EDN: LTNPWW
УДК 543.424

Raman Spectroscopy Method for Identification of Streptococci

Elena V. Timchenko*
Pavel E. Timchenko†
Institute of Computer Science and Cybernetics
Samara, Russian Federation
Artem V. Lyamin‡
Karim A. Kayumov
Center for Genetic Laboratory Technologies of SamSMU
Samara, Russian Federation
Irina V. Bazhutova§
Institute of Professional Education of SamSMU
Samara, Russian Federation
Larisa T. Volova¶
Samara State Medical University
Samara, Russian Federation
Oleg O. Frolov‖
Alena V. Zotova**
Institute of Computer Science and Cybernetics
Samara, Russian Federation
Filipp R. Bazhutov
Lomonosov Moscow State University
Moscow, Russian Federation

Received 10.08.2024, received in revised form 15.09.2024, accepted 24.10.2024

Abstract. In the modern world, various infammatory diseases of the oral cavity have become widespread, in particular periodontitis. Streptococcs are one of the potential participants in the infammatory process. In this work, three streptococcal strains were studied by Raman spectroscopy. As a result, spectral differences were established and criteria for identifying groups of samples were introduced.

Keywords: raman spectroscopy, streptococcus, periodontitis.

Citation: E.V. Timchenko, P.E. Timchenko, A.V. Lyamin, K.A. Kayumov, I.V. Bazhutova, L.T. Volova, O.O. Frolov, A.V. Zotova, F.R. Bazhutov, Raman Spectroscopy Method for Identification of Streptococci, J. Sib. Fed. Univ. Math. Phys., 2025, 18(2), 169–176.
EDN: LTNPWW.

*lazer-optics.timchenko@mail.ru https://orcid.org/0000-0002-0539-7989
†timpavel@mail.ru https://orcid.org/0000-0003-3089-7966
‡a.v.lyamin@samsmu.ru https://orcid.org/0000-0002-5905-1895
§docba@mail.ru https://orcid.org/0000-0003-3200-5538
¶l.t.volova@samsmu.ru https://orcid.org/0000-0002-6599-7732
‖frolov679@mail.ru
**zotova_alena@bk.ru

Introduction

In the modern world, various inflammatory diseases of the oral cavity have become widespread, in particular periodontitis [1, 2]. The main cause of periodontitis and peri-implantitis is tissue infection by microorganisms of the oral cavity. One of the known potential participants in the pathological process is streptococci, which are detected in periodontal pockets in almost 100 percent of cases [3–6]. At the same time, streptococci remain one of the most difficult to identify microorganisms, even when using modern methods. Currently, an actively used physical method for diagnosing microorganisms, including streptococci, is matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOFMS). No new technology for microorganism identifications without problems and the same is true for MALDI TOFMS. Among the most common errors are the inability to perform an accurate differentiation in those microorgan-isms that have a genotypic/protein profile similarity and an absence of reliable data in the database [7]. In this regard, the urgent task is the detection of species identification of streptococci.

As an alternative method for identifying streptococci, the method of Raman spectroscopy (RS), which has found wide application in biomedical practice [8], can be used. RS allows the analysis of vibrational modes of molecules and can distinguish between similar molecules, which gives hope for solving the problem of identifying closely related bacterial species.

Previously, other authors conducted a similar study, but it focused on the species identification of pneumococcus, as the main causative agent of generalized infections (pneumonia and meningitis) [9]. Given the increasing role of streptococci as causative agents of diseases of various localizations, further research in this direction is required. The aim of the study was a spectral study of strains of three closely related species of streptococci Streptococcus mitis, Streptococcus oralis and Streptococcus pneumoniae using Raman spectroscopy for rapid assessment of bacterial strains in the diagnosis of periodontitis.

Materials and methods of research

4 strains of S.mitis, 4 strains of *S.oralis* and 3 strains of *S.pneumoniae* were used as research objects. All isolates were obtained from the clinical material of patients with oral diseases. The growth of crops was obtained on 5 procent blood agar (HiMedia, India), with the addition of mutton blood (HEM LLC, Russia). Incubation of crops was carried out under microaerophilic conditions. Identification was performed using MALDI-ToF mass spectrometry on a Microflex device (Bruker, Germany). For all strains, an additional study was conducted to determine sensitivity to bile and optochine. Daily cultures of streptococcus were suspended in saline solution to obtain an inoculum with a density of 0.5 McFarland units.

For each species, the spectra were obtained: S.pneumoniae (45 Raman spectra), S.oralis (60 Raman spectra), S.mitis (56 Raman spectra). The Raman spectroscopy method described in detail in [10, 11] was used as the main method for analyzing S.mitis, S.oralis, and S.pneumoniae strains. The Raman spectra were captured using a Sharmrock SR-303i spectrograph with an integrated ANDOR DV-420A-OE digital camera (resolution 0.15 nm in wavelength) cooled to –60°C, including a semiconductor laser (LML-785.0RB-04), an optical raman scattering module (PBL 785).

The normalization of the spectra was carried out by the Extended multiplicative signal correction (EMSC) method. The smoothing of the spectra was carried out by the Maximum Likelihood

Estimation Savitzky-Golay filter (MLE-SG) method [12] with the parameter s = 4. To exclude the contribution of autofluorescence to the Raman spectra, a modified method of subtracting the fluorescent component by polynomial approximation Improved Modified Multi-Polynomial Fitting (ExModPoly) with a polynomial degree of 8 was used. The analysis of the Raman spectra of the samples was carried out in the range 450–1800 cm^{-1}.

Results

Fig. 1 shows the averaged Raman spectra of all the studied samples. As can be seen from Fig. 1, the main analytical indicators are manifested at the level of 527 cm^{-1} (S-S di sulfide stretching in proteins, phosphatidylserine or (S-S) gauche-gauche trans (amino acid cysteine)), 621 cm^{-1} (C-C mode of twisting phenylalanine (proteins)), 1280 cm^{-1} (amide III, CH_2, causing vibrations of the glycine backbone or side chains of proline), 1333 cm^{-1} (guanine), 1445 cm^{-1} (CH_2 bending modes, deformation of C-H proteins, deformation of CH_2/CH_3 in lipids), 1525 microns-1 (amide II), 1692 microns-1 (Stretching CO), 1749 $cm^{-1}1$ (C=O, lipids). On the CR line 621 cm^{-1} (C-C twisting mode of phenylalanine (proteins)) The samples of the *S.mitis* group show a noticeable increase in peak intensity. On the CR line of 1280 cm^{-1} corresponding to Amide III, CH_2 wagging vibrations from glycine backbone or proline sidechains, changes in the intensity amplitude of all the studied groups occur. On the CR line of 1445 cm^{-1} (CH_2 bending modes, deformation C-H bending proteins, deformation CH_2/CH_3 in lipids), the group of samples *S.oralis* has the highest intensity. On the Raman scattering line of 1525 cm^{-1} (Amide II), the group of samples it *S.pneumoniae* has the highest intensity.

Fig. 1 shows the averaged Raman spectra of the samples

On the line 1692 cm^{-1} group of samples *S.mitis* it has a noticeably smaller amplitude. Also, on the line 1749 cm^{-1} (C=O, lipids), the group of samples it S.oralis has the highest intensity. Further, in this work, a nonlinear regression analysis of the spectra was carried out, consisting in their decomposition into the sum of asymmetric Gauss lines to increase the information content of the obtained Raman spectra and subsequent analysis using linear discriminant analysis. The amplitude of the a lines was taken as the criterion variable, depending on the values of the independent regressors dx and x0, which determine the initial conditions of the analysis. The composition of spectral lines was determined on the basis of automatic multi-iterative modeling of 161 Raman spectra and tested based on the results of literature analysis. When modeling the spectral contour, the position of x0 and the half-width of the line (HWHM) dx were fixed for the lines used as a template. During the simulation, the line intensity was selected in the range from 0 to the value of the local maximum of the spectrum in the x0 region. HWHM was limited in the range from 1 to 13 cm^{-1}. This made it possible to achieve high stability of the results when modeling the contour and take into account all shifts of the Raman lines. For additional analysis after separation of the spectral lines of the studied samples, the method of linear discriminant analysis in the RS-tool program was chosen. The drawn lines or areas in the LD-1 and LD-2 space can represent class boundaries, which allows you to predict which group a particular sample belongs to based on its LD-1 and LD-2 values. The points in the graph corresponding to group 1 are concentrated in the area where LD-2 is significantly less than 2. This may indicate specific spectral features inherent in this group. For group 2, we notice the concentration of samples in the upper left part (LD-1 $<$ –8 and LD-2 $>$ 1), which also indicates unique characteristics that distinguish it from other groups. Group 3, in turn, occupies the upper right corner, where the values LD-1 $>$–9.

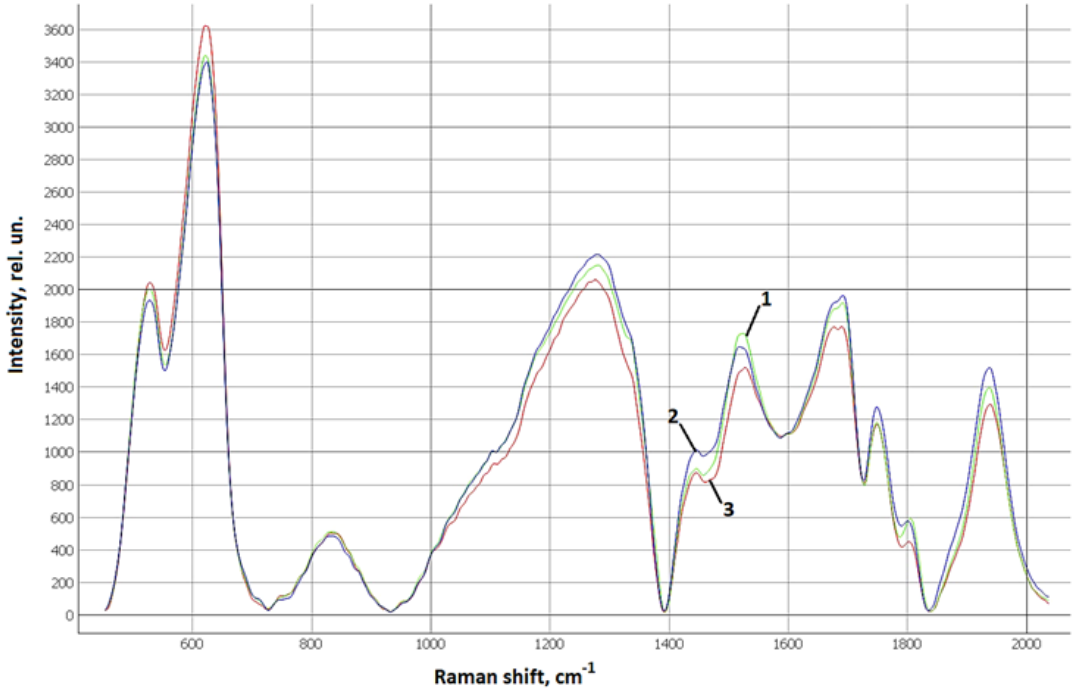

Fig. 1. Averaged Raman spectra of the studied sample groups: 1 — *S.pneumoniae, 2 — S.oralis, 3 — S.mitis*

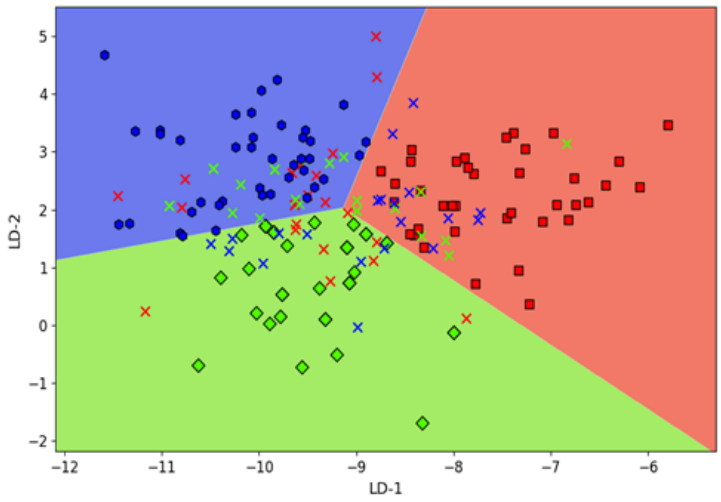

Fig. 2. The results of linear discriminant analysis are a graph of the values of the linear discriminant function

Fig. 3 shows the coeficients of the matrix of the factor structure for the most significant lines of the KR, which have a physical meaning of the correlation between the variables in the model and the discriminating function. The higher the modulo value of LD-1 for a variable, the more it determines the difference in the discriminative model between groups of samples.

It can be noted that the variable with the highest value of SHAP (+ 7.63) is $_k528.73_a$, which indicates its significant influece on the classification. In general, values decrease, which may indicate that variables with high values contribute more to differences between groups than variables with lower values. The smallest value of SHAP (+ 0.22) belongs to the variable $_k1338.65_a$, which indicates its insignificand contribution to the model.

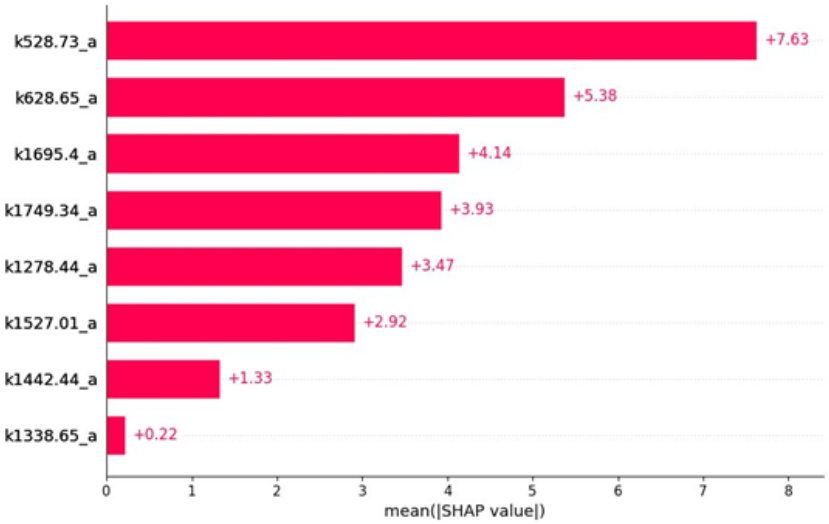

Fig. 3. The contribution of variables to the intensity of lines

The results of the classification of groups are shown in Fig. 4. It can be seen that the number of correctly classified values is approximately equal for each of the groups. For groups 1 and 3, the number of correctly classified values was 9 out of 12, for groups 2–8 out of 12.

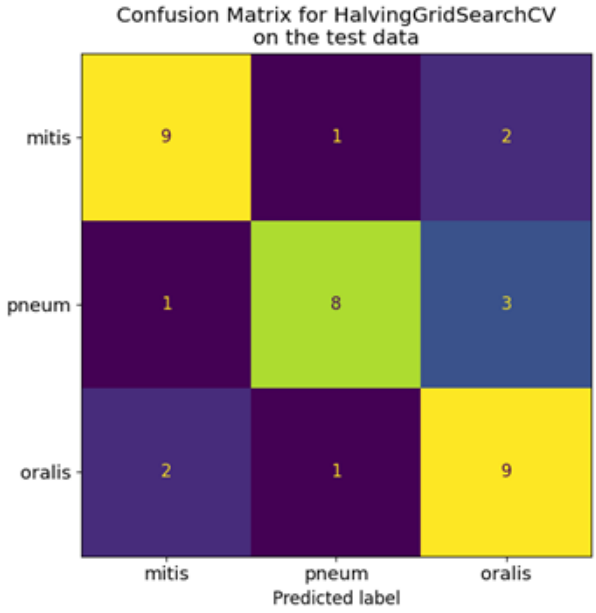

Fig. 4. The Confusion Matrix

Fig. 5 shows the ROC-curves for each sample group. The specificity of the developed algorithm calculated using ROC analysis was 81–91 percent, depending on the defined group. The curves show the ratio of true positive and false positive results, which is important for understanding the accuracy of the classification.
For furthe analysis, AUC (Area Under Curve) values were calculated, which makes it possible to quantify the classification ability of the algorithm. The AUC values for each group were:
mitis (AUC=0.91)
pneunom (AUC=0.89)
oralis(AUC=0.81)
microaverage (AUC=0.86)
macroaverage (AUC=0.87)
The AUC index above 0.8 in all cases indicates the high classification ability of the algorithm, confirming its potential usefulness in practical applications for the diagnosis and classification of varios conditions.

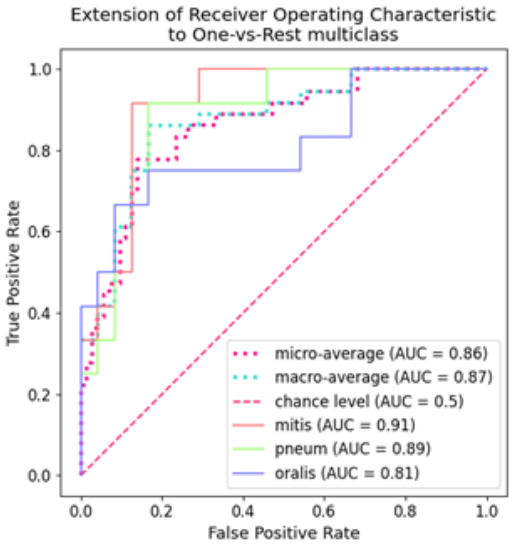

Fig. 5. ROC-curves for all sample groups

Conclusion

In this work, spectral differences of *S.mitis, S.oralis, S.pneumoniae* strains are established. The main spectral differences are visible at 527 cm^{-1} (S-S di sulfide stretching in proteins, Phosphatidylserine or v(S-S) gauche-gauche-trans (aminoacid cysteine)), 621 cm^{-1} (C-C twisting mode of phenylalanine (proteins)), 1280 cm^{-1} (Amide III, CH_2 wagging vibrations from glycine backbone or proline side chains), 1333 cm^{-1} (Guanine), 1445 cm^{-1} (CH_2 bending modes, deformation C-H bending proteins, deformation CH_2/CH_3 in lipids), 1525 cm^{-1} (Amide II), 1692 cm^{-1} (Stretching CO), 1749 cm^{-1} (C=O, lipids). As a result of this study, criteria were introduced for the identification of groups of samples based on the intensity of the lines of the averaged Raman spectra and the conducted discriminant analysis. Thus, for strains of *S.pneumoniae*, the values LD-2 < 2 correspond, for strains *S.oralis* the values LD-1 <-8 and LD-2 >1 correspond, and for strains *S.mitis* the values LD-1 >-9 preferentially. Using ROC analysis, the specificity

of the developed algorithm was calculated, which amounted to 81–91 percent, depending on the defined group. The results obtained will allow further rapid analysis of different types of streptococcal strains using Raman spectroscopy.

References

[1] N.A.Tuneva, B N.V.ogacheva, T U.O.uneva, Problems of dential implantation, *Vyatka medical Bulletin*, **62**(2019), no. 2, 86–93 (in Russian).

[2] A.I.Nikolaev, L.M.Tsepov, Practical therapeutic dentistry, Medpress-inform, 2019.

[3] L.Y.Orekhova, Periodontal diseases, Poly Media Press, 2004, 218–220

[4] I.V.Bazhutova, D.D.Ismatullin, A.V.Lyamin, D.A.Trunin, Species diversity of streptococci in pertodontal diseases, *Immunopathology, allergology, infectology*, (2023), no. 1, 81–86. DOI: 10.14427/jipai.2023.1.81

[5] D.D.Ismatullin, P.V.Bochkareva, I.V.Bazhutova, A.V.Lyamin, R.R.Galieva, A.V.Zhestkov, M.O.Zolotov, Species diversity of streptococci in local and systemic infections, *Clinical laboratory diagnostics*, **68**(2023), no. 10, 597–603 (in Russian). DOI: 10.51620/0869-2084-2023-68-10-597-603

[6] N.A.Tuneva, N.V.Bogacheva, Comparative assessment of microbial contamination in periodontitis and peri-implantitis, *Bulletin of the Perm University. Ser. Biology*, (2021), no. 2, 101–109. DOI: 10.17072/1994-9952-2021-2-101-109 (in Russian).

[7] M.Ercibengoa, MALDI-TOF MS breakpoint limitations for streptococcus mitis group species typing, *Clinical Laboratory*, **63**(2017), no. 11-12, 1935–1938. DOI: 10.7754/Clin.Lab.2017.170619)

[8] A.Carden, R.M.Rajachar, M.D.Morris, D.H.Kohn, Ultrastructural Changes Accompanying the Mechanical Deformation of Bone Tissue: A Raman Imaging Study, *Calcif Tissue Int.*, **72**(2003), 166–175.

[9] S.Berus, E.Witkowska, K.Niciski, E.Sadowy, W.Puzia, P.Ronkiewicz, A.Kamiska, Surface-enhanced Raman scattering as a discrimination method of Streptococcus spp. And alternative approach for identifying capsular types of S. Pneumonia isolates, *Spectrochimica Acta. Part A: Molecular and Biomolecular Spectroscopy*, **233**(2020). DOI: 10.1016/j.saa.2020.118088

[10] P.E.Timchenko, E.V.Timchenko, L.T.Volova, O.O.Frolov, Use of Raman spectroscopy for the assessment of dentin materials during their fabrication, *Journal of Optical Technology*, **88**(2021), no. 9, 485–488. DOI: 10.1364/JOT.88.000485

[11] P.E.Timchenko, E.V.Timchenko, L.T.Volova, O.O.Frolov Spectral Analysis of Organic Components of Demineralized Bone Biografts, *Optics and Spectroscopy (English translation of Optika i Spektroskopiya)*, **126**(2019), no. 6, 769–775. DOI: 10.1134/S0030400X19060262

[12] P.E.Timchenko, E.V.Timchenko, L.T.Volova, M.A.Zybin, O.O.Frolov, G.G.Dolgushov, Optical Assessment of Dentin Materials, *Optical Memory and Neural Networks*, **29**(2020), no. 4, 354–357. DOI: 10.3103/S1060992X20040116

Метод спектроскопии комбинационного рассеяния для идентификации стрептококков

Елена В. Тимченко
Павел Е. Тимченко
Институт информатики и кибернетики
Самара, Российская Федерация
Артём В. Лямин
Карим А. Каюмов
Центр генетических лабораторных технологий СамГМУ
Самара, Российская Федерация
Ирина В. Бажутова
Институт профессионального образования СамГМУ
Самара, Российская Федерация
Лариса Т. Волова
Самарский государственный медицинский университет
Самара, Российская Федерация
Олег О. Фролов
Алёна В. Зотова
Институт информатики и кибернетики
Российская Федерация
Филипп Р. Бажутов
Московский государственный университет имени М. В. Ломоносова
Москва, Российская Федерация

Аннотация. В современном мире широкое распространение приобрели различные воспалительные заболевания полости рта, в частности пародонтит. Одним из потенциальных участников воспалительного процесса являются стрептококки. В данной работе методом спектроскопии комбинационного рассеяния были изучены три штамма стрептококков. В результате были установлены спектральные отличия, введены критерии для идентификации групп образцов.

Ключевые слова: спектроскопия комбинационного рассеяния, стрептококки, пародонтит.