Journal of Dental Research
2020, Vol. 99(8) 969–976

DOI: 10.1177/0022034520914519
journals.sagepub.com/home/jdr

Caries-Associated Biosynthetic Gene Clusters in *Streptococcus mutans*

S.S. Momeni[1*], S.M. Beno[2], J.L. Baker[3], A. Edlund[3], T. Ghazal[4], N.K. Childers[1], and H. Wu[1*]

Abstract

Early childhood caries (ECC) is a chronic disease affecting the oral health of children globally. This disease is multifactorial, but a primary factor is cariogenic microorganisms such as *Streptococcus mutans*. Biosynthetic gene clusters (BGCs) encode small molecules with diverse biological activities that influence the development of many microbial diseases, including caries. The purpose of this study was to identify BGCs in *S. mutans* from a high-caries risk study population using whole-genome sequencing and assess their association with ECC. Forty representative *S. mutans* isolates were selected for genome sequencing from a large-scale epidemiological study of oral microbiology and dental caries in children from a localized Alabama population. A total of 252 BGCs were identified using the antiSMASH BGC-mining tool. Three types of BGCs identified herein—butyrolactone-like, ladderane-like, and butyrolactone-ladderane-like hybrid (BL-BGC)—have not been reported in *S. mutans*. These 3 BGCs were cross-referenced against public transcriptomics data, and were found to be highly expressed in caries subjects. Furthermore, based on a polymerase chain reaction screening for core BL genes, 93% of children with BL-BGC had ECC. The role of BL-BGC was further investigated by examining cariogenic traits and strain fitness in a deletion mutant using in vitro biofilm models. Deletion of the BL-BGC significantly increased biofilm pH as compared to the parent strain, while other virulence and fitness properties remained unchanged. Intriguingly, BL-BGC containing strains produced more acid, a key cariogenic feature, and less biofilm than the model cariogenic strain *S. mutans* UA159, suggesting the importance of this BL-BGC in *S. mutans*–mediated cariogenesity. The structure of any BL-BGC derived metabolites, their functions, and mechanistic connection with acid production remain to be elucidated. Nevertheless, this study is the first to report the clinical significance of a BL-BGC in *S. mutans*. This study also highlights pangenomic diversity, which is likely to affect phenotype and virulence.

Keywords: dental caries, genotypes, butyrolactone-ladderane, transcriptomics, acidity, whole genome sequencing

Introduction

Dental caries is a multifactorial disease with a complex etiology that has not been completely elucidated and understood. Early childhood caries (ECC) is of particular concern, since this form of the disease is characterized by tooth decay dramatically affecting the health and well-being of children under 6 y of age (AAPD 2017). ECC is highly prevalent in children from underserved groups such as minorities, individuals of low socioeconomic status, and those with limited access to dental care (AAPD 2017).

Many factors are thought to play a role in initiation and progression of ECC disease, but the study of microorganism-mediated pathogenesis has uncovered many aspects of the etiology of dental caries. *Streptococcus mutans* is the bacteria most commonly associated with ECC (Loesche 1986; Gross et al. 2012). It has been demonstrated that clinical isolates of *S. mutans* vary considerably in their virulence capabilities, leading researchers to postulate that specific *S. mutans* strains may contribute to the onset of ECC, which may explain health disparities between different populations (Palmer et al. 2013). A few virulence and fitness factors of *S. mutans* related to ECC have been determined, which include the ability to form biofilms, produce organic acids, and withstand various environmental stresses (e.g., aciduric environment, oxidative stress) (Loesche 1986; Busuioc et al. 2009; Bowen and Koo 2011; Palmer et al. 2013).

Many small molecules of bacterial origin are produced by organized genes groups called biosynthetic gene clusters (BGCs). Common types of BGCs harbored by the oral microbiota include polyketide (PKS), nonribosomal peptide synthases (NRPS), hybrid PKS/NRPS, and ribosomally synthesized and posttranslationally modified peptides (RiPPs) that

[1]Department of Pediatric Dentistry, University of Alabama at Birmingham, Birmingham, AL, USA
[2]Department of Psychology, University of Alabama at Birmingham, Birmingham, AL, USA
[3]Genomic Medicine Group, J. Craig Venter Institute, La Jolla, CA, USA
[4]Department of Preventive and Community Dentistry, University of Iowa, Iowa City, IA, USA
*Current affiliation: Department of Integrative Biomedical & Diagnostic Sciences, OHSU School of Dentistry, Portland, OR, USA

A supplemental appendix to this article is available online.

Corresponding Author:
H. Wu, Department of Integrative Biomedical & Diagnostic Sciences, OHSU School of Dentistry, Portland, OR 97239, USA.
Email: wuhu@ohsu.edu

Table. List of Bacterial Strains, Plasmids, and Primers Used in This Study.

	Description		Source
Strain			
UAB-1 to UAB-40	*Streptococcus mutans* clinical strains from Uniontown epidemiological study		Momeni et al. 2016
UA159	Control reference strain		Ajdic et al. 2002
UA140			Qi et al. 2001
Top 10 *Escherichia coli*	Cloning strain		Invitrogen
UAB-10 Δ*but*	*but* mutant		This study
Plasmids			
pGEM-T			Promega
pALH124	Promoterless *aphA3* Kan^r resistant cassette		A. L. Honeyman
pVPT-GFP	*Escherichia coli–Streptococcus* shuttle vector and expression plasmid Ery^r		Zhou et al. (2008)
pVPT-GFP *cbut*	UAB-10 *but* complement		This study
Primers		Sequence (5′ to 3′)	
Butryol 322 F-1	qRT-PCR	AGTTCGAAGGCAGATCAAGC	This study
Butryol 622 R-1		TCAACAAAGGATGCTCTAGGCT	This study
Ladd 146 F-1		TCGGAGGTGAGGTCAGTGT	This study
Ladd 595 R-1		CATACGGTGCACAAGGTTGG	This study
Butryol F-1 (1,000 up)	Knockout	CGGTTTGATGATAAAATTGGTA	This study
Butryol R-1 (1,000 down)		ATAATTAAAACTTATGATCGGCTG	This study
Butyrol Inv R-2		ATTCGGGGATCCAATAATACACACTCACTTTATAATTTTAGCA	This study
Butyrol Inv F-2		CCGTTAGGATCCATGAATAAATTAGAAAAAGAGTCTATAAA AAATC	This study
Butyrol Sal1 F	Complement	CTAGTCGTCGACATGAGGAGTATTCTACAAGATAATTTC	This study
Butyrol Kpn1 R		CTGATTGGTACCTTATTCATCTAATCTTTCTCTTATTTCTTT	This study
Smut3368-F	gtfB	GCCTACAGCTCAGAGATGCTATTCT	Yoshida et al. 2003
Smut3481-R		GCCATACACCACTCATGAATTGA	Yoshida et al. 2003
16s-3	16S	TCCTACGGGAGGCAGCAGT	Nadkarni et al. 2002
16s-4		GGACTACCAGGGTATCTAATCCTGTT	Nadkarni et al. 2002

F, forward primer; qRT-PCR, quantitative reverse transcription polymerase chain reaction; R, reverse primer.

are represented by bacteriocins (i.e., mutacins) as well as other small molecules with various activities, such as biofilm formation, bacterial cell signaling, stress response, and host colonization (Wu et al. 2010; Merritt and Qi 2012; Zvanych et al. 2015; Liu et al. 2016; Aleti et al. 2019). Recent studies have identified novel PKS/NRPS hybrids responsible for the production of mutanicyclin and 3 structural variants of reutericyclin (Hao et al. 2019; Tang et al. 2020). Liu and colleagues identified 355 BGCs from 169 *S. mutans* sequenced genomes available in the NCBI genome database, suggesting clinical *S. mutans* isolates encode a wealth of BGCs (Liu et al. 2016; Tang et al. 2020). Since bacterial metabolites play a critical role in disease and health, there is a need to determine the importance of *S. mutans* BGCs and their products in virulence and fitness features relevant to dental caries.

In this study, whole-genome sequencing of 40 representative *S. mutans* clinical isolates and UA140 was performed from a large-scale epidemiological study of a localized high-caries risk study population. BGCs within these genomes were identified and characterized. A novel butyrolactone-ladderane-like hybrid (BL-BGC) was identified and subsequently found to be associated with ECC by polymerase chain reaction (PCR) screening of *S. mutans* clinical isolates and by expression mapping of this BGC with available transcriptomics data. Knockout of a key cluster gene inhibited acid production in *S. mutans*, suggesting the importance of this BGC in bacterial cariogenicity.

Materials and Methods

Sample Collection and Processing

Bacterial strains and plasmids used in this study are listed in the Table. Additional information on sample population, isolate selection, isolate processing, and institutional review board approvals is available in the Appendix. Briefly, 40 *S. mutans* isolates were selected based on the 34 representative repetitive extragenic palindromic PCR (rep-PCR) genotypes from 13,906 isolates previously identified from an 8-y longitudinal epidemiological study of children and their household family members in a rural population (Uniontown, Perry County, Alabama) (Momeni et al. 2016). UA140 was also sequenced because this strain has been reported to have specific mutacins small molecules (Merritt and Qi 2012). Caries scores (dmfs [decayed, missing, or filled surfaces]) were reported according to World Health Organization (WHO) criteria (WHO 1997). ECC was defined in this study as any visibly detectable caries before age 6 y (i.e., dmfs >0).

Whole-Genome Sequencing

DNA was extracted using the Qiagene DNeasy Ultraclean Microbial Kit (Qiagen) (Moser et al. 2010). Whole-genome sequence drafts were obtained using the Illumina MiSeq platform. Genomes were assembled de novo using SPAdes (v3.9.0)

(Bankevich et al. 2012). Quality control of draft genomes was performed using QUAST (Gurevich et al. 2013). Roary was used for core and pan genome analysis (Page et al. 2015). GenBank accession numbers are listed in Appendix Table 1.

Biosynthetic Gene Cluster Analysis

Biosynthetic gene clusters were identified using the antiSMASH v4.1.0 pipeline (Blin et al. 2017). BGCs and RiPPs were counted and compared with counts of BGCs previously reported by Liu et al. (2016) using the Mann-Whitney test. Liu and colleagues used antiSMASH v3, whereas this study used v4; however, we examined our strains using v3 and found only rare nominal variation in counts between versions. Previously published transcriptomics data by Do et al. (2015) were mapped to the BGCs identified in this study to determine association with caries. A total of 9 caries and 4 healthy subjects from the Do et al. transcriptomics data had sufficient read depth for messenger RNA (mRNA) sequence read mapping.

Prevalence of BL-BGCs in Children

The prevalence of the BL-BGCs was determined using SYBR Green PCR with sequence-specific primers for butyrolactone and ladderane core genes of the BL-BGC (Table). As previously reported, 69 of 96 children originally recruited completed 36 mo of follow-up, and of the 69 children, 68 had at least 1 *S. mutans* isolate recovered (Childers et al. 2017). For BL-BGC prevalence screening, DNA from 3 *S. mutans* isolates for the predominate genotype observed for each child ($n = 63$) from cohort 2 (1 to 5 y of age) of the larger epidemiological study were evaluated. Five children had fewer than 3 isolates available and were excluded. Distribution of *S. mutans* genotypes and caries status are listed in Appendix Table 3. If a child had genotypes G18, G15, or G19 as a minor genotype, then up to 3 additional isolates for each of these minor strains were also evaluated since the representative strain types with these genotypes had BL-BGC. Amplification for both genes was required for inclusion in analysis. Bivariate analyses were conducted to assess the relationships between ECC (count and present/absent) and BL-BGC in either predominate or minor genotype using negative binomial regression.

Mutant and Complement Construction

A key gene of the BL-BGC, the butyrolactone gene (*but*), was selected for mutation. A nonpolar butyrolactone (Δ*but*) mutant was generated by insertional mutagenesis with a kanamycin resistance cassette as described (Peng et al. 2016), and the mutant was complemented (c*but*) using the full-length *but* gene from a replicating plasmid as reported. PCR and sequencing were used to confirm deletion in Δ*but* and complementation in c*but*. Additional method details are available in the Appendix.

Quantitative Reverse Transcription PCR

Expression of the *but* gene was determined by quantitative reverse transcription polymerase chain reaction (qRT-PCR) in wildtype UAB-10 (WT), Δ*but*, and c*but*. Briefly, frozen isolates were plated on Todd Hewitt (TH) agar and grown 48 h. Total RNA was extracted from mid-exponential phase ($OD_{600} = 1.0$) *S. mutans* cultures, using the Direct-zol RNA miniprep kit (Zymo), and DNA digestion was performed using RQ1 RNase-free DNase (Promega). RNA was then purified with the RNase Easy Minikit (Qiagen) and converted to complementary DNA (cDNA) using the iScript cDNA synthesis kit (BioRad). Quality of RNA extracted was confirmed by gel electrophoresis on 1% agarose prior to cDNA synthesis. Approximately 20 ng/μL cDNA was used for qRT-PCR with Fermentas Maxima SYBR Green Master Mix (ThermoFisher) with internal primers for the *but* gene (Appendix Table 1) as well as 16S primers as a reference (Nadkarni et al. 2002). Relative expression values were calculated $2^{-(\text{CT target} - \text{CT 16S reference})}$, where CT is the threshold cycle (Rasmussen et al. 2011).

Biofilm Analysis

For all biofilm analysis, overnight cultures of *S. mutans* were subcultured in Todd Hewitt broth (THB) and grown to early exponential phase ($OD_{600} \approx 0.5$). Following subculture, cells were inoculated at 1:1,000 dilution in THB supplemented with 1% sucrose. Then, 200 μL was added to sterile 96-well flat-bottom plates and incubated at 37°C in CO_2 for 16 h under static conditions. All biofilm experiments were performed with at least 3 technical replicates and 3 biological replicates. *S. mutans* UA159 was used as a control for cross-run comparison as well as a caries reference strain as UA159 has been widely studied and shown to be cariogenic in rat models (Michalek et al. 1975; Garcia et al. 2017; Rainey et al. 2019; Scoffield et al. 2019). Student's *t* test was used for growth curve and biofilm analysis with statistical significance set at $P \leq 0.05$. The following assays and measurements were conducted to assess virulence: biomass (crystal violet), intracellular iodophilic polysaccharide (IPS, glycogen, iodine), pH (pHrodo), glucan (Cascade blue), cell density (Syto9), acid tolerance (pH 2.8), and oxidative stress (hydrogen peroxide 0.2%) (Rampersad 2012; Garcia et al. 2017; Zhou et al. 2017; Rainey et al. 2019). Additional methods for biofilm and growth curve are available in the Appendix.

Results

Representative Strains Contain More Genes and Accessory Genes

Forty *S. mutans* genomes from 38 subjects (31 from children and 7 from adults) and UA140 were sequenced and assembled. Features of these *S. mutans* genomes are summarized in Appendix Table 1. Genome coverage was at least 96% for all strains. Mean genome size and guanine-cytosine content were comparable to the laboratory model strain UA159. The mean number of total genes (2,046) was more than the previously reported for 57 *S. mutans* genomes (1,636) (Cornejo et al. 2013). The core genome consisted of 1,439 genes and the pangenome contained 3,744 accessory genes for the 40 *S. mutans* evaluated (excluding UA140). The data for the core

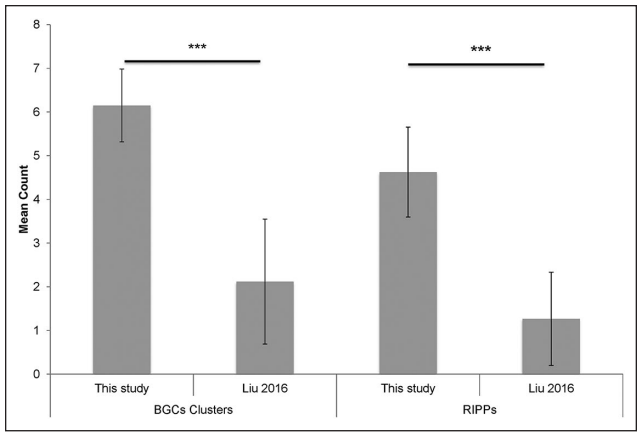

Figure 1. Representative *Streptococcus mutans* strains in this study have a higher number of biosynthetic gene clusters (BGCs) and ribosomally synthesized and posttranslationally modified peptides (RiPPs) per strain as compared with a previous study. Comparison of mean total numbers of BGCs and RiPPs observed in this study as compared with those published by Liu et al. (2016) for *S. mutans*. ***$P > 0.0001$ by Mann-Whitney test.

genome are consistent with previously published data (1,430); however, *S. mutans* strains in the current study have a greater number of accessory genes on average (94/strain here vs. 56/strain previously reported) (Cornejo et al. 2013).

Representative Strains Are Genetically Distinct within This Population

Overall, all 40 genomes displayed pangenomic variations (Appendix Fig. 1). Greater accessory gene variation was observed between strains that were from different rep-PCR genotypes while only minor variations were observed between strains of the same genotype, supporting that these representative rep-PCR genotypes are genetically different.

Representative Strains Have More BGCs and RiPPs

Analysis of the 41 *S. mutans* genomes using antiSMASH identified 252 BGCs (Appendix Table 2). Compared with results reported by Liu et al. (2016), the 40 *S. mutans* strains in the present study had significantly more BGCs clusters overall and more RiPPs per strain than the 169 NCBI strains previously analyzed (Fig. 1) (Liu et al. 2016).

Identification of New Biosynthetic Gene Clusters

BGCs identified were primarily NRPSs and RiPPs (bacteriocins and lantipeptides) (Appendix Table 2). However, 3 classes of BGCs were identified that have not been previously reported in *S. mutans*, including a butyrolactone-like BGC (6 strains), a ladderane-like BGC (1 strain), and a butyrolactone-ladderane-like hybrid BL-BGC (4 strains). Comparative analysis (Appendix Fig. 2) revealed that overall cluster gene arrangement and individual butyrolactone and ladderane genes differ between the 3 new BGC clusters.

New BGCs Were Upregulated in Caries Subjects

Initial analysis of differential BGC expression from the DESeq2 normalized data focused on individual genes from the clusters. The results indicated some genes from ladderane and BL-BGC had a significant increase in expression in subjects with caries versus caries-free subjects (data not shown). Subsequent cluster-mapping analysis using all the genes of the clusters instead of individual genes was limited to the new BGCs representing butyrolactone-like, ladderane-like, and BL-BGC. Three of the butyrolactone-like strains did not have sufficient reads for comparison and were excluded. The increased expression of these 3 new BGCs in samples from caries subjects as compared to caries-free subjects (Fig. 2) indicates these BGCs may be associated with *S. mutans* pathogenesis.

BL-BGC Is Associated with Early Childhood Caries

Sixty-three children (90% with ECC) were examined for prevalence of BL-BGC using PCR screening of the butyrolactone- and ladderane-specific genes (Appendix Table 3). G18 was the most common predominate genotype (11/63, 18%) among these children. Genes for the BL-BGC were detected in 25% of children using the predominate genotypes and 44% of children when minor G18, G19, and G15 were included in analysis. For children with BL-BGC ($n = 28$), 93% had ECC. There were clear trends that BL-BGCs containing predominate and minor genotypes were positively associated with ECC (odds ratio [OR], 1.68; 95% confidence interval [CI], 0.28–9.90; $P > 0.05$).

Inactivation of Butyrolactone Gene Significantly Reduces Biofilm Acidity

To further determine the function of the BL-BGC, the butyrolactone core gene harbored within the gene cluster was inactivated by allelic replacement strategy in a G18 representative strain, UAB-10 (WT), resulting in the Δ*but* strain. The mutant strain was confirmed by PCR and DNA sequencing. Furthermore, qRT-PCR was used to evaluate expression of the core butyrolactone gene. Expression of the butyrolactone gene was detected in UAB-10 and complement (c*but*) with no expression in Δ*but* (Fig. 3).

UAB-10, Δ*but*, and c*but* were then evaluated for biofilm virulence and fitness. No significant differences were observed in biofilm biomass (crystal violet assay) and other biofilm properties, including IPS glycogen, oxidative stress response, and acid tolerance (Appendix Fig. 3A–C, E, F). However, when all UAB-10 derived strains were compared with UA159, a model reference strain for assessing virulence, these strains produced significantly less biofilm and were significantly more acid tolerant. UA159 and UAB-10 have different genetic backgrounds. Future studies to compare UAB-10 with other clinical strains are planned.

The butyrolactone gene knockout (Δ*but*) had significantly higher pH, as indicated by lower fluorescent intensity (Fig. 4A,

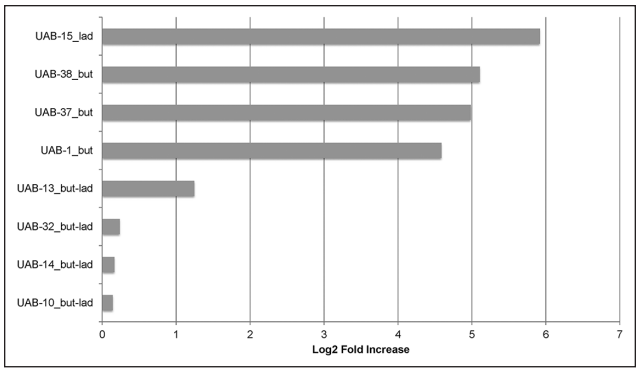

Figure 2. Association of *Streptococcus mutans* biosynthetic gene clusters (BGCs) and dental caries. Log_2 fold change in expression of *S. mutans* butyrolactone-like, ladderane-like, and butyrolactone-ladderane BGCs (BL-BGCs) identified in this study in dental caries by messenger RNA mapping of transcriptomics sequence reads (Do et al. 2015) against BGCs. *but*, butyrolactone-like cluster; *but-lad*, butyrolactone-ladderane-like hybrid cluster; *lad*, ladderane-like cluster.

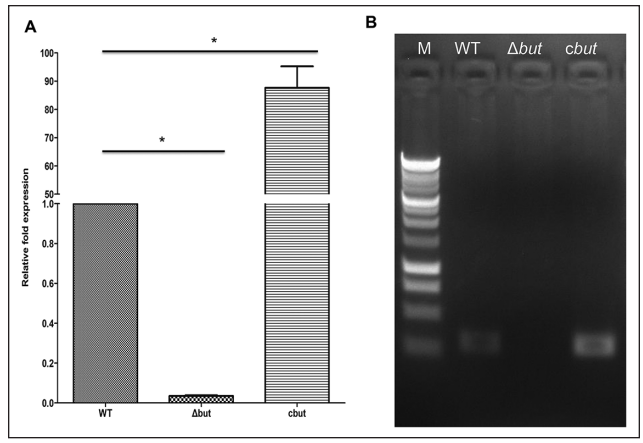

Figure 3. The *Streptococcus mutans but* mutant did not express the butyrolactone gene as determined by quantitative reverse transcription polymerase chain reaction (qRT-PCR). (**A**) Relative expression level of the butyrolactone gene in UAB-10 derivatives using qRT-PCR. Standard error of the mean shown. Data are from 3 independent experiments performed with 3 technical replicates each. (**B**) Visualization of representative RT-PCR products. Marker 1 kb. Statistical analysis was performed using Student's *t* test. $*P < 0.05$.

D). UAB-10 and c*but* demonstrated significantly lower pH when compared with Δ*but* and UA159. No significant difference in cell density was observed between UAB-10 and Δ*but*, but their cell density was significantly lower compared to UA159 (Fig. 4B). The knockout of the butyrolactone gene did not affect glucan. Interestingly, UAB-10 exhibited significantly higher glucan than UA159 (Fig. 4C, E). Both pHrodo red (pH detection) and Cascade blue (glucan detection) are dextran-conjugated fluorescent dyes. A comparable level of glucan was detected in both UAB-10 and Δ*but* by Cascade blue, suggesting the pH increase observed in the butyrolactone gene knockout is real and not due to an increased production of glucans by the mutant. These results suggest BL-BGC is associated with acid production.

Discussion

The whole-genome sequencing data provided by this study validate that the rep-PCR genotypes previously reported were indeed distinct (Appendix Fig. 1) (Momeni et al. 2016). Cornejo et al. (2013) reported a pangenome of 3,296 genes from an international collection of 57 *S. mutans* genomes as compared to the 3,744 genes reported in the present study for 40 *S. mutans* genomes from a localized population. The large amount of genetic variation in the accessory genome suggests that *S. mutans* isolates in this population may have increased caries potential since accessory genes often encode small molecules that can confer survival or competitive advantage (Waterhouse and Russell 2006; Cornejo et al. 2013; Hao et al. 2019; Tang et al. 2020).

Comparison of the number of BGCs and RiPPs identified in this study with those previously published (Liu et al. 2016) indicates that the *S. mutans* strains examined in this study population have significantly more BGCs (Fig. 1). This further supports the theory that strains in this population may have greater cariogenic potential since BGCs encode small molecules, including RiPPs such as bacteriocins, that affect strain fitness and virulence (Waterhouse and Russell 2006; Cornejo et al. 2013; Tang et al. 2020).

The identification of 3 novel BGCs in 11 *S. mutans* strains is particularly interesting. An NCBI search revealed 7 *S. mutans* strains with similar BL-BGC gene clusters currently in the database; however, the BL-BGC has not been discussed in studies related to these strains, nor are caries data available (Cornejo et al. 2013). The individual butyrolactone-like and ladderane-like clusters had minimal homology with strains in NCBI (≤25% and ≤15% homology, respectively). Studies describing the role of butyrolactone-like or ladderane-like small molecules in bacteria are limited, and no data are available on the BL-BGC. Butyrolactone is primarily reported in *Streptomyces* as a hormone-like signaling molecule related to morphological differentiation, quorum sensing, and biofilm formation (Biarnes-Carrera et al. 2015; van der Meij et al. 2017). A similar compound is also found in *Pseudomonas aeruginosa* (*N*-butyryl-L-homoserine lactone) and is related to activation of virulence genes and biofilm formation (Lee et al. 2018). Ladderanes have been described as membrane lipid components belonging exclusively to anammox (anaerobic ammonia-oxidating) bacteria and are unusually high energetic compounds (van Niftrik and Jetten 2012; Javidpour et al. 2016). Discovery of a ladderane-like BGC in *S. mutans* suggests that the oral cavity may have greater potential to harbor diverse BGCs that produce active natural products. In contrast, information regarding the structure, functions, and biosynthesis of metabolites of BL-BGC remains to be elucidated.

The transcriptomics mapping data indicates a significant fold increase in the number of reads mapping to the ladderane-like, butyrolactone-like, and BL-BGCs for caries subjects versus caries-free subjects (Fig. 2). This suggests that these BGCs are associated with caries pathogenicity. Although this finding is limited, since only 9 caries and 4 caries-free subjects from

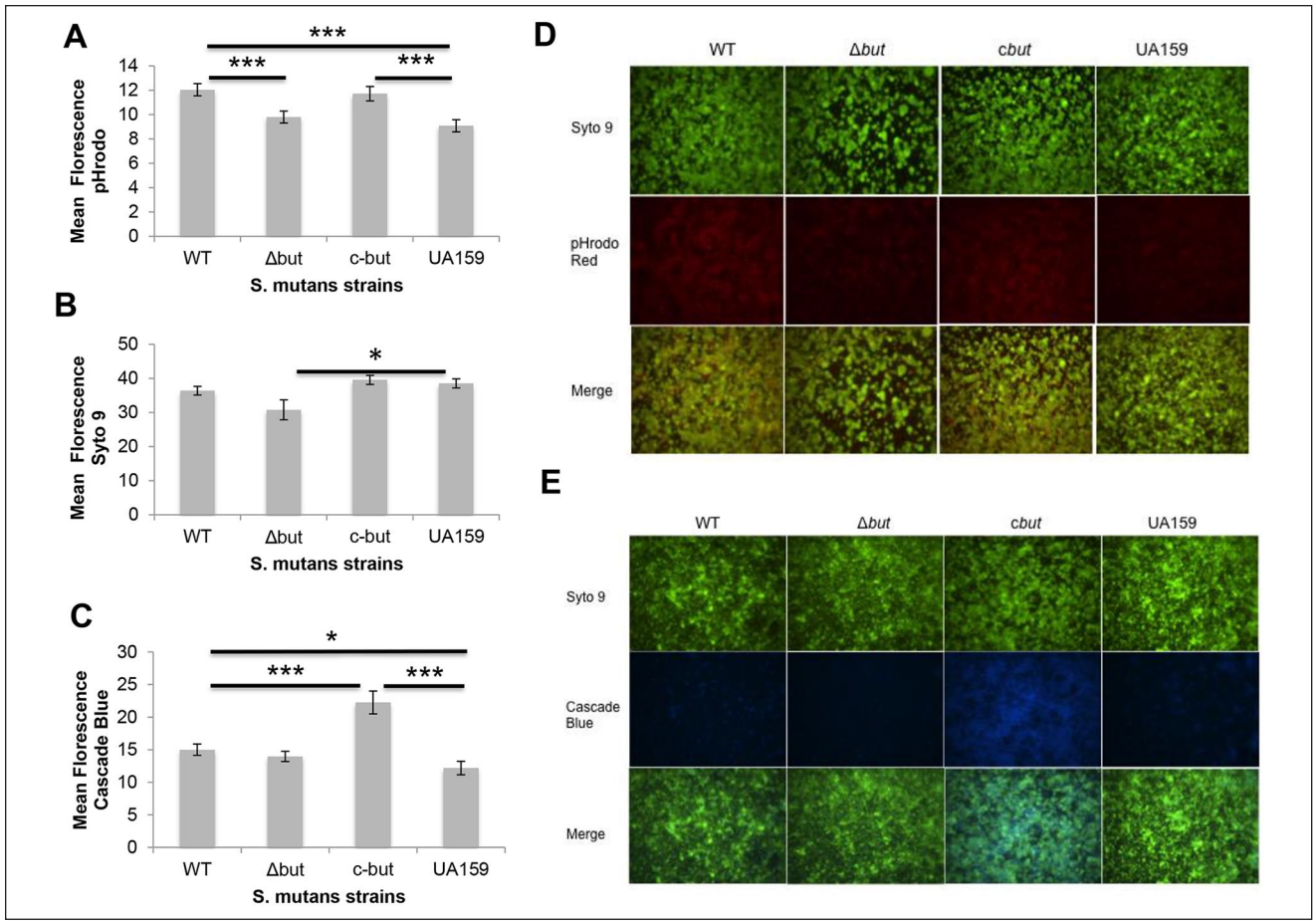

Figure 4. Effects of the *but* deficiency on virulence and fitness properties of *Streptococcus mutans.* Mean fluorescence of biofilm pH (**A**), cell density (**B**), and glucan (**C**) of *S. mutans* UAB-10 WT, Δ*but* mutant, and complement with *S. mutans* UA159 as a control. Lower pH is indicated by higher florescence. Standard error of the mean shown. Results are from 5 independent experiments with 3 technical replicates each. Representative images for visualization of biofilm pH (**D**) and glucan (**E**) using fluorescent microscopy at 16 h of growth. Biofilm was stained using pHrodo red to determine pH change, Cascade blue for glucan, and SYTO9 green for cell density. *$P<0.05$. ***$P<0.001$ by Student's *t* test.

the Do et al. (2015) study had sufficient read depth for analysis, no other meta-transcriptomics libraries were available for comparison at the time of this study.

The BL-BGC-containing strains identified from the 40 whole-genome sequences were distributed to 3 rep-PCR genotypes (G18, G19, G15). Interestingly, G18 is the most prevalent genotype observed among the children in this population and is the most frequently occurring genotype in the overall Uniontown population (Momeni et al. 2016). Although the BL-BGC was observed in only 4 of the 40 sequenced *S. mutans* genomes, the subsequent analysis of the prevalence of the BL-BGC among *S. mutans* isolates from 63 children indicates that the presence of BL-BGC is associated with increased odds of having ECC (Appendix Table 3). These data highlight the importance of screening for new BGCs with larger epidemiological studies to determine their clinical significance. As the present study was from a high-caries risk population, the majority of subjects (90%) had both ECC and *S. mutans*. The low number of caries-free subjects with *S. mutans* is a characteristic of this specific study population, and further study of the BL-BGC among more caries-free subjects is needed.

A key virulence trait of *S. mutans* in dental caries is the ability to form organic acids, especially lactic acid. In this study, *S. mutans* strains with butyrolactone (from the BL-BGC strain, UAB-10) produced significantly more acidic biofilms than UA159 or Δ*but* (Fig. 4). Furthermore, all UAB-10 derived strains produced significantly less biofilm and demonstrated greater acid tolerance than UA159. Together, these findings suggest that UAB-10 has some unique features related to acid production and aciduricity that is not tightly linked to formation of rich biofilms, a hallmark associated with the model cariogenic UA159. The BL-BGC may account for this. Since BGCs containing butyrolactone and ladderane consist of different gene arrangements and the amino acid structures are variable (Appendix Fig. 2), additional studies are needed to determine the significance of these BGCs on caries pathogenesis and strain fitness.

In summary, this study is the first to report the butyrolactone-like BGC, ladderane-like BGC, and BL-BGC in *S. mutans*. The presence of the BL-BGC significantly decreases the biofilm pH of *S. mutans* and concurrently increases the cariogenic potential of these strains by producing more acids. The in vitro

findings are consistent with the observation that children infected with BL-BGC-containing *S. mutans* had greater odds of having ECC. Additional studies are required to establish a mechanistic link between expression of BL-BGC, production of active metabolites, and acidity of biofilms.

Author Contributions

S.S. Momeni, H. Wu, contributed to conception, design, data acquisition, analysis, and interpretation, drafted and critically revised the manuscript; S.M. Beno, J.L. Baker, T. Ghazal, contributed to data analysis and interpretation, critically revised the manuscript; A. Edlund, contributed to conception, design, data analysis, and interpretation, drafted and critically revised the manuscript; N.K. Childers, contributed to conception, design, data acquisition, analysis, and interpretation, critically revised the manuscript. All authors gave final approval and agree to be accountable for all aspects of the work.

Acknowledgments

Special thanks to all the Uniontown Study organizers and participants. Coding assistance was provided by UAB Research Computing. UA140 was generously donated by Peng Zhou (originally Qi lab). This research was supported by the National Institute of Dental and Craniofacial Research (NIDCR) R01DE016684 (N.K. Childers), R01DE022350 (H. Wu), and T-90DE022736 (H. Wu). The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

ORCID iDs

S.S. Momeni https://orcid.org/0000-0002-0883-104X
J.L. Baker https://orcid.org/0000-0001-5378-322X
A. Edlund https://orcid.org/0000-0002-3394-4804

References

AAPD. 2017. Policy on early childhood caries (ECC): classifications, consequences, and preventive strategies. Pediatr Dent. 39(6):59–61.

Ajdic D, McShan WM, McLaughlin RE, Savic G, Chang J, Carson MB, Primeaux C, Tian R, Kenton S, Jia H, et al. 2002. Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. Proc Natl Acad Sci U S A. 99(22):14434–14439.

Aleti G, Baker JL, Tang X, Alvarez R, Dinis M, Tran NC, Melnik AV, Zhong C, Ernst M, Dorrestein PC, et al. 2019. Identification of the bacterial biosynthetic gene clusters of the oral microbiome illuminates the unexplored social language of bacteria during health and disease. mBio. 10(2). pii: e00321-19.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. Spades: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19(5):455–477.

Biarnes-Carrera M, Breitling R, Takano E. 2015. Butyrolactone signalling circuits for synthetic biology. Curr Opin Chem Biol. 28:91–98.

Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, Suarez Duran HG, de Los Santos ELC, Kim HU, Nave M, et al. 2017. antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. Nucleic Acids Res. 45(W1):W36–W41.

Bowen WH, Koo H. 2011. Biology of *Streptococcus mutans*–derived glucosyltransferases: role in extracellular matrix formation of cariogenic biofilms. Caries Res. 45(1):69–86.

Busuioc M, Mackiewicz K, Buttaro BA, Piggot PJ. 2009. Role of intracellular polysaccharide in persistence of *Streptococcus mutans*. J Bacteriol. 191(23):7315–7322.

Childers NK, Momeni SS, Whiddon J, Cheon K, Cutter GR, Wiener HW, Ghazal TS, Ruby JD, Moser SA. 2017. Association between early childhood caries and colonization with *Streptococcus mutans* genotypes from mothers. Pediatr Dent. 39(2):130–135.

Cornejo OE, Lefebure T, Pavinski Bitar PD, Lang P, Richards VP, Eilertson K, Do T, Beighton D, Zeng L, Ahn SJ, et al. 2013. Evolutionary and population genomics of the cavity causing bacteria *Streptococcus mutans*. Mol Biol Evol. 30(4):881–893.

Do T, Sheehy EC, Mulli T, Hughes F, Beighton D. 2015. Transcriptomic analysis of three *Veillonella* spp. present in carious dentine and in the saliva of caries-free individuals. Front Cell Infect Microbiol. 5:25.

Garcia SS, Blackledge MS, Michalek S, Su L, Ptacek T, Eipers P, Morrow C, Lefkowitz EJ, Melander C, Wu H. 2017. Targeting of *Streptococcus mutans* biofilms by a novel small molecule prevents dental caries and preserves the oral microbiome. J Dent Res. 96(7):807–814.

Gross EL, Beall CJ, Kutsch SR, Firestone ND, Leys EJ, Griffen AL. 2012. Beyond *Streptococcus mutans*: dental caries onset linked to multiple species by 16S rRNA community analysis. PLoS One. 7(10):e47722.

Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 29(8):1072–1075.

Hao T, Xie Z, Wang M, Liu L, Zhang Y, Wang W, Zhang Z, Zhao X, Li P, Guo Z, et al. 2019. An anaerobic bacterium host system for heterologous expression of natural product biosynthetic gene clusters. Nat Commun. 10(1):3665.

Javidpour P, Deutsch S, Mutalik VK, Hillson NJ, Petzold CJ, Keasling JD, Beller HR. 2016. Investigation of proposed ladderane biosynthetic genes from anammox bacteria by heterologous expression in *E. coli*. PLoS One. 11(3):e0151087.

Lee DJ, Jo AR, Jang MC, Nam J, Choi HJ, Choi GW, Sung HY, Bae H, Ku YG, Chi YT. 2018. Analysis of two quorum sensing-deficient isolates of *Pseudomonas aeruginosa*. Microb Pathog. 119:162–169.

Liu L, Hao T, Xie Z, Horsman GP, Chen Y. 2016. Genome mining unveils widespread natural product biosynthetic capacity in human oral microbe *Streptococcus mutans*. Sci Rep. 6:37479.

Loesche W. 1986. Role of *Streptococcus mutans* in human dental decay. Microbiol Rev. 50(4):353–380.

Merritt J, Qi F. 2012. The mutacins of *Streptococcus mutans*: regulation and ecology. Mol Oral Microbiol. 27(2):57–69.

Michalek SM, McGhee JR, Navia JM. 1975. Virulence of *Streptococcus mutans*: a sensitive method for evaluating cariogenicity in young gnotobiotic rats. Infect Immun. 12(1):69–75.

Momeni SS, Whiddon J, Cheon K, Ghazal T, Moser SA, Childers NK. 2016. Genetic diversity and evidence for transmission of *Streptococcus mutans* by DiversiLab rep-PCR. J Microbiol Methods. 128:108–117.

Moser SA, Mitchell SC, Ruby JD, Momeni S, Osgood RC, Whiddon J, Childers NK. 2010. Repetitive extragenic palindromic PCR for study of *Streptococcus mutans* diversity and transmission in human populations. J Clin Microbiol. 48(2):599–602.

Nadkarni MA, Martin FE, Jacques NA, Hunter N. 2002. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. Microbiology. 148(Pt 1):257–266.

Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 31(22):3691–3693.

Palmer SR, Miller JH, Abranches J, Zeng L, Lefebure T, Richards VP, Lemos JA, Stanhope MJ, Burne RA. 2013. Phenotypic heterogeneity of genomically-diverse isolates of *Streptococcus mutans*. PLoS One. 8(4):e61358.

Peng X, Zhang Y, Bai G, Zhou X, Wu H. 2016. Cyclic di-AMP mediates biofilm formation. Mol Microbiol. 99(5):945–959.

Qi F, Chen P, Caufield PW. 2001. The group I strain of *Streptococcus mutans*, UA140, produces both the lantibiotic mutacin I and a nonlantibiotic bacteriocin, mutacin IV. Appl Environ Microbiol. 67(1):15–21.

Rainey K, Michalek SM, Wen ZT, Wu H. 2019. Glycosyltransferase-mediated biofilm matrix dynamics and virulence of *Streptococcus mutans*. Appl Environ Microbiol. 85(5). pii: e02247-18.

Rampersad SN. 2012. Multiple applications of alamar blue as an indicator of metabolic function and cellular health in cell viability bioassays. Sensors (Basel). 12(9):12347–12360.

Rasmussen L, White EL, Pathak A, Ayala JC, Wang H, Wu JH, Benitez JA, Silva AJ. 2011. A high-throughput screening assay for inhibitors of bacterial motility identifies a novel inhibitor of the na+-driven flagellar motor and virulence gene expression in *Vibrio cholerae*. Antimicrob Agents Chemother. 55(9):4134–4143.

Scoffield J, Michalek S, Harber G, Eipers P, Morrow C, Wu H. 2019. Dietary nitrite drives disease outcomes in oral polymicrobial infections. J Dent Res. 98(9):1020–1026.

Tang X, Kudo Y, Baker JL, LaBonte S, Jordan PA, McKinnie SMK, Guo J, Huan T, Moore BS, Edlund A. 2020. Cariogenic *Streptococcus mutans*

produces tetramic acid strain-specific antibiotics that impair commensal colonization. ACS Infect Dis. 6(4):563–571.
van der Meij A, Worsley SF, Hutchings MI, van Wezel GP. 2017. Chemical ecology of antibiotic production by actinomycetes. FEMS Microbiol Rev. 41(3):392–416.
van Niftrik L, Jetten MS. 2012. Anaerobic ammonium-oxidizing bacteria: unique microorganisms with exceptional properties. Microbiol Mol Biol Rev. 76(3):585–596.
Waterhouse JC, Russell RR. 2006. Dispensable genes and foreign DNA in *Streptococcus mutans*. Microbiology. 152(Pt 6):1777–1788.
World Health Organization (WHO). 1997. Oral health surveys: basic methods. Geneva (Switzerland): World Health Organization.
Wu C, Cichewicz R, Li Y, Liu J, Roe B, Ferretti J, Merritt J, Qi F. 2010. Genomic island TnSmu2 of *Streptococcus mutans* harbors a nonribosomal peptide synthetase-polyketide synthase gene cluster responsible for the biosynthesis of pigments involved in oxygen and H_2O_2 tolerance. Appl Environ Microbiol. 76(17):5815–5826.
Yoshida A, Suzuki N, Nakano Y, Kawada M, Oho T, Koga T. 2003. 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. 41(9):4438–4441.
Zhou M, Fives-Taylor P, Wu H. 2008. The utility of affinity-tags for detection of a streptococcal protein from a variety of streptococcal species. J Microbiol Methods. 72(3):249–256.
Zhou Y, Yu L, Tao Y, Zhi Q, Lin H. 2017. Genetic polymorphism of scrA gene of *Streptococcus mutans* isolates is not associated with biofilm formation in severe early childhood caries. BMC Oral Health. 17(1):114.
Zvanych R, Lukenda N, Li X, Kim JJ, Tharmarajah S, Magarvey NA. 2015. Systems biosynthesis of secondary metabolic pathways within the oral human microbiome member *Streptococcus mutans*. Mol Biosyst. 11(1):97–104.