

Original Research

DOI : <http://doi.org/10.22438/jeb/44/4/MRN-5061>

Decolorization of p-nitrophenol and draft genome sequence of *Pseudomonas* sp. strain PNPG3: A preliminary report

S.K. Aftabul Alam and P. Saha*^{ORCID}

Department of Microbiology, The University of Burdwan, Burdwan-713 104, India

*Corresponding Author Email : psaha@microbio.buruniv.ac.in*ORCID: <https://orcid.org/0000-0003-2113-3198>

Received: 01.06.2022

Revised: 24.08.2022

Accepted: 06.01.2023

Abstract

Aim: To study decolorization (with concomitant depletion) of p-nitrophenol by a bacterial strain designated as PNPG3 and determination of draft genome sequence of the strain to understand its potential.

Methodology: A comparative study of PNP's decolorization (with concomitant removal) in three different test conditions was undertaken. The experiment was carried out in one-liter volume Schott Duran bottles. Genomic DNA was extracted and draft genome sequence was determined using Illumina HiSeqX platform. Raw reads were assembled and subjected to subsystem classification using several bioinformatics tools (RAST, PATRIC, and NCBI's PGAP pipelines). The genome sequence was deposited at the NCBI Genome database and the strain PNPG3 was also deposited at MTCC, IMTECH, Chandigarh.

Results: The bacterial strain PNPG3 could carry out decolorization with concomitant removal of PNP in all three sets of experiments, including one set, where only distilled water was used. The best decolorization (with concomitant PNP removal) capacity was recorded for set D (with Minimal Medium, MM; PNP, and free cells) followed by set E (MM, PNP, and immobilized cells) and set B (distilled water, PNP and free cells) conditions. The size of the draft genome sequence of the strain PNPG3 was 6,566,321 bp, with 62.26% GC contents. The genome had 6210 protein-coding sequences and 66 tRNA genes. The predictive presence of different types of proteases and siderophore receptors indicated its possible potential for industrial applications and plant growth-promoting activities.

Interpretation: The bacterium *Pseudomonas* sp. strain PNPG3 has the capacity to decolorize p-nitrophenol even in presence of distilled water and it remains viable for up to twelve days. The genome sequence revealed that the strain harbored genes responsible for the metabolism of aromatic compounds, chemotaxis, protease, and siderophore receptors indicating the versatile nature of the strain.

Key words: Decolorization, P-nitrophenol, *Pseudomonas* sp., RAST, Xenobiotic

p-nitrophenol (PNP): a priority pollutant of all ecosystems



Pseudomonas sp strain PNPG3 was isolated from river Ganges, capable of decolorization of PNP



Decolorization based study to determine the efficacy of the strain



Report of draft genome sequence of strain PNPG3

Genome size-6,566,321bp, GC content- 62.26%,
6210 protein-coding sequences and 66 tRNA genes;
Proteases and siderophore receptors present

How to cite : Aftabul Alam, S.K. and P. Saha: Decolorization of p-nitrophenol and draft genome sequence of *Pseudomonas* sp. strain PNPG3: A preliminary report. *J. Environ. Biol.*, **44**, 578-586 (2023).

Introduction

PNP (*p*-nitrophenol) is a phenolic compound that is extensively used in synthetic chemistry with its application in the manufacturing of dye, explosive, pharmaceutical, pesticide, petrochemical, and leather industries (Prakash *et al.*, 1996). Over the past few decades, due to its wide application, persistent, toxic, recalcitrant nature, PNP is considered as one of the well-known priority pollutants by the United States Environmental protection agency (Alam and Saha, 2022b). It is reported to have wide environmental distribution (Nielsen *et al.*, 2011; Sengupta *et al.*, 2018). PNP is an uncoupler of oxidative phosphorylation and is reported to act as a mutagenic and carcinogenic agent (Kitagawa *et al.*, 2004; Arora, 2012). Moreover, this water-soluble, toxic compound is also produced as a hydrolysis product from extensively used organophosphate agrochemicals like parathion, methyl parathion, and herbicides like dinoseb and dinitroresol (Prakash *et al.*, 1996; Labana *et al.*, 2005; Sengupta *et al.*, 2018). The maximum permissible limit of PNP is documented to be 10ng ml⁻¹ in the natural water bodies (Kulkarni and Chaudhari, 2006).

Due to its potential toxicity and other hazardous effects, its presence in the environment (over permissible limit) is a of public concern. Therefore, rapid decontamination of this xenobiotic compound is desirable. Bioremediation of PNP by microorganisms is currently considered as the best possible pollution-free, cost-effective, and eco-friendly approach compared to physical, chemical, and physico-chemical methods (Labana *et al.*, 2005). Ganga is one of the major rivers of India, forming the lifeline of the North-Indian subcontinent (Ghirardelli *et al.*, 2021). The bank of this river supports hundreds of millions of human populations. Over the past few decades, rapid urbanization, population explosion, and the release of industrial wastes into the river water has converted the holy river Ganga into the fifth most polluted river in the world, (Jharia and Tiwari, 2020). Although, till date, no reports are available on PNP from Ganges water, its contamination reports from several other rivers, namely, the El Harrach river of Algeria; river Dene, England; river sediment in Argentina; Huai river, Hai river, and Yellow river of China; (Löser *et al.*, 1998; Kowalczyk *et al.*, 2015; Gemini *et al.*, 2005; Kuang *et al.*, 2020), has indicated possible contamination of river Ganges with PNP. This might be also due to the absence of studies addressing the monitoring of PNP from Ganga water, in literature. Recently, (Alam and Saha, 2022) have reported the presence of PNP-degrading bacteria in river Ganga, Hooghly, West Bengal.

Literature survey have revealed, several PNP-degrading bacteria like *Pseudomonas cepacia* strain RKJ200; *Serratia* sp. strain DS001; *Citriococcus nitrophenolicus* strain PNP1; *Rhodococcus* sp. BUPNP1 (Prakash *et al.*, 1996; Pakala *et al.*, 2007; Nielsen *et al.*, 2011; Sengupta *et al.*, 2018). However, these bacterias have been reported from soil ecosystems. In the Indian scenario, there is no authentic literature on PNP degrading aquatic bacterium, especially from the river Ganges except for the recently reported *Brachy bacterium* sp. strain DNPNG3 (Alam and Saha, 2022). Genomic information is considered the basis for all

phenotypic properties and today genome sequencing has become easier following the introduction of next-generation sequencing (NGS) approaches. To develop a better bioremediation strategy, it is important to understand the full potential and applicability of bacterial strains involved in the biodegradation process. The genome sequence provides a blueprint of all phenotypic potentials (including functional activities) possible with a given organism and thus, is of huge importance. In light of the above, study, a comparative study decolorization of PNP (in three experimental sets and two control sets with laboratory-based aquatic systems), as well as a draft genome sequence of the strain PNPNG3 was undertaken.

Materials and Methods

Selection of bacterium: A PNP decolorizing bacterium, *Pseudomonas* sp. strain PNPNG3 was isolated from a water sample collected from the river Ganga, Chinsura, Hooghly, West Bengal, India, by enrichment method on minimal medium supplemented with 0.5mM PNP as the sole source of carbon. Basal MM was solidified with agar, whenever required. The pH of the medium was adjusted to 7.0 (± 0.2) before sterilization by autoclaving. Growth of the strain on PNP-containing agar plates and its subsequent decolorization of yellow color were considered positive for the PNP utilization phenotype.

Comparative decolorization study in isolated, laboratory-based aquatic systems: To determine the usefulness of the strain PNPNG3 for decolorization (with concomitant removal) of PNP, five experimental setups were made. The content of these five sets were: Set A: one-liter ddH₂O + 0.5mM PNP; Set B: one-liter ddH₂O + 0.5 mM PNP + Strain PNPNG3; Set C: one-liter MM + 0.5mM PNP; Set D: one-liter MM + 0.5mM PNP + strain PNPNG3; Set E: one-liter MM + 0.5mM PNP + immobilized PNPNG3 within calcium-alginate beads. The 0.5mM PNP wherever added were previously filter sterilized (0.22μ pore size, PVDF syringe driven sterile filters from Millipore). Experimental sets B and D were inoculated with a fixed number of cells (3×10⁶ cells ml⁻¹). In order to determine the role of a suitable carrier material, immobilization was carried out with the strain PNPNG3 in calcium alginate beads as reported by Alam and Saha (2022a).

The set E was inoculated with 25 calcium alginate beads each containing 1.4×10⁶ bacterial cells (determined through a viable count experiment). All the experimental setups were incubated at room temperature (in the laboratory that varied from 30°C to 35°C) in a static condition, to simulate natural conditions. From each experimental setup, a requisite amount (2 to 5ml) of the solution was withdrawn aseptically and was processed for determination of depletion of PNP, viable count and the pH (by using Eutech Digital pH 700), after a regular time interval. Depletion in the amount of PNP was monitored by using a UV-Vis spectrophotometer (Shimadzu, Japan) at 410 nm after every 24 hrs of intervals. The percentage of PNP depletion was determined as reported by Sengupta *et al.*, 2015. All the experiments were performed in triplicates for statistically significant results.

Bacterial genomic DNA isolation, whole-genome sequencing, assembly, and annotation: Freshly grown culture (for overnight on TSB, at 28°C, 120 rpm shaking) was taken for extraction of genomic DNA following Pitcher *et al.*, (1989). The purified genomic DNA of strain PNPG3 was sequenced by Illumina HiSeqX with a read length of 151 bp. De novo metagenome assembly was carried out using the metaspades assembler (v3.11.1) with k-mer size 55 using the de-Bruijn graph method (Nurk *et al.*, 2017). The genome was annotated and analyzed using NCBI's PGAP (Tatusova *et al.*, 2016), RAST server (Aziz *et al.*, 2008), and PATRIC (Davis *et al.*, 2020). CGView server was used to visualize a graphical circular map of the entire genome of strain PNPG3 (Grant and Stothard, 2008). All the genes coding for exopolysaccharide, proteases, and siderophore receptor of *Pseudomonas* sp. strain PNPG3, were manually extracted and each of these sequences were used as queries in BLAST_X using the non-redundant protein database to determine the homologous sequences.

Accession number(s): Raw reads of genomic data and nucleotide sequence of PNPG3 strain were deposited to NCBI under accession no- SRR18163134 and JALLKV00 0000000 respectively. The strain PNPG3 was deposited to MTCC, IMTECH, Chandigarh (accession no-MTCC13126).

Statistical analyses: The statistical parameters tested were the correlation coefficient between time and percentage of PNP depletion for experimental sets B, D, and E. The same values were also calculated for the paired t-test.

Results and Discussion

Due to rapid urbanization and industrialization, most of the ecosystems are currently polluted with toxic xenobiotic compounds (Wen *et al.*, 2017). River Ganga is one of the major rivers of India, providing fresh water for consumption and also some source of livelihood for the human population living along its bank (Ghirardelli *et al.*, 2021). Since PNP contamination has been reported in different rivers of different countries, it is expected that PNP might have also contaminated Ganges water, as PNP is now considered as a widespread toxic chemical in the ecosystems. Although, there are reports on the presence of various groups of pesticides, phenolics, and heavy metals (beyond permissible limits), reports on PNP pollution in rivers Ganga totally absent. This may be due to lack of exploration studies in connection with specific PNP residues, in Ganga water. Pure water is now a priceless resource for sustaining life, and therefore, study on contamination and possible remedial activities is of outmost importance. Since the river Ganges happens to be the major source of drinking water for almost all cities of north India. Systematic studies towards the presence of PNP decolorizing bacteria were aimed. Their presence will provide possible indirect clues for the presence of PNP. Thus, attempts were made to isolate PNP decolorizing aquatic bacteria from the river Ganges in Chinsurah, Hooghly district, West Bengal, India. In the present study, although several bacterial strains were obtained by

enrichment culture technique using PNP as a sole source of carbon, *Pseudomonas* sp. strain PNPG3 was selected for further experimental work. A comparative study of decolorization (with concomitant removal) of PNP in isolated laboratory-based aquatic systems was carried out to assess the PNP depletion efficacy and efficiency at a moderately bigger setup. Also, to develop a better, effective PNP biodegradation strategy, parameters affecting growth conditions like pH of the medium was also studied. The comparative decolorization (with concomitant depletion of PNP) revealed that the test experimental sets carried out with MM, PNP, and free cells (Set D), decolorization occurred after 168 hrs and become was complete after 216 hrs. From Fig. 1a, it is evident that after 216 hrs, almost 99.09% PNP was depleted in set D. While, in Set E (test experiment with MM, PNP, and immobilised cells), decolorization occurred after 216 hrs and during this time approximately 98.86% PNP was depleted. With the experimental Set B, (containing ddH₂O and free cells), decolorization occurred after 240 hrs with 94.4% PNP depletion. Control experimental sets showed no decolorization and a very negligible amount of depletion in PNP (Set A- 2.8%; Set C- 4%) (Fig. 1a). These experimental findings indicated that with MM, decolorization (and PNP depletion) efficiency was best. However, with distilled water, the strain showed comparatively less decolorization, but the cells remained viable for longer duration of time (up to twelve days). Since it is an aquatic bacterial strain, it may have inherent adaptive features for survival and protection (maybe through compatible solute production), that might have triggered in distilled water set. This aspect needs to be explored further through experimental evidence.

The experimental results showed that decolorization of the medium occurred at pH 6.8. In the case of B experimental setup, at the onset of the experiment, the pH of the medium was 7.5 and it took 240 hrs to change the pH of the medium to 6.72 and the medium started to decolorize (94.4% PNP depletion). Whereas for the D and E experimental sets, it took 168 and 192 to change the pH of the medium to 6.77 (96.06% PNP depletion) and 6.79 (84.7% PNP depletion), respectively, showing decolorization. Labana *et al.* (2005), while performing a microcosm study with *Arthrobacter protophormiae* RKJ100, reported that complete depletion of PNP in the soil occurred at pH 7.5; at alkaline pH (9.5) complete depletion of PNP took pH (1.5) longer time while at acidic PNP depletion was minimum. In the control sets negligible amount of PNP depletion was observed (Fig. 1a). These result indicated that the rates of PNP depletion by the immobilized cells were slower than the free cells.

The depletion study was carried out at temperatures ranging between from 30°C-35°C, to simulate natural day conditions. Although, Leahy and Colwell (1990) reported that at a lower temperature, biodegradation of hydrocarbon was found to be slower as compared to higher temperatures (30°C-40°C). From the viable count experiment, it was observed that the number of viable cells increased initially and started declining after 144 hrs. To determine the statistical significant of decolorization (with concomitant removal of PNP) a correlation of

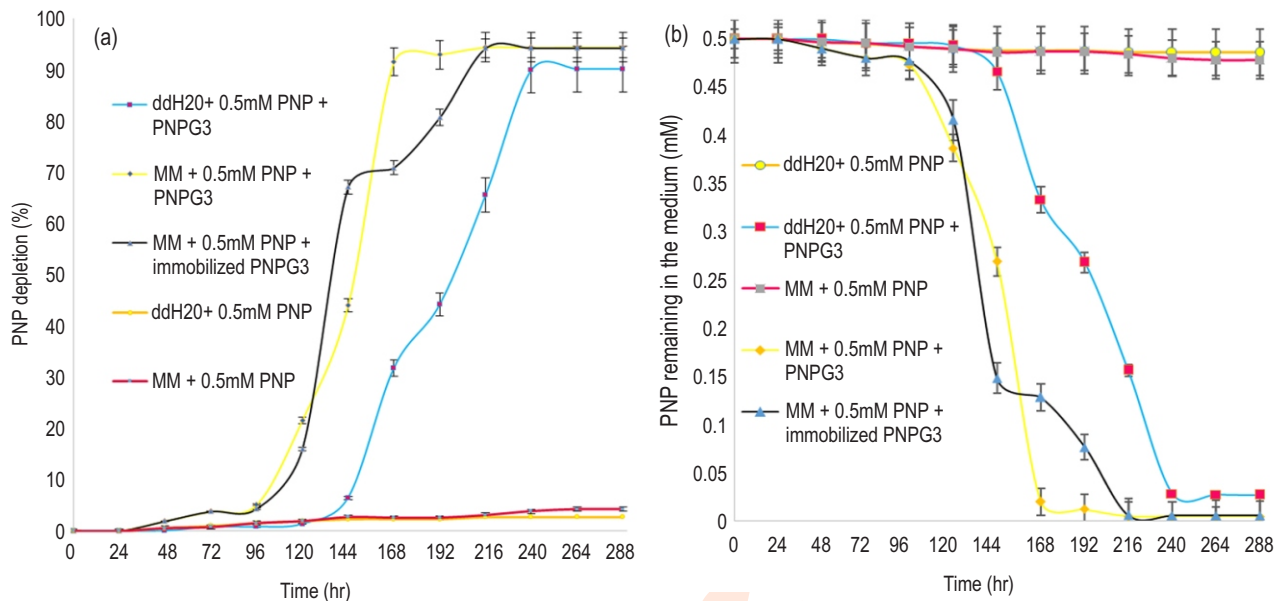


Fig.1: Comparative decolorization studies of p-nitrophenol (a) Percentage depletion and (b). Residual concentration.

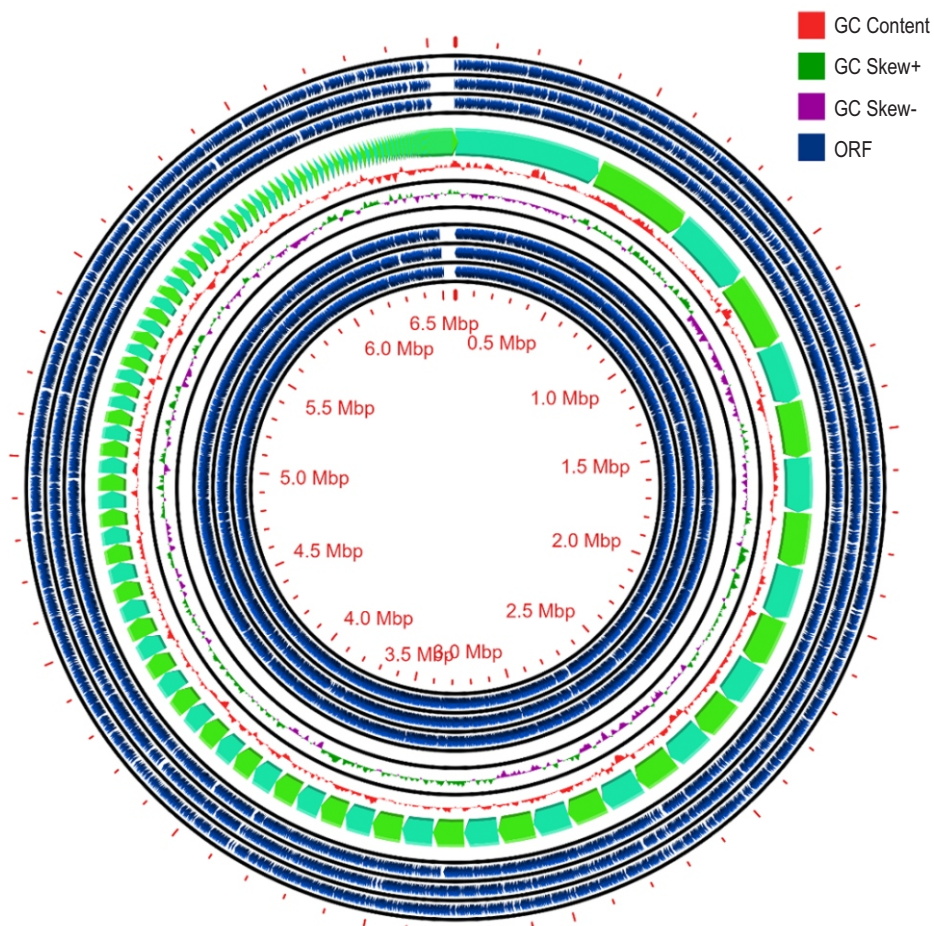


Fig. 2: *Pseudomonas* sp. strain PNPG3 circular view of genome with ORFs (blue color), positive GC skew (green), negative GC skew (violet), and GC content (black).

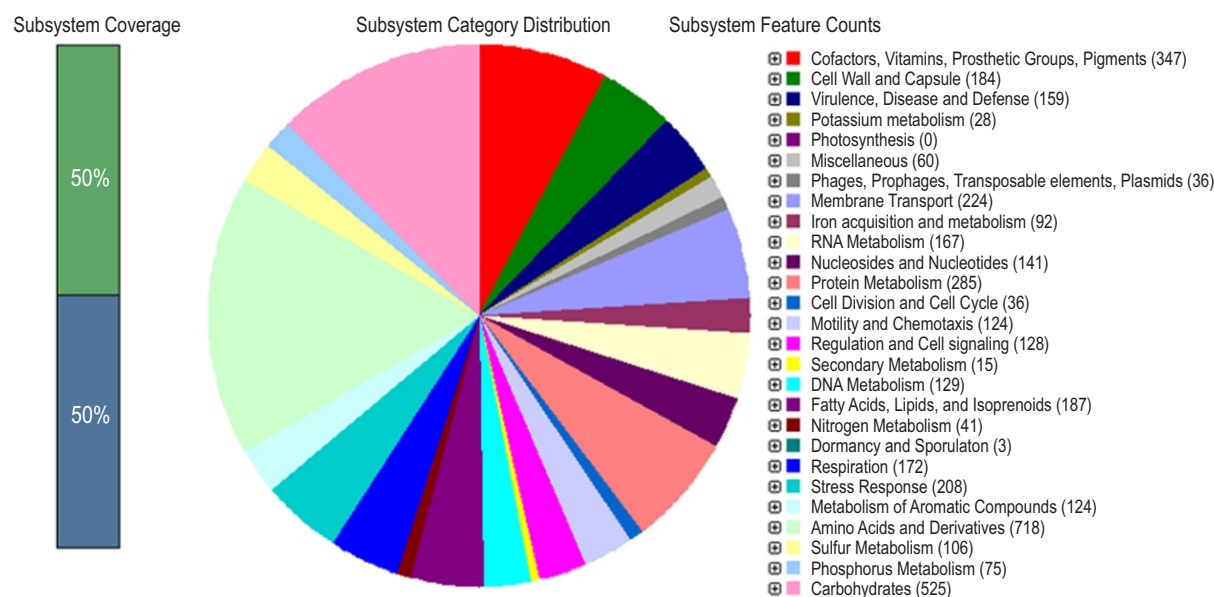


Fig. 3: RAST server-based annotation of the genome of *Pseudomonas* sp. strain PNPG3.

coefficient between time and PNP depletion was performed. The values of correlation of coefficient (r) of these experimental sets were 0.9187, 0.9263 and 0.9393, respectively. For all these values, paired t-test (p) were <0.00001 , indicating highly significant at $p < 0.05$. Genome sequence analysis revealed that the size of the genome was 6.5Mb with a contig number of 400.

The longest contig size was 449121bp, while the shortest contig size was 207bp. The average G+C content was 62.26%. A total of 6210 CDS were predicted to be present, with three 16S rRNA, sixty-six tRNAs, and four ncRNAs. Comparison of genomic parameters of strain PNPG3 with its close phylogenetic relatives are summarized in Table 1. Fig. 2 reveals the graphical map of the circular genome of strain PNPG3. The information is displayed as follows (outside to inside) - outer- (1,2,3 ring) and inner- (7,8,9 rings) are ORFs (blue colour) with forward and reverse strands respectively; ring 4 (aquarium green and asparagus alternate, thick ring)-contigs; ring 5 is GC content (red colour); ring 6 is GC skew (green and violet). The genome positions in Mbp are also shown inside ring 9. The annotation of *Pseudomonas* sp. strain PNPG3 genome using the RAST server predicted a total of 5941 coding sequences which were categorized into 538 subsystems, with 50% of subsystem coverage (Fig.3). RAST server predicted that amino acids and its derivatives contained the highest number of genes (718), followed by carbohydrates (525), cofactors, vitamins, prosthetic groups, pigments (347), protein metabolism (285), membrane transport (224), fatty acids, lipids, and isoprenoids (187), and cell wall and capsule (184). The subsystem category 'metabolism of aromatic compounds' contained 124 genes associated with the degradation of benzoate, p-hydroxybenzoate, chloroaromatic compound, n-phenylalkanoic acid, quinate, catechol, protocatechuate,

phenylacetyl-CoA, homogentisate, n-heterocyclic aromatic compounds, aromatic amine. A spectrum of genes involved in chemotaxis and motility (124) was identified, supporting the motile nature of the strain. A wide variety of amino acid degradation pathway genes indicated that the strain PNPG3 was genetically equipped with the capacity to utilize proteins and amino acids as vital nutrients for its growth and metabolism. The number of genes for the metabolism of mono, oligo, and polysaccharides were recorded as 91, 14, and 5, respectively.

The absence of genes for virulence, and toxins indicated the non-pathogenic nature of the strain. Based on the annotated data, it is speculated that the strain PNPG3 may be used for the degradation of various aromatic compounds. The distribution of genes in various subsystem categories for the genome of strain PNPG3, determined by the seed viewer of RAST, is presented in (Fig. 3). Whereas soil bacterium *Pseudomonas allopuntida* strain PNP harbored more or less similar types of coding DNA sequences (5726) which were categorised into 327 subsystems with 28% of subsystem coverage (Arora *et al.*, 2021). Although the PNP-degrader *Pseudomonas allopuntida* strain PNP (genome accession number- JAGKJH000000000) has a lower number of genes compared to strain PNPG3 for 'metabolism of aromatic compounds (80) and 'motility and chemotaxis (27).

Functional annotation by the RAST (Aziz *et al.*, 2008), NCBI's PGAP (Tatusova *et al.*, 2016) and PATRIC (Davis *et al.*, 2020) revealed the presence of xenobiotic degrading gene clusters and chemotaxis receptor proteins. Different types of proteases like zinc-protease, metalloprotease, serine protease, and aspartic proteases were present within the genome (Table 2). BLAST analysis revealed that most of the proteases were

Table 1: Comparison of genomic features of strain PNPG3 with its close relatives

General features	<i>Pseudomonas</i> sp. Strain PNPG3	<i>Pseudomonas</i> sp. 13159349	<i>Pseudomonas</i> sp. JS425	<i>Pseudomonas</i> sp. Sw144
Country of Isolation	India	China	China	USA
Accession no.	NZ_JALLKV000000000	NZ_CP045553.1	NZ_CP073661	Cp026674
Genome size	6.5Mb	5.9Mb	6.2Mb	5.9Mb
G+C content (%)	62.263	62.50	62.04	61.84
Total no. of CDS	6210	5622	5824	5453
5S rRNA Genes	2	8	8	8
16S rRNA genes	1	7	7	7
23S rRNA genes	1	7	7	7
tRNA	66	74	72	78
nc RNAs	4	4	4	4
Pseudogenes	119	228	195	147

Table 2: Different types of proteases present within the genome of strain PNPG3

Locus tag, PGAP annotation	Protein Product	Functions of protein	Homology	Score	E value	Origin (Accession no)
MJ643_01395 JALLKV010000001	MCK2119249	proteasome-type protease	100%	489	3e-174	<i>Pseudomonas</i> sp; (WP_015270103.1)
MJ643_03440 JALLKV010000003	MCK2119657	ATP-dependent zinc protease	100%	286	1e-97	<i>Pseudomonas</i> sp; (WP_230491514.1)
MJ643_03880 JALLKV010000003	MCK2119743	ATP-dependent zinc protease	100%	313	3e-107	<i>Pseudomonas</i> sp; WP_003257529.1)
MJ643_09170 JALLKV010000009	MCK2120768	ATP-dependent Clp protease adapter ClpS	100%	253	8e-85	<i>Pseudomonas</i> sp; (WP_208209150.1)
MJ643_09175 JALLKV010000009	MCK2120769	ATP-dependent Clp protease ATP-binding subunit ClpA	99.87%	1475	0.0	<i>Pseudomonas putida</i> ; (WP_054573026.1)
MJ643_09350 JALLKV010000009	MCK2120803	protease SohB	99.41%	557	0.0	<i>Pseudomonas</i> sp; (WP_019096524.1)
MJ643_13570 JALLKV010000017	MCK2121615	RIP metalloprotease RseP	100%	809	0.0	<i>Pseudomonas</i> sp; (WP_024086396.1)
MJ643_14090 JALLKV010000018	MCK2121718	metalloprotease PmbA	100%	885	0.0	<i>Pseudomonas</i> sp; (WP_024086335.1)
MJ643_14100 JALLKV010000018	MCK2121720	metalloprotease TldD	100%	912	0.0	<i>Pseudomonas</i> sp; (WP_208209337.1)
MJ643_14495 JALLKV010000019	MCK2121798	aspartic protease	100%	345	7e-120	<i>Pseudomonas</i> sp; (WP_024087816.1)
MJ643_14600 JALLKV010000019	MCK2121819	FtsH protease activity modulator HflK	99.75%	617	0.0	<i>Pseudomonas</i> sp; (WP_015272004.1)
MJ643_14605 JALLKV010000019	MCK2121820	protease modulator HflC	99.65%	461	1e-161	<i>Pseudomonas</i> sp; (WP_063911924.1)
MJ643_17820 JALLKV010000026	MCK2122446	ClpXP protease specificity-enhancing factor	100%	229	1e-74	<i>Pseudomonas</i> sp; (WP_208209567.1)
MJ643_18090 JALLKV010000027	MCK2122499	ATP-dependent Clp protease proteolytic subunit	100%	384	1e-134	<i>Pseudomonas</i> sp; (WP_208209471.1)
MJ643_19830 JALLKV010000033	MCK2122841	M48 family metalloprotease	99.78%	875	0.0	<i>Pseudomonas</i> sp ABFPK; (WP_061552683.1)
MJ643_22470 JALLKV010000042	MCK2123349	ATP-dependent Clp protease ATP-binding subunit ClpX	100%	812	0.0	<i>Pseudomonas taiwanensis</i> Sij9; (ESW40373.1)

Table continue

MJ643_22910 JALLKV010000044	MCK2123435	ATP-dependent zinc metalloprotease FtsH	100%	1219	0.0	<i>Pseudomonas</i> sp; (WP_015271861.1)
MJ643_23100 JALLKV010000045	MCK2123470	ATP-dependent protease subunit HslV	100%	292	1e-98	<i>Pseudomonas</i> sp; (WP_003249309.1)
MJ643_23105 JALLKV010000045	MCK2123471	ATP-dependent protease ATPase subunit HslU	100%	796	0.0	<i>Pseudomonas putida</i> S16; (AEJ15374.1)
MJ643_25980 JALLKV010000059	MCK2124030	rhomboid family intramembrane serine protease	100%	479	1e-168	<i>Pseudomonas</i> sp; (WP_038409144.1)
MJ643_26450 JALLKV010000062	MCK2124116	DegQ family serine endoprotease	100%	956	0.0	<i>Pseudomonas</i> sp; (WP_196184455.1)
MJ643_27490 JALLKV010000070	MCK2124317	Clp protease ClpP	98.25%	587	0.0	<i>Pseudomonas putida</i> ; (UPU90676.1)
MJ643_28115 JALLKV010000075	MCK2124440	Clp protease ClpP	99.65%	590	0.0	<i>Pseudomonas plecoglossicida</i> ; (WP_041506007.1)
MJ643_28910 JALLKV010000083	MCK2124596	CPBP family intramembrane metalloprotease	98.84%	276	1e-89	<i>Pseudomonas</i> sp; (WP_025338304.1)
MJ643_29280 JALLKV010000089	MCK2124668	protease HtpX	100%	597	0.0	<i>Pseudomonas</i> sp; (WP_023661073.1)
MJ643_30170 JALLKV010000112	MCK2124839	SprT family zinc-dependent metalloprotease	100%	311	3e-106	<i>Pseudomonas</i> sp; (WP_024086404.1)

homologous with protein sequences of *Pseudomonas*. All these genes coding for proteases were distributed within different contigs of genome. A total of 26 proteins were identified to contain the properties of protease. The clp protease is an ATP-dependent serine proteases, and Zinc protease is a type of metalloprotease. Within the genome, serine and metalloproteases were found in higher numbers (nine in each case), compared to aspartic protease (one). Serine and metalloproteases represent more than 69% of total proteases. Bacterial proteases are one of the largest groups of industrial enzymes and they are used in the leather, detergent, food, textile, and pharmaceutical industries (Hamza *et al.*, 2017). Besides, they are also used in the bioremediation of industrial, household and medical wastes (Razzaq *et al.*, 2019). Microbial proteins are damaged when a pathogen invades the host cell, therefore, it is presumed that proteases degrade such damaged proteins (Najar *et al.*, 2021). The study also revealed that *Parageobacillus* sp. strain SY1 harbor genes coding for ATP-dependent Clp protease, zinc protease (Najar *et al.*, 2021). Similarly, Santhakalakumari *et al.* (2021) reported that keratin-degrading *Bacillus paralicheniformis* MKU3 harbor genes coding for five different types of proteases like metallo, serine, cysteine, uncharacterized and mixed proteases. Serine and metalloproteases represent more than 70% of total proteases in strain MKU3.

The presence of genes of glycine betaine ABC transporter supported the halotolerant nature of strain PNPG3. In the SEED subsystem of the RAST server, a total of 40 genes classified as related to 'osmotic stress', were identified. Among these 5 genes involved in osmoregulation, 33 genes were

involved in choline and betaine uptake and betaine biosynthesis. Among them, there were seven glycine betaine transporters, which were involved in glycine betaine uptake. Similarly, Ahn *et al.* (2021) when they were working with *Lentibacillus* sp. CBA3610. The strain CBA3610 also contained one gene for osmoregulation, while thirty-three genes were engaged with choline and betaine uptake and betaine biosynthesis like strain PNPG3. Genome sequence analysis revealed that the genome of strain PNPG3 harboured 9 genes encoding TonB-dependent siderophore receptor, indicating the plant growth-promoting nature of the strain. Some other important functions like polysaccharide biosynthesis proteins were also present in the genome. Genome mining revealed that *Pseudomonas* sp. strain PNPG3 genome harboured forty-four genes coding for polysaccharide biosynthesis, out of which 30 genes were encoded for glycosyltransferase enzyme. Stingle *et al.* (1996) reported that the glycosyltransferase enzyme plays a role in transferring sugar to the undecaprenyl-phosphate lipid carrier as the first step in the synthesis of repeating unit in polysaccharide biosynthesis. Cui *et al.* (2016) observed a similar outcome in *Streptococcus thermophilus* strains.

All these phenotypic and genotypic information revealed that the strain PNPG3 is possess PNP decolorization activity in aquatic conditions and several possible industrial applications (proteases). However, a more detailed study on biodegradation using analytical techniques like HPLC and TLC should be carried out to understand and confirm its p-nitrophenol biodegrading potential as well as efficacy.

Acknowledgments

The authors are thankful to the University of Burdwan for providing research infrastructures. SAA is thankful to Dr. Urmimala Sen and Dr. Debdoot Gupta, Department of Microbiology, The University of Burdwan, for sharing lots of valuable scientific ideas related to WGS annotation.

Authors' contribution: S.K. Aftabul Alam: Isolated bacterial strain, performed experiments, analyzed data, prepared tables and/or figures, performed bioinformatics experiments and reviewed drafts of the paper; P. Saha: Conceived and designed the experiments, analyzed data, contributed reagents and reviewed drafts of the paper.

Funding: S.K. Aftabul Alam is grateful for the financial support (received during the initial phase of 1 year, from July 2018 to August 2019) from CSIR, Govt. of India (file no.—09/025(0263)/2019).

Research content: The research content of manuscript is original and has not been published elsewhere.

Ethical approval: The manuscript does not involve any work or studies with animal and/or human participants performed by any of the authors.

Conflict of interest: The authors declare that they do not have any conflict of interest connected to the manuscript.

Data availability: The genome sequence dataset is publicly available under accession number- JALLKV000000000.

Consent to publish: All authors agree to publish the paper in *Journal of Environmental Biology*.

References

- Ahn, S.W., S.H. Lee, H. Son, S.W. Roh and Y. Choi.: Genomic analysis of halophilic bacterium, *Lentibacillus* sp. CBA3610, derived from human feces. *Gut Pathogens*, **13**, 1-6 (2021).
- Alam, S. A. and P. Saha: Biodegradation of p-nitrophenol by a member of the genus *Brachybacterium*, isolated from the River Ganges. *3 Biotech*, **12**, 1-10 (2022).
- Alam, S. A. and P. Saha: Evidence of p-nitrophenol biodegradation and study of genomic attributes from a newly isolated aquatic bacterium *Pseudomonas asiatica* Strain PNP3. *Soil Sediment Contam*, 1-18 (2022).
- Arora, P.K., R.S. Saroj, R. Mishra, R.A. Omar, P. Kumari, A. Srivastava, S.K. Garg and V.P. Singh: Draft genome sequence data of a 4-nitrophenol-degrading bacterium, *Pseudomonas allopuntida* strain PNP. *Data Br.*, **38**, 107390 (2021).
- Arora, P.K.: Metabolism of para-nitrophenol in *Arthrobacter* sp. *SPG. E3 J. Environ. Res. Manag.*, **3**, 52-57 (2012).
- Aziz, R.K., D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, K. Formsma, S. Gerdes, E.M. Glass, M. Kubal, F. Meyer, G.J. Olsen, R. Olson, A.L. Osterman, R.A. Overbeek, L.K. McNeil, D. Paarmann, T. Paczian, B. Parrello, G.D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke and O. Zagnitko: The RAST Server: Rapid annotations using subsystems technology. *BMC Genom.*, **9**, 1-15 (2008).
- Cui Y., T. Xu, X. Qu, T. Hu, X. Jiang, and C. Zhao.: New insights into various production characteristics of *Streptococcus thermophilus* strains. *Int. J. Mol. Sci.*, **17**, 1701 (2016).
- Davis, J.J., A.R. Wattam, R.K. Aziz, T. Brettin, R. Butler, R.M. Butler, T. Brettin, R. Butler, R. M. Butler, P. Chlenski, N. Conrad, A. Dickerman, E. M. Dietrich, J. L. Gabbard, S. Gerdes, A. Guard, R. W. Kenyon, D. Machi, C. Mao, D. Murphy-Olson, M. Nguyen, E. K. Nordberg, G. J. Olsen, R. D. Olson, J. C. Overbeek, R. Overbeek, B. Parrello, G. D. Pusch, M. Shukla, C. Thomas, M. VanOeffelen, V. Vonstein, A. S. Warren, F. Xia, D. Xie, H. Yoo and R. Stevens: The PATRIC Bioinformatics Resource Center: Expanding data and analysis capabilities. *Nucl. Acids Res.*, **48**, 606-612 (2020).
- Gemini, V.L., A. Gallego, V.M. Oliveira, C.E. Gomez, G.P. Manfio and S.E. Korol: Biodegradation and detoxification of p-nitrophenol by *Rhodococcus wratislaviensis*. *Int. Biodeterior. Biodegr.*, **55**, 103-108 (2005).
- Ghirardelli, A., P. Tarolli, M.K. Rajasekaran, A. Mudhatkal, M.G. Macklin and R. Masin: Organic contaminants in Ganga basin: from the Green Revolution to the emerging concerns of modern India. *iScience*, **24**, 102122 (2021).
- Grant, J.R. and P. Stothard: The CGView Server: A comparative genomics tool for circular genomes. *Nucleic Acids Res.*, **36**, 181-184 (2008).
- Hamza, T.A.: Bacterial protease enzyme: Safe and good alternative for industrial and commercial use. *Int. J. Chem. Biomol. Sci.*, **3**, 1-10, (2017).
- Jhariya, D.C. and A.K. Tiwari: Ganga River: A paradox of purity and pollution in India due to unethical practice. *Earth Environ. Sci.*, **597**, 012023 (2020).
- Kitagawa, W., N. Kimura and Y. Kamagata: A novel p-nitrophenol degradation gene cluster from a gram-positive bacterium, *Rhodococcus opacus* SAO101. *J. Bacteriol.*, **186**, 4894-4902 (2004).
- Kowalczyk, A., Ö. Eyice, H. Schäfer, O.R. Price, C.J. Finnegan, R.A. Egmond, L.J. Shaw, G. Barrett and G.D. Bending: Characterization of para-nitrophenol-degrading bacterial communities in river water by using functional markers and stable isotope probing. *Appl. Environ. Microbiol.*, **81**, 6890-6900 (2015).
- Kuang, S., Q. Le, J. Hu, Y. Wang, N. Yu, X. Cao, M. Zhang, Y. Sun, W. Gu, Y. Yang, Y. Zhang, Y. Li, H. Liu and X. Yan: Effects of p-nitrophenol on enzyme activity, histology, and gene expression in *Larimichthys crocea*. *Comp. Biochem. Physiol.*, **228**, 108638 (2020).
- Kulkarni, M. and A. Chaudhari: Biodegradation of p-nitrophenol by *P. putida*. *Bioresour. Technol.*, **97**, 982-988 (2006).
- Labana, S., O. V. Singh, A. Basu, G. Pandey and R. K. Jain: A microcosm study on bioremediation of p-nitrophenol-contaminated soil using *Arthrobacter protophormiae* RKJ100. *Appl. Microbiol. Biotechnol.*, **68**, 417-424 (2005).
- Leahy, J.G. and R.R. Colwell: Microbial degradation of hydrocarbons in the environment. *Microbiol. Rev.*, **54**, 305-315 (1990).
- Löser, C., M.A. Oubelli and T. Hertel: Growth kinetics of the 4-nitrophenol degrading strain *Pseudomonas putida* PNP1. *Acta Biotechnol.*, **18**, 29-41 (1998).
- Najar, I.N., M.T. Sherpa, S. Das and N. Thakur: The draft genome sequence of *Parageobacillus* sp. strain SY1 gives insights into its physiological properties and protease production. *Meta Gene*, **29**, 100894 (2021).
- Nielsen, M.B., K.U. Kjeldsen and K. Ingvorsen: Description of *Citricoccus*

- nitrophenolicus* sp. nov., a para-nitrophenol degrading actinobacterium isolated from a wastewater treatment plant and emended description of the genus *Citricoccus* Altenburger et al. 2002. *Anton. V. Leeu.*, **99**, 489-499 (2011).
- Nurk, S., D. Meleshko, A. Korobeynikov and P.A. Pevzner: meta SPAdes: A new versatile metagenomic assembler. *Genome Res.*, **27**, 824-834 (2017).
- Pakala, S.B., P. Gorla, A.B. Pinjari, R.K. Krovidi, R. Baru, M. Yanamandra, M. Merrick and D. Siddavattam: Biodegradation of methyl parathion and p-nitrophenol: evidence for the presence of a p-nitrophenol 2-hydroxylase in a Gram-negative *Serratia* sp. strain DS001. *Appl. Microbiol. Biotechnol.*, **73**, 1452-1462 (2007).
- Pitcher, D.G., N.A. Saunders and R.J. Owen: Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. *Lett. Appl. Microbiol.*, **8**, 151-156 (1989).
- Prakash, D., A. Chauhan and R.K. Jain: Plasmid-encoded degradation of p-nitrophenol by *Pseudomonas cepacia*. *Biochem. Biophys. Res. Comm.*, **224**, 375-381 (1996).
- Razzaq, A., S. Shamsi, A. Ali, Q. Ali, M. Sajjad, A. Malik and M. Ashraf: Microbial proteases applications. *Front. Bioeng. Biotechnol.*, **7**, 110, (2019).
- SanthaKalaikumari, S., R. Sivakumar, P. Gunasekaran and J. Rajendhran: Whole-genome sequencing and mining of protease coding genes in *Bacillus paralicheniformis* Mku3, and its degradomics in feather meal medium. *Curr. Microbiol.*, **78**, 206-217 (2021).
- Sengupta, K., M. Alam, S. Pailan and P. Saha: Biodegradation of 4-nitrophenol by a *Rhodococcus* species and a preliminary insight into its toxicoproteome based on mass spectrometry analysis. *J. Environ. Biol.*, **40**, 356-362 (2018).
- Sengupta, K., T.K. Maiti and P. Saha: Degradation of 4-nitrophenol in presence of heavy metals by a halotolerant *Bacillus* sp. strain BUPNP2, having plant growth promoting traits. *Symbiosis*, **65**, 157-163 (2015).
- Stingele F., J. Neeser and B. Mollet.: Identification and characterization of the eps (exopolysaccharide) gene cluster from *Streptococcus thermophilus* Sfi6. *J. Bacteriol.*, **178**, 1680-1690 (1996).
- Tatusova, T., M. DiCuccio, A. Badretdin, V. Chetvernin, E.P. Nawrocki, L. Zaslavsky, A. Lomsadze, K.D. Pruitt, M. Borodovsk and J. Ostell: NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.*, **44**, 6614-6624 (2016).
- Wen, Y., G. Schoups and N.V. Giesen: Organic pollution of rivers: Combined threats of urbanization, livestock farming and global climate change. *Sci. Rep.*, **7**, 1-9 (2017).