**OPEN**

Correspondence and
requests for materials
should be addressed to
Y.Y. (yoshi@dent.
kyushu-u.ac.jp)

# Discrimination of the oral microbiota associated with high hydrogen sulfide and methyl mercaptan production

Toru Takeshita[1], Nao Suzuki[2], Yoshio Nakano[3], Masaki Yasui[1], Masahiro Yoneda[2], Yoshihiro Shimazaki[1], Takao Hirofuji[2] & Yoshihisa Yamashita[1]

[1]Section of Preventive Dentistry, Division of Oral Health, Growth and Development, Kyushu University Faculty of Dental Science, Fukuoka, Japan, [2]Section of General Dentistry, Department of General Dentistry, Fukuoka Dental College, Fukuoka, Japan, [3]Department of Chemistry, Nihon University School of Dentistry, Tokyo, Japan.

Both hydrogen sulfide ($H_2S$) and methyl mercaptan ($CH_3SH$) are frequently detected in large amounts in malodorous mouth air. We investigated the bacterial composition of saliva of 30 subjects with severe oral malodor exhibiting extreme $CH_3SH/H_2S$ ratios (high $H_2S$ but low $CH_3SH$ concentrations, n = 14; high $CH_3SH$ but low $H_2S$ concentrations, n = 16) and 13 subjects without malodor, using barcoded pyrosequencing analysis of the 16S rRNA gene. Phylogenetic community analysis with the UniFrac distance metric revealed a distinct bacterial community structure in each malodor group. The $H_2S$ group showed higher proportions of the genera *Neisseria*, *Fusobacterium*, *Porphyromonas* and SR1 than the other two groups, whereas the $CH_3SH$ group had higher proportions of the genera *Prevotella*, *Veillonella*, *Atopobium*, *Megasphaera*, and *Selenomonas*. Our results suggested that distinct bacterial populations in the oral microbiota are involved in production of high levels of $H_2S$ and $CH_3SH$ in the oral cavity.

O ral malodor is a common but distressing condition that may affect interpersonal social communication[1]. Most of the pathologies are found within the mouth and originate mainly from microbial metabolism[1]. The intraoral surfaces are colonized by a diverse array of bacterial species, many of which, especially anaerobes, have an ability to degrade substrates into malodorous compounds[2,3]. Microbial overgrowth due to poor oral hygiene results in the development of oral malodor.

Molecular approaches using the 16S rRNA gene have allowed comprehensive surveys of complex bacterial communities, and the tongue microbiota structure of subjects with severe oral malodor has been determined[4–6]. These studies indicated that the microbiota of subjects with oral malodor shows a greater microbial diversity than that of no-odor subjects, and various microorganisms have been implicated in oral malodor. Our previous study of 240 subjects using terminal restriction fragment length polymorphism (T-RFLP) analysis confirmed that Gram-positive saccharolytic species, including *Streptococcus*, are more predominant in the microbiota of subjects with no odor than in that of malodor patients. Furthermore, our data demonstrated statistically that the global composition of the oral microbiota is correlated with the severity of malodor[7]. However, no single specific bacterial infection has been definitively associated with oral malodor. Key bacterial species to be targeted by therapy remain poorly defined. Some researchers have proposed the "non-specific theory" for the microbial etiology of oral malodor, which suggests that there are multiple causal agents and that many groups can substitute for others[8].

Unpleasant oral odor results from a mixture of various molecules, such as volatile sulfur compounds (VSC), short-chain fatty acids, polyamines, and indoles[8]. The oral malodor-related microbiota is presumed to be a complex consortium of bacterial members involved in production of these compounds. Discrimination of the various production systems might help to identify those microbes important in generation of oral malodor.

Among the various compounds causing oral malodor, two VSCs, hydrogen sulfide ($H_2S$) and methyl mercaptan ($CH_3SH$), are known to be associated with oral malodor, and their concentrations in mouth air are often used for clinical diagnosis. In this study, we aimed to distinguish and identify the oral microbiota structures associated with high $H_2S$ and $CH_3SH$ production. Although malodorous mouth air contains both $H_2S$ and $CH_3SH$ in most cases, some patients exhibit extreme $CH_3SH/H_2S$ ratios. We enrolled two groups of severe halitosis patients with distinct VSC profiles of high $H_2S$ but low $CH_3SH$ concentrations and high $CH_3SH$ but
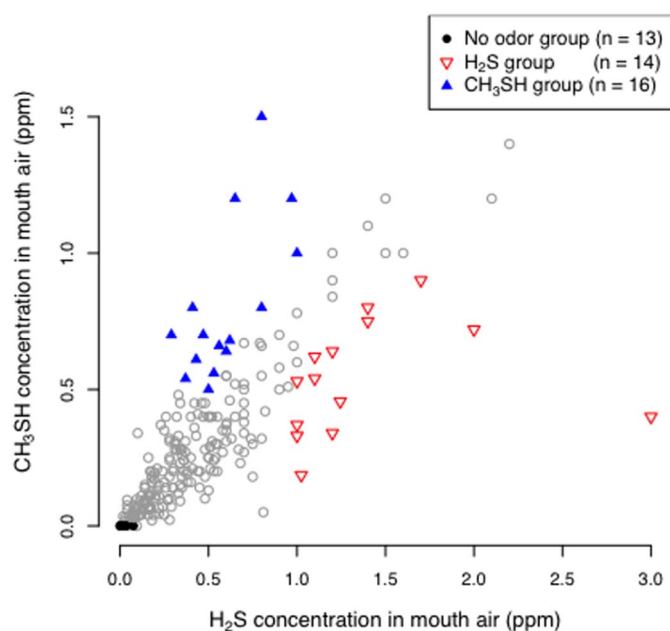
1

**Figure 1 | Hydrogen sulfide (H₂S) and methyl mercaptan (CH₃SH) distribution in mouth air of 240 subjects in our previous study[7].** Subjects with low VSC in mouth air (●, $H_2S \leq 0.075$ ppm and $CH_3SH = 0$ ppm, n = 13), those with high $H_2S$ but low $CH_3SH$ concentrations (∇, $H_2S \geq$ 1 ppm and $CH_3SH/H_2S < 0.6$, n = 14), and those with high $CH_3SH$ but low $H_2S$ concentrations (▲, $CH_3SH \geq 0.5$ ppm and $CH_3SH/H_2S \geq 1.0$, n = 16) were enrolled in this study. The other subjects shown as gray circles were excluded from the present study.

low $H_2S$ concentrations, as well as a group of control subjects with no oral malodor. The bacterial composition of saliva was evaluated using barcoded pyrosequencing analysis of the 16S rRNA gene.

## Results

We investigated the salivary bacterial populations of 43 subjects whose mouth air showed three different VSC profiles: high $H_2S$ but low $CH_3SH$ concentration ($H_2S$ group, n = 14), high $CH_3SH$ but low $H_2S$ concentration ($CH_3SH$ group, n = 16), and no VSC (no-odor group, n = 13) (**Figure 1**). The general and clinical parameters

### Table 1 | General and oral condition of 43 subjects

| | No odor group (n = 13) | H₂S group (n = 14) | CH₃SH group (n = 16) |
|---|---|---|---|
| Age[a] (yr) | 28±11 | 47±15[b] | 57±13[b] |
| Sex[c] (Female (%)) | 6 (46.1) | 5 (35.7) | 9 (56.2) |
| Number of teeth[a] | 27±4 | 26±4 | 25±3 |
| Number of decayed teeth[a] | 1.3±4.6 | 1.4±3.1 | 0.3±0.6 |
| Tongue coating score[a] | 1.2±0.5 | 2.0±0.8[b] | 2.1±0.8[b] |
| Mean pocket depth[a] (mm) | 2.8±0.2 | 3.4±0.9 | 3.4±0.5[b] |
| Amount of saliva[a] (ml/5 min) | 5.7±2.5 | 7.3±3.9 | 6.3±2.5 |

[a]Significant difference between groups were evaluated by Steel-Dwass multiple comparison test.
[b]Significantly higher than no odor group (P < 0.05).
[c]Significant difference were evaluated by Fisher's exact test.

of the study populations are given in Table 1. Age and tongue coating scores of the two malodor groups were significantly higher than those of the no-odor group, whereas no significant difference was observed between the two malodor groups. In addition, the mean periodontal pocket depth of the $CH_3SH$ group was significantly greater than that of the no-odor group.

Whole-genomic DNA was extracted from each saliva sample and the variable regions (V1–V2) of the bacterial 16S rRNA gene were PCR amplified using barcoded universal primers. Pyrosequencing using the Roche 454 FLX instrument produced a dataset consisting of 231,266 high-quality reads with an average length of 344 ± 20 bp (Table 2). Of these, 196,943 were matched with 575 oral bacterial sequences deposited in the Forsyth Institute Human Oral Microbiome (HOMD) database[9], and these sequences were subsequently assigned to 429 oral taxon (OT) defined by HOMD[10]. The remaining 34,323 sequences were assigned to 2,639 operational taxonomic units (OTU) using a 97% pairwise-identity cutoff.

Phylogenetic community analysis with the UniFrac distance metric revealed a distinct overall bacterial community composition in each of the three groups (**Figure 2**). These differences were confirmed statistically in both unweighted (qualitative) and weighted (quantitative) versions of this metric using permutational multivariate analysis

### Table 2 | Summary of pyrosequencing analysis

| No odor group (n = 13) | | H₂S group (n = 14) | | CH₃SH group (n = 16) | |
|---|---|---|---|---|---|
| Subject ID | No. of Reads | Subject ID | No. of Reads | Subject ID | No. of Reads |
| N1 | 4,249 | H1 | 8,661 | M1 | 5,825 |
| N2 | 6,858 | H2 | 7,667 | M2 | 3,720 |
| N3 | 7,044 | H3 | 5,278 | M3 | 2,714 |
| N4 | 5,552 | H4 | 6,310 | M4 | 6,748 |
| N5 | 6,273 | H5 | 3,077 | M5 | 6,412 |
| N6 | 3,912 | H6 | 1,655 | M6 | 2,799 |
| N7 | 7,076 | H7 | 4,811 | M7 | 6,033 |
| N8 | 6,605 | H8 | 5,943 | M8 | 5,564 |
| N9 | 4,673 | H9 | 7,218 | M9 | 4,987 |
| N10 | 3,591 | H10 | 7,087 | M10 | 6,677 |
| N11 | 5,709 | H11 | 4,336 | M11 | 3,649 |
| N12 | 4,417 | H12 | 3,745 | M12 | 3,904 |
| N13 | 5,553 | H13 | 7,031 | M13 | 5,689 |
| | | H14 | 5,590 | M14 | 5,224 |
| | | | | M15 | 3,792 |
| | | | | M16 | 7,608 |
| Average | 5,500±1,230 | | 5,600±1,938 | | 5,084±1,490 |

2
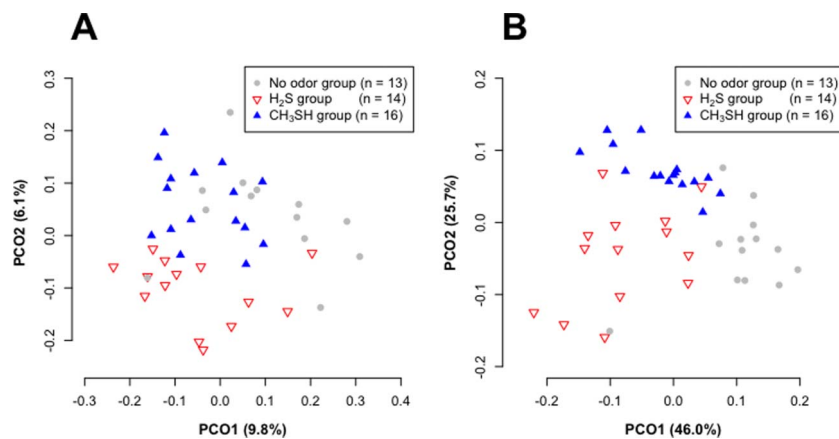
**Figure 2 | Principal coordinate analysis (PCoA) plot showing similarity of composition of 43 salivary microbiotas.** Plots were generated using unweighted (**A**) and weighted (**B**) versions of the Unifrac distance metric. The two components explained 15.9 and 71.7% of the variance, respectively.

of variance (perMANOVA) and analysis of similarities (ANOSIM) analyses ($P < 0.001$, each). The discrimination was stronger with the weighted UniFrac (ANOSIM statistic R = 0.47) than the unweighted metric (R = 0.25), suggesting that the microbiotas of the three groups of subjects had greater differences in community structure than in community membership.
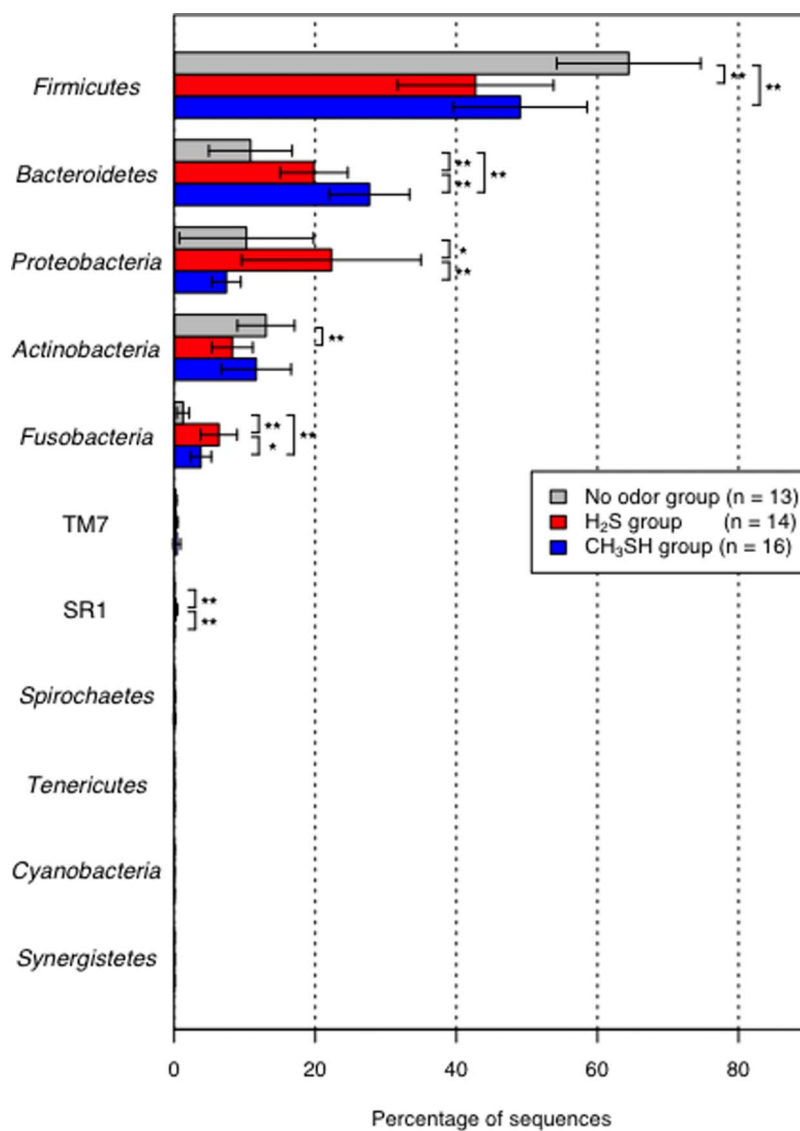


**Figure 3 | The relative abundances of phyla in the salivary microbiota of 43 subjects (mean ± SD).** Significant differences between groups were evaluated by Steel-Dwass test. ** $P < 0.01$, * $P < 0.05$.

The vast majority of the sequences were assigned to five bacterial phyla (*Actinobacteria*, *Bacteroidetes*, *Fusobacteria*, *Firmicutes*, and *Proteobacteria*). TM7, SR1, *Spirochaetes*, *Tenericutes*, *Cyanobacteria*, and *Synergistetes* were also identified but in much lower proportions. Significant differences were observed in the relative proportions of the five major phyla and SR1 among the three groups (**Figure 3**).

The sequences represented 82 different genera (constituting at least 97% of the microbiota of each subject) and 35 upper-level taxa. Of the 82 genera, 58, 66, and 68 were detected in the no-odor, $H_2S$, and $CH_3SH$ groups, respectively. The overall composition of the microbiota was qualitatively fairly uniform, and 49 genera were detected in each group. Eleven genera (*Streptococcus*, *Prevotella*, *Rothia*, *Actinomyces*, *Granulicatella*, *Neisseria*, *Terrahaemophilus*, *Veillonella*, *Gemella*, *Fusobacterium*, and *Haemophilus*) were found in all subjects (**Figure 4**) and constituted 80 to 99% of the microbiota in each individual. On the other hand, the relative abundance of each genus varied among the three groups. Significant differences between the three groups were observed in 19 bacterial genera (**Figure 5**). Compared to the no-odor group, the microbiota of both malodor groups was colonized by significantly lower proportions of *Streptococcus* and *Granulicatella* and higher proportions of *Leptotrichia*, *Peptostreptococcus*, *Eubacterium*, and *Fusobacterium*. On the other hand, the genera *Fusobacterium*, *Neisseria*, *Porphyromonas*, and SR1 genera incertae sedis were significantly more predominant in the $H_2S$ group than in the other two groups. Conversely, the genera *Prevotella*, *Veillonella*, *Atopobium*, *Megasphaera*, and *Selenomonas* were significantly more predominant in the $CH_3SH$ group than in the other two groups. Significant differences were also observed in *Rothia* (no-odor group > $H_2S$ group), *Parvimonas* ($H_2S$ group >

no-odor group), *Peptococcus* ($H_2S$ group > no-odor group), *Solobacterium* ($CH_3SH$ group > no-odor group), and *Dialister* ($CH_3SH$ group > no-odor group).

Of 429 OT and 2,639 OTU, 77, including well-known VSC producers such as *Porphyromonas gingivalis* (OT 619), *P. endodontalis* (OT 273), and *Fusobacterium nucleatum* (OT 200 and OT 202), were found in significantly higher proportions in both malodor groups than in the no-odor group (Fig. 6 and Supplementary Table S1). In contrast, 15 of these were significantly more predominant in the microbiota of the $H_2S$ group compared to the $CH_3SH$ group; conversely, 17 were significantly more predominant in the microbiota of the $CH_3SH$ group compared to the $H_2S$ group (**Figure 6**).

## Discussion
This study revealed the bacterial community structure in the saliva of two groups of severe malodor patients with different $CH_3SH/H_2S$ ratios. The microbiota of the subjects in both malodor groups was characterized by lower proportions of *Streptococcus* and *Granulicatella* and higher proportions of *Leptotrichia*, *Peptostreptococcus*, *Eubacterium*, and *Fusobacterium* compared with the no-odor group. On the other hand, the genera *Neisseria, Fusobacterium, Porphyromonas*, and SR1 and the genera *Prevotella*, *Veillonella*, *Atopobium*, *Megasphaera,* and *Selenomonas* were detected in higher proportions in the $H_2S$ and $CH_3SH$ groups, respectively (**Figure 5**). Many oral bacteria can produce both $H_2S$ and $CH_3SH$[2], and their concentrations in mouth air are generally highly correlated each other. Therefore, previous studies explored the oral malodor-associated microbiota without focusing on the VSC distribution. However, our results clearly indicate that the bacterial community structures
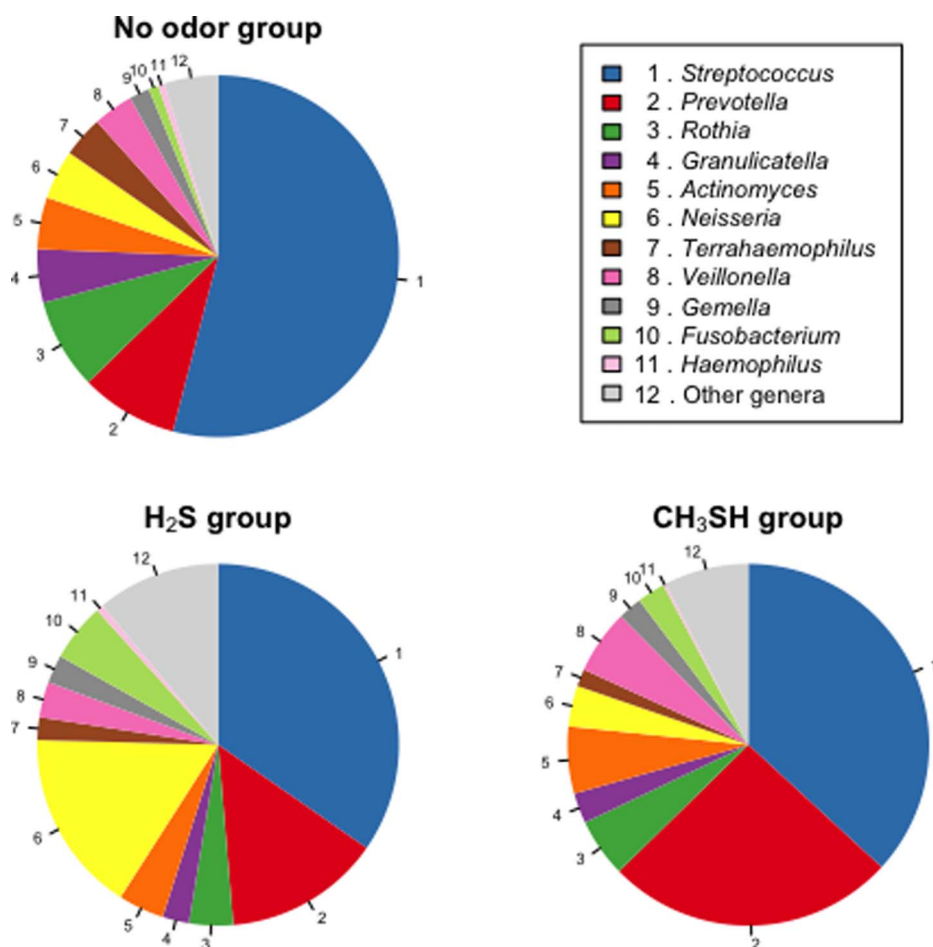


**Figure 4 | The mean abundances of bacterial genera in the no-odor, $H_2S$, and $CH_3SH$ groups, respectively.** Only the 11 genera detected in each subject are shown.
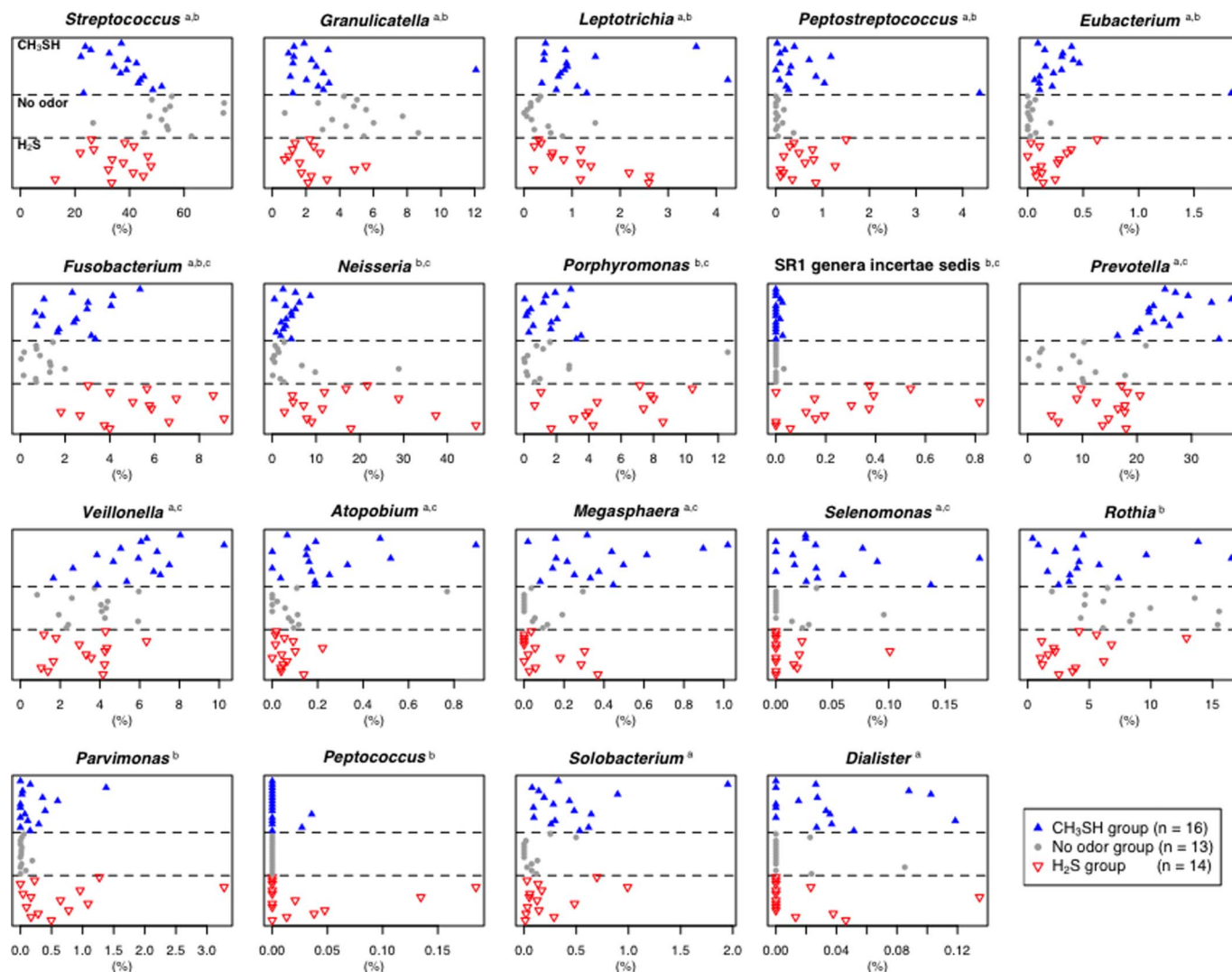
4

**Figure 5 | The relative abundances of 19 bacterial genera that differed significantly between groups.** Statistical differences were evaluated using a Steel-Dwass multiple comparison test ($P < 0.05$). [a]Significant difference between the $CH_3SH$ and no-odor groups. [b]Significant difference between the $H_2S$ and no-odor groups. [c]Significant difference between the $CH_3SH$ and $H_2S$ groups.

of the two malodor groups were phylogenetically different (**Figure 2**), suggesting that distinct bacterial populations are involved in production of high levels of each compound in the oral cavity.

Whereas potential $H_2S$ producers such as *Fusobacterium* and *Porphyromonas* were significantly more predominant in the $H_2S$ group, all characteristic bacterial species, especially in the $CH_3SH$ group were not necessarily heavy producers of each compound. According to Persson *et al.*[2], many *Prevotella* species, including *P. melaninogenica*, have no, or only a low, capacity to produce $CH_3SH$ from serum or methionine. *Veillonella* and *Selenomonas* species degrade neither serum nor methionine into significant amounts of $CH_3SH$, although they are potent $H_2S$ producers. To our knowledge, no report has demonstrated that *Atopobium* and *Megasphaera* species have a high capability to generate $CH_3SH$. VSC production in the oral cavity is considered to result from proteolysis of serum, food debris and other substrates, followed by subsequent breakdown of sulfur-containing amino acids[11]. VSC production is likely to be a complex process that involves many bacterial species including both VSC producing and non-producing strains. Sterer and Rosenberg indicated that even saccharolytic species such as *Streptococcus salivarius* promote VSC production through supply of protein resulting

from glycoprotein putrefaction[11]. Well-known producers of both VSCs, such as *Porphyromonas gingivalis*, *P. endodontalis*, and *Fusobacteirum nucleatum*, were detected in higher proportions in both malodor groups compared with the no-odor group (**Figure 6**), although they represented only small minorities of the total microbiota (Supplementary Table S1). Their $H_2S$ or $CH_3SH$ productivity might be accelerated by synergistic interactions with the surrounding bacterial populations. There is also the possibility that these bacteria are non-participating observers that prefer microbial environments in which production of large amounts of $H_2S$ or $CH_3SH$ occurs. Co-cultivation *in vitro* would assist in identification of the key bacterial combinations. In addition, the dominance of SR1 observed in the $H_2S$ group (**Figures 5 & 6**) is noteworthy, considering that the ecological role of this as-yet-uncultured phylum is not well understood[12]. Development of a method of cultivating this bacterium would clarify its role in high $H_2S$ production.

Our previous study using T-RFLP classified the salivary microbiota of 240 subjects into four patterns. Two of these (clusters II and IV) were equally implicated in being associated with oral malodor[7]. The characteristics of clusters II and IV, which were predicted from their T-RFLP profiles, corresponded well to the bacterial community structures of the $CH_3SH$ and $H_2S$ groups, respectively, in the present study.
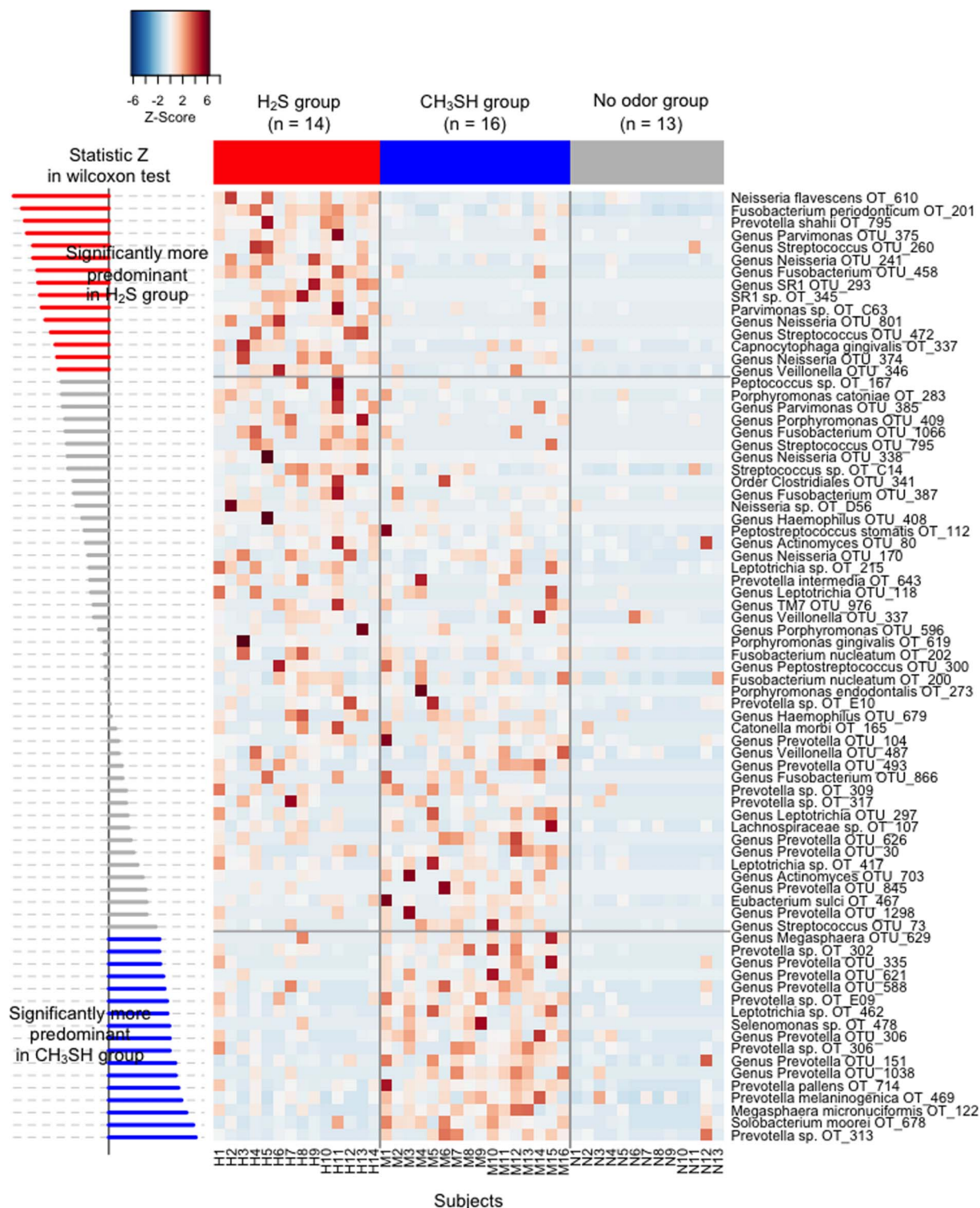
**Figure 6 | Relative abundance distribution of 77 oral taxon (OT) or operational taxonomic units (OTU) that were significantly more abundant in the malodor groups (n = 30) than in the no-odor group (n = 13), evaluated by Wilcoxon rank-sum test ($P < 0.05$).** To show the distribution of the OT or OTU with lower abundance, the relative abundances of each OT or OTU were normalized to have a mean of 0 and standard deviation of 1 (z-score normalization) and represented as the color intensity of each grid (blue, low abundance; red, high abundance). The OT and OTU were ordered according to the Z statistic (shown as a bar plot on the left) in a Wilcoxon rank-sum test comparing the $CH_3SH$ (n = 16) and $H_2S$ (n = 14) groups. The direction of horizontal bars indicates greater predominance in the $H_2S$ than the $CH_3SH$ group (left) and in the $CH_3SH$ than the $H_2S$ group (right). The red and blue bars show significant differences between the two malodor groups.

However, their association with $CH_3SH$ and $H_2S$ was not examined in the previous study. Therefore, we reanalyzed the data regarding $CH_3SH$ and $H_2S$ in each cluster. The results indicated that the adjusted odds ratio for increased $CH_3SH$ levels was higher in cluster II than cluster IV, and conversely, that the odds ratio for increased levels of $H_2S$ were higher in cluster IV than cluster II (data not shown). These findings are

consistent with the relationship between high H$_2$S and CH$_3$SH production and the microbiota structures characterized in this study.

Along with H$_2$S and CH$_3$SH, various other malodorous compounds are often present in malodorous mouth air, such as short-chain fatty acids, polyamines, and indoles[8]. The oral microbiota of the high VSC groups contained not only VSC-producing bacteria, but also those capable of producing other malodorous compounds. Therefore, further subtyping of high VSC-producing microbiota based on the yields of other malodorous compounds might facilitate discrimination and identification of bacterial members involved in their production, if appropriate clinical methods of evaluating these compounds were available. Elucidating the bacterial species involved in production of malodorous compounds other than VSC would help to further narrow the key bacterial species responsible for VSC production.

Consistent with previous studies, we found no single bacterial species whose presence explains oral malodor. However, our results revealed that a large amount of H$_2$S or CH$_3$SH in the oral cavity is produced by characteristic oral microbiota structures, suggesting that suppression or elimination of specific bacterial populations may effectively reduce oral malodor. Although a methodology of altering the stable oral microbiota remains unknown, this may represent a treatment option in addition to the conventional steadfast mouth-cleaning routine that focuses on reducing the total bacterial load in the oral cavity.

## Methods

**Study population.** The subjects were a subgroup of the 240 subjects analyzed in our previous study (Takeshita *et al.* 2010), who visited the Oral Malodor Clinic at Fukuoka Dental College Medical and Dental Hospital in Japan. In this study, we selected 30 subjects with severe oral malodor showing two characteristic VSC profiles: high H$_2$S but low CH$_3$SH concentration (H$_2$S ≥ 1 ppm and CH$_3$SH/H$_2$S < 0.6, n = 14) and high CH$_3$SH but low H$_2$S concentration (CH$_3$SH ≥ 0.5 ppm and CH$_3$SH/H$_2$S ≥ 1.0, n = 16) (**Figure 1**). We also enrolled 13 subjects without oral malodor (H$_2$S ≤ 0.075 ppm and CH$_3$SH = 0 ppm) based on the H$_2$S and CH$_3$SH olfactory thresholds (0.15 and 0.05 ppm, respectively[13]). Their general and clinical conditions were evaluated in the previous study[7]. The Ethics Committee of Fukuoka Dental College and Kyushu University Faculty of Dental Science approved the study design.

**Sample collection and pyrosequencing analysis.** Collection of saliva samples and DNA extraction was performed in the previous study (Takeshita *et al.* 2010). The 16S rRNA genes of each sample were amplified using the following primers: 338R with the 454 Life Sciences adaptor B sequence (5′-CCT ATC CCC TGT GTG CCT TGG CAG TCT CAG TGC TGC CTC CCG TAG GAG T-3′) and 8F with the 454 Life Sciences adaptor A and a subject-specific six-base barcode sequences (5′-CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG NNN NNN AGA GTT TGA TYM TGG CTC AG-3′). PCR amplification was performed as described previously (Takeshita *et al.* 2010). The amplicons were gel-purified using a Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI) according to the manufacturer's instructions. The DNA concentration and quality were assessed using a NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE), and equal amounts of DNA from the 43 subjects were pooled. Pyrosequencing was conducted using a 454 Life Sciences Genome Sequencer FLX instrument (Roche, Basel, Switzerland) at Hokkaido System Science Co., Ltd. (Sapporo, Japan).

**Data analysis and taxonomy assignment.** Sequences were excluded from the analysis using a script written in PHP if they were shorter than 240 bases, or had an average quality score <25, and subsequently removed using a script written in R if they did not include the correct primer sequence, had a homopolymer run >6 nt, or contained ambiguous characters. The remaining sequences were assigned to each subject by examining the six-base barcode sequence. Each sequence was compared using the BLAST algorithm with 1,647 sequences of the oral bacterial 16S rRNA gene (HOMD 16S rRNA RefSeq Extended Version 1.1) deposited in HOMD[9], and assigned to the best BLAST hit with a 98.5% similarity value and minimum coverage of 97%. After excluding assigned sequences, similar sequences were clustered into operational taxonomic units using the complete-linkage clustering tool of the RDP pyrosequencing pipeline[14] at a distance cut off of 0.03, and the representative sequences of each cluster were selected using the Dereplicate request function. The oral bacterial sequences from HOMD, to which at least one sequence was assigned, and the representative sequences from each OTU were aligned using PyNAST[15] and the Greengenes database[16] using a minimum percent identity of 75%. Chimeras were removed from the representative set on the basis of identification as chimeric via Chimera Slayer[17] and verification that the putative chimera appeared in only one sample. After chimera elimination, a relaxed neighbor-joining tree was built using FastTree[18]. To determine the dissimilarity between any pair of bacterial communities, we used the UniFrac metric[19] calculated by Fast UniFrac[20]. The similarity relationship assessed using the unweighted and weighted UniFrac metric was represented in a principal coordinate analysis (PCoA) plot drawn by R. The taxonomy of representative sequences was determined using the RDP classifier with a minimum support threshold of 60% and the RDP taxonomic nomenclature (down to the genus level).

**Statistical analysis.** A Steel-Dwass non-parametric multiple comparison test was performed to compare age, number of teeth, number of decayed teeth, tongue coating score, mean pocket depth, amount of saliva, and the relative abundances of bacterial phyla and genera. Fisher's exact test was conducted to look for differences by sex. Wilcoxon's rank-sum test was performed to compare the relative abundance of each OT or OTU. We used perMANOVA and ANOSIM to test for differences in bacterial community structure among groups of samples. Statistical significance was set at $P < 0.05$. All statistical analyses were conducted using the R software package, version 2.11.1[21].

1. Scully, C., Porter, S. & Greenman, J. What to do about halitosis. *BMJ* **308**, 217–218 (1994).
2. Persson, S., Edlund, M. B., Claesson, R. & Carlsson, J. The formation of hydrogen sulfide and methyl mercaptan by oral bacteria. *Oral Microbiol Immunol* **5**, 195–201 (1990).
3. Loesche, W. J. & Kazor, C. Microbiology and treatment of halitosis. *Periodontol 2000* **28**, 256–279 (2002).
4. Kazor, C. E. *et al.* Diversity of bacterial populations on the tongue dorsa of patients with halitosis and healthy patients. *J Clin Microbiol* **41**, 558–563 (2003).
5. Haraszthy, V. I. *et al.* Identification of oral bacterial species associated with halitosis. *J Am Dent Assoc* **138**, 1113–1120 (2007).
6. Riggio, M. P. *et al.* Molecular identification of bacteria on the tongue dorsum of subjects with and without halitosis. *Oral Dis* **14**, 251–258 (2008).
7. Takeshita, T. *et al.* Relationship between oral malodor and the global composition of indigenous bacterial populations in saliva. *Appl Environ Microbiol* **76**, 2806–2814 (2010).
8. Scully, C. & Greenman, J. Halitosis (breath odor). *Periodontol 2000* **48**, 66–75 (2008).
9. Chen, T. *et al.* The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database (Oxford)* **2010** (2010).
10. Dewhirst, F. E. *et al.* The human oral microbiome. *J Bacteriol* **192**, 5002–5017 (2010).
11. Sterer, N. & Rosenberg, M. Streptococcus salivarius promotes mucin putrefaction and malodor production by Porphyromonas gingivalis. *J Dent Res* **85**, 910–914 (2006).
12. Davis, J. P., Youssef, N. H. & Elshahed, M. S. Assessment of the diversity, abundance, and ecological distribution of members of candidate division SR1 reveals a high level of phylogenetic diversity but limited morphotypic diversity. *Appl Environ Microbiol* **75**, 4139–4148 (2009).
13. Tonzetich, J. & Ng, S. K. Reduction of malodor by oral cleansing procedures. *Oral Surg Oral Med Oral Pathol* **42**, 172–181 (1976).
14. Cole, J. R. *et al.* The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* **37**, D141–145 (2009).
15. Caporaso, J. G. *et al.* PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**, 266–267 (2010).
16. DeSantis, T. Z. *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72** (2006).
17. Haas, B. J. *et al.* Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* **21**, 494–504 (2011).
18. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* **26**, 1641–1650 (2009).
19. Lozupone, C. & Knight, R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71**, 8228–8235 (2005).
20. Hamady, M., Lozupone, C. & Knight, R. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J* **4**, 17–27 (2009).
21. R Development Core Team. *R: A language and environment for statistical computing* (R Foundation for Statistical Computing, Vienna) ⟨http://www.r-project.org⟩ (2010).

## Author contributions

This study was conceived by TT. Samples and clinical data were collected by NS, MY and TH. Laboratory work was done by TT, NS and MY. Analysis was carried out by TT, YN and YS. The manuscript was written by TT and YY. All authors read and approved the manuscript.

## Additional information