Journal of Oral Microbiology

co-action PUBLISHING

ORIGINAL ARTICLE

Concordance of HOMIM and HOMI*NGS* technologies in the microbiome analysis of clinical samples

Jean-Luc C. Mougeot[1]*, Craig B. Stevens[1], Sean L. Cotton[2], Darla S. Morton[1], Keerthana Krishnan[2], Michael T. Brennan[1], Peter B. Lockhart[1], Bruce J. Paster[2] and Farah K. Bahrani Mougeot[1]*

[1]Department of Oral Medicine, Cannon Research Center, Carolinas Health Care System, Charlotte, NC, USA; [2]The Forsyth Institute, Cambridge, MA, USA

Background: Over 700 bacterial species reside in human oral cavity, many of which are associated with local or distant site infections. Extensive characterization of the oral microbiome depends on the technologies used to determine the presence and proportions of specific bacterial species in various oral sites.
Objective: The objective of this study was to compare the microbial composition of dental plaque at baseline using Human Oral Microbe Identification Microarray (HOMIM) and Human Oral Microbe Identification using *Next Generation Sequencing* (HOMI*NGS*) technologies, which are based on 16S rRNA.
Methods: Dental plaque samples were collected from 96 patients at baseline prior to a dental procedure involving manipulation of gingival tissues. The samples were surveyed for 293 and 597 oral bacterial species *via* HOMIM and HOMI*NGS*, respectively, based on 16S rRNA gene sequences. We determined the concordance between the two technologies for common species. Genus level analysis was performed using HOMI*NGS*-specific genus identification capabilities.
Results: HOMI*NGS* detected twice the number of species in the same dental plaque samples compared to HOMIM. For the species detected by both HOMIM and HOMI*NGS*, there was no difference in relative proportions of overall bacterial composition at the species, genus or phylum levels. Additionally, there was no difference in relative proportion for total species per patient between the two technologies.
Conclusion: HOMI*NGS* significantly expanded oral bacterial species identification compared to HOMIM. The genus and species probes, combined in HOMI*NGS*, provided a more comprehensive representation of oral bacterial community, critical for future characterization of oral microbes in distant site infections.

Keywords: *HOMI*NGS; *HOMIM*; *MiSeq*; *ProbeSeq*; *dental plaque*; *oral microbiome*

*Correspondence to: Farah Bahrani Mougeot, Department of Oral Medicine, Cannon Research Center, Carolinas Health Care System, PO Box 32861, Charlotte, NC 28232-2861, USA, Email: farah.mougeot@carolinashealthcare.org; Jean-Luc C. Mougeot, Department of Oral Medicine, Cannon Research Center, Carolinas Health Care System, PO Box 32861, Charlotte, NC 28232-2861, USA, Email: jean-luc.mougeot@carolinashealthcare.org

Received: 12 November 2015; Revised: 10 February 2016; Accepted: 11 February 2016; Published: 8 April 2016

The oral microbiome encompasses over 700 predominant bacterial species, many of which have not been formally named or cultivated (1, 2). These bacterial species reside on various oral surfaces including tongue, buccal mucosa, gingiva, hard palate, and supra- and sub-gingival dental plaque (3, 4).

Numerous oral bacterial species have been associated with oral diseases including dental caries, gingivitis, and periodontitis (5, 6). Some oral bacterial species have increasingly been shown to be associated with systemic diseases such as infective endocarditis (IE) and prosthetic joint infection (PJI) (7). These species are thought to gain access to internal tissues following bacteremia that may result from various manipulations of the oral mucosa, some of which are highly invasive (e.g. tooth extraction) and others are considered less invasive (e.g. tooth brushing or chewing) (8). Oral bacterial niche characterization is, therefore, fundamental for determining predictors of bacteremia and associated systemic disease. Proper characterization of the oral microbiome is dependent upon the technologies used to determine the relative proportions of specific bacterial species in various oral sites.

This study used two semi-quantitative technologies for oral bacterial species identification: the Human

Journal of Oral Microbiology 2016. © 2016 Jean-Luc C. Mougeot et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Citation: Journal of Oral Microbiology 2016, **8**: 30379 - http://dx.doi.org/10.3402/jom.v8.30379
(page number not for citation purpose)

Oral Microbe Identification Microarray (HOMIM) and Human Oral Microbe Identification using *Next Generation Sequencing* (HOMI*NGS*) (9). Both technologies use 16S ribosomal RNA (rRNA) gene sequences for species identification. While HOMIM relies on an *in vitro* hybridization procedure similar to DNA microarray technology, HOMI*NGS* follows an *in silico* hybridization process, in effect processing of unique electronic or 'e'-hybridization events referred to as 'hits'.

Prior to the introduction of HOMIM in 2008, other sequence-based methods, while providing an abundance of 16S rRNA gene data, typically identified taxa at the genus or higher levels (10, 11). HOMIM, for the first time, provided 379 species-level probes capable of identifying 293 predominant oral bacterial species.

HOMI*NGS*, introduced in 2014, provides species-level identification *via* species probes and genus level identification *via* genus probes for those sequences not uniquely e-hybridized to a species probe. Genus probes are useful for detection of species at the genus level that are highly conserved and for which design of specific probe sequences used for e-hybridization is difficult, e.g. *Streptococcus* and *Fusobacterium* species. These probes are also useful to conduct effective genus level analyses. Also, unknown species may be captured by genus probes, providing further opportunities for species probe development. The 672 species probes and 93 genus probes used for HOMI*NGS* are capable of identifying 597 species and 83 genera (due to multiple probes for certain species or genera), respectively. Determination of total hits for a given genus is thereby achieved by adding hits from corresponding species probes to the genus probe hits.

Our objective was to compare baseline oral microbiome of dental plaque using these two technologies.

Materials and methods

Study population

Patients ($n = 96$) were recruited from our hospital-based dental clinic as a subset of a previous clinical bacteremia study (8, 12). All patients had the need for extraction of at least one erupted tooth. Patients were excluded from the study if they had fewer than 10 teeth, had taken a systemic antibiotic within 2 weeks prior to the study, or were immunocompromised. Patient demographics and clinical characteristics are summarized in Table 1 (8). As changes of clinical parameters were not the focus of this study, we kept the description here as basic.

IRB approval and patients' signed consents were obtained for the study.

Sample collection and DNA extraction

Dental plaque samples were collected at least 1 h before any dental manipulation, such as eating, tooth brushing, or tooth extraction. They were acquired by scraping a dental scaler across the supra-and superficial sub-gingival aspect of 1–4 teeth, with the deepest probing depths, and pooled. The sample was suspended in TE buffer (50 mM Tris HCL, 1 mM EDTA pH 7.6), transferred on ice, and stored at -80°C until furthzer analysis. Bacterial DNA extraction was performed using a modification of QIAamp DNA Mini Kit (QIAGEN, Valencia, CA), as described previously (12) and used for simultaneous analysis by HOMIM and HOMI*NGS*.

Table 1. Demographic and clinical characteristics of study cohort

Characteristics[a]	All subjects ($n = 96$)
Mean age in years (SD)	39.8 (12.4)
Male sex	55 [57%]
Ethnicity	
White	26 [27%]
African American	65 [68%]
Hispanic	4 [4%]
Other	1 [1%]
Pocket depth[b]	
Mean (SD)	3.37 (0.97)
Range	1.83–7.89
Calculus index[c]	
Mean (SD)	1.18 (0.72)
Range	0–2.89
Gingival index[d]	
Mean (SD)	1.84 (0.68)
Range	0.17–3.00
Plaque index[e]	
Mean (SD)	1.53 (0.68)
Range	0.07–3.00

[a]Continuous values are mean (SD), discrete values are number [percentage].
[b]Pocket depth: mean periodontal pocket depth for all remaining teeth.
[c]Calculus index: mean score for all tooth scores based on scale: 0 = no calculus; 1 = supra-gingival (not more than 1 mm); 2 = moderate supra-gingival and/or sub-gingival calculus; and 3 = abundance of supra-gingival and sub-gingival calculus.
[d]Gingival index: mean score of all tooth scores based on scale: 0 = normal gingiva; 1 = mild inflammation, a slight change in color and edema, and no bleeding on probing; 2 = moderate inflammation, redness, edema, and bleeding on probing; and 3 = severe inflammation, marked redness and edema, ulcerations, and a tendency toward spontaneous bleeding.
[e]Plaque index: mean score of all tooth areas scores (mesial, distal, facial, and lingual) based on scale: 0 = no plaque in the gingival area; 1 = no plaque visible to the unaided eye, but plaque is visible on the probe after being moved across the gingival crevice; 2 = gingival area covered with a thin to moderately thick layer of plaque visible to the naked eye; and 3 = heavy plaque accumulation and soft debris in the interdental area (8).

Citation: Journal of Oral Microbiology 2016, **8**: 30379 - http://dx.doi.org/10.3402/jom.v8.30379

*HOMIM and HOMI*NGS *platforms*

Bacterial DNA samples were used to survey bacterial identification utilizing HOMIM and HOMI*NGS* technologies. HOMIM uses *in vitro* microarray hybridization technology (13, 14). Briefly, 16S rRNA-based probes were printed on aldehyde-coated slides. 16S rRNA gene sequences were polymerase chain reaction (PCR) amplified and labeled with Cy3-dCTP in a second nested PCR. Hybridization was performed overnight followed by washing. Microarray plate probes were scanned and the resulting fluorescence intensities translated into a 'barcode' format (10). Individual intensity signals were normalized by comparison to the average signals from universal 16S rRNA gene probes (10). Band intensity was scored by approximation using a 0–5 discrete scale, with 0 corresponding to no detection and 5 to highest presence (relative abundance) of a bacterial species. Species were identified based on a BLAST search of the Human Oral Microbiome Database (2).

HOMI*NGS* employs a ProbeSeq program for species detection with modifications as previously described (15). Briefly, 50 ng of genomic DNA was used for each initial PCR. Amplification of the 16S rRNA gene (V3–V4 region) was followed by purification and processing using a modified next generation sequencing method as described by Caporaso et al. (16) using a Miseq (Illumina, Inc., San Diego, CA). ProbeSeq sequence identification used rRNA-based *in silico* probes in a BLAST program to determine the species taxa and frequency (15). HOMI*NGS* follows an *in silico* hybridization process. Species-specific 16S rRNA-based oligonucleotide probes, many of which had originally been designed for HOMIM, were used in a BLAST program, called ProbeSeq for HOMI*NGS*, to identify the frequency of oral bacterial targets. ProbeSeq loads the raw sequence files into a cell array and then loops through the array, one sequence at a time, looking for a 'string' (segment of text) that matches one of the oligomers. When a match is found, a running counter starts to give the total number of probe 'hits'. Partial matches are not considered as a match.

The ProbeSeq process was iterative, that is, the sequences that had not been detected by a single species probe were subsequently processed against genus probes, which consist of two or more species within the genus. All hits were accumulated by species/genus by patient, with a higher hit total interpreted as representing a higher presence of a given species. Modifications of the protocol include the following: DNA was amplified directly in the initial PCR—two rounds of PCR were used in the previous study (15), and chimeric sequences were not removed from analyses for this study. However, in a separate analysis, chimeric sequences ranged from 10 to 15% of the total reads. The relative proportions of detected taxa did not vary significantly (data not shown).

Statistical analysis

HOMIM intensity data (0–5) and HOMI*NGS* hits data (0–300,000 range) were provided as Excel spreadsheets for downstream statistical analysis. HOMIM *in vitro* signal intensity estimations and HOMI*NGS* e-hybridization hits are not directly comparable. Therefore, the data from both technologies were converted into relative proportions to allow semi-quantitative comparison of their outcomes. Common species total abundance per patient between the two technologies was determined based on species probes information. For the common species, Wilcoxon signed-rank test was used to determine the statistical significance of differences found between HOMIM and HOMI*NGS* relative proportions at phylum, genus, and species levels, using a significance level of <0.05 (FDR [false discovery rate] adjusted *p*-value). Statistical analyses including outlier determination were performed using XLSTAT-Pro (version 2014.4.06) and/or SAS Enterprise Guide® (version 6.1).

Results

*HOMIM and HOMI*NGS *capacities and bacterial identification in dental plaque samples*

HOMI*NGS* and HOMIM data were obtained from dental plaque samples of a patient cohort scheduled for single tooth extraction. Demographics and gingival characteristics of these patients are shown in Table 1. HOMI*NGS* doubles the capacity for identification of bacterial species and increases the number of identifiable genera by nearly 85% compared to HOMIM, based on use of species and

Table 2. HOMI*NGS* and HOMIM capacities based on Human Oral Taxonomy (HOT) designation.

Capacity criteria	HOMIM	HOMI*NGS*
Species probes[a] (genera)	379 (92)	672 (158)
Identifiable species[b] (genera)	293 (91)	597 (158)
Genus probes[c] (genera)	N/A	93 (83)
Species + Genus probes[d] (genera)	N/A	765 (159)

[a]Species probes provide three categories of species identification: a unique species that is recognized by a single probe (vast majority), a unique species that is recognized by 2–3 probes, and groups of 2–3 species that are recognized by a single probe.
[b]The number of single species that can be identified by each technology, after processing of multiple recognition occurrences, is presented.
[c]This principle also applies to genus probes.
[d]In HOMI*NGS*, species probes hits may be summed with genus probes hits to account for total genera hits. The combination of species and genus probes can identify 159 genera, as a few species or genus probes do not have matching genus or species probes, respectively.
(genera): Number of corresponding genera.

Table 3. Dental plaque microbiome composition by HOMIM and HOMI*NGS*

a. Comparative analysis of the taxa detected by both HOMIM and HOMI*NGS*

				Common identification comparison	
Taxonomic group	Identifiable species	Identified taxa	Common identification	Raw *p*-value	FDR
HOMIM species	293	244	198	0.093	0.139
HOMI*NGS* species	597	489			
HOMIM genera	91	84	74	0.034	0.102
HOMI*NGS* genera	158	129			
HOMIM phyla	10	10	10	0.695	0.695
HOMI*NGS* phyla	12	12			

b. HOMI*NGS* versus HOMIM added taxa identification capacity

	HOMI*NGS* (% hits)			HOMIM (% intensity)
Taxonomic group[a]	Species probes	Genus probes	All probes	Species probes
Species	100	N/A	100	100
Porphyromonas gingivalis	5.6	N/A	5.6[b]	1.2[b]
Diaiister invisus	3.0	N/A	3[b]	21[b]
Filifactor afocis	2.8	N/A	2.8	1.3
Prevotella melaninogenica	2.7	N/A	2.7	0.1
Rothia dentocariosa	2.6	N/A	2.6	0.6
Remainder species	83.3	N/A	83.3	94.7
Genus	61.3	38.7[c]	100	100
Fusobacterium	2.1	21.3	23.4[d]	2.8[d]
Streptococcus	6.2	7.9	14.1[d]	16.4[d]
Prevotella	7.7	0.7	8.4	4.0
Leptotrichia	4.3	1.1	5.4	0.2
Porphyromonas	4.9	0.2	5.1	2.8
Remainder genera	36.1	7.5	43.6	73.8
Phylum	61.3	38.7	100	100
Firmicutes	22.1	10.5	32.6	64.4
Fusobacteria	6.9	22.5	29.4[e]	3.8[e]
Bacteroidetes	15.8	1.2	17	11.8
Actinobacteria	8.8	2.1	10.9	3.5
Spirochaetes	2.1	1.3	3.4	1.6
Remainder phyla	5.6	1.1	6.7	14.9

The data from patient cohort ($n = 96$) are presented. Gray shaded areas show notable results. (a) The number of taxa (species, genera, and phyla) were determined based on the designation of species and/or genus probes. For the species identified in common by both HOMIM and HOMI*NGS* (based on species probes only), there was no overall difference at species, genus, or phylum level, as determined by Wilcoxon signed-rank test with FDR procedure (FDR > 0.05). (b) Additional HOMI*NGS* species and genus probes capacities are taken into account for comparison. Relative proportions are shown in percentage.

[a]For each taxonomic group, the five taxa with the highest relative proportion are shown.

[b]At species level, relative proportions for *Porphyromonas gingivalis* and *Dialister invisus* show inverse trend between HOMIM and HOMI*NGS*.

[c]At genus level, genus probes account for 38.7% of total hits corresponding to all the genera identified. The largest contributors were *Fusobacterium* and *Streptococcus.* Compared to *Streptococcus*, *Fusobacterium* demonstrated the largest contribution by HOMI*NGS* genus probes for genera undetected by HOMIM.

[d]Relative proportions for *Fusobacterium* and *Streptococcus* show inverse trend between HOMIM and HOMI*NGS*, due to genus probes.

[e]At the phylum level, Firmicutes and Fusobacteria were the largest contributors by HOMI*NGS*, with the latter benefiting most from genus probes hits.

Citation: Journal of Oral Microbiology 2016, **8**: 30379 - http://dx.doi.org/10.3402/jom.v8.30379

genus probes counts (Table 2). In our study of 96 plaque samples, HOMI*NGS* detected 489 species compared to 244 species by HOMIM (Table 3). This is consistent with the doubled capacity provided by HOMI*NGS* technology. A greater number of genera were also detected by HOMI*NGS* (i.e. 129) than HOMIM (i.e. 84), corresponding to 53% increase (Table 3). Detailed examination of the 129 genera revealed that 59 were represented by both species and genus probes, while 68 (i.e. 105 species) were represented by species probes alone. Unaccounted reads, that is, sequence reads that did not uniquely e-hybridize to species or genus probes, ranged from 4.5 to 36.7% [Mean% (SD): 14% (4.9)] of total sequences per sample and were over 20% for nine of the 96 samples (data not shown).

Analysis of the common taxa detected by both HOMIM and HOMINGS

Because HOMI*NGS* detected double the number of species, full dataset statistical comparisons at any taxonomy level would be deemed meaningless, as it would simply highlight the obvious. To assess the concordance between the two technologies, we adjusted the HOMI*NGS* dataset to match those species contained in HOMIM dataset to enable relative proportion of global comparisons at the phylum, genus, and species levels. Both HOMIM and HOMI*NGS* identified 198 species in common, corresponding to 74 genera and 10 phyla (Table 3). Comparison of the relative proportions at three taxonomic levels (e.g. species, genera, and phyla) obtained by HOMIM and HOMI*NGS* showed no significant overall differences when FDR correction was applied for the multiple comparisons (FDR >0.05) (Table 3).

In addition, we compared the total common species (i.e. 198 spp.) relative proportion per patient with regard to total HOMIM intensities or HOMI*NGS* hits for all 96 patients. No significant difference was found ($p=0.617$) (Fig. 1a). These results did not change when removing patient 26 (Pt26), the most prominent outlier according to HOMI*NGS* total species/genus relative proportions per patient ($p=0.805$) (Fig. 1b). Pt26 had the highest number of total hits by HOMI*NGS* (116,385, i.e. 2.75% per total 96 patients) compared to a relatively lower total intensity by HOMIM (i.e. 149 total intensity, i.e. 1.09% per total 96 patients). Further analysis showed that *Prevotella melaninogenica* was responsible for this discrepancy compared to the other 197 common species in Pt26. *P. melaninogenica* had 60,505 hits representing 52% of the total hits of 116,385 (data not shown).

Analysis of the HOMINGS added taxa identification capacity compared to HOMIM

Descriptive analysis of HOMI*NGS* added taxa identification capacity is summarized in Table 3, and Figs. 2 and 3. At the species level, *Porphyromonas gingivalis* was the most represented species by HOMI*NGS* (Table 3). The proportion of this species was nearly twice that of next most represented species, *Dialister invisus*. The opposite ratio was observed for the proportions of these two species detected by HOMIM (Table 3).

At the genus level, HOMI*NGS* genus probes accounted for 38.7% of total hits (Table 3). Due to this added capacity, larger relative proportions compared to HOMIM were apparent for several genera including *Fusobacterium*, *Prevotella*, and *Leptotrichia* (Table 3, Fig. 2). The highest relative proportion was observed for *Fusobacterium,* that is, 23.4% (including 2.1% species probes derived genera) by HOMI*NGS,* compared to 2.8% detected by HOMIM (Table 3). The next four genera with the highest relative proportion detected by HOMI*NGS* were *Streptococcus* (14.1% including 6.2% species probes derived genera) followed by *Prevotella*, *Leptotrichia*, and *Porphyromonas* (Table 3). Of these genera, *Streptococcus* yielded similar proportions by HOMIM (16.4%).

At the phylum level, Firmicutes represented twice the proportion by HOMIM (64.4%) compared to HOMI*NGS* (32.6%) (Table 3, Fig. 3). Firmicutes and Fusobacteria were represented in different proportions by HOMIM alone (64.4 and 3.8%), whereas HOMI*NGS* yielded nearly equal representation of these two phyla (32.6 and 29.4%). Additionally, Fusobacteria were detected at a larger proportion by HOMI*NGS* (29.4%) compared to HOMIM (3.8%) (Table 3, Fig. 3). Similarly, Bacteroidetes and Actinobacteria phyla proportions differed to a large extent using HOMIM (11.8 and 3.5%) and were in similar range (17 and 10.9%) using HOMI*NGS*.

Discussion

Identification of bacterial species from various sites in the mouth and quantification of the relative abundance of each species is a daunting task, given the wide diversity of species known to inhabit the human oral cavity. Both HOMIM and HOMI*NGS* provide semi-quantitative identification of oral microbiome bacterial species.

Using the same dental plaque samples in this study, HOMI*NGS* technology expanded the detection of the number of species (*via* species-level probes) by 100% and the number of genera (*via* species probes and genus probes) by nearly 85%. This represents a significant improvement toward achieving full knowledge of oral bacterial composition at the species level.

Incorporation of genus probes by HOMI*NGS* represents a significant improvement. This capability gives recognition to the fact that bacterial species do not all provide equal unique e-hybridization potential based on the degree of conservation of the V3–V4 region of the 16S rRNA gene. Examples of highly conserved species that are not effectively uniquely e-hybridized to species-level probes are found in several genera including *Fusobacterium* and *Streptococcus*. By using an iterative process to

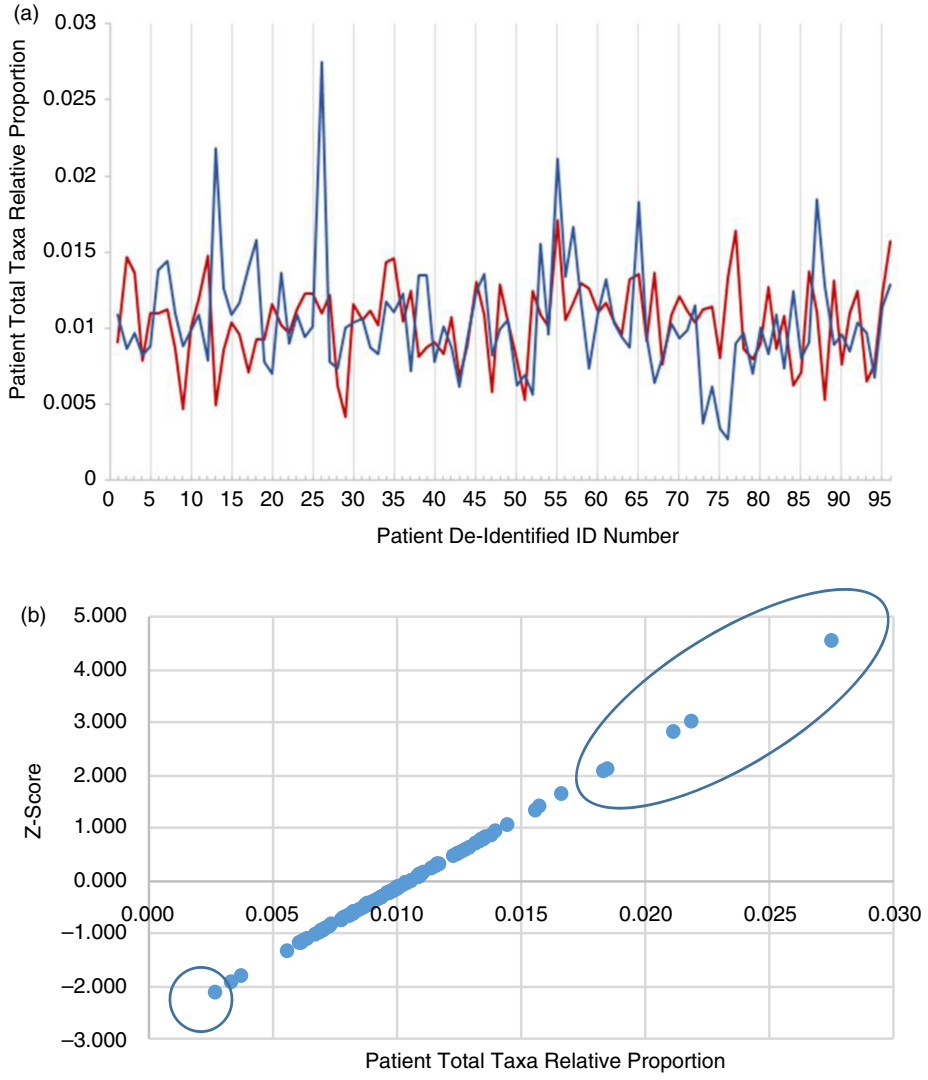

Fig. 1. HOMI*NGS versus* HOMIM comparison of total taxa relative proportion per patient. Total taxa (i.e. species common to both technologies) relative proportion for a patient was calculated by dividing the sum of the intensity scores (HOMIM) or hits (HOMI*NGS*) of all bacterial taxa of the patient by the total number of these values for all 96 patients. (a) Differences in relative taxa proportions are illustrated by patient to patient line chart (HOMI*NGS* [blue]; HOMIM [red]). These differences were not statistically significant overall ($p = 0.617$, Wilcoxon signed-rank test). (b) Grubbs test identified six patients as outliers, with patient 26 (Pt26) representing the most prominent outlier. Removing Pt26 from the analysis also showed that differences in taxa relative proportions were overall not statistically significant ($p = 0.805$).

combine species with genus probe e-hybridization events, HOMI*NGS* technology maximizes reads identification. However, for our dental plaque dataset, there was an average of 14% unaccounted reads per patient. These correspond to sequences which either do not e-hybridize to a target probe or e-hybridize to multiple target probes. These sequences provide opportunities for further species probe refinement and new probe development in order to identify all bacterial species present in oral samples to the lowest detection limit possible.

Identification of species-level information is preferable, but in the absence of this, genus level information is useful and informative. In fact, genus probes altogether accounted for nearly 40% of total hits in this study. Therefore, for genera that have species probes but no genus probes, it is unclear whether there are species remaining unaccounted for within the genus. In addition, within the bulk of genus probes total hits, *Fusobacterium* represented ~20% of hits, followed by, but to a significantly smaller extent, *Streptococcus* (~8%). Therefore, the largest contribution by HOMI*NGS* with the addition of genus probes is the added capacity of detecting more *Fusobacterium* compared to HOMIM. Based on analyses of the samples tested, it seems clear that some key taxa of *Fusobacterium*, *Leptotrichia*, and/or *Actinomyces* are not differentiated at the species level. Additionally,

Citation: Journal of Oral Microbiology 2016, **8**: 30379 - http://dx.doi.org/10.3402/jom.v8.30379

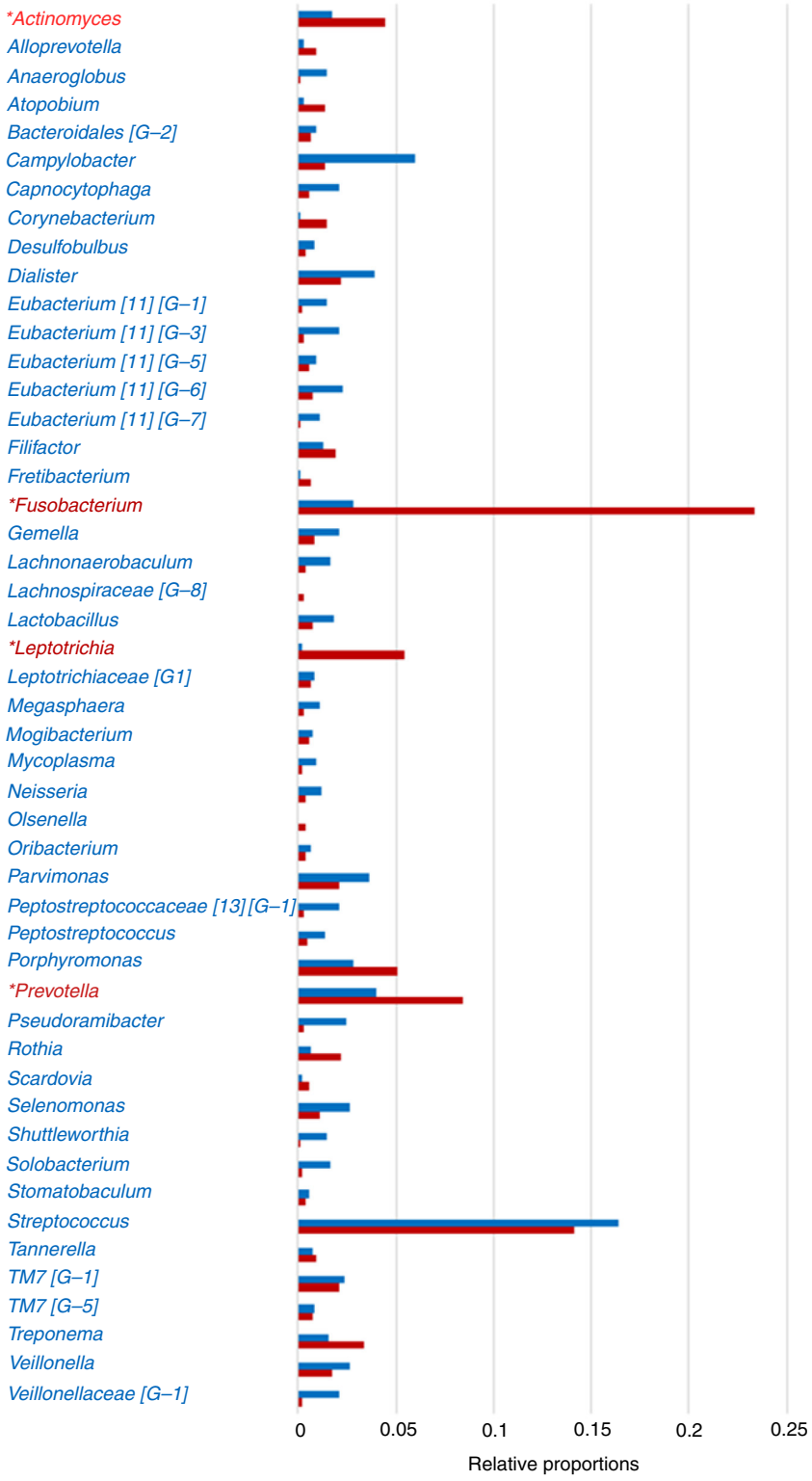

Fig. 2. Dental plaque bacterial genera composition by HOMI*NGS* with added taxa identification capacity compared to HOMIM. *HOMI*NGS* microbiome profiling of dental plaque samples from a patient cohort ($n = 96$) shows a large increase ($> 100\%$) in the detection of *Actinomyces*, *Fusobacterium*, *Leptotrichia,* and *Prevotella* genera compared to HOMIM, likely due to additional genera probes. Relative proportions per genus were calculated based on total hits by HOMI*NGS* or total intensity scores by HOMIM per 96 patients. Genera accounting for 96% of total hits by HOMI*NGS* are represented.

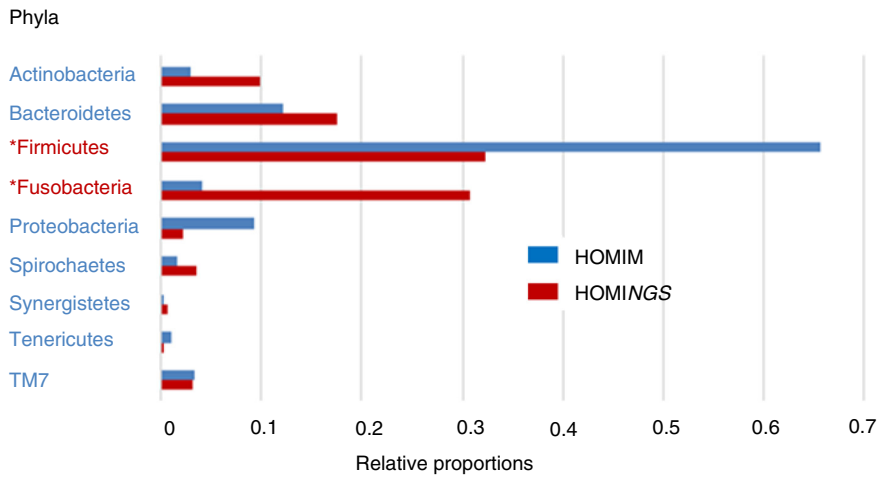

Fig. 3. Dental plaque bacterial phyla composition by HOMI*NGS* with added taxa identification capacity compared to HOMIM. *HOMI*NGS* microbiome profiling of dental plaque samples from a 96 patients cohort shows a large shift in relative proportions for Fusobacteria (3.8% HOMIM vs. 29.4% HOMI*NGS*) and Firmicutes (64.4% HOMIM vs. 32.6% HOMI*NGS*). Relative proportions per phylum were calculated based on total hits by HOMI*NGS* or total intensity score by HOMIM per 96 patients. Chloroflexi, GN02 and SR1 were omitted due to negligible representation.

similar proportions for Firmicutes with Fusobacteria and Bacteroidetes with Actinobacteria were determined by HOMI*NGS*, but not by HOMIM, as probes for many of these taxa were difficult to construct within proper hybridization parameters, for example, GC content and T_m. Nevertheless, HOMIM phylum level data were in overall agreement with the literature (4), namely, Firmicutes was the most predominant phylum. Results obtained by HOMI-*NGS* suggest a more balanced representation of several phyla of dental plaque. However, they might only be representative of the patient population in need for tooth extraction used in our study. Therefore, the results may not be extrapolated to the general population with a better oral hygiene and less dental disease, which clearly may impact the composition of the dental plaque microbiome.

In addition, the introduction of unique e-hybridization events (hit counting) to determine the relative abundance of species and genera is a more precise semi-quantitative process as compared to HOMIM. The fluorescence-based quantification employed in HOMIM provides only a rough estimate of relative proportions on a narrow discrete scale basis (0–5). In contrast, HOMI*NGS* *in silico* matching of target sequence to probe must be exact for a hit to be counted, resulting in no ambiguity of presence or absence of a species. Thus, a low number of hits determined by HOMI*NGS* would most likely correspond to background intensity or no detection by HOMIM. Taking into account the fact that there is no practical limitation to the number of hits that can be counted, hit counting intrinsically provides a more precise and comprehensive representation of the species relative abundance.

These fundamental changes make direct comparisons of HOMI*NGS* semi-quantitative data with those of HOMIM difficult. A fair comparison cannot be achieved because there is no method to convert *in vitro* hybridization intensities accurately for each species into e-hybridization hits. Nonetheless, in the present study, a comparison of relative proportions limited to those species found in both HOMIM and HOMI*NGS* demonstrated concordance between the two technologies in that there were no statistically significant differences noted in the overall results at the phylum, genus, or species taxonomic levels. Concordance between the two technologies was also found regarding the total species relative proportion by patient. In essence, HOMI*NGS* partially confirmed HOMIM data obtained from the same dental plaque sample collection, while differences largely out of range might have been anticipated. In this respect, previous studies that have used HOMIM have been helpful to the extent that differences between disease and control groups can be understood on a global microbial community at the species level to the limit of technology accuracy.

In conclusion, the determination of bacterial species composition in the oral cavity will likely move forward in a significant way with the adoption of HOMI*NGS*. Assuming that further improvements of HOMI*NGS* or other bioinformatic technology would enable detection of all the species present in various oral sites, true quantification may be achieved based on the calibration of bacterial DNA input and knowledge of the bacterial genome sizes for the species detected. Refinement of HOMI*NGS* will provide a better identification tool for

Citation: Journal of Oral Microbiology 2016, **8**: 30379 - http://dx.doi.org/10.3402/jom.v8.30379

the oral microbiome and will be beneficial to the understanding of distant site infections originating from the oral cavity, such as IE and PJIs.

Author contributions

J-L.C. Mougeot designed the study and wrote the manuscript. C.B. Stevens contributed to data analysis and writing of the manuscript. S.L. Cotton and K. Krishnan performed ProbeSeq/HOMI*NGS* analysis. D.S. Morton performed sample processing. M.T. Brennan directed the clinical staff for patient recruitment and samples collection. P.B. Lockhart designed the bacteremia study and conducted the clinical study. B.J. Paster designed and led ProbeSeq/HOMI*NGS* analysis and reviewed the manuscript. F. B. Mougeot designed the study and wrote the manuscript.

Acknowledgements

This study was supported by National Institute of Dental and Craniofacial Research/National Institute of Health grants # R01 DE13559-01, R01 DE021565, and Carolinas HealthCare System (CHS) Research Fund. We thank the residents and clinical coordinators at CMC Dental Clinic for sample collection, Megan Templin for her help in statistical analysis, and P. Hanjra for her help with the database.

Conflict of interest and funding

The authors declare no conflict of interest.

References

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 2010; 2010: baq013.
2. HOMD database: Human Oral Microbiome Database, All Human Microbial Taxa. Available from: http://www.homd.org/modules.php?op=modload&name=HOMD&file=index [cited 26 June 2015].
3. Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner ACR, Yu WH, et al. The human oral microbiome. J Bacteriol 2010; 192: 5002–17.
4. Paster BJ, Boches SK, Galvin JL, Ericson RE, Lau CN, Levanos VA, et al. Bacterial diversity in human subgingival plaque. J Bacteriol 2001; 183: 3770–83.
5. Belstrøm D, Fiehn NE, Nielsen CH, Klepac-Ceraj V, Paster BJ, Twetman S, et al. Differentiation of salivary bacterial profiles of subjects with periodontitis and dental caries. J Oral Microbiol 2015; 7: 27429. doi: http://dx.doi.org/10.3402/jom.v7.27429
6. Shaddox LM, Huang H, Lin T, Hou W, Harrison PL, Aukhil I, et al. Microbiological characterization in children with aggressive periodontitis. J Dent Res 2012; 91: 927–33.
7. Mougeot FB, Saunders SE, Brennan MT, Lockhart PB. Associations between bacteremia from oral sources and distant-site infections: tooth brushing versus single tooth extraction. Oral Surg Oral Med Oral Pathol Oral Radiol 2015; 119: 430–5.
8. Lockhart PB, Brennan MT, Sasser HC, Fox PC, Paster BJ, Bahrani-Mougeot FK. Bacteremia associated with toothbrushing and dental extraction. Circulation 2008; 117: 3118–25.
9. HOMI*NGS*, Forsyth Institute. Available from: http://homings.forsyth.org/index2.html [cited 26 June 2015].
10. Cotton SL, Klepac-Ceraj V, Murphy CM, Kokaras AS, Paster BJ. Species level determination of high-throughput sequencing data using HOMIM Probes. J Dent Res (AADR) 2013; 92: 3828.
11. Segata N, Haake SK, Mannon P, Lemon KP, Waldron L, Gevers D, et al. Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. Genome Biol 2012; 13: R42.
12. Bahrani-Mougeot FK, Paster BJ, Coleman S, Ashar J, Barbuto S, Lockhart PB. Diverse and novel oral bacterial species in blood following dental procedures. J Clin Microbiol 2008; 46: 2129–32.
13. Colombo APV, Bennet S, Cotton SL, Goodson JM, Kent R, Haffajee AD, et al. Comparisons of subgingival microbial profiles of refractory periodontitis, severe periodontitis and periodontal health using the Human Oral Microbe Identification Microarray (HOMIM). J Periodontol 2009; 80: 1421–32.
14. Colombo APV, Bennet S, Cotton SL, Goodson JM, Kent R, Haffajee AD, et al. Impact of periodontal therapy on the subgingival microbiota of severe periodontitis: comparison between good responders and individuals with refractory periodontitis using the Human Oral Microbe Identification Microarray. J Periodontol 2012; 83: 1279–87.
15. Gomes BP, Berber VB, Kokaras AS, Chen T, Paster BJ. Microbiomes of endodontic–periodontal lesions before and after chemomechanical preparation. J Endod 2015; 41: 1975–84. doi: http://dx.doi.org/10.1016/j.joen.2015.08.022
16. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. PNAS 2011; 108(Suppl 1): 4516–22.