npg

ORIGINAL ARTICLE

The ecological proportion of indigenous bacterial populations in saliva is correlated with oral health status

Toru Takeshita[1], Yoshio Nakano[1], Takashi Kumagai[2], Masaki Yasui[1], Noriaki Kamio[1], Yukie Shibata[1], Susumu Shiota[1] and Yoshihisa Yamashita[1]

[1]*Department of Preventive Dentistry, Kyushu University Faculty of Dental Science, Fukuoka, Japan and* [2]*Hiyoshi Dental Office, Sakata, Japan*

To obtain deeper insights into the etiology of oral disease, an understanding of the composition of the surrounding bacterial environments that lead to health or disease is required, which is attracting increasing attention. In this study, the bacterial compositions in the saliva of 200 subjects aged 15–40 years were depicted as peak patterns by terminal restriction fragment length polymorphism (T-RFLP) analysis of 16S rRNA genes. The subjects were classified into three clusters by partitioning around medoids clustering based on their T-RFLP profiles, and the clinical oral health parameters of the clusters were compared. The clustering of the T-RFLP profiles in this study was mainly based on differences in the abundance distribution of the dominant terminal restriction fragments (TRFs) detected in most of the subjects. Predicted from the sizes of the TRFs, the characteristically more predominant members of each were *Prevotella* and *Veillonella* species in cluster I; *Streptococcus* species in cluster II and *Neisseria*, *Haemophilus* or *Aggregatibacter* species and *Porphyromonas* species in cluster III. The parameters associated with periodontal disease were significantly different among the clusters. Clusters I and II had a higher percentage of sites of periodontal pockets greater than 4 mm than cluster III, and cluster I contained sites exhibiting bleeding on probing more often than cluster II or III; no significant differences were observed in other parameters. These results suggest that the abundance distribution of commensal bacteria in saliva is correlated with periodontal health, and might be involved in the susceptibility of an individual to periodontal disease.

Introduction

In the human oral cavity, microbial communities form as biofilms on various saliva-bathed surfaces (Jenkinson and Lamont, 2005). Two major oral diseases, dental caries and periodontal disease, are infectious diseases that are closely associated with some microbes that are specific to these biofilms. Mutans streptococci are major etiologic agents of dental caries (Loesche, 1986). It is also known that *Porphyromonas gingivalis*, *Tannerella forsythia*, *Treponema denticola* and their complex are prime suspects in the development of periodontitis (Socransky *et al.*, 1998). However, the presence of these organisms in dental plaque is not required for the onset of oral disease. The pathogenic bacteria of oral diseases are regulated by numerous interactions with other indigenous bacteria and the host response. An understanding of both the potential virulence factors of the specific bacteria and the composition of the surrounding bacterial environments that lead to health or disease is important to obtain a deeper insight into the etiology of oral diseases.

A large number of previous studies using conventional methods based on cultivation and culture-independent molecular techniques have identified over 700 bacterial species in the human oral cavity (Moore and Moore, 1994; Aas *et al.*, 2005). In particular, molecular approaches based on cloning and sequencing analyses of bacterial 16S rRNA gene sequences have allowed comprehensive surveys of this complex bacterial community, and revealed the bacterial composition of oral microbiota at various sites (Paster *et al.*, 2001, 2002; Becker *et al.*, 2002; Munson *et al.*, 2002; Kazor *et al.*, 2003; Aas *et al.*, 2005, 2007, 2008; Kumar *et al.*, 2005, 2006; de Lillo *et al.*, 2006; Marchini *et al.*, 2007). However,

66

it is difficult to analyze such a large number of samples using this approach because of the labor and costs involved.

One molecular method, terminal restriction fragment length polymorphism (T-RFLP) analysis, is an effective approach for the rapid assessment and comparison of complex bacterial communities (Liu et al., 1997), including oral microbiota (Sakamoto et al., 2003). The advantage of this method is that the composition of the bacterial community can be viewed as a peak pattern that can be easily compared to large numbers of patterns from other samples using statistical methods (Forney et al., 2004). Bacterial communities from numerous subjects can be classified based on the similarity of their compositions, and the bacterial composition pattern associated with a clinical symptom can be explored. Another advantage of T-RFLP analysis is that the bacterial populations can be putatively identified by comparing the sizes of terminal restriction fragments (TRFs) in a profile with those predicted from known 16S rRNA gene sequences. A number of sequences of bacterial 16S rRNA genes detected in the oral cavity are available in public databases. Furthermore, we have been able to reduce the large error in the sizing of DNA fragments in T-RFLP analysis to more accurately assign bacteria to TRFs (Takeshita et al., 2007).

In this study, the bacterial compositions in the saliva of 200 subjects aged 15–40 years were investigated as peak patterns by T-RFLP analysis. We consider saliva the most suitable sample to provide information on the microbial population in the entire oral cavity, because various bacteria that reside on different surfaces can be recovered from it (Denepitiya and Kleinberg, 1982; Mager et al., 2003a; Kononen et al., 2007). Although the microbial populations in the saliva might be only a mixture of the microbial consortia that exist at various sites in the oral cavity, they are relatively stable over time (Sakamoto et al., 2003; Rasiah et al., 2005). In addition, saliva samples can be collected easily and reproducibly. The subjects were classified based on the T-RFLP patterns of microbial populations in their saliva, and various clinical parameters of the subjects in the different clusters were compared. The aim of this study was to identify the characteristic proportions of indigenous bacterial populations in saliva that are correlated with oral health or disease, and to evaluate the clinical efficacy of clustering the T-RFLP profiles derived from salivary bacterial populations.

Materials and methods

Study population
The study population consisted of 98 women and 102 men aged 15–40 years (mean ± s.d., 28.1 ± 6.3 years). Of the 200 subjects, 130 were patients who visited the Hiyoshi Dental Clinic and 70 were undergraduate and graduate students at the Kyushu University Faculty of Dental Science in Japan. All subjects had at least 23 teeth, and none used dentures. All subjects who participated in this study understood the nature of the research project and provided their informed consent. For each subject, the teeth condition (the number of decayed, missing and filled teeth, DMFT), and the periodontal condition (periodontal pocket depth and bleeding on probing, BOP, at the mesiobuccal and mid-buccal sites determined using a periodontal pocket probe) were assessed for all teeth except for the third molars, following stimulated saliva collection. The DMFT value signifies teeth with caries experience, and represents the caries status of the individual. The periodontal pocket is the pathologic space between the gingiva and the tooth root, and its depth is used for the clinical diagnosis of periodontal disease. The depth of these spaces is normally 1–3 mm for periodontally healthy individuals, but deepens as supporting connective tissue and alveolar bone are destroyed by persistent gingival inflammation (Pihlstrom et al., 2005). BOP is bleeding caused by picking the inside of the periodontal pocket with a periodontal pocket probe and is a visible symptom of gingival inflammation, suggestive of an active phase of periodontal disease.

Saliva sample collection and DNA extraction
The subjects were asked to bite on paraffin wax for 5.5 min, and the stimulated saliva sample produced during the last 5 min was collected in sterile plastic tubes. The samples were stored at −80 °C until analysis. Bacterial DNA extraction was performed as described earlier (Takeshita et al., 2007).

T-RFLP analysis
Amplification of 16S rRNA genes. To analyze the TRFs for full-length fragments lacking a restriction site, a primer set and DNA polymerase were selected to produce small PCR amplicons with blunt ends. From each sample, internal regions of 16S rRNA genes of about 800 bp were amplified using the universal forward primer 8F (5′-AGAGTTTGATYMTGGCTCAG-3′; Pei et al., 2004) labeled at the 5′ end with 6-carboxyfluorescein (6-FAM) and the universal reverse primer 806R (5′-GGACTACCRGGGTATCTAA-3′; F16 (Paster et al., 2001) was slightly modified) labeled at the 5′ end with hexachlorofluorescein (HEX). PCR was performed using KOD DNA polymerase (Toyobo, Osaka, Japan) as described earlier (Takeshita et al., 2007), and the 16S rRNA gene amplicons were purified using a Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA) according to the manufacturer's instructions. The DNA was stored at −20 °C until analysis.

Redefinition of the sizes of fragments in the internal standard. As an internal standard, GeneScan-500 ROX standard (Applied Biosystems, Foster City, CA,

USA), which contains 16 6-carboxy-rhodamine (ROX)-labeled DNA fragments and 6 additional DNA fragments (of 541, 600, 663, 730, 799 and 861 bp) labeled at the 5′ end with ROX, was used. The six additional fragments were internal regions of the 16S rRNA gene of *Porphyromonas gingivalis* W83 amplified using the 5′-ROX-labeled forward primer 8F and six non-labeled reverse primers (542m, 601m, 664m, 731m, 806R and 861m; Takeshita *et al.*, 2007), according to the PCR conditions and purification protocol as described above. To improve the accuracy of the observed sizes of the TRFs, the sizes of the DNA fragments in the internal standard were redefined as described earlier (Takeshita *et al.*, 2007). In brief, the internal standard (that is, GeneScan-500 ROX with six additional fragments) was preliminarily subjected to capillary electrophoresis along with a size marker containing TRFs of various lengths derived from the 16S rRNA gene of a bacterium. The size of each fragment in the internal standard was estimated based on the relationship between the size and mobility of the TRFs derived from the 16S rRNA gene of a bacterium (for example, *P. gingivalis* as used here), and the molecular weight (MW) of each fragment was used as the unit of fragment size instead of number of bases.

The FAM size marker was prepared as described earlier (Takeshita *et al.*, 2007), except that D88 was substituted by 8F. The sizes of the 22 ROX-labeled fragments in the internal standard redefined using the FAM size marker (MW of 13 735, 17 287, 25 138, 32 521, 44 808, 47 877, 50 839, 63 396, 77 730, 94 447, 106 290, 109 709, 126 409, 141 718, 154 581, 157 360, 169 129, 187 642, 207 219, 228 308, 249 981 and 269 092) were used to determine the sizes of the 6-FAM-labeled TRFs. The HEX size marker was prepared as follows: seven PCR amplicons were produced using the reverse primer 806R labeled at the 5′ end with HEX and seven non-labeled forward primers (r354m (5′-GAATAACGGGCGATACGAG-3′), r430m (5′-CAATGGGCGAGAGCCTGAA-3′), r489m (5′-ACACTGGTACTGAGACACG-3′), r563m (5′-TC CCATTAGCTTGTTGGTG-3′), r640m (5′-ACGGACT AAAACCGCATAC-3′), r725m (5′-AGCTTGCTAAG GTTGATGG-3′), those were designed in this study, and 8F). In addition, DNA fragments separately digested with a restriction enzyme (*Alu*I, *Aci*I, *Hha*I, *Mse*I or *Sac*II) were produced from the amplicon using 8F- and HEX-labeled 806R. These 12 fragments and HEX-labeled 806R were mixed, and used as the HEX size marker. The sizes of the 22 fragments in the internal standard redefined using HEX size marker (13 759, 17 114, 24 903, 32 251, 44 234, 47 201, 50 063, 62 407, 76 665, 93 226, 105 014, 108 400, 124 576, 139 576, 152 205, 154 916, 166 470, 184 819, 204 342, 225 328, 247 172 and 265 403) were used to determine the sizes of the HEX-labeled TRFs.

T-RFLP. Purified DNA products (3 μl) were digested with 2.5 U of *Hae*III in a total volume of 10 μl for 2.5 h at 37 °C. The restriction digest products (2 μl) were mixed with 10 μl of deionized formamide and 1 μl of the internal standard described above. The samples were denatured and electrophoresed as described earlier (Takeshita *et al.*, 2007). Electropherograms were analyzed with GeneMapper version 4.0 (Applied Biosystems), and the fragment sizes were estimated using the Local Southern method. TRFs with a peak area of less than 0.5% of the total area were excluded from the analysis.

Cluster analysis
Cluster analysis was performed using two T-RFLP profiles obtained per subject using two different fluorescent dyes (6-FAM and HEX). The estimated MWs of all TRFs from each subject were aligned, and TRFs with MWs that differed by 77 or less were considered identical. The aligned T-RFLP profiles composed of the sizes of the TRFs and the percentage of the peak area in each profile were clustered using partitioning around medoids clustering (PAM; Kaufman and Rousseeuw, 1990) with R version 2.6.0 (R Development Core Team, 2007) and the 'cluster' library (Maechler *et al.*, 2005). PAM provides for each subject the cluster to which the subject belongs, the closest neighbor cluster and a silhouette width that measures the degree of fitness of a subject to its cluster. The average silhouette width taken over all objects of the clustering, a measure of the goodness of clustering (Kaufman and Rousseeuw, 1990), indicated that the appropriate number of clusters in this study was three.

The profiles for the 200 subjects were sorted into clusters and visualized as a gel-like image using the R library 'gregmisc' (Warnes, 2008). In addition, the T-RFLP profiles were analyzed by principal components analysis (PCA). PCA allows the visualization of correlations in data sets by compressing information in a low number of dimensions. We obtained a biplot, of which the x and y axes represent the first and second principal components, respectively, to find important components correlated with clustering. The results from all 200 subjects were plotted as dots in this two-dimensional display, and the original variables (each TRF) were represented by arrows. The direction and length of the arrows indicate how each variable contributes to the first two components in the biplot. PCA was performed by using the R library 'ade4' (Dray and Dufour, 2007).

Construction of a database of 16S rRNA gene sequences of oral bacteria and assignment to TRFs
To assign oral bacteria to the peaks in the T-RFLP profiles, nucleotide sequence data for the 16S rRNA genes of bacteria detected in the oral cavity were collected. A total of 1067 sequences of oral bacterial strains identified using cloning and sequencing

analysis (Paster *et al.*, 2001, 2002; Becker *et al.*, 2002; Munson *et al.*, 2002; Kazor *et al.*, 2003; Aas *et al.*, 2005, 2007, 2008; Kumar *et al.*, 2005, 2006; de Lillo *et al.*, 2006; Marchini *et al.*, 2007) were extracted from the Ribosomal Database Project (RDP-II) database (Cole *et al.*, 2007). Of the 1067 sequences in the database, 201 were excluded because the size of the TRF could not be estimated due to insufficient sequence information, resulting in a usable database of 866 sequences from 655 species or phylotypes. Candidate bacterial species corresponding to the TRF sizes were selected from the database using TRFMAW, a modified version of TRFMA (Nakano *et al.*, 2006), for the primer set used in this study. On the basis of our previous work, the matching window was set to a MW of ±600 (Takeshita *et al.*, 2007).

Quantification of total bacteria by real-time PCR

To compare the amount of total bacteria per milliliter of saliva among the clusters, quantitative real-time PCR was performed using a QuantiFast SYBR Green PCR kit (Qiagen, Hilden, Germany) in a StepOne Real-Time PCR System (Applied Biosystems) according to the manufacturer's instructions. The bacterial universal primers 806F (5′-TTAGAT ACCCYGGTAGTCC-3′) and 926R (5′-CCGTCAATT YCTTTGAGTTT-3′) were used. Of the 1067 sequences described above, 902 sequences contained these primer sequences. Nucleotide sequences of some species of *Acinetobacter*, *Actinobaculum*, *Aggregatibacter*, *Campylobacter*, *Capnocytophaga*, *Enterobacter*, *Eubacterium*, *Granulicatella*, *Haemophilus*, *Leptotrichia*, *Mycoplasma*, *Porphyromonas*, *Prevotella*, *Pseudomonas*, *Selenomonas*, *Sphingomonas*, *Streptococcus*, *Treponema*, *Veillonella* and TM7 registered in RDP-II had mismatches with these primers. The cycling conditions were 95 °C for 10 min followed by 40 cycles of 95 °C for 3 s and 60 °C for 30 s. DNA melting curves for the 16S rRNA amplicons were assessed for any putative PCR artifacts or nonspecific PCR products. The relative amounts of total bacteria were calculated using the comparative C_t method, and DNA extracted from *Streptococcus mutans* Xc was cultured as described earlier (Shibata *et al.*, 2003) was used as a real-time PCR control.

Cloning and sequencing analysis

PAM clustering determines the object corresponding to the center of each cluster, or medoid, the dissimilarity of which to all objects in the cluster is minimal. Samples of three subjects corresponding to the medoid of each cluster were analyzed by cloning and sequencing. Internal regions of the 16S rRNA genes were amplified using non-labeled 8F and 806R primers, inserted into the vector pBluescript SK II (+) (Stratagene, La Jolla, CA, USA), and transformed into *Escherichia coli* DH5α (Sambrook

and Russell, 2001) as described earlier (Kumar *et al.*, 2005). The nucleotide sequences of the inserts were determined using the M13 (−40) forward and reverse primers. Sequence analysis was performed as described earlier (Takeshita *et al.*, 2007), and a novel phylotype was defined as a sequence that differed from the closest GenBank entry by >2%. Sequence data for novel species were submitted to the DNA Data Bank of Japan (DDBJ) under accession numbers AB436407–AB436409.

Statistical analyses

The Steel–Dwass test, a non-parametric multiple comparison procedure, was conducted to assess differences in bacterial diversity, age, number of missing teeth, number of DMFT, percentage of sites with periodontal pockets (probing depth deeper than 4 mm), percentage of sites with BOP, number of total bacteria and peak area proportion of each TRF. The degree of correlation between the saliva bacterial composition and gender or smoker distribution was evaluated using Fisher's exact test. The relationship between the percentage of sites with periodontal pockets and the amount of total bacteria in the saliva of subjects and the relationship between the peak area proportion of each of the 12 dominant TRFs and oral health status (percentage of sites with periodontal pockets, BOP and the DMFT value) were assessed with the Pearson's product moment correlation coefficient. The statistical significance was set at $P < 0.05$ to denote a statistically significant difference. All statistical analyses were carried out with R 2.6.0 on Mac OS X.

Results

Selection of restriction enzymes for better phylogenetic resolution of oral bacteria

Multiple distinct species might generate indistinguishable restriction fragments of 16S rRNA gene in T-RFLP analysis. To select restriction enzymes that result in more detailed information on oral bacterial populations, we counted the number of unique combinations of both 5′ TRFs predicted from 16S rRNA gene sequences deposited in our database of 866 oral bacterial strains (655 species or phylotypes). Cleavage with *Hae*III resulted in the greatest variation (361 combinations) of the 12 four-base restriction enzymes tested (Table 1) and was therefore selected for subsequent T-RFLP analysis.

Classification of T-RFLP profiles and their characteristics in each cluster

T-RFLP profiles were generated for saliva samples of 200 subjects aged 15–40 years. The overall profiles contained 110 distinct peaks, 69 peaks in the 6-FAM profiles and 41 peaks in the HEX profiles. The T-RFLP profiles, which consist of the MWs of the TRFs and the percentage of the total peak area for

each, were classified into three clusters using PAM clustering. The profiles were sorted into each cluster and visualized as gel-like images, as shown in Figure 1. Several TRFs were detected as dominant peaks in most subjects, but drastic differences were observed in their abundances among the three clusters (Figure 1 and Table 2). The PCA biplot of the first two principal components (Figure 2) indicated that the clustering of the T-RFLP profiles in this study was particularly influenced by these

dominant TRFs. Subjects of cluster I were localized in the lower left area of this diagram, and the TRFs of no. 19 FAM, no. 33 FAM, no. 16 HEX and no. 22 HEX had large loadings (>0.1 absolute value on the first and second principal components) in this direction. The peak area proportions of these TRFs were significantly greater than those in clusters II and III (Table 2). Our accurate T-RFLP analysis established earlier (Takeshita et al., 2007) allowed us to putatively identify the origins of the TRFs without cloning and sequencing. These TRF combinations corresponded to Prevotella species (no. 33 FAM and no. 22 HEX; P. melaninogenica, P. tannerae and five phylotypes) and Veillonella species (no. 19 FAM and no. 16 HEX; V. dispar, V. parvula and two phylotypes) based on the sizes predicted from the 16S rRNA sequences in the above databases. Subjects of cluster II were localized in the right area of the PCA biplot diagram; the TRFs of no. 47 FAM and no. 32 HEX had large loadings in this direction (Figure 2), and the peak area proportions of these TRFs were significantly greater in cluster II than those in clusters I or III (Table 2). This combination corresponded putatively to Streptococcus species (S. cristatus, S. mitis, S. oligo fermentans, S. oralis, S. parasanguinis, S. peroris, S. pneumoniae, S. salivarius, S. sanguinis, S. suis and 28 phylotypes). Subjects of cluster III were localized at the upper left of the diagram, and TRFs of no. 17 FAM, no. 65 FAM, no. 12 HEX and no. 24

Table 1 Number of unique combinations of both 5′ terminal restriction fragments derived from 866 oral bacterial strains (655 species or phylotypes) using four-base cutters

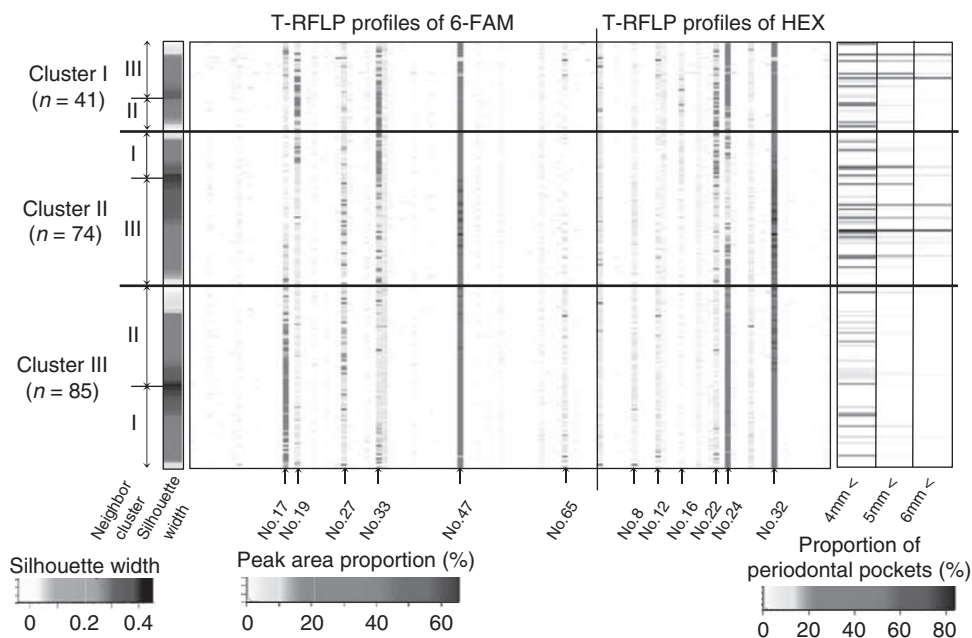| Restriction enzyme	Recognition site	Number of combinations
HaeIII	GGCC	361
AluI	AGCT	345
MspI	CCGG	325
Tsp509I	AATT	310
HhaI	GCGC	296
RsaI	GTAC	295
MaeI	CTAG	273
BstUI	CGCG	258
MseI	TTAA	252
TaqI	TCGA	186
FatI	CATG	178
HpyCH4V	TGCA	128

Figure 1 Terminal restriction fragment length polymorphism (T-RFLP) peak patterns from 200 subjects (rows) visualized as a gel-like image. The patterns were sorted into three clusters, separated based on their nearest neighbor cluster in each cluster, and ordered according to each silhouette width, which measures the degree of fitness of an object to its cluster, within the cluster. The silhouette width of each sample (displayed as gray scale bars) and the nearest neighbor cluster are indicated to the left of the image. The area proportion in each T-RFLP profile (6-carboxyfluorescein (6-FAM) and hexachlorofluorescein (HEX)) of individual peaks (69 detected 6-FAM peaks and 41 detected HEX peaks; columns) is represented as the gray scale intensity in each grid. The percentage of sites with periodontal pockets (pocket depths over 4, 5 or 6 mm, respectively) in all surveyed sites of each subject are displayed as gray scale bars to the right of the image (See online version for color figure).

Table 2 Comparison of peak area proportions of dominant terminal restriction fragments (TRFs) among three clusters

Number of TRF	Peak area proportion (%, average ± s.d.)			Statistical differences[a]
	Cluster I	Cluster II	Cluster III	
No. 17 FAM	7.63 ± 4.24	5.07 ± 3.94	**17.31 ± 5.91**	III > I > II
No. 19 FAM	**14.56 ± 7.89**	7.64 ± 4.32	4.95 ± 3.12	I > II > III
No. 27 FAM	6.69 ± 3.33	9.23 ± 4.06	9.56 ± 4.60	III > I, II > I
No. 33 FAM	**15.53 ± 5.14**	13.16 ± 5.36	8.38 ± 3.74	I > II > III
No. 47 FAM	22.71 ± 5.95	**39.07 ± 7.24**	30.95 ± 6.82	II > III > I
No. 65 FAM	3.74 ± 5.57	1.92 ± 2.80	**5.37 ± 3.67**	III > I > II
No. 8 HEX	2.27 ± 1.52	4.22 ± 2.81	4.78 ± 3.07	III > I, II > I
No. 12 HEX	4.54 ± 4.44	2.76 ± 2.76	**5.25 ± 3.33**	III > I > II
No. 16 HEX	**5.56 ± 3.89**	3.24 ± 2.82	1.15 ± 1.23	I > II > III
No. 22 HEX	**12.75 ± 3.83**	10.44 ± 5.22	5.38 ± 3.33	I > II > III
No. 24 HEX	20.00 ± 5.59	11.26 ± 5.14	**24.03 ± 6.05**	III > I > II
No. 32 HEX	27.98 ± 6.50	**45.22 ± 6.74**	37.67 ± 6.21	II > III > I

Abbreviations: FAM, carboxyfluorescein; HEX: hexachlorofluorescein.
[a]Statistical differences were evaluated among three clusters by the Steel–Dwass test, non-parametric multiple comparison. P-values < 0.05 were considered statistically significant.
Peak area proportions that were significantly greater than the other two clusters were shown in bold.
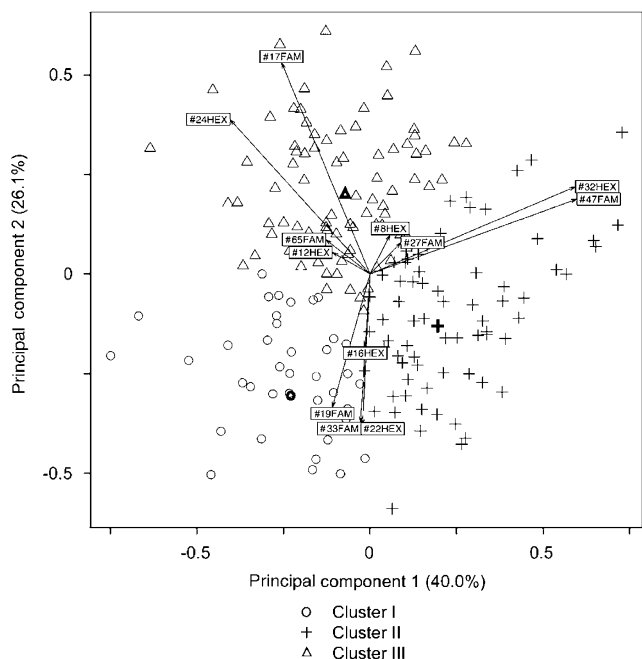

Figure 2 Principal components analysis (PCA) biplot showing the relationship between variables (each terminal restriction fragment (TRF), arrows) and 200 terminal restriction fragment length polymorphism (T-RFLP) profiles classified into three clusters by partitioning around medoids clustering (three types of dots). Only 12 TRFs with large loadings (> 0.1 in absolute value on the first and second principal components) of 110 TRFs were selected and represented. These two components explain 66.1% of the point variability. Each medoid of the three clusters is shown in bold.

genera *Neisseria* (*N. cinerea*, *N. elongata*, *N. subflava*, *N. weaveri* and six phylotypes), *Haemophilus* (*H. influenzae*, *H. parainfluenzae*, *H. pittmania*, *H. quentini* and four phylotypes), and *Aggregatibacter* (*A. actinomycetemcomitans* and *A. segnis*), and the combination of no. 65 FAM and no. 12 HEX corresponded to the genus *Porphyromonas* (eight phylotypes with no identified species). In addition, the TRFs no. 27 FAM and no. 8 HEX had large loadings in the region between clusters II and III (the direction opposite to that of cluster I; Figure 2). The peak area proportions of these TRFs were significantly greater in clusters II and III than in cluster I (Table 2). The combination corresponds to the genus *Rothia* (*R. mucilaginosa* and one phylotype). The peak area proportion of each of the 12 dominant TRFs had no or very low correlation coefficients ($r < 0.25$) with periodontal indices (percentage of sites with periodontal pockets deeper than 4 mm and BOP) and caries status (the DMFT value).

Cloning and sequencing analysis
Samples from three subjects whose T-RFLP profiles represented the medoids of each cluster (M1, M2 and M3) and whose 6-FAM peak patterns are shown in Figure 3 were analyzed by cloning and sequencing. Sequence data for fragments of about 800 bp were obtained for about 150 clones from each sample, for a total of 505 clones. A total of 57 species, including 23 phylotypes, that were identified from these 505 clones were assigned to 23 of the 69 peaks in the 6-FAM T-RFLP profiles, including all of the dominant peaks that were characteristic in each cluster, even though the matching window was set to a MW of ±600 (Table 3). In addition, the peak area proportion of each TRF reflected well the abundance of the number of clones corresponding

HEX had large loadings in this direction (Figure 2). The peak area proportions of these TRFs in cluster III were significantly greater than those in clusters I and II (Table 2). This combination of no. 17 FAM and no. 24 HEX corresponded putatively to the
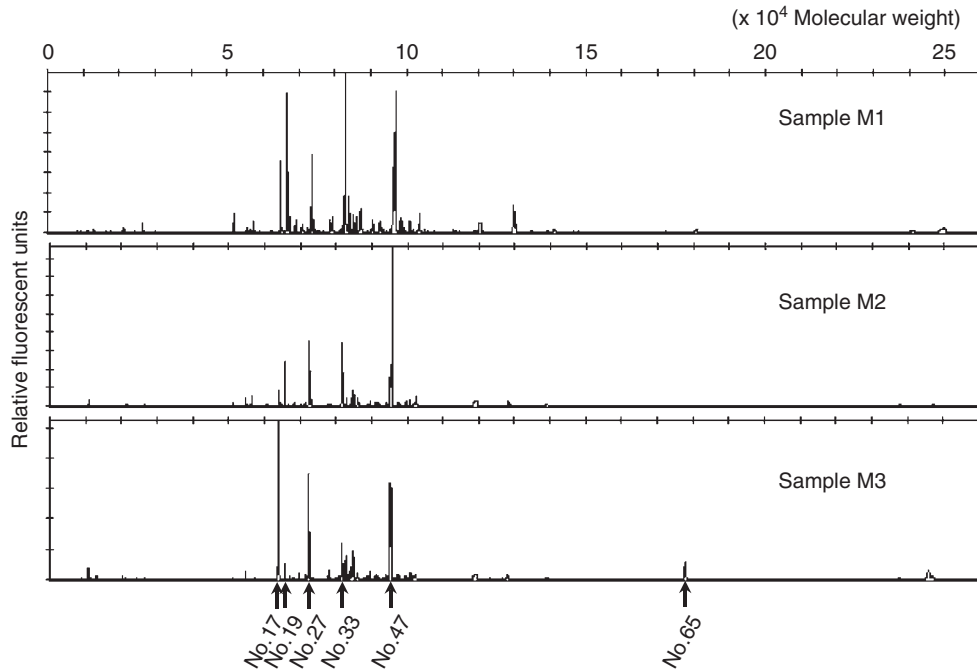
Figure 3 The 6-carboxyfluorescein (6-FAM) terminal restriction fragment length polymorphism (T-RFLP) peak patterns from a sample of subjects whose T-RFLP profiles correspond to the medoids of each cluster. The peak numbers of the characteristic peaks in each cluster described in Figure 1 are indicated under the electropherograms.

to each TRF. The bacterial species corresponding to the six combinations of TRFs (no. 17 FAM and no. 24 HEX, no. 19 FAM and no. 16 HEX, no. 27 FAM and no. 8 HEX, no. 33 FAM and no. 22 HEX, no. 47 FAM and no. 32 HEX, and no. 65 FAM and no. 12 HEX) were predominantly isolated from three subjects.

Comparison of the clinical parameters of the three clusters classified based on T-RFLP profiles
The general and oral clinical parameters of subjects in the three clusters were compared (Table 4). The age, gender and smoker status distribution did not differ significantly among the three clusters, and the parameters representing tooth condition (number of missing teeth and DMFT) also did not differ. However, significant differences were observed in the periodontal tissue status among the clusters. The percentage of sites with BOP in cluster I was significantly greater than those in clusters II and III, and the percentages of sites with periodontal pockets deeper than 4 mm were significantly greater in clusters I and II than in cluster III. The proportions of subjects with no periodontal pockets and those with over 20% periodontal pockets also differed significantly among the clusters. The amount of total bacteria in the saliva of subjects was significantly greater in cluster I than in clusters II and III. However, the percentage of sites with periodontal pockets demonstrated a low correlation with the amount of total bacteria in each cluster ($r = 0.11$ in cluster I; $r = 0.23$ in cluster II; $r = 0.16$ in

cluster III), and there was no significant correlation between the amount of total bacteria and the percentage of sites with periodontal pockets in clusters I and III.

Discussion

In the present study, we clustered the T-RFLP profiles representing saliva bacterial populations into three groups. Significant differences in the periodontal status (percentage of sites with periodontal pockets and BOP) were found among the three clusters, but there was no distinct difference in the prevalence of dental caries (DMFT or number of missing teeth in younger persons) between any two clusters (Table 4). This result was unexpected, as the most commonly used clinical caries activity assessment kits, Dentocult SM and LB (Orion Diagnostica, Espoo, Finland), utilize saliva as an analyte. In contrast, subgingival plaque was concentrated and used as an analyte in the most recently reported studies of the relationships between oral bacteria and periodontal diseases (Kumar *et al.*, 2005, 2006; Ledder *et al.*, 2007). Periodontal disease is a local inflammatory response in gingival tissues that is caused by a microbial community that becomes mature in each gingival crevice. Although periodontal pathogens are also detectable in saliva (von Troil-Linden *et al.*, 1997; Kononen *et al.*, 2007), the inside of periodontal pockets is the most saliva-inaccessible part of the various oral surfaces. Numerous bacteria that are shed from different

Table 3 Species and phylotypes detected in cloning and sequencing analysis of the medoid of each cluster (M1, M2 and M3)

Peak area proportion of each TRF in T-RFLP profile of 6-FAM (%)			TRF number in the aligned profiles in overall subjects[a]		Species/phylotypes	No. of clones			Percentage of clones in each 6-FAM-labeled TRF (total number)		
M1	M2	M3	FAM	HEX		M1	M2	M3	M1	M2	M3
		0.78	1								
		0.58	3								
			4	2	Actinomyces lingnae	1			0.62 (1)	0.57 (1)	
			4	22	Corynebacterium matruchotii		1				
1.13			6	28	Uncultured Oribacterium sp.; EHFS1_S01f; EU071470	1			0.62 (1)		
1.32			8	22	Prevotella pallens	2			1.25 (2)		
	0.85	0.58	9	2	Actinomyces graevenitzii		1				
			9	2	Actinomyces sp. oral clone DR002; AF385517		1			1.14 (2)	
0.58	1.12		11	2	Actinomyces sp. oral clone CT047; AF385504	4	2	1	2.51 (4)	1.14 (2)	0.58 (1)
			13	24	Dialister sp. oral clone BS095; AF287787	1			0.62 (1)		
			17	24	Neisseria mucosa			1			
			17	24	Neisseria sp. J01; DQ409137	7		20			
5.21	2.19	18.53	17	24	Neisseria sp. R-22841; AJ786809	1		11	5.03 (8)	1.14 (2)	19.88 (34)
			17	24	Haemophilus parainfluenzae			2			
			17	24	Terrahaemophilus aromaticivorans		2				
			19	16	Veillonella dispar	22	8	2			
13.89	10.49	2.58	19	16	Veillonella parvula	3	4		22.01 (35)	7.42 (13)	1.16 (2)
			19	24	Veillonella atypica	10	1				
1.42			21								
2.18	0.69	0.82	22	2	Atopobium parvulum	2			1.25 (2)		
		0.83	24								
			27	2	Actinomyces odontolyticus	6	1				
7.38	11.48	9.64	27	19	Rothia dentocariosa	1		1	6.28 (10)	14.85 (26)	13.45 (23)
			27	8	Rothia mucilaginosa	3	25	22			
			28	19	Eubacterium sulci	1			0.62 (1)		
	0.80		30								
			31	5	KUD011 (Leptotrichia trevisanii; LB11; AY029801 90.6%)	1					

Table 3 Continued

Peak area proportion of each TRF in T-RFLP profile of 6-FAM (%)			TRF number in the aligned profiles in overall subjects[a]		Species/phylotypes	No. of clones			Percentage of clones in each 6-FAM-labeled TRF (total number)		
M1	M2	M3	FAM	HEX		M1	M2	M3	M1	M2	M3
2.94	0.75	1.95	31	12	*Leptotrichia hofstadii*			1	3.14 (5)	1.14 (2)	0.58 (1)
			31	12	*Leptotrichia* sp. oral clone GT018; AY349384	4	2				
			33	2	*Gemella haemolysans*			1			
			33	2	*Gemella sanguinis*	4					
16.09	13.22	9.37	33	22	*Prevotella melaninogenica*	5	12		13.83 (22)	10.28 (18)	0.58 (1)
			33	22	*Prevotella* sp. oral clone BE073; AF385551	12	3				
			33	13	*Prevotella* sp. oral clone FM005; AF432133	1	1				
			33	13	*Prevotella* sp. oral clone GI059; AY349397		2				
			34	28	*Eubacterium* sp. oral clone DO008; AF385508		1				
3.96	4.69	6.65	34	28	*Fusobacterium periodonticum*			2	3.14 (5)	6.28 (11)	5.84 (10)
			34	32	*Abiotrophia para-adiacens*	3	7	2			
			34	32	*Streptococcus* sp. oral strain H3-M2; AF385523	2	3	6			
2.47	2.13	0.61	35	—[b]	*Granulicatella paradiacens*		1		0.62 (1)	0.57 (1)	
			35	12	*Leptotrichia* sp. oral clone EI013; AF385571	1					
			36	5	*Leptotrichia* sp. oral clone BU064; AY008309			1			0.58 (1)
0.80	0.99		39								
1.83	0.82		41								
0.88	1.09	0.80	42	2	TM7 phylum sp. oral clone DR034; AF385520	2			1.25 (2)		
			44	—[b]	KUD012 (*Eubacterium yurii* subsp. *margaretiae* 97%)			1	0.62 (1)		0.58 (1)
			44	28	*Eubacterium* sp. oral clone BE088	1					
			47	32	*Streptococcus cristatus*	1		12			
			47	32	*Streptococcus infantis*	1	3	13			
			47	32	*Streptococcus mitis*	11	5	25			
			47	32	*Streptococcus oralis*	8	3	12			
			47	32	*Streptococcus parasanguinis*	8	9	1			
			47	32	*Streptococcus salivarius*	24	70	6			
27.23	41.43	36.19	47	32	*Streptococcus sanguinis*		2	7	35.22 (56)	55.42 (97)	50.29 (86)
			47	32	*Streptococcus* sp. oral clone BE024; AF385550			2			
			47	32	*Streptococcus* sp. oral clone DN025; AF432131		1				
			47	32	*Streptococcus* sp. oral clone FN051; AF432135	2		3			
			47	32	*Streptococcus* sp. oral strain T1-E5; AF385525		3	2			
			47	32	*Streptococcus* sp. oral strain T4-E3; AF385526	1		3			
			47	32	KUD013 (*Streptococcus parasanguinis* (97.4%))		1				

Table 3 Continued

Peak area proportion of each TRF in T-RFLP profile of 6-FAM (%)			TRF number in the aligned profiles in overall subjects[a]		Species/phylotypes	No. of clones			Percentage of clones in each 6-FAM-labeled TRF (total number)		
M1	M2	M3	FAM	HEX		M1	M2	M3	M1	M2	M3
1.52	0.79		49								
1.04	0.55		50	32	Abiotrophia defectiva			1			0.58 (1)
0.64	0.87		51								
	0.79	0.63	52								
		1.17	53	28	uncultured bacterium X112; AF125207			4			2.33 (4)
1.21	1.92		54								
4.37	1.66	1.74	61	13	Prevotella sp. oral clone BI027; AY005064			1			0.58 (1)
		1.53	64								
1.10		5.01	65	12	Porphyromonas sp. oral clone CW034; AY008310	2		5	1.25 (2)		2.92 (5)
	0.69		66								
0.80			68								
100.00	100.00	100.00				159	175	171	100 (159)	100 (175)	100 (171)

Abbreviations: HEX: hexachlorofluorescein; TRF, terminal restriction fragment; T-RFLP, terminal restriction fragment length polymorphism.
[a]Only TRFs detected in T-RFLP profiles of the subjects corresponding to the medoid were selected and represented.
[b]No TRF corresponding to this bacterial species was detected in HEX T-RFLP profiles of overall subjects.

Table 4 Comparisons of clinical data among three groups separated based on T-RFLP profiles

Parameters	Cluster I (n = 41)	Cluster II (n = 74)	Cluster III (n = 85)	Statistical difference[a]
Age (years) (average ± s.d.)	28.9 ± 5.6	28.8 ± 6.6	27.1 ± 6.2	NS[b]
Gender				
Number of women (%)	14 (34.1)	41 (55.4)	43 (50.5)	NS[c]
Smoking				
Number of current smokers (%)	13 (31.7)	16 (21.6)	12 (14.1)	NS[c]
Teeth condition (average ± s.d.)				
Number of missing teeth	0.2 ± 0.6	0.4 ± 1.0	0.2 ± 0.7	NS[b]
Number of DMFT	9.4 ± 6.4	10.1 ± 7.1	7.8 ± 5.4	NS[b]
Periodontal condition (average ± s.d.)				
Percentage of sites bleeding on probing	27.6 ± 18.1	18.9 ± 19.8	14.6 ± 13.5	I > II, I > III[b]
Percentage of sites with periodontal pockets (pocket depth > 4 mm)	10.1 ± 15.4	9.0 ± 15.1	3.7 ± 6.4	I > III, II > III[b]
Number of subjects without periodontal pocket (%)	10 (24.3)	36 (48.6)	55 (64.7)	$P < 0.001$[c]
Number of subjects with periodontal pockets > 20% (%)	9 (21.9)	14 (18.9)	2 (2.3)	$P < 0.001$[c]
Number of total bacteria (× 10^8 deduced CFU per ml)[d]	6.7 ± 7.9	2.6 ± 2.7	1.5 ± 2.0	I > II, I > III[b]

Abbreviations: CFU, colony-forming unit; DMFT, number of decayed, missing and filled teeth; T-RFLP, terminal restriction fragment length polymorphism.
[a]P-values < 0.05 were considered statistically significant. NS indicates lack of significance.
[b]Statistical differences were evaluated among all three clusters by the Steel–Dwass test, non-parametric multiple comparison. P-values < 0.05 were considered statistically significant.
[c]Statistical differences were evaluated by the Fisher's exact test. P-values < 0.05 were considered statistically significant.
[d]Copy number of total bacterial 16S rRNA gene was divided by that of *Streptococcus mutans* UA159 corresponding to one CFU.

surfaces are present in saliva, and the bacterial composition in saliva is reported to be more similar to that of the soft tissues (such as the tongue coating) than those of the supra- and subgingival plaque (Mager *et al.*, 2003b). It is noteworthy that the comprehensive saliva bacterial composition reflects the periodontal health condition in the oral cavity.

The clustering of the T-RFLP profiles in the present study reflected major differences in the abundance distribution of the dominant TRFs common to most of the subjects, rather than the emergence of characteristic TRFs in each cluster. The characteristically more predominant species, predicted from the sizes of the TRFs, were *Prevotella* and *Veillonella* species in cluster I, which is correlated with the active phase of periodontitis; *Streptococcus* species in cluster II, which is linked to periodontitis in the remission phase and *Neisseria*, *Haemophilus* or *Aggregatibacter* species and *Porphyromonas* species in cluster III, which was implicated as a periodontally healthy cluster. Species strongly implicated as periodontal pathogens, including *P. gingivalis*, *T. forsythia* and *T. denticola*, were not assigned to the TRFs that characterized the clusters, but to other minor TRFs, and the peak area proportion of each of the dominant TRFs had little correlation with the periodontal health status. Increases in the numbers of specific bacteria with suspected virulence are often observed in individual periodontal pockets. However, the re-

sults of this study suggest that the abundance distribution of dominant bacteria, or commensal bacteria, which reside in the entire oral cavity, also differs, and is correlated with periodontal health status.

Although comprehensive analyses of oral bacterial communities have been performed in a number of previous studies, few studies have focused on the relationship between the colonization pattern of dominant bacteria in the microbiota and the periodontal condition. One reason for this is that the purpose of these studies was to detect the specific bacterial species that emerge only during disease states or exhibit a positive correlation with disease as putative pathogens, consistent with Koch's postulate. However, there must be numerous interactions among the pathogenic bacteria of oral disease, the other indigenous bacteria and the host. Some investigators have suggested that the balance of these synergistic and antagonistic interactions in the microbial community plays an essential role in the emergence of virulence (Kleinberg, 2002; Marsh, 2005; Kuramitsu *et al.*, 2007). It is likely that this study characterized the bacterial environment surrounding a specific potential pathogen, and that it might affect the periodontal condition through those interactions, which might be involved in susceptibility to periodontal disease. Both *Prevotella* and *Veillonella* species are obligatory anaerobes, and might be derived from the resident bacteria already

present in deepened periodontal pockets. However, this possibility is unlikely, because subjects with or lacking deeper periodontal pockets were both present in the same clusters.

The genus *Veillonella* has been associated with periodontal health through analyses of subgingival plaque (Haffajee *et al.*, 1998; Tanner *et al.*, 1998; Kumar *et al.*, 2005). However, *Veillonella* species were less dominant in cluster III, the cluster that showed good periodontal health, and more dominant in cluster I, the cluster in which the members showed active periodontitis. This contradiction might be explained by differences in the analytes, as saliva was used in our study. On the other hand, it is notable that TRFs that corresponded to certain phylotypes of *Porphyromonas*, but not to *P. gingivalis*, including the *Porphyromonas* sp. oral clone CW034 detected in the sequencing analysis (Table 3), were more dominant in cluster III. It is possible that no virulent *Porphyromonas* species compete with *P. gingivalis* and suppress their colonization or virulence. Further analyses should elucidate the mechanisms of the inter-bacterial interactions in the bacterial consortium of the oral cavity.

Among the clusters, no remarkable differences were observed in the other clinical parameters correlated with periodontitis, including age and smoking habits (Table 4), suggesting that the saliva bacterial composition is an indicator of periodontal health independent of these parameters. Although the amount of total bacteria in cluster I subjects was greater than those in subjects of clusters II and III (Table 4), no significant correlations between the periodontal health status and the amount of total bacteria were observed within cluster I or III. In addition, there was a significant difference in the periodontal health status between subjects in clusters II and III, even though there was no significant difference in the amounts of total bacteria between the two clusters. Taken together, these findings suggest that the saliva bacterial composition is more closely related to periodontal health than the amount of bacteria.

The T-RFLP analysis in this study showed limited phylogenetic resolution at the species level, even though a restriction enzyme was specifically selected to obtain more detailed information (Table 1). To detect the characteristic patterns that were correlated with various symptoms of periodontal disease, the use of several restriction enzymes in combination or other phylogenetic methods might be necessary. In addition, the PAM clustering method is more robust and therefore only identifies global patterns. Other clustering methods that are more efficient at identifying small clusters, such as HOPACH (van der Laan and Pollard, 2003), would be useful in a study with a larger sample size involving various types of periodontal diseases. In this study, the subjects were all 40 years of age or younger, had fewer than five missing teeth and

did not wear dentures. For subjects with a broader range of ages exhibiting a variety of oral health conditions, other clustering methods would be helpful.

This investigation revealed the importance of the proportion of the major indigenous bacteria in saliva with respect to periodontal health status and demonstrated the clinical importance of clustering the T-RFLP profiles of saliva bacterial populations. In the future, this method should be highly useful for evaluating susceptibility to periodontal disease, although the host genetic and immunological factors indicated in recent studies (Hooper and Gordon, 2001; Craig *et al.*, 2002; Donley *et al.*, 2004; Hart *et al.*, 2004) should also be considered, as well as the amounts of specific periodontal pathogens. Future studies of the dynamic interactions among the comprehensive components of oral microbiota should reveal the etiology of periodontal disease.

Acknowledgements

References

Aas JA, Barbuto SM, Alpagot T, Olsen I, Dewhirst FE, Paster BJ. (2007). Subgingival plaque microbiota in HIV positive patients. *J Clin Periodontol* **34**: 189–195.

Aas JA, Griffen AL, Dardis SR, Lee AM, Olsen I, Dewhirst FE *et al.* (2008). Bacteria of dental caries in primary and permanent teeth in children and young adults. *J Clin Microbiol* **46**: 1407–1417.

Aas JA, Paster BJ, Stokes LN, Olsen I, Dewhirst FE. (2005). Defining the normal bacterial flora of the oral cavity. *J Clin Microbiol* **43**: 5721–5732.

Becker MR, Paster BJ, Leys EJ, Moeschberger ML, Kenyon SG, Galvin JL *et al.* (2002). Molecular analysis of bacterial species associated with childhood caries. *J Clin Microbiol* **40**: 1001–1009.

Cole JR, Chai B, Farris RJ, Wang Q, Kulam-Syed-Mohideen AS, McGarrell DM *et al.* (2007). The Ribosomal Database Project (RDP-II): introducing myRDP space and quality controlled public data. *Nucleic Acids Res* **35**: D169–D172.

Craig RG, Boylan R, Yip J, Mijares D, Imam M, Socransky SS *et al.* (2002). Serum IgG antibody response to periodontal pathogens in minority populations: relationship to periodontal disease status and progression. *J Periodontal Res* **37**: 132–146.

de Lillo A, Ashley FP, Palmer RM, Munson MA, Kyriacou L, Weightman AJ *et al.* (2006). Novel subgingival bacterial phylotypes detected using multiple universal polymerase chain reaction primer sets. *Oral Microbiol Immunol* **21**: 61–68.

Denepitiya L, Kleinberg I. (1982). A comparison of the microbial compositions of pooled human dental

plaque and salivary sediment. *Arch Oral Biol* **27**: 739–745.

Donley CL, Badovinac R, Sapir S, Shapira L, Houri Y, Kantarci A *et al.* (2004). IgG antibody levels to *Porphyromonas gingivalis* and clinical measures in children. *J Periodontol* **75**: 221–228.

Dray S, Dufour AB. (2007). The ade4 package: implementing the duality diagram for ecologists. *J Stat Softw* **22**: 1–20.

Forney LJ, Zhou X, Brown CJ. (2004). Molecular microbial ecology: land of the one-eyed king. *Curr Opin Microbiol* **7**: 210–220.

Haffajee AD, Cugini MA, Tanner A, Pollack RP, Smith C, Kent Jr RL *et al.* (1998). Subgingival microbiota in healthy, well-maintained elder and periodontitis subjects. *J Clin Periodontol* **25**: 346–353.

Hart GT, Shaffer DJ, Akilesh S, Brown AC, Moran L, Roopenian DC *et al.* (2004). Quantitative gene expression profiling implicates genes for susceptibility and resistance to alveolar bone loss. *Infect Immun* **72**: 4471–4479.

Hooper LV, Gordon JI. (2001). Commensal host–relationships in the gut. *Science* **292**: 1115–1118.

Jenkinson HF, Lamont RJ. (2005). Oral microbial communities in sickness and in health. *Trends Microbiol* **13**: 589–595.

Kaufman L, Rousseeuw PJ. (1990). *Finding Groups in Data: an Introduction to Cluster Analysis*. Wiley: New York, USA.

Kazor CE, Mitchell PM, Lee AM, Stokes LN, Loesche WJ, Dewhirst FE *et al.* (2003). Diversity of bacterial populations on the tongue dorsa of patients with halitosis and healthy patients. *J Clin Microbiol* **41**: 558–563.

Kleinberg I. (2002). A mixed-bacteria ecological approach to understanding the role of the oral bacteria in dental caries causation: an alternative to *Streptococcus mutans* and the specific-plaque hypothesis. *Crit Rev Oral Biol Med* **13**: 108–125.

Kononen E, Paju S, Pussinen PJ, Hyvonen M, Di Tella P, Suominen-Taipale L *et al.* (2007). Population-based study of salivary carriage of periodontal pathogens in adults. *J Clin Microbiol* **45**: 2446–2451.

Kumar PS, Griffen AL, Moeschberger ML, Leys EJ. (2005). Identification of candidate periodontal pathogens and beneficial species by quantitative 16S clonal analysis. *J Clin Microbiol* **43**: 3944–3955.

Kumar PS, Leys EJ, Bryk JM, Martinez FJ, Moeschberger ML, Griffen AL. (2006). Changes in periodontal health status are associated with bacterial community shifts as assessed by quantitative 16S cloning and sequencing. *J Clin Microbiol* **44**: 3665–3673.

Kuramitsu HK, He X, Lux R, Anderson MH, Shi W. (2007). Interspecies interactions within oral microbial communities. *Microbiol Mol Biol Rev* **71**: 653–670.

Ledder RG, Gilbert P, Huws SA, Aarons L, Ashley MP, Hull PS *et al.* (2007). Molecular analysis of the subgingival microbiota in health and disease. *Appl Environ Microbiol* **73**: 516–523.

Liu WT, Marsh TL, Cheng H, Forney LJ. (1997). Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl Environ Microbiol* **63**: 4516–4522.

Loesche WJ. (1986). Role of *Streptococcus mutans* in human dental decay. *Microbiol Rev* **50**: 353–380.

Maechler M, Rousseeuw PJ, Struyf A, Huberet M. (2005). Cluster analysis basics and extensions. Available from http://cran.r-project.org/web/packages/cluster/index.html.

Mager DL, Haffajee AD, Socransky SS. (2003a). Effects of periodontitis and smoking on the microbiota of oral mucous membranes and saliva in systemically healthy subjects. *J Clin Periodontol* **30**: 1031–1037.

Mager DL, Ximenez-Fyvie LA, Haffajee AD, Socransky SS. (2003b). Distribution of selected bacterial species on intraoral surfaces. *J Clin Periodontol* **30**: 644–654.

Marchini L, Campos MS, Silva AM, Paulino FG, Nobrega FG. (2007). Bacterial diversity in aphthous ulcers. *Oral Microbiol Immunol* **22**: 225–231.

Marsh PD. (2005). Dental plaque: biological significance of a biofilm and community life-style. *J Clin Periodontol* **32**(Suppl. 6): 7–15.

Moore WE, Moore LV. (1994). The bacteria of periodontal diseases. *Periodontol 2000* **5**: 66–77.

Munson MA, Pitt-Ford T, Chong B, Weightman A, Wade WG. (2002). Molecular and cultural analysis of the microflora associated with endodontic infections. *J Dent Res* **81**: 761–766.

Nakano Y, Takeshita T, Yamashita Y. (2006). TRFMA: a web-based tool for terminal restriction fragment length polymorphism analysis based on molecular weight. *Bioinformatics* **22**: 1788–1789.

Paster BJ, Boches SK, Galvin JL, Ericson RE, Lau CN, Levanos VA *et al.* (2001). Bacterial diversity in human subgingival plaque. *J Bacteriol* **183**: 3770–3783.

Paster BJ, Falkler Jr WA, Enwonwu CO, Idigbe EO, Savage KO, Levanos VA *et al.* (2002). Prevalent bacterial species and novel phylotypes in advanced noma lesions. *J Clin Microbiol* **40**: 2187–2191.

Pei Z, Bini EJ, Yang L, Zhou M, Francois F, Blaser MJ. (2004). Bacterial biota in the human distal esophagus. *Proc Natl Acad Sci USA* **101**: 4250–4255.

Pihlstrom BL, Michalowicz BS, Johnson NW. (2005). Periodontal diseases. *Lancet* **366**: 1809–1820.

R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria, available from http://www.r-project.org.

Rasiah IA, Wong L, Anderson SA, Sissons CH. (2005). Variation in bacterial DGGE patterns from human saliva: over time, between individuals and in corresponding dental plaque microcosms. *Arch Oral Biol* **50**: 779–787.

Sakamoto M, Takeuchi Y, Umeda M, Ishikawa I, Benno Y. (2003). Application of terminal RFLP analysis to characterize oral bacterial flora in saliva of healthy subjects and patients with periodontitis. *J Med Microbiol* **52**: 79–89.

Sambrook J, Russell DW. (2001). *Molecular Cloning: A laboratory Manual*, 3rd ed. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, USA.

Shibata Y, Ozaki K, Seki M, Kawato T, Tanaka H, Nakano Y *et al.* (2003). Analysis of loci required for determination of serotype antigenicity in *Streptococcus mutans* and its clinical utilization. *J Clin Microbiol* **41**: 4107–4112.

Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent Jr RL. (1998). Microbial complexes in subgingival plaque. *J Clin Periodontol* **25**: 134–144.

Takeshita T, Nakano Y, Yamashita Y. (2007). Improved accuracy in terminal restriction fragment length polymorphism phylogenetic analysis using a novel

internal size standard definition. *Oral Microbiol Immunol* **22**: 419–428.

Tanner A, Maiden MF, Macuch PJ, Murray LL, Kent Jr RL. (1998). Microbiota of health, gingivitis, and initial periodontitis. *J Clin Periodontol* **25**: 85–98.

van der Laan MJ, Pollard K. (2003). Hybrid clustering of gene expression data with visualization and the bootstrap. *J Stat Plan Inference* **117**: 275–303.

von Troil-Linden B, Alaluusua S, Wolf J, Jousimies-Somer H, Torppa J, Asikainen S. (1997). Periodontitis patient and the spouse: periodontal bacteria before and after treatment. *J Clin Periodontol* **24**: 893–899.

Warnes GR. (2008). *gregmisc: Greg's Miscellaneous Functions* available from http://cran.r-project.org/web/packages/gregmisc/index.html.