Published in final edited form as:

Mol Oral Microbiol. 2018 June ; 33(3): 257–269. doi:10.1111/omi.12222.

Whole genome sequence and phenotypic characterization of a Cbm⁺ serotype e strain of *Streptococcus mutans*

Alejandro Avilés-Reyes^{1,*}, Irlan Almeida Freires^{1,*}, Jessica K. Kajfasz¹, Dicler Barbieri², James H. Miller², José A. Lemos^{1,#}, and Jacqueline Abranches^{1,#}

¹Department of Oral Biology, University of Florida, College of Dentistry, Gainesville, FL

²Center for Oral Biology, University of Rochester School of Medicine and Dentistry, Rochester, NY

Abstract

We report the whole genome sequence (WGS) of the serotype e Cbm⁺ strain LAR01 of *Streptococcus mutans*, a dental pathogen frequently associated with extra-oral infections. The LAR01 genome is a single circular chromosome of 2.1 Mb with a GC content of 36.96%. The genome contains 15 PTS gene clusters, 7 cell wall-anchored (LPxTG) proteins, all genes required for the development of natural competence and genes coding for mutacins VI and K8.

Interestingly, the *cbm* gene is genetically linked to a putative type VII secretion system that has been found in Mycobacteria and few other gram-positive bacteria. When compared to the UA159 type strain, phenotypic characterization of LAR01 revealed increased biofilm formation in the presence of either glucose or sucrose but similar abilities to withstand acid and oxidative stresses. LAR01 was unable to inhibit the growth of *S. gordonii*, which is consistent with the genomic data that indicates absence of mutacins that can kill mitis streptococci. On the other hand, LAR01 effectively inhibited growth of other *S. mutans* strains suggesting that it may be specialized to outcompete strains from its own species. *In vitro* and *in vivo* studies using mutational and heterologous expression approaches revealed that Cbm is a virulence factor of *S. mutans* by mediating binding to extracellular matrix proteins and intracellular invasion. Collectively, the WGS analysis and phenotypic characterization of LAR01 provides new insights on the virulence properties of *S. mutans* and grants further opportunities to understand the genomic fluidity of this important human pathogen.

Introduction

The human pathogen *Streptococcus mutans* is generally regarded as the major etiological agent responsible for the initiation and progression of dental caries, one of the most prevalent infectious diseases in the world ¹. In addition, *S. mutans* is associated with cases of

[#]Corresponding authors: Mailing address: Department of Oral Biology, 1395 Center Drive, Box 100424, University of Florida College of Dentistry, Gainesville, FL 32610, Tel: (352) 273-6672; Fax: (352) 273-8829 jabranches@dental.ufl.edu, jlemos@dental.ufl.edu.

^{*}Both authors contributed equally to this work

DR JACQUELINE ABRANCHES (Orcid ID : 0000-0002-6615-7348)

Conflicts of interest

The authors declare no competing interests

Ethical statement.

This work was approved by the University of Rochester Institutional Review Board number 00042461.

infective endocarditis (IE) ² and a subset of strains expressing collagen-binding proteins (CBPs) like Cnm or Cbm have been implicated in other extra-oral pathologies such as hemorrhagic stroke and cerebral microbleeds ^{3, 4}. Strains of *S. mutans* are classified into four serotypes (*c*, *e*, *f* and *k*) with over 70% of strains isolated from dental plaque belonging to serotype *c*, approximately 20% to serotype *e*, while less than 5% belong to either serotype *f* or *k*. Interestingly, isolates of *S. mutans* that belong to the less prevalent serotypes *e*, *f* and *k* are the ones that most commonly express Cnm and Cbm ⁵⁻⁷.

The virulence of *S. mutans* in dental caries is strongly associated with the expression of glucosyltransferases (GTFs), which are secreted enzymes responsible for the synthesis of extracellular polymers of glucan from sucrose ⁸. A number of elegant studies conducted by different laboratories have shown that secreted GTFs prime the tooth enamel surface with glucans for bacterial adhesion via surface proteins that possess glucan-binding properties, and are primarily responsible for the formation of an extracellular matrix superstructure that anchors and supports the polymicrobial biofilm ⁸⁻¹¹. To colonize extra-oral tissues, *S. mutans* appears to rely on the presence of a small number of surface-associated proteins that tightly interact with extracellular matrix (ECM) proteins such as collagen, laminin and fibronectin. In particular, evidence is now accumulating that the CBPs Cnm and Cbm represent major virulence factors in extra-oral infections ^{3, 12-15}. Expression of either Cnm or Cbm leads to robust adherence to collagen and laminin, invasion of non-phagocytic cells and increased virulence in both invertebrate and vertebrate models of systemic infection ^{5, 12, 16, 17}. Moreover, CBP⁺ strains are more frequently isolated from dental plaque of individuals with bacteremia and IE suggesting a correlation between the expression of these adhesins and the development of extra-oral infections ^{16, 17}. While the distribution of Cbm⁺ strains in oral isolates is lower than 3% ¹⁸, a recent study found that the gene encoding Cbm was detected in 5 out of 8 blood samples of IE patients that tested positive for *S. mutans* DNA whereas the *cnm* gene was detected once ¹⁸.

Like many other bacterial pathogens, *S. mutans* is recognized as a clonal species where multiple lineages with specific genome variations develop over generations thereby providing evolutionary advantages to the population ¹⁹⁻²¹. While in many instances no major changes in the pathogenic potential of the organism are immediately noticed, certain permutations can be associated with decreased or increased virulence ²². For example, genetic alterations of the glucosyltransferase genes *gtfB* and/or *gtfC* can lead to differences in the ability of *S. mutans* to form biofilms in the presence of sucrose ^{23, 24}. Another classic example of such variations and their association with virulence are the *rgp* genes responsible for the synthesis of the rhamnose-glucose polysaccharide (RGP) in *S. mutans* ^{25, 26}. Aside from providing the chemical basis for the grouping of strains into serotypes, the various RGP structures have been shown to differentially contribute to *S. mutans* virulence during systemic infections ^{25, 26}. With the advent of next generation sequencing (NGS) technologies, the number of *S. mutans* strains with available sequences increased from only 1 in 2001 to 187 (<https://www.ncbi.nlm.nih.gov/genome/genomes/856>) as of today. This increase has led to a better understanding of *S. mutans* physiology and evolution and have accelerated the discovery of novel genes involved in virulence ^{24, 27, 28}. Presently, the complete genome sequences of 8 Cnm⁺ strains are publicly available but no Cbm⁺ strain has been sequenced to date. Considering that the pathogenic potential of *S. mutans* extends

beyond the oral cavity, the identification and characterization of virulence factors in strains associated with extra-oral infections is of great relevance. Here, we present the complete and annotated genome sequence of LAR01, a Cbm⁺ serotype *e* strain isolated from the oral cavity. In addition to the WGS, general phenotypic characteristics of LAR01 and the major properties of the *cbm* gene encoded by this strain were also investigated.

Materials and Methods

Bacterial strains and culture conditions

All strains used in this study are listed in Table 1. LAR01 is a *cbm*⁺, serotype *e* strain that was isolated from unstimulated whole saliva of a healthy donor with previous history of caries experience with pulpal involvement that participated in a pilot clinical study conducted at the University of Rochester (IRB approval # 00042461). Strains of *S. mutans* were routinely cultured in brain heart infusion (BHI) medium at 37°C in a humidified 5% CO₂ atmosphere. *Lactococcus lactis* NZ9800 and derivatives were grown in MG17 medium containing 0.5% glucose at 30°C. When required, 1 mg ml⁻¹ kanamycin, 10 µg ml⁻¹ erythromycin or 10 ng ml⁻¹ nisin was added to the growth media. *Streptococcus gordonii* DL-1 was routinely grown in BHI buffered with 25 mM KPO₄ (pH 7.0) at 37°C in 5% CO₂ atmosphere.

Isolation and identification of *S. mutans* LAR01

Saliva was serially diluted and plated on Mitis Salivarius Agar supplemented with sucrose and bacitracin (MSB)²⁹. Plates were incubated at 37°C in a 5% CO₂ atmosphere for 48h. Colonies were screened by PCR using *S. mutans*-specific primers³⁰ and, if positive, with serotype-specific primers as well as *cnn*- and *cbm*-specific primers^{18, 31, 32} (Table 2). Strains UA159 (serotype *c* Cnm⁻), OMZ175 (serotype *f* Cnm⁺), B14 (serotype *e* Cnm⁺), and OM88X (serotype *k* Cbm⁺) were used as controls in the PCR screens.

Chromosomal DNA purification and whole genome sequencing

The Wizard Genomic DNA Purification Kit (Promega) was used to isolate chromosomal DNA from an overnight-grown culture of *S. mutans* LAR01. The manufacturer's instructions for isolation of DNA from Gram Positive bacteria were followed with an additional incubation step with 0.5 U ml⁻¹ mutanolysin and 50 µg ml⁻¹ lysozyme at 37°C for 60 min. The resulting genomic DNA (gDNA) was then purified with the PowerClean Pro DNA Clean-Up Kit (Mo Bio) with a maximum of 20 µg gDNA loaded per column, and eluted in 50 µl of 10 mM Tris, pH 8. Concentration of the resulting gDNA was determined with a NanoDrop (Thermo Scientific) and confirmed by running on an 0.8% agarose gel. Approximately 12.5 µg of total purified gDNA was obtained, and concentrated to 50 ng µl⁻¹ by vacuum centrifugation. Purified *S. mutans* LAR01 gDNA was submitted to the University of Florida Interdisciplinary Center for Biotechnology Research, where concentration was verified and QC performed with a QuBIT fluorometer. Library construction (~20 kb) and full genome sequencing was performed with the PacBio Single-Molecule, Real-Time Sequencing (SMRT) system, using two SMRT cells. Genome assembly and annotation was done using the Celera Assembler and Prokka Annotation softwares, respectively^{33, 34}. The latter employs the Prodigal, RNAmmer, Aragorn, SignalP and Infernal prediction tools to

identify coding sequences (CDS), ribosomal RNA genes (rRNA), transfer and tmRNA genes, signal peptides (at N-terminal of CDS) and non-coding RNA, respectively. The complete genome sequence of strain LAR01 was deposited in the NCBI/GenBank databases under the accession number CP023477.

Construction of recombinant strains

The *cbm* gene from *S. mutans* LAR01 was replaced with a non-polar kanamycin (NPKan) cassette³⁵, which contains a promoterless *aphA3* Km^r gene with upstream stop codons in all three reading frames and without transcription termination sequences to allow transcription readthrough into downstream sequences, via allelic exchange. Briefly, PstI sites were introduced to DNA fragments containing the 5' and 3' flanking regions of *cbm* using the primers listed in Table 2. Upon amplification, PCR products were digested with PstI and then ligated to a PstI-digested NPKan cassette. Natural transformation of LAR01 was achieved using a ComX-inducing peptide (XIP)-based protocol developed for transformation of *S. mutans* UA159³⁶. Briefly, cultures were grown overnight in a chemically defined medium containing 0.5% glucose (CDMG)³⁷, diluted 1:20 in fresh CDMG and grown to OD₆₀₀ 0.1. Cultures were then transformed with 250 ng of the ligation mix in the presence of 100 nM XIP³⁸. Transformants were selected on plates containing kanamycin and the desired mutation confirmed by PCR sequencing of the *cbm* gene and flanking region. To complement the *cbm* mutant, the full length *cbm* gene was amplified from LAR01 using primers containing BamHI (5' primer) and XbaI (3' primer) restriction sites and cloned onto the nisin-inducible shuttle vector pMSP3535³⁹ that had been previously digested with the same restriction enzymes. The resulting plasmid, *pcbm*, was used to transform the *cbm* mutant in the presence of XIP and create the strain C *cbm*. For heterologous expression of *cbm* in *L. lactis*, the *pcbm* was used to electroporate *L. lactis* NZ9800 strain following the protocol described elsewhere¹². All strains (*S. mutans* and *L. lactis*) harboring *pcbm* or the empty pMSP3535 plasmids were selected on agar plates containing erythromycin.

Western blot analysis

Whole cell protein lysates were obtained by homogenization in the presence of 0.1 mm glass beads using a bead-beater (Biospec). Equal amounts of protein lysates (20µg per lane) from different strains were separated on 10% SDS-PAGE and transferred to polyvinylidene fluoride (PVDF) membranes (Millipore). The detection of Cbm was performed using rabbit anti-Cnm polyclonal antibody⁴⁰ diluted 1:2000 in 1X PBS + 0.1% Tween 20 and anti-rabbit horseradish peroxidase (HRP)-coupled antibody (Sigma-Aldrich). Membranes were developed using the enhanced chemiluminescent detection kit (GE Life Sciences).

Biofilm assay

Biofilm development by different *S. mutans* reference strains (LAR01, UA159 and OMZ175) was measured in polystyrene 96-well (flat-bottom) microtiter plates (Costar .595; Corning Inc). Briefly, cultures were grown in CDMG to OD₆₀₀ of 0.5 and used to inoculate (1:100) the wells of a microtiter plate containing 200 µl of CDM medium containing 1% glucose or sucrose. For biofilm formation determination, culture medium was removed after incubation at 37°C for 24h and the wells gently washed with 200 µl deionized water for two times. Subsequently, biofilms were stained with 50 µl of a 0.05% solution of crystal violet

for 5 min at room temperature and washed twice with water. Stained biofilms were then resuspended with a 7% acetic acid solution and OD₅₇₅ measurements were taken. A one-way ANOVA was performed to verify the significance of biofilm production by *S. mutans* strains. *P* values < 0.05 were considered significant.

Stress tolerance assays

The ability of the *S. mutans* strains (LAR01, UA159 and OMZ175) to survive extreme acid and peroxide challenges was determined using established protocols^{24, 41}. For acid killing assays, cells from an overnight culture were diluted 1:25 into BHI broth and incubated to OD₆₀₀ = 0.3 (unadapted cells) or to OD₆₀₀ of 0.2 followed by a 2h incubation in BHI broth that had been acidified with HCl to pH 5.0 (acid-adapted cells). Cells were then washed once with 0.1 M glycine buffer, pH 7.0, and resuspended in one-half of the original volume of 0.1 M glycine buffer, pH 3.0, for up to 45 min. For peroxide killing assays, cells were grown to OD₆₀₀ of 0.3 and treated with 0.2% H₂O₂ (58.8 mM) for up to 45 min. For both assays, cell viability at each time point was expressed as the percentage of viable cells (CFU ml⁻¹) at time zero. A one-way ANOVA was performed to verify significance. *P* values < 0.05 were considered significant.

Growth competition assays

The ability of *S. mutans* strains to inhibit growth of their counterparts was determined using a growth competition assay on solid media. Briefly, overnight cultures of *S. mutans* UA159, OMZ175 and LAR01 were diluted 1:25 in BHI and grown to OD₆₀₀ = 0.3. Then, a 15 µl aliquot of the early colonizer was spotted on a BHI agar plate incubated at 37°C, 5% CO₂ for 24 h. The following day, the plates were exposed to UV light (254 nm, ~3000µwatts/cm²) for 20 min to kill the early colonizer. Then, 15 µl of the late colonizer overnight culture, also grown in BHI, was spotted and the plates were incubated for an additional 24 h. A deferred antagonism assay was further performed to determine the ability of *S. mutans* strains to inhibit growth of *Streptococcus gordonii* and *Lactococcus lactis*, indicator strains for mutacin IV and V, respectively. Cultures of the different *S. mutans* strains were grown in BHI to an OD₆₀₀ of 0.3 when a 15 µl aliquot was spotted onto BHI agar and incubated for 24 h. Following incubation, plates were exposed to UV light for 20 min to ensure that subsequent antagonism was not due to actively growing *S. mutans* cells. Then, 500 µl of an overnight culture of *S. gordonii* DL-1 (mutacin IV sensitive) or *L. lactis* ATCC 11454 (mutacin V sensitive) grown in BHI as previously described was added to 5 ml soft (0.75 %) BHI agar, spread as an overlay and incubated for another 24 h before zones of growth inhibition around the *S. mutans* spots were measured⁴². Growth competition assays were also performed in the presence of aminopeptidase (Sigma-Aldrich) to prove that inhibition zones were due to mutacin activity. For this, early colonizers were spotted as described above and incubated for 24 h. After exposure to UV light for 20 min, aminopeptidase (64 µg) was spotted on the side of the early colonizer and plates were incubated for 1 h at 37°C. Then the late colonizer was spotted exactly over the peptidase site and the plates were incubated for an additional 24 h before visualization.

ECM binding and Human Coronary Artery Endothelial Cell (HCAEC) invasion

For ECM binding assays, 100 μ l of PBS-washed bacterial suspensions (*S. mutans* or *L. lactis*) containing approximately 1×10^9 CFU ml^{-1} were added to each well of a microtiter plate containing immobilized type I collagen from rat tail (Sigma-Aldrich) or mouse laminin (Becton-Dickinson) ^{5, 40, 43}. Adherent cells were stained with 0.05% crystal violet solution and OD₅₇₅ was measured. For HCAEC invasion, primary HCAECs (Lonza, Allendale, NJ) were cultured in endothelial cell basal medium-2 (EBM-2; Lonza) supplemented with EGM-2MV single-use aliquots (Lonza), as suggested by the supplier. The HCAECs were maintained at 37°C in a humidified, 5% CO₂ atmosphere. The cells were harvested by trypsinization and washed in EBM-2 medium. One ml of the suspension containing 10^5 endothelial cells was then seeded per well on 24-well flat-bottom tissue culture plates followed by overnight incubation in the presence of gentamycin at 37°C in a 5% CO₂ atmosphere. Prior to infection, the wells were washed three times with pre-warmed EBM-2 without antibiotics. Overnight bacterial cultures were washed twice in phosphate-buffered saline (pH 7.2), and resuspended in supplemented EBM-2 without antibiotics to obtain bacterial suspensions containing 1×10^7 CFU ml^{-1} of *S. mutans*. One ml of bacterial cell suspensions was used to infect HCAEC wells, in triplicate, for 2 h in the absence of antibiotics. Next, the wells were washed three times with 1 ml of EBM-2, followed by 3 h incubation in 1 ml of EBM-2 containing 300 $\mu\text{g ml}^{-1}$ gentamycin and 50 $\mu\text{g ml}^{-1}$ penicillin G to kill extracellular bacteria. After the incubation period in antibiotics, the wells were washed three times with EBM-2, then HCAECs were lysed for 20 min with 1 ml sterile water. The mixture of lysed HCAECs and *S. mutans* was plated onto BHI agar and incubated for 48 h at 37°C in a 5% CO₂ atmosphere. A one-way ANOVA was performed to verify the significance of binding and invasion. *P* values ≤ 0.05 were considered significant.

Galleria mellonella infection

For the *G. mellonella* killing assays, insects in the final instar larval stage were purchased from Vanderhorst Inc. (St. Marys, Ohio), stored at 4°C in the dark and used within 7 days of shipment. Groups of 15 larvae, ranging from 200 to 300 mg in weight and with no signs of melanization, were randomly chosen and used for subsequent infection. A 25- μ l syringe (Hamilton; Reno, Nevada) was used to inject the hemocoel of each larva via the last left proleg with 5- μ l aliquots containing 1×10^9 CFU/mL of *S. mutans* or *L. lactis* that had been grown overnight in BHI or BHI supplemented with erythromycin (10 $\mu\text{g mL}^{-1}$). Bacterial colony counts on BHI plates were used to confirm initial inocula. Groups injected with saline solution or with heat-inactivated *S. mutans* LAR01 (30 min at 75°C) were used as controls in each experiment. After injection, larvae were incubated at 37°C, and appearance (signs of melanization) and survival were recorded at selected intervals. Larvae were scored as dead when they displayed no movement in response to touch. Kaplan-Meier killing curves were plotted and estimation of differences in survival were compared using the log-rank test. A *P* value ≤ 0.05 was considered significant.

Immunofluorescence labeling of Cnm

Cultures of recombinant *L. lactis* expressing Cbm through *pcbM* were used for cellular localization of Cbm. Cells grown to mid-exponential phase were washed in PBS and

incubated in the presence of anti-rCnmA antibody (1:100) followed by anti-rabbit Alexa-488 conjugate antibody (1:100) with three washes between each antibody exposure. Cell-antibody complexes were visualized using an Olympus BX41 fluorescent microscope and images were captured with the QCapture Software.

Real-time quantitative RT-PCR

RNA extraction and real time RT-qPCR analysis were performed as previously described⁴⁴. Overnight cultures were diluted 1:25 in 5 ml BHI plain or supplemented with antibiotic and grown to OD₆₀₀ 0.5. At that point, cultures were mixed with an equal volume of ice-cold ethanol/acetone solution (1:1), immediately frozen in a dry ice/ethanol bath, and kept at -80°C until ready for RNA extraction. Cells were then harvested, washed twice in TE buffer (10 mM Tris-Cl [pH 8], 1 mM EDTA), and digested with 25 U mutanolysin and 10 mg ml⁻¹ lysozyme for 30 min at 37°C. Protoplasts were lysed with vigorous vortex mixing in 0.35 ml RLT buffer (Qiagen) supplemented with 1% β-mercaptoethanol, and RNA was purified using an RNeasy minikit (Qiagen), including the on-column DNase treatment recommended by the supplier. To further reduce DNA contamination, RNA samples were treated with DNase I (Ambion) at 37°C for 30 min and then repurified using an RNeasy minikit (Qiagen). RNA concentrations were determined using a NanoVue Plus spectrophotometer (GE Life Sciences). Reverse transcription and real-time PCR were carried out as previously described⁴⁵ using *cbm* specific primers (Table 2).

Results

Isolation and whole genome sequencing of *S. mutans* LAR01

The LAR01 isolate was obtained from a saliva sample of a caries experienced adult female subject with no obvious signs of active oral diseases (e.g. dental caries, periodontitis). Upon selection and differentiation on MSB agar, molecular identification using species- and serotype-specific primers confirmed this oral isolate as a serotype *e* *S. mutans* strain (data not shown). In subsequent PCR analysis, LAR01 tested negative for the presence of *cnm* but positive for *cbm*.

The LAR01 genome is a single circular chromosome of 2.1 Mb with a GC content of 36.96% (Fig. 1), which is of comparable size to other *S. mutans* genomes that can range from 1.74 Mb to 2.73Mb with GC content between 35.7 and 37.1%^{24, 27, 46–50}. A total of 1,956 CDSs, 5 rRNA operons, 65 tRNA genes and 10 insertion sequence elements (ISE) were identified in LAR01. All genes identified in LAR01 have been previously identified in other *S. mutans* strains. Previous studies using comparative analysis identified that there is a large degree of genome rearrangements within *S. mutans* strains^{27, 46} and the LAR01 genome was no exception. A gene annotated file is provided as part of the supplemental material (Table S1).

The LAR01 genome displays 15 gene clusters coding for phosphoenolpyruvate sugar-phosphotransferase systems (PTS), including mannose-, fructose- lactose-, mannitol-, maltose-, sucrose-, cellobiose- and ascorbate-PTS. Among the best characterized virulence factors of *S. mutans*, the *gtfB*, *gtfC* and *gtfD* genes encoding for the glucosyltransferases

responsible for the synthesis of the exopolysaccharide glucan from sucrose, and the genes encoding for the glucan-binding proteins GbpA, GbpB and GbpC and GbpD are all present in the LAR01 chromosome. In addition to Cbm, the genome of LAR01 encodes 7 other proteins with cell-anchoring LPxTG motifs, including the widely distributed *dexA*, *fruA*, *gbpC*, *spaP*, *wapA* and *wapE* genes. Also, the gene S_00879 coding for 742 aminoacid LPxTG protein annotated as a putative 5'-nucleotidase precursor with a conserved N-terminal metallophosphatase domain was found in most of the sequenced strains (*nadN*). However, this open reading frame is truncated at amino acid 704, only 2 amino acids apart from the downstream LPxTG motif, in a few sequenced strains, including in UA159 (Smu_1213c; *nadN*) (Fig. S1).

The LAR01 genome was also found to harbor genes associated with competence pathways including the transcriptional regulator *comR*, *comS* (ComX-inducing peptide, XIP), *comC* (competence-stimulating peptide, CSP), the competence factor *comX*, and all late competence genes required for natural transformation, which confirms previous studies that the CSP- and XIP-signaling pathways are widespread in *S. mutans*. Indeed, during genetic manipulations of LAR01 for the construction of *cbm* deletion (*cbm*) and *cbm* complemented strains, we observed that natural transformation of LAR01 increased by several orders of magnitude in the presence of synthetic CSP or XIP (data not shown). In addition, the LAR01 genome harbors two previously characterized mutacins ⁵¹: non-lantibiotic mutacin VI and lantibiotic mutacin K8 (Table 3).

Genetic organization of the *cbm* locus

The *cbm* gene in LAR01 encodes a 56 kDa protein with a conserved N-terminal collagen-binding domain (A-domain) and a threonine-rich C-terminus domain (B-domain) ¹⁸. The Cbm protein of LAR01 shared 80% similarity with the OMZ175 Cnm protein with the largest degree of conservation occurring at the collagen-binding A-domain (Fig. S2). Interestingly, the 3' end of the *cbm* gene in LAR01 is flanked by the same core genes as *cnm* in OMZ175 and in other *cnm*⁺ strains ^{17, 27, 28} (Fig. 2). Specifically, after the *cnm* stop codon is the *pgfS* gene, which encodes for a glycosyltransferase involved in the post-translational modification of Cnm ⁴⁰. When using the glycosylation prediction software NetOGlyc ⁵², like the Cnm B-domain, the threonine-rich B-domain of Cbm protein is predicted to have nearly 100 threonine/serine residues subjected to undergo modification through O-glycosylation (Fig. S3). More recently, we discovered that the three genes immediately downstream *pgfS*, named *pgfM1*, *pgfE* and *pgfM2*, are part of the *pgf* operon and are also involved in the modification of Cnm (data not shown). In LAR01, *cbm* is followed by *pgfS*, *pgfM1* and *pgfE* but a large 29.1 kb DNA insertion encoding 30 additional genes separates *pgfE* from *pgfM2* (Fig. 2). Interestingly, genes coding for components of a type VII seven secretion system (T7SS) are found within this DNA insertion between the core genes *pgfE* and *pgfM2*.

Characterization of virulence-related traits in LAR01

The cariogenic potential of *S. mutans* is strongly associated with its ability to form biofilms on tooth surfaces and to tolerate environmental stresses, particularly acid and oxidative stresses. Here, we used the UA159 (serotype *c*, *cnm*⁻/*cbm*⁻) and OMZ175 strains (serotype

f. cnm⁺) as benchmark to compare the biofilm-forming and stress tolerance capacities of LAR01. Despite the fact that all three strains possess the surface proteins associated with sucrose-independent (AtlA, BrpA, SpaP, WapA, among others) and sucrose-dependent (GtfB, GtfC, GtfD, GbpA, GbpB, among others) biofilm formation^{24, 53}, both LAR01 and OMZ175 showed a modest but statistically significant increased ability to form biofilms in the presence of either glucose or sucrose when compared to the UA159 strain (Fig. 3A). Next, we compared the capacities of each strain to survive acid and H₂O₂ (peroxide) challenges. All three strains showed similar survival rates under the different stress conditions tested (Fig. 3B–C).

The genome of LAR01 encodes genes for mutacin VI and K8 (Table 3). Using UA159 and OMZ175 as benchmark strains, we first tested the capacity of each individual strain to inhibit the growth of another. We found that LAR01 and OMZ175 can inhibit the growth of UA159 whereas LAR01 and OMZ175 modestly inhibit the growth of each other (Fig. 3D). On the other hand, UA159 was unable to inhibit the growth of LAR01 or OMZ175. Next, we investigated the ability of LAR01 to inhibit the growth of *S. gordonii* (a closely-related species), a target of mutacin VI, and *L. lactis*, which serves as indicator strains for wide spectrum mutacins such as mutacin K8. Growth of *S. gordonii* was modestly inhibited by all three strains (Fig. 3E) even though UA159 produces another non-lantibiotic (mutacin IV) active against *S. gordonii*. Finally, UA159 and OMZ175 can also synthesize mutacin V, a non-lantibiotic mutacin active against *L. lactis*. Growth of *L. lactis* was strongly inhibited by both strains with a slightly larger zone of inhibition surrounding OMZ175 when compared to UA159 (Fig. 3F), possibly because OMZ175 also produces mutacin I. Finally, LAR01 could also inhibit *L. lactis* growth (Fig. 3F) albeit not to the same extent as the UA159 and OMZ175 strains suggesting that mutacin K8 may not be as potent as mutacin V against *L. lactis*. Of note, inhibition zones are no longer seen in the presence of aminopeptidase (Fig. S4), thus confirming that mutacins are responsible for the growth inhibition observed.

Characterization of the LAR01 *cbm* strain and heterologous expression of Cbm in *Lactococcus lactis*

Because the collagen-binding domain (CBD) of Cbm shares high levels (80%) of similarity with its Cnm counterpart, we expected the anti-Cnm antibody, raised against the Cnm CBD, to cross-react with Cbm. Indeed, Western-blot analysis identified a band of ~150 kDa in the LAR01 lysate that was absent in the *cbm* strain but restored in the complemented strain, although with lower protein expression levels than those of the wild-type (Fig. 4A). Since the expression plasmid pMSP3535 is leaky and induction of the complemented strain with nisin (16 ng ml⁻¹) did not seem to enhance protein expression (Fig. 4A) nor its ability to bind collagen and laminin (data not shown), all subsequent experiments were performed using non-induced cultures.

Similar to Cnm, strains expressing Cbm have been shown to avidly bind to collagen and laminin and mediate invasion of endothelial cells¹⁷. In fact, LAR01 was able to bind collagen or laminin and invade HCAECs at levels that were similar to those found for the Cnm⁺ OMZ175 strain (Fig. 4B and C). Collagen or laminin binding and HCAECs invasion were drastically reduced in the *cbm* strain but fully or partially restored by *in trans*

complementation using the *pcbm* (Fig. 4B and C). Finally, the mortality rates of *G. mellonella* larvae infected with the *cbm* strain was significantly lower when compared to the parental LAR01 strain (Fig. 4D). In addition, consistent with the previous findings, complementation of *cbm* partially restored the killing behavior of the defective strain in this infection model. Of note, LAR01 was significantly more virulent in the *G. mellonella* model than the Cnm⁺ OMZ175 strain. Lastly, real time quantitative PCR analysis confirms the transcription of *cbm* in LAR01 and in the complemented strain (Fig. S5).

To further attribute to Cbm the virulent-associated properties observed in LAR01, we also introduced *pcbm* into the heterologous non-pathogenic host *L. lactis* NZ9800 creating the Cbm⁺ strain NZ9800+*pcbm*. Expression of Cbm at the cell surface of NZ9800+*pcbm* was confirmed by immunofluorescence labelling using the anti-Cnm antibody (Fig. 5A). As expected, NZ9800+*pcbm* showed increased collagen binding activity and virulence in *G. mellonella* when compared to the parent strain harboring the empty pMSP3535 plasmid (Fig. 5B and C).

Discussion

Cbm has been identified in several *S. mutans* isolates, particularly in serotype *k* strains associated with extra-oral infections (e.g. bacteremia, infective endocarditis)^{17, 26}. While the complete genomes of one serotype *k* strain and 8 *cnm*⁺ strains from serotypes *c*, *e* and *f* are available (<https://www.ncbi.nlm.nih.gov/genome/genomes/856>), no complete genome sequence was available for Cbm⁺ *S. mutans* strains until now. Even though, no unique genes were found in LAR01, we show that the Cbm⁺ LAR01 retained the traits associated with *S. mutans* cariogenicity but it also has the uncommon ability to bind to collagen, invade host endothelial cells and enhance systemic virulence. Thus, it seems that strains expressing CBPs like Cnm and Cbm can have an expanded niche that would allow for colonization of different oral sites as well as provide a colonization advantage in extra-oral sites^{6, 54}.

In addition to the cell characterized surface proteins, LAR01 possesses the cell wall anchored protein S_00879, coding for a putative bifunctional 5' nucleotidase/metallophosphatase (*nadN*) and this gene is found in most *S. mutans* genomes. In few strains like UA159 and KK21, this gene is truncated 2 aminoacids before the LPxTG anchoring domain. By searching the NCBI databases, we found that homologues of this protein are present in closely-related streptococci as well as in staphylococci. A closer look at their deduced amino acid sequences indicated the presence of the LPxTG-anchoring motif in all strains from other species. The consequences of the loss of the cell wall-anchoring motif of this enzyme in a subpopulation of *S. mutans* strains remains to be determined.

The observation that LAR01 harbors a T7SS as part of a 29.1 kb DNA insertion in the *pgf* glycosylation machinery operon was surprising. T7SS was initially identified in *Mycobacterium tuberculosis* but it is also found in other Gram-positive bacteria such as *Staphylococcus aureus*^{55–58}. In *M. tuberculosis*, T7SS was demonstrated to contribute to virulence and host cell membrane lysis⁵⁹. In *S. aureus*, in addition to contribute to the pathogenesis of abscesses and to induce apoptosis in epithelial cells with subsequent release of intracellular *S. aureus*^{60, 61}, this T7SS was also associated with growth inhibition of

competitor bacteria⁶². More recently, genes for this T7SS were identified in a small subset of *S. mutans* strains (8 out of 57)²⁴. However, this is the first time that T7SS genes are found in the genome of a *S. mutans* strain encoding bona-fide CBPs such as Cbm. While strains that contained the genes for the T7SS did not display greater virulence in the *G. mellonella* invertebrate model²⁴, additional studies are necessary to determine if the system is functional and, if so, identify effector proteins secreted through this pathway and their possible association with virulence or competition with other bacteria.

In order to thrive in dental plaque, *S. mutans* must outcompete commensal streptococci (e.g., mitis streptococci) as well as members of its own species⁶³. An important attribute of *S. mutans* is its capacity to produce of inhibitory substances such as lactic acid and bacteriocins. To this end, *S. mutans* strains produce a broad range of lantibiotic and non-lantibiotic bacteriocins, known as mutacins, although distribution of mutacins can greatly vary among strains⁶⁴. Our findings suggest that LAR01 produces mutacin VI and K8. Mutacin VI is non-lantibiotic mutacin widespread in *S. mutans* strains and active against closely-related species such as non-mutans streptococci. However, mutacin K8 is a wide spectrum lantibiotic mutacin found in a small number of strains⁶⁵. While the genome of UA159 does not encode wide-spectrum lantibiotics, mutacin K8 produced by LAR01 and mutacin I produced by OMZ175 are likely responsible for the growth inhibition of UA159. Another important virulence-associated traits of *S. mutans* is its ability to form robust biofilms on tooth surfaces, withstand acidic environment associated with the cariogenic environment and to cope with reactive oxygen species (ROS) that commensal streptococci use to fight back *S. mutans*⁶³. Here, we show that LAR01 is able to withstand acid and oxidative stresses just like our benchmark strains. However, LAR01 was able to form more biofilm biomass than the CBP⁻ strain UA159. Interestingly, Cnm⁺ were previously reported to form less biofilm biomass in general than UA159²⁴ but previous reports on Cbm⁺ *S. mutans* did not evaluate the biofilm phenotype of these strains. Thus, studies with additional strains are required to make an inference on whether Cbm⁺ strains form more biofilms in general.

Our findings confirm that LAR01 encodes a functional CBP that, like Cnm, migrates at a much higher molecular weight than the predicted size. Therefore, the presence of a threonine-rich repeat domain, the prediction that Cbm undergoes glycosylation, the genetic linkage of *cbm* with *pgf* genes, and the aberrant migration of Cbm strongly indicate that, like Cnm, Cbm undergoes post-translational glycosylation. In the Cnm⁺ OMZ175, the *pgf* glycosylation machinery is comprised of 4 genes (*pgfS*, *pgfM1*, *pgfE* and *pgfM2*) (data not shown) and all genes are essential for Cnm glycosylation. The fact that there is a 29.1kb insertion between *pgfE* and *pgfM2* indicates that *pgfM2* may not be essential for Cbm modification and/or that *pgfM2* may also work in concert with different glycosylation machineries. Further studies with *pgf* gene inactivations in LAR01 will reveal which genes are required for Cbm glycosylation. Also, as future genomes of Cbm⁺ strains become available, it would be valuable to determine whether this genetic arrangement is conserved among these strains.

The presence of secretory and protein anchoring enzymes (e.g. SecA, SrtA) combined with the low number of genes encoding secreted and surface-anchored proteins makes *L. lactis* an

ideal host for the individual characterization of surface virulence factors of pathogenic Gram-positive bacteria^{12, 66–68}. In fact, heterologous expression of Cnm in *L. lactis* enabled it to bind to collagen and laminin, invade HCAEC and conferred an advantage in endocardium colonization in an infective endocarditis model¹². Our findings suggest that Cbm, like Cnm, is indeed a surface expressed virulence factor responsible for collagen and laminin binding, invasion of endothelial cells and enhanced pathogenic potential in systemic infections.

Concluding remarks

Although a role for CBPs in caries has been suggested^{54, 69–71}, data associating oral colonization by CBP⁺ *S. mutans* with increased caries risk and poor caries outcomes is still scarce⁷². Nevertheless, oral colonization by this group of strains is associated with extra-oral diseases such as infective endocarditis^{12, 17}, hemorrhagic stroke⁴, cerebral microbleeds^{3, 15}, and nephropathy⁶⁹, indicating that CBP⁺ strains may be well suited to colonize other sites in the body. By binding to ECM proteins and invading human endothelial and epithelial cells, CBP⁺ *S. mutans* could survive within the host by colonizing soft tissues and escaping immune and therapeutic challenges⁶. Also, Cnm was shown to inhibit activation of the classical complement pathway through direct binding to the collagen-like domain of component C1q⁷³, and since Cbm has high degree of homology with Cnm, we speculate similar function for the former. Hence, expression of CBPs by *S. mutans* is an important and underestimated strategy contributing to the establishment and/or progression of recurrent and chronic systemic infections⁶.

In this report, we obtained the WGS of LAR01, the first Cbm⁺ strain of *S. mutans* to be fully sequenced. We also conducted a series of assays to compare the characteristics of LAR01 against other well-characterized strains and showed that Cbm contributed to different tissue colonization strategies and survival. Overall, the WGS of LAR01 provides further information to the pan-genome of *S. mutans* and grants further opportunities to understand the genomic fluidity of this important human pathogen.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Heran Getachew and Dr. Ann Progulke-Fox for kindly providing HCAECs for this study. This work was supported by NIDCR R01 DE022559. A.A-R was supported by NIDCR T90 DE21990 and NHLBI F31 HL124952. D.B.S. was supported by NIDCR T90 DE021985.

References

1. Selwitz RH, Ismail AI, Pitts NB. Dental caries. Lancet. 2007; 369(9555):51–9. [PubMed: 17208642]
2. Abercrombie, GFaSWM. A case of infective endocarditis due to *Streptococcus mutans*. The Lancet. 1928; 190:2340–9.

3. Watanabe I, Kuriyama N, Miyatani F, Nomura R, Naka S, Nakano K, et al. Oral Cnm-positive *Streptococcus mutans* Expressing Collagen Binding Activity is a Risk Factor for Cerebral Microbleeds and Cognitive Impairment. *Sci Rep.* 2016; 6:38561. [PubMed: 27934941]
4. Nakano K, Hokamura K, Taniguchi N, Wada K, Kudo C, Nomura R, et al. The collagen-binding protein of *Streptococcus mutans* is involved in haemorrhagic stroke. *Nat Commun.* 2011; 2:485. [PubMed: 21952219]
5. Abranches J, Miller JH, Martinez AR, Simpson-Haidaris PJ, Burne RA, Lemos JA. The collagen-binding protein Cnm is required for *Streptococcus mutans* adherence to and intracellular invasion of human coronary artery endothelial cells. *Infect Immun.* 2011; 79(6):2277–84. [PubMed: 21422186]
6. Aviles-Reyes A, Miller JH, Lemos JA, Abranches J. Collagen-binding proteins of *Streptococcus mutans* and related streptococci. *Mol Oral Microbiol.* 2017; 32(2):89–106. [PubMed: 26991416]
7. Nakano K, Nomura R, Taniguchi N, Lapirattanakul J, Kojima A, Naka S, et al. Molecular characterization of *Streptococcus mutans* strains containing the *cnm* gene encoding a collagen-binding adhesin. *Arch Oral Biol.* 2010; 55(1):34–9. [PubMed: 20005510]
8. Bowen WH, Koo H. Biology of *Streptococcus mutans*-derived glucosyltransferases: role in extracellular matrix formation of cariogenic biofilms. *Caries Res.* 2011; 45(1):69–86.
9. Koo H, Falsetta ML, Klein MI. The exopolysaccharide matrix: a virulence determinant of cariogenic biofilm. *J Dent Res.* 2013; 92(12):1065–73. [PubMed: 24045647]
10. Banas JA. Virulence properties of *Streptococcus mutans*. *Front Biosci.* 2004; 9:1267–77. [PubMed: 14977543]
11. Banas JA, Vickerman MM. Glucan-binding proteins of the oral streptococci. *Crit Rev Oral Biol Med.* 2003; 14(2):89–99. [PubMed: 12764072]
12. Freires IA, Aviles-Reyes A, Kitten T, Simpson-Haidaris PJ, Swartz M, Knight PA, et al. Heterologous expression of *Streptococcus mutans* Cnm in *Lactococcus lactis* promotes intracellular invasion, adhesion to human cardiac tissues and virulence. *Virulence.* 2017; 8(1):18–29. [PubMed: 27260618]
13. Misaki T, Naka S, Kuroda K, Nomura R, Shiooka T, Naito Y, et al. Distribution of *Streptococcus mutans* strains with collagen-binding proteins in the oral cavity of IgA nephropathy patients. *Clin Exp Nephrol.* 2015; 19(5):844–50. [PubMed: 25492252]
14. Miyatani F, Kuriyama N, Watanabe I, Nomura R, Nakano K, Matsui D, et al. Relationship between Cnm-positive *Streptococcus mutans* and cerebral microbleeds in humans. *Oral Dis.* 2015; 21(7):886–93. [PubMed: 26205098]
15. Tonomura S, Ihara M, Kawano T, Tanaka T, Okuno Y, Saito S, et al. Intracerebral hemorrhage and deep microbleeds associated with cnm-positive *Streptococcus mutans*; a hospital cohort study. *Sci Rep.* 2016; 6:20074. [PubMed: 26847666]
16. Nomura R, Naka S, Nemoto H, Inagaki S, Taniguchi K, Ooshima T, et al. Potential involvement of collagen-binding proteins of *Streptococcus mutans* in infective endocarditis. *Oral Dis.* 2013; 19(4):387–93. [PubMed: 22998492]
17. Nomura R, Naka S, Nemoto H, Otsugu M, Nakamura S, Ooshima T, et al. Potential high virulence for infective endocarditis in *Streptococcus mutans* strains with collagen-binding proteins but lacking PA expression. *Arch Oral Biol.* 2013; 58(11):1627–34. [PubMed: 24112728]
18. Nomura R, Nakano K, Naka S, Nemoto H, Masuda K, Lapirattanakul J, et al. Identification and characterization of a collagen-binding protein, Cbm, in *Streptococcus mutans*. *Mol Oral Microbiol.* 2012; 27(4):308–23. [PubMed: 22759315]
19. 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):416–24. [PubMed: 26443288]
20. Lapirattanakul J, Nakano K, Nomura R, Leelataweewud P, Chalermarp N, Klaophimai A, et al. Multilocus sequence typing analysis of *Streptococcus mutans* strains with the *cnm* gene encoding collagen-binding adhesin. *Journal of medical microbiology.* 2011; 60(Pt 11):1677–84. [PubMed: 21680768]
21. Lapirattanakul J, Nomura R, Nemoto H, Naka S, Ooshima T, Nakano K. Multilocus sequence typing of *Streptococcus mutans* strains with the *cbm* gene encoding a novel collagen-binding protein. *Arch Oral Biol.* 2013; 58(8):989–96. [PubMed: 23506778]

22. Napimoga MH, Kamiya RU, Rosa RT, Rosa EA, Hofling JF, Mattos-Graner R, et al. Genotypic diversity and virulence traits of *Streptococcus mutans* in caries-free and caries-active individuals. *Journal of medical microbiology*. 2004; 53(Pt 7):697–703. [PubMed: 15184543]
23. Mattos-Graner RO, Napimoga MH, Fukushima K, Duncan MJ, Smith DJ. Comparative analysis of Gtf isozyme production and diversity in isolates of *Streptococcus mutans* with different biofilm growth phenotypes. *J Clin Microbiol*. 2004; 42(10):4586–92. [PubMed: 15472313]
24. Palmer SR, Miller JH, Abranches J, Zeng L, Lefebure T, Richards VP, et al. Phenotypic heterogeneity of genomically-diverse isolates of *Streptococcus mutans*. *PloS one*. 2013; 8(4):e61358. [PubMed: 23613838]
25. Nagata E, Okayama H, Ito HO, Yamashita Y, Inoue M, Oho T. Serotype-specific polysaccharide of *Streptococcus mutans* contributes to infectivity in endocarditis. *Oral Microbiol Immunol*. 2006; 21(6):420–3. [PubMed: 17064403]
26. Nakano K, Nomura R, Matsumoto M, Ooshima T. Roles of oral bacteria in cardiovascular diseases--from molecular mechanisms to clinical cases: Cell-surface structures of novel serotype k *Streptococcus mutans* strains and their correlation to virulence. *J Pharmacol Sci*. 2010; 113(2):120–5. [PubMed: 20501965]
27. Song L, Wang W, Conrads G, Rheinberg A, Sztajer H, Reck M, et al. Genetic variability of mutans streptococci revealed by wide whole-genome sequencing. *BMC Genomics*. 2013; 14:430. [PubMed: 23805886]
28. Cornejo OE, Lefebure T, Bitar PD, Lang P, Richards VP, Eilertson K, et al. Evolutionary and population genomics of the cavity causing bacteria *Streptococcus mutans*. *Mol Biol Evol*. 2013; 30(4):881–93. [PubMed: 23228887]
29. Gold OG, Jordan HV, Van Houte J. A selective medium for *Streptococcus mutans*. *Arch Oral Biol*. 1973; 18(11):1357–64. [PubMed: 4518755]
30. Chen Z, Saxena D, Caufield PW, Ge Y, Wang M, Li Y. Development of species-specific primers for detection of *Streptococcus mutans* in mixed bacterial samples. *FEMS Microbiol Lett*. 2007; 272(2):154–62. [PubMed: 17521362]
31. Shibata Y, Ozaki K, Seki M, Kawato T, Tanaka H, Nakano Y, et al. Analysis of loci required for determination of serotype antigenicity in *Streptococcus mutans* and its clinical utilization. *J Clin Microbiol*. 2003; 41(9):4107–12. [PubMed: 12958233]
32. Nakano K, Nomura R, Nakagawa I, Hamada S, Ooshima T. Demonstration of *Streptococcus mutans* with a cell wall polysaccharide specific to a new serotype, k, in the human oral cavity. *J Clin Microbiol*. 2004; 42(1):198–202. [PubMed: 14715753]
33. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014; 30(14):2068–9. [PubMed: 24642063]
34. Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, et al. A whole-genome assembly of *Drosophila*. *Science*. 2000; 287(5461):2196–204. [PubMed: 10731133]
35. Kremer BH, van der Kraan M, Crowley PJ, Hamilton IR, Brady LJ, Bleiweis AS. Characterization of the sat operon in *Streptococcus mutans*: evidence for a role of Ffh in acid tolerance. *J Bacteriol*. 2001; 183(8):2543–52. [PubMed: 11274114]
36. Kim JN, Stanhope MJ, Burne RA. Core-gene-encoded peptide regulating virulence-associated traits in *Streptococcus mutans*. *J Bacteriol*. 2013; 195(12):2912–20. [PubMed: 23603743]
37. van de Rijn I, Kessler RE. Growth characteristics of group A streptococci in a new chemically defined medium. *Infect Immun*. 1980; 27(2):444–8. [PubMed: 6991416]
38. Morrison DA, Khan R, Junges R, Amdal HA, Petersen FC. Genome editing by natural genetic transformation in *Streptococcus mutans*. *J Microbiol Methods*. 2015; 119:134–41. [PubMed: 26481669]
39. Bryan EM, Bae T, Kleerebezem M, Dunny GM. Improved vectors for nisin-controlled expression in gram-positive bacteria. *Plasmid*. 2000; 44(2):183–90. [PubMed: 10964628]
40. Avilés-Reyes A, Miller JH, Simpson-Haidaris PJ, Hagen FK, Abranches J, Lemos JA. Modification of *Streptococcus mutans* Cnm by PgfS contributes to adhesion, endothelial cell invasion, and virulence. *J Bacteriol*. 2014; 196(15):2789–97. [PubMed: 24837294]

41. Wen ZT, Burne RA. LuxS-mediated signaling in *Streptococcus mutans* is involved in regulation of acid and oxidative stress tolerance and biofilm formation. *J Bacteriol.* 2004; 186(9):2682–91. [PubMed: 15090509]
42. Qi F, Kreth J. Methods to Study Antagonistic Activities Among Oral Bacteria. *Methods Mol Biol.* 2017; 1537:203–18. [PubMed: 27924596]
43. Aviles-Reyes A, Miller JH, Simpson-Haidaris PJ, Lemos JA, Abranches J. Cnm is a major virulence factor of invasive *Streptococcus mutans* and part of a conserved three-gene locus. *Mol Oral Microbiol.* 2014; 29(1):11–23. [PubMed: 24103776]
44. Colomer-Winter C, Gaca AO, Lemos JA. Association of Metal Homeostasis and (p)ppGpp Regulation in the Pathophysiology of *Enterococcus faecalis*. *Infect Immun.* 2017; 85(7)
45. Abranches J, Chen YY, Burne RA. Galactose metabolism by *Streptococcus mutans*. *Appl Environ Microbiol.* 2004; 70(10):6047–52. [PubMed: 15466549]
46. Maruyama F, Kobata M, Kurokawa K, Nishida K, Sakurai A, Nakano K, et al. Comparative genomic analyses of *Streptococcus mutans* provide insights into chromosomal shuffling and species-specific content. *BMC Genomics.* 2009; 10:358. [PubMed: 19656368]
47. Biswas S, Biswas I. Complete genome sequence of *Streptococcus mutans* GS-5, a serotype *c* strain. *J Bacteriol.* 2012; 194(17):4787–8. [PubMed: 22887682]
48. Aikawa C, Furukawa N, Watanabe T, Minegishi K, Furukawa A, Eishi Y, et al. Complete genome sequence of the serotype *k* *Streptococcus mutans* strain LJ23. *J Bacteriol.* 2012; 194(10):2754–5. [PubMed: 22535936]
49. Kondo Y, Nishimata H, Hidaka K, Hasuwa T, Moriuchi H, Fujiwara T, et al. Draft Genome Sequence of a Clinical Isolate of *Streptococcus mutans* Strain HM. *Genome Announc.* 2017; 5(33)
50. Meng P, Lu C, Zhang Q, Lin J, Chen F. Exploring the Genomic Diversity and Cariogenic Differences of *Streptococcus mutans* Strains Through Pan-Genome and Comparative Genome Analysis. *Curr Microbiol.* 2017
51. Merritt J, Qi F. The mutacins of *Streptococcus mutans*: regulation and ecology. *Mol Oral Microbiol.* 2012; 27(2):57–69. [PubMed: 22394465]
52. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* 2013; 32(10):1478–88. [PubMed: 23584533]
53. Ajdic D, McShan WM, McLaughlin RE, Savic G, Chang J, Carson MB, et al. Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. *Proc Natl Acad Sci U S A.* 2002; 99(22):14434–9. [PubMed: 12397186]
54. Miller JH, Aviles-Reyes A, Scott-Anne K, Gregoire S, Watson GE, Sampson E, et al. The collagen binding protein Cnm contributes to oral colonization and cariogenicity of *Streptococcus mutans* OMZ175. *Infect Immun.* 2015; 83(5):2001–10. [PubMed: 25733523]
55. Bosserman RE, Champion PA. Esx Systems and the Mycobacterial Cell Envelope: What's the Connection? *J Bacteriol.* 2017; 199(17)
56. Aly KA, Anderson M, Ohr RJ, Missiakas D. Isolation of a membrane protein complex for type VII secretion in *Staphylococcus aureus*. *J Bacteriol.* 2017
57. Abdallah AM, Gey van Pittius NC, Champion PA, Cox J, Luirink J, Vandenbroucke-Grauls CM, et al. Type VII secretion--mycobacteria show the way. *Nat Rev Microbiol.* 2007; 5(11):883–91. [PubMed: 17922044]
58. Ates LS, Houben EN, Bitter W. Type VII Secretion: A Highly Versatile Secretion System. *Microbiol Spectr.* 2016; 4(1)
59. Conrad WH, Osman MM, Shanahan JK, Chu F, Takaki KK, Cameron J, et al. Mycobacterial ESX-1 secretion system mediates host cell lysis through bacterium contact-dependent gross membrane disruptions. *Proc Natl Acad Sci U S A.* 2017; 114(6):1371–6. [PubMed: 28119503]
60. Burts ML, Williams WA, DeBord K, Missiakas DM. EsxA and EsxB are secreted by an ESAT-6-like system that is required for the pathogenesis of *Staphylococcus aureus* infections. *Proc Natl Acad Sci U S A.* 2005; 102(4):1169–74. [PubMed: 15657139]
61. Korea CG, Balsamo G, Pezzicoli A, Merakou C, Tavarini S, Bagnoli F, et al. Staphylococcal Esx proteins modulate apoptosis and release of intracellular *Staphylococcus aureus* during infection in epithelial cells. *Infect Immun.* 2014; 82(10):4144–53. [PubMed: 25047846]

62. Cao Z, Casabona MG, Kneuper H, Chalmers JD, Palmer T. The type VII secretion system of *Staphylococcus aureus* secretes a nuclease toxin that targets competitor bacteria. *Nat Microbiol.* 2016; 2:16183. [PubMed: 27723728]
63. Lemos JA, Quivey RG Jr, Koo H, Abranches J. *Streptococcus mutans*: a new Gram-positive paradigm? *Microbiology.* 2013; 159(Pt 3):436–45. [PubMed: 23393147]
64. Bekal-Si Ali S, Hurtubise Y, Lavoie MC, LaPointe G. Diversity of *Streptococcus mutans* bacteriocins as confirmed by DNA analysis using specific molecular probes. *Gene.* 2002; 283(1–2):125–31. [PubMed: 11867219]
65. Robson CL, Wescombe PA, Klesse NA, Tagg JR. Isolation and partial characterization of the *Streptococcus mutans* type AII lantibiotic mutacin K8. *Microbiology.* 2007; 153(Pt 5):1631–41. [PubMed: 17464078]
66. Morello E, Bermudez-Humaran LG, Llull D, Sole V, Miraglio N, Langella P, et al. *Lactococcus lactis*, an efficient cell factory for recombinant protein production and secretion. *J Mol Microbiol Biotechnol.* 2008; 14(1–3):48–58. [PubMed: 17957110]
67. Danne C, Entenza JM, Mallet A, Briandet R, Debarbouille M, Nato F, et al. Molecular characterization of a *Streptococcus gallolyticus* genomic island encoding a pilus involved in endocarditis. *J Infect Dis.* 2011; 204(12):1960–70. [PubMed: 22043018]
68. Piroth L, Que YA, Widmer E, Panchaud A, Piu S, Entenza JM, et al. The fibrinogen- and fibronectin-binding domains of *Staphylococcus aureus* fibronectin-binding protein A synergistically promote endothelial invasion and experimental endocarditis. *Infect Immun.* 2008; 76(8):3824–31. [PubMed: 18541660]
69. Misaki T, Naka S, Hatakeyama R, Fukunaga A, Nomura R, Isozaki T, et al. Presence of *Streptococcus mutans* strains harbouring the *cnm* gene correlates with dental caries status and IgA nephropathy conditions. *Sci Rep.* 2016; 6:36455. [PubMed: 27811984]
70. Nomura R, Ogaya Y, Nakano K. Contribution of the Collagen-Binding Proteins of *Streptococcus mutans* to Bacterial Colonization of Inflamed Dental Pulp. *PloS One.* 2016; 11(7):e0159613. [PubMed: 27442266]
71. Nomura R, Nakano K, Taniguchi N, Lapidattanakul J, Nemoto H, Gronroos L, et al. Molecular and clinical analyses of the gene encoding the collagen-binding adhesin of *Streptococcus mutans*. *Journal of medical microbiology.* 2009; 58(Pt 4):469–75. [PubMed: 19273643]
72. Esberg A, Sheng N, Marell L, Claesson R, Persson K, Boren T, et al. *Streptococcus mutans* Adhesin Biotypes that Match and Predict Individual Caries Development. *EBioMedicine.* 2017; 24:205–15. [PubMed: 28958656]
73. Kang M, Ko YP, Liang X, Ross CL, Liu Q, Murray BE, et al. Collagen-binding microbial surface components recognizing adhesive matrix molecule (MSCRAMM) of Gram-positive bacteria inhibit complement activation via the classical pathway. *J Biol Chem.* 2013; 288(28):20520–31. [PubMed: 23720782]

Figure 1.

Circular representation of the genome of *Streptococcus mutans* LAR01. Outermost circle indicates the distances from the putative origin of replication. Blue circle represent genes in the forward (+) and red circle represent genes in the reverse strand (-). Secreted genes are represented in green whereas orange indicated genes that are unique to LAR01 in relation to UA159. Innermost circle represents the % G + C with more and less than average shown as light blue and purple, respectively.

Figure 2.

Genetic arrangement of *cbm* locus in LAR01 compared to the *cbm*⁻/*cnm*⁻ strain UA159 and the *cnm*⁺ OMZ175. Like *cnm* in OMZ175, *cbm* is flanked by the *pgf* glycosylation machinery in LAR01. Interestingly, the last two genes of the *pgf* system (*pgfE*-*pgfM2*) are separated by an insertion of 29.1 Kb that contain genes coding for a putative type VII secretion system (T7SS).

Figure 3.

Comparative analysis of the expression of virulence-related traits in LAR01, OMZ175 and UA159. (A) 24h biofilm formation of strains UA159, OMZ175 and LAR01 on the surface of microtiter plates in the presence of glucose or sucrose. (B) Percent survival of unadapted (open symbols) and acid-adapted (closed symbols) cells of UA159, OMZ175 and LAR01 after exposure to pH 3.0. (C) Percent survival of UA159, OMZ175 and LAR01 exposed to 0.2% H₂O₂. (D) Growth competition between *S. mutans* strains. Absence of growth inhibition after addition of aminopeptidase confirmed that inhibition zones were due to mutacin activity. (E–F) Deferred antagonism assays. Cultures of *S. mutans* were grown on BHI agar plates for 24 h incubation and a soft BHI agar overlay containing *S. gordonii* (E) or *L. lactis* (F) cells was incubated for additional 48 h before zone of inhibition was recorded. All experiments were performed at least in triplicate and asterisks indicate statistical significance ($p \leq 0.05$).

Figure 4.

Phenotypic characterization of the *cbm* mutant in LAR01. (A) Western blot analysis using anti-Cnm antibody. (B) Binding to collagen type I and laminin *in vitro*. (C) Invasion of HCAECs and (D) killing of *G. mellonella*. All experiments were performed at least in triplicate and asterisks indicate statistical significance ($p < 0.05$) between *cbm* and the wild-type LAR01 or between *cbm* and the genetically complemented *cbm* ($C\ cbm$).

Figure 5. Heterologous expression of Cbm in *L. lactis* NZ9800. (A) Immunofluorescence microscopy analysis demonstrated that Cbm was effectively expressed and translocated to the cell surface of *L. lactis*. (B) Expression of Cbm in *L. lactis* enables binding to collagen type I. (C) Expression of Cbm in NZ9800 contributes to systemic virulence in *G. mellonella*. (*) $p < 0.05$.

Table 1

Strains used in this study.

Strains	Relevant genotype	Source
<i>S. mutans</i>		
LAR01	Wild-type, serotype <i>e</i> , <i>cbm</i> ⁺	This study
LAR01 <i>cbm</i>	<i>cbm</i> ⁻ , <i>kan</i> ^R	This study
C <i>cbm</i>	complemented <i>cbm</i> , <i>kan</i> ^R <i>erm</i> ^R hosting <i>pcbm</i>	This study
OMZ175	Wild-type, serotype <i>f</i> , <i>cnm</i> ⁺	Laboratory collection
UA159	Wild-type, serotype <i>c</i>	Laboratory collection
B14	Wild-type, serotype <i>e</i> , <i>cnm</i> ⁺	Laboratory collection
OM88X	Wild-type, serotype <i>k</i>	Laboratory collection
<i>S. gordonii</i>		
DL-1	Wild-type	Laboratory collection
<i>L. lactis</i>		
ATCC	Wild-type	Laboratory collection
NZ9800	Wild-type	Laboratory collection
NZ9800- <i>cbm</i>	<i>cbm</i> ⁺ , <i>erm</i> ^R , hosting <i>pcbm</i>	This study

Table 2

Primers used in this study.

Primer	Sequence ^a	Application	Reference
Sm479-F	5'-TCGCGAAAAAGATAACAAACA-3'	Detection of <i>S. mutans</i>	30
Sm479R	5'-GCCCTTCACAGTTGGTTAG-3'	Detection of <i>S. mutans</i>	30
SC-F	5'-CGGAGTGCTTTTACAAGTGCTGG-3'	Detection of serotype <i>c</i>	30
SC-R	5'-AACCACGCCAGCAAACCTTTAT-3'	Detection of serotype <i>c</i>	30
SE-F	5'-CCTGCTTTTCAAGTACCTTTCGCC-3'	Detection of serotype <i>e</i>	30
SE-R	5'-CTGCTTGCCAAGCCCTACTAGAAA-3'	Detection of serotype <i>e</i>	30
SF-F	5'-CCCACAATTGGCTTCAAGAGGAGA-3'	Detection of serotype <i>f</i>	30
SF-R	5'-TGCGAAACCATAAGCATAGCGAGG-3'	Detection of serotype <i>f</i>	30
CEFK-F	5'-ATTCCCGCGTTGGACCATCC-3'	Detection of serotype <i>k</i>	30
K-R	5'-CCAATGTGATTATCCCATCAC-3'	Detection of serotype <i>k</i>	30
cbm-EF	5'-AGCTGAAGTTAGTGTGTAAC ACCTGCTTC-3'	<i>cbm</i> screening	18
cbm-ER	5'-TAGGATCATCAACCTTAGTCAAGTACACGA-3'	<i>cbm</i> screening	18
cbm-F2	5'-AATCTCTTGCATGCTTTACGTTGG-3'	<i>cbm</i> inactivation	This study
cbm-R2	5'-ATTTTGTTCTGCAGATGACTTGTT-3'	<i>cbm</i> inactivation	This study
cbm-F3	5'-AAGTAGTGCTGCAGAACCTGCTTC-3'	<i>cbm</i> inactivation	This study
cbm-R3	5'-ATTTTGTTTCATATGATGACTTGTT-3'	<i>cbm</i> inactivation	This study
cbm-F	5'-GATGGTACCTATGTTGATTG-3'	Real time quantitative PCR	This study
cbm-R	5'-CCGGTAACGTTATGGAGATTATTG-3'	Real time quantitative PCR	This study
cbmBamHI-F	5'-GAAAGGACGGATCCATGAAAA GAA-3'	Complementation and heterologous expression	This study
cbmXbaIR	5'-GTTTTCAATCTAGATCAGCTATG-3'	Complementation and heterologous expression	This study

^aUnderline sequences indicate restriction sites used for cloning purposes.

Table 3Mutacin genes present in *S. mutans*.

Group	LAR01	Comparative producer	Other producer strains
Lantibiotics			
Mutacin I	–	IID3	OMZ175, SF14, U2A, 66-2A, SM1
Mutacin II	–	T8	-
Mutacin III/1140	–	R221	UA787, JH1140
Mutacin K8	+	K8	NN2025
Non-lantibiotics			
Mutacin IV (nlmA)	–	UA159	NVAB, GS-5, 11A1, 81D3, 3SN1, 2ST1, ATCC 25175
Mutacin IV (nlmB)	–	UA159	GS-5, A9, T4, S1B, KK23
Mutacin V (nlmC)	–	UA159	OMZ175, LJ23, S1B, 5DC8, KK21
Mutacin VI (nlmD)	+	UA159	NG8, KCOM1054, GS-5, NN2025