Published in final edited form as:

Clin Oral Investig. 2018 November ; 22(8): 2847–2858. doi:10.1007/s00784-018-2371-8.

Transmission Patterns of *Streptococcus mutans* demonstrated by a combined rep-PCR and MLST approach

Stephanie S. Momeni¹, Jennifer Whiddon², Stephen A. Moser², and Noel K. Childers^{*,1}

¹Department of Pediatric Dentistry, The University of Alabama at Birmingham, The School of Dentistry Building Room 304, 1720 2nd Avenue South, Birmingham, AL 35294-0007, USA

²Department of Pathology, The University of Alabama at Birmingham, West Pavilion Room, 1720 2nd Avenue South, Birmingham, AL 35294-0007, USA

Abstract

Objective—Clinical typing methods of the oral pathogen *Streptococcus mutans* with molecular analysis can be very specific, but expensive. Repetitive extragenic palindromic PCR (rep-PCR) is a relatively inexpensive pre-screening alternative for isolate selection for additional analyses. This study evaluated the prediction accuracy of using rep-PCR to identify *S. mutans* Multilocus Sequence Typing (MLST) sequence types (ST) among children and their family members. Potential *S. mutans* strain sources were evaluated for evidence of transmission.

Material and Methods—Ten dendrograms (rep-PCR), with 20 isolates each of the 10 most common *S. mutans* genotypes, were generated from different subjects. Using a cut-off of 98% similarity, 7–11 isolates each genotype were selected for MLST analysis to determine ST match/no-match.

Results—Overall, rep-PCR was 75% effective at determining MLST ST match/no-match and 100% effective when applied to related individuals. Most genotypes were further differentiated by MLST. MLST ST diversity was greatest for one genotype (Genotype 12, G12) and evidence of transmission among children and their family members was identified by rep-PCR and MLST. Younger children (6 months to 4 years-old) shared ST with their mothers but 50% of older children (5–9 years-old) had ST not identified in their mother. Six ST were shared between different families and probable source members were identified.

Conclusion—This study confirms that rep-PCR offers an affordable option to predict diverse isolates for downstream applications.

Clinical Relevance—Using a combined rep-PCR and MLST approach, it is possible to track probable transmission and strain sources for *S. mutans* genotypes.

*Corresponding Author: Dr. Noel K. Childers, Department of Pediatric Dentistry, University of Alabama at Birmingham, School of Dentistry Room 304, 1530 3rd Avenue South, Birmingham, AL 35294-0007, USA Tel: (+1) 205-934-3230; Fax: (+1) 205-934-7013; nkc@uab.edu.

Compliance with Ethical Standards

Conflicts of Interest: The authors declare they have no conflicts of interest.

Ethical Approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Keywords

genetic diversity; genotyping; multilocus sequence typing; *Streptococcus mutans*; transmission

Introduction

Dental caries is a prevalent global chronic disease affecting between 28 and 82% of children worldwide and is the most common chronic disease among children and adolescents [1,2]. Although many organisms are suspected in the progression of dental caries, the mutans streptococci (*Streptococcus mutans* and *Streptococcus sobrinus*) are the primary bacteria most frequently associated with dental caries initiation. A fundamental objective of the epidemiology of infectious disease is the ability to determine sources of the infection. Studies evaluating the bacterial genotypic diversity of the mutans streptococci, in particular *S. mutans*, are important for the improvement of caries risk assessments and treatment options.

Bacterial genotyping is the foundation of investigations of bacterial strain genetic diversity, microevolution, and epidemiological surveillance studies. Traditionally, genotyping of *S. mutans* has been performed using gel-based typing methods including pulse field gel electrophoresis (PFGE), arbitrarily primed PCR (AP-PCR), and repetitive extragenic palindromic PCR (rep-PCR) [3–8]. Current advances in technology are increasing the application of molecular based typing applications such as multilocus sequence typing (MLST) and whole genome sequencing (WGS) for *S. mutans* strain comparisons [9–13]. It is recommended that gel-based typing methods (e.g. rep-PCR) be confirmed using molecular based typing methods because different bacterial typing methods can yield different results [14,15]. The use of a secondary analysis ensures robustness of the methodology and data generated (e.g., validating lineage assignment). Although costs for molecular based methods continue to decrease, these methods can be significantly more expensive and labor intensive than traditional gel-based typing methods making them less practical for large-scale epidemiological studies as well as less financially feasible. Alternately, it may be more economical to pre-screen isolates prior to employing more expensive testing approaches when appropriate to the research question. For instance, initial analysis using rep-PCR can be utilized to determine strain diversity, revealing isolates that are far more likely to provide diverse data for a comparative genomics study [16,17,15].

The DiversiLab System is an economical option for the investigation of genetic diversity of bacterial isolates using a standardized rep-PCR approach [18]. After PCR, micro fluidics chip technology allows the samples to be visualized using dendrograms, percent similarity and graphic overlays. This method is rapid and economical for high-throughput epidemiological studies. It has been widely used for epidemiological surveillance of clinically relevant bacteria but its application for typing oral bacteria is limited and unique to our laboratory [19–23,4,24].

MLST is a molecular genotyping method that characterizes bacteria based on changes in the molecular sequence of select conserved housekeeping gene loci [25]. This method is considered more discriminative and reproducible than gel based methods because

interpretation of gel based typing approaches are typically based on similarity where as with MLST a single nucleotide change can result in a new sequence type (ST). Processing and data analysis for MLST analysis is more time consuming and more expensive than for rep-PCR.

Studies investigating the genotypic diversity of *S. mutans* frequently include transmission analysis. The current literature offers several possible sources of transmission. Mother-to-child transmission of *S. mutans* has been described as the primary route of transmission [26–28,3,29–31]. Other studies have suggested that *S. mutans* strains are often shared with other household family members including fathers and siblings indicating the potential for intra-familial transmission [29,32,28,33]. Furthermore, extra-familial sharing of *S. mutans* genotypes has been reported for unrelated children within nurseries or school classes [34,35,7,36,24]. There are several limitations to some of these studies, which include using only one genotyping approach (typically gel based); limited numbers of bacterial isolates (small sample sizes, thus decreased statistical power); and a limited focus (e.g., many studies focus only on mother-to-child without looking at other potential sources).

We have reported that rep-PCR is beneficial for large-scale epidemiological studies of *S. mutans*, [4,24] and we predict that it may also effectively be used to determine which isolates are most likely to yield diverse data from additional, more specific, but also more expensive techniques such as MLST or whole genome sequencing. The current study had two aims. The first aim was to determine if rep-PCR could be used to predict *S. mutans* MLST sequence types. The second aim was to assess the practical application of this approach for a specific research question using rep-PCR to predict MLST sequence types for an *S. mutans* rep-PCR genotype (genotype G12) to determine likely transmission of this genotype among and between children and their family members.

Material and Methods

Sample Selection

Isolates used in this study were collected as part of an 8-year longitudinal clinical epidemiological study of *S. mutans* in rural Uniontown, Alabama, United States. This population is primarily African-American, of low socioeconomic status with limited access to dental care. This University of Alabama at Birmingham Institutional Review Board approved study used informed consent from the parent/guardian and wavier of assent due to the young age of children. *S. mutans* isolates used in this study were from 2 cohorts of children and their family members. Cohort 1 (CH1) consisted of school aged children (age 5–6 years at initial recruitment, family ID# less than 300) and Cohort 2 (CH2) were infants (age 12–18 months at initial collection, family ID# 500 and over). Specifics on sample selection for each of the two-part study design are described in more detail below.

Repetitive extragenic palindromic PCR (rep-PCR)

Rep-PCR was performed as part of the larger Uniontown epidemiological study as previously described [4]. Briefly, whole saliva, plaque and/or tongue samples were collected from children and their family members over an 8-year period. Presumptive *S. mutans*

colonies were selected based on morphological appearance from Gold's media [37]. Strains were pure cultured using Todd Hewitt Broth (Becton Dickinson, Sparks, MD) and DNA extracted using the UltraClean microbial DNA isolation kit (bioMérieux, Inc., Durham, NC). Isolates were confirmed as *S. mutans* with SYBR Green PCR using *gfbB*-specific primers [38]. Quality of DNA was assessed (260/280nM optical density) and rep-PCR performed using the DiversiLab system with the *Streptococcus* DNA fingerprinting kit. Amplicons were evaluated using an Agilent 2100 Bio-analyzer (Agilent Technologies, Santa Clara, CA) with microfluidics LabChip (bioMérieux, Inc.). DiversiLab web-based software was used to determine genotypes using dendrograms and electropherogram overlays. An isolate was classified as a new genotype when an isolate differed by 1 major band or 3 minor bands. Major and minor bands were determined based on a 100-fluorescence units cut-off.

Genotypes were previously determined by rep-PCR for ~13,000 *S. mutans* isolates as part of the larger epidemiological study using methods described elsewhere [4]. For the current study, all available isolates for each of the 10 most frequently reported rep-PCR genotypes from children were compiled into a list and randomized in Excel™ using the Rand function. From the randomized list, the first 20 isolates from different children for which sufficient DNA and a strong rep-PCR pattern were available were selected to generate a rep-PCR dendrogram. In rare cases when plaque isolates were not available from 20 children, saliva or tongue isolates were used. In this study, samples types were assumed to be comparable when classified by the same rep-PCR genotype. A total of 200 isolates (19 to 20 children per genotype) were evaluated in the 10 rep-PCR dendrograms representing 108 children. From the 200 isolates in the rep-PCR dendrograms, 97 were selected for MLST analysis.

Each dendrogram consists of a percent similarity tree, sample identification, and virtual gel images. Using an arbitrarily assigned 98% percent similarity cut-off, isolates were predicted to either have a matching or non-matching sequence type. This cut-off was selected based on previous observations from studies in our laboratory using MLST indicating that isolates with <98% percent similarity tended to have different MLST sequence types (data not shown). Therefore, if the node connecting the two isolates was less than 98% similar (major branch), then it was predicted that the two isolates would have a non-matching sequence type. If the connecting node had a greater than 98% similarity (minor branch), then it was predicted that the two isolates would have a matching sequence type. This process was initially performed using two isolates selected from the most extreme branches of the dendrogram, then repeated for 7–9 isolates for various internal branches (i.e., branches within the tree). An overview of the sample analysis is provided in Figure 1. Figure 2 provides an example of this analysis. Prediction accuracy was calculated as the number of correctly predicted sequence types (as either a match or no match) divided by the total sequence types evaluated.

To assess the practical application of this approach to identify evidence of transmission, a pool of all index children and their mothers and fathers with a specific rep-PCR genotype (G12) were selected for analysis. Genotype G12 was selected based on a previous observation in our laboratory that this common genotype group contained the most diverse sequence types by MLST analysis [17,16]. A total of 9 mother-child pairs (M-C), 1 father-child pair (F-C), and 14 children only (i.e., parents did not have G12) groups were used to

generate rep-PCR dendrograms. Dendrograms were generated for all G12 isolates for each parent-child group or for each individual child (when the parents did not have this genotype). For children with less than twenty G12 isolates available, only 1 isolate per child was evaluated. For reports with 20 or more isolates, multiple isolates were selected for analysis consistent with the major branches observed. A total of 700 isolates were evaluated in the rep-PCR dendrograms representing 24 children, 9 mothers and 1 father. Sequence type matches were predicted as described earlier and a total of 75 *S. mutans* isolates were selected from the reports for MLST analysis.

Multilocus Sequence Typing

MLST was performed as previously described [16]. Briefly, primers described by Nakano *et al.* for 8 housekeeping partial gene fragments from *murI*, *tkt*, *glnA*, *aroE*, *gyrA*, *glk*, *glt*, and *lepC* were amplified using a SYBR green real-time PCR approach [11]. Amplicons were purified and sequenced. Manual alignments were performed using CLC DNA Workbench version 5.7.1 (CLC bio USA, Cambridge, MA) to obtain consensus sequences that were cross-referenced with the oral streptococci www.pubmlst.org database for allele and sequence type assignments. New alleles and sequence types were confirmed and added to the PubMLST database (www.PubMLST.org). New alleles were submitted to Genbank.

Following MLST analysis, sequence type data was added to the rep-PCR dendrograms to determine the prediction accuracy of rep-PCR percent similarity trees in predicted MLST sequence type (Fig. 2). First the sequence types of the two extreme isolates were compared. Isolates with the same sequence type were considered a match while isolates with different sequence types were considered a no match. Next, internal branches were evaluated. This approach focused on prediction accuracy between the selected isolates but not between different groups. The number of instances that the prediction agreed or disagreed was recorded to determine prediction accuracy. Given that rep-PCR is based on similarity, further consideration was given to isolates that differed by either a single- or double-locus variants (SLV and DLV, respectively) because up to two allelic differences are considered clonal complexes (i.e., recently diverged, similar strains) [25,11].

Transmission

A subset of the rep-PCR G12 clinical isolates were used to investigate evidence for transmission patterns based on MLST sequence types. The commonality of *S. mutans* strains by MLST sequence types for G12 isolates for mother-child, father-child and children only (children with no parent with G12 genotype) were assessed to track transmission. Individuals sharing the G12 rep-PCR genotype and the same MLST sequence type were considered as mother-to-child transmission, father-to-child or extra-familial transmission (between unrelated children). Subsequently, it was possible to identify potential sources because a number of strains shared among different children appeared to be shared with only one adult in the study. To further provide evidence for the source, a subsequent extended analysis of other household family member was performed. Other family members with the G12 genotype were analyzed with MLST and the sequence types compared. In cases where the child shared the sequence type with the mother, siblings or cousins with the G12 genotype were typically excluded from the analysis because regardless of the sequence type

of the sibling, they would be considered as unlikely sources since the mother already had the shared sequence type. Siblings and cousins were evaluated for cases where the mother and child had different sequence types.

Statistics

A Fisher's exact test was performed to evaluate the ability to correctly predict match or no-match sequence types from the rep-PCR dendrograms (i.e., sensitivity/specificity and predictive values). This analysis included both extreme and internal branches for the top 10 genotypes.

Results

Overall, this study identified six new alleles and 7 new ST were added to the PubMLST Database (Online Resource 1: Table S1). Genbank accession numbers in the range of KX383950-52 and KX852457-59 were assigned for the six new alleles. Summary demographic data for number of individuals and isolates evaluated are listed in Table 1. Rep-PCR dendrograms used in the initial analysis are available in Online Resource 2: Figures 1–10.

Prediction Accuracy

In the initial analysis of the study, using rep-PCR of the 10 most common genotypes to predict MLST sequence types from the two extreme branches resulted in 2 genotypes (G18 and G09) that were connected by branches with greater than 98% (predicted match) while 8 were non-matching (Table 2). Based on this, it was predicted that any internally selected isolates would match the sequence type for the two extreme branch isolates for G18 and G09. While G18 did match this prediction, G09 has other sequence types that were different, but only by SLV and DLV. Thus, using extreme isolates resulted in 90% accuracy in predicting if internal branch sequence types would or would not match.

Analysis of 97 isolates on 52 branch groups (5–7 groups per genotype) resulted in 63% (33/52) prediction accuracy (Table 2). Allowing for SLV and DLV increased the accuracy of prediction to 75% (39/52). Statistical analysis of 62 groups (52 internal and 10 extreme branches combinations) indicated there was a significant predictive relationship using rep-PCR for identifying matching vs. non-matching MLST sequence types (Fisher's exact test, $p < 0.001$) both with and without SLV and DLV included in the analysis (Table 3). The negative predictive value remained the same when SLV and DLV were considered; however the positive predictive value increased to 70%.

Transmission

When evaluating G12 isolates for prediction accuracy, M-C and F-C pairs were evaluated separately from the children only group (Table 4). Statistical analysis was not performed on the data from this portion of the study due to small sample size and because the data would be biased since isolates from children and their parents were more likely to have the same sequence type.

Table 5 summarizes the isolate and MLST data for rep-PCR genotype G12. For the 9 mother-child (M-C) pairs, 6 children shared the same rep-PCR GT and MLST ST with the mother; 3 children had different MLST sequence types from their mothers, although these children shared the rep-PCR genotype. Sequence types were consistent within individuals except for 170-C which had 2 sequence types. The one father-child (F-C) pair did not share any MLST sequence types. For the 14 children only (no parents with G12, indicated in Table 5 section B with a “C” prefix), sequence types were consistent within children except for one child that had two sequence types. Three sequence types were shared among multiple individuals within different families. ST119 (Thailand), ST132 (Thailand), and ST92 (Japan) were isolates matching strains identified in published studies from other countries, while all other isolates were unique to the Uniontown population (United States) [11,39,16,40,17].

Overall for the extended family analysis (Online Resource 1: Table S2), 96 household family members (including mother and father) were available to assess for G12. Of these, only 24 (27.6%) had a rep-PCR genotype G12. Eight of the 9 M-C pairs had at least one additional household family member with G12. MLST analysis was performed for 7 additional family members from 6 families; however, only 3 households (107, 180, and 596) had another adult (assumed a potential source) besides the mother with a G12. In the 14 children only (no parents with G12), only two families had additional family members with G12 (the mean number of family members available was 4). MLST analysis for 5 additional family members in these two families (siblings or cousin) found the same MLST sequence type as the child. Three of the children did not have a mother participating in the study for evaluation while fathers were not available or participated less often in the study overall.

Discussion

Bacterial genotyping is a fundamental instrument in the study of transmission and reservoir determination for epidemiological studies of infectious diseases. Molecular genotyping methods are considered the gold standard for strain typing, but may not be economically feasible for large-scale study [41]. In the present study, the use of rep-PCR dendrograms as a screening tool to predict clinical *S. mutans* MLST sequence type diversity is evaluated. Successful application of this approach would be cost-beneficial for pre-screening isolates for more expensive molecular applications such as whole genome sequencing. Furthermore, this study investigates the combined use of MLST and rep-PCR to evaluate evidence for transmission of *S. mutans*.

Initial analysis using only the two extreme isolates from a dendrogram was 90% accurate in predicting which genotypes would benefit from further analysis (Table 2). This was expected as unrelated individuals were expected to be more likely to have different sequence types. It also supports the need to analyze additional isolates within the dendrogram using an alternate typing method (MLST in this study). For the M-C and F-C pairs, as well as the individual children, evaluated in the second part of the study (i.e., G12), the prediction accuracy of the two extreme isolates was much lower at 40% and 20% respectively. For M-C and individual isolates it is expected that these isolates are more likely to have the same sequence type within a M-C pair or individual. The observed results were low because of

low specificity; the extreme isolates often predicted no match when analysis revealed a match.

For isolates from unrelated subjects, the overall prediction accuracy, when considering both extreme and internal branches, was only 63% and the specificity was 53%, which was lower than anticipated (Table 2, Table 3). However, it is important to note the rep-PCR branching is based largely on percent similarity and there can be minor variations of isolates classified as the same rep-PCR genotype [4]. In contrast, MLST is more sensitive and discriminative because a single base pair change results in a new MLST sequence type. Thus, the analysis of MLST sequence types was expanded to include SLV and DLV so as to be more comparable to the rep-PCR approach. This modification resulted in an increase to 75% prediction accuracy; however, even when SLV and DLV were considered specificity remained low at 63% (Table 2, Table 3). These findings suggest that rep-PCR is very effective at predicting MLST sequence types that will not match. Furthermore, it suggests that when isolates shared a minor branch (>98% similarity) and were predicted to match, the isolates were sometimes further differentiated by MLST.

Based upon the analysis conducted in this study, general guidelines for using rep-PCR to predict MLST sequence type can be made. For a rep-PCR single genotype dendrogram, if isolates are from different individuals and have less than 98% similarity, the current study demonstrates a high possibility these isolates are different and therefore additional testing may be unnecessary. For a rep-PCR single genotype dendrogram with isolates from the same individual and have less than 98% similarity, the data indicates a high probability the isolates are the same, therefore additional testing may only be needed for validation. If a rep-PCR single genotype dendrogram of isolates from different individuals indicates greater than 98% similarity, then the current study suggests these isolates would most likely have the same sequence type; however, additional analysis of select isolates should be performed for validation.

Although most rep-PCR genotypes were further differentiated by MLST, genotype G12 was the most diverse of those analyzed previously in our laboratory. Therefore, we took advantage of this genotype to focus on transmission analysis. Using extreme branches to predict internal branch ST matches for these isolates resulted in a prediction accuracy for G12 M-C pairs of 90% and 75% for G12 individual children (Table 4). It is noteworthy that for the individual children, the reduced predictability was due to false no matches (cases where the sequence type were predicted to not-match but actually matched).

Transmission of rep-PCR genotype G12 was evaluated using the novel combined rep-PCR and MLST approach (Table 4). Few studies have used combined genotyping approaches (e.g. AP-PCR, RFLP, chromosomal DNA fingerprinting) to evaluate genetic diversity and transmission of *S. mutans* in children [36,35,42]. The benefit of using a combined genotyping approach is to confirm strain similarity or diversity by evaluating different portions of the genome. The outcome of a combined analysis is a more robust (supportive, although not conclusive) analysis of genetic diversity and transmission.

Overall, transmission analysis included 24 children and 22 of their family members resulting in 14 different MLST sequence types for the rep-PCR genotype G12 (Table 5, Online Resource 1: Table S2). It was observed that for the 15 individuals with more than one isolate evaluated, only 2 had more than one sequence type (13%, 2/15), indicating that sequence types are typically stable, at least within these individuals, for G12 isolates. Lapirattanakul et al. reported comparable results (10%, 2/20) using MLST to evaluate mother-child pairs [10].

For M-C pairs including infants, sequence types were the same (ID 518, 596, 600). In contrast, M-C pairs for older, school-aged children had 3 pairs for which the child has sequence types that were different from the mother (ID 170, 223, 227). This finding suggests that younger children are more likely to share strains with their mother, whereas older children are more likely to have MLST sequence types that are not shared with their mothers. This finding differs from that reported by Lapirattanakul et al. who found that 75% (9/12) of older children (i.e., 5 years or older) shared a sequence type with their mothers [10]. However, further study with a larger sample size is needed to confirm the present results

Fourteen children had no rep-PCR G12 match with a parent. However, some of these children did have a sequence type in common with some of the M-C or F-C pairs. ST132, ST157, and ST202 were most commonly shared, indicating these strains may be more transmissible. Although it is very difficult to prove transmission conclusively, it appears that the mother from family 180 (180-M) may be the initial source for strain ST157. Although the mother from family 227 (227-M) also shares this sequence type, her child does not carry the sequence type whereas both 180 M-C carry the strain. Thus, this mother and her child seem the most likely candidates for sharing the strain with five other children (159-C, 223-C, 597-C, 608-C, 226-C) in this population. The same conclusion may be drawn for ST202, for which 182 M-C share the sequence type. No other mothers carry this strain yet 4 other children have this strain (170-C, 227-C, 218-C, 509-C). Based on this data, it is possible to track transmission and the potential source of commonly shared *S. mutans* strains using rep-PCR to initially screen isolates for further differentiation with MLST. However, this analysis would benefit from information on inter-family interactions, which were not included in the present study.

Evidence supporting the identified sources is further strengthened through the overall family analysis (Online Resource 1: Table S2). Subsequent MLST analysis for 22 of the family members confirmed the sources noted above and eliminated the other participating adult household members as potential sources. Overall, 6 children shared a sequence type with another family member of which five children shared a sequence type with another child in the family (sibling, cousin, other). This is evidence consistent with child-to-child intra-familial transmission of strains and further supports previous findings reported by our laboratory [24].

Some limitations should be noted. For the prediction accuracy, the data in this study suggests that individuals within the same family are more likely to have the same sequence types so it is probable that the higher prediction accuracies observed in the second part of the study are partly biased. For both parts of the study, not all isolates on a rep-PCR dendrogram were

evaluated so it is possible that these results could have varied if all isolates had been tested. For transmission, in rare instances, not all household members chose to participate in the study and thus other sources not evaluated here may be possible. Also, other individuals in the larger Uniontown Study had G12 strains but were not evaluated in this study since these individuals did not reside with children in this study and any possible interactions were unknown. Furthermore, the combined rep-PCR and MLST approach described may not be applicable to all *S. mutans* strains, such as rep-PCR G18, which consistently have the same rep-PCR and MLST sequence type. Therefore, for some genotypes, alternate approaches would be required.

This study supports the use of rep-PCR to reasonably predict strain diversity for downstream application such as MLST and WGS, providing a more cost effective approach to large-scale epidemiological studies. As regards transmission, some strains of *S. mutans* genotype G12 are commonly shared while other are family specific. Children may acquire additional sequence types from sources, other than their mothers, as they get older. Using a combined rep-PCR and MLST approach, it is possible to track possible transmission and identify potential sources for some strains of *S. mutans*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We especially appreciate all the clinical and laboratory participants of this study; Dr. Mei Han of the UAB Hefflin Genomic Core provided sequencing data; and Dr. Jinthana Lapirattanakul directed the confirmation and registration of isolates with PubMLST.

Funding: This study was supported by National Institute of Dental and Craniofacial Research (NIDCR) Grant DE016684 (Noel Childers). Dr. Momeni is a Dental Academic Research Training (DART) Post-doctoral Fellow under NIDCR Institutional Grant#T-90 DE022736.

References

1. Leong PM, Gussy MG, Barrow SY, de Silva-Sanigorski A, Waters E. A systematic review of risk factors during first year of life for early childhood caries. *Int J Paediatr Dent.* 2013; 23(4):235–250. DOI: 10.1111/j.1365-263X.2012.01260.x [PubMed: 22925469]
2. Global regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet.* 2016; 388(10053):1545–1602. DOI: 10.1016/S0140-6736(16)31678-6 [PubMed: 27733282]
3. Li Y, Caufield PW. The fidelity of initial acquisition of mutans streptococci by infants from their mothers. *J Dent Res.* 1995; 74(2):681–685. [PubMed: 7722065]
4. Moser SA, Mitchell SC, Ruby JD, Momeni S, Osgood RC, Whiddon J, Childers NK. Repetitive extragenic palindromic PCR for study of *Streptococcus mutans* diversity and transmission in human populations. *J Clin Microbiol.* 2010; 48(2):599–602. JCM.01828-09 [pii]. DOI: 10.1128/JCM.01828-09 [PubMed: 20007400]
5. Lynch DJ, Villhauer AL, Warren JJ, Marshall TA, Dawson DV, Blanchette DR, Phipps KR, Starr DE, Drake DR. Genotypic characterization of initial acquisition of *Streptococcus mutans* in American Indian children. *J Oral Microbiol.* 2015; 7:27182.doi: 10.3402/jom.v7.27182 [PubMed: 25840611]

6. Mitchell SC, Ruby JD, Moser S, Momeni S, Smith A, Osgood R, Litaker M, Childers N. Maternal transmission of mutans *Streptococci* in severe-early childhood caries. *Pediatr Dent*. 2009; 31(3): 193–201. [PubMed: 19552223]
7. Baca P, Castillo AM, Liebana MJ, Castillo F, Martin-Platero A, Liebana J. Horizontal transmission of *Streptococcus mutans* in schoolchildren. *Med Oral Patol Oral Cir Bucal*. 2012; 17(3):e495–500. [PubMed: 22143733]
8. Gilbert K, Joseph R, Vo A, Patel T, Chaudhry S, Nguyen U, Trevor A, Robinson E, Campbell M, McLennan J, Houran F, Wong T, Flann K, Wages M, Palmer EA, Peterson J, Engle J, Maier T, Machida CA. Children with severe early childhood caries: streptococci genetic strains within carious and white spot lesions. *J Oral Microbiol*. 2014; :6.doi: 10.3402/jom.v6.25805
9. Do T, Gilbert SC, Clark D, Ali F, Fatturi Parolo CC, Maltz M, Russell RR, Holbrook P, Wade WG, Beighton D. Generation of diversity in *Streptococcus mutans* genes demonstrated by MLST. *PLoS One*. 2010; 5(2):e9073.doi: 10.1371/journal.pone.0009073 [PubMed: 20140210]
10. Lapidattanakul J, Nakano K, Nomura R, Hamada S, Nakagawa I, Ooshima T. Demonstration of mother-to-child transmission of *Streptococcus mutans* using multilocus sequence typing. *Caries Res*. 2008; 42(6):466–474. 000170588 [pii]. DOI: 10.1159/000170588 [PubMed: 18997467]
11. Nakano K, Lapidattanakul J, Nomura R, Nemoto H, Alaluusua S, Gronroos L, Vaara M, Hamada S, Ooshima T, Nakagawa I. *Streptococcus mutans* clonal variation revealed by multilocus sequence typing. *J Clin Microbiol*. 2007; 45(8):2616–2625. JCM.02343-06 [pii]. DOI: 10.1128/JCM.02343-06 [PubMed: 17567784]
12. Song L, Wang W, Conrads G, Rheinberg A, Sztajer H, Reck M, Wagner-Dobler I, Zeng AP. Genetic variability of mutans streptococci revealed by wide whole-genome sequencing. *BMC Genomics*. 2013; 14:430.doi: 10.1186/1471-2164-14-430 [PubMed: 23805886]
13. Cornejo OE, Lefebure T, Pavinski Bitar PD, Lang P, Richards VP, Eilertson K, Do T, Beighton D, Zeng L, Ahn SJ, Burne RA, Siepel A, Bustamante CD, Stanhope MJ. Evolutionary and Population Genomics of the Cavity Causing Bacteria *Streptococcus mutans*. *Mol Biol Evol*. 2013; 30(4):881–893. DOI: 10.1093/molbev/mss278 [PubMed: 23228887]
14. Foley SL, White DG, McDermott PF, Walker RD, Rhodes B, Fedorka-Cray PJ, Simjee S, Zhao S. Comparison of subtyping methods for differentiating *Salmonella enterica* serovar Typhimurium isolates obtained from food animal sources. *J Clin Microbiol*. 2006; 44(10):3569–3577. 44/10/3569 [pii]. DOI: 10.1128/JCM.00745-06 [PubMed: 17021084]
15. van Belkum A, Tassios PT, Dijkshoorn L, Haeggman S, Cookson B, Fry NK, Fussing V, Green J, Feil E, Gerner-Smidt P, Brisse S, Struelens M. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clin Microbiol Infect*. 2007; 13(Suppl 3):1–46. DOI: 10.1111/j.1469-0691.2007.01786.x
16. Momeni SS, Whiddon J, Moser SA, Cheon K, Ruby JD, Childers NK. Comparative genotyping of *Streptococcus mutans* by repetitive extragenic palindromic polymerase chain reaction and multilocus sequence typing. *Mol Oral Microbiol*. 2013; 28(1):18–27. DOI: 10.1111/omi.12002 [PubMed: 23194334]
17. Momeni SS, Whiddon J, Cheon K, Moser SA, Childers NK. Assessment of two multilocus sequence typing (MLST) schemes available for *Streptococcus mutans*. *Arch Oral Biol*. 2015; 60(12):1769–1776. DOI: 10.1016/j.archoralbio.2015.09.012 [PubMed: 26439181]
18. Healy M, Huong J, Bittner T, Lising M, Frye S, Raza S, Schrock R, Manry J, Renwick A, Nieto R, Woods C, Versalovic J, Lupski JR. Microbial DNA typing by automated repetitive-sequence-based PCR. *J Clin Microbiol*. 2005; 43(1):199–207. 43/1/199 [pii]. DOI: 10.1128/JCM.43.1.199-207.2005 [PubMed: 15634972]
19. Kim GH, Breidt F, Fratafico P, Oh DH. Acid Resistance and Molecular Characterization of *Escherichia coli* O157:H7 and Different Non-O157 Shiga Toxin-Producing *E. coli* Serogroups. *J Food Sci*. 2015; 80(10):M2257–2264. DOI: 10.1111/1750-3841.12996 [PubMed: 26375176]
20. Mutlu Sariguzel F, Berk E, Koc AN, Sav H, Demir G. Investigation of the relationship between virulence factors and genotype of *Candida* spp. isolated from blood cultures. *J Infect Dev Ctries*. 2015; 9(8):857–864. DOI: 10.3855/jidc.5359 [PubMed: 26322878]
21. Werner G, Fleige C, Neumann B, Bender JK, Layer F, Klare I. Evaluation of DiversiLab(R), MLST and PFGE typing for discriminating clinical *Enterococcus faecium* isolates. *J Microbiol Methods*. 2015; 118:81–84. DOI: 10.1016/j.mimet.2015.08.019 [PubMed: 26341611]

22. Ford CD, Lopansri BK, Haydoura S, Snow G, Dascomb KK, Asch J, Bo Petersen F, Burke JP. Frequency, risk factors, and outcomes of vancomycin-resistant *Enterococcus* colonization and infection in patients with newly diagnosed acute leukemia: different patterns in patients with acute myelogenous and acute lymphoblastic leukemia. *Infect Control Hosp Epidemiol*. 2015; 36(1):47–53. DOI: 10.1017/ice.2014.3 [PubMed: 25627761]
23. Aguadero V, Gonzalez Velasco C, Vindel A, Gonzalez Velasco M, Moreno JJ. Evaluation of rep-PCR/DiversiLab versus PFGE and spa typing in genotyping methicillin-resistant *Staphylococcus aureus* (MRSA). *Br J Biomed Sci*. 2015; 72(3):120–127. [PubMed: 26510268]
24. Momeni SS, Whiddon J, Cheon K, Ghazal T, Moser SA, Childers NK. Genetic Diversity and Evidence for Transmission of *Streptococcus mutans* by DiversiLab rep-PCR. *J Microbiol Methods*. 2016; 128:108–117. DOI: 10.1016/j.mimet.2016.07.010 [PubMed: 27432341]
25. Maiden MC. Multilocus sequence typing of bacteria. *Annu Rev Microbiol*. 2006; 60:561–588. DOI: 10.1146/annurev.micro.59.030804.121325 [PubMed: 16774461]
26. Teanpaisan R, Chaethong W, Piwat S, Thitasomakul S. Vertical transmission of mutans streptococci and lactobacillus in Thai families. *Pediatr Dent*. 2012; 34(2):e24–29. [PubMed: 22583873]
27. Klein MI, Florio FM, Pereira AC, Hofling JF, Goncalves RB. Longitudinal study of transmission, diversity, and stability of *Streptococcus mutans* and *Streptococcus sobrinus* genotypes in Brazilian nursery children. *J Clin Microbiol*. 2004; 42(10):4620–4626. DOI: 10.1128/JCM.42.10.4620-4626.2004 [PubMed: 15472319]
28. Kozai K, Nakayama R, Tedjosongko U, Kuwahara S, Suzuki J, Okada M, Nagasaka N. Intrafamilial distribution of mutans streptococci in Japanese families and possibility of father-to-child transmission. *Microbiol Immunol*. 1999; 43(2):99–106. [PubMed: 10229263]
29. Douglass JM, Li Y, Tinanoff N. Association of mutans streptococci between caregivers and their children. *Pediatr Dent*. 2008; 30(5):375–387. [PubMed: 18942596]
30. de da Silva Bastos VA, Freitas-Fernandes LB, Fidalgo TK, Martins C, Mattos CT, de Souza IP, Maia LC. Mother-to-child transmission of *Streptococcus mutans*: a systematic review and meta-analysis. *J Dent*. 2015; 43(2):181–191. DOI: 10.1016/j.jdent.2014.12.001 [PubMed: 25486222]
31. Binks C, Duane B. Mother-to-child transmission of *Streptococcus mutans*. *Evid Based Dent*. 2015; 16(2):39–40. DOI: 10.1038/sj.ebd.6401089 [PubMed: 26114782]
32. Tedjosongko U, Kozai K. Initial acquisition and transmission of mutans streptococci in children at day nursery. *ASDC J Dent Child*. 2002; 69(3):284–288. 234–285. [PubMed: 12613313]
33. Redmo Emanuelsson IM, Wang XM. Demonstration of identical strains of mutans streptococci within Chinese families by genotyping. *Eur J Oral Sci*. 1998; 106(3):788–794. [PubMed: 9672101]
34. Domejean S, Zhan L, DenBesten PK, Stamper J, Boyce WT, Featherstone JD. Horizontal transmission of mutans streptococci in children. *J Dent Res*. 2010; 89(1):51–55. DOI: 10.1177/0022034509353400 [PubMed: 19918090]
35. Liu Y, Zou J, Shang R, Zhou XD. Genotypic diversity of *Streptococcus mutans* in 3- to 4-year-old Chinese nursery children suggests horizontal transmission. *Arch Oral Biol*. 2007; 52(9):876–881. DOI: 10.1016/j.archoralbio.2007.03.004 [PubMed: 17466259]
36. Alves AC, Nogueira RD, Stipp RN, Pampolini F, Moraes AB, Goncalves RB, Hofling JF, Li Y, Mattos-Graner RO. Prospective study of potential sources of *Streptococcus mutans* transmission in nursery school children. *J Med Microbiol*. 2009; 58(Pt 4):476–481. DOI: 10.1099/jmm.0.005777-0 [PubMed: 19273644]
37. Gold OG, Jordan HV, Van Houte J. A selective medium for *Streptococcus mutans*. *Arch Oral Biol*. 1973; 18(11):1357–1364. [PubMed: 4518755]
38. Yoshida A, Suzuki N, Nakano Y, Kawada M, Oho T, Koga T. Development of a 5' nuclease-based real-time PCR assay for quantitative detection of cariogenic dental pathogens *Streptococcus mutans* and *Streptococcus sobrinus*. *J Clin Microbiol*. 2003; 41(9):4438–4441. [PubMed: 12958287]
39. Lapidattanakul J, Nakano K, Nomura R, Leelataweewud P, Chalermarp N, Klaophimai A, Srisatjaluk R, Hamada S, Ooshima T. Multilocus sequence typing analysis of *Streptococcus*

- mutans* strains with the *cnm* gene encoding collagen-binding adhesin. J Med Microbiol. 2011; 60(Pt 11):1677–1684. DOI: 10.1099/jmm.0.033415-0 [PubMed: 21680768]
40. Momeni SS, Whiddon J, Cheon K, Moser SA, Childers NK. Assessment of clonality and serotypes of *Streptococcus mutans* among children by multilocus sequence typing. Eur J Oral Sci. 2015; 123(6):416–424. DOI: 10.1111/eos.12221 [PubMed: 26443288]
41. Foxman B. Molecular tools and infectious disease epidemiology. Elsevier Academic Press; 2012.
42. Li Y, Caufield PW. Arbitrarily primed polymerase chain reaction fingerprinting for the genotypic identification of *mutans* streptococci from humans. Oral Microbiol Immunol. 1998; 13(1):17–22. [PubMed: 9573817]

Figure 1.

Summary of Approach for Prediction of MLST Sequence Types from a Rep-PCR Dendrogram based on Percent Similarity for *S. mutans* Isolates for Different Children. Approach was applied to the 10 most frequently occurring rep-PCR genotypes using 20 *S. mutans* isolates from 20 different children. ^a For isolates on equidistant branches occurring on the most distant branches, either may be selected (i.e. branches can rotate). ^b Internal branches may include the two extreme isolates. At a minimum 5 different branches were compared. On occasion, additional MLST data was available for more isolates allowing selection of additional branches (up to 7) for analysis.

Figure 2. Example of How a Rep-PCR Dendrogram was used to Predict MLST Sequence Types (ST). Key number indicates sample number for dendrogram. The red node connects the two extreme isolates used for initial analysis (indicated by letter A in key, e.g., samples 2 and 20). The blue nodes indicate branches selected for internal branch analysis (e.g., samples 2 and 4 were analyzed for the first internal branch, indicated by the letter B in the key). Connecting nodes <98% were predicted to be MLST ST no match, while connecting nodes ≥98% were predicted to be an MLST ST match. GT = rep-PCR genotype, ST = MLST sequence type

Table 1

Summary demographics of children and family members evaluated in this study.

PART I				
	Cohort 1 (ID 100–255)	Cohort 2 (ID 500–608)	Total	
No. of Children	51	57	108	
Male/Females	27/24	33/23	60/47	
Mean Age (years) ^a	7.1	2.3		
Age Range (years) ^a	5.0–9.6	0.5–4.7		
No. Isolates analyzed with MLST	53	44	97	

PART II				
	Children	Mothers	Fathers	Other Family
No. of Subjects	24	9	1	14
Cohort 1	14	6	1	10
Cohort 2	10	3	0	4
Male/Females	11/13	6	1	-
Mean Age (years) ^a	5.3	-	-	-
Age Range (years) ^a	1.2–9.4	-	-	-
No. Isolates analyzed with MLST	61	12	2	12

^a Age was determined based on period of collection for the isolates used for MLST analysis. Subjects with multiple isolates from different periods were included in the average whilst isolates from the same individual and period were excluded.

Table 2

Prediction accuracy for two extreme isolates and internal branch groups for the ten most common rep-PCR genotypes.

Extreme Branches (2 isolates)				Internal Branches (2–3 isolates/group)			
Genotype	rep-PCR Predicted match	MLST ST Actual match ^a	rep-PCR Predicted match ^b	No. of isolates analyzed by MLST	No. of Isolate Groups	MLST ST for group predicted correctly (%)	SLV, DLV MLST ST for group predicted correctly (%)
G18	Yes	Yes	Yes	10	5	5 (100)	5 (100)
G09	No	Yes	No	9	5	2 (40)	4 (80)
G01a	No	No	No	10	5	4 (80)	5 (100)
G12	No	No	No	10	5	3 (60)	4 (80)
G01	No	No	No	11	5	3 (60)	4 (80)
G13	No	No	No	11	7	4 (57)	4 (57)
G06	No	No	No	9	5	2 (40)	2 (40)
G07	No	No	No	9	5	4 (80)	5 (100)
G23	No	No	No	9	5	2 (40)	2 (40)
G22	No	No	No	9	5	4 (80)	4 (80)
Predicted accurately			9 of 10	97	52	33 (63)	39 (75)
Totals							

^aCorrect predictions are bold.

^brep-PCR Predictions represented here are for the dendrogram as a whole (derived from 98% cut-off and results of extreme branch analysis), not individual groups.

Total combinations represented is 62 (52 from internal branch groups and 10 from extremes for each genotype) SLV; single locus variant, DLV; double locus variant

Table 3

Summary statistics of using rep-PCR to predict matching and non-matching MLST sequence types.

Actual Match/No-Match			
	Actual Match	Actual No-match	<i>p</i>-value^a
Predicted Match	22	18	0.0004
Predicted No Match	2	20	
Sensitivity	91.7%		
Specificity	52.6%		
Positive Predictive Value	55.0%		
Negative Predictive Value	90.9%		

Match/No-Match Including SLV and DLV			
	Actual Match	Actual No-match	<i>p</i>-value^a
Predicted Match	28	12	<0.0001
Predicted No Match	2	20	
Sensitivity	93.3%		
Specificity	62.5%		
Positive Predictive Value	70.0%		
Negative Predictive Value	90.9%		

^a *p*-value from Fisher's Exact test.

Table 4

Prediction accuracy for children only and child-parent pairs with rep-PCR genotype G12.

A. Parent-child Pairs						
Family member ID	No. of Isolates	Predicted extreme ST match	Actual extremes ST match	Internal isolates ST match extreme ST	Predicted parent- child match	Actual parent- child match
107-C	1	-	-	-	Yes	Yes
107-M	1	-	-	-		
170-C	4	No	No	Yes	No	No
170-M	1	-	-	-		
180-C	1	-	-	-	Yes	Yes
180-M	1	-	-	-		
182-C	6	No	Yes	Yes	Yes	Yes
182-M	2	No	Yes	-		
223-C	1	-	-	-	Yes	No
223-M	1	-	-	-		
226-C	2	Yes	Yes	-	Yes	No
226-F	2	No	Yes	-		
227-C	3	Yes	Yes	Yes	No	No
227-M	1	-	-	-		
518-C	5	Yes	Yes	Yes	Yes	Yes
518-M	1	-	-	-		
596-C	4	No	Yes	Yes	Yes	Yes
596-M	1	-	-	-		
600-C	6	No	Yes	Yes	Yes	Yes
600-M	3	No	Yes	Yes		
Predicted Correctly (%)				40.0%	100.0%	90% M-C 0% F-C

B. Child Only ^a					
Family Member ID	No. of isolates	Predicted extreme ST match	Actual extreme ST match	Internal isolates ST match	Internal isolates ST match extreme ST
159-C	4	No	Yes	Yes	Yes
219-C	2	Yes	Yes	Yes	-
241-C	6	No	Yes	Yes	No
597-C	3	No	Yes	Yes	Yes
608-C	4	No	Yes	Yes	Yes
Predicted Correctly (%)				20.0%	75.0%

^aFor child only table, only 5 of the 14 children are shown because all others had only one isolate evaluated with MLST. Family ID numbers <500 and >500 indicate Cohort 1 (older children) and Cohort 2 (younger children), respectively. Bold line separates age groups.

C; Child, M; mother, F; Father, ST; MLST sequence type

Commonality of MLST sequence types for children only and parent-child pairs with rep-PCR genotype G12.

Table 5

A. Mother-Child Pairs with G12			
ID	Total G12 Isolates	Isolates Analyzed	ST
107-C	1	1	
107-M	2	1	
170-C	24	4	157 (3), 202 (1)
170-M	13	1	212
180-C	15	1	
180-M	1	1	
182-C	74	6	
182-M	4	2	
223-C	10	1	
223-M	5	1	
227-C	42	3	
227-M	11	1	
518-C	35	5	156
518-M	1	1	156
596-C	32	4	192
596-M	7	1	192
600-C	92	6	
600-M	4	3	
Totals	373	43	

B. Child Only with G12			
ID	Total G12 Isolates	Isolates Analyzed	ST
150-C	1	1	175
159-C	113	4	
218-C	3	1	
219-C	16	2	
222-C	1	1	119 ^a
241-C	106	6	166
			181 (2), 132 ^a (4)
228-C	1	1	150
502-C	1	1	92 ^b
509-C	12	1	
528-C	1	1	
588-C	1	1	
597-C	24	3	
601-C	1	1	
608-C	22	4	
C. Father-Child Pairs with G12			
226-C	15	2	
226-F	9	2	
			132 ^a
Totals	327	32	

Sequence types (ST) shared between unrelated individuals are colored. Numbers in parenthesis indicate number of isolates with this MLST sequence type when more than one sequence types was detected.
C = child, M = Mother, F = Father

^aMatches a sequence type previously reported from Thailand

^bMatches a sequence type previously reported from Japan