Bionatura
Latin American Journal of Biotechnology and Life Sciences

Bionatura Issue 1 Vol 8 No 1 2023

Article

# Identification of clinical Proteus isolates by using *rpoB* gene sequence analysis

Hussein Ali Mutlag and Intesar N. Khelkal
[1] Biology Department, College of Science, Al-Mustansiriyah University, Iraq.
* Correspondence: Intesarnkshaibani2021@Uomustansiriyah.edu.iq, Husseinalsaleem84@gmail.com
*Available from: http://dx.doi.org/10.21931/RB/CSS/2023.08.01.74*

**ABSTRACT:** Some bacterial genera and species could not be identified by 16S rRNA sequencing, *so the rpoB* gene has been suggested to be the alternative because the sequence of *the rpoB* gene is more distinguished than the 16S rRNA gene to discriminate various species of bacteria. Thus, this study has been designed to show the usefulness *of the rpoB* gene in accurately identifying Proteus clinical isolates. Fifty bacterial isolates were collected and identified phenotypically according to the culture, microscopic examination and biochemical tests. VITEK 2 compact system was used to confirm identification. Genotypic identification was performed after DNA extraction for 10 selected isolates and amplification with *rpoB* gene-specific primer and gel electrophoresis. The products were detected with a (1090 bp) molecular size band, which was sent for Sanger sequencing using an ABI3730XL automated DNA sequencer, and data were analyzed and compared with standard sequences in GenBank. It has been concluded that identifying and discriminating Proteus species by *rpoB* gene sequence analysis is more correct and accurate.

**Keywords:** 16S rRNA, *rpoB* gene, *Proteus spp.* identification, sequencing.

## Introduction

Sequencing of the 16S ribosomal RNA gene has been initiated for analyzing and identifying bacterial groups since the 1970s, aiming for the 16S rRNA gene, which has regions that are both highly conserved and widely variable (hypervariable) among bacterial species. The highly conserved regions allow for the design of "universal" primers in PCR to target and amplify the 16S rRNA sequence, while the hypervariable regions permit differentiating among diverse bacterial clades. These characters allow 16S rRNA sequencing tests to capture nearly all bacterial species, which can then be compared to large 16S rRNA databases to conclude their characteristics [1].

Among other core bacterial genes, the rpoB gene has been recognized as one of the little possible candidates suitable for bacterial phylogenetic analyses, and so sequence analysis of the RNA polymerase ß subunit has been suggested as a powerful tool for universal bacterial genetic identification; the usage of the rpoB gene offers many potential benefits above standard 16SrRNA gene-based methods [2]. It codes for 1342 amino acids, rendering it the second-largest polypeptide in the bacterial cell ( Goldstein,2014), and this gene is also the site of mutations responsible for resistance to rifampin [3].

The sequence of the rpoB gene is more distinguished than the 16S rRNA gene in discriminating various species of bacteria owing to the divergence levels of the rpoB gene sequence that is explicitly higher than the 16S rRNA gene. Furthermore, the partial rpoB gene sequence displays the exact reading frame, leading to the easy confirmation of sequence accuracy. Thus, the rpoB gene is an excellent tool for bacterial identification [4].

Primers and probes targeting the rpoB gene have been used for the specific Detection and phylogenetic analysis of numerous bacterial groups of the family Enterobacteriaceae [5]. Proteus vulgaris and P. penneri could not be identified by the 16S rRNA gene, so another indicator has been suggested, including the beta subunit of RNA polymerase, rpoB [6].

## Materials

Bacterial isolates: Fifty bacterial isolates suspected to belong to Proteus spp. have been collected from different clinical sources: stool, urine, ear and wound swabs.

*Markers*

Promega/USA provided DNA ladder (100-1500 bp).

*Oligonucleotide Primers*

The RPO B-specific primer F: AACCAGTTCCGCGTTGGCCTGG R: CCTGAACAACACGCTCGGA was provided by Macrogen /Korea [5].

*Standard Strain*

Standard strain Proteus mirabilis ATCC 29245 was obtained from Al-Kindy Hospital/Baghdad.

*Methods*

Identification of bacterial isolates

Isolation and identification of bacterial isolates were performed according to standard bacteriological techniques [7]. All colonies suspected as Proteus in primary cultures were purified by subculture on different media. Species were identified phenotypically according to the morphological features on culture media, microscopic examination, and biochemical tests [8, 9]. VITEK 2 compact system was used to confirm identification.

Detection of rpoB gene

DNA extraction and rpoB amplification

According to the protocol of ABIO pure extraction, Genomic DNA was isolated from the bacterial growth of all fifty bacterial isolates.DNA was quantified, and primer stock was prepared. PCR master was mixed, and primer solutions were stirred by vortex at room temperature. The components of the tubes were mixed by vortex and then placed in the PCR apparatus.

Gel electrophoresis

The integrity of extracted DNA was examined by 1% agarose gel electrophoresis and viewed using a UV trans illuminator. Gel electrophoresis was adopted to confirm the presence of amplification. The horizontal agarose gel was cast, and DNA was loaded.

Sequencing of the rpoB gene

The PCR product from amplifying the rpoB gene for both the forward and reverse strands of (10) selected species was sent for Sanger Sequencing using an ABI3730XL automated DNA sequencer, by Macrogen Corporation – Korea. The results were analyzed and edited by using GENEIOUS software. The edited sequence was compared with the database using BLAST (Basic Local Alignment Search Tool) to detect the close relation with submitted sequences.

## Results

*Identification of bacterial isolates*

Several conventional biochemical tests were done to characterize the suspected Proteus isolates, and the results indicated that these isolates were primarily identified as Proteus spp. To confirm the identification results, VITEK 2 Compact automated system results showed that 49 isolates were identified as P. mirabilis (98%), showed a confidence value of 99-96% (excellent identification) and only one isolate was identified as P. Hauser (2 % )

*Detection of rpoB*

The rpoB gene encodes the β subunit of bacterial RNA polymerase, also known as housekeeping genes. It has emerged as a core gene candidate for phylogenetic analyses and identification of bacteria, especially when studying closely related isolates.

To detect the rpoB gene in Proteus isolates, DNA was extracted from all samples, and conventional PCR was carried out in order to amplify this gene by using a specific rpoB primer; when the two aligned together to form a visible band by gel electrophoresis, all of these isolates (100%) indicated by the presence of band with molecular size ( 1090 bp ) when compared with ladder; as shown in figure 1.

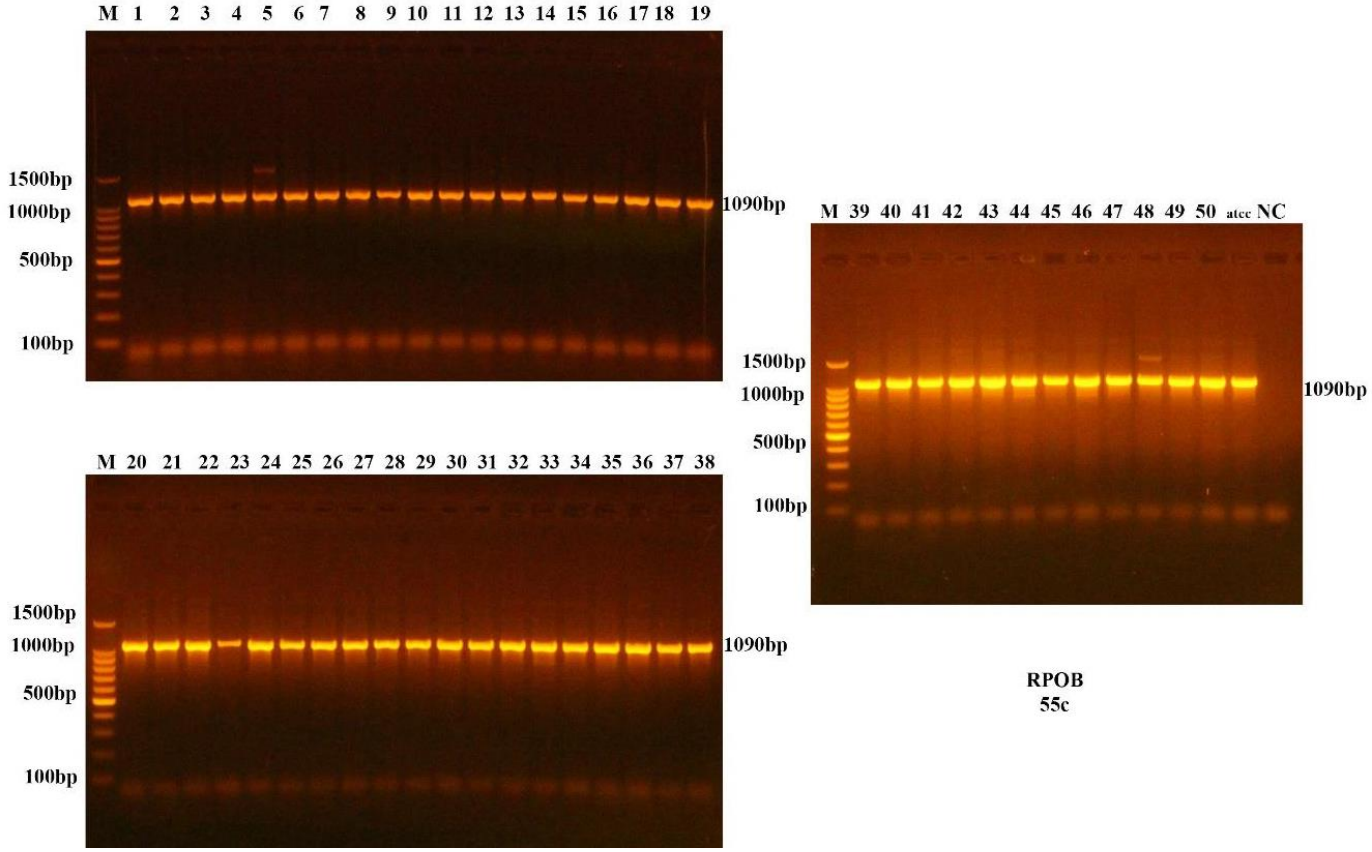


**Figure 1. The amplification results of *the rpoB* gene of *Proteus spp.* Samples (1-51 ) were fractionated on 1.5% agarose gel electrophoresis and stained with Eth. Br. M: 100bp ladder marker.**

*Proteus mirabilis ATCC 29245*

*Sequence analysis of rpoB gene (genotypic identification)*

Sequencing of the rpoB gene using the Sanger method has been performed for ten Proteus isolates figure 2 and the sequences that obtained were deposited in the GenBank with the following accession no: LC699880, LC699881, LC699882, LC699883, LC699884, LC699885, LC699886, LC699887, LC699888, LC699889, the obtained DNA sequences were compared with standard sequences in GenBank also with numerical chain maximum identification ratio.

tgcattaatg ccacaagata tgatcaatgc aaaaccgatt tctgctgcag ttaaagagtt
ctttggctct agccagttat cgcagtttat ggatcagaat aacccgctgt cagaaattac
ccataaacgt cgtatttctg cattaggccc tggtggtctg actcgtgaac gtgcaggctt
cgaagttcga gacgtacacc caactcacta tggtcgtgtg tgtccaatcg aaacacctga
aggtccaaat atcggtctga ttaactcatt atctgtttat gcacagacta acgagtacgg
attcctagaa acgccttacc gtgttgtaga aaacaacgct gtaacagatg aaatccacta
tctgtctgca attgaagaag gtaacttcat cattgctcag gctaactccg tattagatga
tgatggccac ttcgtagaag aattagtaac ttgccgtcat aaaggtgagt caagtttatt
tagtcgtgat caagttcagt acatggacgt ttcaactcaa caggttgttt ctgtcggtgc
ttcactgatc ccattccttg aacacgatga cgccaaccgt gcattgatgg gtgcgaacat

gcaacgtcag gctgttccta cacttcgtgg tgataaacca ctggtaggta caggtatgga
acgtgctgta gcggttgact ccggtgttac tgcggtagca aaacgtggtg gtactgttca
gtatgttgat gcttctcgta tcgttatcaa agtaaacgaa gatgagactt atgctggtga
agcaggcatt gatatttaca gcttaactaa atatacccgc tctaaccaaa acacatgtat
taaccaaaca ccttgtgttt ctttaggtga acctgttgaa cgtggtgatg tact
A
tgcattgatg ccacaagata tgatcaacgc aaaaccgatt tctgctgctg tcaaagaatt
ctttggctct agtcagttat cacagtttat ggatcagaat aacccgctat cagaaattac
gcacaaacgt cgtatttctg cattaggccc aggtggtttg acccgtgaac gtgcaggctt
tgaagttcga gacgtacacc caactcacta cggtcgtgtg tgcccaatcg aaacgcctga
aggtccaaac atcggtctta tcaactcatt atctgtttat gcacagacta acgagtacgg
tttcctagaa acgccttacc gtgttgttga aaacaacgct gtaacagatg aaattcacta
tctgtctgca atcgaagaag gtaacttcat cattgctcag gctaactccg tattagacga
tgacaatcac tttgttgaag aattagttac ttgccgtcat aaaggtgagt caagcttatt
tagtcgtgac caagttcagt acatggacgt ttcaacccaa caggttgttt ctgtcggtgc
ttcactgatc ccattccttg aacacgatga cgccaaccgt gcattgatgg gtgcgaacat
gcaacgtcag gctgttccta cacttcgtgg tgataaacca ctagtaggta caggtatgga
acgtgcagta gcggttgact ccggtgttac tgcggttgct aaacgtggtg gtgttgttca
gtatgttgat gcttctcgta tcgttatcaa agtaaacgaa gaagagactt acgctggtga
agcaggcatt gatatttaca gcttgactaa atacacccgc tctaaccaaa acacatgtat
taaccaaact ccttgtgttt ctttaggtga acctgttgaa cgtggtgatg tactt
B

**Figure 2: Sequences of rpoB gene: A- Proteus mirabilis HIMUS1 B- Proteus terrae HIMUS7.**

The sequences have been given a similarity range between 98.76 % - and 100% in Table 1. According to rpoB gene sequences in GenBank, the ten strains have been recognized into two species, P.mirabilis and P.terrae.

| | CP032663 | CP055009 | CP063440 | CP065039 | CP066833 | CP070569 | CP085481 | CP089317 | DQ836273 | LC699880 | LC699881 | LC699882 | LC699883 | LC699884 | LC699885 | LC699886 | LC699887 | LC699888 | LC699889 | LN809885 | MG386493 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CP032663 | | 94.1% | 94.1% | 94.1% | 94.1% | 94.1% | 98.5% | 94.1% | 97.9% | 94.1% | 93.9% | 94.1% | 94.1% | 93.9% | 94.1% | 93.9% | 97.8% | 93.9% | 94.1% | 98.4% | 98.6% |
| CP055009 | 94.1% | | 100% | 100% | 100% | 100% | 93.7% | 100% | 93.9% | 100% | 99.7% | 100% | 99.8% | 99.7% | 100% | 99.9% | 94.9% | 99.7% | 99.8% | 94.0% | 94.3% |
| CP063440 | 94.1% | 100% | | 100% | 100% | 100% | 93.7% | 100% | 93.9% | 100% | 99.7% | 100% | 99.8% | 99.7% | 100% | 99.9% | 94.9% | 99.7% | 99.8% | 94.0% | 94.3% |
| CP065039 | 94.1% | 100% | 100% | | 100% | 100% | 93.7% | 100% | 93.9% | 100% | 99.7% | 100% | 99.8% | 99.7% | 100% | 99.9% | 94.9% | 99.7% | 99.8% | 94.0% | 94.3% |
| CP066833 | 94.1% | 100% | 100% | 100% | | 100% | 93.7% | 100% | 93.9% | 100% | 99.7% | 100% | 99.8% | 99.7% | 100% | 99.9% | 94.9% | 99.7% | 99.8% | 94.0% | 94.3% |
| CP070569 | 94.1% | 100% | 100% | 100% | 100% | | 93.7% | 100% | 93.9% | 100% | 99.7% | 100% | 99.8% | 99.7% | 100% | 99.9% | 94.9% | 99.7% | 99.8% | 94.0% | 94.3% |
| CP085481 | 98.5% | 93.7% | 93.7% | 93.7% | 93.7% | 93.7% | | 93.7% | 98.9% | 93.7% | 93.6% | 93.7% | 93.7% | 93.6% | 93.7% | 93.6% | 98.0% | 93.6% | 93.7% | 98.6% | 98.9% |
| CP089317 | 94.1% | 100% | 100% | 100% | 100% | 100% | 93.7% | | 93.9% | 100% | 99.7% | 100% | 99.8% | 99.7% | 100% | 99.9% | 94.9% | 99.7% | 99.8% | 94.0% | 94.3% |
| DQ836273 | 97.9% | 93.9% | 93.9% | 93.9% | 93.9% | 93.9% | 98.9% | 93.9% | | 93.9% | 93.8% | 93.9% | 93.9% | 93.8% | 93.9% | 93.8% | 97.8% | 93.8% | 93.9% | 98.1% | 98.3% |
| LC699880 | 94.1% | 100% | 100% | 100% | 100% | 100% | 93.7% | 100% | 93.9% | | 99.7% | 100% | 99.8% | 99.7% | 100% | 99.9% | 94.9% | 99.7% | 99.8% | 94.0% | 94.3% |
| LC699881 | 93.9% | 99.7% | 99.7% | 99.7% | 99.7% | 99.7% | 93.6% | 99.7% | 93.8% | 99.7% | | 99.7% | 99.7% | 100% | 99.7% | 99.8% | 94.7% | 99.8% | 99.9% | 93.9% | 94.1% |
| LC699882 | 94.1% | 100% | 100% | 100% | 100% | 100% | 93.7% | 100% | 93.9% | 100% | 99.7% | | 99.8% | 99.7% | 100% | 99.9% | 94.9% | 99.7% | 99.8% | 94.0% | 94.3% |
| LC699883 | 94.1% | 99.8% | 99.8% | 99.8% | 99.8% | 99.8% | 93.7% | 99.8% | 93.9% | 99.8% | 99.7% | 99.8% | | 99.7% | 99.8% | 99.7% | 94.9% | 99.7% | 99.8% | 94.0% | 94.3% |
| LC699884 | 93.9% | 99.7% | 99.7% | 99.7% | 99.7% | 99.7% | 93.6% | 99.7% | 93.8% | 99.7% | 100% | 99.7% | 99.7% | | 99.7% | 99.8% | 94.7% | 99.8% | 99.9% | 93.9% | 94.1% |
| LC699885 | 94.1% | 100% | 100% | 100% | 100% | 100% | 93.7% | 100% | 93.9% | 100% | 99.7% | 100% | 99.8% | 99.7% | | 99.9% | 94.9% | 99.7% | 99.8% | 94.0% | 94.3% |
| LC699886 | 93.9% | 99.9% | 99.9% | 99.9% | 99.9% | 99.9% | 93.6% | 99.9% | 93.8% | 99.9% | 99.8% | 99.9% | 99.7% | 99.8% | 99.9% | | 94.7% | 99.8% | 99.9% | 93.9% | 94.1% |
| LC699887 | 97.8% | 94.9% | 94.9% | 94.9% | 94.9% | 94.9% | 98.0% | 94.9% | 97.8% | 94.9% | 94.7% | 94.9% | 94.9% | 94.7% | 94.9% | 94.7% | | 94.7% | 94.9% | 98.1% | 98.1% |
| LC699888 | 93.9% | 99.7% | 99.7% | 99.7% | 99.7% | 99.7% | 93.6% | 99.7% | 93.8% | 99.7% | 99.8% | 99.7% | 99.7% | 99.8% | 99.7% | 99.8% | 94.7% | | 99.9% | 93.9% | 94.1% |
| LC699889 | 94.1% | 99.8% | 99.8% | 99.8% | 99.8% | 99.8% | 93.7% | 99.8% | 93.9% | 99.8% | 99.9% | 99.8% | 99.8% | 99.9% | 99.8% | 99.9% | 94.9% | 99.9% | | 94.0% | 94.3% |
| LN809885 | 98.4% | 94.0% | 94.0% | 94.0% | 94.0% | 94.0% | 98.6% | 94.0% | 98.1% | 94.0% | 93.9% | 94.0% | 94.0% | 93.9% | 94.0% | 93.9% | 98.1% | 93.9% | 94.0% | | 99.8% |
| MG386493 | 98.6% | 94.3% | 94.3% | 94.3% | 94.3% | 94.3% | 98.9% | 94.3% | 98.3% | 94.3% | 94.1% | 94.3% | 94.3% | 94.1% | 94.3% | 94.1% | 98.1% | 94.1% | 94.3% | 99.8% | |

**Table 1. Similarity ratio in sequences of rpoB gene between Proteus strains in the current study and reference strain. Strains in this table were prepared by Sanger sequencing using ABI3730XL, automated DNA sequences, by GENEIOUS software.**

Strains in this study have differences in the rpoB gene sequences of the same species belonging to P.mirabilis, and other strains of P.mirabilis have the same rpoB gene sequence.

It was observed that the difference between the rpoB gene sequence of P.mirabilis strains ranged between ( 0.01 – 0.03 % ) and the difference between the rpoB gene sequence of P. mirabilis and P. terrae ( 5.1 – 5.3 % ).

Sequencing of the entire rpoB gene might be necessary when describing a new bacterial species, using a delineation cut-off of less than 97.7 % or a bacterial subspecies, using a delineation cut-off of more (300 – 750-bp) rpoB gene sequencing is sufficient for accurate identification of most clinical isolates.

## Discussion

With the 16S rRNA gene, rpoB has helped describe new bacterial species and upgrade bacterial community analysis. Furthermore, these housekeeping genes are generally present as single copies in bacterial genomes, regulating the overestimation of operational taxonomical units (OTUs) in microbial assemblages [3]. The gene detection results are similar to those of a previous local study by [2] who isolated P.mirabilis from UTI samples and recognized the rpoB gene with the same molecular size band (1090 bp). Moreover**, c**onfirmation of the bacterial identification was accomplished by molecular genotypic identification. Sequence analysis of the RNA polymerase ß subunit encoding gene (rpoB) has been proposed as a novel tool for bacterial identification. Also, the partial rpoB gene sequence displays the exact reading frame leading to the easy verification of sequence accuracy [10]. [4] have been indicated the reliability and importance of the rpoB gene as a vital tool for bacterial identification by their experiment; A portion of the coding region of the rpoB gene from two species of Proteus was amplified with specific primers. Strain no. 36, which was identified as P. Hauser by VITEK 2; it appeared that this isolate belonged to the species P. terrae after conducting a gene sequencing for bacterial isolates under study because the VITEK 2 compact system could not detect the species P.terrae according to the instructions for the device that is identifying only four species; P.mirabilis, P. vulagaris, P. penner and P Hauser 11. [12] isolated the novel species P .terrae from the soil; identification was achieved according to rpoB gene sequence analysis. P .terrae showed high similarity with the respective type strains of P. vulgaris and P. penneri. Partial sequences might even demonstrate a more significant percentage of difference among isolates than sequencing of the entire gene [13].

## Conclusion

In conclusion, rpoB sequencing proved able to characterize the different species of the genus Proteus on a molecular basis. However, our results provide further evidence for the existence of

Genetic differences within Proteus spp.

## Acknowledgments

Thanks to the bacteriology labs. Of Al-Kindi Teaching Hospital and some other hospitals in Baghdad and Dyala for their assistance in providing the samples. Also, we thank the Biology Department / College of Science / Mustansiriyah University (http://uomustansiriyah.edu.iq) / Baghdad for advice and support.

## References


1 Lu, J., & Salzberg, S. L. Ultrafast and accurate 16S rRNA microbial community analysis using Kraken 2. *Microbiome,* **2020**; *8(1*), 1-11.

2 Jamel, A. N.; Allami, R. H., & Hamza, S. J. Molecular study of *rpoB* gene In *Proteus mirabilis* isolated from urinary tract infection from different hospitals in Baghdad. *Plant Archives,* **2020**; *20(1),* 237-.2383.

3 Ogier, J. C.; Pages, S.; Galan, M.; Barret, M., & Gaudriault, S. *rpoB*, a promising marker for analyzing the diversity of bacterial communities by amplicon sequencing. *BMC microbiology* , **2019**; *19(1),* 1-16.

4 Rungsirivanich, P.; Inta, A.; Tragoolpua, Y., & Thongwai, N. Partial rpoB gene sequencing identification and probiotic potential of Floricoccus penangensis ML061-4 isolated from Assam tea (Camellia sinensis var. assamica). *Scientific Reports,* **2019**,*9(1),* 1-12.

5 Giammanco, G.; Pignato, S.; Grimont, P. A., Grimont, F., and Giammanco, G. M. Genotyping of the genus Proteus by *rpoB* sequence analysis. *Ital J Public Health.* **2011**. [serial on internet]. Available from*: ijphjournal. it/article/download/5992/5737.*

6 Mohammed, G.J.; Kadhim, M.J., and Hameed, I.H. Proteus species: Characterization and herbal antibacterial: A review. *Int J Pharmacogn Phytochem Res.* **2016**, *8(11):*1844–54.

7 Pathirana, H. N. K. S.; De Silva, B. C. J.; Wimalasena, S. H. M. P.; Hossain, S., & Heo, G. J. Comparison of virulence genes in *Proteus species* isolated from human and pet turtle. *Iranian journal of veterinary research,* **2018**, *19(1), 48.*

8 Forbes, B. A.; Daniel, F. S., and Alice, S. W. B. Scott's. Diagnostic Microbiology. 12th ed., Mosby Elsevier Company. USA. **2007**. pp62-465.

9 Brooks, G. F.; Carroll, K. C.; Butel, J. S.; Morse, S. A., and Mietzner, T. A. Medical Microbiology. 26th edit.pp46-48.

10 Lin, J. N.; Lai, C. H.; Yang, C. H., & Huang, Y. H. (2019). Elizabethkingia infections in humans: from genomics to clinics. *Microorganisms,* **2013**, *7(9),* 295.

11 BioMérieux. VITEK 2 Identification Cards | bioMérieux United Kingdom & Ireland [Internet]. Available from: https://www.biomerieux.co.uk/product/vitek-2- identification-cards. **2021**.

12 -Behrendt, U., Augustin, J., Spröer, C., Gelbrecht, J., Schumann, P. and Ulrich, A. Taxonomic characterization of Proteus terrae sp. nov., an N2O-producing, nitrate-ammonifying soil bacterium. *Antonie Van Leeuwenhoek,* **2015**, *108(6),* 1457-1468.

13 Adékambi, T.; Drancourt, M., and Raoult, D. The rpoB gene as a tool for clinical. **2009**.

14 microbiologists. *Trends Microbiol.* 17(1):37–45.


Received: May 15, 2023/ Accepted: June 10, 2023 / Published: June 15, 2023
Citation: Mutlag, H.A.; Khelkal, I.N. Identification of clinical Proteus isolates by using rpoB gene sequence analysis. Revis Bionatura 2023;8 (1) 74. http://dx.doi.org/10.21931/RB/CSS/2023.08.01.74