16S rRNA gene-based metagenomic analysis identifies a novel bacterial co-prevalence pattern in dental caries

Sri Nisha Jagathrakshakan[1], Raghavendra Jayesh Sethumadhava[1], Dhaval Tushar Mehta[2], Arvind Ramanathan[2]

[1]Department of Prosthodontia, Sree Balaji Dental College and Hospital, Bharath University, Narayanapuram, Pallikaranai, Chennai, Tamil Nadu, India,
[2]Department of Human Genetics Laboratory, Central Research Facility, Sree Balaji Medical and Dental College and Hospital, Bharath University, Narayanapuram, Pallikaranai, Chennai, Tamil Nadu, India

Correspondence: Dr. Arvind Ramanathan
Email: drarvindram@yahoo.co.in

ABSTRACT

Objective: To identify the prevalence of acidogenic and nonacidogenic bacteria in patients with polycaries lesions, and to ascertain caries specific bacterial prevalence in relation to noncaries controls. **Materials and Methods:** Total genomic DNA extracted from saliva of three adults and four children from the same family were subjected to 16S rRNA gene sequencing analysis on a next generation sequencer, the PGS-Ion Torrent. Those bacterial genera with read counts > 1000 were considered as significant in each of the subject and used to associate the occurrence with caries. **Results and Conclusion:** Sequencing analysis indicated a higher prevalence of *Streptococcus, Rothia, Granulicatella, Gemella, Actinomyces, Selenomonas, Haemophilus* and *Veillonella* in the caries group relative to controls. While higher prevalence of *Streptococcus, Rothia* and *Granulicatella* were observed in all caries samples, the prevalence of others was observable in 29–57% of samples. Interestingly, *Rothia* and *Selenomonas*, which are known to occur within anaerobic environments of dentinal caries and subgingival plaque biofilms, were seen in the saliva of these caries patients. Taken together, the study has identified for the first time a unique co-prevalence pattern of bacteria in caries patients that may be explored as distinct caries specific bacterial signature to predict cariogenesis in high-risk primary and mixed dentition age groups.

Key words: *Actinomyces*, bacterial prevalence in polycaries lesion, dental caries, Gemella, *Granulicatella, Rothia, Selenomonas, Streptococcus*

INTRODUCTION

Carious lesion of the oral cavity is one of the most common disease among the primary, mixed dentition and adult age groups.[1] Cariogenesis itself is the outcome of multifactorial events that includes plaque microflora, susceptibility of host immune system, dietary habits, and most importantly maintenance of oral hygiene.[2] Plaque formation *per se* is conceived as a physiologically normal event, which otherwise is disrupted and removed by effective brushing techniques. However, when maintenance of oral hygiene suffers the bacterial flora within the plaque begins to mature with early and late colonizers,[3,4] which in the availability of carbohydrate rich substrates becomes preferentially populated by acidogenic and acid-tolerant species.[5,6]

Higher prevalence of *Streptococcus mutans* has frequently been associated with cariogenesis though not all caries patients test positive for *S. mutans*.[1,6] While *S. mutans* is known to occur at higher

prevalence, several other acidogenic species have also been reported to co-prevail within the niche areas such as supra- and sub-gingival plaque substance and dentinal caries. Some of the most commonly reported bacteria include *Lactobacillus, Granulicatella, Gemella, Haemophilus, Veillonella, Actinomyces, Rothia, Selenomonas, Atopobium, Neisseria, Eikenella, Fusobacterium, Leptotrichia*, and *Enterococcus*.[1,2,7-10] Several studies have indicated co-existence of above bacteria in various combinations, though the detection depended on the source of sample saliva, mature plaque or dentinal caries from primary, secondary or mixed dentition age groups.[11,12]

In the background of these data, we designed the present study to specifically investigate the bacterial prevalence in the patients with carious lesions from within the same family. The advantage of investigating subjects from within the same family is that, it provides a well-controlled sample collection and assay system without any bias that otherwise may arise as a result of variation in dietary intake and maintenance of oral hygiene. The subjects included three brothers and their four children with polycaries lesions, all of whom were pure vegetarians and had fairly similar knowledge on oral health maintenance. As 16S rRNA gene-based metagenomics assay provides a rapid and broad spectrum analysis platform to reliably identify the microbial flora in a given sample in a quantitative manner,[13,14] we used this technique to identify the existence of bacterial flora in the saliva of these subjects in relation to four unrelated control subjects without carious lesions.

MATERIALS AND METHODS

Patient samples and DNA extraction
Unstimulated saliva samples from seven patients – three brothers of age 52, 45 and 36 respectively, and their four children of age 22, 16, 13 and 7 respectively, all with polycaries lesions and DMFT > 5, and four non-caries healthy control subjects of age 55, 45, 22 and 16 with DMFT < 1 were analyzed in the study. The caries samples were designated as CA1 to CA7 according to the above age order for ease of identification. Ethical approval for the study was granted by the institutional ethics committee. The saliva samples were collected in 15 ml pyrogen free tubes (Tarsons, India) and stored in the freezer until being used. At the time of DNA extraction, the saliva samples were thawed to room temperature and 1 ml of the saliva was centrifuged in a fresh tube at 2000 rpm for 15 min at room temperature to pellet the bacterial cells. After discarding the supernatant, the cell pellet was processed for DNA extraction with bacterial genome extraction kit (Cat # NA2110-1KT, Sigma-Aldrich).

16S rRNA amplification and sequencing
Polymerase chain reaction (PCR) amplification of the 16S rRNA gene hypervariable region V6 was performed on 100 ng of total DNA with a pool of six degenerate forward and reverse primers, the 5′ ends of which were tagged with an adaptor sequences as per Ion Torrent-sequencing library preparation protocol to barcode each of the sample in the subsequent emulsion PCR step (second PCR). These primers are designed to efficiently amplify the V6 hypervariable region of all aerobic, anaerobic, culturable and non-culturable bacteria present in any given sample, due to selective sequence ambiguity. PCR amplification of the V6 region was performed under the following conditions: Initial denaturation at 94°C for 4 min, followed by 35 cycles of 94°C for 30 s, 58°C for 30 s, 72°C for 30 s, and a final extension at 72°C for 5 min. The PCR amplicon libraries thus generated were estimated for size using TrackIt 1 kb Plus DNA Ladder (PN 10488-085; Life Technologies) on 1.5% agarose gel, and purified subsequently using gel elution kit (Cat #GE28-9034-70, Sigma-Aldrich) to remove primer dimers and non-specific amplification products. Following purification, the eluted PCR amplicon libraries were quantified with Qubit dsDNA HS assay kit (Qubit V2, Vienna). The amount of DNA fragments per microliter of the eluted PCR amplicon library was calculated from the total concentration and average size of each amplicon library, which was then diluted to obtain a final concentration of 2.8×10^8 DNA molecules as described earlier.[15] Emulsion PCR (second PCR) was carried out to generate bar-coded libraries using the Ion Xpress Plus fragment library kit (Catalog #4471269; Life Technologies) according to instructions of the manufacturer. The mixed emulsion was amplified under the following conditions: 94°C for 6 min, followed by 40 cycles at 94°C for 30 s, 58°C for 30 s, and 72°C for 90 s and then 5 cycles at 94°C for 30 s and 68°C for 6 min. This amplification step results in clonal amplification of first PCR product that upon sequencing reveals their percent prevalence. The clonal library thus generated was sequenced on a 314 chip using the Ion Torrent Personal Genome Machine (PGM) system with Ion Sequencing 200 kit (Cat #4482006, Life Technologies) according to the supplier's instructions. The DNA sequence of each of the clone from the library was processed and filtered with PGM software to remove low homology, ambiguous and polyclonal sequences.

Sequence selection parameter

The raw sequences were processed under stringent conditions to identify high confidence target species. The following filtration criteria are used to remove low quality reads: (1) Length of target sequence <150 bp (optimal chosen length was 150–200 bp), (2) reads with more than three errors with the adapter sequences, (3) reads with more than two errors or not matching with the degenerated PCR primers, (4) sequences with ambiguous bases (long repeats of homopolymers >6 bp), and (5) reads with <95% confidence interval. In order to calculate diversity, the 16S rRNA Operational Taxonomic Units (OTUs) were clustered independently for each of the sample at 97% identity threshold using QIIME and a reference dataset from the Ribosomal Database Project.[16] OTUs were assigned, and hierarchical clustering of most common and abundant taxa was performed using UPGMA clustering (Unweighted Pair Group Method with Arithmetic mean, also known as average linkage).

RESULTS

16S rRNA gene sequencing analysis

To determine the bacterial diversity in the saliva of patients with dental caries, we investigated the total genomic DNA extracted from the saliva of three adults and four children with polycaries lesions along with four caries-free controls. The DNA samples were processed for 16S rRNA gene sequencing on the next generation sequencing platform, Ion Torrent PGM system. The target sequence for the 16S rRNA gene sequencing was limited to the V6 hypervariable region corresponding to positions 984–1047 of *Escherichia coli* as it has been shown to discriminate and unambiguously map the amplicons at the genus level.[17] Besides, the V6 region is also frequently analyzed to determine the bacterial population in environmental and medical samples.[17] The PGM system expresses the prevalence of all bacterial genera present in the analyzed sample as read counts. Accordingly, a higher read count for a specific bacterial genus implies a higher population of that particular bacterium within the sample flora. Data analysis showed read counts in the range of 39,000–1, 8,000 and 2000–3500 in caries and noncaries control samples respectively, which indicated a higher bacterial population in caries samples relative to controls [Figure 1]. Variation of values was observed in the read counts across the samples, which indeed implied that the total bacterial population varied in each of the samples. However, the obtained read counts represented true bacterial population in respective
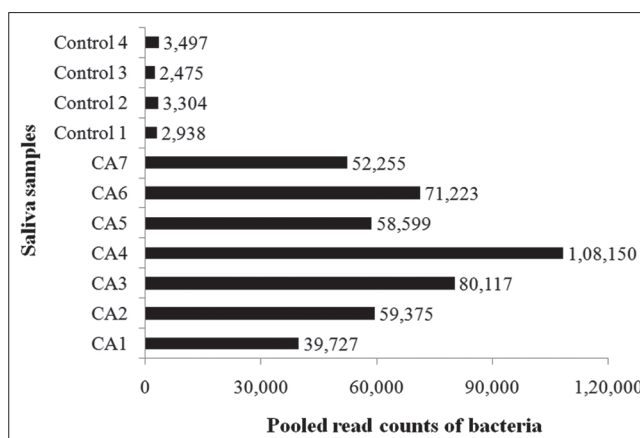

Figure 1: Comparison of the total read counts of eight bacterial genus in caries and control samples as expressed by the Personal Genome Machine system

samples as the sequence data was filtered based on two criteria: (1) the bacterial sequence should have a 97% homology with the respective reference sequence maintained at the public database (ncbi.nlm.nih.gov/nucleotide), and (2) the read count of bacterial genus should be >1000. When a bacterial genus met the above criteria, it was considered as highly prevalent in the sample. Those bacterial genera with read counts <1000 were regarded as not prevalent.

Bacterial community highly prevalent in caries samples

The dominant bacteria identified in the saliva of caries patients based on the above criteria included a total of eight bacterial genera from three different phyla, *Streptococcus*, *Granulicatella*, *Gemella*, *Veillonella*, *Selenomonas* [*Firmicutes*], *Rothia* and *Actinomyces* (Actinobacteria) and *Haemophilus* (Proteobacteria) [Table 1]. Stratification according to their respective read counts showed a higher range for all identified bacterial genera in caries patients relative to controls. *Streptococcus*, *Rothia* and *Granulicatella* (mentioned as *SRG* here after) were highly prevalent in all caries samples (CA1 to CA7, 100%). These were followed by *Gemella*, *Actinomyces* and *Selenomonas* in 57% (4/7), 43% (3/7) and 29% (2/7) of caries samples respectively [Table 2]. Samples from non-caries control group had read counts <400 for the above six bacterial genera and hence were regarded as not prevalent. The read count observed for *Streptococcus* was highest among all, which indicated this genus to be the most dominant in the salivary bacterial flora of caries patients. *Haemophilus* and *Veillonella* showed read counts of >1000 in both caries and non-caries samples. While *Haemophilus* was higher in 71% (5/7) and 75% (3/4) of caries and non-caries samples respectively, *Veillonella* was higher

in 100% (7/7) and 75% (3/4) of caries and non-caries controls respectively [Table 2].

Comparison of the co-prevalence of identified bacterial genera in caries and control groups

We next analyzed to ascertain the co-prevalence of bacteria in the caries group, as the bacterial community associated with cariogenesis and its progression is known to be complex. CA1, CA2 and CA3 had a higher prevalence of *SRG* + *Gemella*, CA4 had higher prevalence of *SRG* + *Actinomyces*, CA5 had higher prevalence of *SRG* + *Actinomyces* + *Selenomonas*, CA6 had higher prevalence of *SRG* + *Gemella* + *Actinomyces*, and CA7 had higher prevalence of *SRG* + *Selenomonas*. Though *Haemophilus* and *Veillonella* were found in both caries and control groups, their comparative prevalence was several folds higher in caries group relative to controls [Figure 2a and b], and hence was considered to be associated with cariogenesis in respective samples. This indicated that the co-prevalence of the bacteria varied among the caries samples even though the patients were from within the same family.

DISCUSSION

The present study describes the utilization of recent but well established 16S rRNA gene-based

metagenome analysis of salivary bacterial flora in two groups– patients with polycaries lesions and DMFT index >5 and non-caries control group with DMFT <1, to understand: (1) prevalence pattern of bacterial hierarchy, and (2) co-prevalence and comparative prevalence of bacterial genera among the groups. Sequence analysis identified a dominant pattern of bacteria from three phyla – *Firmicutes*, *Actinobacteria* and *Proteobacteria* [Table 1]. Since taxonomic resolution down to species-level is impossible in 16S rRNA gene-based metagenomic analysis,[15] we were not able to determine the species of bacteria. Of those identified, *SRG* showed a higher prevalence in all caries samples, while the rest were found to be associated in 29–57% of them [Table 2]. The prevalence of other bacterial genera, besides *SRG*, was similar in at least two members of the family. This selective prevalence of bacteria in only a few members of the family may be due to their independent exposures to environmental and habitual factors or competing bacterial flora within the niche areas of the oral cavity.

Several earlier studies have reported co-prevalence of *Streptococcus*, *Rothia*, *Granulicatella*, *Gemella*, *Actinomyces*, *Selenomonas*, *Haemophilus* and *Veillonella* in caries samples,[1,2,7-10] very similar to the observation made in the present study. But, what makes our study significant is that the co-prevalence of the bacteria was observed in the saliva samples of caries samples, in contrary to plaque or dentinal caries samples reported in other studies. It is important to note that *Rothia* and *Selenomonas* have been specifically identified in dentinal caries and subgingival biofilms,[12,18,19] where the environment is anaerobic. Hence, the finding of *Rothia* and *Selenomonas* in the saliva of caries samples was quite unexpected. We believe that these two bacteria may have entered the salivary flow from dislodged subgingival plaque or dentinal carious

Table 1: Bacterial genera with read counts >1000 identified in the caries and control group of samples

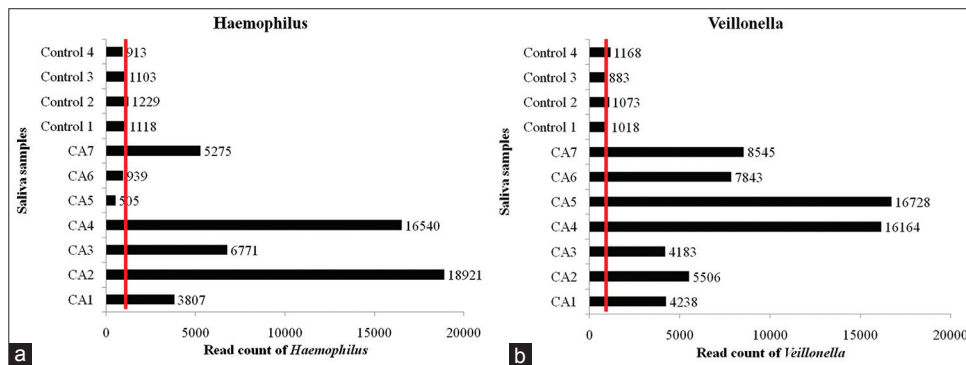| Phylum	Genus	Acidogenicity
Firmicutes	*Streptococcus*	Acidogenic
Firmicutes	*Granulicatella*	Acidogenic
Firmicutes	*Gemella*	Acidogenic
Firmicutes	*Veillonella*	Nonacidogenic
Firmicutes	*Selenomonas*	Acidogenic
Actinobacteria	*Rothia*	Acidogenic
Actinobacteria	*Actinomyces*	Acidogenic
Proteobacteria	*Haemophilus*	Acidogenic

Figure 2: Read counts of *Haemophilus* and *Veillonella* relative to controls. The read counts for both of these bacteria were several folds higher in, (a) *Haemophilus* and (b) *Veillonella* positive caries samples than the controls. Bold red line indicates the read count of 1000. Respective read counts of *Haemophilus* and *Veillonella* in individual samples are mentioned along each bar

Table 2: Actual read counts of bacterial genus expressed by PGM system for each of the caries and control samples											
Bacterial species and read counts	CA1	CA2	CA3	CA4	CA5	CA6	CA7	Control 1	Control 2	Control 3	Control 4
Streptococcus	20691	25839	45869	61294	31446	42770	23620	315	349	48	392
Rothia	5228	4792	15990	3830	3312	2746	6361	123	191	5	86
Granulicatella	3295	2606	5100	7961	1060	13793	5753	40	337	27	378
Gemella	1473	1434	1111	534	34	1212	304	73	37	2	169
Actinomyces	453	216	797	1096	2200	1808	982	224	53	362	308
Selenomonas	542	61	296	731	3314	112	1415	27	35	45	83
Haemophilus	3807	18921	6771	16540	505	939	5275	1118	1229	1103	913
Veillonella	4238	5506	4183	16164	16728	7843	8545	1018	1073	883	1168

substance, while the other six bacteria may have originated from dental biofilms, gingival crevicular fluid and/or supragingival plaque. Nevertheless, a careful in-depth analysis of the prevalence of these two bacteria in the saliva, subgingival plaque and dentinal caries substance on a larger sample size is essential to identify the extent of their occurrence in saliva, and confirm whether or not, it could be used as an indicator to predict the severity of cariogenesis process.

Of the caries specific bacteria that were identified in the present study, all except *Veillonella* are acidogenic and hence may be expected to promote cariogenesis [Table 2]. *Veillonella*, on the other hand, is an acid tolerant genus and is known to utilize lactate as its energy source. Hence, it is only logical to expect higher prevalence of *Veillonella* in locus of high lactate producers.[20] In line with this, we found higher prevalence of *Veillonella* in the caries than the control group [Figure 2], which indeed strongly correlated with the corresponding higher occurrence of acidogenic bacteria in each of them [Table 2]. Studies from other laboratories have also identified higher prevalence of *Veillonella* in caries group in association with other acidogenic genera.[12,20] Given this strong correlation between *Veillonella* and acidogenic bacteria, we believe that *Veillonella* may be an opportunistic bacterium. On the other side, since the read counts of *Veillonella* were <1000 in 75% of control samples it may also be interpreted, with caution however that *Veillonella* in controls may act to ameliorate the caries process in the absence higher acidogenic population, which otherwise was found in the caries group.[21] Careful analysis of the data also indicated read counts >500 for *Gemella, Actinomyces, Selenomonas* and *Haemophilus* in a few samples in the caries group [Table 2], which suggested the possibility that these bacteria may be just beginning to populate in order to co-inhabit the community in these samples. However, it remains to be investigated further as the biofilm and bacterial community within may be expected to be mature in established carious

lesions and hence for a new genera to compete and populate an existent niche may not be straight forward.

Since the age groups of patients analyzed in the present study were in a broad range 7–52 years, it was not possible to relate age wise distribution or prevalence of the bacterial genera. Nevertheless, the overall prevalence pattern of *SRG* was identical in all the family members, which is suggestive of a novel caries specific bacterial signature. The confidence of the above being used as a predictor of cariogenesis in high-risk group, however, depends on the outcome of similar analysis on a larger sample size. Taken together, the study has given a clear snap shot of the bacterial community at family and genera level along with their co-existent pattern in the saliva of caries patients.

CONCLUSION

The 16S rRNA gene-based metagenomics study has identified for the first time co-prevalence of the bacterial genera – *Streptococcus, Rothia, Granulicatella, Gemella, Actinomyces, Selenomonas, Haemophilus* and *Veillonella* in the saliva of patients with polycaries lesions. The finding gains significance as the co-prevalence of *Streptococcus, Rothia, Granulicatella* was observed in all caries patients, but not control subjects and hence is likely to be strongly associated with cariogenesis. The higher prevalence of other identified bacteria needs to be explored along with *Streptococcus, Rothia* and *Granulicatella* on a higher sample size to determine categorically their extent of association with cariogenesis, and determine their usage as distinct caries specific bacterial signature to predict cariogenesis in high risk group.

ACKNOWLEDGMENT

REFERENCES

1. 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.
2. Torlakovic L, Klepac-Ceraj V, Ogaard B, Cotton SL, Paster BJ, Olsen I. Microbial community succession on developing lesions on human enamel. J Oral Microbiol 2012;4:doi:10.3402/jom.v4i0.16125.
3. Rao AP, Austin RD. Serotype specific polymerase chain reaction identifies a higher prevalence of *Streptococcus mutans* serotype k and e in a random group of children with dental caries from the Southern region of India. Contemp Clin Dent 2014;5:296-301.
4. Hassan H, Lingström P, Carlén A. Plaque pH in Caries-Free and Caries-Active Young Individuals before and after Frequent Rinses with Sucrose and Urea Solution. Caries Res 2014;49:18-25.
5. Kanasi E, Dewhirst FE, Chalmers NI, Kent R Jr, Moore A, Hughes CV, *et al*. Clonal analysis of the microbiota of severe early childhood caries. Caries Res 2010;44:485-97.
6. Pannu P, Gambhir R, Sujlana A. Correlation between the salivary *Streptococcus mutans* levels and dental caries experience in adult population of Chandigarh, India. Eur J Dent 2013;7:191-5.
7. 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.
8. Munson MA, Banerjee A, Watson TF, Wade WG. Molecular analysis of the microflora associated with dental caries. J Clin Microbiol 2004;42:3023-9.
9. Chhour KL, Nadkarni MA, Byun R, Martin FE, Jacques NA, Hunter N. Molecular analysis of microbial diversity in advanced caries. J Clin Microbiol 2005;43:843-9.
10. Dammaschke T, Jung N, Harks I, Schafer E. The effect of different root canal medicaments on the elimination of *Enterococcus faecalis ex vivo*. Eur J Dent 2013;7:442-8.
11. Palmer CA, Kent R Jr, Loo CY, Hughes CV, Stutius E, Pradhan N, *et al*. Diet and caries-associated bacteria in severe early childhood caries. J Dent Res 2010;89:1224-9.
12. 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.
13. Alcaraz LD, Belda-Ferre P, Cabrera-Rubio R, Romero H, Simón-Soro A, Pignatelli M, *et al*. Identifying a healthy oral microbiome through metagenomics. Clin Microbiol Infect 2012;18 Suppl 4:54-7.
14. Gross EL, Leys EJ, Gasparovich SR, Firestone ND, Schwartzbaum JA, Janies DA, *et al*. Bacterial 16S sequence analysis of severe caries in young permanent teeth. J Clin Microbiol 2010;48:4121-8.
15. Jünemann S, Prior K, Szczepanowski R, Harks I, Ehmke B, Goesmann A, *et al*. Bacterial community shift in treated periodontitis patients revealed by ion torrent 16S rRNA gene amplicon sequencing. PLoS One 2012;7:e41606.
16. Milani C, Hevia A, Foroni E, Duranti S, Turroni F, Lugli GA, *et al*. Assessing the fecal microbiota: An optimized ion torrent 16S rRNA gene-based analysis protocol. PLoS One 2013;8:e68739.
17. Huse SM, Dethlefsen L, Huber JA, Mark Welch D, Relman DA, Sogin ML. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. PLoS Genet 2008;4:e1000255.
18. Shin JH, Shim JD, Kim HR, Sinn JB, Kook JK, Lee JN. Rothia dentocariosa septicemia without endocarditis in a neonatal infant with meconium aspiration syndrome. J Clin Microbiol 2004;42:4891-2.
19. Drescher J, Schlafer S, Schaudinn C, Riep B, Neumann K, Friedmann A, *et al*. Molecular epidemiology and spatial distribution of *Selenomonas spp*. in subgingival biofilms. Eur J Oral Sci 2010;118:466-74.
20. Gross EL, Beall CJ, Kutsch 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.
21. Arif N, Sheehy EC, Do T, Beighton D. Diversity of *Veillonella spp*. from sound and carious sites in children. J Dent Res 2008;87:278-82.

Access this article online

Quick Response Code:

Website: www.eurjdent.com

Source of Support: Internal funding from Bharath University.
Conflict of Interest: None declared