Comparative genomics reveals the acquisition of a novel transposon Tn.*arsmer*S12 by the plant growth-promoting *Pantoea eucrina* OB49 in polluted environments

Abdelmalek Lekired[1], Hafsa Cherif-Silini[2], Allaoua Silini[2], Hamza Ben Yahia[1], and Imene-Hadda Ouzari[1*]

[1]Laboratory of Microorganisms and Active Biomolecules, MBA-LR03ES03, Faculty of Sciences of Tunis, University of Tunis El Manar, Tunis, Tunisia.

[2]Laboratory of Applied Microbiology, Department of Microbiology, Faculty of Natural and Life Sciences, Ferhat Abbas University, 19000, Setif, Algeria.

* Corresponding author: OUZARI Hadda-Imene: ouzari.imene@gmail.com / imene.ouzari@fst.utm.tn

Abstract

Heavy metal-tolerant plant growth-promoting bacteria (PGPB) have gained popularity in bioremediation in recent years. A genome-assisted study of a heavy metal-tolerant PGPB *Pantoea eucrina* OB49 isolated from the rhizosphere of wheat grown on a heavy metal-contaminated site is presented. Comparative pan-genome analysis indicated that OB49 acquired heavy metal resistance genes through horizontal gene transfer. On contigs S10 and S12, OB49 has two *arsRBCH* operons that give arsenic resistance. On the S12 contig, an *arsRBCH* operon was discovered in conjunction with the *merRTPCADE* operon, which provides mercury resistance. OB49 contains singleton proteins that are found on mobile genetic elements (MGEs) that are involved in resistance to salinity, osmotic stress, silver, cadmium, copper, arsenic, and mercury. In conclusion, *P. eucrina* OB49, which has PGP properties and acquires MGEs to adapt to a polluted environment, might be used as a biofertilizer in contaminated soils for bioremediation.

Keywords: *Pantoea eucrina*, Heavy metals, Genome comparison, ars operon, Transposon, Bioremediation

1. Introduction

The contamination of the aquatic and terrestrial environment by heavy metals and xenobiotic pollutants is highly increasing in the world, especially in the last decade due to industrial development and anthropogenic activity [1]**.** Heavy metals are among the toxic elements that can enter the food chain by absorption at the plant roots by a process called translocation [2]**.** Plant growth-promoting bacteria (PGPB) that have the potential to tolerate high levels of heavy metals are used in bioremediation [3]. Therefore, detoxification of soils irrigated by heavy metal contaminated wastewater using heavy metal tolerant PGP bacteria is a necessity as an application that should take place in bioremediation processes. PGPBs are generally free-living soil bacteria, although they can also be symbiotic with plant roots or endophytic [4]. The most used group known for their PGP potential are *Actinobacteria* [5], *Rhizobia* [6], *Pseudomonads* [7], *Bacillus* [8,9], and *Enterobacteria*, which include the *Pantoea* genus [10]. *Pantoea* is a genus in the *Erwiniaceae* family that was introduced by Gavini et al. [11] where *Pantoea agglomerans* as the type species. The name has since been expanded as other species have been recognized [12]. We noticed *Pantoea agglomearans* as the most studied and used PGP species [13–16]. They are known for their potentiality for phosphate solubilization [17], nitrogen fixation, production of secondary metabolites, siderophores, and biomolecules useful in biocontrol [18]. Horizontal gene transfer (HGT) allows bacteria to incorporate advantageous genes into their genomes, which is crucial for bacterial adaptation to changing environments. Horizontally transmitted genes are often grouped into genomic islands, which are required for adaptation to varied habitats. Several mechanisms can promote bacterial adaptive evolution, including horizontally acquired genes, gene duplication-amplification processes, and interactions among various species within communities [19]. A pan-genome is a complete collection of genes from different strains of a species. Moreover, HGT causes substantial variability in the genes that each strain possesses, and pan-genome studies are commonly employed to examine species evolution and changes in gene function. Pan-genome analyses identify a group of core genes shared by all strains, accessory genes shared by two or more strains, and unique genes unique to a single strain [20,21]. Bacteria can generally integrate a variety of beneficial genes into their genomes to adapt to varied habitats, resulting in substantial intraspecific genome diversity and genetic redundancy. These accessory genes have a crucial role in responsiveness to environmental change as part of adaptive evolution. [22]. Using the next-generation sequencing technology, researchers are increasingly moving toward whole-genome sequencing. This technique has enabled scientists to provide comprehensive insights into the genetic determinants and metabolic pathways involved in various life processes and adaptations [23,24]. Many mysteries have been answered as a result of genome annotation and proteogenomic comparisons with other organisms. In this study, *Pantoea eucrina* OB49 was identified from wheat cultivated on a site polluted by mixed effluents in Algeria's Setif province. *In vitro,* OB49 demonstrated great resistance to arsenic, zinc, copper, chromium, and cadmium, as well as the ability to solubilize phosphate in both solid and liquid media and to synthesize siderophores. Furthermore, in the presence of zinc and cadmium, this strain increased shoot length in the seedling test.

This work includes *in-vitro* and *in-silico* genome analysis of *Pantoea eucrina* OB49, emphasizing genetic factors such as mercury and arsenic genes and operons, as well as processes involved in heavy metals tolerance and PGP characteristics. The high tolerance and adaptability of OB49 were explored utilizing Orthovenn2 and Pan-genome analysis, and Multigeneblast, which allows us to determine the proteins involved in tolerance and alleviating abiotic stress, particularly for mercury and arsenic. These findings prompted us to investigate the relevance of mobile genetic elements (MGEs) in this adaptation. When comparing the proteins of the *Pantoea* genus, it was found that the *P. agglomerans* strain P5 [25], which is commercially used as a biofertilizer is the most similar strain to the OB49 strain in terms of protein clusters. Based on these findings, OB49 might be a prospective bioremediation alternative for heavy-metal-contaminated soils and wastewaters, as well as a biofertilizer candidate.

2. Material and methods

2.1 Isolation of bacteria

Pantoea eucrina OB49 strain was isolated from rhizospheric soil of wheat (*Triticum durum*) cultivated in soil irrigated with industrial and municipal wastewater in Setif (36.0972 N, 5.383 E), Algeria. Physico-chemical proprieties of soil are presented in **(**Table S1**).** Serial dilutions of 1g of rhizospheric soil were carried out in a saline solution (0.85% NaCl (w/v)) and 1 ml of each dilution was spread on Petri dishes of Tryptic Soy agar medium (TSA, BioLife, Italy), then morphologically distinctive colonies were selected for further application.

2.2 Heavy metal tolerance test

In solid medium: To determine the ability of OB49 to tolerate cadmium, copper, arsenic, zinc, and chromium, inoculums of 10 µl of fresh bacterial culture were applied to TSA medium supplemented with increasing concentrations of heavy metals and incubated for 48h at 30C°. Serial two-fold dilutions were prepared from stock solutions of heavy metals. Cadmium chloride ($CdCl_2$) concentrations ranged from 0 mg l^{-1} to 200 mg l^{-1}, Copper chloride ($CuCl_2$), Zinc sulfate ($ZnSO_4$), and potassium dichromate ($K_2Cr_2O_7$) from 0 to 1600 mg l^{-1}, and disodium hydrogen-arsenate (Na_2HAsO_4) from 0 to 8000 mg l^{-1}.

In liquid medium: Heavy metal bacterial tolerance was carried out in a liquid medium supplemented with appropriate concentrations of chromium, copper, zinc, and cadmium. Serial dilutions were carried out from stock solution ranging from (0 to 3200 mg l^{-1}) for Potassium dichromate $K_2Cr_2O_7$, Copper chloride ($CuCl_2$), and Zinc sulfate ($ZnSO_4$), and from (0 to 320 mg l^{-1}) for Cadmium chloride ($CdCl_2$) in microplates. Cultures were incubated for 24h at 30C° and optical density at 600 nm was measured. Tests were applied in triplicates.

2.3 Seedlings under heavy metal stress

For surface sterilization, wheat seeds were washed and disinfected with 70% ethanol for 15 seconds, then with a 5% sodium hypochlorite solution for 15 minutes, followed by several rinses with distilled water and settled in Petri dishes containing wet blotting paper with gradually increasing concentrations of $CdCl_2$ (0–150 mg l^{-1}) and $ZnSO_4$ (0-3200 mg l^{-1}). The inoculated seeds were tested under zinc and cadmium ED concentrations. The Petri dishes were then placed in the dark for five days at 25°C.

2.4 DNA extraction, library preparation, and genome sequencing

DNA of Gram-negative bacteria was extracted using the MagCore Genomic DNA Bacterial kit (RBC Biosciences Corp., Taïwan). The quality and quantity of DNA were verified using the Qubit dsDNA HS Assay kit (Invitrogen) and NanoDrop 2000c UV–Vis Spectrophotometer (Thermo Scientific, USA). The library was prepared using the DNA flex prep kit from Illumina according to the manufacturer's recommendations. Briefly, 200 ng of high-quality DNA was used for tagmentation. Then, dual indexation (i5-i7) was ligated to samples. The final library (4 nM) was quantified and denatured with NaOH, and sequenced on the Illumina Miseq instrument (Illumina, San Diego, CA, USA) using 2×250 paired-end reads).

2.5 Genome assembly, annotation, and phylogeny

Sequencing quality control was performed for reads with an average size of 250 bp using FastQC (version 0.11.8) and poor quality reads were trimmed. *De Novo* assembly was carried out using the Spades-3.13.1 tool [26]**.** The genome was

annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). ContEst16S (Contamination Estimator by 16S) was used to assess for contamination and retrieve the 16s rDNA fragment from the draft genome [27]. The Whole Genome Shotgun project *Pantoea eucrina* strain OB49 has been deposited at DDBJ/ENA/GenBank under the accession JAFCXS000000000. The version described in this paper is version JAFCXS010000000. The BioProject accession number is PRJNA690006.

2.6 Whole-Genome Average Nucleotide Identity Analysis

The Average Nucleotide Identity (ANI) matrix was calculated with the standalone OrthoANIu tool using a modified algorithm by Yoon et al., [28] to measure the overall genome relatedness between OB49 and the other *Pantoea* genomes. An ANI value of 95% was set as the cut-off for species demarcation.

2.7 Pan-Genome Construction

The *Pantoea* genomes assigned as representative genomes retrieved from NCBI were re-annotated to obtain *gff3* files format on the Galaxy server using Prokka v.1.13.3 [29]. The pan-genome was computed using the default settings in Roary v.3.11.2 [30]. As calculated in Roary, the *Pantoea* pan-genome was divided into the following categories: core (gene families present in 99–100% of the genomes); soft-core (gene families present in 95–99% of the genomes); shell (gene families present in 15–95% of the genomes); cloud genes.

2.8 Orthologous gene clusters with OrthoVenn2

The OrthoVenn2 web server [31] was used to compare genomes and identify orthologous gene clusters based on coding sequences (CDSs). The orthologous clusters were identified using the default parameters, with a 1^{e-5} e-value cut-off for all protein similarity comparisons and a 1.5 inflation value for the generation of orthologous clusters. PATRIC server (3.9.6) was used for genome annotation analysis [32].

2.9 Mobile genetic elements analysis

The MobileElementFinder tool [33] was used to detect mobile genetic elements (MGEs). MultiGeneBLAST [34] was used to blast unmapped sequences against Genbank Database, The database was built by downloading over 50 Gigabytes of data, including 688 gbbct files (gbbct1 to gbbct99) of complete bacterial sequences/loci from the Genbank Bacterial Division (December 2021). (NCBI). The specified region of the contig S12 containing the arsenic and mercury operons was used as the query sequence against the constructed database for homologous cluster sequence similarity at the level of the entire gene cluster with a 30% sequence identify cut-off and 250 blastp hits mapped per query sequence using the MultiGeneBLAST program.

2.10 Statistical analysis

The results were statistically processed using the R software (R.4.0). ANOVA one-way test was used and when significant effects are detected, the groups were compared using a Post-hoc Tukey's HSD test using the *agricolae* r package. The level of significance used for all statistical tests is 5% ($p<0.05$). The toxicity test (dose-response curve) was performed by the *DRC* package [35]; the graphs were created with ggplot2 on RStudio.

3. Results

3.1 Heavy metal tolerance and minimal inhibitory concentration (MIC)

Pantoea eucrina OB49 exhibits high tolerance to heavy metals, with resistance to arsenic in solid medium reaching 6000 mg l^{-1}. For copper, zinc, and chromium, tolerance level was up to 400 mg l^{-1}, and 25 mg l^{-1} for cadmium. However, in liquid media, (**Fig. 1**) The ANOVA test combined with Tukey's HSD test revealed no significant effect of zinc and copper on the growth of OB49 up to 3200 mg l^{-1} and 800 mg l^{-1}, respectively. In contrast, the effect of chromium was significant up to 400 mg l^{-1}, while the effect of cadmium was observed up to 80 mg l^{-1}. (Table S2**).**

Fig. 1 here

3.2 Heavy metal's effective dose curves

The dose-response curves were created using the *DRC* package of the effective dose with four parameters (Slope, Lower Limit, Upper Limit, and ED_{50}) of Copper (**Fig. 2, a**), Zinc (**Fig. 2, b**), Cadmium (**Fig. 2, c)**, and Chromium (**Fig. 2, d**) on the growth of OB49 showed that heavy metal toxicity has a variable effect. The logarithm of the heavy metal dose is plotted on the X-axis and the response (optical density) is plotted on the Y-axis. OB49 is more sensitive to cadmium than other heavy metals, with the concentration of $CdCl_2$ that inhibits the growth of 50% of bacterial cells (ED_{50}) being 72.65 mg l^{-1} and ED_{95} being 149.23 mg l^{-1}. The ED_{50} for zinc, copper, and chromium, on the other hand, are 3293 mg l^{-1}, 1618 mg l^{-1}, and 1118 mg l^{-1}, respectively. Furthermore, zinc and chromium have a linear response in which the optical density decreases with increasing heavy metal concentrations. However, the effect of Cd and Cu plots a sigmoidal curve when ED_{50} is defined as the inflection point of the curve. The lower and upper parameters are shown in (Table S3**).**

Fig. 2 here

3.3 Seedlings under heavy metal stress

In the presence of cadmium and zinc, the growth-promoting potential of strain OB49 was also confirmed in vivo by measuring the shoot and root length of wheat seedlings in the dark for 5 days and in the light for up to 15 days. The wheat shoot length of inoculated seeds was increased by 15.64%, 15.79%, and 4.18% under cadmium, zinc, and control conditions, respectively. (**Fig. 3a**). In contrast, there was a lack of a significant effect of OB49 on root length in all treatments (**Fig. 3b**). This result might be related to the quantity of auxin produced by OB49 in the presence of these heavy metals (**Table 1**).

Table 1 here

Fig. 3 here

3.4 Genome Features and Phylogeny

The *Pantoea eucrina* OB49 genome was assembled and found to be 4.039986 bp in length, distributed over 51 scaffolds, with an N50 of 244935 bp. The NCBI Prokaryotic genome annotation pipeline (PGAP) feature annotation produced a total of 3854 genes, with 3757 CDSs, 3701 coding genes, and 97 RNA genes distributed as 5, 4, 3, (5s, 16s, 23s) (Table S4). The distribution of genome annotation features is shown in **(Fig. S1)** including, from outside to inside, contigs, CDS on the forward strand, and RNA genes [36]. The colors indicate the different subsystems corresponding to the identified genes **(Fig. S2)**.

The 16s-based phylogenetic tree shows that the OB49 isolate belongs to *Pantoea eucrina*. Using the RAxML program, the Codon Tree approach selects single-copy PATRIC Cross-genus families (PGfams) and analyzes aligned proteins and DNA encoding single-copy genes. The codon tree showed that the OB49 isolate belongs to the species *Pantoea eucrina*

(Fig. 4). These results are confirmed by the OrthoANI matrix of 11 most related species of the genus *Pantoea* that were selected using the dDDH method and showed that *Pantoea eucrina* XL123 is the most related species to OB49 with an OrthoANI value of 99.83%.

Fig. 4 here

3.5 Genomes comparison using CGview server

The genomes used for the comparative analysis were retrieved from The National Center for Biotechnology Information (NCBI) based on their relatedness to *P. eucrina* and their PGP potential. The comparison was done using Circular genome visualization (CGview server) [37]**.** The genomes from the outside to the inside are *Pantoea eucrina* strain OB49 (JAFCXS010000000), *P. eucrina* strain xl123 (CP083448), *P. eucrina* strain LMG 5346 (NZ_MIPP01000000), *P. vagans* strain Pa (NZ_MUJJ01000000), *Pantoea* sp. Strain 3.5.1 (NZ_JMRT02000000), *P. eucrina* strain Russ (NZ_MAYN01000000), *P. ananatis* strainYJ76 (NZ_CP022427), *P. ananatis* strain NN08200 (CP035034), *P. agglomerans* strain P5 (NZ_NGNU02000000), *P. agglomerans* strain 190 (JNGC00000000), and *P. agglomerans* strain 3 (NZ_LVHW00000000). **(Fig. 5).**

Fig. 5 here

Comparison of OB49 genome features with other *Pantoea* species using the CGview tool revealed the existence of unaligned regions distributed across the genome; these regions are mainly distributed on contigs JAFCXS010000010.1 (S10), JAFCXS010000012.1 (S12), JAFCXS010000015.1 (S15) and JAFCXS010000024.1 (S24).

3.6 Pan-genome analysis

A pan-genome study was undertaken to further explore the chosen *Pantoea* genomes. The *Pantoea* pan-genome has 72083 gene orthologs, of which 63 are shared by the core (28) and soft-core (26) genomes, respectively, while 3538 and 68425 clusters constitute the shell and cloud genomes. Figure 6 depicts the effect of adding a new genome on the total pan-genome size. The curve rises with each new genome added, yet the number of genomes investigated does not reach a plateau. It shows also the effect of adding genomes on the number of new genes in the pan-genome (Fig. 6a), the number of the unique genes (Fig. 6b), the number of conserved genes against the number of total genes (Fig. 6c), and the unique genes against the number of genes (Fig. 6d). These findings suggest that the *Pantoea* pan-genome is open.

Fig. 6 here

3.7 COG comparison of *Pantoea eucrina* OB49 to other *Pantoea* species

The comparison of orthologous cluster groups reveals that *Pantoea* PGPB species form 5361 clusters, 2737 orthologous clusters (at least two species), and 2624 single-copy gene clusters. Whereas *Pantoea eucrina* OB49 has the fewest proteins (3666) and clusters (3320) compared to all species, it has the second-highest number of singletons (374), trailing *Pantoea vagans* 3.5.1, which has 473 singletons **(Table 2).**

There are 2699 clusters shared by all species. *P. eucrina* OB49, on the other hand, has 122 distinct clusters, each of which contains 132 proteins. According to GO enrichment annotation, biological process, and metabolic process are present in 14 percent of the cases, followed by the cellular metabolic process at 13%, nitrogen compound metabolic process and

cellular process at 9%, and macromolecule metabolic process, heterocycle metabolic process, cellular aromatic compound metabolic process, primary metabolic process, and nucleobase-containing compound metabolic process at 8 % or less.

Table 2 here

3.8 Singletons clusters of *P. eucrina* OB49 and PGPB *Pantoea* strains

The comparison of clusters reveals that *Pantoea eucrina* OB49 has proteins involved in resistance to various abiotic stresses such as salinity (Glycine betaine transporter), osmotic stress protection (Probable metabolite transport protein *CsbC*), silver resistance (Probable transcriptional regulatory protein *SilR*), cadmium resistance (Probable cadmium-transporting ATPase), and phosphate metabolism (Alkaline phosphatase *phoA*). As well as proteins involved in UV protection and mutation (MucA, MucB).

Bacteriophage *P1* and *P2* proteins are also present in the genome. These proteins are involved in encapsidation (Probable terminase, large subunit *pacB*) which prevents degradation of viral DNA by the host *EcoB* and *EcoK* restriction-modification antiviral defense systems (Defense against restriction protein A and B), and enhances IS1 element recombination (Recombination enhancement function protein, Recombinase *cre*), and an Endolysin K with transglycosylase activity that degrades host peptidoglycans which lead to the programmed host cell lysis releasing the mature viral particles. In addition, proteins involved in recombination, transposable element insertion, and plasmid stabilization have been identified; these proteins include (Insertion element iso-IS1d protein InsA, tyrosine recombinase XerD, type-2 restriction enzyme EcoRII, transposase InsF for insertion sequence *IS3A*, transposon Tn2501 resolvase *tnpR*, recombination-associated protein *RdgC*, transposase for transposon Tn1721, putative transposase y4zB, resolvase, and regulatory protein rop).

3.9 Transposons detection and identification

We performed an extensive analysis of mobile genetic elements (MGEs) in the *Pantoea eucrina* OB49 genome. We found that OB49 carries a plethora of transposons of different types; some of them were identified as IS3 family transposases, where, Tn3 and IS1 are detected on contigs S10 and S24. While, the transposase family such as Tn3-like element Tn3, IS4-like element ISApu2, IS110-like element IS4321, Tn3-like element TnAs1, IS4-like element ISApu1 are detected on the contig S12. These transposons harbor genes that encode for tolerance to antibiotics and heavy metals, including class A broad-spectrum beta-lactamase TEM-1 (bla_{TEM-1B}), arsenic, mercury, tellurium, silver, and copper as shown in **(**Table S5**)**. Using the MobileElementFinder tool [33], the contig S12 has four MGEs identified as Tn2 transposon of type Unit transposon and located at position (32966-37915) with a length of 4950 bp in the reverse strand with 99.82% sequence identity and 9 substitutions. This transposon harbored the bla_{TEM-1B} gene that codes for phenotypic resistance to amoxicillin, ampicillin, ticarcillin, piperacillin, and cephalothin. The second transposon was the ISApu1 transposon from the IS4 family and ISH8 group, it was identified as a composite transposon and situated at 66094-69223 position in the forward strand. An insertion sequence IS was identified as ISApu1 which belongs to the IS4 family and the ISH8 group. This IS was found to be nested into the ISApu1 transposon and situated at 67840-69223 in the forward strand with 1384 bp length. This analysis also showed that OB49 bears two virtually identical *arsRBCH* operons putatively encoding resistance to inorganic arsenic species, one of them is situated on the S10 and the second one on the S12 contig**.** A transposable element (Tn.arsmerS12) carrying the *arsRBCH* and *merRTPCADE* operons was located between (34150-69223) on the S12 and was found flanked by two inverted (IR) sequences (**Figs 7 and 8**). The IR sequences were identified using the einverted tool. The Multigeneblast showed perfect similarity (100%) of the region located at (57935-72310) on the S12 contig to *Enterobacter hormaechei* strain EB_P9_L5_03. 19 plasmid pIMPIncH12_331kb which mainly hosts

the *merRTPCADE* operon, recombinase family protein, Tn3-like element TnAs1 family transposase, IS4-like element ISApu2 family transposase, hypothetical protein, zinc-binding alcohol dehydrogenase family protein, heme-binding protein, DUF1330 domain-containing protein, and TlpA family protein disulfide reductase. On the other hand, the *arsRBCH* operon blast shows no perfect hit to any ars cluster in the database. While the best similarity results were detected for the chromosome of, the *Enterobacter hormaechei* subsp. *xiangfangensis* strain 34978 and plasmid pJNQH579-2 of the *Klebsiella variicola* strain. The Blast results for best hits are shown in (Table S6). The blast of the complete transposon situated at (34150-69223) showed no similarity to any other fragment or region from the database, which suggests that this sequence is a novel transposon that carries *arsRBCH* and the arsenic resistance proteins with the *merRTPCADE* operon on the same transposable element (**Fig. 7**).

Fig. 7 here

Fig. 8 here

4. Discussion

The genus *Pantoea* has commonly recognized as a plant growth-promoting bacterium (PGPB) and has industrial features including bioremediation and the degradation of herbicides and other toxic compounds [38]. The well-known *Pantoea* PGPB species are *P. agglomerans, P. ananatis, and P. vagans*. although *P. ananatis* was reported as a phytopathogen that infects agronomic crops such as maize, rice, and onion [39,40]. In contrast, some other studies have demonstrated its biocontrol ability [41]. The most important PGPB trait among *Pantoea* species is its ability to solubilize phosphorus by converting insoluble forms of phosphorus (P) to accessible forms [17,42,43] in solid or liquid media, highlighting the importance of this attribute in plant growth.

In this study, we investigated for the first time *Pantoea eucrina* species as a plant-growth-promoting and heavy metal tolerant bacterium. This species displays the following properties; colonies are round, convex, and smooth with entire margins; Indole is not produced; Phenylalanine deaminase reaction is weakly positive as described by Brady et al., [44]. OB49 strain was isolated from the rhizosphere of wheat (*Triticum durum*) grown on a site contaminated with heavy metals and irrigated with industrial and municipal wastewater in Setif, Algeria (Table S1). In vitro OB49 exhibited the ability to grow under abiotic stress, such as heavy metals, salt, and drought. This strain was identified as *Pantoea eucrina* OB49 using 16S RNA and average nucleotide identity (ANI method) and by the genome to genome hybridization (dDDH), OB49 is a gram-negative, rod-shaped bacteria and has paprika color and belongs to the *Enterobacteriaceae* group.

The minimal inhibitory concentration (MIC) study of heavy metals showed that OB49 had a high tolerance to arsenic As(V) up to 6000 mg l^{-1}, 25 mg l^{-1} of cadmium, 400 mg l^{-1} of copper, 400 mg l^{-1} of chromium, and 400 mg l^{-1} of zinc. Y. Feng Zhang et al., [46] reported an endophytic bacterium *P. agglomerans* JP3-3 isolated from Cu-tolerant plants grown on Cu mine wasteland in China which could grow under 2.2 mM (450 mg l^{-1}) of $CdSO_4$ and 1.6 mM (255 mg l^{-1}) of $CuSO_4$ that would be used in phytoremediation applications. Furthermore, Wang et al., [47] described *Pantoea* sp IMH that exhibited high tolerance to As(V) 20 mM, which harbored two ars systems - ars1 (arsR1B1C1H1) and ars2 (arsR2B2C2H2) - with low sequence homology and two arsC-like genes. These findings suggest that OB49 is considered a high tolerant heavy metals bacteria. On the other hand, the test of toxicity using the dose-response curve in the liquid medium has shown that the (estimated dose) ED_{50} for Cd is 72.65 mg l^{-1}, this concentration could allow 50% of *Pantoea* cells to grow in $CdCl_2$ contaminated medium. The ED_{50} for zinc, copper, and chromium, on the other hand, are 3293 mg l^{-1}, 1618 mg l^{-1}, and 1118 mg l^{-1}, respectively. With a toxicity order of; $CdCl_2$> $K_2Cr_2O_7$> $CuCl_2$> $ZnSO_4$ >Na_2HAsO4.

To offer information on genomic elements associated with heavy metal resistance and the genetic determinants, the whole genome of the OB49 strain was sequenced and annotated. In this regard, we found that OB49 could use many mechanisms to cope with heavy metal stress, among them proteins of the transport system. Some of these genes encode transport proteins, for instance, Multidrug efflux system MdtABC-TolC, inner-membrane proton/drug antiporter MdtB (RND type) (TC 2.A.6.2.12), and Copper/silver efflux RND transporter, transmembrane protein CusA (TC 2.A.6.1.3) **(**Table S7**)**. These proteins are involved in resistance to drugs and different heavy metals such as copper, arsenic, cadmium, lead, mercuric, and zinc. According to Su et al., [48] Gram-negative bacteria use tripartite efflux complexes of the resistance-nodulation-cell division (RND) family to expel a variety of harmful substances from the cell. Where copper ions (Cu^{+}) and Ag^{+} were removed from the cytoplasm and periplasm of the cell by the Cus efflux system. Whereas, drug resistance depends on the simultaneous presence of all three proteins (MdtA/MdtB/MdtC) that also contribute to copper and zinc resistance [49,50]. In accordance with the present results, previous studies have demonstrated that *Pantoea agglomerans* C1 tolerates up to 100 mM As(V) equivalent to 7400 mg l^{-1}, whose mechanism was confirmed through genome analysis by the existence of the Arsenite/antimonite: H+ antiporter ArsB (TC 3.A.4.1.1) [51]. Moreover, the ATP-binding protein ZnuC (TC 3.A.1.15.5), a component of the ZnuABC transporter, was increased only in planktonic cultures of *Halobacterium salinarum* exposed to Ni^{2+}, these findings may suggest its contribution to heavy metal resistance [52]. In addition to the transport system, OB49 also has an enzymatic process to deal with other heavy metals such as lead (Pb), cadmium (Cd), zinc (Zn), and mercury (Hg). The copper transport ATPase; copper translocation P-type ATPase (TC 3.A.3.5.4) are also involved in heavy metal resistance [53]. The P-type Cu^{2+} transporter (EC 7.2.2.9) undergoes covalent phosphorylation during the transport cycle. The enzyme of the thermophilic archaeon *Archaeoglobus fulgidus* is involved in the extrusion of copper from the cell [54].

OB49 produces secondary metabolites such as siderophores of type hydroxamate (*fhuE, fhuB, fhuD, fhuC,* and *fhuA*), which can chelate iron to make it accessible to the roots of the plant and contribute to heavy metal stress elevation [55]. Dimkpa [8] reported that *Streptomyces acidiscabies* E13 promoted cowpea growth under nickel contamination by binding iron and nickel.

The *Pantoea* genome comparative analysis with Orthovenn2 and CGview server indicated, that OB49 has partial phage sequences spread on the contigs JAFCXS010000010 and JAFCXS010000015, However, JAFCXS010000015 contig was identified as phage region using PHASTER (PHAge Search Tool Enhanced Release) [57].

Orthovenn2 results showed that *P. agglomerans* P5 shared the highest number of clusters with *P. eucrina* OB49. These clusters included genes involved in the resistance to heavy metals, xenobiotic degradation, phosphate regulation (Phosphate regulon transcriptional regulatory protein *PhoB*), and resistance to abiotic stress. The majority of them are plasmid-encoded proteins; these are Copper resistance protein *CpoA, CpoD, CpoB, CpoC,*and tellurium resistance proteins *TerA, TerW, TerC, TerB, TerD, TerZ, TerE*. Some of these clusters are chromosomes located as a cellular hyperosmotic response (GO:0071474) and Cysteine protease StiP that may play a role in regulating cell morphology in response to stressful conditions which likely cause oxidative damage. Reichert et al., [58] have suggested that this protease is required for normal cell morphology and resistance to tellurite.

As mentioned in the results section, the MGEs were detected in the genome and harbor genes that confer resistance to antibiotics and heavy metals. Knowing that OB49 was isolated from contaminated soil irrigated with mixed effluent and that the MGEs are frequently carriers of hydrocarbon degradation genes and heavy metals tolerance, their dissemination has been proposed as a key component in rapid adaptation to pollution [59]. Based on the nature (open/closed) and size of the pan-genome we can guess about an organism's lifestyle and capacity to acquire exogenous DNA, and so gain new beneficial functions. A large and open pan-genome is typically linked with organisms living in a community where horizontal gene transfer is common [60]. The genus *Pantoea* has an open genome as reported by De Maayer et al., [18]

for *P. ananatis,* concerning our Pan-genome analysis we suggest that *Pantoea* has an open genome and we suggest that OB49 acquired heavy metal resistance genes for adaptation following this prolonged exposure to polluted wastewater. This view is supported by Altimira *et al.*, [61] who suggested that bacterial communities of agricultural soils from central Chile subjected to long-term Cu-pollution have been adapted by acquiring Cu genetic determinants like the *copA* gene in plasmids of four Cu-resistant strains.

The co-existence of two copies of the ars operon, which has been observed when the same genetic arrangement appears twice in the same chromosome, is detected in some situations, as was reported in *Pseudomonas putida* KT2440 [62] and subsequently detected in *Pantoea* sp. IMH, a highly arsenic-tolerant bacterium [47]. Therefore, given the high tolerance of OB49 to heavy metals particularly arsenic, the comparative analysis led us to investigate the likely cause of this tolerance. In *P. eucrina* OB49, the ars1RBCH operon was detected in the S12 contig and was found associated with the *merRTPCADE* operon in a Tn3-like transposon that seems hypothetically described for the first time, the ars2RBCH operon that is located on the S10 contig is commonly found in chromosomes or plasmids of *Pantoea* and *Enterobacteria* species.

The two ars operons are identical in terms of genetic arrangement and consist of a self-repressive transcriptional regulator (*arsR*), a membrane transporter that extrudes As^{+3} from the cell (*arsB*), an arsenate reductase (*arsC*) for the conversion of As^{+4} to $As^{+3,}$ and an *arsH* gene of unknown function but also important for arsenic resistance. Interestingly, the pairwise alignment showed differences between these two operons, where genetic variations were found. This allows us to suggest that these two operons might function in different ways. Another point that needs to be emphasized is that *ars1RBCH* of (S12) had no perfective similarity to the sequences in the entire database. On the other hand, the existence of the complete *merRTPCADE* operon in association with the ars operon suggests that OB49 acquired this MGE from different bacteria to cope with mercury and arsenic stress. This transposon (Tn.arsmerS12) belongs to the Tn3 family transposon that carries the $bla_{TEM\text{-}1B}$ passenger gene [63] and appears to be a union of two types of transposons. The first one is a composite transposon that harbors the mer operon and the second is a unitary transposon that harbors the ars operon along with arsenic resistance proteins to form a complete transposon, which carries genes that confer resistance to arsenic and mercury simultaneously. These findings allow us to suggest that the mer operon may be involved in a phenotypical resistance to mercury (Hg). According to Barkay and Wagner-Döbler [64], the genes of *mer* operon are transmitted by horizontal transfer. As well, Nakaya et al., [65] reported that Tn21, a composite transposon first identified on plasmid NR1 from *Shigella flexneri* in Japan, carries the mer operon.

According to Páez-Espino et al, [62], the contribution of each ars operon to the arsenic resistance phenotype of KT2440 was not additive, as either cluster was sufficient to confer a high level of resistance to the metalloid on the cells. Similarly, the influence of temperature seems to be the fundamental element that favors the coexistence of the two identical operons. Wang et al, [47] reported that pH variations do not influence the expression of arsC1 or arsC2, whereas temperature influences the expression of ars1 which turns out to be the predominant contributor. In agreement with the explanation of Paez and Wang and by observing the climate of the Setif region, where the temperature shows a great divergence between summer and winter, we could suggest that temperature has influenced this adaptability of OB49 to such conditions.

Figure 8 illustrated the Tn.arsmerS12 map annotation, which includes the mer operon, which contains six structural genes *merTPCADE*. Meanwhile, the Tn21 mer operon is comprised of five structural genes (*merTPCAD*), from which *MerR* is divergently transcribed. *MerTPCAD* participates in mercury detoxification via a process that, contrary to popular belief, include mercury import into the cell. *MerP* binds mercury in the periplasm and transports it to *MerT* (a membrane-bound transporter), which then transports Hg^{2+} to *MerA*. *MerA* converts Hg^{2+} to the more volatile Hg^{0}, which may then diffuse through the cell membrane and into the environment [66]. Helman et al., [66] reported that *merD* serves as a co-regulator of this system and in essence resets the response by allowing *MerR* without Hg^{2+} bound to be made, thereby repressing

transcription again. Sone et al., [67] suggested that among the four transporters studied, *MerC* showed more potential to transport Hg(II) across bacterial membrane than *MerE*, *MerF,* and *MerT*. In the same context, Ohshiro et al., [68] suggested that *MerC*, *MerE*, *MerF,* and *MerT* are broad-spectrum heavy metal transporters that mediate both mercury and cadmium transport into cells and that *MerP* accelerates the cadmium transport ability of *MerC*, *MerE*, *MerF,* and *MerT.*

By blasting the transposon that hosts both the *ars1RBCH* operon and the *merRTPCADE* operon in the Genbank database we detected transport proteins that extrude small molecules and ions from the cell via the inner membrane such as the fluoride efflux transporter CrcB that participates in the fluoride resistance described in the *E. cloacae* FRM which coded by an operon of six genes working together among them the universal stress proteins [69]. The Na+/H+ antiporter NhaA, the inorganic anion transporter of the SulP family, and the major facilitator superfamily transporter (MFS) contribute to mitigating osmotic stress [70–72]. These genes have been found associated with the ars1RBCH operons in a significant number of species with more or less different genetic arrangements from one species to another. Which may suggest their contribution to arsenic resistance since this arrangement could affect the expression of the ars operon under different conditions.

5. Conclusion

This work provides insight into the heavy metal tolerance and PGP properties of *Pantoea eucrina* OB49 through in-vitro and in silico analyses. We showed that OB49 is able to survive under high concentrations of heavy metals; this strain has PGP characteristics that allow it to enhance plant development in contaminated sites. Pan-genomic and comparative analysis revealed that the pan-genome of *P. eucrina* OB49 is open, allowing it to acquire MGEs that harbour heavy metal resistance genes through horizontal gene transfer. Our results suggest the identification of a new transposon that harbours genes for resistance to arsenic, fluoride and mercury. Based on the present results, *P. eucrina* strain OB49 may be involved in an ecological alternative for heavy metal remediation and growth promotion of wheat grown in metal-polluted soils.

Data availability

The Whole Genome Shotgun project *Pantoea eucrina* strain OB49has been deposited at DDBJ/ENA/GenBank under the accession JAFCXS000000000. The version described in this paper is version JAFCXS010000000. The BioProject accession number is PRJNA690006. Seven supplementary tables and two supplementary figures are available with the online version of this article.

Funding

This work received no specific grant from any funding agency

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contribution

AL contributed to all parts of the experimental work, formal analysis, data curation, and wrote the original draft of the manuscript. **HC-S and AS** contributed to resource provision and the manuscript review. **HBY** has participated in

experimental work and data curation. **H-IO** has contributed to the conceptualization, resources, reviewing, and supervision of the work. All authors have read and approved the final manuscript.

Acknowledgments

The authors acknowledge support from the Tunisian Ministry of Higher Education and Scientific Research in the ambit of the laboratory project LR03ES03.

References

[1] J. li Liu, J. Yao, F. Wang, N. Min, J. hai Gu, Z. fu Li, G. Sunahara, R. Duran, T. Solevic-Knudsen, K.A. Hudson-Edwards, L. Alakangas, Bacterial diversity in typical abandoned multi-contaminated nonferrous metal(loid) tailings during natural attenuation, Environ. Pollut. 247 (2019) 98–107. https://doi.org/10.1016/j.envpol.2018.12.045.

[2] L.E. de-Bashan, J.P. Hernandez, Y. Bashan, The potential contribution of plant growth-promoting bacteria to reduce environmental degradation - A comprehensive evaluation, Appl. Soil Ecol. 61 (2012) 171–189. https://doi.org/10.1016/j.apsoil.2011.09.003.

[3] P.R. Patel, S.S. Shaikh, R.Z. Sayyed, Dynamism of PGPR in bioremediation and plant growth promotion in heavy metal contaminated soil, (2016).

[4] S. Timmusk, L. Behers, J. Muthoni, A. Muraya, A.C. Aronsson, Perspectives and challenges of microbial application for crop improvement, Front. Plant Sci. 8 (2017) 1–10. https://doi.org/10.3389/fpls.2017.00049.

[5] Y. Gong, J.L. Bai, H.T. Yang, W. Di Zhang, Y.W. Xiong, P. Ding, S. Qin, Phylogenetic diversity and investigation of plant growth-promoting traits of actinobacteria in coastal salt marsh plant rhizospheres from Jiangsu, China, Syst. Appl. Microbiol. 41 (2018) 516–527. https://doi.org/10.1016/j.syapm.2018.06.003.

[6] Y. Bashan, L.E. De-Bashan, S.R. Prabhu, J.P. Hernandez, Advances in plant growth-promoting bacterial inoculant technology: Formulations and practical perspectives (1998-2013), Plant Soil. 378 (2014) 1–33. https://doi.org/10.1007/s11104-013-1956-x.

[7] L.P. Singha, N. Sinha, P. Pandey, Rhizoremediation prospects of Polyaromatic hydrocarbon degrading rhizobacteria, that facilitate glutathione and glutathione-S-transferase mediated stress response, and enhance growth of rice plants in pyrene contaminated soil, Ecotoxicol. Environ. Saf. 164 (2018) 579–588. https://doi.org/10.1016/j.ecoenv.2018.08.069.

[8] C. Dimkpa, T. Weinand, F. Asch, Plant-rhizobacteria interactions alleviate abiotic stress conditions, Plant, Cell Environ. 32 (2009) 1682–1694. https://doi.org/10.1111/j.1365-3040.2009.02028.x.

[9] C. Leal, F. Fontaine, A. Aziz, C. Egas, C. Clément, P. Trotel-Aziz, Genome sequence analysis of the beneficial *Bacillus subtilis* PTA-271 isolated from a *Vitis vinifera* (cv. Chardonnay) rhizospheric soil: assets for sustainable biocontrol, Environ. Microbiomes. 16 (2021) 1–14. https://doi.org/10.1186/s40793-021-00372-3.

[10] H. Chauhan, D.J. Bagyaraj, G. Selvakumar, S.P. Sundaram, Novel plant growth promoting rhizobacteria-Prospects and potential, Appl. Soil Ecol. 95 (2015) 38–53. https://doi.org/10.1016/j.apsoil.2015.05.011.

[11] F. Gavini, J. Mergaert, A. Beji, C. Mielcarek, D. Izard, K. Kersters, J. De Ley, Transfer of *Enterobacter agglomerans* (Beijerinck 1888) Ewing and Fife 1972 to Pantoea gen. nov. as *Pantoea agglomerans* comb. nov. and description of *Pantoea dispersa* sp. nov., Int. J. Syst. Bacteriol. 39 (1989) 337–345. https://doi.org/10.1099/00207713-39-3-337.

[12] J. lian Gao, J. Xue, H. Yan, S. Tong, M. Sayyar Khan, L. wei Wang, X. jie Mao, X. Zhang, J. guang Sun, *Pantoea endophytica* sp. nov., novel endophytic bacteria isolated from maize planting in different geographic

regions of northern China, Syst. Appl. Microbiol. 42 (2019) 488–494. https://doi.org/10.1016/j.syapm.2019.06.001.

[13] S.G. Dastager, C.K. Deepa, S.C. Puneet, C.S. Nautiyal, A. Pandey, Isolation and characterization of plant growth-promoting strain *Pantoea* NII-186. from Western Ghat Forest soil, India, Lett. Appl. Microbiol. 49 (2009) 20–25. https://doi.org/10.1111/j.1472-765X.2009.02616.x.

[14] C. Chen, K. Xin, H. Liu, J. Cheng, X. Shen, Y. Wang, L. Zhang, *Pantoea alhagi*, a novel endophytic bacterium with ability to improve growth and drought tolerance in wheat, Sci. Rep. 7 (2017) 1–14. https://doi.org/10.1038/srep41564.

[15] H. Cherif-Silini, B. Thissera, A.C. Bouket, N. Saadaoui, A. Silini, M. Eshelli, F.N. Alenezi, A. Vallat, L. Luptakova, B. Yahiaoui, S. Cherrad, S. Vacher, M.E. Rateb, L. Belbahri, Durum wheat stress tolerance induced by endophyte *Pantoea agglomerans* with genes contributing to plant functions and secondary metabolite Arsenal, Int. J. Mol. Sci. 20 (2019). https://doi.org/10.3390/ijms20163989.

[16] C.L. Brady, S.N. Venter, I. Cleenwerck, K. Vandemeulebroecke, P. De Vos, T.A. Coutinho, Transfer of *Pantoea citrea*, *Pantoea punctata* and *Pantoea terrea* to the genus *Tatumella* emend. as *Tatumella citrea* comb. nov., *Tatumella punctata* comb. nov. and *Tatumella terrea* comb. nov. and description of *Tatumella morbirosei* sp. nov., Int. J. Syst. Evol. Microbiol. 60 (2010) 484–494. https://doi.org/10.1099/ijs.0.012070-0.

[17] Q. Chen, S. Liu, Identification and Characterization of the Phosphate-Solubilizing Bacterium *Pantoea* sp. S32 in Reclamation Soil in Shanxi, China, Front. Microbiol. 10 (2019) 1–12. https://doi.org/10.3389/fmicb.2019.02171.

[18] P. De Maayer, W. Chan, E. Rubagotti, S.N. Venter, I.K. Toth, P.R.J. Birch, T.A. Coutinho, Analysis of the *Pantoea ananatis* pan-genome reveals factors underlying its ability to colonize and interact with plant, insect and vertebrate hosts, BMC Genomics. 15 (2014) 404. https://doi.org/10.1186/1471-2164-15-404.

[19] H. Wu, D. Wang, F. Gao, Toward a high-quality pan-genome landscape of *Bacillus subtilis* by removal of confounding strains, Brief. Bioinform. 22 (2021) 1951–1971. https://doi.org/10.1093/bib/bbaa013.

[20] D. Medini, C. Donati, H. Tettelin, V. Masignani, R. Rappuoli, The microbial pan-genome, Curr. Opin. Genet. Dev. 15 (2005) 589–594. https://doi.org/10.1016/j.gde.2005.09.006.

[21] G. Vernikos, D. Medini, D.R. Riley, H. Tettelin, Ten years of pan-genome analyses, Curr. Opin. Microbiol. 23 (2015) 148–154. https://doi.org/10.1016/j.mib.2014.11.016.

[22] Z.K. Yang, H. Luo, Y. Zhang, B. Wang, F. Gao, Pan-genomic analysis provides novel insights into the association of *E.coli* with human host and its minimal genome, Bioinformatics. 35 (2019) 1987–1991. https://doi.org/10.1093/bioinformatics/bty938.

[23] C. Knief, Analysis of plant microbe interactions in the era of next generation sequencing technologies, Front. Plant Sci. 5 (2014) 1–23. https://doi.org/10.3389/fpls.2014.00216.

[24] A. Aigle, J.I. Prosser, C. Gubry-Rangin, The application of high-throughput sequencing technology to analysis of amoA phylogeny and environmental niche specialisation of terrestrial bacterial ammonia-oxidisers, Environ. Microbiomes. 14 (2019) 1–10. https://doi.org/10.1186/s40793-019-0342-6.

[25] V.J. Shariati, M.A. Malboobi, Z. Tabrizi, E. Tavakol, P. Owilia, M. Safari, Comprehensive genomic analysis of a plant growth-promoting rhizobacterium *Pantoea agglomerans* strain P5 /631/208/212/748 /631/208/464 /38/43 /45 article, Sci. Rep. 7 (2017) 1–12. https://doi.org/10.1038/s41598-017-15820-9.

[26] S. Nurk, A. Bankevich, D. Antipov, A. Gurevich, A. Korobeynikov, A. Lapidus, A. Prjibelsky, A. Pyshkin, A. Sirotkin, Y. Sirotkin, R. Stepanauskas, J. McLean, R. Lasken, S.R. Clingenpeel, T. Woyke, G. Tesler, M.A. Alekseyev, P.A. Pevzner, Assembling genomes and mini-metagenomes from highly chimeric reads, Lect. Notes

Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 7821 LNBI (2013) 158–170. https://doi.org/10.1007/978-3-642-37195-0_13.

[27] I. Lee, M. Chalita, S.M. Ha, S.I. Na, S.H. Yoon, J. Chun, ContEst16S: An algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences, Int. J. Syst. Evol. Microbiol. 67 (2017) 2053–2057. https://doi.org/10.1099/ijsem.0.001872.

[28] S.H. Yoon, S. min Ha, J. Lim, S. Kwon, J. Chun, A large-scale evaluation of algorithms to calculate average nucleotide identity, Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol. 110 (2017) 1281–1286. https://doi.org/10.1007/s10482-017-0844-4.

[29] T. Seemann, Prokka: Rapid prokaryotic genome annotation, Bioinformatics. 30 (2014) 2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

[30] A.J. Page, C.A. Cummins, M. Hunt, V.K. Wong, S. Reuter, M.T.G. Holden, M. Fookes, D. Falush, J.A. Keane, J. Parkhill, Roary: Rapid large-scale prokaryote pan genome analysis, Bioinformatics. 31 (2015) 3691–3693. https://doi.org/10.1093/bioinformatics/btv421.

[31] L. Xu, Z. Dong, L. Fang, Y. Luo, Z. Wei, H. Guo, G. Zhang, Y.Q. Gu, D. Coleman-Derr, Q. Xia, Y. Wang, OrthoVenn2: A web server for whole-genome comparison and annotation of orthologous clusters across multiple species, Nucleic Acids Res. 47 (2019) W52–W58. https://doi.org/10.1093/nar/gkz333.

[32] A.R. Wattam, J.J. Davis, R. Assaf, S. Boisvert, T. Brettin, C. Bun, N. Conrad, E.M. Dietrich, T. Disz, J.L. Gabbard, S. Gerdes, C.S. Henry, R.W. Kenyon, D. Machi, C. Mao, E.K. Nordberg, G.J. Olsen, D.E. Murphy-Olson, R. Olson, R. Overbeek, B. Parrello, G.D. Pusch, M. Shukla, V. Vonstein, A. Warren, F. Xia, H. Yoo, R.L. Stevens, Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center, Nucleic Acids Res. 45 (2017) D535–D542. https://doi.org/10.1093/nar/gkw1017.

[33] M.H.K. Johansson, V. Bortolaia, S. Tansirichaiya, F.M. Aarestrup, A.P. Roberts, T.N. Petersen, Detection of mobile genetic elements associated with antibiotic resistance in *Salmonella enterica* using a newly developed web tool: MobileElementFinder, J. Antimicrob. Chemother. 76 (2021) 101–109. https://doi.org/10.1093/JAC/DKAA390.

[34] M.H. Medema, E. Takano, R. Breitling, Detecting sequence homology at the gene cluster level with multigeneblast, Mol. Biol. Evol. 30 (2013) 1218–1223. https://doi.org/10.1093/molbev/mst025.

[35] C. Ritz, F. Baty, J.C. Streibig, D. Gerhard, Dose-response analysis using R, PLoS One. 10 (2015) 1–13. https://doi.org/10.1371/journal.pone.0146021.

[36] J.J. Davis, A.R. Wattam, R.K. Aziz, 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, R. Stevens, The PATRIC Bioinformatics Resource Center: Expanding data and analysis capabilities, Nucleic Acids Res. 48 (2020) D606–D612. https://doi.org/10.1093/nar/gkz943.

[37] P. Stothard, D.S. Wishart, Circular genome visualization and exploration using CGView, Bioinformatics. 21 (2005) 537–539. https://doi.org/10.1093/bioinformatics/bti054.

[38] A.M. Walterson, J. Stavrinides, *Pantoea*: Insights into a highly versatile and diverse genus within the *Enterobacteriaceae*, FEMS Microbiol. Rev. 39 (2015) 968–984. https://doi.org/10.1093/femsre/fuv027.

[39] T. Weller-Stuart, P. De Maayer, T. Coutinho, *Pantoea ananatis* : genomic insights into a versatile pathogen, Mol. Plant Pathol. 18 (2017) 1191–1198. https://doi.org/10.1111/mpp.12517.

[40] C.L. Brady, T. Goszczynska, S.N. Venter, I. Cleenwerck, P. de Vos, R.D. Gitaitis, T.A. Coutinho, *Pantoea allii*

sp. nov., isolated from onion plants and seed, Int. J. Syst. Evol. Microbiol. 61 (2011) 932–937. https://doi.org/10.1099/ijs.0.022921-0.

[41] L. Wu, R. Liu, Y. Niu, H. Lin, W. Ye, L. Guo, X. Hu, Whole genome sequence of *Pantoea ananatis* R100, an antagonistic bacterium isolated from rice seed, J. Biotechnol. 225 (2016) 1–2. https://doi.org/10.1016/j.jbiotec.2016.03.007.

[42] L.N. Castagno, M.J. Estrella, A.I. Sannazzaro, A.E. Grassano, O.A. Ruiz, Phosphate-solubilization mechanism and in vitro plant growth promotion activity mediated by *Pantoea eucalypti* isolated from Lotus tenuis rhizosphere in the Salado River Basin (Argentina), J. Appl. Microbiol. 110 (2011) 1151–1165. https://doi.org/10.1111/j.1365-2672.2011.04968.x.

[43] A.D. Suleimanova, A. Beinhauer, L.R. Valeeva, I.B. Chastukhina, N.P. Balaban, E. V. Shakirov, R. Greiner, M.R. Sharipova, Novel Glucose-1-Phosphatase with High Phytase Activity and Unusual Metal Ion Activation from Soil Bacterium *Pantoea* sp. Strain 3.5.1, Appl. Environ. Microbiol. 81 (2015) 6790–6799. https://doi.org/10.1128/AEM.01384-15.

[44] C.L. Brady, I. Cleenwerck, S.N. Venter, K. Engelbeen, P. De Vos, T.A. Coutinho, Emended description of the genus *Pantoea*, description of four species from human clinical samples, *Pantoea septica* sp. nov., *Pantoea eucrina* sp. nov., *Pantoea brenneri* sp. nov. and *Pantoea conspicua* sp. nov., and transfer of *Pectobacterium cypripedii* (Hor, Int. J. Syst. Evol. Microbiol. 60 (2010) 2430–2440. https://doi.org/10.1099/ijs.0.017301-0.

[45] Y. feng Zhang, L. yan He, Z. jin Chen, Q. ya Wang, M. Qian, X. fang Sheng, Characterization of ACC deaminase-producing endophytic bacteria isolated from copper-tolerant plants and their potential in promoting the growth and copper accumulation of *Brassica napus*, Chemosphere. 83 (2011) 57–62. https://doi.org/10.1016/j.chemosphere.2011.01.041.

[46] Y.F. Zhang, L.Y. He, Z.J. Chen, W.H. Zhang, Q.Y. Wang, M. Qian, X.F. Sheng, Characterization of lead-resistant and ACC deaminase-producing endophytic bacteria and their potential in promoting lead accumulation of rape, J. Hazard. Mater. 186 (2011) 1720–1725. https://doi.org/10.1016/j.jhazmat.2010.12.069.

[47] L. Wang, X. Zhuang, G. Zhuang, C. Jing, Arsenic resistance strategy in *Pantoea* sp. IMH: Organization, function and evolution of ars genes, Sci. Rep. 6 (2016) 1–9. https://doi.org/10.1038/srep39195.

[48] C.C. Su, F. Long, E.W. Yu, The cus efflux system removes toxic ions via a methionine shuttle, Protein Sci. 20 (2011) 6–18. https://doi.org/10.1002/pro.532.

[49] K. Nishino, E. Nikaido, A. Yamaguchi, Regulation of multidrug efflux systems involved in multidrug and metal resistance of *Salmonella enterica* serovar *typhimurium*, J. Bacteriol. 189 (2007) 9066–9075. https://doi.org/10.1128/JB.01045-07.

[50] N. Baranova, H. Nikaido, The BaeSR two-component regulatory system activates transcription of the yegMNOB (mdtABCD) transporter gene cluster in Escherichia coli and increases its resistance to novobiocin and deoxycholate, J. Bacteriol. 184 (2002) 4168–4176. https://doi.org/10.1128/JB.184.15.4168-4176.2002.

[51] F. Luziatelli, A.G. Ficca, M. Cardarelli, F. Melini, A. Cavalieri, M. Ruzzi, Genome sequencing of *Pantoea agglomerans* C1 provides insights into molecular and genetic mechanisms of plant growth-promotion and tolerance to heavy metals, Microorganisms. 8 (2020) 1–17. https://doi.org/10.3390/microorganisms8020153.

[52] S. Völkel, S. Hein, N. Benker, F. Pfeifer, C. Lenz, G. Losensky, How to Cope With Heavy Metal Ions: Cellular and Proteome-Level Stress Response to Divalent Copper and Nickel in *Halobacterium salinarum* R1 Planktonic and Biofilm Cells, Front. Microbiol. 10 (2020) 1–15. https://doi.org/10.3389/fmicb.2019.03056.

[53] O. Lewinson, A.T. Lee, D.C. Rees, A P-type ATPase importer that discriminates between essential and toxic transition metals, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 4677–4682.

https://doi.org/10.1073/pnas.0900666106.
[54] S. Jayakanthan, S.A. Roberts, A. Weichsel, J.M. Argüello, M.M. McEvoy, Conformations of the apo-, substrate-bound and phosphate-bound ATP-binding domain of the Cu(II) ATPase CopB illustrate coupling of domain movement to the catalytic cycle, Biosci. Rep. 32 (2012) 443–453. https://doi.org/10.1042/BSR20120048.
[55] P. Gusain, R. Paliwal, V. Singh, Rhizoremediation of cadmium-contaminated soil associated with hydroxamate siderophores isolated from Cd-resistant plant growth–promoting *Dietzia maris* and *Lysinibacillus* strains, Int. J. Phytoremediation. 19 (2017) 290–299. https://doi.org/10.1080/15226514.2016.1225281.
[56] C.O. Dimkpa, D. Merten, A. Svatoš, G. Büchel, E. Kothe, Siderophores mediate reduced and increased uptake of cadmium by *Streptomyces tendae* F4 and sunflower (*Helianthus annuus*), respectively, J. Appl. Microbiol. 107 (2009) 1687–1696. https://doi.org/10.1111/j.1365-2672.2009.04355.x.
[57] D. Arndt, J.R. Grant, A. Marcu, T. Sajed, A. Pon, Y. Liang, D.S. Wishart, PHASTER: a better, faster version of the PHAST phage search tool, Nucleic Acids Res. 44 (2016) W16–W21. https://doi.org/10.1093/nar/gkw387.
[58] B. Reichert, A.J. Dornbusch, J. Arguello, S.E. Stanley, K.M. Lang, C.P. Lostroh, M.A. Daugherty, *Acinetobacter baylyi* long-term stationary-phase protein stiP is a protease required for normal cell morphology and resistance to tellurite, Can. J. Microbiol. 59 (2013) 726–736. https://doi.org/10.1139/cjm-2013-0517.
[59] R.G. Taketani, N.O. Franco, A.S. Rosado, J.D. van Elsas, Microbial community response to a simulated hydrocarbon spill in mangrove sediments, J. Microbiol. 48 (2010) 7–15. https://doi.org/10.1007/s12275-009-0147-1.
[60] L. Rouli, V. Merhej, P.E. Fournier, D. Raoult, The bacterial pangenome as a new tool for analysing pathogenic bacteria, New Microbes New Infect. 7 (2015) 72–85. https://doi.org/10.1016/j.nmni.2015.06.005.
[61] F. Altimira, C. Yẫez, G. Bravo, M. González, L.A. Rojas, M. Seeger, Characterization of copper-resistant bacteria and bacterial communities from copper-polluted agricultural soils of central Chile, BMC Microbiol. 12 (2012). https://doi.org/10.1186/1471-2180-12-193.
[62] A.D. Páez-Espino, G. Durante-Rodríguez, V. de Lorenzo, Functional coexistence of twin arsenic resistance systems in Pseudomonas putida KT2440, Environ. Microbiol. 17 (2015) 229–238. https://doi.org/10.1111/1462-2920.12464.
[63] G. Lima-Mendez, D.O. Alvarenga, K. Ross, B. Hallet, L. Van Melderen, A.M. Varani, M. Chandler, Toxin-antitoxin gene pairs found in Tn3 family transposons appear to be an integral part of the transposition module, MBio. 11 (2020). https://doi.org/10.1128/mBio.00452-20.
[64] T. Barkay, I. Wagner-Döbler, Microbial transformations of mercury: Potentials, challenges, and achievements in controlling mercury toxicity in the environment, Adv. Appl. Microbiol. 57 (2005) 1–52. https://doi.org/10.1016/S0065-2164(05)57001-1.
[65] R. Nakaya, A. Nakamura, Y. Murata, Resistance transfer agents in *Shigella*, Biochem. Biophys. Res. Commun. 3 (1960) 654–659. https://doi.org/10.1016/0006-291X(60)90081-4.
[66] J.D. Helmann, S. Soonsanga, S. Gabriel, Metalloregulators: Arbiters of Metal Sufficiency, in: Mol. Microbiol. Heavy Met., Springer Berlin Heidelberg, Berlin, Heidelberg, n.d.: pp. 37–71. https://doi.org/10.1007/7171_2006_073.
[67] Y. Sone, R. Nakamura, H. Pan-Hou, T. Itoh, M. Kiyono, Role of MerC, MerE, MerF, MerT, and/or MerP in resistance to mercurials and the transport of mercurials in escherichia coli, Biol. Pharm. Bull. 36 (2013) 1835–1841. https://doi.org/10.1248/bpb.b13-00554.
[68] Y. Ohshiro, S. Uraguchi, R. Nakamura, Y. Takanezawa, M. Kiyono, Cadmium transport activity of four

mercury transporters (MerC, MerE, MerF and MerT) and effects of the periplasmic mercury-binding protein MerP on Mer-dependent cadmium uptake, FEMS Microbiol. Lett. 367 (2020) 1–6. https://doi.org/10.1093/femsle/fnaa177.

[69] X. Liu, J. Tian, L. Liu, T. Zhu, X. Yu, X. Chu, B. Yao, N. Wu, Y. Fan, Identification of an operon involved in fluoride resistance in *Enterobacter cloacae* FRM, Sci. Rep. 7 (2017) 1–9. https://doi.org/10.1038/s41598-017-06988-1.

[70] N. Xu, L. Wang, H. Cheng, Q. Liu, J. Liu, Y. Ma, In vitro functional characterization of the na+/h+ antiporters in *Corynebacterium glutamicum*, FEMS Microbiol. Lett. 363 (2015) 1–7. https://doi.org/10.1093/femsle/fnv237.

[71] D. Drew, R.A. North, K. Nagarathinam, M. Tanabe, Structures and General Transport Mechanisms by the Major Facilitator Superfamily (MFS), Chem. Rev. 121 (2021) 5289–5335. https://doi.org/10.1021/acs.chemrev.0c00983.

[72] M. Ludwig, D.A. Bryant, Synechococcus sp. Strain PCC 7002 transcriptome: Acclimation to temperature, salinity, oxidative stress, and mixotrophic growth conditions, Front. Microbiol. 3 (2012) 1–14. https://doi.org/10.3389/fmicb.2012.00354.

Figure captions

Fig. 1 *Pantoea* growth (A_{600} nm) response under increased heavy metals concentrations

Fig. 2 Dose-response curves of the effective dose of heavy metals, $CuCl_2$ (A), $ZnSO_4$ (B), $CdCl_2$ (C), and $K_2Cr_2O_7$ (D) represented as the logarithm of increasing concentrations.

Fig. 3 Shoots (a) and roots (b) length of inoculated and uninoculated wheat seeds under cadmium and zinc concentration. Using the Tukey's HSD test, bars labeled with various letters are significantly different across treatments at $P < 0.05$.

Fig. 4 Phylogenetic Tree of OB49 genome and *Pantoea* genomes created using the Codon Tree method

Fig. 5 Protein sequence-based genome comparison using bidirectional BLASTP of *Pantoea* species using the circular genome viewer (CGview)

Fig. 6 Pan-genome estimation for the genus *Pantoea.* **a)**. Number of new genes in the pan-genome, **b)** Number of the unique genes, **c)** Number of conserved genes against the number of total genes, **d)** Unique genes against the number of genes.

Fig. 7 Transposon map of Tn.arsmerS12. **a**) Full length of Tn.arsmerS12. **b**) Left end of Tn.arsmerS12. **c**) Right end of Tn.arsmerS12.

Fig. 8 Transposon map annotation of Tn.arsmerS12 that harbors *arsRBCH* and *merRTPCADE* operons

Supplementary Figures captions

Figure S1 Circular graphical display of the distribution of the genome annotations

Figure S2 Subsystem annotation features of *P. eucrina* OB49 genome