Analyst

PAPER

View Article Online
View Journal | View Issue

Cite this: *Analyst*, 2026, **151**, 137

Metabolic fingerprinting of periodontal bacteria: a multi-scale mass spectrometry and vibrational spectroscopy approach

Jawaher Albahri, [a,b] Daniel Smaje,[a] Yun Xu, [a] Steven Robinson,[c] Heather Allison,[e] Kathryn A. Whitehead [d] and Howbeer Muhamadali *[a]

Received 7th May 2025,
Accepted 7th November 2025

DOI: 10.1039/d5an00507h

rsc.li/analyst

Periodontitis is a prevalent condition characterised by progressive destruction of the supporting periodontium around teeth, resulting from prolonged interactions between the host immune system and bacterial infection. Classifying periodontal bacteria remains challenging due to their high diversity, culturability, strain level variability, biofilm complexity, and methodological constraints. Fourier Transform Infrared (FT-IR) spectroscopy was employed to investigate the biochemical composition of five oral *Streptococcus* spp., *Porphyromonas gingivalis, Actinomyces israelii, Fusobacterium nucleatum,* and *Parvimonas micra* at microbial-community level. In parallel, Matrix-Assisted Laser Desorption/Ionisation Time-of-Flight (MALDI-TOF) mass spectrometry, the gold standard analytical platform in clinical settings, was used for a comprehensive understanding of cellular protein/peptide molecules, identifying approximately 20 *m/z* features associated with the clustering pattern. Particularly, *P. gingivalis* was distinctly clustered within higher *m/z* range (8000–16 500), primarily driven by features at 9329, 9800, and 11 029 *m/z*, highlighting the potential of this window for its specific identification. Additionally, Optical Photothermal Infrared (O-PTIR) spectroscopy, combined with chemometrics, was also utilised at single-cell level. Classification patterns across IR techniques and MALDI-MS were comparable, revealing significant variation in protein and lipid profiles among the studied strains. *P. gingivalis*, *F. nucleatum* and *A. israelii* exhibited distinct clustering, while *S. anginosus* and *S. oralis* showed significant variation within the *Streptococcus* spp. The overall clustering pattern was primarily attributed to spectral information within amides (1500–1800 cm^{-1}) and fatty acids (2800–3050 cm^{-1}) regions, alongside vibrational bands observed between (900–1200 cm^{-1}), indicative of polysaccharides and phosphate-containing compounds. The submicron spatial resolution demonstrated by O-PTIR, suggests promising potential for direct application to clinical samples without the need for prolonged culturing.

Introduction

Periodontitis is a globally pervasive condition characterised by the progressive inflammation of the periodontal tissues, resulting from a complex interaction between microbial infection and the host immune response.[1] Nevertheless, the progression of periodontitis is contingent upon dysbiotic alterations in the oral microbiome in response to the breakdown products of periodontal tissues and the efficacy of anti-bacterial mechanisms.[2] According to the complex theory, oral bacteria associated with subgingival plaque formation are classified into five colour-labelled groups based on their presence in tooth colonisation and periodontitis severity.[3] The red complex bacteria *Porphyromonas gingivalis, Tanerrela forsythia,* and *Treponema denticola* are known as the keystone pathogens in chronic periodontitis, and they are strict anaerobes and late colonisers in subgingival biofilm formation, appearing in later stages of periodontitis.[4] Clinically, the impacts of periodontitis are not restricted to the oral cavity. As it has been explored in our previous review,[5] periodontal pathogens and their toxins can disseminate to distal parts of the body, having been reported in conjunction with some systemic diseases such as Alzheimer's disease,[6] Parkinson's disease,[7] chronic kidney disease,[8] and diabetes,[9] highlighting its broader clinical relevance. Beyond oral health, periodontitis imposes economic burden, with

[a]*Centre for Metabolomics Research, Department of Biochemistry, Cell and Systems Biology, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK. E-mail: Howbeer.muhamad-ali@liverpool.ac.uk*
[b]*Department of Pharmaceutical Chemistry, College of Pharmacy, King Khalid University, Abha 62529, Saudi Arabia*
[c]*Materials Innovation Factory, University of Liverpool, Liverpool L7 3NY, UK*
[d]*Microbiology at Interfaces, Department of Life Sciences, Manchester Metropolitan University, Chester St, Manchester M1 5GD, UK*
[e]*Department of Clinical Infection, Microbiology, and Immunology. Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, L69 7ZB, UK*

global costs, including treatment expenses and productivity loss, estimated to be in the billions annually.[10]

The application of appropriate detection and identification methods provides valuable information on the early diagnosis of periodontitis, assisting in designing preventive and therapeutic intervention strategies, which have been demonstrated by recent studies.[11,12] Although the traditional microbial phenotyping methods including Gram staining, microscopy tests, colony morphology, and analytical profile index (API) tests are still valuable, they are often time-consuming, subjective, have specificity limitations and may not always differentiate closely related species.[13] Many methods require up to 48 hours to acquire reliable results for these fastidious bacteria.[14] Such methods require complementary techniques for fast, precise, and reliable results. Vibrational spectroscopy and mass spectrometry techniques, such as metabolic fingerprinting approaches, have attracted a lot of interest for such applications. These approaches are generally quick and reproducible, requiring minimal sample preparation, with the potential to provide simultaneous information on a diverse range of biomolecules, including lipids, proteins, nucleotides, and carbohydrates.[15,16]

Fourier Transform Infrared (FT-IR) spectroscopy, in particular, has been widely used for bacterial identification and classification.[17–19] In this technique, the absorption of IR light by distinct functional groups induces changes in bond vibrations involving stretching or bending, consequently causing specific variations in the dipole moment.[20] The resultant distinctive spectral profile displays unique characteristics, forming a metabolic fingerprint, which can be utilised for the identification or discrimination of samples, as has been reported previously.[21,22] FT-IR has been extensively utilised across various fields, such as the pharmaceutical industry,[23,24] clinical diagnostics,[25,26] and even environmental research.[27] However, the diffraction limit of FT-IR restricts its application to bulk measurements, requiring large cell clusters to obtain sufficient signal.[28–30] Optical Photothermal Infrared (O-PTIR) spectroscopy, on the other hand, is a novel vibrational technique providing IR measurements and contactless imaging with the superior advantage of providing sub-micron spatial resolution in comparison to traditional FT-IR microscopy. It has excellent spectral sensitivity and the ability to collect high-quality spectra within a very short time with minimal sample preparation required.[31] Such characteristics have facilitated the exploration of numerous metabolomics investigations at the single-cell level, such as human cellular specimen differentiation,[32] bacterial cell discrimination,[33] and detection of antimicrobial resistance in bacteria.[34] In contrast to FT-IR spectroscopy, the capabilities of O-PTIR are constrained to the measurement of infrared spectra exclusively within the fingerprint region (950–1800 cm^{-1}). This limitation arises from the coverage of the utilised quantum cascade laser (QCL), which selectively covers this specific spectral range.[18] However, recent advancements in new O-PTIR platforms have integrated QCL with tuneable Optical Parametric Oscillator (OPO) lasers, providing continuous tunability across a wider range of wavelengths.[35]

On the other hand, Matrix-Assisted Laser Desorption Ionisation Mass Spectrometry (MALDI-MS) is a well-established technique that has been widely used for bacterial identification and discrimination at the microbial community level.[36,37] It is a soft ionisation technique in which a laser induces ionisation of bacterial samples with the assistance of the appropriate matrix, such as sinapinic acid (SA), α-cyano-4-hydroxycinnamic acid (HCCA), or 2,5-dihydroxybenzoic acid (DHB).[38] The generated molecular ions then travel inside the Time Of Flight (TOF) mass analyser, where they separately reach the detector based on their mass-to-charge ratio (m/z).[39] In contrast to the vibrational spectroscopic techniques, MALDI-TOF-MS spectral results provide information mainly about the protein and peptide content within the bacterial cells.[40] For bacterial identification, the generated protein profile can be matched to reference profiles stored in organism-specific databases.[41] The reproducibility, speed and ease of sample preparation have made MALDI-TOF-MS the gold-standard analytical technique for bacterial identification in clinical settings, through platforms such as Bruker's Biotyper.

In the present study, *P. gingivalis*, a representative of the red complex pathogens, alongside, eight strains representatives of the other microbial complexes were investigated. These included the facultative anaerobes *Streptococcus anginosus*, *Streptococcus constellatus, Streptococcus gordonii*, *Streptococcus sanguinis*, and *Streptococcus oralis*, and three obligate anaerobes including *Actinomyces israelii*, *Fusobacterium nucleatum*, and *Parvimonas micra.* Thus, our primary objective was to conduct a fingerprint analysis to characterise the selected nine periodontal bacteria employing mid-infrared spectroscopy in addition to comparison with MALDI-TOF-MS to provide information about the cellular protein/peptide molecules. To the best of our knowledge, this study marks the first characterisation of the selected bacteria by combining FT-IR, and O-PTIR spectroscopic methodologies in addition to MALDI-TOF-MS. This pioneering approach is designed to comprehensively investigate the biochemical composition of these bacterial strains at both the microbial community and single-cell levels.

Experimental methods

Growth conditions and sample preparation

S. anginosus (NCTC 11169), *S. constellatus* (NCTC 11063), *S. gordonii* (NCTC 7870), *S. sanguinis* (NCTC 10904), *S. oralis* (NCTC 11427), *A. israelii* (NCTC 6831), *F. nucleatum* (NCTC 10562), *P. gingivalis* (NCTC 11834), and *P. micra* (NCTC 11808) were purchased from National Collection of Type Cultures (NCTC) in lyophilised form and aseptically reconstituted and cultured following their instructions and recommended growth conditions. To eliminate any potential bias and changes in metabolic fingerprints of the cells associated with the components of different culturing media, fastidious anaerobic agar (Neugen culture media, UK) supplemented with 5% defibrinated horse blood (TCS Biosciences, UK) was used as the control growth medium to support the growth of

all nine strains in this study. All culture plates were kept in an anaerobic chamber for 24 h before use. The strains were sub-cultured anaerobically three times to achieve axenic colonies. A single colony was streaked to make three biological replicates for the FT-IR analysis and five biological replicates for MALDI-TOF-MS analysis. The plates were anaerobically incubated at 37 °C for 24 h for the five *Streptococcus* spp., 48 h for *P. micra*, 72 h for *A. israelii* and *F. nucleatum*, and a week for *P. gingivalis* using Oxoid™ anaerobic jars. The anaerobic atmosphere inside the jars was frequently checked using Oxoid™ Resazurin Anaerobic Indicator.

FT-IR spectroscopic analysis

Samples were prepared following the established pipeline used in our previous work, which yielded spectra within the linear range.[17,34] Bacterial biomass was collected from the blood agar plate using a 10 μL inoculation loop (Sigma-Aldrich, UK) and resuspended in 1 mL of sterile physiological saline solution (0.9% NaCl). The optical density of the bacterial slurries was recorded at 600 nm wavelength ($OD_{600\ nm}$) using a 6705 UV/Vis spectrophotometer (Jenway, UK). The samples were then centrifuged and washed using the sterile saline solution twice at 4 °C and a speed of 5000 *g* for 6 min to remove the supernatant and any residuals from the culture media. After the final washing step, the supernatant was removed, and the bacterial pellets were kept for further analysis. Samples were randomised to ensure unbiased allocation. All samples were normalised on the same day to an $OD_{600\ nm}$ of 20 using sterile saline solution, before spotting 20 μL of each onto an FT-IR silicon plate (Bruker Ltd, UK), which was thoroughly washed and dried as previously outlined.[42] The plate was dried overnight in a desiccator and then placed in a 50 °C oven for 45 min before performing the FT-IR analysis. FT-IR spectral data were collected in absorbance mode within the mid-infrared region (600–4000 cm^{-1}), with 4 cm^{-1} spectral resolution, using a Bruker Invenio FTIR spectrometer (Bruker Ltd, UK). In total, 12 spectra were obtained for each strain, resulting from three biological replicates and four technical replicates to provide a representative of the microbial population and detect any variations in terms of sample heterogeneity. FT-IR data collection was conducted *via* OPUS spectroscopy software (Bruker Ltd, UK). All spectral data were pre-processed following the optimised workflow reported in our previous studies.[17,36,43] Since the instrument was not N_2 purged, the peaks associated with atmospheric CO_2 was removed by replacing these vibrations (2400–2275 cm^{-1}) with a linear trend. The spectral data were then scaled using the extended multiplicative signal correction (EMSC) method to reduce the variation attributed to the scattering effect, before performing data analysis.[44]

MALDI-TOF-MS analysis

Sample preparation optimisation

Sinapinic acid (Sigma Aldrich, UK) was used as the matrix, and mixing method was employed for sample deposition. Sample preparation was optimised to determine the optimal amount of bacterial biomass to provide peaks with acceptable reproducibility and signal-to-noise ratio. *Escherichia coli* MG1655 was used to optimise the sample preparation method. Bacterial cultures were prepared following the methods described in the growth condition section above. The bacterial biomass was then carefully collected using a sterile 20 μL plastic loop and resuspended into 1 mL of deionised (DI) water, which was vortexed until it became homogenous, then 100 μL was used to measure the OD_{600} using a 6705 UV/Vis spectrophotometer (Jenway, UK), while the remaining sample was washed twice using DI water as described above. All bacterial pellets were then stored at −80 °C freezer until further analysis. The optimisation investigated twelve different conditions, including four different ODs ranging from the lowest value of 5 up to 20. In parallel, three different ratios of sample to the matrix were tested, including mixing equal volumes (1 : 1), doubling the sample volume (2 : 1), and doubling the matrix volume (1 : 2).

Periodontal bacterial samples

In total, five biological replicates were prepared and stored at −80 °C freezer until further analysis. The defrosted pellets were normalised to $OD_{600\ nm}$ of 10 (optimal bacterial load) using DI water containing 0.1% trifluoroacetic acid (TFA). The matrix was prepared by dissolving 10 mg of sinapinic acid in 500 μL of acetonitrile and 500 μL of water containing 2% TFA. An aliquot (20 μL) from the normalised bacterial sample was then mixed with 20 μL of the matrix and vortexed for 20 s. All samples were randomised to ensure unbiased distribution. Then, 2 μL of the bacterial mixture was spotted on a microtiter plate 384-well polished steel plate (Bruker Ltd, UK) and left to dry at room temperature for 1 h.

Samples were analysed using a Bruker Autoflex Speed time-of-flight mass spectrometer equipped with a Bruker smartbeamTM-II laser (355 nm wavelength). The instrument was operated in positive ion mode, and linear TOF was used over the range from 1000 to 80 000 *m*/*z*. The full-width at half maximum (FWHM) was determined for representative peaks at low, medium, and high ends within the *m*/*z* range containing detectable peaks (2000 to 17 000 *m*/*z*) (Table S1). Each sample had five biological replicates, and from each biological replicate, five analytical replicates were prepared. The results of this analysis generated 225 MALDI-TOF-MS spectra: 9 bacteria × 5 biological replicates × 5 technical replicates. The instrument was calibrated using five standard proteins (MSCAL3) from Sigma Aldrich, UK including (molecular weight provided in parentheses) insulin (5630), cytochrome c (12 185), apomyoglobin (16 722), aldolase (39 212), and albumin (66 430). A mixture of the five proteins was prepared as per the supplier's recommendation. Then aliquots (2 μL) of the (1 : 1) sinapinic acid : standards mixture were spotted after every five samples on the plate. Mass accuracy achieved after calibration was calculated using (eqn (1)).[45] This included the standards covering the 2000–1700 *m*/*z* range (Fig. S1), yielding an overall average error of around 168 ppm (Table S2), which falls within the

accepted mass error tolerance in peptide mass fingerprinting.[46] MS data were collected using Bruker flexAnalysis software. All MS spectra were baseline corrected, then subjected to morphological score correction and filtered using signal-to-noise ratio threshold of 10, followed by vector normalisation.

$$\text{Mass error (ppm)} = \left|\frac{m_{\text{measured}} - m_{\text{reference}}}{m_{\text{reference}}}\right| \times 10^6 \quad (1)$$

O-PTIR single-cell analysis

Samples were prepared by taking 100 μL aliquots of the bacterial slurries from the FT-IR samples and diluted with 900 μL of DI water. The samples were then centrifuged and washed twice with DI water at 4 °C and 5000 *g* for 6 min before undergoing serial dilutions with DI water to achieve single-cell distribution of the sample. Aliquots of 2 μL of each sample were then spotted on a calcium fluoride (CaF_2) slide in a descending order from the highest to the lowest concentration to determine the best dilution to achieve single bacterial cells. The samples were then dried in a 50 °C oven for 30 min before carrying out the O-PTIR analysis. Single-point O-PTIR spectra of up to 55 individual bacterial cells from each sample spot were collected in the reflection mode within the mid-infrared region (950–1800 cm^{-1}, and 2 cm^{-1} spectral resolution) using Schwarzschild objective (40×, 0.78 NA) on a mIRage microscope (Photothermal Spectroscopy Corp., Santa Barbara, CA, USA), equipped with tunable QCL and a continuous-wave (CW) 532 nm laser, which acted as pump and probe beams, respectively. Single-frequency images were obtained by tuning the QCL to vibrational modes associated with amide I at (1655 cm^{-1}) and polysaccharides at (1061 cm^{-1}). Probe power was adjusted to 39% with an imaging pixel size of 0.2 μm and a settle time of 10 min. O-PTIR data were filtered by removing all the vibrational bands with intensities below 0.1 at the amide I peak (1655 cm^{-1}), resulting in five to eight spectra for each strain generated from different single cells. The resulted spectra were then subjected to Gaussian smoothing using a window size of 20 data points and autoscaled.

Multivariate statistical analysis

Data processing and analysis were carried out using MATLAB® (version r2024a, The MathWorks Inc., UK). All data were subjected to the Principal Component Analysis (PCA), an unsupervised method used to minimise the dimensionality of the multivariate data while retaining most of its variability without prior knowledge of samples classification in the dataset.[47,48] Then, a supervised method of Principal Component-Discriminant Function Analysis (PC-DFA) was employed to analyse the data. In this method, DFA is applied to principal components (PCs) to find linear combinations that best discriminate between predefined classes.[48] However, in this study PC-DFA was used as a semi-supervised method,[17] where the algorithm was provided only with information regarding the replicates, without specifying the genus or species of the samples, therefore each biological replicate was defined as a separate class.

Results and discussion

Discrimination at microbial-community level

FT-IR fingerprint analysis. Infrared spectra of all nine periodontal microorganisms were collected using FT-IR spectroscopy to investigate the variance in their biochemical composition in the bulk population. The results of FT-IR analysis showed variation in three main regions of the IR spectra, including fingerprint (900–1200 cm^{-1}), amides (1500–1700 cm^{-1}), and lipid (2800–3000 cm^{-1}) regions (Fig. S2, and S3). The PCA scores plot of the FT-IR spectral data collected from these samples displayed four distinct clusters according to PC1 and PC2 axes, with a combined total explained variance (TEV) of 83.1% (Fig. 1A). The clustering patterns observed were different to Socransky's complex theory,[3] which classified subgingival bacteria based on DNA hybridisation profiles. While Socransky's model groups *P. gingivalis* and *F. nucleatum*, into the red and orange complexes, respectively, in the present study, these two species were clustered closely together on the negative side of PC1 accounting for 70.94% of the TEV in the dataset. *A. israelii*, belonging to the blue complex, was postioned on the positive side of PC1 axis. *S. sanguinis*, *S. gordonii*, and *S. oralis*, which are all from the yellow complex, were found clustered in the middle together, as expected. The clustering pattern of the three isolates was consistent with their assignment to the *Streptococcus mitis* group, as members of this group reliably form well-defined clusters in both gene-based and whole-genome phylogenetic analyses.[49,50] Such clustering supports the conclusion that the isolates share a common genetic background characteristic of the mitis group. However, two other strains from the orange complex, *S. constellatus* and *P. micra* were clustered with bacteria from the yellow complex. Furthermore, *S. anginosus* was distinctly separated based on PC1 from the other *Streptpcoccus* spp. These differences may reflect the distinct biological features captured by FT-IR biochemical profiling compared to DNA-based association patterns. While the PCA is an unsupervised approach, the results showed tight clustering of the data within the same class, emphasising the high reproducibility of the FT-IR results. According to the PC1 loadings plot, accounting for 70.94% of the TEV, vibrational bands attributed to polysaccharides due to C–O stretching at 1061 cm^{-1} influenced the separation of *S. anginosus* and *A. israelii* on the positive side of the PC1 axis. In contrast, the spectral features that significantly contributed to the discrimination of *P. gingivalis* and *F. nucleatum* along the negative axis of PC1, included the following main vibrations: the amide I at 1634 cm^{-1} and amide II at 1543 cm^{-1}, attributed to C═O stretching and N–H bending coupled with C–N stretching, respectively. Other significant peaks include the fatty acids vibrational bands at 2961 cm^{-1} of

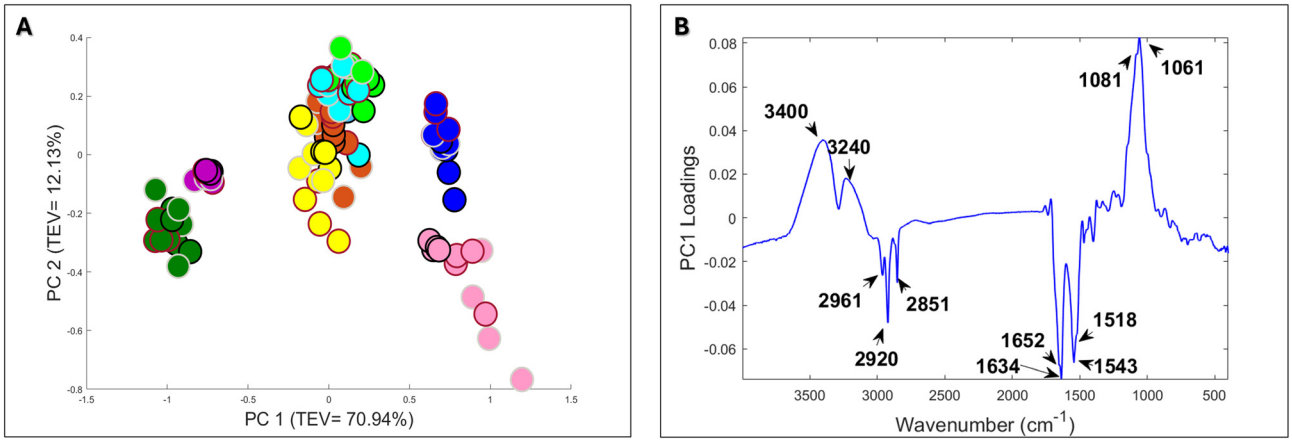

Fig. 1 (A) Shows PCA score plot obtained using processed FT-IR data of the nine periodontal bacteria cultured under the same growth conditions using fastidious anaerobic agar supplemented with 5% horse defibrinated blood. Key colours used in figures: *S. anginosus* (●), *S. constellatus* (●), *S. gordonii* (●), *S. oralis* (●), *S. sanguinis* (●), *P. micra* (●), *A. israelii* (●), *F. nucleatum* (●), and *P. gingivalis* (●). Different outline colours represent different biological replicates. (B) Represents PC1 loadings of FT-IR spectral data including fatty acids region (2800–3050 cm^{-1}), amides region (1500–1800 cm^{-1}), mixed region (1250–1450 cm^{-1}), polysaccharides region (900–1200 cm^{-1}), and fingerprint region (600–900 cm^{-1}).

C–H asymmetric stretching, and 2920, and 2851 cm^{-1} assigned to C–H asymmetric, and symmetric stretching of the methylene group, respectively (Fig. 1B). Since the samples are mixed-species, the corresponding loading plots reflect variations in the overall biochemical composition of the bacterial cells, with separation occurring along PC axes based on differences in general biomolecular features. Due the observed variation among the *Streptococcus* strains, and to further investigate the variation at the species level, the FT-IR dataset of the *Streptococcus* spp. was analysed separately and detailed in a later section. Furthermore, since all the nine strains were grown on FAA supplemented with 5% defibrinated horse blood, to investigate the potential effects of growth medium on the detected clustering patterns, another FT-IR fingerprint analysis was performed where all the bacteria were grown on different media as per the supplier recommendation and the results were compared with FAA with 5% defibrinated horse blood as the control condition (Fig. S4). The results showed a similar clustering pattern across all conditions, indicating that the observed clustering is attributable to phenotypic differences rather than background-dependent effects.

MALDI-TOF-MS analysis. MALDI is a soft ionisation technique which has been widely used for proteins measurements in bacterial cells.[38] Since there are no universal matrix or standard technique for the deposition of sample/matrix mixture, the MALDI-TOF-MS sample analysis protocols were tested using *E. coli* MG1655 as a representative strain for method optimisation. To optimise MALDI-TOF-MS spectral quality and ensure reproducibility, bacterial biomass normalised by optical density (OD_{600}) across a range of 5 to 20 were evaluated. This range was selected to examine the influence of cell density on ionisation efficiency and signal-to-noise ratio. At lower ODs (*e.g.*, 5), the protein content was insufficient for reliable detection, whereas very high ODs (*e.g.*, 20) led to matrix suppression and ion suppression effects, resulting in inconsistent signal intensities and peak distortion (Fig. S5–S7). Empirical assessment identified an intermediate OD_{600} of 10 as optimal, providing sufficient biomass for consistent ionisation while minimising suppression artefacts. This normalisation produced reproducible, high-intensity MALDI-TOF-MS profiles, enabling robust and accurate comparative analyses between samples. The results showed that the spectral data collected from samples with $OD_{600\ nm}$ of 10 using a mixing ratio of 1 : 1 was the most reproducible method, with a relative error of 22–33% (Table S3). These findings are in agreement with the literature since the mixing ratio of 1 : 1 is the most frequently applied ratio in several MALDI-TOF-MS studies for bacterial characterisation.[36,51] The optimised method was used as the standard approach to characterise the periodontal bacteria at microbial community level using MALDI-TOF-MS analysis, where 225 spectra were collected and processed as mentioned above.

To evaluate the impact of signal-to-noise (S/N) threshold on the detection of low-abundance peptide peaks, the data was processed and analysed using a S/N cutoff of 5 as well as 10. While the lower threshold (*i.e.* S/N = 5) yielded generally good reproducibility (Fig. S8), to minimise the influence of noise particularly for low-intensity signals near the detection limit, as shown in (Fig. S9), a higher threshold (*i.e.* S/N = 10) was selected for the data analysis. After data pre-processing and filtering, only 187 spectra remained. *F. nucleatum* spectral data was the most affected, which reduced from 25 to just 4, collected from three distinct biological replicates. All of the analysed periodontal bacteria displayed signals in the *m*/*z* range between 2000 to 17 000, except for *P. micra* and *S. constellatus*, which had limited *m*/*z* ranges up to ~7000 and ~9000, respectively. Among the MS spectra of *Streptococcus* spp., only *S. oralis* had a significant peak at ~11 000 *m*/*z* (Fig. S10). The PC-DFA scores plot of the whole *m*/*z* range (2000–17 000) using 6 PCs (Fig. 2A), with a TEV of 89.07%, was comparable to the IR find-

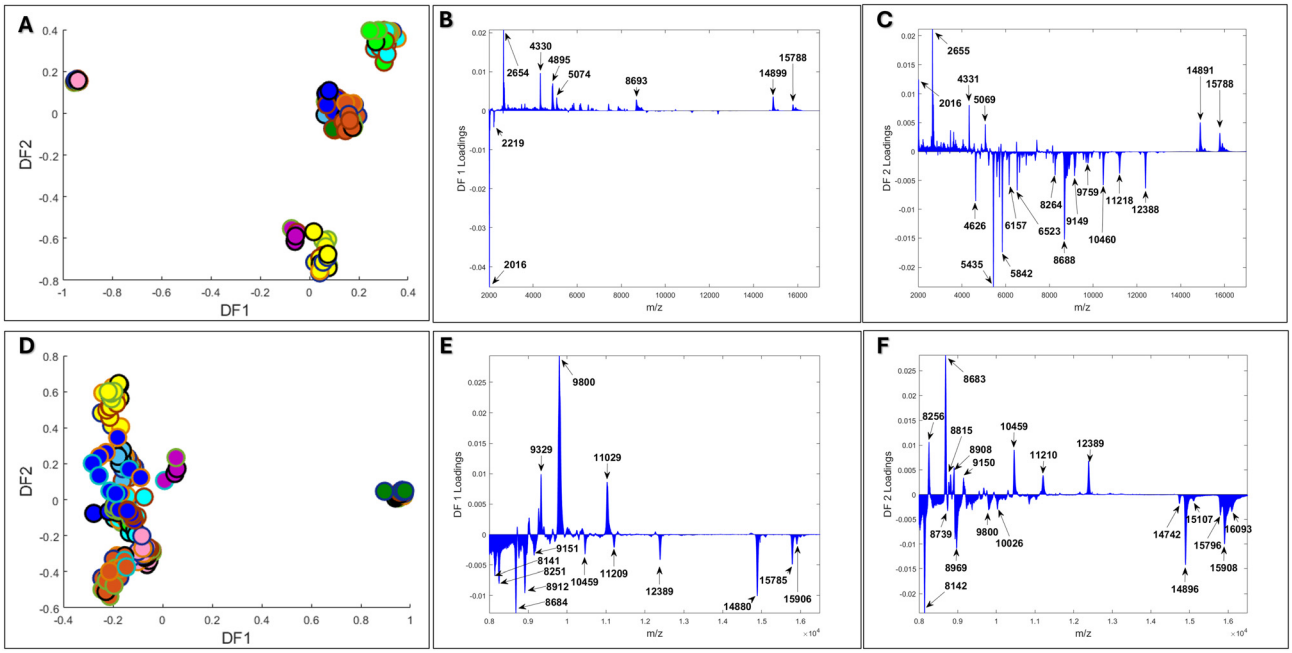

Fig. 2 (A) Shows PC-DFA scores plot of the complete MALDI-TOF-MS mass spectral data (2000–17 000 *m*/*z*), using 6 PCs (TEV = 89.07%), while (B) and (C) represent DF1, and DF2 loadings plot associated with the full range. However, (D) shows PC-DFA score plot of the limited high range (8000–16 500 *m*/*z*) using 4 PCs (TEV = 83.93%), whereas (E), and (F) represent DF1, and DF2 loadings at this range, respectively. All the significant loadings are presented at (Table S2). In (A) and (D), different outline colours were used to represent different biological replicates as each biological replicate was introduced as a separate class (total number of classes 45). Key colours used in figures: *S. anginosus* (●), *S. constellatus* (●), *S. gordonii* (●), *S. oralis* (●), *S. sanguinis* (●), *P. micra* (●), *A. israelii* (●), *F. nucleatum* (●), and *P. gingivalis* (●).

ings (Fig. 1A). Scree plots were used to select the number of the PCs included (Fig. S11). *A. israelii* was separated from the other samples dominating the negative side of DF1, whereas *F. nucleatum* and *S. oralis* were clustered on the negative side of DF2 (Fig. 2A–C). According to DF1 loadings, *m*/*z* of 2016 and 2219 were found as the two significant variables that influenced the separation of *A. israelii*. However, twelve *m*/*z* values, including 4626, 5435, 5842, 6157, 6523, 8264, 8688, 9149, 9759, 10 460, 11 218, and 12 388 were associated with the clustering pattern of *F. nucleatum* and *S. oralis* along DF2 axis (Table S4). To further investigate the variance between the strains, we restricted the data only to the higher *m*/*z* window of 8000 to 16500, and subjected this limited range to PC-DFA using 6 PCs, with a total explained variance of 83.93% (Fig. 2D–F). Using this new range of data, *P. gingivalis* displayed a clear separation from the other strains according to DF1 axis, which is in agreement with the IR results. Since this clustering pattern was not seen when the whole range was analysed, it indicates the importance of this region for differentiating *P. gingivalis* from the remaining strains. Three *m*/*z* values at 9329, 9800, and 11 029 were found influencing the clustering of *P. gingivalis* at the higher *m*/*z* range. Moreover, eight *m*/*z* values including 8256, 8683, 8815, 8908, 9150, 9150, 10 459, 11 210, and 12 389 could be linked to the separation of *S. oralis* on the positive side of DF2.

Ribosomal proteins are the most frequently detected proteins in the bacterial MALDI fingerprint because of their widespread distribution inside the bacterial cell and in high abundance.[52] These proteins remain unaffected by the phenotypic changes during cultivation, enhancing the reproducibility of MALDI as a technique suitable for clinical laboratory applications.[37] The conserved nature of ribosomal protein sequences at the genus level, sometimes even at the species level, allows MALDI to readily detect small variations, enabling differentiation between closely related microorganisms.[40] The mass range of 2000 to 20 000 *m*/*z* has been previously reported as being associated with ribosomal proteins, with few housekeeping proteins and is frequently used for bacteria identification at the species level.[53,54] Thus, the *m*/*z* values highlighted above could be attributed to ribosomal proteins and could be considered for use in organism-specific databases in MALDI-MS profiling.

Discrimination at single-cell level

O-PTIR fingerprint analysis. Unlike FT-IR spectra, the O-PTIR spectra displayed lower signal-to-noise ratio. This was of course as expected, as the O-PTIR analysis was carried out on single bacterium where the amount of biomass being analysed was significantly lower, around 1 pg of material. Therefore, following the filtration step outlined in the methods section, approximately five to eight spectral data points were obtained for each strain (Fig. S12). While the clustering pattern of PCA scores plot, with combined total explained variance of 74.75% within the dataset (Fig. 3A), is

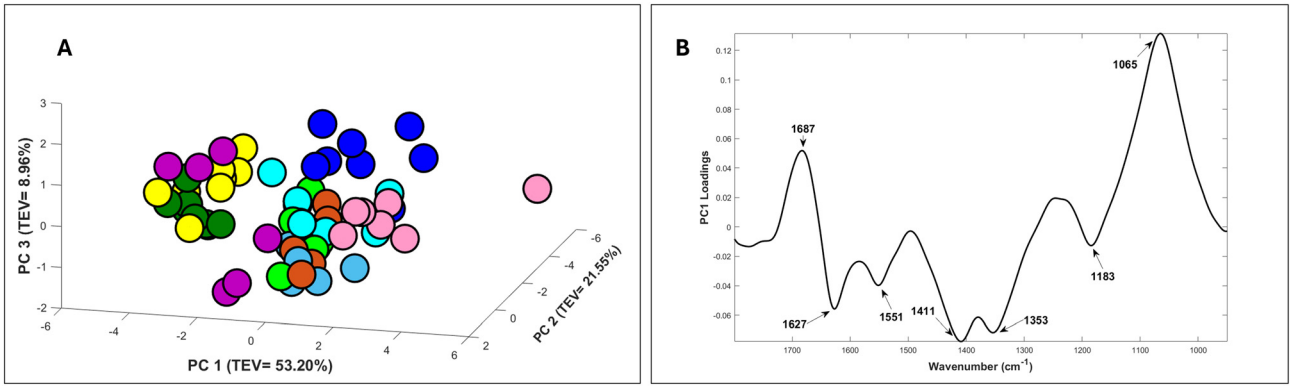

Fig. 3 (A) Shows PCA score plot of the processed O-PTIR spectral data of the nine periodontal bacteria at single-cell level within the fingerprint region (950–1800 cm^{-1}) as multiple principle components were plotted PC1 *vs.* PC2 *vs.* PC3, where different colours represent different species. Key colours used in figures: *S. anginosus* (●), *S. constellatus* (●), *S. gordonii* (●), *S. oralis* (●), *S. sanguinis* (●), *P. micra* (●), *A. israelii* (●), *F. nucleatum* (●), and *P. gingivalis* (●). (B) represents the PC1 loadings plot, which accounts for the maximum variance in the dataset including fingerprint region (950–1800 cm^{-1}).

not as distinct as that observed in the FT-IR data, the clustering patterns observed remains comparable to those obtained from the FT-IR analysis (Fig. 1A). *P. gingivalis* and *F. nucleatum* along with *S. oralis* separated on the negative side of PC1 influenced by vibrational bands at 1353 cm^{-1} and 1411 cm^{-1}, which are attributed to C═O symmetric stretching of COO– group in amino acids and fatty acids, in addition to N–H bending coupled with C–N stretching at the amide II absorption band at 1551 cm^{-1} and C═O stretching at amide I band at 1655 cm^{-1} (Fig. 3B). Due to the higher spatial resolution of O-PTIR, this technique can detect within population heterogeneity, which may contribute to the broader spread of clusters and the less distinct separation compared to FT-IR findings. The observed heterogeneity within the community, captured by O-PTIR, may also explain why *F. nucleatum* clustered in a different region and the results appeared less reproducible than in the FT-IR data. Importantly, the vibrational bands corresponding to N–H bending coupled with C–N stretching at the amide II band (1551 cm^{-1}) and C═O stretching at the amide I band (1655 cm^{-1}) were found to be in agreement with the FT-IR results, while the vibrational bands at 1353 cm^{-1} and 1411 cm^{-1} were not idetified as significant by the FT-IR PCA loadings. However, it is important to note their presence alongside the aforementioned amide bands in O-PTIR PC-loadings, as they could be attributed to single-cell variability contributing to the overall clustering pattern.

Recent studies have demonstrated the lower reproducibility of O-PTIR clustering patterns compared to FT-IR findings[18,34]. In FT-IR, the spectral data represent an averaged signal from the microbial community of each sample, as it captures a larger area with lower spatial resolution. In contrast, the higher spatial resolution of the O-PTIR technique enables the detection of heterogeneity between individual cells, preserving important insights that could be lost when averaging the response from a broader bacterial community. To investigate further the metabolic fingerprints at the bulk and single-cell levels, the relative ratios of the key vibrational features contributing to the clustering patterns observed by FT-IR and O-PTIR using baseline corrected data were compared (Fig. 4). The FT-IR relative ratios of peak intensities of proteins (1654 cm^{-1}), lipid esters (1740 cm^{-1}), and fatty acids (2920 cm^{-1}), to polysaccharides (1061 cm^{-1}) were calculated and presented as box whisker plots (Fig. 4A–C). These findings suggested that *P. gingivalis* and *F. nucleatum* had the highest protein and lipid to polysaccharide ratios when compared to the other strains. At the single-cell level, the O-PTIR findings also displayed similar trends, where the relative ratios of protein (1655 cm^{-1}) to polysaccharide (1063 cm^{-1}) intensities indicated that *P. gingivalis*, *F. nucleatum*, and *S. oralis* had the highest protein to polysaccharide ratios (Fig. 4D). This is perhaps not surprising as both *P. gingivalis* and *F. nucleatum* are Gram-negative anaerobes, and the high ratios of protein and lipid could be attributed to the nature of the structure of the Gram-negative bacterial cells when compared with Gram-positive ones.

However, the capabilities of our current O-PTIR setup were limited to measuring infrared spectra exclusively within the fingerprint region (950–1800 cm^{-1}), due to constraints imposed by the instrument setup and the spectral range of the QCL, which selectively covers this region.[18] As a result, lipid ratios at the single-cell level were calculated only using the peak intensity of the lipid ester vibrational band at 1740 cm^{-1}, with *P. gingivalis* exhibiting the highest relative ratio (Fig. 4E).

Next, the targeted single-frequency imaging capabilities of the O-PTIR setup was utilised to image the bacterial cells. These single-frequency photothermal images serve as label-free biochemical maps, acquired in a contactless manner. By tuning the laser to specific vibrational bands, the images revealed the relative cellular distribution of key biomolecules such as proteins, lipids, and polysaccharides. The results of O-PTIR images collected at amide I (1655 cm^{-1}) and polysaccharides (1061 cm^{-1}) vibrational bands (Fig. 5) were in agree-

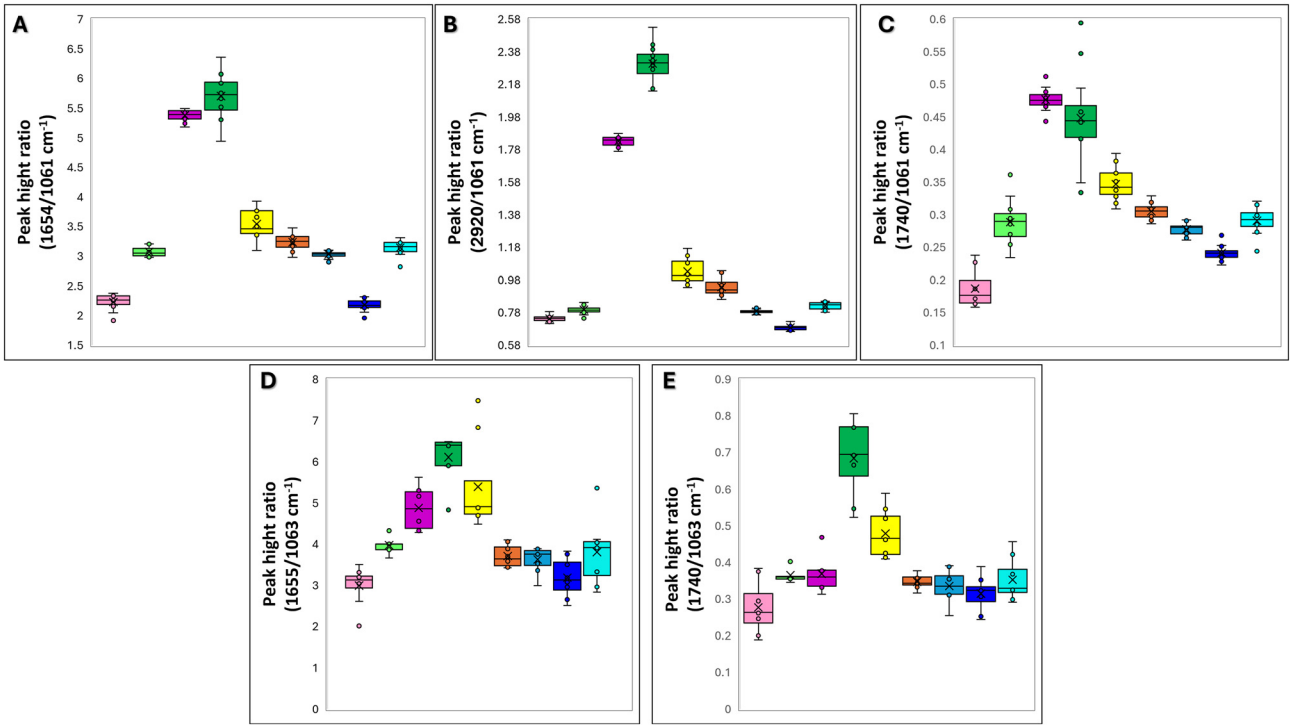

Fig. 4 Represents the calculated relative ratios of bacterial cell biomolecules as (A), (B), and (C) show Box whisker plot using FT-IR spectral data, (A) shows the relative ratio of proteins at amide I (C═O stretching at 1654 cm^{-1}) to polysaccharides (C–O stretching at 1061 cm^{-1}). (B) and (C) Represent relative ratios of lipids at (CH_2 asymmetric stretching at 2029 cm^{-1}), and lipid esters (C═O stretching at 1740 cm^{-1}), respectively, to polysaccharides vibrational band at 1061 cm^{-1}. However, (D) and (E) show Box whisker plot using O-PTIR spectral data. (D) and (E) Plots respectively show the relative ratio of proteins at amide I (C═O stretching at 1655 cm^{-1}), and lipid esters (C═O stretching at 1740 cm^{-1}) to polysaccharides (C–O stretching at 1061 cm^{-1}) to compare the variances in the relative cellular content of proteins, lipids, and polysaccharides at single-cell level. Key colours used in figures: *S. anginosus* (●), *S. constellatus* (●), *S. gordonii* (●), *S. oralis* (●), *S. sanguinis* (●), *P. micra* (●), *A. israelii* (●), *F. nucleatum* (●), and *P. gingivalis* (●).

ment with the calculated relative ratio results. Variation of protein content between periodontal bacteria can be seen across the images that were collected. Although both *P. gingivalis* and *F. nucleatum* displayed the highest protein content as previously determined by the relative ratio calculations, it could be that the coccobacillus morphological nature of *P. gingivalis* made its protein content less visually prominent in comparison to the spindle-shaped cells of *F. nucleatum.*

Furthermore, potential variation of the cellular protein content was assessed by O-PTIR imaging at amide I, where higher protein abundance was observed in *S. gordonii* and *S. oralis* in comparison to the remaining *Streptococcus* spp. (Fig. 5E–I). However, since the method was applied for qualitative analysis, future studies involving quantitative image-based evaluation is recommended. Over the years, there have been challenges in accurately identifying oral *Streptococcus* spp. and categorising them within the phylogenetic microbiota. The oral streptococci involved in this study were recently phylogenomically categorised into discrete groups: *S. anginosus* and *S. constellatus* (*S. anginosus* group – SAG), *S. oralis*, *S. sanguinis* and *S. gordonii* (mitis group).[49,55] Furthermore, literature has demonstrated challenges in classifying SAG species, leading to limitations in comprehending their role as potential pathogens.[56] To investigate these potential variations further, the FT-IR and O-PTIR spectral data of the streptococcal samples were analysed separately. Due to restricted spectral range of the O-PTIR data, the FT-IR data was also limited to the region of (950–1800 cm^{-1}) to allow for a fair comparative analysis. The PCA scores plot of FT-IR and O-PTIR data, using this limited range, accounting for combined total explained variance 80.7 and 67.29%, respectively (Fig. 6), displayed clear separation of *S. anginosus* and *S. oralis* from all other streptococcal strains. In the PC-DFA score plot of the FT-IR spectral data (Fig. 6A), *S. anginosus* was clustered on the positive side of PC1, while *S. oralis* was clustered on the negative side, displaying separate clusters compared to other streptococci, according to PC2 axis. According to PCA loadings plot (Fig. 6B, and C), vibrational bands at 1066 cm^{-1} and 1123 cm^{-1} influenced the separation of *S. anginosus* on the positive side of PC1, while vibrational bands at 1697 cm^{-1}, 1624 cm^{-1}, 1523 cm^{-1}, and 1491 cm^{-1} were the main contributors to the separation of *S. oralis* from the remaining streptococci. Similarly, for O-PTIR, *S. anginosus* and *S. oralis* were separated from other strains on the negative side of PC1 axis (Fig. 6D). However, both were clearly separated from one

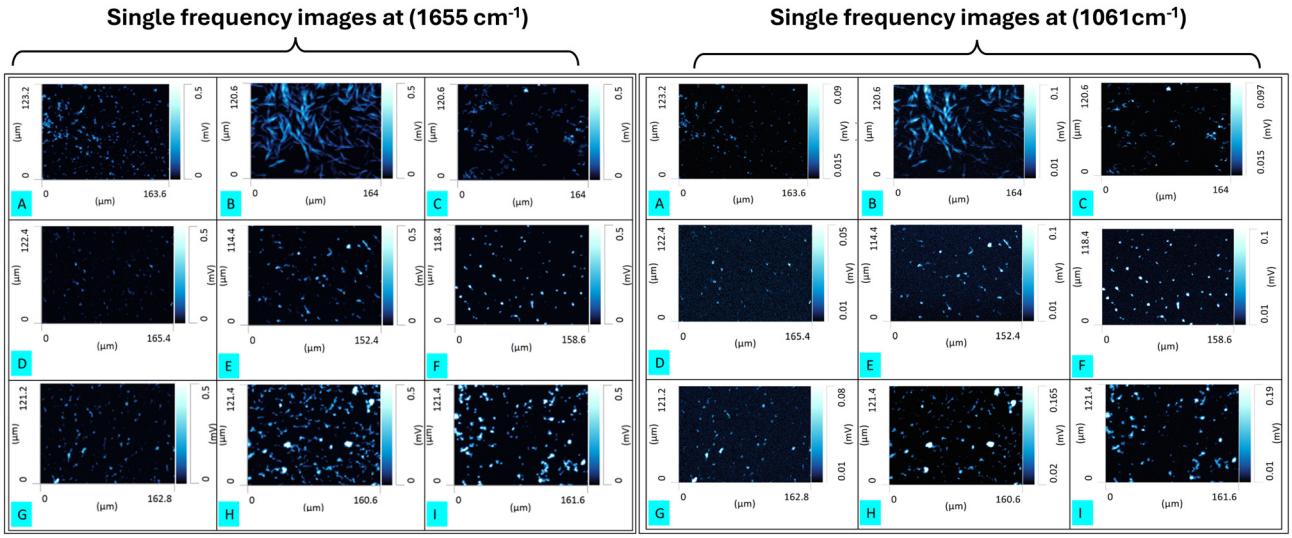

Fig. 5 Single-frequency O-PTIR images using amide I vibration at 1655 and polysaccharides vibration at 1061 cm^{-1} showing the relative concentration of proteins and polysaccharides of periodontal bacteria at single-cell level (A) *P. gingivalis*, (B) *F. nucleatum*, (C) *A. israelii*, (D) *P. micra*, (E) *S. sanguinis*, (F) *S. anginosus*, (G) *S. constellatus*, (H) *S. gordonii*, and (I) *S. oralis*.

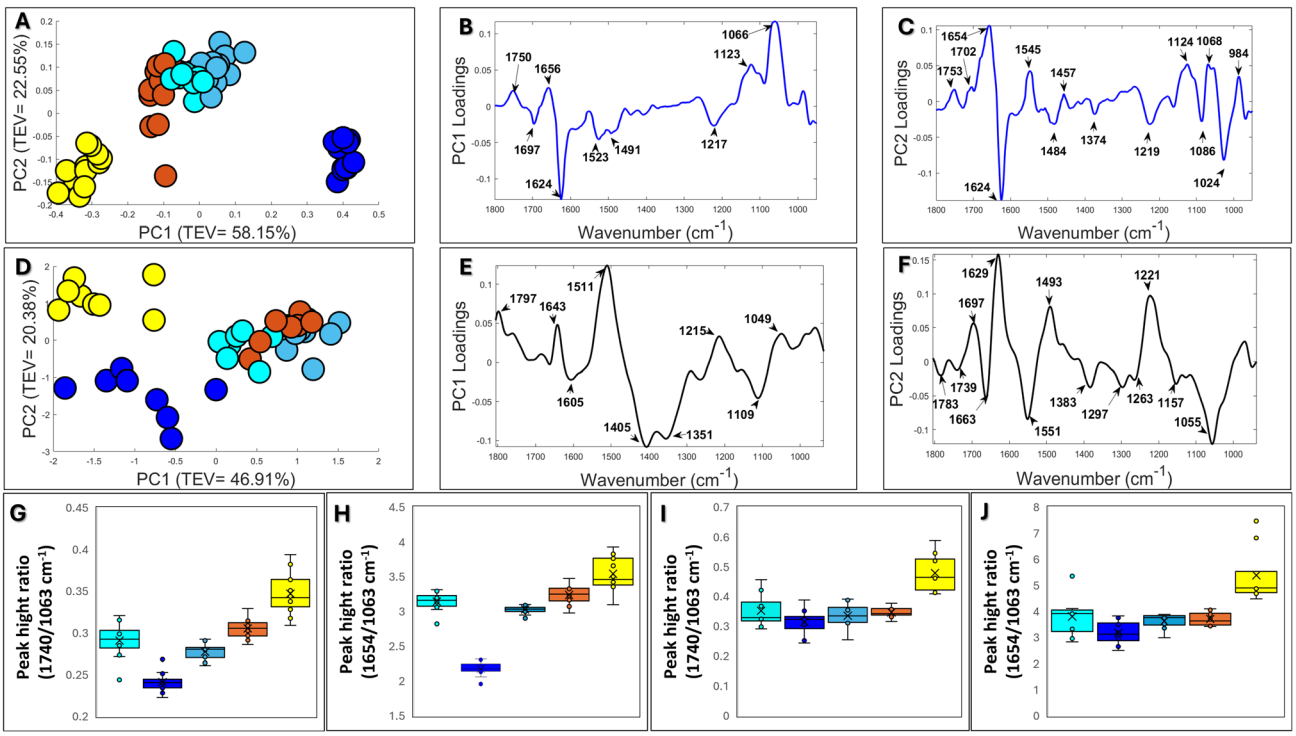

Fig. 6 Comparative discrimination of streptococcal strains using the vibrational spectroscopy techniques data. (A) Represent PCA score plot of the processed FT-IR data using fingerprint region (950–1800 cm^{-1}), while (B) and (C) show PC1 and PC2 loadings plots, respectively. However, (D) represent PCA score plot of the processed O-PTIR data using fingerprint region (950–1800 cm^{-1}), while (E) and (F) show PC1 and PC2 loadings, respectively. (G), and (H) Show Box whisker plot of FT-IR spectral raw data. (G), and (H) Plots respectively show the relative ratio of lipid esters (C═O stretching at 1740 cm^{-1}), and proteins at amide I (C═O stretching at 1654 cm^{-1}) to polysaccharides (C–O stretching at 1061 cm^{-1}) highlighting the variances in the cellular content of proteins, lipids, and polysaccharides at microbial community level (FT-IR data). However, (I) and (J) plots show the variance of the relative ratio of proteins, and lipid esters, respectively, to polysaccharides at single-cell level (O-PTIR data). Key colours used in figures: *S. anginosus* (●), *S. constellatus* (●), *S. gordonii* (●), *S. oralis* (●), and *S. sanguinis* (●).

another based on PC2 axis. According to PC1 and PC2 loadings, the vibrational bands that influenced their separation were found aligned with the FT-IR findings. The trends observed by both IR-based methods displayed comparable results, with *S. anginosus* showing the lowest lipids to polysaccharides ratio, whereas *S. oralis* presented the highest ratio. In addition, according to both, FT-IR and O-PTIR findings, *S. oralis* exhibited the highest relative ratio of protein to polysaccharide among the streptococcal species, while *S. anginosus* exhibited the lowest ratio. These findings are consistent with our MALDI-MS results (Fig. 2), where *S. oralis* clustered separately from the other streptococcal samples and exhibited a notably enhanced MALDI-MS spectral profile (Fig. S10H). It is noteworthy that the clustering pattern observed was consistent with the findings of a previous FT-IR analysis of oral streptococci conducted by Van der Mei *etal.*, in which *S. sanguinis* and *S. gordonii* where grouped together alongside three isolates of *S. mitis.*[57]

Biochemical profiling through employing traditional and advanced analytical techniques has been widely utilised in microbial identification and classification, providing insights into metabolic capabilities and phenotypic traits. The IR-based findings in this study highlighted the biochemical diversity among periodontal bacteria, aligning with literature reports on their distinct cellular compositions. The observed spectral variations and classification patterns reflect key differences in macromolecular content, including proteins, lipids, and polysaccharides, which underpin species-specific metabolic adaptations and functional roles within the periodontal microbiome. However, since FT-IR spectroscopy provides information primarily about functional groups rather than specific molecular identities, it remains challenging to assign the exact biomolecules responsible for the observed clustering. This highlights a gap in the current understanding and emphasising the need for complementary techniques to achieve a more precise molecular characterisation. Another limitation of this study is the nature of O-PTIR measurements requiring manual data collection, which in fact poses practical limitations. In our study, this resulted in the acquisition of data from only 55 cells, and due to the low signal-to-noise ratio, we were able to include only eight individual cells per species. Nevertheless, within this dataset, we can clearly observe heterogeneity within the community, as exemplified by *F. nucleatum*. In contrast, this heterogeneity is not apparent in the FT-IR results, as the larger beam size averages out these intra-population variations. However, given the high spatial resolution of O-PTIR, future studies should include a larger number of individual cells to better account for intra-species heterogeneity. On the other hand, the single-cell sensitivity demonstrated by O-PTIR suggests promising potential for direct application to clinical samples, such as saliva. As it has been previously described by Barros *et al.*, chronic periodontitis is associated with increased permeability of the junctional epithelium and a significant rise in gingival crevicular fluid (GCF) flow rate.[58] GCF leakage into saliva makes it a valuable target for the detection of the subgingival pathogens.[5] Such bacteria are known for their fastidious nature and typically require prolonged culturing, which complicates traditional detection methods. By enabling biochemical profiling without extensive cultivation, O-PTIR could significantly enhance the rapid detection of such pathogens. However, further method optimisation is needed to address the complexity of the clinical samples, including minimising background interference and distinguishing between multiple bacterial species within mixed populations. With these advancements, O-PTIR could become a powerful tool for rapid and less invasive clinical diagnostics.

Conclusion

This study explored the biochemical composition differences among nine periodontal bacteria including *P. gingivalis*, *S. anginosus*, *S. constellatus*, *S. gordonii*, *S. sanguinis*, *S. oralis*, *A. israelii*, *F. nucleatum*, and *P. micra*, which were representatives of some of the subgingival plaque complexes. Combined with chemometrics, they were analysed at the microbial-community level using FT-IR spectroscopy, and MALID-TOF-MS to allow for classification and differentiation of these clinically significant species. To investigate the heterogenicity between single cells within the community, and explore the tantalising potential of bypassing the prolonged culturing requirements of some of these anaerobic microorganisms, O-PTIR spectroscopy was also employed. The findings displayed substantial variations in biochemical composition across the strains studied, with three exhibiting notably higher ratios of protein and lipid to polysaccharides than the remaining strains *P. gingivalis*, *F. nucleatum*, and *S. oralis*, while *S. anginosus* demonstrated the lowest observed ratio. IR clustering patterns were found comparable to MALDI-MS results. Furthermore, according to MALID-MS, around twenty *m*/*z* features were found to influence the classification pattern of *A. israelii*, *P. gingivalis*, *F. nucleatum*, and *S. oralis*, increasing the potential of using the defined *m*/*z* values as biomarkers for their identification. Such results suggest future research to identify the molecular constituents underlying the observed *m*/*z* values. It is recommended to employ methods that could offer deeper insights into the variability of protein and lipid profiles among these bacteria, such as using liquid chromatography mass spectrometry (LC-MS) for lipid profiling and proteomic pipelines for protein analysis. Overall, this study advances our understanding of the biochemical diversity present in periodontitis-associated bacteria and broadens the analytical approaches available for future research on their pathogenicity, which may guide future targeted therapeutic strategies.

Author contributions

JA: experimental design, sample collection and preparation, MALDI, FT-IR and O-PTIR data analysis, data interpretation, and manuscript writing; DS: O-PTIR data collection; SR: MALDI training and instrument automation; YX: assistance

with MALDI data analysis; HA and KAW: data interpretation, and manuscript preparation; HM: principal investigator, experimental design, data interpretation, and manuscript preparation. All authors reviewed and approved the final manuscript.

Conflicts of interest

The authors declare that they have no competing interests.

Data availability

All data generated in this study will be made available upon request.

The Supplementary Information file associated with this article contains MALDI-TOF calibration and optimisation data, *m/z* spectra for periodontal bacteria, as well as FTIR and O-PTIR spectral datasets. All files are available *via* the Royal Society of Chemistry website under the article's DOI and support's the findings presented in the main manuscript. See DOI: https://doi.org/10.1039/d5an00507h.

Acknowledgements

HM would like to thank the University of Liverpool, and Analytical Chemistry Trust Fund (ACTF) and Community for Analytical Measurement Science (CAMS) for funding and support (600310/22/09). JA was funded by a PhD scholarship from King Khalid University, Kingdom of Saudi Arabia. MALDI-TOF-MS work made use of shared equipment located at the Materials Innovation Factory, created as part of the UK Research Partnership Innovation Fund (Research England) and co-funded by the Sir Henry Royce Institute. We would like to acknowledge Roy Goodacre, Najla AlMasoud, and members of Centre of Metabolomics Research for their valuable advice throughout this work.

References

1 F. A. Caranza, M. G. Newman and I. Glickman's, in *Clinical periodontology*, Philadelphia, 1996, p. 470.
2 M. S. Tonetti, H. Greenwell and K. S. Kornman, Staging and grading of periodontitis: Framework and proposal of a new classification and case definition, *J. Periodontol.*, 2018, **89**(Suppl 1), S159–S172.
3 S. S. Socransky, A. D. Haffajee, M. A. Cugini, C. Smith and R. L. Kent, Microbial complexes in subgingival plaque, *J. Clin. Periodontol.*, 1998, **25**(2), 134–144.
4 C. M. da Silva-Boghossian, R. M. do Souto, R. R. Luiz and A. P. V. Colombo, Association of red complex, A. actinomycetemcomitans and non-oral bacteria with periodontal diseases, *Arch. Oral Biol.*, 2011, **56**(9), 899–906.
5 J. Albahri, H. Allison, K. A. Whitehead and H. Muhamadali, The role of salivary metabolomics in chronic periodontitis: bridging oral and systemic diseases, *Metabolomics*, 2025, **21**(1), 24.
6 S. Gaur and R. Agnihotri, A lzheimer's disease and chronic periodontitis: Is there an association?, *Geriatr. Gerontol. Int.*, 2015, **15**(4), 391–404.
7 I. Olsen, D. B. Kell and E. Pretorius, Is Porphyromonas gingivalis involved in Parkinson's disease?, *Eur. J. Clin. Microbiol. Infect. Dis.*, 2020, **39**(11), 2013–2018.
8 S. Lapo, C. Leonardo, B. Luigi, N. Michele, S. Sergio, C. Calogero Lino, *et al.*, Association between chronic kidney disease and periodontitis. A systematic review and metanalysis, *Oral Dis.*, 2023, **29**(1), 40–50.
9 I. L. Chapple, R. Genco and Working Group 2 of the Joint EFP/AAP Workshop*, Diabetes and periodontal diseases: consensus report of the Joint EFP/AAP Workshop on Periodontitis and Systemic Diseases, *J. Clin. Periodontol.*, 2013, **40**(Suppl 14), S106–S112.
10 J. Botelho, V. Machado, Y. Leira, L. Proença, L. Chambrone and J. J. Mendes, Economic burden of periodontitis in the United States and Europe: An updated estimation, *J. Periodontol.*, 2022, **93**(3), 373–379.
11 L. Sedghi, V. DiMassa, A. Harrington, S. V. Lynch and Y. L. Kapila, The oral microbiome: Role of key organisms and complex networks in oral health and disease, *Periodontology 2000*, 2021, **87**(1), 107–131.
12 L. Abusleme, A. Hoare, B. Y. Hong and P. I. Diaz, Microbial signatures of health, gingivitis, and periodontitis, *Periodontology 2000*, 2021, **86**(1), 57–78.
13 S. S. Arbefeville, T. T. Timbrook and C. D. Garner, Evolving strategies in microbe identification-a comprehensive review of biochemical, MALDI-TOF MS and molecular testing methods, *J. Antimicrob. Chemother.*, 2024, **79**(Supplement_1), i2–i8.
14 I. Gajic, J. Kabic, D. Kekic, M. Jovicevic, M. Milenkovic, D. Mitic Culafic, *et al.*, Antimicrobial Susceptibility Testing: A Comprehensive Review of Currently Used Methods, *Antibiotics*, 2022, **11**(4), 427.
15 Z. Movasaghi, S. Rehman and D. I. ur Rehman, Fourier Transform Infrared (FTIR) Spectroscopy of Biological Tissues, *Appl. Spectrosc. Rev.*, 2008, **43**(2), 134–179.
16 M. Wenning and S. Scherer, Identification of microorganisms by FTIR spectroscopy: perspectives and limitations of the method, *Appl. Microbiol. Biotechnol.*, 2013, **97**(16), 7111–7120.
17 S. Ahmed, J. Albahri, S. Shams, S. Sosa-Portugal, C. Lima, Y. Xu, *et al.*, Rapid Classification and Differentiation of Sepsis-Related Pathogens Using FT-IR Spectroscopy, *Microorganisms*, 2024, **12**(7), 1415.
18 C. Lima, H. Muhamadali and R. Goodacre, Simultaneous Raman and Infrared Spectroscopy of Stable Isotope Labelled, *Sensors*, 2022, **22**(10), 3928.
19 D. Helm, H. Labischinski, G. Schallehn and D. Naumann, Classification and identification of bacteria by Fourier-transform infrared spectroscopy, *J. Gen. Microbiol.*, 1991, **137**(1), 69–79.

20 M. A. Mohamed, J. Jaafar, A. F. Ismail, M. H. D. Othman and M. A. Rahman, Chapter 1 – Fourier Transform Infrared (FTIR) Spectroscopy, in *Membrane Characterization*, ed. N. Hilal, A. F. Ismail, T. Matsuura and D. Oatley-Radcliffe, Elsevier, 2017, pp. 3–29.
21 A. A. Enders, N. M. North, C. M. Fensore, J. Velez-Alvarez and H. C. Allen, Functional Group Identification for FTIR Spectra Using Image-Based Machine Learning Models, *Anal. Chem.*, 2021, **93**(28), 9711–9718.
22 X. Zhang, F. Yang, J. Xiao, H. Qu, N. F. Jocelin, L. Ren, *et al.*, Analysis and comparison of machine learning methods for species identification utilizing ATR-FTIR spectroscopy, *Spectrochim. Acta, Part A*, 2024, **308**, 123713.
23 A. El Orche, A. Cheikh, J. B. Johnson, O. Elhamdaoui, S. Jawhari, F. M. El Abbes, *et al.*, A Novel Approach for Therapeutic Drug Monitoring of Valproic Acid Using FT-IR Spectroscopy and Nonlinear Support Vector Regression, *J. AOAC Int.*, 2023, **106**(4), 1070–1076.
24 H. Ouhaddouch, A. Cheikh, M. O. B. Idrissi, M. Draoui and M. Bouatia, FT-IR Spectroscopy Applied for Identification of a Mineral Drug Substance in Drug Products: Application to Bentonite, *J. Spectrosc.*, 2019, **2019**(1), 2960845.
25 D. Finlayson, C. Rinaldi and M. J. Baker, Is Infrared Spectroscopy Ready for the Clinic?, *Anal. Chem.*, 2019, **91**(19), 12117–12128.
26 B. Yang, C. Chen, C. Chen, H. Cheng, Z. Yan, F. Chen, *et al.*, Detection of breast cancer of various clinical stages based on serum FT-IR spectroscopy combined with multiple algorithms, *Photodiagn. Photodyn. Ther.*, 2021, **33**, 102199.
27 A. Badawi and M. G. Althobaiti, Effect of Cu-doping on the structure, FT-IR and optical properties of Titania for environmental-friendly applications, *Ceram. Int.*, 2021, **47**(8), 11777–11785.
28 G. Della Ventura, F. Bellatreccia, A. Marcelli, M. Cestelli Guidi, M. Piccinini, A. Cavallo, *et al.*, Application of micro-FTIR imaging in the Earth sciences, *Anal. Bioanal. Chem.*, 2010, **397**(6), 2039–2049.
29 J. Reffner, Advances in infrared microspectroscopy and mapping molecular chemical composition at submicrometer spatial resolution, *Spectroscopy*, 2018, **33**(9), 12–17.
30 J. Xu, X. Li, Z. Guo, W. E. Huang and J. X. Cheng, Fingerprinting Bacterial Metabolic Response to Erythromycin by Raman-Integrated Mid-Infrared Photothermal Microscopy, *Anal. Chem.*, 2020, **92**(21), 14459–14465.
31 O. Klementieva, C. Sandt, I. Martinsson, M. Kansiz, G. K. Gouras and F. Borondics, Super-Resolution Infrared Imaging of Polymorphic Amyloid Aggregates Directly in Neurons, *Adv. Sci.*, 2020, **7**(6), 1903004.
32 M. Kansiz, L. M. Dowling, I. Yousef, O. Guaitella, F. Borondics and J. Sulé-Suso, Optical Photothermal Infrared Microspectroscopy Discriminates for the First Time Different Types of Lung Cells on Histopathology Glass Slides, *Anal. Chem.*, 2021, **93**(32), 11081–11088.
33 C. Lima, H. Muhamadali, Y. Xu, M. Kansiz and R. Goodacre, Imaging Isotopically Labeled Bacteria at the Single-Cell Level Using High-Resolution Optical Infrared Photothermal Spectroscopy, *Anal. Chem.*, 2021, **93**(6), 3082–3088.
34 S. Shams, C. Lima, Y. Xu, S. Ahmed, R. Goodacre and H. Muhamadali, Optical photothermal infrared spectroscopy: A novel solution for rapid identification of antimicrobial resistance at the single-cell level, *Front. Microbiol.*, 2023, **14**, 1077106.
35 A. Spadea, J. Denbigh, M. J. Lawrence, M. Kansiz and P. Gardner, Analysis of Fixed and Live Single Cells Using Optical Photothermal Infrared with Concomitant Raman Spectroscopy, *Anal. Chem.*, 2021, **93**(8), 3938–3950.
36 H. Muhamadali, D. Weaver, A. Subaihi, N. AlMasoud, D. K. Trivedi, D. I. Ellis, *et al.*, Chicken, beams, and Campylobacter: rapid differentiation of foodborne bacteria via vibrational spectroscopy and MALDI-mass spectrometry, *Analyst*, 2016, **141**(1), 111–122.
37 N. AlMasoud, H. Muhamadali, M. Chisanga, H. AlRabiah, C. A. Lima and R. Goodacre, Discrimination of bacteria using whole organism fingerprinting: the utility of modern physicochemical techniques for bacterial typing, *Analyst*, 2021, **146**(3), 770–788.
38 B. Fuchs, R. Süss and J. Schiller, An update of MALDI-TOF mass spectrometry in lipid research, *Prog. Lipid Res.*, 2010, **49**(4), 450–475.
39 S. Sauer and M. Kliem, Mass spectrometry tools for the classification and identification of bacteria, *Nat. Rev. Microbiol.*, 2010, **8**(1), 74–82.
40 A. E. Clark, E. J. Kaleta, A. Arora and D. M. Wolk, Matrix-assisted laser desorption ionization-time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology, *Clin. Microbiol. Rev.*, 2013, **26**(3), 547–603.
41 C. S. Stîngu, A. C. Rodloff, H. Jentsch, R. Schaumann and K. Eschrich, Rapid identification of oral anaerobic bacteria cultivated from subgingival biofilm by MALDI-TOF-MS, *Oral Microbiol. Immunol.*, 2008, **23**(5), 372–376.
42 H. Muhamadali, Y. Xu, R. Morra, D. K. Trivedi, N. J. Rattray, N. Dixon, *et al.*, Metabolomic analysis of riboswitch containing *E. coli* recombinant expression system, *Mol. BioSyst.*, 2016, **12**(2), 350–361.
43 T. Tengsuttiwat, N. N. Kaderbhai, J. Gallagher, R. Goodacre and H. Muhamadali, Metabolic Fingerprint Analysis of Cytochrome, *Front. Microbiol.*, 2022, **13**, 874247.
44 H. Martens, J. P. Nielsen and S. B. Engelsen, Light scattering and light absorbance separated by extended multiplicative signal correction. application to near-infrared transmission analysis of powder mixtures, *Anal. Chem.*, 2003, **75**(3), 394–404.
45 J. H. Gross, *Mass Spectrometry: A Textbook*, Springer Nature, Berlin, Heidelberg, 2nd edn, 2006.
46 K. C. Parker, Scoring methods in MALDI peptide mass fingerprinting: ChemScore, and the ChemApplex program, *J. Am. Soc. Mass Spectrom.*, 2002, **13**(1), 22–39.
47 H. Hotelling, Analysis of a complex of statistical variables into principal components, *J. Educ. Psychol.*, 1933, **24**(6), 417.

48 P. S. Gromski, H. Muhamadali, D. I. Ellis, Y. Xu, E. Correa, M. L. Turner, *et al.*, A tutorial review: Metabolomics and partial least squares-discriminant analysis–a marriage of convenience or a shotgun wedding, *Anal. Chim. Acta*, 2015, **879**, 10–23.

49 K. Imai, R. Nemoto, M. Kodana, N. Tarumoto, J. Sakai, T. Kawamura, *et al.*, Rapid and Accurate Species Identification of Mitis Group Streptococci Using the MinION Nanopore Sequencer, *Front. Cell. Infect. Microbiol.*, 2020, **10**, 11.

50 W. Zheng, T. K. Tan, I. C. Paterson, N. V. Mutha, C. C. Siow, S. Y. Tan, *et al.*, StreptoBase: An Oral Streptococcus mitis Group Genomic Resource and Analysis Platform, *PLoS One*, 2016, **11**(5), e0151908.

51 R. McGalliard, H. Muhamadali, N. AlMasoud, S. Haldenby, V. Romero-Soriano, E. Allman, *et al.*, Bacterial discrimination by Fourier transform infrared spectroscopy, MALDI-mass spectrometry and whole-genome sequencing, *Future Microbiol.*, 2024, **19**(9), 795–810.

52 V. Ramakrishnan and S. W. White, Ribosomal protein structures: insights into the architecture, machinery and evolution of the ribosome, *Trends Biochem. Sci.*, 1998, **23**(6), 208–212.

53 N. Singhal, M. Kumar, P. K. Kanaujia and J. S. Virdi, MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis, *Front. Microbiol.*, 2015, **6**, 791.

54 P. R. Murray, What is new in clinical microbiology-microbial identification by MALDI-TOF mass spectrometry: a paper from the 2011 William Beaumont Hospital Symposium on molecular pathology, *J. Mol. Diagn.*, 2012, **14**(5), 419–423.

55 V. P. Richards, S. R. Palmer, P. D. Pavinski Bitar, X. Qin, G. M. Weinstock, S. K. Highlander, *et al.*, Phylogenomics and the dynamic genome evolution of the genus Streptococcus, *Genome Biol. Evol.*, 2014, **6**(4), 741–753.

56 M. Pilarczyk-Zurek, I. Sitkiewicz and J. Koziel, The Clinical View on, *Front. Microbiol.*, 2022, **13**, 956677.

57 H. C. van der Mei, D. Naumann and H. J. Busscher, Grouping of oral streptococcal species using fourier-transform infrared spectroscopy in comparison with classical microbiological identification, *Arch. Oral Biol.*, 1993, **38**(11), 1013–1019.

58 S. P. Barros, R. Williams, S. Offenbacher and T. Morelli, Gingival crevicular fluid as a source of biomarkers for periodontitis, *Periodontology 2000*, 2016, **70**(1), 53–64.