Development of a New PCR Test System for the Detection of Bacteria Streptococcus Mitis in Biological Samples

Sofia Matskevich

Received November 13, 2024
Accepted December 16, 2024
Electronic access December 31, 2024

There is a significant portion of bacteria in the human oropharynx. One of them - Streptococcus mitis - is a common commensal that is frequently met in oral cavity; however, it could cause some serious diseases such as endocarditis, bacteraemia and septicaemia in immunocompromised patients and impact oral tissue. That could lead to destroying consequences, but at the same time it is highly challenging to identify the source of the illness and to differentiate bacteria to cure it. That is why the PCR-test system was chosen to be created as it has a lot of benefits and advantages over other tools for its speed and accuracy. During the analysis was concluded that the most precise way to identify the bacteria is by using testing system based on Nested PCR that gives better results as it is repeated twice, providing an opportunity to improve the process and exactness of identifying Streptococcus mitis in diverse biological samples. Results of the work suggest that there is a high precision in this method, so effective primers for Nested PCR were designed and checked by conducting the whole experiment of identifying Streptococcus mitis in biological samples.

Introduction

The oral microbiome rests within biofilms throughout the oral cavity, forming an ecosystem that maintains health when in equilibrium. Human oropharynx contains the second largest microbiota in the body after gastrointestinal tract: more than 772 prokaryotic species[1]. These elusive habitants impact human health and could cause serious diseases. That is why it is crucial to figure out the whole composition of oral microflora and be able to identify organisms inside. The analysis of the microbiome and its genomes will pave the way for more effective therapeutic and diagnostic techniques and, ultimately, contribute to the development of personalized medicine and personalized dental medicine.

Streptococcus mitis (S.mitis) - is one of the most common organisms in the human oral cavity which was identified in the research[2] and presented in Fig.1. Streptococci are one of the first bacteria identified as a pathogenic. The Mitis group is a normal commensal; moreover, it could be beneficial as some but not all strains of S.mitis are characterized by immunoglobulin A1 (IgA1) protease production and the ability to bind salivary α-amylase[3], both of which have been suggested to confer ecological advantages. However, it could effectively colonize the human oropharynx. S.mitis can escape from its niche in immunocompromised patients, particularly immediately after tissue transplants, and in neutropenic cancer patients[4] and cause a variety of infectious diseases including endocarditis, bacteraemia and septicaemia.

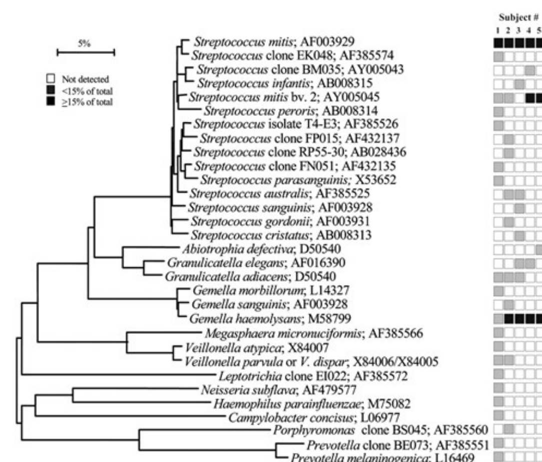

Fig. 1 Bacterial profiles of the buccal epithelium of healthy subjects[2].

Previously, S.mitis was identified using a matrix-assisted laser desorption/ionization-time of flight mass spectrometry

© The National High School Journal of Science 2024

(MALDI-TOF MS). The study showed that S.mitis was the predominant bacteria in the pharynx[5]. MALDI-TOF MS can be problematic and less specific for some species belonging to the Mitis group[6], it also necessitates highly specialized equipment and often requires a pure microbial culture. In the context of analyzing biological samples like saliva, obtaining a pure culture is time-consuming and challenging, thus limiting the practicality and speed of this method. S.mitis was also identified by suppression subtractive hybridization based on pheA gene[7]. While Suppression subtractive hybridization is a highly effective method for separating DNA molecules, it can only distinguish two closely related DNA samples. Therefore, this method is ineffective for identifying a specific bacterial species in a biological sample containing a large number of different bacteria. Fluorescence in situ hybridization (FISH), another method used to detect S. mitis, yielded positive results in the conducted research[8]. However, this method has certain limitations. For example, it has a limited throughput and lower sensitivity compared to PCR. S. mitis identification utilized the OmniLog ID System[9], which is not considered part of the current "gold standard"(culture-based, PCR, MALDI-TOF MS, FISH, NGS). However, it remains a viable method, though its time-consuming nature should be acknowledged. The method requires pre-experimental cultivation and an incubation period that can extend to several days. Thus, after a detailed analysis of existing test systems and their shortcomings, the decision was made to develop primers for a PCR-based test system.

Several PCR systems exist, including Conventional-PCR and Real-time-PCR. However, the Nested-PCR test system was chosen because of its higher specificity. Although it is more time-consuming, Nested-PCR has been shown to be 100 times more sensitive than Conventional-PCR[10]. The use of two pairs of oligonucleotides allows for a higher number of cycles, thereby increasing the sensitivity of the PCR[11].

The genome of S. mitis is very similar to that of S. pneumoniae, S. pseudopneumoniae, and S. oralis. The nucleotide sequence of the pheA gene of S.mitis is markedly different from other bacteria. The 275-bp sequence that precedes licD3, comprising pheA and a transcription factor gene in S. mitis, is in a different location in the genome than it is in the genomes of S. pneumoniae and other closely related species. This operon was unique to the S.mitis species. In contrast, S.pneumoniae, S.pseudopneumoniae, and S.oralis had a transcription factor gene and the licD3 gene but lacked the pheA gene[7]. The pheA found in S.mitis showed relatively low homology to the corresponding gene in S.pneumoniae: 32% query sequence identity and 90% nucleotide identity, respectively[7] . Therefore, the pheA gene was selected as the target. The pheA gene encodes PDT, found in many bacteria as a terminal enzyme in the aromatic amino acid biosynthesis pathway that catalyzes the conversion of prephenate to phenylpyruvate with the elimination of water and carbon

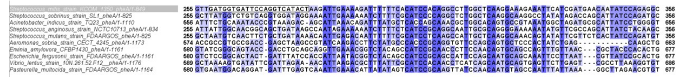

Fig. 2 Result of multiple alignment of gene pheA. The image was generated using JalView. The primer PheA_N2_fwd is highlighted in the image.

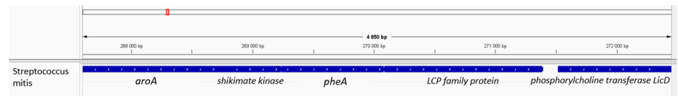

Fig. 3 Visualization of the genomic locus with pheA gene of the Streptococcus mitis LMG 14557 strain. The image was generated using the IGV Genome Browser.

dioxide[11]. So S.mitis specific pheA PCR primers were designed.

In the course of this study, a PCR test system based on nested PCR technology was developed and validated, allowing the detection of S.mitis in biological samples.

Results

Multiple alignment and the analysis of the conservativeness of the pheA gene

The conservativeness of the bacterial pheA gene has been analyzed. Multiple alignment of nucleotide sequences for pheA was constructed from the 10 most common representatives of the pheA gene (Fig. 2.). Based on the data, we concluded that the pheA gene exhibits sufficient conservation using percentage identity analysis in JalView.

Bioinformatic analysis of bacterial genome locus sequences.

The genomic locus of the Streptococcus mitis LMG 14557 strain was analyzed (Fig. 3.). The close proximity and similar functions of pheA and the neighboring aroA gene may indicate that they are part of a functional unit within the same operon. They both are involved in aromatic amino acid biosynthesis and folate metabolism[12]. This suggests that the expression of these genes is co-regulated.

Primer design and analysis of their thermodynamic characteristics.

A PCR test system based on Nested PCR technology has been developed. Two pairs of specific primers were designed for the amplification of the pheA gene. Their thermodynamic characteristics were analyzed, and additional specificity validation was performed. Using NCBI BLAST, it was shown that the designed second pair of primers, which is more specific, efficiently anneals only to the genome of bacteria of the species Streptococcus mitis.

© The National High School Journal of Science 2024

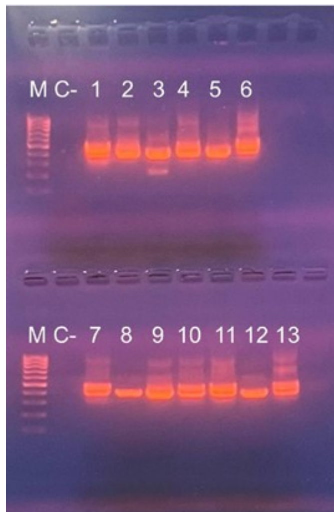

Fig. 4 Results of electrophoresis of PCR products obtained during amplification of gene pheA.

Percentage identity	Streptococcus_mitis_LMG_14557:269988-269495
F_A1_G4409_H08_23_03	93%
F_A2_G4409_A09_03_03	94%
F_A3_G4409_B09_06_03	92%
F_A4_G4409_C09_09_03	92%

Table 1. Comparison of forward PCR product sequencing data to the pheA gene nucleotide sequence in Streptococcus mitis LMG 14557 strain.

Nested PCR and Electrophoresis of PCR products in agarose gel.

Nested PCR was performed using the developed PCR test system. 13 biological saliva samples were analyzed. Electrophoresis yielded 13 positive results, all of which were the expected length. They were visualized on an agarose gel (Fig. 4.).

The control reaction (C-) showed a negative result, indicating the reliability of the obtained results. The pheA PCR products obtained during amplification, No. 3, No. 5, No. 8, No. 12, were analyzed. Products No.1, No.2, No.4, No.6, No.7, No.9, No.10, No.11, No.12 have unclear results that could be blurry or contain extra products of amplified closely related strains. Clear products No. 3, No. 5, No. 8, No. 12 were purified from the gel and sequenced by Sanger. Results of sequenced products were analyzed using NCBI BLAST. Percentage identity of sequences is shown in Table 1 and Table 2.

Analysis of the sequencing data using NCBI BLAST showed that the DNA purified from the gel is indeed the product of amplification of the pheA gene of S.mitis. Multiple sequence alignment of the sequencing results with the original pheA gene sequence confirms that the correct genomic region was amplified (Fig. 5.).

Thus, we successfully obtained and validated the PCR test

Percentage identity	Streptococcus_mitis_LMG_14557:269988-269495
R_A5_G4409_D09_12_03	93%
R_A6_G4409_E09_15_03	95%
R_A7_G4409_F09_18_03	92%
R_A8_G4409_G09_21_03	93%

Table 2. Comparison of reverse PCR product sequencing data to the pheA gene nucleotide sequence in Streptococcus mitis LMG 14557 strain.

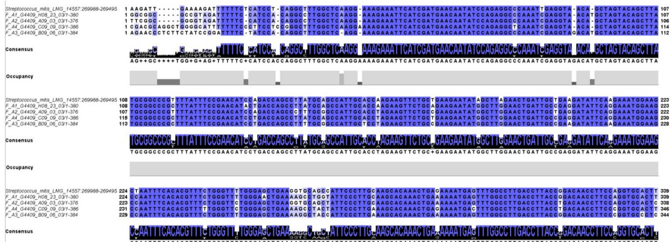

Fig. 5 Result of multiple alignment of the sequencing results with gene pheA. The image was generated using JalView.

system on biological samples.

Discussion

The oral cavity is one of the most densely populated areas with bacteria microbiota. Streptococcus mitis is one of the most widespread organisms[2], which is commensal and crucial in maintaining oral as well as systemic health. However, it could cause some serious infections[13]. That is why there is a need to obtain an efficient test system. Nested PCR technology was chosen as beneficial for increasing the sensitivity and specificity of PCR. It is particularly useful when amplifying a specific member of a group.

The pheA gene in S.mitis has a unique nucleotide sequence, distinct from other bacteria. Moreover, bioinformatic analysis indicated that the pheA gene exhibits a sufficient level of conservation. This led us to develop specific pheA PCR primers that would only target S.mitis, as a portion of the genome was selected that is distinct from other bacteria. Primer validation was performed using 13 saliva samples. This sample size was limited by equipment constraints and the availability of participants. Validation of sequenced products using NCBI BLAST showed high percentage identity, but it did not reach a 100% as the quality of the Sanger sequencing is reduced at the beginning and end of the sequence. However, during amplification, our PCR test showed several positive results of varying lengths, suggesting potential cross-reactivity with other closely related strains. To accurately determine the species of the bacteria from which the additional signal was obtained, it will be necessary to perform a longer separation of PCR products in an agarose gel and re-sequence them.

Such a system could be used to identify Streptococcus mitis

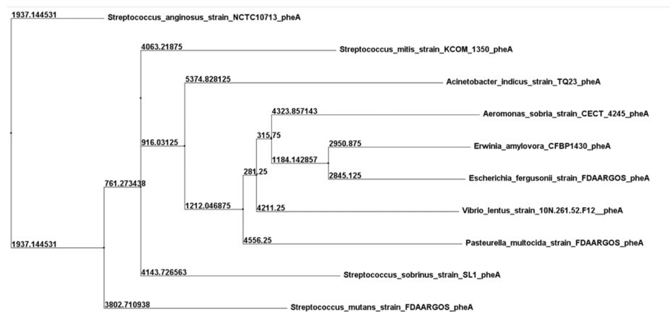

Fig. 6 The neighbor-joining tree calculated from the multiple alignment.

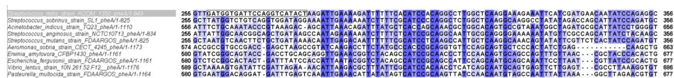

Fig. 7 Multiple alignment performed in the JalView program.

in other biological samples in the format of commercial kits for clinics and laboratories.

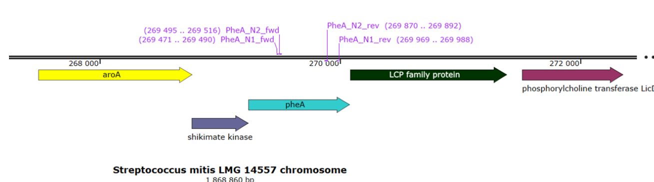

Fig. 8 Location of primers selected for the creation of a PCR test system on the genome.

Primer name	5'-3' sequence	Primer length, bp	Amplicon length, bp	Tm, °C	GC comp, %	Max hairpin, ΔG	Max homodimer, ΔG	Max heterodimer, ΔG	Reference
F1	CGTTCAGCCTATTCATCAGC	20	478	60.7	50	0.28	-3.61 kcal/mol	-3.07 kcal/mol	GCF_031191225.1: 269471-269490
R1	AGTGGACCAAGTCCTTATCA	20		60.4	45	-0.45	-5.02 kcal/mol		GCF_031191225.1: 269969-269988
F2	GATGGTGGTTCCAGGTCATACT	22	353	63.5	50	-1.88	-5.02 kcal/mol	-5.02 kcal/mol	GCF_031191225.1: 269495-269516
R2	CAATTCCTCTCCAAGCAAAGGTC	23		63.5	47.8	-0.8	-5.36 kcal/mol		GCF_031191225.1: 269870-269892

Table 4. Thermodynamic characteristics of the selected primers.

Materials and Methods

Multiple alignment and the analysis of the conservativeness of the pheA gene.

10 most common representatives from different phylogenetic groups (Fig. 6.) of pheA gene were selected for the construction of multiple alignment of nucleotide sequences of pheA (Supplement - Table 3, S1). Multiple alignment of nucleotide sequences was performed using the online resource [14] using the MAFFT algorithm. The alignment result was visualized in the JalView program (Fig. 7.)[15].

Bioinformatic analysis of bacterial genome locus sequences.

The reference genome of the most prevalent Streptococcus mitis strain was selected to analyze the genome structure and study the locus carrying the pheA gene. Genome GCF_031191225.1 Streptococcus mitis LMG 14557 was downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/) and visualized in the IGV Genome Browser[16].

Primer design and analysis of their thermodynamic characteristics.

The GCF_031191225.1 Streptococcus mitis LMG 14557 genome was also visualized in SnapGene Viewer (SnapGene software (www.snapgene.com) for convenience. To create a PCR test system based on "nested"PCR technology (Nested PCR) using the online resource IDT's PrimerQuest Tool (https://eu.idtdna.com/Primerquest/Home/Index) , two pairs of primers were selected for each of the two genes. The first pair of primers was used for the first round of amplification, during which a longer PCR product was synthesized. The second pair of primers annealed to sites within the amplicon obtained in the first round. Accordingly, during the second round of amplification, a shorter PCR product was synthesized. This system has increased sensitivity and specificity due to the presence of two sequential rounds of amplification. The obtained primers were mapped onto the genome in SnapGene Viewer (Fig. 8.).

Analysis of the thermodynamic characteristics of the selected primers was performed using the online resource IDT's OligoAnalyzer Tool (https://eu.idtdna.com/calc/analyzer). Among the analyzed parameters are the melting temperature of primer homo- and heterodimers, the energy of formation of potential secondary structures, etc. Examples of potential secondary structures identified during primer analysis are shown in Fig. 9. The primers and their characteristics are shown in Table 4.

Additional validation of second pair of primers specificity was performed using NCBI BLAST (&

PAGE_TYPE=BlastSearch&LINK_LOC=blasthome). The data set shows 100% accuracy of the selected primers (Fig. 10).

Nested PCR.

The work was conducted under sterile conditions using necessary equipment, alcohol sterilization, and a UV lamp. 13 biological saliva samples were collected among 13 teenagers

© The National High School Journal of Science 2024

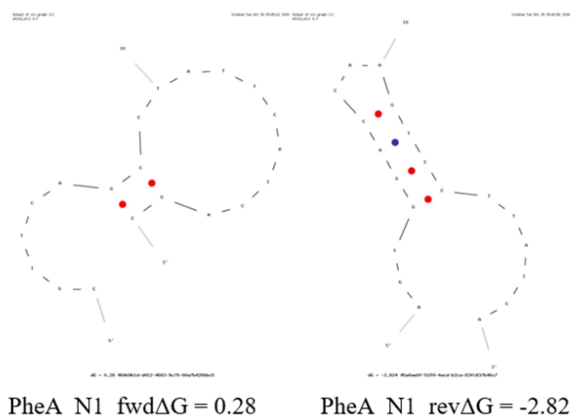

Fig. 9 Potential intramolecular secondary structures of primers.

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Streptococcus sp. SN-1 DNA, complete genome	Streptococcus s...	44.1	44.1	100%	0.082	100.00%	2111706	AP028929.1
☑	Streptococcus mitis strain S022-V7-A3 chromosome, complete genome	Streptococcus ...	44.1	44.1	100%	0.082	100.00%	2033396	CP067992.1
☑	Streptococcus mitis strain LMG 14557 chromosome, complete genome	Streptococcus ...	44.1	44.1	100%	0.082	100.00%	1868860	CP133471.1
☑	Streptococcus sp. 116-D4 DNA, complete genome	Streptococcus s...	44.1	44.1	100%	0.082	100.00%	1943757	AP021887.1
☑	Streptococcus mitis NCTC 12261 chromosome, complete genome	Streptococcus ...	44.1	44.1	100%	0.082	100.00%	1868883	CP028414.1
☑	Streptococcus mitis strain FDAARGOS_684 chromosome, complete genome	Streptococcus ...	44.1	44.1	100%	0.082	100.00%	1868857	CP046335.1
☑	Streptococcus mitis strain FDAARGOS 1456 chromosome, complete genome	Streptococcus ...	44.1	44.1	100%	0.082	100.00%	1868859	CP077259.1

Fig. 10 Additional specificity validation

in the Primakov school and stored 2 hours in the temperature 8°C. They were prepared for analysis using the developed PCR system. First, they were heat-treated for 30 minutes at 95°C to disrupt the cell walls and membranes of the bacteria present in the sample. Then, PCR mixture was prepared for each sample for amplification of the pheA gene. During the preparation of the mixtures for the first round of amplification, the following were used: components of the Encyclo Plus PCR kit ("Evrogen", Russia), as well as primer pairs PheA_N1_fwd, PheA_N1_rev. Components used in the PCR are shown in the Table 5.

Control mixtures were also prepared, in which deionized water free from nucleases was added as a template instead of the biological sample. The resulting PCR mixtures were placed in a BIO-RAD C1000 Touch amplifier ("BioRad", USA). Amplification was carried out according to the PCR protocol presented in Fig. 11.

In the second round of amplification, the corresponding PCR products obtained in the first round were used as a template. When preparing the mixtures for the second round, 2 μl of the mixture containing the PCR product of the first stage of amplification was added, as well as the second pair of primers for amplification of the internal region of the PCR product (PheA_N2_fwd, PheA_N2_rev). The resulting mixtures were placed in the amplifier and amplified according to the PCR protocol established in advance, which differed from the protocol of the first stage by a higher primer annealing temperature (63°C), presented in Fig. 12.

Component	Reaction with Encyclo buffer	Reaction with Encyclo Red buffer
Deionized water, nuclease-free	14.5 µl	12 µl
10X Encyclo buffer	2.5 µl	-
5X Encyclo Red buffer	-	5 µl
dNTP mix (10 mM each)	0.5 µl	0.5 µl
PCR-primer 1	1 µl	1 µl
PCR-primer 2	1 µl	1 µl
DNA-matrix	5 µl	5 µl
50X Encyclo polymerase Mix	0.5 µl	0.5 µl
Final volume	25 µl	25 µl

Table 5. Components used in the PCR.

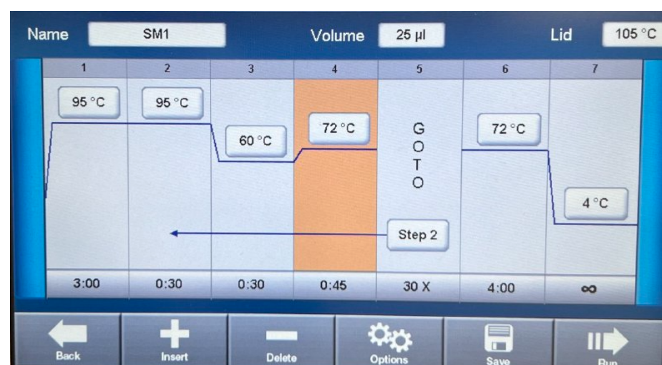

Fig. 11 Protocol for the first round of amplification of PheA gene.

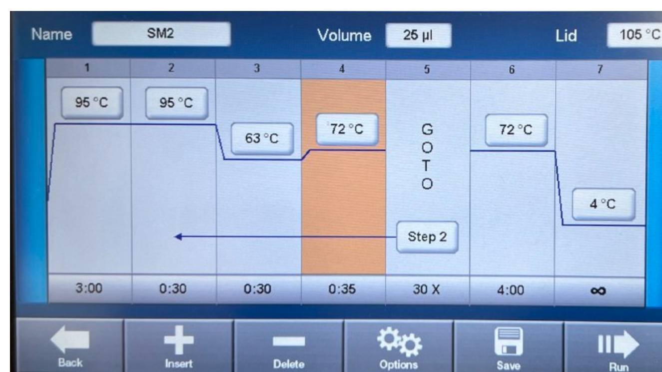

Fig. 12 Protocol for the second round of amplification of pheA gene.

Electrophoresis of PCR products in agarose gel.

Electrophoretic separation of the PCR products obtained in the second round of amplification was performed in a 2% agarose gel and DNA marker Step 100 ("biolabmix", Russia) was used. Ethidium bromide was used as an intercalating dye to visualize the PCR products in the gel under a UV lamp. After electrophoretic separation, the clearest PCR products were purified from the agarose gel for further sequencing. The Cleanup Mini kit ("Evrogen", Russia) was used to purify DNA from the agarose gel. The entire purification process was performed according to the protocol attached to the kit[17]. Next, the concentration of the purified PCR product was measured using a NanoDrop 2000 instrument ("Thermo Scientific", USA), and then it was sequenced by Sanger.

Acknowledgements

I am grateful to Primakov School and Larisa Zhiganova for technical assistance during primer synthesis.

References

1 D. Verma, P. K. Garg and A. K. Dubey, *Arch Microbiol*, 2018, **200**, 525–540.

2 J. A. Aas, B. J. Paster, L. N. Stokes, I. Olsen and F. E. Dewhirst, *Journal of Clinical Microbiology*, 2005, **43**, 5721–5732.

3 J. Hohwy, J. Reinholdt and M. Kilian, *Infection and Immunity*, 2001, **69**, 6055–6063.

4 A. M. Whatmore, A. Efstratiou, A. P. Pickerill, K. Broughton, G. Woodard, D. Sturgeon, R. George and C. G. Dowson, *Infection and Immunity*, 2000, **68**, 1374–1382.

5 T. Oishi, T. Muratani, T. Tanaka, M. Sato, K. Urara, K. Ouchi, S. Iwata, T. Matsumoto and C. Nakahama, *Japanese journal of infectious diseases*, 2021, **74**, 450–457.

6 E. Sadowy and W. Hryniewicz, *European journal of clinical microbiology & infectious diseases*, 2020, **39**, 2247–2256.

7 H. K. Park, H. T. Dang, S. C. Myung and W. Kim, *Applied and environmental microbiology*, 2012, **78**, 3004–3009.

8 L. Shang, D. Deng, S. Roffel and S. Gibbs, *Contact Dermatitis*, 2020, **83**, 347, 60.

9 R. Tomazin, T. Cerar Kišek, T. Janko, T. Triglav, K. Strašek Smrdel, V. Cvitković Špik, A. Kukec, J. Mulec and T. Matos, *Microorganisms*, 2024, **12**, 1427.

10 D. Kim, G. Park, H. S. Kim, J. Y. Lee, G. P. Neupane, S. Graves and J. Stenos, *Journal of Clinical Microbiology*, 2011, **49**, year.

11 M. R. Green and J. Sambrook, *Cold Spring Harb Protoc*, 2019.

12 H. G. Griffin and M. J. Gasson, *Molecular Genetics and Genomics*, 1995, **246**, 119–127.

13 B. A. Catto, M. R. Jacobs and D. M. Shlaes, *Archives of internal medicine*, 1987, **147**, 885–888.

14 F. Madeira, M. Pearce, A. R. N. Tivey, P. Basutkar, J. Lee, O. Edbali, N. Madhusoodanan, A. Kolesnikov and R. Lopez, *Nucleic Acids Research*, 2022, **50**, W276–W279.

15 A. M. Waterhouse, J. B. Procter, D. M. Martin, M. Clamp and G. J. Barton, *Bioinformatics (Oxford, England)*, 2009, **25**, 1189–1191.

16 J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz and J. P. Mesirov, *Nature Biotechnology*, 2011, **29**, 24–26.

17 B. Hong, L. Wang, E. Lammertyn, N. Geukens, L. Van Mellaert, Y. Li and J. Anné, *Microbiology (Reading, England)*, 2005, **151**, 3137–3145.

© The National High School Journal of Science 2024

Supplementary materials

Table 3, S1

pheA
>NZ_CP012646.1:c1615757-1614909 Streptococcus mitis strain KCOM 1350 (= ChDC B183) chromosome, complete genome
>NZ_LR134283.1:290478-291311 Streptococcus anginosus strain NCTC10713 chromosome 1, complete sequence
>NZ_CP077404.1:c1083085-1082261 Streptococcus mutans strain FDA ARGOS 1458 chromosome, complete genome
>NZ_CP029490.1:796502-797326 Streptococcus sobrinus strain SL1 chromosome, complete genome
>NZ_CDBW01000016.1:c27766-26594 Aeromonas sobria strain CECT 4245, whole genome shotgun sequence
>NZ_MCXR02000001.1:2871963-2873138 Vibrio lentus strain 10N.261.52.F12 10N26152F12_0, whole genome shotgun sequence
>NZ_CP045198.1:c1359832-1358723 Acinetobacter indicus strain TQ23 chromosome, complete genome
>NZ_CP020405.2:1536031-1537194 Pasteurella multocida strain FARGOS 218 chromosome, complete genome
>NC_013961.1:c913230-912070 Erwinia amylovora CFBP1430, complete sequence
>NZ_CP083638.1:2408089-2409249 Escherichia fergusonii strain FDAARGOS_1499 chromosome, complete genome