Heliyon 10 (2024) e37114

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Genomic insights of the emerging human pathogen *Proteus appendicitidis* sp. nov

Yao Peng [a,1], Yanting Wang [b,1], Xingxing Kang [c,**], Xunchao Cai [a,*]

[a] Department of Gastroenterology and Hepatology, Shenzhen University General Hospital, Shenzhen University, Shenzhen, 518055, China
[b] Department of Biochemical Pharmacy, School of Pharmacy, Naval Medical University, Shanghai, 200433, China
[c] Department of Bioinformatics, School of Medical Informatics, Xuzhou Medical University, Xuzhou, 221004, China

ARTICLE INFO

Keywords:
Proteus appendicitidis
Antimicrobial resistant genes
Virulence
Genomic insights
Human pathogen

ABSTRACT

Objectives: *Proteus* are known as opportunistic human pathogens that can cause a variety of infections. *Proteus appendicitidis* is a novel *Proteus* species associated with appendicitis, whereas their genomic characteristics and virulence potential remain understudied. This study aims to compare the genomic features of *P. appendicitidis* to that of the close *Proteus* species, and to assess its virulence-factor encoding capacity as an emerging pathogen.

Methods: Genomes similar to that of *P. appendicitidis* HZ0627^T were retrieved from the PATRIC-v3.6.10 web-server using the implanted Similar Genome Finder tool. Average nucleotide identity (ANI) between HZ0627^T and the retrieved genomes was calculated using FastANI-v1.33. Core-genome sequences were extracted using Roary-v3.13.0, and core-genomic tree was constructed using FastTree-v2.1.11. Virulence-factor encoding capacity was predicted using PathoFact-v1.0.

Results: Two previously unclassified *Proteus* sp. strains were reclassified as *P. appendicitidis*. Strains phylogenomically close to *P. appendicitidis* were clustered into five species, three of which were previously categorized under *P. vulgaris* biogroup 3. Remarkably, *Proteus* genomosp. 6 was identified as the closest species to *P. appendicitidis*, exhibiting ANI values ranging from 94.45 % to 94.95 % against HZ0627^T. Genome annotation revealed shared genomic features and antimicrobial resistance (AMR) genes between *P. appendicitidis* and its phylogenetic neighbors. Additionally, *P. appendicitidis* is hypothesized to share infection mechanisms with *Proteus* genomosp. 6, as evidenced by the encoding of numerous virulence factors implicated in cell lysis and membrane pore-formation in the genome of both species.

Conclusions: This study provides genomic insights of *P. appendicitidis* sp. nov. and its taxonomic relatives, shedding light on their evolutionary relationships, pathogenic mechanisms, and AMR profiles. The findings are significant for the development of targeted therapeutic interventions against infections caused by this emerging pathogen.

The genus *Proteus* is a group of Gram-negative bacteria within the Morganellaceae family [1], known for their opportunistic pathogenicity and remarkable adaptability across diverse ecological habitats, such as soil, aquatic environments, and the intestinal

* Corresponding author.
** Corresponding author.
E-mail addresses: kxxouc@163.com (X. Kang), cxch1224@126.com (X. Cai).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e37114
Received 20 July 2024; Accepted 27 August 2024
Available online 28 August 2024

tracts of both animals and humans [2,3]. Bacteria belonging to this genus are complicated by the high degree of genetic diversity and are associated with a range of human infections [4]. Recently, we have described *Proteus appendicitidis* sp. nov. as a noteworthy addition to this genus. The type strain, designated HZ0627^T, was isolated from the appendiceal pus of a patient diagnosed with appendicitis in China. This strain demonstrated resistance to multiple antibiotics, including ampicillin, cephazolin, nitrofurantoin, cefaclor, cefadroxil, and cefuroxime [5]. The strain's physiological characteristics, along with its chemotaxonomic characteristics, marker-gene-based phylogeny, genomic-based phylogeny, and other genomic characteristics, set it apart from other recognized *Proteus* species [5]. Despite the initial characterization of the type strain, the genomic features of other *P. appendicitidis* strains have yet to be explored. This includes an in-depth analysis of their genome-based phylogenetic relationships, their capacity to encode virulence factors, and their antimicrobial resistance profiles in comparison with other closely related *Proteus* species. In this study, we conducted a comprehensive genomic analysis of *P. appendicitidis* to elucidate its taxonomic position, pathogenic potential, and antimicrobial resistance profile, thereby enhancing our understanding of this emerging pathogen.

Here, we retrieved genome sequences with top 50 genome distance (GD) to strain HZ0627^T from all public genomes of the PATRIC web-server using the implanted Similar Genome Finder tool (https://www.bv-brc.org/app/GenomeDistance) with default parameters (Table S1). Coding sequences of the genomes were predicted and annotated using Prokka-v1.14.5 with default parameters. Subsequently, aligned core genome sequences were extracted separately from the resulting gff files using Roary-v3.11.2 [6]. These alignments were then utilized to construct the core-genomic tree by employing FastTree-v2.1.11 with the generalized time-reversible (GTR) model. The average nucleotide identity (ANI) between all the collected sequences was calculated using fastANI-v1.33 with default parameters [7]. The ANI matrix and GD matrix were visualized using the pheatmap-v1.0.12 R package. Finally, the core-genomic tree

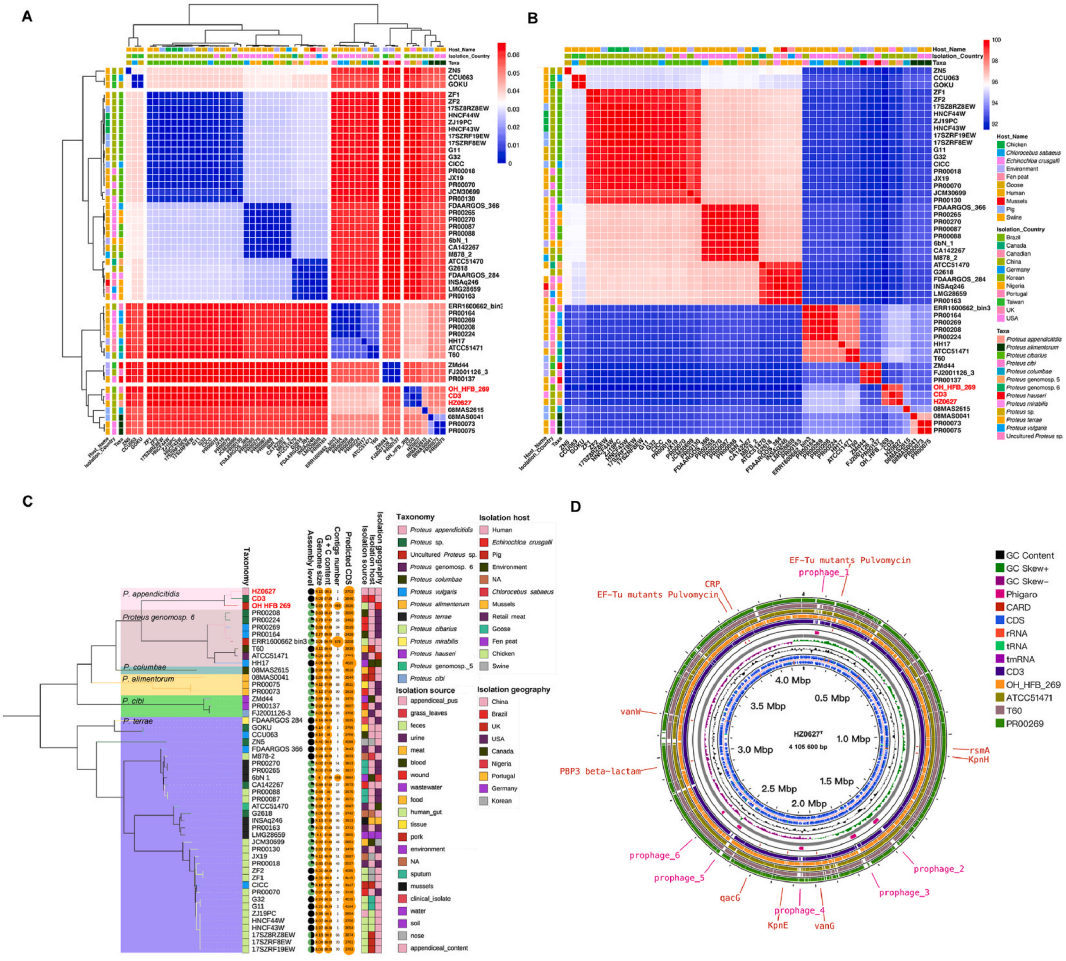

Fig. 1. Genome-based characteristics of the top 50 *Proteus* spp. strains close to *P. appendicitidis*. (A) and (B) displayed the genome distance (GD) and whole-genome-based average nucleotide identity (ANI) of the top 50 *Proteus* spp. strains close to the type strain HZ0627^T, respectively. (C) Core-genomic tree of the *Proteus* spp. strains. Metadata beside the tree were retrieved from the PATRIC web-server, DSMZ or NCBI database. (D) Circular genome sequences view of the *P. appendicitidis* strains (referenced by strain HZ0627^T) and three close *Proteus* genomosp. 6 strains. From inner to outer tracks, CDSs on the reverse strand, CDSs on the forward strand, GC content, GC skew, Prophage (Phigaro), CARD antibiotic resistance genes, and five aligned genome sequences. The RNA sequences are displayed in the CDSs track. The circular view was constructed using the Proksee web-server (https://proksee.ca/).

with metadata features was visualized using iTol-v6 [8]. Notably, the genome similarity analysis indicated that two strains, CD3 and OH_HFB_269, previously classified as *Proteus* sp., exhibit GD and ANI values of 0.0089 and 99.14, and 0.0155 and 98.53, respectively, to HZ0627^T (Fig. 1A and B). These values fall below the proposed thresholds (GD, 0.05; ANI: 95–96 %) for species delineation, suggesting that these strains should be reclassified as *P. appendicitidis*. Interestingly, they were isolated from grass leaves and human feces, distinct from that of the HZ0627^T (Fig. 1C). Another group of strains, including T60, ATCC51471, PR00208, PR00164, PR00269, PR00224, ERR1600662_bin3, and HH17, showed high genomic similarities with HZ0627^T by referencing ANI (94.45–94.95 %) and GD (0.047–0.057) values. Although these strains are currently classified into various taxonomic groups such as *Proteus* genomosp. 6, *P. columbae*, and *P. vulgaris*, our whole-genome ANI comparison suggests that they should be reclassified into a single *Proteus* species represented by *Proteus* genomosp. 6 ATCC51471 (Fig. 1B). By integrating the results of ANI and the core-genomic tree, we identified other four distinct species clusters close to *P. appendicitidis*, including *P. columbae*, *P. alimentorum*, *P. cibi*, and *P. terrae* (i.e., *P. terrae* subsp. *terrae* and *P. terrae* subsp. *cibarius*) (Fig. 1C). The metadata demonstrated that strains correspond to these genomes predominantly isolated from China and the United States, isolated from various environments (e.g., human tissues, body-fluids, feces, sewage, and grass leaves) and hosts (e.g., humans, pigs, plants, and the environment), and are primarily associated with clinical settings and infections. Interestingly, the genomic sequence similarity based on GD and ANI did not exhibit clustering patterns related to the isolation source, host, or geography. This suggests a strong environmental adaptability of these strains, underscoring the importance of monitoring their dissemination (Fig. 1C). Apart from the recently identified *P. columbae* and *P. alimentorum*, these taxa close to *P. appendicitidis* were previously assigned to the *P. vulgaris* biogroup 3, which have been reported sharing highly similar phenotypic characteristics [9]. Furthermore, *P. appendicitidis* shares a very similar genomic size, G + C content, and predicted CDSs with these taxa (Fig. 1C). Additionally, the three *P. appendicitidis* genomes harbor the same antimicrobial resistant (AMR) genes as the three most closely related strains from *Proteus* genomosp. 6, which encode multidrug efflux, beta-lactam antibiotic resistance, etc. (Fig. 1D–Table S2). However, the *P. appendicitidis* genomes encode a higher number of prophages compared to the three *Proteus* genomosp. 6 strains, which may contribute to the genetic and phenotypic divergence of *P. appendicitidis* from closely related *Proteus* species (Fig. 1D–Table S2).

Furthermore, virulence factor coding capacity of *P. appendicitidis* and *Proteus* genomosp. 6 were predicted using PathoFact v1.0 with default parameters [10]. The results revealed that *P. appendicitidis* encodes a multitude of virulence factors, with hemolysin (resulting in cell lysis and tissue damage), enterotoxin B (involved in membrane pore-formation), the phosphotransferase family (engaged in nutrient competition), and the phosphatase family (participating in gene modification or regulation) being the principal and core categories, shared with all *Proteus* genomosp. 6 strains (Fig. S1). Of note, all of the *P. appendicitidis* strains encode the Zinc dependent phospholipase C (resulting in membrane pore-formation), which is only present in two *Proteus* genomosp. 6 strains (Fig. S1).

In conclusion, we refined the taxonomy of 50 *Proteus* strains with the closest genome distance to *P. appendicitidis* at the genomic level. We found that two strains previously classified as *Proteus* sp. should be reclassified as *P. appendicitidis*. Additionally, five distinct *Proteus* species clusters were identified close to *P. appendicitidis*, among which *Proteus* genomosp. 6 is the closest. We disclosed that *P. appendicitidis* shares similar genomic features and AMR genes with these close taxa. Its genome also encodes numerous virulence factors and shares them with *Proteus* genomosp. 6, mainly involved in cell lysis and membrane pore-formation. We recognize the inherent limitations of our study. Firstly, the selection of genomes based on genome similarity to strain HZ0627^T might not capture all relevant strains, particularly those with more divergent sequences that could provide additional insights into the diversity and evolution of *P. appendicitidis*; Secondly, the prediction of virulence factors from genome annotations necessitates experimental validation to ascertain their contribution to pathogenicity. Despite these limitations, our study has unveiled the genomic traits of *Proteus appendicitidis* sp. nov., particularly concerning its evolution, pathogenicity, and antimicrobial resistance, which are pivotal for comprehending its implications for public health and for guiding the development of effective treatment strategies against infections caused by this species. We anticipate that our findings will catalyze additional research endeavors to corroborate our results and to explore the broader research landscape of *P. appendicitidis* and its close relatives.

Funding

The authors acknowledge the Shenzhen Science and Technology Program (JCYJ20230808105208017), the National Natural Science Foundation of China (82103992), the Shuangchuang Ph.D Award of Jiangsu Province (No.JSSCBS20211252), the Initializing Fund of Xuzhou Medical University (No.D2020013), and the Natural Science Foundation of Jiangsu Higher Education Institutions of China (No.22KJB310021).

Ethical approval

Not required.

International standard randomized controlled trial number

Not applicable.

Data availability

All of the genome sequences analyzed in this study are available at the NCBI GenBank database. Detailed accessions of each genome sequence are listed at Table S1.

CRediT authorship contribution statement

Yao Peng: Funding acquisition, Writing – original draft. **Yanting Wang:** Formal analysis, Funding acquisition, Methodology, Writing – review & editing. **Xingxing Kang:** Funding acquisition, Supervision, Writing – review & editing. **Xunchao Cai:** Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Yao Peng reports financial support was provided by Shenzhen Science and Technology Program, Shenzhen, China. Yanting Wang reports financial support was provided by National Natural Science Foundation of China, China. Xingxing Kang reports financial support was provided by Shuangchuang Ph.D Award of Jiangsu Province, Jiangsu, China. Xingxing Kang reports financial support was provided by Initializing Fund of Xuzhou Medical University, Xuzhou, China. Xingxing Kang reports financial support was provided by Natural Science Foundation of Jiangsu Higher Education Institutions of China, Jiangsu, China. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e37114.

References

[1] M. Adeolu, S. Alnajar, S. Naushad, R. Gupta, Genome-based phylogeny and taxonomy of the 'Enterobacteriales': proposal for Enterobacterales ord. nov. divided into the families Enterobacteriaceae, Erwiniaceae fam. nov., Pectobacteriaceae fam. nov., Yersiniaceae fam. nov., Hafniaceae fam. nov., Morganellaceae fam. nov., and Budviciaceae fam. nov, Int. J. Syst. Evol. Microbiol. 66 (12) (2016) 5575–5599.

[2] P. Han, L. Luo, M. Shahid, M. Li, S. Wang, D.B. Nobrega, et al., Prevalence, genetic diversity and antimicrobial resistance of *Proteus mirabilis* isolated from dogs hospitalized in Beijing, Pak. Vet. J. 40 (1) (2020) 61–66.

[3] Q. Kang, X. Wang, J. Zhao, Z. Liu, F. Ji, H. Chang, et al., Multidrug-resistant *Proteus mirabilis* isolates carrying *blaOXA-1* and *blaNDM-1* from wildlife in China: increasing public health risk, Integr. Zool. 16 (6) (2021) 798–809.

[4] F.A. Bakkar, M. Zubair, Molecular study based on *Proteus mirabilis* contamination in chicken meat in selected districts of Kerala, India Int J Dev Res 9 (6) (2019) 28367–28369.

[5] J. He, Z. Zhang, H. Qu, G. Chen, J. Zhou, W. Zhou, et al., *Proteus appendicitidis* sp. nov., isolated from the appendiceal pus of an appendicitis patient in Yongzhou, China, Arch. Microbiol. 206 (4) (2024) 1–8.

[6] A.J. Page, C.A. Cummins, M. Hunt, V.K. Wong, S. Reuter, M.T. Holden, et al., Roary: rapid large-scale prokaryote pan genome analysis, Bioinformatics 31 (22) (2015) 3691–3693.

[7] C. Jain, L.M. Rodriguez-R, A.M. Phillippy, K.T. Konstantinidis, S. Aluru, High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries, Nat. Commun. 9 (1) (2018) 5114.

[8] Letunic I., Bork P., Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool, Nucleic Acids Res. 52 (W1) (2024) W78–W82.

[9] C.M. O'Hara, F.W. Brenner, A.G. Steigerwalt, B.C. Hill, B. Holmes, P.A. Grimont, et al., Classification of *Proteus vulgaris* biogroup 3 with recognition of *Proteus hauseri* sp. nov., nom. rev. and unnamed *Proteus* genomospecies 4, 5 and 6, Int. J. Syst. Evol. Microbiol. 50 (5) (2000) 1869–1875.

[10] L. de Nies, S. Lopes, S.B. Busi, V. Galata, A. Heintz-Buschart, C.C. Laczny, et al., PathoFact: a pipeline for the prediction of virulence factors and antimicrobial resistance genes in metagenomic data, Microbiome 9 (2021) 1–14.