UPPSALA
UNIVERSITET

Isolation of *Streptococcus salivarius* from human oral samples and *In vivo* recombination cloning of EAL 2 of *Streptococcus uberis* C6344

Thamida Tauhid

Master Degree Project in Infection Biology, 45 credits. Semester Spring 2022
Department: Microbiology, Tumor and Cell Biology, Karolinska Institutet
Supervisor: Prof. Ute Römling

Abstract

The second messenger cyclic diguanylate monophosphate (c-di-GMP) has been proven to be a central regulator for physiological and metabolic processes including biofilm formation and sessile to motile transitioning (1,2). The synthesis and degradation of c-di-GMP are regulated by GGDEF- respectively EAL-domain proteins. Recently, c-di-GMP has been discovered in the Gram-positive *Streptococcus* genus including *Streptococcus gallolyticus,* which showed to have diguanylate cyclase activity (3). Characterisation of the c-di-GMP network in other *Streptococcus* is of relevance. Hence, the aim of this project was the assessment of the GGDEF- and EAL domains from the animal pathogenic *Streptococcus uberis* and *Streptococcus henryi*. *In vivo* recombination cloning was used for the analysis of the GGDEF, EAL and GGDEF-EAL domain proteins from *S. uberis* and *S. henryi*. The cloning was unsuccessful for most of the domain proteins, except, for EAL 2 of *S. uberis*. However, analysis of the sequencing results for the cloned EAL 2 presented mutations. Further studies testing alternative cloning methods should be applied. Research regarding probiotic streptococci is also of interest. Therefore, isolation of *Streptococcus salivarius* from human oral samples using Streptococcus Selection Agar was conducted. Isolation of *S. salivarius* from human saliva and tongue samples was successful using Streptococcus Selection Agar. Other *Streptococcus* spp., *Lactobacillus*, *Staphylococcus*, and additional bacterial species were also isolated.

Key words:

Streptococcus uberis C6344; *Streptococcus henryi*; c-di-GMP; GGDEF domain protein; EAL domain protein; *In vivo* recombination cloning; *Streptococcus salivarius*; Streptococcus Selection Agar.

Popular scientific summary

Bacteria are single-celled microorganisms that exist in various environments, including volcanic vents, human microflora, and the hospital environment. They come in all shapes and sizes and can be both hostile and friendly. An example of one hostile bacterium is the one commonly known to cause sore throat, the Group A *Streptococcus*. Another harmful bacterium that causes disease in farm animals is *Streptococcus uberis*. However, there are plentiful friendly and beneficial bacteria including *Streptococcus salivarius*, that are part of the human oral flora. For years scientists have researched the characteristics, mechanisms, and abilities of these microorganisms by studying their genes and proteins. A particular set of proteins called GGDEF diguanylate cyclase and EAL phosphodiesterase constitute the most abundant bacterial domain families, as the proteins containing these domains control the supply of an essential molecule known as cyclic diguanylate monophosphate (c-di-GMP).

In this study, these two classes of domains from bacteria including *Streptococcus uberis* and two other potentially pathogenic bacteria, named *Streptococcus henryi* and *Pseudomonas aeruginosa* were analysed. One way of studying the function of these genes and proteins is by a method called cloning which utilizes small circular double-stranded DNA, called plasmids. Cloning is used to express the relevant gene, which can be used to analyse its function. The cloning was only successfully performed for the gene encoding EAL phosphodiesterase proteins of *Streptococcus uberis* and was cloned into *E. coli* DH5α. However, changes in the genes were detected in the clone, hence further analysis could not be performed.

Hence, our quest for understanding bacteria continued and finding *Streptococcus salivarius* from various sources was carried out. Saliva, tongue, and faeces samples from humans were taken and environmental samples including water, wood, cheese, moss, and horse faeces were collected. The study was conducted using a media specifically made for growing *Streptococcus* bacteria. The results concluded that *Streptococcus salivarius* from human saliva and tongue samples could be isolated using this media and several other bacterial species were also isolated. Nevertheless, additional studies are needed to better understand these microorganisms.

List of abbreviations

A	Adenosine
BLASTN	Basic Local Alignment Search Tool
bp	Base pair
c-di-GMP	Cyclic dimeric guanosine monophosphate
DGC	Diguanylate cyclase
DMSO	Dimethyl sulfoxide
E. coli	*Escherichia coli*
E. faecalis	*Enterococcus faecalis*
G	Guanine
LB	Luria-Bertani
MH-F	Mueller-Hinton agar fastidious
MUSCLE	Multiple Sequence Comparison by Log- Expectation
NCBI	National Center for Biotechnology information
ORF	Open reading frame
P. aeruginosa	*Pseudomonas aeruginosa*
PAS	Per-ARNT-Sim
PCR	Polymerase chain reaction
PDE	Phosphodiesterase
rpm	Revolutions per minute
rRNA	Ribosomal ribonucleic acid
S. typhimurium	*Salmonella typhimurium*
S. aureus	*Staphylococcus aureus*
S. gallolyticus	*Streptococcus gallolyticus*
S. henryi	*Streptococcus henryi*
S. parauberis	*Streptococcus parauberis*
S. pyogenes	*Streptococcus pyogenes*
S. salivarius	*Streptococcus salivarius*
S. uberis	*Streptococcus uberis*
T	Tyrosine
XRD	X-ray diffraction analysis

Table of Contents

Introduction

Streptococcus is a genus with spherical Gram-positive cells consisting of pathogenic and probiotic bacteria, found in humans, animals and the environment (4,5). An example of one pathogenic *Streptococcus* is the opportunistic pathogen, *Streptococcus gallolyticus* which is part of the natural flora in the rumen of ruminants and can cause ruminal acidosis (6). The bacterium is also part of the human gut microbiome and is commonly associated with human colorectal cancer, endocarditis and bacteriemia (7–10). There are some speculations that there is a transmission of the bacteria between humans and farm animals via environmental factors or through diet (11,12).

Some streptococcal species found in animals are *Streptococcus uberis, Streptococcus parauberis* and *Streptococcus henryi*. *S. uberis* is a mastitis-causing pathogen that harms different cattle animals (13). The fish pathogen *S. parauberis* is known for causing disease in marine life and aquaculture (14,15). *S. henryi* is one of the lesser-known streptococci and it's involved in equine laminitis (16). However, not all streptococci cause life-threatening conditions, for instance, the probiotic *Streptococcus salivarius*. This probiotic bacterium mostly found in the oral cavity and gastrointestinal tract is a lactic acid bacterium that is prosperous for gut health (17,18). Some studies have proven the benefit of *S. salivarius* probiotic abilities against colonisation of pathogenic bacteria for instance *Streptococcus pyogenes*, in the oral cavity (19,20).

Bacteria including *Streptococcus* and *Lactobacillus* can be found in other habitats aside from humans and animals, including nutrient-rich environments like diary food, plants, and water (5,21,22). The presence of bacteria in food is most often a man-made habitat, as bacteria like *Streptococcus* and *Lactobacillus* are utilized for fermentation and probiotic purposes. However, studies have shown the transmission of bacteria from external habitats including water or wild animals to humans or domesticated animals (23). One example is the discovery of the transmission of the wild fish bacteria *Streptococcus iniae* to farm cultured fish (24). Hence, research comparing bacteria in macro-, and microenvironments and exploring the repertoire of genes and proteins in genera like *Streptococcus* are of interest and importance.

Phenotypic characteristics, the bacterial genome and biochemical assays are some elements used for bacterial classification and identification (25). One study used the 16S ribosomal RNA (RNA) for the classification of *S. henryi* (16). 16S rRNA of bacteria can be used for the identification and classification of bacteria due to its abundance in prokaryotic microorganisms and due to the sequence withstanding time (26). The 16S rRNA sequence consists of conserved and variable regions, where the variable regions are of interest for the differentiation of bacteria on species level or higher. The rRNA molecules are highly conserved due to their important role in the bacterial life cycle (27). Primers for amplification of 16S rRNA are often designed to include the different variable regions.

Furthermore, analysis of different signalling systems and essential molecules in bacteria is an area of interest. Second messenger systems have an essential role in biological processes, including various intracellular signalling, in bacteria. Second messengers are smaller molecules for example hydrophilic cyclic nucleotide molecules like cyclic dimeric guanosine monophosphate (c-di-GMP) (Figure 1). Exposure to extracellular molecules called first messengers, for example, hormones will trigger an intracellular signalling cascade involving the second messengers. This type of cellular communication using first and second messengers is an important feature in biological processes, because of its rapid amplification of signals (28–32).

Cyclic and linear nucleotide molecules are frequently used by bacteria, in signal transduction systems as second messengers. The most abundant intracellular second messenger molecule is the c-di-GMP (Figure 1), which was first identified as an allosteric activator and was discovered in 1987. C-di-GMP is a central regulator for several physiological and metabolic processes, including biofilm formation, motile to sessile transition, nutritional acquisition and more (1,2). The regulation of the synthesis of c-di-GMP is conducted by the GGDEF domain proteins. EAL- or HD-GYP domain proteins are responsible for the degradation of the cyclic molecule. GGDEF refers to the amino acid sequence, glycine-glycine-aspartate-glutamate-phenylalanine (33). EAL refers to the amino acid sequence, glutamate-alanine-leucine (28).

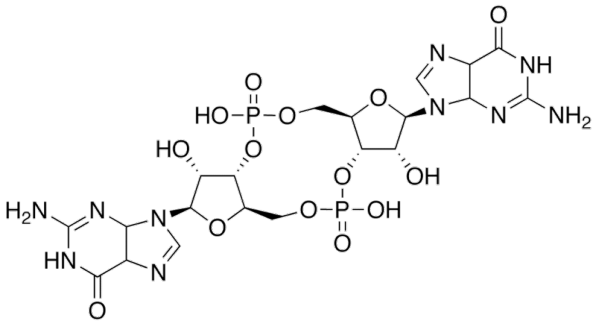

Figure 1. Chemical structure of c-di-GMP.
From By Jypx3 - Own work, Public Domain, https://commons.wikimedia.org/w/index.php?curid=8715696. Retrieved through open access.

The GGDEF domain proteins has diguanylate cyclase (DGC) catalytic activity while the EAL- and HD-GYP domain proteins have phosphodiesterase (PDE) activity (3,28,33). The activity of DGC entails the catalysis of a chemical reaction containing two guanine triphosphate molecules resulting in the synthesis of c-di-GMP (29). PDE:s activity performs the hydrolysis of phosphodiester bonds present in the c-di-GMP, resulting in the degradation of the cyclic nucleotide molecule into pGpG and/or GMP (Figure 2) (34). The regulation of DGC:s and PDE:s is accomplished by different N-terminal sensory domains, for example, the Per-ARNT-Sim (PAS) domain (35,36).

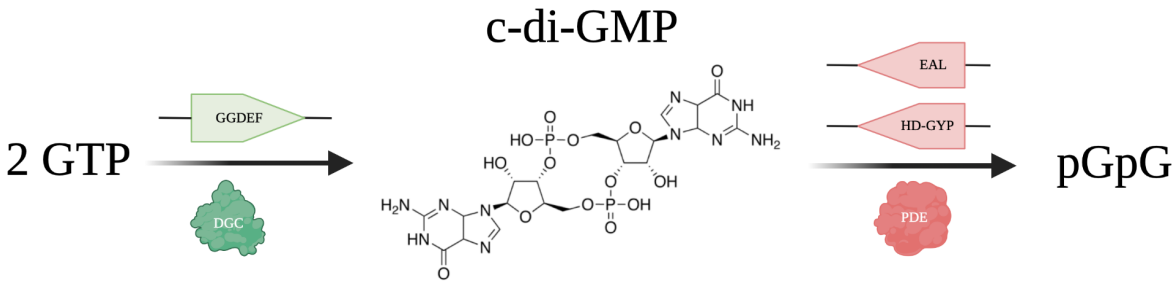

Figure 2. Chemical reaction of c-di-GMP.
Synthesis by GGDEF domain protein, diguanylate cycles (DGC), of c-di-GMP using two guanosine triphosphate molecules. Degradation of c-di-GMP by EAL- or HD-GYP domain protein, phosphodiesterase resulting in the linear diguanylate. The illustration was created with Biorender.com.

Several studies have investigated the cyclic dinucleotide-based second messenger signalling system in Gram-negative bacteria for instance in *Pseudomonas aeruginosa*, *Aeromonas veronii* and *Salmonella enterica serovar Typhimurium* (37–39). Recently the area regarding the GGDEF- and EAL domain proteins in Gram-positive bacteria is being explored and investigated. One study has proven that the biofilm in *Staphylococcus epidermidis* is regulated

by GGDEF domain proteins independent from c-di-GMP (40). In contrast, c-di-GMP was recently discovered in *S. gallolyticus* and it was proven that the PAS-GGDEF domain protein in the bacteria had catalytic DGC activity (3). Another cyclic dinucleotide molecule, cyclic di-adenosine monophosphate (41), has been investigated in several streptococcal species and, for example, shown to be involved in cellular processes connected to stress management in *S. pyogenes* (42). Thus, research regarding second messengers including c-di-GMP in bacteria, particularly, *Streptococcus* is of significance.

One method of studying the different signalling systems is by analysing the gene of interest, through mutagenesis and using various assays. Cloning can be used to express a certain gene and assess its function by conducting assays of choice. There are many different methods available for cloning bacterial genes, for example, conventional restriction enzyme cloning. The disadvantages of this method are the long incubation time and dependence on the availability of restriction endonucleases.

Another method is *in vivo* recombinant cloning which has been proven to be easier, faster, and cheaper than restriction enzyme cloning (43). The method requires chemically treated host bacteria, primers giving overlapping sequences and properly amplified PCR products. Some of the most used chemo-competent cells are *Escherichia coli* DH5α and TOP10 (44). Using chemicals including calcium chloride ($CaCl_2$) (45), entail alteration of the bacterial membranes, making them more permeable. The cells are then heat-shocked to facilitate the uptake of foreign DNA (46,47). One vector frequently used for this method is the pBAD28 vector. The vector contains two antibiotic cassettes, for ampicillin resistance and chloramphenicol resistance. The operon of arabinose and the regulatory gene *araC* has a promotor with the name of $P_{BAD,}$ which is included in most pBAD vectors (48). The primers used in this method, add overlapping segments onto the amplified gene and vector, enabling recombination of the gene insert into the plasmid (Figure 3). Furthermore, it is essential to use high-fidelity polymerases when performing this method, increasing the efficiency of the method (43,49).

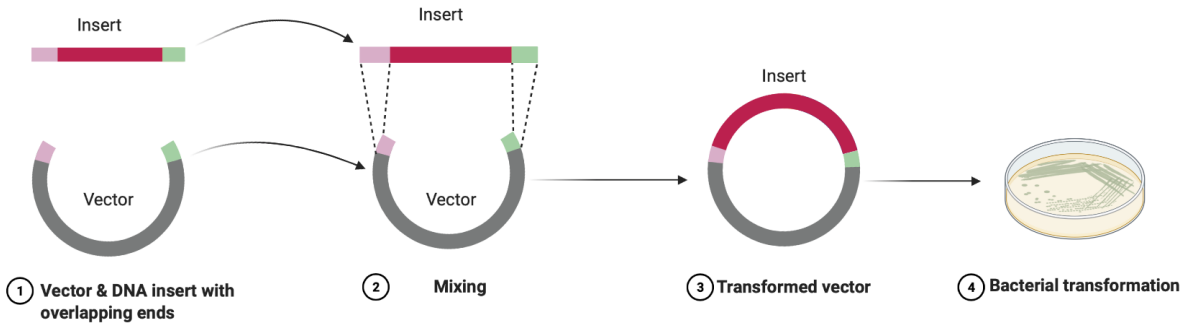

Figure 3. The process of *in vivo* recombination cloning

Illustration showcasing the process of *in vivo* recombination cloning independent of restriction enzymes. 1. Amplified vector and gene insert having overlapping segments. 2. The vector and the insert get mixed. 3. The insert and the vector can recombine. 4. The vector gets transformed into bacteria and is grown on selective agar. The illustration was created with Biorender.com

Aim

The study consisted of 4 projects. (1) Isolation of probiotic *S. salivarius* from human oral samples and investigation of the occurrence of *Streptococcus* species in various environmental habitats using Streptococcus Selection Agar. (2) Cloning and assessment of the GGDEF domain proteins and EAL domain proteins from *S. henryi* and *S. uberis*. Cloning and assessment of the GGDEF-EAL domain of *P. aeruginosa* PT31 were conducted as a side project. (3) Introduction of point mutations in the PAS domain, one containing the transmembrane and one without it, of *S. gallolyticus* UCN34, for the assessment of the effect of the mutation on the protein functionality.

Materials & Methods

Sample collection for isolation of 16S rRNA of streptococcal and related species

The collection of human samples included saliva, tongue, faeces, and ear. An ethical permit was not required as the samples were handled anonymously. All of them were grown on Streptococcus Selection Agar (HIMEDIA)(50). The agar was prepared according to the manufacturer's instructions and plated in divided petri dishes. The plates were incubated inside a moist chamber within an incubator at 37°C for 2-5 days.

Saliva from 7 volunteers was collected by letting them spit on the agar plate. The tongue samples were collected from 3 volunteers, by using sterile swabs or letting the volunteer lick the agar plate and making sure to not include saliva. Random areas of the tongue were included in the sample collection. Except for one plate where the plate was licked, making sure that the imprint represented the whole tongue and the areas were the colonies grew on the plate would represent the front, the middle, and the back parts of the tongue. From one volunteer an ear sample was obtained by using a swab and a faeces sample was collected and grown on an agar plate. The environmental samples included decayed wood, garden pond water, pasteurized hard cheese from a local farm, moss, and horse manure from a pasture. The first sample collection of the garden pond water was conducted during winter when the pond was partially frozen.

Amplification of 16S rRNA of streptococci and related species

The three variable regions V7, V8 and V9 of the 16S rRNA for *Streptococcus* (Figure 4) were amplified through colony-PCR from colonies grown on the Streptococcus Selective Agar.

Name:	*Sequence:*
V7	CCTTTGTTGCCAGCGGTCCGGCCGGGAACTCAAAGGAGA
V8	CGCATACAAAGAGAAGCGACCTCGCGAGAGCAAGCGGACCTCATAAAGTGCGTC
V9	GTAGGTAGCTTAACCTTCGGGAGGGCGCTTA

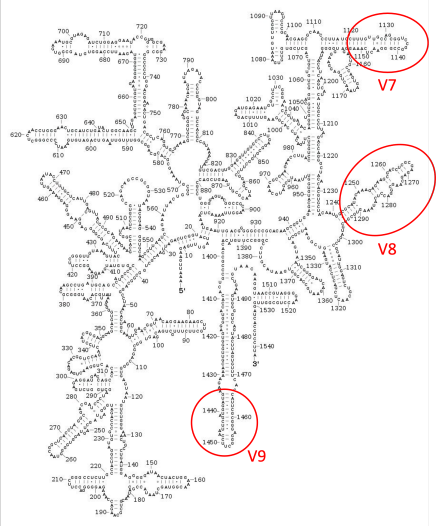

Figure 4. Sequence and 16S rRNA structure of *E. coli*

Figure showcasing the sequence of the variable regions V7, V8 and V9 for *E. coli*. 16S rRNA structure showcasing the three relevant variable regions. 16S rRNA structure taken from

http://rna.ucsc.edu/rnacenter/xrna/xrna_gallery.html. Retrieved through open access and adapted by highlighting the three variable regions relevant for this work.

Colony-PCR was conducted by resuspending one colony from the plate with Milli Q-water. The suspension was heated at 95℃ for 10 min before being centrifuged at 12 000 x g (Eppendorf) for 10 min. The supernatant was then used in PCR, which was performed using one primer pair (Table I), amplifying the V7-V9 region (Figure 1), following the PCR conditions for Taq DNA polymerase (5 U/mL) (New England BioLabs). The annealing temperature was at 48℃. All PCR reactions, mentioned in the study (Agilent Technologies SureCycle) were visualized through gel electrophoresis (Standard Power Pack P25, Biometra) using 1% agarose gel (Sigma-Aldrich). Visualisation of nucleic acids was done with GelRed or GelGreen (1 µL/mL) (Merck Millipore). The size of the bands was compared to GeneRuler 1kb Plus DNA ladder (Thermo Scientific). Pictures of the gels were taken with Molecular Image Gel Doc XR+ (Biorad) and Image Lab Software (Biorad) was used to document and annotate images. If successful amplification of the 16S rRNA was confirmed on a gel, then the PCR products were purified using QIAquick PCR Purification Kit (Qiagen), before being sent for Sanger sequencing.

Table I. List of primers used for 16S rRNA amplification of isolates

No.	***Nucleotide sequence (5'-3')***	***Description***
1.	CACTCTAGCGAGACTGCCG	Forward primer for amplification and sequencing of 16S rRNA of *Streptococcus* species
2.	CGGTTACCTTGTTACGACTT	Reverse primer for amplification and sequencing of 16S rRNA of *Streptococcus* species

Bacterial strains and growth condition

All the bacterial strains used in this project were provided by supervisor Ute Römling. *S. uberis* C6344 (51), *S. henryi* (16) and *P. aeruginosa* PT31 (52) were first grown on Mueller-Hinton agar fastidious (MH-F). Chemo-competent *E. coli* DH5α and TOP10 were grown on Luria-Bertani (LB) agar plates or in LB media. LB agar plates supplemented with ampicillin (100 µg/mL) (Sigma) were used for *E. coli* pBAD28, the selection of recombinant strains, *S. gallolyticus* UCN34 pRS02100 and pRS02100ΔTM. All the strains were grown at 37℃ for approximately 24h. All experiments were conducted in a BSL-2 lab and bacteria were handled

in a laminar airflow bench (SCANLAF Mars, Labogene). Ethical approval to work with genetically modified bacterium has been obtained.

Condition for PCR reactions and *in vivo* cloning of GGDEF-EAL domain proteins in *S. henryi*, *S. uberis* C6344, *P. aeruginosa* PT31

PCR conditions for Phusion DNA polymerase and Taq DNA polymerase were set following the protocol from New England Biolabs (53–57). The only modifications were regarding the annealing temperature and elongation time which were according to the primer pair used and the gene size of the product (Table VI).

Amplification of GGDEF-EAL domain of *S. henryi*, *S. uberis* C6344, *P. aeruginosa* PT31 and the pBAD28 vector were performed by colony-PCR, the same protocol as the amplification of 16S rRNA, using Phusion DNA polymerase (2 U/mL) (Thermo Scientific). *P. aeruginosa* PT31 and the pBAD28 vector were amplified using primer pairs 6-9. The GGDEF domain with DGC activity and two EAL domains of *S. uberis* C6344 and amplification of pBAD28 was conducted using primer pairs 10-21. One EAL domain and two GGDEF domain of *S. henryi* and pBAD28 vector were amplified using primer pair 22-33 (Table II). The PCR procedure introduced a Shine-Dalgarno sequence, a hexa-histidine-tag and part of the multiple cloning site of pBAD28 to the open reading frame (ORF) of the respective bacterium. The amplification of the vector introduced an overhang corresponding to the amplified genes products of the respective bacterium (Figure 3) (Table II). Dimethyl sulfoxide (DMSO) (3%) (FINNZYMES) and 5X Phusion GC Buffer (1X) (Thermo Scientific) had to be added to the PCR reaction with *P. aeruginosa* PT31 and its vector amplification, due to the high GC value. Taq DNA polymerase (5 U/mL) (New England BioLabs) was used if the reaction had failed with the previously mentioned polymerase.

Table II. List of primers used for the cloning of GGDEF-EAL domain proteins from *S. uberis*, *S. henryi* & *P. aeruginosa*

No.	*Nucleotide sequence (5'-3')*	*Description*
3.	GTCTATAATCACGGCAGAAAAGTCCAC	Forward control and sequencing primer binding to pBAD28
4.	CTGTTTTATCAGACCGCTTCTGC	Reverse control and sequencing primer binding to pBAD28

5.	GTCTCTATAATCACGGCAGAAAAG	Forward control and sequencing primer binding to pBAD28
6.	CTAGCGAATTCGAGCTCGGTACCCGGGGATCCTCTAGATGGCATGACGAGTACCGAGTGC	Forward primer for cloning of PT31 GGDEF-EAL domain into pBAD28
7.	CGCCAAAACAGCCAAGCTTTCAATGGTGATGGTGATGGTGATCCGGCTTCTGGTCTGGCG	Reverse primer for cloning of PT31 GGDEF-EAL domain into pBAD28
8.	CGACCGTCTGGAAACTGCACTCGGTACTCGTCATGCCATCTAGAGGATCCCCGGGTACCG	Forward primer for amplification of pBAD28 for cloning of PT31
9.	GACCAGAAGCCGGATCACCATCACCATCACCATTGAAAGCTTGGCTGTTTTGGCGGATGA	Reverse primer for amplification of pBAD28 for cloning of PT31
10.	GAATTCGAGCTCGGTACCCGGGGATCCTCTAGAAAGAAAGATAAAGGAAAAGAATGTCAT	Forward primer for cloning of C6344 EAL1 domain into pBAD28
11.	GCTTGCATGCTTAATGGTGATGGTGATGGTGTTGATTTAATCTTAAGACTTTAGGAGATT	Reverse primer for cloning of C6344 EAL1 domain into pBAD28
12.	AATTGGTCTAATGACATTCTTTTCCTTTATCTTTCTTTCTAGAGGATCCCCGGGTACCG	Forward primer for amplification of pBAD28 for cloning of C6344 EAL1
13.	TTAAGATTAAATCAACACCATCACCATCACCATTAAGCATGCAAGCTTGGCTGTTTTGGC	Reverse primer for amplification of pBAD28 for cloning of C6344 EAL1
14.	TTTGGGCTAGCGAATTCGAGCTCGGTACCGACCATTTATTGGAAAAAGTATGTTTAA	Forward primer for cloning of C6344 DGC domain into pBAD28
15.	GCCAAGCTTGCATGCTTAATGGTGATGGTGATGGTGGGATAAAGACAGATGCACTCTCTT	Reverse primer for cloning of C6344 DGC domain into pBAD28
16.	TTTTAAACATACTTTTTCCAATAAATGGTCGGTACCGAGCTCGAATTCGCTAG	Forward primer for amplification of pBAD28 for cloning of C6344 DGC
17.	GCATCTGTCTTTATCCCACCATCACCATCACCATTAAGCATGCAAGCTTGGCTGTTTTGG	Reverse primer for amplification of pBAD28 for cloning of C6344 DGC
18.	GAGCTCGGTACCCGGGGATCCTCTAGATTAAAGAAGGAATAAAAAGATGCTTATTGAAA	Forward primer for cloning of C6344 EAL2 domain into pBAD28
19.	CCAAAACAGCCAAGCTTAATGGTGATGGTGATGGTGGTTTTCCTTGAGGAGTGGTTTTCC	Reverse primer for cloning of C6344 EAL2 domain into pBAD28

20.	CAAGATTTCAATAAGCATCTTTTTATTCCTTCTTTAATCTAGAGGATCCCCGGGTACCG	Forward primer for amplification of pBAD28 for cloning of C6344 EAL2
21.	CACTCCTCAAGGAAAACCACCATCACCATCACCATTAAGCTTGGCTGTTTTGGCGGATGA	Reverse primer for amplification of pBAD28 for cloning of C6344 EAL2
22.	CGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTGGGAGGTTTTATTTAAATGAGTGA	Forward primer for cloning of *S. henryi* EAL domain into pBAD28
23.	GCCAAGCTTGCATGCTTAATGGTGATGGTGATGGTGTTTTATATTCTCTATTTTATCAGC	Reverse primer for cloning of *S. henryi* EAL domain into pBAD28
24.	TTCATTGAGTTATCACTCATTTAAATAAAACCTCCCACTCTAGAGGATCCCCGGGTACCG	Forward primer for amplification of pBAD28 for cloning of *S. henryi* EAL
25.	ATAGAGAATATAAAACACCATCACCATCACCATTAAGCATGCAAGCTTGGCTGTTTTGGC	Reverse primer for amplification of pBAD28 for cloning of *S. henryi* EAL
26.	CGAATTCGAGCTCGGTACCCGGGGATCCTCTAGATAGAATCGTTTTGTGGAAAATGGCTT	Forward primer for cloning of *S. henryi* GGDEF1 domain into pBAD28
27.	CAGCCAAGCTTTCAATGGTGATGGTGATGGTGTTTTTTTGTTTTTCCTCTCATCTTGTAT	Reverse primer for cloning of *S. henryi* GGDEF1 domain into pBAD28
28.	AAACAATCAGCAAAGCCATTTTCCACAAAACGATTCTATCTAGAGGATCCCCGGGTACCG	Forward primer for amplification of pBAD28 for cloning of *S. henryi* GGDEF1
29.	GGAAAAACAAAAAAACACCATCACCATCACCATTGAAAGCTTGGCTGTTTTGGCGGATGA	Reverse primer for amplification of pBAD28 for cloning of *S. henryi* GGDEF1
30.	GAATTCGAGCTCGGTACCCGGGGATCCTCTAGATTTACTAGGAGGTTTTCTATGACATTA	Forward primer for cloning of *S. henryi* GGDEF2 domain into pBAD28
31.	GCCAAAACAGCCAAGCTTAATGGTGATGGTGATGGTGCAATATCGTACAGCCAGCTCCGC	Reverse primer for cloning of *S. henryi* GGDEF2 domain into pBAD28
32.	ATCGTTTCTAATAATGTCATAGAAAACCTCCTAGTAAATCTAGAGGATCCCCGGGTACCG	Forward primer for amplification of pBAD28 for cloning of *S. henryi* GGDEF2
33.	GGCTGTACGATATTGCACCATCACCATCACCATTAAGCTTGGCTGTTTTGGCGGATGA	Reverse primer for amplification of pBAD28 for cloning of *S. henryi* GGDEF2

Site-directed mutagenesis of pRS02100 & pRS02100ΔTM

Site-directed mutagenesis of the PAS domain in the GGDEF domain of *S. gallolyticus* UCN34 pRS02100 and pRS02100ΔTM was performed using Q5 Site-Directed Mutagenesis Kit (New England BioLabs). The plasmids, pRS02100 and pRS02100ΔTM were isolated using GenElute™ Plasmid Miniprep Kit (Sigma). The isolated plasmids were amplified using seven primer pairs to introduce the mutations (Figure 4, Table III). Mutations were suggested to be made for further analysis and assessment of the functionality of the PAS domain.

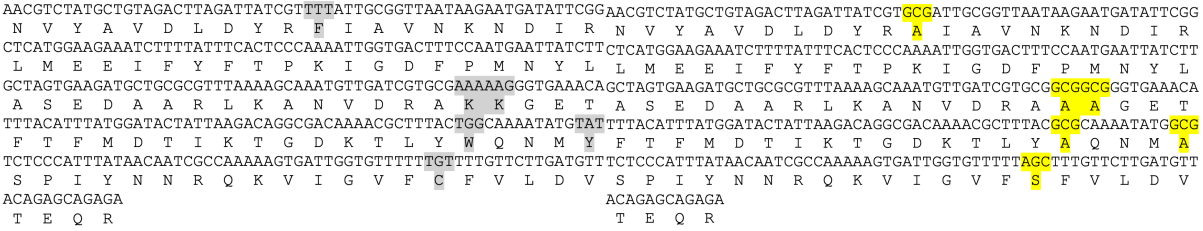

Figure 4. Mutations introduced to pRS02100 and pRS02100ΔTM.
Right side: Original nucleotide and amino acid sequence of the *S. gallolyticus* UCN34 PAS domain with grey highlight indicating where the mutation will undergo. Left side: Nucleotide and amino acid sequence with yellow highlight displaying the point mutations.

Table III. List of primers used for the mutagenesis of GGDEF UCN34 pRS02100 and pRS02100ΔTM

No.	*Primer name*	*Nucleotide sequence (5'-3')*
34.	GGDEF_UCN34_F92A_FW	AGATTATCGTGCGATTGCGGTTAATAAGAAT
35.	GGDEF_UCN34_F92A_RV	AAGTCTACAGCATAGACGTTAATAAG
36.	GGDEF_UCN34_C92A_FW	GATTGGTGTTTTTAGCTTTGTTCTTGATGTTACAG
37.	GGDEF_UCN34_C92A_RV	ACTTTTTGGCGATTGTTATAAATGG
38.	GGDEF_UCN34_W157A_FW	AAACGCTTTACGCGCAAAATATGTATTCTCCCATT
39.	GGDEF_UCN34_W157A_RV	TGTCGCCTGTCTTAATAGTATC
40.	GGDEF_UCN34_Y161A_FW	GGCAAAATATGGCGTCTCCCATTTATAACAATCGC
41.	GGDEF_UCN34_Y161A_RV	AGTAAAGCGTTTTGTCGCCTGT
42.	GGDEF_UCN34_KK137_8AA_FW	ATCGTGCGGCGGCGGGTGAAACATTTACATTTATG
43.	GGDEF_UCN34_KK137_8AA_RV	CAACATTTGCTTTTAAACGCGC
44.	GGDEF_EFK29175_D310A_FW	TATGGCGGTGCGGAATTCGTCATCATTTTTCGTG
45.	GGDEF_EFK29175_D310A_RV	GCGGAAGATCTGACCGTGAGTG
46.	GGDEF_003642201_D298A_FW	CACCGGTGGTGCGGAATTCAATGTCCTGTTTCCCG
47.	GGDEF_003642201_D298A_RV	CGATACAGCTTGACCTTCTCG

Preparation of chemo-competent DH5α and TOP10

E. coli DH5α or TOP10 was streaked out on an LB agar plate and incubated overnight at 37°C. DH5α or TOP10 was inoculated from the plate in LB medium and incubated overnight at 37°C with shaking at 200 revolutions per minute (rpm). The overnight culture was then inoculated into fresh LB medium (1:100 dilution) and incubated with the same conditions as before until optical density (OD) has reached 0.375, at 600 nm. The flask was incubated on ice for 10 min before being centrifuged for 7 min at 3000 rpm (Allegra X-12, Beckman Coulter). The cells were then washed with ice-cold 0.1 M $CaCl_2$ and again centrifuged for 5 min at 2500 rpm. Resuspension of the cells in ice-cold 0.1 M $CaCl_2$ was conducted and then incubated on ice for 30 min. Lastly, a final centrifugation step with the same conditions as previously was done before the cells were resuspended in ice-cold 0.1 M $CaCl_2$. The cells were left overnight on ice and then allocated as aliquots containing 15% glycerol for cryopreservation at -80°C.

***In vivo* cloning of GGDEF-EAL domain *S. uberis*, *S. henryi* & *P. aeruginosa* & pRS02100 and pRS02100ΔTM**

The amplified gene products in combination with the amplified pBAD28 vector (51) (1:1) were transformed *in vivo* into chemo-competent *E.coli* DH5α and TOP10. PCR products of pRS02100 and pRS02100ΔTM were ligated through a KLD treatment (New England BioLabs), which included treatment with kinase, ligase and DpnI, before being transformed into *E.coli* DH5α and TOP10. The transformation was conducted by thawing the chemo-competent cells on ice before adding the amplified PCR product and pBAD28 vector to the cells. The cells were then incubated on ice for 30 min, before being heat-shocked at 42°C for 30 sec and then SOC medium was immediately added. Afterwards, the mixtures were incubated for 2h at 37°C with heavy shaking (250 rpm). The transformation solution was centrifuged for 10 min at 8000 x g and part of the supernatant was removed. 250 µl of the supernatant was kept for resuspending the pellet before they were plated on LB agar plates supplemented with ampicillin (100 µg/mL)(Sigma). Selective colonies from the ampicillin supplemented plates were controlled by PCR amplification of the multiple cloning site of pBAD28 and visualized with gel electrophoresis (Table II). If the transformation was successful, then plasmid isolation was transacted using GenElute^{TM} Plasmid Miniprep Kit (Sigma).

Sanger sequencing and data analysis of *in vivo* cloned isolates of *S. uberis*, pRS02100 & pRS02100ΔTM and 16S rRNA amplified isolates

Isolated plasmids and purified PCR products were controlled on a gel for confirmation of successful *in vivo* cloning and amplification. Only confirmed constructs containing the plasmid with the gene insert were sent for Sanger sequencing using NightXpress Mix2Seq Kit (Eurofins Genomics). Sequencing data of the *in vivo* cloned EAL 2 domain of *S. uberis* C6344 and mutagenesis of GGDEF for pRS02100 and pRS02100ΔTM and the 16S rRNA of *Streptococcus* from the different isolates were processed and trimmed using the SnapGene Software (Insightful Science; available at snapgene.com). The sequences were trimmed according to the quality seen on the electropherogram, meaning sequences with a lot of background signals was removed or excluded. The processed data was then aligned to control sequences (Appendix A-C) using Basic Local Alignment Search Tool (BLASTN) 2.13.0 from National Center for Biotechnology Information (NCBI).

The trimmed sequencing data of the 16S rRNA was used for comparison with the nucleotide database using BLASTN 2.13.0. Multiple Sequence Comparison by Log- Expectation (MUSCLE) alignment of all processed sequences was performed on Unipro UGENE v. 42.0 and was manually controlled for irregularities, including random errors made by the sequencing process, such as additional nucleotides. MEGAX software was used to display the aligned sequences using the bootstrap method Tamura-Nei model, in a phylogenetic tree.

Results

Isolation of *S. salivarius* from human oral samples

The purpose of growing human and environmental samples on Streptococcus Selection Agar was to isolate *S. salivarius* from human oral samples and investigate which *Streptococcus* species exist in the environment. All human and environmental samples (Table IV) were effectively amplified using the primer pair (primer no. 1-2, Table I) for 16S rRNA isolation (Figure 6). The colonies grown on the Streptococcus Selection Agar were colourless, white, lilac, or violet coloured. The morphology of the colonies varied from sample to sample. It included colonies having a matt or glossy complexion, some being dry while others had a sticky

consistency (Table IV). The majority of the original plates had between 5-50 colonies, while the re-streaked plates grew lawns of colonies.

Table IV. Source and colony morphology of the isolates

Sample name	***Source***	***Colony morphology***
Saliva sample 1	Saliva	White/lilac
Saliva sample 2a	Saliva	Colourless
Saliva sample 2b	Saliva	Colourless
Saliva sample 3a	Saliva	White
Saliva sample 3b	Saliva	White
Saliva sample 4a	Saliva	N/A
Saliva sample 4b	Saliva	N/A
Saliva sample 5	Saliva	Sticky white/lilac
Saliva sample 6	Saliva	White/lilac
Saliva sample 7	Saliva	Dry colourless
Feces sample 1	Feces	White
Tongue sample 8	Tongue	White
Tongue sample 1a	Tongue	Glossy white/lilac
Tongue sample 1b	Tongue	Glossy hite/lilac
Tongue sample 2a	Tongue	Matt white
Tongue sample 2b	Tongue	Matt white
Tongue sample 2c	Tongue	Matt white
Tongue sample 3a	Tongue	White/violet
Tongue sample 3b	Tongue	White/violet
Tongue sample 4a	Tongue	Colourless
Tongue sample 4b	Tongue	Colourless
Tongue sample 5 (Localisation: F1)	Front part of the tongue	White/violet
Tongue sample 5 (Localisation: F2)	Front part of the tongue	White/violet
Tongue sample 5 (Localisation: M)	Middle part of the tongue	White

Tongue sample 5 (Localisation: B1)	Back part of the tongue	White/violet
Tongue sample 5 (Localisation: B2)	Back part of the tongue	White/violet
Tongue sample 6	Tongue	White
Tongue sample 7	Tongue	N/A
Ear sample 1a	Ear	White/lilac
Ear sample 1b	Ear	White/lilac
Horse feces sample 1a	Horse feces from pasture	Violet
Horse feces sample 1b	Horse feces from pasture	Violet
Wood sample 1a	Wood	Sticky white
Wood sample 1b	Wood	Sticky white
Wood sample 1c	Wood	Sticky white
Water sample 1a	Water	N/A
Water sample 1b	Water	N/A
Water sample 2a	Water	White
Water sample 2b	Water	White
Water sample 3a	Water	Dry white/lilac
Water sample 3b	Water	Dry white/lilac
Moss sample 1	Moss	Colourless
Cheese sample 1a	Cheese	White
Cheese sample 1b	Cheese	White

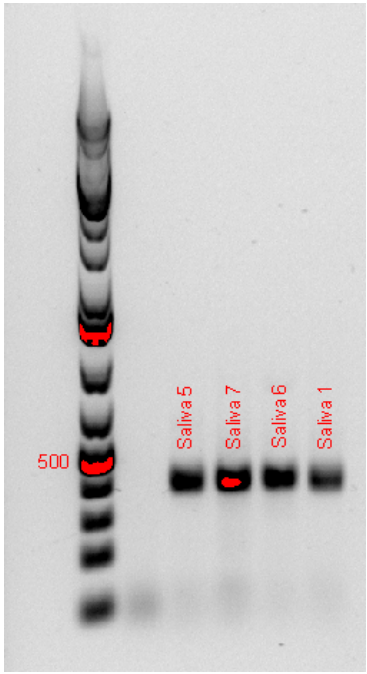

Figure 6. Gel electrophoresis picture of successfully amplified 16S rRNA
Successful amplification of 16S rRNA of saliva samples 1, 5, 6 and 7 (Table V).

Amplification of the three variable regions V7-V9 of the 16S rRNA (Figure 1) in the human and environmental isolates was conducted and sent for sequencing. The BLASTN analysis of the sequencing data presented that three saliva, and twelve tongue isolates were closely related to *S. salivarius* out of 27 human oral samples. Three saliva and three tongue isolates had 100% similarity to an *S. salivarius* strain with the accession no. MT611793.1. A fourth tongue sample was similar to the same strain; however, the similarity was 97,93%. Eight other tongue samples had 100% similarity to an *S. salivarius* strain with the accession no. CP053998.1. Other *Streptococcus* species were found to have similarities to some of the samples such as *Streptococcus agalactiae* and *Streptococcus pneumoniae*. The environmental samples were closely related to various kinds of species, including *Pediococcus*, *Staphylococcus* and *Lactococcus,* although no *Streptococcus* were found (Table V).

Table V. List of species from BLAST analysis of isolated sequences

Sample name:	***Species name:***	***Percent. Identity:***
Saliva sample 1	*Streptococcus anginosus*	100%
Saliva sample 2a	*Streptococcus salivarius*	100%
Saliva sample 2b	*Streptococcus salivarius*	100%

Saliva sample 3a	*Streptococcus salivarius*	100%
Saliva sample 3b	*Streptococcus agalactiae*	100%
Saliva sample 4a	*Streptococcus pneumoniae*	99,56%
Saliva sample 4b	*Streptococcus pneumoniae*	99,56%
Saliva sample 5	*Lacticaseibacillus paracasei*	100%
Saliva sample 6	*Lactobacillus equigenerosi*	100%
Saliva sample 7	*Lactiplantibacillus pentosus*	98,89%
Feces sample 1	*Enterococcus faecalis*	100%
Tongue sample 8	*Streptococcus salivarius*	100%
Tongue sample 1a	*Streptococcus pneumoniae*	100%
Tongue sample 1b	*Streptococcus pneumoniae*	100%
Tongue sample 2a	*Streptococcus salivarius*	100%
Tongue sample 2b	*Streptococcus salivarius*	100%
Tongue sample 2c	*Streptococcus agalactiae*	100%
Tongue sample 3a	*Streptococcus salivarius*	100%
Tongue sample 3b	*Streptococcus salivarius*	100%
Tongue sample 4a	*Streptococcus salivarius*	100%
Tongue sample 4b	*Streptococcus salivarius*	100%
Tongue sample 5 (Localisation: F1)*	*Streptococcus salivarius*	100%
Tongue sample 5 (Localisation: F2)*	*Streptococcus salivarius*	100%
Tongue sample 5 (Localisation: M)*	*Enterococcus faecium*	100%
Tongue sample 5 (Localisation: B1)*	*Streptococcus salivarius*	100%
Tongue sample 5 (Localisation: B2)*	*Enterococcus faecalis*	100%
Tongue sample 6	*Streptococcus salivarius*	100%
Tongue sample 7	*Streptococcus salivarius*	97,93%
Ear sample 1a	*Staphylococcus epidermidis*	100%
Ear sample 1b	*Staphylococcus epidermidis*	100%
Horse feces sample 1a	*Pediococcus pentosaceus*	100%
Horse feces sample 1b	*Enterococcus hirae*	98,57%
Wood sample 1a	*Staphylococcus epidermidis*	100%
Wood sample 1b	*Staphylococcus epidermidis*	100%
Wood sample 1c	*Staphylococcus epidermidis*	100%
Water sample 1a	*Lactococcus lactis subsp. cremoris*/*Lactococcus cremoris*	100%
Water sample 1b	*Lactococcus lactis*	100%
Water sample 2a	*Carnobacterium viridans*	100%

Water sample 2b	*Latilactobacillus sakei*	100%
Water sample 3a	*Staphylococcus saprophyticus*	100%
Water sample 3b	*Staphylococcus saprophyticus*	100%
Moss sample 1	*Latilactobacillus sakei*	100%
Cheese sample 1a	*Enterococcus faecalis*	100%
Cheese sample 1b	*Enterococcus faecalis*	100%

*Localisation of tongue sample 5: F = Front, M = Middle and B = Back

A phylogenetic tree was made using the aligned sequences which exhibited the close relationship between the 15 sequences related to *S. salivarius*, that had identical sequences. Two tongue samples, one faeces sample and two cheese samples were also identical to each other, all of them having similarities to *Enterococcus* species. One of the ear samples and all the wood samples were also identical to each other. One water sample and the moss sample were also identical to each other (Figure 7). In conclusion, *S. salivarius* was isolated from human saliva and tongue samples and no streptococcal species was isolated from the environmental samples.

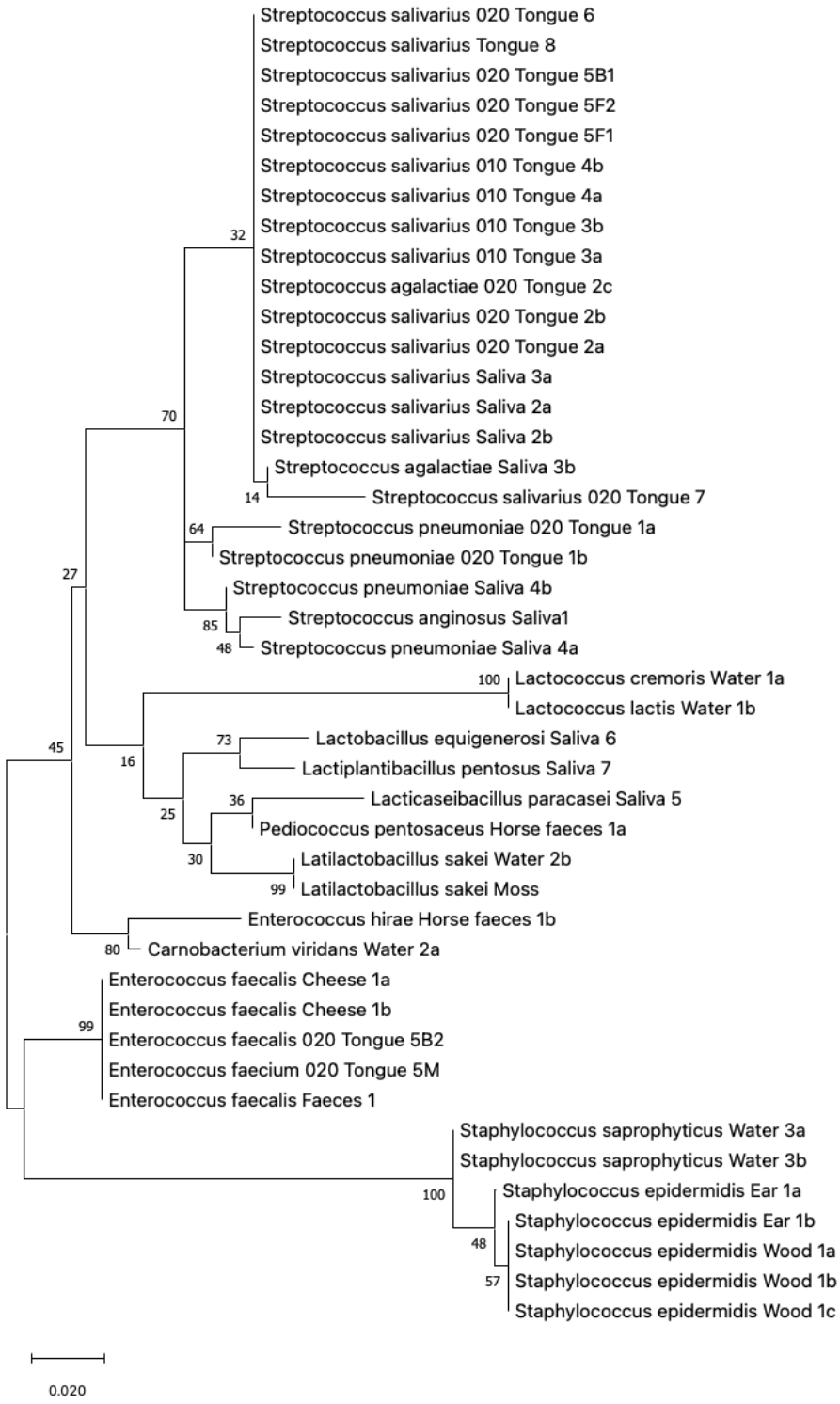

Figure 7. Phylogenetic tree of sequencing data from human and environmental samples.
Phylogenetic tree showcasing the similarities and common ancestors of sequencing data from human and environmental samples.

In vivo recombination cloning of EAL2 domain protein of in *S. uberis* C6344

The aim was to *in vivo* clone the GGDEF-EAL domain proteins of *S. henryi*, *S. uberis* C6344 and *P. aeruginosa* PT31 and then continue with the assessment of the functionality of the domain proteins. Amplification for all the genes and vectors proved to be difficult and had to be optimized. The PCR amplification resulted in 13 achieved PCR reactions out of 14 (Table VI). GGDEF-EAL domain proteins of both *S. uberis* C6344 and *P. aeruginosa* PT31 were successfully amplified with Phusion DNA polymerase (Thermo Scientific), albeit there were some difficulties with the amplification of the pBAD28 vector for *P. aeruginosa* PT31. The amplification of two GGDEF domain proteins of *S. henryi* was successful with the use of Taq DNA polymerase (New England BioLabs), however, amplification of the *S. henryi* EAL domain protein was not achievable.

Table VI. List of optimized annealing temperature for the amplification of gene and vector

PCR target	Primer no.*	Polymerase	Annealing temperature (°C)
PT31 GGDEF-EAL	6-7	Phusion DNA polymerase	65
pBAD28 for cloning of PT31 GGDEF-EAL	8-9	Phusion DNA polymerase	70
C6344 EAL1	10-11	Phusion DNA polymerase	65
pBAD28 for cloning of C6344 EAL1	12-11	Phusion DNA polymerase	65
C6344 DGC	14-15	Phusion DNA polymerase	60
pBAD28 for cloning of C6344 DGC	16-17	Phusion DNA polymerase	65
C6344 EAL2	18-19	Phusion DNA polymerase	60
pBAD28 for cloning of C6344 EAL2	20-21	Phusion DNA polymerase	65
S. henryi EAL	22-23	Taq DNA polymerase	N/A
pBAD28 for cloning of OF *S. henryi* EAL1	24-25	Taq DNA polymerase	N/A

S. S. henryi GGDEF1	26-27	Taq DNA polymerase	63
pBAD28 for cloning of *S. henryi* GGDEF1	28-29	Taq DNA polymerase	67
S. henryi GGDEF2	30-31	Taq DNA polymerase	65
pBAD28 for cloning of *S. henryi* GGDEF2	32-33	Taq DNA polymerase	64

*Primer no. refers to the primer no. on Table II.

All the successfully amplified PCR products that had been cloned into the pBAD28 and transformed into chemo-competent *E.coli* DH5α and TOP10 resulted in either no colonies or very few colonies growing on the ampicillin supplemented plates. The majority of the results from the PCR confirmation of the agar-grown colonies exhibited false-positive results i.e., bands at approximately 400 base pairs (bp) on the gel indicating the presence of the pBAD28 vector without the gene insert being present (Figure 8).

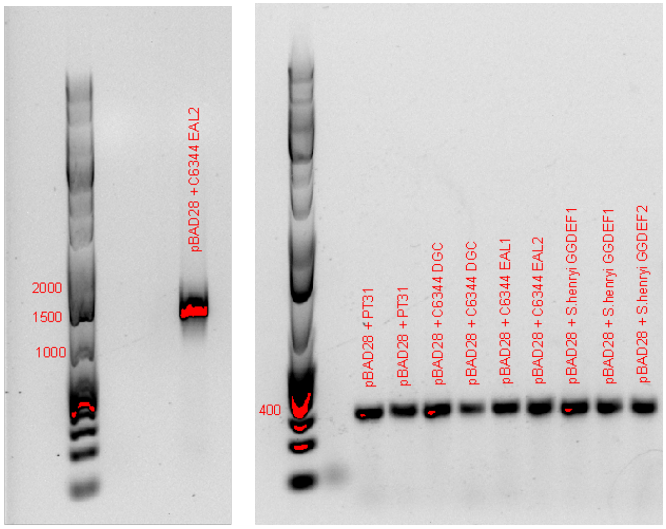

Figure 8. Gel electrophoresis picture of positive and false positive bands from transformations.
Right picture: Positive result, insert present in the vector, showcasing a band at approximately 1500 bp. Left picture: False positive result showcasing 9 bands with the size of approximately 400 bp.

Only two *in vivo* cloning experiments of *S. uberis* C6344 EAL2 in *E.coli* DH5α resulted in transformations, with the gene being present inside the pBAD28 vector. Analysis of the electropherogram of the first clone revealed that the *S. uberis* C6344 EAL2 sequenced with the reverse primer had a cleaner sequence, meaning it had less background signals and sequencing errors than the sequencing with the forward primer, therefore only the reverse primer sequence

was deemed reliable. Alignment with the control sequence of *S. uberis* C6344 EAL2 (Appendix A), disclosed two mutations, one at position 284, where the isolated plasmid sequence had a tyrosine (T), while the control sequence had adenosine (A). The second mutation was at position 610, the isolated plasmid sequence having guanine (G), while the control sequence had an A (Figure 9). The mutation at position 284, had the sequence TTC which translates to phenylalanine, while the control sequence had ATC which corresponds to isoleucine. The sequenced GAA, position 610-612, converts to the amino acid glutamate while the control sequence AAA translates to lysine.

The electropherogram of the second clone revealed that the sequencing had a higher quality sequence, meaning less background signals, with the use of the forward primer only, hence it was deemed to be more reliable than the sequence using the reverse primer. Alignment with the control sequence displayed four mutations. The first mutation was at position 610, where the isolated plasmid sequence had a G instead of an A (Figure 10), which was also seen on the first clone (Figure 9), resulting in the amino acid change from lysine to glutamate. The second mutation can be seen at position 630, the isolated plasmid sequence displaying a T, while the control sequence had an A. The mutation would have been translated to tyrosine, due to the sequence being TAT, instead of asparagine as the control sequence had. The third and fourth mutations were at positions 648 and 651, where the isolated plasmid had G, while the control sequence had a T respectively an A (Figure 10). The cloned sequence GCA at position 648-650 would have been translated to alanine, while the control sequence had TCA, which corresponds to serine. The mutation at position 651 corresponds to glutamate, where the control sequence has a lysine, similar to the mutation at position 610.

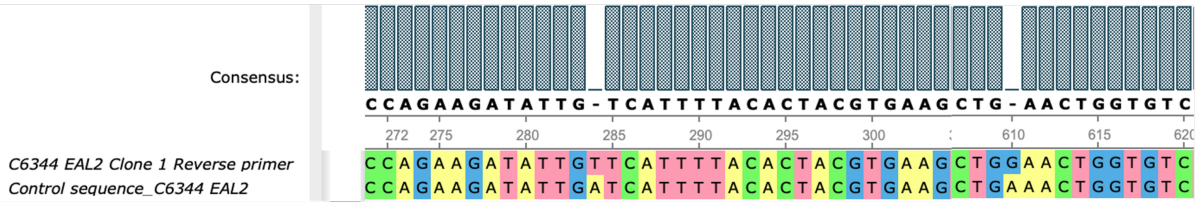

Figure 9. MUSCLE alignment of clone 1

MUSCLE alignment of the isolated plasmid of the first cloned *S. uberis* C6344 EAL2, displaying the two mutations. Alignment made with Unipro UGENE v. 42.0.

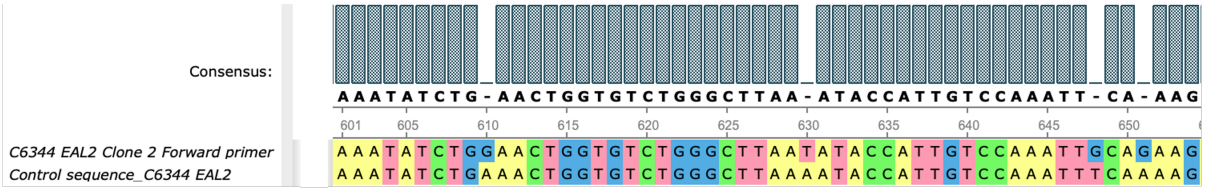

Figure 10. MUSCLE alignment of clone 2

MUSCLE alignment of the isolated plasmid of the second cloned *S. uberis* C6344 EAL2, displaying the four mutations. Alignment made with Unipro UGENE v. 42.0.

Site-directed mutagenesis and *in vivo* cloning of *S. gallolyticus* UCN34 pRS02100 and pRS02100ΔTM

The objective was to perform site-directed mutagenesis of *S. gallolyticus* UCN34 pRS02100 and pRS02100ΔTM and perform an assessment of the effect on the functionality of the PAS domain. Amplification of *S. gallolyticus* UCN34 pRS02100 and pRS02100ΔTM with the mutagenesis primers resulted in 10 successful PCR amplifications out of 14. Amplification with primer pair 44-47 was not achievable (Table III). *In vivo* cloning in *E.coli* DH5α was achieved for PCR products amplified with four primer pairs (Table III, primer no. 34-41) in both pRS02100 and pRS02100ΔTM, i.e., 8 in total. Analysis of the electropherogram of the sequencing results displayed background signals and lower quality. All alignments of the sequenced data and the control sequence of both pRS02100 and pRS02100ΔTM (Appendix B & C) presented multiple mutations without the desired mutations being present. The outcome of the transformation with the 6 other amplicons resulted in either false positive colonies or empty agar plates. Transformation with *E.coli* TOP10 was also performed and resulted in false-positive colonies (Figure 11).

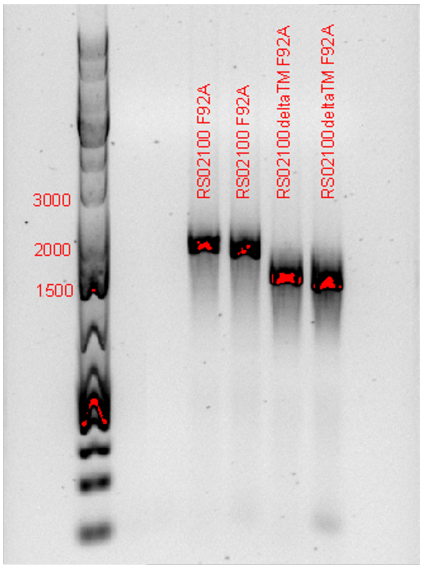

Figure 11. Gel electrophoresis picture of positive bands from a successful transformation.
Positive result showcasing bands at approx. 2000 bp and 1700 bp, for pRS02100 respectively pRS02100ΔTM, amplified with the primers GGDEF_UCN34_F92A_FW and GGDEF_UCN34_F92A_RV (primer no. 34-35, Table III).

Discussion

Isolation of 16S rRNA of *Streptococcus* and related species

Isolation of streptococci from environmental samples including dairy products using a selective media has been performed in previous studies (58,59). Using selective agar to isolate streptococcal species has been accomplished in previous studies, where they have used Mitis-salivarius agar or blood agar (60,61). Very few studies have used the selective media Streptococcus Selection Agar (HIMEDIA)(50) for the isolation of *Streptococcus* and related species. Hence in this study, the selective media was used to investigate the possibility of isolating probiotic *S. salivarius* from human oral samples and investigate the isolation of *Streptococcus* and *Lactobacillus* from external environmental samples.

Isolation of *S. salivarius* from human oral samples was possible using the Streptococcus Selection Agar. Other streptococci and related species from the human samples were also isolated using this media. However, no streptococcal species were isolated from the external environmental samples. The possibility of isolating other genera and species was confirmed by the technical data of the agar from the manufacturer's site. The technical data form stated that

the growth of *E. faecalis*, *S. pyogenes* and *S. aureus* was possible on this agar (53). Both *Enterococcus* and *Staphylococcus* were grown on the plates. Other genera including *Lactococcus*, *Latilactobacillus* and *Pediococcus* were able to grow on the agar, although further determination of the sequences must be performed. Sequencing data used for BLASTN identified the three decayed wood samples as *Staphylococcus epidermidis*, a skin flora bacterium (62). Hence, possible contamination of the samples could not be ruled out, as the sample collection was not performed under completely sterile conditions. The third water sample was identified as another pathogenic human bacterium, *Staphylococcus saprophyticus*. Few studies have recovered *Staphylococcus saprophyticus* from water, with one study isolating the bacterium from recreational waters (63). Similarly, completely sterile conditions could not be confirmed during the sample collection.

The identification of pathogenic and in particular non-pathogenic *Streptococcus* and other related species could be of interest, as these could be used for the identification of novel GGDEF and EAL domain proteins and additional genes and proteins of interest, including novel biofilm genes. Identification of GGDEF domain proteins in different streptococcal species could further expand the knowledge regarding the functionality of the domain proteins in both pathogenic and non-pathogenic streptococci. Hence, these isolated streptococci could further be used to expand the inventory of streptococcal GGDEF domain protein.

The phylogenetic analysis showcased how the different isolates were interconnected and some even identical to each other, on the 16S rRNA level. Twelve tongue samples and three saliva samples were presented to be completely identical on the phylogenetic tree. All the identical tongue and saliva samples were identified as *S. salivarius*, except one tongue sample that had been identified as *S. agalactiae*. This result showcases the importance of using additional sequencing primers that covers other variable regions or whole-genome sequencing, which could properly identify the isolates with higher reliability. Using these additions would strengthen the accuracy of the phylogenetic tree and increase the soundness of conclusions made from it. The diversity within the sample collection could have been broader, including more tongue samples or samples from other dairy products than just cheese. Though more extensive work could have been done, the results from this small study may be a stepping stone for more studies regarding probiotic bacteria and their function. Future studies may include

research regarding the presence of novel GGDEF domains and extending the repertoire of GGDEF domain proteins in *Streptococcus*.

Method of choice - *in vivo* cloning

In vivo recombination cloning of GGDEF-EAL domain protein of *S. henryi*, *S. uberis* C6344 and *P. aeruginosa* PT31 was not achievable, the only exception being *S. uberis* C6344 EAL2 that had mutations. Site-directed mutagenesis of *S. gallolyticus* UCN34 pRS02100 and pRS02100ΔTM was partly achieved for 10 experiments out of 14 and *in vivo* recombination cloning in *E.coli* DH5α was successful in some cases, however, unwanted mutations were present.

The method of *in vivo* recombination cloning has been proven to be very effective in previous studies using either the cloning hosts NEB5α or DH5α (43,49). Factors essential to the method were cloning hosts, primers and minimizing random errors in PCR. The cloning hosts used for the transformation in this study were *E.coli* DH5α and TOP10 which have been proven to be highly efficient for cloning purposes (49,64). Primers adding overlapping segments onto the gene and plasmid PCR product are essential to the method, as the overlapping segments are used for the recombination between the insert and the plasmid (Figure 3) (49). The primers used in this study added overlapping segments to both the gene and plasmid PCR products, therefore facilitating recombination.

The likelihood of successful transformation without mutations could be higher when the relevant PCR products have a high concentration and a low number of random errors for example mismatched base pairs caused by the polymerase. The use of high-fidelity polymerases can minimize the random errors and, in this study, Phusion DNA polymerase (Thermo Scientific) was used for the amplification. Phusion DNA polymerase is composed of Taq DNA polymerase and a proofreading enzyme, minimizing random error (65–68). All amplifications conducted using Phusion DNA polymerase were successful after optimization of annealing temperature. The amplification of the EAL domain of *S. henryi* was unachievable, which may be because of the use of Taq DNA polymerase (New England BioLabs) or due to the size of the gene which was approximately 3500 bp. More than one primer pair could have

been used for the amplification, which may have facilitated the process and increased the likelihood of successful amplification.

Too high a concentration of PCR products may have been used, most likely resulting in a high number of false-positive colonies. A low concentration of PCR products may have resulted in no colonies on the supplemented LB agar plates. Analysis of the concentration of PCR products before the transformation had not been conducted, therefore future studies should apply Nanodrop analysis. The recommended concentration of the PCR products is between 10 pg – 100 ng (70). DpnI digestion has been proven to minimize false-positive colonies by digesting methylated DNA, which could have been added to minimize the high frequency of false-positive colonies (49). The use of KLD enzyme treatment, kinase, ligase and DpnI treatment, in the site-directed mutagenesis of *S. gallolyticus* UCN34 pRS02100 and pRS02100ΔTM, may have been effective as 8 experiments of *in vivo* cloning were successful.

Other cloning methods for instance restriction enzyme digestion and ligation cloning could have been implemented to exclude the possibility of the *in vivo* cloning method being the problem. Although restriction enzyme cloning may have drawbacks for example long incubation time and dependence on restriction enzymes, this type of method has been proven to be effective. For instance, a recent work from this laboratory by Liu et al. had cloned GGDEF domain proteins of *S. gallolyticus* and *Proteus mirabilis* by using restriction enzyme digestion before the transformation (3). Furthermore, additional steps including using synthetically made gene of *S. uberis* C6344 EAL2 suitable for transformation into *E.coli* could have been implemented as the unsuccessful transformation may have been due to the introduction of unorthodox gene products. In this case, the streptococcal domain protein might have not been suitable for cloning, hence a modified version of the gene may be the solution. Alternative cloning hosts or performing whole-genome sequencing on the transformed isolated plasmid to confirm if additional mutations were present, could have been additional steps to be conducted.

The amplification of *S. gallolyticus* UCN34 pRS02100 and pRS02100ΔTM using the primers for the site-directed mutagenesis had more success than the amplification for *S. henryi*, *S. uberis* C6344 and *P. aeruginosa* PT31. The amplification that was not successful may be due to the length of the primers (Table III, primer no 42-45). The *in vivo* cloning of the two plasmids

was more prosperous than the cloning of the GGDEF-EAL domain for the three bacterial species. However, the unwanted mutations being presented in the sequencing data confirmed that the cloning was unsuccessful. The lower quality in some of the electropherograms may have been an indication of low plasmid concentration. Nanodrop should have been used to confirm the plasmid concentration before they were sent for Sanger sequencing. Low plasmid concentration could have also been solved by using other methods for plasmid isolation including alkaline lysis instead of using a kit.

Future experiments including an assessment of the domain protein could be conducted after successful *in vivo* cloning. Such experiments could include Congo red (CR) binding assay for assessment of biofilm and swimming motility assay. A prior study has used CR binding assay and swimming motility assay to assess the DGCs of *S. gallolyticus*, which showed promotion of red, dry, and rough morphotype of *Salmonella typhimurium* using CR binding assay. The DGCs of *S. gallolyticus* proved to suppress the motility activity of *S. typhimurium* using the swimming motility assay (3).

The conclusion

S. salivarius was successfully isolated from human oral samples using Streptococcus Selective Agar, which has not been seen before. Those isolates can be used to assess the role of the GGDEF domain diguanylate cyclase in this species. Other bacterial genera were isolated using the agar and the occurrence of streptococcus in the environmental samples could not be seen. However, further determination of the bacterial identification is needed. *In vivo* recombination cloning of EAL 2 of *S. uberis* C6344 was possible, however, point mutations generating amino acid change were present in the sequence. *In vivo* cloning of other GGDEF-EAL domain proteins of *S. henryi*, *S. uberis* C6344 and *P. aeruginosa* PT31 were unsuccessful. Site-directed mutagenesis and *in vivo* cloning of *S. gallolyticus* UCN34 pRS02100 and pRS02100ΔTM did also fail. However, further optimization and other choices of methods could be future steps to conduct.

Acknowledgements

I would like to sincerely thank my supervisor Ute Römling for providing me with a project. I would like to express my gratitude to Analucia Diaz Lostao for welcoming, helping and supporting me at the beginning of my project. I would also like to extend my appreciation to Lianying Cao (Hillary) for helping, assisting, and teaching me at the lab. I am deeply indebted to Azreen Salahuddin and Simone Brant-Lundin for letting me rant, for giving me support when things went wrong and for motivating me to continue. Special thanks to my examiner Catharina Svensson for assisting and guiding me and answering my questions. Lastly, I would like to extend my deepest appreciation and gratitude to my family and closest friends for being there for me during hardships and prosperity.

References

1. Simm R, Morr M, Kader A, Nimtz M, Römling U. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility: Cyclic di-GMP turnover by GGDEF and EAL domains. Mol Microbiol. 2004 Jul 9;53(4):1123–34.

2. Jones CJ, Utada A, Davis KR, Thongsomboon W, Zamorano Sanchez D, Banakar V, et al. C-di-GMP Regulates Motile to Sessile Transition by Modulating MshA Pili Biogenesis and Near-Surface Motility Behavior in Vibrio cholerae. Parsek MR, editor. PLOS Pathog. 2015 Oct 27;11(10):e1005068.

3. Liu Y, Lee C, Li F, Trček J, Bähre H, Guo RT, et al. A Cyclic di-GMP Network Is Present in Gram-Positive *Streptococcus* and Gram-Negative *Proteus* Species. ACS Infect Dis. 2020 Oct 9;6(10):2672–87.

4. Gray BM, Stevens DL. Streptococcal Infections. In: Brachman PS, Abrutyn E, editors. Bacterial Infections of Humans [Internet]. Boston, MA: Springer US; 2009 [cited 2022 May 13]. p. 743–82. Available from: http://link.springer.com/10.1007/978-0-387-09843-2_35

5. Mundt JO. The ecology of the streptococci. Microb Ecol. 1982 Dec;8(4):355–69.

6. Chen L, Liu S, Wang H, Wang M, Yu L. Relative significances of pH and substrate starch level to roles of Streptococcus bovis S1 in rumen acidosis. AMB Express. 2016 Dec;6(1):80.

7. Schlegel L, Grimont F, Ageron E, Grimont PAD, Bouvet A. Reappraisal of the taxonomy of the Streptococcus bovis/Streptococcus equinus complex and related species: description of Streptococcus gallolyticus subsp. gallolyticus subsp. nov., S. gallolyticus subsp. macedonicus subsp. nov. and S. gallolyticus subsp. pasteurianus subsp. nov. Int J Syst Evol Microbiol. 2003 May 1;53(3):631–45.

8. Yusuf E. A Book Review on: Sherris Medical Microbiology â€“ International Edition, 6th Edition. Front Cell Infect Microbiol [Internet]. 2015 Apr 10 [cited 2022 May 27];5. Available from: http://journal.frontiersin.org/article/10.3389/fcimb.2015.00034/abstract

9. Hungate RE. The Rumen and its Microbes [Internet]. Elsevier; 1966 [cited 2022 May 27]. Available from: https://linkinghub.elsevier.com/retrieve/pii/C2013012555X

10. Jean SS, Teng LJ, Hsueh PR, Ho SW, Luh KT. Bacteremic Streptococcus bovis infections at a university hospital, 1992-2001. J Formos Med Assoc Taiwan Yi Zhi. 2004 Feb;103(2):118–23.

11. Boleij A, Tjalsma H. The itinerary of Streptococcus gallolyticus infection in patients with colonic malignant disease. Lancet Infect Dis. 2013 Aug;13(8):719–24.

12. Dumke J, Hinse D, Vollmer T, Schulz J, Knabbe C, Dreier J. Potential Transmission Pathways of Streptococcus gallolyticus subsp. gallolyticus. Beall B, editor. PLOS ONE. 2015 May 15;10(5):e0126507.

13. Oliver SP, Pighetti GM, Almeida RA. Mastitis Pathogens | Environmental Pathogens. In: Encyclopedia of Dairy Sciences [Internet]. Elsevier; 2011 [cited 2022 May 27]. p. 415–21. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780123744074003010

14. Park MA, Kwon MG, Hwang JY, Jung SH, Kim DW, Park JY, et al. Genome Sequence of Streptococcus parauberis Strain KCTC11980, Isolated from Diseased *Paralichthys olivaceus*. Genome Announc. 2013 Oct 31;1(5):e00780-13.

15. Ghittino C, Latini M, Agnetti F, Panzieri C, Lauro L, Ciappelloni R, et al. Emerging Pathologies in Aquaculture: Effects on Production and Food Safety. Vet Res Commun. 2003;27:471–9.

16. Milinovich GJ, Burrell PC, Pollitt CC, Bouvet A, Trott DJ. Streptococcus henryi sp. nov. and Streptococcus caballi sp. nov., isolated from the hindgut of horses with oligofructose-induced laminitis. Int J Syst Evol Microbiol. 2008 Jan 1;58(1):262–6.

17. Hakalehto E, Vilpponen-Salmela T, Kinnunen K, von Wright A. Lactic Acid Bacteria Enriched from Human Gastric Biopsies. ISRN Gastroenterol. 2011 Jul 20;2011:1–4.

18. Kechagia M, Basoulis D, Konstantopoulou S, Dimitriadi D, Gyftopoulou K, Skarmoutsou N, et al. Health Benefits of Probiotics: A Review. ISRN Nutr. 2013 Jan 2;2013:1–7.

19. Burton JP, Wescombe PA, Moore CJ, Chilcott CN, Tagg JR. Safety Assessment of the Oral Cavity Probiotic *Streptococcus salivarius* K12. Appl Environ Microbiol. 2006 Apr;72(4):3050–3.

20. Tagg JR, Dierksen KP. Bacterial replacement therapy: adapting ‘germ warfare’ to infection prevention. Trends Biotechnol. 2003 May;21(5):217–23.

21. Duar RM, Lin XB, Zheng J, Martino ME, Grenier T, Pérez-Muñoz ME, et al. Lifestyles in transition: evolution and natural history of the genus Lactobacillus. FEMS Microbiol Rev. 2017 Aug 1;41(Supp_1):S27–48.

22. Dan T, Jin R, Ren W, Li T, Chen H, Sun T. Characteristics of Milk Fermented by Streptococcus thermophilus MGA45-4 and the Profiles of Associated Volatile Compounds during Fermentation and Storage. Molecules. 2018 Apr 11;23(4):878.

23. Doron S, Gorbach SL. Bacterial Infections: Overview. In: International Encyclopedia of Public Health [Internet]. Elsevier; 2008 [cited 2022 Jun 18]. p. 273–82. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780123739605005967

24. Zlotkin A, Hershko H, Eldar A. Possible Transmission of *Streptococcus iniae* from Wild Fish to Cultured Marine Fish. Appl Environ Microbiol. 1998 Oct;64(10):4065–7.

25. Baron EJ. Chapter 3. Classification. In: Baron S, editor. Medical Microbiology [Internet]. 4th ed. Galveston (TX): University of Texas Medical Branch at Galveston; 1996 [cited 2022 Jun 5]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK8406/

26. Janda JM, Abbott SL. 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. J Clin Microbiol. 2007 Sep;45(9):2761–4.

27. Coenye T, Vandamme P. Intragenomic heterogeneity between multiple 16S ribosomal RNA operons in sequenced bacterial genomes. FEMS Microbiol Lett. 2003 Nov;228(1):45–9.

28. Aono S. The Dos Family of Globin-Related Sensors Using PAS Domains to Accommodate Haem Acting as the Active Site for Sensing External Signals. In: Advances in Microbial Physiology [Internet]. Elsevier; 2013 [cited 2022 May 27]. p. 273–327. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780124076938000078

29. Chan C, Paul R, Samoray D, Amiot NC, Giese B, Jenal U, et al. Structural basis of activity and allosteric control of diguanylate cyclase. Proc Natl Acad Sci. 2004 Dec 7;101(49):17084–9.

30. Römling U, Gomelsky M, Galperin MY. C-di-GMP: the dawning of a novel bacterial signalling system: C-di-GMP signalling in bacteria. Mol Microbiol. 2005 Jun 22;57(3):629–39.

31. Kodis EJ, Smindak RJ, Kefauver JM, Heffner DL, Aschenbach KL, Brennan ER, et al. First Messengers. In: John Wiley & Sons, Ltd, editor. First Messengers [Internet]. 1st ed. Wiley; 2012 [cited 2022 Jun 18]. Available from: https://onlinelibrary.wiley.com/doi/10.1002/9780470015902.a0024167

32. Pollard TD, Earnshaw WC, Lippincott-Schwartz J, Johnson. Second Messengers. In: Cell Biology [Internet]. Elsevier; 2017 [cited 2022 Jun 18]. p. 443–62. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780323341264000268

33. Ausmees N, Mayer R, Weinhouse H, Volman G, Amikam D, Benziman M, et al. Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity. FEMS Microbiol Lett. 2001 Oct;204(1):163–7.

34. Hengge R. Principles of c-di-GMP signalling in bacteria. Nat Rev Microbiol. 2009 Apr;7(4):263–73.

35. Taylor BL, Zhulin IB. PAS Domains: Internal Sensors of Oxygen, Redox Potential, and Light. Microbiol Mol Biol Rev. 1999 Jun;63(2):479–506.

36. Hoffman EC, Reyes H, Chu FF, Sander F, Conley LH, Brooks BA, et al. Cloning of a Factor Required for Activity of the Ah (Dioxin) Receptor. Science. 1991 May 17;252(5008):954–8.

37. Meissner A, Wild V, Simm R, Rohde M, Erck C, Bredenbruch F, et al. Pseudomonas aeruginosa cupA-encoded fimbriae expression is regulated by a GGDEF and EAL domain-dependent modulation of the intracellular level of cyclic diguanylate. Environ Microbiol. 2007 Oct;9(10):2475–85.

38. Rahman M, Simm R, Kader A, Basseres E, RÃ¶mling U, MÃ¶llby R. The role of c-di-GMP signaling in an *Aeromonas veronii* biovar *sobria* strain. FEMS Microbiol Lett. 2007 Aug;273(2):172–9.

39. Le Guyon S, Simm R, Rehn M, Römling U. Dissecting the cyclic di-guanylate monophosphate signalling network regulating motility in *Salmonella enterica* serovar Typhimurium: c-di-GMP signalling regulating motility in *S* . sv. Typhimurium. Environ Microbiol. 2015 Apr;17(4):1310–20.

40. Holland LM, O'Donnell ST, Ryjenkov DA, Gomelsky L, Slater SR, Fey PD, et al. A Staphylococcal GGDEF Domain Protein Regulates Biofilm Formation Independently of Cyclic Dimeric GMP. J Bacteriol. 2008 Aug;190(15):5178–89.

41. McDonough KA, Rodriguez A. The myriad roles of cyclic AMP in microbial pathogens: from signal to sword. Nat Rev Microbiol. 2012 Jan;10(1):27–38.

42. Fahmi T, Faozia S, Port GC, Cho KH. The Second Messenger c-di-AMP Regulates Diverse Cellular Pathways Involved in Stress Response, Biofilm Formation, Cell Wall Homeostasis, SpeB Expression, and Virulence in Streptococcus pyogenes. Freitag NE, editor. Infect Immun. 2019 Jun;87(6):e00147-19.

43. Huang F, Spangler JR, Huang AY. In vivo cloning of up to 16 kb plasmids in E. coli is as simple as PCR. Isalan M, editor. PLOS ONE. 2017 Aug 24;12(8):e0183974.

44. Watson JF, García-Nafría J. In vivo DNA assembly using common laboratory bacteria: A re-emerging tool to simplify molecular cloning. J Biol Chem. 2019 Oct;294(42):15271–81.

45. Bubeck P, Winkler M, Bautsch W. Rapid cloning by homologous recombination *in vivo*. Nucleic Acids Res. 1993;21(15):3601–2.

46. Liu J, Chang W, Pan L, Liu X, Su L, Zhang W, et al. An Improved Method of Preparing High Efficiency Transformation Escherichia coli with Both Plasmids and Larger DNA Fragments. Indian J Microbiol. 2018 Dec;58(4):448–56.

47. Chan WT, Verma CS, Lane DP, Gan SKE. A comparison and optimization of methods and factors affecting the transformation of *Escherichia coli*. Biosci Rep. 2013 Dec 1;33(6):e00086.

48. Guzman LM, Belin D, Carson MJ, Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol. 1995 Jul;177(14):4121–30.

49. Jacobus AP, Gross J. Optimal Cloning of PCR Fragments by Homologous Recombination in Escherichia coli. Korolev S, editor. PLOS ONE. 2015 Mar 16;10(3):e0119221.

50. HIMEDIA Techinal Data Streptococcus Selection Agar M304 [Internet]. HIMEDIA Techinal Data Streptococcus Selection Agar M304. 2011. Available from: https://www.himedialabs.com/TD/M304.pdf

51. Hossain M, Egan SA, Coffey T, Ward PN, Wilson R, Leigh JA, et al. Virulence related sequences; insights provided by comparative genomics of Streptococcus uberis of differing virulence. BMC Genomics. 2015 Dec;16(1):334.

52. Römling U, Wingender J, Müller H, Tümmler B. A major Pseudomonas aeruginosa clone common to patients and aquatic habitats. Appl Environ Microbiol. 1994 Jun;60(6):1734–8.

53. Chester N, Marshak DR. Dimethyl Sulfoxide-Mediated Primer Tm Reduction: A Method for Analyzing the Role of Renaturation Temperature in the Polymerase Chain Reaction. Anal Biochem. 1993 Mar;209(2):284–90.

54. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, et al. Enzymatic Amplification of β-Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia. Science. 1985 Dec 20;230(4732):1350–4.

55. Powell LM, Wallis SC, Pease RJ, Edwards YH, Knott TJ, Scott J. A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. Cell. 1987 Sep;50(6):831–40.

56. Henke W, Herdel K, Jung K, Schnorr D, Loening SA. Betaine Improves the PCR Amplification of GC-Rich DNA Sequences. Nucleic Acids Res. 1997 Oct 1;25(19):3957–8.

57. Sarkar G, Kapelner S, Sommer SS. Formamide can dramatically improve the specificity of PCR. Nucleic Acids Res. 1990;18(24):7465–7465.

58. Sawant AA, Pillai SR, Jayarao BM. Evaluation of Five Selective Media for Isolation of Catalase-Negative Gram-Positive Cocci from Bulk Tank Milk. J Dairy Sci. 2002 May;85(5):1127–32.

59. Shani N, Isolini D, Marzohl D, Berthoud H. Evaluation of a new culture medium for the enumeration and isolation of Streptococcus salivarius subsp. thermophilus from cheese. Food Microbiol. 2021 May;95:103672.

60. Takada K, Hayashi K, Sasaki K, Sato T, Hirasawa M. Selectivity of Mitis Salivarius agar and a new selective medium for oral streptococci in dogs. J Microbiol Methods. 2006 Sep;66(3):460–5.

61. Conrads G, Westenberger J, Lürkens M, Abdelbary MMH. Isolation and Bacteriocin-Related Typing of Streptococcus dentisani. Front Cell Infect Microbiol. 2019 Apr 16;9:110.

62. Otto M. Staphylococcus epidermidis — the "accidental" pathogen. Nat Rev Microbiol. 2009 Aug;7(8):555–67.

63. Sousa VS de, da-Silva AP de S, Sorenson L, Paschoal RP, Rabello RF, Campana EH, et al. *Staphylococcus saprophyticus* Recovered from Humans, Food, and Recreational Waters in Rio de Janeiro, Brazil. Int J Microbiol. 2017;2017:1–11.

64. García-Nafría J, Watson JF, Greger IH. IVA cloning: A single-tube universal cloning system exploiting bacterial In Vivo Assembly. Sci Rep. 2016 Jul;6(1):27459.

65. Johnson KA. The kinetic and chemical mechanism of high-fidelity DNA polymerases. Biochim Biophys Acta BBA - Proteins Proteomics. 2010 May;1804(5):1041–8.

66. Joyce CM, Benkovic SJ. DNA Polymerase Fidelity: Kinetics, Structure, and Checkpoints. Biochemistry. 2004 Nov 1;43(45):14317–24.

67. Tindall KR, Kunkel TA. Fidelity of DNA synthesis by the Thermus aquaticus DNA polymerase. Biochemistry. 1988 Aug 1;27(16):6008–13.

68. Eckert KA, Kunkel TA. DNA polymerase fidelity and the polymerase chain reaction. Genome Res. 1991 Aug;1(1):17–24.

69. Chang AY, Chau VW, Landas JA, Pang Y. Preparation of calcium competent Escherichia coli and heat-shock transformation [Internet]. Preparation of calcium competent Escherichia coli and heat-shock transformation. 2017. Available from: https://jemi.microbiology.ubc.ca/sites/default/files/Chang%20et%20al%20JEMI-methods%20Vol%201%20pg%2022-25.pdf

Appendix

A. Sequence of *S. uberis* C6344 EAL2:

GTCCTACACGAAAAAGCTTAAAGAAGGAATAAAAAGATGCTTATTGAAATCTTGTTA
CTCATTGCAATTATTTTTACTATAGTTAGTATTGTTACAATTTTGATTAGTTATATT
GGGGCCAAAAGACAAAATGCTATCTACCCAAAAACTATGAAAAATATCGAAAATTAC

TCATTTCATTTTCAGAAAATTATAGATAGAGATGGTCGGGTAAGTGGCTTTGAAGCC
CTACTTAGGAAATATAATGAAGAAAATAAGAACTGGTCTTTACCAGAAGATATTGAT
CATTTTACACTACGTGAAGTTATCTACCTACTTCATAAAAGTCTACTAAAAAAGGAA
TACCCTAATGGTTTTCTAGCTATCAATATTAGCTTAAAACAACTGGTTGATCCTCGT
TATGTCTACTTTATCAAATGGTTAAAAGGGGTCATTTACCCAATGAAAGTCCGTATA
GAATTTCATATTGCTTCAACTCAATTTATTAGCCCTTGGACCAGCTGGCGCCTGAAA
AAGAATTTAAAAGTAAGTAAAGACTTAGGCGTAGAAGTCATCTTGGAACAAATAAGT
CCAGACAAAAACTATTATCAAAAGGTTAAAAAATATCTGAAACTGGTGTCTGGGCTT
AAAATACCATTGTCCAAATTTCAAAAGAAAAATGATGAAGAATGGTATTTTAAGAAC
ATAGGTGACTGGGTGAGACTAAGTCAGTTAAATCAAATAAGTATTGATTTAACTGAG
ATTGAGTCTACAGAAGATATGTCTTTAGCAGATCAATTAGATATGTCTAACCGCCAA
GGTTATTACATTGGAAAACCACTCCTCAAGGAAAACCACCATCACCATCACCATTAA

B. Sequence of *Streptococcus Gallolyticus* UCN34 GGDEF domain – pRS02100:

GTGAGGCTGTAATGGTATTAGGAAAAGTAAGTGATTTTTTAATCAAGCATCAAAGAT
TGTTAACGTACTTAAGTATTTTTGCTACGTTAATGTTATTAATGCGCTTAGTGTATG
ATGATGTTTTGATTGACCATGATAATCCCATCGTTTTAGCAATCTTGATAGGAATAC
TAGTTCTTTGGACCCTAGCATCTTACCTTAATCGAAGACAGCATATGCTATCCCACT
TTATCTTACAGAGTTCGAAACTTATTAACGTCTATGCTGTAGACTTAGATTATCGTT
TTATTGCGGTTAATAAGAATGATATTCGGCTCATGGAAGAAATCTTTTATTTCACTC
CCAAAATTGGTGACTTTCCAATGAATTATCTTGCTAGTGAAGATGCTGCGCGTTTAA
AAGCAAATGTTGATCGTGCGAAAAAGGGTGAAACATTTACATTTATGGATACTATTA
AGACAGGCGACAAAACGCTTTACTGGCAAAATATGTATTCTCCCATTTATAACAATC
GCCAAAAAGTGATTGGTGTTTTTTGTTTTGTTCTTGATGTTACAGAGCAGAGACTTC
ATGAACTTGAAATACAGAGGATGGCATACGAAGATGTTTTAACACGGGTTCATAATC
GACGATATATTGAGCTTGCTTTTGAGGAATGTCTGACACGTAAAGAAGAACAAATTA
CGGTTATTATATCTGATTTGGATAAATTCAAGGAAGCAAATGATACCTTTGGACATG
CTACGGGAGATAAGATTCTAATTGAATTTGGTGATATTTTAACGAAGATAATGCCTG
AAAGTGCTGTAATAGCAAGGTTAGGTGGAGATGAATTTGCGGTTTTATTACCGGAAG
TTTCTGAAAATCAAGCTGAATTTTTAATAAAACTTGTTCAAGCAGAAATGACTGTCA
AAGATATGGGTGTCACAGCATCTCTAGGTGCTTATACAGATTCTTATCAATCACATA
AAACCTTCGTTGATTTTTGTGCAATGGCGGATAAGAAAATGTACGAAAATAAAAGTC
AGAAAGGATAA

C. Sequence of *Streptococcus Gallolyticus* UCN34 GGDEF domain without the transmembrane – pRS02100ΔTM:

AATCGAAGACAGCATATGCTATCCCACTTTATCTTACAGAGTTCGAAACTTATTAAC
GTCTATGCTGTAGACTTAGATTATCGTTTTATTGCGGTTAATAAGAATGATATTCGG
CTCATGGAAGAAATCTTTTATTTCACTCCCAAAATTGGTGACTTTCCAATGAATTAT
CTTGCTAGTGAAGATGCTGCGCGTTTAAAAGCAAATGTTGATCGTGCGAAAAAGGGT
GAAACATTTACATTTATGGATACTATTAAGACAGGCGACAAAACGCTTTACTGGCAA
AATATGTATTCTCCCATTTATAACAATCGCCAAAAAGTGATTGGTGTTTTTTGTTTT
GTTCTTGATGTTACAGAGCAGAGACTTCATGAACTTGAAATACAGAGGATGGCATAC
GAAGATGTTTTAACACGGGTTCATAATCGACGATATATTGAGCTTGCTTTTGAGGAA
TGTCTGACACGTAAAGAAGAACAAATTACGGTTATTATATCTGATTTGGATAAATTC
AAGGAAGCAAATGATACCTTTGGACATGCTACGGGAGATAAGATTCTAATTGAATTT
GGTGATATTTTAACGAAGATAATGCCTGAAAGTGCTGTAATAGCAAGGTTAGGTGGA
GATGAATTTGCGGTTTTATTACCGGAAGTTTCTGAAAATCAAGCTGAATTTTTAATA

AAACTTGTTCAAGCAGAAATGACTGTCAAAGATATGGGTGTCACAGCATCTCTAGGT
GCTTATACAGATTCTTATCAATCACATAAAACCTTCGTTGATTTTTGTGCAATGGCG
GATAAGAAAATGTACGAAAATAAAAGTCAGAAAGGA

All the sequences were used for comparison with the cloned products.