Effect of alkaline growth pH on the expression of cell envelope proteins in *Fusobacterium nucleatum*

Peter S. Zilm,1 Alex Mira,2† Christopher J. Bagley3‡ and Anthony H. Rogers1

1Oral Microbiology Laboratory, School of Dentistry, The University of Adelaide, Adelaide 5005, Australia
2Department of Genomics and Health, Center for Advanced Research in Public Health (CSISP), Valencia, Spain
3Adelaide Proteomics Centre, Hanson Institute and affiliate of the School of Medicine, The University of Adelaide, Adelaide, Australia

*Fusobacterium nucleatum* is a Gram-negative anaerobic organism that plays a central role in the development of periodontal diseases. The progression of periodontitis is associated with a rise in pH of the gingival sulcus which promotes the growth and expression of virulence factors by periodontopathic bacteria. We have previously reported that the expression of specific cytoplasmic proteins is altered by a shift in growth pH. In the present study we have compared cell envelope protein expression of *F. nucleatum* during chemostat growth at pH 7.2 and 7.8. From a total of 176 proteins resolved from the cell envelope, 15 were found to have altered expression in response to an increase in growth pH and were identified by MS. Uregulated proteins included an outer membrane porin which has been identified as playing a role in virulence, a periplasmic chaperone which assists in the folding of outer membrane proteins, and a transporter thought to be involved with iron uptake. Proteins downregulated at pH 7.8 were consistent with our previous findings that the bacterium reduces its catabolism of energy-yielding substrates in favour of energy-storage pathways. Among the downregulated proteins, two transporters which are involved in the uptake of C4 dicarboxylates and phosphate were identified. A putative protease and an enzyme associated with the metabolism of glutamate were also identified. A high proportion of the cell envelope proteins suggested by these data to play a role in the organism's response to alkaline growth pH may have arisen by lateral gene transfer. This would support the hypothesis that genes that provide an ability to adapt to the changing conditions of the oral environment may be readily shared between oral bacteria.

INTRODUCTION

Periodontal diseases are widespread and do not have a single causative microbial agent. Periodontitis is, however, associated with the proliferation of the anaerobic pathogenic species that belong to the so-called ‘red complex’ (Socransky et al., 1998). These organisms normally inhabit healthy subgingival sites within the oral cavity, but the allogenic factors that cause an increase in their proportions during the transition from health to disease are poorly understood. It is believed that, in the early stages, a combination of host factors and the presence of *Fusobacterium nucleatum* and *Prevotella intermedia* leads to the alkalization of the gingival sulcus which is conducive to the proliferation of more acid-sensitive (red complex) bacteria (Socransky & Haffajee, 2005; Takahashi, 2003). Resolving the magnitude of the pH increase in the gingival crevice has proved inconclusive, as Vroom et al. (1999), using pH-sensitive dyes, showed that plaque organisms in a biofilm create a ‘mosaic of micro-

**Abbreviations:** ACN, acetonitrile; CE, cell envelope; 2DE, 2D electrophoresis; ESI, electron spray ionization; FA, formic acid; HGR, habitat gene reservoir; HGT, horizontal gene transfer; IPG, immobilized pH gradient; TBP, tributyl phosphine.

A supplementary figure, showing estimated ribosome-binding strengths for gene FN0689 and its two flanking genes, and two supplementary tables, showing up- or downregulated proteins from the cell envelope of *Fusobacterium nucleatum* following chemostat growth at pH 7.8 relative to pH 7.2, and potential ecoparalogues in the *F. nucleatum* ATCC 25586 genome, are available with the online version of this paper.
environments’ which have differing pH values. Gradients in the biofilm are not uniform and adjacent zones, with significantly different pHS, are detected. Bickel & Cimasoni (1985) showed that the pH in periodontal pockets may be as high as 9.0, but the interaction of reducing compounds with the metallic antimy pH electrodes used may have led to artefactual results (Eggert et al., 1991). Using glass micro-electrodes, Eggert and co-workers reported that 7% of subjects with periodontal disease had sites measuring pH 7.6 or higher (Eggert et al., 1991), and concluded that there was no correlation between increased pH and periodontal pockets compared with healthy sites. However, metabolism of proteins derived from gingival crevicular fluid by plaque inhabitants would be expected to promote the alkalinization of the environment, a notion supported by the findings of Marsh et al. (1993) that the growth pH optimum of Porphyromonas gingivalis is pH 7.5 and that alkaline pH is also associated with an increased proteolytic activity of the organism.

As with many other oral bacteria of importance, the genome of F. nucleatum has been sequenced (Kapral et al., 2002, 2003; Karpathy et al., 2007), and this has facilitated the investigation of the proteome in response to environmental changes. We have previously investigated the response of F. nucleatum to changes in external pH (Zilm et al., 2007) and reported the regulation of cytoplasmic proteins associated with nutrient metabolism and energy storage. We have also demonstrated that, at a growth pH above 8.2, F. nucleatum co-adheres and forms homogeneous biofilms (Zilm & Rogers, 2007). These investigations suggest that the organism has a coordinated response to an alkaline environment.

An organism’s regulation of cell envelope (CE) proteins in response to environmental stresses has been termed the extracytoplasmic stress response (ESR) (Rowley et al., 2006). The ESR to environmental alkalinization has, to our knowledge, not been investigated in F. nucleatum and could be considered the first step in understanding the signal transduction pathways which lead to changes in gene expression.

The sequencing of the genomes of oral bacteria has also led to the investigation of niche-specific genes which are the result of lateral transfer between species. These genes have been described as the habitat gene reservoir (HGR), and are retained if they increase the recipient’s adaptation to the environment (Legault et al., 2006). In contrast to other oral organisms, between 20 and 45% of F. nucleatum genes are reported to have been acquired by horizontal gene transfer (HGT), maybe as a result of its ability to aggregate with many plaque inhabitants (Mira et al., 2004). The transcription and subsequent translation of genes into functional proteins, however, depends on the taxonomic ‘distance’ between donor and recipient and on the ‘new’ genomic context of the recipient, which may have a different set of promoters, ribosome-binding sites and regulatory regions (Mira, 2008; Sorek et al., 2007). In terms of functionality, it is therefore important to demonstrate protein expression of laterally transferred genes associated with the HGR.

Another challenge associated with the study of gene expression in response to a single change in environmental conditions has been to eliminate the influence of other confounding factors. In some of the early and most insightful studies on bacterial ecology in dental plaque, continuous culture was used in an attempt to simulate in vivo growth conditions (Ellwood & Hunter, 1976; Marsh et al., 1993, 1994; McDermid et al., 1988; McKee et al., 1986; Rogers et al., 1986, 1991, 1992; Tempest, 1969). The post-genomic era has seen the resurgence of this method for the culture of bacteria and its advantages for functional genomic studies have been reviewed (Hoskisson & Hobbs, 2005).

In the present study, we aim to provide an insight into how F. nucleatum copes with the alkalinization of the gingival sulcus by examining the change in expression of CE proteins. We have also undertaken a genomic comparison of the genes found to be regulated by external pH to establish whether they also constitute part of the HGR shared by other plaque inhabitants.

**METHODS**

**Materials and reagents.** Amino acids, endonucleases and Bacterial Protease Inhibitor Cocktail were purchased from Sigma Aldrich, and DeStreak reagent from GE Healthcare. All SDS-PAGE and 2D electrophoresis (2DE) chemicals, solubilization buffer, immobilized pH gradient (IPG) strips, IEF equipment, densitometer and image analysis software (PD-Quest 7.2) used for proteomic analysis were purchased from Bio-Rad Laboratories. Brain heart infusion broth and anaerobic blood agar plates were purchased from Oxoid.

**Micro-organism.** F. nucleatum ATCC 10953, subspecies polymorphum (Drink et al., 1990), was maintained on anaerobic blood agar plates, and the preparation of starter cultures for chemostat inoculation has been described previously (Zilm et al., 2007).

**Growth media and culture conditions.** The organism was grown under nitrogen limitation in a chemically defined growth medium (CDM) based on that of van der Hoeven et al. (1985). The concentrations of the energy-yielding amino acids glutamic acid (40 mM), lysine and histidine (10 mM each), and glucose (20 mM) were raised to increase growth yields. Thioglycolic acid (0.5 g l⁻¹) was added to maintain a low redox potential, as was Tween 80 (0.2 g l⁻¹) to aid cell dispersion.

F. nucleatum was grown anaerobically using a model C-30 BioFlo Chemostat (New Brunswick Scientific) with a culture volume of 365 ml, as described previously (Rogers et al., 1991). The flow rate of the medium was set at 27.5 ml h⁻¹ (D=0.075 h⁻¹), giving an estimated generation time of 9.2 h, which is typical of natural ecosystems such as dental plaque (Hamilton et al., 1979; Socorsky et al., 1977). Anaerobic growth conditions and sampling of the culture, under steady-state growth conditions, have been described previously (Zilm et al., 2007). Growth pH was maintained at either 7.2 or 7.8 by the automatic addition of 2 M KOH.

**Cell harvesting.** After equilibration, cell culture samples were raised to increase growth yields. Thioglycollic acid (0.5 g l⁻¹) was added to maintain a low redox potential, as was Tween 80 (0.2 g l⁻¹) to aid cell dispersion.

F. nucleatum was grown anaerobically using a model C-30 BioFlo Chemostat (New Brunswick Scientific) with a culture volume of 365 ml, as described previously (Rogers et al., 1991). The flow rate of the medium was set at 27.5 ml h⁻¹ (D=0.075 h⁻¹), giving an estimated generation time of 9.2 h, which is typical of natural ecosystems such as dental plaque (Hamilton et al., 1979; Socorsky et al., 1977). Anaerobic growth conditions and sampling of the culture, under steady-state growth conditions, have been described previously (Zilm et al., 2007). Growth pH was maintained at either 7.2 or 7.8 by the automatic addition of 2 M KOH.

**Cell harvesting.** After equilibration, cell culture samples were harvested over five consecutive days and pooled, as described previously (Zilm et al., 2007).
CE sample preparation for proteomic analysis. Cell lysis in the presence of endonucleases (DNase and RNase) and protease inhibitor cocktail was performed as described previously (Zilm et al., 2007). Disrupted cell suspensions were centrifuged (20 000 g, 30 min at 4 °C) to collect the CE. The CE fraction was then washed twice in sodium phosphate buffer (0.05 M, pH 7.5) and resuspended in an aliquot (1 ml) of either solubilization buffer [5 M electrophoresis grade urea, 2 M thiourea, 40 mM Tris, 2 % CHAPS, 2 % sulfobetaine 3-10 (SB 3-10), 2 mM tributyl phosphine (TBP), 0.2 % (w/v) ampholytes] or SDS gel buffer (Laemmli, 1970). Samples were left for 1 h at room temperature before carefully (to avoid foaming) aspirating the solution with a fine-gauge needle. Samples were then clarified by centrifugation (20 000 g, 60 min at 15 °C) to remove insoluble material and stored at −80 °C.

2DE. Protein quantification was performed using an RC DC protein assay kit (Bio-Rad Laboratories) in accordance with the manufacturer’s instructions. IEF was performed on 11 cm pre-cast IPG strips with pH ranges of 4–7 and 7–10 using a Protein IEF cell. Briefly, 0.35 mg cell protein was cup-loaded (anode end) onto an 11 cm IPG strip which had been passively rehydrated for 20 h in 0.2 ml solubilization buffer containing DeStreak reagent (15 mg ml−1) instead of TBP. IEF was run using a customized method aimed at optimizing protein separation. Briefly, the voltage was gradually increased using linear ramping to 4000 V over 7.5 h as follows: 150 V for 1 h, 300 V for 3 h, 600 V for 1.5 h, 1200 V for 1.5 h, followed by rapid ramping to 4000 and 8000 V over 1.5 and 1 h, respectively. Focusing occurred for 8000 V h with a 50 μA per strip current limit and the temperature was maintained at 20 °C. After 8000 V had been achieved, the strips were maintained at 500 V until the IEF strips were removed and placed at −20 °C. Between three and six replicate IPG strips per sample were run concurrently. After IEF, proteins were separated in the second dimension, as previously described (Zilm et al., 2007).

1D electrophoresis (SDS-PAGE). CE proteins solubilized in SDS buffer were boiled for 90 s and allowed to cool before being separated by SDS-PAGE using discontinuous gels (12 % T 3.3 % C resolving gel contained a minimum of three gels. Analysis sets containing proteins were extracted from the gel pieces with 50 % acetonitrile (ACN), trypsin (sequencing grade, Promega) per sample. Tryptic peptides were digested for 3 h at 37 °C. The peptides were then diluted to 5 μl with FA30 (seven parts 0.1 % FA, three parts ACN).

MALDI MS (MS and MS/MS). One microlitre of each digested protein sample was applied to a 600 μm AnchorChip (Bruker Daltonic) following the method of Zhang et al. (2005). MALDI-TOF mass spectra were acquired at random locations over the matrix surface spot using a Bruker Ultraflex III MALDI-TOF/TOF mass spectrometer (Bruker Daltonic) in reflectron mode and controlled by flexControl software (version 3, Bruker Daltonic). External calibration was performed using peptide standards (Bruker Daltonic). Between three and six of the most highly abundant sample ions (not trypsin or keratin) were selected for MALDI-TOF/TOF analysis, which was performed in LIFT mode using the same spot on the target. Mass spectra data acquired by MS and MS/MS were analysed using flexAnalysis (version 3, Bruker Daltonic), and then exported to BioTools (version 3.1, Bruker Daltonic), and the MS and corresponding MS/MS data were combined and used to interrogate an in-house Mascot database search engine (version 2.2, Matrix Science; http://www.matrixscience.com) using the following parameters:

- **Taxonomy:** Eubacteria
- **Database:** NCBI non-redundant 20080214
- **Enzyme:** trypsin
- **Fixed modifications:** carboxamidomethyl
- **Variable modifications:** oxidation of Met
- **Mass tol. MS:** 50 p.p.m.
- **MS/MS tol.:** 0.5 Da
- **Missed cleavages:** 1
- **Protein identification:** based upon the MOWSE and probability scores generated by the software. Based on the combined MS/MS data, samples that returned a positive ‘hit’ were submitted independently to Mascot. Predicted molecular masses and isoelectric points of identified proteins were calculated using Compute pI/Mw from the Expert Protein Analysis System (http://au.expasy.org/tools/pi_tool.html).

Liquid chromatography-electron spray ionization (LC-ESI) MS. Samples which did not give sufficient spectra using MALDI MS for accurate protein identification were further analysed using LC-ESI ion trap MS. Samples (2.5 μl) were diluted to 6 μl with 3 % ACN and 0.1 % FA, and 5 μl of each sample was introduced into an Agilent Protein ID Chip column assembly (40 nl trap column with 0.075 × 43 mm C-18 analytical column) housed in an Agilent HPLC-Chip Cube Interface which was connected to an HCTultra 3D ion-trap mass spectrometer (Bruker Daltonic). After the column had been equilibrated with 4 % ACN/0.1 % FA at a flow rate of 0.5 μl min−1, the samples were eluted over 32 min with an ACN gradient (4–31 %).

Ionizable species (300 < m/z < 1200) were trapped, and one or two of the most intense ions eluted were fragmented by collision-induced dissociation. Peak detection of MS and MS/MS spectra was achieved using DataAnalysis software (version 3.4, Bruker Daltonic) and the data were imported into BioTools. MS/MS data were searched as described above, but with an MS mass tolerance and MS/MS tol. of 0.3 and 0.4 Da, respectively, and a peptide charge of 1+, 2+ and 3+.


Prediction of bacterial protein subcellular localization (Gardy et al., 2005) was derived from http://www.psort.org/psortb/index.html using the Gram-negative option. Protein localization was based on scores out of 10.

Predicted β-barrel integral outer membrane proteins (BOMP) (Berven et al., 2004) were identified (http://www.bioinfo.no/tools/bomp) with an E-value of less than 1e−10.
Comparative genomic analysis. The presence of paralogous genes was determined by amino acid sequence similarity among all proteins in *F. nucleatum*, considering that two ORFs were part of a gene family if (i) their BLAST E-value was lower than $10^{-5}$, (ii) the amino acid identity was higher than 30 %, and (iii) the sequence alignment was at least 60 % of the total protein length, following the method of Pushker et al. (2004).

Potential HGT events were detected by a combination of three methods based on sequence similarity, phylogenetic trees and DNA compositional features (Mira, 2008).

Sequence similarity. Protein sequences of *F. nucleatum* subsp. *nucleatum* ATCC 25586 were retrieved from ftp://ftp.ncbi.nih.gov/genomes/Bacteria and a BLASTP (Altschul et al., 1997) search of each protein was performed against the non-redundant database. Top hits for each protein were recorded after filtering those with E-values $>10^{-5}$ and sequence identity and length lower than 30 and 50 %, respectively. All recorded hits were then assigned to one of eight categories as belonging to the Bacteroides, Firmicutes, $\alpha$-, $\beta$-, $\gamma$-Proteobacteria, $\delta$-, $\varepsilon$-Proteobacteria, Spirochaetes, other Bacteria, Archaea or Eukaryotes/no hit. Hits to Firmicutes (the group to which *Fusobacterium* appears to be most closely related) were refined by further BLASTP analysis between *F. nucleatum* and all sequenced bacteria available from this group. If the gene had a homologue in only one genus from all the available low-GC Gram-positive species, it was considered an HGT event from/to this group. If it was present in more than one genus it was considered vertically inherited and consistent with the ribosomal phylogeny.

Phylogenetic trees. For each *F. nucleatum* gene, the protein sequences of up to 50 best BLAST hits (E-value $<10^{-5}$) were retrieved. All sequences were then automatically aligned with CLUSTAL W (Thompson et al., 1994) with default parameters. A neighbour-joining tree with 1000 bootstrap replicates was generated from the resulting alignment. Phylogenetic incongruence was detected in well-resolved trees (trees should have more than four hits), and bootstrap values at the nodes chosen for a decision on taxonomic assignment should be $>500$ following the protocol of Mira et al. (2004).

DNA compositional features. Significant deviations in codon usage, amino acid composition and G+C content were considered evidence of potential HGT (Garcia-Vallve et al., 2000). Significant differences in oligonucleotide use between a given region and the rest of the genome were detected using the OligoWeb resource (http://insilico.ehu.es/oligoweb) and also considered indicative of potential HGT.

RESULTS

Growth of *F. nucleatum* in a chemostat

We have previously reported changes in cellular yields and nutrient utilization by *F. nucleatum* when grown at pH 7.2 and 7.8 in a chemostat (Zilm et al., 2007). Briefly, optimal cellular yields (700 mg protein ml$^{-1}$) were obtained at pH 7.4, while growth at pH 7.2 and 7.8 resulted in decreased cellular yields of 10 and 13 %, respectively. Growth above pH 7.2 was also characterized by the consumption of all exogenous nutrients from which *F. nucleatum* derives energy (Rogers et al., 1991).

2DE of CE proteins expressed by *F. nucleatum*

Representative 2DE and SDS-PAGE gels are shown in Fig. 1. The separation of proteins by SDS-PAGE (Fig. 1a) proved inadequate in resolving all proteins, and better resolution was achieved with 2DE (Fig. 1b, c).

Based on staining with Coomassie blue R250, a total of 176 protein ‘spots’ were used for 2DE analysis, and included
proteins within the isoelectric point (pI) range of 4–10 and a molecular mass of 10–80 kDa. In silico representation of the F. nucleatum proteome indicates that the majority of cellular proteins are within this range (Zilm et al., 2007). From a total of ~2361 proteins (subsp. polymorphum), 272 proteins have been annotated as transporters and ~20 % of all genes (subsp. nucleatum) are translated into proteins associated with the CE (http://www.oralgen.lanl.gov/).

The degree of solubility of the CE fraction was also investigated. Of the total protein associated with the CE, 11.5 mg was initially solubilized in 2DE buffer as described above. Protein assay of the insoluble proteins, collected by centrifugation, revealed that ~78 % of the protein was solubilized (data not shown).

Proteins regulated by alkaline pH

Gels were arranged into two matchsets (pI 4–7 and 7–10) containing three to five replicate gels from each growth pH. Gels were normalized using the total density of all detected spots in the matchset gels. The correlation coefficient between all member gels ranged from 0.69 to 0.86. Proteins were considered to be regulated by pH if, after normalization, the average mean spot density between replicate groups of gels representing growth at pH 7.2 was significantly ($P<0.01$ or $0.01>P<0.05$) different from those representing growth at pH 7.8 (Supplementary Table S1). Quantification was based on a synthetically derived, Gaussian image of the matchset member gels, and 16 proteins were identified as showing significant regulation (Table 1) in response to growth pH. Fig. 2 shows the relative position of each regulated protein, and spot quantities (shown in the margins) provide a qualitative assessment of the level of up- or downregulation and the standard error between each group.

Identification of regulated proteins

Regulated protein spots were excised from the 2DE gels and, following in-gel digestion, were initially identified by MALDI-TOF MS (Table 1). If MOWSE (molecular mass search) scores were below or close to the cut-off threshold, individual peptide ions were analysed using ion-trap MS. Protein identifications were also supported by comparing observed and predicted pIs and molecular masses. All proteins showed relatively small variations in molecular masses and pIs (Supplementary Table S1). Due to the sensitivity of the MS analysis, some spots did not give a single positive identification (Table 1), and two protein spots (spot pairs, SSP3202/3201) were identified as the same protein having two locations on the gel (Fig. 2). The relative proportion of proteins within spots containing more than one protein was determined by calculation of the Exponentially Modified Protein Abundance Index (emPAI) score (Supplementary Table S1).

Two programs (PSORTb and BOMP) were used to predict the location of proteins in the CE. Spot 3802 was the only regulated protein that contained a single β-barrel structure (BOMP analysis), representative of proteins found in the outer membrane (data not shown). PSORTb analysis of regulated proteins predicted that six proteins (including 3802) were associated with the CE (Table 1). Of these, two proteins (5204 and 3301) were predicted to be located in the bacterial periplasm and one (3104) was predicted to be associated with the cytoplasmic membrane. The remaining four proteins (0202, 0203, 1102, 3002) were predicted to be associated with the CE, but PSORTb was unable to further resolve their location. The relative hydrophobicity of regulated proteins was determined by their GRAVY scores (Table 1). Two proteins (3101 and 3104) with positive GRAVY values were considered hydrophobic, while the remaining proteins had negative GRAVY values and were thus considered hydrophilic.

In response to an increase in growth pH to 7.8, five proteins (0202, 0203, 3101, 3202/3201 and 3802) were upregulated. Of these, spots 0202, 0203, 3101 and 3802 were upregulated more than threefold, and the remaining proteins showed a minimum twofold and maximum 10-fold downregulation in response to growth at pH 7.8 (Table 1).

Analysis of the F. nucleatum (subsp. nucleatum) proteome shows ~20 % of proteins as 'hypothetical', with unknown functions (http://www.oralgen.lanl.gov/).

Table 1 shows that seven of the proteins (42 %) were identified as being 'hypothetical', but are now shown to be true proteins and associated with the organism’s coordinated response to increasing external pH.

Comparative genomic analysis

The proteins shown to respond to alkaline growth pH were investigated further to determine whether the genome contained paralogous sequences (i.e. genes showing significant sequence similarity that could be involved in related functions) and whether they had been acquired by HGT (Table 2). Many trees were difficult to resolve because of a low phylogenetic signal or because the protein gave an insufficient number of hits in public databases. Among those that could be resolved, only FN1439 and FN1549 showed a phylogeny congruent with vertical inheritance, and several indicated clear transfers from phylogenetically unrelated species (Table 2). These genes also showed signs of HGT in their sequence composition, in agreement with previous work that indicates that CE genes are over-represented among potential HGT events in F. nucleatum (Mira et al., 2004). Genes FN1252 and FN1856 appeared to be transferred to/from other oral species, despite the low number of oral bacteria currently sequenced.

The paralogues analysis showed that four of the 15 proteins under study had related ORFs in the genome. Some of them are proteins of unknown function, and their sequence similarity to the ones identified in the present work would therefore suggest that they are also located in the CE. It has
Table 1. Up- or downregulated proteins isolated from the CE of *F. nucleatum* following growth in a chemostat at pH 7.8 relative to pH 7.2

<table>
<thead>
<tr>
<th>SSP*</th>
<th>Mean spot quantity ratio†</th>
<th>P†</th>
<th>GRAVY score§</th>
<th>PSORTb location</th>
<th></th>
<th>NCBI accession no.</th>
<th>Gene ID#</th>
<th>Protein identification**</th>
</tr>
</thead>
<tbody>
<tr>
<td>0202</td>
<td>3.2</td>
<td>9.1×10⁻⁶</td>
<td>-0.77</td>
<td>Non-cytoplasmic</td>
<td>1483236202</td>
<td>FNP_0033</td>
<td>Hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>0203</td>
<td>44.7</td>
<td>1.4×10⁻¹⁸</td>
<td>-0.29</td>
<td>Non-cytoplasmic</td>
<td>148323088</td>
<td>FNP_0530</td>
<td>34 kDa membrane antigen precursor</td>
<td></td>
</tr>
<tr>
<td>1001</td>
<td>0.5</td>
<td>9.0×10⁻⁷</td>
<td>-0.37</td>
<td>Unknown</td>
<td>148323877</td>
<td>FNP_1344</td>
<td>Hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1102</td>
<td>0.5</td>
<td>4.8×10⁻⁷</td>
<td>-0.82</td>
<td>Non-cytoplasmic</td>
<td>148322846</td>
<td>FNP_0281</td>
<td>Hypothetical membrane-spanning protein</td>
<td></td>
</tr>
<tr>
<td>2101</td>
<td>0.5</td>
<td>1.8×10⁻⁸</td>
<td>-0.67</td>
<td>Cytoplasm</td>
<td>148322305</td>
<td>FNP_2161</td>
<td>Hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>2505‡‡</td>
<td>0.2</td>
<td>3.1×10⁻⁸</td>
<td>-0.20</td>
<td>CE</td>
<td>148323554</td>
<td>FNP_1008</td>
<td>Hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>3502</td>
<td>0.54</td>
<td>9.2×10⁻⁸</td>
<td>-0.30</td>
<td>Periplasm</td>
<td>148323082</td>
<td>FNP_0524</td>
<td>Dicarboxylate:proton TRAP-T family tripartite ATP-independent transporter binding protein</td>
<td></td>
</tr>
<tr>
<td>3002‡‡</td>
<td>0.3</td>
<td>3.4×10⁻⁴</td>
<td>-0.86</td>
<td>Non-cytoplasmic</td>
<td>148322844</td>
<td>FNP_0279</td>
<td>Hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>3101‡‡</td>
<td>6.7</td>
<td>3.1×10⁻³</td>
<td>0.014</td>
<td>Cytoplasm</td>
<td>148323151</td>
<td>FNP_0594</td>
<td>Hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>3104‡‡</td>
<td>0.5</td>
<td>3.1×10⁻³</td>
<td>0.19</td>
<td>Cytoplasmic membrane</td>
<td>148323515</td>
<td>FNP_0969</td>
<td>Butyrate-acetoacetate CoA transferase subunit B</td>
<td></td>
</tr>
<tr>
<td>3202</td>
<td>1.9</td>
<td>8.1×10⁻⁴</td>
<td>-0.68</td>
<td>Cytoplasm</td>
<td>148322857</td>
<td>FNP_0293</td>
<td>Pepidyl prolly cis-trans isomerase</td>
<td></td>
</tr>
<tr>
<td>3201</td>
<td>2.23</td>
<td>1.5×10⁻⁴</td>
<td>-0.68</td>
<td>Cytoplasm</td>
<td>148322857</td>
<td>FNP_0293</td>
<td>Pepidyl prolly cis-trans isomerase</td>
<td></td>
</tr>
<tr>
<td>3301‡‡</td>
<td>0.1</td>
<td>7.3×10⁻⁸</td>
<td>-0.3</td>
<td>Periplasm</td>
<td>148323082</td>
<td>FNP_0524</td>
<td>Dicarboxylate:proton (H⁺) TRAP-T family tripartite, ATP-independent transporter binding protein, C4 dicarboxylate binding protein</td>
<td></td>
</tr>
<tr>
<td>4302‡‡</td>
<td>0.4</td>
<td>2.8×10⁻³</td>
<td>-0.07</td>
<td>Cytoplasm</td>
<td>14832296</td>
<td>FNP_2151</td>
<td>Stomatin-like protein</td>
<td></td>
</tr>
<tr>
<td>5204</td>
<td>0.2</td>
<td>4.0×10⁻²</td>
<td>-0.01</td>
<td>Periplasm</td>
<td>148323210</td>
<td>FNP_0657</td>
<td>Phosphonate ABC superfamily ATP binding cassette transporter binding protein</td>
<td></td>
</tr>
<tr>
<td>3802</td>
<td>3.2</td>
<td>4.1×10⁻²</td>
<td>-0.44</td>
<td>Outer membrane</td>
<td>148323388</td>
<td>FNP_2196</td>
<td>OmplIP family outer membrane porin</td>
<td></td>
</tr>
</tbody>
</table>

*Standard SPot numbers assigned by PDQuest analysis software; refer to Fig. 2.
†Ratio of protein quantity (after normalization) for each protein identified from *F. nucleatum* grown at pH 7.8 compared with growth at pH 7.2 (mean of 4–5 replicates).
§Significance level using Student’s *t* test.
||Predicted subcellular location of identified bacterial proteins assigned using PSORTb v2.0 (http://www.psort.org/). For CE, expected location includes periplasm, outer membrane and cytoplasmic membrane.
¶National Center for Biotechnology Information.
#Gene identification code from genome-sequencing project (http://www.oralgen.lanl.gov/).
**Protein name or description of function from annotations in the NCBI database.
‡‡Proteins identified using MALDI-TOF/TOF. All others were identified using ESI ion trap MS.

Recently been shown that paralogous proteins are in most cases very divergent, with only a few cases above 50% amino acid identity (Sanchez-Perez et al., 2008). These instances of highly similar paralogues have been related to pairs of proteins performing the same function under different environmental conditions. The presence of these so-called ‘ecoparalogues’ is more frequent in species undergoing severe environmental fluctuations and is more apparent among CE proteins (which are extracellularly exposed) (Sanchez-Perez et al., 2008). It is therefore interesting that two potential pairs of ecoparalogues were found among the proteins under study (Table 2). Given that fluctuations in external pH could influence the functionality of CE proteins, a list of highly similar paralogues in the *F. nucleatum* genome was identified. This shows that there are a significant number of potential ecoparalogues, that many of them display important differences in pI, and that they may be localized in the CE (Supplementary Table S2).

**DISCUSSION**

The recent genomic sequencing of *F. nucleatum* subsp. *polymorphum* (FNP) has confirmed the phenotypic heterogeneity seen amongst *F. nucleatum* subspecies and casts serious doubt on the validity of differentiating disease-related subspecies based on the electrophoretic migration...
of two intracellular enzymes (glutamine dehydrogenase and 2-oxoglutarate reductase) (Gharbia & Shah, 1992; Karpathy et al., 2007; Morris et al., 1996). *F. nucleatum* subsp. *polymorphum* represents a separate phylogenetic branch which includes several human pathogens (Conrads, 2002; Gmur et al., 2006; Goldstein et al., 1995). At a species level, its metabolic versatility enables it to create and survive conditions conducive to the establishment of pathogenic obligate anaerobes (Diaz et al., 2002). Increasing alkalinity of the gingival sulcus has been identified as one of the disease-associated conditions potentially initiated by *F. nucleatum* (Takahashi, 2003, 2005). In this study we have examined changes in the CE proteome of *F. nucleatum* subsp. *polymorphum* caused by a change in growth pH from 7.2 to 7.8.

One of the problems when dealing with the separation of CE proteins using conventional proteomic technologies is the relatively poor solubility of hydrophobic membrane proteins, the abundance of which is usually far less than that of hydrophilic cytoplasmic proteins found in whole-cell lysates. In dealing with these problems, we have partially purified the CE of *F. nucleatum* by centrifugation, and the solubility of proteins has been optimized by using a solution containing multiple chaotropes and strong detergents. Unlike our previous study (Zilm et al., 2007), in which we looked at the protein expression of cytoplasmic proteins within the pI range 4–7, we have currently achieved a greater proteome coverage by separating CE proteins within the pI range 4–10 (Zilm et al., 2007).

Parenthetically, it is worth noting that two hydrophobic, pH-regulated proteins (positive GRAVY score, Table 1) were separated using our 2DE protocol. 1D SDS-PAGE has the advantage that most hydrophobic membrane proteins, with the exception of proteins containing many transmembrane (TMR) regions, solubilize in strong ionic detergents such as SDS (Cordwell, 2006). The disadvantage of SDS-PAGE is that the separation of complex mixtures is difficult when determining the level of expression of individual proteins (Fig. 1).

Nevertheless, the separation of complex mixtures of proteins by 2DE does not completely eliminate the potential for multiple proteins to be separated in the same conditions. The Gaussian images (Fig. 2) illustrate the changes in protein expression by comparing the proteomes of *F. nucleatum* grown at pH 7.2 and 7.8.

![Gaussian image of proteins expressed by *F. nucleatum* grown in a chemostat at pH 7.2 and 7.8 and separated by 2D-PAGE using a pI range of 4–7 (a) and 7–10 (b). Numbered spots represent proteins which were significantly regulated by culture pH. The level of expression of numbered spots is represented by quantity graphs shown in the left and right margins. The right-hand columns of each pair represent growth at pH 7.8 and the left-hand columns represent growth at pH 7.2. Results represent the mean of 3–5 replicates.](http://mic.sgmjournals.org)
Table 2. Phylogenetic characteristics of genes regulated by alkalinity

<table>
<thead>
<tr>
<th>SSP accession no.</th>
<th>NCBI Gene ID</th>
<th>Phylogenetic tree*</th>
<th>DNA composition†</th>
<th>Paralogues (%)‡</th>
<th>F. nucleatum subsp. polymorphum homologue (%)‡</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>202</td>
<td>19704771</td>
<td>FN1439 Consistent with 16S–23S rDNA phylogeny</td>
<td>Normal</td>
<td>FN0228 (32)</td>
<td>148322602</td>
<td>Clear native gene</td>
</tr>
<tr>
<td>203</td>
<td>19704587</td>
<td>FN1252 Clear HGT from Proteobacteria</td>
<td>GC content different, excess of proline</td>
<td>148323088 (96)</td>
<td>HGR gene (transferred to many oral species)</td>
<td></td>
</tr>
<tr>
<td>1001</td>
<td>19704024</td>
<td>FN0689 Uncertain</td>
<td>GC3 different</td>
<td>148323877 (90)</td>
<td>Paralogous genes are contiguous and may recombine</td>
<td></td>
</tr>
<tr>
<td>66475006</td>
<td></td>
<td></td>
<td></td>
<td>FN0688</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3101</td>
<td>19704527</td>
<td>FN1192 No tree§</td>
<td>GC1 different</td>
<td>148323151 (94)</td>
<td>Could be not annotated in other genomes due to very short length. No BLAST hits to viral databases</td>
<td></td>
</tr>
<tr>
<td>2505</td>
<td>19705204</td>
<td>FN1899 No tree§</td>
<td>GC content different</td>
<td>FN1590 (90)</td>
<td>148323554 (93)</td>
<td>No BLAST hits to viral databases</td>
</tr>
<tr>
<td>3502</td>
<td>34762380</td>
<td>FNV2041 Present in Fn, Fv, Fp. Low similarity to Proteobacteria and one clostridium</td>
<td>Unusual codon usage and amino acid frequencies</td>
<td>148322305 (84)</td>
<td>Potential transfer</td>
<td></td>
</tr>
<tr>
<td>3002</td>
<td>19703594</td>
<td>FNO249 No tree§</td>
<td>Normal</td>
<td>148322844 (90)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3104</td>
<td>19705161</td>
<td>FN1856 Base of Firmicutes. Top hit to P. gingivalis</td>
<td>GC1 different</td>
<td>148323515 (atoA, 95)</td>
<td>Likely a native gene transferred towards Porphyromonas</td>
<td></td>
</tr>
<tr>
<td>3202</td>
<td>57117487</td>
<td>FN0263 Tree uncertain, but Fn within proteobacterial cluster</td>
<td>GC1 and GC2 different</td>
<td>148322857 (87)</td>
<td>Potential HGT</td>
<td></td>
</tr>
<tr>
<td>3201</td>
<td>3301</td>
<td>19704593 FN1258 Likely HGT from Proteobacteria</td>
<td>Normal</td>
<td>148323082 (97)</td>
<td>Universally distributed protein</td>
<td></td>
</tr>
<tr>
<td>4302</td>
<td>3451881</td>
<td>FN1549 Consistent with 16S–23S rDNA phylogeny</td>
<td>Normal</td>
<td>148322296 (95)</td>
<td>Clear native gene</td>
<td></td>
</tr>
<tr>
<td>5204</td>
<td>19704470</td>
<td>FN1135 Ambiguous. At the base of Firmicutes and Proteobacteria</td>
<td>Normal</td>
<td>148323210 (96)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3802</td>
<td>19705216</td>
<td>FN1911 No tree§</td>
<td>Normal</td>
<td>148322338 (95)</td>
<td>Many repeats belonging to different antigen families</td>
<td></td>
</tr>
<tr>
<td>1102</td>
<td>19703596</td>
<td>FN0251 Ambiguous. More similar to Proteobacteria but gene also present in Firmicutes</td>
<td>GC content different; excess of methionine and glutamic acid</td>
<td>148322846 (93)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Phylogenetic assignment based on a maximum-likelihood tree performed with the 50 BLAST top hits of each gene. Fn, Fv and Fp represent Fusobacterium nucleatum, F. nucleatum subsp. vincentii and F. polymorphum, respectively.

†Significantly different GC content, codon use or amino acid composition is indicated, following Garcia-Vallve et al. (2000).

‡Figures in parentheses represent the percentage amino acid similarity. Genes in bold type indicate potential ecoparalogues (proteins performing the same function optimally under different environmental conditions, such as differences in pH). Blank cells indicate that the gene is a singleton.

§No tree could be drawn because no hits were retrieved in the database outside the Fusobacteria.
‘spot’ (Camposstrini et al., 2005). In the present study, if more than one protein was identified in each spot, the relative abundance of each protein was determined by calculating the emPAI score (Exponentially Modified Protein Abundance Index) from the MS/MS data, and the assumption was made that the most abundant protein would account for the observed regulation (Ishihama et al., 2005). Supplementary Table S1 shows the calculated emPAI scores for each identified protein. In some cases (spots 3502, 3201, 3301, 5204), identified proteins also matched homologues in F. nucleatum subsp. vincentii (FNV) and F. nucleatum (FNN) (Supplementary Table S1). Proteins identified as being regulated by increasing pH are shown in Table 1. With the exception of spot 0203, all regulated proteins matched best to the subspecies polymorphum used in this study. Spot 0203 was unique in that the peptides of the identified protein matched better to the homologue from subspecies nucleatum. The reason for this is unclear but may be a result of the heterogeneous nature of the F. nucleatum genome.

Proteins downregulated by increasing pH
A rise in pH from 7.2 to 7.8 was associated with the downregulation of two transport proteins. One of these, C4 dicarboxylate binding protein (3301), is an extracytoplasmic solute receptor (ECSR)-dependent uptake system which, unlike many prokaryote transport proteins, uses a proton motive force instead of ATP to drive solute uptake. Described as a tripartite ATP-independent periplasmic (TRAP-T) transporter for C4 dicarboxylates (e.g. malate, succinate and fumarate), the transporter has been well characterized in Rhodobacter capsulatus (Forward et al., 1997), and consists of an ECSR and two integral membrane proteins. Homologues of R. capsulatus proteins have been found in many Gram-negative bacteria, and SSP 3301 represents the periplasmic component of the TRAP-T uptake system found in F. nucleatum. The 10-fold downregulation of this protein appears to be correlated with the rise in alkalinity of the environment, and the reduced uptake of C4 dicarboxylates may be related to the general decrease in ATP synthesis and fall in cell yield associated with growth at pH 7.8 (Zilm et al., 2007). Hypothetically, the downregulation may also be associated with a reduced proton gradient, which may help the organism to maintain a neutral cytoplasmic pH during an increase in external pH.

The multiple-component uptake system for the transport of phosphonate (SSP 5204) was also downregulated at pH 7.8. Unlike TRAP-T transporters, this uptake system consists of a periplasmic binding protein, and the translocation of phosphonate across the inner membrane is coupled with ATP hydrolysis. Phosphonate uptake represents one of three ways in which prokaryotes acquire phosphorus, and the downregulation of the phosphonate-binding protein (PBP) is consistent with the decreased ATP synthesis associated with growth at high pH (Rogers et al., 1991). The reduced uptake and dissociation of phosphonates may also contribute to the maintenance of a neutral periplasmic pH, despite an elevated external pH.

The reduced ability to acquire substrates required for ATP synthesis and growth at pH 7.8 was also reflected in the 2.5-fold downregulation of the stomatin-like protein (SSP 4302) (Table 1). Annotated as an integral membrane protease (http://www.oralgen.lanl.gov/), this protein represents one of the little-known proteins associated with the transport and degradation of proteins in F. nucleatum (Kapatral et al., 2002). Its function and potential substrates are unknown.

A rise in pH also produced twofold downregulation of butyrate-acetoacetate CoA transferase (SSP 3104, Table 1), a metabolic enzyme associated with energy production from amino acids such as glutamate (Barker, 1981). The short-chain fatty acid thioesters butanoyl-CoA and acetoacyl-CoA are the preferred substrates, and the products of the reaction (butyric acid and acetoacetate) cause acidification of the environment. The downregulation of this enzyme is consistent with the overall decrease in metabolic activity previously reported in F. nucleatum when grown at pH 7.8 (Zilm et al., 2007). The decreased production of butyric acid would promote the alkalinization of the environment caused by the fermentation of amino acids derived from proteins. Importantly, butyric acid production is considered one of several virulence factors in fusobacteria, since butyric acid may slow wound healing by causing mitotic arrest in human gingival fibroblasts (Bartold et al., 1991).

Five proteins (SSP 1001, 1102, 2101, 2505, 3002) whose expression was downregulated in response to growth pH were represented in the F. nucleatum subsp. polymorphum genome but were identified as having no known function. With the exception of spot 2505, all of the hypothetical proteins had an observed molecular mass of 20 kDa or less, and their expression may play an important role in the organism’s response to pH. The potential function and identification of these proteins will form the basis for future investigations. An interesting case is that of SSP 1001, because the gene could be an adhesin and is flanked by two paralogous ORFs of high similarity (Table 1), both of which lack ribosome-binding sites (Supplementary Fig. S1) and may therefore not be translated. This could be indicative of intragenomic recombination between the gene encoding SSP 1001 and its two flanking ORFs, as has been shown in cases of antigens undergoing combinatorial gene conversion among contiguous ORFs (Santoyo & Romero, 2005).

Proteins upregulated by increasing pH
Two hypothetical proteins (SSP 0202 and 3101) were upregulated in response to growth at pH 7.8. PSORTb analysis predicted that one protein (3101), which had a positive GRAVY score, was associated with the cytoplasm.
The cellular location of spot 0202 could not be predicted using the PSORTb algorithm, although the detection of a signal peptide suggests that the protein is transported across the cytoplasmic membrane.

The outer membrane protein (OMP) SSP 3802 (Table 1) was upregulated about threefold and belongs to the Omp IP family of porins. It has been identified as one of the 132 proteins that play a role in bacterial virulence (Karpathy et al., 2007) and is a putative surface antigen containing variable number repeat (VNR) motifs (http://www.oralgen.lanl.gov/). Five of the 14 predicted outer membrane proteins identified in F. nucleatum subsp. nucleatum (Kapatral et al., 2002) have molecular masses above 188 kDa and so homologues in subsp. polymorphum would not be resolved by the 2DE gels used in this study.

The approximately twofold upregulation of the predicted peptidyl prolyl cis-trans isomerase (PPI) (SSP 3202 and 3201, Table 1) may affect the secretion of outer membrane proteins. Although PSORTb analysis predicted that the location of the protein was cytoplasmic, PPIs are a family of soluble proteins found in the periplasm of Gram-negative bacteria that act as chaperones for the folding of proteins such as outer membrane proteins and pilus components (Matsuzaki et al., 2000).

In response to growth at pH 7.8, spot 0203 (Table 1) displayed the greatest upregulation (44 fold) and is identified as a pathogen-specific membrane antigen (http://www.oralgen.lanl.gov/). As with many proteins associated with the F. nucleatum proteome, it is uncharacterized but is thought to be a high-affinity Fe^{2+} transport protein found in the CE. Iron is an essential element required for the growth of periodontal pathogens such as P. gingivalis. The onset of periodontal disease is associated with destruction of host tissues, and the subsequent bleeding releases iron in the form of haem and haem proteins (Ruby & Goldner, 2007). Perhaps the increased ability to sequester iron from the environment is in response to competition for iron caused by the proliferation of pathogenic species. We have also previously reported (Zilm et al., 2007) the upregulation (71-fold) of cytoplasmic flavodoxin, which is usually upregulated in response to iron limitation. It appears, therefore, that proteins associated with iron metabolism and uptake form part of the organism’s co-ordinated response to elevated pH.

### Proteins acquired by HGT

The phylogenetic analysis shows that only two of the up- or downregulated proteins present a tree pattern that is congruent with the ribosomal phylogeny (Table 2), whereas using the same protocol over 50% of all CE proteins were found to be native. Some proteins are short and do not show significant sequence similarity hits, suggesting that they could be of phage origin (transferred by phage transduction), currently underrepresented in public databases. Between 30 and 50% of genes in Table 2 show some evidence for HGT, especially from Proteobacteria, and have significant hits to phylogenetically unrelated oral species such as P. gingivalis and Treponema denticola. Thus, we suggest that these HGR genes could be good targets for a vaccine, as their cognate proteins are extracellularly exposed, highly expressed and present in several pathogenic species of the subgingival dental plaque.

### Conclusion

We have examined the proteomic changes associated with the CE when F. nucleatum is grown at physiological (pH 7.2) and elevated pH (7.8), thought to be consistent with the pH of the diseased gingival sulcus. The results presented allow gene expression to be compared with the vast amount of published physiological data and follow on from our previous examination of cytoplasmic proteins regulated by growth pH (Zilm et al., 2007). Growth at pH 7.8 produced changes in CE proteins which were associated with decreased uptake of exogenous C4 dicarboxylates and phosphorus. Reduced metabolic activity was also seen in the downregulation of a membrane protease and an enzyme associated with amino acid fermentation. This overall decrease in cell metabolic activity was also consistent with the expression of cytoplasmic proteins during growth at pH 7.8 (Zilm et al., 2007). Interestingly, the increased expression of an outer membrane porin at pH 7.8 was also associated with the upregulation of a periplasmic chaperone responsible for the correct folding of proteins as they transit the periplasm destined for the outer membrane. It should be emphasized that, although pH and other physicochemical properties may be controlled intracellularly, their effect will be important for the proteins exposed to the environment. Thus, it is interesting that many proteins predicted to be located in or on the outer membrane have paralogous genes of high sequence similarity (Supplementary Table S2) that may be performing the same function under different environmental conditions (for instance under pH fluctuations). Two of these proteins appear to be down-regulated at pH 7.8 (Table 2), and further studies could indicate whether this is an extended phenomenon and whether their paralogous counterparts are overexpressed under the same circumstances.

Unfortunately, the annotation of sequenced genomes is not keeping pace with the rate at which prokaryote genomes are being sequenced. Indeed, this study has identified a number of hypothetical proteins which are represented in the F. nucleatum genome but currently have no known function. Thus, the potential role of these proteins in response to elevated growth pH remains to be resolved.

### ACKNOWLEDGEMENTS

This work was supported by the Australian Dental Research Foundation.
REFERENCES


Edited by: M. A. Curtis