Identification and Functional Annotation of Hypothetical Proteins of Pan-Drug-Resistant *Providencia rettgeri* Strain MRSN845308 Toward Designing Antimicrobial Drug Targets

Dipta Chandra Pal (iD), Tasnimul Arabi Anik, Atiq Abrar Rahman and S M Mahfujur Rahman

Department of Microbiology, University of Dhaka, Dhaka, Bangladesh.

ABSTRACT: *Providencia rettgeri* has increasingly been responsible for several infections, including urinary tract, post-burn wounds, neonatal sepsis, and others. The emergence of drug-resistant isolates of *P rettgeri*, accompanied by intrinsic and acquired antibiotic resistance, has exacerbated the challenge of treating such infections, necessitating the development of novel therapeutics. Hypothetical proteins (HPs) form a major portion of cellular proteins and can be targeted by these novel therapeutics. In this study, 410 HPs from a pan-drug-resistant (PDR) *P rettgeri* strain (MRSN845308) were functionally annotated and characterized by physicochemical properties, localization, virulence, essentiality, druggability, and functionality. Among 410 HPs, the VirulentPred 2.0 tool and VICMpred combinedly predicted 33 HPs as virulent, whereas 48 HPs were highly interacting proteins based on the STRING v12 database. BlastKOALA and eggNOG-mapper v2.1.12 predicted 13 HPs involved in several metabolic pathways like Riboflavin metabolism and Lipopolysaccharide biosynthesis. Overall, 83 HPs were selected as primary drug targets; however, only 80 remained after nonhomology searching and essentiality analysis. In addition, all were detected as novel drug targets according to DrugBank 5.1.12. Considering the potential of membrane and extracellular proteins, 29 HPs (extracellular, outer, and inner membrane) were selected based on the combined prediction from PSORTb v3.0.3, CELLO v.2.5, BUSCA, SOSUIGramN, and PSLpred. According to the prevalence of those HPs in different strains of *P rettgeri* sequences in National Center for Biotechnology Information Identical Protein Groups (NCBI-IPG), 5 HPs were selected as final drug targets. In addition, 5 other HPs annotated as transporter proteins were also added to the list. As no crystal structures of our targets are present, 3-dimensional structures of selected HPs were predicted by the AlphaFold Server powered by AlphaFold 3. Our findings might facilitate a better understanding of the mechanism of virulence and pathogenesis, and up-to-date annotations can make uncharacterized HPs easy to identify as targets for novel therapeutics.

KEYWORDS: Pan-drug-resistant, *Providencia rettgeri*, hypothetical protein, HPs, drug targets, therapeutics

Introduction

Providencia rettgeri is a Gram-negative opportunistic bacterium belonging to the genus "*Providencia.*" It is now classified under the *Morganellacea*e family, although it previously belonged to the *Enterobacteriaceae* family. Like some members of the *Enterobacteriaceae* family, such as *Klebsiella pneumoniae*, *Acenetobacter baumannii*, and *Serratia marcescens*, *P rettgeri* is commonly involved in hospital-acquired infections.[1,2] *Providencia rettgeri* mainly causes travelers' diarrhea and catheter-related urinary tract infections.[3,4] Recently, it has also been reported in several other infections, including post-burn wound infections, meningitis, and neonatal sepsis.[5-7] The increasing involvement in different types of infections is particularly a problem to public health as drug-resistance is nothing unfamiliar to *P rettgeri*. It has intrinsic resistance to colistin and tigecycline as well as first-generation cephalosporins making it insensitive to many antibiotic treatments.[8,9]

Other than intrinsic resistance, the acquisition of antibiotic resistance has also been well-documented in *P rettgeri*. In 2000, carbapenem-resistant *P rettgeri* was first identified in western Japan. Since then, various antibiotic-resistance genes carrying *P rettgeri* strains have been identified worldwide.[10] A *P rettgeri* isolate harboring the *blaNDM-1* gene was reported in Israel in 2013.[11] These isolates, carrying resistance genes, have been involved in several outbreaks in recent years. *Providencia rettgeri* strain co-producing New Delhi metallo-β-lactamase-1 (NDM-1) carbapenemase and Pseudomonas Extended Resistance-1 (PER-1) extended-spectrum β-lactamase (ESBL) was documented for causing nosocomial outbreaks during 2016 to 2017 in South Korea.[9] Similarly, in 2020, an outbreak mediated by multidrug-resistant (MDR) *P rettgeri* has come into concern in South-eastern Nigeria.[12]

The transfer of resistance through plasmids can lead to more frequent outbreaks of *P rettgeri* as plasmids carrying antibiotic-resistant genes are frequently detected in *P rettgeri*. The presence of plasmids carrying *blaIMP-11* or *blaIMP-70* was detected by sequencing *P rettgeri* isolates.[10] In Japan, unique plasmids carrying the *blaIMP-70* and *blaCTX-M-253* genes were found in MDR *P rettgeri* strains, as reported in 2022.[13] The availability of plasmid loaded with antibiotic-resistance genes elevates the risk of MDR *P rettgeri* turning into pan-drug-resistant (PDR) under persistent selective pressure resulting from the widespread use of antibiotics. Recently, a PDR *P rettgeri* MRSN845308 was noted to cause severe fatality in a
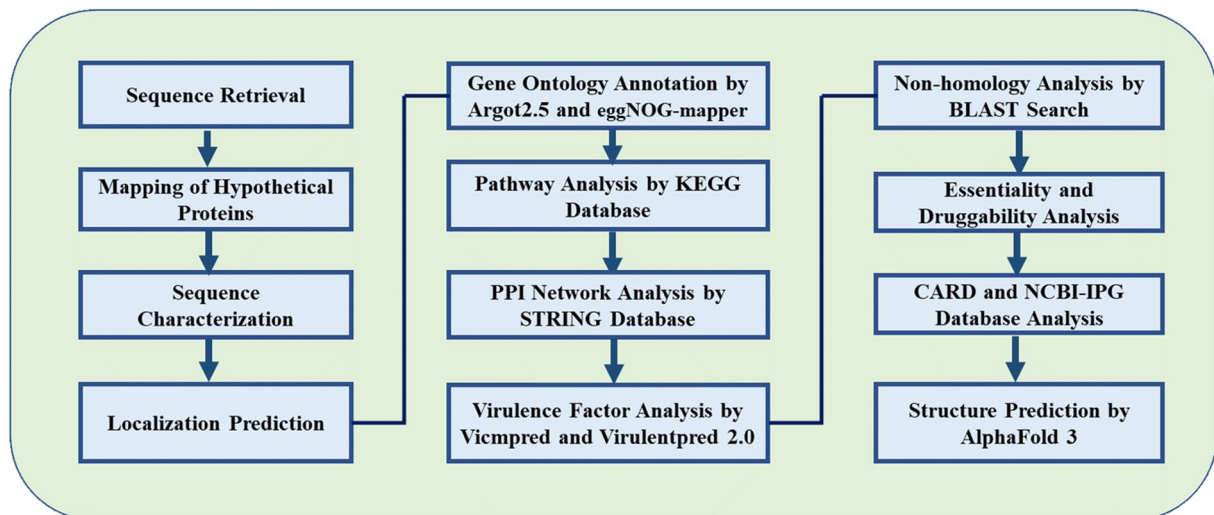
Figure 1. Flowchart of methodology.

COVID-19 patient.[8] The strain's extensive resistance to all classes of antibiotics resulted in multiorgan failure, ultimately leading to the patient's death on hospital day 66. If the emergence of such PDR *P rettgeri* is left unnoticed, the current trend of drug failure in the treatment of *P rettgeri* is unlikely to change.

Hence, a study was designed with the aim of identifying novel therapeutic targets for the treatment of PDR *P rettgeri* infections. New drugs that are effective against PDR isolates would likely be functional against MDR and Extensively Drug-Resistant (XDR) isolates. Therefore, the best practice in designing drugs or finding drug targets would be taking PDR isolates as target pathogens. Besides, most practices to unearth drug targets are based on analysis of defined or known proteins leaving a significant portion of targets unexplored. To our concern, no study thus far has explored the possibility that the hypothetical proteins (HPs) of *P rettgeri* can potentially be the targets of developing and already developed drugs. Considering the potential, we have explored the HPs of the recently sequenced PDR *P rettgeri* strain MRSN845308 for drug targets. In this study, we employed a range of computational algorithms to deduce the potential functions of HPs from the PDR *P rettgeri* strain MRSN845308. Our analysis includes in silico assessments of these HPs, focusing on both the molecular and structural levels, aiming to identify promising therapeutic targets for drug design and development. Through computational methodologies, we sought to elucidate the roles and characteristics of these proteins, paving the way for targeted approaches in combating antibiotic resistance and advancing antimicrobial strategies.

Materials and Methods

Sequence retrieval

The complete genome of *P rettgeri* strain MRSN845308 was retrieved on April 10, 2024, from the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm. nih.gov/) with accession number JAEEFV000000000.1. According to the Average Nucleotide Identity (ANI)-based NCBI taxonomy check, the strain is attributed to *Providencia hangzhouensis*. The sequences of HPs were mined using the keyword "hypothetical protein" by comparing with the reference (GCF_016512695.1) annotation (referenced on December 2, 2023). Their Uniprot accession IDs were retrieved using the Uniprot database. Proteins with more than 100 amino acids were considered for further analysis (Figure 1).

Physicochemical properties of hypothetical proteins

The physicochemical characteristics of all the HPs were investigated using Expasy's ProtParam server (https://web.expasy. org/protparam/).[14] This analysis facilitated the determination of key parameters, including molecular weight, isoelectric point,[15] instability index,[16] aliphatic index, and grand average of hydropathicity (GRAVY).[17] The GRAVY score, calculated as the sum of hydropathy values divided by the number of residues in the sequence, indicates the overall hydrophobicity of a peptide or protein.

Cellular and subcellular localization and protein classification

The topology of membrane proteins was predicted using TOPCONS, a web server designed for consensus prediction of membrane protein topology (https://topcons.cbr.su.se/ pred/).[18] In addition, CCTOP (https://cctop.ttk.hu),[19] TMHMM v.2.0 (https://services.healthtech.dtu.dk/services/ TMHMM-2.0/),[20] and the SignalP 5.0 server (https://services.healthtech.dtu.dk/services/SignalP-5.0/)[21] were used to identify the cellular locations of these proteins.

To determine the subcellular localization of the HPs, we employed PSLpred (https://webs.iiitd.edu.in/raghava/pslpred/

submit.html), which integrates Position-Specific Iterated BLAST (PSI-BLAST) and 3 SVM (Support Vector Machine) modules based on residue compositions, dipeptides, and physicochemical properties,[22] BUSCA (https://busca.biocomp. unibo.it),[23] CELLO v.2.5 (http://cello.life.nctu.edu.tw),[24] SOSUIGramN (https://harrier.nagahama-i-bio.ac.jp/sosui/ sosuigramn/sosuigramn_submit.html),[25] and PSORTb version 3.0.3 (https://www.psort.org/psortb/).[26]

Functional domain/motif prediction

Functional domain prediction aims to identify conserved segments within proteins, as these domains often represent functional units. In the case of the HPs, we used several publicly available databases for domain prediction, including ScanProsite (https://prosite.expasy.org/scanprosite/),[27] NCBI-CDD (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb. cgi),[28] InterProScan 5 (https://www.ebi.ac.uk/interpro/search/ sequence/),[29] and PFP-FunDSeqE (http://www.csbio.sjtu. edu.cn/bioinf/PFP-FunDSeqE/).[30]

Function prediction

Accurately annotating protein function is essential for comprehending molecular-level life processes. In our study, we predicted the Gene Ontology (GO) terms, specifically Molecular Function and Biological Process, associated with the HPs. With this aim, Argot2 (Annotation Retrieval of Gene Ontology Terms) (https://www.medcomp.medicina.unipd.it/ Argot2/)[31] and eggNOG-mapper v2.1.12 (http://eggnog-mapper.embl.de) were used.[32]

Pathway analysis of hypothetical proteins

The HPs were analyzed using the tool BlastKOALA (https:// www.kegg.jp/blastkoala/) from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database to predict the pathways regulated by these proteins.[33]

Interactional network analysis of hypothetical proteins

We conducted a protein interaction network analysis to identify HPs with significant protein-protein interactions (PPIs) using the STRING database version 12 (https://string-db.org).[34]

Virulence protein prediction

Identifying virulent proteins within bacterial sequences is valuable for assessing pathogenic potential and elucidating the intricate mechanisms of virulence in pathogenesis. To achieve the desired goal, we employed the VirulentPred 2.0 tool (https://bioinfo.icgeb.res.in/virulent2/index.html/) and VICMpred (https://webs.iiitd.edu.in/raghava/vicm-pred/index.html) to identify virulence factors among the HPs.[35,36]

Nonhomology analysis by Basic Local Alignment Search Tool for proteins human proteome

In our analysis, HPs from *P rettgeri* present in the metabolic pathways, proteins with higher PPIs, and proteins involved in virulence were compared with the proteome of *Homo sapiens* (human; taxid 9606) using BLASTp (Basic Local Alignment Search Tool for proteins) (https://blast.ncbi.nlm.nih.gov/Blast. cgi). It allows us to identify similarities and differences in protein sequences between the bacterial HPs and the human proteome. Proteins that showed BLAST hits with an expectation value (E-value) less than 0.001 were considered homologous to the human proteome. Conversely, proteins without significant BLAST hits (E-value ⩾ 0.001) were deemed specific to *P rettgeri*, suggesting potential species-specific functions or roles.[37]

Essentiality analysis

Essentiality analysis for the selected nonhomologous HPs was performed using BLASTp against a Database of Essential Genes (DEG; http://origin.tubic.org/deg/public/index.php/ blast/bacteria).[38]

Druggability analysis

We assessed the druggability of nonhomologous HPs by investigating their potential to interact with known drug candidates. Each nonhomologous HP underwent a homology search using the DrugBank 5.1.12 Target collection database (https:// go.drugbank.com/structures/search/bonds/sequence#results).[39]

Homologs of known antibiotic–resistance determinants

The Comprehensive Antibiotic Resistance Database (CARD) (https://card.mcmaster.ca/analyze/blast) was used to search for homologs of known antibiotic-resistance determinants of HPs.[40]

Abundance of hypothetical proteins

To identify and prioritize potential drug targets among HPs, we conducted an abundance analysis using the NCBI Identical Protein Groups database (IPG) (https://www.ncbi.nlm.nih.gov/ipg).[41]

Structural modeling of functionally important hypothetical proteins

In our study, protein structural prediction was conducted using the AlphaFold Server powered by AlphaFold 3 (https://alpha-foldserver.com).[42]

Results

Selection of hypothetical proteins

Using *P rettgeri* strain MRSN845308 from the NCBI Genome database, 573 proteins were found to be HPs containing specific
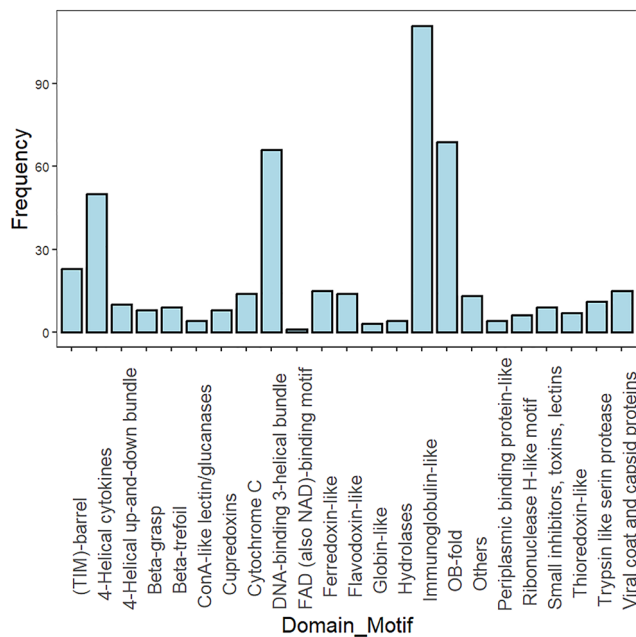
Figure 2. Domain/motif prediction by PFP-FunDSeqE.

Uniprot accession IDs. Among all the HPs, 410 proteins were found to be longer than 100 amino acids and were therefore considered suitable for further analysis.

Functional domain/motif and physicochemical characteristics prediction

Based on the NCBI-CDD database, we identified 111 proteins with conserved domains (Supplementary File 1.1). InterProScan shows 67 proteins have functional domains (Supplementary File 1.2). ScanProsite, which matches proteins against the Prosite collection of motifs, identified 44 functional regions, and others showed no hit. PFP-FunDSeqE predicted protein fold pattern with functional domain for all the proteins, as shown in Figure 2 (Supplementary File 2). Physicochemical characteristics showed that 183 proteins were found to be basic ($pI \geq 7$), whereas 227 proteins were found to be acidic ($pI < 7$). The GRAVY, amino acid composition, molecular weight, instability index, and aliphatic index were also calculated; the details are given in Supplementary File 3.

Cellular and subcellular localization of hypothetical proteins

Several bioinformatics tools were used to determine the presence of transmembrane helices. TOPCONS predicted 23 proteins with transmembrane helices. Specifically, the TMHMM algorithm, known for its accuracy in distinguishing between soluble and membrane proteins using a Hidden Markov Model, identified 108 HPs containing transmembrane helices. CCTOP identified 99 proteins with transmembrane segments,

and SignalP identified 82 HPs containing transmembrane residues (Supplementary File 4).

For subcellular localization, we employed PSORTb, CELLO v.2.5, BUSCA, SOSUIGramN, and PSLpred tools. By taking consensus from all, we have identified 281 sequences in the cytoplasm, 45 proteins in the extracellular space, 63 as inner membrane proteins, 6 as outer membrane proteins, and 15 predicted to be in the periplasmic space (Supplementary File 5).

Functional classification of hypothetical proteins

For functional annotation at the cellular, molecular, and biological levels, Argot2 was used. It predicted biological processes of 372 HPs (Figure 3), cellular functions of 366 HPs (Figure 4), and molecular functions of 406 HPs (Figure 5). More than 24 broad classes of molecular functions were identified at the molecular level. The most significant group of proteins was predicted to be involved in transferase activity, followed by DNA, ATP, Ion, other binding proteins, and hydrolase activity. In biological processes, 17 broad categories were identified, with most proteins primarily associated with DNA-associated metabolic processes, transportation, signal transduction, folding, biosynthesis, cell division, and stress resistance. With respect to cellular components, 15 broad functional categories were identified, as presented in Figure 3 (Supplementary File 6.1).

Pathway analysis of functionally characterized proteins

The functional roles of HPs in various pathways were predicted using the tool BlastKOALA and KO (KEGG ORTHOLOGY) from the KEGG database and with eggNOG-mapper v2.1.12 to elucidate their regulatory functions (Supplementary File 6.2). Our analysis identified 13 proteins involved in different metabolic pathways (Table 1).

Protein-protein interaction analysis

Proteins exhibiting highly interactive functional associations were investigated using the STRING database, identifying 48 proteins with strong interacting relationships at a confidence level exceeding 0.400. Among them, 4 highest interacting clusters were shown in Figure 6 along with their details provided in Table 1 (Supplementary File 7).

Virulence factor analysis of hypothetical proteins

To identify virulent proteins associated with bacterial pathogenesis, VirulentPred 2.0 tool and VICMpred were used. Based on the prediction from both tools, 33 HPs were classified as potential virulence factors (Supplementary File 8).
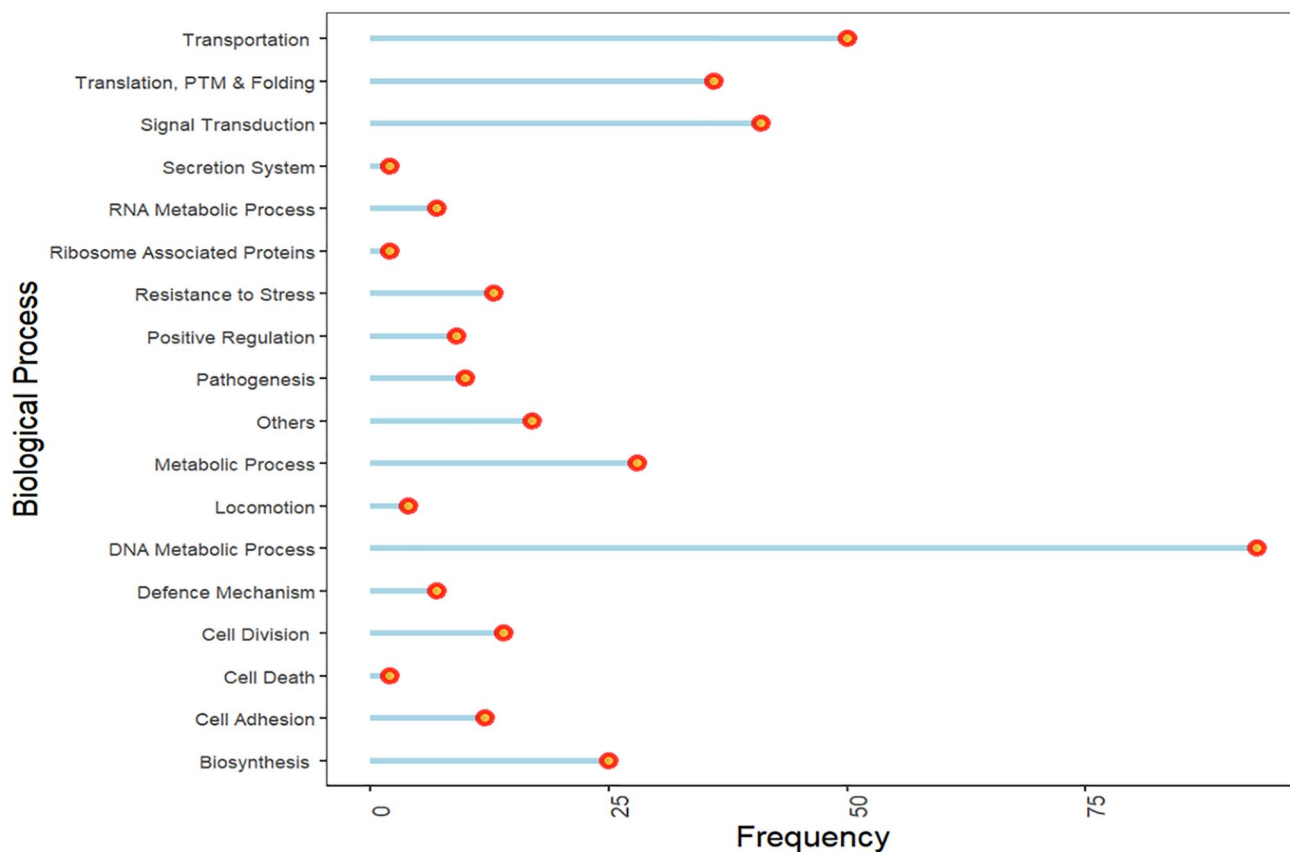
Figure 3. Figure representing the Gene Ontology of hypothetical proteins: biological process.

Selection of hypothetical proteins as putative drug targets

A total of 83 HPs, including 13 HPs associated with metabolic pathways, 33 HPs associated with virulence proteins, and 48 HPs associated with PPIs were primarily selected as drug targets. Among them, 3 proteins were duplicated in both pathway and PPI, 7 duplicated in virulent and PPI, and 1 duplicated in virulent and pathway. Homology to humans was predicted using BLASTp against the human proteome (taxid: 9606) available in the NCBI database. Among the 83 HPs, only 2 showed homology with the reference proteome (E-value < 0.001).[43]

Essentiality and druggability analysis

Among the 81 nonhomologous HPs identified, 80 were determined to be essential for pathogen survival based on a similarity search conducted in the DEG database (Supplementary File 9). The 80 essential nonhomologous HPs were then assessed for druggability through a similarity search against targets listed in the DrugBank database. On manual inspection, it was found that none of the 80 essential nonhomologous HPs exhibited similarity to a target in the DrugBank database (E-value = 0.00001), classifying them as novel drug targets (Supplementary File 9).

Homologs of known antibiotic–resistance determinants and their abundance

According to the CARD database, 66 of the total 80 essential nonhomologous novel HPs showed similarity to different antibiotic-resistance determinants. In addition, the abundance of those 80 proteins in different strains of *P rettgeri* based on NCBI-IPG database searching has been recorded in Supplementary File 9.

Structural modeling of functionally important hypothetical proteins

The identification and structural characterization of HPs are essential for selecting potential targets in drug design.[44] Among the 80 essential novel nonhomologous HPs analyzed, 29 proteins were primarily selected based on their subcellular locations (extracellular, outer membrane, and inner membrane). Subsequently, from these 29 proteins, the top 5 most abundant HPs (MBJ9971398.1, MBJ9971402.1, MBJ9971395.1, MBJ9971147.1, and MBJ9973330.1) were identified from the NCBI-IPG database and their tertiary structures were modeled using AlphaFold Server powered by AlphaFold 3. In addition, 5 other HPs (MBJ9971209.1, MBJ9972277.1, MBJ9972284.1, MBJ9973148.1, and MBJ9973050.1) involved in various transporter activities were also selected for structure
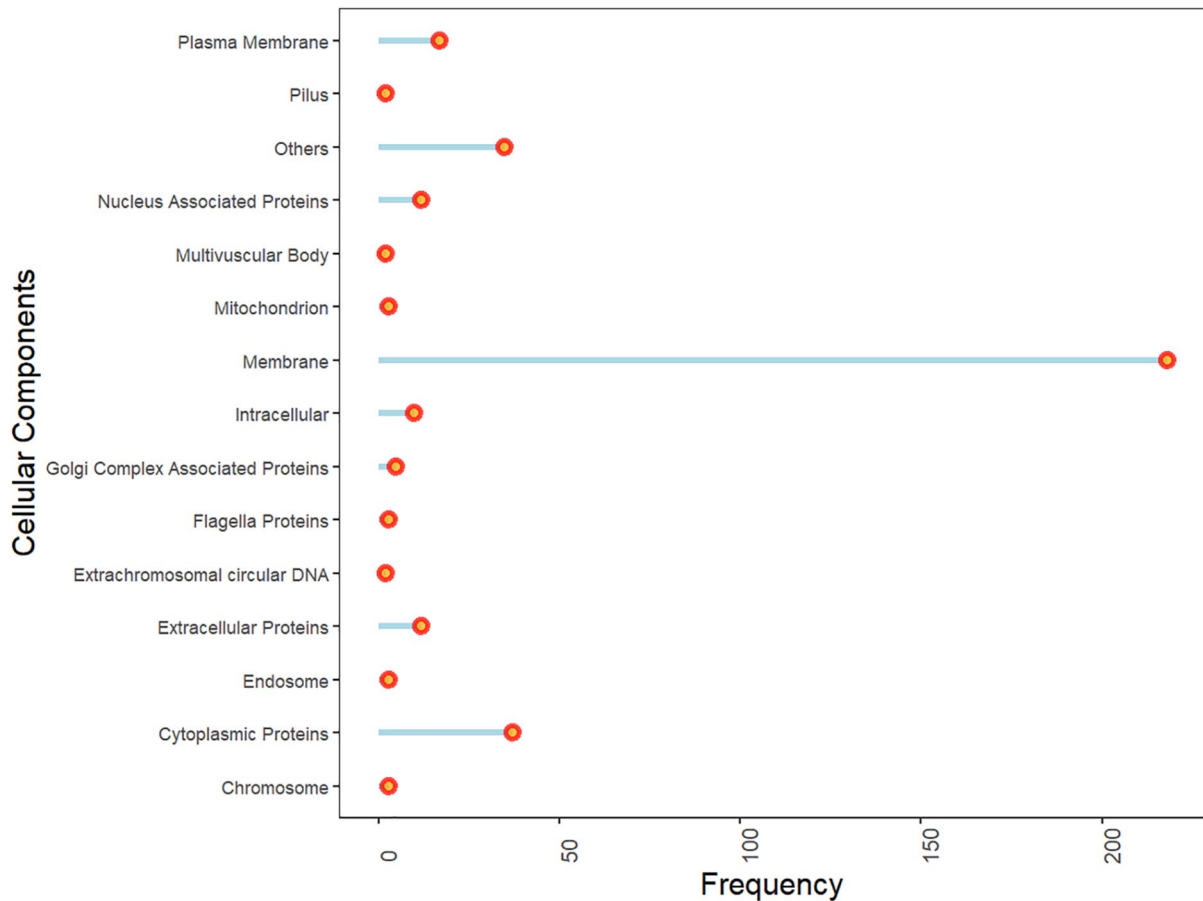
Figure 4. Figure representing the Gene Ontology of hypothetical proteins: cellular components.

prediction (Table 2 and Figure 7). The details of their physico-chemical characteristics are provided in Table 3.

Discussion

In this study, our target was to characterize HPs of a PDR *P rettgeri* strain MRSN845308, with an emphasis on identifying potential drug targets. The sequence of the selected strain was obtained from the NCBI Genome database. The genome of the strain encodes 4,405 proteins, including 3832 (87%) defined proteins and 573 (13%) uncharacterized (hypothetical) proteins. In a study conducted by Wang et al,[45] the HP was around 30% in a strain of *P rettgeri* annotated by Prokka. The lower percentage of HPs in the selected strain is due to the use of the Prokaryotic Genome Annotation Pipeline (PGAP) by NCBI, which has up-to-date and vast databases for annotation.[44] Hypothetical proteins shorter than 100 amino acids, categorized as pseudogenes, were omitted from this analysis to enhance the accuracy of functional predictions. Identifying functional domains can be invaluable in selecting drug targets. Proteins identified as bacteriophage proteins, including viral coat and capsid proteins, were excluded from the downstream analysis to ensure that the results are specific to bacterial proteins. The functional or conserved domains were explored by the NCBI-CDD database, InterProScan, PFP-FunDSeqE, and ScanProsite.

The cellular, molecular, and biological roles of the HPs were predicted by Argot 2. This website predicted cellular functions, biological processes, and molecular functions of 410 HPs. There was a high prevalence of binding proteins that attach to DNA, RNA, protein, ions, or other metabolites. DNA-binding proteins play an important role in survival and stress tolerance and may also be responsible for virulence.[46,47] Metal-binding proteins can serve as cofactors in a range of functions, such as controlling metabolism, regulating signal transport, and maintaining metal homeostasis.[48] Five HPs were annotated as transporter proteins. Transporter proteins that act as efflux pumps can help bacteria prevent the entry of antibiotics.[49] Some proteins were identified as outer membrane porins, which are crucial for the survival of Gram-negative bacteria. Several antibiotics enter the cell through the pores formed by porins, and by regulating these pores, bacteria can acquire drug-resistance. Their presence is also crucial to maintain membrane integrity.[50] Some HPs function as enzymes contributing to virulence and antibiotic resistance. They also play key roles in metabolic pathways, the synthesis of cell organelles, and bacterial multiplication, among other cellular processes.[51-53] We also
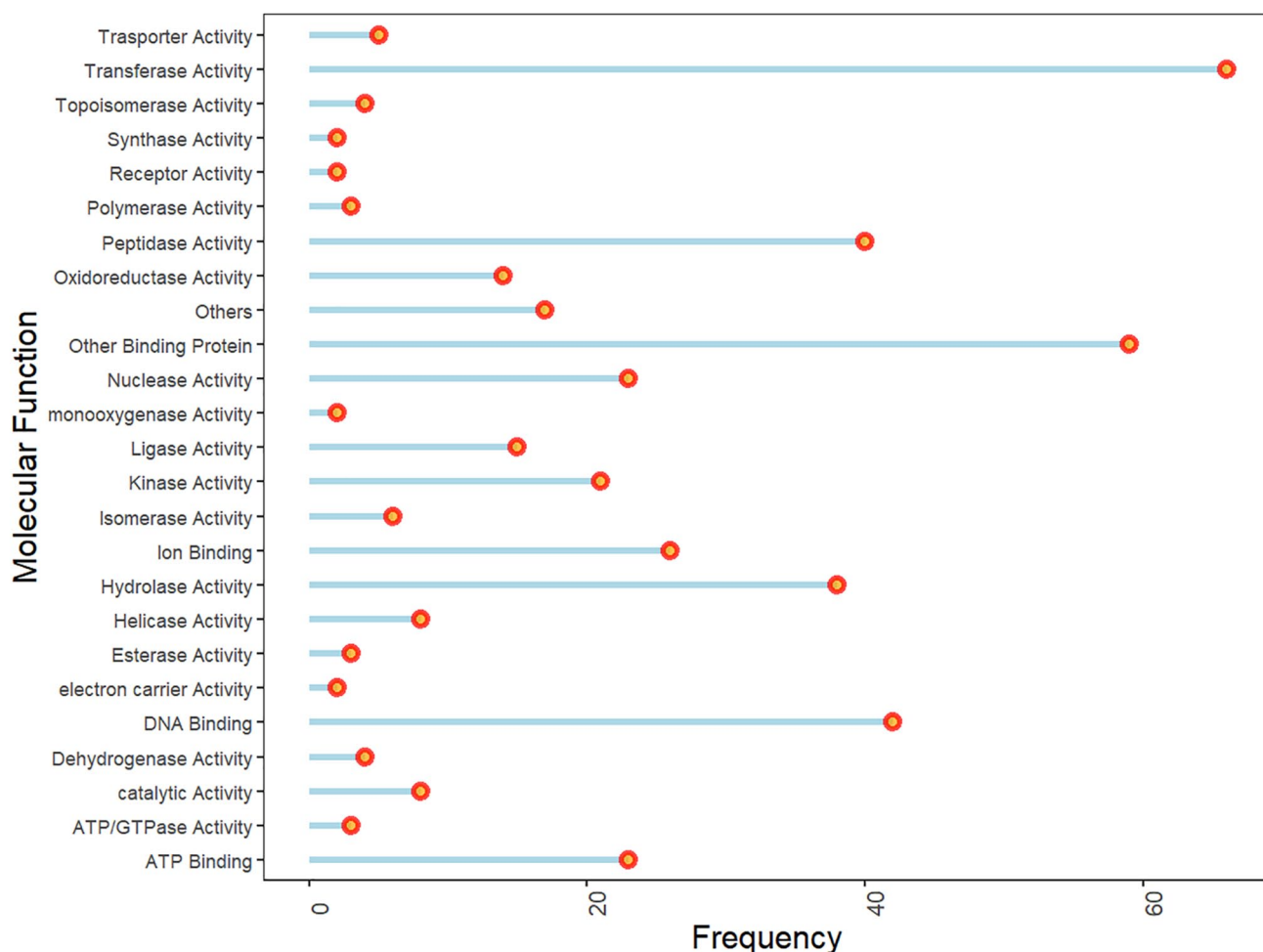
Figure 5. Figure representing the Gene Ontology of hypothetical proteins: molecular function.

Table 1. Hypothetical proteins interacting with functionally important proteins.

SEQID	CLUSTER	INTERACTING PROTEIN
MBJ9972745.1	CL:8030	Mixed, including tRNA endonuclease-like domain superfamily, and GDSL-like Lipase/Acylhydrolase family
MBJ9971369.1	CL:4568	Mixed, including Type III secretion needle MxiH, YscF, SsaG, EprI, PscF, EscF, and signal transduction response regulator, C-terminal effector
MBJ9972750.1	CL:8028	Mostly uncharacterized, including tRNA endonuclease-like domain superfamily, and bacterial regulatory protein, arsR family
MBJ9972779.1	CL:8309	Mixed, including RecT family, and coiled coil

found several proteins characterized as toxins, inhibitors, and immunogens. These proteins can play a significant role in bacterial pathogenesis.

Pathway analysis provides additional insights into the potential roles of these proteins within biological processes. A pathway analysis based on predictions from the KEGG database and eggNOG-mapper (v2.1.12) revealed that 13 HPs are involved in different metabolic pathways. Two important metabolic pathways involve riboflavin metabolism and lipopolysaccharide biosynthesis, affected by MBJ9973193.1 and MBJ9973676.1, respectively. In Riboflavin metabolism, MBJ9973193.1 HP may play a vital role in the synthesis or utilization of riboflavin (Vitamin B_2), which is essential for cellular metabolism. In Lipopolysaccharide biosynthesis, the identified MBJ9973676.1 may be involved in the biosynthesis of lipopolysaccharides, crucial components of Gram-negative bacterial outer membranes. Nonhomology searching suggests that the 2 HPs associated with these pathways share no
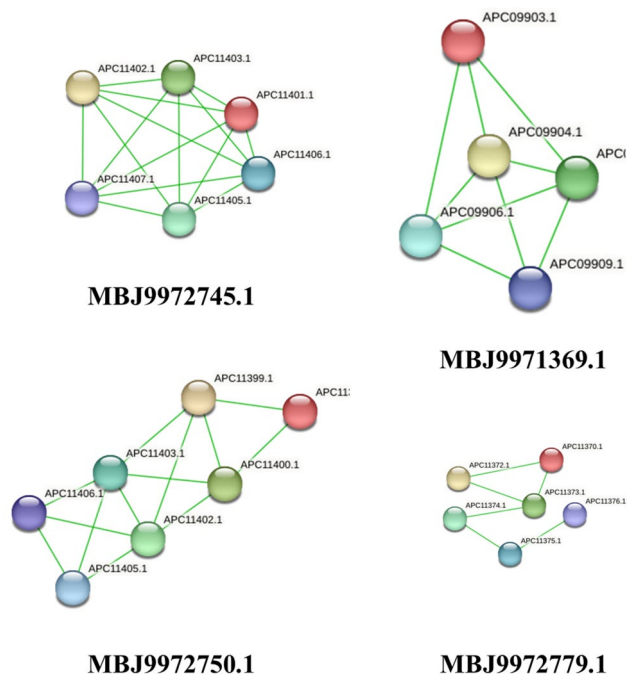
Figure 6. Protein-protein interaction of hypothetical proteins.

Table 2. Hypothetical proteins annotated as transporter by Argot2.

SEQUENCE ID	TRANSPORTER ACTIVITY
MBJ9971209.1	Protein-N(PI) phosphohistidine-mannitol phosphotransferase system transmembrane transporter activity
MBJ9972277.1	Ferrous iron transmembrane transporter activity
MBJ9972284.1	Transporter activity
MBJ9973148.1	Inorganic phosphate transmembrane transporter activity
MBJ9973050.1	Hydrogen ion transmembrane transporter activity; protein transporter activity

antibacterial therapies. Hence, MBJ9971369.1 can also be a potential target for drugs attacking T3SS systems.[43] Other HPs with a high number of interactions belong to families such as signal transduction response regulators, the tRNA endonuclease-like domain superfamily, protein secretion and bacterial export proteins, DNA integration, and transmembrane transport proteins. Biological networks serve as pivotal starting points for studying and addressing numerous human diseases.[55] Gram-negative pathogens often undergo evolutionary adaptations to enhance their virulence within the host environment, manifesting in alterations to traits such as motility, cell adhesion, and evasion of host immune responses. A consensus sequence analysis conducted by VICMpred and VirulentPred classified 43 (7.05% of HPs) proteins as potential virulence factors. These virulent proteins are involved in bacterial pathogenesis, including responses to stress and signaling pathways, as well as cellular functions such as enzymatic activity and transportation. The targeting of virulence factors has been increasingly recognized as a promising therapeutic approach against bacterial pathogenesis, offering an alternative to conventional antibiotics or as adjuncts to enhance host immune responses.[56]

Gram-negative pathogens often undergo evolutionary adaptations to enhance their virulence within the host environment, manifesting in alterations to traits such as motility, cell adhesion, and evasion of host immune responses. A consensus sequence analysis conducted by VICMpred and VirulentPred classified 33 proteins (8.1%) as potential virulence factors. A slightly higher percentage (18.02%) of virulence proteins were classified from HPs of *Orientia tsutsugamushi* in a study by Imam et al.[57] These virulent proteins are involved in bacterial pathogenesis, including responses to stress and signaling pathways, as well as cellular functions such as enzymatic activity and transportation. The targeting of virulence factors has been increasingly recognized as a promising therapeutic approach against bacterial pathogenesis, offering an alternative to conventional antibiotics or as adjuncts to enhance host immune responses.[56]

A total of 83 HPs from interaction, virulence, and pathway analysis was our primary drug target. Targets similar to human proteins may induce immunogenic reactions and, hence, often are eliminated from the lists of potential drug targets. Among the 83 HPs, only 2 showed homology with human proteins and thus were not considered in further analysis. Necessarily, the drug targets should be essential to the target pathogens, and as such, the essentiality of the 81 HPs was determined by the DEG database. A total of 80 HPs were found to be essential for pathogen survival. The HPs that match any target with similar biological functions from the list of nonhomologous HPs in the DrugBank database indicate potential druggability. Alternatively, if no match is found, the target is considered a "Novel target," suggesting a unique opportunity for drug discovery and development.[58] All were detected as novel targets,

homology to human proteins. Hence, they can be targeted by drugs without affecting metabolic pathways in humans.

Protein-protein interaction analysis revealed that 48 proteins had strong interacting relationships with a confidence level above 0.400. This threshold sets a minimum required interaction score, ensuring that only interactions with higher confidence are included in the predicted network. Lower scores mean more interaction but also higher false positive results. The HP MBJ9971369.1 was found to interact with the type III secretion (T3SS) systems, suggesting a role in virulence. The T3SSs are protein transport nanomachines that play a role in Gram-negative bacterial virulence and pathogenesis.[54] Owing to its role in host-pathogen interaction, the T3SS has been an attractive target for the development of new
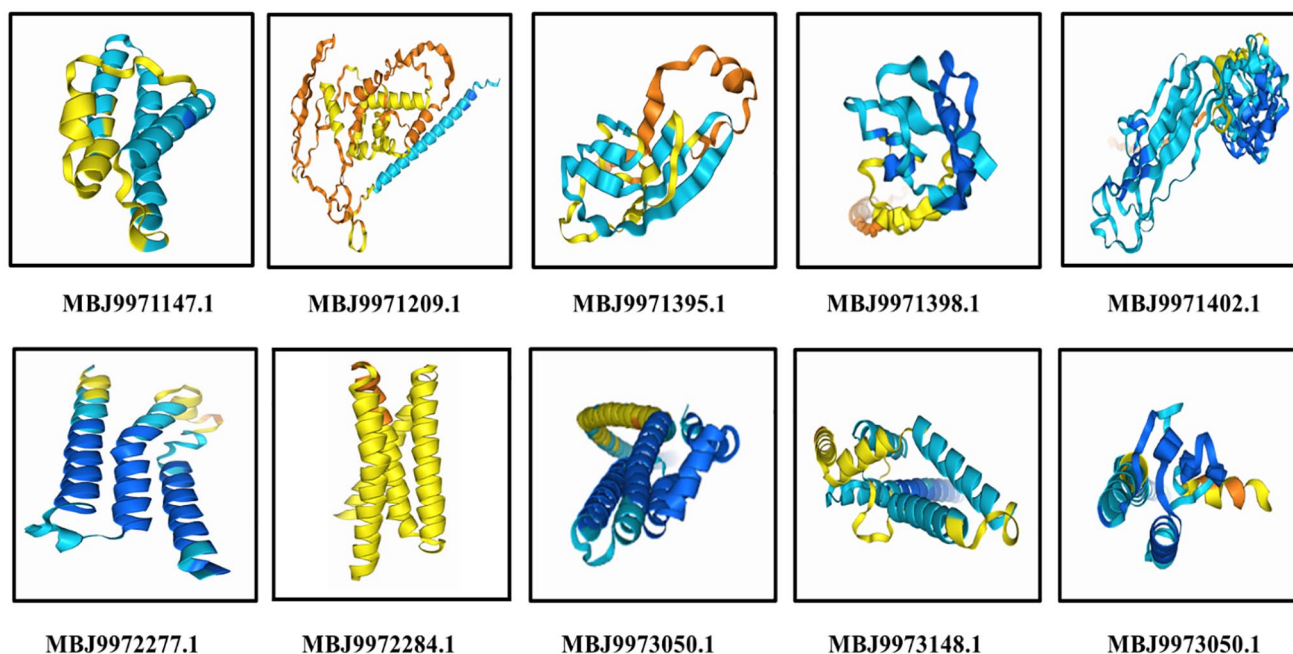
Figure 7. The 3D structure of the 10 target proteins are modeled by AlphaFold Server powered by AlphaFold 3.

Table 3. Physicochemical characteristics of the selected drug targets.

SEQUENCE ID	MOLECULAR WEIGHT (KDA)	ISO-ELECTRIC POINT (PI)	IN-STABILITY INDEX	GRAVY	ALIPHATIC INDEX	SUBCELLULAR LOCALIZATION	DOMAIN/MOTIF
MBJ9971398.1	36.742	9.170	31.826	−0.136	0.006	Extracellular	Carbohydrate binding
MBJ9971402.1	14.763	6.256	35.286	0.108	0.007	Outer membrane	Four-helical cytokines
MBJ9971395.1	15.597	5.871	48.003	−0.112	0.007	Extracellular	Immunoglobulin-like
MBJ9973330.1	21.584	6.152	40.116	−0.210	0.008	Inner membrane	Four-helical cytokines
MBJ9971147.1	15.006	9.581	46.087	0.785	0.010	Inner membrane	Cupredoxins
MBJ9971209.1	37.483	9.592	35.553	−0.566	0.007	Cytoplasm	OB-fold
MBJ9972277.1	13.908	9.047	35.509	0.537	0.008	Inner membrane	Ferredoxin-like
MBJ9972284.1	16.219	9.894	42.935	0.633	0.010	Inner membrane	Four-helical cytokines
MBJ9973148.1	20.518	7.508	46.926	0.261	0.008	Cytoplasm	Four-helical cytokines
MBJ9973050.1	32.633	6.244	50.009	0.105	0.008	Cytoplasm	Four-helical cytokines

which suggest that a large number of HPs as novel drug targets are still unexplored.

The localization of HPs in the cell is also an important factor when considering drug accessibility. Although intracellular drug targets may be less accessible, the extracellular targets are more exposed and have more potential to serve as drug aims. Besides, the extracellular proteins are often critical components of the biofilm matrix, and few attempts have been taken to target them for antimicrobial drugs.[59-61] Membrane proteins (inner or outer) represent a substantial proportion of current drug targets due to their roles in signal transduction, transport, bacterial pathogenicity, and drug-resistance.[62] Extracellular and outer membrane proteins have an extra advantage as being viable targets for vaccine development.[63] Considering the potential of membrane and extracellular proteins, 29 HPs (extracellular, outer, and inner membrane) were selected based

on the combined prediction from PSORTb, CELLO v.2.5, BUSCA, SOSUIGramN, and PSLpred. The prevalence of HPs in the *P rettgeri* population is also an important consideration for drug target evaluation. The targeted HPs abundance was compared with the mean abundance of the selected HPs using the NCBI-IPG database. The NCBI-IPG database is a resource within the NCBI that aims to simplify protein searches by grouping together protein sequences that are 100% identical. Based on the high abundance in NCBI-IPG, the top 5 HPs were selected as drug targets that are available in most *P rettgeri* isolates. The MBJ9971398.1 was the most prevalent one and was present in 216 strains, followed by MBJ9971402.1 (189 strains), MBJ9971395.1 (150 strains), MBJ9971147.1 (119 strains), and MBJ9973330.1 (111 strains). These highly abundant HPs can be targeted for developing broad-spectrum drugs that will be effective against most strains of *P rettgeri*.

Transporter proteins can be potential drug targets and have been targeted by antibacterial vaccines and antibiotics.[64,65] However, 5 HPs annotated as transporter proteins by Argot 2 were not found in HPs associated with virulence, metabolic pathways, and PPI. Considering their potential as drug targets, the HPs annotated as transporter proteins were also added to the list. The 5 transporter proteins, including MBJ9973148.1, MBJ9972277.1, MBJ9972284.1, MBJ9971209.1, and MBJ9973050.1, were present in 42, 15, 10, 2, and 2 strains, respectively. Although these transporter proteins are not that abundant, they can be targeted for developing strain-specific drugs. As the selected HPs have no crystal structure, determining their 3-dimensional (3D) structure is required to understand their function and suitability as targets for drug design. Accordingly, their 3D structures were modeled using the AlphaFold Server powered by AlphaFold 3. The recently developed artificial intelligence (AI)-based AlphaFold 3 can predict the structure and interactions of proteins and other molecules with unprecedented accuracy.[42] All these analyses in our work might help facilitate narrowing down targets and accelerate the drug screening process against *P rettgeri* infections.

Conclusions

This study facilitates the prompt identification of unknown functions associated with HPs using a variety of bioinformatics web tools and databases. The rigorous functional and spatial characterization of HPs in this study might unveil promising targets for therapeutic intervention. The 3D modeling of these targets will help identify the optimal binding sites that can be targeted for drug design by molecular docking and simulation. Virtual screening against these target proteins might be effective in the discovery of novel therapeutic compounds against *P rettgeri*. Further studies are, hence, required to address the therapeutic potential of these HPs and validate their suitability as drug targets through experimental approaches.

Author Contributions

DCP conceived and designed the experiment. DCP and TAA conducted the primary investigation and literature review. DCP, TAA, AAR, and SMMR performed data collection. DCP and TAA conducted data validation, formal analysis, and interpreted the results. DCP, TAA, AAR, and SMMR wrote the manuscript. DCP and TAA primarily revised and edited the manuscript. All authors reviewed and approved the final manuscript.

ORCID iD

Dipta Chandra Pal [iD] https://orcid.org/0009-0006-9542-1095

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

1. Stock I, Wiedemann B. Natural antibiotic susceptibility of *Providencia stuartii*, *P. rettgeri*, *P. alcalifaciens* and *P. rustigianii* strains. *J Med Microbiol*. 1998;47:629-642. doi:10.1099/00222615-47-7-629
2. Abdallah M, Balshi A. First literature review of carbapenem-resistant Providencia. *New Microbes New Infect*. 2018;25:16-23. doi:10.1016/j.nmni.2018.05.009
3. Yoh M, Matsuyama J, Ohnishi M, et al. Importance of Providencia species as a major cause of travellers' diarrhoea. *J Med Microbiol*. 2005;54:1077-1082. doi:10.1099/jmm.0.45846-0
4. O'Hara CM, Brenner FW, Miller JM. Classification, identification, and clinical significance of Proteus, Providencia, and Morganella. *Clin Microbiol Rev*. 2000;13:534-546. doi:10.1128/CMR.13.4.534-546.2000
5. Huff ML, Blome-Eberwein S. *Providencia rettgeri* infection compromising post-burn recovery: a lesson in the importance of follow-up care. *Cureus*. 2022;14:e25450. doi:10.7759/cureus.25450
6. Maiti TK, Singh VK, Pandey P. *Providencia rettgeri*: an unusual cause of central nervous system infections. *Am J Med Sci*. 2013;346:158-159. doi:10.1097/MAJ.0b013e318294f998
7. Sharma D, Sharma P, Soni P. First case report of *Providencia rettgeri* neonatal sepsis. *BMC Res Notes*. 2017;10:536. doi:10.1186/s13104-017-2866-4
8. Mc Gann P, Geringer MR, Hall LR, et al. Pan-drug resistant *Providencia rettgeri* contributing to a fatal case of COVID-19. *J Med Microbiol*. 2021;70:001406. doi:10.1099/jmm.0.001406
9. Shin S, Jeong SH, Lee H, Hong JS, Park MJ, Song W. Emergence of multidrug-resistant *Providencia rettgeri* isolates co-producing NDM-1 carbapenemase and PER-1 extended-spectrum β-lactamase causing a first outbreak in Korea. *Ann Clin Microbiol Antimicrob*. 2018;17:20. doi:10.1186/s12941-018-0272-y
10. Iwata S, Tada T, Hishinuma T, et al. Emergence of carbapenem-resistant *Providencia rettgeri* and *Providencia stuartii* producing IMP-type metallo-β-lactamase in Japan. *Antimicrob Agents Chemother*. 2020;64:e00382-20. doi:10.1128/AAC.00382-20
11. Gefen-Halevi S, Hindiyeh MY, Ben-David D, et al. Isolation of genetically unrelated blandm-1-positive *Providencia rettgeri* strains in Israel. *J Clin Microbiol*. 2013;51:1642-1643. doi:10.1128/JCM.00381-13
12. Ugwu MC, Ndunagu LU, Ejikwugwu CP, Esimone CO. First phenotypic and molecular characterization of clinical isolates of drug resistant *Providencia rettgeri* in Imo state, South-Eastern Nigeria. *Asian J Biomed Pharm Sci*. 2020;10:1-5. doi:10.35841/2249-622x.70.12507
13. Watanabe M, Nakano R, Tanouchi A, et al. Emergence and evolution of unique plasmids harboring bla IMP-70 and bla CTX-M-253 in multidrug-resistant *Providencia rettgeri*. *Microbiol Spectr*. 2022;10:e0120422. doi:10.1128/spectrum.01204-22
14. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*. 2003;31:3784-3788. doi:10.1093/nar/gkg563
15. Gill SC, von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem*. 1989;182:319-326. doi:10.1016/0003-2697(89)90602-7

16. Guruprasad K, Reddy BV, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng*. 1990;4:155-161. doi:10.1093/protein/4.2.155

17. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol*. 1982;157:105-132. doi:10.1016/0022-2836(82)90515-0

18. Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res*. 2015;43:W401-W407. doi:10.1093/nar/gkv485

19. Dobson L, Reményi I, Tusnády GE. CCTOP: a Consensus Constrained TOPology prediction web server. *Nucleic Acids Res*. 2015;43:W408-W412. doi:10.1093/nar/gkv451

20. Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*. 2001;305:567-580. doi:10.1006/jmbi.2000.4315

21. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol*. 2019;37:420-423. doi:10.1038/s41587-019-0036-z

22. Bhasin M, Garg A, Raghava GPS. PSLpred: prediction of subcellular localization of bacterial proteins. *Bioinformatics*. 2005;21:2522-2524. doi:10.1093/bioinformatics/bti309

23. Savojardo C, Martelli PL, Fariselli P, Profiti G, Casadio R. BUSCA: an integrative web server to predict subcellular localization of proteins. *Nucleic Acids Res*. 2018;46:W459-W466. doi:10.1093/nar/gky320

24. Yu CS, Lin CJ, Hwang JK. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci*. 2004;13:1402-1406. doi:10.1110/ps.03479604

25. Imai K, Asakawa N, Tsuji T, et al. SOSUI-GramN: high performance prediction for sub-cellular localization of proteins in Gram-negative bacteria. *Bioinformation*. 2008;2:417-421. doi:10.6026/97320630002417

26. Yu NY, Wagner JR, Laird MR, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics*. 2010;26:1608-1615. doi:10.1093/bioinformatics/btq249

27. de Castro E, Sigrist CJA, Gattiker A, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res*. 2006;34:W362-W365. doi:10.1093/nar/gkl124

28. Wang J, Chitsaz F, Derbyshire MK, et al. The conserved domain database in 2023. *Nucleic Acids Res*. 2023;51:D384-D388. doi:10.1093/nar/gkac1096

29. Jones P, Binns D, Chang HY, et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics*. 2014;30:1236-1240. doi:10.1093/bioinformatics/btu031

30. Shen H, Bin Chou KC. Predicting protein fold pattern with functional domain and sequential evolution information. *J Theor Biol*. 2009;256:441-446. doi:10.1016/j.jtbi.2008.10.007

31. Falda M, Toppo S, Pescarolo A, et al. Argot2: a large scale function prediction tool relying on semantic similarity of weighted Gene Ontology terms. *BMC Bioinformatics*. 2012;13:S14. doi:10.1186/1471-2105-13-S4-S14

32. Cantalapiedra CP, Hern andez-Plaza A, Letunic I, Bork P, Huerta-Cepas J. EggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol*. 2021;38:5825-5829. doi:10.1093/molbev/msab293

33. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol*. 2016;428:726-731. doi:10.1016/j.jmb.2015.11.006

34. Szklarczyk D, Kirsch R, Koutrouli M, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res*. 2023;51:D638-D646. doi:10.1093/nar/gkac1000

35. Sharma A, Garg A, Ramana J, Gupta D. VirulentPred 2.0: an improved method for prediction of virulent proteins in bacterial pathogens. *Protein Sci*. 2023;32:e4808. doi:10.1002/pro.4808

36. Saha S, Raghava GP. VICMpred: an SVM-based method for the prediction of functional proteins of gram-negative bacteria using amino acid patterns and composition. *Genomics Proteomics Bioinformatics*. 2006;4:42-47. doi:10.1016/S1672-0229(06)60015-6

37. Altschul SF, Madden TL, Schäffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25:3389-3402. doi:10.1093/nar/25.17.3389

38. Luo H, Lin Y, Liu T, et al. DEG 15, an update of the Database of Essential Genes that includes built-in analysis tools. *Nucleic Acids Res*. 2021;49:D677-D686. doi:10.1093/nar/gkaa917

39. Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res*. 2018;46:D1074-D1082. doi:10.1093/nar/gkx1037

40. McArthur AG, Waglechner N, Nizam F, et al. The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother*. 2013;57:3348-3357. doi:10.1128/AAC.00419-13

41. Sayers EW, Bolton EE, Brister JR, et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res*. 2022;50:D10-D17. doi:10.1093/nar/gkab1112

42. Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024;630:493-500. doi:10.1038/s41586-024-07487-w

43. Uddin R, Masood F, Azam SS, Wadood A. Identification of putative non-host essential genes and novel drug targets against *Acinetobacter baumannii* by in silico comparative genome analysis. *Microb Pathog*. 2019;128:28-35. doi:10.1016/j.micpath.2018.12.015

44. Li W, O'Neill KR, Haft DH, et al. RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. *Nucleic Acids Res*. 2021;49:D1020-D1028. doi:10.1093/nar/gkaa1105

45. Wang GH, Brucker RM. Genome sequence of *Providencia rettgeri* NVIT03, isolated from *Nasonia vitripennis*. *Microbiol Resour Announc*. 2019;8:e01157-18. doi:10.1128/mra.01157-18

46. Wang G, Maier RJ. A novel DNA-binding protein plays an important role in *Helicobacter pylori* stress tolerance and survival in the host. *J Bacteriol*. 2015;197:973-982. doi:10.1128/JB.02489-14

47. Kaito C, Morishita D, Matsumoto Y, Kurokawa K, Sekimizu K. Novel DNA binding protein SarZ contributes to virulence in *Staphylococcus aureus*. *Mol Microbiol*. 2006;62:1601-1617. doi:10.1111/j.1365-2958.2006.05480.x

48. Kumar S. Prediction of metal ion binding sites in proteins from amino acid sequences by using simplified amino acid alphabets and random forest model. *Genomics Inform*. 2017;15:162-169. doi:10.5808/gi.2017.15.4.162

49. Webber MA, Piddock LJ. The importance of efflux pumps in bacterial antibiotic resistance. *J Antimicrob Chemother*. 2003;51:9-11. doi:10.1093/jac/dkg050

50. Choi U, Lee CR. Distinct roles of outer membrane porins in antibiotic resistance and membrane integrity in *Escherichia coli*. *Front Microbiol*. 2019;10:953. doi:10.3389/fmicb.2019.00953

51. Zeng X, Lin J. Beta-lactamase induction and cell wall metabolism in Gram-negative bacteria. *Front Microbiol*. 2013;4:128. doi:10.3389/fmicb.2013.00128

52. Pancholi V, Chhatwal GS. Housekeeping enzymes as virulence factors for pathogens. *Int J Med Microbiol*. 2003;293:391-401. doi:10.1078/1438-4221-00283

53. Egorov AM, Ulyashova MM, Rubtsova MY. Bacterial enzymes and antibiotic resistance. *Acta Naturae*. 2018;10:33-48. doi:10.32607/20758251-2018-10-4-33-48

54. Deng W, Marshall NC, Rowland JL, et al. Assembly, structure, function and regulation of type III secretion systems. *Nat Rev Microbiol*. 2017;15:323-337. doi:10.1038/nrmicro.2017.20

55. Kushwaha SK, Shakya M. Protein interaction network analysis—approach for potential drug target identification in *Mycobacterium tuberculosis*. *J Theor Biol*. 2010;262:284-294. doi:10.1016/j.jtbi.2009.09.029

56. Fleitas Martínez O, Cardoso MH, Ribeiro SM, Franco OL. Recent advances in anti-virulence therapeutic strategies with a focus on dismantling bacterial membrane microdomains, toxin neutralization, quorum-sensing interference and biofilm inhibition. *Front Cell Infect Microbiol*. 2019;9:74. doi:10.3389/fcimb.2019.00074

57. Imam N, Alam A, Ali R, et al. In silico characterization of hypothetical proteins from *Orientia tsutsugamushi* str. Karp uncovers virulence genes. *Heliyon*. 2019;5:e02734. doi:10.1016/j.heliyon.2019.e02734

58. Crowther GJ, Shanmugam D, Carmona SJ, et al. Identification of attractive drug targets in neglected-disease pathogens using an in silico approach. *PLoS Negl Trop Dis*. 2010;4:e804. doi:10.1371/journal.pntd.0000804

59. Park AJ, Surette MD, Khursigara CM. Antimicrobial targets localize to the extracellular vesicle-associated proteome of *Pseudomonas aeruginosa* grown in a biofilm. *Front Microbiol*. 2014;5:464. doi:10.3389/fmicb.2014.00464

60. Rosas NC, Lithgow T. Targeting bacterial outer-membrane remodelling to impact antimicrobial drug resistance. *Trends Microbiol*. 2022;30:544-552. doi:10.1016/j.tim.2021.11.002

61. Gadar K, McCarthy RR. Using next generation antimicrobials to target the mechanisms of infection. *NPJ Antimicrob Resist*. 2023;1:11. doi:10.1038/s44259-023-00011-6

62. Huang P, Wang Z, Cai K, et al. Targeting bacterial membrane proteins to explore the beneficial effects of natural products: new antibiotics against drug resistance. *Curr Med Chem*. 2021;29:2109-2126. doi:10.2174/0929867328666621 0614121222

63. Wang Y, Wang X, Ali F, et al. Comparative extracellular proteomics of *Aeromonas hydrophila* reveals iron-regulated secreted proteins as potential vaccine candidates. *Front Immunol*. 2019;10:256. doi:10.3389/fimmu.2019.00256

64. Garmory HS, Titball RW. ATP-binding cassette transporters are targets for the development of antibacterial vaccines and therapies. *Infect Immun*. 2004;72:6757-6763. doi:10.1128/IAI.72.12.6757-6763.2004

65. Pahil KS, Gilman MSA, Baidin V, et al. A new antibiotic traps lipopolysaccharide in its intermembrane transporter. *Nature*. 2024;625:572-577. doi:10.1038/s41586-023-06799-7