

# Distinct composition of the oral indigenous microbiota in South Korean and Japanese adults

Toru Takeshita<sup>1</sup>, Kazuki Matsuo<sup>1,2</sup>, Michiko Furuta<sup>1</sup>, Yukie Shibata<sup>1</sup>, Kaoru Fukami<sup>1</sup>, Yoshihiro Shimazaki<sup>3</sup>, Sumio Akifusa<sup>4</sup>, Dong-Hung Han<sup>5</sup>, Hyun-Duck Kim<sup>5</sup>, Takeshi Yokoyama<sup>2</sup>, Toshiharu Ninomiya<sup>6</sup>, Yutaka Kiyohara<sup>6</sup>, Yoshihisa Yamashita<sup>1\*</sup>

<sup>1</sup>Section of Preventive and Public Health Dentistry, Division of Oral Health, Growth and Development and <sup>2</sup>Section of Dental Anesthesiology, Division of Maxillofacial Diagnostic and Surgical Sciences, Kyushu University Faculty of Dental Science, Fukuoka, Japan, <sup>3</sup>Department of Preventive Dentistry and Dental Public Health, School of Dentistry, Aichi-Gakuin University, Nagoya, Japan, <sup>4</sup>Department of Health Management, School of Oral Health Sciences, Kyushu Dental College, Kitakyushu, Japan, <sup>5</sup>Department of Preventive and Social Dentistry, School of Dentistry, Seoul National University, Seoul, Korea, <sup>6</sup>Department of Environmental Medicine, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

Correspondence and requests for materials should be addressed to Y.Y.

Fig. S1

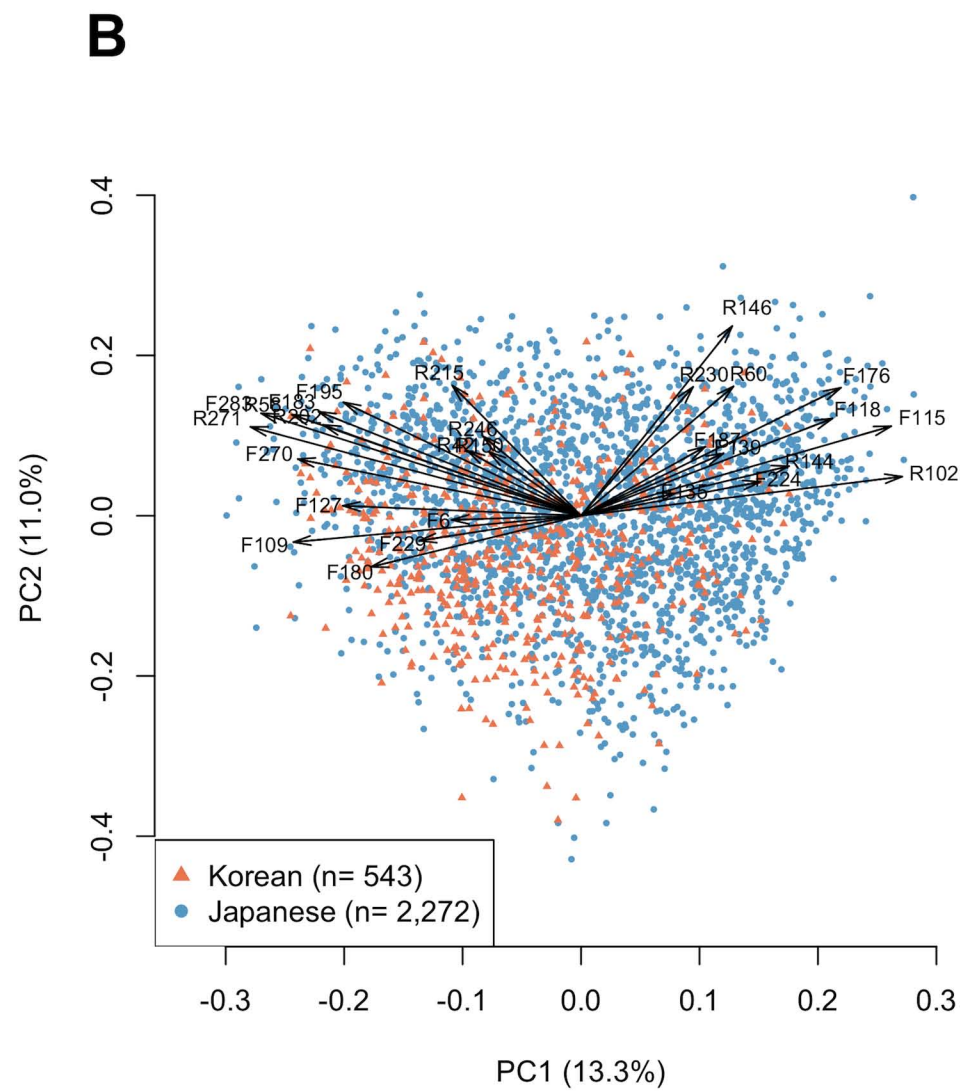
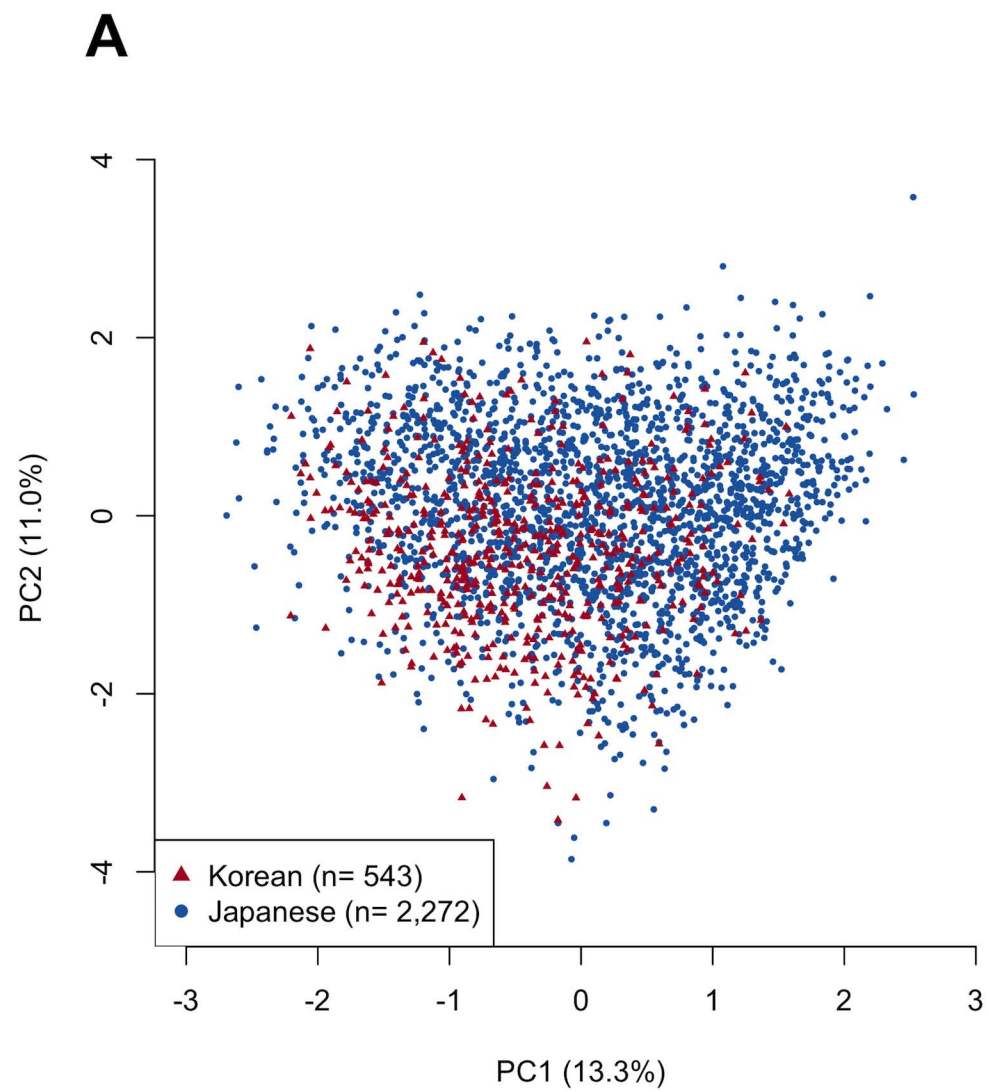


Fig. S2

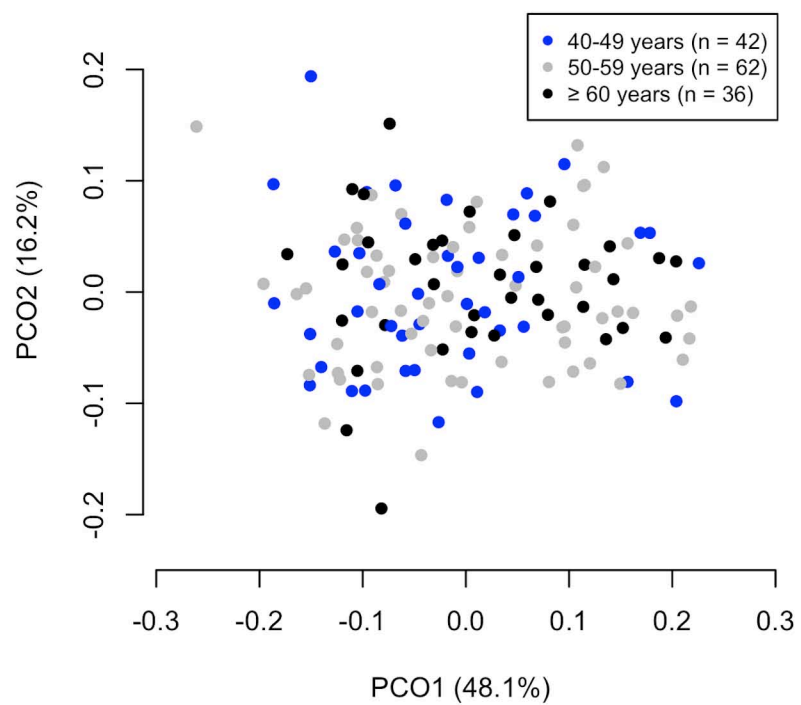
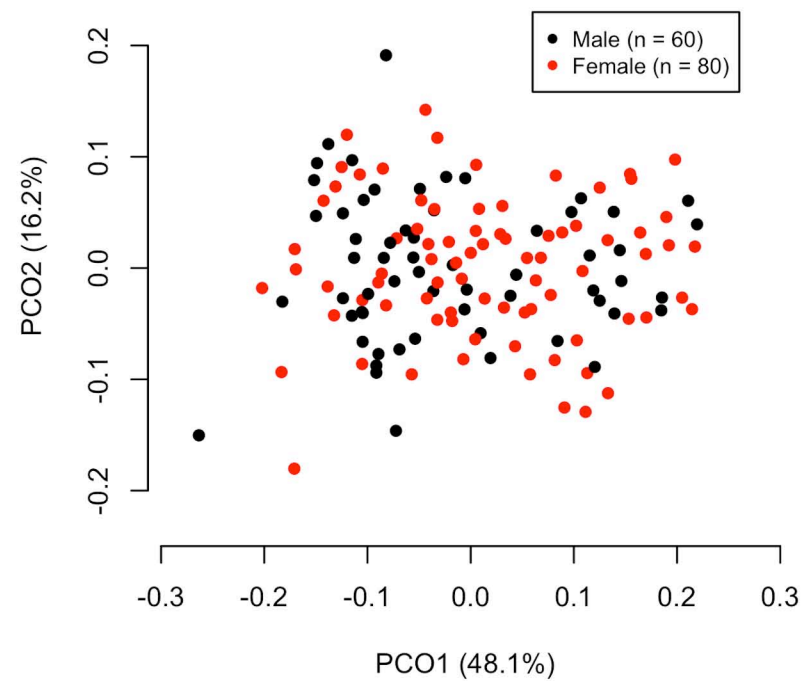
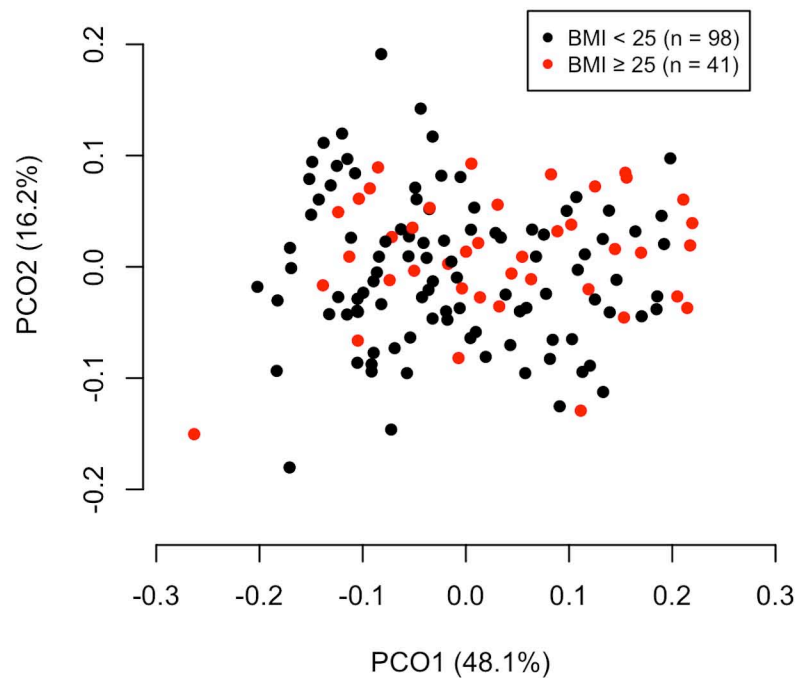
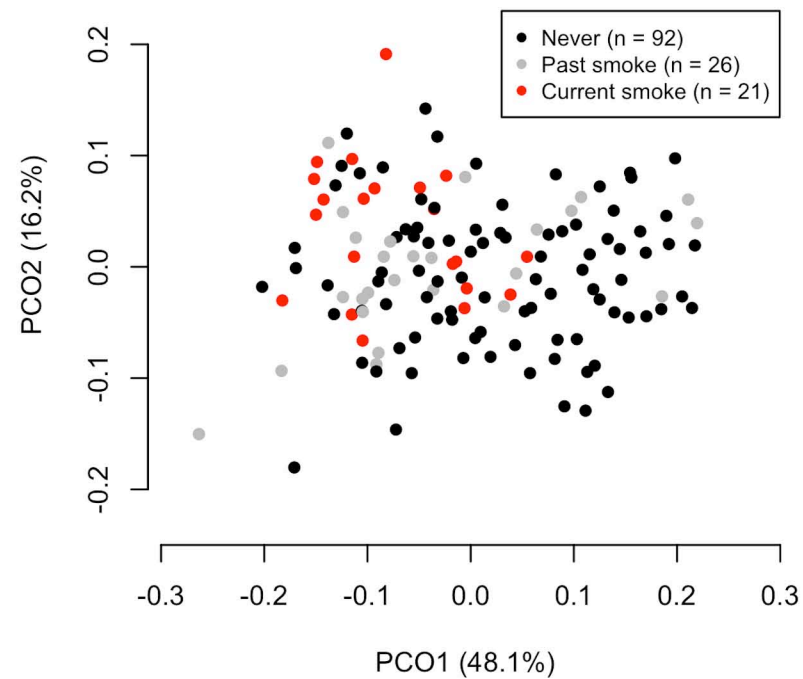
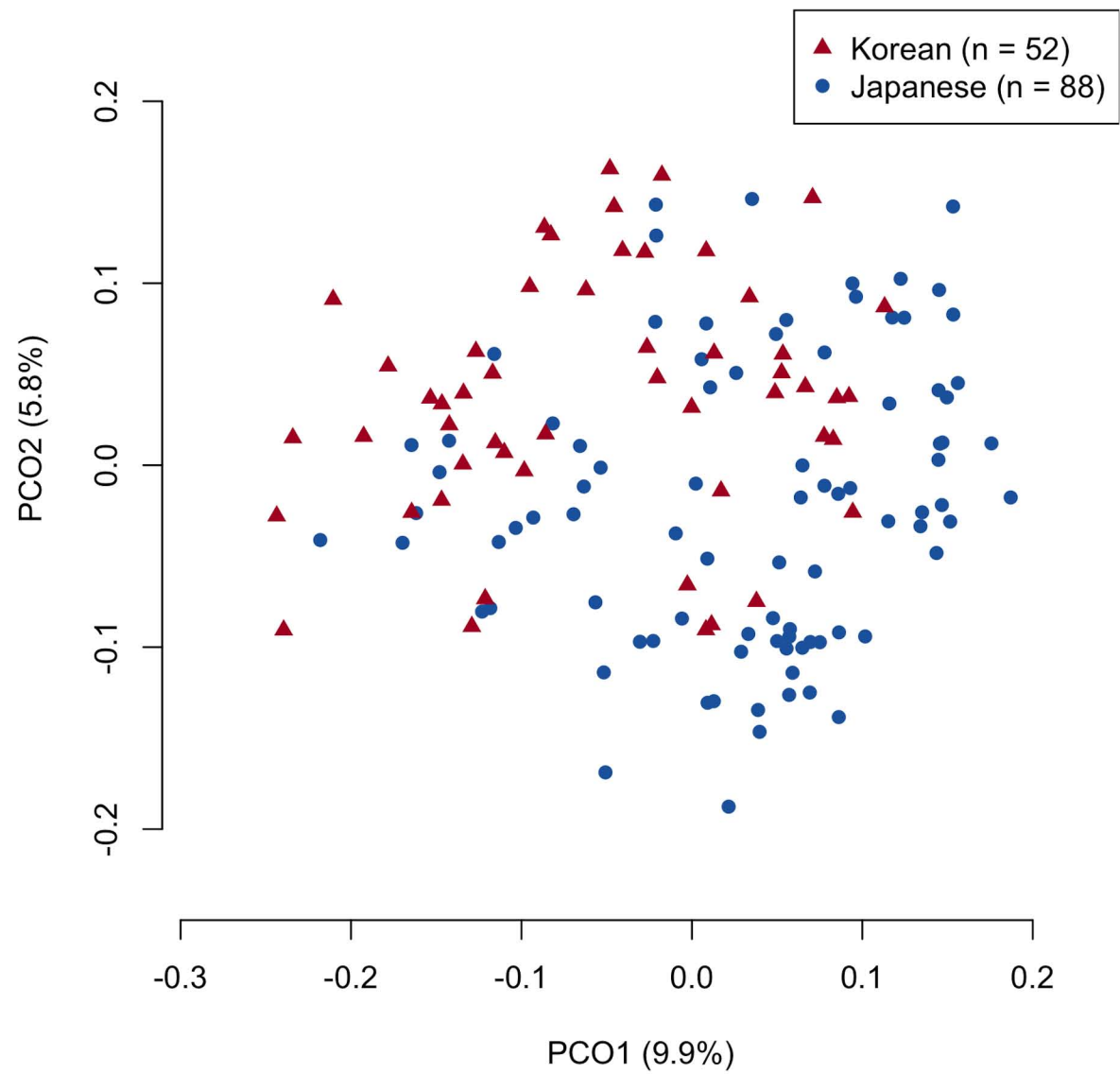
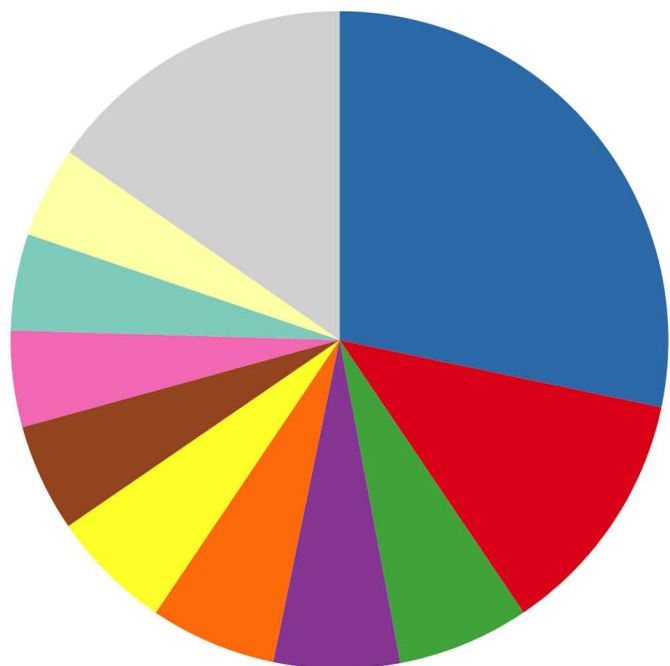
**A****B****C****D**

Fig. S3



Japanese (n = 88)



Korean (n = 52)

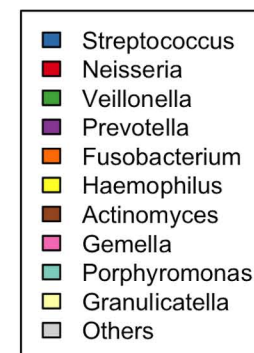
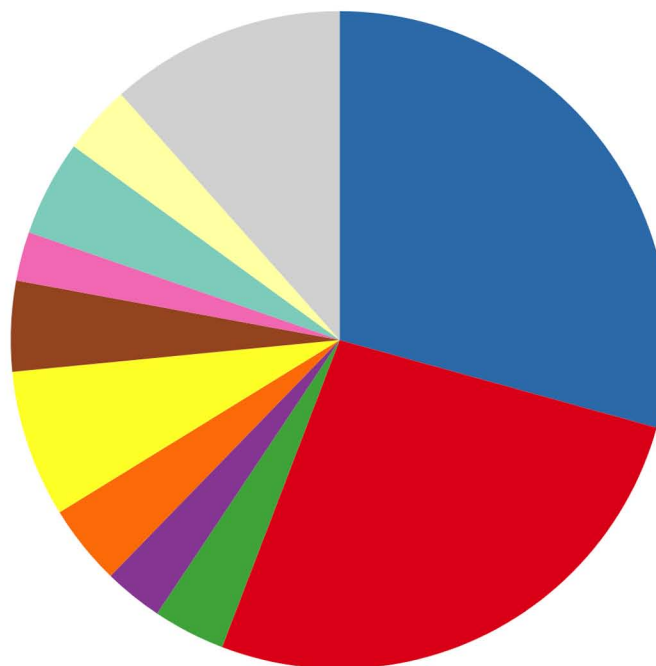
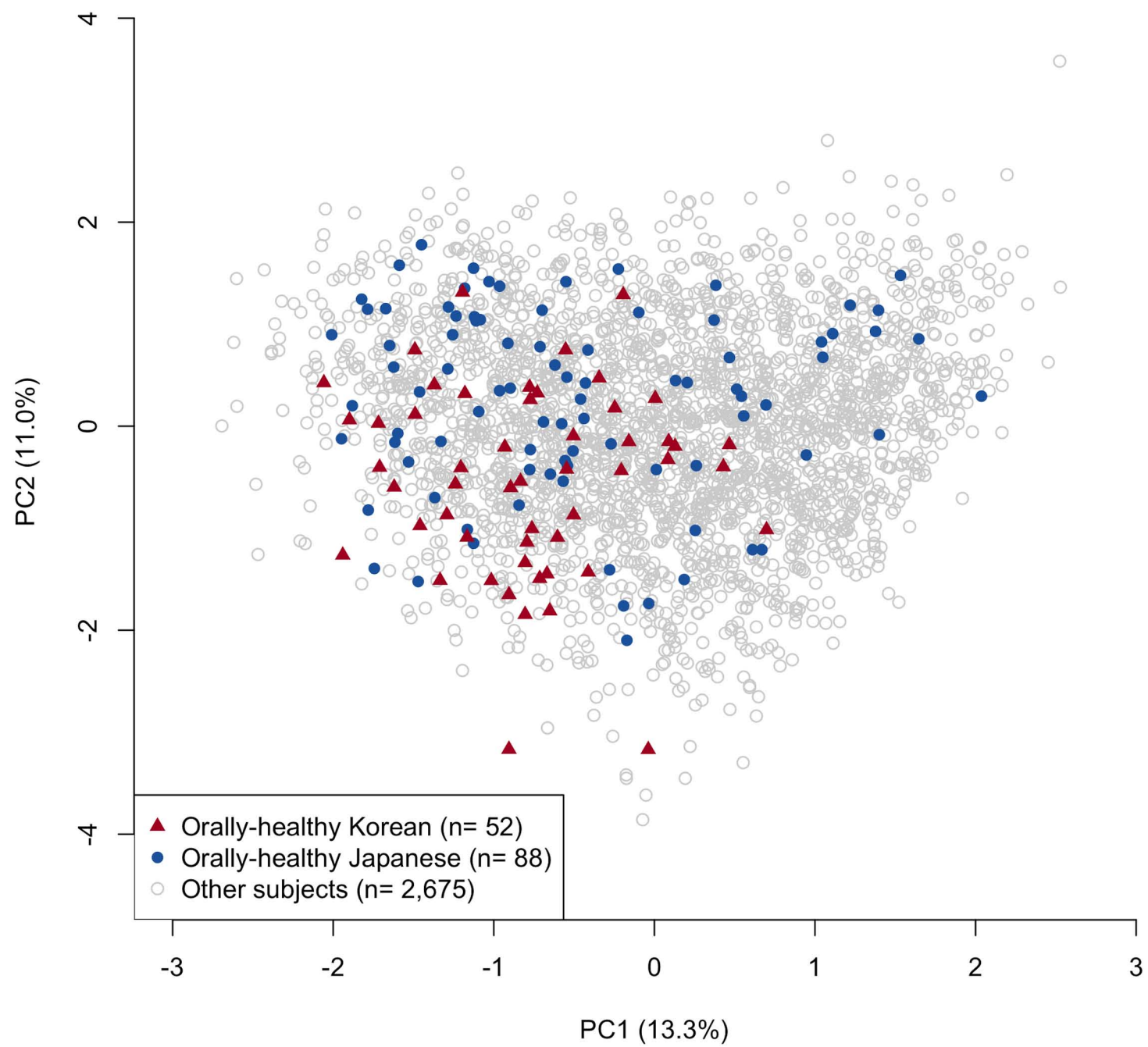


Fig. S5



## Supplementary Figure Legend

**Figure S1** **A** Principal component analysis (PCA) diagram showing similarity relationships among 2,815 T-RFLP profiles. The T-RFLP profiles of subjects in the two countries are depicted using different colors and shapes. **B** Loading plot for the first two components based on 2,815 T-RFLP profiles. Only 28 TRFs with large factor loadings ( $> 0.2$  in absolute value) on the first principal component were selected, and these are indicated by arrows.

**Figure S2** Principal coordinate analysis (PCoA) plot showing similarity relationships among bacterial community samples from 140 orally healthy individuals (52 South Koreans and 88 Japanese) using the weighted UniFrac distance metric. The two components explained 48.1 and 16.2% of the variance, respectively. **A** Samples are depicted using different colors and shapes according to their age. **B** Samples are depicted using different colors and shapes according to their gender. **C** Samples are depicted using different colors and shapes according to body mass index ( $\text{BMI} \geq 25$ , obesity group;  $< 25$ , normal group). One Korean subject had no data about BMI. **D** Samples are depicted using different colors and shapes according to smoking experience. One Korean individual had no data about smoking status.

**Figure S3** Principal coordinate analysis (PCoA) plot showing similarity relationships among bacterial community samples from 140 orally healthy individuals (52 South Koreans and 88 Japanese) using the unweighted UniFrac distance metric. The two components explained 9.9 and 5.8% of the variance, respectively.

**Figure S4** Mean genera abundances in the salivary microbiota of orally healthy individuals in Japan and South Korea.

**Figure S5** Principal component analysis (PCA) diagram showing similarity relationships among 2,815 T-RFLP profiles also described in Figure S1. The T-RFLP profiles of orally healthy individuals in each cohort, who involved in pyrosequencing analysis are depicted using different colors and shapes. These orally healthy individuals were localized in the negative direction of PC1 on a PCA diagram that included all individuals, suggesting that their microbiomes comprised higher proportions of *Neisseria* and *Haemophilus* and lower proportions of *Prevotella* and *Veillonella* compared to subjects that had oral health problems. The peak area proportions of terminal restriction fragments corresponding to *Neisseria* and *Haemophilus*, *Prevotella* or *Veillonella* were described in Table S5.



## Supplemental Tables

Table S1. General and oral health conditions of each study population. [n (%)]

	Hisayama (n = 2,272)	Yangpyeong (n = 543)
Age*	59.9 ± 10.3	61.1 ± 8.9
Male	1011 (44.5)	198 (36.5)
Dental condition		
Number of present teeth*	22.6 ± 7.5	24.4 ± 6.2
≥ 24	1405 (61.8)	378 (69.6)
< 24	867 (38.2)	165 (30.4)
Denture use	626 (27.6)	91 (16.8)
Number of decayed, missing and filled teeth*	19.3 ± 6.6	10.3 ± 7.2
≤ 10	252 (11.1)	322 (59.3)
> 10	2020 (88.9)	221 (40.7)
Number of decayed teeth*	0.7 ± 1.6	1.7 ± 2.2
No teeth	1598 (70.3)	232 (42.7)
≥ 1 teeth	674 (29.7)	311 (57.3)
Periodontal condition		
% teeth with PD of ≥ 4mm*†	10.7 ± 19.4	3.6 ± 10.9
No teeth	1341 (61.2)	442 (82.0)
≥ 1 teeth	851 (38.8)	97 (18.0)
% teeth with CAL of ≥ 5mm*†	9.7 ± 18.9	4.7 ± 11.6
No teeth	1441 (65.7)	402 (74.6)
≥ 1 teeth	751 (34.3)	137 (25.4)
% sites of BOP*†	19.8 ± 23.3	18.8 ± 23.2
≤ 20	1433 (65.4)	368 (68.3)
> 20	759 (34.6)	171 (31.7)

\* Mean ± SD.

† Edentulous individuals were excluded; Hisayama (n = 2,192), Yanpyeong (n = 539).

PD, pocket depth; CAL, clinical attachment loss; BOP, bleeding on probing.

Table S2. Bacterial candidates corresponding to the TRF combinations.

TRF combination	Genus	Species (Oral Taxon ID)
Negative loading for PC1		
F6–R58	<i>Flavobacteriales</i> [G-1]	2 oral taxon (318, 321)
F6–R150	<i>Stenotrophomonas</i>	<i>S. maltophilia</i> (663)
F6–R215	<i>Bergeyella</i>	2 oral taxon (322, 900)
	<i>Bacteroidetes</i> [G-3]	1 oral taxon (365)
F6–R246	<i>Bergeyella</i>	1 oral taxon (319)
F109–R150	<i>Neisseria</i>	<i>N. oralis</i> (14), <i>N. weaveri</i> (92), <i>N. subflava</i> (476), <i>N. elongata</i> (598), <i>N. flava</i> (609), <i>N. flavescens</i> (610), <i>N. lactamica</i> (649), <i>N. meningitidis</i> (669), <i>N. mucosa</i> (682), <i>N. pharyngis</i> (729), <i>N. sicca</i> (764) and 2 oral taxon (15, 16)
	<i>Haemophilus</i>	<i>H. aegyptius</i> (535), <i>H. ducreyi</i> (821), <i>H. influenzae</i> (641), <i>H. parainfluenzae</i> (718), <i>H. haemolyticus</i> (851) and 2 oral taxon (35, 36, 908)
	<i>Aggregatibacter</i>	<i>A. actinomycetemcominans</i> (531), <i>A. aphrophilus</i> (545), <i>A. paraphrophilus</i> (720), <i>A. segnis</i> (762) and 3 oral taxon (458, 512 513, 898)
	<i>Eikenella</i>	1 oral taxon (11)
	<i>Ralstonia</i>	<i>R. pickettii</i> (854) and 1 oral taxon (406)
F127–R150	<i>Neisseria</i>	<i>N. bacilliformis</i> (13), <i>N. polysaccharea</i> (737) and one oral taxon (20)
	<i>Eikenella</i>	<i>E. corrodens</i> (577)
	<i>Kingella</i>	<i>K. oralis</i> (706)
	<i>Simonsiella</i>	<i>S. muelleri</i> (683)

F127–R215	<i>Paenibacillus</i>	1 oral taxon (48)
F183–R58	<i>Parvimonas</i>	<i>P. micra</i> (111) and 2 oral taxon (110, 393)
F183–R215	<i>Streptococcus</i>	<i>S. anginosus</i> (543)
	<i>Parvimonas</i>	<i>P. micra</i> (111)
F195–R150	<i>Treponema</i>	2 oral taxon (251, 256)
F229–R150	<i>Treponema</i>	1 oral taxon (254)
F270–R58	<i>Porphyromonas</i>	<i>P. catoniae</i> (283) and 2 oral taxon (279, 284)
F283–R271	<i>Porphyromonas</i>	<i>P. gingivalis</i> (619) and 2 oral taxon (285, 395)

Positive loading for PC1

F115–R102	<i>Veillonella</i>	<i>V. dispar</i> (160), <i>V. parvula</i> (161) and <i>V. denticariosi</i> (887)
F115–R230	<i>Mitsuokella</i>	<i>M. multacida</i> (684)
F118–R60	<i>Prevotella</i>	<i>P. baroniae</i> (553)
F118–R102	<i>Dialister</i>	<i>D. micraerophilus</i> (843)
	<i>Anaeroglobus</i>	<i>A. geminatus</i> (121)
F118–R230	<i>Dialister</i>	<i>D. pneumonisintes</i> (736) and 1 oral taxon (502)
F135–R144	<i>Corynebacterium</i>	<i>C. durum</i> (595)
F176–R60	<i>Prevotella</i>	<i>P. oulorum</i> (288), <i>P. maculosa</i> (289), <i>P. denticola</i> (291), <i>P. salivae</i> (307), <i>P. melaninogenica</i> (469), <i>P. bivia</i> (562), <i>P. dentalis</i> (583), <i>P. loescheii</i> (658), <i>P. fusca</i> (782), <i>P. shahii</i> (795), <i>P. scopos</i> (885) and 10 oral taxon (293, 306, 309, 313, 314, 317, 396, 443, 472, 820)
F176–R146	<i>Prevotella</i>	<i>P. histicola</i> (298) and <i>P. melaninogenica</i> (469)

	<i>Alloprevotella</i>	<i>A. rava</i> (302) and 6 oral taxon (308, 473, 474, 912, 913, 914)
F176–R230	<i>Alloiococcus</i>	<i>A. otitis</i> (831)
F187–R230	<i>Lactobacillus</i>	<i>L. salivarius</i> (756)

---

Bacterial candidates were selected from 831 oral bacterial 16S rRNA gene sequences (HOMD 16S rRNA RefSeq Version 13.2) deposited in deposited in the Human Oral Microbiome Database<sup>1</sup>.

## Reference

1. Chen T, Yu WH, Izard J, Baranova OV, Lakshmanan A, Dewhirst FE. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. Database (Oxford), baq013 (2010).

Table S3. Summary of pyrosequencing analysis.

Sample name	Tag sequence	Number of reads	Number of high quality reads
Japanese individuals (n = 88)			
HS001	AGACGCACTC	7,385	3,670
HS002	AGCACTGTAG	5,715	2,963
HS003	TCTCTATGCG	7,363	3,907
HS004	TGATACGTCT	6,938	3,408
HS005	ATAGAGTACT	7,233	3,598
HS006	CGACGTGACT	9,177	4,719
HS007	TACACGTGAT	7,167	3,826
HS008	TGACGTATGT	7,540	3,821
HS009	TGTGAGTAGT	7,389	3,411
HS010	ACTCGCGCAC	8,101	4,377
HS011	ATCTACTGAC	8,679	4,226
HS012	CGAGACGCGC	9,607	5,268
HS013	CTCTACGCTC	6,882	3,447
HS014	TCACTCATAAC	6,479	3,155
HS015	TCATCGAGTC	8,367	4,428
HS016	TGAGTGACGC	7,174	3,355
HS017	ATGTACGATG	4,884	2,553
HS018	CACTCGCACG	6,357	3,276
HS019	CTCGCACGCG	8,349	4,140
HS020	CTGTGCGTCG	5,868	2,789
HS021	TATACATGTG	5,483	2,555
HS022	TCTGATCGAG	4,684	2,394
HS023	ACGAGTGCGT	8,244	3,270
HS024	ATATCGCGAG	8,611	3,197
HS025	CTCGCGTGTC	8,687	3,459
HS026	TCGTCGCTCG	7,701	3,193
HS027	ACGCGAGTAT	9,253	3,557
HS028	ACACATACGC	8,909	3,569
HS029	AGTGTATGTC	7,949	2,859
HS030	CATACTCTAC	8,320	3,077

HS031	CGAGACGCGC	8,777	3,267
HS032	CGTCGATCTC	8,553	3,540
HS033	CTCTACGCTC	8,119	3,218
HS034	TAGACTGCAC	7,905	3,190
HS035	ATGTGTCTAG	4,781	1,896
HS036	CGATCTGTCTG	6,456	2,548
HS037	CTATGTACAG	7,706	2,966
HS038	ACGAGTGCCT	8,718	3,401
HS039	AGACGCACTC	8,807	2,969
HS040	AGCACTGTAG	8,649	3,169
HS041	ATCAGACACG	5,947	2,318
HS042	ATATCGCGAG	6,943	2,801
HS043	CTCGCGTGTC	9,127	3,153
HS044	TCTCTATGCG	8,464	3,636
HS045	TGATACGTCT	6,800	2,549
HS046	CATAGTAGTG	8,852	3,331
HS047	CGAGAGATAC	9,533	4,184
HS048	CGTCTAGTAC	9,556	3,898
HS049	TCTACGTAGC	7,023	2,856
HS050	TGTACTACTC	6,925	2,661
HS051	ACGACTACAG	8,573	3,269
HS052	CGTAGACTAG	8,757	3,150
HS053	TACTCTCGTG	7,490	2,451
HS054	TAGAGACGAG	6,817	2,622
HS055	TCGTCGCTCG	6,285	2,566
HS056	ACATACGCGT	7,334	2,967
HS057	ACGCGAGTAT	7,657	3,124
HS058	ACTACTATGT	8,078	3,169
HS059	ACTGTACAGT	8,618	3,208
HS060	AGACTATACT	8,178	3,071
HS061	AGCGTCGTCT	6,453	2,615
HS062	AGTACGCTAT	7,725	2,804
HS063	ATAGAGTACT	6,522	2,559
HS064	CACGCTACGT	7,771	3,077
HS065	CAGTAGACGT	7,450	2,710

HS066	CGACGTGACT	5,823	2,205
HS067	TACACACACT	6,815	2,783
HS068	TACACGTGAT	6,213	2,362
HS069	TACAGATCGT	5,950	2,341
HS070	TACGCTGTCT	7,163	3,174
HS071	TAGTGTAGAT	5,749	2,226
HS072	TCGATCACGT	7,397	2,865
HS073	TCGCACTAGT	7,859	3,224
HS074	TCTAGCGACT	7,668	2,772
HS075	TCTATACTAT	7,767	3,241
HS076	TGACGTATGT	7,491	2,697
HS077	TGTGAGTAGT	5,937	2,270
HS078	ACACATACGC	6,974	3,130
HS079	ACAGTCGTGC	6,523	2,733
HS080	ACATGACGAC	7,553	2,840
HS081	ACGACAGCTC	6,358	2,372
HS082	ACGTCTCATC	7,148	2,987
HS083	ACTCATCTAC	8,192	3,165
HS084	ACTCGCGCAC	7,618	3,003
HS085	AGAGCGTCAC	6,964	2,918
HS086	AGCGACTAGC	7,149	2,906
HS087	AGTAGTGATC	6,072	2,484
HS088	AGTGACACAC	5,779	2,634
South Korean individuals (n = 52)			
YP001	AGTGTATGTC	6,313	2,801
YP002	ATAGATAGAC	6,066	2,687
YP003	ATATAGTCGC	6,043	2,531
YP004	ATCTACTGAC	7,958	3,208
YP005	CACGTAGATC	7,659	3,235
YP006	CACGTGTCGC	6,713	2,849
YP007	CATACTCTAC	8,124	3,682
YP008	CGACACTATC	7,426	3,061
YP009	CGAGACGCGC	7,220	3,127
YP010	CGTATGCGAC	7,162	2,977
YP011	CGTCGATCTC	8,338	3,388

YP012	CTACGACTGC	8,532	3,225
YP013	CTAGTCACTC	7,728	3,522
YP014	CTCTACGCTC	8,444	3,304
YP015	CTGTACATAC	7,335	3,253
YP016	TAGACTGCAC	6,978	3,084
YP017	TAGCGCGCGC	6,364	2,879
YP018	TAGCTCTATC	7,905	3,615
YP019	TATAGACATC	6,284	2,626
YP020	TATGATACGC	8,752	3,642
YP021	TCACTCATAAC	6,120	2,482
YP022	TCATCGAGTC	7,596	3,222
YP023	TCGAGCTCTC	5,745	2,454
YP024	TCGCAGACAC	5,456	2,045
YP025	TCTGTCTCGC	5,256	2,300
YP026	TGAGTGACGC	4,425	1,574
YP027	TGATGTGTAC	5,146	2,149
YP028	TGCTATAGAC	6,601	2,847
YP029	TGCTCGCTAC	6,283	2,410
YP030	ACGTGCAGCG	5,791	2,236
YP031	ACTCACAGAG	6,235	2,487
YP032	AGACTCAGCG	5,498	2,182
YP033	AGAGAGTGTG	6,433	2,590
YP034	AGCTATCGCG	7,446	2,837
YP035	AGTCTGACTG	6,898	3,042
YP036	AGTGAGCTCG	5,624	2,189
YP037	ATAGCTCTCG	6,885	3,219
YP038	ATCACGTGCG	7,238	3,205
YP039	ATCGTAGCAG	7,898	3,561
YP040	ATCGTCTGTG	7,050	2,909
YP041	ATGTACGATG	6,557	2,209
YP042	ATGTGTCTAG	5,536	2,237
YP043	CACACGATAG	8,228	3,476
YP044	CACTCGCACG	6,711	2,637
YP045	CAGACGTCTG	6,417	2,311
YP046	CAGTACTGCG	5,901	2,178



YP047	CGACAGCGAG	7,204	3,084
YP048	CGATCTGTCG	7,239	2,702
YP049	CGCGTGCTAG	5,786	2,112
YP050	CGCTCGAGTG	6,508	2,760
YP051	CGTGATGACG	6,260	2,660
YP052	CTATGTACAG	6,926	2,683

---

Pyrosequencing analyses of the 140 samples were performed in three runs (pool 1, HS001–022; pool 2, HS023–037; pool 3, HS038–088 and YP001–052).

Table S4. Statistics in PERMANOVA for the weighted UniFrac analysis which correspond to the results shown in Fig 1A and S3.

	F value	Partial $\eta^2$	P-value
Geographic difference			
Korean (n = 52) vs. Japanese (n = 88)	31.2	0.18	< 0.001
Age			
40–49 years (n = 42) vs. $\geq$ 50 years (n = 98)	1.32	0.009	0.23
Gender			
Female (n = 60) vs. Male (n = 80)	2.52	0.018	0.051
Obesity (Body mass index, BMI) <sup>a</sup>			
$\geq$ 25 (n = 41) vs. < 25 (n = 98)	1.65	0.011	0.14
Smoking status <sup>a</sup>			
Never (n = 92) vs. Past and Current (n = 47)	9.18	0.062	< 0.001
Never and Past (n = 118) vs. Current (n = 21)	7.85	0.054	< 0.001
Geographic difference in non-smoker <sup>b</sup>			
Korean (n = 45) vs. Japanese (n = 47)	21.4	0.19	< 0.001

<sup>a</sup> One Korean individual had no data about smoking status and BMI.

<sup>b</sup> Geographic difference was assessed after exclusion of individuals with smoking history (n = 92).

Table S5. Comparison of peak area proportions of terminal restriction fragments (TRFs) corresponding to *Neisseria*, *Haemophilus*, *Prevotella* and *Veillonella* species in orally-healthy individuals with those in the other individuals.

TRFs	Peak area proportion of each TRFs					
	Individuals in Japan			Individuals in South Korea		
	Healthy individuals (n = 88)	Other individuals (n = 2,184)	<i>P</i> -value <sup>a</sup>	Healthy individuals (n = 52)	Other individuals (n = 491)	<i>P</i> -value <sup>a</sup>
TRFs corresponding to <i>Neisseria</i> and <i>Haemophilus</i> species <sup>b</sup>						
F109	11.5 ± 6.0	8.5 ± 6.2	<0.001	21.3 ± 6.4	18.0 ± 7.2	0.001
F127	1.5 ± 1.2	0.8 ± 1.1	<0.001	1.8 ± 1.3	0.8 ± 1.2	<0.001
R150	21.1 ± 6.9	18.0 ± 6.7	<0.001	27.4 ± 5.9	23.1 ± 7.3	<0.001
TRFs corresponding to <i>Veillonella</i> species <sup>b</sup>						
F115	13.2 ± 8.3	14.8 ± 8.1	0.008	7.1 ± 6.0	9.5 ± 7.5	0.001
R102	3.9 ± 4.0	5.1 ± 3.7	<0.001	1.9 ± 2.5	3.4 ± 3.7	<0.001
TRFs corresponding to <i>Prevotella</i> species <sup>b</sup>						
F118	0.3 ± 0.9	0.9 ± 1.6	<0.001	0.2 ± 0.6	0.4 ± 1.0	0.17
F176	8.9 ± 4.5	10.9 ± 4.8	<0.001	7.8 ± 4.3	9.7 ± 5.0	0.01
R60	3.3 ± 1.6	3.7 ± 2.0	0.10	3.6 ± 2.0	3.7 ± 2.4	0.87
R146	10.0 ± 4.3	10.8 ± 4.4	0.05	7.7 ± 4.4	9.2 ± 4.8	0.02

<sup>a</sup> *P*-values were calculated by Wilcoxon rank-sum test.

<sup>b</sup> Details of bacterial candidates corresponding to T-RF combinations (F, 6-FAM labeled and R, HEX labeled) were described in Table S2.