REVIEW ARTICLE

Metatranscriptomics reveals overall active bacterial composition in caries lesions

Aurea Simón-Soro, Miriam Guillen-Navarro and Alex Mira*

FISABIO Foundation, Center for Advanced Research in Public Health, Valencia, Spain

Background: Identifying the microbial species in caries lesions is instrumental to determine the etiology of dental caries. However, a significant proportion of bacteria in carious lesions have not been cultured, and the use of molecular methods has been limited to DNA-based approaches, which detect both active and inactive or dead microorganisms.

Objective: To identify the RNA-based, metabolically active bacterial composition of caries lesions at different stages of disease progression in order to provide a list of potential etiological agents of tooth decay.

Design: Non-cavitated enamel caries lesions ($n = 15$) and dentin caries lesions samples ($n = 12$) were collected from 13 individuals. RNA was extracted and cDNA was constructed, which was used to amplify the 16S rRNA gene. The resulting 780 bp polymerase chain reaction products were pyrosequenced using Titanium-plus chemistry, and the sequences obtained were used to determine the bacterial composition.

Results: A mean of 4,900 sequences of the 16S rRNA gene with an average read length of 661 bp was obtained per sample, giving a comprehensive view of the active bacterial communities in caries lesions. Estimates of bacterial diversity indicate that the microbiota of cavities is highly complex, each sample containing between 70 and 400 metabolically active species. The composition of these bacterial consortia varied among individuals and between caries lesions of the same individuals. In addition, enamel and dentin lesions had a different bacterial makeup. Lactobacilli were found almost exclusively in dentin cavities. Streptococci accounted for 40% of the total active community in enamel caries, and 20% in dentin caries. However, *Streptococcus mutans* represented only 0.02–0.73% of the total bacterial community.

Conclusions: The data indicate that the etiology of dental caries is tissue dependent and that the disease has a clear polymicrobial origin. The low proportion of mutans streptococci detected confirms that they are a minority and questions its importance as the main etiological agent of tooth decay. Future experimental work should be performed to confirm the cariogenicity of the identified bacteria.

Keywords: *16S rRNA*; *pyrosequencing*; Streptococcus mutans; *polymicrobial disease*; *tissue-dependent hypothesis*; *caries etiology*

*Correspondence to: Alex Mira, FISABIO Foundation, Center for Advanced Research in Public Health, Avda. Cataluña 21, 46020 Valencia, Spain, Email: mira_ale@gva.es

To access the supplementary material for this article, please see Supplementary files under Article Tools online

It has been estimated that approximately 50% of oral bacteria have not been cultured to date (1). Classical studies based on microbial culture established mutans streptococci and lactobacilli as the main causative agents of dental caries (2, 3). However, other microbial species were also isolated from caries lesions and have been related to the disease, including bifidobacteria and *Scardovia* (4, 5). Furthermore, the application of molecular cloning and Sanger sequencing to study carious lesions at different stages of the disease revealed that although *Streptococcus mutans* levels correlated with disease severity, it could not be always amplified by polymerase chain reaction (PCR) whereas other bacteria were present, including *Prevotella*, *Atopobium*, and *Propionibacterium* (6). When the metagenomic DNA from individual dentin caries samples was directly sequenced without the need for PCR, the genus *Veillonella* appeared dominant within a surprisingly diverse community (7), underlining the varying nature of microbial composition in cavities. The application of pyrosequencing to PCR products of the 16S rDNA gene has become an extremely powerful approach, revealing that cavities are extraordinarily diverse ecosystems (8)

where *S. mutans* accounts at most for 1.6% of the carious lesion bacterial community (9).

A drawback of these DNA-based studies is that the PCR step may amplify DNA from inactive or even dead microorganisms, making it necessary to determine the functional bacteria that effectively contribute to the disease (10). A way to achieve this is to perform the 16S gene amplification starting from RNA material, given that the amount of rRNA material in bacterial cells is known to be related to their degree of metabolic activity (11, 12). In the current work, we have performed PCR amplification of RNA extracted from enamel and dentin caries lesions, after a reverse-transcription step (13). The obtained PCR products were then pyrosequenced with the aim of characterizing the active bacterial composition of cavities.

Materials and methods

Sample collection

All donors signed a written informed consent and the sampling procedure was approved by the Ethics Committee from the DGSP-CSISP (Valencian Health Authority), with reference 10/11/2009. All donors attended the University of Santiago Dental Clinic, had not been treated with antibiotics or antifungals in the previous 6 months, had all 28 teeth present (excluding third molars) and had not suffered from any systemic disease. Clinical data are shown in Supplementary Table 1. All caries lesions sampled were active lesions as assessed by their texture and colour, and active white spot lesions were identified because they appear chalky white, opaque, and rough (14). All enamel caries collected ($n = 15$) were non-cavitated ('white spot' lesions) and were collected with sterile spoon excavators. Supragingival dental plaque samples in caries-bearing individuals were taken 24 hours after tooth-brushing from vestibular and palatine surfaces of the teeth using the same procedure (15). Teeth were not dried before sampling. Unstimulated saliva samples were collected by drooling, as previously described (15). Open dentin caries lesions ($n = 6$) were sampled directly with a sterile spoon excavator, after removing the top, biofilm layer in contact with the oral environment. All carious teeth were isolated with rubber dam to reduce the risk of saliva contamination. Unexposed dentin cavities ('hidden dentin lesions', $n = 6$) were assessed radiographically, and a water-cooled diamond bur in an air-turbine handpiece was used to drill the enamel. Water circulation was stopped right before reaching dentine tissue to minimize contamination, and the dentine lesion was hand excavated with a sterile spoon excavator. Radiographic images revealed no evidence of pulp necrosis in any dentin carious lesion sampled. Samples were eluted in 500 µl of PBS buffer and immediately frozen at $-80°C$. RNA concentrations obtained ranged from 12 to 120 µg/µl. All samples were diluted to 12 µg/µl before PCR was performed.

RNA extraction, PCR amplification and sequencing

RNA was extracted by a combination of physical and chemical lysis. Carious lesions samples were suspended in 500 µl sterile saline solution containing 0.1 and 0.5 mm glass beads, and subject to 50 Hz beating for 2 min in a Tissuelyzer II (QIAGEN, Limburg, The Netherlands) followed by 5 min on ice three times. Chemical lysis, RNA extraction, and DNAse treatment were then performed using the RNA/DNA Masterpure Extraction kit (Epicentre, Hamburg, Germany) following the manufacturer instructions with the addition of a lysozyme treatment at 37°C for 30 min.

Single-stranded cDNA was constructed with the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems, Grand Island, NY) in 20 µl reactions, with several modifications, following Benítez-Páez et al. (13). Universal bacterial primers 8F (5′-TCAGAGTTTGATCMT GGCTCAG-3′) and 788R (5′-GGCCVGGGTATCTAA TCC-3′) were used to partly amplify the 16S rRNA gene from the single-stranded cDNA in two 50 µl reactions, following the PCR and purification conditions described by Simón-Soro et al. (15). In three cases, there was enough carious material to obtain both DNA and RNA. In those three individuals, DNA was also extracted from dental plaque and drooling saliva following reference (15) and used for comparison. Purified PCR products were measured in a Modulus fluorimeter (Turner Biosystems, Madison, WI) and mixed in equimolar amounts in two pools of 14 samples, which were sequenced in 1/8 of a plate in the GS-FLX pyrosequencer (Roche, Basel, Switzerland) with Titanium-plus chemistry.

Sequence analysis

To increase accuracy in taxonomic assignment, only reads longer than 400 bp were selected. Chimeric PCR products were filtered out using the software Uchime, (16) and the reads were end-trimmed and quality filtered following Cabrera-Rubio et al. (17). Sequences were separated by the sample-specific 8-bp barcode and assigned to a genus using the Ribosomal Database Project classifier, with an 80% confidence threshold (18). Sequences were clustered at 97% nucleotide identity over 90% sequence alignment length, and rarefaction curves were obtained using the RDP pyrosequencing pipeline. An attempt was made to assign the reads to the species taxonomic level. In order to do this, a curated database was constructed with the full-length 16S sequences of all species present in the Ribosomal Database Project that belonged to the genera *Streptococcus*, *Lactobacillus*, and *Veillonella*. A BlastN (19) was then performed against this database with the reads that had been previously assigned to the above-mentioned genera and that were >500 bp. The top hit from each sequence comparison was selected if the alignment length was >500 bp, and the sequence identity $>99\%$.

Results and discussion

After end-trimming and quality filtering, a total of 132,599 sequences were obtained for 15 enamel lesions and 12 dentin lesions. Chimeras reached 1.51% of the total and were filtered out. An average of 4,911 reads were obtained per sample, with a mean length of 661 bp.

Reads were clustered at 97% sequence identity, which is the consensus threshold for bacterial species boundaries (20). These Operational Taxonomic Units (OTUs) can therefore be used to estimate the number of species in a sample. Enamel caries lesions were the least diverse, with a median of 177.7 bacterial species, whereas the estimates for open and hidden dentin cavities were 250.7 and 201.2, respectively (Fig. 1). Bacterial diversity levels varied not only between individuals but also between caries samples from the same individual, as also shown by rarefaction curves where the sequencing effort is plotted against the estimated number of species (Supplementary Fig. 1). The data suggest that white spot lesions appear to be a very restrictive niche, whereas open dentin cavities are the most diverse, even though only the inner layer of the lesion was selected for RNA extraction. Hidden dentin lesions were less diverse than deep dentin lesions from open dentin cavities, suggesting that the latter have a supply or microorganisms from the oral cavity. Open dentin lesions displayed also the highest variability, in accordance to their exposure to the salivary environment. The existence of such a high level of diversity even in the active fraction of the bacterial community confirms that the high number of organisms detected in caries lesions is not due to dead or inactive species and that dental caries is a polymicrobial disease, where multispecies microbial consortia are metabolically active in the lesions.

Active bacterial composition was significantly different between enamel and dentin cavities (Unifrac distance $p < 0.001$), suggesting that these microbial communities are tissue dependent. Streptococci, *Rothia*, *Leptotrichia*, and *Veillonella*, for instance, were at higher levels in enamel carious lesions, whereas *Lactobacillus*, *Shlegelella*, *Pseudoramibacter*, and *Atopobium* appeared to be clearly associated with dentin lesions (Fig. 2, top). There is also a high number of minority species (found at <1% proportion, indicated as 'Other' in Fig. 2) that were exclusively found in enamel lesions and a few of them were only found in dentin carious lesions. The latter included *Tannerella*, *Olsenella*, *Filifactor*, and *Treponema*. Given
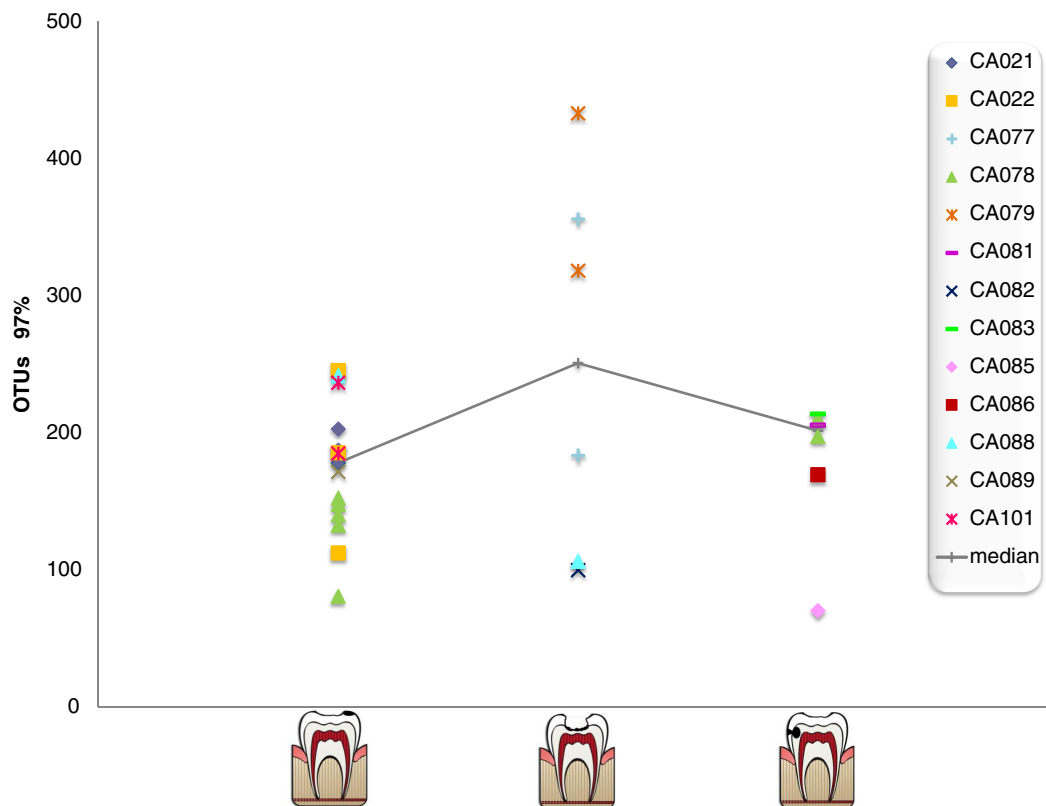

Fig. 1. Active bacterial diversity in caries lesions. Data show the estimated number of bacterial species (Operational Taxonomic Units, or OTUs), as calculated by the Chao Richness Index on the sequences clustered at 97% nucleotide identity. Patients' codes are displayed on the right panel and samples from the same individual are depicted with the same symbol. Data are shown for non-cavitated enamel caries (left), open dentin caries (middle), and unexposed, hidden caries (right). Medians from each group are shown for reference. Note that several caries samples are available for a few individuals.

Fig. 2. Taxonomic composition of active bacteria in caries samples as determined by pyrosequencing of the 16S rRNA gene. Graphs show the proportion of bacterial genera found at >1% of the total (top panel) and the proportion of different streptococcal species (lower panel), calculated as the means of all carious samples ($n = 15$ for enamel lesions, $n = 6$ for open dentin lesions, and $n = 6$ for closed dentin lesions). Data are shown for non-cavitated enamel caries (left), open dentin caries (middle), and unexposed, hidden caries (right).

the high frequency of streptococci, an effort was made to identify streptococcal sequences at the species level (Fig. 2, lower panel). *Streptococcus sanguinis* increased significantly in dentin cavities, whereas *S. mitis* was more abundant in enamel lesions. In relation to *S. mutans*, which is probably the most studied caries-associated species, a dramatically low proportion was found in all samples, ranging from 0.73% in enamel lesions to 0.48% in open dentin and 0.02% in hidden dentin lesions. The low proportion detected confirms that this species is a minority (8, 9) and questions its importance as the main etiological agent of tooth decay (21). Also, bacterial counts of lactobacilli frequently used to predict caries risk in diagnostic tests may not be informative given that they are virtually absent in enamel lesions, and this would imply that they are probably not involved in caries initiation.

There were three cases in which DNA from the same caries lesions, together with saliva and supragingival dental plaque DNA from sound teeth surfaces, could also be obtained. These pilot data show that when the DNA-based bacterial composition of a lesion is compared to the RNA-based composition of the same individual lesion, a clear difference was observed (Fig. 3, inner circles). The former, metagenomic approach, shows the total bacterial community, whereas the metatranscriptomic approach describes the active players in that community and we propose the latter procedure is therefore a closer approximation to the disease etiology. Other samples taken from the oral cavity of the same individual, like supragingival dental plaque or unstimulated saliva, show a different bacterial composition (Fig. 3, outer circles), where the detection of cariogenic bacteria is hampered by the predominance of microorganisms from other buccal niches (15). For instance, *Megasphaera* is found in saliva but passed virtually undetected in caries lesions, and *Porphyromonas* was found at considerable proportions in saliva and supragingival dental plaque, but is a minority in enamel lesions.

Conclusion

The number of active species in lesions found in the current work is lower than in previous studies based on ribosomal DNA (from 206 in white spot lesions and 379 in
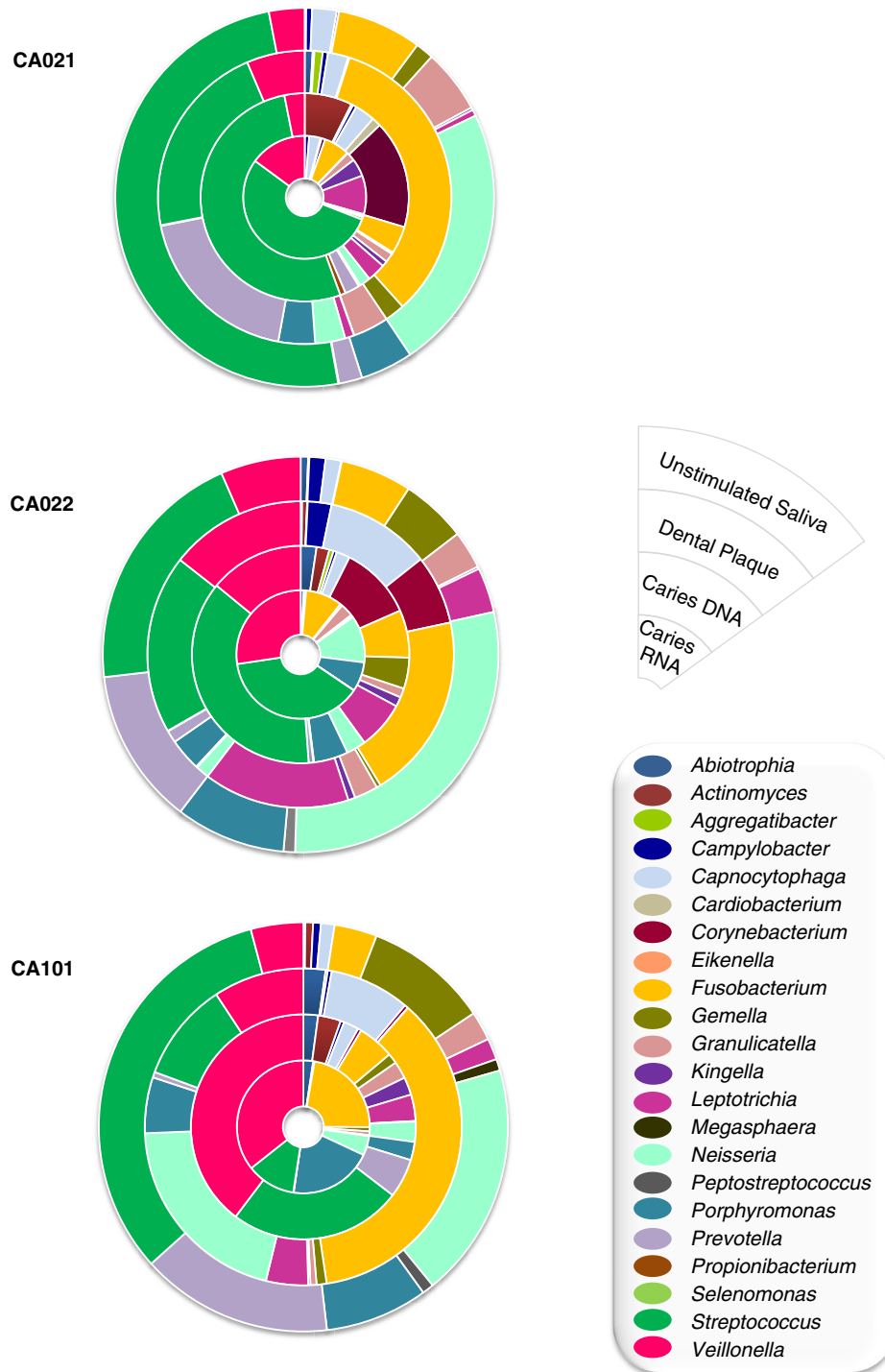
Fig. 3. Bacterial composition in different oral samples from three caries-bearing individuals. Data show the proportion of bacterial genera for unstimulated saliva (DNA-based), supragingival dental plaque of sound teeth surfaces (DNA-based), enamel caries lesions (DNA-based) and enamel caries lesions (RNA-based), from the outer to the inner circles.

dentin lesions (9) to 177 and 201 [this study], respectively). This list can be taken as a first approximation of the repertoire of microorganisms potentially involved in caries initiation and progression, as the taxonomic description of caries lesions is a vital step in determining the etiology

of the disease. Several taxonomic groups are nevertheless likely to be under-represented, such as G + C rich bacteria including *Actinomyces* and *Bifidobacterium*, which are known to be poorly amplified by 'universal' primers (22). Mutans streptococci, however, are readily amplified with

the primers used in this and other studies (21), and their low proportion is probably not a PCR artifact, as also proposed by other authors in DNA-based studies (8, 9). It has also to be borne in mind that fungal species have also been detected in caries lesions, and proposed to contribute to cariogenicity (23). Thus, the use of bacterial primers cannot detect fungal organisms, and it would be necessary to perform high-throughput sequencing analyses of fungal species to gain insights into the diversity and contribution of these microorganisms to dental caries. Nevertheless, the varying polymicrobial nature of cavities shown in the current manuscript and the currently accepted ecology-based hypothesis of caries disease (24, 25) underline that the functional output of the microbial community is probably more important than its species-composition in order to understand and combat the disease. Thus, in the future, the application of high-throughput direct sequencing to the RNA extracted from oral samples (13) will provide an opportunity to identify not only the active microbial composition but also the expressed genetic repertoire underlying disease initiation and progression. From an applied point of view, the polymicrobial etiology of dental caries underlines that diagnostic and therapeutic strategies directed at single species are likely to be unsuccessful.

Acknowledgements

Conflict of interest and funding

References

1. Aas JA, Paster BJ, Stokes LN, Olsen I, Dewhirst FE. Defining the normal bacterial flora of the oral cavity. J Clin Microbiol 2005; 43: 5721–32.
2. Loesche WJ, Rowan J, Straffon LH, Loos PJ. Association of Streptococcus mutants with human dental decay. Infect Immun 1975; 11: 1252–60.
3. Badet C, Thebaud NB. Ecology of lactobacilli in the oral cavity: a review of literature. Open Microbiol J 2008; 2: 38–48.
4. Mantzourani M, Fenlon M, Beighton D. Association between Bifidobacteriaceae and the clinical severity of root caries lesions. Oral Microbiol Immunol 2009; 24: 32–7.
5. Tanner AC, Mathney JM, Kent RL, Chalmers NI, Hughes CV, Loo CY, et al. Cultivable anaerobic microbiota of severe early childhood caries. J Clin Microbiol 2011; 49: 1464–74.
6. Aas JA, Griffen AL, Dardis SR, Lee AM, Olsen I, Dewhirst FE, et al. Bacteria of dental caries in primary and permanent teeth in children and young adults. J Clin Microbiol 2008; 46: 1407–17.
7. Belda-Ferre P, Alcaraz LD, Cabrera-Rubio R, Romero H, Simón-Soro A, Pignatelli M, et al. The oral metagenome in health and disease. ISME J 2012; 6: 46–56.
8. Gross EL, Beall CJ, Ktusch SR, Firestone ND, Leys EJ, Griffen AL. Beyond Streptococcus mutans: dental caries onset linked to multiple species by 16S rRNA community analysis. PLoS One 2012; 7: e47722.
9. Simón-Soro A, Belda-Ferre P, Cabrera-Rubio R, Alcaraz LD, Mira A. A tissue-dependent hypothesis of dental caries. Caries Res 2013; 47: 591–600.
10. Nyvad B, Crielaard W, Mira A, Takahashi N, Beighton D. Dental caries in a molecular microbiological perspective. Caries Res 2013; 47: 89–102.
11. Gentile G, Giuliano L, D'Auria G, Smedile F, Azzaro M, De Domenico M, Yakimov MM. Study of bacterial communities in Antarctic coastal waters by a combination of 16S rRNA and 16S rDNA sequencing. Environ Microbiol 2006; 8: 2150–61.
12. Amann RI, Ludwig W, Schleifer KH. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev 1995; 59: 143–69.
13. Benítez-Páez A, Belda-Ferre P, Simón-Soro A, Mira A. Microbiota diversity and gene expression dynamics in the human oral biofilm. BMC Genomics 2014; 15: 311.
14. Nyvad B, Machiulskiene V, Baelum V. Reliability of a new caries diagnostic system differentiating between active and inactive caries lesions. Caries Res 1999; 33: 252–60.
15. Simón-Soro A, Tomás I, Cabrera-Rubio R, Catalan MD, Nyvad B, Mira A. Microbial geography of the oral cavity. J Dent Res 2013; 92: 616–21.
16. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 2011; 27: 2194–200.
17. Cabrera-Rubio R, Garcia-Núñez M, Setó L, Antó JM, Moya A, Monsó E, et al. Microbiome diversity in the bronchial tracts of patients with chronic obstructive pulmonary disease. J Clin Microbiol 2012; 50: 3562–8.
18. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 2007; 73: 5261–7.
19. Gertz EM, Yu YK, Agarwala R, Schäffer AA, Altschul SF. Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST. BMC Biol 2006; 4: 41.
20. Yarza P, Richter M, Peplies J, Euzeby J, Amann R, Schleifer KH, et al. The All-Species Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. Syst Appl Microbiol 2008; 31: 241–50.
21. Munson MA, Banerjee A, Watson TF, Wade WG. Molecular analysis of the microflora associated with dental caries. J Clin Microbiol 2004; 42: 3023–9.
22. Cabrera-Rubio R, Collado MC, Laitinen K, Salminen S, Isolauri E, Mira A. The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. Am J Clin Nutr 2012; 96: 544–51.
23. Falsetta ML, Klein MI, Colonne PM, Scott-Anne K, Gregoire S, Pai CH, et al. Symbiotic relationship between Streptococcus mutans and Candida albicans synergizes virulence of plaque biofilms in vivo. Infect Immun 2014; 82: 1968–81.
24. Marsh PD. Are dental diseases examples of ecological catastrophes? Microbiology 2003; 149: 279–94.
25. Takahashi N, Nyvad B. The role of bacteria in the caries process: ecological perspectives. J Dent Res 2011; 90: 294–303.