Gut Pathogens

GENOME REPORT Open Access

Genome-based analyses from four clinically-isolated strains refined the taxonomy of *Proteus* genomosp. 6 and revealed their underestimated role in gastrointestinal diseases

Zhiyun Zhang[1,2], Jing He[3,4], Xueqing Liu[1,2], Linqian Huang[5], Zhou Zeng[6], Yao Peng[7*] and Xunchao Cai[7*]

Abstract

Background *Proteus* spp. have long been recognized for their role in urinary tract infections, while recent evidence disclosed their implications in gastrointestinal diseases. Despite this, the taxonomy of clinically-derived *Proteus* spp., particularly those from gastrointestinal samples, remains understudied and is frequently mis-assigned, which limits our understanding of infections caused by these species.

Results Four *Proteus* strains (i.e., DFP240708, LHD240705, TSJ240517 and WDL240414) were isolated from the appendiceal pus of patients with acute appendicitis, whole-genome average nucleotide identity (ANI) analysis identified all of them as *Proteus* genomosp. 6, different from that identified using the automated bacterial identification instrument (VITEK®-32). Based on ANI and the core-genomic phylogenetic tree, we found that 87.5% of clinically-related strains previously identified as *P. columbae* should be re-classified as *Proteus* genomosp. 6. Additionally, the *Proteus* genomosp. 6 genomes all carry one or more beta-lactam resistance genes, but none carry aminoglycoside resistance genes, and antibiotic susceptibility testing conducted on the four strains isolated in this study confirmed these findings. Among the genomes analyzed, only four (two from this study (TSJ240517 and WDL240414)) carried virulence genes, specifically the *hlyA*, *hlyB*, and *hlyD* genes encoding hemolysin.

Conclusion Our study highlights inaccuracies in the taxa classification of *Proteus* species under clinical settings, underscoring the necessity of using genomic-based taxonomic assignment methods. We revealed that the prevalence of *Proteus* genomosp. 6 in clinical infections has likely been underestimated. Furthermore, given the resistance-gene absence and their sensitivity to aminoglycosides, aminoglycosides may serve as a promising first-line treatment option for infections caused by this species.

Keywords *Proteus* genomosp. 6, *Proteus columbae*, Acute appendicitis, Antibiotic resistance, Virulence factor

*Correspondence:
Yao Peng
pengy2021@szu.edu.cn
Xunchao Cai
cxch1224@126.com

Full list of author information is available at the end of the article

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Background

The genus *Proteus*, belonging to the Morganellaceae family, is composed of Gram-negative rods known for their adaptability and metabolic flexibility 1, 2. Among them, *Proteus mirabilis*, *Proteus vulgaris*, and *Proteus penneri* are the most frequently identified species in human infections 3. *P. mirabilis* is the most frequent cause of human disease, recognized as causes of urinary tract infections and serious hospital-onset invasive disease 4, 5; *P. vulgaris* occurs less frequently, it is associated with complicated infections and has potential for treatment failure as a result of inducible beta-lactamases 4; *P. penneri* has rarely been reported in clinical human studies and phenotypically close to *P. mirabilis* 4, 6. While the improved culture-independent detection methods (e.g., high-throughput 16 S rRNA sequencing, shotgun sequencing) has enhanced detection sensitivity for low-abundance populations (e.g., *Proteus*) in gut microbiome, which implicates *Proteus* spp. in the pathophysiology of various gastrointestinal diseases, including gastroenteritis, appendicitis, and Crohn's disease 7, 8, 9. Their implications in gastrointestinal diseases have been reported both correlative and causitive 5, 10, thereby challenging the prevailing notion of their innocuous bystander role within the gut microbiome 11.

Currently, 13 taxa within the *Proteus* genus are validly published under the International Code of Nomenclature of Prokaryotes (ICNP) rules 12. Additionally, three genomospecies (*Proteus* genomosp. 4, 5, and 6) remain unclassified due to their high phenotypic similarity to other close *Proteus* species (e.g., *P. vulgaris*, *Proteus hauseri*, *P. penneri*) 3. Of note, these three genomospecies were found asociated with human infections, whereas their pathogenicity has not yet been determined 3. For instance, among the 11 reported *Proteus* genomosp. 6 isolates, seven were recovered from human urine, two from wounds, one from stool, and one from unknown human sources 3. We previously isolated a *Proteus* strain from the appendiceal pus of a patient with appendicitis and identified it as a novel species, *Proteus appendicitidis*, based on both genomic and phenotypic evidence, suggesting a potential association between *Proteus* species and acute appendicitis 12. Therefore, further investigation of *Proteus* species derived from the diseased appendices may serve to substantiate our previous findings and enhance our understanding of the interplay between *Proteus* species and human health.

Here, we isolated four *Proteus* strains from patients diagnosed with acute appendicitis. We subsequently assembled their whole-genome sequences using next-generation sequencing (NGS) technology, refined their taxonomic classification through genome-based phylogenetic analysis, and performed comparative genomics to identify genes associated with virulence.

Methods

Specimen collection and bacteria isolation

Four patients were diagnosed with acute appendicitis with peritonitis or abscess, and were accompanied by microbial infection based on clinical blood test diagnosis, which revealed elevated levels of C-reactive protein, increased white blood cell counts, and a heightened neutrophil percentage. During the appendectomy or ultrasound-guided drainage, appendiceal pus was collected from the patients for the isolation of the causative microorganisms. Briefly, appendiceal pus samples were collected and immediately transferred to the department of clinical laboratory for microbial culture. Subsequently, 100 µL aliquots of the pus from each sample were evenly spread onto blood agar plates or Luria–Bertani Agar (1.5%) (LBA) plates and incubated separately under anaerobic and aerobic conditions at 37 °C for 48 h. Colonies that appeared on the plates were subsequently re-streaked to ensure the acquisition of pure cultures.

Bacterial identification and antimicrobial susceptibility testing

The obtained pure cultures were subjected to taxonomic identification using an automated bacterial identification instrument (VITEK®-32; bioMérieux, Marcy-l'Étoile, France) as per the manufacturer's instructions. One single *Proteus* strain was exclusively isolated from each sample, with no other aerobes or strict anaerobes recovered (see Table S1 for the detailed identification outcomes). Antimicrobial susceptibility testing was conducted by employing the VITEK®-2 v9.02.3 automated system with the AST-GN13 card and AST-N33 card. The minimum inhibitory concentration (MIC) and antimicrobial susceptibility of the isolates on the tested antimicrobials were determined as per the Clinical and Laboratory Standards Institute guidelines (CLSI 2023).

Genome sequencing, assembly, and annotation

Bacterial cells were harvested from LBA plates by gentle washing with 1 mL of sterile phosphate-buffered saline (PBS). Bacteria cells were then pelleted by centrifugation at 10 000 × g for 5 min for DNA extraction. Genomic DNA was promptly extracted using the TaKaRa MiniBEST Bacteria Genomic DNA Extraction Kit (Takara Bio Inc., Shiga, Japan), strictly adhering to the manufacturer's protocol. The quality of the extracted DNA was assessed employing the Synergy HTX Multi-Mode Reader (BioTek, USA). Whole-genome sequencing was conducted on the Illumina NovaSeq™ (Illumina, Inc., CA, USA) platform. The sequencing library was prepared with a 350 bp insert size and sequenced using the pair-end strategy (PE150). Quality control of the raw sequencing reads was meticulously performed using fastp v0.20.1 [13], which involved adapter trimming and

the removal of low-quality reads, ensuring that over 95% of the reads had a quality score higher than Q20 (Phred-score). Genome assembly was performed using Unicycler v0.4.9b 14, and contig sequences shorter than 200 bp (required by the NCBI genome database in the submission process) were removed from the assembled genomes.

Genome-based taxonomy classification and virulence prediction

The GTDB-Tk software, specifically employing the gtdbtk_wf workflow, was utilized to conduct a rigorous validation of the taxonomic assignments for the isolated strains 15. To identify genomes with high sequence similarity to those of the four sequenced strains, the Similar Genome Finder tool integrated within the PATRIC v3.6.10 web-server was employed. The 50 top similar genome sequences, along with their associated metadata, were retrieved from the NCBI genome database. Subsequently, the whole-genome average nucleotide identity (ANI) between the assembled genomes and these similar genomes was calculated using fastANI 16. Furthermore, the core genome sequences of these genomes were extracted and aligned using the Roary (v3.13.0) pan genome pipeline, configured with the parameters "-e -v -n", which invokes MAFFT for multiFASTA alignment (a 95% minimum sequence identity were used to cluster orthologs) 17. A maximum-likelihood core-genomic phylogenetic tree was then constructed, employing the generalized time-reverse model with FastTree (v2.1.11) utilizing the parameter "-gtr" 18. The virulence coding gene and antimicrobial resistance gene carried by these genomes were predicted using the ABRicate v1.00 (https://github.com/tseemann/abricate), by searching against the pre-calculated NCBI antimicrobial resistance (AMR) database (AMRFinderPlus database) and the "vfdb" virulence database 19.

Quality assurance

Single colonies of the four strains were subjected to three successive re-streaking on LB agar plates t to ensure the isolation of pure colonies. For shotgun whole-genome sequencing, low-quality reads were removed as part of quality control. Moreover, the completeness and contamination of the assembled genome sequences were meticulously evaluated using CheckM v1.0.12 [20]. CheckM v1.0.12 analysis revealed that all assembled genomes had 100% completeness and 0% contamination.

Results and discussion

Clinical features of patients

Of the four patients diagnosed with acute appendicitis, three underwent appendectomy (laparoscopic appendectomy) and one underwent ultrasound-guided drainage (Table 1). Bacterial cultures were recovered from each of the postoperative specimens and were identified as *Proteus* spp. (Table S1). Notably, three patients (Case 1, 3, 4) exhibited severe appendicitis with complications such as peritonitis, abscess formation, or chronic inflammation. One (Case 2) presented with delayed treatment due to a psychiatric comorbidity (schizophrenia), leading to advanced complications (perforation, abscess, and intestinal obstruction) (Table 1). These findings underscore the relevance of *Proteus* spp. in complicated appendicitis and suggest its potential role in secondary infections. Further genomic and phenotypic analyses of the isolates would be conducted in the future to elucidate their pathogenicity and resistance mechanisms.

General genome features

NGS of the four strains yielded 8,869,464 to 9,831,412 raw reads. After quality control and adaptor removal, 8,835,548 to 9,774,520 clean reads were obtained (Table S2). Subsequently, draft genomes were generated for the four strains through genomic assembly. The number of contigs in the draft genomes was as follows: DFP240708, 59 contigs; TSJ240517, 42 contigs; LHD240705, 63 contigs; and WDL240414, 56 contigs. The genome sizes ranged from 3.85 to 4.11 Mbp. The GC contents were 38.00%, 38.14%, 38.02%, and 38.20%, respectively. The number of coding DNA sequences (CDSs) encoded were 3,482, 3,656, 3,536, and 3,783, respectively. Notably, strain WDL240414 had the largest genome and the highest number of CDSs (Table S2). GTDBtk analysis revealed that all four strains belongto *Proteus* genomosp. 6 (ANI cutoff used for species delineation is 95%)rather

Table 1 Clinical features of patients with acute appendicitis

Strains	Patients	Host Age/Sex	Appendicitis Type	Complications	Key Inflammatory Markers	Surgery Performed
WDL240414	Case 1	52/F	Suppurative with peritonitis	Intestinal adhesions	WBC: 15.06×10^9/L, CRP: 49.9 mg/L	Laparoscopic appendectomy + adhesiolysis
DFP240708	Case 2	33/M	Abscess with obstruction	Pneumonia, schizophrenia	WBC: 21.05×10^9/L, CRP: 99.9 mg/L	Ultrasound-guided drainage
LHD240705	Case 3	15/M	Chronic abscess (delayed surgery)	Intestinal adhesions	WBC: Normal, CRP: 12.2 mg/L	Laparoscopic appendectomy + adhesiolysis
TSJ240517	Case 4	10/F	Acute with localized peritonitis	Rheumatic comorbidity	WBC: 16.88×10^9/L, ASO: 541.7 IU/mL	Laparoscopic appendectomy

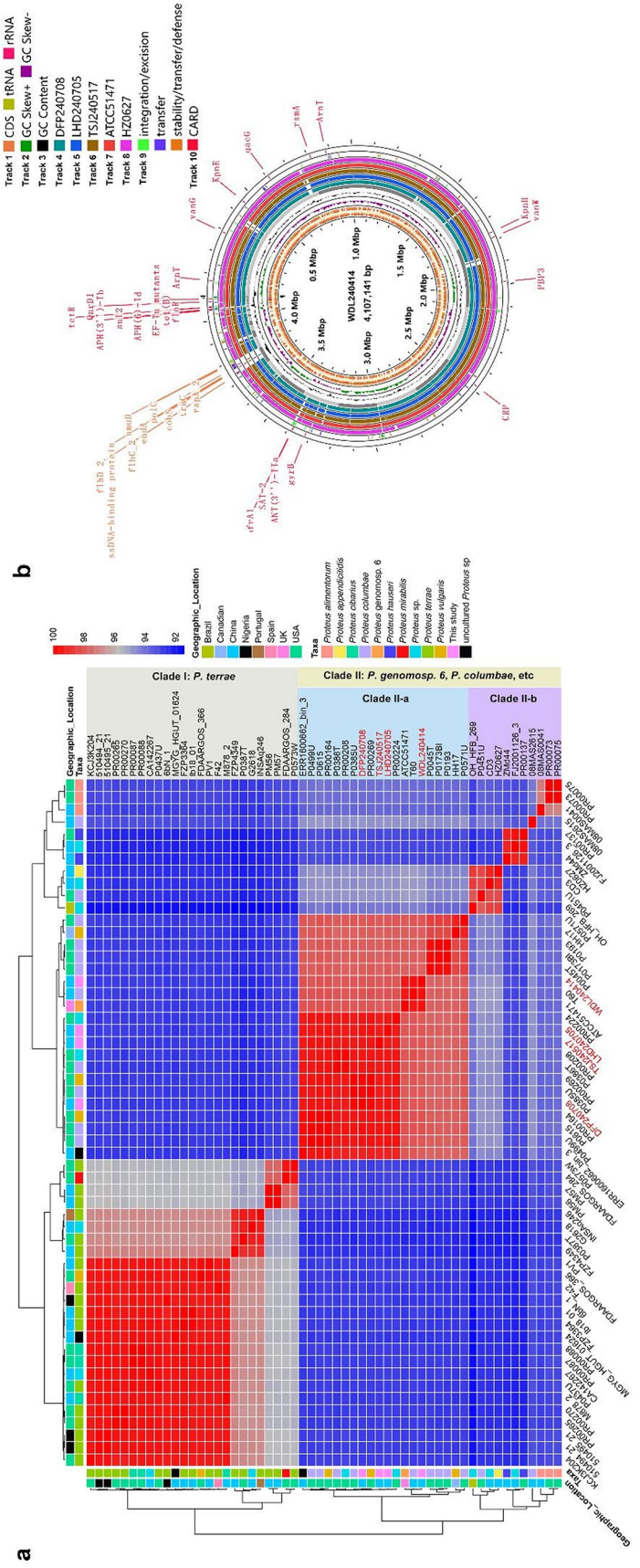

Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 Genome-based characteristics of the top 50 *Proteus* spp. strains close to the isolated *Proteus* strains. (**A**) The whole-genome-based average nucleotide identity (ANI) of the analyzed *Proteus* spp. strains. (**B**) Circular genome sequence view of the four isolated *Proteus* strains (referenced by strain WDL240414), the type strain (ATCC51471) of *Proteus* genomosp. 6, and the type strain (HZ0627) of *Proteus appendicitidis*. From inner to outer tracks, CDSs on the reverse strand, CDSs on the forward strand, GC skew, GC content, DFP240708, LHD240705, TSJ240517, ATCC52471, HZ0627, mobile genetic elements, CARD antibiotic resistance genes. The RNA sequences are displayed in the CDSs track. The circular view was constructed using the Proksee webserver (https://proksee.ca/)

than *Proteus vulgaris* or *Proteus hauseri*, as identified by the VITEK®-32 system (Table S1). The misidentification by the VITEK®-32 system is primarily due to the high phenotypic similarity among certain *Proteus* species, such as nearly indistinguishable biochemical profiles (e.g., urease activity, carbohydrate fermentation patterns, and swarming behavior) between phylogenetically-close *Proteus* spp. 3. These overlapping traits challenge phenotype-based identification systems that rely on conventional biochemical assays, emphasizing the necessity of genomic methods to resolve such taxonomic ambiguities within the *Proteus* genus.

Taxonomy assignment and characterization

Using the Similar Genome Finder tool integrated in the PATRIC web-server (https://www.bv-brc.org/app/GenomeDistance), we identified the 50 closest genomes in public databases to the four strains isolated in our research. After removing duplicates, a total of 52 strains were included for comparative genomic analysis with our strains, with the majority isolated from China (22 strains) and the United States (26 strains) (Table S3). ANI comparisons grouped these genomes into two major clades. Clade I predominantly consisted of *P. terrae*, including its two subspecies (i.e., *P. terrae* sub. *terrae* and *P. terrae* sub. *cibarius*). Clade II could be further divided into two sub-clades, namely Clade II-a and Clade II-b (Fig. 1a). Clade II-a was mainly constituted by *Proteus* genomosp. 6, while clade II-b encompassed *P. columbae*, *P. appendicitidis*, *P. hauseri*, and *P. alimentorum* (Fig. 1a). The four strains isolated in our study clustered within clade II-a, belonging to *Proteus* genomosp. 6. ANI values indicated that their closest species were *P. columbae* and *P. appendicitidis*. Notably, based on the ANI threshold for species delineation, a high rate of misidentification among *P. columbae* isolates was revealed. Of the 10 *P. columbae* strains analyzed (including eight clinically-related isolates), eight were reassigned to *Proteus* genomosp. 6, forming a distinct clade (Clade II-a; Fig. 1a). Critically, seven out of the eight clinically-related *P. columbae* strains (87.5%) were reclassified as *Proteus* genomosp. 6. This finding underscores that the prevalence of *Proteus* genomosp. 6 in clinical settings may have been significantly underestimated, highlighting the necessity to use genomic-based taxonomic approaches in clinical microbiology. This finding was further supported by the core-genomic phylogenetic tree (Figure S1). As a genomospecies rather than a formally named species, its phenotypic characteristics are typically excluded from comparative analyses during the nomenclature and classification of novel *Proteus* species, which has resulted in incomplete resolution of phenotypic distinctions between *Proteus* genomosp. 6 and recently described species such as *P. columbae* (first identified in 2018) [1]. In this study, only the type strain 08MAS2615^{T} was confirmed as the authentic *P. columbae* (Fig. 1a). By performing a literature review [1, 2], we found that the phenotypic profiles of *P. columbae* 08MAS2615^{T} and *Proteus* genomosp. 6 (11 strains) share numerous metabolic traits. For instance, both *P. columbae* 08MAS2615^{T} and all *Proteus* genomosp. 6 strains exhibited positive reactions for indole production, sucrose utilization, and maltose utilization, while being negative for L-rhamnose utilization. However, *P. columbae* 08MAS2615^{T} was negative for salicin and aesculin ferric citrate utilization, whereas 9% of *Proteus* genomosp. 6 strains exhibited positive reactions for these substrates. Additionally, *P. columbae* 08MAS2615^{T} was positive for methyl α-D-glucopyranoside and trehalose utilization, whereas only 10% and 20% of *Proteus* genomosp. 6 strains exhibited positive reactions for these substrates, respectively. Furthermore, the four *Proteus* genomosp. 6 strains isolated in this study were positive for sucrose and maltose utilization but negative for trehalose utilization (see Table S4). These findings underscore the phenotypic similarity between *P. columbae* and *Proteus* genomosp. 6, as well as the variability in substrate utilization among different *Proteus* genomosp. 6 strains, thereby complicating their differentiation based solely on phenotypic assays.

Using the genome of strain WDL240414 as a reference, a circle genome map comparison with the genomes of the four strains isolated in our study, as well as *Proteus* genomosp. 6 ATCC51471 and *P. appendicitidis* HZ0627, showed extremely high genomic sequence similarity. These strains encode multiple drug resistance genes, including *arnT* (peptide antibiotic resistance), *vanG* (glycopeptide resistance), *qacG* (disinfectant resistance), and *kpnE* (multidrug resistance) (Fig. 1b). Additionally, the genome of strain WDL240414 also possessed some unique gene sequences, such as *dfrA1*, *SAT-2*, *umuD*, and *endA*, encoding diaminopyrimidine antibiotic resistance, nucleoside antibiotic resistance, translation error-prone DNA polymerase V and transcriptional activator, respectively. These unique sequences are likely associated with

genomic mobile elements (e.g., integration, excision, and transfer), suggesting that they may have been acquired through horizontal gene transfer events (Fig. 1b).

In Silico identification of AMR- and virulence-coding genes

Among the 56 analyzed genomes, 20 classes of antibiotic resistance genes (annotated via the AMRFinderPlus database) were predicted, conferring resistance to 12 antibiotics (Table S5). Of these, 12 classes were identified in *Proteus* genomosp. 6 genomes, encoding resistance to quinolones, beta-lactams, tetracyclines, streptomycin, streptothricin, chloramphenicol, trimethoprim, kanamycin, sulfonamides, florfenicol, gentamicin, and hygromycin. Conversely, eight classes were absent, including those encoding resistance to rifamycins, aminoglycosides (e.g., amikacin/kanamycin/ tobramycin), bleomycin, carbapenems, cephalosporins, and macrolides (Table S5). Except *Proteus alimentorum* 08MAS0041, all *Proteus* strains analyzed in this study carried beta-lactam antibiotic resistance genes (Fig. 2). Clinical antibiotic resistance testing of the four strains in this study revealed variable resistance profiles to beta-lactam, cephalosporin, quinolone, and sulfonamide antibiotics, while uniform sensitivity to aminoglycoside antibiotics was observed (Table S5). For instance, all four strains exhibited resistance to cefuroxime but were sensitive to ceftazidime, cefepime, gentamicin, tobramycin and amikacin. Strain LHD240705 showed resistance to cefotaxime, whereas the other three strains were sensitive to cefotaxime (Table S5). All four isolated strains exhibited resistance to at least one beta-lactam or cephalosporin, likely due to the presence of the class A beta-lactamase gene *hugA*. (Table S5, Fig. 2). The HugA is known to hydrolyze cefuroxime, leading to resistance against this antibiotic 21. This has been observed in clinical isolates, such as *Proteus penneri*, where the *hugA* gene and its regulator gene *hugR* are present, and regulatory changes may lead to increased production of the enzyme, confering resistance to cefuroxime but not ceftazidime, cefepime and imipenem 21. Morever, their susceptibility to the tested aminoglycosides aligns with their antibiotic resistance gene profiles on aminoglycosides (Table S5, Fig. 2). A 2019 systematic review supports aminoglycosides as a clinically effective first-line treatment for *Proteus* infections. For instance, among 33 *Proteus* spp. isolates from urinary tract infections (UTIs) analyzed in that study, only two exhibited resistance to aminoglycosides 22. As such, in the clinical management of infections caused by these bacteria, the use of beta-lactam and cephalosporin antibiotics should be avoided when possible, while aminoglycoside antibiotics, such as amikacin, may serve as the preferred therapeutic agents. Moreover, the genome of this strain was found to carry the highest number of resistance genes among the *Proteus* genomosp. 6 strains, including genes conferring resistance to beta-lactams, quinolones, trimethoprim, streptomycin, streptothricin, tetracyclines, sulfonamides, and chloramphenicol/florfenicol (Fig. 2). Despite the absence of known carbapenem resistance genes, strain WDL240414 exhibited resistance to imipenem and multidrug resistance, including cephalosporins (cefuroxime), quinolones (levofloxacin, ciprofloxacin), sulfonamides (trimethoprim), beta-lactams (ampicillin, cefazolin), and nitrofurans (nitrofurantoin). The imipenem resistance of strain WDL240414 may indicate the presence of porin mutations or efflux-mediated resistance, requiring further investigation. Its resistance to levofloxacin, ciprofloxacin, trimethoprim, and sulfamethoxazole aligns with their antibiotic resistance gene profiles (Table S5). A Kruskal-Wallis non-parametric test was conducted to assess differences in AMR gene profiles across six *Proteus* species ($p = 5.68 \times 10^{-5}$), revealing statistically significant differences among them. Pairwise comparisons revealed that no significant differences between *P. terrae* and *Proteus* genomosp. 6, wheras *Proteus* genomosp. 6 harbored 38.85%, 47.51%, 73.75%, and 73.75% more AMR genes than *P. hauseri*, *P. columbae*, *P. appendicitidis* and *P. alimentorum*, respectively ($p < 0.05$). Notably, *Proteus* genomosp. 6 exhibited a significantly higher AMR gene burden compared to its phylogenetically close *Proteus* species. This genomic divergence in AMR coding capacity may reflect adaptive evolution under selective pressures (e.g., antibiotic exposure), highlighting the need for functional studies to validate genotype-phenotype linkages and assess clinical implications of these resistance determinants, highlighting the imperative for accurate taxa identification and monitoring of its antibiotic resistance in clinical settings.

Among the 56 genomes, only four strains carried virulence genes (*hlyA*, *hlyB*, *hlyD*) encoding hemolysin (Table S6), a toxin known to damage host cell membranes. Hemolysin has been identified in various pathogenic Gram-negative bacteria, including *Proteus* spp., *Morganella* spp., and *Escherichia coli*23, 24. Two of these four strains, namely TSJ240517 and WDL240414, were isolated and identified in this study, carrying the complete *hlyABD* gene cluster (Table S6), hemolysin may thus represent one of their virulence factors. Hemolysins (e.g., HlyA, HlyB, HlyD) in *Proteus* spp. are pore-forming toxins that lyse erythrocytes and other host cells, facilitating iron acquisition and tissue invasion 10. In *Proteus* genomosp. 6, hemolysin production may enhance pathogenicity by damaging epithelial barriers, triggering inflammation, and promoting bacterial dissemination. While direct evidence linking hemolysins to appendicitis specifically is limited, studies in related *P. mirabilis* suggest hemolysins contribute to virulence in urinary tract and wound infections 10. The absence of these genes in some strains could reflect niche-specific adaptation or

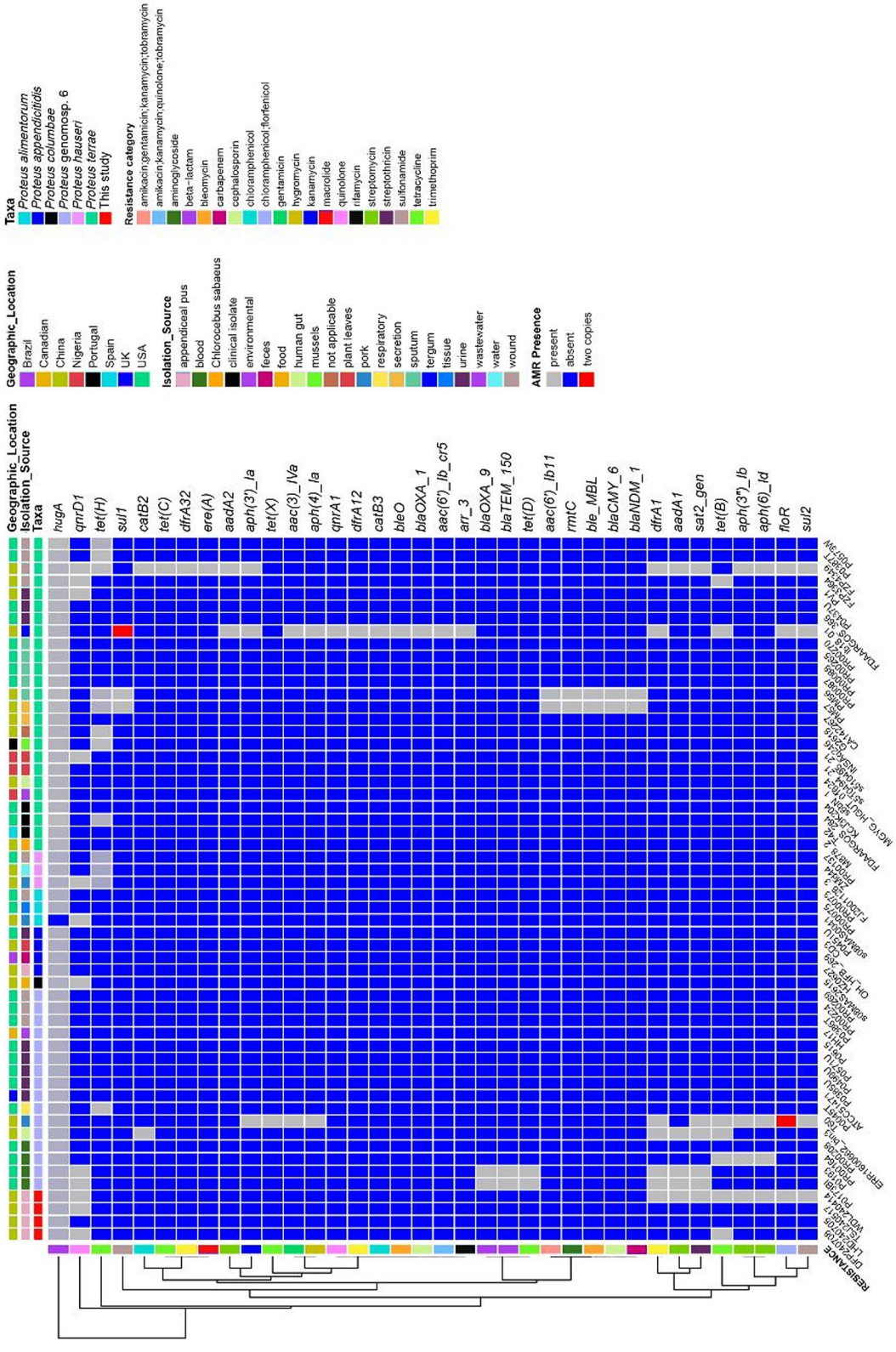

Fig. 2 AMR gene profile of the *Proteus* spp. phylogenetically close to *Proteus* genomosp. 6. AMR gene classes are from the pre-calculated AMRFinderPlus database, certain resistance gene classes may confer resistance to multiple antibiotic types (e.g., the class "amikacin/kanamycin/quinolone /tobramycin" encodes resistance to aminoglycosides [e.g., amikacin, kanamycin, tobramycin] and quinolones) (See Table S5 for details.)

attenuated virulence. In appendicitis, hemolysin-positive strains might exacerbate mucosal damage or secondary infections, though further mechanistic studies are needed to confirm their role.

Conclusions

Our comprehensive genomic analysis of four *Proteus* strains isolated from patients with acute appendicitis has provided novel insights into the taxonomy and genomic characteristics of *Proteus* genomosp. 6. The prevalence of *Proteus* genomosp. 6 in clinical infections has been likely underestimated, as evidenced by the reclassification of numerous strains previously identified as *P. columbae*. This species may play a more prominent role in gastrointestinal infections, such as acute appendicitis, than previously recognized. Antibiotic resistance gene prediction and antibiotic resistance testing suggests that aminoglycosides may be a more effective first-line treatment option for infections caused by this species. Moreover, virulence genes prediction revealed that hemolysin may be a key virulence factor for *Proteus* genomosp. 6. Our findings highlight the need to reconsider treatment strategies for *Proteus* genomosp. 6 infections, suggesting aminoglycosides as a preferred therapeutic option. Additionally, the generalizability of these findings should be interpreted cautiously due to the limited sample size, which may constrain extrapolation to broader patterns of genomic diversity, antimicrobial susceptibility or clinical heterogeneity within this emerging *Proteus* lineage. Future research should focus on further characterizing the virulence factors and antibiotic resistance mechanisms of *Proteus* genomosp. 6 to better understand its pathogenic potential and inform clinical practices.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13099-025-00701-8.

Supplementary Material 1

Supplementary Material 2

Acknowledgements
Not applicable.

Author contributions
XCC, JH, and YP contributed to genome analysis and manuscript preparation. ZYZ and XQL contributed to the bacterial isolation, data collection, and manuscript preparation. LQH and ZZ contributed to data collection and manuscript editing. XCC and YP contributed to the conception and design of the study. All authors read and approved the final manuscript.

Funding
This research was supported by the Shenzhen Science and Technology Program (grant number JCYJ20230808105208017), the Yongzhou City Instructive Science and Technology Plan Project (grant number 2023YZ091, 2024YZ062).

Data availability
The complete genome data of Proteus sp. DFP240708, LHD240705, TSJ240517 and WDL240414 has been deposited in GenBank, with the BioProject Accession Number PRJNA1222607.

Declarations

Ethics approval and consent to participate
The study was approved by the Ethics Committee of the Central Hospital of Yongzhou (June 23, 2022; IRB No. 2022062301) and was compliant with all relevant ethical regulations.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Clinical Laboratory, The Central Hospital of Yongzhou, Yongzhou 425000, China
[2]Department of Clinical Laboratory, Yongzhou Hospital, University of South China, Yongzhou 425000, China
[3]Department of Pharmacy, The Central Hospital of Yongzhou, Yongzhou 425000, China
[4]Department of Pharmacy, Yongzhou Hospital, University of South China, Yongzhou 425000, China
[5]College of Life Science, Northeast Agricultural University, Harbin 150030, China
[6]School of Pharmacy, Hunan University of Chinese Medicine, Changsha 410208, China
[7]Department of Gastroenterology and Hepatology, Shenzhen University General Hospital, Shenzhen University, Xueyuan R.D. 1298, Nanshan District, Shenzhen 518071, China

Received: 12 February 2025 / Accepted: 15 April 2025
Published online: 15 May 2025

References

1. Garrity G. Bergey's manual® of systematic bacteriology: volume 2: the Proteobacteria, part B: the Gammaproteobacteria. Springer Science & Business Media. 2007.
2. Manos J, Belas R. The genera proteus, Providencia, and morganella. Prokaryotes. 2006;6:245–69.
3. O'Hara CM, Brenner FW, Steigerwalt AG, 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. 2000;50:1869–75.
4. Laupland KB, Edwards F, Harris PN. Proteus species bloodstream infections: comparative epidemiology of three species. Diagn Microbiol Infect Dis. 2024;109:116286.
5. Ioannou P, Vougiouklakis G. Infective endocarditis by Proteus species: a systematic review. Germs. 2020;10:229.
6. Pandey A, Verma H, Asthana AK, et al. Extended spectrum beta lactamase producing Proteus Penneri: a rare missed pathogen? Indian J Pathol Microbiol. 2014;57:489–91.
7. Zilberstein B, Quintanilha AG, Santos MA, et al. Digestive tract microbiota in healthy volunteers. Clinics. 2007;62:47–56.
8. Mukhopadhya I, Hansen R, El-Omar EM, et al. IBD—what role do Proteobacteria play? Nat Reviews Gastroenterol Hepatol. 2012;9:219–30.
9. Li N, Cai Q, Miao Q, et al. High-throughput metagenomics for identification of pathogens in the clinical settings. Small Methods. 2021;5:2000792.
10. Hamilton AL, Kamm MA, Ng SC, et al. Proteus spp. As putative Gastrointestinal pathogens. Clin Microbiol Rev. 2018;31. https://doi.org/10.1128/cmr.00085–00017.
11. Yatsunenko T, Rey FE, Manary MJ et al. Human gut microbiome viewed across age and geography. nature. 2012;486:222–227.

12. He J, Zhang Z, Qu H, et al. Proteus appendicitidis Sp. nov., isolated from the appendiceal pus of an appendicitis patient in Yongzhou, China. Arch Microbiol. 2024;206:171.
13. Chen S, Zhou Y, Chen Y, et al. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:i884–90.
14. Wick RR, Judd LM, Gorrie CL, et al. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017;13:e1005595.
15. Parks DH, Chuvochina M, Waite DW, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol. 2018;36:996–1004.
16. Jain C, Rodriguez-R LM, Phillippy AM, et al. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun. 2018;9:5114.
17. Page AJ, Cummins CA, Hunt M, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31:3691–3.
18. Price MN, Dehal PS, Arkin AP. FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS ONE. 2010;5:e9490.
19. Zhou S, Liu B, Zheng D, et al. VFDB 2025: an integrated resource for exploring anti-virulence compounds. Nucleic Acids Res. 2025;53:D871–7.
20. Parks DH, Imelfort M, Skennerton CT, et al. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25:1043–55.
21. Liassine N, Madec S, Ninet B, et al. Postneurosurgical meningitis due to Proteus penneri with selection of a ceftriaxone-resistant isolate: analysis of chromosomal class A β-lactamase HugA and its LysR-type regulatory protein HugR. Antimicrob Agents Chemother. 2002;46:216–9.
22. Goodlet KJ, Benhalima FZ, Nailor MD. A systematic review of single-dose aminoglycoside therapy for urinary tract infection: is it time to resurrect an old strategy? Antimicrob Agents Chemother. 2019;63:02165–02118. https://doi.org/10.1128/aac
23. Minnullina L, Misheeva P, Mukhtarova G, et al. Features of hemolysin biosynthesis by Morganella morganii. Bull Exp Biol Med. 2023;176:181–6.
24. Zhao H, Lee J, Chen J. The hemolysin A secretion system is a multi-engine pump containing three ABC transporters. Cell. 2022;185:3329–40.e3313.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.