td>
<td>0.994–0.997</td>
<td>Subspecies <em>polymorphum</em> 99.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ATCC 25586 Subspecies <em>nucleatum</em> ATCC 25586</td>
<td></td>
<td>Subspecies <em>polymorphum</em> JCM 12990</td>
<td>0.992</td>
<td>Subspecies <em>polymorphum</em> 99.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. nucleatum</em> subsp.</td>
<td>1, 3, 12, 13, 15 A2</td>
<td>Subspecies <em>fusiforme</em></td>
<td>100/100 Subspecies <em>fusiforme</em> DSM 19508 (type strain)</td>
<td>Subspecies <em>fusiforme</em> DSM 19508 (type strain)</td>
<td>0.992</td>
<td>Subspecies <em>fusiforme</em> 99.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>fusiforme</em> DSM 19508</td>
<td></td>
<td>DSM 19508 Subspecies <em>fusiforme</em> DSM 19508</td>
<td></td>
<td>Subspecies <em>fusiforme</em> DSM 19508 (type strain)</td>
<td>0.992</td>
<td>Subspecies <em>fusiforme</em> 99.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7, 8 A2, C2</td>
<td>Subspecies <em>fusiforme</em></td>
<td>100/100 Subspecies <em>fusiforme</em> DSM 19508 (type strain)</td>
<td>Subspecies <em>fusiforme</em> DSM 19508 (type strain)</td>
<td>0.992</td>
<td>Subspecies <em>fusiforme</em> 99.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DSM 19508 Subspecies <em>fusiforme</em> DSM 19508</td>
<td></td>
<td>Subspecies <em>fusiforme</em> DSM 19508 (type strain)</td>
<td>0.992</td>
<td>Subspecies <em>fusiforme</em> 99.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DMS 20699, NCTC 11464 Subspecies <em>fusiforme</em> DSM 19508 (type strain)</td>
<td></td>
<td>Subspecies <em>fusiforme</em> DSM 19508 (type strain)</td>
<td>0.992</td>
<td>Subspecies <em>fusiforme</em> 99.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C. González-Marin, D. A. Spratt and R. P. Allaker
PCR allowed a notable improvement in the sensitivity, with a detection limit of $10^3$ cells ml$^{-1}$, whilst the direct approach only allowed detection when samples contained at least $10^5$ cells ml$^{-1}$ (data not shown). A double band was observed in the gels for *F. nucleatum* subsp. *polymorphum*, possibly corresponding to the medium-sized fragment (calculated fragment of 662 bp as amplified with primers FnpF and 456R) and the long fragment (775 bp) sequences, but no fragment was observed for the short fragment expected to be 596 bp, whilst for *F. nucleatum* subsp. *animalis*, *nucleatum*, *fusiforme* and *vincentii*, only one band was observed, corresponding to the long fragment. A preferential amplification has been proposed previously by Conrads et al. (2002) for similar observations; however, strains showing one band that has been cloned previously have demonstrated amplification of all two or three ITS sizes. It is therefore likely that the agarose gel did not allow a good resolution. The primers were demonstrated to be specific for *F. nucleatum* subsp. *polymorphum* (Fig. 2b) and also confirmed that the strain was only present in the oral cavity and the neonatal samples, and not in the vaginal sample from case study 10. Using this simplified technique, ten paired neonatal–maternal samples known to contain *F. nucleatum* were also investigated and the results are shown in Table 4.

### DISCUSSION

In this study, we introduced a simplified tool to investigate the origin of *F. nucleatum* strains in neonates presenting with adverse pregnancy outcomes. For this purpose, we used the 16S–23S rRNA gene ITS sequence. Previously, based on the 16S rRNA gene, we observed that subspecies of *F. nucleatum* in neonates matched better with the strains amplified from the maternal oral cavity (F. *nucleatum* subsp. *polymorphum*, *vincentii* and *nucleatum*) in contrast with *F. nucleatum* subsp. *animalis*, which was more commonly observed in vaginal samples (Gonzales-Marin et al., 2011). It has been demonstrated that *Fusobacterium* spp. present with unique sequences of the 16S rRNA gene and substantial intragenic heterogeneity that allows species differentiation (Strauss et al., 2008). Although this was also observed for the *F. nucleatum* subspecies (Lawson et al., 1991), they exhibited relatively high levels of sequence similarity. The level of resolution provided by the 16S rRNA gene may not provide enough evidence to help support a translocation theory.

One main difficulty we identified during our previous and present studies was the inability to determine precisely the bacterial subspecies using the different libraries for sequence comparison, as they showed no useful agreement for our purposes. In this study, we present the inaccuracies produced when clone sequences, obtained from known strains of *F. nucleatum* and containing 1290 bp of the 16S rRNA, were compared against three commonly used libraries (GenBank, RDP and HOMD). There is a definite need to increase the existing sequence information and
make it available in all databases to form a broader range of fully characterized \textit{F. nucleatum} strains. Preferably, the whole genome should be obtained to ensure that the different operons are described individually. In this study, only \textit{F. nucleatum} subsp. \textit{nucleatum} was identified by the three libraries, as the genome sequence is available for this strain \cite{Kapatral_2002}, although this was not the case with subsp. \textit{polymorphum} whose genome has also been characterized \cite{Karpathy_2007}. Certainly, bacterial genome data will become accessible exponentially as a result of the developing genome projects and emerging next-generation sequencing; however, as part of these advances, more precise and simplified techniques are also necessary to evaluate the strains of \textit{F. nucleatum} in a greater number of samples, such as in clinical trials. Ideally, these techniques should have the potential of being used as a highly reliable tool to determine the precise origin of the bacteria.

Using the ITS region, several studies have investigated species/subspecies and phylogenetic relationships from a range of important clinical isolates including the genus \textit{Fusobacterium} \cite{Christensen_1999, Narayanan_2001, Conrads_2002, Sadeghifard_2006}. In general, these studies demonstrated a markedly higher heterogeneity of this region compared with the 16S rRNA gene to differentiate closely related taxa. The mosaic pattern of the ITS described for other genera/species has revealed that the presence or absence of one or more regions (also known as block types) were the reason for the length differences \cite{Sadeghifard_2006}. Similar patterns were observed in this study for all the \textit{F. nucleatum} subspecies, where the operon’s variability was based mainly on the presence or absence of these block types.

The expected size of the fragment amplified in this study was in accordance with that observed by Conrads \textit{et al.} \cite{Conrads_2002} (~1860–2020 bp). However, our data differed with respect to the band pattern described by them, in which three to four bands for \textit{F. nucleatum} subsp. \textit{nucleatum}, three bands for the \textit{F. nucleatum} subsp. \textit{animalis}, \textit{fusiforme} and \textit{vincentii} and one band for \textit{F. nucleatum} subsp. \textit{polymorphum} were observed. In their study, using different primers and conditions, the band sizes were calculated by approximation in an agarose gel and, as they suggested, probably only the shorter fragments were selectively amplified (lengths varied between 121 and 151 bp). Calculations of the fragment length in a gel have obvious disadvantages. The data presented in this study revealed more accuracy, as the lengths were confirmed by sequencing. To avoid controversy.

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|}
\hline
\textbf{\textit{F. nucleatum} subspecies} & \textbf{Long} & \textbf{Medium} & \textbf{Short} \\
\hline
& A1 & A2 & B & C1 & C2 \\
\hline
\textit{polymorphum} & - & - & AATAA, AATGT* & AATAA & - \\
\textit{nucleatum} & AATAA & - & - & AATAA & - \\
\textit{animalis} & AATGT* & - & - & AATGT* & - \\
\textit{fusiforme} & TATAA & - & - & TATAA, AAGAA & - \\
\textit{vincentii} & - & AAAAA & - & - & AAAAA \\
\hline
\end{tabular}
\caption{Sequences of the highly variable segment of the ITS (nt 28–32) amplified from the five strains of \textit{F. nucleatum} and arranged by fragment length.}
\end{table}

The sizes of the ITS were determined as: long: A1, 331 bp; A2, 322 bp; medium: B, 222 bp; and short: C1, 156 bp; C2, 147 bp. –, Sequence not present.

*AATGT was observed in both the long and short fragments of \textit{F. nucleatum} subsp. \textit{nucleatum}, but only in the medium-sized fragment of \textit{F. nucleatum} subsp. \textit{polymorphum}.
regarding the start and end points of the ITS region, in our study this was calculated by alignment with the ITS regions available for all the operons of the F. nucleatum subsp. nucleatum ATCC 25586 strain.

Case study 10 was analysed thoroughly in this report by sequence comparison of the F. nucleatum clones from the three sites investigated. Based on previous case reports (Han et al., 2006, 2010), a 100 % match was expected between the site of origin and the distant site. However, our results did not demonstrate a complete match; instead a number of variations were observed between the sequences. There could be different reasons for the variations. One possibility could be due to genetic changes within strains, as the 16S–23S ITS has a non-coding function and therefore is less susceptible to selection pressure and so is able to accumulate a higher percentage of mutations when compared with rRNA (Tyrrell et al., 1997). However, a certain degree of variation was also observed in the more conserved 16S and 23S rRNA gene fragments also amplified by us. Another possibility may be due to PCR or sequencing bias and chimaeric formation molecules (Wang & Wang, 1997). This was shown to be unlikely, as a complete match was observed when comparing the sequences obtained in this study from F. nucleatum subsp. nucleatum ATCC 25586 with the published information. Finally, the low number of strains available for investigation represented a limitation for this study. Regardless of not finding a 100 % match between the sites, when we compared the 5 bp strain-specific sequence among the three samples, the fragment contained in the gastric aspirate was only observed in the oral sample for this case study. It is possible that this fragment is not constant at the subspecies level, as other fragments were also observed in the clinical samples, which suggests that these could correspond to other strains. There is a need to identify and evaluate other F. nucleatum strains. Nonetheless, this fragment was shown to be persistent within the operons of the same strains.

The number of clones analysed in total for case study 10 was 52, which signified a relatively costly method. This would not be a practical approach for a clinical trial and is the main reason for the necessity for a simplified technique. ITS

<table>
<thead>
<tr>
<th>Sample</th>
<th>ITS region fragment length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Long</td>
</tr>
<tr>
<td>A1</td>
<td>B</td>
</tr>
<tr>
<td>GA10</td>
<td>AATAA (Fnp)</td>
</tr>
<tr>
<td>OC10</td>
<td>TATAT</td>
</tr>
<tr>
<td>V10</td>
<td>ACAAT</td>
</tr>
</tbody>
</table>

Table 3. Sequences of the highly variable segment of the ITS (nt 28–32) amplified from the samples obtained from case study 10 and arranged by fragment length

Table 4. Prevalence of F. nucleatum subsp. polymorphum in neonatal gastric aspirates and the respective maternal oral and vaginal samples per case study and relevant clinical characteristics

+ , Positive detection; −, negative detection; NA, sample not available.

<table>
<thead>
<tr>
<th>Case no.</th>
<th>Prevalence of F. nucleatum subsp. polymorphum</th>
<th>Clinical characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neonatal gastric aspirates</td>
<td>Oral cavity</td>
</tr>
<tr>
<td>1</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>NA</td>
<td>+</td>
</tr>
<tr>
<td>10</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>16</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>18</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>53</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>54</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>55</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>62</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>162</td>
<td>–</td>
<td>+</td>
</tr>
</tbody>
</table>

*Periodontal disease is described as follows: total number of sites with pockets (>4 mm), maximum depth of the pockets.
sequence variability has been exploited previously in a case report of a 39-week stillbirth to determine the origin of *F. nucleatum* (Han et al., 2010). *F. nucleatum* was demonstrated for the first time to originate from the maternal oral cavity, as an identical clone was identified in both the placenta/fetus and the mother’s subgingival plaque, but not in her supragingival plaque, vagina or rectum. Nevertheless, the species was identified as *F. nucleatum* subsp. *animalis*, which our findings have suggested to be more likely of vaginal origin. It is possible that the same strain might have been present in the vagina or gut at levels that were undetectable using the end-point PCR. We have observed previously that, when using a more sensitive quantitative PCR approach, *F. nucleatum* was amplified from samples that were initially negative with end-point PCR (Gonzales-Marin et al., 2011). Hence, it was important to ensure a high detection limit by using a nested PCR technique.

A primer specific for the given strain of *F. nucleatum* subsp. *polymorphum* was constructed. The reason for focusing on *F. nucleatum* subsp. *polymorphum* was based on the higher percentage of cases identifying this subspecies in our previous study, and this was confirmed using the 5 bp strain-specific fragment of case study 10, a sample known to contain high levels of *F. nucleatum*. *F. nucleatum* subsp. *polymorphum* has also been reported as the cause of a fatal sepsis in a patient with macroglobulinaemia (Goldstein et al., 1995) and has demonstrated stronger adhesion to collagen compared with other strains (Xie et al., 1991), supporting the possibility that this opportunistic strain can be involved in pregnancy complications. The fragment amplified to study the *F. nucleatum* subsp. *polymorphum* strain could be used similarly to detect other strains, with the exception of *F. nucleatum* subsp. *vincentii* and *fusiforme*, which both presented the same variable sequence (AAAAA). However, there is ongoing discussion with regard to classifying these as the same subspecies (Dzink et al., 1990; Gharbia & Shah, 1992; Kim et al., 2010), which may explain the homogeneity of this fragment.

Finally, we tested a set of ten neonatal–maternal samples. Primer amplification helped to exclude the possibility that the matching clones were not amplified by chance during a cloning analysis. The samples were analysed previously on the basis of the 16S rRNA gene and were known to contain *F. nucleatum*. It was observed that, in two cases, *F. nucleatum* subsp. *polymorphum* was present in the neonatal sample. In both cases, the strain was also present in the respective maternal oral sample but not in the vaginal sample. Both were cases of preterm birth, and in both cases, the mothers were observed to present one and four sites, respectively, with probing depths of 4–5 mm, as observed in a full-mouth periodontal chart. Interestingly, a negative result was observed for the presence of *F. nucleatum* subsp. *polymorphum* in vaginal samples for all the cases included in this study, even those reporting the practice of oro-genital contact during pregnancy, excluding the possibility that this pathogen may be coming from the partner’s oral cavity through oral–genital transmission.

As suggested previously by Gharbia et al. (1990), *F. nucleatum* subsp. *polymorphum* is more commonly observed in periodontally healthy sites, whilst *F. nucleatum* subsp. *nucleatum* is observed mainly in disease. Our findings support the hypothesis that destructive periodontal disease may not be strictly necessary for bacterial translocation to occur (Han et al., 2006, 2010). It is likely that the sole presentation of gingival inflammation, as observed frequently in pregnancy gingivitis, may suffice to allow opportunistic pathogens such as *F. nucleatum* to translocate and invade the amniotic cavity. However, the presence of periodontitis or the prevalence of bacteria in the maternal oral cavity does not necessarily indicate that a pregnancy will result in a complication. Therefore, a link with an additional susceptibility component should be also evaluated. Finally, studies investigating *F. nucleatum* associated with amniotic fluid infection and pregnancy complications should consider possible subspecies/strain variability. Interestingly, the evidence accumulated by a group in Cleveland, USA, (Han et al., 2000, 2004; Liu et al., 2007; Ikegami et al., 2009) strongly demonstrated that *F. nucleatum* can invade the amniotic tissues and induce stillbirth and fetal death in mice. These investigations were often carried out with wild-type *F. nucleatum* 12230, which is NCBI registered as *F. nucleatum* subsp. *polymorphum*.

In conclusion, the highly variable fragment in the ITS region analysed in a more sensitive nested PCR technique could be used to simplify bacterial identification and the determination of the origin of *F. nucleatum* in a clinical study. Specifically, *F. nucleatum* subsp. *polymorphum* should be further evaluated in order to determine its association with adverse pregnancy outcomes.

**ACKNOWLEDGEMENTS**

We would like to thank the UNESCO–L’OREAL partnership for supporting this study through the International Fellowship For Women-In-Science awarded to C. G. Consumables for this study were supported with departmental funds. Special thanks to the staff of the hospital’s laboratory, neonatal and maternity wards at Barts and The London NHS Trust for their support, especially to Dr Michael Millar, Dr Stephen Kempley and Miss Anita Sanghi.

**REFERENCES**


