MOLECULAR MICROBIOLOGICAL ANALYSIS OF DENTAL CARIES IN THE PRIMARY AND PERMANENT DENTITIONS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Erin L. Gross, B.S.

*****

The Ohio State University
2009

Dissertation Committee:                        Approved by
Dr. Ann Griffen, Advisor

Dr. Eugene Leys

Dr. Daniel Janies

Dr. Judith Schwartzbaum                        Advisor

Oral Biology Graduate Studies
ABSTRACT

Dental caries is the most common disease of childhood, and is especially a problem for low-income and minority populations. It is a complex, polymicrobial disease and we do not currently have a complete understanding of its etiology. Culture-based studies have identified *Streptococcus mutans* and *Lactobacillus* species as important to the caries process. These species are known to ferment dietary carbohydrates, generating the lactic acid that demineralizes tooth structure. More recent studies using molecular methods have made it possible to identify oral microbiome inhabits without the use of selective culture techniques.

The goals of these studies were to use open-ended molecular techniques to identify bacteria that are associated with dental caries and health in three populations: incipient early childhood caries of infants and toddlers, severe caries in the primary dentition, and severe caries in the permanent dentition. Plaque samples were collected from the surface of intact enamel for the healthy control groups and from four types of sites in caries subjects (when present): intact enamel, white spot lesions, cavitated lesions, and dentin. 16S cloning and sequencing using universal PCR primers were used to identify the bacteria present in plaque samples. Maximum likelihood phylogenetic
trees with bootstrap cutoffs were used to define taxa. Data were analyzed using PROC MIXED.

In incipient early childhood caries, *S. mutans, S. vestibularis salivarius, S. sobrinus, S. parasanguinis~oralis* and *Veillonella atypica dispar parvula* were associated with caries. In severe early childhood caries, *S. mutans, S. vestibularis salivarius, and S. parasanguinis~oralis* were again associated with caries. Additionally, *Lactobacillus, Propionibacterium, Mitsuokella, and Parascardovia* species that were rarely seen in incipient lesions were associated with caries. In severe lesions of the permanent teeth, *Lactobacillus* species became very important. Only *Lactobacillus* and *Propionibacterium* FMA5 were significantly associated with caries in the permanent dentition. In all three studies, many health-associated species were identified, some whose role in the oral microbiome is well-established in the literature, and many others that have not been previously associated with health. In general the data supported an ecological plaque hypothesis, which states that caries is the result of an ecological shift, with caries pathogens eliminating acid-sensitive species from a healthy biofilm.

Having a complete understanding of the microbes associated with caries and health during childhood is imperative to developing interventions and treatments. Based on the typical microbial profiles seen in children, screening strategies may be developed to identify children who may be susceptible to dental caries. This screening tool could identify over-represented potential pathogens or the loss of important health-associated species.
Dedicated to Ryan, Mom, Daddy, Kari, and Devon for your constant support and unending love
ACKNOWLEDGMENTS

I would like to thank my advisors, Dr. Ann Griffen and Dr. Eugene Leys, for inspiring my interest in clinical research. They have been a source of constant support and invaluable advice on everything from dentistry and science to future career opportunities to healthy living and home repairs. I am happy to count them among my friends, and appreciate everything that they have done for me throughout my years in the lab. Without them this dissertation would not be possible.

I would also like to thank Dr. Daniel Janies for his support in learning the bioinformatics methods used in this research, and Dr. Judith Schwartzbaum for her contribution to the statistical analyses of this investigation.

I am grateful to Kashmira Asnani, Noah Firestone, and Dr. Bastienne Vriesendorp for their contributions to the completion of this project. Thank you to Dr. Stacey Kutsch, Dr. Jaclyn Martin, and Dr. Stephen Gasparovich, who collected the clinical samples. I would also like to thank Dr. Bruce Paster for the sequencing that was done in his lab.

I would like to thank all of my friends in the lab, especially Dr. Purnima Kumar and Dr. Christina Igboin, for their help and support.
This research was supported by grants DE16125 and F30DE019339 from the National Institute of Dental and Craniofacial Research.
VITA

March 15, 1982. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Born – Salem, OH

2005. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Bachelor of Science
Biology
The Ohio State University
Columbus, OH

2003 – present. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Dental Student
The Ohio State University
College of Dentistry

2005 – present. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Oral Biology
The Ohio State University

FIELDS OF STUDY

Major Field: Oral Biology
TABLE OF CONTENTS

Abstract ............................................................... ii
Dedication ............................................................ iv
Acknowledgments ..................................................... v
Vita ................................................................. vii
List of Tables ........................................................ ix
List of Figures ........................................................ x

Chapters:

1. INTRODUCTION ..................................................... 1
2. MOLECULAR MICROBIOLOGICAL ANALYSIS OF INCIPIENT EARLY
   CHILDHOOD CARIES ............................................... 17
3. MOLECULAR MICROBIOLOGICAL ANALYSIS OF SEVERE CARIES
   IN PRIMARY TEETH ................................................ 45
4. MOLECULAR MICROBIOLOGICAL ANALYSIS OF SEVERE CARIES
   IN PERMANENT TEETH ........................................... 70
5. CONCLUSIONS ....................................................... 98

References ........................................................... 104
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>34</td>
</tr>
<tr>
<td>3.1</td>
<td>60-61</td>
</tr>
<tr>
<td>4.1</td>
<td>86</td>
</tr>
</tbody>
</table>

All S-OTUs identified in incipient early childhood caries.
All S-OTUs identified in severe early childhood caries.
All S-OTUs identified in severe caries of the permanent dentition.
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Distribution of S-OTUs by phyla for all samplesType title here</td>
<td>35</td>
</tr>
<tr>
<td>2.2</td>
<td>PROC MIXED estimates</td>
<td>36</td>
</tr>
<tr>
<td>2.3</td>
<td>Mean levels for S-OTUs greater than 0.2% of total clones</td>
<td>37-38</td>
</tr>
<tr>
<td>2.4</td>
<td>Mean Shannon Diversity Indices by sample type</td>
<td>39</td>
</tr>
<tr>
<td>2.5</td>
<td>Subjects identified according to their primary acid-producers</td>
<td>40</td>
</tr>
<tr>
<td>2.6</td>
<td>Hierarchical Cluster Analysis</td>
<td>41-42</td>
</tr>
<tr>
<td>2.7</td>
<td>Levels of <em>Streptococcus mutans</em> in arrested versus progressed samples</td>
<td>43</td>
</tr>
<tr>
<td>2.8</td>
<td>Four S-OTUs were significantly different between arrested and progressed samples</td>
<td>44</td>
</tr>
<tr>
<td>3.1</td>
<td>Distribution of S-OTUs by phyla for all samplesType title here</td>
<td>62</td>
</tr>
<tr>
<td>3.2</td>
<td>PROC MIXED estimates</td>
<td>63</td>
</tr>
<tr>
<td>3.3</td>
<td>Mean levels for S-OTUs greater than 0.2% of total clones</td>
<td>64-65</td>
</tr>
<tr>
<td>3.4</td>
<td>Mean Shannon Diversity Indices by sample type</td>
<td>66</td>
</tr>
<tr>
<td>3.5</td>
<td>Subjects identified according to their primary acid-producers</td>
<td>67</td>
</tr>
<tr>
<td>3.6</td>
<td>Hierarchical Cluster Analysis</td>
<td>68-69</td>
</tr>
<tr>
<td>4.1</td>
<td>Distribution of S-OTUs by phyla for all samplesType title here</td>
<td>87</td>
</tr>
<tr>
<td>4.2</td>
<td>PROC MIXED estimates</td>
<td>88-89</td>
</tr>
</tbody>
</table>
4.3 Mean levels for S-OTUs greater than 0.2% of total clones ....... 90
4.4 Mean Shannon Diversity Indices by sample type .............. 91
4.5 Subjects identified according to their primary acid-producers .... 92
4.6 Cumulative caries scores ............................................... 93
4.7 Hierarchical Cluster Analysis ........................... 94
4.8 Hierarchical Cluster Analysis of all *Lactobacillus* species ... 95
4.9 Comparison of levels of lactobacilli and *S. mutans* ............ 96-97
CHAPTER 1

INTRODUCTION

1.1 Dental Caries

Although oral health has improved in the United States, dental caries is still the most common chronic disease of childhood, and is five times more common than asthma (104). By the time children reach the age of 17, nearly 80% have at least one cavity or filling (104). Dental caries is especially prevalent among low-income populations, including minorities and rural populations, that do not receive regular dental care (74). In fact, dental care is the most prevalent unmet health care need among America’s children (71).

The term “early childhood caries” describes caries of the primary dentition that follows a characteristic pattern: the maxillary incisors are most severely infected and the mandibular incisors are unharmed, with varying involvement of the other teeth (85). An increased risk for early childhood caries is attributed to “inappropriate” drinking behaviors, with sugar consumption leading to colonization by the mutans streptococci (105). Sugars, such as those found in fruit juices and milk, are the preferred substrate of the oral bacteria that are capable of producing the acid that demineralizes teeth (85).
Early childhood caries has serious consequences, with children needing to be treated under general anesthesia. Several studies have found that children that have been treated for early childhood caries are likely to have subsequent caries experience, despite attempts at preventative care (3, 9, 90).

Early investigations by Stephan et al (97, 98) elucidated the relationship between sugar consumption and acid production by oral microbes. He made pH measurements directly on the surfaces of 3- to 4-day-old plaque covered teeth following a glucose solution rinse, observing the resulting pH fall, then continued measurements at 10-minute intervals until the pH returned to normal. It was observed that the most drastic drop occurred in caries-active subjects, and the resulting pH was low enough to decalcify teeth. Using in vitro studies of oral bacteria, he was able to produce similar pH falls using pure cultures. It was observed that when glucose was limited, not all bacteria were able to maintain pH levels low enough to cause caries because they further metabolized the acid when glucose was no longer available. However, in the presence of excess amounts of glucose, most oral bacteria could replace acid as fast as they could consume it, and could maintain the pH levels necessary to demineralize enamel.

While recent investigations of the oral microbiome have begun to make clear the true diversity of dental plaque, historically, studies focused on two microbes that had been isolated from the oral cavity and cultured. Two groups of bacteria most consistently associated with caries are the mutans streptococci and the lactobacilli.
1.2 Oral Streptococci

Though there are many streptococci in the oral cavity that are capable of producing lactic acid from fermentable carbohydrates, the focus here will be on the mutans streptococci, specifically *Streptococcus mutans*. It has been shown that *S. mutans* is acquired early in life. Some studies have reported *S. mutans* acquisition with the eruption of the primary incisors before a child is even a year old (reviewed in (91)), while another found acquisition to occur closer to two years of age (16). *S. mutans* is often detected in children with no history of caries, which is a challenge to its role as the main caries-causing organism. One study of children in rural Sudan identified *S. mutans* in 96% of subjects, but diagnosed caries in only 15% (15). A review noted that mutans streptococci are similarly distributed in Africa, Europe, and North America, but caries experience is very different among the populations (109). Recently Saxena *et al* (89) suggested that strain differences of *S. mutans* correlate with disease status.

Despite the evidence that *S. mutans* does not cause caries in all individuals that harbor it, it has many virulence factors that make it ideally suited to cause dental caries. Perhaps most importantly, it is highly acidogenic and aciduric (23, 44), fermenting dietary carbohydrates to produce lactic acid. *S. mutans* is capable of producing enough acid to reach the critical pH between 5.0 and 5.5 necessary for the demineralization of enamel (54). It is able to withstand an acid environment because of the action of an ATPase that actively pumps protons out of the cell (82). *S. mutans* can also form sticky glucans that allow it to adhere to the tooth and form a biofilm (13).

Besides its virulence properties, evidence for the role of *S. mutans* in tooth decay can also be found in animal and human studies. A study not specific to *S. mutans* showed
that caries could occur in a germ-free hamster inoculated with streptococci from a caries-active hamster (29). Culture-based studies of human dental plaque have found higher levels of *S. mutans* in subjects with caries than those that were healthy (61, 64). *S. mutans* has been isolated from plaque associated with carious lesions and carious dentin (56). It has also been isolated from the saliva of caries-active children (113). Several longitudinal studies in children have associated *S. mutans* with progressive dental decay (11, 20, 55, 66).

### 1.3 Oral Lactobacilli

Another group of bacterial species that has been associated with dental caries historically is the lactobacilli. Unlike *S. mutans*, which is acquired early in life and is part of the normal resident flora, lactobacilli seem to be exogenous pathogens that are not associated with the teeth unless it is at a diseased site. Lactobacilli do not have good adhesion properties, so they have to have a suitable niche in order to colonize (106), such as the carious lesion. It was shown *in vitro* that lactobacilli are poor biofilm formers (27). Even though they are poorly adhesive, lactobacilli have several virulence properties, including the ability to make and tolerate lactic acid (6). Besides acid formation, lactobacilli use other antimicrobial mechanisms to give them a competitive advantage over other species present in the biofilm. Some lactobacilli produce bacteriocins and others produce antimicrobial substances that currently are not understood (99). Strains isolated from children and adolescents have been shown to inhibit the growth of *S. mutans* (94).
Lactobacilli have been shown to be associated with caries in humans, but have also been shown to cause caries in animal models. Lactobacilli were used to produce caries in gnotobiotic rats, using a strain that had been isolated from a caries-active rat (28), and a strain that was isolated from a human carious lesion (86). Many culture-based studies have confirmed the association of lactobacilli with caries. Lactobacilli have been isolated from caries-associated plaque and carious dentin in children (56). When compared to healthy sites, lactobacilli have been found at higher levels in carious lesions (61) and in caries-associated plaque (64). Longitudinal studies in children have associated lactobacilli with progressive dental decay (11, 55, 66).

1.4 Molecular Methods of Bacterial Identification

The studies referenced in the discussion of S. mutans and lactobacilli above used culture-based techniques to identify the bacteria associated with caries. Culture techniques are targeted and are limited in the species that they can identify, since bacteria have specific requirements for growth that may not be satisfied by the media being used in an experiment. Culture-independent molecular techniques have been used in many investigations of caries and health in children. Checkerboard assays have been used to determine the relative abundance of many common oral bacteria at different sites on healthy or carious teeth, and comparisons were made to determine association with health and disease (1, 7, 21, 22). PCR was used to detect S. mutans and S. sobrinus in children, with the results being compared to caries status (75). Denaturing gel gradient electrophoresis has also been used to study the microbiota associated with severe early childhood caries (50).
With the exception of Li et al (50), the molecular studies described above were targeted to species that were expected as part of the oral flora. In order to study the true diversity of sites in the oral cavity, open-ended molecular techniques must be used.

Unlike other techniques that target species bacterial species, open-ended techniques like 16S cloning and sequencing allow identification of all species present in a community. 16S cloning and sequencing were used here for phylogenetic study of caries and health in children. Ribosomal genes have regions that are necessarily highly conserved and can be used as targets for universal PCR primers that amplify everything present in a sample (37). They also have sufficient diversity to identify many oral bacteria to the species level. Based on this study design, a quantitative analysis of the bacteria associated with caries and health in children was possible.

1.5 The Ecological Plaque Hypothesis

In early studies looking for the causative agent of dental caries, specific species could not be associated with disease. This lead to the conclusion that caries was caused by a mixed bacterial infection, often called the “non-specific plaque hypothesis” (43, 54). However, in subsequent work the mutans streptococci and lactobacilli were identified as specific pathogens associated with the majority of dental decay because of their associations with disease and their acidogenic and aciduric properties (54). This new hypothesis became known as the “specific plaque hypothesis.”

The next stage in understanding the complex microbial etiology of caries was the realization that dental caries is the result of an ecological shift from a healthy, stable community to a cariogenic community (63). This has been termed the “ecological plaque
hypothesis” (63). In response to the environmental challenge of fermentable carbohydrates, acidogenic bacteria are able to gain a foothold in the community and eliminate acid-sensitive species that were in low-abundance (62). Dental caries is not caused by specific bacteria, but by any bacteria capable of contributing to an acidic environment (62). The results of the studies presented here lend support to the ecological plaque hypothesis, since many species were identified as caries-associated, and diversity decreased as health-associated species were lost.

1.6 Bioinformatics Methods

In previous microbiological studies in this laboratory, sequences were identified by comparing them to GenBank, a database that contains thousands of 16S rDNA sequences hosted by the National Center for Biotechnology Information (NCBI). Recently it has become very difficult to use GenBank, however, because of the rapidly increasing number of inaccurately or incompletely characterized sequences present. Frequently sequences compared to GenBank using the Basic Alignment Search Tool (BLAST) (4) would return results that included more than one named bacterial species, but it could not be determined which named species was truly the closest match. It was not uncommon to submit a BLAST search to GenBank and receive only uncultured and unnamed bacteria in the results. This was especially a problem for the studies presented here because each study had many thousands of bacterial clones that needed to be identified, and manually reviewing the BLAST results for each sequence, looking for satisfactory matches, was arduous work.
In order to combat this problem, a database of oral microbiome sequences was developed for use in this laboratory. This database was extensively curated using specific criteria and can be searched locally. By restricting the sequences in the database to only high-quality 16S oral bacterial sequences, searching against the database results in a match that can be accepted with a high level of confidence. Phylogenetics is the study of the evolutionary relationships of genes and organisms (41). Phylogenetic analysis was used here to compare 16S rDNA sequences from known oral bacteria, in order to define supported species-level operational taxonomic units (S-OTUs) for the oral microbiome database. Comparison of the sequences under study can then be made to the database of S-OTUs, rather than the database of species in GenBank. The phylogenetic development of this database is described in the following sections.

**Sequence Selection and Multiple Sequence Alignment**

Sequences were selected for incorporation in the database from the literature and from unpublished studies completed in this laboratory. Past studies of the oral microbiome were collected, and the accession numbers for oral bacterial sequences were identified and used to retrieve the sequences from GenBank. Additional oral bacterial species were identified from as yet unpublished studies done in this laboratory, and a representative sequence was found in GenBank to be deposited in the database.

Once these sequences were collected in the database, a multiple sequence alignment was generated for each phylum or genus with ClustalW, which is among the most widely used multiple alignment programs (47). ClustalW uses a progressive alignment strategy that aligns closely related sequences first. This works well for
sequences that are closely related (103), such as the 16S ribosomal gene. A disadvantage of this heuristic approach is that gaps and errors that are incorporated in the earlier steps of the alignment are not corrected, affecting the quality of the final alignment, especially for divergent sequences (116). However, Clustal has several advantages; it is accurate and relatively fast, and easy-to-use (18). Each multiple sequence alignment was examined in Mesquite (58). Inaccurately aligned regions were manually edited, and the 5’ and 3’ ends were trimmed to minimize the inclusion of unidentified bases (Ns).

**Phylogenetic Analysis**

The edited and trimmed multiple sequence alignment was used as the input for phylogenetic analysis. Character-based methods were chosen over distance-based methods because the information in the actual character sequence is maintained. Three commonly used character-based phylogenetic methods were considered for this purpose, parsimony, maximum likelihood with Bayesian analysis, and maximum likelihood with bootstrapping. Since many sequences were already identified by a species name, we were able to compare the three methods for their ability to reproduce groupings based on previous speciation. The taxonomic groupings generated by each method were similar, but the phylogeny suggested by maximum likelihood with bootstrapping were most consistent with existing speciation.

Parsimony searches for the simplest explanation of the data, which is the tree that requires the fewest evolutionary changes. However, parsimony does not consider back mutations. Bayesian phylogenetic analysis is based on maximum likelihood methods, but it generates both a tree and measures of support based on a posterior probability (48). It
is gaining popularity because it takes much less time than the traditional maximum likelihood method described below.

Maximum likelihood estimation is considered by many workers to be superior to distance-based methods and parsimony. Maximum likelihood was used for the development of this oral microbiome database. Swofford et al (100) reported that even though parsimony can outperform maximum likelihood in specific situations, this does not outweigh parsimony’s potential to derive an incorrect tree. Whelan et al (114) reported that in simulation studies maximum likelihood generally outperformed distance and parsimony methods. They listed the major disadvantage of distance methods being the loss of sequence data when differences are converted to distances, and of parsimony to be its inability to handle parallel and convergent evolution as sequence divergence increases. Maximum likelihood, however, maintains sequence information since it is a character-based method, and uses base substitution models that consider multiple substitutions and the probabilities of specific changes occurring (36).

The major drawback to using maximum likelihood methods is the computational time necessary to evaluate all possible trees, so RAxML, a heuristic program for maximum-likelihood based inference was used (96). RAxML optimizes a starting parsimony tree comprising all sequences—which is likely to be better than a random or distance-based tree—that undergoes a series of rearrangement steps until no better tree can be found (96). RAxML was implemented using the web server RAxML BlackBox (95). An outgroup that was closely-related to the phylum or genus of interest was chosen and the GTRGAMMA model of nucleotide substitution was used, based on a model fit test.
Viewing the Tree and Assigning Supported Operational Taxonomic Units

The best-scoring maximum likelihood tree with support values from RAxML BlackBox was imported into Dendroscope (38) to be viewed. The sequences were divided into supported species-level operational taxonomic units (S-OTUs) based on several criteria. First the topology and previous naming of species was considered. The topology often grouped sequences with more than one species name, and each sequence was considered to be of good quality. In this circumstance, all previous species names were maintained in the name of the S-OTU. Sometimes the topology suggested misnaming of specific GenBank sequences, because they clustered with species of another name. The topology was used to define clusters based on the second criterion: distance. All sequences within a cluster shared greater than 98% sequence homology. When comparing the distance between two clusters suggested by the topology, the furthest-neighbor (or complete linkage) clustering method was used. Based on furthest-neighbor, if the farthest pair of sequences from each cluster is less than 98% similar, then the clusters were divided into two S-OTUs. Using this method, S-OTUs were divided into more ranks than with other clustering methods (84). The final criterion used to define clusters was the bootstrap support values calculated in RAxML. Bootstrapping is a statistical procedure in which the data is randomly resampled and the tree is redrawn, for many iterations. The measure is of how often the group is recovered. (26). A bootstrap value of at least 70% was considered to be statistical support for the separation of the clade from adjacent sequences. Bootstrap values of 70% have been previously determined to correspond to 95% probability of occurrence (35). A combination of
topology and previous naming, distance, and bootstrap support was used to determine the divisions among S-OTUs.

In defining S-OTUs, there were instances when all three criteria of topology and previous naming, distance, and bootstrap support could not be met. There were instances when the topology of the tree and previous naming suggested several clades, but by the distance criteria and bootstrap support the clades were combined into one S-OTU, and other instances where previous naming suggested one group, but distance and bootstrap support defined separate S-OTUs. An example of the former was the important group *Streptococcus mitis pneumoniae infantis–oralis*. Although previous naming would indicate the separation of the species *S. mitis* and *S. pneumoniae*, and the topology of the *Streptococcus* tree put all of the *S. pneumoniae* sequences in a tight cluster, there was not enough distance between any two 16S sequences in the group to separate them. They could not be distinguished on the basis of similarity of their 16S sequences, so they were combined into one S-OTU. An example of the latter from the *Streptococcus* tree is the division of the *Streptococcus anginosus* sequences into two S-OTUs, A and B. Although all of the sequences shared the name *S. anginosus*, two clades were supported by bootstrap values, and the farthest pair of sequences was greater than 2% different than each other.

After the S-OTUs were defined, the database was used for a BLAST (Basic Local Alignment Search Tool) search (4) of the clinical study sequences. The BLAST program was downloaded from the NCBI and hosted locally. If an input sequence did not match the local database, then it was examined manually for the presence of unidentified bases (Ns) or chimeric features, and if the sequence was high-quality, then it was blasted
against GenBank to determine whether it was actually an oral sequence that needed to be added to the database. A high-quality sequence that was found more than once and did not match anything in the local database or GenBank was considered to be a novel S-OTU, and was added to the database as such.

### 1.7 Statistical Analyses

The studies to be presented here had a repeated measures design, because several samples were collected from a single subject. Data from the same subject is highly correlated, and the appropriate statistical approach must consider the covariance of within-subjects data (53). PROC MIXED, which was implemented in SAS 9.1, is the best approach, because it directly addresses the covariance structure by modeling it and can also handle missing data (52). These studies have missing data in the form of missing samples within subjects, and S-OTUs that are not represented within samples. PROC MIXED has been used in many repeated-measures study designs in the literature, including longitudinal studies (32, 34, 78) and studies in which the repeated measurement was something besides time (1, 87). The result of the PROC MIXED analysis was a series of estimates of the amount of increase or decrease in levels of specific S-OTUs as caries progressed and the statistical tests associated with the estimates.

For multiple independent test statistics, such as the results from the PROC MIXED analysis, the false discovery rate (FDR) has to be controlled. The false discovery rate is the expected proportion of false positive tests (those that are declared significant) among all rejected hypotheses (8). Because it controls the proportion of errors, the FDR is preferred because it has greater power than those multiple comparison
procedures that eliminate making even one error in rejecting true hypotheses (8). The FDR correction has been used most often in a variety of genetic studies, such as those studying large numbers of genes or identifying single nucleotide polymorphisms (57, 76, 92, 93). It was also used recently to identify health care providers whose performance was different than the majority (42). All FDR-adjusted p-values were calculated in SAS 9.1.

It was important to explore the composition of the bacterial communities that made up the human oral microbiome within this study, but it was desired that this description go beyond just the number of S-OTUs present. The Shannon Diversity Index is a concept that considers not only S-OTU richness, but also the relative abundance of each S-OTU in a community, and is the best measure of the diversity of a community (80). Evenness is an important part of the Shannon Diversity Index, since a more even distribution of S-OTUs increases diversity (80), just as increasing the number of S-OTUs can. The Shannon Diversity Index has been used to study environmental communities such as tree species in a human-impacted forest (40), microbial composition of water in wetlands (39), bacterial diversity in a permafrost active layer soil (51), fish communities (67), and bird communities (83). It has recently been used to study the human microbiome, including the gut microbiota (31, 112) and cultured oral microbial communities (79). It is ideally suited to study the composition of supragingival dental plaque. The formula used to compute the Shannon Diversity Index was

$$H' = -\Sigma p_i \ln p_i,$$

where \( p_i \) equals the relative abundance of each of the \( i \)th species (60).

Because we were interested in finding subsets of subjects with similar microbial profiles, TIGR Multiexperiment Viewer (MeV) was used to perform hierarchical cluster
Hierarchical cluster analysis has been used to study bacterial communities in several studies (19, 30, 59). MeV made it easy to follow the approach of Eisen et al (25), who used hierarchical cluster analysis to study gene expression data in yeast. Namely, MeV allowed us to input the data organized in a way inherent to the data, that is, S-OTUs were ordered by their PROC MIXED estimates. MeV also made it possible to represent the values in the data using a color-scale, which is intuitively easy to understand. The distance metric used for these analyses was Euclidean Distance and the linkage method was average linkage.

1.8 Outline of the Studies

The purpose of this work was to identify, using open-ended molecular techniques, the bacteria that are associated with dental caries and oral health during childhood, in both the primary and permanent dentitions. In order to identify bacterial sequences with confidence an oral microbiome database was built.

Molecular Microbiological Analysis of Early Childhood Caries

The purpose of this study was to identify the bacteria associated with the earliest stages of dental caries and health in infants and toddlers. Caries subjects had a relatively “new” infection that had yet to cause significant tooth destruction. Caries subjects were asked to return for longitudinal sampling, so that disease progression could be monitored over time.
Molecular Microbiological Analysis Of Severe Caries In Primary Teeth

The purpose of this study was to identify bacteria associated with a severe, established infection and health in the primary dentition. Subjects were older than in the previous study, and had several teeth that were badly damaged.

Molecular Microbiological Analysis Of Severe Caries In Permanent Teeth

The purpose of this study was to identify bacteria associated with severe caries and health, now in the permanent dentition. Subjects had several teeth with caries that affected many surfaces.
CHAPTER 2

MOLECULAR MICROBIOLOGICAL ANALYSIS OF INCIPIENT EARLY CHILDHOOD CARIES

2.1 Introduction

Dental caries is the most common chronic disease of childhood, and is especially a problem for children of lower socioeconomic status and minorities (74). Early childhood caries is a rampant infection of the primary dentition that progresses rapidly, and can be defined by a typical pattern of caries experience (105). It is most often associated with frequent sugar intake, either by prolonged bottle use or frequent access to drinks (24). The consequences of early childhood caries are severe, with many children needing to be put under general anesthesia for treatment of their lesions. It has been associated with early colonization by mutans streptococci (105), but other factors of its microbial etiology are not completely understood.

Many investigations have focused on the biofilm communities associated with caries of the primary dentition. Some have used culture-based techniques to identify bacterial species associated with caries (61, 66), while others have used newer molecular techniques (1, 7, 21, 22). These studies did not focus on incipient lesions, however. A
study of infants and toddlers with incipient lesions sampled saliva and used culture-based techniques to target bacteria for identification (113), while a longitudinal study of initially caries-free subjects also used saliva and culture methods (20). This study used an open-ended molecular approach to determine the bacteria that are associated with caries and healthy biofilms in incipient early childhood caries, and sampled longitudinally to track changes over time.

2.2 Materials and Methods

The purposes of this study were to identify the oral bacteria associated with the earliest stages of dental caries and health in infants and toddlers using an open-ended molecular approach, and to track microbial changes over time.

Clinical Methods

Subject recruitment

Subjects were recruited at the Nationwide Children’s Hospital Dental Clinic in Columbus, Ohio. Subjects with white spot lesions on at least two primary maxillary incisors were recruited, as well as age-, race-, and gender-matched healthy controls that were caries-free and had no existing restorations. Caries subjects were asked to return for longitudinal sampling at four to six week intervals. General exclusionary criteria included (i) age greater than 36 months, (ii) need for SBE prophylaxis, (iii) antibiotic use in the past 30 days, (iv) professional cleaning and/or fluoride application in the past 30 days, and (v) any cavitated lesions greater than 1.0 mm or existing restorations. Only one
child per family was included in each group. Informed consent was obtained from the parents of all subjects. This study was approved by the Institutional Review Board.

**Sampling and clinical data collection**

Dental plaque was collected by swiping the tooth or lesion surface with a coarse endodontic paper point. Each sample was obtained by pooling plaque collected from multiple teeth, placed in a sterile 1.5-ml microcentrifuge tube and frozen for storage. Dental plaque was sampled from the intact enamel of healthy subjects. If a subject was determined to have white spot lesions or frank cavitations, then plaque was collected separately from the surfaces of each of three types of sites available: (i) healthy, intact enamel, (ii) white spot lesions, and (iii) cavitated lesions. For healthy subjects the number of plaque samples was one, and for caries subjects the total number of samples was two or three, depending on the type of lesions that were present.

The teeth present were recorded on the tooth scoring form for healthy subjects and the caries scoring form for caries subjects. For caries subjects all carious surfaces were scored and recorded according to the worst presentation on the surface by the following codes:

- w—visible white spot lesions
- c—surface cavitation, not extending to pulp
- p—requires pulp therapy

Parents completed a brief survey regarding the history of potential caries risk factors for their children. The responses to this survey were used to determine the antibiotic and medication history, fluoride status, oral hygiene practices, and exposure to
cigarette smoke for each subject. They also provided an open-ended account of each subject’s drinking (including whether the subject was receiving breast milk) and snacking preferences, including types of drinks and snacks and how frequently they were consumed. Some caries subjects returned for follow-up visits so that caries progression could be monitored. At each subsequent visit plaque samples were collected as described above, with the addition of the cavitated lesion sample if it were now present. Of the subjects who did not return to the Baby Clinic for longitudinal sampling, the progression status for all but one subject was determined by checking the subject’s chart for further treatment received at Nationwide Children’s Hospital.

Toothbrush prophylaxis and fluoride varnish application were performed for each subject at the initial visit and at each subsequent visit. Parents received extensive oral hygiene instructions, including a demonstration of the knee-to-knee position that could facilitate oral hygiene practices at home. Parents were counseled on the potential causes of the caries experienced by their child and on the detrimental effects of frequent consumption of sugary drinks and snacks.

**Laboratory Methods**

**Sample preparation**

Bacterial DNA was isolated using a bead beater. Samples were placed in 300 µl of TE buffer and then beaten with 0.25 g of 0.1 mm glass beads for 60 seconds at 4,800 rpm in a Biospec Products bead beater. The bacterial DNA was purified using glass beads as previously described (49) and frozen until analysis.
PCR amplification for clonal analysis

The 16S rRNA genes were amplified from the purified bacterial DNA using universal primers A17 (5’-GTT TGA TCC TGG CTC AG-3’) and 317 (5’-AAG GAG GTG ATC CAG GC-3’) (Biosynthesis, Lewisville, TX). These primers encompass hypervariable regions V5-V9. PCR conditions were as previously described (45). The products of PCR amplification were examined by electrophoresis in 1% agarose and purified using the QIAquick PCR Purification Kit (QIAGEN, Valencia, CA).

Cloning and sequencing

Cloning and sequencing of the 16S amplicons generated by PCR were as previously described (45).

Bacterial 16S rDNA sequence identification

Each clone was identified by comparing it to a local, curated oral microbiome database of supported operational taxonomic units (S-OTUs) using the National Center for Biotechnology Information’s Basic Local Alignment Search Tool (BLAST). The development of this database was described in Chapter 1. Briefly, oral microbiome 16S rDNA sequences were identified for input into the database from the published literature, unpublished sequences from our lab, and GenBank. Sequences were aligned to each other using Clustal and the resulting alignment was manually examined and trimmed of its incompletely sequenced 5’ and 3’ ends. The edited multiple sequence alignment was used as the input for a maximum likelihood phylogenetic tree search. Based on the topology of the phylogenetic tree, distances between clades, and bootstrap support, S-
OTUs were identified. S-OTUs often encompass sequences from more than one previously named species, because the sequences are too closely related to be distinguishable at the 16S level. If an S-OTU contained sequences from more than one previously named species, then the name of each species was preserved in the final S-OTU designation.

**Data Management and Statistical Analysis**

Sequences were collected into a database to be analyzed for this study. Each clone was identified by its S-OTU, the type of site from which it was collected (healthy, intact enamel, white spot, and cavitated), its match number, the time at which it was collected (baseline, follow-up, follow-up 2), and whether the subject progressed to a more severe caries stage or arrested at the baseline caries experience. The match number was the same for all samples from the same subject, plus the paired healthy control. S-OTU levels were calculated as a percent of the total bacterial population for each sample. A PROC MIXED analysis was performed using baseline samples for this repeated measures dataset. The covariance structure was the default various components structure. The type of site from which each sample was collected was expressed as a level of severity in the PROC MIXED analysis on a scale from one to four, with one being the least severe and four being the most severe form of caries. Using this scale, healthy control samples were assigned a value of 1, intact enamel of caries subjects samples were assigned a value of 2, white spot lesion were assigned a value of 3, and cavitated lesion samples were assigned a value of 4. The PROC MIXED analysis used a linear model to calculate an estimate of the percent of increase or decrease of each S-OTU as caries
progressed from severity one through four, determined the significance of the estimate at \( \alpha = 0.05 \), and was followed by a False Discovery Rate correction for multiple comparisons. Mean bacterial levels and 95\% confidence intervals were determined for the most prevalent S-OTUs, and the percentage of cultivated S-OTUs was determined for each phylum present in the study. The Shannon Diversity Index was computed for each sample using the formula \( H' = -\sum p_i \ln p_i \), where \( p_i \) equals the relative abundance of each of the \( i \)th species (60), and was analyzed by PROC MIXED for baseline samples. Post hoc tests were used to elucidate differences in the Shannon Diversity Index among sample types. The primary acid-producer was identified for each white spot sample and cluster analyses were performed for each sample type to examine patterns among subjects and were expressed as heat maps. Progressed and arrested subjects were compared by chi-squared analysis for the presence of \( S. \) mutans, and by t-test for levels of \( S. \) mutans.

2.3 Results

Thirty-six caries subjects and 36 healthy controls were sampled for this study, with 21 of the caries subjects returning for a follow-up visit, and two returning for a second follow-up. The number of clones per sample ranged from 44 to 102, with the average being 53.7 clones per sample, for a total of 9,554 clones analyzed. Subjects were between twelve and 36 months old (mean age was 24 months old). A t-test showed no difference between the mean ages of the caries and healthy groups. Both the caries group and the healthy group were 20\% Hispanic and 80\% non-Hispanic. The caries groups had 21 females and 15 males, while the healthy group had 18 females and 18 males, but a chi-squared test showed that this difference was not significant. In the caries group 14
subjects identified their race as white, 13 identified as African-American, 6 were unknown, 3 indicated Asian and zero chose more than one race. In the healthy group 15 subjects identified their race was white, 14 identified as African-American, 5 were unknown, zero were Asian, and two identified as more than one race. The chi-squared test showed that these differences were not significant.

Table 2.1 lists in alphabetical order every S-OTU that was identified in this open-ended study, with percentage of total clones and its rank in order of decreasing prevalence. A total of 107 S-OTUs were identified among every sample collected, with Streptococcus mutans accounting for more than 18% of the clones. Also found at high levels were the Streptococcus mitis pneumoniae infantis–oralis group, with 14.14% of clones, and the Veillonella atypica dispar parvula group, with 11.76% of total clones. Using an open-ended approach allowed us to identify one previously unidentified phylotype, which appears to be a Capnocytophaga species. All clones represented six bacterial phyla, as seen in Figure 2.1. Two phyla, Fusobacteria and Spirochaetes, were found at such low levels that they are not visible in this figure. Overall only 2.7% of S-OTUs were uncultivated. Means for individual S-OTUs were compared between healthy control samples and intact enamel samples of caries subjects. S. mutans was found at higher levels in healthy enamel of caries subjects than that of healthy subjects. C. gingivalis–granulosa, S. intermedius constellatus–anginosus, Kingella oral taxon 499, and C. hominis were found at lower levels in healthy enamel of caries subjects than that of healthy subjects.

S-OTUs that made up greater than 0.2% of total clones were considered for the PROC MIXED analysis. Figure 2.2 displays the linear PROC MIXED estimates of the
percentage of increase or decrease in levels of each S-OTU as caries progressed from healthy controls through cavitated lesions, with 95% confidence intervals. Thirty-five of the 46 S-OTUs analyzed had negative estimates (levels decreased as caries severity increased), while nine had positive estimates (levels increased as caries severity increased). Seven of those with negative estimates were highly significant, that is they had \( p \)-values less than 0.01. Another six S-OTUs had \( p \)-values less than 0.05, including *Streptococcus mitis pneumoniae infantis–oralis*, which is the second most abundant S-OTU in the study. Among S-OTUs with positive estimates, only *Streptococcus mutans* was highly significant, but four other S-OTUs had \( p \)-values less than 0.05. Figure 2.3 displays the means for all sample types for the S-OTUs found at greater than 0.2%. It is evident that many of the S-OTUs that significantly decreased as caries progressed are found at relatively low levels in health.

The Shannon Diversity Index was computed for every baseline sample and was analyzed by PROC MIXED. The Shannon Diversity Index had a significant negative estimate \( (p<1\times10^{-6}) \), indicating that S-OTU diversity was lost as caries progressed. *Post hoc* tests were performed in JMP to determine the differences in the Shannon Diversity Indices among the stages of caries progression, and the results can be seen in Figure 2.4. The mean of the Shannon Diversity Index for the healthy control samples was significantly higher than the means for any of the caries samples, so healthy samples were more diverse relative to caries samples. There was also a significant difference between the intact enamel samples of caries subjects and cavitated samples; even within caries subjects, diversity decreases as caries severity increases.
Figure 2.5 shows percentage of subjects of the progressed and arrested groups, identified according to their primary acid-producers. The most abundant acid-producer in each baseline white spot sample was identified among those S-OTUs that had significant positive estimates, *S. mutans*, *S. vestibularis salivarius*, *S. sobrinus*, and *S. parasanguinis*–oralis. Two subjects had equal numbers of *S. vestibularis salivarius* and *S. parasanguinis*–oralis. If none of the significant acid-producers were found at greater than 2% in a sample, then known acid-producers were chosen from among the S-OTUs that were not significant, with *S. mitis pneumoniae infantis*–oralis and *L. mirabilis* being identified for some samples. Figure 2.6 is a heat map generated in the Multi-Experiment Viewer for baseline white spot samples. Using both of these figures, patterns of microbial profiles can be identified among subsets of subjects. A clear pattern of S-OTUs that were important in arresting caries progression was not elucidated, since arrested samples had the same variability in primary acid-producers as progressed samples. However, it was evident in the cluster analysis that most of the arrested samples were not a part of the cluster that had high levels of *Streptococcus mutans*. To determine whether there was a difference between arrested and progressed samples based on *S. mutans* history, the two groups were compared using the chi-squared analysis based on the presence of any clones of *S. mutans* in white spot samples at baseline (data not shown). This difference was not significant. However, a t-test comparing the levels of *S. mutans* in white spot samples at baseline was significant, with progressed samples having higher levels of *S. mutans* than arrested samples.

Data is not shown for the following analyses of potential caries risk factors. A t-test compared the number of courses of antibiotics taken throughout the subjects’
lifetimes, and indicated that antibiotic history did not predict whether a subject had caries. Chi-squared analysis showed that exposure to environmental smoking (by parents or anyone else who is regularly around the subject) did not predict the presence of the subject in either the healthy or caries group. The chi-squared analysis comparing healthy and caries subjects based on fluoride exposure was not significant.

2.4 Discussion

This is the first time this approach has been used to determine which bacteria are related to caries and health in incipient early childhood caries. Several studies have used molecular techniques to study early childhood caries (7, 21, 22), but the approach was not open-ended for all subjects, and they did not specifically seek children with the earliest stage of caries (white spot lesions). Other studies focused on incipient caries, but used saliva and culture techniques to identify bacteria (113), or they used an adult population and identified bacteria using culture techniques (88, 108, 110). Another study of early childhood caries (61) sampled subjects with severe caries and used culture techniques to identify bacteria. A longitudinal study that began sampling when 10- to 16-month-old subjects were caries free also used culture techniques to identify bacteria (66), and another longitudinal study of infants and toddlers who were initially caries-free used culture techniques to identify bacteria in saliva (20). This is the first study that has used open-ended 16S cloning and sequencing to study incipient caries in infants and toddlers and sample longitudinally, allowing us to track microbial changes over time.

Among the advantages to using an open-ended study design is the ability to detect species that have not yet been cultivated, as well as those that have never been identified.
In this study we identified 107 S-OTUs among all samples, including one previously unidentified species. These represented six bacterial phyla, and nearly 75% of the total clones were Firmicutes. It was found that less than 3% of total clones were uncultivated species. This is very different from previous oral microbiome studies done by others that found more than 50% uncultivated species (1, 2, 45). This is most likely a direct consequence of the new S-OTU designations that were used to identify sequences for this study, based on the phylogenetic analysis described in Chapter 1. In this analysis, many oral clones from GenBank were found to be present in the same clades as cultivated species, so their S-OTU names were based on the cultivated species present. Also, many oral clone sequences from GenBank were not included in our database because they were determined to be artifactual sequences.

Forty-four of the S-OTUs were analyzed by PROC MIXED (Figure 2.1). These S-OTUs were present at levels greater than 0.2% of total clones, except Total Lactobacillus, which was included because of its importance in the studies presented in Chapters 3 and 4. Thirty-five S-OTUs had negative PROC MIXED estimates, meaning they decreased as caries progressed from healthy controls through cavitated samples. Those that were significant, and should be considered potential health-associated S-OTUs were the Streptococcus mitis pneumoniae infantis–oralis group, Streptococcus sanguinis, the Neisseria flava mucosa pharyngis sicca group, Neisseria elongata, Corynebacterium durum, Kingella denitrificans, Capnocytophaga gingivalis–granulosa, Kingella oral taxon 499, Streptococcus cristatus, Granulicatella adiacens, Campylobacter rectus, Cardiobacterium hominis, and Lachnospiraceae oral taxon 107. Because the S-OTU designation combines many species names, it is difficult to compare the results for
Streptococcus mitis pneumoniae infantis–oralis to the literature. However, recent investigations have associated either S. mitis or S. oralis (or both) with health (1, 21, 22). This study agrees with many others that have associated S. sanguinis with health (7, 21, 22, 61). It has also been shown previously that the presence of S. sanguinis in the oral cavity of infants significantly delayed colonization by S. mutans (17). S. cristatus (1, 21, 22, 81) and C. gingivalis–granulosa (1) have also been associated previously with health. The role of the rest of the S-OTUs in the oral biofilm is less clear, and requires further study.

The PROC MIXED analysis also identified Streptococcus mutans, Streptococcus vestibularis salivarius, Veillonella atypica dispar parvula, Streptococcus sobrinus, and Streptococcus parasanguinis–oralis as potential caries pathogens, since these S-OTUs had a significant positive estimate. The role of S. mutans in caries biofilms has been long established (13, 54, 102, 107), and it was associated with caries in many studies of early childhood caries (7, 21, 22, 61). Culture-based studies of incipient lesions in adult populations did not separate S. mutans from the other major mutans streptococci, S. sobrinus, and found them to be associated with caries (88, 108). A targeted PCR study found that preschool children with both S. mutans and S. sobrinus had a higher incidence of caries than those with S. mutans alone (75). The mutans streptococci are known to be acidogenic, aciduric, and have good adherence and biofilm-forming properties, all of which contribute to their virulence as caries pathogens (54, 107). S. vestibularis salivarius is a group of species that have been analyzed individually by other studies, so it is difficult to compare it to results in the literature. S. vestibularis is a known acid-producer, but is not usually associated with caries (23). Historically, S. salivarius has not
been associated with caries (102), but recent molecular studies report mixed results regarding S. salivarius, which was associated with caries by Aas et al. (1) and Becker et al. (7), but with health by Corby et al. (21). Mixed results can also be found for S. parasanguinis; it has been associated with both caries (1, 7) and health (21, 22) in molecular studies. Veillonella have been associated with caries in both culture-based and molecular studies (1, 7, 61, 66). Veillonella species are lactic acid metabolizers and may facilitate the growth of S. mutans. It has been shown in vitro that combinations of S. mutans and V. alcalescens have higher acid production than monoclonies (72).

The Shannon Diversity data presented in Figure 2.4 and the results of PROC MIXED analysis indicate that there was an overall decrease in S-OTU diversity as caries progressed from healthy controls through cavitated lesions for baseline samples. This is support for the ecological plaque hypothesis, which states that environmental pressure can create an imbalance of resident species, resulting in a shift to a predominantly acid-producing community (62). Another aspect of the ecological plaque hypothesis is that a specific species is not implicated in disease (62); any species that can contribute to a decrease in pH may contribute to early childhood caries. This concept is also supported by the identification of five bacterial S-OTUs that increased as caries progressed in this study. An in vitro study showed that as the pH of mixed bacterial communities supplemented with glucose decreased, the proportions of acidogenic species increased and the proportions of acid-sensitive species decreased (12). This is directly corroborated by the data presented here. As S. mutans and others increased throughout the progression of caries, many health-associated S-OTUs were lost.
The predominant acid-producer was determined for each white spot sample and is presented in Figure 2.5. Representative subjects for each of the acid-producers identified were found in both the progressed group and the arrested group. The heat map presented in Figure 2.6 was an intuitive alternative to viewing the data for white spot lesions. A combination of the data presented in Figures 2.5 and 2.6 allowed us to find subsets of subjects that had similar microbial profiles, to determine whether there were specific profiles that contributed to an arrested or progressed caries experience. Several clusters of subjects were obvious, including those that had high levels of *S. mutans*, *S. vestibularius salivarius*, or *S. sobrinus*. Some subjects did not have a statistically significant acid-producer, and for them, other known acid-producers were identified. For some subjects *S. mitis pneumoniae infantis–oralis* or *Lautropia mirabilis* were the most abundant acid producers. Strains of *S. mitis* and *S. oralis* have been shown *in vitro* to produce acid at a high velocity and to continue producing acid as pH decreases (23). Those subjects for whom *S. mitis pneumoniae infantis–oralis* was the major acid-producer were compared to the rest of the subjects based on the levels of the specific species within the *S. mitis pneumoniae infantis–oralis* OTU (data not shown). A strain of *S. mitis* and a strain of *S. pneumoniae* were found at higher levels in these subjects. Perhaps with further study, they would be found to be highly acidogenic strains from within this typically health-associated group. *Lautropia* are known glucose-fermenters (115), which would allow them to contribute to caries pathogenesis. Several subjects also had relatively high levels of *Neisseria flava mucosa pharyngis sicca*, in addition to the species identified as their major acid contributor. An *in vitro* study of bacterial mixtures found that a mixture of *S. oralis* and *N. sicca* could produce pH levels below 3.5 (43), so
Neisseria may also be contributing to the caries environment. Even though a specific pattern could not be determined that predicted arrested subjects, it was noted that only two out of 17 subjects with S. mutans were arrested. This indicates that for this study population, S. mutans is an important predictor for future progression of caries.

The longitudinal study design presented here allowed us to explore the differences between subjects whose caries progressed to cavitated lesions and subjects who arrested at the white spot stage. To explore the role of S. mutans in caries subjects who arrested or progressed, two tests were performed. A chi-squared analysis compared the arrested and progressed subjects based on the simple presence of any S. mutans clones in white spot lesions at baseline. This test was not significant (data not shown), so the groups were compared based on the levels of S. mutans in white spot lesions at baseline. The resulting t-test was significant, with progressed subjects having higher levels of S. mutans than arrested samples (Figure 2.7). This, combined with the knowledge that S. mutans were also present in healthy subjects (this study, (15)), suggests that there is a critical level of S. mutans that must be reached before the plaque becomes cariogenic.

Progressed and arrested subjects were also compared by t-tests based on levels of individual S-OTUs. S-OTUs were compared at baseline for each stage of caries. Significant results from these analyses are presented in Figure 2.8. Four S-OTUs were significantly different between progressed and arrested samples in intact enamel and white spot samples. S. mutans was significantly higher in both intact enamel and white spots of subjects who progressed to cavitated lesions. This is not surprising since S. mutans is highly acidogenic and aciduric. However, there were three S-OTUs that were
actually higher in arrested subjects: *Neisseria flavescens, Selenomonas AA024,* and *Eikenella corrodens.* Interestingly, none of these S-OTUs were associated with health according to the PROC MIXED analysis. They should be considered potential health-associated S-OTUs since they may be contributing to the stable and non-progressive biofilm community for the arrested subjects.

The study presented here provided a unique perspective in the study of early childhood caries. Using an open-ended molecular approach allowed the identification of several health-associated species that have not been previously identified as potentially helpful organisms, and the confirmation of other studies that have implicated caries pathogens. Longitudinal sampling made possible the comparison of subjects who arrested at the white spot stage of caries, and those who progressed to more severe cavitated lesions. The differences between these populations may elucidate important factors in the caries process that can be exploited to develop interventions and therapies. Ideally, this data could be used to develop a screening strategy that could be used clinically to predict the chance of future progression to caries. Identifying different communities of bacteria that have the potential to produce caries or maintain homeostasis, then screening for over-representation or under-representation of any number of key species, could make it possible to find caries-susceptible individuals and treat them using prevention strategies. The most important aspect of this screening may be the ability to detect the loss of low-level, acid-sensitive species, whose loss may signal a shift to a more cariogenic biofilm. Much work still needs to be done to elucidate the full spectrum of important community members, especially with deeper mining of samples from a longitudinal study with longer intervals of sampling.
<table>
<thead>
<tr>
<th>Rank</th>
<th>Supported OTU</th>
<th>Clone %</th>
<th>Rank</th>
<th>Supported OTU</th>
<th>Clone %</th>
</tr>
</thead>
<tbody>
<tr>
<td>91</td>
<td>Abiotrophia defectiva</td>
<td>0.02</td>
<td>96</td>
<td>Mogibacterium diversum neglectum</td>
<td>0.01</td>
</tr>
<tr>
<td>55</td>
<td>Achromobacter xylosoxidans</td>
<td>0.10</td>
<td>97</td>
<td>Moraxella clone AM420053</td>
<td>0.01</td>
</tr>
<tr>
<td>71</td>
<td>Actinobaculum sp A11</td>
<td>0.04</td>
<td>47</td>
<td>Neisseria AP065</td>
<td>0.15</td>
</tr>
<tr>
<td>58</td>
<td>Actinomycetes georgiae</td>
<td>0.09</td>
<td>20</td>
<td>Neisseria elongata</td>
<td>0.78</td>
</tr>
<tr>
<td>33</td>
<td>Actinomycetes gerencseriae</td>
<td>0.28</td>
<td>5</td>
<td>Neisseria flavia mucosa pharyngis sicca</td>
<td>5.41</td>
</tr>
<tr>
<td>38</td>
<td>Actinomycetes massiliensis</td>
<td>0.22</td>
<td>21</td>
<td>Neisseria flavescens</td>
<td>0.78</td>
</tr>
<tr>
<td>39</td>
<td>Actinomycetes odontolyticus linguae</td>
<td>0.22</td>
<td>50</td>
<td>Neisseria meningitidis-polysaccharae</td>
<td>0.13</td>
</tr>
<tr>
<td>9</td>
<td>Actinomycetes viscosus naeslundii</td>
<td>2.77</td>
<td>67</td>
<td>NOVEL Capnocytophaga SK117 1 E 052</td>
<td>0.07</td>
</tr>
<tr>
<td>62</td>
<td>Atopobium parvulum</td>
<td>0.07</td>
<td>98</td>
<td>Obobacterium sinus</td>
<td>0.01</td>
</tr>
<tr>
<td>78</td>
<td>Bifidobacterium dentium</td>
<td>0.03</td>
<td>85</td>
<td>Obobacterium sp A0068</td>
<td>0.03</td>
</tr>
<tr>
<td>40</td>
<td>Campylobacter concisus</td>
<td>0.22</td>
<td>77</td>
<td>Parvimonas micra</td>
<td>0.04</td>
</tr>
<tr>
<td>23</td>
<td>Campylobacter gracilis</td>
<td>0.72</td>
<td>51</td>
<td>Parvimonas catoniae</td>
<td>0.13</td>
</tr>
<tr>
<td>37</td>
<td>Campylobacter rectus</td>
<td>0.24</td>
<td>86</td>
<td>Parvimonas CW034</td>
<td>0.03</td>
</tr>
<tr>
<td>92</td>
<td>Capnocytophaga DS022</td>
<td>0.01</td>
<td>88</td>
<td>Preotella olyron</td>
<td>0.01</td>
</tr>
<tr>
<td>19</td>
<td>Capnocytophaga gingivalis-granulosa</td>
<td>0.82</td>
<td>101</td>
<td>Preotella veroralis melainogenica</td>
<td>0.01</td>
</tr>
<tr>
<td>72</td>
<td>Capnocytophaga leadbetteri</td>
<td>0.04</td>
<td>103</td>
<td>Preotella veroralis melainogenica</td>
<td>0.01</td>
</tr>
<tr>
<td>93</td>
<td>Capnocytophaga sp ChDC OS44</td>
<td>0.01</td>
<td>87</td>
<td>Propionibacterium BM085</td>
<td>0.03</td>
</tr>
<tr>
<td>59</td>
<td>Capnocytophaga sp</td>
<td>0.09</td>
<td>36</td>
<td>Propionibacterium propionicus A</td>
<td>0.25</td>
</tr>
<tr>
<td>35</td>
<td>Campylobacter hominis</td>
<td>0.25</td>
<td>102</td>
<td>Raistonia insidiosa</td>
<td>0.01</td>
</tr>
<tr>
<td>63</td>
<td>Cardobacterium variabilis</td>
<td>0.07</td>
<td>32</td>
<td>Rothia aerea</td>
<td>0.33</td>
</tr>
<tr>
<td>64</td>
<td>Catonella morbi</td>
<td>0.07</td>
<td>68</td>
<td>Rothia dentocariosa</td>
<td>1.05</td>
</tr>
<tr>
<td>65</td>
<td>Centipeda periodontii</td>
<td>0.07</td>
<td>88</td>
<td>Rothia mucilaginosa</td>
<td>0.07</td>
</tr>
<tr>
<td>80</td>
<td>Clostridiales oral taxon 075</td>
<td>0.07</td>
<td>42</td>
<td>Selenomonas infelix</td>
<td>0.22</td>
</tr>
<tr>
<td>17</td>
<td>Corynebacterium durum</td>
<td>0.94</td>
<td>44</td>
<td>Selenomonas AA024</td>
<td>0.42</td>
</tr>
<tr>
<td>12</td>
<td>Corynebacterium matruchotii</td>
<td>1.42</td>
<td>48</td>
<td>Selenomonas AJ036</td>
<td>0.03</td>
</tr>
<tr>
<td>79</td>
<td>Dialister invisus</td>
<td>0.03</td>
<td>57</td>
<td>Selenomonas dianae</td>
<td>0.10</td>
</tr>
<tr>
<td>31</td>
<td>Efukella correonella</td>
<td>0.33</td>
<td>89</td>
<td>Selenomonas DO042</td>
<td>0.03</td>
</tr>
<tr>
<td>73</td>
<td>Enterococcus casseliflavus</td>
<td>0.04</td>
<td>103</td>
<td>Selenomonas EW051a</td>
<td>0.01</td>
</tr>
<tr>
<td>66</td>
<td>Escherichia coli</td>
<td>0.07</td>
<td>61</td>
<td>Selenomonas EW051a</td>
<td>0.01</td>
</tr>
<tr>
<td>56</td>
<td>Eubacterium BE088</td>
<td>0.01</td>
<td>53</td>
<td>Selenomonas lueggei</td>
<td>0.12</td>
</tr>
<tr>
<td>64</td>
<td>Eubacterium brachy</td>
<td>0.01</td>
<td>48</td>
<td>Selenomonas F1050</td>
<td>0.15</td>
</tr>
<tr>
<td>94</td>
<td>Eubacterium DO008</td>
<td>0.15</td>
<td>42</td>
<td>Selenomonas infelix</td>
<td>0.22</td>
</tr>
<tr>
<td>27</td>
<td>Eubacterium IR009</td>
<td>0.45</td>
<td>24</td>
<td>Selenomonas noxia</td>
<td>0.66</td>
</tr>
<tr>
<td>80</td>
<td>Eubacterium sp CH017</td>
<td>0.03</td>
<td>44</td>
<td>Selenomonas spotigena</td>
<td>0.18</td>
</tr>
<tr>
<td>60</td>
<td>Eubacterium sp E1074</td>
<td>0.09</td>
<td>104</td>
<td>Staphylococcus hominis</td>
<td>0.01</td>
</tr>
<tr>
<td>74</td>
<td>Eubacterium yurii subsp margaretiae yurii</td>
<td>0.04</td>
<td>90</td>
<td>Streptococcus anginosus A</td>
<td>0.03</td>
</tr>
<tr>
<td>52</td>
<td>Eubacterium yurii subsp schlitika</td>
<td>0.12</td>
<td>105</td>
<td>Streptococcus anginosus B</td>
<td>0.01</td>
</tr>
<tr>
<td>81</td>
<td>Fusobacterium nucleatum</td>
<td>0.03</td>
<td>22</td>
<td>Streptococcus cristatus</td>
<td>0.76</td>
</tr>
<tr>
<td>15</td>
<td>Gemella haemolytica</td>
<td>1.08</td>
<td>30</td>
<td>Streptococcus downei</td>
<td>0.37</td>
</tr>
<tr>
<td>28</td>
<td>Gemella morbillorum</td>
<td>0.42</td>
<td>7</td>
<td>Streptococcus gordonii</td>
<td>4.02</td>
</tr>
<tr>
<td>75</td>
<td>Gemella sanguinis</td>
<td>0.04</td>
<td>25</td>
<td>Streptococcus intermedius constellatus-anginosus</td>
<td>0.64</td>
</tr>
<tr>
<td>14</td>
<td>Granulicatella adiacens</td>
<td>1.26</td>
<td>2</td>
<td>Streptococcus mitis pneumoniae infantis-oralis</td>
<td>14.14</td>
</tr>
<tr>
<td>34</td>
<td>Granulicatella elegans</td>
<td>0.27</td>
<td>30</td>
<td>Streptococcus mitis pneumoniae infantis-oralis</td>
<td>14.14</td>
</tr>
<tr>
<td>82</td>
<td>Haemophilus parainfluenzae group A</td>
<td>0.03</td>
<td>7</td>
<td>Streptococcus mitis pneumoniae infantis-oralis</td>
<td>14.14</td>
</tr>
<tr>
<td>13</td>
<td>Kingella denitrificans</td>
<td>1.33</td>
<td>1</td>
<td>Streptococcus mutans</td>
<td>18.64</td>
</tr>
<tr>
<td>26</td>
<td>Kingella oral taxon 499</td>
<td>0.51</td>
<td>106</td>
<td>Streptococcus oligofermentans sinensis</td>
<td>0.01</td>
</tr>
<tr>
<td>18</td>
<td>Kingella oralis</td>
<td>0.91</td>
<td>11</td>
<td>Streptococcus parasanginis-oralis</td>
<td>1.51</td>
</tr>
<tr>
<td>41</td>
<td>Lachnospiraceae Oral Taxon 107</td>
<td>0.22</td>
<td>54</td>
<td>Streptococcus peroni</td>
<td>0.12</td>
</tr>
<tr>
<td>69</td>
<td>Lachnospiraceae sp C1</td>
<td>0.06</td>
<td>4</td>
<td>Streptococcus sanguinis</td>
<td>6.53</td>
</tr>
<tr>
<td>95</td>
<td>Lactobacillus fermentum</td>
<td>0.01</td>
<td>8</td>
<td>Streptococcus sobrinus</td>
<td>2.99</td>
</tr>
<tr>
<td>76</td>
<td>Lactobacillus gasseri johnsonii</td>
<td>0.04</td>
<td>6</td>
<td>Streptococcus vestibularis salivarus</td>
<td>4.92</td>
</tr>
<tr>
<td>83</td>
<td>Lactobacillus vaginalis</td>
<td>0.03</td>
<td>107</td>
<td>Treponema socranskii subsp buccale paredis socranskii</td>
<td>0.01</td>
</tr>
<tr>
<td>70</td>
<td>Lautropia AP009</td>
<td>0.06</td>
<td>3</td>
<td>Veillonella atypica dispar parvula</td>
<td>11.76</td>
</tr>
<tr>
<td>10</td>
<td>Lautropia mirabilis</td>
<td>2.39</td>
<td>45</td>
<td>Veillonella sp oral clone HB016</td>
<td>0.18</td>
</tr>
<tr>
<td>84</td>
<td>Megasphaera micronuconiformis</td>
<td>0.03</td>
<td>45</td>
<td>Veillonella sp oral clone HB016</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Table 2.1. All S-OTUs identified in incipient early childhood caries, with percentage of total clones. Rank puts them in order of decreasing overall prevalence.
Figure 2.1. Distribution of S-OTUs by phyla for all samples.
Figure 2.2. PROC MIXED estimates of percentage of increase or decrease in bacterial levels as caries progresses, with 95% confidence intervals. *P*-values were corrected by the False Discovery Rate. * p < 0.05. ** p < 0.01.
Figure 2.3. Mean levels for S-OTUs greater than 0.2% of total clones, with 95% confidence intervals. From left to right within an S-OTU, bars indicate levels for healthy control, intact enamel, white spot, cavitated, and dentin samples. S-OTUs that are green have significant negative estimates, and S-OTUs that are red have significant positive estimates. The upper limit of the 95% confidence interval for the cavitated sample of S. mutans, which isn’t shown, is 59.931.
Figure 2.3

- Campylobacter gracilis
- Selenomonas noxia
- Streptococcus intermedius constellatus-anginosus
- Kingella oral taxon 499
- Eubacterium IR009
- Gemella morbillorum
- Selenomonas AA024
- Streptococcus downei
- Eikenella corrodens
- Rothia aeria
- Actinomyces gerencseriae
- Granulicatella elegans
- Cardiobacterium hominis
- Propionibacterium propionicus A
- Campylobacter rectus
- Actinomyces massiliensis
- Actinomyces odontolyticus linguae
- Campylobacter concisus
- Lachnospiraceae Oral Taxon 107
- Selenomonas infelix
- Mitsuokella oral taxon 521
- Total Lactobacillus

- Streptococcus mutans
- Streptococcus mitis pneumoniae infantis-oralis
- Veillonella atypica dispar parvula
- Streptococcus sanguinis
- Neisseria flava mucosa pharyngis sicca
- Streptococcus vestibularis salivarius
- Streptococcus gordonii
- Streptococcus sobrinus
- Actinomyces viscosus naeslundii
- Lautropia mirabilis
- Streptococcus parasanguinis-oralis
- Corynebacterium matruchotii
- Kingella denitrificans
- Granulicatella adiacens
- Gemella haemolysans
- Rothia dentocariosa
- Corynebacterium durum
- Kingella oralis
- Capnocytophaga gingivalis-granulosa
- Neisseria elongata
- Neisseria flavescens
- Streptococcus cristatus
Figure 2.4. Mean Shannon Diversity Indices by sample type, with 95% confidence intervals. All paired post hoc tests showed significance with the exception of the comparisons between intact enamel and white spot and between white spot and cavitation. The dashed line represents the significant PROC MIXED estimate, which had a value of -0.25.
Figure 2.5. Subjects identified according to their primary acid-producers. The most abundant acid-producer in each baseline white spot sample was identified among those S-OTUs that had significant positive estimates, *S. mutans, S. vestibularis salivarius, S. sobrinus,* and *S. parasanguinis–oralis.* Two subjects (“Mix” above) had equal numbers of *S. vestibularis salivarius* and *S. parasanguinis–oralis.* If none of the significant acid-producers were found at greater than 2% in a sample, then known acid-producers were chosen from among the S-OTUs that were not significant, with *S. mitis pneumoniae infantis–oralis* and *L. mirabilis* being identified for some samples.
Figure 2.6. Hierarchical cluster analysis of white spot samples. S-OTUs that were the primary acid producers in at least one subject are in decreasing order of their PROC MIXED estimates. The color of each cell represents the percent of total clones within that sample for the S-OTU of interest, according to the scale above. Samples that are marked by an asterisk were arrested, and the remaining samples are progressed (except for sample 36, which is unknown because the subject did not return for follow-up).
Figure 2.6
Figure 2.7. Levels of *Streptococcus mutans* in arrested versus progressed samples in baseline white spot samples. The difference is significant ($p = 0.0401$). The mean of each group is represented by the horizontal bar through the middle of each diamond. The upper and lower points of the diamonds represent the 95% confidence intervals.
Figure 2.8. Four S-OTUs were significantly different between arrested and progressed samples when intact enamel and white spot samples at baseline were compared.
CHAPTER 3

MOLECULAR MICROBIOLOGICAL ANALYSIS OF SEVERE CARIES IN PRIMARY TEETH

3.1 Introduction

Dental caries is the most common chronic disease of childhood, and is especially a problem for children of lower socioeconomic status and minorities (74). Early childhood caries is a rampant infection of the primary dentition that progresses rapidly, and can be defined by a typical pattern of caries experience (105). It is most often associated with frequent sugar intake, either by prolonged bottle use or frequent access to drinks (24). The consequences of early childhood caries are severe, with many children needing to be put under general anesthesia for treatment of their lesions. Despite complete restorative treatment and increased preventive measures, children with early childhood caries remain susceptible to future caries experience (3, 90). It has been associated with early colonization by mutans streptococci (105), but other factors of its microbial etiology are not completely understood. In order to develop the most effective treatments for early childhood caries, a thorough understanding of its progression is
imperative, including maximum understanding of the biofilm associated with healthy primary teeth.

Many investigators have studied early childhood caries, but most have not used the approach that will be presented here. Both culture-based (11, 61, 66) and molecular methods (1, 7, 21, 22, 75) have been used to study primary dentition biofilms, and all agreed on the association of S. mutans with early childhood caries. However, since most of these studies used targeted methods for bacterial identification, they were unable to identify new potential pathogens, and did not focus on health-associated bacterial communities (except (1)). In order to more deeply study the oral microbiome and its association with caries and healthy biofilms in early childhood caries, we used an open-ended molecular approach.

3.2 Materials and Methods

The purpose of this study was to identify the oral bacteria associated with severe dental caries and health in the primary dentition using an open-ended molecular approach.

Clinical Methods

Subject recruitment

Subjects were recruited at the Nationwide Children’s Hospital Dental Clinic in Columbus, Ohio. Subjects with severe caries were recruited who satisfied the inclusion requirement of the presence of at least four teeth with smooth surface caries, at least two of which involved the dental pulp. Age-, race-, and gender-matched healthy controls that
were caries-free and had no existing restorations were also recruited. General exclusionary criteria included (i) age greater than six years, (ii) need for SBE prophylaxis, (iii) antibiotic use in the past 30 days, (iv) professional cleaning in the past 30 days, and (v) presence of permanent teeth. Only one child per family was included in each group. Consent was obtained from the parents of all subjects. This study was approved by the Institutional Review Board.

**Sampling and clinical data collection**

Dental plaque was collected by swiping the tooth or lesion surface with a dental explorer and wiping the plaque onto a coarse endodontic paper point. Each sample was obtained by pooling plaque collected from multiple teeth, placed in a sterile 1.5-ml microcentrifuge tube and frozen for storage. Dental plaque was sampled from the intact enamel of healthy subjects. If a subject was determined to have dental caries, then plaque was collected separately from the surfaces of each of three types of sites available: (i) healthy, intact enamel, (ii) white spot lesions, and (iii) cavitated lesions. Carious dentin was sampled from a tooth that required pulp therapy. For healthy subjects the number of plaque samples was one, and for caries subjects the total number of samples was four.

The teeth present were recorded on the tooth scoring form for healthy subjects and the caries scoring form for caries subjects. For caries subjects existing restorations were recorded and all carious surfaces were scored and recorded according to the worst presentation on the surface by the following codes:

- w—visible white spot lesions
- c—surface cavitation, not extending to pulp
 Parents completed a brief survey regarding the history of potential caries risk factors for their children. The answers to this survey were used to determine the antibiotic and medication history, fluoride status, oral hygiene practices, and exposure to cigarette smoke for each subject. They also provided an open-ended account of each subject’s drinking and snacking preferences, including types of drinks and snacks and how frequently they were consumed.

**Laboratory Methods**

**Sample preparation**

Bacterial DNA was isolated using a bead beater. Samples were placed in 300 µl of TE buffer and then beaten with 0.25 g of 0.1 mm glass beads for 60 seconds at 4,800 rpm in a Biospec Products bead beater. The bacterial DNA was purified using glass beads as previously described (49) and frozen until analysis.

**PCR amplification for clonal analysis**

The 16S rRNA genes were amplified from the purified bacterial DNA using universal primers A17 (5’-GTT TGA TCC TGG CTC AG-3’) and 317 (5’-AAG GAG GTG ATC CAG GC-3’) (Biosynthesis, Lewisville, TX). PCR conditions were as previously described (45). The products of PCR amplification were examined by electrophoresis in 1% agarose and purified using the QIAquick PCR Purification Kit (QIAGEN, Valencia, CA).
**Cloning and sequencing**

Cloning of the 16S amplicons generated by PCR was as previously described (45). Approximately half of the total clones for this study were sent to the Forsyth Institute in Boston, MA to be sequenced according to the methods used in Paster *et al* (77), while the other half were sequenced at the Plant-Microbe Genomics Facility at The Ohio State University as described by Kumar *et al* (45). Clones sequenced at the Forsyth Institute included hypervariable regions V1-V3, while those sequenced at The Ohio State University included hypervariable regions V5-V9.

**Bacterial 16S rDNA sequence identification**

Each clone was identified by comparing it to a local, curated oral microbiome database of supported operational taxonomic units (S-OTUs) using the National Center for Biotechnology Information’s Basic Local Alignment Search Tool (BLAST). The development of this database was described in Chapter 1. Briefly, oral microbiome 16S rDNA sequences were identified for input into the database from the published literature, unpublished sequences from our lab, and GenBank. Sequences were aligned to each other using Clustal and the resulting alignment was manually examined and trimmed of its incompletely sequenced 5’ and 3’ ends. The edited multiple sequence alignment was used as the input for a maximum likelihood phylogenetic tree search. Based on the topology of the phylogenetic tree, distances among clades, and bootstrap support, supported operational taxonomic units (S-OTUs) were identified. S-OTUs often encompass sequences from more than one previously named species, because the
sequences are too closely related to be distinguishable at the 16S level. If an S-OTU contained sequences from more than one previously named species, then the name of each species was preserved in the final S-OTU designation.

Data Management and Statistical Analyses

A database of total clones was collected to be analyzed for this study. Each clone was identified by its S-OTU, the type of site from which it was collected (healthy, intact enamel, white spot, cavitated, and dentin) and its match number. The match number was the same for all samples from the same subject, plus the paired health control. S-OTU levels were calculated as a percent of the total bacterial population for each sample. A PROC MIXED analysis was performed for this repeated measures dataset in SAS 9.1. The covariance structure was the default various components structure. The type of site from which each sample was collected was expressed as a level of severity in the PROC MIXED analysis on a scale from 1 to 5, with 1 being the least severe and 5 being the most severe form of caries. Using this scale, healthy control samples were assigned a value of 1, intact enamel samples were assigned a value of 2, white spot lesion samples were assigned a value of 3, cavitated lesion samples were assigned a value of 4, and dentin samples were assigned a value of 5. The PROC MIXED analysis used a linear model to calculate an estimate of the percent of increase or decrease of each S-OTU as caries progressed from severity one through five, and determined the significance of the estimate at $\alpha = 0.05$, and was followed by a False Discovery Rate correction for multiple tests. Mean bacterial levels and 95% confidence intervals were determined for the most prevalent S-OTUs, and the percentage of cultivated S-OTUs was determined for each
phylum present in the study. T-tests were used to compare means of healthy control samples and intact enamel samples in caries subjects, followed by a False Discovery Rate correction. The Shannon Diversity Index was computed for each sample using the formula $H' = -\Sigma p_i \ln p_i$, where $p_i$ equals the relative abundance of each of the $i$th species (60), and was analyzed by PROC MIXED. T-tests and paired t-tests were computed in JMP and used as post hoc tests. The primary acid-producer was identified for each sample, and hierarchichal cluster analyses were performed for each sample type using the TM4 MultiExperiment Viewer to examine patterns among subjects and were expressed as heat maps. The analyses were done using the Euclidean Distance metric and Average Linkage clustering.

3.3 Results

Thirty-one caries subjects and 29 healthy controls were sampled for this study. Between 22 and 66 clones were collected per sample type, with the average being 47 clones per sample, for a total of 6,985 clones sequenced. In the caries group 15 females and 16 males were sampled, while the healthy control group contained 15 females and 14 males, a difference that was not significant by chi-squared analysis. The caries group had 22 subjects who identified their race as white, seven that identified as African-American, and two that reported more than one race. The healthy control group contained 21 subjects who identified themselves as white, seven that identified as African-American, and one that reported more than one race. The groups appear to be equal with regards to race based on chi-squared analysis. Every subject identified as non-Hispanic. Subjects’
ages ranged from 2 to 6 years old (mean age was 4.1), and the mean ages of the caries
and healthy groups were not significantly different by a t-test.

A total of 130 S-OTUs were identified among all of the samples in the study, and
are presented in Table 3.1 in alphabetical order with percentages of total clones and rank
in order of decreasing prevalence. The most abundant S-OTU was *Streptococcus mutans*,
with more than 20% of the total clones. Other major S-OTUs included *Veillonella
atypica dispar parvula*, and the *Streptococcus mitis pneumoniae infantis–oralis* group.
Six phyla are represented by the clones and are presented in Figure 3.1. *Spirochaetes* and
Fusobacteria are found at levels too low to be seen on the graph. Only 4.2% of the total
clones are uncultivated.

Fifty S-OTUs with levels at least 0.2% of total clones were analyzed by PROC
MIXED. All Lactobacillus S-OTUs were combined into the category “Total
Lactobacillus” for the PROC MIXED analysis. Figure 3.2 contains the estimates from
PROC MIXED, with 95% confidence intervals. Thirty-three S-OTUs had negative
estimates and decreased as caries progressed. Seventeen S-OTUs had positive estimates,
increasing as caries progressed. Many health- and caries-associated candidates were
identified by those that had significant estimates after the False Discovery Rate
adjustment was applied to the results of the analysis. *Streptococcus mutans*, Total
*Lactobacillus*, *Streptococcus parasanguinis–oralis*, and *Propionibacterium* FMA5 had
highly significant positive estimates (p<0.01), while several additional S-OTUs with
positive estimates had p-values less than 0.05. Among those S-OTUs with negative
estimates, most were significant according to the PROC MIXED analysis. The
significant negative S-OTUs represented several genera, including *Propionibacterium,
*Eubacterium, Corynebacterium, Actinomyces, Kingella, Eikenella, Campylobacter, Neisseria, Selenomonas, Abiotrophia, Capnocytophaga, Gemella,* and *Streptococcus.* Two S-OTUs with highly significant negative estimates ($p<0.01$) that also occurred at relatively high levels in the study population were *S. mitis pneumoniae infantis-*oralis and *Streptococcus sanguinis.* Figure 3.3 depicts the relative abundance of all S-OTUs greater than 0.2% of total clones, with colors indicating caries or health association. It is clear from this graph that many health-associated candidates are found at low levels in healthy subjects. We can also see that healthy subjects and healthy sites in a caries mouth contained *Streptococcus mutans,* while other disease-associated S-OTUs like *Lactobacillus, Propionibacterium FMA5,* and *Parascardovia denticolens* were only found at caries sites.

The Shannon Diversity Index was computed for each sample and was analyzed by PROC MIXED. The resulting negative estimate was significant ($p<1\times10^{-8}$), and post hoc tests found the differences among sample types. Figure 3.4 shows that the means for both healthy sites were significantly higher than those for caries sites. Overall, bacterial diversity decreased as caries progressed. The PROC MIXED estimate for the linear decrease as caries progressed is indicated by the dashed line in Figure 3.4.

In Figure 3.5, subjects were identified according to their primary acid-producers. The most abundant acid-producer in each dentin sample was identified among those S-OTUs that had significant positive estimates. Most subjects had one predominant acid-producer, but four subjects had equal numbers of two acid-producers. Mixed samples included lactobacilli and *S. vestibularis salivarius,* lactobacilli and *S. mutans,* *Propionibacterium FMA5* and *S. mutans,* and *S. vestibularis salivarius* and *S.
parasanguinis-oralis. Heat maps and cluster analyses were generated for each sample type in the Multi-Experiment Viewer. The heat maps for white spots, cavitated lesions, and dentin are displayed in Figure 3.6. When the three heat maps are considered together, a transition is visible as caries progresses. At the white spot stage the samples were divided into many clusters, and many subjects did not have a major acid-producer. As caries progresses to the cavitated stage, many samples became dominated by Streptococcus mutans, fewer clusters were obvious, and only a few subjects did not have any clones representing a major acid-producer. Finally in the dentin lesions, there were still many subjects with high levels of Streptococcus mutans, but there were also some patients who shifted to a predominantly Lactobacillus infection. Every subject had some level of at least one of the caries-associated S-OTUs S. mutans, Total Lactobacillus, or S. vestibularis salivarius.

Data is not shown for the following analyses of potential caries risk factors. A t-test compared the number of courses of antibiotics taken throughout the subjects’ lifetimes, and indicated that antibiotic history did not predict whether a subject had caries. Chi-squared analyses showed that exposure to environmental smoking (by parents or anyone else who was regularly around the subject) did not predict the presence of the subject in either the healthy or caries group. The chi-squared analysis comparing healthy and caries subjects based on fluoride exposure was not significant, but was suspect because of expected values that were too low.
3.4 Discussion

Early childhood caries continues to be a public health problem, despite some understanding of its progression and continued preventive efforts by dentists. It is critical that a complete picture of the caries and healthy biofilms in the primary dentition be understood, so that new therapies can be developed and implemented in children. This study used a unique approach to study severe early childhood caries as compared to previous investigations (1, 7, 11, 21, 22, 61, 66, 75), using an open-ended molecular approach to potentially identify any bacteria present in the samples. Using this technique, 130 S-OTUs were found, representing six bacterial phyla. Interestingly, fewer than 5% of the total clones were uncultivated. Other studies of the oral microbiome have reported uncultivated species of at least 50% (1, 2, 45). This contradictory finding is likely directly related to the use of the new S-OTU designations described in Chapter 1. In previous studies by this laboratory, simple distance criteria were applied to identify bacteria, rather than the robust approach used here. We also have a better awareness of artifactual sequence among the oral clones in GenBank and discarded them from our database. Using statistical support as described, most oral clones from GenBank were combined with named, cultivated species within S-OTUs. Many may have been identified as uncultivated oral clones if GenBank and simple distance criteria had been used, but they were assigned to a cultivated S-OTU with the new naming system.

This study identified several bacterial S-OTUs as potential caries pathogens. *S. mutans* has been associated with caries in both culture-based and molecular studies (1, 7, 11, 21, 22, 61, 66, 75), and was also associated with caries here. Other S-OTUs that were made up of species that have been associated with early childhood caries in other
investigations included lactobacilli, *S. vestibularis salivarius, S. parasanguinis–oralis* and *Propionibacterium* FMA5 (1, 7, 21, 22). While it is unknown whether *P. denticolens* or *Mitsuokella* sp. oral taxon 131 have been associated with early childhood caries previously, it is known that they are capable of producing acids, so their role in caries pathogenesis is certainly probable.

This study found 22 S-OTUs that had significant negative PROC MIXED estimates. Many of these S-OTUs were those that were found at relatively low levels in healthy communities, but were lost as one of the acid-producers identified above took hold at the site. While it is often difficult to compare these results to the literature because of the new S-OTU designations, species that are a part of significant S-OTUs have been associated with health in previous investigations. Species that have consistently been associated with health in many studies of early childhood caries include *S. sanguinis*, strains of *S. mitis* and *S. oralis*, *S. cristatus*, and *G. morbillorum* (1, 7, 21, 22). *S. sanguinis*, which was highly significant (*p*<0.01) and had the most negative estimate, has been shown to delay colonization of *S. mutans* in infants (17). It is very important to study health-associated species as potential targets for caries prevention and treatment. If a health-promoting community could somehow be stabilized in susceptible hosts, then perhaps it could prevent the overgrowth of the major acid-producers that place subjects on a slippery slope toward a lifetime of caries experience.

The Shannon Diversity data presented in Figure 3.4 and the results of PROC MIXED analysis indicate that there was an overall decrease in S-OTU diversity as caries progressed from healthy controls through cavitated lesions. This is in agreement with recent investigations that found decreasing diversity with the progression of severe caries
in children. Arif et al (5) found that the diversity of the genus Veillonella was lower in caries lesions than at caries-free sites, and Li et al (50) found significantly lower diversity in dental plaque from caries subjects when comparing the Shannon Diversity Indices of that group to caries-free children. Together these data support the ecological plaque hypothesis described by Marsh (62, 63), which attributes dental caries to a shift from a stable microbial community to one that is dominated by species that have the capacity to contribute to disease. In this study S. mutans, lactobacilli, S. vestibularis salivarius, S. parasanguinis oralis, Propionibacterium FMA5, Mitsuokella sp. oral taxon 131, and Parascardovia denticolens are found at higher levels in disease, even though diversity has decreased. They have become the dominant bacteria in plaque associated with disease, and each can contribute to dental caries by producing the acid that demineralizes enamel. They were able to dominate infections at the expense of the many acid-sensitive health-associated species identified (62).

Figure 3.4 also shows that there was no significant difference between the means of the Shannon Diversity Index for the healthy control samples and the intact enamel samples of caries subjects. Not only was there no difference between the Shannon Diversity Indices for the samples, but there were also no significant differences between means of individual S-OTUs for healthy sites in controls and caries subjects. Based on these observations, the PROC MIXED analysis was also computed without the data from the healthy samples, that is, it was run with severity levels two through five only. However, this resulted in fewer significant S-OTUs being identified. Because important information was clearly lost by omitting the healthy samples, they were included for the final analyses.
Several S-OTUs were identified here as contributors to tooth decay. This is support for the heterogeneous etiology of dental caries. Combined, Figures 3.5 and 3.6 show that the major acid-producer in a subject can be one of several different pathogenic bacteria, or a mixture of several. It can be seen in the heat map for white spot samples (Figure 3.6) that many subjects do not have a major acid-producer at this stage of caries. Yet, by the time caries progressed to dentin, every subject had some detectable level of either *S. mutans*, *S. vestibularis salivarius*, or lactobacilli. The microbial profiles in dentin are simple, with the biofilms being dominated by only a few S-OTUs for most subjects. Most subjects had a predominantly *S. mutans* infection, while smaller clusters were predominantly lactobacilli, *S. vestibularis salivarius*, *S. sobrinus*, or a mixture of S-OTUs.

A small subset of subjects had a mixture of low levels of these main acid-producers, and had higher levels of *S. mitis pneumoniae infantis–oralis*. It is known that some strains of *S. mitis* and *S. oralis* can produce lactic acid at a rate greater than that of *S. mutans* under the right conditions and can continue at low pH (23), but because strains have been combined as the S-OTU *S. mitis pneumoniae infantis–oralis*, strain specific patterns could not be detected. In dentin samples nearly every subject had *Veillonella atypica dispar parvula* within the community. *Veillonella atypica dispar parvula* had a positive PROC MIXED estimate, but it was not significant. As utilizers of lactic acid, however, *Veillonella* species have previously been associated with caries. *Veillonella alcalescens* has been shown in vitro to produce more acid with *S. mutans* than either species produced alone (72) and other *Veillonella* species were associated with caries in primary teeth in a recent molecular study (1).
The results of this open-ended molecular study of severe caries in the primary dentition have corroborated many other investigations that have identified *S. mutans* and lactobacilli among the primary caries pathogens. It has also identified new potential caries pathogens, including other streptococci, *Propionibacterium* FMA5, *Mitsuokella* sp. oral taxon 131, and *Parascardovia denticolens*. There were many S-OTUs that were lost as caries progressed, and it was shown using the Shannon Diversity Index that S-OTU diversity decreases in severe caries. Further study is necessary to determine whether any of these S-OTUs can be exploited as targets for interventions and therapies for early childhood caries. Based on the microbial patterns elucidated by the identification of primary acid-producers and the cluster analyses, it is probable that screening criteria for caries risk could be developed. Subjects could be screened for several potential pathogens or health-associated species, and it could be determined whether any of these species are over-represented or under-represented considering the age of the patient. The most successful implementation of this screening may be having the ability to detect the loss of acid-sensitive species, which signals the ecological shift to a more cariogenic biofilm. Being able to accurately predict future caries experience could help in the determination of the appropriate course of interventions and therapies for the patient.
Table 3.1. All S-OTUs identified in severe early childhood caries, with percentage of total clones. Rank puts them in order of decreasing overall prevalence.
<table>
<thead>
<tr>
<th>Rank</th>
<th>Supported OTU</th>
<th>% of Clones</th>
<th>Rank</th>
<th>Supported OTU</th>
<th>% of Clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Abiotrophia deflectiva</td>
<td>1.32</td>
<td>83</td>
<td>Lautropia AP009</td>
<td>0.14</td>
</tr>
<tr>
<td>85</td>
<td>Achromobacter xylosoxidans</td>
<td>0.04</td>
<td>41</td>
<td>Lautropia FX006</td>
<td>0.31</td>
</tr>
<tr>
<td>86</td>
<td>Acidaminococcaceae/[G-1]/sp oral taxon 155</td>
<td>0.04</td>
<td>9</td>
<td>Lautropia mirabilis</td>
<td>2.39</td>
</tr>
<tr>
<td>58</td>
<td>Actinobaculum EL030</td>
<td>0.17</td>
<td>119</td>
<td>Leptotrichia shahii</td>
<td>0.01</td>
</tr>
<tr>
<td>68</td>
<td>Actinobaculum sp A11</td>
<td>0.10</td>
<td>57</td>
<td>Megagamera micronucleiformis</td>
<td>0.19</td>
</tr>
<tr>
<td>60</td>
<td>Actinomyces dentalis</td>
<td>0.16</td>
<td>35</td>
<td>Mitsuokella sp oral taxon 151</td>
<td>0.40</td>
</tr>
<tr>
<td>39</td>
<td>Actinomyces georgiae</td>
<td>0.31</td>
<td>80</td>
<td>Mitsuokella sp oral taxon 251</td>
<td>0.06</td>
</tr>
<tr>
<td>33</td>
<td>Actinomyces gerencseriae</td>
<td>0.40</td>
<td>120</td>
<td>Mogibacterium timaeum pumilum neglectum diversum vascul</td>
<td>0.01</td>
</tr>
<tr>
<td>95</td>
<td>Actinomyces graevenitzii</td>
<td>0.03</td>
<td>107</td>
<td>Moraxellaceae sp</td>
<td>0.03</td>
</tr>
<tr>
<td>48</td>
<td>Actinomyces IO077</td>
<td>0.26</td>
<td>50</td>
<td>Neisseria AP085</td>
<td>0.24</td>
</tr>
<tr>
<td>37</td>
<td>Actinomyces IP073</td>
<td>0.36</td>
<td>70</td>
<td>Neisseria bacilliformis</td>
<td>0.10</td>
</tr>
<tr>
<td>42</td>
<td>Actinomyces massiliensis</td>
<td>0.30</td>
<td>20</td>
<td>Neisseria elongata</td>
<td>0.79</td>
</tr>
<tr>
<td>52</td>
<td>Actinomyces odontolyticus lingnai</td>
<td>0.23</td>
<td>5</td>
<td>Neisseria flav a mucosa pharyngis sicca</td>
<td>4.94</td>
</tr>
<tr>
<td>6</td>
<td>Actinomyces viscosus nasi inund</td>
<td>3.31</td>
<td>22</td>
<td>Neisseria flavescens</td>
<td>0.73</td>
</tr>
<tr>
<td>96</td>
<td>Aggregatibacter segnis</td>
<td>0.03</td>
<td>53</td>
<td>Neisseria meningitidis–polysacchara</td>
<td>0.23</td>
</tr>
<tr>
<td>74</td>
<td>Alobiporum parvulum</td>
<td>0.09</td>
<td>90</td>
<td>Olsenella profusa</td>
<td>0.04</td>
</tr>
<tr>
<td>111</td>
<td>Alobiporum riae</td>
<td>0.01</td>
<td>121</td>
<td>Orbibacterium sinus</td>
<td>0.01</td>
</tr>
<tr>
<td>97</td>
<td>Bilobobacterium scardovii denti</td>
<td>0.03</td>
<td>91</td>
<td>Orbibacterium sp group B</td>
<td>0.04</td>
</tr>
<tr>
<td>45</td>
<td>Campylobacter concisus</td>
<td>0.27</td>
<td>54</td>
<td>Parascardovia denticolens</td>
<td>0.23</td>
</tr>
<tr>
<td>12</td>
<td>Campylobacter curvus</td>
<td>0.01</td>
<td>71</td>
<td>Parvimonas micra</td>
<td>0.10</td>
</tr>
<tr>
<td>27</td>
<td>Campylobacter gracilis</td>
<td>0.63</td>
<td>122</td>
<td>Peptostreptococcus stomatis</td>
<td>0.01</td>
</tr>
<tr>
<td>31</td>
<td>Campylobacter showlee rectus</td>
<td>0.59</td>
<td>81</td>
<td>Porphyromonas catoniae</td>
<td>0.06</td>
</tr>
<tr>
<td>19</td>
<td>Capnophytophaga gingivalis–granulosa</td>
<td>0.67</td>
<td>123</td>
<td>Prevotella oris</td>
<td>0.01</td>
</tr>
<tr>
<td>87</td>
<td>Capnophytophaga leakei</td>
<td>0.04</td>
<td>17</td>
<td>Propionibacterium Bl085</td>
<td>0.07</td>
</tr>
<tr>
<td>76</td>
<td>Cardiohydrobacter hominis</td>
<td>0.07</td>
<td>32</td>
<td>Propionibacterium FM05</td>
<td>0.46</td>
</tr>
<tr>
<td>98</td>
<td>Centipeda periodontii</td>
<td>0.03</td>
<td>47</td>
<td>Propionibacterium propionicus A</td>
<td>0.27</td>
</tr>
<tr>
<td>43</td>
<td>Corynebacterium durum</td>
<td>0.29</td>
<td>124</td>
<td>Pseudomonas synxantha fulgida bibanensis</td>
<td>0.01</td>
</tr>
<tr>
<td>15</td>
<td>Corynebacterium matruchotii</td>
<td>0.09</td>
<td>72</td>
<td>Pseudoramibacter alcalolyticus</td>
<td>0.10</td>
</tr>
<tr>
<td>99</td>
<td>Cryptobacterium curtum</td>
<td>0.03</td>
<td>61</td>
<td>Rothia aeria</td>
<td>0.16</td>
</tr>
<tr>
<td>101</td>
<td>Dialister ivitus</td>
<td>0.03</td>
<td>51</td>
<td>Rothia dentocariosa</td>
<td>0.24</td>
</tr>
<tr>
<td>44</td>
<td>Eikenella corrodens</td>
<td>0.29</td>
<td>108</td>
<td>Rothia mucilaginosa</td>
<td>0.03</td>
</tr>
<tr>
<td>113</td>
<td>Eschecchia coli</td>
<td>0.01</td>
<td>73</td>
<td>Scardovia inopinata</td>
<td>0.10</td>
</tr>
<tr>
<td>114</td>
<td>Eubacterium brachy</td>
<td>0.01</td>
<td>82</td>
<td>Scardovia sp C1</td>
<td>0.06</td>
</tr>
<tr>
<td>101</td>
<td>Eubacterium BU014</td>
<td>0.03</td>
<td>29</td>
<td>Selenomonas AA024</td>
<td>0.60</td>
</tr>
<tr>
<td>115</td>
<td>Eubacterium CH017</td>
<td>0.01</td>
<td>109</td>
<td>Selenomonas AJ036</td>
<td>0.03</td>
</tr>
<tr>
<td>102</td>
<td>Eubacterium DO006</td>
<td>0.03</td>
<td>125</td>
<td>Selenomonas CS024</td>
<td>0.01</td>
</tr>
<tr>
<td>55</td>
<td>Eubacterium EI074</td>
<td>0.21</td>
<td>62</td>
<td>Selenomonas dianae</td>
<td>0.16</td>
</tr>
<tr>
<td>25</td>
<td>Eubacterium sp group C</td>
<td>0.09</td>
<td>83</td>
<td>Selenomonas DO042</td>
<td>0.06</td>
</tr>
<tr>
<td>88</td>
<td>Eubacterium yuri subs p margaretiae</td>
<td>0.04</td>
<td>110</td>
<td>Selenomonas DY027</td>
<td>0.03</td>
</tr>
<tr>
<td>103</td>
<td>Eubacterium yuri subs p yuri</td>
<td>0.03</td>
<td>126</td>
<td>Selenomonas EW051a</td>
<td>0.01</td>
</tr>
<tr>
<td>65</td>
<td>Firmicutes F005</td>
<td>0.11</td>
<td>92</td>
<td>Selenomonas EY047</td>
<td>0.04</td>
</tr>
<tr>
<td>116</td>
<td>Fusobacterium nucleatum</td>
<td>0.01</td>
<td>93</td>
<td>Selenomonas flueggei</td>
<td>0.04</td>
</tr>
<tr>
<td>38</td>
<td>Gemella haemolysans</td>
<td>0.36</td>
<td>78</td>
<td>Selenomonas FT050</td>
<td>0.07</td>
</tr>
<tr>
<td>14</td>
<td>Gemella morbillorum</td>
<td>1.06</td>
<td>21</td>
<td>Selenomonas infelix</td>
<td>0.77</td>
</tr>
<tr>
<td>75</td>
<td>Gemella sanguinis</td>
<td>0.09</td>
<td>12</td>
<td>Selenomonas noxia</td>
<td>1.25</td>
</tr>
<tr>
<td>18</td>
<td>Granulicatella adiacens</td>
<td>0.92</td>
<td>56</td>
<td>Selenomonas sp</td>
<td>0.20</td>
</tr>
<tr>
<td>79</td>
<td>Granulicatella elegans</td>
<td>0.06</td>
<td>127</td>
<td>Shuttlworthia sateles</td>
<td>0.07</td>
</tr>
<tr>
<td>104</td>
<td>Haemophilus genomas P3 oral clone MB3_C38</td>
<td>0.03</td>
<td>84</td>
<td>Streptococcus anginosus B</td>
<td>0.06</td>
</tr>
<tr>
<td>66</td>
<td>Haemophilus parainfluenzae group A</td>
<td>0.11</td>
<td>64</td>
<td>Streptococcus australis–sanguinis</td>
<td>0.14</td>
</tr>
<tr>
<td>105</td>
<td>Johnsonella ignava</td>
<td>0.03</td>
<td>10</td>
<td>Streptococcus cristas</td>
<td>1.39</td>
</tr>
<tr>
<td>117</td>
<td>Johnsonella sp</td>
<td>0.01</td>
<td>30</td>
<td>Streptococcus downei</td>
<td>0.60</td>
</tr>
<tr>
<td>28</td>
<td>Kingella denitrificans</td>
<td>0.63</td>
<td>8</td>
<td>Streptococcus gordoni</td>
<td>2.48</td>
</tr>
<tr>
<td>13</td>
<td>Kingella oralis</td>
<td>1.17</td>
<td>17</td>
<td>Streptococcus intermedius constellatus–anginosus</td>
<td>0.96</td>
</tr>
<tr>
<td>106</td>
<td>Kingella sp</td>
<td>0.03</td>
<td>3</td>
<td>Streptococcus mitis pneumoniae infantis–oralis</td>
<td>13.41</td>
</tr>
<tr>
<td>89</td>
<td>Lachnospiraceae C1_D0016</td>
<td>0.04</td>
<td>1</td>
<td>Streptococcus mutans</td>
<td>20.11</td>
</tr>
<tr>
<td>26</td>
<td>Lactobacillus casei paracasei</td>
<td>0.66</td>
<td>128</td>
<td>Streptococcus oligofermentans sinensis</td>
<td>0.01</td>
</tr>
<tr>
<td>118</td>
<td>Lactobacillus crispatus</td>
<td>0.01</td>
<td>24</td>
<td>Streptococcus parasaurophilis–oralis</td>
<td>0.70</td>
</tr>
<tr>
<td>49</td>
<td>Lactobacillus fermentum</td>
<td>0.26</td>
<td>67</td>
<td>Streptococcus peroris</td>
<td>0.11</td>
</tr>
<tr>
<td>16</td>
<td>Lactobacillus gassen johnsonii</td>
<td>0.59</td>
<td>4</td>
<td>Streptococcus sanguinis</td>
<td>6.64</td>
</tr>
<tr>
<td>69</td>
<td>Lactobacillus oris</td>
<td>0.10</td>
<td>23</td>
<td>Streptococcus sobrinus</td>
<td>0.72</td>
</tr>
<tr>
<td>40</td>
<td>Lactobacillus plantarum</td>
<td>0.31</td>
<td>7</td>
<td>Streptococcus vestibularis salivarius</td>
<td>3.01</td>
</tr>
<tr>
<td>34</td>
<td>Lactobacillus rhamnosus–casei</td>
<td>0.40</td>
<td>129</td>
<td>Tannerella BU063</td>
<td>0.01</td>
</tr>
<tr>
<td>59</td>
<td>Lactobacillus salivarius</td>
<td>0.17</td>
<td>94</td>
<td>Treponema sororositii subs p buccale paretis sororositii</td>
<td>0.04</td>
</tr>
<tr>
<td>46</td>
<td>Lactobacillus ultunensis</td>
<td>0.27</td>
<td>2</td>
<td>Veillonella alpica dispar pastura</td>
<td>13.67</td>
</tr>
<tr>
<td>36</td>
<td>Lactobacillus vaginalis</td>
<td>0.37</td>
<td>130</td>
<td>Veillonella HB016</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 3.1
Figure 3.1. Distribution of S-OTUs by phyla for all samples.
**Streptococcus mutans**  
**Total Lactobacillus**  
*Streptococcus vestibularis salivarius*  
Veillonella atypica dispar parvula  
**Streptococcus parasanguinis-oralis**  
**Propionibacterium FMA5**  
*Mitsuokella sp oral taxon 131*  
Streptococcus sobrinus  
*Parascardovia denticolens*  
Actinomyces IP073  
Lautropia FX006  
Actinomyces odontolyticus linguae  
Rothia dentocariosa  
Neisseria elongata  
Actinomyces IO077  
Gemella haemolysans  
Neisseria meningitidis-poly saccharea  
Selenomonas sputigena  
Streptococcus downei  
Granulicatella adiacens  
Actinomyces gerencseriae  
Actinomyces massiliensis  
Campylobacter concisus  
*Propionibacterium propionicus A*  
**Eubacterium E1074**  
Neisseria AP085  
**Corynebacterium durum**  
*Actinomyces georgiae*  
*Kingella denitrificans*  
**Eikenella corrodens**  
*Campylobacter gracilis*  
**Eubacterium sp group C**  
Actinomyces viscosus naeslundii  
*Kingella oralis*  
*Neisseria flavescens*  
*Campylobacter showae rectus*  
**Selenomonas AA024**  
**Selenomonas infelix**  
Lautropia mirabilis  
Selenomonas noxia  
**Corynebacterium matruchotii**  
**Abiotrophia defectiva**  
**Capnocytophaga gingivalis-granulosa**  
**Streptococcus intermedius constellatus-anginosus**  
**Gemella morbillo rum**  
Streptococcus gordonii  
**Streptococcus cristatus**  
*Neisseria flava mucosa pharyngis sicca*  
**Streptococcus mitis pneumoniae infants-oralis**  
**Streptococcus sanguinis**

Figure 3.2. PROC MIXED estimates of percentage of increase or decrease in bacterial levels as caries progressed, with 95% confidence intervals. *P*-values were corrected using the False Discovery Rate. * p < 0.05. ** p < 0.01.
Figure 3.3. Mean levels for S-OTUs greater 0.2% of total clones, with 95% confidence intervals. From left to right within an S-OTU, bars indicate levels for healthy control, intact enamel, white spot, cavitated, and dentin samples. S-OTUs that are green have significant negative estimates, and S-OTUs that are red have significant positive estimates.
Figure 3.3
Figure 3.4. Mean Shannon Diversity Indices by sample type, with 95% confidence intervals. The means for healthy controls and intact enamel are significantly higher than the other three caries samples, as indicated by the brackets. There are no other significant differences. The dashed line represents the significant PROC MIXED estimate, which had a value of -0.1964.
Figure 3.5. Subjects identified according to their primary acid-producers. The most abundant acid-producer in each dentin sample was identified among those S-OTUs that had significant positive estimates. Four subjects (“Mix” above) had equal numbers of two acid-producers. Mixed samples included lactobacilli and *S. vestibularis salivarius*, lactobacilli and *S. mutans*, *Propionibacterium* FMA5 and *S. mutans*, and *S. vestibularis salivarius* and *S. parasanguinis–oralis*. 

![Bar graph showing acid-producers](image)
Figure 3.6. Hierarchical cluster analysis of white spot, cavitated, and dentin samples. S-OTUs that were the primary acid producers in at least one subject are in decreasing order of their PROC MIXED estimates. The color of each cell represents the percent of total clones within that sample for the S-OTU of interest, according to the scale above.
Figure 3.6

Dentin  Cavitated  White spot

S. mutans  Total Lactobacillus  S. vestibularis salivarius  S. sobrinus

Sample number
CHAPTER 4

MOLECULAR MICROBIOLOGICAL ANALYSIS OF SEVERE CARIES IN PERMANENT TEETH

4.1 Introduction

Dental caries is the most common chronic disease of childhood, and is especially a problem for children of lower socioeconomic status and minorities (74). The etiologic roles of Streptococcus mutans and Lactobacillus species in dental caries were established based on culture studies (107), but culture studies limited the ability to study the role of other species present in the biofilm. Recently molecular methods have been used to more completely study several oral microbiome niches: subgingival plaque (77); periodontal pockets (45); caries and healthy biofilms in early childhood caries (7), twins in the primary dentition (21), and root caries in the elderly (81); advanced dentin lesions (14, 70); and the normal bacterial flora of several sites in the oral cavity (2). There was a very recent study that used molecular techniques to identify health- and caries-associated bacterial species in primary and permanent teeth in children and young adults (1).

The purpose of this study was to identify the oral bacteria associated with severe dental caries and health in the young permanent dentition using an open-ended molecular
Our ultimate goal is to combine the results of this investigation with those from similar investigations of early childhood caries and severe caries of the primary dentition, constructing a natural history of caries progression from the incipient stages through the severe destruction of an established infection.

4.2 Materials and Methods

Clinical Methods

Subject recruitment

Subjects were recruited at the Nationwide Children’s Hospital Dental Clinic in Columbus, Ohio. The inclusion requirement for the caries group was the presence of at least three permanent teeth with multi-surface lesions with at least one smooth surface involved, and at least one of those teeth had a vital pulp. Age-, race-, and gender-matched healthy controls that were caries-free and had no existing restorations were also recruited. General exclusionary criteria for either group included (i) age greater than 16 years, (ii) need for SBE prophylaxis, (iii) antibiotic use in the past 30 days, and (iv) professional cleaning in the past 30 days. Only one child per family was included in each group. Consent was obtained from the parents of all subjects, and assent was obtained from subjects who were at least nine years old. This study was approved by the Institutional Review Board.

Sampling and clinical data collection

Dental plaque was collected by swiping the tooth or lesion surface with a dental explorer and wiping the plaque onto a coarse endodontic paper point. Each sample was
obtained by pooling plaque collected from multiple teeth, placed in a sterile 1.5-ml microcentrifuge tube and frozen for storage. Dental plaque was sampled from the intact enamel of healthy subjects. If a subject was determined to have dental caries, then plaque was collected separately from the surfaces of each of three types of sites available: (i) healthy, intact enamel, (ii) white spot lesions, and (iii) cavitated lesions. Carious dentin was sampled from a tooth that required pulp therapy. For healthy subjects the number of plaque samples is one, and for caries subjects the total number of samples four.

The teeth present were recorded on the tooth scoring form for healthy subjects and the caries scoring form for caries subjects. For caries subjects existing restorations were recorded and all carious surfaces were scored and recorded according to the worst presentation on the surface by the following codes:

- w—visible white spot lesions
- c—surface cavitation, not extending to pulp
- p—requires pulp therapy
- x—extraction indicated

Parents completed a brief survey regarding the history of potential caries risk factors for their children. The answers to this survey were used to determine the antibiotic and medication history, fluoride status, oral hygiene practices, and exposure to cigarette smoke for each subject. They also provided an open-ended account of each subject’s drinking and snacking preferences, including types of drinks and snacks and how frequently they were consumed.
Laboratory Methods

Sample preparation

Bacterial DNA was isolated using a bead beater. Samples were placed in 300 µl of TE buffer and then beaten with 0.25 g of 0.1 mm glass beads for 60 seconds at 4,800 rpm in a Biospec Products bead beater. The bacterial DNA was purified using the QIAGEN QIAamp® DNA Mini Kit according to the manufacturer’s instructions and frozen until analysis.

PCR amplification for clonal analysis

The 16S rRNA genes were amplified from the purified bacterial DNA using universal primers A17 (5’-GTT TGA TCC TGG CTC AG-3’) and 317 (5’-AAG GAG GTG ATC CAG GC-3’) (Biosynthesis, Lewisville, TX). PCR conditions were as previously described (45). The products of PCR amplification were examined by electrophoresis in 1% agarose and purified using the QIAquick PCR Purification Kit (QIAGEN, Valencia, CA).

Cloning and sequencing

Cloning and sequencing of the 16S amplicons generated by PCR were as previously described (45). Amplicons encompassed the hypervariable regions V5-V9.

Bacterial 16S rDNA sequence identification

Each clone was identified by comparing it to a local, curated oral microbiome database of supported operational taxonomic units (S-OTUs) using the National Center
for Biotechnology Information’s Basic Local Alignment Search Tool (BLAST). The development of this database was described in Chapter 2. Briefly, oral microbiome 16S rDNA sequences were identified for input into the database from the published literature, unpublished sequences from our lab, and GenBank. Sequences were aligned to each other using Clustal and the resulting alignment was manually examined and trimmed of its incompletely sequenced 5’ and 3’ ends. The edited multiple sequence alignment was used as the input for a maximum likelihood phylogenetic tree search. Based on the topology of the phylogenetic tree, distances between clades, and bootstrap support, S-OTUs were identified. S-OTUs often encompassed sequences from more than one previously named species, because the sequences were too closely related to be distinguishable at the 16S level. In this case the name of each species was preserved in the final S-OTU designation.

Data Management and Statistical Analyses

Sequences were collected into a database to be analyzed for this study. Each clone was identified by its S-OTU, the type of site from which it was collected (healthy, intact enamel, white spot, cavitated, and dentin), and its match number. The match number was the same for all samples from the same subject, plus the paired healthy control. S-OTU levels were calculated as a percent of the total bacterial population for each sample. PROC MIXED analysis was performed for this repeated measures dataset. The covariance structure was the default various components structure. The type of site from which each sample was collected was expressed as a level of severity in the PROC MIXED analysis on a scale from one to four, with one being the least severe and four
being the most severe form of caries. Using this scale, healthy control samples were assigned a value of 1, intact enamel samples were assigned a value of 2, white spot lesion samples were assigned a value of 3, cavitated lesion samples were assigned a value of 4, and dentin samples were assigned a value of 5. The PROC MIXED analysis used a linear model to calculate an estimate of the percent of increase or decrease of each S-OTU as caries progressed from severity one through five, determined the significance of the estimate at $\alpha = 0.05$, and was followed by a False Discovery Rate correction for multiple comparisons. Mean bacterial levels and 95% confidence intervals were determined for the most prevalent S-OTUs, and the percentage of cultivated S-OTUs was determined for each phylum present in the study. The Shannon Diversity Index was computed for each sample, using the formula $H' = -\sum p_i \ln p_i$, where $p_i$ equals the relative abundance of each of the $i$th species (60), and was analyzed by PROC MIXED. Post hoc tests were used to elucidate differences in the Shannon Diversity Index among sample types. The primary acid-producer was identified for each dentin sample. Cluster analyses were performed for each sample type to examine patterns of acid-producers among subjects and were expressed as heat maps. Subjects were additionally divided on the basis of citric acid (such as Mountain Dew) consumption, fluoride exposure, environmental smoke exposure, and antibiotic history. Chi-squared analyses were used to determine presence in the caries or healthy groups based on these divisions. T-tests were also used to compare mean cumulative caries scores and mean levels of specific bacterial S-OTUs for the groups. The cumulative caries score was computed by weighting each type of caries present, summing the products, and dividing by the number of teeth present.
4.3 Results

Twenty-one caries subjects and 18 healthy controls were recruited for this study. The number of clones analyzed per sample ranged from 47 to 60, with an average of 52 clones per sample. The total number of clones for all samples was 5,299. Both the caries and healthy groups contained two Hispanic subjects, but the caries groups had 19 non-Hispanics and the healthy group had only 16. This difference was not significant by Chi-squared analysis. The caries group contained 12 white subjects, six African-Americans, one Native American, and two subjects who identified their race as “Other.” The healthy group had 10 Caucasions, six African-Americans, and two “Other.” Chi-squared analysis indicated that these differences were not significant. Ages ranged from seven to 16 years old (mean age was 13) and the difference in means between the caries and healthy groups was not significant by t-test.

Table 4.1 lists the 144 S-OTUs identified in this study in alphabetical order, with percentage of clones and their overall rank. The most abundant S-OTU in this study was *Veillonella atypica dispar parvula*, although it was given the rank 2. All of the individual *Lactobacillus* S-OTUs were combined into one “Total *Lactobacillus*” group for subsequent statistical analyses, so for most analyses Total *Lactobacillus* was the most abundant S-OTU. The 144 S-OTUs could be divided into seven bacterial phyla, shown in Figure 4.1. *Spirochaetes*, *Deinococcus-Thermus*, and *Fusobacteria* were found at such low levels that they appear to be zero on the graph. Overall, only 6.6% of total clones were uncultivated.
PROC MIXED was used to analyze data for those S-OTUs that were greater than 0.2% of the total clones. The resulting linear estimates of the percent of increase or decrease as caries progressed are depicted in Figure 4.2. Thirty-seven S-OTUs had negative estimates (levels decreased as caries severity increased), and nine of those were significant. Of 19 S-OTUs that had positive estimates (levels increased as caries severity increased), only Total Lactobacillus and Propionibacterium FMA5 were significant. *Streptococcus mutans* was not significant in this study. Figure 4.3 graphs the mean levels of every S-OTU that was analyzed by PROC MIXED, and provides insight into the relative abundance of many common oral microbiome inhabitants within this population of subjects. It is clear in Figure 4.3 that Total Lacobacillus is much more abundant than other S-OTUs in this study population, and that many S-OTUs are found at relatively low levels, no matter the severity of disease.

The Shannon Diversity Index was computed for each sample and was analyzed by PROC MIXED. The resulting negative estimate was significant, and *post hoc* t-tests and paired t-tests were used to identify differences among sample types. Figure 4.4 shows that the mean for healthy sites was significantly higher than those for caries sites. There was also a significant difference between intact enamel and dentin samples. Overall, bacterial diversity decreased as caries progressed.

In Figure 4.5, subjects were identified according to their primary acid-producers. The most abundant acid-producer in each dentin sample was identified among those S-OTUs that had significant positive estimates and *S. mutans*. One subject had equal numbers of two acid-producers, *S. mutans* and Total Lactobacillus. For those subjects without a significant S-OTU greater than 2% of clones, a known acid-producer was
selected from among the non-significant S-OTUs. Here they are *A. viscosus naeslundii, C. matruchotii*, and *S. inopinata*. Heat maps and cluster analyses were generated for each sample type in the Multi-Experiment Viewer. The heat maps for white spot, cavitated, and dentin lesions are displayed in Figure 4.6. It is clear when examining the three heat maps together that for many people, caries shifts from a *Streptococcus* dominated infection to a *Lactobacillus* dominated infection. Several distinct profiles are evident in dentin samples, including *S. mutans* or *Lactobacillus* dominated, and *S. mutans* and *Lactobacillus* together. In order to look more closely at the subjects who had high levels of *Lactobacillus*, subjects were divided into those who regularly and frequently consumed citric acid (like Mountain Dew) and those that did not (data not shown). For cavitated samples there was no difference in the means comparing Total *Lactobacillus* or *Streptococcus mutans* for consumers and non-consumers. There was also no difference between the two groups in the means of the cumulative caries score.

Data is not shown for the following analyses of potential caries risk factors. Chi-squared analysis showed that exposure to environmental smoking (by parents or anyone else who is regularly around the subject) did not predict the presence of the subject in either the healthy or caries group. However, significantly more subjects in the caries group reported that their primary water sources did not contain fluoride. *T*-tests showed that cumulative caries scores, *Streptococcus mutans* levels, and Total *Lactobacillus* levels could not be predicted by antibiotic use or environmental smoke exposure. Subjects were also divided by those that regularly consumed citric acid and those that did not, and this designation did not predict presence in the caries or healthy groups by chi-squared analysis. *T*-tests also failed to find a difference in mean cumulative caries scores, levels
of Total *Lactobacillus*, or levels of *S. mutans* based on citric acid consumption. Figure 4.6 shows that there are no significant differences among the means of cumulative caries scores for any subjects based on the acid-producers present.

### 4.4 Discussion

The association of *Streptococcus mutans* and *Lactobacillus* species with dental caries is well-established. However, previous studies have used cultivation and targeted molecular methods that limit the possibility of identifying other potential caries pathogens. In this study an open-ended design using 16S cloning and sequencing made possible a more thorough investigation of the microbial communities associated with disease and health in children and adolescents with severe caries of the permanent dentition. The results presented here provide evidence for the heterogeneous and complex nature of oral biofilms.

Table 4.1 lists 144 S-OTUs that were identified in the supragingival dental plaque of this study population, including three previously unidentified phylotypes. S-OTUs belonged to seven phyla, and nearly 80% belonged to the Gram-positive phylum Firmicutes. In contrast to previous investigations of the oral microbiome that reported at least 50% uncultivated species (2, 45, 81), less than 7% of the S-OTUs identified were uncultivated. Admittedly, the previous investigations studied different populations. This difference can also be attributed to the naming system developed for this study. *Lactobacillus* S-OTUs represented nearly 14% of the total clones and were combined for statistical analyses. Twelve *Lactobacillus* S-OTUs were combined because they all behave the same way; each S-OTU had a positive PROC MIXED estimate (data not
shown) and was only found in caries subjects. It can be seen in Figure 4.8 that *Lactobacillus* communities vary among caries subjects, with different S-OTUs being the dominant species in different subjects. Some lactobacilli were present only at low levels. It was previously observed that lactobacilli may not be functionally different (70), so they were combined because analyzing them individually failed to capture the significance of the group as a whole (data not shown).

The oral cavity contains several unique ecological niches, and supragingival biofilms from subjects with caries and those who are healthy are among them. The relationships among oral microbes in healthy biofilms are not completely understood, but it is known that they are complex, with many organisms depending on the others for various vital functions. In the caries biofilm, on the other hand, organisms that are acid-producers or have other antimicrobial properties are predominant, and they threaten microbial homeostasis of the biofilm. While it is important to identify the S-OTUs present in healthy and caries biofilms, another important measure of the nature of these communities is diversity, which considers more than just the number of species present. The Shannon Diversity Index was calculated for each sample, which considers not only the number of species present in the sample, but also the distribution of those species (80), making it an ideal measurement for communities in the oral cavity. It was seen when the Shannon Diversity Indices were compared by PROC MIXED across sample types that diversity decreased as caries progressed, with the lowest diversity in dentin samples (means shown in Figure 4.3). As acid-producers like *S. mutans* and lactobacilli increased in the biofilms, an ecological shift occurred, with many acid-sensitive species being lost (62). In fact in the dentin samples of two different subjects, lactobacilli
accounted for 100% of S-OTU richness, having made the biofilm uninhabitable for health-related microbes.

Nearly two-thirds of the S-OTUs analyzed by PROC MIXED analysis had negative estimates, meaning that they decreased as caries progressed. Nine of these estimates were significant; perhaps these species can be targeted as health-associated species that maintain the ecological diversity necessary to prevent acid-producers gaining a foothold in the community. Significant S-OTUs were *Streptococcus mitis pneumoniae infantis−oralis*, *Corynebacterium matruchotii*, *Streptococcus gordonii*, *Streptococcus cristatus*, *Capnocytophaga gingivalis−granulosa*, *Eubacterium* sp. group C, *Campylobacter showae rectus*, and *Lachnospiraceae C1 DO016*. Surprisingly, *S. sanguinis* was not identified as health-associated, although it has been in many investigations (7, 21, 73). Many of the identified S-OTUs have been associated previously with oral health, but the role of others is not well understood. Because of the new criteria that were used here to group *S. mitis pneumoniae infantis−oralis*, it is difficult to compare this finding to the existing literature. Recent investigations have found strains of *S. mitis*, *S. oralis* to be associated with health (1, 21). *C. matruchotii* is associated with calculus formation (69). *S. gordonii* has been associated with sound surfaces in a study of nursing bottle caries (61), and has been found to reduce the effects of many *S. mutans* virulence factors (46, 111). *S. cristatus* has been associated with health in many studies (1, 21, 81). *C. gingivalis−granulosa* and *C. showae rectus* are groups of species that have been considered individually in previous studies, and have been associated with oral health (1). The roles of *Eubacterium* sp. group C and *Lachnospiraceae C1 DO016* are unknown.
Of the remaining positive S-OTUs, only Total *Lactobacillus* and *Propionibacterium* FMA5 had statistically significant PROC MIXED estimates. *Lactobacillus* species have long been associated with dental caries. They are acidogenic and aciduric, and were the first organisms to be identified as potential caries pathogens (107). They have been found to cause caries when inoculated into gnotobiotic rats (28, 86), but are normally considered to be associated with the progression of caries (102, 107). Lactobacilli have been previously detected in carious dentin (14, 56). *Propionibacterium* FMA5 was recently associated with dental caries in primary and permanent teeth (1) and with root caries in elderly patients (81).

It is of interest to note that *S. mutans* did not have a statistically significant positive estimate by PROC MIXED, meaning that it did not significantly increase as caries progressed to dentin. This may be explained by examining the means of *S. mutans* at each stage of the caries progression (Figure 4.3). PROC MIXED uses a linear model to calculate significance, but mean levels of *S. mutans* did not consistently increase as caries progressed; the mean level in healthy controls samples was actually higher than that in the intact enamel of caries subjects, and was nearly at the same level as the mean for white spot lesions. *S. mutans* also decreased from the surface cavitation samples to the dentin samples. It is possible that lactobacilli, which reached high levels in dentin samples, began to eliminate *S. mutans* in dentin samples. It has been shown that lactobacilli have antimicrobial properties, and can inhibit the growth of *S. mutans* (94, 99).

Caries subjects were compared on the basis of their microbial profiles, potentially identifying a clinically relevant difference in the presence of specific acid-producers.
Figure 4.5 shows the percentage of subjects that had high levels of specific acid-producers in dentin samples. Though most subjects could be identified by high levels of *Lactobacillus* or *S. mutans*, four subjects had little to no lactobacilli or *S. mutans* present. *Actinomyces viscosus naeslundii*, *Propionibacterium FMA5*, *Corynebacterium matruchotii*, and *Scardovia inopinata* each were found at high levels for one of these subjects. *Propionibacterium FMA5* was associated with caries in this study, and *A. viscosus naeslundii*, *S. inopinata* have been shown to produce lactic acid (33, 68, 101).

The identification of *C. matruchotii* is difficult to interpret. It was significantly associated with health in this study and is important in calculus formation as part of the normal flora (69). However, some *Corynebacterium* species have been shown to produce lactic acid (10, 65), and an oral clone was associated with primary caries in a recent molecular study (1). In this one sample, it was the most abundant S-OTU besides *Veillonella* species, which are known to be lactic-acid utilizers, so perhaps its role in caries progression is just undetermined. Though they were not significant for the entire population, it is probable that *A. viscosus naeslundii*, *S. inopinata*, and *C. matruchotii* were the important acid-producers in certain subjects. Further study is required to understand the contribution these bacteria make to the progression of caries.

The cluster analysis of dentin samples presented in Figure 4.7 provides an alternative view of the subsets of caries subjects that have developed very different microbial profiles than the others. The data presented in Figures 4.5 and 4.7 provide additional evidence for the heterogeneous etiology of dental caries. At the white spot stage, several subjects do not have a significant acid-producer. As caries progressed to cavitated lesions, there were subjects with high levels of *S. mutans*, and subjects with
high levels of *Lactobacillus* were identified. Finally at the dentin stage, many subjects have a predominantly *Lactobacillus* infection, and since many of these approached levels of greater than 90% lactobacilli, few other species are present. They have simply been lost during the takeover by lactobacilli. Another subset of subjects has a predominantly *S. mutans* infection. Many of these also have relatively high levels of *Veillonella atypica dispar parvula*. *Veillonella atypica dispar parvula* is a lactic-acid utilizer that may facilitate the growth of *S. mutans*. It has been shown *in vitro* that combinations of *S. mutans* and *V. alcalescens* have higher acid production than monoclonies (72). Some subjects had a mixed *Lactobacillus* and *S. mutans* infection. Finally there was a subset of subjects that did not have any lactobacilli or *S. mutans*.

After identifying the primary caries pathogen for each subject, they were divided by *Lactobacillus, S. mutans, Lactobacillus* and *S. mutans*, and Neither, and analyzed for differences in the mean cumulative score for each group. There were no differences among any of these groups (Figure 4.6), meaning that subjects with specific microbial profiles did not have a more severe clinical caries experience than others. To further examine the relationship between *S. mutans* and *Lactobacillus* as caries progressed, levels of each were compared directly to each other at each stage of caries (Figure 4.9).

It can be seen in the figure that white spot lesions have high levels of *S. mutans* and virtually no *Lactobacillus* present. In cavitated lesions, both *S. mutans* and *Lactobacillus* have increased, but by the time caries reached dentin, *S. mutans* began to decline, while *Lactobacillus* continued to increase. This seems to corroborate previous reports that associate *Lactobacillus* species with progressive caries, but not initiation. Because they
are weakly adherent, lactobacilli depend on *S. mutans* and other acid-producers to initiate the caries lesions that are their ideal niches.

Using this open-ended molecular approach, we have been able to identify bacterial S-OTUs that are present in the supragingival biofilm of young permanent teeth. Three S-OTUs were previously unidentified. Several species were found to significantly decrease as caries progressed, while *Lactobacillus* and *Propionibacterium* FMA5 increased in deeper lesions. The community diversity decreased as caries progressed, with dentin samples having the lowest measure of diversity. Subjects could be identified by their primary caries pathogen, but there was no difference among them based on caries experience. Using the data provided by the identification of acid-producers and cluster analyses, it is obvious that caries has heterogeneous etiology, with several S-OTUs contributing to disease. Ideally this information will be used to develop a screening mechanism that identifies caries-susceptible subjects based on their microbial profiles. If the potential for caries development can be identified early, perhaps it can be arrested with the appropriate interventions and therapies.
<table>
<thead>
<tr>
<th>Rank</th>
<th>Supported OTU</th>
<th>% of Clones</th>
<th>Rank</th>
<th>Supported OTU</th>
<th>% of Clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Abiotrophia definita</td>
<td>2.15</td>
<td>86</td>
<td>Haemophilus parainfluenzae A</td>
<td>0.11</td>
</tr>
<tr>
<td>45</td>
<td>Actinomycobacter yoxosidans</td>
<td>0.42</td>
<td>135</td>
<td>Johnsonglota sp</td>
<td>0.02</td>
</tr>
<tr>
<td>66</td>
<td>Acidaminococcaceae G-1 sp</td>
<td>0.21</td>
<td>46</td>
<td>Kingella dentifirica</td>
<td>0.42</td>
</tr>
<tr>
<td>106</td>
<td>Actinobacterium ELO30</td>
<td>0.04</td>
<td>68</td>
<td>Lachnospiraceae bacterium</td>
<td>0.21</td>
</tr>
<tr>
<td>92</td>
<td>Actinomyces denticilis</td>
<td>0.08</td>
<td>35</td>
<td>Kingella oralis</td>
<td>0.53</td>
</tr>
<tr>
<td>117</td>
<td>Actinomyces georgii</td>
<td>0.02</td>
<td>56</td>
<td>Lachnospiraceae C1_DO016</td>
<td>0.21</td>
</tr>
<tr>
<td>59</td>
<td>Actinomyces genomei</td>
<td>0.23</td>
<td>10</td>
<td>Lactobacillus casei paracasei</td>
<td>2.81</td>
</tr>
<tr>
<td>60</td>
<td>Actinomyces IOO77</td>
<td>0.23</td>
<td>51</td>
<td>Lactobacillus crispatus</td>
<td>0.34</td>
</tr>
<tr>
<td>118</td>
<td>Actinomyces israeli</td>
<td>0.02</td>
<td>38</td>
<td>Lactobacillus delbrueckii</td>
<td>0.51</td>
</tr>
<tr>
<td>119</td>
<td>Actinomyces maculigenae</td>
<td>0.02</td>
<td>36</td>
<td>Lactobacillus fermentum</td>
<td>0.53</td>
</tr>
<tr>
<td>61</td>
<td>Actinomyces odoratinus</td>
<td>0.36</td>
<td>6</td>
<td>Lactobacillus gassen johnsoni</td>
<td>4.62</td>
</tr>
<tr>
<td>107</td>
<td>Actinomyces IP073</td>
<td>0.04</td>
<td>53</td>
<td>Lactobacillus oris</td>
<td>0.32</td>
</tr>
<tr>
<td>15</td>
<td>Actinomyces viscosus naestrandii</td>
<td>1.43</td>
<td>11</td>
<td>Lactobacillus parabuchneri kefri</td>
<td>0.32</td>
</tr>
<tr>
<td>120</td>
<td>Aggregatibacter segnis</td>
<td>0.02</td>
<td>42</td>
<td>Lactobacillus plantarum</td>
<td>0.45</td>
</tr>
<tr>
<td>72</td>
<td>Aino paavulum</td>
<td>0.17</td>
<td>20</td>
<td>Lactobacillus thornii murielis</td>
<td>1.19</td>
</tr>
<tr>
<td>84</td>
<td>Atopobium rimea</td>
<td>0.11</td>
<td>29</td>
<td>Lactobacillus salivarius</td>
<td>0.85</td>
</tr>
<tr>
<td>121</td>
<td>Bacteroidesides MCE3_262</td>
<td>0.02</td>
<td>47</td>
<td>Lactobacillus ultunensis</td>
<td>0.42</td>
</tr>
<tr>
<td>74</td>
<td>Bifidobacterium scardovi</td>
<td>0.15</td>
<td>14</td>
<td>Lactobacillus vaginalis</td>
<td>1.68</td>
</tr>
<tr>
<td>69</td>
<td>Campylobacter concisus</td>
<td>0.19</td>
<td>17</td>
<td>Lactobacillus murielis</td>
<td>1.19</td>
</tr>
<tr>
<td>122</td>
<td>Campylobacter curvus</td>
<td>0.02</td>
<td>40</td>
<td>Megaspheara microfolicum</td>
<td>0.49</td>
</tr>
<tr>
<td>13</td>
<td>Campylobacter gracilis</td>
<td>1.92</td>
<td>80</td>
<td>Megaspheara sp group A</td>
<td>0.13</td>
</tr>
<tr>
<td>43</td>
<td>Campylobacter showcase rectus</td>
<td>0.43</td>
<td>110</td>
<td>Megaspheara sp group B</td>
<td>0.04</td>
</tr>
<tr>
<td>28</td>
<td>Capnocytophaga gingivalis--granulosa</td>
<td>0.87</td>
<td>111</td>
<td>Mitsuokella sp</td>
<td>0.04</td>
</tr>
<tr>
<td>93</td>
<td>Capnocytophaga leadbetteri</td>
<td>0.08</td>
<td>98</td>
<td>Mitsuokella sp oral taxon 131</td>
<td>0.51</td>
</tr>
<tr>
<td>123</td>
<td>Capnocytophaga ochracea A</td>
<td>0.02</td>
<td>87</td>
<td>Mitsuokella sp oral taxon 521</td>
<td>0.11</td>
</tr>
<tr>
<td>94</td>
<td>Capnocytophaga spitzigera</td>
<td>0.08</td>
<td>89</td>
<td>Mogibacterium timidum pumilum</td>
<td>0.09</td>
</tr>
<tr>
<td>75</td>
<td>Cardiobacterium hominis</td>
<td>0.15</td>
<td>15</td>
<td>Mogibacterium timidum diversum vesuci</td>
<td>0.21</td>
</tr>
<tr>
<td>99</td>
<td>Cardiobacterium valvarum</td>
<td>0.06</td>
<td>137</td>
<td>Moraxellaceae sp</td>
<td>0.02</td>
</tr>
<tr>
<td>85</td>
<td>Catonella morti</td>
<td>0.11</td>
<td>112</td>
<td>Neisseria AP085</td>
<td>0.04</td>
</tr>
<tr>
<td>124</td>
<td>Catonella AH153</td>
<td>0.02</td>
<td>90</td>
<td>Neisseria baillfordii</td>
<td>0.09</td>
</tr>
<tr>
<td>70</td>
<td>Centipeda periodontii</td>
<td>0.19</td>
<td>41</td>
<td>Neisseria elongata</td>
<td>0.47</td>
</tr>
<tr>
<td>125</td>
<td>Corynebacterium argenticolorens</td>
<td>0.02</td>
<td>8</td>
<td>Neisseria flavica mucosa</td>
<td>3.38</td>
</tr>
<tr>
<td>54</td>
<td>Corynebacterium durnum</td>
<td>0.28</td>
<td>33</td>
<td>Neisseria flavicua</td>
<td>0.74</td>
</tr>
<tr>
<td>9</td>
<td>Corynebacterium matruchoti</td>
<td>3.04</td>
<td>33</td>
<td>Neisseria flavicua</td>
<td>0.74</td>
</tr>
<tr>
<td>126</td>
<td>Cryptobacterium curtum</td>
<td>0.02</td>
<td>138</td>
<td>NOVEL Capnocytophaga</td>
<td>0.67</td>
</tr>
<tr>
<td>44</td>
<td>Dialister invisus</td>
<td>0.43</td>
<td>43</td>
<td>SK117 1 E 052</td>
<td>0.02</td>
</tr>
<tr>
<td>127</td>
<td>Dialister microaerophilus</td>
<td>0.02</td>
<td>139</td>
<td>NOVEL Lachnospiraceae RF01</td>
<td>0.02</td>
</tr>
<tr>
<td>128</td>
<td>Dialister pneumosintes</td>
<td>0.02</td>
<td>71</td>
<td>NOVEL Selenomonas SG006</td>
<td>0.19</td>
</tr>
<tr>
<td>108</td>
<td>Dialister sp</td>
<td>0.11</td>
<td>114</td>
<td>Olsenella genomosp C1</td>
<td>0.23</td>
</tr>
<tr>
<td>62</td>
<td>Eikenella corrodens</td>
<td>0.23</td>
<td>73</td>
<td>Olsenella profusa</td>
<td>0.17</td>
</tr>
<tr>
<td>106</td>
<td>Enterococcus casseliflavus gallinarum</td>
<td>0.06</td>
<td>96</td>
<td>Onibacterium anus</td>
<td>0.08</td>
</tr>
<tr>
<td>23</td>
<td>Eubacterium sp group C</td>
<td>0.98</td>
<td>88</td>
<td>Paracardiova dentocatenis</td>
<td>0.11</td>
</tr>
<tr>
<td>129</td>
<td>Eubacterium BU014</td>
<td>0.02</td>
<td>34</td>
<td>Parvimonas micra</td>
<td>0.55</td>
</tr>
<tr>
<td>55</td>
<td>Eubacterium DOG008</td>
<td>0.28</td>
<td>113</td>
<td>Peptococcus MCE10_265</td>
<td>0.04</td>
</tr>
<tr>
<td>67</td>
<td>Eubacterium EIO74</td>
<td>0.21</td>
<td>91</td>
<td>Peptostreptococcus stomaticus</td>
<td>0.02</td>
</tr>
<tr>
<td>130</td>
<td>Eubacterium sulci infirmum</td>
<td>0.02</td>
<td>102</td>
<td>Porphyromonas catoniae</td>
<td>0.06</td>
</tr>
<tr>
<td>131</td>
<td>Eubacterium yuni subsrp</td>
<td>0.02</td>
<td>103</td>
<td>Prevotella palens</td>
<td>0.06</td>
</tr>
<tr>
<td>396</td>
<td>Eubacterium yuni subsrp, ichtima</td>
<td>0.02</td>
<td>104</td>
<td>Propionibacterium BN053</td>
<td>0.06</td>
</tr>
<tr>
<td>101</td>
<td>Eubacterium yuni subsrp yuni</td>
<td>0.06</td>
<td>64</td>
<td>Propionibacterium propionicus A</td>
<td>0.23</td>
</tr>
<tr>
<td>79</td>
<td>Firmicutes oral clone F056</td>
<td>0.13</td>
<td>140</td>
<td>Propionibacterium alcalolyticus</td>
<td>0.02</td>
</tr>
<tr>
<td>133</td>
<td>Fusobacterium nucleatum</td>
<td>0.02</td>
<td>76</td>
<td>Roffia aerea</td>
<td>0.15</td>
</tr>
<tr>
<td>30</td>
<td>Gemella haemolytica</td>
<td>0.85</td>
<td>57</td>
<td>Roffia dentocatena</td>
<td>0.28</td>
</tr>
<tr>
<td>25</td>
<td>Gemella morbilomorosa</td>
<td>0.91</td>
<td>114</td>
<td>Roffia mucilaginosa</td>
<td>0.04</td>
</tr>
<tr>
<td>109</td>
<td>Gemella sanginis</td>
<td>0.04</td>
<td>16</td>
<td>Scardovia inopinata</td>
<td>1.30</td>
</tr>
<tr>
<td>31</td>
<td>Granulicatella adelicis</td>
<td>0.83</td>
<td>141</td>
<td>Scardovia sp C1</td>
<td>0.02</td>
</tr>
<tr>
<td>134</td>
<td>Granulicatella elegans</td>
<td>0.02</td>
<td>22</td>
<td>Selenomonas AA024</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 4.1. All S-OTUs identified in severe caries of the permanent dentition, with percentage of total clones. Rank puts them in order of decreasing overall prevalence.
Figure 4.1. Distribution of S-OTUs by phylum for all samples.
Figure 4.2. PROC MIXED estimates of percentage of increase or decrease in bacterial levels as caries progresses, with 95% confidence intervals. *p* < 0.05. **p** < 0.01.
Figure 4.2
Figure 4.3. Mean levels for S-OTUs greater than 0.2% of total clones, with 95% confidence intervals. From left to right within an S-OTU, bars indicate levels for healthy control, intact enamel, white spot, cavitated, and dentin samples. S-OTUs that are green have significant negative PROC MIXED estimates, and S-OTUs that are red have significant positive PROC MIXED estimates. The upper limit for the 95% confidence interval bar for the dentin sample of Total Lactobacillus, which isn’t shown, is 58.165.
Figure 4.4. Mean Shannon Diversity Indices by sample type, with 95% confidence intervals. The paired *post hoc* tests showed that the mean for healthy controls was significantly higher than the means for white spot, cavitated, and dentin samples. There was also a significant difference between the means for intact enamel and dentin samples. The dashed line represents the significant PROC MIXED estimate, which had a value of -0.127.
Figure 4.5. Subjects identified according to their primary acid-producers. The most abundant acid-producer in each dentin sample was identified among those S-OTUs that had significant positive estimates and *S. mutans*. One subject (“Mix” above) had equal numbers of two acid-producers, *S. mutans* and Total Lactobacillus. For those subjects without a significant S-OTU greater than 2% of clones, a known acid-producer was selected from among the non-significant S-OTUs. Here they are *A. viscosus naeslundii*, *C. matruchotii*, and *S. inopinata*. 
Figure 4.6. Cumulative caries scores for subjects based on the primary acid producer identified in their samples. The mean for each group is represented by the horizontal bar through the middle of each diamond. The upper and lower points of the diamonds represent the 95% confidence intervals.
Figure 4.7. Hierarchical cluster analysis of white spot, cavitated, and dentin samples. S-OTUs that were the primary acid producers in at least one subject are in decreasing order of their PROC MIXED estimates. The color of each cell represents the percent of total clones within that sample for the S-OTU of interest, according to the scale above.
Figure 4.8. Hierarchical cluster analysis of all *Lactobacillus* species present in the study. The color of each cell represents the percent of total clones within that sample for the S-OTU of interest, according to the scale above.
Figure 4.9. Comparison of levels of lactobacilli and S. mutans. Total Lactobacillus and S. mutans percentages were compared to each other in white spot, cavitated, and dentin samples.
Figure 4.9

- White spot
- Cavitated
- Dentin
The data presented here have corroborated much of what exists in the literature about dental caries and health in children, but have also made new contributions. Using open-ended molecular techniques, each study identified new potential caries- and health-associated S-OTUs, in addition to confirming what was already known based on targeted methods. More than 90% of total clones from all three studies were cultivated species. The studies began to elucidate subsets of subjects with microbial profiles that contributed to caries, including S-OTUs that were not significantly associated with disease based on statistical analyses. In each study, diversity decreased as caries progressed, which lends support to the ecological plaque hypothesis. The ecological plaque hypothesis does not identify specific bacteria as caries pathogens, but considers caries communities to be ecological shifts from normal communities to communities dominated by any number of organisms that are capable of contributing to disease. While many studies have focused on finding the cause of caries, fewer have identified potentially healthy counterparts. Health-associated microbes may play an important role in treatment of caries in the future by being used to stabilize the normal biofilm communities. Having a complete
understanding of the complex etiology of dental caries in children will facilitate the
development of the best interventions and therapies.

Our study of incipient early childhood caries identified five S-OTUs that were
associated with caries, including *S. mutans*, and all had the ability to either produce or
utilize acid. *S. mutans* was the most abundant species in the study, and was the primary
acid-producer in most of the subjects at baseline. Interestingly, the study also identified
13 S-OTUs that were associated with health, many of which had not been previously
associated with a healthy biofilm community. More study will be required to determine
the actual contribution of these microbes to the oral community, and their potential as
therapeutic targets. This study used a longitudinal study design to follow the progression
of incipient lesions, so data was available for whether subjects’ caries became arrested or
progressed. Comparing arrested to progressed samples for levels of *S. mutans* found that
those that progressed had significantly higher levels of *S. mutans* than those that arrested.
While other S-OTUs were identified as the primary acid-producers for subjects in this
study, the importance of *S. mutans* cannot be overlooked. Once the biofilm reached a
critical level of *S. mutans*, caries progressed rapidly to cavitation.

By studying severe early childhood caries we were able to identify more S-OTUs
that were associated with caries. While incipient lesions had only five significant S-
OTUs, the severe lesions presented here contributed additional species to the caries-
associated microbiota. Severe lesions contained *Lactobacillus, Proportionibacterium,
Mitsuokella*, and *Parascardovia* species that were rarely seen in incipient lesions.
Additionally, subjects were divided by their primary acid-producer, which strengthened
the evidence for a heterogeneous etiology for childhood caries. A majority of the
subjects had high levels of *S. mutans*, but many did not have detectable levels. For other subjects their infections were dominated by *Lactobacillus* species, *S. vestibularis salivarius*, *S. sobrinus*, or mixtures of these.

In our study of severe caries and health in the young permanent dentition, *Lactobacillus* and *Propionibacterium* FMA5 were again associated with caries, and many S-OTUs were associated with health. Though *S. mutans* was not significant by PROC MIXED analysis, it was clear by examining the means for each stage of disease that it was still important in this population. Besides *Lactobacillus*- and *S. mutans*-dominated infections, subsets of subjects were also highly colonized by several S-OTUs that were not associated with caries in any study presented here, but are known to be acid-producers. One subject each had high levels of *Acinomyces viscosus naeslundii*, *Scardovia inopinata*, and *Corynebacterium matruchotii*. Though population-wide they have not been associated with caries, these may be additional targets for continued study on their contribution to dental disease. Importantly a comparison of cumulative caries scores was made among subsets of subjects with infections made up of *Lactobacillus*, *S. mutans*, *Lactobacillus* and *S. mutans*, or neither. There was no significance found in this comparison, which may mean that subjects do not have a more severe presentation of caries based on the types of bacteria present.

The importance of these studies is the continued need for a more complete understanding of the causes of caries in children, and what makes some children unsuscetible. It is only once we have this information that the ideal interventions and treatments for this serious childhood disease can be developed. Based on what has been presented here, we believe that a screening mechanism for caries can be developed, based
on the S-OTUs that have been identified here as important in the caries process. This screening would not measure levels of only one bacteria, such as *S. mutans* or *Lactobacillus*, but would look for over- or under-represented bacteria from among a group of S-OTUs that are known to contribute to health or disease. For instance, it would include all of the S-OTUs determined here to be the primary acidogens in subsets of people, since measuring only *S. mutans* would not identify patients as caries-susceptible who never had a single clone. If caries experience could be accurately predicted, then interventions and treatments could be tailored to the needs of individuals.

While these studies have filled a void in the literature for open-ended molecular investigation of caries in children, a disadvantage of the general design was the sequencing of only 50 clones per sample. With such a limited sample size, species that were present at very low levels would be unlikely to be detected, and as we saw through these investigations, it may be the species present in low levels that help to maintain microbial homeostasis in the oral cavity. It is important to identify these, including previously unidentified species, so that further investigation can determine their contributions to oral health. At the time this study was designed, 16S cloning and sequencing was an excellent option for open-ended investigation of the oral microbiome. However, new technologies are currently being developed for extremely high throughput sequencing reactions. Using 454 sequencing, each study presented here could have been completed in one sequencing reaction, resulting in hundreds or thousands of sequences per sample. In the future this technique will be used to deeply mine the human oral microbiome.
Another disadvantage was the inadequate demographic, medical, and diet data supplied by the subjects and their parents. Often questions were incompletely answered or unanswered. These data, especially that regarding fluoride exposure, environmental smoke exposure, antibiotic history, and dietary habits, were used to make comparisons between groups of subjects that had different experiences of certain caries risk factors. Unfortunately, because there is a large potential for data to be inaccurate, conclusions from these comparisons cannot be made confidently. Diet especially is known to be intimately linked to caries experience, but since questions regarding diet were open-ended, they were often answered ambiguously. Perhaps in future studies, surveys for this kind of information would best be implemented using a multiple-choice format. If open-ended responses are highly desired, then questions should be asked by a clinician with good interpersonal and interview skills, rather than left for the subject to answer independently.

Currently these studies have been analyzed independently of each other. Future analyses will combine the data from all three studies to gain an understanding of the entire natural history of caries progression. Analyzing the complete dataset will allow us to elucidate potential contributions of age and primary or permanent dentition status to caries experience. It will also allow us to determine more completely the microbial profiles that contribute to dental caries and health in childhood. Ideally, this data could be used with further studies to develop treatment strategies to be used in clinical practice. If we can use these data to determine normal levels of specific bacteria at different ages, we could develop screening criteria to predict future caries experience. The key to such a screening may be the health-associated acid-sensitive species that are lost as caries...
progresses. An element of the screening mechanism may be the detection of a loss of important bacteria that signal impending severe caries.
REFERENCES


