Current Genetics (2022) 68:481–503
https://doi.org/10.1007/s00294-022-01245-z

ORIGINAL ARTICLE

Whole genome sequencing and comparative genomic analyses of *Pseudomonas aeruginosa* strain isolated from arable soil reveal novel insights into heavy metal resistance and codon biology

Jayanti Saha[1] · Sourav Dey[1] · Ayon Pal[1]

Received: 28 March 2022 / Revised: 14 May 2022 / Accepted: 6 June 2022 / Published online: 28 June 2022

Abstract

Elevated concentration of non-essential persistent heavy metals and metalloids in the soil is detrimental to essential soil microbes and plants, resulting in diminished diversity and biomass. Thus, isolation, screening, and whole genomic analysis of potent strains of bacteria from arable lands with inherent capabilities of heavy metal resistance and plant growth promotion hold the key for bio remedial applications. This study is an attempt to do the same. In this study, a potent strain of *Pseudomonas aeruginosa* was isolated from paddy fields, followed by metabolic profiling using FTIR, metal uptake analysis employing ICP–MS, whole genome sequencing and comparative codon usage analysis. ICP–MS study provided insights into a high degree of Cd uptake during the exponential phase of growth under cumulative metal stress to Cd, Zn and Co, which was further corroborated by the detection of *cadA* gene along with *czcCBA* operon in the genome upon performing whole-genome sequencing. This potent strain of *Pseudomonas aeruginosa* also harboured genes, such as *copA, chrA, znuA, mgtE, corA*, and others conferring resistance against different heavy metals, such as Cd, Zn, Co, Cu, Cr, etc. A comparative codon usage bias analysis at the genomic and genic level, whereby several heavy metal resistant genes were considered in the backdrop of two housekeeping genes among 40 *Pseudomonas* spp. indicated the presence of a relatively strong codon usage bias in the studied strain. With this work, an effort was made to explore heavy metal-resistant bacteria (isolated from arable soil) and whole genome sequence analysis to get insight into metal resistance for future bio remedial applications.

Keywords *Pseudomonas aeruginosa* · Heavy metal resistance · Plant growth promotion · Whole genome sequencing · Codon usage bias · Metabolic profiling · Fourier transform infrared spectroscopy · Inductively coupled plasma mass spectrometry

Introduction

Bioaccumulation of pollutants, especially persistent heavy metals and metalloids in soil due to anthropogenic activities, including mal-agricultural practices, has been proven to inhibit carbon biomass accumulation at the lower soil levels leading to diminished biodiversity of the essential bacterial population (Šmejkalová et al. 2003; Giller et al. 2009). However, giving credence to the "microbial infallibility" hypothesis (O'Malley and Walsh 2021), soil-borne bacteria, having been exposed to a metalliferous soil environment over a period of time, have established resistance mechanisms not only to mitigate heavy metals and metalloids from contaminated soil but also promote plant growth (Jadhav et al. 2010). Thus, from the perspective of amelioration of heavy metal contaminants in polluted soil, especially in arable lands, it becomes necessary to explore avenues for employing these metal resistant soil-borne indigenous eubacterial species because of their inherent competitive advantage over non-indigenous bacteria (Kumar and Gopal 2015). In this regard, the soil of arable lands requires much more attention for global food security (Alengebawy et al. 2021). Though polluted sites are regarded as the principal sink for obtaining metal-resistant bacteria, in this study, the focus was deliberately placed on arable lands, which are not classified as polluted sites. The detection of heavy

Communicated by Michael Polymenis.

✉ Ayon Pal
ayonpal.ruc@gmail.com

[1] Microbiology and Computational Biology Laboratory, Department of Botany, Raiganj University, Raiganj, West Bengal 733134, India

metal resistant bacteria from such land may indicate "silent" arable land contamination and help understand the prevalence of heavy metal resistant bacteria in arable land. Earlier studies have provided insights into heavy metal toxicity reduction by different soil-borne eubacterial genera, such as *Pseudomonas* (Chong et al. 2012), *Acinetobacter* (Helal 2016), *Achromobacter* (Stanbrough et al. 2013), *Alcaligenes* (Abbas et al. 2015), *Burkholderia* (Jiang et al. 2008), *Bacillus* (Çolak et al. 2011), and others through different metabolic activities including accumulation and transformation of toxic metals. However, there is a need to extend the search for indigenous soil-dwelling heavy metal resistant bacteria for bio remedial purposes. Therefore, the present study was aimed to explore heavy metal resistant indigenous bacterial strains from arable lands with the potential to promote plant growth and gain insight into the biology of heavy metal uptake and resistance. This was further followed up with whole-genome sequencing and a comprehensive comparative genomics-based codon usage analysis to understand better the mechanistic forces shaping codon usage of the genes conferring resistance to heavy metals. Codon usage bias (CUB), i.e., preference in the usage of one synonymous codon over others is mainly observed due to three significant factors, namely, mutation, selection and random drift, and plays a significant role in genome evolution (Sharp and Matassi 1994; Plotkin and Kudla 2011; Parvathy and Udayasuriyan 2022). The isolation of bacteria during this study was done from arable land of Uttar Dinajpur district of West Bengal, India, cultivating 'Tulaipanji' variety of rice. 'Tulaipanji' is an exclusive aromatic indigenous rice variety endemic to this region (Mondal et al. 2013) coveted with a geographical indications (GI) tag by the Government of India (Registrar 2017). As the microbial diversity of arable land from this region is largely unexplored, this offers an excellent opportunity to explore the bacteria resident here for environmental remediation and plant growth promotion.

A potent *Pseudomonas aeruginosa* strain resistant to multiple heavy metals and plant growth-promoting potential was obtained during this study.The genus *Pseudomonas* represents a well-known and versatile Gram negative bacterial group (Aguilar-Barajas et al. 2010). It can uptake and resist toxic metals at significantly higher concentrations while subsequently converting them into relatively less toxic forms (Mohamed and Abo-Amer 2012; Wasi et al. 2013; Chellaiah 2018), paving the way for effective bioremediation. Several species belonging to this genus such as *P. aeruginosa*, *P. fluorescens, P. putida, P. mendocina, P. rhizophila, P. veronii* and *P. stutzeri* (Chong et al. 2012; Manara et al. 2012; Alhasawi et al. 2015; Méndez et al. 2017; Hassen et al. 2018) have been reported to demonstrate competence in degrading toxic non-biodegradable contaminants. In addition, it has been shown that heavy metal resistant bacteria play an indispensable role in promoting plant growth (Khanna et al. 2019; Ren et al. 2019; Mesa-Marín et al. 2020). In recent years, some studies have also revealed the complete genome sequence of plant growth-promoting *Pseudomonas* spp. under heavy metal stressed conditions (Kang et al. 2020). Apart from this, whole genome sequences of some highly Cu resistant soil-dwelling (Havryliuk et al. 2020), hydrocarbon and heavy metal degrading rhizospheric *Pseudomonas* species (Chlebek et al. 2022) have also been reported.

In this study, metabolic fingerprinting based on Fourier transform infrared (FTIR) spectroscopy and high resolution inductively coupled plasma mass spectra analysis (ICP–MS) was performed to study the strain's heavy metal resistance and uptake potential, respectively. Because of the growing importance of genomics within the conventional dimensions of environmental microbiology, a comparative genomic analysis of whole-genome data of forty different *Pseudomonas* spp. belonging to soil and plant-based habitats was also carried out to compare and comprehend the codon usage bias (CUB) pattern of both genome and genes related to heavy metal resistance.

Materials and methods

Isolation and initial characterization of heavy metal resistant bacterial isolate

A potent multi heavy metal resistant bacterial isolate was obtained from the rhizospheric soil of paddy fields (cultivating indigenous 'Tulaipanji' rice variety) located at Raiganj block of West Bengal (near Dhurail; 25° 42′ 19″ N, 88° 7′ 3″ E). The soil samples were collected under aseptic conditions and processed to isolate metal-resistant bacteria. Initially, the soil samples were incubated in sterilized metal-supplemented (100 µg/mL metal concentration) Tryptone Soya Broth (TSB) medium in different Erlenmeyer flasks (for each metal) at 35 °C for 48 h. A total of nine heavy metal salts were used to determine resistance against nine different heavy metals (Cd, Zn, Ni, Pb, Fe, Co, Cu, As, and Cr). Next, standard serial dilution followed by the pour plate method was used to isolate bacterial colonies on culture plates. The heavy metal resistant bacterial isolates were obtained in axenic cultures by transferring each of the single bacterial colonies that appeared on the incubated culture plates into freshly prepared Tryptone Soya Agar (TSA) slants. The details of the different heavy metal salts used in this study are given in Supplementary Table S1. Many bacterial isolates showing growth in a variety of heavy metals, such as Cd, Ni, Cr, Co, Zn, Fe(II), Pb, Cu, As, etc., were obtained during the initial screening. Minimum inhibitory concentration (MIC) values of all the resistant bacterial isolates for each of the studied heavy metals were determined, followed by the characterization of biochemical attributes.

The most enduring bacterial isolate, MR41, exhibiting an elevated level of metal tolerance to most heavy metals, was selected for further studies regarding plant growth promotion, heavy metal uptake, and subsequent whole genome sequencing, followed by comparative genomic analysis.

Morphological and biochemical characterization

Scanning Electron Microscopy (SEM) and Gram staining were done to elucidate the morphology of the isolate MR41. For SEM analysis, primary fixation of the bacterial cell mass was done in 2.5% glutaraldehyde solution, followed by washing in phosphate buffer saline (PBS). After that, dehydration was carried out in graded ethanol series (30%, 50%, 70%, 80%, 90%, 95%, and 100%) followed by gold sputter coating. The prepared specimen was visualized under the scanning electron microscope (JEOL, model no. JSM-IT100).

Biochemical tests such as starch hydrolysis, lipid hydrolysis, catalase, oxidase, urease, nitrate reduction, indole production, gelatine liquefaction, and citrate utilization were carried out for characterization of MR41 (Cappuccino and Welsh 2017).

In vitro screening tests for detecting plant growth promotion potential

The isolate MR41 was screened in vitro to assess plant growth promotion ability using tests for phosphate solubilisation (Nautiyal 1999), zinc solubilisation (Kamran et al. 2017), hydrogen cyanide (HCN) production (Lorck 1948), indole acetic acid (IAA) production (Bric et al. 1991), ammonia production (Cappucino and Sherman 1992), and siderophore production (Schwyn and Neilands 1987).

Molecular identification and phylogenetic analysis of the isolate based on 16S rRNA gene

Molecular identification of the bacterial isolate MR41 was carried out by sequencing the 16S rRNA gene. For this, genomic DNA was initially isolated by the standard phenol–chloroform method (Lever et al. 2015), followed by PCR amplification of the 16S rRNA gene using universal primers 16F27 [5′-CCA GAG TTT GAT CMT GGC TCA G-3′] and 16R1492 [5′-TAC GGY TAC CTT GTT ACG ACT T-3′] (Frank et al. 2008). The amplified 16S rRNA gene PCR product was purified by PEG–NaCl precipitation and directly sequenced on an ABI® 3730XL automated DNA sequencer as per the manufacturer's instructions. Sequencing was carried out from both ends using additional internal primers to read at least twice for each position. Assembly was performed utilizing the Lasergene package (Burland 2000) followed by identification using the EzBioCloud database (Yoon et al. 2017). The obtained sequence was submitted to the GenBank repository, and accession number was obtained.

To infer on phylogenetic affinity of the isolate with other closely related intragenic type strains, a comprehensive molecular phylogenetic analysis was carried out utilizing the 16S rRNA gene sequence of MR41 with that of other strains possessing close resemblance, revealed through nucleotide BLAST (McGinnis and Madden 2004). Moreover, the List of Prokaryotic Names with Standing in Nomenclature (LPSN) hosted at https://lpsn.dsmz.de (Euzéby 1997; Parte 2014, 2018; Parte et al. 2020) was thoroughly searched for the verification of the derived type strains. Next, the obtained 16S rRNA gene sequences of the studied type strains, including MR41, were subjected to a multiple sequence alignment (MSA) (Pirovano and Heringa 2008) using Clustal Omega (Sievers and Higgins 2014). Phylogeny was inferred using the maximum likelihood (ML) method (Yang 2007) after a thorough model test employing MEGA X software (Kumar et al. 2018). The model with the lowest Bayesian Information Criterion (BIC) value was used for inferring phylogeny (Posada and Buckley 2004; Luo et al. 2010; Saha et al. 2019). The bootstrap consensus tree was inferred from 1000 replicates (Felsenstein 1985).

Study of bacterial growth in normal and heavy metal stressed condition

The isolate MR41 was subjected to growth in normal TSB (HiMedia) as well as in TSB supplemented with heavy metals (pH 7.2) to study and compare the growth profile in the absence and presence of different heavy metals. For this, the TSB medium was supplemented with different heavy metals at their maximum tolerable concentrations (MTCs) for MR41, such as Cd (500 µg/mL), Co (200 µg/mL), Ni (1400 µg/mL), Zn (1100 µg/mL), and Cr (100 µg/mL). Each set of growth experiments was carried out in triplicate. The O. D. values were taken every 2 h interval at 600 nm wavelength of light using a visible spectrophotometer (Systronics, type 166). The O. D. values were used to construct growth curves of the isolate MR41.

Cumulative metal stress detection and simultaneous growth kinetics analysis

Two and three different heavy metals in equal proportions were supplemented in each set of culture media to detect double or triple cumulative metal stress tolerance. The growth of the isolate MR41 was monitored using the pour plate method. Bacterial growth was assessed in different concentrations starting from 50 µg/mL and up to the limit, where no growth was observed.

The effect of triple cumulative metal stress was studied using the heavy metal combinations of Ni:Cd:Zn and Co:Zn:Cd in equal proportions, and further corroborated by growth curve analysis in TSB medium. Based on the response of the isolate MR41 to triple cumulative metal stress in solid media, the combination of Co:Cd:Zn (showing comparatively vigorous growth) was used for growth study in liquid TSB. The medium was supplemented with the three heavy metals Co:Cd:Zn at MTCs (50 µg/mL in equimolar proportions), and the growth of MR41 was measured in terms of O. D. at 600 nm at 2 h interval for constructing the growth curve.

FTIR spectroscopy based metabolic profiling in presence and absence of heavy metals

FTIR spectroscopy is an effective and swift analytical technique for investigating the compositional changes occurring within bacterial cells in response to various stress factors, including heavy metals (Kepenek et al. 2020). It effectively determines the structural changes in the bacterial cell upon molecular binding with heavy metal ions and provides information regarding the nature of heavy metal interactions (Faghihzadeh et al. 2016). FTIR analysis of MR41 cell mass grown in different heavy metals and under normal conditions (without metal) was carried out to detect the functional groups present in the bacterial cell surface and involved in innate bacterial resistance during metal-induced stress. Initially, the isolate MR41 was grown separately in TSB containing Cd, Co, Ni, Zn, and Cr at the concentration of 500 µg/mL, 200 µg/mL, 1400 µg/mL, 1100 µg/mL and 100 µg/Ml, respectively, along with the control. MR41 was also inoculated in Co:Cd:Zn supplemented (50 µg/mL in equimolar proportions) medium (pH 7.2). After 24 h of growth at 35 °C, the cell pellets from each culture were separated by centrifugation using a mini centrifuge (REMI RM 02 Plus) in 2 mL Eppendorf tubes (RCF = 2000*g* for 20 min). These pellets were dried in a hot air oven for 24 h at 45–50 °C after washing thoroughly with 1X PBS buffer solution (pH 7.4). The resulting dry masses were powdered using agate mortar and pestle. Finally, the powdered samples were subjected to FTIR analysis. For each sample, about 1 mg was mixed with 100 mg of KBr to make a 13 mm die pellet. FTIR analysis was carried out on a Vertex 80 FTIR system (Bruker, Germany) at SAIF, IIT Bombay and the spectral data were recorded. Sample scanning was carried out 64 times for each sample. OPUS 8.2 software was used for processing the spectral data. The spectral data were collected over the range 450–4000 cm^{-1} with 4 cm^{-1} spectral resolution (Afzal et al. 2017). The average spectral data were baseline corrected and subsequently normalized with respect to specific bands.

Genomic DNA isolation, whole genome shotgun sequencing and assembly

Sample DNA extraction was done by Qiagen DNeasy Blood and Tissue Kits, and the library was prepared using KAPA Hyper Preparation Kit according to the manufacturer's protocols (Gautam et al. 2019). Paired-end high quality paired reads of length 151 bp were sequenced on the Illumina NovaSeq 6000 platform. Genome sequencing was carried out at Bionivid Technology Pvt. Ltd., Bangalore, India. Bacterial genome assembly was done in steps including quality control (QC) and pre-processing of raw reads, primary assembly, assembly validation and final draft assembly. Initially, QC and pre-processing of FASTQ files such as adaptor trimming and quality filtering were carried out to provide clean data for downstream analysis using the Fastp tool (V 0.20.0) (Chen et al. 2018) with the parameters; length required 75, length limit 201, and phred score (Q30). After that, de novo assembly was done using the tool SPAdes (V 3.13.0) (Bankevich et al. 2012) with the minimum length 200 bp to produce long contiguous pieces of sequence (contigs) from the high quality (HQ) reads. Then, from the constructed primary genome assembly, the 16S rRNA gene sequence was predicted utilising the tool Barrnap 0.9, which detects the location of ribosomal RNA genes present within the genome. In the next step of assembly validation, the filtered HQ reads of the isolate MR41 were aligned to the respective assembled genome using the aligner Bowtie2 (V 2.2.2) (Langmead and Salzberg 2012) to determine the percentage of HQ reads utilized in making the draft genome assembly. Finally, the assembled primary genome and its nearest reference found in NCBI were used for assembly polishing using the web-based tool MeDuSa (Bosi et al. 2015), which rearranges the assembled genome contigs into scaffolds. The raw data was generated for a depth of 250× to cover the entire genome.

Gene prediction, annotation, and functional characterization

The draft genome was subjected to gene prediction and annotation using the Rapid Annotation using Subsystem Technology (RAST) server (Aziz et al. 2008; Overbeek et al. 2014; Brettin et al. 2015), a fully automated service for annotating complete or nearly complete genomes and providing high-quality annotations. The parameters such as genetic code 11, and E value cutoff for selection CDSs 1e-20 were considered. Moreover, prophage sequences and genomic islands were screened and detected in the draft genome employing the tool PHASTER (PHAge Search Tool Enhanced Release) (Arndt et al. 2016) and IslandViewer4 (Bertelli et al. 2017), respectively.

Calculation of average nucleotide identity (ANI)

For identifying bacteria, determination of ANI serves as one of the most essential and widely used parameters (Figueras et al. 2014). An ANI value above 95% between two genomes indicates their close relatedness, i.e., they belong to the same species (Jain et al. 2018). ANI was calculated using Chun-Lab's online ANI Calculator (Yoon et al. 2017).

Estimation of metal uptake through high-resolution ICP–MS

The metal uptake capability of MR41 was estimated at different phases of growth such as lag, log and stationary phase through ICP–MS, an analytical technique through which elements can be detected and measured in any biological fluid (Wilschefski and Baxter 2019). During cumulative metal stress detection, MR41 was observed to exhibit prolific growth in the presence of triple cumulative metals Co:Zn:Cd (50 µg/mL each, pH 7.2). ICP–MS was thus used to detect the metal accumulation by MR41 cells under triple cumulative metal stress Co:Zn:Cd (50 µg/mL each). Cultures were drawn at different growth phases and processed after digestion with concentrated HNO_3. The cultures were centrifuged using a mini centrifuge (REMI RM 02 Plus) in 2 mL Eppendorf tubes (RCF = 2000 *g* for 20 min). Then, the cell pellets were thoroughly washed with PBS buffer solution (1X, pH 7.4) three times, and concentrated HNO_3 was added to the cell pellets. The supernatants were also separately mixed with concentrated HNO_3 (for peptone digestion). The final volume was made up to 10 mL with sterile distilled water (Challaraj Emmanuel et al. 2011). The clear and transparent solutions of supernatants and pellets were analyzed using a high-resolution ICP–Mass Spectrometer (Make: Element XR, Model: Thermo Fisher Scientific, Germany) at SAIF, IIT Bombay.

Comparative CUB analysis of MR41 with reference to other Pseudomonads

The isolate MR41 showed marked resistance against a variety of heavy metals. The genes responsible for conferring resistance to all these heavy metals were considered for CUB analysis, and their orthologs from other *Pseudomonas* spp. were retrieved from the NCBI database. A total of 39 different species of *Pseudomonas* from soil and plant parts with available whole genome sequences in NCBI were retrieved (Supplementary Table S2) to compare their codon biology with MR41. Moreover, LPSN (Parte et al. 2020) was also thoroughly consulted to assure that the genomes of the taxonomically valid type strains were included in the study. The orthologs with improper annotation data and other errors were deliberately kept out of this study. Apart from the heavy metal resistance genes, two housekeeping genes, *rpoB* and *trpB*, were also considered for comparative CUB analysis due to their antiquity and functional constancy. The *rpoB* gene is a constituent of the bacterial information processing system that encodes the beta subunit of the bacterial RNA polymerase enzyme, while *trpB* encodes the beta subunit of the tryptophan synthase enzyme. Both these genes are highly conserved in bacteria.

Genome and genic level analysis of codon usage pattern

Different codon usage parameters such as Nc (Wright 1990), guanine and cytosine content on the first (GC1), second (GC2) and third position (GC3) of a codon, codon adaptation index or CAI (Sharp and Li 1987) were calculated to comprehend the pattern of codon usage. The tools employed for this purpose include CodonW (Peden 1999), ACUA (Vetrivel et al. 2007) and in house PERL scripts. The Nc is one of the best measures of codon usage, and its value ranges between 20 to 61, where a lower value indicates higher codon usage bias and vice versa. GC content plays a crucial role in genome evolution (Sharp et al. 2005). GC3 content, along with GC1 and GC2, mediates a vital role in genomic and genic organisation (Song et al. 2017). The CAI, another important index for calculating potential gene expression, ranges from 0 to 1 (Sharp and Li 1987). Relative synonymous codon usage (RSCU), which is a measure of the ratio between the observed and expected values of synonymous codons, was calculated using the online CAIcal server (freely available at http://genomes.urv.es/CAIcal) (Puigbò et al. 2008). An RSCU value of 1 indicates no bias, while a value > 1 indicates biased codon usage. After estimating the genomic codon usage parameters, genic CUB data were mined from the genomic data to comprehend the codon usage pattern of the different genes related to heavy metal resistance considered in this study.

Nc plot, neutrality plot and parity rule 2 plot (PR2) analysis

Nc plots (Wright 1990) depicting the correlation between Nc and GC3 were plotted on a graph to detect variations in intra or interspecific synonymous codon usage patterns of genome and genes. These plots help elucidate the mechanistic forces affecting CUB (Sun et al. 2016; Wang et al. 2018).

A neutrality plot depicts the interrelationship between GC12 (the average of GC1 and GC2) and GC3 to investigate the mutation–selection equilibrium in shaping codon usage bias (Zhang et al. 2007). If GC12 is completely neutral with GC3, the slope of the regression line will be 1 (i.e., all the points lie on the diagonal line directing no selection pressure), while a slope of zero indicates

no effect of directional mutational pressure (i.e., all the points are present on the line parallel to the abscissa) (Sueoka 1988). The regression curve thus obtained helps determine the degree of neutrality.

Another plot used for evaluating the impact of mutation and selection on codon usage is the PR2 plot. It is a plot of AT bias [A3/(A3 + T3)] as the ordinate and GC bias [G3/(G3 + C3)] as the abscissa (Parvathy and Udayasuriyan 2022). In the centre of such plots, A = T and G = C (where both coordinates are 0.5) signify no bias between mutation and selection pressure (Sueoka 1995; Nasrullah et al. 2015; Parvathy et al. 2022).

Correspondence analysis (CoA) based on RSCU

The trend in codon usage variation among genes was investigated utilizing CoA (Greenacre and Hastie 1987), a multivariate statistical method based on RSCU values of 59 sense codons. CoA depicts the distribution of synonymous codons and orthologous genes in a multidimensional space. This generates a series of orthogonal axes to identify the trends interpreting data variation (Liu 2006; Suzuki et al. 2008; Nie et al. 2014).

The entire workflow adopted in this study is schematically represented in Fig. 1.

Results

Initial characterization of the potent isolate

The bacterial isolate MR41 was designated as the most potent one, since it could resist multiple heavy metals. The MIC values of MR41 was found to be quite higher for the metals Ni(II) (1500 µg/mL), Cd (600 µg/mL), Zn (1200 µg/mL), Fe(II) (1400 µg/mL), Cu (600 µg/mL), and As (500 µg/mL). Culture plates depicting in vitro resistance of MR41 in TSA medium supplemented with different heavy metals (Cd, Fe, Ni, Zn) at 500 µg/mL concentration are shown in Fig. 2.

Morphological and biochemical characterization of MR41

Gram staining and SEM analysis of the isolate MR41 revealed a Gram-negative, small rod-shaped bacterium. The length and breadth of the rod-shaped cells were found to range from 1.4 to 2.2 µm and 0.61 to 0.72 µm, respectively. The SEM micrograph of MR41 is shown in Supplementary Fig. S1. In terms of biochemical attributes, MR41 responded positively to the tests for lipid hydrolysis, citrate utilization, gelatine liquefaction, nitrate reduction, indole production, catalase activity and negatively to starch hydrolysis, oxidase and urease test. The results of some of these tests are depicted in Supplementary Fig. S2.

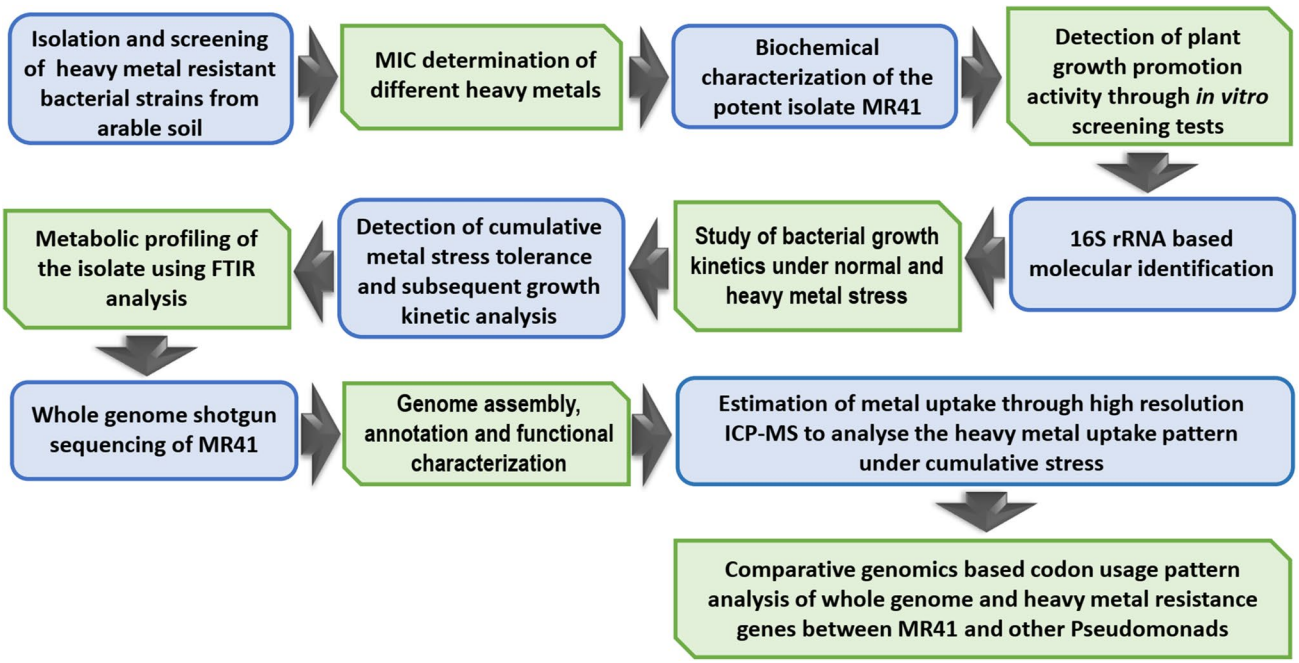

Fig. 1 Graphical representation of the entire workflow used in the present study

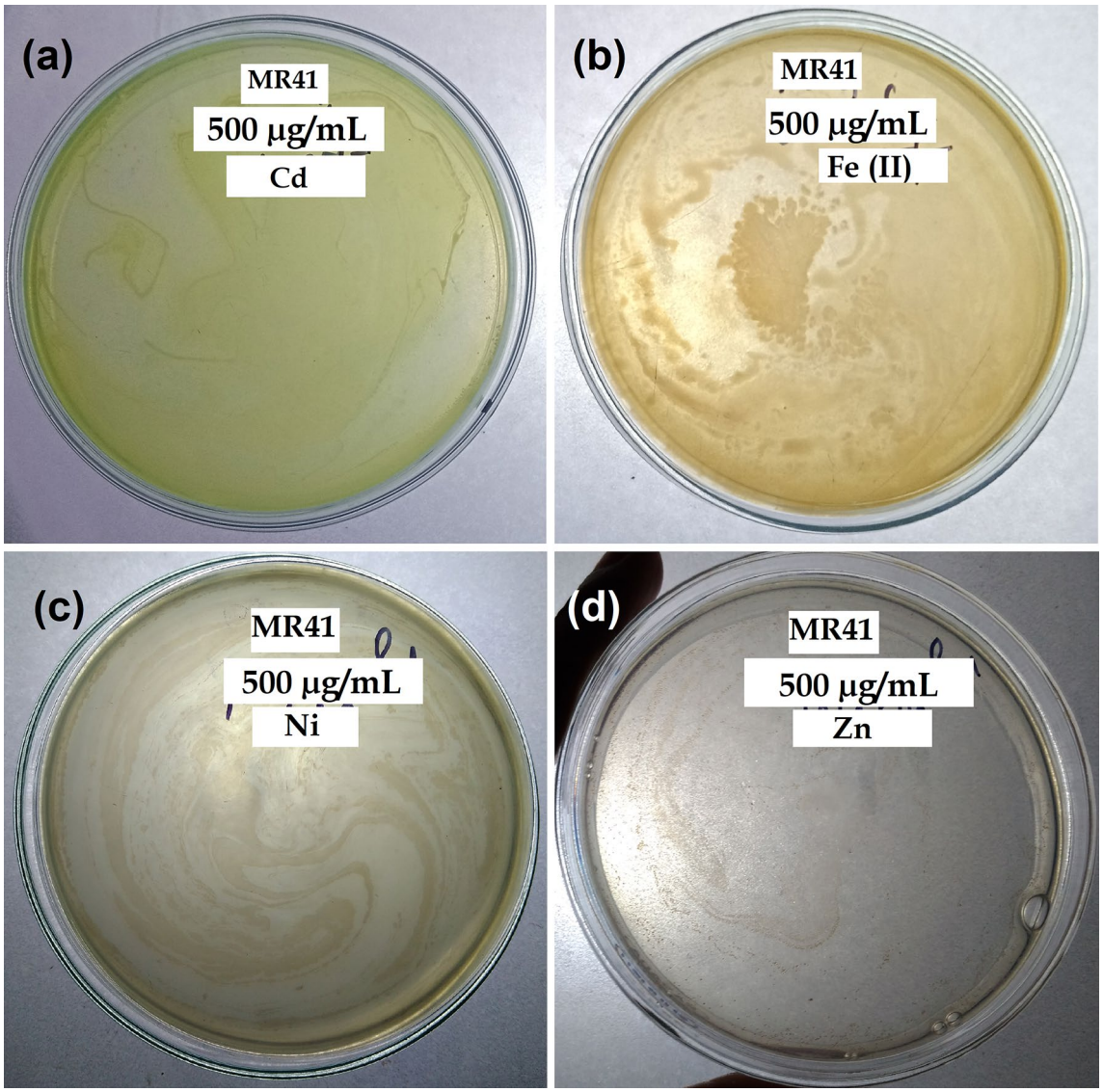

Fig. 2 Culture plates depicting in vitro resistance of MR41 in different heavy metal supplemented media at 500 µg/mL concentration for **a** Cd, **b** Fe (II), **c** Ni, and **d** Zn

Based on Bergey's Manual of Systematic Bacteriology, 2nd edition, volume 2, part B, (page no 355, table BXII.**γ**.112), all the biochemical and morphological features provisionally identified the isolate as a strain of *Pseudomonas* sp.

Screening for in vitro plant growth promotion activity

The isolate MR41 responded positively to all the screening tests for plant growth promotion assays, including phosphate solubilisation, siderophore production, HCN production, zinc solubilisation, IAA, and ammonia production. The response of MR41 to the phosphate solubilisation test in the NBRIP medium with 0.025% bromophenol blue is shown in Supplementary Fig. S3.

16S rRNA based molecular identification and phylogenetic analysis

Molecular identification of MR41 using 16S rRNA gene sequence analysis revealed the bacterium to be identical to the type strain *Pseudomonas aeruginosa* JCM 5962^{T}. The accession number for the 16S rRNA gene sequence submitted in GenBank is ON491428.

Kimura 2-parameter model (Kimura 1980) with gamma (G) distributed rate variation among sites along with significant proportion of invariable sites (I) or K2 + G + I

model possessing the lowest Bayesian Information Criterion (BIC) value was found to be the best model for inferring phylogeny. Molecular phylogenetic analysis also revealed the close resemblance of MR41 with the type strain *Pseudomonas aeruginosa* JCM 5962^{T}, depicted

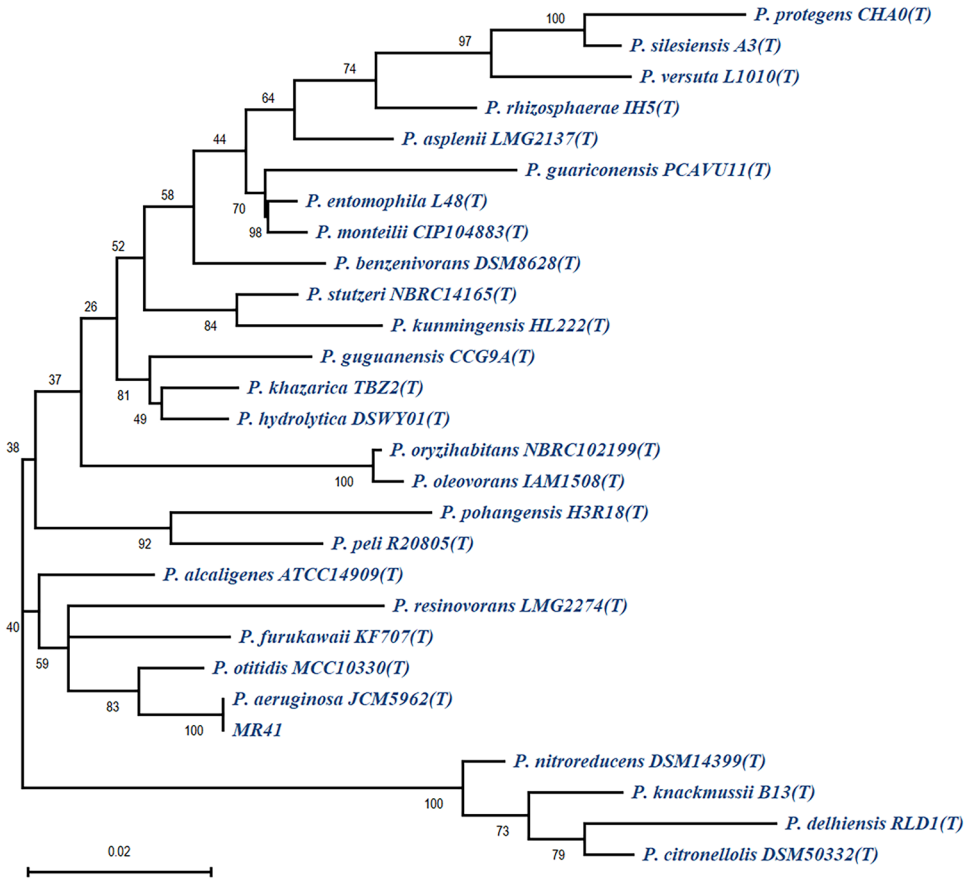

Fig. 3 Molecular phylogenetic tree based on 16S rRNA gene depicting the relationship of the isolate MR41 with other type strains of *Pseudomonas* spp. inferred using Maximum Likelihood method and Kimura 2-parameter model with 1000 bootstrap replicates. A discrete Gamma distribution along with rate variation model was used to model the evolutionary rate differences among sites

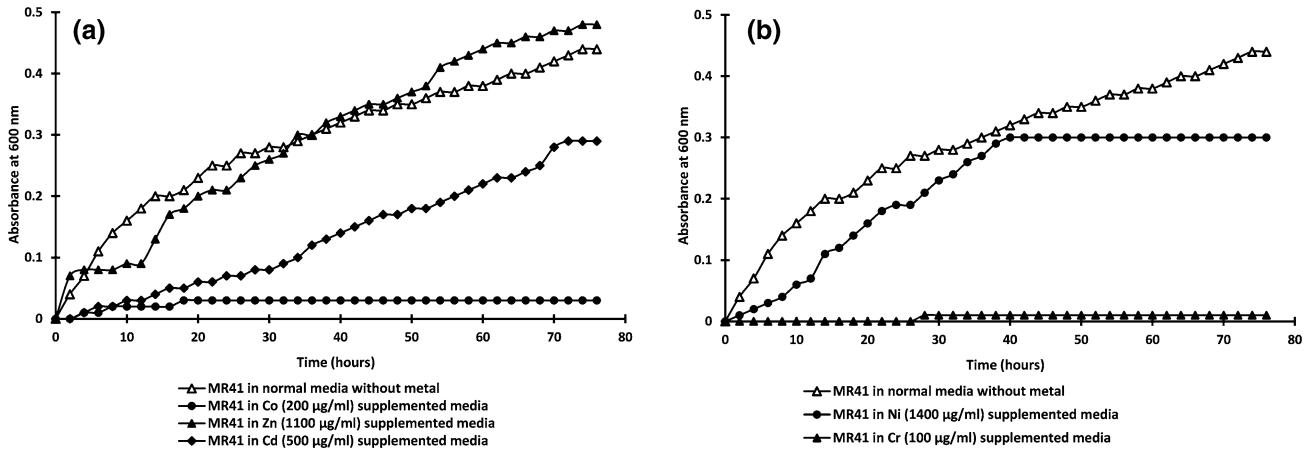

Fig. 4 Growth pattern of the isolate MR41 in normal TSB medium as well as in Cd (500 µg/mL), Co (200 µg/mL), Zn (1100 µg/mL), Ni (1400 µg/mL), and Cr (100 µg/mL) supplemented TSB medium at MTCs in 37 °C. **a** Depicting growth in presence of Cd, Co and Zn including normal and **b** depicting growth of MR41 in presence of Ni and Cr against normal medium

in the form of proper dichotomous branching with 100% bootstrap support as seen in the phylogenetic tree (Fig. 3).

Effect of Cd, Co, Ni, Cr, and Zn on the growth MR41

The heavy metals Cd, Ni, Cr and Co were found to inhibit growth, as was evident by the prolonged lag phase in the growth curve shown in Fig. 4. The analysis of growth revealed the stimulatory effect of Zn on MR41 as depicted in Fig. 4. The bacterial growth curve for Zn supplemented medium was found to be in accordance with the control medium, where no heavy metals were added. The growth curve also suggests that the inhibitory effect of Cd is relatively greater than that of Ni on the growth of MR41.

Study of cumulative metal stress on the growth of MR41

The isolate MR41 responded positively to the double cumulative metal combinations (two different heavy metals mixed) of Zn:Ni, Zn:Cd, Zn:Cu and Zn:Cr combinations (500 µg/mL of each). In the case of triple cumulative metal stress (three different heavy metals), MR41 demonstrated significant growth in Co:Zn:Cd supplemented culture plate (50 µg/mL each) with discrete colonies. These results were corroborated by simultaneous growth kinetics analysis in the TSB medium. The growth curve in Fig. 5 represents a sigmoid curve with a prolonged lag period of 22 h, indicated by the plateau phase on the curve.

Metabolic profiling using FTIR in the absence and presence of different heavy metals

The differential spectral profile of the isolate MR41 subjected to growth in the presence of different heavy metals was derived using an FTIR study. In the control set (bacteria grown without any heavy metal), the derived

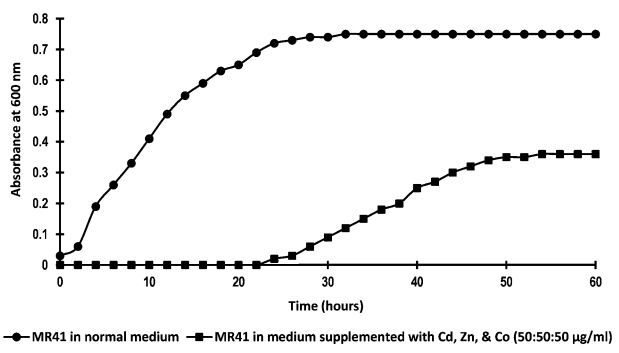

Fig. 5 Growth curve of MR41 in both normal and triple cumulative metal stressed condition (Cd, Zn, and Co supplemented TSB with 50 µg/mL concentration of each metal) incubated at 37 °C

spectrum showed two characteristic peaks, respectively at 3422.37 cm^{-1} and 3304.90 cm^{-1}, due to the stretching of the O–H bond of carboxylic acid and N–H bond of amino groups present on the bacterial cell surface (Choudhary and Sar 2009). Similarly, in the case of Co (200 µg/mL) and cumulative Co:Cd:Zn (50:50:50 µg/mL) incorporated samples, two successive peaks were detected at the same spectral range, i.e., at 3421.12 cm^{-1} and 3296.33 cm^{-1} for Co, whereas 3412.83 cm^{-1} and 3303.17 cm^{-1} for triple cumulative metals (Co:Cd:Zn) loaded sample. A change in this region with a single sharp peak at 3425.02 cm^{-1}, 3433.23 cm^{-1}, 3433.78 cm^{-1} and 3422.91 cm^{-1}, respectively, for Cd, Zn, Ni and Cr loaded bacterial cell mass also revealed the involvement of hydroxyl and amino groups in the process of metal binding to the bacterial surface. Other complex absorption peaks were observed in all the metal accumulated samples, including the control, at the 2853–2960 cm^{-1} region due to the asymmetric stretching of –C–H bond of the $–CH_2$ groups associated with CH_3 group (Beech et al. 1999). Protein related bands were also prominent in all the samples in the range of 1650–1656 cm^{-1} and 1540–1545 cm^{-1}, respectively, for –C=O of amide I and –NH/–C=O combination of amide II bonds (Kazy et al. 1999). The appearance of a peak in the region of 1650–1656 cm^{-1} (amide I absorption region) may be due to the alpha helical secondary structure of protein though amino sugars present in the bacterial cell wall may form the peak in the same region (Kazy et al. 2006; Choudhary and Sar 2009). In the control set, the sharp peak at 1447.39 cm^{-1} appeared due to the presence of carboxyl groups, while a shift to a lower energy level was found upon metal binding, i.e., 1408.79 cm^{-1} for Cd, 1443.67 cm^{-1} for Zn, 1403.78 cm^{-1} for Co, and 1403.79 cm^{-1} for Ni binding. In the case of Cr uptake, such peaks were not found. In Co:Cd:Zn loaded samples, the splitting of the peak at 1441.35 cm^{-1} and 1406.35 cm^{-1} was observed. Due to the symmetric stretching of COO– vibration, sharp peaks were found in the sample after metal binding in comparison to the control, i.e., at 1385.22 cm^{-1} for Zn, 1366.37 cm^{-1} for Ni, 1367.64 cm^{-1} for Co:Cd:Zn loading, while two successive sharp peaks were observed in the same region for the metal Cd (1385.77 cm^{-1}, 1308.39 cm^{-1}) and Cr (1383.94 cm^{-1}, 1366.91 cm^{-1}). Moreover, in each of the metal loaded samples along with the control, strong peaks at 1222–1230 cm^{-1} and 1051–1083 cm^{-1} region appeared due to vibrations of carboxyl and phosphate groups present in the different cellular components of bacteria, such as peptidoglycan and phospholipids (Kazy et al. 2006). One of the interesting observations in the overall analysis was the appearance of two completely different sharp peaks in Ni and Cr loaded samples, respectively, at 1714.06 cm^{-1} and 1714.14 cm^{-1} wavenumbers compared to the control. It might be due to the presence of the carbonyl group (Hu et al. 2007) formed due to the interaction of metals with cellular proteins in the

bacterial cell. In both treated and untreated control samples, the fingerprint region falling between wavenumbers 900–600 cm^{-1} (Faghihzadeh et al. 2016) contained a highly complex series of absorptions. Figure 6 represents the comparative FTIR spectral pattern of MR41 in response to different heavy metals.

Whole genome shotgun sequence of MR41

The whole genome sequence of MR41 revealed several genomic features depicted in Table 1. A total of 11,579,290 quality filtered reads with an average read length of 151 bp were obtained. The final assembled genome of MR41 consists of 6,211,544 bases arranged into 15 contigs (N50 and maximum sequence length of 6,195,387 bp) with an average GC content of 66.54%. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAJEJW000000000. The version described in this paper is version JAJEJW010000000.

Gene prediction, annotation and functional characterization

A total of 5948 genes were predicted in the final assembled MR41 genome, including 5883 protein coding genes and 1340 hypothetical genes, while the total number of rRNA and tRNA genes were 34 and 138, respectively. A circular representation of the entire MR41 genome generated using the interactive genome visualization tool DNAPlotter (Carver et al. 2008) is given in Fig. 7. The RAST annotation data of MR41 has been provided as a Supplementary file (MR41 annotation.csv). Analysis of the MR41 genome revealed the presence of several genes involved in metal homeostasis and transport. Different genes associated with heavy metal resistance such as *czcA, czcB, czcC, czcD, chrA, copA, copB*, *cadA*, *corA*, *corC*, *znuA*, *zur*, *fur*, etc. were detected in the MR41 genome. Table 2 describes the genes present in the MR41 genome involved in heavy metal resistance. Figure 8 depicts the complete subsystem category distribution along with the feature counts. Moreover, two prophage containing regions have been detected and identified in the respective genome of MR41 (Supplementary Table S3). A total of seven (07) genomic islands containing 102 protein encoding genes (including 38 hypothetical genes) were predicted in the genome.

Table 1 Characteristic features of MR41 genome

Metric	MR41 genome
Number of genes predicted	5948
Number of protein coding genes	5883
Number of characterized proteins	4478
Number of hypothetical/putative proteins	1340
Number of protein non-coding genes	65
Number of rRNA genes	34
Number of tRNA genes	138
Number of proteins with pathway annotation	45

ANI calculation

OrthoANIu value revealed 99.37% identity between the MR41 genome and the whole genome sequence of *Pseudomonas aeruginosa* JCM 5962 (the nearest relative of MR41 depicted through 16S rRNA gene-based molecular

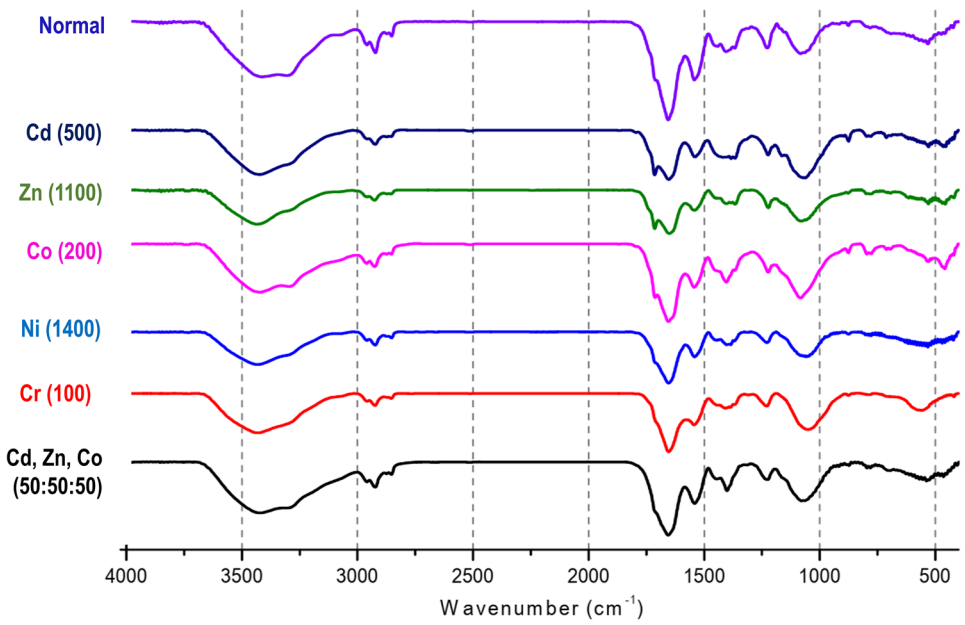

Fig. 6 Stacked FTIR plot of MR41 under normal and heavy metal stress. The relative concentrations of heavy metals for MR41 growth are given in parentheses

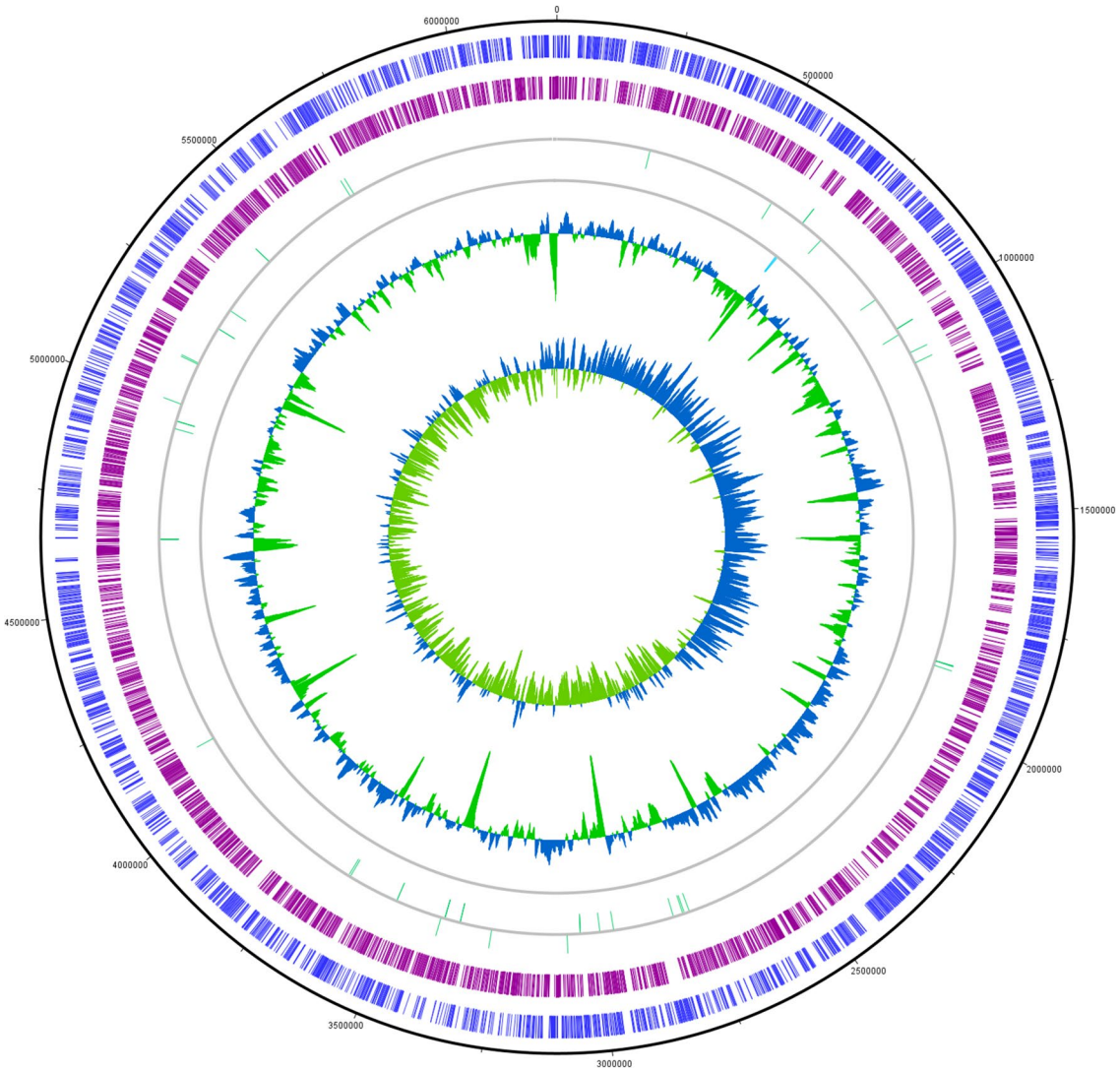

Fig. 7 Circular representation of MR41 whole genome using DNA-Plotter interactive genome visualization tool (Carver et al. 2008). From the outside to the inside—Track 1 (blue) represents the forward strand CDSs, Track 2 (magenta) reverse strand CDSs, Track 3 and 4 (green) represents forward and reverse strand tRNA genes, respectively, Track 5 (cyan) represents all rRNA genes, Track 6 and 7 represents variation in GC content and GC skew for each gene in the genome, respectively (green = below average, while blue = above average)

identification), which was the greatest among all the *Pseudomonas* strains present in the NCBI database.

Estimation of metal uptake utilizing ICP–MS during different phases of growth under triple cumulative metal stress

The *czcCBA* operon encoding the CzcCBA proton driven RND type efflux pump was detected in the MR41 genome (Fig. 9). This pump is commonly found in *Pseudomonas*, governing intrinsic resistance to the heavy metals Co/Zn/Cd (czc) (Perron et al. 2004). ICP–MS was used to detect the metal uptake pattern of MR41 under concurrent stress to the heavy metals Co, Cd and Zn (50:50:50 µg/mL each). Co uptake was found to be the minimum, as it was found in relatively higher amounts in the supernatant than in the cell pellet. The maximum amount of Co was detected in the supernatant during the stationary phase. A general trend was observed throughout the growth of MR41 in batch cultures, where during the early log phase, after a prolonged period of lag, the Cd and Zn accumulation were found to be higher in the cell pellets (32 h of growth), followed by the gradual decrease of Cd and Zn in cell pellets in mid logarithmic phase (36 h of growth). This is most probably due to the active efflux of these metal ions in the medium during active growth. Increased

Table 2 Genes involved in heavy metal resistance in the MR41 genome

Sl. No.	Gene	Product	Function
1.	*arsB*	Arsenic Transporter/arsenical efflux pump membrane protein ArsB/ACR3 family arsenite efflux transporter	Arsenic resistance
2.	*arsC*	Arsenate Reductase (glutaredoxin/Thioredoxin)	Arsenic resistance
3.	*arsH*	Arsenical Resistance Protein	Arsenic resistance
4.	*arsR*	transcriptional regulator, arsR family	Arsenic resistance
5.	*chrA*	Chromate efflux transporter	Chromium resistance
6.	*copA*	Copper resistance system multicopper oxidase	Copper resistance
7.	*copB*	Copper resistance protein B	Copper resistance
8.	*czcA*	cusA/czcA family heavy metal efflux RND transporter	Cd, Zn, Co resistance
9.	*cadA*	Cadmium translocating p type ATPase	Cd, Zn, Pb, Cu resistance
10.	*znuA*	Zn ABC transporter substrate binding protein	Zn transport
11.	*znuB*	Zn ABC transporter permease subunit	Zn transport
12.	*znuC*	Zn ABC transporter ATP binding protein	Zn transport
13.	*zur*	Zinc uptake transcriptional repressor	Zn resistance
14.	*Fur*	Ferric iron uptake transcriptional regulator	Fe resistance
15.	*corC*	hlyc/corC family transporter, magnesium and cobalt efflux protein	Co, Mg, Ni resistance
16.	*corA*	Mg and Co transporter cor A family protein	Co, Mg, Ni resistance
17.	*mgtE*	Mg/Co/Ni transporter	Mg, Co, Ni resistance
18.	*czcB*	Cobalt/zinc/cadmium efflux RND transporter, membrane fusion protein	Cd, Zn, Co resistance
19.	*czcC*	Cobalt/zinc/cadmium efflux RND transporter, outer membrane protein	Cd, Zn, Co resistance
20.	*czcD*	Cobalt/zinc/cadmium resistance protein CzcD	Cd, Zn, Co resistance
21.	*FieF*	Ferrous-iron efflux pump FieF	Iron transport
22.	*FeoC*	Ferrous iron-sensing transcriptional regulator FeoC	Iron transport
23.	*FeoB*	Ferrous iron transporter FeoB	Iron transport
24.	*FeoA*	Ferrous iron transporter-associated protein FeoA	Iron transport
25.	*HitB*	Iron (III)-transport system permease HitB	Iron transport

amounts of Cd and Zn were again detected in the cell pellets during the stationary phase of growth. These results have been graphically shown in Fig. 10.

Comparative genomic features of Pseudomonads considered in this study

On analysing the GC content of the genomes considered in this study, the genomic GC content of MR41 was found to be 66.54%. Among all the strains, *P. otitidis* WP8-S7-CRE-03 (isolation source—soil, given in Supplementary Table S2) displayed the maximum GC content (67.1%). In terms of genome size, *P. chlororaphis* PCLAR04, a soil isolate, demonstrated the largest genome with 7,305,652 bp sequences, while MR41 had a genome of 6,211,544 bp. The total number of genes present in MR41 was 5948, while that of *P. otitidis* WP8-S7-CRE-03 and *P. chlororaphis* PCLAR04 were 5778 and 6647, respectively. The entire data are depicted in Table 3.

Comparative CUB analysis of MR41 with other *Pseudomonas* spp.

A total of 17 genes associated with resistance against the heavy metals such as, Cd, Ni, Zn, Co, Cr, As, Cu, Fe, Pb with orthologs in the genome of MR41 and other Pseudomonads were selected for CUB analysis. These genes include *arsB, arsC, arsH, arsR* which are involved with arsenic resistance; *copA* and *copB* which are associated with copper resistance; *czcA, czcD, cadA*, for Cd, Co and Zn resistance and others. Detailed attributes of these 19 genes are given in Supplementary Table S4.

Genomic and genic codon usage pattern analysis

The mean genomic Nc (Nc_{Gen}) of all the organisms included in this study ranged from 33.2 to 48.9 (Table 3). MR41 demonstrated the lowest Nc_{Gen} of 33.2 with a standard deviation (SD) of ± 5.56, indicating the highest genomic codon bias among all *Pseudomonas* strains. The highest Nc_{Gen} of

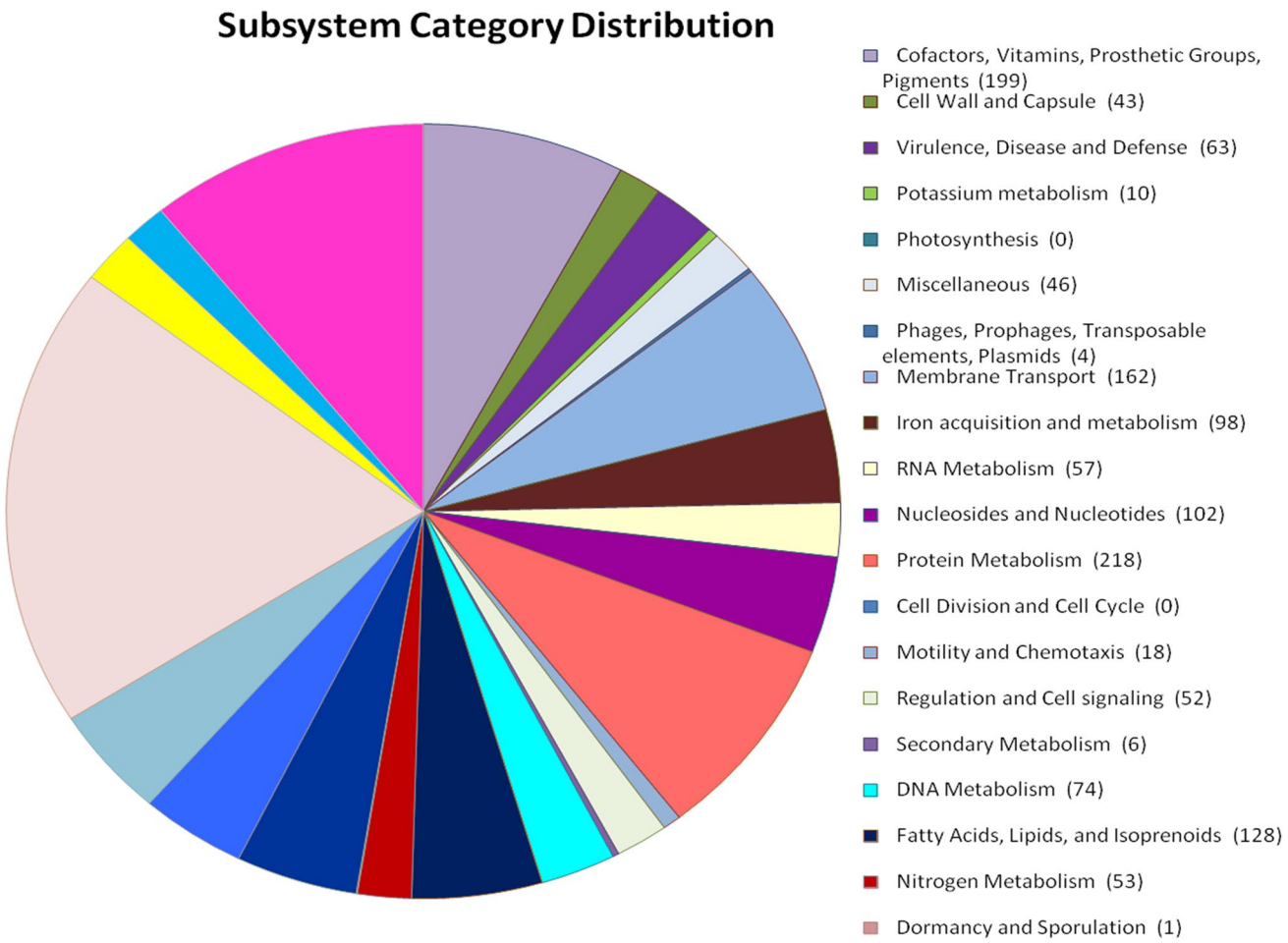

Fig. 8 Complete subsystem category distribution of MR41 genome (circular presentation)

48.9 ± 6.67 was depicted by *P. cannabina* pv. *alisalensis*, isolated from Japanese radish, causing bacterial leaf blight. In general, this organism serves as a foliar plant pathogen in the plant families of Brassicaceae and Poaceae (Sakata et al. 2021). About 20% of the organisms included in this study exhibited Nc_{Gen} below 40, suggesting relatively higher codon bias within these genomes. The mean genomic GC3 ($GC3_{Gen}$) content of the organisms ranged from 59% (in *P. cannabina* pv. alisalensis) to 86% (MR41).

The correlation between the different codon usage parameters was determined to comprehend the codon usage pattern further. Correlation using Spearman's rank order revealed a significant negative correlation between Nc_{Gen} and $GC3_{Gen}$. In contrast, a positive correlation was detected between $GC3_{Gen}$ and genomic GC at $p < 0.01$ level for all the organisms studied. The isolate MR41 also exhibited a significant negative correlation between Nc_{Gen} and $GC3_{Gen}$ ($\rho = -0.75$, $p < 0.01$) and a significant positive correlation between $GC3_{Gen}$ and genomic GC ($\rho = 0.342$, $p < 0.01$). The results of Spearman's rank order correlation analysis are given in Table 3. Codon usage pattern analysis of the 17 coding sequences associated with heavy metal resistance (Table 4), showed that the Nc values of the genes *cadA* (28.09), *arsR* (30.19), *zur* (27.08), *mgtE* (31.09) and the housekeeping gene *rpoB* (28.77) were the lowest in MR41 compared to the other bacterial strains. These genes were also found to depict elevated GC3 in MR41, of which *cadA* (0.92), *arsR* (0.88) and *rpoB* (0.84) were the highest among all the concerned *Pseudomonas* spp. The genes *zur* (0.92) and *mgtE* (0.85) also showed high GC3 values. All of these indicate greater codon bias existing in the MR41 genes.

In terms of codon usage pattern, MR41 demonstrated proximity to *P. otitidis* WP8-S7-CRE-03. *P. otitidis* was initially isolated from clinical specimens of patients with ear infection. However, its widespread distribution in different non-clinical environments such as soil and wastewater has also been reported (Miyazaki et al. 2020). The lowest Nc values were observed in majority of *P. otitidis* WP8-S7-CRE-03 heavy metal resistance genes, including *czcA* (27.24), *znuA* (27.32), *znuB* (25.67), *znuC* (25.49), *chrA* (29.88), *arsC* (25.49), *corC* (26.76) along with the housekeeping gene *trpB* (25.49). All these genes also depicted

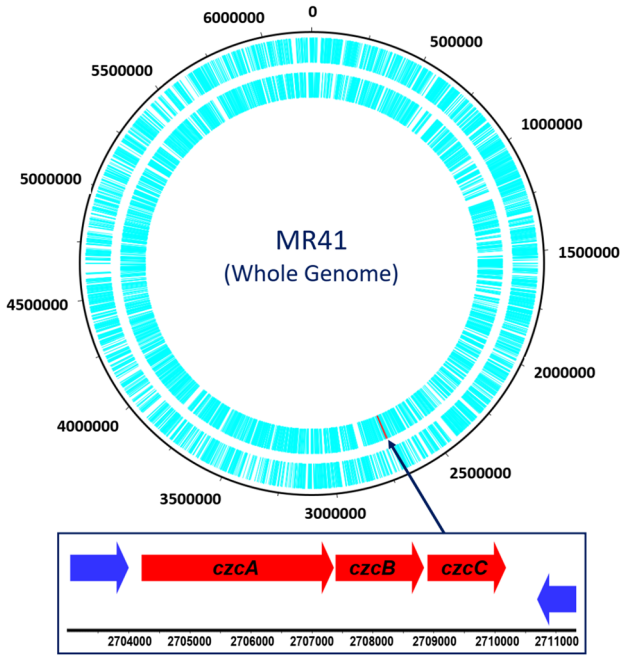

Fig. 9 Schematic representation of *czcCBA* operon with respect to MR41 whole genome

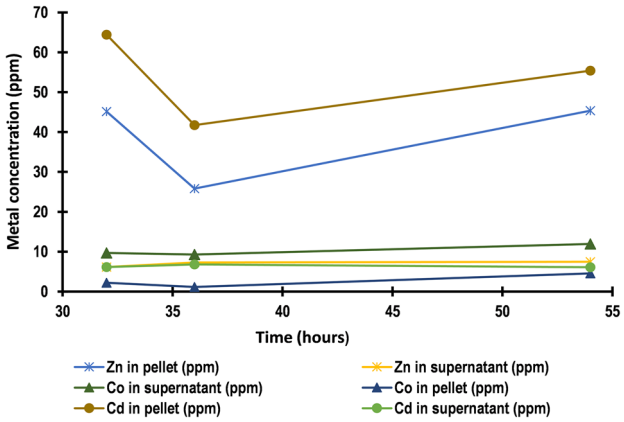

Fig. 10 Metal content profile in different growth phases of MR41 in triple cumulative metal stressed condition (Cd:Zn:Co = 50:50:50) revealed through ICP–MS analysis

relatively higher GC3 values. The Nc and GC3 of the genes *arsB*, *arsH*, *fur*, *corA*, *copA*, and *copB* from MR41 showed resemblance to other soil-dwelling strains. The lowest Nc for the gene *arsH* (28.91), *fur* (32.83), and *copB* (30.12) was observed in the soil dwellers *P. oleovorans* T9AD, *P. nitroreducens* HBP1 and *P. chlororaphis* PCLAR04, respectively. For the genes *arsH*, *fur*, and *copB*, MR41 demonstrated lower Nc values such as those of the other *Pseudomonas* spp. suggesting strong codon bias. *P. oleovorans* prevails in diverse habitats, and the strain T9AD was isolated from marine environments (Wang et al. 2021). *P. nitroreducens* HBP1 has been reported as sewage sludge isolate possessing biocide degradation ability (Carraro et al. 2020). *P. chlororaphis* largely occurs in soil and is used as a potential biocontrol agent against plant diseases due to its ability to synthesize phenazine (Morohoshi et al. 2017). For all the genes studied, a strong anti-correlation was observed between Nc and GC3, while a positive correlation between CAI and GC3 at $p < 0.01$ level of significance was found to exist (Table 4). Among all the studied genes, the *trpB* gene demonstrated the strongest negative correlation between Nc and GC3 ($\rho = -0.948$). Genes such as *czcA* ($\rho = -0.934$), *mgtE* ($\rho = -0.944$), *corC* ($\rho = -0.94$), *copB* ($\rho = -0.926$), *cadA* ($\rho = -0.891$) along with both the housekeeping genes showed significant negative correlation between Nc and GC3 at $p < 0.01$ significance level. Nc was significantly negatively correlated with CAI in all the studied genes, suggesting the relative influence of selection over mutation (Parvathy et al. 2022). The gene *copA* displayed the strongest anti-correlation between Nc and CAI ($\rho = -0.833$, $p < 0.01$) (Table 4).

Nc plot, neutrality plot and PR2 plot analysis

The Nc plots of all the studied *Pseudomonas* genomes depicted right centric aggregation of the coding sequences, though slight mid centric aggregation was observed in some species along with right shift. MR41 also demonstrated a typical right centric Nc plot and scattered clustering of majority of the genes from mid to right centric region lying below the parabolic line of trajectory. In contrast, few genes lied on and above the expected curve (Fig. 11a). This observation is an indication of the fact that translational selection along with mutation pressure to some extent determines the codon usage pattern of MR41 genome. On the other hand, all the selected genes exhibited similar pattern of right centric aggregation in Nc plots. Almost all the genes with low Nc values were situated below the null hypothesis curve, suggesting that translational selection is the primary force influencing codon usage variation. Both the housekeeping genes under study (*rpoB* and *trpB*) and few genes such as *corC, mgtE*, etc. displayed relatively tight clustering of the coding sequences in the Nc plots, as shown in Fig. 12.

The neutrality plots help determine the extent of variation between mutation pressure and natural selection. In MR41, $GC3_{Gen}$ was anti-correlated with genomic GC2 ($r = -0.121$, $p < 0.01$) and genomic GC12 ($r = -0.108$, $p < 0.01$), suggesting low mutation bias or high conservation of GC contents across the MR41 genome. The slope of the neutrality plot (Fig. 11b) revealed that mutation pressure only accounts for 5.83% of MR41 codon usage, and natural selection and other factors accounted for the remaining 94.17%, suggesting a dominant role of natural selection over mutation. Similarly, for all the studied genes, the correlation coefficient depicted the relative neutrality within the range of 2.86–23.9%

Table 3 Codon usage parameters of the studied *Pseudomonas* spp. strains

Sl no.	Organism name	Nc_{Gen}	Standard deviation of Nc_{Gen}	$GC3_{Gen}$	Standard deviation of $GC3_{Gen}$	Genomic GC content (%)	Genome size (bp)	Total genes	Nc:GC3	GC3:GC
1.	*Pseudomonas aeruginosa*	33.2	5.56	0.86	0.08	66.54	6,211,544	5948	−0.75	0.342
2.	*Pseudomonas asplenii*	41.8	7.60	0.72	0.12	61.2	6,529,636	5762	−0.806	0.453
3.	*Pseudomonas avellanae*	45.05	6.95	0.67	0.11	58.6	6,242,845	5957	−0.777	0.518
4.	*Pseudomonas azotoformans*	41.43	7.55	0.73	0.12	61.1	6,884,339	6370	−0.816	0.448
5.	*Pseudomonas brassicacearum*	47.96	7.42	0.61	0.14	60.8	6,738,544	6023	−0.652	0.199
6.	*Pseudomonas cannabina* pv. *alisalensis*	48.99	6.67	0.59	0.13	58.8	6,103,677	5511	−0.644	0.269
7.	*Pseudomonas chlororaphis*	39.36	8.29	0.76	0.13	62.6	7,305,652	6647	−0.836	0.527
8.	*Pseudomonas congelans*	44.49	6.81	0.68	0.10	59.2	5,843,414	5100	−0.757	0.456
9.	*Pseudomonas corrugata*	43.39	7.22	0.71	0.06	60.5	6,332,970	5621	−0.787	0.427
10.	*Pseudomonas fluorescens*	42.12	7.51	0.72	0.12	60.5	6,957,239	6356	−0.791	0.449
11.	*Pseudomonas fragi*	42.78	6.79	0.69	0.10	59.3	5,072,729	4645	−0.716	0.5
12.	*Pseudomonas frederiksbergensis*	43.73	7.14	0.69	0.11	59.4	6,595,804	5981	−0.753	0.439
13.	*Pseudomonas furukawaii*	36.84	8.56	0.80	0.14	65.5	6,183,134	5813	−0.876	0.403
14.	*Pseudomonas koreensis*	40.86	7.14	0.72	0.10	59.9	6,444,290	7141	−0.723	0.555
15.	*Pseudomonas lactis*	48.14	7.53	0.60	0.14	60	6,179,543	5504	−0.619	0.192
16.	*Pseudomonas lurida*	41.66	7.30	0.72	0.12	60.8	6,100,532	5567	−0.77	0.365
17.	*Pseudomonas mendocina*	45.05	7.69	0.65	0.15	62.7	5,120,146	5081	−0.769	0.154
18.	*Pseudomonas monteilii*	40.3	7.80	0.73	0.12	62	5,802,791	5492	−0.821	0.334
19.	*Pseudomonas mosselii*	44.5	9.06	0.66	0.17	64.4	5,742,165	5233	−0.746	0.207
20.	*Pseudomonas nitroreducens*	36.71	8.06	0.79	0.13	64.7	6,874,121	7116	−0.865	0.321
21.	*Pseudomonas oleovorans*	37.48	8.17	0.78	0.14	64.5	5,623,977	5298	−0.858	0.393
22.	*Pseudomonas orientalis*	47.67	7.79	0.61	0.14	60.6	5,919,814	5276	−0.677	0.187
23.	*Pseudomonas oryzihabitans*	37.63	6.98	0.78	0.12	65.1	4,834,356	4459	−0.815	0.266
24.	*Pseudomonas otitidis*	33.84	8.36	0.83	0.14	67.1	6,198,093	5778	−0.855	0.372
25.	*Pseudomonas plecoglossicida*	39.63	8.32	0.75	0.12	62.4	6,233,254	5771	−0.844	0.442
26.	*Pseudomonas poae*	41.55	7.33	0.72	0.11	60.5	6,530,734	5877	−0.773	0.424
27.	*Pseudomonas pohangensis*	43.17	6.88	0.68	0.11	59.6	3,769,689	3527	−0.807	0.548
28.	*Pseudomonas protegens*	45.54	8.30	0.64	0.16	63.4	6,868,303	6260	−0.726	0.197
29.	*Pseudomonas rhizosphaerae*	40.47	7.32	0.75	0.11	62	4,688,635	4214	−0.815	0.356
30.	*Pseudomonas rhodesiae*	42.42	7.68	0.71	0.12	60.3	6,109,631	5507	−0.792	0.479
31.	*Pseudomonas savastanoi*	46.04	6.58	0.66	0.09	58.1	6,156,114	5759	−0.739	0.466
32.	*Pseudomonas silesiensis*	43.37	7.53	0.70	0.11	59.6	6,823,539	6166	−0.782	0.517
33.	*Pseudomonas simiae*	42.3	7.33	0.71	0.11	60.3	6,199,521	5679	−0.785	0.368
34.	*Pseudomonas stutzeri*	40.87	6.63	0.74	0.08	60.1	6,263,601	5589	−0.716	0.564
35.	*Pseudomonas syringae*	45.01	7.02	0.68	0.11	58.9	6,191,793	5393	−0.769	0.496

Table 3 (continued)

Sl no.	Organism name	Nc_{Gen}	Standard deviation of Nc_{Gen}	$GC3_{Gen}$	Standard deviation of $GC3_{Gen}$	Genomic GC content (%)	Genome size (bp)	Total genes	Nc:GC3	GC3:GC
36.	*Pseudomonas syringae pv. tomato*	48.98	6.77	0.60	0.12	58.4	6,397,126	5919	−0.653	0.328
37.	*Pseudomonas veronii*	41.94	8.02	0.72	0.12	60.7	7,110,596	6618	−0.821	0.537
38.	*Pseudomonas versuta*	45.17	6.45	0.66	0.10	58.2	5,149,788	4671	−0.684	0.5
39.	*Pseudomonas viridiflava*	44.11	6.77	0.69	0.10	59.2	5,997,220	5297	−0.763	0.409
40.	*Pseudomonas yamanorum*	44.4	7.19	0.68	0.11	59	6,943,411	6670	−0.775	0.479

Table 4 Spearman's rank correlation analysis involving the different codon usage parameters of the studied genes

Sl. no.	Name of the genes	Nc:GC3	Nc:CAI	CAI:GC3
1.	*rpoB*	−0.93	−0.0732	0.161
2.	*cadA*	−0.891	−0.78	0.628
3.	*czcA*	−0.934	−0.815	0.677
4.	*znuA*	−0.828	−0.667	0.551
5.	*znuB*	−0.85	−0.811	0.773
6.	*znuC*	−0.93	−0.788	0.824
7.	*zur*	−0.871	−0.811	0.789
8.	*fur*	−0.687	−0.522	0.75
9.	*corA*	−0.889	−0.505	0.337
10.	*corC*	−0.94	−0.387	0.392
11.	*mgtE*	−0.944	−0.577	0.497
12.	*copA*	−0.883	−0.833	0.74
13.	*copB*	−0.926	−0.702	0.694
14.	*chrA*	−0.895	−0.751	0.817
15.	*arsB*	−0.915	−0.664	0.551
16.	*arsH*	−0.805	−0.719	0.68
17.	*arsC*	−0.852	−0.736	0.669
18.	*arsR*	−0.712	−0.56	0.316
19.	*trpB*	−0.948	−0.726	0.625

showing the limited influence of mutation pressure. Supplementary Fig. S4 illustrates such minor role of mutation (contrarily dominant role of selection) in shaping codon usage bias of some of the genes such as *rpoB, trpB, arsR*, *zur* in *Pseudomonas* spp.

Analysis of the PR2 plots showed that the mean values for AT bias were a bit higher than the mean values for GC bias in all the studied *Pseudomonas* spp. including MR41. In MR41 (Fig. 11c), the values for GC and AT biases were, respectively, 0.509 and 0.569, which indicated more frequent use of A and T than G and C, suggesting relatively intense selection pressure in shaping codon bias. Except for *chrA*, this trend was observed in all the other studied genes, i.e., relatively more frequent usage of AT than GC. PR2 plots of some of the studied genes are shown in Supplementary Fig. S5.

CoA based on RSCU

CoA on RSCU revealed variations and trends in codon usage among the different genes. The first axis captured significant variation in all the genes studied and was the central explanatory axis for interpreting codon usage variation among genes. Of all the genes, in *czcA*, the first axis captured the maximum variation (49.14%). Similarly, in *rpoB* and *copA*, the first axis captured 42.45% and 36.59% of the total variation, respectively. Supplementary Fig. S6 depicts such a

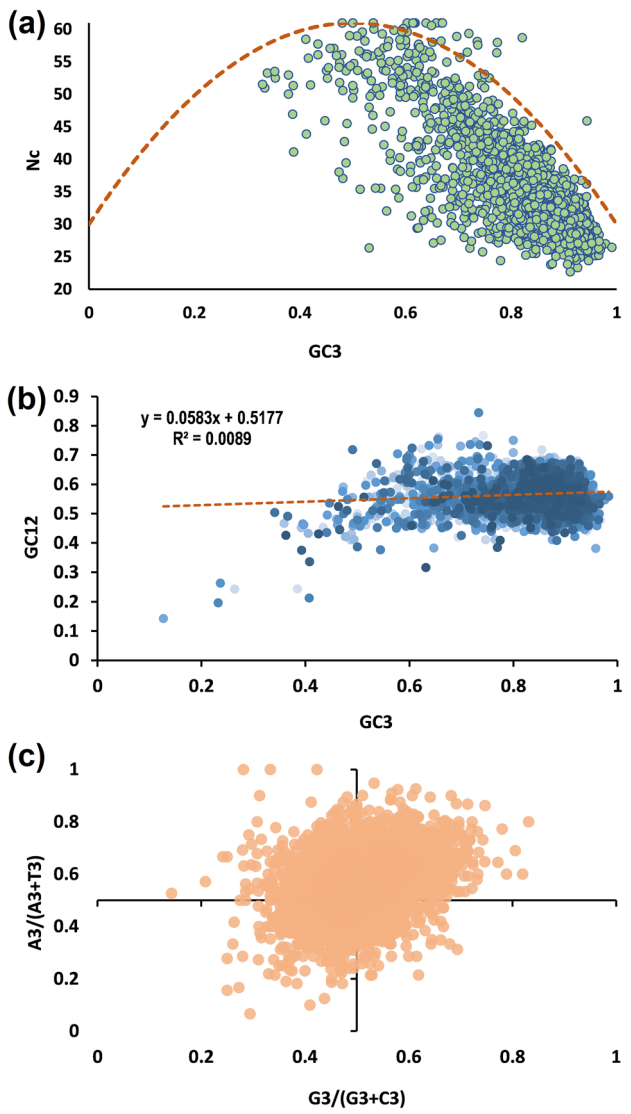

Fig. 11 Graphs depicting **a** Nc plot, **b** neutrality plot, and **c** PR2 bias (parity rule 2) plot of MR41 genome

pattern of variation based on the genes. Moreover, it was observed that in each of the CoA plots, the distribution of orthologs from different *Pseudomonas* species differs along the axes. For example, in *rpoB*, these are widely scattered along the positive quadrant of the first axis, whereas in *czcA*, most of the points displayed aggregation in the intercept region of the first and second axes (shown in Fig. 13).

Discussion

In this study, a potent multi heavy metal resistant bacterial isolate designated MR41 was obtained from soil samples of paddy cultivating arable land. The isolate demonstrated resistance against multiple heavy metals either singly or cumulatively at relatively high concentrations. Initial 16S rRNA gene-based phylogeny followed by whole genome sequencing showed the isolate is identical to *Pseudomonas aeruginosa* JCM 5962^{T}. Furthermore, whole genome sequence analyses of MR41 revealed several genomic properties of the isolate regarding its multiple heavy metal resistance. During this study, the growth kinetics of MR41, along with its metal uptake capability, was studied elaborately in the presence of a variety of heavy metals, such as Cd, Zn, and Co. This was done to evaluate the resistance and uptake potentiality of the 'czc' system of resistance, the most characteristic proton-driven RND type efflux pump found in *Pseudomonas* involved in Cd, Zn, and Co resistance (Perron et al. 2004). Different studies have been put forward (Zeng et al. 2009; Fashola et al. 2016; Izrael-Živković et al. 2019) to decipher Cd resistance and its concomitant removal (Chellaiah 2018) by *P. aeruginosa*.

Metabolic profiling of MR41 using FTIR confirmed the presence of different functional groups, such as O–H, C–O, N–H, and C–H, which varied not only in response to the presence or absence but also on the type of heavy metal, inducing stress. In comparison to that of the normal, variable functional groups, including amide I and amide II region were assigned upon the adsorption of metal ions onto the bacterial cell mass in heavy metal-containing culture as depicted in other studies (Hu et al. 2007; Afzal et al. 2017; Kepenek et al. 2020; Sodhi et al. 2020). During growth, Zn was found to induce growth up to a certain level compared to the other heavy metals. This finding is well in line with other studies (Hassen et al. 1998; Kuffner et al. 2008; Pederick et al. 2015). In the case of triple cumulative metal stress, the derived growth curve showed a prolonged lag phase. This suggests that, in the initial phase of growth, the bacteria prepares itself for growth to survive in a highly stressed environment (Rolfe et al. 2012; Bertrand 2019; Hamill et al. 2020).

Analysis of the genome sequence data depicted the presence of many genes associated with heavy metal resistance in accordance with the bacterial response during in vitro screening tests regarding metal tolerance. The genome size of MR41 was found to be 6,211,544 base pairs coding for 5883 protein CDSs, signifying a substantially larger genome than the other Pseudomonads included in this study. The detection of prophage in the genome of MR41 was in line with some other studies reporting the genome of metabolically versatile *Pseudomonas* sp. (Nelson et al. 2002; Berger et al. 2021). Analyzing the sequence data of the MR41 whole genome, the genes *czcA, czcB, czcC, czcD,* and the Cd resistance gene *cadA* were detected in MR41. During the ICP–MS study, the Cd affinity of the isolate was found to be higher than the other metals. This indicates a correlation between the czc system and cad system, which is strongly involved in Cd resistance of MR41. Such interrelationship

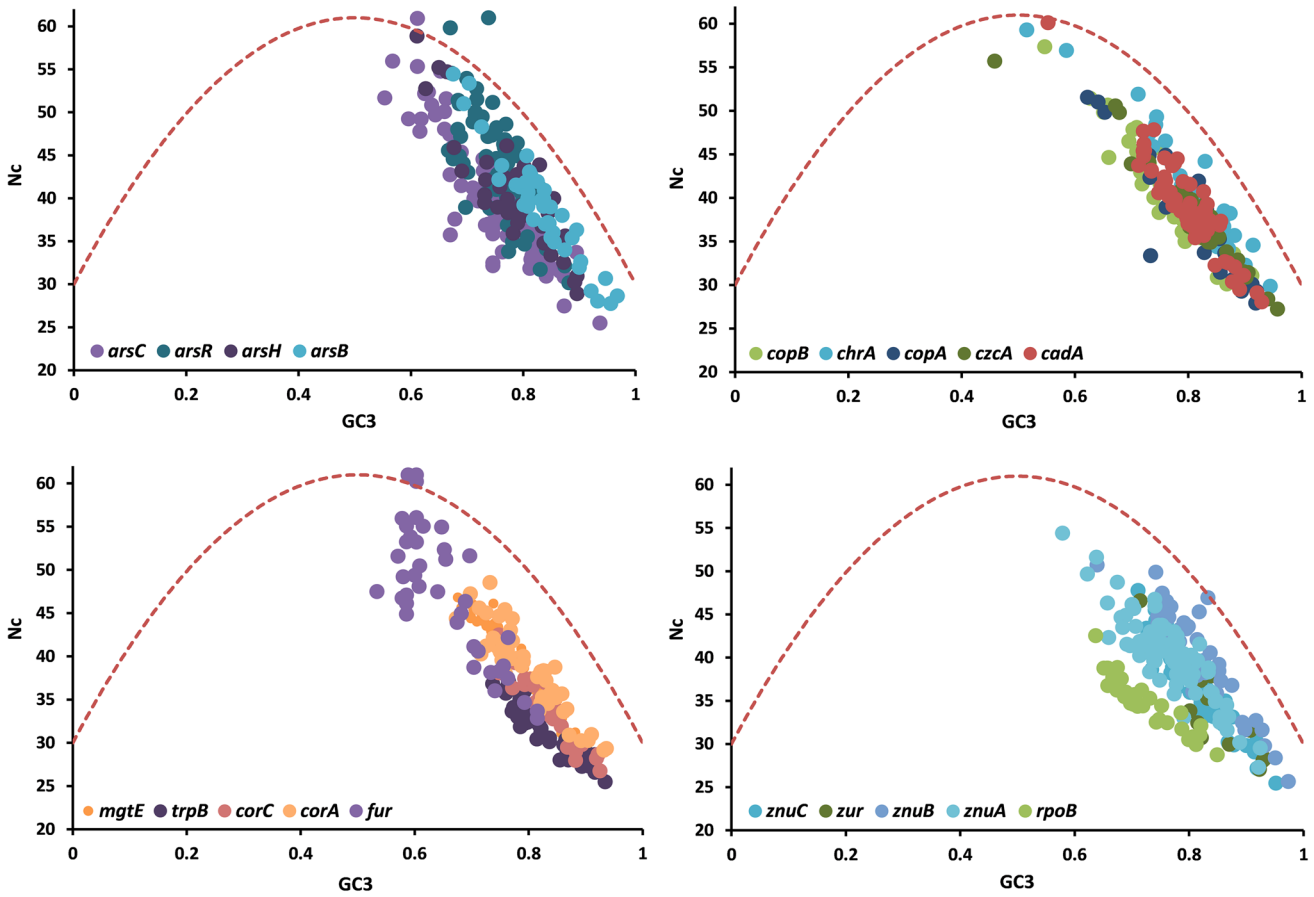

Fig. 12 Nc plot of all the genes involved in heavy metal resistance considered in this study along with the housekeeping genes *trpB* and *rpoB*

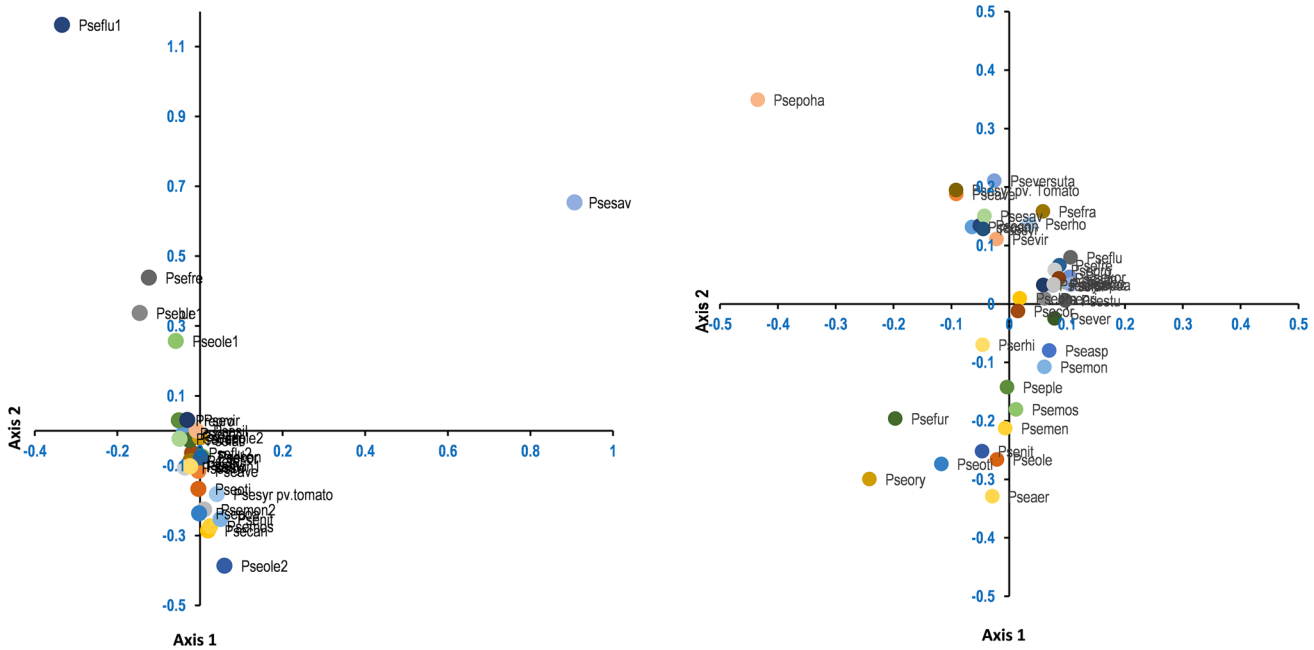

Fig. 13 Correspondence analysis (CoA) on RSCU of **a** *rpoB* and **b** *czcA* gene

between czc system and cad system has also been established in other studies (Liu et al. 2021). Moreover, a cross-link between Cd and Zn resistance was observed in MR41, exhibiting an elevated level of Zn tolerance, which signifies that P-type ATPase *cadA*, along with *czcABC* genes, affects both Zn and Cd tolerance in MR41 (Liu et al. 2015; Ducret et al. 2020).

In the present study, 17 heavy metal resistant genes, and two housekeeping genes, were selected for comparative codon usage analysis. The biased usage of codons for the genes *cadA, czcA, znuA, zur, mgtE*, *corA*, including the housekeeping gene *rpoB*, was detected. Moreover, MR41 depicted the lowest mean genomic Nc and the highest mean genomic GC3, suggesting relatively strong codon bias among all the organisms under study. MR41 also displayed relatively higher CUB for both the genome and genes under study. Nc plots, along with neutrality and PR2 plots, depicted natural selection as the primary force shaping codon usage variation. The members showing proximity to MR41 in terms of codon usage patterns were found to occupy the same habitat, i.e., soil. Simultaneously, the ICP–MS study reported the lowest biosorption capacity for the metal Co in this study, which points towards the reduced affinity for Co (Bazzi et al. 2020). This feature was found to vary with the growth phase of the bacteria (Raja et al. 2006). The increased uptake of Cd by MR41 revealed in this study may be exploited for devising strategies to carry out Cd bioremediation (Lin et al. 2016; Chellaiah 2018).

The present study emphasizes multi-metal resistant, arable soil-dwelling natural indigenous bacteria in heavy metal uptake and removal efficiency. The morphological, biochemical and molecular analyses showed the most potent heavy metal resistant MR41 obtained in this study, sharing maximum identity with the type strain *Pseudomonas aeruginosa* JCM 5962^{T}. The whole genome sequence of MR41 (obtained from arable soil) revealed the presence of many genes such as *czcA, czcD, czcB, czcC, cadA*, *znuA, zur, mgtE*, *copA, copB*, and others conferring resistance against a variety of heavy metals. The isolate also responded positively to different in vitro screening tests for detecting plant growth promotion activity. Metabolic profiling through FTIR depicted differential spectral patterns due to the presence of various functional groups such as O–H, C–O, N–H, and C–H on metal binding with the bacterial cell wall. Cd uptake efficiency of the strain grown under cumulative Cd, Zn, and Co stress was revealed utilising ICP–MS. This was further corroborated by the presence of *czcABC* gene cluster and cad system involved in Cd resistance. A comprehensive and comparative codon usage study demonstrated the MR41 genome to be strongly biased compared to all the studied Pseudomonads. The results of the codon usage analysis could be utilized further to understand the evolution of genes conferring heavy metal resistance in natural environments.

Conclusions

The present study is a novel attempt to characterize and sequence the whole genome of a soil-dwelling bacterium from arable lands of Uttar Dinajpur that exhibits considerable resistance to multiple heavy metals. Future research that incorporates large scale whole genome sequencing, comparative genome, transcriptome and metagenome-based study of soil-borne bacteria from various agricultural fields may help us comprehend the microbial composition of the region's arable lands. In addition, it is necessary to investigate the molecular processes of heavy metal resistance in soil bacteria to develop effective bioremediation agents for polluted locations. Future research is also necessary to comprehend the plant growth-promoting function of heavy metal-resistant arable soil bacteria, particularly through in vivo tests. Moreover, the production of bioconsortia of such potent bacteria to mitigate heavy metal contamination is another crucial area that requires further investigation.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00294-022-01245-z.

Acknowledgements The authors are grateful to Bionivid Technology Pvt. Ltd., Bangalore, India for timely sequencing of the genome and help in assembly, and specially Mr. J. Godwin for his help. The authors would also like to acknowledge DST, Govt. of India, SAIF-IIT Bombay for providing their sophisticated analytical services, and USIC-University of North Bengal for SEM imaging.

Author contributions JS: conceptualization, methodology, investigation, writing—original draft, formal analysis, data curation, and validation. SD: investigation, data curation, writing—original draft, formal analysis, and data curation. AP: conceptualization, supervision, investigation, methodology, writing—original draft, review and editing, validation, visualization, and resources.

Funding None of the authors have received any funding.

Availability of data and materials All data generated or analyzed during this study are included in this published article and its supplementary information files.

Code availability Not applicable.

Declarations

Conflict of interest All the authors declare that they have no conflict of interest.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication All the authors consented to publish the article.

References

Abbas S, Ahmed I, Iida T, Lee Y-J, Busse H-J, Fujiwara T, Ohkuma M (2015) A heavy-metal tolerant novel bacterium, *Alcaligenes pakistanensis* sp. nov., isolated from industrial effluent in Pakistan. Antonie Leeuwenhoek 108:859–870

Afzal AM, Rasool MH, Waseem M, Aslam B (2017) Assessment of heavy metal tolerance and biosorptive potential of *Klebsiella variicola* isolated from industrial effluents. AMB Express 7:184

Aguilar-Barajas E, Ramírez-Díaz MI, Riveros-Rosas H, Cervantes C (2010) Heavy metal resistance in *Pseudomonas*. In: Ramos JL, Alain F (eds) *Pseudomonas*: molecular microbiology, infection and biodiversity. Springer, Dordrecht, pp 255–282

Alengebawy A, Abdelkhalek ST, Qureshi SR, Wang M-Q (2021) Heavy metals and pesticides toxicity in agricultural soil and plants: ecological risks and human health implications. Toxics 9:42

Alhasawi A, Costanzi J, Auger C, Appanna ND, Appanna VD (2015) Metabolic reconfigurations aimed at the detoxification of a multimetal stress in *Pseudomonas fluorescens*: Implications for the bioremediation of metal pollutants. J Biotechnol 20:38–43

Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS (2016) PHASTER: a better faster version of the PHAST phage search tool. Nucleic Acids Res 44:16–21

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA et al (2008) The RAST Server: rapid annotations using subsystems technology. BMC Genom 9:75

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS et al (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477

Bazzi W, Abou Fayad AG, Nasser A, Haraoui L-P, Dewachi O, Abou-Sitta G et al (2020) Heavy metal toxicity in armed conflicts potentiates AMR in *A. baumannii* by selecting for antibiotic and heavy metal co-resistance mechanisms. Front Microbiol 11:68

Beech I, Hanjagsit L, Kalaji M, Neal AL, Zinkevich V (1999) Chemical and structural characterization of exopolymers produced by *Pseudomonas* sp. NCIMB 2021 in continuous culture. Microbiology 145:1491–1497

Berger C, Rückert C, Blom J, Rabaey K, Kalinowski J, Rosenbaum MA (2021) Estimation of pathogenic potential of an environmental *Pseudomonas aeruginosa* isolate using comparative genomics. Sci Rep 11:1370

Bertelli C, Laird MR, Williams KP, Lau BY, Hoad G, Winsor GL, Brinkman FSL (2017) IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res 45:30–35

Bertrand RL (2019) Lag Phase Is a Dynamic, Organized, Adaptive, and Evolvable Period That Prepares Bacteria for Cell Division. J Bacteriol 201:e00697-18

Bosi E, Donati B, Galardini M, Brunetti S, Sagot MF, Lió P et al (2015) MeDuSa: a multi-draft based scaffolder. Bioinformatics 31:2443–2451

Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ et al (2015) RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365

Bric JM, Bostock RM, Silverstone SE (1991) Rapid in situ assay for indoleacetic acid production by bacteria immobilized on a nitrocellulose membrane. Appl Environ Microbiol 57:535–538

Burland TG (2000) DNASTAR's lasergene sequence analysis software. In: Misener S, Krawetz SA (eds) bioinformatics methods and protocols. Humana Press, Totowa, pp 71–91

Cappuccino JG, Welsh C (2017) Microbiology a laboratory manual. Pearson, Upper Saddle River

Cappucino JC, Sherman N (1992) Microbiology: a laboratory manual. Benjamin/Cummings Publishing Company, New York

Carraro N, Sentchilo V, Polák L, Bertelli C, van der Meer JR (2020) Insights into mobile genetic elements of the biocide-degrading bacterium *Pseudomonas nitroreducens* HBP-1. Genes 11:930

Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J (2008) DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 25:119–120

Challaraj Emmanuel ES, Vignesh V, Anandkumar B, Maruthamuthu S (2011) Bioaccumulation of cerium and neodymium by *Bacillus cereus* isolated from rare earth environments of Chavara and Manavalakurichi, India. Indian J Microbiol 51:488–495

Chellaiah ER (2018) Cadmium (heavy metals) bioremediation by *Pseudomonas aeruginosa*: a minireview. Appl Water Sci 8:154

Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:884–890

Chlebek D, Płociniczak T, Gobetti S, Kumor A, Hupert-Kocurek K, Pacwa-Płociniczak M (2022) Analysis of the genome of the heavy metal resistant and hydrocarbon-degrading rhizospheric *Pseudomonas qingdaonensis* ZCR6 strain and assessment of its plant-growth-promoting traits. Int J Mol Sci 23:214

Chong TM, Yin WF, Mondy S, Grandclément C, Dessaux Y, Chan KG (2012) Heavy-metal resistance of a France vineyard soil bacterium *Pseudomonas mendocina* strain S5.2 revealed by whole-genome sequencing. J Bacteriol 194:6366–6366

Choudhary S, Sar P (2009) Characterization of a metal resistant *Pseudomonas* sp. isolated from uranium mine for its potential in heavy metal (Ni^{2+}, Co^{2+}, Cu^{2+}, and Cd^{2+}) sequestration. Bioresour Technol 100:2482–2492

Çolak F, Atar N, Yazıcıoğlu D, Olgun A (2011) Biosorption of lead from aqueous solutions by *Bacillus* strains possessing heavy-metal resistance. Chem Eng J 173:422–428

Ducret V, Gonzalez MR, Leoni S, Valentini M, Perron K (2020) The CzcCBA efflux system requires the CadA P-type ATPase for timely expression upon zinc excess in *Pseudomonas aeruginosa*. Front Microbiol 11:911

Euzéby JP (1997) List of bacterial names with standing in nomenclature: a folder available on the internet. Int J Syst Bacteriol 47:590–592

Faghihzadeh F, Anaya NM, Schifman LA, Oyanedel-Craver V (2016) Fourier transform infrared spectroscopy to assess molecular-level changes in microorganisms exposed to nanoparticles. Nanotechnol Environ Eng 1:1

Fashola MO, Ngole-Jeme VM, Babalola OO (2016) Heavy metal pollution from gold mines: environmental effects and bacterial strategies for resistance. Int J Env Res Public Health 13:1047

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Figueras MJ, Beaz-Hidalgo R, Hossain MJ, Liles MR (2014) Taxonomic affiliation of new genomes should be verified using average nucleotide identity and multilocus phylogenetic analysis. Genome Announc 2:e00927-00914

Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ (2008) Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol 74:2461–2470

Gautam SS, Kc R, Leong KW, Mac Aogáin M, O'Toole RF (2019) A step-by-step beginner's protocol for whole genome sequencing of human bacterial pathogens. J Biol Methods 6:e110

Giller KE, Witter E, McGrath SP (2009) Heavy metals and soil microbes. Soil Biol Biochem 41:2031–2037

Greenacre M, Hastie T (1987) The geometric interpretation of correspondence analysis. J Am Stat Assoc 82:437–447

Hamill PG, Stevenson A, McMullan PE, Williams JP, Lewis ADR, Sudharsan S et al (2020) Microbial lag phase can be indicative of or independent from cellular stress. Sci Rep 10:5948

Hassen A, Saidi N, Cherif M, Boudabous A (1998) Effects of heavy metals on *Pseudomonas aeruginosa* and *Bacillus thuringiensis*. Bioresour Technol 65:73–82

Hassen W, Neifar M, Cherif H, Najjari A, Chouchane H, Driouich RC et al (2018) *Pseudomonas rhizophila* S211, a new plant growth-promoting rhizobacterium with potential in pesticide-bioremediation. Front Microbiol. https://doi.org/10.3389/fmicb.2018.00034

Havryliuk O, Hovorukha V, Patrauchan M, Youssef NH, Tashyrev O (2020) Draft whole genome sequence for four highly copper resistant soil isolates *Pseudomonas lactis* strain UKR1, *Pseudomonas panacis* strain UKR2, and *Pseudomonas veronii* strains UKR3 and UKR4. Current Research in Microbial Sciences 1:44–52

Helal M (2016) Multiple Heavy Metal and Antibiotic Resistance of *Acinetobacter baumannii* Strain HAF – 13 Isolated from Industrial Effluents. Am J Microbiol Res 4:26–36

Hu C, Guo J, Qu J, Hu X (2007) Photocatalytic degradation of pathogenic bacteria with AgI/TiO_2 under visible light irradiation. Langmuir 23:4982–4987

Izrael-Živković L, Beškoski V, Rikalović M, Kazazić S, Shapiro N, Woyke T et al (2019) High-quality draft genome sequence of *Pseudomonas aeruginosa* san ai, an environmental isolate resistant to heavy metals. Extremophiles 23:399–405

Jadhav JP, Kalyani DC, Telke AA, Phugare SS, Govindwar SP (2010) Evaluation of the efficacy of a bacterial consortium for the removal of color, reduction of heavy metals, and toxicity from textile dye effluent. Bioresour Technol 101:165–173

Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S (2018) High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114

Jiang C-Y, Sheng X-F, Qian M, Wang Q-Y (2008) Isolation and characterization of a heavy metal-resistant *Burkholderia sp.* from heavy metal-contaminated paddy field soil and its potential in promoting plant growth and heavy metal accumulation in metal-polluted soil. Chemosphere 72:157–164

Kamran S, Shahid I, Baig DN, Rizwan M, Malik KA, Mehnaz S (2017) Contribution of zinc solubilizing bacteria in growth promotion and zinc content of wheat. Front Microbiol 8:2593

Kang S-M, Asaf S, Khan AL, Lubna A, Khan A, Mun B-G et al (2020) Complete genome sequence of *Pseudomonas psychrotolerans* CS51, a plant growth-promoting bacterium, under heavy metal stress conditions. Microorganisms 8:382

Kazy SK, Sar P, Asthana RK, Singh SP (1999) Copper uptake and its compartmentalization in *Pseudomonas aeruginosa* strains: chemical nature of cellular metal. World J Microbiol Biotechnol 15:599–605

Kazy SK, Das SK, Sar P (2006) Lanthanum biosorption by a *Pseudomonas* sp.: equilibrium studies and chemical characterization. J Ind Microbiol Biotechnol 33:773–783

Kepenek ES, Severcan M, Gozen AG, Severcan F (2020) Discrimination of heavy metal acclimated environmental strains by chemometric analysis of FTIR spectra. Ecotoxicol Environ Saf 202:110953

Khanna K, Jamwal VL, Gandhi SG, Ohri P, Bhardwaj R (2019) Metal resistant PGPR lowered Cd uptake and expression of metal transporter genes with improved growth and photosynthetic pigments in *Lycopersicon esculentum* under metal toxicity. Sci Rep 9:5855

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Kuffner M, Puschenreiter M, Wieshammer G, Gorfer M, Sessitsch A (2008) Rhizosphere bacteria affect growth and metal uptake of heavy metal accumulating willows. Plant Soil 304:35–44

Kumar BL, Gopal DVRS (2015) Effective role of indigenous microorganisms for sustainable environment. 3 Biotech 5:867–876

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359

Lever MA, Torti A, Eickenbusch P, Michaud AB, Šantl-Temkiv T, Jørgensen BB (2015) A modular method for the extraction of DNA and RNA, and the separation of DNA pools from diverse environmental sample types. Front Microbiol. https://doi.org/10.3389/fmicb.2015.00476

Lin X, Mou R, Cao Z, Xu P, Wu X, Zhu Z, Chen M (2016) Characterization of cadmium-resistant bacteria and their potential for reducing accumulation of cadmium in rice grains. Sci Total Environ 569–570:97–104

Liu Q (2006) Analysis of codon usage pattern in the radioresistant bacterium Deinococcus radiodurans. BioSyst 85:99–106

Liu P, Chen X, Huang Q, Chen W (2015) The Role of CzcRS Two-Component Systems in the Heavy Metal Resistance of Pseudomonas putida X4. Int J Mol Sci 16:17005–17017

Liu H, Zhang Y, Wang Y, Xie X, Shi Q (2021) The Connection between Czc and Cad systems involved in cadmium resistance in *Pseudomonas putida*. Int J Mol Sci 22:9697

Lorck H (1948) Production of Hydrocyanic Acid by Bacteria. Physiol Plant 1:142–146

Luo A, Qiao H, Zhang Y, Shi W, Ho SY, Xu W et al (2010) Performance of criteria for selecting evolutionary models in phylogenetics: a comprehensive study based on simulated datasets. BMC Evol Biol 10:242–242

Manara A, DalCorso G, Baliardini C, Farinati S, Cecconi D, Furini A (2012) *Pseudomonas putida* response to cadmium: changes in membrane and cytosolic proteomes. J Proteome Res 11:4169–4179

McGinnis S, Madden TL (2004) BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res 32:W20–W25

Méndez V, Fuentes S, Morgante V, Hernández M, González M, Moore E, Seeger M (2017) Novel hydrocarbonoclastic metal-tolerant *Acinetobacter* and *Pseudomonas* strains from Aconcagua river oil-polluted soil. J Soil Sci Plant Nutr 17:1074–1087

Mesa-Marín J, Pérez-Romero JA, Redondo-Gómez S, Pajuelo E, Rodríguez-Llorente ID, Mateos-Naranjo E (2020) Impact of plant growth promoting bacteria on *Salicornia ramosissima* ecophysiology and heavy metal phytoremediation capacity in estuarine soils. Front Microbiol 11:2148

Miyazaki K, Hase E, Maruya T (2020) Complete genome sequence of *Pseudomonas otitidis* Strain MrB4, isolated from Lake Biwa in Japan. Microbiol Resour Announc 9:e00148-00120

Mohamed RM, Abo-Amer AE (2012) Isolation and characterization of heavy-metal resistant microbes from roadside soil and phylloplane. J Basic Microbiol 52:53–65

Mondal G, Ghosh M, Mazumdar D, Biswas A (2013) Base-line survey for Tulaipanji rice production status in Uttar Dinajpur district of West Bengal, India. J Crop Weed 9:148–150

Morohoshi T, Yamaguchi T, Xie X, Wang WZ, Takeuchi K, Someya N (2017) Complete genome sequence of *Pseudomonas chlororaphis* subsp. *aurantiaca* reveals a triplicate quorum-sensing mechanism for regulation of phenazine production. Microbes Environ 32:47–53

Nasrullah I, Butt AM, Tahir S, Idrees M, Tong Y (2015) Genomic analysis of codon usage shows influence of mutation pressure, natural selection, and host features on Marburg virus evolution. BMC Evol Biol 15:174

Nautiyal CS (1999) An efficient microbiological growth medium for screening phosphate solubilizing microorganisms. FEMS Microbiol Lett 170:265–270
Nelson KE, Weinel C, Paulsen IT, Dodson RJ, Hilbert H, Martins dos Santos VAP et al (2002) Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. Environ Microbiol 4:799–808
Nie X, Deng P, Feng K, Liu P, Du X, You FM, Weining S (2014) Comparative analysis of codon usage patterns in chloroplast genomes of the Asteraceae family. Plant Mol Biol Rep 32:828–840
O'Malley MA, Walsh DA (2021) Rethinking microbial infallibility in the metagenomics era. FEMS Microbiol Ecol. https://doi.org/10.1093/femsec/fiab092
Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T et al (2014) The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res 42:D206-214
Parte AC (2014) LPSN—list of prokaryotic names with standing in nomenclature. Nucleic Acids Res 42:D613-616
Parte AC (2018) LPSN—list of prokaryotic names with standing in nomenclature (bacterio.net), 20 years on. Int J Syst Evol Microbiol 68:1825–1829
Parte AC, Carbasse JS, Meier-Kolthoff JP, Reimer LC, Göker M (2020) List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. Int J Syst Evol Microbiol 70:5607–5612
Parvathy ST, Udayasuriyan V (2022) Codon usage bias. Mol Biol Rep 49:539–565
Parvathy ST, Udayasuriyan V, Bhadana V (2022) Codon usage bias. Mol Biol Rep 49:539–565
Peden JF (1999) Analysis of codon usage. Department of Genetics, University of Nottingham, Nottingham
Pederick VG, Eijkelkamp BA, Begg SL, Ween MP, McAllister LJ, Paton JC, McDevitt CA (2015) ZnuA and zinc homeostasis in *Pseudomonas aeruginosa*. Sci Rep 5:13139
Perron K, Caille O, Rossier C, van Delden C, Dumas J-L, Köhler T (2004) CzcR-CzcS, a two-component system involved in heavy metal and carbapenem resistance in *Pseudomonas aeruginosa*. J Biol Chem 279:8761–8768
Pirovano W, Heringa J (2008) Multiple sequence alignment. Methods Mol Biol 452:143–161
Plotkin JB, Kudla G (2011) Synonymous but not the same: the causes and consequences of codon bias. Nat Rev Genet 12:32–42
Posada D, Buckley TR (2004) model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst Biol 53:793–808
Puigbò P, Bravo IG, Garcia-Vallve S (2008) CAIcal: a combined set of tools to assess codon usage adaptation. Biol Direct 3:38
Raja CE, Anbazhagan K, Selvam GS (2006) Isolation and characterization of a metal-resistant *Pseudomonas aeruginosa* strain. World J Microbiol Biotechnol 22:577–585
Registrar GI (2017) Geographical indications journal no. 97. In: Govt. of India
Ren XM, Guo SJ, Tian W, Chen Y, Han H, Chen E et al (2019) Effects of plant growth-promoting bacteria (PGPB) inoculation on the growth antioxidant activity cu uptake and bacterial community structure of rape (*Brassica napus* L.) grown in Cu-contaminated agricultural soil. Front Microbiol 10:1455
Rolfe MD, Rice CJ, Lucchini S, Pin C, Thompson A, Cameron AD et al (2012) Lag phase is a distinct growth phase that prepares bacteria for exponential growth and involves transient metal accumulation. J Bacteriol 194:686–701
Saha, J., Saha, B.K., Pal Sarkar, M., Roy, V., Mandal, P., and Pal, A. (2019) Comparative Genomic Analysis of Soil Dwelling Bacteria Utilizing a Combinational Codon Usage and Molecular Phylogenetic Approach Accentuating on Key Housekeeping Genes. *Front Microbiol* **10**.
Sakata N, Ishiga T, Ishiga Y (2021) *Pseudmonas cannabina* pv. *alisalensis* TrpA is required for virulence in multiple host plants. Front Microbiol 12:659734
Schwyn B, Neilands JB (1987) Universal chemical assay for the detection and determination of siderophores. Anal Biochem 160:47–56
Sharp PM, Li WH (1987) The codon adaptation index-a measure of directional synonymous codon usage bias and its potential applications. Nucleic Acids Res 15:1281–1295
Sharp PM, Matassi G (1994) Codon usage and genome evolution. Curr Opin Genet Dev 4:851–860
Sharp PM, Bailes E, Grocock RJ, Peden JF, Sockett RE (2005) Variation in the strength of selected codon usage bias among bacteria. Nucleic Acids Res 33:1141–1153
Sievers F, Higgins DG (2014) Clustal Omega, accurate alignment of very large numbers of sequences. Methods Mol Biol 1079:105–116
Šmejkalová M, Mikanová O, Borůvka L (2003) Effects of heavy metal concentrations on biological activity of soil micro-organisms. Plant Soil Environ 49:321–326
Sodhi KK, Kumar M, Singh DK (2020) Multi-metal resistance and potential of *Alcaligenes* sp. MMA for the removal of heavy metals. SN Appl Sci 2:1885
Song H, Gao H, Liu J, Tian P, Nan Z (2017) Comprehensive analysis of correlations among codon usage bias, gene expression, and substitution rate in *Arachis duranensis* and *Arachis ipaënsis* orthologs. Sci Rep 7:14853
Stanbrough R, Chuaboonmee S, Palombo EA, Malherbe F, Bhave M (2013) Heavy metal phytoremediation potential of a heavy metal resistant soil bacterial isolate *Achromobacter* sp. strain AO22. APCBEE Proc 5:502–507
Sueoka N (1988) Directional mutation pressure and neutral molecular evolution. Proc Natl Acad Sci USA 85:2653
Sueoka N (1995) Intrastrand parity rules of DNA base composition and usage biases of synonymous codons. J Mol Evol 40:318–325
Sun S, Xiao J, Zhang H, Zhang Z (2016) Pangenome evidence for higher codon usage bias and stronger translational selection in core genes of *Escherichia coli*. Front Microbiol 7:1180
Suzuki H, Brown CJ, Forney LJ, Top EM (2008) Comparison of correspondence analysis methods for synonymous codon usage in bacteria. DNA Res 15:357–365
Vetrivel U, Arunkumar V, Dorairaj S (2007) ACUA: a software tool for automated codon usage analysis. Bioinformation 2:62–63
Wang L, Xing H, Yuan Y, Wang X, Saeed M, Tao J et al (2018) Genome-wide analysis of codon usage bias in four sequenced cotton species. PLoS ONE 13:e0194372
Wang SZ, Cruaud C, Aury JM, Vallenet D, Poulain J, Vacherie B et al (2021) Complete genome sequences of two *pseudomonas* species isolated from marine environments of the Pacific Ocean. Microbiol Resour Announc 10:e01062-e11019
Wasi S, Tabrez S, Ahmad M (2013) Use of Pseudomonas spp. for the bioremediation of environmental pollutants: a review. Environ Monit Assess 185:8147–8155
Wilschefski SC, Baxter MR (2019) Inductively coupled plasma mass spectrometry: introduction to analytical aspects. Clin Biochem Rev 40:115–133
Wright F (1990) The "effective number of codons" used in a gene. Gene 87:23–29
Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586–1591

Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, Seo H, Chun J (2017) Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol 67:1613–1617

Zeng X-X, Tang J-X, Liu X-D, Jiang P (2009) Isolation, identification and characterization of cadmium-resistant *Pseudomonas aeruginosa* strain E_1. J Cent South Univ Technol 16:416–421

Zhang WJ, Zhou J, Li ZF, Wang L, Gu X, Zhong Y (2007) Comparative analysis of codon usage patterns among mitochondrion, chloroplast and nuclear genes in *Triticum aestivum* L. J Integr Plant Biol 49:246–254