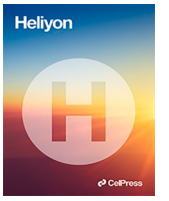
Heliyon 10 (2024) e39436

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Genomic insights into heterotrophic nitrifying-aerobic denitrifying bacteria from petroleum terminal effluents

Lutecia Rigueira Medina [a], Lívia Carneiro Fidélis Silva [a], Helena Santiago Lima [a], Pedro Marcus Pereira Vidigal [b], Alex Gazolla de Castro [a], Maira de Paula Sousa [c], Rodrigo Suhett de Souza [c], Sérgio Oliveira de Paula [d], Cynthia Canêdo da Silva [a,*]

[a] Department of Microbiology, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil
[b] Department of Biochemistry, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil
[c] Petrobras - Research and Development Center, Rio de Janeiro, RJ, Brazil
[d] Department of General Biology, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil

ARTICLE INFO

Keywords:
Comparative genomic analysis
Nitrogen removal
HN-AD
Differential expressed genes

ABSTRACT

Heterotrophic nitrification-aerobic denitrification (HN/AD) is a single-organism process that converts ammonia into nitrogen gas under strictly aerobic conditions, playing a crucial role in biological ammonia removal from industrial wastewater. Despite several studies, significant knowledge gaps remain about the genes involved in the process. This study aimed to characterize the genomes of four HN/AD bacterial strains, *Pseudomonas stutzeri* UFV5, *Pseudomonas balearica* UFV3, *Rhodococcus ruber* UFV2, and *Gordonia amicalis* UFV4, and identify potential genes involved in the HN/AD process. Results revealed that shared genes of these strains were primarily involved in amino acid and protein biosynthesis. The two *Pseudomonas* strains had more genes linked to nitrogen metabolism than the others. Additionally, four strains showed a significant number of hypothetical proteins and genes related to oxidative stress. Notably, no common nitrogen metabolism genes were found among the strains, indicating a lack of a shared HN/AD pathway. However, comparing these genomes with previous transcriptomic data of the *P. stutzeri* UFV5 identified nine shared proteins as potential HN/AD pathway candidates. This study enhances our understanding of the genomes of these HN/AD-capable bacterial strains and identifies nine candidate genes as markers for the HN/AD process.

1. Introduction

Biological ammonia removal can be achieved through various metabolic pathways. The heterotrophic nitrification-aerobic denitrification (HN/AD) pathway has gained prominence due to the unique characteristics of the microorganisms involved. These microorganisms are aerobic, utilize organic matter as a carbon source, and exhibit tolerance to high salinity conditions [1]. Despite its potential advantages, the HN/AD process remains less understood compared to the well-established sequential process of autotrophic nitrification followed by anaerobic denitrification.

Autotrophic nitrification, carried out under strictly aerobic conditions in two steps, involves the oxidation of ammonia to nitrite,

* Corresponding author. Av. Peter Henry Rolfs, s/n, University Campus, Laboratory of Applied Environmental Microbiology, Viçosa, 36570, Minas Gerais, Brazil.
E-mail address: ccanedo@ufv.br (C.C. da Silva).

https://doi.org/10.1016/j.heliyon.2024.e39436
Received 10 October 2024; Accepted 14 October 2024
Available online 18 October 2024

which is subsequently oxidized to nitrate in the second step. In the first step, three enzymes are involved: AmoA (ammonia monooxygenase), which converts ammonia to hydroxylamine; Hao (hydroxylamine oxidoreductase), which transforms hydroxylamine into nitrite; and Nar (nitrite oxidoreductase), responsible for the conversion of nitrite to nitrate [2,3]. Different autotrophic bacteria and archaea, including *Nitrosococcus, Nitrosospira, Nitrosomonas,* and the phylum *Thaumarchaeota*, act in the first step, while bacteria such as *Nitrobacter, Nitrospira, Nitrospina*, and *Nitrococcus* function in the second step [4].

In anaerobic denitrification, microorganisms use the nitrate produced in nitrification as a substrate and reduce it to nitrite via the *nap* gene, which encodes the periplasmic nitrate reductase enzyme, composed of the *nap*FDAGHBC operon, where NapA is the catalytic unit, NapAB is the functional unit, and NapC is responsible for electron transfer from quinone to NapAB [5]. Subsequently, nitrite reductase converts nitrite to nitric oxide, encoded by the *nirK* and *nirS* genes [6]. Next, the enzyme nitric oxide reductase (Nor) reduces nitric oxide to nitrous oxide [6]. Finally, the *nos* gene, which encodes the periplasmic nitrous oxide reductase enzyme, catalyzes nitrous oxide reduction to gaseous nitrogen [7]. Different bacterial genera carry out this process, including *Pseudomonas, Achromobacter, Alcaligenes, Bacillus, Micrococcus, Proteus*, *klebsiella,* and many others [8–12].

Unlike the anaerobic denitrification, the HN/AD (Heterotrophic Nitrification-Aerobic Denitrification) pathway is not yet fully understood. In addition to the *nos* gene, which reduces nitrous oxide to gaseous nitrogen [10,11], studies evaluating gene expression during the pathway induction of the HN/AD process have identified genes associated with nitrification and denitrification, such as *nirS*, *nirK*, *hao*, *napA*, *narGHJI*, and homologs of nitric oxide reductase *FlRd* and *FlRd-red*, present in this genus, using polymerase chain reaction (PCR) specific for these genes [9,13,14]. Bacteria from various genera, such as Klebsiella*, Paracoccus*, *Pseudomonas*, *Bacillus*, *Alkali*, *Acinetobacter*, *Pallidus*, *Halomonas*, *Rhizobium*, *Trichomonas*, *Sulthiobacter*, *Providenensis*, *Achromobacter*, *Rhodococcus*, *Gordonia*, *Arthrobacter*, and *Agrobacterium*, have been identified as capable of performing the HN/AD process [12–18].

Despite the limited knowledge about the metabolic pathways involved in the HN/AD process, it is understood that these microorganisms have a faster growth rate compared to autotrophic nitrifiers, leading to increased efficiency of biological ammonia removal. They are also able to remove ammonia under a single atmospheric condition and are more tolerant to environmental stresses. These characteristics make the presence of HN/AD special in the biological sludge of wastewater treatment plants [1].

The exploration and production of oil in pre-salt layers have resulted in the generation of a large amount of highly saline effluents [15]. This has made the process of ammonia removal very challenging, if not impossible, for autotrophic nitrifiers. As a result, HN/AD microorganisms have become prominent in this new scenario. Therefore, to better understand and control this process in wastewater treatment reactors and other systems, this study aimed to characterize and analyze the genome of four HN/AD bacteria: *Pseudomonas stutzeri* UFV5, *Pseudomonas balearica* UFV4, *Rhodococcus ruber* UFV2, and *Gordonia amicalis* UFV3, and to identify proteins involved in the HN/AD process.

2. Methods

2.1. Heterotrophic nitrifying and aerobic denitrifying bacterial strains

Four bacterial strains isolated from an activated sludge wastewater treatment system sample at the Maritime Terminal for oil exploration were studied. These strains were identified as *Pseudomonas stutzeri* UFV5 (accession number KY616652.1), *Pseudomonas balearica* UFV3 (accession number KX495190.2), *Rhodococcus ruber* UFV2 (accession number CP024890.1), and *Gordonia amicalis* UFV4 (accession number KU904410.1), as described in SILVA et al. [16].

2.2. Quality of reads

The sequencing reads were assessed for quality using FastQC version 0.11.9 (https://github.com/s-andrews/FastQC). Subsequently, adapter sequences were automatically detected and removed with TrimGalore version 0.6.7 [17], applying a Phred cutoff value of 25. Adapters were further removed using *Trimmomatic* version 0.39 [18] with specific parameters. The trimmed reads were then used for genome *contig* assembly.

2.3. Genome assembly

The genome assembly process used the “isolate” method of SPAdes version 3.15.3 [19], with odd k-mers from 21 to 127. *Reads* were sorted to form high-quality contigs, which were grouped into scaffolds using the “*scaffold*” method of RagTag 2.1.0 (https://github.com/malonge/RagTag). The scaffolds were then aligned and corrected against a reference genome using the BWA algorithm version 0.7.17 [20] and the Picard toolkit version 2.26.2 (https://github.com/broadinstitute/picard).

To ensure high-quality assemblies, the Pilon *software* version 1.24 [21] was used with up to five polishing iterations, correcting bases and incorrect assemblies to produce more contiguous genomes. The "*scaffolds*" files correspond to the sequences obtained after running SPAdes. Assembly quality was assessed using *assembly-stats* version 1.0.1, and scaffold coverage was calculated using the BBMap package version 38.76 (https://sourceforge.net/projects/bbmap).

2.4. Gene annotation and sequences deposition of the strains

After assembly and quality control, the genomes of the four HN/AD strains were submitted for deposition in the National Center for Biotechnology Information (NCBI) database (https://www.ncbi.nlm.nih.gov/guide/howto/submit-data).

Genome annotation was performed by NCBI, and Rapid Annotation was performed using the Subsystem Technology (RAST) web server (https://rast.nmpdr.org/rast.cgi) [22]. These annotations were compared with a FIGfam (protein families) collection, which assigns gene functions. RAST enables metabolic reconstruction by providing an overview of genome subsystems annotated by the SEED viewer (http://rast.nmpdr.org/) [23].

The genomes were deposited in NCBI under the project titled "Comparative Genomics of Heterotrophic Nitrifiers for Detection and Characterization of Genes Involved in the HN/AD Metabolic Pathway," project accession number: PRJNA851409. Deposits are available in the following databases: JANJEW000000000 – *P. stutzeri* UFV5, JANIEO000000000 – *P. balearica* UFV3, JANJEU000000000 – *R. ruber* UFV2, and JANJEV000000000 – *G. amicalis* UFV4. Submission of HN/AD strains can be accessed through the following accession numbers: SUB11715879 – *P. stutzeri* UFV5, SUB11715788 – *P. balearica* UFV3, SUB11715992 – *R. ruber* UFV2, and SUB11715916 – *G. amicalis* UFV4.

2.5. Orthologous genes

Comparative analysis of the four HN/AD bacterial strains was conducted using the OrthoVenn3 web service (https://orthovenn3.bioinfotoolkits.net/). OrthoVenn3 analyzed the similarity and clustering of orthologous genes among the evaluated strains. Genomic sequences encoding proteins were formatted in FASTA format, and the program compared orthologous genes among the HN/AD strains.

For these analyses, the OrthoMCL algorithm was employed for orthologous gene identification, with an E-*value* of $1{,}0 \times 10^{-2}$ and an inflation index of 1.5 to compare all protein similarities. Additionally, the Markov clustering algorithm was utilized to generate orthologous clusters [24].

2.6. Comparative genomics of HN/AD strains and differentially expressed genes

A comparative genomics analysis was conducted using the genomes of the strains in this study against differentially expressed genes described in SILVA et al. [25]. These authors conducted a transcriptomic study of the *Pseudomonas stutzeri* UFV5 strain and identified 29 differentially expressed genes under HN/AD pathway induction conditions. In addition to transcriptomic data, gene sequences (KEGG: Kyoto Encyclopedia of Genes and Genomes) encoding enzymes involved in the autotrophic nitrification and anaerobic denitrification (conventional process) of ammonia biological removal were included. Genomic sequences of bacteria such as *Nitrosomonas europaea* (NCBI accession number AL954747.1), capable of autotrophic nitrification, and *Micrococcus luteus* (NCBI accession number CP082331.1), a bacterium that does not remove ammonia, were also included as a negative control.

The BLASTn analysis compared the differentially expressed genes [25] with the nucleotide sequences of four studied bacteria was done. To ensure accurate translation, a BLASTp comparison was also conducted, comparing the amino acid sequences of the bacteria with the differentially expressed genes. This approach helped identify any errors during protein translation. Two other bacterial genomes, from *Nitrosomonas europaea* (involved in autotrophic nitrification) and *Micrococcus luteus* (which does not perform nitrification), were used as controls. Proteins meeting the criteria of identity ≥30 %, coverage ≥70 %, and *e*-value $\geq 10^{-5}$ from the BLASTp analysis were selected. Coverage was calculated by dividing the *query* size by the protein size and multiplying by 100. The selected proteins then proceeded to the annotation step for determining their functions using PFAM (http://pfam.xfam.org/) [26].

Table 1
Characterization and comparison of genome annotations of HN/AD bacterial strains derived from nitrifying sludge samples from an oil exploration effluent treatment plant.

	P. stutzeri UFV5	P. balearica UFV3	R. ruber UFV2	G. amicalis UFV4
Size (pb)	*4,559,141*	*4,554,350*	*5,563,789*	*5,080,713*
CDS [a]	*4208*	*4212*	*5284*	*4378*
GC (%) [b]	*64*	*64*	*70.5*	*67.3*
RNAs [c]	*66*	*59*	*63*	*62*
Pseudo Genes	*29*	*30*	*69*	*208*
tRNA [d]	*55*	*55*	*53*	*48*
Plasmids	*2*	*2*	-	-
Phages	*2*	*2*	-	-
N50 (pb) [e]	*740,842*	*742,452*	*432,576*	*190,949*
L50 [f]	*3*	*2*	*4*	*10*

[a] Number of protein-coding DNA sequences (CDS).
[b] Percentage (%) of G + C in the total nucleotide sequence.
[c] Number of RNAs.
[d] Number of transporter RNAs.
[e] Represents the contig size.
[f] Represents the quantity of contigs.

3. Results

3.1. Characterization of HN/AD strain genomes

The annotation of the four genomes of HN/AD bacterial strains derived from nitrifying sludge samples from an oil exploration effluent treatment plant yielded the following results, as represented in Table 1. The genome sizes of *P. stutzeri* UFV5 and *P. balearica* UFV3 were 4,559,141 bp and 4,554,350 bp, respectively, falling within the range of sizes (3923 bp to 7329 bp) observed for 117 *Pseudomonas* genomes cataloged in the NCBI. Both strains of the *Pseudomonas* genus exhibited a 64 % GC content. The *G. amicalis* UFV4 strain presented a genome size of 5,080,713 bp containing 67.3 % GC content, a result similar to that observed for the 49 reference genomes, which ranged from 3233 bp to 5962 bp. The genome of *R. ruber* UFV2 is the largest among those analyzed, with 5,563,789 base pairs and a higher GC content of 70.5 %. This falls within the expected size range, as the 133 reference genomes available in the NCBI ranged in size from 3944 base pairs to 9867 base pairs.

The statistical metrics N50 and L50 were utilized to assess the total number of base pairs (bp) in the assembly and the number of

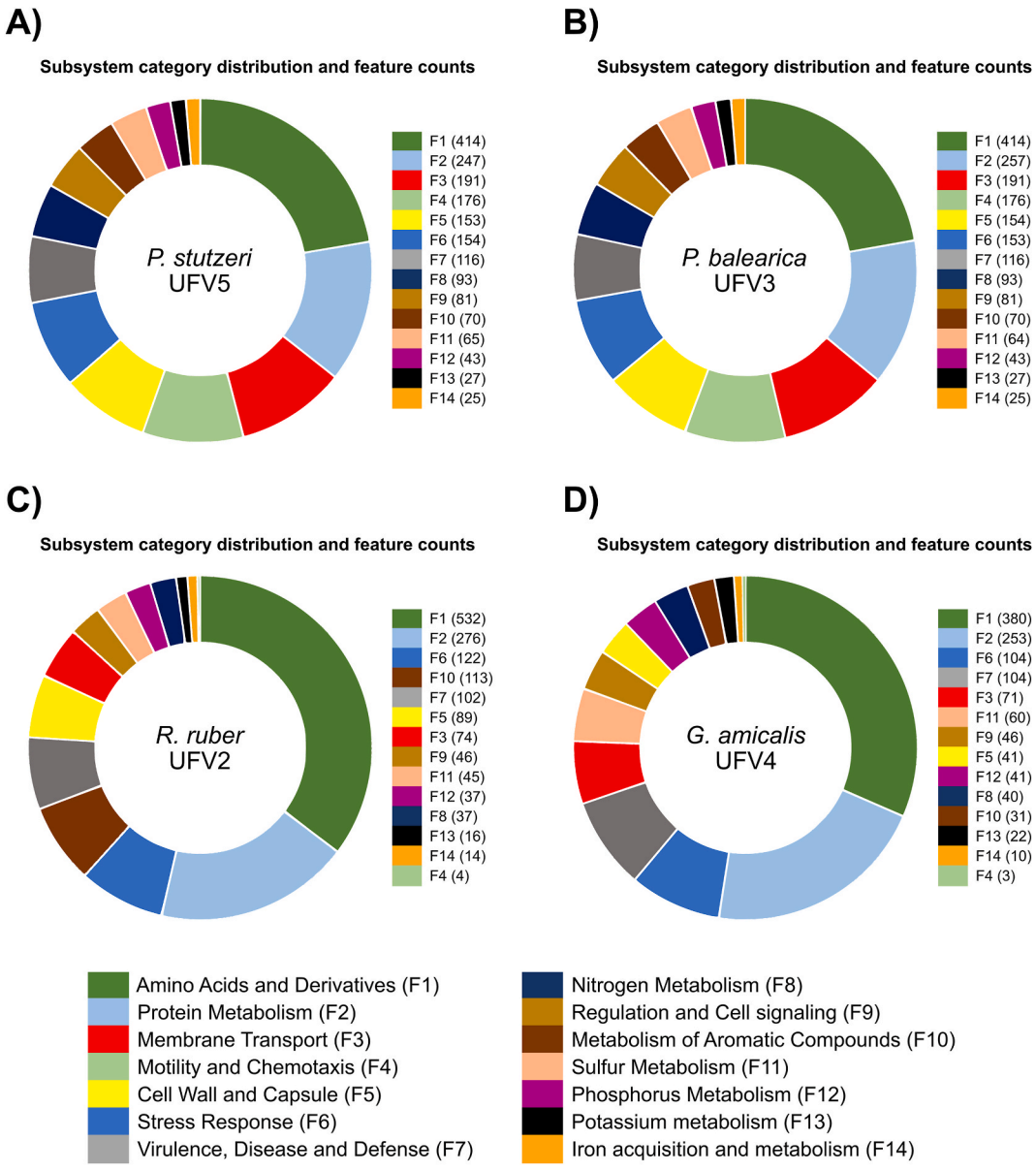


Fig. 1. Subsystem annotation by Rapid Annotation using Subsystem Technology (RAST). The graph shows the categories of subsystems of protein-coding genes. The number of genes assigned to each functional category is represented in parentheses. A) *Pseudomonas stutzeri* UFV5, B) *Pseudomonas balearica* UFV3, C) *Rhodococcus rubber* UFV2, e D) *Gordonia amicalis* UFV4.

contigs generated in this study for each genome, thus evaluating assembly quality. N50 evaluates *contig* continuity, while L50 evaluates the number of *contigs*. According to the applied statistical metrics, the strains exhibited excellent assembly and annotation quality, with N50 resulting in a higher number compared to the L50 value, which is desirable and indicates excellent genome assembly quality. In the case of *P. stutzeri* UFV5, it was observed that 50 % of the total bases assembled were contained in 3 *contigs* (L50) of size 740,842 bp (N50). For *P. balearica* UFV3, N50 was 740,852 bp and L50 was 2 contigs. In *R. ruber* UFV2, N50 was 432,576 bp and L50 was 4 *contigs*. Finally, in *G. amicalis* UFV4, N50 was 190,949 bp and L50 was 10 *contigs*.

3.2. The RAST subsystems of HN/AD strains

The RAST characterization indicated that 58 % of *Pseudomonas* genes were related to protein-coding subsystems, while 46 % were *R. ruber* UFV2 and *G. amicalis* UFV4 (Fig. 1). The remaining genes were categorized into general subsystems such as "Carbohydrates," "Cofactors, Vitamins, Prosthetic Groups, Pigments," "RNA metabolism," "Fatty Acids, Lipids, and Isoprenoids," "Respiration," "DNA Metabolism," "Nucleosides and Nucleotides," and others. However, these were not graphically represented to emphasize significant differences among the strains. *P. stutzeri* UFV5 has 4208 coding DNA sequences (CDS), with 52 % (2177) in 495 subsystems, including 140 with unknown functions. The remaining 48 % (2031) were not classified, with 1029 as hypothetical proteins. *P. balearica* UFV3 has 4305 CDS, with 52 % (2180) in 495 subsystems, including 139 hypothetical proteins. The remaining 47 % (2032) were unclassified, with 1030 hypothetical proteins. Both strains had two phage gene sequences (Fig. 1B).

R. ruber UFV2 has 4999 CDS, with 38 % (1879) in 403 subsystems, including 72 hypothetical functions. The remaining 62 % (3120) were unclassified, with 1924 unknown functions (Fig. 1C). *G. amicalis* UFV4 has 4378 CDS, with 37 % (1616) in 377 subsystems, including 76 hypothetical functions. The remaining 63 % (2762) were unclassified (Fig. 1D).

Regarding the genes involved in nitrogen metabolism and stress-related genes, highlighted by Silva et al. [25] for their role in the HN/AD process, *P. stutzeri* UFV5 and *P. balearica* UFV3 had 93 nitrogen metabolism genes and 154 stress response genes. *R. ruber* UFV2 had 37 nitrogen metabolism genes and 112 stress response genes. *G. amicalis* UFV4 had 40 nitrogen metabolism genes and 104 stress response genes (Table 2).

The *Pseudomonas* strains contained a higher number of denitrification genes, while *R. ruber* UFV2 and *G. amicalis* UFV4 had similar numbers of ammonia assimilation genes. All four HN/AD strains had more oxidative stress genes than ammonia assimilation genes, with the *Pseudomonas* species having the highest number.

3.3. Orthologous and unique genes

The analysis in OrthoVenn3 identified between the 14,749 proteins, a total of 6157 protein clusters in the four HN/AD bacterial strains. Among these, 855 clusters represented 3545 proteins were shared among the four genomes. These shared proteins were involved in the following biological processes: 16.4 % in the metabolic process; 15.1 % in biological processes; 14.6 % in cellular metabolism; 11.6 % in nitrogen compound metabolic processes; 9.3 % in primary metabolic processes; 7.4 % in cellular processes; 7.1 % in heterocycle metabolic processes; 6.9 % in organic acid processes; 6.7 % in cellular aromatic compound metabolic processes; and 4.9 % in nucleobase-containing compound metabolic processes.

In the analysis of exclusive protein clusters, 145 unique clusters were detected for *R. ruber* UFV2, 93 for *G. amicalis* UFV4, 2 clusters for *P. balearica* UFV3, and no exclusive clusters for *P. stutzeri* UFV5 (Fig. 2A). The *Pseudomonas* strains exhibited 3101 common clusters. Meanwhile, the *R. ruber* UFV2 and *G. amicalis* UFV4 strains shared 1732 clusters, as they belong to the same taxonomic order.

In *R. ruber* UFV2, 2868 clusters were identified, of which 1728 were single-copy protein clusters. In *G. amicalis* UFV4, 2773 clusters were identified, with 1322 being single-copy protein clusters. The *Pseudomonas* strains had a higher number of clusters, but a significantly lower number of single-copy genes compared to the other two strains. *P. balearica* UFV3 presented 4187 clusters, of which 10 were single-copy, while *P. stutzeri* UFV5 had 4185 clusters, with 8 being single-copy clusters (Fig. 2B).

Table 2
Genes from Nitrogen Metabolism and Stress Response Subsystems. On the horizontal axis are the two subsystems, and on the vertical axis are the subcategories found. The numbers correspond to the quantity of genes involved in each subsystem in the HN/AD strains (*Rhodococus ruber* UFV2; *Pseudomonas balearica* UFV3, *Gordonia amicalis* UFV4 e *Pseudomonas stutzeri* UFV5).

		P. stutzeri UFV5	*P. balearica* UFV3	*R. ruber* UFV2	*G. amicalis* UFV4
Nitrogen Metabolism	Ammonia assimilation	16	16	15	17
	Nitrosative stress	4	4	1	1
	Denitrification	33	33	4	4
	Cyanate hydrolysis	4	4	0	5
	Dissimilatory nitrite reductase	13	13	0	0
	Nitrate and nitrite ammonification	23	23	8	13
Stress Response	Periplasmic Stress	6	6	1	0
	Oxidative stress	72	72	23	57
	Osmotic stress	15	14	1	10
	Heat shock	16	16	0	14
	Cold shock	4	4	0	3
	Detoxification	9	9	3	15
	no subcategory	32	32	8	5

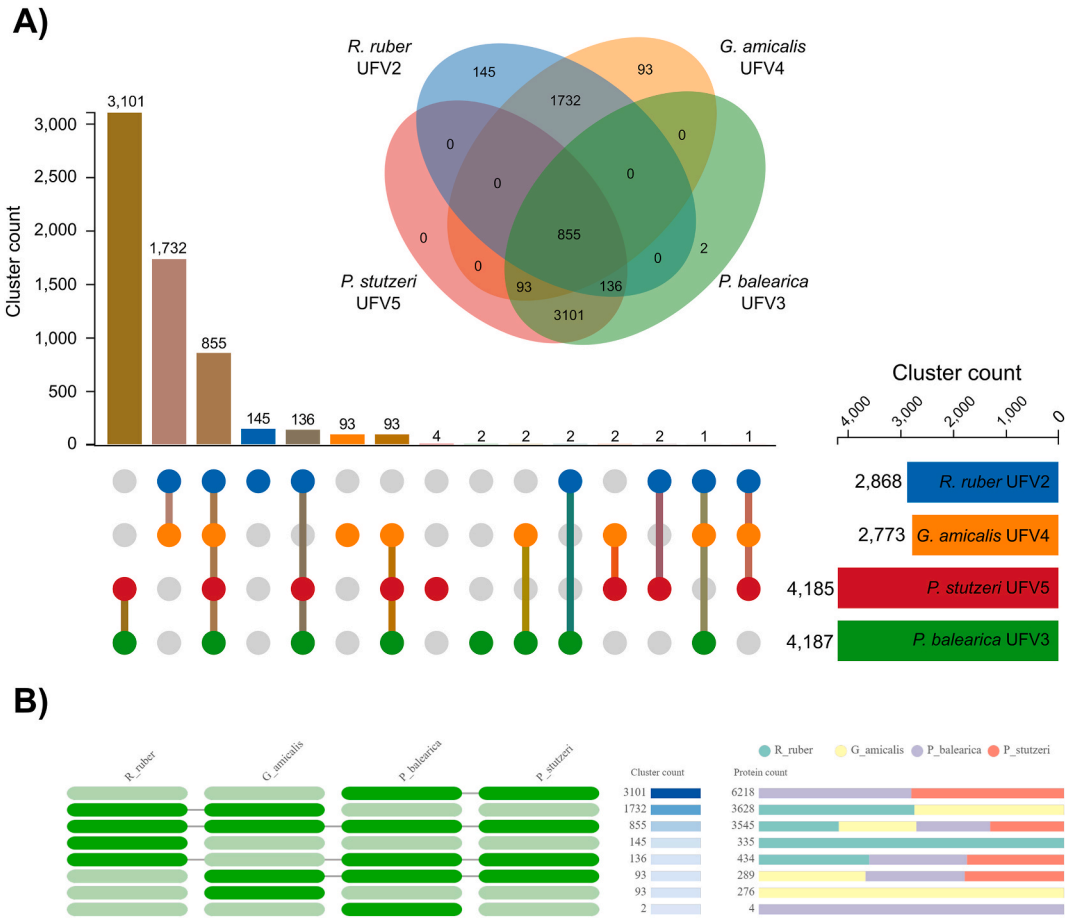


Fig. 2. Representation of OrthoVenn3 analysis results among the four HN/AD strains. (A) Visual representation of cluster sharing and unique clusters among the strains. (B) Graphical visualization of orthology cluster analysis results among the four strains and the number of proteins (reduce the Venn diagram size).

3.4. Comparative genomics

In the comparative genomics analysis, no genes involved in autotrophic nitrification were found in the evaluated genomes. However, genes involved in anaerobic denitrification were identified in the *Pseudomonas* strains, which are already described as anaerobic denitrifiers (Fig. 3). The detected genes included *nap*, which encodes the periplasmic nitrate reductase enzyme; *nirS*, encoding the nitrite reductase enzyme; *Nor*, encoding the nitric oxide reductase enzyme; and the *nos* gene, encoding the periplasmic nitrous oxide reductase enzyme, with the latter observed only in *P. balearica* UFV3 (Fig. 3B). The genes involved in the anaerobic denitrification process were not found in Gram-positive strains, *R. ruber* UFV2, and *G. amicalis* UFV4.

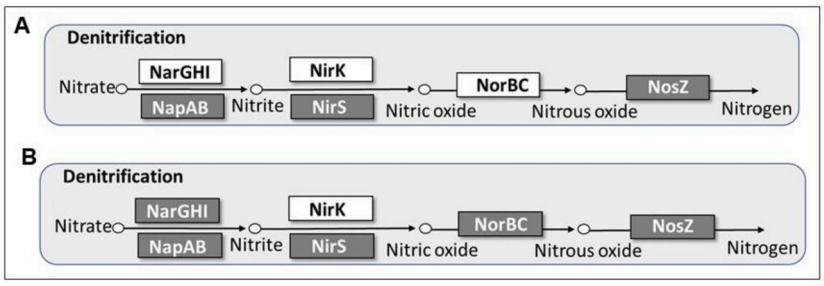


Fig. 3. Comparative genomics of HN/AD strains against genes involved in conventional anaerobic denitrification. According to KEGG: *Pseudomonas stutzeri* UFV5 **(A)** and *Pseudomonas balearica* UFV3 **(B)**, The genes filled with (■) were found in the HN/AD strain of this study.

The genomes were also compared with 35 proteins encoded by differentially expressed genes identified in the transcriptome of *Pseudomonas stutzeri* UFV5 studied by Silva et al. (2020) in a HN/AD pathway induction in culture medium containing ammonia (Table 3). Comparative genomics revealed that of the 35 proteins, six (**WP_013982320.1, WP_013983005.1, WP_013982283.1, WP_013983347.1, WP_010594313.1,** and **WP_005260096.1)** were observed in at least three studied HN/AD strains and were not

Table 3
Differentially expressed genes found in the transcriptome of *P. stutzeri* UFV5 (Silva et al., 2020) and in other HN/AD strains evaluated in this study. The presence or absence of these genes in the HN/AD strains (*Rhodococcus ruber* UFV2, *Pseudomonas balearica* UFV3, *Gordonia amicalis* UFV4, and *Pseudomonas stutzeri* UFV5) and the bacteria Nitrosomonas europaea and Micrococcus luteus, which were used as controls, was assessed. The former is involved in the autotrophic nitrification process, while the latter does not perform nitrification. The squares marked (■) indicate that the genes encoding the proteins were found in the genomes of the strains evaluated in this study.

		HN/AD					
Gene code in the proteome	Pfam annotation	*P. stutzeri* UFV5	*P. balearica* UFV3	*R. ruber* UFV2	*G. amicalis* UFV4	*N. europaea*	*M. luteus*
WP_013982320.1	**Protein of unknown function**	■	■	■			
WP_003246255.1	"Cold-shock" DNA-binding domain	■	■	■	■	■	
WP_013983373.1	BFD-like [2Fe-2S] binding domain	■	■	■	■	■	
WP_003283670.1	NADH dehydrogenase	■	■	■	■	■	■
WP_003293422.1	"Cold-shock" DNA-binding domain	■	■	■	■	■	■
WP_041771612.1	50S ribosomal protein L28	■	■	■	■	■	■
WP_026006463.1	Putative quorum-sensing-regulated virulence factor	■	■				
WP_013983644.1	Iron-sulphur cluster assembly	■	■				
WP_013983005.1	**Protein of unknown function**	■	■	■			
WP_013982942.1	'Cold-shock' DNA-binding domain	■	■	■	■	■	■
WP_013982283.1	**Protein of unknown function**	■	■	■			
WP_013982024.1	Protein of unknown function	■	■				
WP_011912999.1	Sulfurtransferase TusA	■	■	■		■	
WP_011912373.1	30S ribosomal protein S16	■	■	■	■		■
WP_003298367.1	50S ribosomal protein L35	■	■	■	■	■	■
WP_003284185.1	50S ribosomal protein L33	■	■	■	■	■	■
WP_003282211.1	**"Cold-shock" DNA-binding domain**	■	■	■	■	■	
WP_003281814.1	50S ribosomal protein L36	■	■	■	■	■	■
WP_041771783.1	(Na+)-NQR maturation NqrM	■	■				
WP_013984129.1	Protein of unknown function	■	■				
WP_013983984.1	Prokaryotic Cytochrome C oxidase subunit IV	■	■				
WP_013983347.1	**Bacterial regulatory proteins, tetR family**	■	■		■		
WP_013982479.1	Protein of unknown function	■	■				
WP_013982009.1	DNA-binding protein	■	■				
WP_011913772.1	Protein of unknown function	■	■				
WP_011913664.1	Heme exporter protein D	■	■				
WP_011913223.1	Protein of unknown function	■	■				
WP_011912146.1	**Protein of unknown function**	■	■	■	■	■	
WP_011911750.1	Phosphate-starvation-inducible E	■	■				
WP_003294452.1	Alanine-zipper, major outer membrane lipoprotein	■	■				
WP_003293518.1	50S ribosomal protein L32	■	■			■	
WP_003290642.1	30S ribosomal protein S21	■	■				
WP_003281834.1	**50S ribosomal protein L29**	■	■	■	■	■	
WP_010594313.1	**Glutamine synthetase**	■	■	■	■		
WP_005260096.1	**Aminomethyltransferase**	■	■	■	■		

found in the control genomes of the *Micrococcus luteus* or the *Nitrosomonas europaea*, indicating the potential relation of these genes with HN/AD process of the studied strains (Table 3). In addition, three more genes encoding proteins **WP_003282211.1, WP_011912146.1**, and **WP_003281834.1,** were found in three of the HN/AD bacterial strains and *N. europaea* (Table 3). Despite being observed in the genome of a bacterium that performs autotrophic nitrification, these three genes could still be related to the HN/AD process. Thus, based on comparative genomics, nine genes encoding proteins with potential relation in the HN/AD process have been selected in this study.

4. Discussion

The four bacterial strains that carry out the HN/AD process were isolated from activated sludge of petroleum effluent treatment, and two of them, *G. amicalis* and *P. balearica* were physiologically characterized by Silva and collaborators [16]. *R. ruber* UFV2 and *G. amicalis* UFV4, are aerobic, Gram-positive, and belong to the phylum Actinomycetota. However, *R. ruber* UFV2 belongs to the family Nocardiaceae, while *G. amicalis* UFV4 belongs to the family Gordoniaceae. *R. ruber* are opportunistic pathogenic bacteria and were described as HN/AD bacteria by CHEN and collaborators [28], showing significant potential for use in the bioremediation of environmental pollutants [29]. *G. amicalis* UFV4 was first described as HN/AD by our research group [16]. *Pseudomonas stutzeri* UFV5 and *Pseudomonas balearica* UFV3, both belonging to the class Gammaproteobacteria, are facultative anaerobes, Gram-negative, and rod-shaped [27]. *Pseudomonas* species are versatile in different environments and have been studied in fields ranging from medicine to polluted environments [10].

The genome characterization of four strains revealed that the most abundant subsystems were the same, those related to amino acid and protein synthesis. The third most abundant subsystem varied among the strains. In *Pseudomonas*, the membrane transport subsystem was also abundant. This might be because Gram-negative bacteria have a more complex outer membrane, thus having a greater variety of specialized transport proteins performing various functions [30].

Regarding genes related to nitrogen metabolism, a higher number was observed for *Pseudomonas* strains (93 genes), while less than half were found for *R. ruber* UFV3 (37 genes) and *G. amicalis* UFV4 (40 genes). This highlights that each evaluated genus presented its characteristic set of genes related to nitrogen metabolism, especially in the denitrification stage for *Pseudomonas*. It is worth noting that *Pseudomonas* strains can use nitrate as an electron source (denitrification) under anoxic conditions. This is because they are facultatively anaerobic bacteria [31]. Another subsystem that caught our attention in the four strains was the stress response, which included genes linked to detoxification. According to previous studies [16,25,32,33], the HN/AD process may act as a detoxification mechanism for microorganisms under conditions of high ammonia concentration. It is already known that effluents from Brazilian oil terminal treatment plants have a high concentration of ammonia [15], which corroborates this hypothesis.

The analysis of similar and common ancestral genes showed that the 855 clusters shared by four HN/AD strains potentially represent a central set of functionalities essential for these microorganisms, associated with fundamental biological processes, and indicate the conservation of genes throughout the evolution of these bacteria. As expected, the clusters among the *Pseudomonas* strains showed a higher identity, and no exclusive protein clusters were found between *Pseudomonas* and *G. amicalis* UFV4 or *R. ruber UFV2*. However, exclusive protein clusters were observed for *G. amicalis* UFV4 and *R. ruber* UFV2.

These two strains, *G. amicalis* UFV4 and *R. ruber* UFV2, exhibited the highest percentage of unique copy orthologous genes, with 60 % for *R. ruber* UFV2 and 48 % for *G. amicalis* UFV4. On the other hand, this percentage was less than 10 % in the *Pseudomonas* strains. Unique copy orthologs are crucial as they imply a direct evolution with the genome, used in various phylogenetic analyses, as they capture the entire genome evolution as single copies [35].

The comparative genomics conducted among the strains in search of genes involved in the conventional process of ammonia removal (autotrophic nitrification and anaerobic denitrification) revealed a surprising finding. Only genes involved in the process of anaerobic denitrification, known to be present in the *Pseudomonas* genus, were found. Our results corroborate the previous study [16], which had already observed by PCR that only genes related to anaerobic denitrification were present in the *Pseudomonas* strains evaluated in this study, and that *R. ruber* UFV2 and *G. amicalis* UFV4 did not present genes involved in the conventional process.

However, our results differed from those of the study by CHEN [36], who analyzed a species of *Rhodococcus* sp. S2 and found three genes involved in the conventional denitrification process: *NarG, NxrA*, and *NirBD*. This author suggested that these genes are essential for this species' the HN/AD process.

Although previous studies have suggested that anaerobic denitrification involved genes in the HN/AD process [9], our work did not find these genes in the Gram-positive strain genomes. In addition, Silva and collaborators [25] also did not find anaerobic denitrification genes being expressed at *Pseudomonas stutzeri* UFV5 [25] at a concentration of 1300 mg of ammonia. These results suggested the involvement of other genes in the HN/AD process of the four strains studied.

In this way, we identified nine common proteins between the four strains and the differentially expressed genes of *P. stutzeri* UFV5 transcriptome as potentially related to the HN/AD process. The protein WP_013982320.1 has an unknown function, and WP_013982283.1 is a protein of the DUF2945 domain, which belongs to the Tudor domain, including chromo, MBT, and PWWP domains [34]. However, its function is still unclear, although it is involved in positive and negative regulation of transcription [35]. The protein WP_013983347.1 functions as a transcriptional regulator (TFRs) [36], and it has been described for the genera *Pseudomonas* and *Gordonia*. The protein WP_010594313.1, glutamine synthetase, functions to convert ammonia into glutamine [12], and the protein WP_005260096.1, aminomethyltransferase, catalyzes the transfer of ammonia to the aminomethyl portion of glycine [37]. TRAN and colleagues [38] have previously observed glutamine synthetase and aminomethyltransferase proteins in the genus *Pseudomonas*; however, this study was the first to observe the genes encoding these enzymes in HN/AD strains of the genus *Gordonia* and *Rhodococcus*. The last two proteins, WP_011912146.1 and WP_003281834.1, are ribosomal proteins involved in the protein translation

process. SILVA et al. [25] explained that during the degradation of ammonia by *P. stutzeri* UFV5, there was an increase in the expression of proteins of the cellular translation machinery; therefore, these proteins could also be candidates for possible relation with HN/AD process.

Previous studies with the strains of this work [16,25], inferred that there may not be a unique metabolic pathway for the HN/AD process between the different microbial groups. As previously stated, some literature works suggested that the HN/AD process may be a bacterial response to ammonia toxicity, i.e. a detoxification mechanism, when in high ammonia concentrations [16,25,32,33].

However, in the present study were found nine common genes, even though these proteins are not closely involved with the HN/AD process, they may be related to the pathway, and these genes may be used as genetic markers for monitoring and better understanding the ecology of this group in various environmental or constructed systems, as well as identifying new microorganisms capable of performing the HN/AD process. Future work involving the gene expression of the four strains will be developed to validate the relationship of the nine genes to the HN/AD process.

5. Conclusion

Our research provided new insights into understanding the heterotrophic nitrification and aerobic denitrification (HN/AD) process by analyzing four strains able to carry out this process. Our findings suggest that the absence of genes responsible for autotrophic nitrification in the genomes of these strains, and the absence of anaerobic denitrification genes in the *G. amicalis* UFV4 and *R. ruber* UFV2 indicates the involvement of other genes in the HN/AD process of this strains. By comparing the four genomes with previous results of *P. stutzeri* UFV5 transcriptome, we identified nine common proteins, including cold-shock DNA-binding domain proteins, bacterial regulatory proteins from the TetR family, the ribosomal protein L29 of the 50S subunit, glutamine synthetase, aminomethyltransferase, and proteins with unknown function. The validation of these candidates' proteins related to the HN/AD process will open the possibility of contributing to the identification of genetic markers that can be used to find new HN/AD microorganisms, as well as monitoring them in different environments. Furthermore, a better understanding of this mechanism can be used to improve biological ammonia removal.

CRediT authorship contribution statement

Lutecia Rigueira Medina: Conceptualization, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft. **Lívia Carneiro Fidélis Silva:** Investigation, Methodology, Supervision, Writing – review & editing. **Helena Santiago Lima:** Investigation, Methodology. **Pedro Marcus Pereira Vidigal:** Formal analysis, Methodology, Software. **Alex Gazolla de Castro:** Methodology, Visualization, Writing – review & editing. **Maira de Paula Sousa:** Resources, Supervision. **Rodrigo Suhett de Souza:** Resources, Supervision. **Sérgio Oliveira de Paula:** Resources, Supervision, Writing – review & editing. **Cynthia Canêdo da Silva:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Funding

This work was supported by the PETROBRAS [grant numbers 8603–4600391375]; FAPEMIG (Fundação de Amparo à Pesquisa de Minas Gerais), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Código de Financiamento 001), and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnologia).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:Lutecia Rigueira Medina reports financial support was provided by Minas Gerais State Foundation of Support to the Research. Cynthia Canedo da Silva reports a relationship with Petrobras that includes: funding grants. 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.

Acknowledgment

The authors are grateful to PETROBRAS S.A. for the financial support and to FAPEMIG - Fundação de Amparo à Pesquisa do Estado de Minas Gerais, Brazil, CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil and CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil for granting scholarships.

References

[1] L.C.F. Silva, H.S. Lima, A. Sartoratto, MP de Sousa, A.P.R. Torres, RS de Souza, et al., Effect of salinity in heterotrophic nitrification/aerobic denitrification performed by acclimated microbiota from oil-produced water biological treatment system, Int. Biodeterior. Biodegrad. 130 (February) (2018) 1–7, https://doi.org/10.1016/j.ibiod.2018.03.009.

[2] X. Huang, W. Li, D. Zhang, W. Qin, Bioresource Technology Ammonium removal by a novel oligotrophic *Acinetobacter* sp. Y16 capable of heterotrophic nitrification – aerobic denitrification at low temperature, Bioresour Technol [Internet] 146 (2013) 44–50, https://doi.org/10.1016/j.biortech.2013.07.046.

[3] Y. Liu, G.M. Ai, M.R. Wu, S.S. Li, L.L. Miao, Z.P. Liu, *Photobacterium* sp. NNA4, an efficient hydroxylamine-transforming heterotrophic nitrifier/aerobic denitrifier, J Biosci Bioeng [Internet] 128 (1) (2019) 64–71, https://doi.org/10.1016/j.jbiosc.2018.12.014.

[4] B. Vanparys, P. Bodelier, P. De Vos, Validation of the correct start codon of norX/nxrX and universality of the norAXB/nxrAXB gene cluster in *Nitrobacter* species, Curr. Microbiol. 53 (3) (2006) 255–257, https://doi.org/10.1007/s00284-006-0161-z.

[5] R.F. Vazoller, L.P. Gomes, V.M.Z. Baldochi, D.M.F. Villas-Bôas, R.J. Badra, J. Povinelli, Biodegradability potential of two experimental landfills in Brazil, Brazilian J Microbiol 32 (2) (2001) 87–92, https://doi.org/10.1590/S1517-83822001000200003.

[6] B. Ji, H. Wang, K. Yang, Tolerance of an aerobic denitrifier (Pseudomonas stutzeri) to high O_2 concentrations, Biotechnol. Lett. 36 (2014) 719–722, doi.10.1007/s10529-013-1417-x.

[7] B. Kraft, M. Strous, H.E. Tegetmeyer, Microbial nitrate respiration – genes, enzymes, and environmental distribution, J Biotechnol [Internet] 155 (1) (2011) 104–117, https://doi.org/10.1016/j.jbiotec.2010.12.025.

[8] B.E. Rittmann, W.E. Langeland, Simultaneous denitrification with nitrification in single-channel oxidation ditches, J. Water Pollut. Control Fed. 57 (4) (1985) 300–308.

[9] P. Jin, Y. Chen, R. Yao, Z. Zheng, Q. Du, New insight into the nitrogen metabolism of simultaneous heterotrophic nitrification-aerobic denitrification bacterium in mRNA expression, J Hazard Mater [Internet] 371 (January) (2019) 295–303, https://doi.org/10.1016/j.jhazmat.2019.03.023.

[10] S. Sharma, H. Pathak, *Pseudomonas* in biodegradation, Int J Pure Appl Biosci. 2 (1) (2014) 213–222.

[11] C. Marazioti, M. Kornaros, G. Lyberatos, Kinetic modeling of a mixed culture of *Pseudomonas denitrificans* and *Bacillus subtilis* under aerobic and anoxic operating conditions, Water Res. 37 (2003) 1239–1251, https://doi.org/10.1016/S0043-1354(02)00463-3.

[12] Y. Zhao, J. Gao, S. Su, X. Shan, S. Li, H. Liu, et al., Regulation of the activity of maize glutamate dehydrogenase by ammonium and potassium, Biosci. Biotechnol. Biochem. 85 (2) (2020) 262–271, https://doi.org/10.1093/bbb/zbaa020.

[13] S.K. Padhi, S. Tripathy, S. Mohanty, N.K. Maiti, Aerobic and heterotrophic nitrogen removal by *Enterobacter cloacae* CF-S27 with efficient utilization of hydroxylamine, Bioresour Technol [Internet] 232 (2017) 285–296, https://doi.org/10.1016/j.biortech.2017.02.049.

[14] S.K. Padhi, S. Tripathy, R. Sen, A.S. Mahapatra, S. Mohanty, N.K. Maiti, Characterization of heterotrophic nitrifying and aerobic denitrifying *Klebsiella pneumoniae* CF-S9 strain for bioremediation of wastewater, Int Biodeterior Biodegrad [Internet] 78 (2013) 67–73, https://doi.org/10.1016/j.ibiod.2013.01.001.

[15] L. Quartaroli, C.M. Silva, L.C.F. Silva, H.S. Lima, S.O. Paula, R.S. Dias, et al., Effect of the gradual increase of salt on stability and microbial diversity of granular sludge and ammonia removal, J Environ Manage 248 (April) (2019), https://doi.org/10.1016/j.jenvman.2019.109273.

[16] L.C.F. Silva, H.S. Lima, TA. de O. Mendes, A. Sartoratto, M. de Paula Sousa, R.S. Souza, et al., Heterotrophic nitrifying/aerobic denitrifying bacteria: ammonium removal under different physical-chemical conditions and molecular characterization, J Environ Manage 248 (April) (2019), https://doi.org/10.1016/j.jenvman.2019.109294.

[17] F. Krueger, F. James, P. Ewels, E. Afyounian, B. Schuster-Boeckler, 30 FelixKrueger/TrimGalore: v0.6.7- DOI via Zenodo (2021) 26–37.

[18] A.M. Bolger, M. Lohse, B. Usadel, Genome analysis Trimmomatic : a flexible trimmer for Illumina sequence data 30 (15) (2014) 2114–2120, https://doi.org/10.1093/bioinformatics/btu170.

[19] A. Bankevich, S. Nurk, D. Antipov, Alexey A. Gurevich, D. Mikhail, Alexander S. Kulikov, Lesin Vm, Nikolenko Si, et al., SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, J. Comput. Biol. 19 (May 2014) (2012) 455–457, https://doi.org/10.1089/cmb.2012.0021.

[20] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform 25 (14) (2009) 1754–1760, https://doi.org/10.1093/bioinformatics/btp324.

[21] B.J. Walker, T. Abeel, T. Shea, M. Priest, A. Abouelliel, S. Sakthikumar, et al., Pilon : an integrated tool for comprehensive microbial variant detection and genome assembly improvement, PLoS One 9 (11) (2014), https://doi.org/10.1371/journal.pone.0112963.

[22] R.K. Aziz, D. Bartels, A.A. Best, M. Dejongh, T. Disz, R.A. Edwards, et al., The RAST Server : Rapid Annotations Using Subsystems Technology, vol. 15, 2008, pp. 1–15, https://doi.org/10.1186/1471-2164-9-75.

[23] R. Overbeek, T. Begley, R.M. Butler, J.V. Choudhuri, H. Chuang, M. Cohoon, et al., The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes, Nucleic Acids Res. 33 (17) (2005) 5691–5702, https://doi.org/10.1093/nar/gki866.

[24] J. Sun, F. Lu, Y. Luo, L. Bie, L. Xu, Y. Wang, OrthoVenn3: an integrated platform for exploring and visualizing orthologous data across genomes, Nucleic Acids Res. 51 (W1) (2023) W397–W403, https://doi.org/10.1093/nar/gkad313.

[25] L.C.F. Silva, H.S. Lima, TA. de O. Mendes, A. Sartoratto, M.P. Sousa, R.S. de Souza, et al., Physicochemical characterization of *Pseudomonas stutzeri* UFV5 and analysis of its transcriptome under heterotrophic nitrification/aerobic denitrification pathway induction condition, Sci. Rep. 10 (1) (2020) 1–13, https://doi.org/10.1038/s41598-020-59279-7.

[26] R.D. Finn, P. Coggill, R.Y. Eberhardt, S.R. Eddy, J. Mistry, A.L. Mitchell, et al., The Pfam protein families database : towards a more sustainable future, Nucleic Acids Res. 44 (December 2015) (2016) 279–285, https://doi.org/10.1093/nar/gkv1344.

[27] S. Sah, R. Singh, Phylogenetical coherence of Pseudomonas in unexplored soils of Himalayan region, 3 Biotech 6 (2) (2016) 1–10, https://doi.org/10.1007/s13205-016-0493-8.

[28] P. Chen, J. Li, Q.X. Li, Y. Wang, S. Li, T. Ren, et al., Simultaneous heterotrophic nitrification and aerobic denitrification by bacterium *Rhodococcus* sp. CPZ24, Bioresour. Technol. 116 (April 2008) (2012) 266–270, https://doi.org/10.1016/j.biortech.2012.02.050.

[29] N.H. Yao, Y.N. Du, J.X. Xiong, Y. Xiao, H.H. He, Z.F. Xie, et al., Microbial detoxification of 3,5-xylenol via a novel process with sequential methyl oxidation by *Rhodococcus* sp, CHJ602. Environ Res [Internet] 220 (68) (2023) 115258, https://doi.org/10.1016/j.envres.2023.115258.

[30] T.J. Silhavy, D. Kahne, S. Walker, The bacterial cell envelope1 T. J. Silhavy, D. Kahne and S. Walker, Cold Spring Harb Perspect Biol [Internet] 2 (2010) 1–16. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2857177/pdf/cshperspect-PRK-a000414.pdf.

[31] M.S.M. Jetten, M. Strous, K.T. Van De Pas-Schoonen, J. Schalk, U.G.J.M. Van Dongen, A.A. Van De Graaf, et al., The anaerobic oxidation of ammonium, FEMS Microbiol. Rev. 22 (5) (1998) 421–437, https://doi.org/10.1111/j.1574-6976.1998.tb00379.x.

[32] P. Jin, Y. Chen, R. Yao, Z. Zheng, Q. Du, New insight into the nitrogen metabolism of simultaneous heterotrophic nitrification-aerobic denitrification bacterium in mRNA expression, J Hazard Mater [Internet] 371 (December 2018) (2019) 295–303, https://doi.org/10.1016/j.jhazmat.2019.03.023.

[33] L.C.F. Silva, Nitrificação Heterotrófica/Desnitrificação Aeróbia: Caracterização de Isolados e Investigação das Vias Metabólicas, 2018.

[34] S. Maurer-stroh, N.J. Dickens, L. Hughes-davies, T. Kouzarides, F. Eisenhaber, C.P. Ponting, The tudor domain ' royal family ', TRENDS Biochem Sci. 28 (2) (2003) 69–74, https://doi.org/10.1016/S0968-0004(03)00004-5.

[35] J.C. Eissenberg, Molecular Biology of the Chromo Domain : an Ancient Chromatin Module Comes of Age, vol. 275, Elsevier, 2001, pp. 19–29, https://doi.org/10.1016/s0378-1119(01)00628-x.

[36] Z. Yu, S.E. Reichheld, A. Savchenko, J. Parkinson, A.R. Davidson, A comprehensive analysis of structural and sequence conservation in the TetR family transcriptional regulators, J Mol Biol [Internet] 400 (4) (2010) 847–864, https://doi.org/10.1016/j.jmb.2010.05.062.

[37] K. Okamura-ikeda, H. Hosaka, N. Maita, K. Fujiwara, A.C. Yoshizawa, A. Nakagawa, et al., Crystal structure of aminomethyltransferase in complex with dihydrolipoyl-H-protein of the Glycine cleavage system implications for recognition of lipoyl protein substrate, disease-related, J Biol Chem [Internet] 285 (24) (2010) 18684–18692, https://doi.org/10.1074/jbc.M110.110718.

[38] T.T. Tran, N.J. Bott, R Van Gelderen, N.T. Nguyen, Comparative genome analysis of two heterotrophic nitrifying *Pseudomonas putida* strains isolated from freshwater shrimp ponds in soc trang province, Fermentation 8 (2022) 336, https://doi.org/10.3390/fermentation8070336 [Internet].